David P. Bazett-Jones · Graham Dellaire
Editors

The Functional Nucleus

Springer

The Functional Nucleus

David P. Bazett-Jones • Graham Dellaire
Editors

The Functional Nucleus

Springer

Editors
David P. Bazett-Jones
Genetics and Genome Biology
The Hospital for Sick Children
Toronto, Ontario
Canada

Graham Dellaire
Dalhousie University
Halifax, Nova Scotia
Canada

ISBN 978-3-319-38880-9 ISBN 978-3-319-38882-3 (eBook)
DOI 10.1007/978-3-319-38882-3

Library of Congress Control Number: 2016951228

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Contents

Part I
Nuclear Periphery

Human Diseases Related to Nuclear Envelope Proteins

Howard J. Worman

Abstract The nuclear envelope has traditionally been looked at as a barrier separating the nucleus and cytoplasm and a complex organelle that disassembles and precisely reassembles during mitosis. However, the combination of cell biological discoveries localizing proteins to the nuclear envelope and human genetic investigations identifying disease-causing genes has show that the nuclear envelope must have tissue-selective functions beyond those general ones. Mutations in genes encoding proteins of the nuclear lamina, nuclear membranes, nuclear pore complexes and perinuclear space have been linked to a wide range of human diseases, sometimes called laminopathies or nuclear envelopathies, that often affect specific tissues and organ system. Genetic manipulations in model organisms and experiments on cultured cells have begun to decipher how mutations in genes encoding broadly expressed nuclear envelope proteins cause diseases. This research has even identified potential treatments for these rare diseases that impact on human health.

1 Introduction to the Nuclear Envelope

The nuclear envelope is composed of the nuclear membranes, nuclear lamina and nuclear pore complexes (Fig. 1). Traditionally, the nuclear envelope has been considered a barrier separating the contents of the nucleus from those of the cytoplasm, with transport between these subcellular compartments in interphase occurring through the pore complexes. Additionally, the nuclear envelope has been a focus of cell biologists studying the cell cycle, as it disassembles at the start of mitosis and precisely reassembles in the daughter cells. More recently, however, the nuclear envelope has been inferred to have tissue-selective functions based initially on discoveries in human genetics linking mutations in genes encoding several of its widely expressed protein components to disease.

H.J. Worman (✉)
Department of Medicine and Department of Pathology and Cell Biology, College of Physicians and Surgeons, Columbia University, New York, NY 10032, USA
e-mail: hjw14@columbia.edu

D.P. Bazett-Jones, G. Dellaire (eds.), *The Functional Nucleus*,
DOI 10.1007/978-3-319-38882-3_1

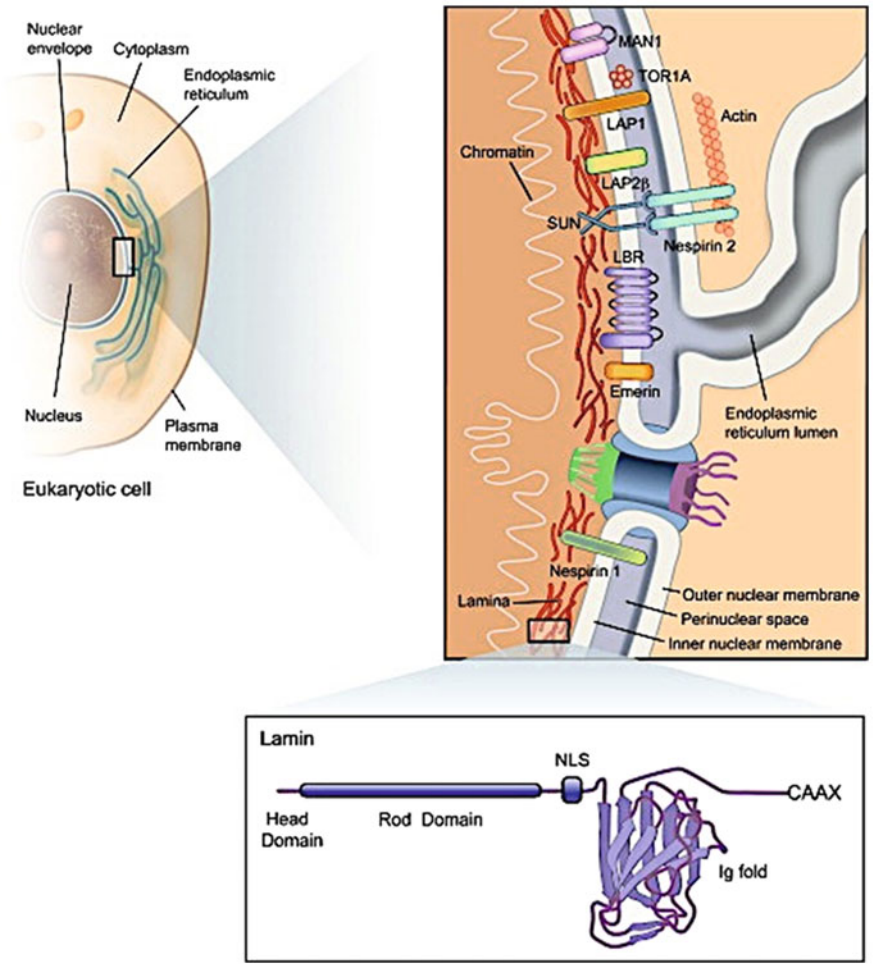

Fig. 1 The Nuclear Envelope. The nuclear envelope separates the contents of the nucleus from those of the cytoplasm and is composed the nuclear membranes, nuclear pore complexes, and nuclear lamina. The nuclear membranes are interconnected but divided into three morphologically distinct domains: the outer nuclear membrane, which is directly continuous with the rough endoplasmic reticulum, the inner nuclear membrane and the pore membranes, which connect the inner and outer membranes at the nuclear pore complexes (one pore complex is shown in this diagram). Integral proteins such as nesprin-2, a component of the Linker of Nucleoskeleton and Cytoskeleton (LINC) complex, preferentially concentrate in the outer nuclear membrane and bind to cytoskeletal filaments such as actin. The inner nuclear membrane is separated from the outer nuclear membrane by the perinuclear space, a continuation of the endoplasmic reticulum lumen, which may contain secreted proteins such as torsinA (TOR1A). Certain transmembrane proteins concentrate in the inner nuclear membrane in interphase cells. Many of these proteins bind to the nuclear lamina and chromatin. A few integral proteins that are components of nuclear pore complexes similarly concentrate within the pore membranes in interphase cells. Most of the integral proteins of the nuclear membranes are expressed to some degree in all somatic cells and tissues. The nuclear lamina is a meshwork of intermediate filaments on the inner aspect of the inner

The nuclear membranes are interconnected but divided into three morphologically distinct domains: outer, inner and pore. The outer nuclear membrane is directly continuous with the rough endoplasmic reticulum and they generally share integral proteins. It similarly contains ribosomes on its cytoplasmic surface. However, integral proteins called nesprins in mammalian cells, which are components of the Linker of Nucleoskeleton and Cytoskeleton (LINC) complex, preferentially concentrate in the outer nuclear membrane. The inner nuclear membrane is separated from the outer nuclear membrane by the perinuclear space, a continuation of the endoplasmic reticulum lumen. The inner and outer nuclear membranes are connected at the nuclear pore complexes by the pore membranes. Certain transmembrane proteins concentrate in the inner nuclear membrane in interphase cells. Many of these bind to the nuclear lamina and chromatin. A few integral proteins that are components of nuclear pore complexes similarly concentrate within the pore membranes in interphase cells. Most of the integral proteins of the nuclear membranes are expressed to some degree in most somatic cells and tissues.

Active and passive transport between the nucleus and cytoplasm in interphase occurs through the nuclear pore complexes. These are not simply "holes" through the nuclear envelope but macromolecular complexes composed of multiple copies of thirty or more proteins, most of which are called nucleoporins. The majority of nucleoporins appear to be expressed in all mammalian tissues.

Primarily localized at the inner aspect of the inner nuclear membrane is the nuclear lamina. The lamina is a meshwork of intermediate filament proteins called lamins. Like cytoplasmic intermediate filament proteins, lamins have alpha-helical rod domains flanked by variable head and tail domains. The tail domains of lamins contain a nuclear localization signal and an immunoglobulin-like fold. Most human lamin proteins have an cysteine-aliphatic-aliphatic-any amino acid (CAAX) motif at their carboxyl-terminus, which is a signal for a series of biochemical reactions leading to farnesylation and carboxymethylation of the cysteine. The lamin proteins exist as dimers that polymerize to form the higher-ordered 10 nm filaments of the nuclear lamina. Lamins have been reported to interact with myriad nuclear proteins, including integral proteins of the inner nuclear membrane, and with chromatin.

In humans, three genes encoding lamins: *LMNA*, *LMNB1* and *LMNB2*. *LMNA* encodes the A-type lamins, lamin A and lamin C (and a germ cell isoform lamin

Fig. 1 (continued) nuclear membrane composed of proteins called lamins. The lamina is associated with integral proteins of the inner nuclear membrane and representative examples MAN1, lamina-associated polypeptide-1 (LAP1), the SUN proteins, lamina-associated polypeptide-2β (LAP2 2β), lamin B receptor (LBR), emerin, and a nesprin-1 isoform are shown. A schematic of the structure of a lamin molecule is shown in the lower inset (not to scale). They have α-helical rod domains and head and tail domains that vary in sequence among members of intermediate filament protein family. Within the tail domain of lamins, there is a nuclear localization signal (NLS) and an immunoglobulin-like fold (Ig fold). Most lamins contain a CAAX motif that acts as a signal for farnesylation and carboxymethylation. Reprinted from *Developmental Cell*, Volume 17 /Edition 5, William T. Dauer and Howard J. Worman, The nuclear envelope as a signaling node in development and disease, Pages 626–638, Copyright 2009, with permission from Elsevier

C2). *LMNB1* encodes lamin B1 and *LMNB2* lamin B2 (and a germ cell isoform lamin B3). Lamins A and C are expressed in most, although not all, terminally differentiated cells and absent from pluripotent stem cells and early embryos. Lamin B1 and lamin B2 are expressed to some degree in virtually all somatic cells.

A relatively recently described structure of the nuclear envelope is the LINC complex, which connects the lamina to the cytoskeleton by forming a bridge across the inner and outer nuclear membrane. The LINC complex is evolutionarily conserved and composed of outer nuclear membrane KASH proteins and inner nuclear membrane SUN proteins. In mammals, the KASH proteins are most often referred to as nesprins and the SUN proteins as SUNs. The KASH and SUN domains of these proteins interact within the perinuclear space. SUNs interact with lamins and other nuclear proteins. Nesprins, via variable cytoplasmic domains, bind directly or indirectly to cytoskeletal filaments. The nucleocytoskeletal connections established by the LINC complex participate in various cellular processes, including nuclear positioning and mechanotransduction.

Mutations in genes encoding proteins of all these nuclear envelope components can cause diseases that are often tissue-specific despite the fact that most of the proteins are widely expressed. In some cases, different mutations in the same gene can cause different diseases. In others, mutations in genes encoding different nuclear envelope proteins case lead to the same phenotypes. As a group, these diseases have been referred to as laminopathies, particularly those resulting from mutations in *LMNA*, or nuclear envelopathies, a nomenclature that places focus on the cell biological origin of the disorders rather than the clinical features. We will review these inherited diseases grouped by the portion of the nuclear envelope in which the affected protein resides.

2 Nuclear Membranes

2.1 Inner Nuclear Membrane

X-linked Emery-Dreifuss muscular dystrophy was the first human disease shown to be caused by mutations in a gene encoding a nuclear envelope protein, an integral protein of the inner nuclear membrane. The disease is classically characterized by the triad of (1) early contractures of the elbows, Achilles tendons and postcervical muscles, (2) slowly progressive muscle wasting and weakness with a humeroperoneal distribution in the early stages and (3) cardiomyopathy usually first presenting as heart block (Emery 1989). In 1994, Toniolo and colleagues (Bione et al. 1994) used positional cloning to link X-linked Emery-Dreifuss muscular dystrophy to mutations in a previously uncharacterized gene now with the official symbol *EMD* encoding a protein they named emerin. The protein was predicted to have a single transmembrane domain and the gene was expressed in a wide range of human tissues. Two years after its discovery, emerin was localized

to the inner nuclear membrane (Nagano et al. 1996; Manilal et al. 1996). Most disease-causing *EMD* mutations lead to lack of emerin expression. While Emery-Dreifuss muscular dystrophy is the most common resulting phenotype, variations such as limb-girdle muscular dystrophy and cardiac conduction defects with minimal skeletal muscle disease can occur (Astejada et al. 2007). Emerin has been extensively studied and basic knowledge about its biochemistry, binding partners, localization, posttranslational modifications and roles in development are available (Berk et al. 2013). However, it is still not clear how loss of function of this widely expressed protein causes muscular dystrophy and cardiomyopathy. Emerin's protein binding partners include A-type lamins, genetic alterations in which also have been clearly linked to muscular dystrophy and cardiomyopathy (see below).

Since the discovery showing that mutations in the gene encoding emerin cause Emery-Dreifuss muscular dystrophy, mutations in genes of several other transmembrane proteins of the inner nuclear membrane have been linked to human diseases. Lamin B receptor or LBR is a polytopic integral inner nuclear membrane protein that binds to B-type lamins and chromatin proteins (Worman et al. 1988, 1990; Ye and Worman 1996). In addition, LBR contains a Δ(14)-sterol reductase domain (Holmer et al. 1998; Li et al. 2015). Heterozygous mutations in *LBR* generally cause Pelger-Huet anomaly, a benign condition of hyposegmented neutrophil nuclei (Hoffmann et al. 2002). In contrast, homozygous LBR mutations can cause hydrops-ectopic calcification-"moth-eaten" or Greenberg skeletal dysplasia, an in utero lethal disorder characterized by fetal hydrops, short limbs and abnormal chondro-osseous calcification (Waterham et al. 2003). This heterozygous versus homozygous relationship of LBR mutations correlating with benign and lethal conditions, respectively, is not strictly accurate. Homozygotes, heterozygotes and compound heterozygotes with different mutation can fall in a continuum ranging from isolated Pelger-Huet anomaly to Pelger-Huet with mild skeletal dysplasia to Greenberg skeletal dysplasia (Borovik et al. 2013). This phenotypic variability may be determined by how mutations differentially affect the structural functions—B-type lamin and chromatin protein binding—or sterol reductase activity of LBR (Clayton et al. 2010).

MAN1 was originally identified as an integral protein of the inner nuclear membrane recognized by autoantibodies from a patient with an ill-defined collagen vascular disease (Lin et al. 2000). It has two putative transmembrane segments—and two putative nucleoplasmic domains—as well as a LEM motif shared with emerin, LAP2 and some other nuclear proteins. The second nucleoplasmic domain of MAN1 binds to rSmads, transcription factors that mediate signaling by transforming growth factor-beta family members, and inhibits their gene regulatory activities (Raju et al. 2003; Hellemans et al. 2004; Lin et al. 2005; Pan et al. 2005). Structural biology studies indicate that this occurs by MAN1 competing with transcription factors for binding to Smad2 and Smad3 and facilitating their dephosphorylation by PPM1A, which also binds to MAN1 (Bourgeois et al. 2013). Heterozygous mutations in the *LEMD3* gene encoding MAN1 leading to loss of function cause osteopoikilosis, Buschke-Ollendorff syndrome and non-sporadic melorheostosis (Hellemans et al. 2004). These disorders are sclerosing bone

dysplasia characterized by heterogeneously increased bone density; Buschke-Ollendorff syndrome additionally affects the skin. The bone and skin phenotypes in patients with heterozygous loss of function of MAN1 are consistent with the know consequences of excessive transforming growth factor-beta signaling in these tissues. MAN1 has also been shown to regulate circadian rhythmicity in *Drosophila* and mice but abnormalities in this function have yet to be linked to human disease (Lin et al. 2014).

Mutations in genes encoding other integral proteins of the inner nuclear that bind to A-type lamins and emerin have been linked to muscular dystrophy and cardiomyopathy. So far, these linkages are based on isolated cases rather than large series of patients or informative families. LAP1 is a protein with at least three isoforms arising by alternative RNA splicing that binds to lamins (Senior and Gerace 1988; Foisner and Gerace 1993). LAP1 also interacts with emerin and its depletion from mouse skeletal and cardiac muscle leads to muscular dystrophy and cardiomyopathy (Shin et al. 2013, 2014). Mutations in the *TOR1AIP1* gene encoding LAP1 have been linked to muscular dystrophy and cardiomyopathy in one family and to cardiomyopathy and dystonia in an additional isolated case (Kayman-Kurekci et al. 2014; Dorboz et al. 2014). LAP2 encoded by the *TMPO* gene binds to lamins and has both transmembrane and nucleoplasmic isoforms arising by alternative RNA splicing (Foisner and Gerace 1993; Dechat et al. 2000). An amino acid substitution in the non-membrane alpha isoform has been reported in two brothers with dilated cardiomyopathy but segregation of the risk allele to only affected subjects was not demonstrated within the family (Taylor et al. 2005). *TMEM43* encodes an integral inner nuclear membrane protein called LUMA that interacts with emerin (Bengtsson and Otto 2008). *TMEM43* mutations have been reported in two subjects with Emery-Dreifuss muscular dystrophy-like phenotypes but segregation of the risk allele to only affected individuals within the families was not demonstrated (Liang et al. 2011). Amino acid substitutions in SUN1 and SUN2 have similarly been reported in subjects with Emery-Dreifuss muscular dystrophy-like phenotypes but again without data from families showing clear segregation of the mutations (Meinke et al. 2014).

The inner nuclear membrane proteome shows variability between tissues (Korfali et al. 2012), suggesting that certain proteins may have cell type-specific functions. This could potentially explain tissue-specific diseases resulting from mutations in their genes. However, most of the inner nuclear membrane proteins linked to disease so far are fairly widely expressed. Possible pathogenic mechanisms that can account for this include the role of selectively expressed binding partners involved in tissue-specific functions and functional redundancies of selectively expressed proteins. The nature of the tissue itself may also determine if loss of a protein's function has deleterious consequences. For example, if a protein plays a role in cellular stability or nuclear positioning, its loss of function may have greater consequences in striated muscle than in other tissues.

2.2 Outer Nuclear Membrane

Mutations in genes encoding nesprins have been linked to human disease. Nesprins are the mammalian KASH domain proteins of the LINC complex. Different nesprins bind to different cytoskeletal elements; for example, the high molecular mass nesprin-1G and nesprin-2G bind directly to actin whereas other nesprin-1 and nesprin-2 isoforms interact with microtubules indirectly via binding to kinesin or dynein (Chang et al. 2015b). As a result of these interactions with dynamic cytoskeletal elements, nesprins function in moving and positioning the nucleus in cells (Gundersen and Worman 2013). They further function in transmitting forces from the outside to the inside of the nucleus (Lombardi et al. 2011; Guilluy et al. 2014). Hence, defects in nuclear positioning or mechanotransduction may be involved in the pathogenesis of diseases caused by mutations in genes encoding nesprins.

Mutations in the *SYNE1* gene encoding nesprin-1 cause autosomal recessive cerebellar ataxia (Gros-Louis et al. 2007; Dupré et al. 2007). Nesprin-1 isoforms are highly expressed in Purkinje cells and it is possible that loss may disrupt nuclear positioning in these cells leading to death or dysfunction (Gros-Louis et al. 2007). In one large consanguineous family, recessive mutation in *SNYE1* has been linked to arthrogryposis multiplex congenita, a disorder characterized by decreased fetal movements, delay in motor milestones and progressive motor decline (Attali et al. 2009). While high molecular mass nesprin-1 isoforms are localized to the outer nuclear membrane via their interactions with SUNs and comprise the LINC complex, it is possible that isoforms affected by these disease-causing *SNYN1* mutations are smaller ones that can reach the inner nuclear membrane. Sequence variants in *SYNE1* as well as *SNYE2*, encoding nesprin-2 have additional been reported in patients with Emery-Dreifuss muscular dystrophy-like phenotypes and dilated cardiomyopathy; however, there have only been a few case reports without clear segregation of the *SYNE1* mutants to only affected family members (Zhang et al. 2007; Puckelwartz et al. 2010).

Homozygosity for a mutation in *SYNE4* leading to truncation of nesprin-4 has been described in two families of Iraqi Jewish ancestry with progressive high-frequency hearing loss (Horn et al. 2013). Nesprin-4 binds to kinesin and is expressed in selected tissues including salivary gland, exocrine pancreas, bulbourethral gland, mammary tissue and hair cells of the inner ear (Roux et al. 2009; Horn et al. 2013). As nuclei of the inner ear's outer hair cells are improperly positioned in mice lacking the protein, it has been hypothesized that nesprin-4-mediated nuclear positioning in sensory epithelial cells is critical for maintenance of normal hearing.

3 Perinuclear Space

Autosomal dominant DYT1 dystonia, an early-onset disorder characterized by progressive problems with movement, is associated with abnormal concentration of an endoplasmic reticulum protein in the perinuclear space of the nuclear envelope. DYT1 dystonia is caused by an in-frame deletion in the *TOR1A* gene encoding the AAA-ATPase torsinA (Ozelius et al. 1997). TorsinA in normally localized diffusely throughout the lumen of the endoplasmic reticulum but the disease-associated variant, which has deletion of a glutamic acid residue, concentrates in the perinuclear space (Goodchild and Dauer 2004; Gonzalez-Alegre and Paulson 2004; Naismith et al. 2004). TorsinA binds to and is activated by LAP1 and an endoplasmic reticulum protein LULL1; the disease-associated variant preferentially binds to the domain of inner nuclear membrane protein LAP1 that is localized to the lumen of the perinuclear space (Goodchild and Dauer 2005; Brown et al. 2014; Sosa et al. 2014). Mice with conditional deletion of torsinA from the central nervous system and mice with brain over-expression of the disease-causing variant both develop abnormal twisting movements, suggesting that loss of torsinA function underlies pathogenesis (Liang et al. 2014). Neurons from germ line knockout mice and homozygous knock-in mice expressing the disease-causing variant contain morphologically abnormal nuclear membranes, further suggesting that loss of function underlies pathogenesis; however, non-neuronal cell types appear to be unaffected (Goodchild et al. 2005). TorsinB, a homologous protein, is expressed at high levels in non-neuronal cells, likely providing protection (Kim et al. 2010).

4 Nuclear Pore Complex

Several inherited disorders have been linked to mutations in genes encoding nuclear pore complex proteins. An initial analysis of the protein composition of the nuclear pore complex from rat liver identified 29 nucleoporins and 18 associated proteins (Cronshaw et al. 2002). One of the proteins identified in this initial analysis was the WD-repeat protein ALADIN. The human gene encoding ALADIN was initially discovered because it is mutated in triple-A or Allgrove syndrome (Tullio-Pelet et al. 2000). Triple-A syndrome is inherited as an autosomal recessive disorder characterized by adrenocorticotropin hormone-resistant adrenal insufficiency, achalasia and alacrima (Allgrove et al. 1978). Disease-associated variants of ALLADIN fail to localize to nuclear pore complexes (Cronshaw and Matunis 2003). While ALADIN is widely expressed in different tissues, its association with triple-A syndrome was the first datum to suggest that an individual nucleoporin can have tissue-specific functions.

Mutations in two other genes encoding nucleoporins have been linked to disorders of the central nervous system. A missense mutation in the gene encoding

NUP62 has been shown to cause recessive infantile bilateral striatal necrosis in eight Israeli Bedouin families with 12 affected and 39 unaffected individuals (Basel-Vanagaite et al. 2006). Infantile bilateral striatal necrosis is characterized by symmetrical degeneration of the caudate nucleus, putamen and sometimes the globus pallidus. Dominantly inherited mutations in the gene encoding Ran binding protein 2, another nucleoporin, have been linked to susceptibility to infection-triggered acute necrotizing encephalopathy (Neilson et al. 2009; Singh et al. 2015). This disease is a rapidly progressive encephalopathy occurring after common viral infections such as influenza and parainfluenza.

A recessively inherited missense mutation in gene encoding nucleoporin NUP155 has been reported to segregate with atrial fibrillation in one family (Zhang et al. 2008). The association is supported by the fact that mice lacking one copy of the *Nup155* gene develop cardiac arrhythmias (Zhang et al. 2008). As is the case with ALADIN, the disease-associated NUP155 variant does not properly localize to nuclear pore complexes. NUP155 appears to be differentially expressed across different tissues, with relatively higher levels in heart and skeletal muscle (Zhang et al. 1999).

5 Nuclear Lamina

5.1 A-type Lamins

LMNA encodes the A-type lamins with lamin A and lamin C being the main isoforms isoforms expressed in most differentiated somatic cells (Lin and Worman 1993). The proteins are identical for the first 566 amino acids. Alternative splicing in the RNA encoding by exon 10 generates lamin C, with six unique carboxyl-terminal amino acids, and prelamin A, with 98 unique carboxyl-terminal amino acids. Prelamin A is a transiently expressed precursor protein that is processed to lamin A (Sinensky et al. 1994). Like lamin B1 and lamin B2, prelamin A has a CAAX motif at its carboxyl-terminus. The motif is a signal for the following series of reactions: (1) farnesylation of the cysteine catalyzed by protein farnesyltransferase, (2) proteolysis of the –AAX residues (serine-isoleucine-methionine in the case of prelamin A) catalyzed by RCE1 and, for prelamin A, ZMPSTE24 and (3) carboxylmethylation of the cysteine catalyzed by isoprenylcysteine carboxyl methyltransferase (Sinensky et al. 1994; Worman et al. 2009). The posttranslationally-modified prelamin A is then the substrate for a second endoproteolytic cleavage catalyzed by the zinc metallopeptidase ZMPSTE24 that that removes the carboxyl-terminal 15 amino acids, including the modified cysteine, to generate mature lamin A (Fig. 2).

Mutations *LMNA* were first shown to cause autosomal dominant Emery-Dreifuss muscular dystrophy (Bonne et al. 1999). Since then, mutations in the gene have been linked to over a dozen diseases that have been classified as distinct clinical entities. These can be broadly grouped into disorders primarily affecting striated

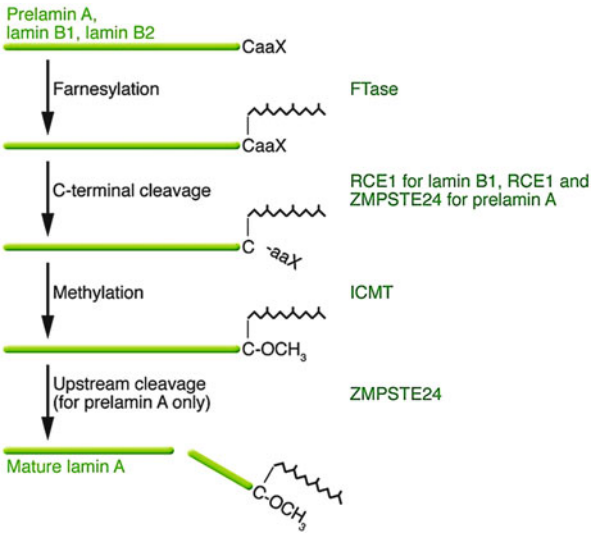

Fig. 2 Processing of Prelamin A. Prelamin A, like lamin B1 and lamin B2, has a CAAX motif at its carboxyl-terminus. The CAAX motif is a signal for the following series of reactions: (1) farnesylation of the cysteine catalyzed by protein farnesyltransferase (FTase), (2) proteolysis of the –AAX residues catalyzed by RCE1 for lamin B1 and RCE1 and ZMPSTE24 for prelamin A and (3) carboxylmethylation of the cysteine catalyzed by isoprenylcysteine carboxyl methyltransferase (ICMT). The modified prelamin A is then the substrate for a second endoproteolytic cleavage catalyzed by ZMPSTE24 that that removes the carboxyl-terminal 15 amino acids, including the modified cysteine, to generate mature lamin A. Republished with permission of the American Society for Clinical Investigation, from *Journal of Clinical Investigation*, Howard J. Worman, Loren G. Fong, Antoine Muchir and Stephen G Young, Laminopathies and the long strange trip from basic cell biology to therapy, Volume 119, Edition 7, 2009; permission conveyed through Copyright Clearance Center, Inc.

muscle, adipose tissue, peripheral nerve or involving multiple organ systems (Dauer and Worman 2009) (Fig. 3).

5.1.1 Striated Muscle

Autosomal dominant Emery-Dreifuss muscular dystrophy is clinically similar to the X-linked form of the disease caused by mutations in the gene encoding emerin (described above). Soon after mutations in *LMNA* were shown to cause autosomal dominant Emery-Dreifuss muscular dystrophy, they were reported to cause dilated cardiomyopathy and conduction-system disease in the absence of significant skeletal myopathy (Fatkin et al. 1999). *LMNA* mutations were then reported to cause limb-girdle muscular dystrophy with dilated cardiomyopathy (Muchir et al. 2000). Subsequently, it was reported that a single point mutation in *LMNA* in the same family could result in a phenotype of dilated cardiomyopathy with minimal to no skeletal myopathy, Emery-Dreifuss-like muscle involvement or limb girdle-like muscle involvement (Brodsky et al. 2000). It is now clear that *LMNA* mutations

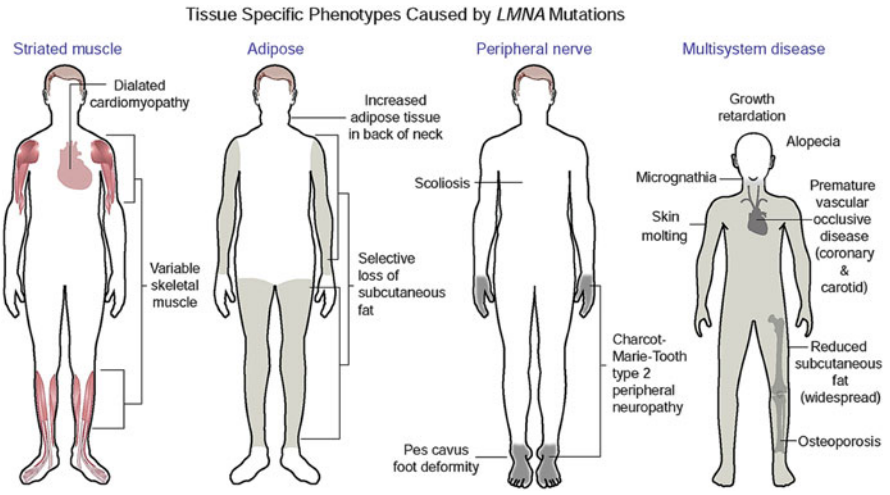

Fig. 3 *LMNA* Mutations Cause Tissue Specific Phenotypes Primarily Involving Either Striated Muscle, Adipose Tissue, Peripheral Nerve or Multiple Organ Systems. Most autosomal dominant mutations in *LMNA* cause dilated cardiomyopathy with variable skeletal muscle involvement. The diagram shows the classical Emery-Dreifuss muscular dystrophy phenotype with a scapulohumeral-peroneal distribution of skeletal muscle involvement and tendon contractures. The same mutations can result in cardiomyopathy with different types of skeletal muscle involvement (see Fig. 4). Other autosomal dominant missense mutations, mostly those in exon 8 leading to a change in the surface charge of the immunoglobulin fold in the tail domain, cause Dunnigan-type partial lipodystrophy, characterized by loss of subcutaneous fat from the extremities, excessive fat accumulation in the neck and face and development of insulin resistance and diabetes mellitus. The autosomal recessive R298C *LMNA* mutation causes a Charcot-Marie-Tooth type 2 peripheral neuropathy, with phenotypic variability but most often characterized by a stocking-glove sensory neuropathy, an associated pes cavus foot deformity and additional features such as scoliosis. Multisystem diseases caused *LMNA* mutations include Hutchinson-Gilford progeria syndrome with signs including growth retardation, micrognathia, reduced subcutaneous fat, alopecia, osteoporosis, skin mottling and early-onset vascular occlusive disease. Other *LMNA* mutations can cause variant progeriod syndromes with some of the same features or mandibuloacral dysplasia, with a combination of progeroid features and partial lipodystrophy. Reprinted from *Developmental Cell*, Volume 17 /Edition 5, William T. Dauer and Howard J. Worman, The nuclear envelope as a signaling node in development and disease, Pages 626–638, Copyright 2009, with permission from Elsevier

cause a spectrum of striated muscle diseases, including a comparatively severe congenital muscular dystrophy, with dilated cardiomyopathy and conduction system abnormalities as the most common feature (Lu et al. 2011) (Fig. 4).

LMNA mutations that cause striated muscle generally lead to single amino acid substitutions, short in-frame deletions or splicing alterations throughout the molecules or truncations that lead to haploinsufficiency of the proteins. Mice with depletion of A-type lamins develop cardiomyopathy and muscular dystrophy (Sullivan et al. 1999). This suggests that loss of A-type lamin function causes striated muscle disease. However, this may occur as a "dominant-negative-type" effect. For example, certain lamin A variants that are expressed in patients with the disease

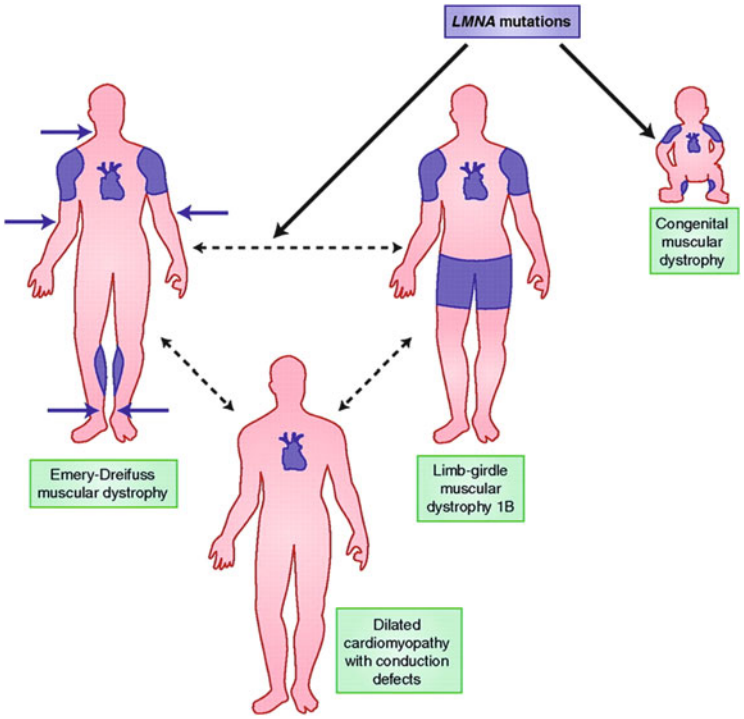

Fig. 4 Spectrum of Striated Muscle Diseases Caused by *LMNA* Mutations. Emery-Dreifuss muscular dystrophy, limb-girdle muscular dystrophy 1B and isolated cardiomyopathy are clinically described disorders presenting in childhood or adulthood caused by *LMNA* mutations. The affected skeletal muscle groups affected in these disorders are shaded and arrows indicate the location of contractures that are characteristic of Emery-Dreifuss muscular dystrophy. These specific diseases are actually a spectrum of the same disorder, cardiomyopathy with variable skeletal muscle involvement, which can have overlapping phenotypes (indicated by dashed arrows) and be caused by the same *LMNA* mutations. Some *LMNA* mutations cause congenital muscular dystrophy with cardiomyopathy, which presents in infants or very young children. Figure from Lu et al. (2011)

cause structural alterations in the nuclear lamina that may be detrimental for processes such as mechanotransduction or cell signaling (Worman et al. 2009).

While it remains unclear exactly how alterations in A-type lamins cause striated muscle disease, experimental findings suggest that abnormalities in cellular stress responses and stress-related signaling pathways play a role. Fibroblasts lacking A-type lamins have defective mechanical properties and strain-induced signaling when subjected to stress (Lammerding et al. 2004). Disease-causing alterations in A-type lamins also lead to impaired signaling by the mechanosensitive transcription factor megakaryoblastic leukemia 1 secondary to altered dynamics of cytoskeletal actin (Ho et al. 2013). Cells lacking A-type lamins or expressing disease-associated lamin A variants, as well as striated muscle from knock in mice, have abnormal hyperactivation of stress-induced MAP kinases (Muchir et al. 2007, 2009).

Blocking their activities has beneficial effects on heart and skeletal muscle function in knock in mice (Wu et al. 2011; Muchir et al. 2012, 2013). In addition, hyperactivation of AKT-mTOR signaling, associated with defective autophagy, occurs in mice lacking A-type lamins and knock in mice expressing a disease-causing variant and inhibitors of mTOR have beneficial effects (Choi et al. 2012; Ramos et al. 2012). In migrating fibroblasts and myoblasts, striated muscle disease-associated lamin A defects also block normal actin-dependent nuclear positioning, a process that may be essential in regenerating skeletal muscle (Folker et al. 2011; Chang et al. 2015a).

5.1.2 Lipodystrophy

Mutations primarily concentrated in exon 8 of *LMNA* cause Dunnigan-type familial partial lipodystrophy (Cao and Hegele 2000; Shackleton et al. 2000; Speckman et al. 2000). This autosomal dominant disorder is characterized by loss of fat from the extremities around the time of puberty with a subsequent increase in central adiposity, the development of insulin resistance and in some cases skeletal muscle hypertrophy. Most subjects then develop diabetes mellitus and hepatic steatosis. Most mutations causing Dunnigan-type familial partial lipodystrophy lead to amino acid substitutions of arginine residue at positions 482 in lamin A and lamin C. These lead to a change in the surface charge of the immunoglobulin-like fold in the tail domains of these proteins (Dhe-Paganon et al. 2002; Krimm et al. 2002). In contrast, amino acid substitutions in the immugloblulin-like fold associated with striated muscle disease are predicted to cause a more dramatic disruption protein structure. Mice lacking A-type lamins have no evidence of lipodystrophy (Cutler et al. 2002). However, transgenic mice overexpressing a disease-associated lamin A variant develop signs of lipodystrophy (Wojtanik et al. 2009). Overexpression of disease-associated variants also decreases the ability of cultured fibroblasts to differentiate into adipocytes (Boguslavsky et al. 2006). Hence, the Dunnigan-type partial lipodystrophy-causing amino acid substitutions likely alter a particular function of lamin A and lamin C, perhaps specific to peripheral adipocytes. In addition to the classical Dunnigan-type phenotype, mutations outside of the portion of the gene encoding the immunoglobulin-like fold have been associated with variant lipodystrophy phenotypes (Caux et al. 2003; Decaudain et al. 2007; Mory et al. 2012).

5.1.3 Peripheral Neuropathy

A homozygous arginine to cysteine substation at amino acid residue 298 in lamin A and C cause a Charcot-Marie-Tooth type 2 peripheral neuropathy (De Sandre-Giovannoli et al. 2002). This peripheral neuropathy-causing mutation has been described in several unrelated Algerian families with variability in phenotype, age of onset and severity (Tazir et al. 2004). Sciatic nerves of mice without A-type

lamins have a reduction of axon density and nonmyelinated axons, findings similar to humans with Charcot-Marie-Tooth type 2 disease (De Sandre-Giovannoli et al. 2002). This suggests that loss the disease-causing mutations cause loss of a particular activity of A-type lamins necessary for proper peripheral nerve structure and function.

6 Multisystem Disorders

In addition to the tissue-selective diseases, disorders affecting multiple organ systems are caused by some *LMNA* mutations. These diseases have features of progeria, which means that they are characterized by physical signs and symptoms suggestive of premature or accelerated aging. However, it must be cautioned that these disorders are not rigorous mimics of physiological aging. They do not have all of the key features of physiological aging such as dementia, other forms of neurodegeneration or atherosclerosis correlating with high serum LDL cholesterol concentrations.

Hutchinson-Gilford progeria syndrome is a rare disorder characterized by short stature, low body mass, alopecia, sclerotic skin, joint contractures, osteolysis and facial features that partially resemble those of aged persons (DeBusk 1972; Merideth et al. 2008). Affected children generally die in the second decade of life from ischemic heart disease and stokes that occur secondary to vascular occlusions. However, blood total cholesterol, LDL and HDL cholesterol and triglyceride concentrations are similar to those in control children and aspects of blood vessel wall pathology are different than those in atherosclerosis associated with physiological aging (Stehbens et al. 1999; Gordon et al. 2005; Olive et al. 2010).

Hutchinson-Gilford progeria syndrome is caused by de novo point mutations in exon 11 of *LMNA*. The mutations optimize an alternative splice donor site that results in an in-frame deletion of 50 amino acids near the carboxyl-terminus of prelamin A (Eriksson et al. 2003; de Sandre-Giovannoli et al. 2003). The truncated prelamin A variant, which has been called progerin, undergoes farnesylation and carboxyl methylation like prelamin A but the second site for ZMPSTE24-catalyzed cleavage is eliminated, preventing further processing to lamin A. The farnesylated progerin accumulates in cells and induces alterations in nuclear envelope morphology (Goldman et al. 2004). However, similar alterations in nuclear envelope morphology are also seen with expression of many other disease-associated lamin A variants in cultured cells, making it unclear how they specifically relate to defects that occur in Hutchinson-Gilford progeria syndrome.

Considerable experimental evidence supports the hypothesis that accumulation of farnesylated progerin is the "upstream" pathogenic defect in Hutchinson-Gilford progeria syndrome. Deletion of *Zmpste24* generates a progeroid phenotype in mice (Bergo et al. 2002; Pendás et al. 2002). Similarly, homozygous and compound heterozygous loss of function of *ZMPSTE24* in humans leading to accumulation of farnesylated prelamin A causes restrictive dermopathy, a neonatal lethal progeroid

syndrome (Moulson et al. 2005; Navarro et al. 2005). The "toxic" effects of unprocessed prelamin A was first demonstrated by crossing *Zmpste24* null mice to mice heterozygous for A-type lamin deficiency, which partially ameliorated the progeroid phenotype (Fong et al. 2004). Subsequently, treatment of *Zmpste24* null mice with a pharmacological inhibitor of protein farnesyltransferase was shown to ameliorate the progeroid phenotype (Fong et al. 2006). Treatment with a protein farnesyltransferase inhibitor was subsequently shown to improve the disease phenotype in mice with a targeted Hutchinson-Gilford progeria syndrome mutation and later in a BAC transgenic mouse model that expresses progerin (Yang et al. 2006; Capell et al. 2008). Protein farnesyltransferase inhibitors further reverse the abnormal nuclear morphology in cells expressing progerin or lacking ZMPSTE24 activity (Yang et al. 2005; Toth et al. 2005; Capell et al. 2005; Mallampalli et al. 2005; Glynn and Glover 2005). These findings have lead to a clinical trial of a protein farnesyltransferase inhibitor in children with Hutchinson-Gilford progeria syndrome, with reported improvement in some clinical parameters (Gordon et al. 2012). While somewhat encouraging, because of the uncontrolled, un-blinded design of the trial and the absence of a correlation between inhibition of protein farnesylation and clinical response, it is difficult to confidently interpret the results (Young et al. 2013).

While accumulation of farnesylated progerin in Hutchinson-Gilford progeria syndrome and farnesylated prelamin A in restrictive dermopathy are likely upstream defects in these diseases, the resulting "downstream" pathogenic abnormalities are not clearly understood. Myriad experimental findings have been reported but their precise role in pathogenesis are unclear. Cells expressing progerin have abnormal biomechanical properties including decreased viability and increased apoptosis under repetitive mechanical strain (Verstraeten et al. 2008). Progerin also appears to interfere with normal tissue stem cell function. It activates downstream effectors of notch signaling and alters the differentiation potential of mesenchymal stem cells (Scaffidi and Misteli 2008). Deficiency of ZMPSTE24 also alters the proliferative capacity of tissue stem cells, with concurrent alterations in various cellular signaling pathways including Wnt (Espada et al. 2008). Expression of another farnesylated prelamin A variant in mice affects Wnt signaling and extracellular matrix abnormalities (Hernandez et al. 2010). Dysfunction of sirtuin 1, which binds to lamin A, may also be involved in the adult stem cell decline that occurs with progerin or prelamin A expression (Liu et al. 2012). Unprocessed prelamin A and progerin further alter DNA damage responses and repair, resulting in genomic instability (Liu et al. 2005, 2006). Treatment of cells expression progerin with rapamycin, an mTOR inhibitor, also improves nuclear morphology, delays the onset of cellular senescence and enhances the autophagic degradation of progerin (Cao et al. 2011). Sulforaphane, an antioxidant derived from cruciferous vegetables similarly enhances progerin clearance by autophagy and reverses cellular phenotypic changes (Gabriel et al. 2015).

While expression farnesylated prelamin A and progerin are certainly involved in the pathogenesis of Hutchinson-Gilford progeria syndrome and restrictive dermopathy, atypical progeroid syndromes can result form mutations in *LMNA*

that do not apparently lead to accumulanytion of a farnesylated prelamin A variants (Chen et al. 2003; Csoka et al. 2004). In mice, progeroid phenotypes similarly occur with expression of at least one non-farnesylated prelamin A variant; however, do not occur with expression of similar non-farnesylated variant (Yang et al. 2008, 2011). Mandibuloacral dysplasia, an autosomal recessive disease with features of progeria as well as lipodystrophy, is also caused by an amino acid substitution in lamins A and C (Novelli et al. 2002). A similar mandibuloacral dysplasia phenotype occurs with homozygous or compound heterozygous mutations in *ZMPSTE24* that lead to only partial loss of function with residual enzyme activity (Agarwal et al. 2003; Ben Yaou et al. 2011; Barrowman et al. 2012).

6.1 B-type Lamins

Duplication of the *LMNB1* gene with increased expression of lamin B1 in brain causes adult-onset autosomal dominant leukodystrophy (Padiath et al. 2006). This is a slowly progressive neurological disorder characterized widespread myelin loss in the central nervous system. Transgenic mice with either generalized or oligodendrocyte-selective overexpression of lamin B1 develop similar phenotypes to the human disease (Heng et al. 2013). These mice have decreased expression of proteolipid protein, which plays a critical function in myelination, a decreased binding of the Yin Yang 1 transcriptional activator to its gene. Mice with deletion of either *Lmnb1* or *Lmnb2*, respectively encoding lamin B1 and lamin B2, have neurodevelopmental abnormalities with defective neuronal migration in the mice lacking lamin B1 (Coffinier et al. 2010, 2011). However, neurodevelopmental human diseases have not been linked to the orthologous human genes to date.

7 Conclusions

In few other instances have human genetics and cell biology come together than in understanding human diseases related to nuclear envelope proteins. While tremendous progress has been made in their genetics, how cell biological defects resulting from inherited alterations in these proteins contribute to pathology is only more slowly emerging. Research in cultured cells and vertebrate model organisms has provided several insights into pathogenic defects, some of which can be targeted by drugs that have been shown to be beneficial at least in model organisms. Further research is needed to understand this fascinating group of human diseases that affect the nuclear envelope and to translate the discoveries made in the laboratory to the patient.

Acknowledgements The author is currently supported by grants from the United States National Institutes of Health (AR048997, NS059352, HD070713), the Muscular Dystrophy Association (MDA 294537) and Los Angeles Thoracic and Cardiovascular Foundation (CRV 2011-873R1).

References

Agarwal AK, Fry's JP, Auchus R et al (2003) Zinc metalloproteinase, ZMPSTE24, is mutated in mandibuloacral dysplasia. Hum Mol Genet 12:1995–2001

Allgrove J, Clayden GS, Grant DB et al (1978) Familial glucocorticoid deficiency with achalasia of the cardia and deficient tear production. Lancet 1:1284–1286

Astejada MN, Goto K, Nagano A et al (2007) Emerinopathy and laminopathy clinical, pathological and molecular features of muscular dystrophy with nuclear envelopathy in Japan. Acta Myol 26:159–164

Attali R, Warwar N, Israel A et al (2009) Mutation of SYNE-1, encoding an essential component of the nuclear lamina, is responsible for autosomal recessive arthrogryposis. Hum Mol Genet 18:3462–3469

Barrowman J, Wiley PA, Hudon-Miller SE et al (2012) Human ZMPSTE24 disease mutations: residual proteolytic activity correlates with disease severity. Hum Mol Genet 21:4084–4093

Basel-Vanagaite L, Muncher L, Straussberg R et al (2006) Mutated nup62 causes autosomal recessive infantile bilateral striatal necrosis. Ann Neurol 60:214–222

Bengtsson L, Otto H (2008) LUMA interacts with emerin and influences its distribution at the inner nuclear membrane. J Cell Sci 121:536–548

Ben Yaou R, Navarro C, Quijano-Roy S et al (2011) Type B mandibuloacral dysplasia with congenital myopathy due to homozygous ZMPSTE24 missense mutation. Eur J Hum Genet 19:647–654

Bergo MO, Gavino B, Ross J et al (2002) Zmpste24 deficiency in mice causes spontaneous bone fractures, muscle weakness, and a prelamin A processing defect. Proc Natl Acad Sci U S A 99:13049–13054

Berk JM, Tifft KE, Wilson KL (2013) The nuclear envelope LEM-domain protein emerin. Nucleus 4:298–314

Bione S, Maestrini E, Rivella S et al (1994) Identification of a novel X-linked gene responsible for Emery-Dreifuss muscular dystrophy. Nat Genet 8:323–327

Boguslavsky RL, Stewart CL, Worman HJ (2006) Nuclear lamin A inhibits adipocyte differentiation: implications for Dunnigan-type familial partial lipodystrophy. Hum Mol Genet 15:653–663

Bonne G, Di Barletta MR, Varnous S et al (1999) Mutations in the gene encoding lamin A/C cause autosomal dominant Emery-Dreifuss muscular dystrophy. Nat Genet 21:285–288

Borovik L, Modaff P, Waterham HR et al (2013) Pelger-huet anomaly and a mild skeletal phenotype secondary to mutations in LBR. Am J Med Genet A 161A:2066–2073

Bourgeois B, Gilquin B, Tellier-Lebègue C et al (2013) Inhibition of TGF-β signaling at the nuclear envelope: characterization of interactions between MAN1, Smad2 and Smad3, and PPM1A. Sci Signal 6:ra49

Brodsky GL, Muntoni F, Miocic S et al (2000) Lamin A/C gene mutation associated with dilated cardiomyopathy with variable skeletal muscle involvement. Circulation 101:473–476

Brown RS, Zhao C, Chase AR et al (2014) The mechanism of Torsin ATPase activation. Proc Natl Acad Sci U S A 111:E4822–E4831

Cao H, Hegele RA (2000) Nuclear lamin A/C R482Q mutation in Canadian kindreds with Dunnigan-type familial partial lipodystrophy. Hum Mol Genet 9:109–112

Cao K, Graziotto JJ, Blair CD et al (2011) Rapamycin reverses cellular phenotypes and enhances mutant protein clearance in Hutchinson-Gilford progeria syndrome cells. Sci Transl Med 3:89ra58

Capell BC, Erdos MR, Madigan JP et al (2005) Inhibiting farnesylation of progerin prevents the characteristic nuclear blebbing of Hutchinson-Gilford progeria syndrome. Proc Natl Acad Sci U S A 102:12879–12884

Capell BC, Olive M, Erdos MR et al (2008) A farnesyltransferase inhibitor prevents both the onset and late progression of cardiovascular disease in a progeria mouse model. Proc Natl Acad Sci U S A 105:15902–15907

Caux F, Dubosclard E, Lascols O et al (2003) A new clinical condition linked to a novel mutation in lamins A and C with generalized lipoatrophy, insulin-resistant diabetes, disseminated leukomelanodermic papules, liver steatosis, and cardiomyopathy. J Clin Endocrinol Metab 88:1006–1013

Chang W, Antoku S, Östlund C et al (2015a) Linker of nucleoskeleton and cytoskeleton (LINC) complex-mediated actin-dependent nuclear positioning orients centrosomes in migrating myoblasts. Nucleus 6:77–88

Chang W, Worman HJ, Gundersen GG (2015b) Accessorizing and anchoring the LINC complex for multifunctionality. J Cell Biol 208:11–22

Chen L, Lee L, Kudlow BA et al (2003) LMNA mutations in atypical Werner's syndrome. Lancet 362:440–445

Choi JC, Muchir A, Wu W et al (2012) Temsirolimus activates autophagy and ameliorates cardiomyopathy caused by lamin A/C gene mutation. Sci Transl Med 4:144ra102

Clayton P, Fischer B, Mann A et al (2010) Mutations causing Greenberg dysplasia but not Pelger anomaly uncouple enzymatic from structural functions of a nuclear membrane protein. Nucleus 1:354–366

Coffinier C, Chang SY, Nobumori C et al (2010) Abnormal development of the cerebral cortex and cerebellum in the setting of lamin B2 deficiency. Proc Natl Acad Sci U S A 107:5076–5081

Coffinier C, Jung HJ, Nobumori C et al (2011) Deficiencies in lamin B1 and lamin B2 cause neurodevelopmental defects and distinct nuclear shape abnormalities in neurons. Mol Biol Cell 22:4683–4693

Cronshaw JM, Krutchinsky AN, Zhang W et al (2002) Proteomic analysis of the mammalian nuclear pore complex. J Cell Biol 158:915–927

Cronshaw JM, Matunis MJ (2003) The nuclear pore complex protein ALADIN is mislocalized in triple A syndrome. Proc Natl Acad Sci U S A 100:5823–5827

Csoka AB, Cao H, Sammak PJ et al (2004) Novel lamin A/C gene (LMNA) mutations in atypical progeroid syndromes. J Med Genet 41:304–308

Cutler DA, Sullivan T, Marcus-Samuels B et al (2002) Characterization of adiposity and metabolism in Lmna-deficient mice. Biochem Biophys Res Commun 291:522–527

Dauer WT, Worman HJ (2009) The nuclear envelope as a signaling node in development and disease. Dev Cell 17:626–638

DeBusk FL (1972) The Hutchinson-Gilford progeria syndrome. Report of 4 cases and review of the literature. J Pediatr 80:697–724

Decaudain A, Vantyghem MC, Guerci B et al (2007) New metabolic phenotypes in laminopathies: LMNA mutations in patients with severe metabolic syndrome. J Clin Endocrinol Metab 92:4835–4844

Dechat T, Korbei B, Vaughan OA et al (2000) Lamina-associated polypeptide 2alpha binds intranuclear A-type lamins. J Cell Sci 113:3473–3484

De Sandre-Giovannoli A, Chaouch M, Kozlov S et al (2002) Homozygous defects in LMNA, encoding lamin A/C nuclear-envelope proteins, cause autosomal recessive axonal neuropathy in human (Charcot-Marie-Tooth disorder type 2) and mouse. Am J Hum Genet 70:726–736

De Sandre-Giovannoli A, Bernard R, Cau P et al (2003) Lamin A truncation in Hutchinson-Gilford progeria. Science 300:2055

Dhe-Paganon S, Werner ED, Chi YI et al (2002) Structure of the globular tail of nuclear lamin. J Biol Chem 277:17381–17384
Dorboz I, Coutelier M, Bertrand AT et al (2014) Severe dystonia, cerebellar atrophy, and cardiomyopathy likely caused by a missense mutation in TOR1AIP1. Orphanet J Rare Dis 9:174
Dupré N, Gros-Louis F, Chrestian N (2007) Clinical and genetic study of autosomal recessive cerebellar ataxia type 1. Ann Neurol 62:93–98
Emery AE (1989) Emery-Dreifuss syndrome. J Med Genet 26:637–641
Eriksson M, Brown WT, Gordon LB et al (2003) Recurrent de novo point mutations in lamin A cause Hutchinson-Gilford progeria syndrome. Nature 423:293–298
Espada J, Varela I, Flores I et al (2008) Nuclear envelope defects cause stem cell dysfunction in premature-aging mice. J Cell Biol 181:27–35
Fatkin D, MacRae C, Sasaki T et al (1999) Missense mutations in the rod domain of the lamin A/C gene as causes of dilated cardiomyopathy and conduction-system disease. N Engl J Med 341:1715–1724
Foisner R, Gerace L (1993) Integral membrane proteins of the nuclear envelope interact with lamins and chromosomes, and binding is modulated by mitotic phosphorylation. Cell 73:1267–1279
Folker ES, Östlund C, Luxton GW et al (2011) Lamin A variants that cause striated muscle disease are defective in anchoring transmembrane actin-associated nuclear lines for nuclear movement. Proc Natl Acad Sci U S A 108:131–136
Fong LG, Frost D, Meta M et al (2006) A protein farnesyltransferase inhibitor ameliorates disease in a mouse model of progeria. Science 311:1621–1623
Fong LG, Ng JK, Meta M et al (2004) Heterozygosity for Lmna deficiency eliminates the progeria-like phenotypes in Zmpste24-deficient mice. Proc Natl Acad Sci U S A 101:18111–18116
Gabriel D, Roedl D, Gordon LB et al (2015) Sulforaphane enhances progerin clearance in Hutchinson-Gilford progeria fibroblasts. Aging Cell 14:78–91
Glynn MW, Glover TW (2005) Incomplete processing of mutant lamin A in Hutchinson-Gilford progeria leads to nuclear abnormalities, which are reversed by farnesyltransferase inhibition. Hum Mol Genet 14:2959–2969
Gonzalez-Alegre P, Paulson HL (2004) Aberrant cellular behavior of mutant torsinA implicates nuclear envelope dysfunction in DYT1 dystonia. J Neurosci 24:2593–2601
Goodchild RE, Dauer WT (2004) Mislocalization to the nuclear envelope: an effect of the dystonia-causing torsinA mutation. Proc Natl Acad Sci U S A 101:847–852
Goodchild RE, Dauer WT (2005) The AAA+ protein torsinA interacts with a conserved domain present in LAP1 and a novel ER protein. J Cell Biol 168:855–862
Goodchild RE, Kim CE, Dauer WT (2005) Loss of the dystonia-associated protein torsinA selectively disrupts the neuronal nuclear envelope. Neuron 48:923–932
Goldman RD, Shumaker DK, Erdos MR et al (2004) Accumulation of mutant lamin A causes progressive changes in nuclear architecture in Hutchinson-Gilford progeria syndrome. Proc Natl Acad Sci U S A 101:8963–8968
Gordon LB, Harten IA, Patti ME et al (2005) Reduced adiponectin and HDL cholesterol without elevated C-reactive protein: clues to the biology of premature atherosclerosis in Hutchinson-Gilford Progeria Syndrome. J Pediatr 146:336–341
Gordon LB, Kleinman ME, Miller DT et al (2012) Clinical trial of a farnesyltransferase inhibitor in children with Hutchinson-Gilford progeria syndrome. Proc Natl Acad Sci U S A 109:16666–16671
Gros-Louis F, Dupré N, Dion P et al (2007) Mutations in SYNE1 lead to a newly discovered form of autosomal recessive cerebellar ataxia. Nat Genet 39:80–85
Guilluy C, Osborne LD, Van Landeghem L et al (2014) Isolated nuclei adapt to force and reveal a mechanotransduction pathway in the nucleus. Nat Cell Biol 16:376–381
Gundersen GG, Worman HJ (2013) Nuclear positioning. Cell 152:1376–1389

Hellemans J, Preobrazhenska O, Willaert A et al (2004) Loss-of-function mutations in LEMD3 result in osteopoikilosis, Buschke-Ollendorff syndrome and melorheostosis. Nat Genet 36:1213–1218

Heng MY, Lin ST, Verret L et al (2013) Lamin B1 mediates cell-autonomous neuropathology in a leukodystrophy mouse model. J Clin Invest 123:2719–2729

Hernandez L, Roux KJ, Wong ES et al (2010) Functional coupling between the extracellular matrix and nuclear lamina by Wnt signaling in progeria. Dev Cell 19:413–425

Ho CY, Jaalouk DE, Vartiainen MK et al (2013) Lamin A/C and emerin regulate MKL1-SRF activity by modulating actin dynamics. Nature 497:507–511

Hoffmann K, Dreger CK, Olins AL et al (2002) Mutations in the gene encoding the lamin B receptor produce an altered nuclear morphology in granulocytes (Pelger-Huët anomaly). Nat Genet 31:410–414

Holmer L, Pezhman A, Worman HJ (1998) The human lamin B receptor/sterol reductase multigene family. Genomics 54:469–476

Horn HF, Brownstein Z, Lenz DR et al (2013) The LINC complex is essential for hearing. J Clin Invest 123:740–750

Kayman-Kurekci G, Talim B, Korkusuz P et al (2014) Mutation in TOR1AIP1 encoding LAP1B in a form of muscular dystrophy: a novel gene related to nuclear envelopathies. Neuromuscul Disord 24:624–633

Kim CE, Perez A, Perkins G et al (2010) A molecular mechanism underlying the neural-specific defect in torsinA mutant mice. Proc Natl Acad Sci U S A 107:9861–9866

Korfali N, Wilkie GS, Swanson SK et al (2012) The nuclear envelope proteome differs notably between tissues. Nucleus 3:552–564

Krimm I, Östlund C, Gilquin B et al (2002) The Ig-like structure of the C-terminal domain of lamin A/C, mutated in muscular dystrophies, cardiomyopathy, and partial lipodystrophy. Structure 10:811–823

Lammerding J, Schulze PC, Takahashi T et al (2004) Lamin A/C deficiency causes defective nuclear mechanics and mechanotransduction. J Clin Invest 113:370–378

Li X, Roberti R, Blobel G (2015) Structure of an integral membrane sterol reductase from *Methylomicrobium alcaliphilum*. Nature 517:104–107

Liang WC, Mitsuhashi H, Keduka E et al (2011) TMEM43 mutations in Emery-Dreifuss muscular dystrophy-related myopathy. Ann Neurol 69:1005–1013

Lin F, Blake DL, Callebaut I et al (2000) MAN1, an inner nuclear membrane protein that shares the LEM domain with lamina-associated polypeptide 2 and emerin. J Biol Chem 275:4840–4847

Lin F, Morrison JM, Wu W et al (2005) MAN1, an integral protein of the inner nuclear membrane, binds Smad2 and Smad3 and antagonizes transforming growth factor-beta signaling. Hum Mol Genet 14:437–445

Lin F, Worman HJ (1993) Structural organization of the human gene encoding nuclear lamin A and nuclear lamin C. Biol Chem 268:16321–16326

Lin ST, Zhang L, Lin X et al (2014) Nuclear envelope protein MAN1 regulates clock through BMAL1. Elife 3, e02981

Liu B, Ghosh S, Yang X et al (2012) Resveratrol rescues SIRT1-dependent adult stem cell decline and alleviates progeroid features in laminopathy-based progeria. Cell Metab 16:738–750

Liu B, Wang J, Chan KM et al (2005) Genomic instability in laminopathy-based premature aging. Nat Med 11:780–785

Liu Y, Rusinol A, Sinensky M et al (2006) DNA damage responses in progeroid syndromes arise from defective maturation of prelamin A. J Cell Sci 119:4644–4649

Lombardi ML, Jaalouk DE, Shanahan CM et al (2011) The interaction between nesprins and sun proteins at the nuclear envelope is critical for force transmission between the nucleus and cytoskeleton. J Biol Chem 286:26743–26753

Lu JT, Muchir A, Nagy PL et al (2011) LMNA cardiomyopathy: cell biology and genetics meet clinical medicine. Dis Model Mech 4:562–568

Mallampalli MP, Huyer G, Bendale P et al (2005) Inhibiting farnesylation reverses the nuclear morphology defect in a HeLa cell model for Hutchinson-Gilford progeria syndrome. Proc Natl Acad Sci U S A 102:14416–11421

Manilal S, Nguyen TM, Sewry CA et al (1996) The Emery-Dreifuss muscular dystrophy protein, emerin, is a nuclear membrane protein. Hum Mol Genet 5:801–808

Meinke P, Mattioli E, Haque F et al (2014) Muscular dystrophy-associated SUN1 and SUN2 variants disrupt nuclear-cytoskeletal connections and myonuclear organization. PLoS Genet 10, e1004605

Merideth MA, Gordon LB, Clauss S et al (2008) Phenotype and course of Hutchinson-Gilford progeria syndrome. N Engl J Med 358:592–604

Mory PB, Crispim F, Freire MB et al (2012) Phenotypic diversity in patients with lipodystrophy associated with LMNA mutations. Eur J Endocrinol 167:423–431

Moulson CL, Go G, Gardner JM et al (2005) Homozygous and compound heterozygous mutations in ZMPSTE24 cause the laminopathy restrictive dermopathy. J Invest Dermatol 125:913–919

Muchir A, Bonne G, van der Kooi AJ et al (2000) Identification of mutations in the gene encoding lamins A/C in autosomal dominant limb girdle muscular dystrophy with atrioventricular conduction disturbances (LGMD1B). Hum Mol Genet 9:1453–1459

Muchir A, Pavlidis P, Decostre V et al (2007) Activation of MAPK pathways links LMNA mutations to cardiomyopathy in Emery-Dreifuss muscular dystrophy. J Clin Invest 117:1282–1293

Muchir A, Kim YJ, Reilly SA et al (2013) Inhibition of extracellular signal-regulated kinase 1/2 signaling has beneficial effects on skeletal muscle in a mouse model of Emery-Dreifuss muscular dystrophy caused by lamin A/C gene mutation. Skelet Muscle 3:17

Muchir A, Wu W, Choi JC et al (2012) Abnormal p38α mitogen-activated protein kinase signaling in dilated cardiomyopathy caused by lamin A/C gene mutation. Hum Mol Genet 21:4325–4333

Muchir A, Wu W, Worman HJ (2009) Reduced expression of A-type lamins and emerin activates extracellular signal-regulated kinase in cultured cells. Biochim Biophys Acta 1792:75–81

Nagano A, Koga R, Ogawa M et al (1996) Emerin deficiency at the nuclear membrane in patients with Emery-Dreifuss muscular dystrophy. Nat Genet 12:254–259

Naismith TV, Heuser JE, Breakefield XO et al (2004) TorsinA in the nuclear envelope. Proc Natl Acad Sci U S A 101:7612–7617

Navarro CL, Cadiñanos J, De Sandre-Giovannoli A et al (2005) Loss of ZMPSTE24 (FACE-1) causes autosomal recessive restrictive dermopathy and accumulation of Lamin A precursors. Hum Mol Genet 14:1503–1513

Neilson DE, Adams MD, Orr CM et al (2009) Infection-triggered familial or recurrent cases of acute necrotizing encephalopathy caused by mutations in a component of the nuclear pore, RANBP2. Am J Hum Genet 84:44–51

Novelli G, Muchir A, Sangiuolo F et al (2002) Mandibuloacral dysplasia is caused by a mutation in LMNA-encoding lamin A/C. Am J Hum Genet 71:426–431

Olive M, Harten I, Mitchell R et al (2010) Cardiovascular pathology in Hutchinson-Gilford progeria: correlation with the vascular pathology of aging. Arterioscler Thromb Vasc Biol 30:2301–2309

Ozelius LJ, Hewett JW, Page CE et al (1997) The early-onset torsion dystonia gene (DYT1) encodes an ATP-binding protein. Nat Genet 17:40–48

Padiath QS, Saigoh K, Schiffmann R et al (2006) Lamin B1 duplications cause autosomal dominant leukodystrophy. Nat Genet 38:1114–1123

Pan D, Estévez-Salmerón LD, Stroschein SL et al (2005) The integral inner nuclear membrane protein MAN1 physically interacts with the R-Smad proteins to repress signaling by the transforming growth factor-β superfamily of cytokines. J Biol Chem 280:15992–16001

Pendás AM, Zhou Z, Cadiñanos J et al (2002) Defective prelamin A processing and muscular and adipocyte alterations in Zmpste24 metalloproteinase-deficient mice. Nat Genet 31:94–99

Puckelwartz MJ, Kessler EJ, Kim G et al (2010) Nesprin-1 mutations in human and murine cardiomyopathy. Mol Cell Cardiol 48:600–608

Raju GP, Dimova N, Klein PS et al (2003) SANE, a novel LEM domain protein, regulates bone morphogenetic protein signaling through interaction with Smad1. J Biol Chem 278:428–437
Ramos FJ, Chen SC, Garelick MG et al (2012) Rapamycin reverses elevated mTORC1 signaling in lamin A/C-deficient mice, rescues cardiac and skeletal muscle function, and extends survival. Sci Transl Med 4:144ra103
Roux KJ, Crisp ML, Liu Q et al (2009) Nesprin 4 is an outer nuclear membrane protein that can induce kinesin-mediated cell polarization. Proc Natl Acad Sci U S A 106:2194–2199
Scaffidi P, Misteli T (2008) Lamin A-dependent misregulation of adult stem cells associated with accelerated ageing. Nat Cell Biol 10:452–459
Senior A, Gerace L (1988) Integral membrane proteins specific to the inner nuclear membrane and associated with the nuclear lamina. J Cell Biol 107:2029–2036
Shackleton S, Lloyd DJ, Jackson SN et al (2000) LMNA, encoding lamin A/C, is mutated in partial lipodystrophy. Nat Genet 24:153–156
Sinensky M, Fantle K, Trujillo M et al (1994) The processing pathway of prelamin A. J Cell Sci 107:61–67
Shin JY, Le Dour C, Sera F et al (2014) Depletion of lamina-associated polypeptide 1 from cardiomyocytes causes cardiac dysfunction in mice. Nucleus 5:260–459
Shin JY, Méndez-López I, Wang Y et al (2013) Lamina-associated polypeptide-1 interacts with the muscular dystrophy protein emerin and is essential for skeletal muscle maintenance. Dev Cell 26:591–603
Singh RR, Sedani S, Lim M et al (2015) RANBP2 mutation and acute necrotizing encephalopathy: 2 cases and a literature review of the expanding clinico-radiological phenotype. Eur J Paediatr Neurol 19:106–113
Sosa BA, Demircioglu FE, Chen JZ et al (2014) How lamina-associated polypeptide 1 (LAP1) activates Torsin. Elife 3, e03239
Speckman RA, Garg A, Du F et al (2000) Mutational and haplotype analyses of families with familial partial lipodystrophy (Dunnigan variety) reveal recurrent missense mutations in the globular C-terminal domain of lamin A/C. Am J Hum Genet 66:1192–1198
Stehbens WE, Wakefield SJ, Gilbert-Barness E et al (1999) Histological and ultrastructural features of atherosclerosis in progeria. Cardiovasc Pathol 8:29–39
Sullivan T, Escalante-Alcalde D, Bhatt H et al (1999) Loss of A-type lamin expression compromises nuclear envelope integrity leading to muscular dystrophy. J Cell Biol 147:913–920
Taylor MR, Slavov D, Gajewski A et al (2005) Thymopoietin (lamina-associated polypeptide 2) gene mutation associated with dilated cardiomyopathy. Hum Mutat 26:566–574
Tazir M, Azzedine H, Assami S et al (2004) Phenotypic variability in autosomal recessive axonal Charcot-Marie-Tooth disease due to the R298C mutation in lamin A/C. Brain 127:154–163
Toth JI, Yang SH, Qiao X et al (2005) Blocking protein farnesyltransferase improves nuclear shape in fibroblasts from humans with progeroid syndromes. Proc Natl Acad Sci U S A 102:12873–12878
Tullio-Pelet A, Salomon R, Hadj-Rabia S et al (2000) Mutant WD-repeat protein in triple-A syndrome. Nat Genet 26:332–335
Verstraeten VL, Ji JY, Cummings KS et al (2008) Increased mechanosensitivity and nuclear stiffness in Hutchinson-Gilford progeria cells: effects of farnesyltransferase inhibitors. Aging Cell 7:383–393
Waterham HR, Koster J, Mooyer P et al (2003) Autosomal recessive HEM/Greenberg skeletal dysplasia is caused by 3 beta-hydroxysterol delta 14-reductase deficiency due to mutations in the lamin B receptor gene. Am J Hum Genet 72:1013–1017
Wojtanik KM, Edgemon K, Viswanadha S et al (2009) The role of LMNA in adipose: a novel mouse model of lipodystrophy based on the Dunnigan-type familial partial lipodystrophy mutation. J Lipid Res 50:1068–1079
Worman HJ, Evans CD, Blobel G (1990) The lamin B receptor of the nuclear envelope inner membrane: a polytopic protein with eight potential transmembrane domains. J Cell Biol 111:1535–1542

Worman HJ, Fong LG, Muchir A et al (2009) Laminopathies and the long strange trip from basic cell biology to therapy. J Clin Invest 119:1825–1836

Worman HJ, Yuan J, Blobel G et al (1988) A lamin B receptor in the nuclear envelope. Proc Natl Acad Sci U S A 85:8531–8534

Wu W, Muchir A, Shan J et al (2011) Mitogen-activated protein kinase inhibitors improve heart function and prevent fibrosis in cardiomyopathy caused by mutation in lamin A/C gene. Circulation 123:53–61

Yang SH, Andres DA, Spielmann HP et al (2008) Progerin elicits disease phenotypes of progeria in mice whether or not it is farnesylated. J Clin Invest 118:3291–3300

Yang SH, Bergo MO, Toth JI et al (2005) Blocking protein farnesyltransferase improves nuclear blebbing in mouse fibroblasts with a targeted Hutchinson-Gilford progeria syndrome mutation. Proc Natl Acad Sci U S A 102:10291–10296

Yang SH, Chang SY, Ren S et al (2011) Absence of progeria-like disease phenotypes in knock-in mice expressing a non-farnesylated version of progerin. Hum Mol Genet 20:436–444

Yang SH, Meta M, Qiao X et al (2006) A farnesyltransferase inhibitor improves disease phenotypes in mice with a Hutchinson-Gilford progeria syndrome mutation. J Clin Invest 116:2115–2121

Ye Q, Worman HJ (1996) Interaction between an integral protein of the nuclear envelope inner membrane and human chromodomain proteins homologous to Drosophila HP1. J Biol Chem 271:14653–14656

Young SG, Yang SH, Davies BS et al (2013) Targeting protein prenylation in progeria. Sci Transl Med 5:171ps3

Zhang Q, Bethmann C, Worth NF (2007) Nesprin-1 and -2 are involved in the pathogenesis of Emery Dreifuss muscular dystrophy and are critical for nuclear envelope integrity. Hum Mol Genet 16:2816–2833

Zhang X, Chen S, Yoo S et al (2008) Mutation in nuclear pore component NUP155 leads to atrial fibrillation and early sudden cardiac death. Cell 135:1017–1027

Zhang X, Yang H, Corydon MJ et al (1999) Localization of a human nucleoporin 155 gene (NUP155) to the 5p13 region and cloning of its cDNA. Genomics 57:144–151

Part II
Nuclear Bodies

The Nucleolus: Structure and Function

Marie-Line Dubois and François-Michel Boisvert

Abstract The nucleolus is the largest nuclear organelle and is the primary site of ribosome subunit biogenesis in eukaryotic cells. It is assembled around arrays of ribosomal DNA genes, forming specific chromosomal features known as nucleolar organizing regions (NORs) which are the sites of ribosomal DNA transcription. While the nucleolus main activity involve different steps of ribosome biogenesis, the presence of proteins with no obvious relationship with ribosome subunit production suggests additional functions for the nucleolus, such as regulation of mitosis, cell cycle progression, stress response and biogenesis of multiple ribonucleoprotein complexes. The many novel factors and separate classes of proteins identified within the nucleolus support this view that the nucleolus may perform additional functions beyond its known role in ribosome subunit biogenesis. Here we review our knowledge of the nucleolar functions and will provide a detailed picture of how the nucleolus is involved in many cellular pathways.

Keywords Nucleolus • Ribosome biogenesis • rRNA transcription

1 Introduction

Nucleoli are present in almost every eukaryotic cell type and represent the most prominent compartment of the cell nucleus. The primary function of the nucleolus consists in ribosomal RNA (rRNA) transcription, rRNA processing and ribosome subunit assembly (Hernandez-Verdun et al. 2010; Pederson 2011; Raska et al. 2006). Nucleoli assemble at the end of mitosis around the tandemly repeated clusters of rDNA genes forming a subnuclear compartment that locally recruits the specific transcription and processing machineries that are responsible for generating ribosome subunits (Hernandez-Verdun 2011; Raska et al. 2006). The process of assembling a ribosome subunit requires the initial transcription of the ribosomal DNA (rDNA) genes by the RNA polymerase I. Because these rDNA genes are

M.-L. Dubois • F.-M. Boisvert (✉)
Department of Anatomy and Cell Biology, Université de Sherbrooke, Sherbrooke, Québec, Canada
e-mail: fm.boisvert@usherbrooke.ca

D.P. Bazett-Jones, G. Dellaire (eds.), *The Functional Nucleus*,
DOI 10.1007/978-3-319-38882-3_2

arranged in arrays of tandem repeats, it results in local concentration of proteins involved in different aspect of transcription, processing and assembly of rRNA into ribosomes. In higher eukaryotes, three sub-nucleolar compartments can be distinguished by their distinct morphology using electron and light microscopy: The fibrillar centres (FC) are surrounded by the dense fibrillar component (DFC) and the granular component (GC), in which the FC and DFC are embedded. The composition of these sub-compartments is tightly linked to sequential steps in ribosome biogenesis [reviewed in (Olson and Dundr 2005)]. The FC contains unengaged RNA polymerase I transcription factors, whereas the DFC contains mostly pre-RNA processing factors, indicating specialization of these compartments. Transcription occurs at the boundary of the FC and DFC (Raska et al. 2006), and the transcribed rRNA is then moving to the GC compartment for further maturation and assembly into ribosomes (Figs. 1 and 2).

In many cell types, only a subset of rDNA genes are transcriptionally active, even though inactive rDNA are still assembled into nucleoli. The initial 47S ribosomal RNA (rRNA) precursor transcript transcribed by RNA pol I is subsequently cleaved to form the mature 28S, 18S and 5.8S rRNAs which is then post-transcriptionally modified through interaction with small nucleolar ribonucleoproteins (snoRNPs) and additional protein processing factors (Henras et al. 2015). Finally, the processed and modified rRNAs are assembled with the many ribosomal proteins, prior to interaction with the export machinery and transport to the cytoplasm.

While the nucleolar function in ribosomal biogenesis is well characterized, many additional functions of the nucleolus besides ribosomal biogenesis have been uncovered, suggesting many important cellular functions for the nucleolus (Boisvert et al. 2007; Pederson and Powell 2015). This was made particularly evident following several reports of proteomic analyses to characterize the nucleolar proteome in human, mouse, as well as other organisms (Andersen et al. 2002, 2005; Kar et al. 2011; Pendle et al. 2005; Scherl et al. 2002). Nucleolar proteins identified in these studies have shown that over 70 % are not involved in the production of ribosome subunits, consistent with the nucleolus performing additional cellular functions (Boisvert et al. 2007; Lam and Trinkle-Mulcahy 2015; Pederson and Powell 2015). A further dimension to these studies has been added by recent studies that have characterized the dynamic protein composition of the nucleolus following treatment with actinomycin D (Andersen et al. 2005), etoposide (Boisvert et al. 2010) or through the cell cycle (Ly et al. 2014).

Additionally, the nucleolus has been linked to multiple forms of diseases involving a wide range of mechanisms including cancer (Tsai and Pederson 2014), viral infections (Hiscox 2007) and neurodegenerative diseases (Parlato and Kreiner 2013), affecting either ribosome biogenesis, nucleolar structure or other functions.

This chapter provides an overview of the characteristics of nucleolar organisation, discusses non-ribosomal functions of the nucleolus, and will present some of

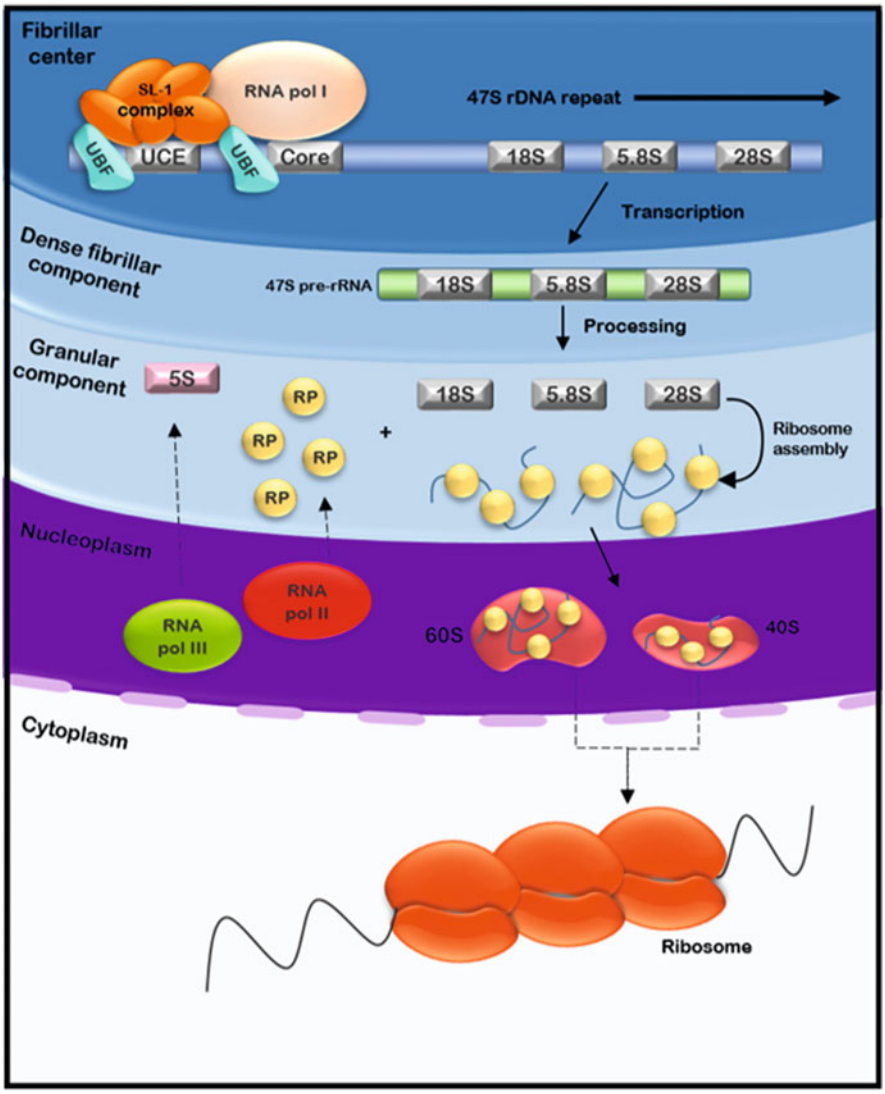

Fig. 1 Ribosome biogenesis. Transcription of ribosomal DNA by RNA polymerase I occurs at the fibrillary centres or at the boundary with the dense fibrillary component. The 47S pre-ribosomal RNA transcripts are then further processed by specific cleavages and post-transcriptional modifications by small nucleolar ribonucleoproteins (snoRNPs) in the dense fibrillary component. Assembly of the rRNA with the ribosomal proteins then occurs in the granular component of the nucleolus prior to export of the 40 and 60S subunits to the cytoplasm for final assembly into the functional ribosome

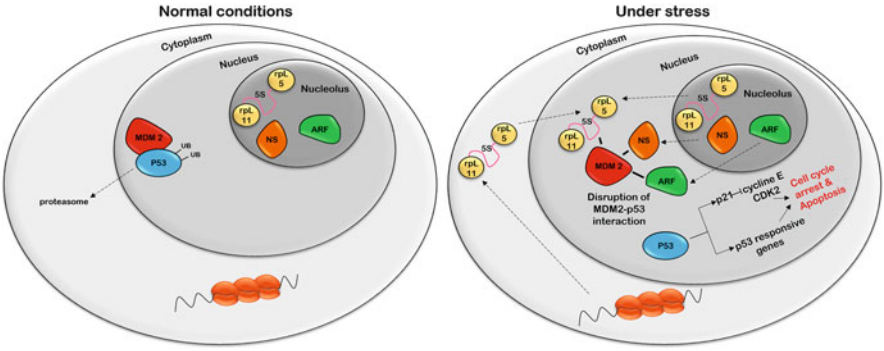

Fig. 2 Nucleolus under stress. The nucleolus plays a central role during the cellular response to stress. Under normal conditions, p53 is kept at very low level by proteasomal mediated degradation through ubiquitination by MDM2. Following activation of stress responses, oncogene activation or DNA damage, the p14ARF tumor suppressor normally located in the nucleolus associates with HDM2 and sequesters it within the nucleolus. This segregation prevents the ubiquitination of p53 mediated by HDM2 resulting in increased levels of cellular p53 resulting in cell cycle arrest and apoptosis

the latest findings regarding their regulation and dynamic behaviour, as well as its implications in cancers and diseases.

2 The Nucleolar Organisation

2.1 Nucleolar Organiser Regions (NORs)

The rDNA genes are arranged in arrays of head-to-tail tandem repeats, termed nucleolar organizer regions (NORs). In the human genome, approximately 400 copies of 43-kb repeat units are distributed along all acrocentric chromosomes (chromosomes 13, 14, 15, 21 and 22) to form NORs (Henderson et al. 1972), which are assumed to have resulted from interchromosomal recombination (Gonzalez and Sylvester 1997; Worton et al. 1988). While the rDNA repeats are mostly arranged into canonical head-to-tail repeats, around a third of them are arranged into palindromic or non-canonical repeats (Caburet et al. 2005). During metaphase, these clusters contain so called r-chromatin (for ribosomal genes complexed with proteins involved in rDNA transcription) and are constituted of 60–80 nm fibers showing a twisted loop organization as visualized by electron tomography (Heliot et al. 1997). These RNA Polymerase I associated transcription factors remain associated with NORs throughout mitosis as well (Roussel et al. 1996).

Each of the 43 kb repeats includes a 13–14 kb segment coding for the rRNA sequence and are separated from the next transcription unit by 30 kb intergenic spacers (IGS). These spacers contains regulatory elements such as the gene promoter, repetitive enhancer elements and terminator sequences. Polar replication fork barriers

(RFBs) within the transcription termination element between the individual repeats ensure the stalling of the bidirectionally running replication fork opposite to the direction of transcription (Brewer and Fangman 1988; Gerber et al. 1997; Hernandez et al. 1993; Little et al. 1993; Wiesendanger et al. 1994). Additional elements located within the IGS including small, non-coding RNAs such promoter RNAs (pRNAs) have been recently identified. These 150–250 nucleotides long RNAs are transcribed from a promoter within the IGS and involved in epigenetic mechanisms acting on the rDNA locus (Mayer et al. 2006, 2008; Santoro et al. 2010).

In many cell types, only a subset of rDNA genes are transcriptionally active, even though inactive rDNA are still assembled into nucleoli. (Akhmanova et al. 2000; Strohner et al. 2001; Sullivan et al. 2001). The difference results from different chromatin states of rDNA repeats that correspond with the state of the nucleosomes within the rDNA (Dammann et al. 1993; Langst et al. 1998; Li et al. 2006b; Sogo et al. 1984). Inactive rDNA has been shown to be inaccessible for psoralen crosslinking and exhibits regularly spaced nucleosomes, while active regions of rDNA is accessible for crosslinking and lack nucleosomes (Conconi et al. 1989; Dammann et al. 1993). These different chromatin states are usually stable throughout the cell cycle (Conconi et al. 1989). The rate of production of rDNA can therefore be regulated either by increasing the transcription rate of active genes and/or by activation of silent genes. In general, rapid changes in rRNA expression, i.e., in response to nutrient status or growth factor signalling, will result in a change in the transcription rate of rDNA genes that are already active (Grummt and Pikaard 2003; Russell and Zomerdijk 2005; Stefanovsky et al. 2006), whereas slower changes such as during development or differentiation results from a change in the number of genes that are actively transcribed (Haaf et al. 1991).

2.2 rDNA Transcription

The RNA Polymerase I does not require a TATA box sequence in the promoter, but instead relies on regulatory sequences that are divided in two functional elements: a core element (CORE) next to the transcription start site located between −45 to +20, and an upstream control element (UCE) located between −200 to −107 (Haltiner et al. 1986; Learned et al. 1986). Transcription of rRNA genes requires the formation of a preinitiation complex (PIC) that is composed of the RNA Polymerase I, the upstream binding factor (UBF) and the promoter selectivity factor (SL1) at the rDNA promoter [reviewed in (Grummt and Pikaard 2003; Moss et al. 2007; Russell and Zomerdijk 2005)]. Upon dimerization, the upstream binding factor UBF binds the UCE and the CORE element to recruit the SL1 protein complex composed of the TATA-binding protein (TBP) and TBP-associated factors (TAFs), TAF_I 110/95; TAF_I 68; TAF_I 48 (Comai et al. 1992; Heix et al. 1997; Zomerdijk et al. 1994). The UBF dimer allows the UCE and CORE elements to come into contact by introducing loops in the upstream region into a structure called the enhanceosome to allow the binding of SL1 and formation of a stable PIC

(Bazett-Jones et al. 1994; Bell et al. 1988). By interacting with UBF and the RNA Pol I component TIF-1A, TAFs recruit RNA Pol I to the rDNA promoter (Miller et al. 2001; Moorefield et al. 2000). UBF not only binds to the promoter region, but also throughout the whole rDNA repeat (Mais et al. 2005; O'Sullivan et al. 2002) where it acts as a transcriptional activator of RNA Pol I, modulator of elongation in response to growth factor signalling and as an anti-repressor by replacing histone H1 (Kermekchiev et al. 1997; Kuhn and Grummt 1992). Thus, regulation of UBF is central to the regulation of rDNA transcription by RNA Polymerase I. Important insight into the role of UBF in ribosomal chromatin remodelling and nucleolar formation has come from the studies of pseudoNORs [reviewed in (Prieto and McStay 2008)]. PseudoNORs are artificial, high-affinity UBF-binding arrays that are transcriptionally silent due to the lack of a promoter, but behave like active NORs. For example, they exhibit consistent under condensation throughout the cell cycle and are highly enriched in the RNA Pol I machinery through UBF (Mais et al. 2005; Prieto and McStay 2007) that stays associated with rDNA after nucleoli disassembly in mitosis (Roussel et al. 1993). Therefore, UBF is believed to establish a chromatin structure that allows rapid re-initiation of rRNA transcription after mitosis and thereby promotes nucleolar formation. This is also supported by studies that show rDNA gene silencing upon UBF deletion, suggesting a regulatory role for UBF in determining the number of active rDNA genes (Sanij and Hannan 2009; Sanij et al. 2008).

As RNA Polymerase I starts the elongation process of the rRNA, UBF and SL1 remain associated with the promoter region to allow recruitment of another RNA Polymerase I so that each rDNA gene can be transcribed multiple times simultaneously, contrary to RNA Polymerase II genes. This gives rise to the Christmas tree like structure as visualized by electron microscopy (Miller and Beatty 1969). Termination occurs through the binding of TTF-I at the 3′ end of the transcribed region, bending the termination site and, with the help of the transcript-releasing factor PTRF, will induce RNA Polymerase I to dissociate from the DNA and the new transcript.

2.3 Processing of rRNA and Ribosome Assembly

The production of human 40S and 60S ribosomal subunits prior to export into the cytoplasm is a complex mechanism that requires a large proportion of the cellular energy for production (Warner 1999). Indeed, the size and organisation of the nucleolus reflects the scale of this process (Sirri et al. 2008). The regulation of ribosome biogenesis is crucial for cellular growth and proliferation (Tschochner and Hurt 2003) and it is upregulated in the majority of cancers to accommodate the increase rate of cellular growth and metabolism (Ruggero and Pandolfi 2003). The two ribosomal subunits includes 79 major ribosomal proteins (RPs) as well as four mature ribosomal RNAs (Yusupova and Yusupov 2014). The smaller ribosomal subunit (40S) includes the 18S rRNA, whereas the large subunit (60S) includes the

28S, 5.8S as well as the RNA Polymerase III transcribed 5S rRNA (Fedoriw et al. 2012).

The rRNA precursor is transcribed by RNA Pol I at the FC/DFC border as one long 47S transcript (~13 kb in humans), which includes the sequences for the 18S, 5.8S and 28S rRNA in that order. The additional sequences are called external transcribed spacers called ETS at both ends of the rRNA (5′-ETS and 3′-ETS), as well as internal transcribed spacers (ITS1 and ITS2) within the rRNA precursor. Sequential cleavages at specific sites termed A′, A0, 1, 2 and 3 will generate the 18S rRNA and requires distinct small nucleolar RNAs (snoRNAs) (Henras et al. 2015). Cleavage at the T1 site by the U8 snoRNA will remove the 3′-ETS. Site 3 cleavage near the end of ITS1 between the 18S and 5.8S rRNAs will generate a 32S intermediate which will be further processed (Preti et al. 2013). Cleavage at the 4′ site will generate the 28S rRNA and will results in a 5.8S rRNA with a longer 3′-end that will be further processed. The maturing rRNAs will be assembled within a 90S pre-ribosome complex which accumulates in the DFC (Henras et al. 2015; Tschochner and Hurt 2003). The pre-ribosome contains the 47S rRNA, 5S rRNA, ribosomal proteins and ~150 non-ribosomal proteins, including factors involved in processing and maturation, for example endo- and exonucleases, pseudouridine synthases, methyltransferases, RNA chaperones, GTPases and AAA-ATPases helicases (Tschochner and Hurt 2003). The pre-ribosome is subsequently separated into pre-40S and pre-60S subunits in the GC. These subunits are exported to the cytoplasm, where they undergo further processing to form the mature small 40S and large 60S ribosome subunits. The 40S subunit contains 18S rRNA and ~33 ribosomal proteins, whereas the 60S subunit is composed of the 28S and 5.8S rRNAs, the RNA Pol III-transcribed 5S rRNA and ~49 ribosomal proteins. Interestingly, around two thirds of the ribosomal protein genes are duplicated in humans. Although it has been assumed that these copies are largely redundant, recent work has suggested that these copies exhibit functional specificity (Komili et al. 2007).

3 Other Functions for the Nucleolus

3.1 RNA Complexity in the Nucleolus

While the nucleolus is known to have a major role in coordinating the processing and maturation of rRNAs, there is now extensive literature demonstrating that the nucleolus is also involved in the processing and maturation of several different families of RNA. For example, the nucleolus has been suggested to be a site of covalent RNA modifications and protein assembly of multiple ribonucleoprotein complexes, such as the spliceosomal small nuclear RNPs, telomerase and several other small RNAs transcribed by RNA polymerase III, such as 5S rRNA, some

tRNAs, RNAse P RNA, the signal recognition particle (SRP) RNA and now also miRNA (Gerbi et al. 2003; Lam and Trinkle-Mulcahy 2015).

The signal recognition particle (SRP) is a ribonucleoprotein complex responsible for the recognition of the N-terminal signal peptide sequence on nascent proteins and for proper targeting of proteins onto a receptor on the cytoplasmic face of the endoplasmic reticulum (Walter and Johnson 1994). This complex is formed by 6 proteins and a ~300 nucleotide RNA (Walter and Blobel 1982). Recent studies have shown that both the RNA and proteins from the SRP can be found to transit through the nucleolus of mammalian cells prior to SRP export to the cytoplasm (Jacobson and Pederson 1998). These results suggest that the nucleolus is the site of assembly and processing of the SRP complex prior to their cytoplasmic export, and that the RNA could be modified/matured within the nucleolus.

In addition to the above roles for the nucleolus in RNA modification and maturation, there are several other observations linking the nucleolus to RNA processing. The RNase P RNA, a component of the pre-tRNA processing enzyme RNase P, has been found in both the nucleolus and nucleoplasm (Jacobson et al. 1997), suggesting that some pre-tRNA processing happens within the nucleolus. An alternative possibility is that the nucleolus plays a role in the assembly of RNase P ribonucleoprotein complex. The nucleolus contains all the trans-acting factors that are responsible for the synthesis of the eight 2′-O-methylated nucleotides and three pseudouridine residues carried by the mammalian U6 spliceosomal small nuclear RNA (Ganot et al. 1999). These findings demonstrate a trafficking pathway in which the U6 spliceosomal RNA cycles through the nucleolus to undergo nucleolar RNA-directed processing. Interestingly, like 5S rRNA, these RNA (tRNA, RNase P RNA and U6 spliceosomal small nuclear RNA) are all transcribed outside the nucleolus by RNA polymerase III and subsequently transit through the nucleolus. This suggest a possible common maturation process shared by RNA pol III transcripts that could occur within the nucleolus. ADAR1 and ADAR2 are editing enzymes that deaminate adenosine to inosine in long double stranded RNA duplexes and specific pre-mRNA transcripts. Live microscopy experiments demonstrate that ADAR1 and ADAR2 are in constant flux in and out of the nucleolus (Desterro et al. 2003). Furthermore, it was shown that ADAR2- but not ADAR1-mediated RNA editing occurs within the nucleolus, indicating a role for the nucleolus in the regulation of RNA editing (Vitali et al. 2005).

More recently, evidence has started to emerge demonstrating a role for the nucleolus in the regulation of small interfering RNA (siRNA). The finding that many proteins involved in siRNA processing, including RDR2, DCL3, AGO4, and NRPD1b (the largest subunit of RNA Pol IVb) were identified with siRNAs within the nucleolus in plant cells suggest that processing of endogenous nuclear siRNAs, and possibly RISC storage or sequestration, occurs within the nucleolus (Li et al. 2006a; Pontes et al. 2006). It was also reported in mammalian cells that a microRNA (miR-206) had been found to co-localize with the 28S rRNA in the granular component of the nucleolus, implying that this miRNA associates early with the ribosome subunits (Politz et al. 2006). Several other miRNAs have also been identified within the nucleolus, further supporting these observations (Bai

et al. 2014a, b; Li et al. 2013). It will be interesting to determine whether multiple forms of miRNAs arise within the nucleolus and whether they either function in nucleolar processes or leave the nucleolus to regulate downstream cellular events, such as protein translation.

3.2 Mitosis and Cell Cycle Regulation

The nucleolus is a dynamic structure which is disassembled/re-assembly at each cell division. The first step in prophase is initiate by the phosphorylation of components of the rDNA-transcription machinery by cyclin B1 and CDK1 (Heix et al. 1998). The phosphorylation of these components triggers a repression of rRNA transcription but the machinery such as UBF remains localized to the NORs (Dundr et al. 2000). In contrast, during the disassembly of nucleolus, the rRNA-processing machinery does not remain associated with nucleolar regions and move to the cytoplasm or become attached to the surface of condensed chromosomes at the perichromosomal region (PR) (Gautier et al. 1992). During anaphase, the processing proteins remains attached to the PR while proteins within the cytoplasm become packaged into nucleolar-derived foci (NDF). Cyclin B1/CDK1 levels decrease during late anaphase and early telophase, which results in the reactivation of rRNA transcription by loss of hypherphosphorylation of the transcription machinery (Sirri et al. 2000). During the G1 phase, the rRNA-processing proteins are release in a specific order starting with proteins such as fibrillarin which plays a role early in the rRNA processing followed by proteins involved in late stage of processing (Leung et al. 2004). The dynamic of nucleoli during cell cycle involves several stages which are tightly regulated and still not fully understood. Interestingly, the rDNA-transcription and rRNA-processing proteins might be regulated independently during the cell cycle.

During the interphase, the nucleoli remains dynamic. The nucleolar proteome has been shown to contain approximately 4000 proteins which can change localization (Boisvert et al. 2010). This pool of proteins is not all involved in ribosome biogenesis and several proteins have been shown to interact with the nucleolus at different stages of the cell cycle, further underlining a potential role in cell cycle regulation. In fact, the nucleolus and cell cycle regulation are influenced mutually. The most striking situation where the cell cycle can alter the ribosome biogenesis is in response to DNA damage (Jordan and Carmo-Fonseca 1998). Two proteins involved in DNA damage response, DNA-PK and PARP1, have been identified to be responsible for this inhibition (Calkins et al. 2013). Additionally, the RNA Polymerase I can be inhibited by DNA lesions pathway dependent of ATM, NBS1 and MDC1 (Kruhlak et al. 2007). The nucleolus can also regulates key cell cycle checkpoints as well. One of the mechanisms is by modulation of post-translational modifications such as sumoylation and phosphorylation of proteins involved in the cell cycle. For example, SENP5, a SUMO-specific protease found within the nucleolus is involved in sumoylation of proteins that affect progression

through cell division (Di Bacco et al. 2006). Another example is the sequestration in nucleoli of the telomerase reverse transcriptase, the RNP enzyme that adds telomeric sequences (Wong et al. 2002). The telomerase is retained in the nucleolus until the telomeres are replicated at stages of S phase.

3.3 Stress Sensor

The ribosome biogenesis is the mechanism that requires the largest amount of cell's energy, consuming ~80 % of total cell's energy (Schmidt 1999). Considering the large amount of effort needed for the biogenesis of ribosomes, it is surprising that one of the strategies used to preserve energy homeostasis is decreasing ribosome biosynthesis under stress conditions. In response to stress, such as nutrient deprivation, oxidative stress or drastic change in temperature, several steps in ribosome biogenesis is altered. These stresses can cause a change in the localization of nucleolar proteins, downregulation of RNA polymerase I activity, pre-rRNA processing and transport, or changes in chromatin structure (Grummt 2013). Many drug have also been shown to alter the integrity and the activities of the nucleolus. For example, UV irradiation leads to segregation of nucleolar components and ends with the complete disintegration of the nucleolus (Govoni et al. 1994).

The alteration of any of the steps in ribosome biogenesis, such as transcription, processing and assembly of the 40S and 60S subunits, results in the activation of nucleolar stress pathways leading to senescence or apoptosis. The activation of this pathway culminates in the stabilization of the p53 protein (Rubbi and Milner 2003) by disruption of its interaction with MDM2, which can no longer add ubiquitin on p53 and target it for proteasomal degradation. MDM2 can be targeted by three proteins in the nucleolus, alternative reading frame protein $p14^{ARF}$ (ARF), nucleostemin (NS) and ribosomal proteins (RPs), in response to nucleolar stresses. ARF is a tumor suppressor protein usually weakly expressed (Brady et al. 2004). Under different cellular stress conditions, an increase in the expression of ARF results in binding to MDM2 and disruption of the interaction between MDM2 and p53 (Weber et al. 1999). Nucleostemin is a nucleolar protein and its function in this structure is not well understood. However, following stress in the nucleolus, the protein delocalizes to the nucleoplasm and bind MDM2 to prevent the association with p53 (Tsai 2011). Finally, several ribosomal proteins have been found to be involved in p53 stabilization such as rpL11, rpL23, rpL5, rpL7 and rpL26. However recent study demonstrate that only rpL5 and rpL11 are essential for this function (Fumagalli et al. 2012). In addition, the 5s rRNA is also involved in the formation of the rpL5-rpL11-MDM2 complex (Donati et al. 2013). The change of localization of ribosomal proteins has long been considerate passive and a result of nucleolar disruption. However, the protein rpL11 can move to the nucleoplasm without loss of nucleolar integrity (Fumagalli et al. 2009), indicating that the relocalization of ribosomal proteins may be a regulated mechanism. For example, the localization of

rpL11 is regulated by NEDDylation, a ubiquitin-like molecule (Sundqvist et al. 2009). Another method to retain the ribosomal proteins within the nucleolus is through their interaction with PICT1 who is degraded in response to nucleolar stress (Maehama et al. 2014; Sasaki et al. 2011). It is clear that cells react in response to environmental stress by decreasing ribosome biogenesis to decrease energy consumption and preserve homeostasis but the nucleolus is also directly involved in the regulation of cellular response to different type of stresses.

4 The Nucleolus and Diseases

4.1 Cancer and Genomic Instability

Several years before the discovery of the role of the nucleolus in the synthesis of ribosomes, cytological analysis had already made a connection between the size of nucleoli and cancer (Pinease et al. 1896). It was observed that tumor cells had an increase in the number and size of nucleoli. Thereafter, numerous studies confirms these observations and concluded that these abnormalities could be used as a marker for the aggressiveness of malignancies (Derenzini et al. 2009) and was directly related to cell grow rate (Derenzini et al. 1998, 2000). In addition, the biogenesis of ribosome is under control of key cellular growth and proliferation signaling pathways which are known to be frequently mutated in several types of cancer such MYC, RAS, phosphatidylinositol-3-kinase (PI3K), tumor suppressors including TP53 (p53), retinoblastoma protein (Rb) and PTEN (Moss 2004). Irregular shape and deregulation of ribosome synthesis rate has long been considered a consequence of cancer. However, new data suggest that dysfunction of rRNA synthesis is not only a consequence cellular transformation but is also required for the survival of tumor cells and can even initiate the tumor transformation. This is support by two evidences: first, perturbation in ribosome biogenesis activated stress pathways and second, the specific inhibition of RNA polymerase I transcription leads to tumor cell death.

The first evidence is that the cell were demonstrated to monitors closely ribosome synthesis and disruption of this process leads to the activation of a ribosomal surveillance pathway (Zhang et al. 2003). This stress pathway leads to the accumulation of p53 through a non-genotoxic activation, as described in the section stress sensor. The second evidence is based on the observation that deregulation of ribosome synthesis is essential for the survival of cancer and small molecule inhibitors of RNA Polymerase I transcription can, in tumors cells, selectively activate the nucleolar stress pathway (Bywater et al. 2012). Many drugs already used in cancer therapy affect, at least in part, ribosome biogenesis or rRNA processing such as doxorubicin, flavopiridol and roscovitine (Burger et al. 2010). However, these treatments induce a general nucleolar disruption and affect many other pathways resulting in several side effects. By contrast, a new generation of

small molecules is being developed that selectively inhibit RNA Polymerase I transcription. The CX-3543 (first generation) and CX-5461 (second generation) bind GC-rich sequence, found in large proportion in rRNA and prevents RNA Polymerase I transcription (Balasubramanian et al. 2011). This treatment specifically induces p53 non-genotoxic stress pathway activation by release of ribosomal proteins from the nucleolus and only affect cancer cells and not in normal cells (Bywater et al. 2012). These targeted transcription therapies are less genotoxic for normal cells reducing risk of secondary cancer associated with classical chemotherapeutic agent (Godley and Larson 2008). Recent studies also show another possible pathway by which RNA Polymerase I transcription inhibitors are also effective independent of p53 activation (Peltonen et al. 2014). While the mechanisms involved remain unknown, these data underline the potential for selective RNA Polymerase I inhibition for cancer treatment whether p53 is present or mutated.

4.2 Viral Infections

Viruses are obligate intracellular parasites and used host cell for genome replication, protein expression and assembly of new virus particles. Several types of virus involved the nucleolus for effective infection. In fact, the RNA virus, retroviruses and DNA viruses interact with/or alter the nucleolus when they infect cells (Hiscox 2002). The infection results in change of morphology and in the proteome of the nucleolus (Dove et al. 2006). The association between viral proteins and the nucleolus results from three type of direct interaction : with rDNA, with nucleolar RNA (consisting mainly of rRNA) or with nucleolar protein components (Carmo-Fonseca et al. 2000). The viruses affect the nucleosome homeostasis at many level. First, the viral components can co-localize with the nucleolus. Second, the virus can use nucleolar proteins to allow his own proliferation and finally, viral infection can result in changes in the localization of nucleolar proteins.

Following infection with certain different viruses, such as coronavirus and arterivirus, some normally nuclear proteins are found within the nucleolus (Hiscox et al. 2001; Rowland et al. 1999). Changes in the protein distribution can be either through specific nucleolar-trafficking signal present in the viral proteins (Rowland and Yoo 2003) or viral proteins trafficking through the nucleolus can associate with cellular proteins and recruit them to a different localization such as the hepatitis delta antigen which requires nucleolin association for its nucleolar localization (Lee et al. 1998). The nucleolar localization of viral proteins is also important for an effective infection. For example the disruption of nucleolar localization of Semliki Forest virus non-structural protein nsP2 results in a reduction in neurovirulence (Fazakerley et al. 2002).

The three nucleolar proteins that have been most studied during viral infection are B23, fibrillarin and Nucleolin. The B23 protein acts in several functions associated with the nucleolus, such as ribosome assembly, nucleocytoplasmic

shuttling, possibly regulating transcription of rDNA and recent study have indicated is implication in the p53 regulation (Boulon et al. 2010; Hiscox 2002). During infection with HIV, B23 facilitates the nuclear import of Rev proteins promoting virus mRNA trafficking (Szebeni et al. 1997). B23 protein can also stimulates the replication of adenovirus as well (Okuwaki et al. 2001). Fibrillarin is involved in pre-rRNA processing, pre-rRNA methylation, and ribosome assembly (Tollervey et al. 1993). During adenovirus and coronavirus infection, fibrillarin redistribution leads to decrease in RNA Polymerase I transcription (Puvion-Dutilleul and Christensen 1993). Nucleolin has been involved in the regulation of gene expression, chromatin remodeling, DNA recombination and replication, RNA synthesis, rRNA processing, mRNA stabilization, cytokinesis and apoptosis (Mongelard and Bouvet 2007). During infection with HIV, nucleolin promotes DNA replication process and stimulates IRES-mediated translation of the poliovirus genome (Izumi et al. 2001). Nucleolin has also a membrane fraction which is involved in viral infection by facilitating the virus attachment (Nisole et al. 2002).

Finally, viral infection can results in relocalization of nucleolar protein to other cellular compartments. Nucleolin, B23 and fibrillarin change their localization after cell are infected with adenovirus or HIV (Matthews 2001). The mechanisms that are responsible for this delocalization are unknown, but the displacement of nucleolar proteins changes the nucleolar, nuclear or cytoplasmic pool of these proteins. The interaction between the nucleolus and viral components is thus an interesting potential therapeutic target. These associations can be used for the development of new therapies against viral infection. For example, the HIV Rev protein localizes in the nucleolus and is involved in viral mRNA trafficking (Dundr et al. 1995). A nucleolar localizing Rev response element (RRE) decoy in infected cells results in a significant inhibition of the replication of HIV-1 in cell culture (Michienzi et al. 2006).

4.3 Neurodegenerative Disorders

While the increased size of the nucleolus is associated with cell proliferation and a potential marker for cancer, the reduction in size of nucleoli is correlated with completely different disorders. It was reported patients affected by Alzheimer's disease and Parkinson's disease show a reduction in nucleolar size in different part of the brain (Mann et al. 1988; Mann and Yates 1982). In fact, the dysregulation of different nucleolar aspects is associated with several human neuropathological conditions (Kinderman and LaVelle 1976).

In Alzheimer's disease, not only the size of the nucleolus is affected but the organization of NOR is also altered (Donmez-Altuntas et al. 2005). The aberrant structure is associated with a diminution of mature rRNA products and is found as an early event in patients affected (Ding et al. 2005). This is further confirmed by recent studies which demonstrate that rRNA promoter region is hypermethylated in early Alzheimer's disease, leading to a diminution in transcription of rRNA

(Pietrzak et al. 2011). Reducing nucleolar transcription may participate in the decline of the rRNA component of brain ribosomes contributing to Alzheimer associated synapse loss and dementia. In Parkinson's disease, nucleolar integrity was also reported to be disrupted in the dopaminergic neurons of patients (Rieker et al. 2011). In support of this observation, two proteins which are associated with Parkinson's disease, Alpha-synuclein and DJ-1, are capable of interacting with nucleolin (Jin et al. 2007), although the exact function of this interaction is not known and could also be a consequence of an accumulation of DNA damage and activation of oxidative stress (Markesbery and Lovell 2006; Rieker et al. 2011).

The Huntington's disease and spinocerebellar ataxias belong to a group of polyglutamine disease which are cause by additional repetition of CAG in particular genes. The repeat of these nucleotides alters the function of the affected proteins leading to toxicity (Orr and Zoghbi 2007). However, the toxicity is not only caused by the gain of function of the proteins. Recent study showed that mutant transcripts can also contribute to the adverse effect (Tsoi et al. 2012). The mutant transcripts can alter rDNA transcription and induced apoptosis by activating nucleolar stress pathways. In amyotrophic lateral sclerosis or frontotemporal dementia, an abortive transcripts of C9orf72 gene is produce with several repetitions of the hexanucleotide repeat region GGGGCC (Haeusler et al. 2014). This abortive transcript migrates to the nucleolus and binds Nucleolin. This association results in a mislocalized Nucleolin and likely contributes to the nucleolar stress activation observed (Kwon et al. 2014).

5 Summary

It has become increasingly apparent that the cellular function for the nucleolus goes well beyond ribosome subunit production. Multiple lines of investigation have demonstrated a wide range of function including cell cycle regulation, stress responses and maturation and biogenesis of a wide range of ribonucleoprotein complexes. Several recently developed technologies are driving forward our understanding of the extent and relevance of the nucleolus in these cellular processes, and the dynamic nature of the nucleoli underline the importance of studying the structure and function of this nuclear organelle under a wide range of conditions. Considering the large number of proteins identified in the nucleolus in proteomic experiments that have no known functions or that are still poorly characterized, there will be likely further functions that will be uncovered in the future associated with the nucleolus.

Cell growth and proliferation is critically dependent on an efficient supply of ribosomes to maintain protein synthesis levels. Therefore, the nucleolus is emerging as a key centre of cell growth regulation and it is not surprising that its activity is influenced by a wide range of signaling events that can modulate the efficiency of rRNA expression and ribosome subunit assembly and transport. Several examples where disruption of nucleolar components and activities result in human disease,

including inherited genetic disorders and predisposition to cancer, directly reflecting the importance on cell function of disrupting mechanisms that occur in the nucleolus. The link between the nucleolus and regulation of such important cellular function demonstrate the potential as a therapeutic target for cancer treatment, viral infection and neurodegenerative diseases.

Acknowledgments Work in the Boisvert laboratory is supported by the Canadian Institutes of Health Research (CIHR), and by the National Sciences and Engineering Research Council of Canada (NSERC).

References

Akhmanova A, Verkerk T, Langeveld A, Grosveld F, Galjart N (2000) Characterisation of transcriptionally active and inactive chromatin domains in neurons. J Cell Sci 113 (Pt 24):4463–4474

Andersen JS, Lam YW, Leung AK, Ong SE, Lyon CE, Lamond AI, Mann M (2005) Nucleolar proteome dynamics. Nature 433:77–83

Andersen JS, Lyon CE, Fox AH, Leung AK, Lam YW, Steen H, Mann M, Lamond AI (2002) Directed proteomic analysis of the human nucleolus. Curr Biol 12:1–11

Bai B, Liu H, Laiho M (2014a) Small RNA expression and deep sequencing analyses of the nucleolus reveal the presence of nucleolus-associated microRNAs. FEBS Open Biol 4:441–449

Bai B, Yegnasubramanian S, Wheelan SJ, Laiho M (2014b) RNA-Seq of the nucleolus reveals abundant SNORD44-derived small RNAs. PLoS One 9, e107519

Balasubramanian S, Hurley LH, Neidle S (2011) Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? Nat Rev Drug Discov 10:261–275

Bazett-Jones DP, Leblanc B, Herfort M, Moss T (1994) Short-range DNA looping by the Xenopus HMG-box transcription factor, xUBF. Science 264:1134–1137

Bell SP, Learned RM, Jantzen HM, Tjian R (1988) Functional cooperativity between transcription factors UBF1 and SL1 mediates human ribosomal RNA synthesis. Science (New York, NY) 241:1192–1197

Boisvert FM, Lam YW, Lamont D, Lamond AI (2010) A quantitative proteomics analysis of subcellular proteome localization and changes induced by DNA damage. Mol Cell Proteomics 9:457–470

Boisvert FM, van Koningsbruggen S, Navascues J, Lamond AI (2007) The multifunctional nucleolus. Nat Rev Mol Cell Biol 8:574–585

Boulon S, Westman BJ, Hutten S, Boisvert FM, Lamond AI (2010) The nucleolus under stress. Mol Cell 40:216–227

Brady SN, Yu Y, Maggi LB Jr, Weber JD (2004) ARF impedes NPM/B23 shuttling in an Mdm2-sensitive tumor suppressor pathway. Mol Cell Biol 24:9327–9338

Brewer BJ, Fangman WL (1988) A replication fork barrier at the 3′ end of yeast ribosomal RNA genes. Cell 55:637–643

Burger K, Muhl B, Harasim T, Rohrmoser M, Malamoussi A, Orban M, Kellner M, Gruber-Eber A, Kremmer E, Holzel M, Eick D (2010) Chemotherapeutic drugs inhibit ribosome biogenesis at various levels. J Biol Chem 285:12416–12425

Bywater MJ, Poortinga G, Sanij E, Hein N, Peck A, Cullinane C, Wall M, Cluse L, Drygin D, Anderes K, Huser N, Proffitt C, Bliesath J, Haddach M, Schwaebe MK, Ryckman DM, Rice WG, Schmitt C, Lowe SW, Johnstone RW, Pearson RB, McArthur GA, Hannan RD (2012)

Inhibition of RNA polymerase I as a therapeutic strategy to promote cancer-specific activation of p53. Cancer Cell 22:51–65
Caburet S, Conti C, Schurra C, Lebofsky R, Edelstein SJ, Bensimon A (2005) Human ribosomal RNA gene arrays display a broad range of palindromic structures. Genome Res 15:1079–1085
Calkins AS, Iglehart JD, Lazaro JB (2013) DNA damage-induced inhibition of rRNA synthesis by DNA-PK and PARP-1. Nucleic Acids Res 41:7378–7386
Carmo-Fonseca M, Mendes-Soares L, Campos I (2000) To be or not to be in the nucleolus. Nat Cell Biol 2:E107–112
Comai L, Tanese N, Tjian R (1992) The TATA-binding protein and associated factors are integral components of the RNA polymerase I transcription factor, SL1. Cell 68:965–976
Conconi A, Widmer RM, Koller T, Sogo JM (1989) Two different chromatin structures coexist in ribosomal RNA genes throughout the cell cycle. Cell 57:753–761
Dammann R, Lucchini R, Koller T, Sogo JM (1993) Chromatin structures and transcription of rDNA in yeast *Saccharomyces cerevisiae*. Nucleic Acids Res 21:2331–2338
Derenzini M, Montanaro L, Trere D (2009) What the nucleolus says to a tumour pathologist. Histopathology 54:753–762
Derenzini M, Trere D, Pession A, Govoni M, Sirri V, Chieco P (2000) Nucleolar size indicates the rapidity of cell proliferation in cancer tissues. J Pathol 191:181–186
Derenzini M, Trere D, Pession A, Montanaro L, Sirri V, Ochs RL (1998) Nucleolar function and size in cancer cells. Am J Pathol 152:1291–1297
Desterro JM, Keegan LP, Lafarga M, Berciano MT, O'Connell M, Carmo-Fonseca M (2003) Dynamic association of RNA-editing enzymes with the nucleolus. J Cell Sci 116:1805–1818
Di Bacco A, Ouyang J, Lee HY, Catic A, Ploegh H, Gill G (2006) The SUMO-specific protease SENP5 is required for cell division. Mol Cell Biol 26:4489–4498
Ding Q, Markesbery WR, Chen Q, Li F, Keller JN (2005) Ribosome dysfunction is an early event in Alzheimer's disease. J Neurosci Off J Soc Neurosci 25:9171–9175
Donati G, Peddigari S, Mercer CA, Thomas G (2013) 5S ribosomal RNA is an essential component of a nascent ribosomal precursor complex that regulates the Hdm2-p53 checkpoint. Cell Rep 4:87–98
Donmez-Altuntas H, Akalin H, Karaman Y, Demirtas H, Imamoglu N, Ozkul Y (2005) Evaluation of the nucleolar organizer regions in Alzheimer's disease. Gerontology 51:297–301
Dove BK, You JH, Reed ML, Emmett SR, Brooks G, Hiscox JA (2006) Changes in nucleolar morphology and proteins during infection with the coronavirus infectious bronchitis virus. Cell Microbiol 8:1147–1157
Dundr M, Leno GH, Hammarskjold ML, Rekosh D, Helga-Maria C, Olson MO (1995) The roles of nucleolar structure and function in the subcellular location of the HIV-1 Rev protein. J Cell Sci 108(Pt 8):2811–2823
Dundr M, Misteli T, Olson MO (2000) The dynamics of postmitotic reassembly of the nucleolus. J Cell Biol 150:433–446
Fazakerley JK, Boyd A, Mikkola ML, Kaariainen L (2002) A single amino acid change in the nuclear localization sequence of the nsP2 protein affects the neurovirulence of Semliki Forest virus. J Virol 76:392–396
Fedoriw AM, Starmer J, Yee D, Magnuson T (2012) Nucleolar association and transcriptional inhibition through 5S rDNA in mammals. PLoS Genet 8, e1002468
Fumagalli S, Di Cara A, Neb-Gulati A, Natt F, Schwemberger S, Hall J, Babcock GF, Bernardi R, Pandolfi PP, Thomas G (2009) Absence of nucleolar disruption after impairment of 40S ribosome biogenesis reveals an rpL11-translation-dependent mechanism of p53 induction. Nat Cell Biol 11:501–508
Fumagalli S, Ivanenkov VV, Teng T, Thomas G (2012) Suprainduction of p53 by disruption of 40S and 60S ribosome biogenesis leads to the activation of a novel G2/M checkpoint. Genes Dev 26:1028–1040
Ganot P, Jady BE, Bortolin ML, Darzacq X, Kiss T (1999) Nucleolar factors direct the 2′-O-ribose methylation and pseudouridylation of U6 spliceosomal RNA. Mol Cell Biol 19:6906–6917

Gautier T, Robert-Nicoud M, Guilly MN, Hernandez-Verdun D (1992) Relocation of nucleolar proteins around chromosomes at mitosis. A study by confocal laser scanning microscopy. J Cell Sci 102(Pt 4):729–737

Gerber JK, Gogel E, Berger C, Wallisch M, Muller F, Grummt I, Grummt F (1997) Termination of mammalian rDNA replication: polar arrest of replication fork movement by transcription termination factor TTF-I. Cell 90:559–567

Gerbi SA, Borovjagin AV, Lange TS (2003) The nucleolus: a site of ribonucleoprotein maturation. Curr Opin Cell Biol 15:318–325

Godley LA, Larson RA (2008) Therapy-related myeloid leukemia. Semin Oncol 35:418–429

Gonzalez IL, Sylvester JE (1997) Beyond ribosomal DNA: on towards the telomere. Chromosoma 105:431–437

Govoni M, Farabegoli F, Pession A, Novello F (1994) Inhibition of topoisomerase II activity and its effect on nucleolar structure and function. Exp Cell Res 211:36–41

Grummt I (2013) The nucleolus-guardian of cellular homeostasis and genome integrity. Chromosoma 122:487–497

Grummt I, Pikaard CS (2003) Epigenetic silencing of RNA polymerase I transcription. Nat Rev 4:641–649

Haaf T, Hayman DL, Schmid M (1991) Quantitative determination of rDNA transcription units in vertebrate cells. Exp Cell Res 193:78–86

Haeusler AR, Donnelly CJ, Periz G, Simko EA, Shaw PG, Kim MS, Maragakis NJ, Troncoso JC, Pandey A, Sattler R, Rothstein JD, Wang J (2014) C9orf72 nucleotide repeat structures initiate molecular cascades of disease. Nature 507:195–200

Haltiner MM, Smale ST, Tjian R (1986) Two distinct promoter elements in the human rRNA gene identified by linker scanning mutagenesis. Mol Cell Biol 6:227–235

Heix J, Vente A, Voit R, Budde A, Michaelidis TM, Grummt I (1998) Mitotic silencing of human rRNA synthesis: inactivation of the promoter selectivity factor SL1 by cdc2/cyclin B-mediated phosphorylation. EMBO J 17:7373–7381

Heix J, Zomerdijk JC, Ravanpay A, Tjian R, Grummt I (1997) Cloning of murine RNA polymerase I-specific TAF factors: conserved interactions between the subunits of the species-specific transcription initiation factor TIF-IB/SL1. Proc Natl Acad Sci U S A 94:1733–1738

Heliot L, Kaplan H, Lucas L, Klein C, Beorchia A, Doco-Fenzy M, Menager M, Thiry M, O'Donohue MF, Ploton D (1997) Electron tomography of metaphase nucleolar organizer regions: evidence for a twisted-loop organization. Mol Biol Cell 8:2199–2216

Henderson AS, Warburton D, Atwood KC (1972) Location of ribosomal DNA in the human chromosome complement. Proc Natl Acad Sci U S A 69:3394–3398

Henras AK, Plisson-Chastang C, O'Donohue MF, Chakraborty A, Gleizes PE (2015) An overview of pre-ribosomal RNA processing in eukaryotes. Wiley Interdiscip Rev RNA 6:225–242

Hernandez-Verdun D (2011) Assembly and disassembly of the nucleolus during the cell cycle. Nucleus 2:189–194

Hernandez-Verdun D, Roussel P, Thiry M, Sirri V, Lafontaine DL (2010) The nucleolus: structure/function relationship in RNA metabolism. Wiley Interdiscip Rev RNA 1:415–431

Hernandez P, Martin-Parras L, Martinez-Robles ML, Schvartzman JB (1993) Conserved features in the mode of replication of eukaryotic ribosomal RNA genes. EMBO J 12:1475–1485

Hiscox JA (2002) The nucleolus—a gateway to viral infection? Arch Virol 147:1077–1089

Hiscox JA (2007) RNA viruses: hijacking the dynamic nucleolus. Nat Rev Microbiol 5:119–127

Hiscox JA, Wurm T, Wilson L, Britton P, Cavanagh D, Brooks G (2001) The coronavirus infectious bronchitis virus nucleoprotein localizes to the nucleolus. J Virol 75:506–512

Izumi RE, Valdez B, Banerjee R, Srivastava M, Dasgupta A (2001) Nucleolin stimulates viral internal ribosome entry site-mediated translation. Virus Res 76:17–29

Jacobson MR, Cao LG, Taneja K, Singer RH, Wang YL, Pederson T (1997) Nuclear domains of the RNA subunit of RNase P. J Cell Sci 110(Pt 7):829–837

Jacobson MR, Pederson T (1998) Localization of signal recognition particle RNA in the nucleolus of mammalian cells. Proc Natl Acad Sci U S A 95:7981–7986

Jin J, Li GJ, Davis J, Zhu D, Wang Y, Pan C, Zhang J (2007) Identification of novel proteins associated with both alpha-synuclein and DJ-1. Mol Cell Proteomics 6:845–859

Jordan P, Carmo-Fonseca M (1998) Cisplatin inhibits synthesis of ribosomal RNA in vivo. Nucleic Acids Res 26:2831–2836

Kar B, Liu B, Zhou Z, Lam YW (2011) Quantitative nucleolar proteomics reveals nuclear re-organization during stress-induced senescence in mouse fibroblast. BMC Cell Biol 12:33

Kermekchiev M, Workman JL, Pikaard CS (1997) Nucleosome binding by the polymerase I transactivator upstream binding factor displaces linker histone H1. Mol Cell Biol 17:5833–5842

Kinderman NB, LaVelle A (1976) Ultrastructural changes in the developing nucleolus following axotomy. Brain Res 108:237–247

Komili S, Farny NG, Roth FP, Silver PA (2007) Functional specificity among ribosomal proteins regulates gene expression. Cell 131:557–571

Kruhlak M, Crouch EE, Orlov M, Montano C, Gorski SA, Nussenzweig A, Misteli T, Phair RD, Casellas R (2007) The ATM repair pathway inhibits RNA polymerase I transcription in response to chromosome breaks. Nature 447:730–734

Kuhn A, Grummt I (1992) Dual role of the nucleolar transcription factor UBF: trans-activator and antirepressor. Proc Natl Acad Sci U S A 89:7340–7344

Kwon I, Xiang S, Kato M, Wu L, Theodoropoulos P, Wang T, Kim J, Yun J, Xie Y, McKnight SL (2014) Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. Science 345:1139–1145

Lam YW, Trinkle-Mulcahy L (2015) New insights into nucleolar structure and function. F1000Prime Rep 7:48

Langst G, Becker PB, Grummt I (1998) TTF-I determines the chromatin architecture of the active rDNA promoter. EMBO J 17:3135–3145

Learned RM, Learned TK, Haltiner MM, Tjian RT (1986) Human rRNA transcription is modulated by the coordinate binding of two factors to an upstream control element. Cell 45:847–857

Lee CH, Chang SC, Chen CJ, Chang MF (1998) The nucleolin binding activity of hepatitis delta antigen is associated with nucleolus targeting. J Biol Chem 273:7650–7656

Leung AK, Gerlich D, Miller G, Lyon C, Lam YW, Lleres D, Daigle N, Zomerdijk J, Ellenberg J, Lamond AI (2004) Quantitative kinetic analysis of nucleolar breakdown and reassembly during mitosis in live human cells. J Cell Biol 166:787–800

Li CF, Pontes O, El-Shami M, Henderson IR, Bernatavichute YV, Chan SW, Lagrange T, Pikaard CS, Jacobsen SE (2006a) An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in *Arabidopsis thaliana*. Cell 126:93–106

Li J, Langst G, Grummt I (2006b) NoRC-dependent nucleosome positioning silences rRNA genes. EMBO J 25:5735–5741

Li ZF, Liang YM, Lau PN, Shen W, Wang DK, Cheung WT, Xue CJ, Poon LM, Lam YW (2013) Dynamic localisation of mature microRNAs in Human nucleoli is influenced by exogenous genetic materials. PLoS One 8, e70869

Little RD, Platt TH, Schildkraut CL (1993) Initiation and termination of DNA replication in human rRNA genes. Mol Cell Biol 13:6600–6613

Ly T, Ahmad Y, Shlien A, Soroka D, Mills A, Emanuele MJ, Stratton MR, Lamond AI (2014) A proteomic chronology of gene expression through the cell cycle in human myeloid leukemia cells. Elife 3, e01630

Maehama T, Kawahara K, Nishio M, Suzuki A, Hanada K (2014) Nucleolar stress induces ubiquitination-independent proteasomal degradation of PICT1 protein. J Biol Chem 289:20802–20812

Mais C, Wright JE, Prieto JL, Raggett SL, McStay B (2005) UBF-binding site arrays form pseudo-NORs and sequester the RNA polymerase I transcription machinery. Genes Dev 19:50–64

Mann DM, Marcyniuk B, Yates PO, Neary D, Snowden JS (1988) The progression of the pathological changes of Alzheimer's disease in frontal and temporal neocortex examined both at biopsy and at autopsy. Neuropathol Appl Neurobiol 14:177–195

Mann DM, Yates PO (1982) Pathogenesis of Parkinson's disease. Arch Neurol 39:545–549
Markesbery WR, Lovell MA (2006) DNA oxidation in Alzheimer's disease. Antioxid Redox Signal 8:2039–2045
Matthews DA (2001) Adenovirus protein V induces redistribution of nucleolin and B23 from nucleolus to cytoplasm. J Virol 75:1031–1038
Mayer C, Neubert M, Grummt I (2008) The structure of NoRC-associated RNA is crucial for targeting the chromatin remodelling complex NoRC to the nucleolus. EMBO Rep 9:774–780
Mayer C, Schmitz KM, Li J, Grummt I, Santoro R (2006) Intergenic transcripts regulate the epigenetic state of rRNA genes. Mol Cell 22:351–361
Michienzi A, De Angelis FG, Bozzoni I, Rossi JJ (2006) A nucleolar localizing Rev binding element inhibits HIV replication. AIDS Res Ther 3:13
Miller G, Panov KI, Friedrich JK, Trinkle-Mulcahy L, Lamond AI, Zomerdijk JC (2001) hRRN3 is essential in the SL1-mediated recruitment of RNA Polymerase I to rRNA gene promoters. EMBO J 20:1373–1382
Miller OL Jr, Beatty BR (1969) Visualization of nucleolar genes. Science 164:955–957
Mongelard F, Bouvet P (2007) Nucleolin: a multiFACeTed protein. Trends Cell Biol 17:80–86
Moorefield B, Greene EA, Reeder RH (2000) RNA polymerase I transcription factor Rrn3 is functionally conserved between yeast and human. Proc Natl Acad Sci U S A 97:4724–4729
Moss T (2004) At the crossroads of growth control; making ribosomal RNA. Curr Opin Genet Dev 14:210–217
Moss T, Langlois F, Gagnon-Kugler T, Stefanovsky V (2007) A housekeeper with power of attorney: the rRNA genes in ribosome biogenesis. Cell Mol Life Sci 64:29–49
Nisole S, Krust B, Hovanessian AG (2002) Anchorage of HIV on permissive cells leads to coaggregation of viral particles with surface nucleolin at membrane raft microdomains. Exp Cell Res 276:155–173
O'Sullivan AC, Sullivan GJ, McStay B (2002) UBF binding in vivo is not restricted to regulatory sequences within the vertebrate ribosomal DNA repeat. Mol Cell Biol 22:657–668
Okuwaki M, Iwamatsu A, Tsujimoto M, Nagata K (2001) Identification of nucleophosmin/B23, an acidic nucleolar protein, as a stimulatory factor for in vitro replication of adenovirus DNA complexed with viral basic core proteins. J Mol Biol 311:41–55
Olson MO, Dundr M (2005) The moving parts of the nucleolus. Histochem Cell Biol 123:203–216
Orr HT, Zoghbi HY (2007) Trinucleotide repeat disorders. Annu Rev Neurosci 30:575–621
Parlato R, Kreiner G (2013) Nucleolar activity in neurodegenerative diseases: a missing piece of the puzzle? J Mol Med (Berl) 91:541–547
Pederson T (2011) The nucleolus. Cold Spring Harb Perspect Biol 3
Pederson T, Powell K (2015) Thoru Pederson: spotting novel roles for the nucleolus. J Cell Biol 208:384–385
Peltonen K, Colis L, Liu H, Jaamaa S, Zhang Z, Af Hallstrom T, Moore HM, Sirajuddin P, Laiho M (2014) Small molecule BMH-compounds that inhibit RNA polymerase I and cause nucleolar stress. Mol Cancer Ther 13:2537–2546
Pendle AF, Clark GP, Boon R, Lewandowska D, Lam YW, Andersen J, Mann M, Lamond AI, Brown JW, Shaw PJ (2005) Proteomic analysis of the Arabidopsis nucleolus suggests novel nucleolar functions. Mol Biol Cell 16:260–269
Pietrzak M, Rempala G, Nelson PT, Zheng JJ, Hetman M (2011) Epigenetic silencing of nucleolar rRNA genes in Alzheimer's disease. PLoS One 6, e22585
Politz JC, Zhang F, Pederson T (2006) MicroRNA-206 colocalizes with ribosome-rich regions in both the nucleolus and cytoplasm of rat myogenic cells. Proc Natl Acad Sci U S A 103:18957–18962
Pontes O, Li CF, Nunes PC, Haag J, Ream T, Vitins A, Jacobsen SE, Pikaard CS (2006) The Arabidopsis chromatin-modifying nuclear siRNA pathway involves a nucleolar RNA processing center. Cell 126:79–92

Preti M, O'Donohue MF, Montel-Lehry N, Bortolin-Cavaille ML, Choesmel V, Gleizes PE (2013) Gradual processing of the ITS1 from the nucleolus to the cytoplasm during synthesis of the human 18S rRNA. Nucleic Acids Res 41:4709–4723

Prieto JL, McStay B (2007) Recruitment of factors linking transcription and processing of pre-rRNA to NOR chromatin is UBF-dependent and occurs independent of transcription in human cells. Genes Dev 21:2041–2054

Prieto JL, McStay B (2008) Pseudo-NORs: a novel model for studying nucleoli. Biochim Biophys Acta 1783:2116–2123

Puvion-Dutilleul F, Christensen ME (1993) Alterations of fibrillarin distribution and nucleolar ultrastructure induced by adenovirus infection. Eur J Cell Biol 61:168–176

Raska I, Shaw PJ, Cmarko D (2006) Structure and function of the nucleolus in the spotlight. Curr Opin Cell Biol 18:325–334

Rieker C, Engblom D, Kreiner G, Domanskyi A, Schober A, Stotz S, Neumann M, Yuan X, Grummt I, Schutz G, Parlato R (2011) Nucleolar disruption in dopaminergic neurons leads to oxidative damage and parkinsonism through repression of mammalian target of rapamycin signaling. J Neurosci Off J Soc Neurosci 31:453–460

Roussel P, Andre C, Comai L, Hernandez-Verdun D (1996) The rDNA transcription machinery is assembled during mitosis in active NORs and absent in inactive NORs. J Cell Biol 133:235–246

Roussel P, Andre C, Masson C, Geraud G, Hernandez-Verdun D (1993) Localization of the RNA polymerase I transcription factor hUBF during the cell cycle. J Cell Sci 104(Pt 2):327–337

Rowland RR, Kervin R, Kuckleburg C, Sperlich A, Benfield DA (1999) The localization of porcine reproductive and respiratory syndrome virus nucleocapsid protein to the nucleolus of infected cells and identification of a potential nucleolar localization signal sequence. Virus Res 64:1–12

Rowland RR, Yoo D (2003) Nucleolar-cytoplasmic shuttling of PRRSV nucleocapsid protein: a simple case of molecular mimicry or the complex regulation by nuclear import, nucleolar localization and nuclear export signal sequences. Virus Res 95:23–33

Rubbi CP, Milner J (2003) Disruption of the nucleolus mediates stabilization of p53 in response to DNA damage and other stresses. EMBO J 22:6068–6077

Ruggero D, Pandolfi PP (2003) Does the ribosome translate cancer? Nat Rev Cancer 3:179–192

Russell J, Zomerdijk JC (2005) RNA-polymerase-I-directed rDNA transcription, life and works. Trends Biochem Sci 30:87–96

Sanij E, Hannan RD (2009) The role of UBF in regulating the structure and dynamics of transcriptionally active rDNA chromatin. Epigenetics 4:374–382

Sanij E, Poortinga G, Sharkey K, Hung S, Holloway TP, Quin J, Robb E, Wong LH, Thomas WG, Stefanovsky V, Moss T, Rothblum L, Hannan KM, McArthur GA, Pearson RB, Hannan RD (2008) UBF levels determine the number of active ribosomal RNA genes in mammals. J Cell Biol 183:1259–1274

Santoro R, Schmitz KM, Sandoval J, Grummt I (2010) Intergenic transcripts originating from a subclass of ribosomal DNA repeats silence ribosomal RNA genes in trans. EMBO Rep 11:52–58

Sasaki M, Kawahara K, Nishio M, Mimori K, Kogo R, Hamada K, Itoh B, Wang J, Komatsu Y, Yang YR, Hikasa H, Horie Y, Yamashita T, Kamijo T, Zhang Y, Zhu Y, Prives C, Nakano T, Mak TW, Sasaki T, Maehama T, Mori M, Suzuki A (2011) Regulation of the MDM2-P53 pathway and tumor growth by PICT1 via nucleolar RPL11. Nat Med 17:944–951

Scherl A, Coute Y, Deon C, Calle A, Kindbeiter K, Sanchez JC, Greco A, Hochstrasser D, Diaz JJ (2002) Functional proteomic analysis of human nucleolus. Mol Biol Cell 13:4100–4109

Schmidt EV (1999) The role of c-myc in cellular growth control. Oncogene 18:2988–2996

Sirri V, Roussel P, Hernandez-Verdun D (2000) In vivo release of mitotic silencing of ribosomal gene transcription does not give rise to precursor ribosomal RNA processing. J Cell Biol 148:259–270

Sirri V, Urcuqui-Inchima S, Roussel P, Hernandez-Verdun D (2008) Nucleolus: the fascinating nuclear body. Histochem Cell Biol 129:13–31

Sogo JM, Ness PJ, Widmer RM, Parish RW, Koller T (1984) Psoralen-crosslinking of DNA as a probe for the structure of active nucleolar chromatin. J Mol Biol 178:897–919
Stefanovsky V, Langlois F, Gagnon-Kugler T, Rothblum LI, Moss T (2006) Growth factor signaling regulates elongation of RNA polymerase I transcription in mammals via UBF phosphorylation and r-chromatin remodeling. Mol Cell 21:629–639
Strohner R, Nemeth A, Jansa P, Hofmann-Rohrer U, Santoro R, Langst G, Grummt I (2001) NoRC--a novel member of mammalian ISWI-containing chromatin remodeling machines. EMBO J 20:4892–4900
Sullivan GJ, Bridger JM, Cuthbert AP, Newbold RF, Bickmore WA, McStay B (2001) Human acrocentric chromosomes with transcriptionally silent nucleolar organizer regions associate with nucleoli. EMBO J 20:2867–2874
Sundqvist A, Liu G, Mirsaliotis A, Xirodimas DP (2009) Regulation of nucleolar signalling to p53 through NEDDylation of L11. EMBO Rep 10:1132–1139
Szebeni A, Mehrotra B, Baumann A, Adam SA, Wingfield PT, Olson MO (1997) Nucleolar protein B23 stimulates nuclear import of the HIV-1 Rev protein and NLS-conjugated albumin. Biochemistry 36:3941–3949
Tollervey D, Lehtonen H, Jansen R, Kern H, Hurt EC (1993) Temperature-sensitive mutations demonstrate roles for yeast fibrillarin in pre-rRNA processing, pre-rRNA methylation, and ribosome assembly. Cell 72:443–457
Tsai RY-L (2011) New frontiers in nucleolar research: nucleostemin and related proteins. In: Olson MOJ (ed) The nucleolus, protein reviews. Springer, New York, pp 301–320
Tsai RY, Pederson T (2014) Connecting the nucleolus to the cell cycle and human disease. FASEB J 28:3290–3296
Tschochner H, Hurt E (2003) Pre-ribosomes on the road from the nucleolus to the cytoplasm. Trends Cell Biol 13:255–263
Tsoi H, Lau TC, Tsang SY, Lau KF, Chan HY (2012) CAG expansion induces nucleolar stress in polyglutamine diseases. Proc Natl Acad Sci U S A 109:13428–13433
Vitali P, Basyuk E, Le Meur E, Bertrand E, Muscatelli F, Cavaille J, Huttenhofer A (2005) ADAR2-mediated editing of RNA substrates in the nucleolus is inhibited by C/D small nucleolar RNAs. J Cell Biol 169:745–753
Walter P, Blobel G (1982) Signal recognition particle contains a 7S RNA essential for protein translocation across the endoplasmic reticulum. Nature 299:691–698
Walter P, Johnson AE (1994) Signal sequence recognition and protein targeting to the endoplasmic reticulum membrane. Annu Rev Cell Biol 10:87–119
Warner JR (1999) The economics of ribosome biosynthesis in yeast. Trends Biochem Sci 24:437–440
Weber JD, Taylor LJ, Roussel MF, Sherr CJ, Bar-Sagi D (1999) Nucleolar Arf sequesters Mdm2 and activates p53. Nat Cell Biol 1:20–26
Wiesendanger B, Lucchini R, Koller T, Sogo JM (1994) Replication fork barriers in the Xenopus rDNA. Nucleic Acids Res 22:5038–5046
Wong JM, Kusdra L, Collins K (2002) Subnuclear shuttling of human telomerase induced by transformation and DNA damage. Nat Cell Biol 4:731–736
Worton RG, Sutherland J, Sylvester JE, Willard HF, Bodrug S, Dube I, Duff C, Kean V, Ray PN, Schmickel RD (1988) Human ribosomal RNA genes: orientation of the tandem array and conservation of the 5′ end. Science (New York, NY) 239:64–68
Yusupova G, Yusupov M (2014) High-resolution structure of the eukaryotic 80S ribosome. Annu Rev Biochem 83:467–486
Zhang Y, Wolf GW, Bhat K, Jin A, Allio T, Burkhart WA, Xiong Y (2003) Ribosomal protein L11 negatively regulates oncoprotein MDM2 and mediates a p53-dependent ribosomal-stress checkpoint pathway. Mol Cell Biol 23:8902–8912
Zomerdijk JC, Beckmann H, Comai L, Tjian R (1994) Assembly of transcriptionally active RNA polymerase I initiation factor SL1 from recombinant subunits. Science (New York, NY) 266:2015–2018

Pre-mRNA Splicing and Disease

Michael R. Ladomery and Sebastian Oltean

Abstract The splicing of pre-mRNA was discovered in the late 1970s. It soon became apparent that pre-mRNA can be *alternatively spliced*, providing eukaryotes with an important additional mechanism with which to regulate gene expression. Alternative splicing is prevalent and often highly complex. It occurs co-transcriptionally and can be influenced by the chromatin environment, promoter structure and speed of transcription. Splice isoforms of genes often have radically different biological properties; it is therefore very important to ensure that they are expressed in the appropriate ratio. In humans, more than 94 % of genes are alternatively spliced, and one in six disease causing mutations cause changes in splicing. Ongoing research has painted an increasingly complex picture of the nature of alternative splicing regulation. We start by describing the discovery of splicing, the splicing reaction, and the spatial organization of splicing. We discuss the nature and extent of alternative splicing, following by its regulation. We then describe several examples that clearly illustrate the involvement of alternative splicing in disease. Lastly we discuss ways in which alternative splicing can be manipulated, suggesting new therapies in the battle against cancer and other important diseases.

Keywords Pre-mRNA splicing • Splice sites • snRNPs • Spliceosome • Splice factors • Splice factor compartments • Alternative splicing • Splice factor kinases • Splicing mutations • Splicing therapeutics

M.R. Ladomery (✉)
Faculty of Health and Applied Sciences, University of the West of England, Frenchay Campus, Coldharbour Lane, Bristol BS16 1QY, UK
e-mail: Michael.Ladomery@uwe.ac.uk

S. Oltean (✉)
School of Physiology and Pharmacology, Dorothy Hodgkin Building, Whitson Street, Bristol BS1 3NY, UK
e-mail: Sebastian.Oltean@bristol.ac.uk

D.P. Bazett-Jones, G. Dellaire (eds.), *The Functional Nucleus*,
DOI 10.1007/978-3-319-38882-3_3

1 Introduction

1.1 The Discovery of Splicing

In the 1960s and 1970s researchers were interested in working out the physical structure and properties of eukaryotic genes. They were baffled by a number of intriguing observations. They were aware that the nucleus contained many species of RNA, including what was described as 'heterogenous nuclear RNA' (hnRNA). HnRNA, packaged by a complex assortment of RNA-binding proteins, could be seen in ultracentrifugation experiments to sediment in rather large ribonucleoprotein (RNP) particles. According to sedimentation rates hnRNP appeared to be significantly larger when compared to cytoplasmic mRNP by a factor of five. HnRNA shared two features with mRNA: a trimethylated cap at the 5′ end and a poly(A) tail at the 3′ end: it therefore appeared to be functionally related to mRNA. Furthermore, experiments aimed at measuring the kinetics of hybridization of nuclear RNA to cytoplasmic mRNA showed that hnRNA is much more kinetically complex. This means that hnRNA includes more sequence compared to mRNA, representing about 25 % of the genome. It was therefore clear that mRNA and hnRNA were related; but somehow hnRNA contains extra RNA sequence that is lost once it is exported into the cytoplasm as mRNA. We can now explain these observations: hnRNA is in fact the unspliced nuclear precursor of mRNA. Consequently, the term hnRNA is no longer used and has been supplanted by the term pre-mRNA (precursor messenger RNA).

Pre-mRNA splicing was first discovered in the context of experiments performed on adenoviruses. Adenoviruses were (and still are) very popular experimental models used to study gene expression because they infect human cell lines efficiently, generating large copy numbers of viral genomes and transcripts. The discovery of pre-mRNA splicing, in 1977, was attributed to two groups, led by Phil Sharp at the Massachussets Institute of Technology and Richard Roberts at Cold Spring Harbour Laboratory. They were interested in visualizing the hybridisation of RNA transcripts to genomic DNA through a technique known as R-looping. Specifically, they wanted to map adenoviral mRNAs onto the corresponding genes in the viral genome. They denatured viral DNA by using formamide at high temperature. RNA (in this case adenoviral messenger RNA) was then added and allowed to renature forming a DNA:RNA heteroduplex. The heteroduplex was visualized by electron microscopy. When they tried this with the mRNA corresponding to the adenoviral *hexon* gene (it encodes a viral coat protein), to their surprise, instead of perfect pairing of the two molecules, they saw a number of loops of DNA that were apparently unable to hybridise to the mRNA. The loops corresponded to parts of the *hexon* gene that were unable to hybridise to the mRNA (Fig. 1). Thus, what we now call introns and exons had been discovered (Berget et al. 1977; Chow et al. 1977). As a result of their pioneering work, Phil Sharp and Richard Roberts obtained a Nobel Prize in Physiology or Medicine in 2003 for their discovery of 'split genes'.

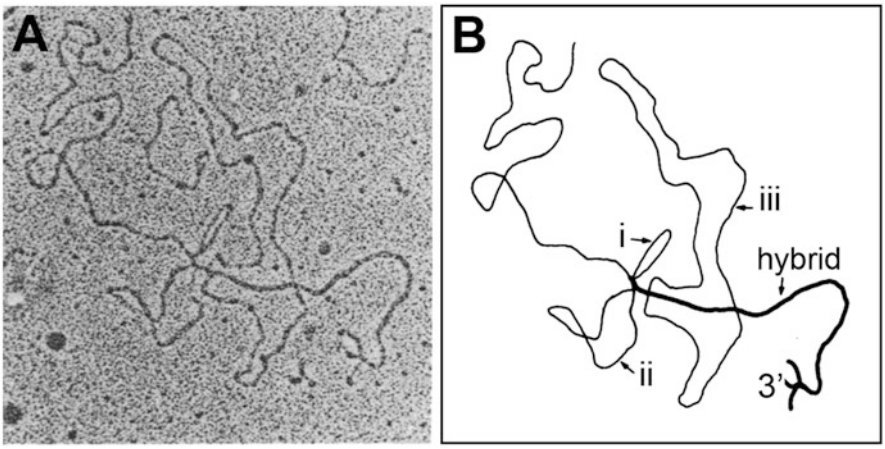

Fig. 1 The discovery of pre-mRNA splicing in the *hexon* gene. Messenger RNA derived from the adenoviral *hexon* gene was hybridized to the corresponding genomic DNA. The cartoon is an interpretation of the electron micrograph. The loops labeled i, ii and iii indicate genomic DNA that is not hybridized. The loops correspond to the introns of the *hexon* gene. Figure adapted from Berget et al. (1977), Fig. 4GH

Note that despite the prefix 'in', introns are in fact the parts of transcripts that are physically removed from the pre-mRNA—this is the process we call splicing. Instead, exons are precisely joined together to form mature mRNA. Be aware that there are other types of splicing (including self-splicing introns and tRNA splicing)—but in this chapter we will focus exclusively on pre-mRNA splicing.

At first glance, it is not clear why there is a to splice pre-mRNA. It is energetically costly, due to the complex machinery required, as we will see in the next section. However, the point is that exons can be skipped, lengthened or shortened, and introns retained—this is the process called *alternative splicing*. Alternative splicing is an increasingly important aspect of gene expression that greatly increases proteomic complexity. Its evolutionary benefit lies in the ability to generate a plethora of additional proteins from a relatively limited set of genes, throughout multicellular organisms and particularly in mammals (Keren et al. 2010).

1.2 The Splicing Reaction

The beginnings and ends of exons and introns are defined by splice sites. The 5′ splice site defines the end of an exon and the beginning of the intron; and the 3′ splice site defines the end of an intron and the beginning of the next exon. Splice sites are evolutionarily conserved; there is some variation in the detail, but the core sequence of splice sites is conserved indicating a common evolutionary origin

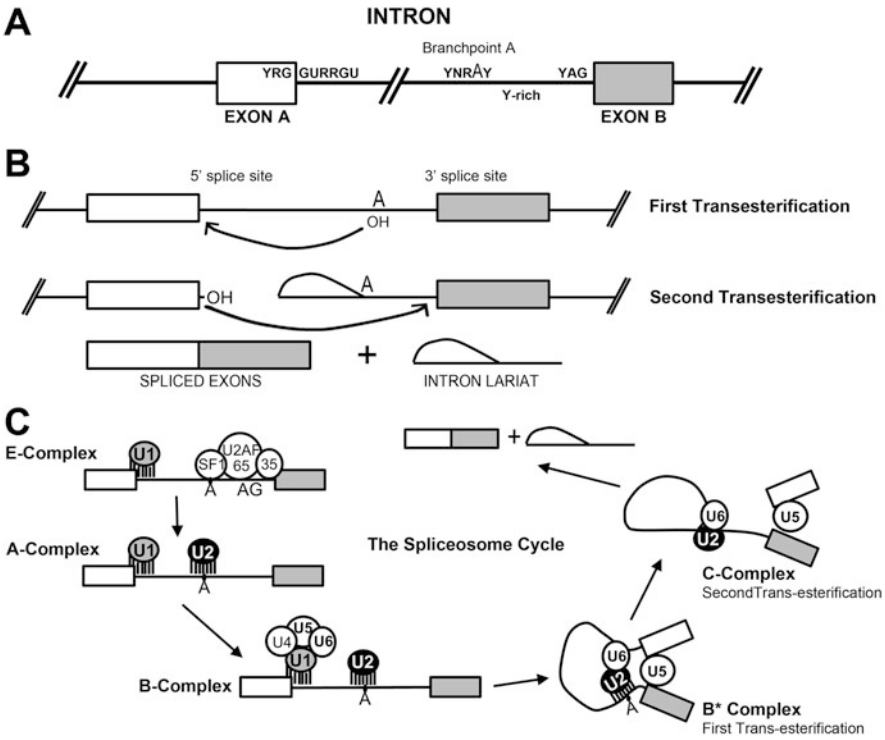

Fig. 2 Mechanisms and conservation of pre-mRNA splicing. (**a**) Exon sequences are indicated by *boxes* and introns by *lines*. The 5′ splice site is also known as the splice donor site, and the 3′ splice site as the splice acceptor site. Conserved sequences are indicated (R, purine; Y, pyrimidine; N, any nucleotide). These splice sites are recognized by the major spliceosome (following the 'GU-AG' rule). The typical size of an exon is around 150–300 nt, whereas introns generally range from 100 to 10,000 nt. The branchpoint A is generally about 20–40 nt upstream of the 3′ splice site. (**b**) The splicing reaction consists of two transesterification reactions. In the first transesterification the free 2′-OH of the branchpoint A nucleophilically attacks the 5′-splice site, detaching the upstream exon. In the second transesterification, the free hydroxyl at the 5′-splice site attacks the 3′-splice site, joining the exons and releasing the circularized intron lariat. (**c**) The spliceosome catalyses the splicing reactions. In the early (E) complex, U1 snRNP base pairs across the 5′-splice site; meanwhile the splice factor SF1 binds to the branchpoint adenosine (A) site, U2AF65 to the pyrimidine tract, and U2AF35 to the AG at the end of the intron. U2 snRNP then base pairs across the branchpoint A, making it bulge out (A complex). The tri-snurposome (U4/U5•U6) then joins the U1 snRNP (B complex). This is followed by a series of rearrangements in which the U5 snRNP is key because it contacts both the upstream and downstream exon, bringing them together to facilitate the first transesterification (B* complex). The second transesterification takes place in the C complex. The intron lariat is degraded and the snRNPs are released ready for further rounds of splicing (the spliceosome cycle)

(Fig. 2a). Another conserved sequence surrounds the 'branchpoint adenosine (A)' towards the 3′ end of introns; it is followed by a pyrimidine-rich tract. The splicing reaction consists of two transesterification reactions. In the first transesterification, the branchpoint A attacks the 5′ splice site, detaching the upstream exon. In the

second transesterification, the free 3′ end of the upstream exon attacks the 3′ splice site, forming a covalent bond that joins the exons together. The circularized intron, at this stage known as the "lariat", is released and degraded (in other words, recycled) in the nucleus (Fig. 2b, c). The 5′ splice site is also known as the 'splice donor site' and the 3′ splice site as the 'splice acceptor site'.

The precise joining of exons is by no means a trivial molecular event. The spliceosome, an assembly of ribonucleoprotein complexes (RNP) has evolved to catalyse the splicing reaction. Depending on the stage of assembly, the 60S spliceosome is 40–60 nm in diameter and is up to 5MDa in mass (Azubel et al. 2004). The core components of the spliceosome are the snRNPs (pronounced "snurps"); they consist of small nuclear RNAs (snRNAs) and several proteins. Some proteins are common to all snRNPs (core proteins); whereas others are snRNP-specific. The snRNPs start the splicing reactions by defining the location of the splice sites. Early in spliceosome assembly, the U1 snRNP base-pairs across the 5′ splice site. The splice factor U2AF65 (U2 auxiliary factor) binds to the pyrimidine-rich tract and facilitates the recruitment of U2 snRNP. U2 snRNP base-pairs across the intronic branchpoint A, causing the branchpoint A to bulge out (this facilitates the first transesterification reaction). Next, three other snRNPs join, initially bound together in the U4/U5•U6 tri-snurposome. The U6 snRNP interacts with the U2 snRNP and the 5′ splice site, and the U1 and U4 snRNPs are then released. The U5 snRNP is key to the splicing reaction because it pairs with both exons; it contains RNA helicase activity that also facilitates the splicing reaction. This is now the catalytically active spliceosome, and the first transesterification starts with the nucleophilic attack by the free 2′-OH of the branchpoint A. As well as these five snRNPs, several additional splice factors help assemble the spliceosome and facilitate the two transesterification reactions. The process of formation and disassembly of the spliceosome is known as the spliceosome cycle (Fig. 2c).

It is worth noting briefly that there is additional complexity in the machinery of splicing. Most introns begin with GU and end with AG: this is known as the 'GU-AG' rule. However, in a minority of cases, introns follow the 'AU-AC' rule because they begin with AU and end in AG. Introns that carry these sequences are spliced out in a similar way, albeit more slowly, but by the *minor spliceosome*. The minor spliceosome includes a different set of snRNPs (U11, U12, U4atac and U6atac snRNPs); however, the U5 snRNP remains the same.

1.3 The Splicing Code

Considering the complexity of the sequences present in pre-mRNAs that are often tens or even hundreds of thousands of bases long, it is quite remarkable that splicing can occur so precisely (Black 1995). Furthermore, pre-mRNA is packaged by proteins, and, like all RNA, forms complex secondary structures due to the tendency of single-stranded RNA to base-pair. Secondary structure can in fact

influence the recognition of splice sites (Hiller et al. 2007). Somehow, in this morass of ribonucleoprotein, the splicing machinery has to find the correct splice sites. This is by no means a trivial problem! What is clear, is that ultimately specific sequences, present in the pre-mRNA, define where splicing will occur. This is known as the *splicing code* (Fu 2004; Wang and Burge 2008).

There are two models to explain how the spliceosome determines where splicing will occur. According to the intron definition model, splice factors and the U1 and U2 snRNPs assemble across introns. This is followed by the assembly of the complete spliceosome. On the other hand, the exon definition model states that the first step is to define exons. In exon definition both the 5′ splice sites (bound by U1 snRNP) and the 3′ splice site (defined by the splice factor U2AF65 binding to the pyrimidine tract and the U2 snRNP across the branchpoint A must first be defined, before the spliceosome can proceed. The advantage of exon definition is that it ensures that isolated sequences that might be similar to 5′ or 3′ splice sites but that are not part of a legitimate exon are not recognized. It turns out that the exon definition model applies particularly in species, including humans, in which exons are on average a lot smaller than introns.

1.4 Co-transcriptional Nature of Splicing

As RNA polymerase II moves along a gene transcribing it, a nascent pre-mRNA is formed (the transcript). By definition, the pre-mRNA must remain attached to the gene until transcription terminates. This raises the question: is splicing cotranscriptional or posttranscriptional? Originally it was thought that splicing is posttranscriptional. This means that the spliceosomes only assemble on the transcripts once they have become physically detached from the gene, presumably after the RNA polymerase has finished the job. A second, alternative model is that splice factors do bind the nascent transcript, and initiate and finish splicing even before transcription is completed. A third model is called co-transcriptional commitment, in which splice factors and snRNPs bind to the definitive splice sites committing their use after transcription occurs.

This question was answered through a very elegant experiment conducted on the *dystrophin* gene (Tennyson et al. 1995). Dystrophin is a protein that connects the cytoskeleton of muscle cells to the extracellular matrix. The statistics of the gene are quite striking. It is one of the longest human genes, about 2.5 Mbp in total. The pre-mRNA takes a full 16 h to transcribe—that is because it is 2.4 Mb long. It contains a whopping 79 exons and the resulting mRNA is 14 kb long (well above average). Cultured muscle precursor cells (which don't express *dystrophin*) were made to differentiate into muscle cells (which do express it). At this point the first wave of transcription started, and from then it was possible to track the formation of spliced products. Spliced exons at the 5′ end of the transcript could be detected 12 h before the transcription of exon 68 had occurred. This experiment demonstrated quite clearly the cotranscriptional nature of pre-mRNA splicing.

1.5 Splicing Factor Compartments

Several interconnected processes occur in the nucleus, and pre-mRNA splicing is no exception. These processes involve a very complex array of proteins and RNA molecules that need to be localized, stored, and released to sites of active transcription where they are most needed. It seems unlikely that splicing should rely on random collisions of hundreds of proteins that are uniformly distributed in the nucleus. Researchers have examined the physical location of splice factors in the eukaryotic nucleus. It turns out that the nucleus contains splicing factor compartments (SFCs), often referred to as 'nuclear speckles' due to their appearance under the microscope. These SFCs can be visualized with antibodies against specific splice factors by using immunofluorescence microscopy and immunogold electron microscopy. SFCs are thought to be held together by protein interactions. It is estimated that the SFCs can occupy about 20 % of the nuclear volume, a staggering proportion. What do they contain? As well as splice factors, SFCs also contain proteins that are involved in transcription and other aspects of pre-mRNA processing (such as 3′ end formation, cleavage and polyadenylation). When examined by electron microscopy, the SFCs are described as interchromatin granule clusters. These are thought to be the sites of storage of splice factors. However, SFCs are also adjacent to sites of transcription. Specifically they associate with perichromatin fibrils, sites where nascent transcripts are accumulating from actively transcribed genes. The implication is that the SFCs act as reservoirs that can be called to deliver a high concentration of splice factors to sites of active transcription (Misteli 2000; Lamond and Spector 2003).

In most cases, the localization of splice factors is observed in fixed (therefore dead) cells. This only provides a static picture of localization. On the other hand, it is possible to express splice factors fused to green fluorescent protein (GFP), which allows their diffusion to be monitored by photobleaching and recovery in living cells, a technique known as fluorescence recovery after photobleaching (FRAP). Using FRAP, remarkable insights regarding the dynamics of SFCs has been uncovered. For example, the well-studied splice factor, SRSF1 (also known as ASF/SF2) is estimated to spend on average 45 s in an SFC, and that more than 10,000 SRSF1 molecules depart SFCs *each second* per nucleus! Thus it seems that splice factors move around the nucleus rapidly, and are recruited from SFCs to sites of active transcription (Misteli 2008). What then regulates the release of splice factors such as SRSF1? The activity and localization of several splice factors can be regulated by reversible phosphorylation (Stamm 2008). For example, the splice factor kinase CLK1 hyperphosphorylates SRSF1 in SFCs, facilitating its release from these subnuclear domains.

Note that components of the pre-mRNA splicing machinery are not only detected in SFCs. SnRNPs can also be detected in Cajal bodies (Stanek et al. 2008). Cajal bodies are spherical structures found in the nuclei of diving cells, generally associated with nucleoli. They have several functions; and in the

context of splicing they contribute to the biogenesis of mature snRNPs and the reassembly of the U4/U5•U6 trisnurposome.

2 Extent and Regulation of Alternative Splicing

2.1 First Discoveries of Alternative Splicing

The splicing mechanism described at the beginning of this chapter is commonly referred to as 'constitutive splicing', in other words a process that occurs in all multi-exon genes. However, not long after the characterization of this process, several observations suggested that exons are not always joined together in a predictable fashion and a multi-exon gene may give rise to several mature RNA transcripts, a process now referred to as 'alternative splicing'.

For example analysis of 5′ termini of cytoplasmic mRNA species from polyoma virus (following infection of mouse cells) showed that at least seven species of different length are present, that they are encoded in the viral genome not the host and that they occur through an as yet uncharacterized posttranscriptional process, possibly via splicing (Flavell et al. 1979). Also, in one of the first examples of analysis of alternatively spliced variants that today is considered trivial, De Noto et al. compared the gene sequence of human growth hormone (hGH) with cloned cDNA species. They found that several variants exist resulting from 3′ alternative splice sites that give rise to different polypeptides and therefore increase diversity in the function of the hGH (DeNoto et al. 1981).

2.2 Modes of Alternative Splicing

It is clear that alternative splicing may occur through several molecular mechanisms (see Fig. 3). The main categories are: (1) cassette exon—where a certain exon is either included or excluded from the transcript; (2) mutually exclusive exons, where two neighbouring exons alternate in their inclusion in the mRNA depending on cell type or various conditions; (3) alternative 5′ or 3′ splice sites, when additional splice sites may be present in exons; (4) retained introns, when an intron is not removed in the transcript. Note that alternative splicing can affect all parts of an mRNA: not just the open reading frame (changing the nature of the encoded protein) but also the 5′ and 3′ UTRs (untranslated regions) affecting mRNA translation, localization, and stability.

Additional subclasses of alternative splicing have been described including: alternative poly(A) sites, when several functional poly(A) sites may be present usually in introns and differential usage results in transcripts of various length at the 3′ termini; and alternative promoters and first exon usage, although these are not a

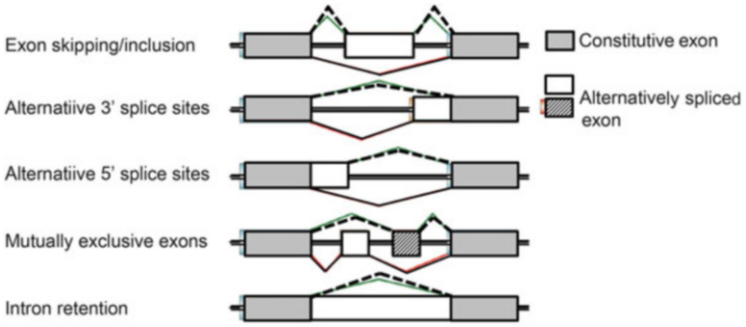

Fig. 3 Main modes of alternative splicing. Constitutive exons are not alternatively spliced (*light shading*). In alternative splicing, exons can be entirely skipped (cassette exons). The precise boundaries of exons can also change in the presence of alternative 5′ or 3′ splice sites. Another mode of alternative splicing is mutually exclusive exons in which only one or the other is included, but never both. Lastly, introns can be retained. Adapted from Matlin et al. (2005)

purely alternative splicing mechanisms but result from the coordinated regulation of transcription and splicing. Added to these additional categories there are rarer mechanisms such as cryptic splice sites; splice sites that exist in exons and introns are very seldom used and only in particular conditions.

2.3 Regulation of Alternative Splicing Through Splicing Enhancers and Silencers

At the level of RNA the regulation of alternative use of splice sites is accomplished by both trans- and cis-acting factors. Trans-acting factors are splice factors that are usually involved in constitutive splicing as well. Some of the most important groups of splice factors involved in alternative splicing are serine-arginine rich proteins (SR-proteins) and heteronuclear ribonuclear proteins (hnRNPs).

Splice factors bind cis-acting regions of the pre-mRNA that can be situated both in exons and introns. Upon binding, these factors act to repress or activate a certain splice and therefore they are called ‘enhancers’ or ‘silencers’ (i.e. exonic splicing enhance/silencers and intronic splicing enhancer/silencers) (Fig. 4). The final fate of an exon to be included or skipped in the mature transcript depends on the combined action of enhancers and silencers and their effect is dependent on the concentration of the trans-acting factors able to bind them.

In general SR-proteins are considered activators while hnRNPs usually function as repressors. However, depending on the local context of a particular splice site and the contribution of regulatory sequences and factors that bind them roles may be reversed. For example, an SR-protein may act as a repressor such as SRSF1, which can repress adenoviral transcription by binding to a region close to the branch point in the adenovirus transcription I1 unit (Kanopka et al. 1998).

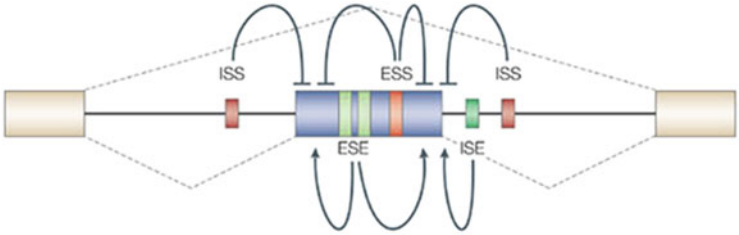

Fig. 4 Regulatory elements in exons or introns may activate or repress the inclusion of an exon. In the same way as promoter elements help regulate transcription, sequences in pre-mRNA regulate splicing. Regulatory sequences can either enhance or repress (silence) the inclusion of an exon or use of a particular splice site; they are located both within exons and introns. *ESE* exonic splice enhancer, *ESS* exonic splice silencer, *ISE* intronic splice enhancer, *ISS* intronic splice silencer. Adapted from Matlin et al. (2005)

2.4 *Effects of Transcription, Promoters and Chromatin on Alternative Splicing*

While our tendency to simplify things for a better understanding of the molecular universe of the cell has viewed initially that various gene regulation processes (e.g. transcription, RNA processing, translation) happen somehow separate in time and space it has become clear in the last 10 years that they are closely linked and interconnected. Several aspects of these processes take place concomitantly, have common protein regulators and influence each other.

As discussed in Sect. 2.4, the nascent RNA that is transcribed by RNA polymerases is at the same time being processed by splice factors that are recruited co-transcriptionally. Promoter structure may recruit splice factors to the transcription machinery and therefore affect splicing outcomes. Transcription elongation rate also affects alternative splicing; a fast elongation rate and RNA formation may not allow for certain factors to bind to splice sites and result in skipping of an exon while a slow elongation rate may result in the opposite effect (Braunschweig et al. 2013). Chromatin organization and nucleosome positioning has been found to influence alternative splicing with certain correlations observed in genome-wide studies. For instance exons that are more frequently included in alternative splicing are more highly enriched in nucleosomes; and nucleosome density varies along with the splice site strength (Luco et al. 2011).

Finally, epigenetic modification may impact on alternative splicing outcomes. Several mechanisms have been described. For instance if a certain histone methylation pattern exists an adaptor protein may bind to a particular methylation site. This in turn recruits a particular splice factor and therefore influence alternative splicing by providing the necessary factor to bind a certain exonic splicing enhancer (Luco et al. 2011).

2.5 *Integration of Splicing into Cell Signaling and Regulatory Networks*

It is considered common knowledge that transcription factors are integrated within signaling pathways in the cell and their activation (or repression) with resulting up-regulation (or down-regulation) of gene expression, represents one of the most important ways in which cells respond to changes and signals from the outside environment. However, one must also keep in mind that splice factors are also integrated in signaling pathways, and their activity is heavily regulated. As such, their regulation has a very similar outcome as the regulation of transcription factors (i.e. changes in the protein repertoire occur, and therefore cell functions).

In the usual paradigm, splice factors are inactive in the cytoplasm, they are phosphorylated by protein kinases, become active and shuttle (often chaperoned) into the nucleus where they affect outcomes of alternative splicing. While several protein kinases are not very specific and affect many intracellular processes and functions, several distinct examples have emerged in the last years as being more frequently involved in splicing regulation than other processes—e.g SR-protein kinases (SRPKs) or cyclin-dependent like kinases (CLKs) (Zhou et al. 2012; Naro and Sette 2013). While examples of virtually all canonical signaling pathways may be found to be involved in splice factor regulation and isoform choice (Oltean and Bates 2014) a recent report from Fu and colleagues in San Diego has demonstrated a central role played by the EGFR/PI3kinase/Akt signalling axis in regulating SRPK1 and SR-protein phosphorylation, and thus having a rather global effect on alternative splicing modulation (Zhou et al. 2012).

2.6 *Extent of Alternative Splicing Genome-Wide and Conservation Across Species*

While alternative splicing has been described more than 30 years ago and heavily studied mechanistically, it has not been clear until recently how extensive it is genome-wide or how conserved splicing machinery is across phyla. Two recent studies using genome-wide RNAseq have demonstrated that more than 94 % of genes are alternatively spliced in humans highlighting an immense combinatorial power of ~22,000 genes (each with 7–8 exons on average) to generate incredible protein diversity (Pan et al. 2008; Wang et al. 2008).

With the growing list of fully sequenced genomes for a plethora of plant, fungal, vertebrate and invertebrate species, it has been clear that differences in species phenotypes cannot be accounted for by the gene repertoire alone. Moreover, mRNA levels of different genes measured in various organs and tissues do not vary much during evolution of vertebrates, making unlikely that the level of gene expression is a major driver in species diversity. It has been long postulated, and partially proven by others, that alternative splicing is the main engine for species and organ-specific

phenotypic differences. In a recent paper, Blencowe and coworkers (Barbosa-Morais et al. 2012) have analyzed high-throughput RNAseq data from several organs across several species and demonstrated that the further an organism is in evolutionary terms from humans and primates the less complex the alternative splicing repertoire. In vertebrates, the conservation of alternative splicing in neural tissues is higher in comparison to other organs highlighting its importance in the evolution and development of the vertebrate nervous system. Also, another important conclusion was that alternative splicing patterns that are species-specific are coded mainly by the sequences rather than by the protein or factors in the cell that act on them; this was demonstrated by using a partial human chromosome in mouse cells that human sequences splice to human splice variants, even when expressed in mouse cells (Barbosa-Morais et al. 2012). These findings have very important implications when studying for instance alternative splicing events associated with human disease in transgenic mice using constructs specific for the human sequence.

3 Pre-mRNA Splicing and Disease

3.1 Nature and Frequency of Splicing Mutations

While classically mutations in coding regions of genes have been studied in detail due to the direct effect on encoded proteins and phenotype, mutations in non-coding regions of DNA may be equally important. From a splicing perspective mutations present in splice sites or any regulatory intronic or exonic regions may affect the outcome of a particular splicing event, result in changes in the protein sequence or length, change in the composition of a particular protein splice isoforms or lack of the protein due to the RNA being targeted for degradation (for example through nonsense-mediated decay).

Indeed, while the exact numbers vary among studies due to analysis methods and databases used, between 10 and 15 % of pathogenic mutations involve splicing-related elements (Lewandowska 2013; Gamazon and Stranger 2014; Hartmann et al. 2008). Mutation at the splice donor or acceptor site are the most common while mutations in the polypyrimidine tract or the branch point are the most rare.

A comprehensive analysis of mutations affecting splice factors or proteins that are components of the spliceosomal machinery is lacking. However, it is worth mentioning recent published studies highlighting a high frequency of mutations in genes involved in splicing in some hematological malignancies including myelodisplastic syndromes. Mutations in the splice factor *SF3B1* are seen in 20 % of cases, while in other types of myeloid syndromes such as acute or chronic myeloid leukemia in less than 5 % of cases (Ogawa 2012).

3.2 Examples of Aberrant Splicing in Cancer

Among various diseases, cancer is by far the most researched in respect to how faulty splicing contributes to pathogenesis. There is a myriad of alternative splicing isoforms reported to be associated (with various degrees of causality) with different aspects of the oncogenic process (Oltean and Bates 2014; Ladomery 2013; Ghigna et al. 2013). We will discuss a few examples that may form various paradigms for how aberrant splicing contributes to cancer.

One of the several hallmarks of cancer (Hanahan and Weinberg 2011) and essential property for tumours to be able to grow and survive is the ability to form new vessels through angiogenesis. It has been known for several years that there are two classes of molecules in the organism with opposing actions—pro- and anti-angiogenic—that regulate a fine balance for maintaining adequate rates for vessel formation from embryonic development to the adult life. These opposing actions may be achieved through de-regulation of alternative splicing as well. More than 10 years ago, an alternative splicing event in the vascular endothelial growth factor (VEGF)—a major regulator of angiogenesis—had been described that resulted in the $VEGF_{xxx}b$ family of isoforms from an 3′ alternative splice site in the terminal exon (Bates et al. 2002). At the protein level this splicing event translates into a different sequence of amino-acids at the C-terminus, which confers antagonistic properties and is therefore anti-angiogenic. There are numerous examples of cancers with decreased proportion of the anti-angiogenic VEGF isoforms or studies showing that reversal of pro-to anti-angiogenic VEGF isoforms results in decrease in tumour growth in animal models (Harper and Bates 2008).

Another essential property of cancer cells is the ability to migrate at distant sites and metastasize. This is achieved through several mechanisms, including epithelial-to-mesenchymal (EMT) transition; a plastic property of cells to switch between epithelial, static phenotypes to mesenchymal, migratory ones. Splicing may be again hijacked to nurture these properties as in the example of two mutually exclusive exons in fibroblast growth factor receptor 2 (FGFR2)- exon IIIb (or 8) found almost exclusively in epithelial cells and IIIc (or 9) found in mesenchymal cells. The alternate exons change the coding regions in the extracellular domain of the receptor, making it responsive to different FGF ligands. This results in different levels of activation of the receptor and therefore differential modulation of signaling pathway and cellular properties. Again, numerous studies have shown that the mesenchymal IIIc exon is associated with aggressive cancers (Carstens et al. 1997) or that overexpression of the epithelial IIIb isoform decreases tumour growth in animal models (Shoji et al. 2014; Yasumoto et al. 2004).

Apoptosis (programmed cell death) is another important hallmark of cancer; a process through which developmentally superfluous or irreparably mutated cells are physically eliminated. Deregulation of the signals that control apoptosis can contribute to the process of carcinogenesis. Apoptosis occurs through two pathways: the extrinsic and intrinsic pathway, the latter mediated through mitochondria. In the intrinsic pathway, free cytochrome c is released into the cytosol, resulting in

accumulation of a complex known as the apoptosome. The apoptosome in turn activates a series of proteases known as caspases. The executor caspases 3, 6 and 7 activate apoptosis. What is most striking is that several of the genes that encode the machinery of apoptosis are alternatively spliced into pro- and antiapoptotic (prosurvival) isoforms (Miura et al. 2012). The apoptosome itself contains two proteins, APAF1 and caspase 9, that include alternatively spliced isoforms with radically different properties. APAF1 (apoptosis-activating factor 1) expresses a pro-apoptotic splice isoforms (the full length APAF1-XL) and alternative splice isoforms (including APAF1-ALT) that impede the induction of apoptosis (Ogawa et al. 2003). When it is predominantly expressed, the pro-apoptotic isoform of APAF-1 activates procaspase 9. However caspase 9 is also alternatively spliced. A series of cassette exons can be skipped in the caspase 9 pre-mRNA, resulting in the expression of a shorter isoform known as caspase 9b, a dominant negative anti-apoptotic splice variant. The regulation of caspase 9 alternative splicing has been examined. The splice factor SRSF1 contributes to its regulation by binding to an intronic splice enhancer (Shultz et al. 2011). The same study shows that the effect of SRSF1 on caspase 9 splicing can even affect the chemotherapeutic sensitivity of non-small cell lung cancer cells. This suggests that it might be desirable to design drugs that can alter alternative splicing patterns by targeting SRSF1 or the protein kinases that regulate SRSF1 activity (SRPK1 and CLK1). Such drugs could be used in conjunction with established chemotherapies.

These examples clearly illustrate how each of the hallmarks of cancer are powerfully affected by alternative splicing. In many ways, dysregulated alternative splicing might itself be considered a hallmark of cancer (Ladomery 2013). Accordingly, one would predict that in cancer the machinery that regulates alternative splicing goes completely astray, much as the stability of the genome and the regulation of cell signaling goes out of control. Metastatic cancer is thought to arise after a series of gene mutations accumulate—a notable example being the multistage carcinogenesis model of colorectal cancer (the adenoma-carcinoma sequence). In this model a series of lesions occur; starting for example, with loss of function of the *APC* tumour suppressor gene, followed by activation of oncogenic *Ras*, and inactivation of the tumour suppressor gene *TP53*. Note that *TP53*, the most widely studied tumour suppressor, expresses a bewildering array of splice isoforms, greatly complicating its function (Surget et al. 2013). Each of these genes is alternatively spliced in such a way as to express isoforms that could, if overexpressed, potentially contribute to the progression of cancer (Ladomery 2013). Hence the range of mutations that could result in colorectal cancer could include mutations that affect splice sites of these genes, or mutations that affect the expression and activity of key splice factors.

3.3 Examples of Aberrant Splicing in Other Diseases

Knowing the almost global extent of alternative splicing across the human genome it is not surprising that the number of studies describing aberrant splicing in a large variety of diseases is continuously increasing. Among the most dramatically affected by aberrant splicing are several neurological disorders. For example, myotonic dystrophy, one of the most common muscular dystrophies in adults, is characterized by abnormal CUG or CUGC repeats in the mRNA of the *DMPK* gene. These repeats sequester a splicing factor known as muscle-blind-like-1 (MBNL-1), and therefore affect alternative splicing of various other transcripts targeted by MBNL-1 (Fugier et al. 2011). Spinal muscular atrophy (SMA) is an autosomal recessive disease and the most common genetic cause of mortality in infants. It is due to mutations or deletions that affect the *SMN1* gene, which encodes a protein with crucial functions in RNA metabolism called SMN. There is a nearly identical gene to *SMN1*, called *SMN2*, that cannot compensate for the protein because of a silent mutation in exon 7 that provokes exon skipping resulting in a truncated non-functional protein (Naryshkin et al. 2014). Several reports have identified at least six genes that have aberrant splicing and are thought to be involved in pathogenesis of Parkinson's disease (Fu et al. 2013).

Many genes involved in the pathogenesis of diabetes have been reported to express splice isoforms associated with advanced diabetes: for example a splice variant of the soluble receptor for advanced glycation end products (sRAGE) has been reported to be associated with the severity of diabetic nephropathy (Gohda et al. 2008); and T-cell factor 7- like 2 (TCF7L2) splice variants affect beta-cell function (Le Bacquer et al. 2011). There is in vivo data from animal studies that particular splice isoforms are causal to diabetes; for example soluble cytotoxic T-lymphocyte-associated protein 4 (sCTLA-4) is associated with acceleration of progression of type 1 diabetes (when knocked down in a mouse model) (Gerold et al. 2011).

The ever-increasing spectrum of diseases associated with aberrant splicing includes cardiovascular diseases. For example, dilated cardiomyopathy has been reported to be associated with aberrant splicing isoforms of titin, a protein important for the structure of the sarcomeres in striated muscles. The cause of faulty titin splicing is mutations in RBM20, an RNA-binding protein that works as a master regulator of alternative splicing in cardiac muscle (Guo et al. 2012).

3.4 Manipulating Splicing in Therapy

As there is a growing number of chronic diseases associated with aberrant pre-mRNA splicing, including diabetes and cancer, splicing has emerged as an exciting new therapeutic target. Several strategies have been developed to try to reverse faulty splicing for therapeutic purposes. The technique most employed so

far is the use of anti-sense oligos (ASO), also called splicing-switching oligos (SSOs). The general principle is to design ASOs that bind to regulatory sequences; if they bind to splice sites or enhancer elements they will inhibit inclusion of an exon; whereas if they bind splicing silencers they will activate exon inclusion. So far SSOs have been proved very promising, with several of them in clinical trials, including ASO-based treatments for Duchenne muscular dystrophy (DMD) and SMA (Singh and Cooper 2012).

More recently a specialized bifunctional targeting oligonucleotide has been designed, called targeted oligonucleotide enhancers of splicing (TOES). the typical TOES molecule has two parts: the first is the annealing ASO that targets the splice-site, and the second is the 'tail' oligo, which is not complementary to the target mRNA but instead serves to bind and recruit trans-factors to allow their local accumulation, and thus promote the correct splicing event. This strategy has also proved to be effective in correcting the splicing defect in SMA (Owen et al. 2011).

Use of trans-splicing has also been developed as a therapeutic tool—so-called SMaRT method (spliceosome mediated RNA trans-splicing). The principle is to create a hybrid RNA resulting from splicing between a 5′ splice site of an endogenous RNA and 3′ splice site of an artificial construct that provides a corrected exon (Rodriguez-Martin et al. 2009).

Finally, there is a growing number of small molecule inhibitors that have been shown to affect splicing. An interesting example is amiloride. This drug is a well-known diuretic used to regulate the ion channels within the renal tubules of the kidney. However, it was identified in a screen of small molecules that amiloride potently affects splicing of several genes involved in apoptosis and moreover to be able to decrease tumour growth in animal models (Ding et al. 2012). Recently, a class of small molecule compounds called SPHINX, were shown to inhibit SRPK1, the major kinase responsible for SR-protein phosphorylation, which in turn inhibited VEGF splicing and angiogenesis in a model of ocular neovascularization (Gammons et al. 2013) as well as melanoma xenografts growth (Gammons et al. 2014).

4 Summary

Alternative splicing of pre-mRNA has emerged as an absolutely key process in gene expression and cellular homeostasis. Even now, novel splice isoforms of human genes are still being discovered and characterized. Throughout eukaryotes, evolutionarily conserved mechanisms underlying alternative splicing allow genomes to express a diverse variety of proteins from a relatively limited number of genes, often having radically different biological properties. It is then not surprising that dysregulated splicing results in disease. Aberrant alternative splicing is now associated with cancer, neurological disorders, and diabetes. A better understanding of alternative splicing will, without a doubt, lead to novel treatment avenues.

Acknowledgments We would like to acknowledge Dr. Graham Dellaire (Dalhousie University, Halifax, Canada) for his aid in editing and preparing the figures in this chapter. This work was supported by operating grants from BBSRC (BB/J007293/1) to Sebastian Oltean and The Bristol Urological Institute (BUI256) to Michael Ladomery.

References

Azubel M, Wolf SG, Sperling J, Sperling R (2004) Three-dimensional structure of the native spliceosome by cryo-electron microscopy. Mol Cell 15(5):833–839

Barbosa-Morais NL, Irimia M, Pan Q, Xiong HY, Gueroussov S, Lee LJ et al (2012) The evolutionary landscape of alternative splicing in vertebrate species. Science 338(6114): 1587–1593

Bates DO, Cui TG, Doughty JM, Winkler M, Sugiono M, Shields JD et al (2002) VEGF165b, an inhibitory splice variant of vascular endothelial growth factor, is down-regulated in renal cell carcinoma. Cancer Res 62(14):4123–4131

Berget SM, Moore C, Sharp PA (1977) Spliced segments at the 5' terminus of adenovirus 2 late mRNA. Proc Natl Acad Sci USA 74(8):3171–3175

Black DL (1995) Finding splice sites within a wilderness of RNA. RNA 1(8):763–771

Braunschweig U, Gueroussov S, Plocik AM, Graveley BR, Blencowe BJ (2013) Dynamic integration of splicing within gene regulatory pathways. Cell 152(6):1252–1269

Carstens RP, Eaton JV, Krigman HR, Walther PJ, Garcia-Blanco MA (1997) Alternative splicing of fibroblast growth factor receptor 2 (FGF-R2) in human prostate cancer. Oncogene 15(25): 3059–3065

Chow LT, Gelinas RE, Broker TR, Roberts RJ (1977) An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. Cell 12(1):1–8

DeNoto FM, Moore DD, Goodman HM (1981) Human growth hormone DNA sequence and mRNA structure: possible alternative splicing. Nucleic Acids Res 9(15):3719–3730

Ding Y, Zhang H, Zhou Z, Zhong M, Chen Q, Wang X et al (2012) u-PA inhibitor amiloride suppresses peritoneal metastasis in gastric cancer. World J Surg Oncol 10:270

Flavell AJ, Cowie A, Legon S, Kamen R (1979) Multiple 5' terminal cap structures in late polyoma virus RNA. Cell 16(2):357–371

Fu XD (2004) Towards a splicing code. Cell 119(6):736–738

Fu RH, Liu SP, Huang SJ, Chen HJ, Chen PR, Lin YH et al (2013) Aberrant alternative splicing events in Parkinson's disease. Cell Transplant 22(4):653–661

Fugier C, Klein AF, Hammer C, Vassilopoulos S, Ivarsson Y, Toussaint A et al (2011) Misregulated alternative splicing of BIN1 is associated with T tubule alterations and muscle weakness in myotonic dystrophy. Nat Med 17(6):720–725

Gamazon ER, Stranger BE (2014) Genomics of alternative splicing: evolution, development and pathophysiology. Hum Genet 133(6):679–687

Gammons MV, Federov O, Ivison D, Du C, Clark TL, Hopkins C et al (2013) Topical anti-angiogenic SRPK1 inhibitors reduce choroidal neovascularization in rodent models of exudative-AMD. Invest Ophthalmol Vis Sci 54(9):6052–6062

Gammons MV, Lucas R, Dean R, Coupland SE, Oltean S, Bates DO (2014) Targeting SRPK1 to control VEGF-mediated tumour angiogenesis in metastatic melanoma. Br J Cancer 111(3): 477–485

Gerold KD, Zheng P, Rainbow DB, Zernecke A, Wicker LS, Kissler S (2011) The soluble CTLA-4 splice variant protects from type 1 diabetes and potentiates regulatory T-cell function. Diabetes 60(7):1955–1963

Ghigna C, Riva S, Biamonti G (2013) Alternative splicing of tumor suppressors and oncogenes. Cancer Treat Res 158:95–117

Gohda T, Tanimoto M, Moon JY, Gotoh H, Aoki T, Matsumoto M et al (2008) Increased serum endogenous secretory receptor for advanced glycation end-product (esRAGE) levels in type 2 diabetic patients with decreased renal function. Diabetes Res Clin Pract 81(2):196–201

Guo W, Schafer S, Greaser ML, Radke MH, Liss M, Govindarajan T et al (2012) RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing. Nat Med 18(5):766–773

Hanahan D, Weinberg RA (2011) Hallmarks of cancer: the next generation. Cell 144(5):646–674

Harper SJ, Bates DO (2008) VEGF-A splicing: the key to anti-angiogenic therapeutics? Nat Rev Cancer 8(11):880–887

Hartmann L, Theiss S, Niederacher D, Schaal H (2008) Diagnostics of pathogenic splicing mutations: does bioinformatics cover all bases? Front Biosci 13:3252–3272

Hiller M, Zhang Z, Backofen R, Stamm S (2007) Pre-mRNA secondary structure influence exon recognition. PLoS Genet 3(11):e204

Kanopka A, Muhlemann O, Petersen-Mahrt S, Estmer C, Ohrmalm C, Akusjarvi G (1998) Regulation of adenovirus alternative RNA splicing by dephosphorylation of SR proteins. Nature 393(6681):185–187

Keren H, Lev-Maor G, Ast G (2010) Alternative splicing and evolution: diversification, exon definition and function. Nat Rev Genet 11(5):435–455

Ladomery M (2013) Aberrant alternative splicing is another hallmark of cancer. Int J Cell Biol 2013:463786

Lamond AI, Spector DL (2003) Nuclear speckles: a model for nuclear organelles. Nat Rev Mol Cell Biol 4(8):605–612

Le Bacquer O, Shu L, Marchand M, Neve B, Paroni F, Kerr Conte J et al (2011) TCF7L2 splice variants have distinct effects on beta-cell turnover and function. Hum Mol Genet 20(10): 1906–1915

Lewandowska MA (2013) The missing puzzle piece: splicing mutations. Int J Clin Exp Pathol 6(12):2675–2682

Luco RF, Allo M, Schor IE, Kornblihtt AR, Misteli T (2011) Epigenetics in alternative pre-mRNA splicing. Cell 144(1):16–26

Matlin AJ, Clark F, Smith CW (2005) Understanding alternative splicing: towards a cellular code. Nat Rev Mol Cell Biol 6(5):386–398

Misteli T (2000) Cell biology of transcription and pre-mRNA splicing: nuclear architecture meets nuclear function. J Cell Sci 113(11):1841–1849

Misteli T (2008) Physiological importance of RNA and protein mobility in the cell nucleus. Histochem Cell Biol 129(1):5–11

Miura K, Fujibuchi F, Unno M (2012) Splice variants in apoptotic pathway. Exp Oncol 34(3): 212–217

Naro C, Sette C (2013) Phosphorylation-mediated regulation of alternative splicing in cancer. Int J Cell Biol 2013:151839

Naryshkin NA, Weetall M, Dakka A, Narasimhan J, Zhao X, Feng Z et al (2014) Motor neuron disease. SMN2 splicing modifiers improve motor function and longevity in mice with spinal muscular atrophy. Science 345(6197):688–693

Ogawa S (2012) Splicing factor mutations in myelodysplasia. Int J Hematol 96(4):438–442

Ogawa T, Shiga K, Hashimoto S, Kobayashi T, Horii A, Furukawa T (2003) APAF-1-ALT, a novel alternative splicing form of APAF-1, potentially causes impeded ability of undergoing DNA damage-induced apoptosis in the LNCaP human prostate cancer cell line. Biochem Biophys Res Commun 306(2):537–543

Oltean S, Bates DO (2014) Hallmarks of alternative splicing in cancer. Oncogene 33(46):5311–8

Owen N, Zhou H, Malygin AA, Sangha J, Smith LD, Muntoni F et al (2011) Design principles for bifunctional targeted oligonucleotide enhancers of splicing. Nucleic Acids Res 39(16): 7194–7208

Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet 40(12): 1413–1415

Rodriguez-Martin T, Anthony K, Garcia-Blanco MA, Mansfield SG, Anderton BH, Gallo JM (2009) Correction of tau mis-splicing caused by FTDP-17 MAPT mutations by spliceosome-mediated RNA trans-splicing. Hum Mol Genet 18(17):3266–3273

Shoji K, Teishima J, Hayashi T, Ohara S, McKeehan WL, Matsubara A (2014) Restoration of fibroblast growth factor receptor 2IIIb enhances the chemosensitivity of human prostate cancer cells. Oncol Rep 32(1):65–70

Shultz JC, Goehe RW, Murudkar CS, Wijesinghe DS, Mayton EK, Hawkins AK, Mukerjee P, Pinkerman RL, Park MA, Chalfant CE (2011) SRSF1 regulates the alternative splicing of caspase 9 via a novel intronic splicing enhancer affecting the chemotherapeutic sensitivity of non-small cell lung cancer cells. Mol Cancer Res 9(7):889–900

Singh RK, Cooper TA (2012) Pre-mRNA splicing in disease and therapeutics. Trends Mol Med 18(8):472–482

Stamm S (2008) Regulation of alternative splicing by reversible protein phosphorylation. J Biol Chem 283(3):1223–1227

Stanek D, Pridalova-Hnilicova J, Novotny I, Huranova M, Blazikova M, Wen X, Sapra AK, Neugebauer KM (2008) Spliceosomal small nuclear ribonucleoprotein particles repatedly cycle through Cajal bodies. Mol Biol Cell 19(6):2534–2543

Surget S, Khoury MP, Bourdon JC (2013) Uncovering the role of p53 splice variants in human malignancy: a clinical perspective. Onco Targets Ther 7:57–68

Tennyson CN, Klamut HJ, Worton RG (1995) The human *dystrophin* gene requires 16 hours to be transcribed and is cotranscriptionally spliced. Nat Genet 9(2):184–190

Wang Z, Burge CB (2008) Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. RNA 14(5):802–813

Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C et al (2008) Alternative isoform regulation in human tissue transcriptomes. Nature 456(7221):470–476

Yasumoto H, Matsubara A, Mutaguchi K, Usui T, McKeehan WL (2004) Restoration of fibroblast growth factor receptor2 suppresses growth and tumorigenicity of malignant human prostate carcinoma PC-3 cells. Prostate 61(3):236–242

Zhou Z, Qiu J, Liu W, Zhou Y, Plocinik RM, Li H et al (2012) The Akt-SRPK-SR axis constitutes a major pathway in transducing EGF signaling to regulate alternative splicing in the nucleus. Mol Cell 47(3):422–433

Acute Promyelocytic Leukaemia: Epigenetic Function of the PML-RARα Oncogene

Julia P. Hofmann and Paolo Salomoni

Abstract Acute promyelocytic leukaemia is a myeloid neoplasm characterized by expansion of promyelocytic progenitors. Its main driver is the oncogenic fusion protein generated by the (t15;17) chromosomal translocation between the promyelocytic leukaemia (PML) and retinoic acid receptor α (RARα) genes. Being PML-RARα the primary trigger of APL, it represents an excellent model to study neoplastic transformation in the haematopoietic system. Importantly, epigenetic changes imposed by and/or associated with PML-RARα have been implicated not only in promoting/sustaining the tumour phenotype, but also in influencing therapy response. In this chapter we will discuss the existing literature on chromatin remodelling driven by PML-RARα and its impact on APL pathogenesis and therapy.

1 Acute Promyelocytic Leukaemia

Acute promyelocytic leukaemia (APL), a variant of acute myeloid leukaemia (AML), is a neoplastic myeloid disorder that is characterized by accumulation of undifferentiated blood cells at the promyelocyte stage in patient's blood samples (de The and Chen 2010; Puccetti and Ruthardt 2004). This high number of blast cells is accompanied by a reduced number of mature blood cells, therefore causing a high risk of bleeding due to coagulopathy and thrombocytopenia. APL can occur at any age with a median age-of-onset between 40 and 50 years, with peadiatric APL occurring more rarely compared to adult APL. The increasing knowledge on the molecular causes of APL have effectively improved therapy over the last 40 years, thus turning the previously fatal disease into a highly curable disease with remission rates over 90 % (de The and Chen 2010). Based on molecular analysis, APL is clinically classified by the detection of different oncofusion-proteins as a result of chromosomal translocations between the RARα gene and

J.P. Hofmann • P. Salomoni (✉)
Samantha Dickson Brain Cancer Unit, University College London Cancer Institute, London, United Kingdom
e-mail: p.salomoni@ucl.ac.uk

D.P. Bazett-Jones, G. Dellaire (eds.), *The Functional Nucleus*,
DOI 10.1007/978-3-319-38882-3_4

the genes for PML (de The et al. 1990a), PLZF (Chen et al. 1993), NuMA (Wells et al. 1997), NPM (Redner et al. 1996), STAT5b (Arnould et al. 1999), PRKAR1A (Catalano et al. 2007), FIPL1 (Kondo et al. 2008), BCOR (Yamamoto et al. 2010) or OBFC2A (Won et al. 2013). In this chapter we will focus on the PML-RARα fusion protein and its influence on epigenetics in cancer pathogenesis.

2 The PML-RARα Oncogene

In 95 % of cases, APL blasts present the reciprocal chromosomal translocation (t15;17) indicating that the resulting PML-RARα fusion is the likely driver of transformation in these cells. Three different chromosomal break points for PML-RARα are known of which the most frequently (~70 %) occurring break point (bcr1) results in the longest PML-RARα transcript, bcr3 (~20 % of cases) in the shortest transcript and the rarely (~10 %) occurring bcr2 in an intermediate long transcript (Melnick and Licht 1999). All three PML-RARα isoforms contain the functional domains of both PML and RARα and we will first give a brief overview of the normal functions of PML and RARα to point out how both proteins functionally contribute to APL pathogenesis.

2.1 PML

The human *PML* gene is encoded on chromosome 15 and seven different isoforms exist with variations in the central and C-terminal part of PML (Salomoni et al. 2008; Jensen et al. 2001; Nisole et al. 2013; Condemine et al. 2006). Three different domains of PML - the RING, B-box, and coiled-coil (RBCC) motif - classify PML as a tripartite motif (TRIM) protein family member and exhibit functional importance for the formation of higher multimers (Jensen et al. 2001; Reymond et al. 2001). The α-helical coiled-coil domain allows self-association of PML molecules (Burkhard et al. 2001), the RING domain and two B-boxes are Zn-binding domains, but do not confer direct DNA binding ability to PML (Reymond et al. 2001). Interestingly, the RING domain has E3 ligase activity (Joazeiro and Weissman 2000). These functional domains of PML are important for protein-protein interactions to form macromolecular structures in the nucleus, termed PML nuclear bodies (PML-NBs) (Salomoni et al. 2008; Nisole et al. 2013; Borden et al. 1995). More than 70 proteins—such as p53, DAXX, pRb, SP100, HP-1—are known to be associated with PML-NBs either by direct or indirect interaction with PML (Dellaire and Bazett-Jones 2004; Salomoni et al. 2008; Wolyniec et al. 2013). Most importantly, post-translational modifications such as phosphorylation, acetylation, ubiquitination and small ubiquitin-like modifier (SUMO)-ylation regulate protein-protein interactions in PML-NBs and affect the dynamics of PML-NB functions (Dellaire and Bazett-Jones 2004; Fu et al. 2005;

Wolyniec et al. 2013). Structural dissociation of NB components is observed in $PML^{-/-}$ cells, showing that PML displays the integral part for NB assembly that is based on PML SUMOylation (Ishov et al. 1999). PML has a SUMO-interacting motif (SIM) and three lysine residues for SUMOylation that are critical for the recruitment of other PML-NB proteins, for instance through the SIM domain of partner proteins such as death-domain associated protein 6 (DAXX) (Salomoni 2013; de The et al. 2012; Jeanne et al. 2010). Furthermore, SUMOylation is important for PML's degradation as shown by using arsenic trioxide that induces ubiquitination and proteasomal degradation of PML via poly-SUMOylation (Jeanne et al. 2010; Lallemand-Breitenbach et al. 2001). Importantly, through its ability to control cellular senescence, apoptosis, and viral infection, PML has been implicated in regulation of tumour suppression (for more exhaustive details see recent review articles) (Dellaire and Bazett-Jones 2004; Bernardi and Pandolfi 2007; Salomoni et al. 2008; de The et al. 2012). A number of these functions have been linked to the ability of PML to recruit nuclear factors involved in tumourigenesis to NBs (Salomoni et al. 2008). For instance, PML leads to p53 and pRb localization to NBs, thus resulting in their activation (Salomoni et al. 2008). However, PML-NB-independent function of PML in regulation of p53 has been reported (Bernardi et al. 2004).

Notably, PML has been implicated in regulation of stem cell fate in multiple tissues (for more exhaustive details see recent review article) (Salomoni 2009). For instance, PML controls the self-renewing capacity of stem/progenitor cells in both the haematopoietic system and the nervous system via mechanisms involving mTOR, fatty acid oxidation and pRb (Regad et al. 2009; Ito et al. 2008, 2012). Recent studies have suggested that within established tumours PML is able to regulate cancer stem cell self-renewal and cancer cell survival via its ability to restrict G1/S progression and control lipid metabolism (Ito et al. 2008, 2012; Carracedo et al. 2012). Overall, these data suggest that PML plays a pleiotropic role in cancer pathogenesis.

2.2 RARα

Retinoic acid receptors are nuclear hormone receptors existing as RARα, β and γ form with RARα being expressed in most tissue types (de The et al. 1990b; Kastner and Chan 2001). The two functional domains of RARα are a DNA binding domain at the N-terminal domain and a ligand binding domain at the C-terminal domain (Kastner and Chan 2001). RARα forms a heterodimer with retinoid X receptor (RXR) and the RARα/RXRα dimers bind to retinoic acid response elements (RAREs) at gene promoter regions in DNA that are characterized by two 5′-(A/G)G(G/T)TCA-3′ core sequence motifs (de The et al. 1990b; Melnick and Licht 1999). The direct repeat (DR) orientation of this motif is interspaced by a variable number of nucleotides and wild-type RARα/RXRα dimers mainly bind to classical RAREs interspaced by 2 or 5 nucleotides (DR2, DR5) (de The

et al. 1990b). RARα/RXRα regulate transcription through gene repression together with co-repressor complexes (see below), while binding of its ligands all-trans- and 9-cis-retinoic acid (RA) causes release of the co-repressors and binding of co-activators (Kastner and Chan 2001). Based on the fact that RARα is fusion partner in all APL oncogenes, RARα was hypothesised to be a major regulator of granulocyte differentiation (Kastner and Chan 2001). However, genetic mouse models on RARα knock-out are viable exhibiting only few defects (Lufkin et al. 1993) and show no impairment of neutrophil maturation, whereas antagonists of RA inhibit granulopoiesis in wild-type mice (Kastner et al. 2001). These in vivo data indicate that RARα functions as a negative-regulator of granulocyte differentiation, but promotes differentiation upon ligand binding (Kastner and Chan 2001).

2.3 *Pharmacological Targeting of PML-RARα Leads to Disease Remission*

The characteristic differentiation block of myeloid cells in APL was identified early as a target for specialized therapy and the development of a combination therapy consisting of chemotherapeutical reagents in combination with all-trans-retinoic acid (ATRA) or arsenic trioxide (As_2O_3) has turned out to be highly effective treatment for APL patients (Castaigne et al. 1990; Lo-Coco et al. 2010; Estey et al. 2006; Sanz et al. 2010; Chen et al. 1997). Genome-wide studies have shown that pharmacological doses of ATRA cause release of PML-RARα from its target sites on chromatin and subsequently induce transcription and a profound terminal differentiation of promyelocytes that is contributing to disease remission in patients (Martens et al. 2010; de The and Chen 2010). In contrast to ATRA As_2O_3 causes only partial differentiation of promyelocytes, but induces apoptosis, thus suggesting that terminal differentiation of promyelocytes is not the key event for APL remission (Chen et al. 1997). Indeed, in vivo experiments with mice could show that not the differentiation induced by the synergy of ATRA and As_2O_3 is sufficient to eradicate APL, but the abolishment of leukaemia initiating cells (LICs) (Nasr et al. 2008). LICs are quiescent cells with the intrinsic property for self-renewal and can re-initiate disease, when not completely eradicated by treatment (Wojiski et al. 2009; Guibal et al. 2009). How can the effects of ATRA and As_2O_3 be explained on the molecular level? Extensive studies in vitro and in vivo could show that both ATRA and As_2O_3 cause degradation of PML-RARα, with ATRA degrading PML-RARα via the RARα moiety (Zhu et al. 1999), while As_2O_3 degrades PML-RARα via the PML moiety (Zhu et al. 1997). In detail, ATRA degrades the oncofusion protein via a cyclic AMP (cAMP)-dependent phosphorylation site in PML-RARα, whereas As_2O_3 degrades PML-RARα in a SUMO-dependent, ubiquitin/proteasome-mediated way and induces senescence by restoration of normal PML-NBs (Zhu et al. 1997; Lallemand-Breitenbach et al. 2001; Nasr et al. 2008). For instance, eradication of LICs in mice is delayed upon the

proteasome inhibitor bortezomib, suggesting that degradation of PML-RARα by both ATRA and As_2O_3 is important for disease remission not only in mice (Nasr et al. 2008), but most likely also in humans. Of note, the therapeutic combination of ATRA and As_2O_3 is highly effective and can dispense the need for DNA damaging reagents (Estey et al. 2006). These findings present correlative evidence that PML-RARα degradation is required for disease remission and point to the oncofusion protein as the driver of APL.

2.4 PML-RARα Mouse Models

Experimental evidence that PML-RARα is driving APL comes from a number of studies based on genetically engineered mouse models with endogenous expression of human PML-RARα, transplantation of haematopoietic progenitor cells (HPCs) transduced with PML-RARα into syngenic recipient mice or from xenograft studies, which all lead to APL-like disease development (Kogan 2007). However, mouse models with human PML-RARα suggest that in the initial phase of PML-RARα induced leukaemia formation, PML-RARα primarily increases the self-renewing capacity of haematopoietic stem cells without major changes in differentiation (Wojiski et al. 2009). The observed latency has been correlated to different expression levels of PML-RARα in the different models (Westervelt et al. 2003; Kogan 2007). Furthermore, expression of the human PML-RARα gene rearrangement on a murine background may contribute to delayed onset, as introduction of a murine PML-RARα was recently reported to shorten the latency in mice (Korf et al. 2014). The long latency and incomplete penetrance observed with these mouse models may also indicate the need for secondary genetic or epigenetic events during leukaemogenesis (Chan et al. 2006; Zimonjic et al. 2000; Le Beau et al. 2002). Hence, as in APL patients secondary mutations such as activated Ras and Flt3 as well as genome-wide changes of the epigenetic landscape are found, these events were remodelled and analysed in cell lines and mice (Chan et al. 2006; Kelly et al. 2002; Walter et al. 2004). For instance, in a mouse model combination of PML-RARα and Flt3 with internal tandem duplications significantly shortens the latency period (Kelly et al. 2002). Epigenetic changes may also be involved in promoting transformation and have been analysed in cell culture models, as discussed in detail below. Overall, besides the clear promotion of pathogenesis by secondary genetic and epigenetic events, disease eradication by pharmacological targeting of PML-RARα and leukaemia induction in PML-RARα mouse models clearly identify PML-RARα as the driving oncogene in APL.

2.5 PML-RARα Mechanisms of Action

How does PML-RARα cause transformation of haematopoietic cells? First, PML-RARα has a dominant-negative effect on wild-type PML function through impairment of PML-NB formation (Salomoni and Pandolfi 2002). The oncogenic fusion protein disrupts the normal PML-NB structure and causes formation of smaller PML entities most likely via multimerisation with the heterozygous normal PML protein through the PML coiled-coil domain (Koken et al. 1994; Daniel et al. 1993; Minucci et al. 2000). In microscopic immunostaining, these changes are visible as a microspeckled pattern of PML with increased number of PML entities compared to untransformed control cells (Daniel et al. 1993; de The and Chen 2010). Is NB disruption required for leukaemogenesis? Several studies suggest that this may be the case. For instance, PML loss results in increased incidence and decreased onset of APL-like disease in mice (Rego et al. 2001). Further, Hugues de The's group has shown that derepression of the PML/p53 axis in APL cells is critical for the response to therapy (Ablain et al. 2014). On the other hand, other reports have been arguing for a less prominent role of PML/PML-RARα interaction and NB disruption in APL. In this respect, PML has been proposed to serve only as a multimerisation domain. This view is in part corroborated by the existence of other APL translocations which involve RARα partners that similarly to PML act as multimerisation platforms (Licht 2006; Sternsdorf et al. 2006; Kwok et al. 2006). Further, PML loss is not sufficient to induce leukaemia (Wang et al. 1998). Therefore, dominant-negative effects on RARα may also contribute to APL formation (Melnick and Licht 1999). As mentioned above, RARα likely works as a negative regulator of granulocyte differentiation in the absence of ligand, but it becomes a positive regulator upon binding to RA (Kastner and Chan 2001). In contrast, PML-RARα is resistant to physiological levels of ATRA most likely via multimerisation and resulting increase in co-repressor density at target promoters (see below) (He et al. 1998; Lin et al. 1998; Grignani et al. 1998; Minucci et al. 2000; Perez et al. 1993). Supporting a role for the repressive function of RARα in APL, it was reported that overexpression of RARα is sufficient to cause a differentiation block in the absence of RA in vitro (Onodera et al. 1995; Du et al. 1999). However, a dominant-negative RARα is unable to cause APL in mice (Kogan et al. 2000; Matsushita et al. 2006). Furthermore, forced multimerisation of RARα is able to transform cells in vitro, but it does not cause a full-blown leukaemia in mice (Sternsdorf et al. 2006). Overall, these data suggest that (1) multimerisation/repressive function of oncogenic RARα is required but not sufficient to cause leukaemia in vivo and (2) PML (and other translocation partners) likely plays an important role in disease pathogenesis. However, it is conceivable that in addition to its dominant-negative activities PML-RARα may also act as a gain-of-function mutant. For instance, through multimerisation with RXRα PML-RARα binds not only to classical DR2 and DR5 motifs, but binds to multiple different RARE motifs (Zeisig et al. 2007; Kamashev et al. 2004). This relaxed binding specificity of PML-RARα/RXRα

likely contributes to APL formation by keeping the chromatin of these target genes in a repressed state, while at the same time inducing a subset of genes that would normally not be actively expressed (Zeisig et al. 2007; Martens et al. 2010; Hoemme et al. 2008; Zhu et al. 2007). Changes in transcription levels of PML-RARα/RXRα target genes, but not wild-type RARα/RXRα targets contribute to an overall aberrant transcription profile of HPCs (Lin and Evans 2000; de The and Chen 2010). Resulting is a block of differentiation of HPCs at the promyelocyte stage and a lack of mature myeloid cells in APL patients (de The and Chen 2010). The important role of PML-RARα/RXRα interaction is underlined by the fact that a PML-RARα mutant defective in RXRα binding inhibits APL formation in mice (Zhu et al. 2007) and loss of RXR causes terminal differentiation of blast cells in vivo suggesting that promoter clearance of PML-RARα/RXRα induces terminal differentiation (Vitaliano-Prunier et al. 2014). Additional gain-of-function properties of PML-RARα are discussed below.

3 PML-RARα Epigenetic Function

Epigenetic changes are occurring frequently in many types of cancer and contribute to the global cellular changes of transformed cells (Dawson and Kouzarides 2012). Maintenance and tight regulation of the epigenetic profile is important for global genome functionality. Epigenetic modifications include methylation of DNA at CpG dinucleotides as well as methylation and acetylation of histone tails and chromatin loading of histone variants (Dawson and Kouzarides 2012) (see Box 1). Epigenetic changes in APL have been routinely analysed in primary APL patient blasts or in patient derived cell lines, and there is a large body of literature implicating epigenetic remodelling in APL pathogenesis. In particular, PML-RARα has been reported to affect both histone and DNA modifications directly via interaction with epigenetic enzymes and chromatin remodelling complexes and indirectly via regulation of gene expression (see Fig. 1). We now summarise and discuss the available literature on this subject.

3.1 Histone Modifications

3.1.1 Histone Acetylation

As mentioned before, one of PML-RARα's dominant-negative effects on RARα is its enhanced repressive effect that is potentiated through multimerisation and the resulting increase in binding sites for co-repressors. Bound to chromatin PML-RARα forms, likewise wild-type RARα/RXRα, complexes with nuclear repressors such as nuclear receptor corepressor (NCoR), silencing mediator for retinoid and thyroid hormone receptors (SMRT) and histone deacetylases

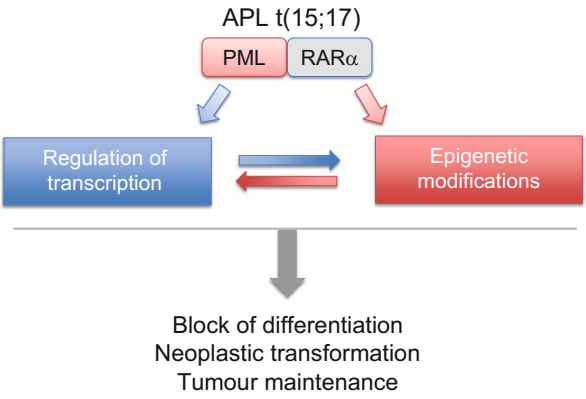

Fig. 1 PML-RARα regulates transcription (both activation and repression) via its ability to act as a dominant-negative and gain-of-function mutant. It is also clear that PML-RARα is able to promote epigenetic modifications, which in turn are believed to contribute to transcriptional changes. However, transcriptional activation or repression by PML-RARα is predicted to affect the chromatin status of its target genes. Further, some of the epigenetic changes found in APL cells are in part due to ability of PML-RARα to control the transcription of epigenetic modifiers. Overall, the molecular changes caused by PML-RARα are believed to synergistically contribute to transformation of haematopoietic progenitors and potentially tumour progression

(HDACs), thus repressing basal transcription of RARα target genes (Lin et al. 1998; Grignani et al. 1998; Minucci et al. 2000; Rice and de The 2014). Likely due to enhanced co-repressor recruitment by PML-RARα physiological concentrations of RA are not sufficient to release the bound co-repressors and thus the RARα-typical binding of transcriptional co-activators such as CBP/p300 histone acetyltransferases (HATs) is inhibited and can only be achieved through pharmacological concentrations of ATRA (Di Croce et al. 2004; Lin and Evans 2000). PML-RARα's interaction with HDACs and HATs directly links APL with epigenetic modifications that presumably contribute to APL pathogenesis (see Fig. 2). The histone marks of APL cells are markedly altered in comparison to normal blood cells. Most prevalent is a general hypoacetylation of chromatin compared to control cells with clearly reduced acetylation levels of histones H3 and H4 in NB4, U937 and primary APL patient cells (Hoemme et al. 2008; Martens et al. 2010; Saeed et al. 2012; Nouzova et al. 2004; Schoofs et al. 2013). Upon Zn-induced expression of PML-RARα, increased chromatin deacetylation is observed in U937 cells (Martens et al. 2010) underlining the hypothesis that hypoacetylation is caused by the presence of PML-RARα at its specific target genes and its association with HDACs (Lin et al. 1998; Grignani et al. 1998). Further epigenetic influence on leukaemogenesis potentially results from depletion of HDAC from its normal binding sites due to the enhanced recruitment by PML-RARα, thereby changing PML-RARα unrelated pathways. The prevalent hypoacetylation of APL cells indicates that HDAC may be the most important co-repressor of PML-RARα and may substantially contribute to pathogenesis, however, a RARα-HDAC fusion protein is not sufficient to induce leukaemia in mice (Matsushita et al. 2006).

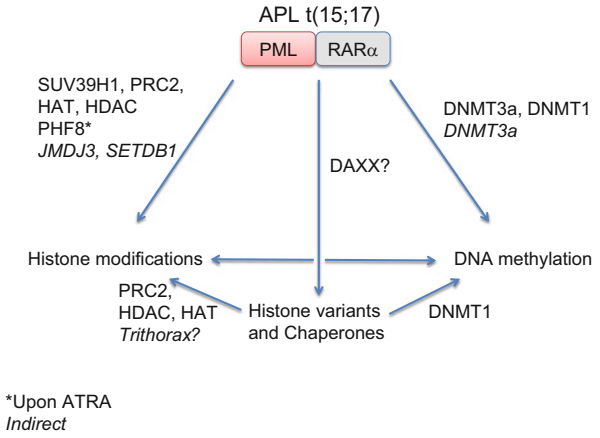

Fig. 2 PML-RARα controls chromatin status by multiple mechanisms: (1) it promotes histone modifications via its ability to interact with (or control the expression of) histone-modifying enzymes (see also Table 1); (2) it interacts with the H3.3 chaperone DAXX thus potentially affecting H3.3 loading in APL cells; (3) it associates with DNA methyltransferases. These modifications are not predicted to work in isolation, as for instance histone modifications and DNA methylation are closely interconnected and H3.3 deposition has been shown to affect the function of the Polycomb repressive complex 2. Finally, DAXX can interact with DNMT1, suggesting that its recruitment by PML-RARα may influence DNA methylation

Thus, the role of deacetylation in APL pathogenesis remains to be fully investigated. Nonetheless, there are a number of clinical trials based on the HDAC inhibitor vorinostat that have been completed or are currently ongoing.

3.1.2 Histone Methylation

Besides H3/H4 acetylation/deacetylation, methylation/demethylation of lysine residues of H3 tail contribute to chromatin remodelling and subsequent transcriptional alterations. Methylation at H3 tails affects mainly residues K4, K9, K27 and K36, as further summarised in Table 1. Interestingly, bioinformatic cluster analysis revealed a subcluster that linked the active H3K4me3 mark to promoter regions in NB4 cells, while another subcluster was associated with repressive chromatin marks such as H3K9me3 and H4K20me3 (Saeed et al. 2012). Moreover, several subclusters were associated with non-promoter regions and could display enhancer clusters, thus being indicative of an influence of epigenetic changes at non-promoter regulatory regions in APL pathogenesis (Saeed et al. 2012). Contradictory data exist regarding the histone methylation status of PML-RARα target sites. For instance, in the study of Martens and colleagues levels of H3K9me3 were found to be low at PML-RARα target sites and overall levels of H3K27me3 and DNA methylation were low (Martens et al. 2010). In contrast Hoemme and colleagues found an increase of H3K9me3 at PML-RARα targeting promoters as

well as increases of H3K4me3 at most of the target genes in U937 cells (Hoemme et al. 2008). Interestingly, the increase in H3K9me3 levels were linked to changes in mRNA expression levels and were corresponding to repressed chromatin (Hoemme et al. 2008). The opposing data observed for H3 methylation marks may be explained by the use of different cell lines (Martens et al. 2010) and ChIP-chip versus ChIP-seq techniques. Furthermore, different expression levels of PML-RARα may influence the number of PML-RARα binding sites and epigenetic modifiers at each target may influence the methylation pattern. Of note, PML-RARα target sites include several chromatin modifying enzymes such as JMDJ3 that is involved in H3K27 demethylation, the H3K9 methylating enzyme SETDB1 and the DNMT3a DNA methyltransferase (see also below) (Martens et al. 2010). Changes in transcription levels of these enzymes may well explain the aberrant methylation profile of APL cells. Even though many of the observed histone modifications in APL might result of secondary events due to PML-RARα regulating the expression of epigenetic modifying enzymes, direct interaction of PML-RARα with the cellular epigenetic modifying machinery has been shown (see Fig. 2). PML-RARα interacts via the PML moiety with SUV39H1 that contains a C-terminal SET domain and functions as a histone methyltransferase for H3K9 in heterochromatin organization (Carbone et al. 2006). In Zn-inducible U937 cells PML-RARα binding induces H3K9me3 marks at the RARβ2 promoter correlating with a repressed chromatin state (Carbone et al. 2006). RARβ2, the most studied target of PML-RARα/RXRα, has a well characterized RARE and a CpG island in its promoter and is important for the differentiation of HPCs (Di Croce et al. 2002). Interestingly, complex formation of PML-RARα with SUV39H1 depends on the multimerisation properties of PML-RARα (Carbone et al. 2006). Thus, PML-RARα multimers provide the basis for the interaction with chromatin modifiers that cooperate in the differentiation block by inducing a heterochromatin-like environment (Carbone et al. 2006). The impact of PML-RARα multimerisation on gene silencing is further shown by the recruitment of polycomb repressive complex 2 (PRC2) (Villa et al. 2007). Polycomb group (PcG) proteins establish bivalent histone marks to alter the chromatin state of target cells between active and repressed state by methylation of H3K27 (Mills 2010). Multimerised PML-RARα can form a complex with the PRC2/3/4 components SUZ12, EZH2 and EED, subsequently targeting the PRC2 complex to PML-RARα target sites such as RARβ2 promoter (Villa et al. 2007). Upon PRC2 complex binding H3K27me3 levels increased at the RARβ2 promoter with concomitant decrease of H3K27me1 and H3K27ac (Villa et al. 2007). Interestingly, DNA methylation levels changed together with the methylation status of H3K27 and indicate a cross-talk between DNA and histone methylation (Villa et al. 2007; Widschwendter et al. 2007; Dawson and Kouzarides 2012). In accordance with hypermethylation of PcG sites in various cancers (Widschwendter et al. 2007; Mills 2010), SUZ12 binding sites were hypermethylated in APL cells (Schoofs et al. 2013). An aberrant PRC2-methyltransferase cross-talk at specific genes was hypothesised to contribute to neoplastic transformation by impairing normal differentiation (Widschwendter et al. 2007). A further correlation of reduced DNA methylation and decreased

Table 1 List of selected histone modifications along with information about the respective epigenetic enzymes responsible for writing or erasing these marks

Histone residue	Chemical modification	Histone modifying enzyme (gene, homo sapiens)
H2A, H2B, H3, H4	Deacetylation	HDAC3, HDAC7
H2A, H3, H4	Deacetylation	SIRT1
H3, H4	Deacetylation	SIRT2
H2A, H2B, H3, H4	Acetylation	CBP/p300
H3, H4	Acetylation	GCN5, PCAF, MOZ/MORF
H2A, H4	Acetylation	TIP60
H4	Acetylation	HBO1
H3K4	Methylation	KMT2B, KMT2D, SMYD1, SMYD2, SMYD3, Set1/Ash2, MLL1/MLL, MLL/SET1, WHSC1L1, PRDM9, SETD1A, SETD1B, SETD7, SETMAR
H3K4	Demethylation	KDM1A, KDM1B, KDM2B, KDM4A, KDM5A, KDM5B, KDM5C, KDM5D, NO66
H3K9	Methylation	SUV39H1, SUV39H2, PRC2/EED-EZH2, GLP1/EHMT1, G9a/EHMT2, SETDB1, SETDB2
H3K9	Demethylation	KDM1A, KDM3A, KDM3B, KDM4A, KDM4B, KDM4C, KDM4D, KDM4E, KDM7A, MINA, PHF2, PHF8, HR, JMJD1C
H3K27	Methylation	PRC2/EED-EZH2, PRC2/EED-EZH1, EHMT1, EHMT2, WHSC1, WHSC1L1
H3K27	Demethylation	KDM6A, KDM7A, PHF8
H3K36	Methylation	SMYD2, ASH1L, NSD1, SETD2, SETD3, SETMAR
H3K36	Demethylation	KDM2A, KDM2B, KDM4A, KDM4C, KDM8, NO66
H4K20	Methylation	NSD1, PRDM6, SETD8, SUV420H1, SUV420H2
H4K20	Demethylation	KDM7A, PHF8

H3K27me3 levels could also be observed when the interaction of PML-RARα with the NuRD complex was impaired due to knock-down of NuRD complex components, hence indicating a functional repressive link between NuRD complex and PRC2 (Morey et al. 2008). However, the influence of PRC2 on the epigenome of APL patients remains to be further elucidated as H3K27me3 levels were found to be low at PML-RARα target sites in NB4 cells (Martens et al. 2010). Overall, genome-wide changes of histone tail modifications and PML-RARα's direct interaction with chromatin modifying enzymes can be found in PML-RARα expressing cells, but additional studies are necessary to characterize its impact on APL pathogenesis.

3.2 Chromatin-Remodelling Factors and Histone Variants

Further impact of PML-RARα on the epigenome might result from its interaction with the co-repressor DAXX that binds via its SIM domain to lysine 160 in PML (de The and Chen 2010; Salomoni 2013). Chromatin remodelling via DAXX could be established through the interaction with other epigenetic enzymes, such as CBP acetyltransferases (Kuo et al. 2005), HDAC II (Hollenbach et al. 2002) and DNMT1 (Puto and Reed 2008; Zhang et al. 2013) (see Fig. 2). Interestingly, mice expressing mutant PML-RARα K160R have deficient PML-RARα/DAXX interaction and develop only a myeloproliferative syndrome, but no APL, whereas a DAXX-PML-RARα K160R fusion protein induces APL, thus showing the important influence of DAXX in APL (Zhu et al. 2005). Additionally, a multimerisation-prone DAXX-RARα fusion protein can cause transformation, inhibition of transcription and differentiation of HPCs in vitro (Zhou et al. 2006). Of note, recent reports identified DAXX as a histone chaperone that loads the histone variant H3.3 on chromatin in conjunction with ATRX, suggesting that PML-RARα could cause chromatin remodelling through histone variant loading (Salomoni 2013; Drane et al. 2010). H3.3 is encoded by two genes, *H3F3A* and *H3F3B*, and differs from canonical H3 histones in replication independence, chaperone choice, chromatin localization and post-translational modifications (Szenker et al. 2011; Salomoni 2013; Skene and Henikoff 2013). Studies from our group and others suggest that H3.3 acts as an important carrier of epigenetic information, with implications for transcription and telomere maintenance (Goldberg et al. 2010; Drane et al. 2010; Michod et al. 2012; Adam et al. 2013; Pchelintsev et al. 2013). Deposition of H3.3 is partly associated with active chromatin and a recent report showed H3.3 loading in mouse embryonic stem cells at RA-response genes (Chen et al. 2013; Henikoff 2008). In resting state, H3.3 accumulates at enhancers of RA-response genes, but is depleted upon RA treatment. The depletion may open chromatin for RARα/RXRα binding and during the subsequent gene activation H3.3 accumulates at the promoter. This implicates that H3.3 may be important in gene activation of RA-inducible genes (Chen et al. 2013). Further support for the hypothesis that inducible genes are regulated by H3.3 deposition comes from a study from our group, which implicated loading of H3.3 at regulatory elements of immediate early genes, such as Fos and Jun, in regulation of transcription upon neuronal activation by calcium signalling (Michod et al. 2012). Very interestingly, Fos and Jun were shown to be dysregulated in PML-RARα expressing HPCs in comparison to preleukaemic myeloid cells (Yuan et al. 2007), suggesting that a PML-RARα/DAXX/H3.3 interaction might be implicated in this transcriptional dysregulation. In normal cells, DAXX is required for localisation of H3.3 to PML-NBs (Delbarre et al. 2013) however these structures are destroyed in APL cells thereby potentially changing the normal H3.3/PML-NB biology. A recent report showed that PML-RARα expression causes disruption of the normal PML/DAXX/ATRX complex (Korf et al. 2014). Further transcriptional dysregulation may result from histone modifications. For instance, H3.3 is enriched in the H3K4me3 active

mark (Henikoff 2008) and PML-RARα targets were shown to be enriched in H3K4me3 (Hoemme et al. 2008). Furthermore, H3.3 deposition by the chaperone HIRA has been proposed to promote H3K27 trimethylation at Polycomb targets (Banaszynski et al. 2013). This indicates that PML-RARα expressing progenitor cells may sustain bivalent status of genes that normally lose the H3K4me3 and/or H3K27me3 marks during differentiation (Cui et al. 2009). The fact that H3.3 is found at PcG targets and PcG targets often contain RAREs suggests that PML-RARα/DAXX-mediated histone variant loading may change the activity of bivalent loci as it is observed in haematopoietic neoplasms (Muntean and Hess 2012; Maze et al. 2014). Notably, H3.3 and/or DAXX/ATRX are found mutated in different human neoplasms, such as glioblastoma, AML and bone tumours, thus providing strong evidence for a role of this histone variant and its loading machinery in tumourigenesis (Schwartzentruber et al. 2012; Wu et al. 2012; Behjati et al. 2013; Ding et al. 2012) [for complete information refer to Yuen and Knoepfler (2013), Skene and Henikoff (2013)].

H3.3 is frequently found associated with the H2A histone variant H2A.Z in heterotypic nucleosomes (Jin et al. 2009). In this respect a bioinformatic cluster analysis of restriction enzyme accessible regions in NB4 cells showed one cluster with enrichment of H2A.Z and acetylated H2A.Z at highly accessible regions of chromatin (Saeed et al. 2012). The separation of H2A.Z as a separate cluster is indicative for an important role of histone variant replacement in APL. The second cluster showed obvious differences in histone marks and could be subdivided in five groups (Saeed et al. 2012). One group identified increased levels of H2A.Z, H2A.Zac, H3ac and RNAPII at promoter regions (Saeed et al. 2012). Further subgroups showed that genes involved in proliferation were preferentially enriched in H2A.Zac and H3ac levels, while genes involved in signal transduction and communication showed decreased H2A.Zac and H3ac levels (Saeed et al. 2012). These findings indicate that histone variant loading could play a role in increased cell proliferation and inhibited differentiation of PML-RARα expressing progenitor cells, for instance via changes of histone tail modifications at bivalent loci. Overall, the existing literature provides evidence for an important role of PML-RARα/DAXX-mediated chromatin remodelling via interaction with epigenetic modifiers and histone variant replacement, thereby affecting transcription and contributing to APL pathogenesis.

3.3 DNA Methylation

Changes in DNA methylation have been reported in many cancer types and have been implicated in disease pathogenesis and/or used for diagnosis/prognosis (see Box 1) (Dawson and Kouzarides 2012). APL patient samples exhibit a genome-wide aberrant methylation pattern as well as specific methylation profiles that allow discrimination from other myeloid leukaemias based on methylation profile clustering (Schoofs et al. 2013; Figueroa et al. 2010). Several ChIP-sequencing,

ChIP-chip and microarray studies on NB4 and U937 cells revealed almost 3000 PML-RARα binding sites (Martens et al. 2010; Wang et al. 2010; Schoofs et al. 2013) and an increase in methylation of CpG islands of NB4 cells over normal peripheral blood mononuclear cells (Nouzova et al. 2004). The key questions are (1) whether PML-RARα is involved in regulation of DNA methylation in APL and (2) if changes in DNA methylation contribute to APL pathogenesis. Contradictory data exist regarding the influence of PML-RARα binding to DNA and its ability to recruit epigenetic modifiers on the hypermethylation profile observed in APL patient cells. This is most likely due to the fact that primary studies on the effect of PML-RARα on CpG methylation were performed on only a subset of selected gene promoters such as RARβ2. Physical interaction of DNMT3a and DNMT1 with PML-RARα was shown by co-immunoprecipitations at the RARβ2 promoter in PML-RARα-inducible U937 and NB4 cells (Di Croce et al. 2002) (see Fig. 2). In line with an enrichment of DNMTs at the RARβ2 promoter, binding of PML-RARα to RARβ2 promoter was shown to be accompanied by an increase in CpG methylation in PML-RARα-inducible U937, NB4 and APL patient cells, while healthy control cells showed no increased methylation (Di Croce et al. 2002; Fazi et al. 2005). In contrast, down-regulation of DNMTs results in increased expression of RARβ2 (Fazi et al. 2005). Another set of proteins associated with methylated DNA in cancer are the methyl-binding proteins (MBDs) (see Box 1) (Dawson and Kouzarides 2012). MBD2 and MBD3 are associated with the nucleosome remodelling and histone deacetylase (NuRD) complex and all together contribute to the repression and remodelling of methylated genes and chromatin (Villa et al. 2004; Morey et al. 2008). PML-RARα is associated with all the components of the NuRD complex and down-regulation of MBD3 in NB4 cells disrupts the NuRD complex formation resulting in increased differentiation of NB4 cells (Morey et al. 2008). Concomitantly, CpG methylation is reduced at the RARβ2 promoter in NB4 and U937 cells (Morey et al. 2008). Based on these findings, interactions of PML-RARα with de-novo DNA methylating proteins as well as methyl binding proteins were supposed to contribute to disease initiation and the global transcriptional silencing at PML-RARα target promoters (Di Croce et al. 2002). However, recent publications on a genome-wide scale have revealed a different picture of methylation at PML-RARα binding sites. Even though on a genome-wide scale APL patient samples are hypermethylated compared to control cells, this methylation profile does not correlate with PML-RARα binding sites (Schoofs et al. 2013). Aberrant methylation occurs across all chromosomes, but is especially enriched to chromosome ends (Schoofs et al. 2013). Moreover, the hypermethylated pattern is significantly overrepresented in gene bodies, while gene promoters are underrepresented (Schoofs et al. 2013). Based on the study of Schoofs and colleagues, PML-RARα is associated with open chromatin and blocks its DNA binding sites from hypermethylation (Schoofs et al. 2013). Hence, inhibition of DNA methylation could be an effect of PML-RARα binding to its target promoters thereby rather blocking than inducing recruitment of DNMTs. Furthermore, Schoofs et al. also report that (1) DNA methylation changes are minimal at disease presentation and (2) PML-RARα knock-in mice display little DNA

methylation variations in preleukaemic cells (Schoofs et al. 2013). Therefore, methylation seems not to be an initial driver of APL, but might contribute to the disease phenotype at later stages (Schoofs et al. 2013). However, transplantation of bone marrow cells overexpressing PML-RARα together with DNMT3a1 into irradiated recipients reduces the latency in leukaemia development and results in greater penetrance compared to leukaemia induced solely by PML-RARα (Subramanyam et al. 2010). Enhanced methylation at the RARβ promoter indicates a cooperative effect of PML-RARα with epigenetic modifiers in this in vivo model (Subramanyam et al. 2010). Such contradictory results may be explained by a combination of both, a methylation of PML-RARα target promoters by PML-RARα's direct interaction with DNMTs and a methylation pattern established by secondary events caused through PML-RARα gene silencing. Even though DNA methylation seems not to be implicated in leukaemia establishment, analysis of DNA methylation is thus an important investigation in APL pathology, especially as hypermethylation correlates with more aggressive disease (Schoofs et al. 2013). Aberrant methylation of genes, such as p15 and p16, are known to be associated with poor prognosis for APL patients and an increased risk of relapse (Teofili et al. 2003). Given this negative correlation of DNA methylation on patient survival, it is of importance to analyse the causative events of DNA modifications independent of its missing link to leukaemia initiation.

4 Is Chromatin Remodelling a Viable Option for APL Therapy?

The most profound epigenetic effect of ATRA treatment on APL cell lines and patient samples is a marked increase of acetylation of histones H3 and H4, while changes in the methylation profile are minor (Nouzova et al. 2004; Fazi et al. 2005; Villa et al. 2007; Morey et al. 2008; Martens et al. 2010; Schoofs et al. 2013). Pharmacological doses of ATRA release both the NuRD complex and HDAC, subsequently recruiting HATs such as CBP/p300 (Morey et al. 2008; Hoemme et al. 2008). Interestingly, genes that are induced upon ATRA treatment display higher levels of H3 acetylation compared to genes that are repressed upon ATRA (Martens et al. 2010). Mainly genes related to differentiation, development and signal transduction are induced by ATRA, while genes associated with cell metabolism are repressed (Martens et al. 2010; Schoofs et al. 2013; Hoemme et al. 2008; Di Croce et al. 2002). Increased levels of H3K9ac and H3K14ac were shown at PML-RARα target sites, while H3K27me3, H3K9me3 and DNA methylation were relatively stable in comparison with untreated cells (Martens et al. 2010). However, studies performed on single gene promoters such as RARβ2 revealed that RA treatment reduces DNMT expression and correlates with decreased DNA methylation and H3K27me3, again indicating the previously discussed cross-talk between DNA and histone methylation (Di Croce et al. 2002). The fact that changes in

H3K27me3 levels were minor suggests that ATRA treatment has no positive effect through epigenetic modifications. However, the homeobox (HOX) cluster of genes that is known to regulate haematopoietic differentiation is enriched in H3K27me3 after ATRA treatment indicating that gene specific changes in methylation occur at bivalent loci (Martens et al. 2010). Very interestingly, upon ATRA treatment in NB4 cells PML-RARα interacts via the RARα moiety with PHF8, a JmjC domain containing protein that preferentially demethylates H3K9me1 and H3K9me2 (Arteaga et al. 2013). Recruitment of PHF8 to the RARβ2 promoter resulted in reduced H3K9me2 repressive mark, but increased H3K4me3 and H3K9ac active marks together with increased mRNA levels of RARβ2 (Arteaga et al. 2013). Based on all these findings chemical reagents were analysed in multiple studies in regards of their influence on cellular differentiation in cell culture models. For example, the DNMT inhibitor 5-aza-dC can act synergistically with RA in reduction of the methylation status at RARβ2 promoter in NB4 cells (Di Croce et al. 2002). Further studies with 5-Aza-dC and the histone deacetylase inhibitor Trichostatin A in NB4 and APL patient cells revealed synergistic effects of the compounds together with ATRA, while single reagents were not effective (Fazi et al. 2005). Of note, 5-aza-dC treatment reduced binding of PRC2 components at RARβ2 promoter concomitantly with a decrease in H3K27me2 and H3K27me3 (Villa et al. 2007). Therefore, combinations of the classical APL therapeutics with demethylating agents and histone deacetylase inhibitors could be valuable additions to patient therapy due to the reversibility of epigenetic marks (Martens et al. 2010; Lin et al. 1998; Tabe et al. 2006). Especially HDAC inhibitors are of interest for drug development due to the major changes in acetylation, when combined with ATRA. Future studies will show if combinations with epigenetic modifying agents will further improve the outcome of standard ATRA and As_2O_3 leukaemia therapy, especially for patients undergoing relapse.

5 Conclusions and Outstanding Questions

Overall, the epigenetic modifications triggered either directly or indirectly by PML-RARα have been proposed to play an important role in APL pathogenesis and the reversion of APL epigenetic state is of prevalent interest. However, there are a number of outstanding questions that could improve our understanding of the epigenetic function of PML-RARα and other oncogenic fusion proteins found in AML:

– Is PML-RARα epigenetic function involved in disease initiation? We believe there is lack of studies directly assessing the role of epigenetic reprogramming during early phases of APL pathogenesis, i.e. the preleukaemic phase. For instance, one could hypothesise that epigenetic changes are necessary for the transition from the preleukaemic state to full blown leukaemia. On the other hand, one could argue that PML-RARα is primarily a transcription factor (either

repressor or activator) and that the interaction with epigenetic modifiers contributes only in part to the transcriptional changes and neoplastic transformation.

- Are the APL cells addicted to PML-RARα with respect to their epigenetic status? It is conceivable that a number of epigenetic changes could be triggered by PML-RARα during the first steps of neoplastic transformation but they could become PML-RARα–independent in established tumours. This would imply that in patients not responding to frontline therapy epigenetic changes could contribute to disease dormancy during treatment with agents inducing degradation of the fusion protein.
- Are H3.3 and other histone variants involved in APL? Findings from Hugues de The's group indicate that interaction with the H3.3 chaperone DAXX is important for transformation in vitro (Zhou et al. 2006). What remains to be established is whether DAXX and its chaperone activity are themselves involved in transformation and/or tumour progression. In this respect, it is possible that other chromatin-remodelling functions attributed to DAXX, such as interaction with DNMT1 (Puto and Reed 2008), are required for PML-RARα epigenetic role. Further, as DAXX has been reported to promote loading of CenH3 at euchromatin in cancer cells (Lacoste et al. 2014), it would be interesting to test whether PML-RARα causes alterations of CenH3 deposition as well. Finally, it is possible that H3.3-driven chromatin remodelling may play a role in other AML subtypes.
- Does PML-RARα have other functions in regulation of genome topology that could be relevant for transcriptional regulation and transformation? It is becoming clear that genome organisation within the mammalian nucleus plays an important role in regulation of transcription and genome maintenance (Cavalli and Misteli 2013). In this respect, PML-RARα disrupts the PML-NBs and forms microspeckled domains that have not been studied in depth. It cannot be excluded that these novel domains could be involved in genome organisation [see also in Torok et al. (2009)]. This is an aspect of PML-RARα function that has not been investigated, but recent technological advances in the field have made analysis of genome topology changes possible.
- Would APL fusion proteins other than PML-RARα and PLZF-RARα (not discussed in this review; see Boukarabila et al. (2009), Spicuglia et al. (2011)] carry epigenetic functions relevant for disease pathogenesis? Related to this, would other fusion proteins found in leukaemia work by similar epigenetic mechanisms and for instance affect histone variant loading? In this respect, MLL fusion proteins have been reported to reprogramme the epigenome via mechanisms involving regulation of bivalency, which is clearly linked to loading of the H3.3 and H2A.Z histone variants.

Box 1: Epigenetics

- Epigenetic modifications include de-/acetylation and de-/methylation of histone tail residues, histone variant replacement by chromatin remodelling factors and DNA methylation.
- Histone tail modifications occur at lysine or arginine residues in H2A, H2B, H3 and H4. Different methyltransferases, acetyltransferases, demethylases and deacetylases are responsible for chemical modification of specific residues (see Table 1). Acetylation changes the charge of chromatin, thereby leading to a more accessible chromatin conformation and a general association with active transcription, while histone deacetylases revert this effect. Methylated lysine residues can be active or repressive regulators of transcription depending on the specific residue and the number of bound methyl-groups (mono-, di- or trimethylation). Methylated H3K9 and H3K27 are associated with gene repression, while methylated H3K4 is associated with active genes. Bivalent loci contain both active H3K4me3 and repressive H3K27me3 marks and play an important role in stem cell differentiation. Specific chromatin readers bind to the distinct histone marks thereby regulating gene expression, for instance PHF8 binds to H3K4me3, HP1α to H3K9me3, polycomb repressive complex 2 subunit Suz12 to H3K27me3. Polycomb group and Trithorax group proteins regulate transcription through establishment/eradication of epigenetic marks, with Polycomb group proteins being associated with repression and Trithorax group proteins with gene activation (Barski et al. 2007; Henikoff 2008; Bartke et al. 2010).
- Different variants of histones—such as H2A.Z/ H2A.X for H2A and H3.3/ CenH3 for H3—exist that differ from the canonical proteins by a few amino acid changes. However, these sequence variations confer slightly different chemical behaviour, thereby influencing the nucleosome stability, chromatin accessibility and transcriptional state. Different chaperone complexes exist that load histone variants onto chromatin, for instance HIRA/ASF1 and ATRX/DAXX for H3.3 (Henikoff 2008; Maze et al. 2014).
- DNA methylation primarily occurs at CpG dinucleotides by covalent binding of a methyl-group to cytosine. The DNA methyltransferase (DNMT) protein family includes DNMT1, DNMT3a and DNMT3b. DNMT1 is linked to maintenance of DNA methylation, while DNMT3a and DNMT3b are de-novo DNA methylating proteins. MBD domain proteins, Kaiso and Kaiso-like proteins and SRA domain proteins are known to specifically bind to methylated DNA, inducing a repressive or silencing transcriptional state via interaction with other protein complexes (Villa et al. 2004; Di Croce et al. 2004; Clouaire and Stancheva 2008).

(continued)

– Most cancer cells have a global reduction of methylated CpGs in comparison to normal cells; however cancer cells exhibit an increased methylation of CpGs across promoter regions (Villa et al. 2004; Dawson and Kouzarides 2012).

Acknowledgements We thank our collaborator Hugues de The for support and critical discussion. PS lab is supported by the Worldwide Cancer Research, the European Research Council, Medical Research Council, The Brain Tumour Charity. JPH is recipient of a Research Fellowship from the Deutsche Forschungsgemeinschaft (German Research Foundation).

References

Ablain J, Rice K, Soilihi H, de Reynies A, Minucci S, de The H (2014) Activation of a promyelocytic leukemia-tumor protein 53 axis underlies acute promyelocytic leukemia cure. Nat Med 20(2):167–174. doi:10.1038/nm.3441

Adam S, Polo SE, Almouzni G (2013) Transcription recovery after DNA damage requires chromatin priming by the H3.3 histone chaperone HIRA. Cell 155(1):94–106. doi:10.1016/j.cell.2013.08.029

Arnould C, Philippe C, Bourdon V, Gr goire MJ, Berger R, Jonveaux P (1999) The signal transducer and activator of transcription STAT5b gene is a new partner of retinoic acid receptor alpha in acute promyelocytic-like leukaemia. Hum Mol Genet 8(9):1741–1749

Arteaga MF, Mikesch JH, Qiu J, Christensen J, Helin K, Kogan SC, Dong S, So CW (2013) The histone demethylase PHF8 governs retinoic acid response in acute promyelocytic leukemia. Cancer Cell 23(3):376–389. doi:10.1016/j.ccr.2013.02.014, S1535-6108(13)00070-6 [pii]

Banaszynski LA, Wen D, Dewell S, Whitcomb SJ, Lin M, Diaz N, Elsasser SJ, Chapgier A, Goldberg AD, Canaani E, Rafii S, Zheng D, Allis CD (2013) Hira-dependent histone H3.3 deposition facilitates PRC2 recruitment at developmental loci in ES cells. Cell 155 (1):107–120. doi:10.1016/j.cell.2013.08.061

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K (2007) High-resolution profiling of histone methylations in the human genome. Cell 129(4):823–837. doi:10.1016/j.cell.2007.05.009

Bartke T, Vermeulen M, Xhemalce B, Robson SC, Mann M, Kouzarides T (2010) Nucleosome-interacting proteins regulated by DNA and histone methylation. Cell 143(3):470–484. doi:10.1016/j.cell.2010.10.012

Behjati S, Tarpey PS, Presneau N, Scheipl S, Pillay N, Van Loo P, Wedge DC, Cooke SL, Gundem G, Davies H, Nik-Zainal S, Martin S, McLaren S, Goodie V, Robinson B, Butler A, Teague JW, Halai D, Khatri B, Myklebost O, Baumhoer D, Jundt G, Hamoudi R, Tirabosco R, Amary MF, Futreal PA, Stratton MR, Campbell PJ, Flanagan AM (2013) Distinct H3F3A and H3F3B driver mutations define chondroblastoma and giant cell tumor of bone. Nat Genet 45 (12):1479–1482. doi:10.1038/ng.2814, ng.2814 [pii]

Bernardi R, Pandolfi PP (2007) Structure, dynamics and functions of promyelocytic leukaemia nuclear bodies. Nat Rev Mol Cell Biol 8(12):1006–1016. doi:10.1038/nrm2277

Bernardi R, Scaglioni PP, Bergmann S, Horn HF, Vousden KH, Pandolfi PP (2004) PML regulates p53 stability by sequestering Mdm2 to the nucleolus. Nat Cell Biol 6(7):665–672. doi:10.1038/ncb1147

Borden KL, Boddy MN, Lally J, O'Reilly NJ, Martin S, Howe K, Solomon E, Freemont PS (1995) The solution structure of the RING finger domain from the acute promyelocytic leukaemia proto-oncoprotein PML. EMBO J 14(7):1532–1541

Boukarabila H, Saurin AJ, Batsche E, Mossadegh N, van Lohuizen M, Otte AP, Pradel J, Muchardt C, Sieweke M, Duprez E (2009) The PRC1 Polycomb group complex interacts with PLZF/RARA to mediate leukemic transformation. Genes Dev 23(10):1195–1206. doi:10.1101/gad.512009, 23/10/1195 [pii]

Burkhard P, Stetefeld J, Strelkov SV (2001) Coiled coils: a highly versatile protein folding motif. Trends Cell Biol 11(2):82–88, S0962-8924(00)01898-5 [pii]

Carbone R, Botrugno OA, Ronzoni S, Insinga A, Di Croce L, Pelicci PG, Minucci S (2006) Recruitment of the histone methyltransferase SUV39H1 and its role in the oncogenic properties of the leukemia-associated PML-retinoic acid receptor fusion protein. Mol Cell Biol 26 (4):1288–1296. doi:10.1128/MCB.26.4.1288-1296.2006, 26/4/1288 [pii]

Carracedo A, Weiss D, Leliaert AK, Bhasin M, de Boer VC, Laurent G, Adams AC, Sundvall M, Song SJ, Ito K, Finley LS, Egia A, Libermann T, Gerhart-Hines Z, Puigserver P, Haigis MC, Maratos-Flier E, Richardson AL, Schafer ZT, Pandolfi PP (2012) A metabolic prosurvival role for PML in breast cancer. J Clin Invest 122(9):3088–3100. doi:10.1172/jci62129

Castaigne S, Chomienne C, Daniel MT, Ballerini P, Berger R, Fenaux P, Degos L (1990) All-trans retinoic acid as a differentiation therapy for acute promyelocytic leukemia. I. Clinical results. Blood 76(9):1704–1709

Catalano A, Dawson MA, Somana K, Opat S, Schwarer A, Campbell LJ, Iland H (2007) The PRKAR1A gene is fused to RARA in a new variant acute promyelocytic leukemia. Blood 110 (12):4073–4076. doi:10.1182/blood-2007-06-095554

Cavalli G, Misteli T (2013) Functional implications of genome topology. Nat Struct Mol Biol 20 (3):290–299. doi:10.1038/nsmb.2474

Chan IT, Kutok JL, Williams IR, Cohen S, Moore S, Shigematsu H, Ley TJ, Akashi K, Le Beau MM, Gilliland DG (2006) Oncogenic K-ras cooperates with PML-RAR alpha to induce an acute promyelocytic leukemia-like disease. Blood 108(5):1708–1715. doi:10.1182/blood-2006-04-015040

Chen Z, Brand NJ, Chen A, Chen SJ, Tong JH, Wang ZY, Waxman S, Zelent A (1993) Fusion between a novel Kruppel-like zinc finger gene and the retinoic acid receptor-alpha locus due to a variant t(11;17) translocation associated with acute promyelocytic leukaemia. EMBO j 12 (3):1161–1167

Chen GQ, Shi XG, Tang W, Xiong SM, Zhu J, Cai X, Han ZG, Ni JH, Shi GY, Jia PM, Liu MM, He KL, Niu C, Ma J, Zhang P, Zhang TD, Paul P, Naoe T, Kitamura K, Miller W, Waxman S, Wang ZY, de The H, Chen SJ, Chen Z (1997) Use of arsenic trioxide (As2O3) in the treatment of acute promyelocytic leukemia (APL): I. As2O3 exerts dose-dependent dual effects on APL cells. Blood 89(9):3345–3353

Chen P, Zhao J, Wang Y, Wang M, Long H, Liang D, Huang L, Wen Z, Li W, Li X, Feng H, Zhao H, Zhu P, Li M, Wang QF, Li G (2013) H3.3 actively marks enhancers and primes gene transcription via opening higher-ordered chromatin. Genes Dev 27(19):2109–2124. doi:10.1101/gad.222174.113, gad.222174.113 [pii]

Clouaire T, Stancheva I (2008) Methyl-CpG binding proteins: specialized transcriptional repressors or structural components of chromatin? Cell Mol Life Sci 65(10):1509–1522. doi:10.1007/s00018-008-7324-y

Condemine W, Takahashi Y, Zhu J, Puvion-Dutilleul F, Guegan S, Janin A, de The H (2006) Characterization of endogenous human promyelocytic leukemia isoforms. Cancer Res 66 (12):6192–6198. doi:10.1158/0008-5472, CAN-05-3792. 66/12/6192 [pii]

Cui K, Zang C, Roh TY, Schones DE, Childs RW, Peng W, Zhao K (2009) Chromatin signatures in multipotent human hematopoietic stem cells indicate the fate of bivalent genes during differentiation. Cell Stem Cell 4(1):80–93. doi:10.1016/j.stem.2008.11.011

Daniel MT, Koken M, Romagne O, Barbey S, Bazarbachi A, Stadler M, Guillemin MC, Degos L, Chomienne C, de The H (1993) PML protein expression in hematopoietic and acute promyelocytic leukemia cells. Blood 82(6):1858–1867

Dawson MA, Kouzarides T (2012) Cancer epigenetics: from mechanism to therapy. Cell 150 (1):12–27. doi:10.1016/j.cell.2012.06.013, S0092-8674(12)00762-3 [pii]

de The H, Chen Z (2010) Acute promyelocytic leukaemia: novel insights into the mechanisms of cure. Nat Rev Cancer 10(11):775–783. doi:10.1038/nrc2943, nrc2943 [pii]

de The H, Chomienne C, Lanotte M, Degos L, Dejean A (1990a) The t(15;17) translocation of acute promyelocytic leukaemia fuses the retinoic acid receptor alpha gene to a novel transcribed locus. Nature 347(6293):558–561. doi:10.1038/347558a0

de The H, Vivanco-Ruiz MM, Tiollais P, Stunnenberg H, Dejean A (1990b) Identification of a retinoic acid responsive element in the retinoic acid receptor beta gene. Nature 343 (6254):177–180. doi:10.1038/343177a0

de The H, Le Bras M, Lallemand-Breitenbach V (2012) The cell biology of disease: acute promyelocytic leukemia, arsenic, and PML bodies. J Cell Biol 198(1):11–21. doi:10.1083/jcb.201112044, jcb.201112044 [pii]

Delbarre E, Ivanauskiene K, Kuntziger T, Collas P (2013) DAXX-dependent supply of soluble (H3.3-H4) dimers to PML bodies pending deposition into chromatin. Genome Res 23 (3):440–451. doi:10.1101/gr.142703.112

Dellaire G, Bazett-Jones DP (2004) PML nuclear bodies: dynamic sensors of DNA damage and cellular stress. Bioessays 26(9):963–977. doi:10.1002/bies.20089

Di Croce L, Raker VA, Corsaro M, Fazi F, Fanelli M, Faretta M, Fuks F, Lo Coco F, Kouzarides T, Nervi C, Minucci S, Pelicci PG (2002) Methyltransferase recruitment and DNA hypermethylation of target promoters by an oncogenic transcription factor. Science 295 (5557):1079–1082. doi:10.1126/science.1065173, 295/5557/1079 [pii]

Di Croce L, Buschbeck M, Gutierrez A, Joval I, Morey L, Villa R, Minucci S (2004) Altered epigenetic signals in human disease. Cancer Biol Ther 3(9):831–837, 1103 [pii]

Ding L, Ley TJ, Larson DE, Miller CA, Koboldt DC, Welch JS, Ritchey JK, Young MA, Lamprecht T, McLellan MD, McMichael JF, Wallis JW, Lu C, Shen D, Harris CC, Dooling DJ, Fulton RS, Fulton LL, Chen K, Schmidt H, Kalicki-Veizer J, Magrini VJ, Cook L, McGrath SD, Vickery TL, Wendl MC, Heath S, Watson MA, Link DC, Tomasson MH, Shannon WD, Payton JE, Kulkarni S, Westervelt P, Walter MJ, Graubert TA, Mardis ER, Wilson RK, DiPersio JF (2012) Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. Nature 481(7382):506–510. doi:10.1038/nature10738, nature10738 [pii]

Drane P, Ouararhni K, Depaux A, Shuaib M, Hamiche A (2010) The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. Genes Dev 24(12):1253–1265. doi:10.1101/gad.566910, gad.566910 [pii]

Du C, Redner RL, Cooke MP, Lavau C (1999) Overexpression of wild-type retinoic acid receptor alpha (RARalpha) recapitulates retinoic acid-sensitive transformation of primary myeloid progenitors by acute promyelocytic leukemia RARalpha-fusion genes. Blood 94(2):793–802

Estey E, Garcia-Manero G, Ferrajoli A, Faderl S, Verstovsek S, Jones D, Kantarjian H (2006) Use of all-trans retinoic acid plus arsenic trioxide as an alternative to chemotherapy in untreated acute promyelocytic leukemia. Blood 107(9):3469–3473. doi:10.1182/blood-2005-10-4006, 2005-10-4006 [pii]

Fazi F, Travaglini L, Carotti D, Palitti F, Diverio D, Alcalay M, McNamara S, Miller WH Jr, Lo Coco F, Pelicci PG, Nervi C (2005) Retinoic acid targets DNA-methyltransferases and histone deacetylases during APL blast differentiation in vitro and in vivo. Oncogene 24 (11):1820–1830. doi:10.1038/sj.onc.1208286, 1208286 [pii]

Figueroa ME, Lugthart S, Li Y, Erpelinck-Verschueren C, Deng X, Christos PJ, Schifano E, Booth J, van Putten W, Skrabanek L, Campagne F, Mazumdar M, Greally JM, Valk PJ, Lowenberg B, Delwel R, Melnick A (2010) DNA methylation signatures identify biologically

distinct subtypes in acute myeloid leukemia. Cancer Cell 17(1):13–27. doi:10.1016/j.ccr.2009.11.020, S1535-6108(09)00420-6 [pii]

Fu C, Ahmed K, Ding H, Ding X, Lan J, Yang Z, Miao Y, Zhu Y, Shi Y, Zhu J, Huang H, Yao X (2005) Stabilization of PML nuclear localization by conjugation and oligomerization of SUMO-3. Oncogene 24(35):5401–5413. doi:10.1038/sj.onc.1208714, 1208714 [pii]

Goldberg AD, Banaszynski LA, Noh KM, Lewis PW, Elsaesser SJ, Stadler S, Dewell S, Law M, Guo X, Li X, Wen D, Chapgier A, DeKelver RC, Miller JC, Lee YL, Boydston EA, Holmes MC, Gregory PD, Greally JM, Rafii S, Yang C, Scambler PJ, Garrick D, Gibbons RJ, Higgs DR, Cristea IM, Urnov FD, Zheng D, Allis CD (2010) Distinct factors control histone variant H3.3 localization at specific genomic regions. Cell 140(5):678–691. doi:10.1016/j.cell.2010.01.003, S0092-8674(10)00004-8 [pii]

Grignani F, De Matteis S, Nervi C, Tomassoni L, Gelmetti V, Cioce M, Fanelli M, Ruthardt M, Ferrara FF, Zamir I, Seiser C, Lazar MA, Minucci S, Pelicci PG (1998) Fusion proteins of the retinoic acid receptor-alpha recruit histone deacetylase in promyelocytic leukaemia. Nature 391(6669):815–818. doi:10.1038/35901

Guibal FC, Alberich-Jorda M, Hirai H, Ebralidze A, Levantini E, Di Ruscio A, Zhang P, Santana-Lemos BA, Neuberg D, Wagers AJ, Rego EM, Tenen DG (2009) Identification of a myeloid committed progenitor as the cancer-initiating cell in acute promyelocytic leukemia. Blood 114(27):5415–5425. doi:10.1182/blood-2008-10-182071, blood-2008-10-182071 [pii]

He LZ, Guidez F, Tribioli C, Peruzzi D, Ruthardt M, Zelent A, Pandolfi PP (1998) Distinct interactions of PML-RARalpha and PLZF-RARalpha with co-repressors determine differential responses to RA in APL. Nat Genet 18(2):126–135. doi:10.1038/ng0298-126

Henikoff S (2008) Nucleosome destabilization in the epigenetic regulation of gene expression. Nat Rev Genet 9(1):15–26. doi:10.1038/nrg2206, nrg2206 [pii]

Hoemme C, Peerzada A, Behre G, Wang Y, McClelland M, Nieselt K, Zschunke M, Disselhoff C, Agrawal S, Isken F, Tidow N, Berdel WE, Serve H, Muller-Tidow C (2008) Chromatin modifications induced by PML-RARalpha repress critical targets in leukemogenesis as analyzed by ChIP-Chip. Blood 111(5):2887–2895. doi:10.1182/blood-2007-03-079921, blood-2007-03-079921

Hollenbach AD, McPherson CJ, Mientjes EJ, Iyengar R, Grosveld G (2002) Daxx and histone deacetylase II associate with chromatin through an interaction with core histones and the chromatin-associated protein Dek. J Cell Sci 115(Pt 16):3319–3330

Ishov AM, Sotnikov AG, Negorev D, Vladimirova OV, Neff N, Kamitani T, Yeh ET, Strauss JF III, Maul GG (1999) PML is critical for ND10 formation and recruits the PML-interacting protein daxx to this nuclear structure when modified by SUMO-1. J Cell Biol 147(2):221–234

Ito K, Bernardi R, Morotti A, Matsuoka S, Saglio G, Ikeda Y, Rosenblatt J, Avigan DE, Teruya-Feldstein J, Pandolfi PP (2008) PML targeting eradicates quiescent leukaemia-initiating cells. Nature 453(7198):1072–1078. doi:10.1038/nature07016, nature07016 [pii]

Ito K, Carracedo A, Weiss D, Arai F, Ala U, Avigan DE, Schafer ZT, Evans RM, Suda T, Lee CH, Pandolfi PP (2012) A PML-PPAR-delta pathway for fatty acid oxidation regulates hematopoietic stem cell maintenance. Nat Med 18(9):1350–1358. doi:10.1038/nm.2882

Jeanne M, Lallemand-Breitenbach V, Ferhi O, Koken M, Le Bras M, Duffort S, Peres L, Berthier C, Soilihi H, Raught B, de The H (2010) PML/RARA oxidation and arsenic binding initiate the antileukemia response of As2O3. Cancer Cell 18(1):88–98. doi:10.1016/j.ccr.2010.06.003

Jensen K, Shiels C, Freemont PS (2001) PML protein isoforms and the RBCC/TRIM motif. Oncogene 20(49):7223–7233. doi:10.1038/sj.onc.1204765

Jin C, Zang C, Wei G, Cui K, Peng W, Zhao K, Felsenfeld G (2009) H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. Nat Genet 41(8):941–945. doi:10.1038/ng.409

Joazeiro CA, Weissman AM (2000) RING finger proteins: mediators of ubiquitin ligase activity. Cell 102(5):549–552, S0092-8674(00)00077-5 [pii]

Kamashev D, Vitoux D, De The H (2004) PML-RARA-RXR oligomers mediate retinoid and rexinoid/cAMP cross-talk in acute promyelocytic leukemia cell differentiation. J Exp Med 199 (8):1163–1174. doi:10.1084/jem.20032226, jem.20032226 [pii]

Kastner P, Chan S (2001) Function of RARalpha during the maturation of neutrophils. Oncogene 20(49):7178–7185. doi:10.1038/sj.onc.1204757

Kastner P, Lawrence HJ, Waltzinger C, Ghyselinck NB, Chambon P, Chan S (2001) Positive and negative regulation of granulopoiesis by endogenous RARalpha. Blood 97(5):1314–1320

Kelly LM, Kutok JL, Williams IR, Boulton CL, Amaral SM, Curley DP, Ley TJ, Gilliland DG (2002) PML/RARalpha and FLT3-ITD induce an APL-like disease in a mouse model. Proc Natl Acad Sci U S A 99(12):8283–8288. doi:10.1073/pnas.122233699

Kogan SC (2007) Mouse models of acute promyelocytic leukemia. Curr Top Microbiol Immunol 313:3–29

Kogan SC, Hong SH, Shultz DB, Privalsky ML, Bishop JM (2000) Leukemia initiated by PMLRARalpha: the PML domain plays a critical role while retinoic acid-mediated transactivation is dispensable. Blood 95(5):1541–1550

Koken MH, Puvion-Dutilleul F, Guillemin MC, Viron A, Linares-Cruz G, Stuurman N, de Jong L, Szostecki C, Calvo F, Chomienne C et al (1994) The t(15;17) translocation alters a nuclear body in a retinoic acid-reversible fashion. EMBO J 13(5):1073–1083

Kondo T, Mori A, Darmanin S, Hashino S, Tanaka J, Asaka M (2008) The seventh pathogenic fusion gene FIP1L1-RARA was isolated from a t(4;17)-positive acute promyelocytic leukemia. Haematologica 93(9):1414–1416. doi:10.3324/haematol.12854

Korf K, Wodrich H, Haschke A, Ocampo C, Harder L, Gieseke F, Pollmann A, Dierck K, Prall S, Staege H, Ma H, Horstmann MA, Evans RM, Sternsdorf T (2014) The PML domain of PML-RARalpha blocks senescence to promote leukemia. Proc Natl Acad Sci USA 111 (33):12133–12138. doi:10.1073/pnas.1412944111, 1412944111 [pii]

Kuo HY, Chang CC, Jeng JC, Hu HM, Lin DY, Maul GG, Kwok RP, Shih HM (2005) SUMO modification negatively modulates the transcriptional activity of CREB-binding protein via the recruitment of Daxx. Proc Natl Acad Sci USA 102(47):16973–16978. doi:10.1073/pnas.0504460102

Kwok C, Zeisig BB, Dong S, So CW (2006) Forced homo-oligomerization of RARalpha leads to transformation of primary hematopoietic cells. Cancer Cell 9(2):95–108. doi:10.1016/j.ccr.2006.01.005, S1535-6108(06)00027-4 [pii]

Lacoste N, Woolfe A, Tachiwana H, Garea AV, Barth T, Cantaloube S, Kurumizaka H, Imhof A, Almouzni G (2014) Mislocalization of the centromeric histone variant CenH3/CENP-A in human cells depends on the chaperone DAXX. Mol Cell 53(4):631–644. doi:10.1016/j.molcel.2014.01.018

Lallemand-Breitenbach V, Zhu J, Puvion F, Koken M, Honore N, Doubeikovsky A, Duprez E, Pandolfi PP, Puvion E, Freemont P, de The H (2001) Role of promyelocytic leukemia (PML) sumolation in nuclear body formation, 11S proteasome recruitment, and As2O3-induced PML or PML/retinoic acid receptor alpha degradation. J Exp Med 193(12):1361–1371

Le Beau MM, Bitts S, Davis EM, Kogan SC (2002) Recurring chromosomal abnormalities in leukemia in PML-RARA transgenic mice parallel human acute promyelocytic leukemia. Blood 99(8):2985–2991

Licht JD (2006) Reconstructing a disease: What essential features of the retinoic acid receptor fusion oncoproteins generate acute promyelocytic leukemia? Cancer Cell 9(2):73–74. doi:10.1016/j.ccr.2006.01.024, S1535-6108(06)00032-8 [pii]

Lin RJ, Evans RM (2000) Acquisition of oncogenic potential by RAR chimeras in acute promyelocytic leukemia through formation of homodimers. Mol Cell 5(5):821–830, S1097-2765(00)80322-6 [pii]

Lin RJ, Nagy L, Inoue S, Shao W, Miller WH Jr, Evans RM (1998) Role of the histone deacetylase complex in acute promyelocytic leukaemia. Nature 391(6669):811–814. doi:10.1038/35895

Lo-Coco F, Avvisati G, Vignetti M, Breccia M, Gallo E, Rambaldi A, Paoloni F, Fioritoni G, Ferrara F, Specchia G, Cimino G, Diverio D, Borlenghi E, Martinelli G, Di Raimondo F, Di

Bona E, Fazi P, Peta A, Bosi A, Carella AM, Fabbiano F, Pogliani EM, Petti MC, Amadori S, Mandelli F (2010) Front-line treatment of acute promyelocytic leukemia with AIDA induction followed by risk-adapted consolidation for adults younger than 61 years: results of the AIDA-2000 trial of the GIMEMA Group. Blood 116(17):3171–3179. doi:10.1182/blood-2010-03-276196, blood-2010-03-276196 [pii]

Lufkin T, Lohnes D, Mark M, Dierich A, Gorry P, Gaub MP, LeMeur M, Chambon P (1993) High postnatal lethality and testis degeneration in retinoic acid receptor alpha mutant mice. Proc Natl Acad Sci USA 90(15):7225–7229

Martens JH, Brinkman AB, Simmer F, Francoijs KJ, Nebbioso A, Ferrara F, Altucci L, Stunnenberg HG (2010) PML-RARalpha/RXR Alters the Epigenetic Landscape in Acute Promyelocytic Leukemia. Cancer Cell 17(2):173–185. doi:10.1016/j.ccr.2009.12.042, S1535-6108(10)00005-X [pii]

Matsushita H, Scaglioni PP, Bhaumik M, Rego EM, Cai LF, Majid SM, Miyachi H, Kakizuka A, Miller WH Jr, Pandolfi PP (2006) In vivo analysis of the role of aberrant histone deacetylase recruitment and RAR alpha blockade in the pathogenesis of acute promyelocytic leukemia. J Exp Med 203(4):821–828. doi:10.1084/jem.20050616, jem.20050616 [pii]

Maze I, Noh KM, Soshnev AA, Allis CD (2014) Every amino acid matters: essential contributions of histone variants to mammalian development and disease. Nat Rev Genet 15(4):259–271. doi:10.1038/nrg3673

Melnick A, Licht JD (1999) Deconstructing a disease: RARalpha, its fusion partners, and their roles in the pathogenesis of acute promyelocytic leukemia. Blood 93(10):3167–3215

Michod D, Bartesaghi S, Khelifi A, Bellodi C, Berliocchi L, Nicotera P, Salomoni P (2012) Calcium-dependent dephosphorylation of the histone chaperone DAXX regulates H3.3 loading and transcription upon neuronal activation. Neuron 74(1):122–135. doi:10.1016/j.neuron.2012.02.021, S0896-6273(12)00188-2 [pii]

Mills AA (2010) Throwing the cancer switch: reciprocal roles of polycomb and trithorax proteins. Nat Rev Cancer 10(10):669–682. doi:10.1038/nrc2931, nrc2931 [pii]

Minucci S, Maccarana M, Cioce M, De Luca P, Gelmetti V, Segalla S, Di Croce L, Giavara S, Matteucci C, Gobbi A, Bianchini A, Colombo E, Schiavoni I, Badaracco G, Hu X, Lazar MA, Landsberger N, Nervi C, Pelicci PG (2000) Oligomerization of RAR and AML1 transcription factors as a novel mechanism of oncogenic activation. Mol Cell 5(5):811–820, S1097-2765(00)80321-4 [pii]

Morey L, Brenner C, Fazi F, Villa R, Gutierrez A, Buschbeck M, Nervi C, Minucci S, Fuks F, Di Croce L (2008) MBD3, a component of the NuRD complex, facilitates chromatin alteration and deposition of epigenetic marks. Mol Cell Biol 28(19):5912–5923. doi:10.1128/MCB.00467-08, MCB.00467-08 [pii]

Muntean AG, Hess JL (2012) The pathogenesis of mixed-lineage leukemia. Annu Rev Pathol 7:283–301. doi:10.1146/annurev-pathol-011811-132434

Nasr R, Guillemin MC, Ferhi O, Soilihi H, Peres L, Berthier C, Rousselot P, Robledo-Sarmiento-M, Lallemand-Breitenbach V, Gourmel B, Vitoux D, Pandolfi PP, Rochette-Egly C, Zhu J, de The H (2008) Eradication of acute promyelocytic leukemia-initiating cells through PML-RARA degradation. Nat Med 14(12):1333–1342. doi:10.1038/nm.1891, nm.1891 [pii]

Nisole S, Maroui MA, Mascle XH, Aubry M, Chelbi-Alix MK (2013) Differential Roles of PML Isoforms. Front Oncol 3:125. doi:10.3389/fonc.2013.00125

Nouzova M, Holtan N, Oshiro MM, Isett RB, Munoz-Rodriguez JL, List AF, Narro ML, Miller SJ, Merchant NC, Futscher BW (2004) Epigenomic changes during leukemia cell differentiation: analysis of histone acetylation and cytosine methylation using CpG island microarrays. J Pharmacol Exp Ther 311(3):968–981. doi:10.1124/jpet.104.072488, jpet.104.072488 [pii]

Onodera M, Kunisada T, Nishikawa S, Sakiyama Y, Matsumoto S, Nishikawa S (1995) Overexpression of retinoic acid receptor alpha suppresses myeloid cell differentiation at the promyelocyte stage. Oncogene 11(7):1291–1298

Pchelintsev NA, McBryan T, Rai TS, van Tuyn J, Ray-Gallet D, Almouzni G, Adams PD (2013) Placing the HIRA histone chaperone complex in the chromatin landscape. Cell Rep 3 (4):1012–1019. doi:10.1016/j.celrep.2013.03.026

Perez A, Kastner P, Sethi S, Lutz Y, Reibel C, Chambon P (1993) PMLRAR homodimers: distinct DNA binding properties and heteromeric interactions with RXR. EMBO j 12(8):3171–3182

Puccetti E, Ruthardt M (2004) Acute promyelocytic leukemia: PML/RARalpha and the leukemic stem cell. Leukemia 18(7):1169–1175. doi:10.1038/sj.leu.2403367

Puto LA, Reed JC (2008) Daxx represses RelB target promoters via DNA methyltransferase recruitment and DNA hypermethylation. Genes Dev 22(8):998–1010. doi:10.1101/gad.1632208

Redner RL, Rush EA, Faas S, Rudert WA, Corey SJ (1996) The t(5;17) variant of acute promyelocytic leukemia expresses a nucleophosmin-retinoic acid receptor fusion. Blood 87 (3):882–886

Regad T, Bellodi C, Nicotera P, Salomoni P (2009) The tumor suppressor Pml regulates cell fate in the developing neocortex. Nat Neurosci 12(2):132–140. doi:10.1038/nn.2251, nn.2251 [pii]

Rego EM, Wang ZG, Peruzzi D, He LZ, Cordon-Cardo C, Pandolfi PP (2001) Role of promyelocytic leukemia (PML) protein in tumor suppression. J Exp Med 193(4):521–529

Reymond A, Meroni G, Fantozzi A, Merla G, Cairo S, Luzi L, Riganelli D, Zanaria E, Messali S, Cainarca S, Guffanti A, Minucci S, Pelicci PG, Ballabio A (2001) The tripartite motif family identifies cell compartments. EMBO J 20(9):2140–2151. doi:10.1093/emboj/20.9.2140

Rice KL, de The H (2014) The acute promyelocytic leukaemia success story: curing leukaemia through targeted therapies. J Intern Med 276(1):61–70. doi:10.1111/joim.12208

Saeed S, Logie C, Francoijs KJ, Frige G, Romanenghi M, Nielsen FG, Raats L, Shahhoseini M, Huynen M, Altucci L, Minucci S, Martens JH, Stunnenberg HG (2012) Chromatin accessibility, p300, and histone acetylation define PML-RARalpha and AML1-ETO binding sites in acute myeloid leukemia. Blood 120(15):3058–3068. doi:10.1182/blood-2011-10-386086, blood-2011-10-386086 [pii]

Salomoni P (2009) Stemming out of a new PML era? Cell Death Differ 16(8):1083–1092. doi:10.1038/cdd.2009.63

Salomoni P (2013) The PML-Interacting Protein DAXX: Histone Loading Gets into the Picture. Front Oncol 3:152. doi:10.3389/fonc.2013.00152

Salomoni P, Pandolfi PP (2002) The role of PML in tumor suppression. Cell 108(2):165–170, S0092867402006268 [pii]

Salomoni P, Ferguson BJ, Wyllie AH, Rich T (2008) New insights into the role of PML in tumour suppression. Cell Res 18(6):622–640. doi:10.1038/cr.2008.58, cr200858 [pii]

Sanz MA, Montesinos P, Rayon C, Holowiecka A, de la Serna J, Milone G, de Lisa E, Brunet S, Rubio V, Ribera JM, Rivas C, Krsnik I, Bergua J, Gonzalez J, Diaz-Mediavilla J, Rojas R, Manso F, Ossenkoppele G, Gonzalez JD, Lowenberg B (2010) Risk-adapted treatment of acute promyelocytic leukemia based on all-trans retinoic acid and anthracycline with addition of cytarabine in consolidation therapy for high-risk patients: further improvements in treatment outcome. Blood 115(25):5137–5146. doi:10.1182/blood-2010-01-266007, blood-2010-01-266007 [pii]

Schoofs T, Rohde C, Hebestreit K, Klein HU, Gollner S, Schulze I, Lerdrup M, Dietrich N, Agrawal-Singh S, Witten A, Stoll M, Lengfelder E, Hofmann WK, Schlenke P, Buchner T, Hansen K, Berdel WE, Rosenbauer F, Dugas M, Muller-Tidow C (2013) DNA methylation changes are a late event in acute promyelocytic leukemia and coincide with loss of transcription factor binding. Blood 121(1):178–187. doi:10.1182/blood-2012-08-448860, blood-2012-08-448860 [pii]

Schwartzentruber J, Korshunov A, Liu XY, Jones DT, Pfaff E, Jacob K, Sturm D, Fontebasso AM, Quang DA, Tonjes M, Hovestadt V, Albrecht S, Kool M, Nantel A, Konermann C, Lindroth A, Jager N, Rausch T, Ryzhova M, Korbel JO, Hielscher T, Hauser P, Garami M, Klekner A, Bognar L, Ebinger M, Schuhmann MU, Scheurlen W, Pekrun A, Fruhwald MC, Roggendorf W, Kramm C, Durken M, Atkinson J, Lepage P, Montpetit A, Zakrzewska M,

Zakrzewski K, Liberski PP, Dong Z, Siegel P, Kulozik AE, Zapatka M, Guha A, Malkin D, Felsberg J, Reifenberger G, von Deimling A, Ichimura K, Collins VP, Witt H, Milde T, Witt O, Zhang C, Castelo-Branco P, Lichter P, Faury D, Tabori U, Plass C, Majewski J, Pfister SM, Jabado N (2012) Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. Nature 482 (7384):226-231. doi: 10.1038/nature10833. nature10833 [pii]

Skene PJ, Henikoff S (2013) Histone variants in pluripotency and disease. Development (Cambridge, England) 140(12):2513–2524. doi:10.1242/dev.091439

Spicuglia S, Vincent-Fabert C, Benoukraf T, Tiberi G, Saurin AJ, Zacarias-Cabeza J, Grimwade D, Mills K, Calmels B, Bertucci F, Sieweke M, Ferrier P, Duprez E (2011) Characterisation of genome-wide PLZF/RARA target genes. PLoS One 6(9), e24176. doi:10.1371/journal.pone.0024176

Sternsdorf T, Phan VT, Maunakea ML, Ocampo CB, Sohal J, Silletto A, Galimi F, Le Beau MM, Evans RM, Kogan SC (2006) Forced retinoic acid receptor alpha homodimers prime mice for APL-like leukemia. Cancer Cell 9(2):81–94. doi:10.1016/j.ccr.2005.12.030, S1535-6108(06)00025-0 [pii]

Subramanyam D, Belair CD, Barry-Holson KQ, Lin H, Kogan SC, Passegue E, Blelloch R (2010) PML-RAR{alpha} and Dnmt3a1 cooperate in vivo to promote acute promyelocytic leukemia. Cancer Res 70(21):8792–8801. doi:10.1158/0008-5472.CAN-08-4481, 0008-5472.CAN-08-4481 [pii]

Szenker E, Ray-Gallet D, Almouzni G (2011) The double face of the histone variant H3.3. Cell Res 21(3):421–434. doi:10.1038/cr.2011.14

Tabe Y, Konopleva M, Contractor R, Munsell M, Schober WD, Jin L, Tsutsumi-Ishii Y, Nagaoka I, Igari J, Andreeff M (2006) Up-regulation of MDR1 and induction of doxorubicin resistance by histone deacetylase inhibitor depsipeptide (FK228) and ATRA in acute promyelocytic leukemia cells. Blood 107(4):1546–1554. doi:10.1182/blood-2004-10-4126, 2004-10-4126 [pii]

Teofili L, Martini M, Luongo M, Diverio D, Capelli G, Breccia M, Lo Coco F, Leone G, Larocca LM (2003) Hypermethylation of GpG islands in the promoter region of p15(INK4b) in acute promyelocytic leukemia represses p15(INK4b) expression and correlates with poor prognosis. Leukemia 17(5):919–924. doi:10.1038/sj.leu.2402907, 2402907 [pii]

Torok D, Ching RW, Bazett-Jones DP (2009) PML nuclear bodies as sites of epigenetic regulation. Front Biosci (Landmark Ed) 14:1325–1336, 3311 [pii]

Villa R, De Santis F, Gutierrez A, Minucci S, Pelicci PG, Di Croce L (2004) Epigenetic gene silencing in acute promyelocytic leukemia. Biochem Pharmacol 68(6):1247–1254. doi:10.1016/j.bcp.2004.05.041, S0006295204003831 [pii]

Villa R, Pasini D, Gutierrez A, Morey L, Occhionorelli M, Vire E, Nomdedeu JF, Jenuwein T, Pelicci PG, Minucci S, Fuks F, Helin K, Di Croce L (2007) Role of the polycomb repressive complex 2 in acute promyelocytic leukemia. Cancer Cell 11(6):513–525. doi:10.1016/j.ccr.2007.04.009, S1535-6108(07)00116-X [pii]

Vitaliano-Prunier A, Halftermeyer J, Ablain J, de Reynies A, Peres L, Le Bras M, Metzger D, de The H (2014) Clearance of PML/RARA-bound promoters suffice to initiate APL differentiation. Blood. doi:10.1182/blood-2014-03-561852

Walter MJ, Park JS, Lau SK, Li X, Lane AA, Nagarajan R, Shannon WD, Ley TJ (2004) Expression profiling of murine acute promyelocytic leukemia cells reveals multiple model-dependent progression signatures. Mol Cell Biol 24(24):10882–10893. doi:10.1128/MCB.24.24.10882-10893.2004, 24/24/10882 [pii]

Wang ZG, Delva L, Gaboli M, Rivi R, Giorgio M, Cordon-Cardo C, Grosveld F, Pandolfi PP (1998) Role of PML in cell growth and the retinoic acid pathway. Science 279 (5356):1547–1551

Wang K, Wang P, Shi J, Zhu X, He M, Jia X, Yang X, Qiu F, Jin W, Qian M, Fang H, Mi J, Xiao H, Minden M, Du Y, Chen Z, Zhang J (2010) PML/RARalpha targets promoter regions containing

PU.1 consensus and RARE half sites in acute promyelocytic leukemia. Cancer Cell 17 (2):186–197. doi:10.1016/j.ccr.2009.12.045, S1535-6108(10)00008-5 [pii]

Wells RA, Catzavelos C, Kamel-Reid S (1997) Fusion of retinoic acid receptor alpha to NuMA, the nuclear mitotic apparatus protein, by a variant translocation in acute promyelocytic leukaemia. Nat Genet 17(1):109–113. doi:10.1038/ng0997-109

Westervelt P, Lane AA, Pollock JL, Oldfather K, Holt MS, Zimonjic DB, Popescu NC, DiPersio JF, Ley TJ (2003) High-penetrance mouse model of acute promyelocytic leukemia with very low levels of PML-RARalpha expression. Blood 102(5):1857–1865. doi:10.1182/blood-2002-12-3779

Widschwendter M, Fiegl H, Egle D, Mueller-Holzner E, Spizzo G, Marth C, Weisenberger DJ, Campan M, Young J, Jacobs I, Laird PW (2007) Epigenetic stem cell signature in cancer. Nat Genet 39(2):157–158. doi:10.1038/ng1941, ng1941 [pii]

Wojiski S, Guibal FC, Kindler T, Lee BH, Jesneck JL, Fabian A, Tenen DG, Gilliland DG (2009) PML-RARalpha initiates leukemia by conferring properties of self-renewal to committed promyelocytic progenitors. Leukemia 23(8):1462–1471. doi:10.1038/leu.2009.63, leu200963 [pii]

Wolyniec K, Carney DA, Haupt S, Haupt Y (2013) New Strategies to Direct Therapeutic Targeting of PML to Treat Cancers. Front Oncol 3:124. doi:10.3389/fonc.2013.00124

Won D, Shin SY, Park CJ, Jang S, Chi HS, Lee KH, Lee JO, Seo EJ (2013) OBFC2A/RARA: a novel fusion gene in variant acute promyelocytic leukemia. Blood 121(8):1432–1435. doi:10.1182/blood-2012-04-423129

Wu G, Broniscer A, McEachron TA, Lu C, Paugh BS, Becksfort J, Qu C, Ding L, Huether R, Parker M, Zhang J, Gajjar A, Dyer MA, Mullighan CG, Gilbertson RJ, Mardis ER, Wilson RK, Downing JR, Ellison DW, Baker SJ (2012) Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. Nat Genet 44(3):251–253. doi:10.1038/ng.1102, ng.1102 [pii]

Yamamoto Y, Tsuzuki S, Tsuzuki M, Handa K, Inaguma Y, Emi N (2010) BCOR as a novel fusion partner of retinoic acid receptor alpha in a t(X;17)(p11;q12) variant of acute promyelocytic leukemia. Blood 116(20):4274–4283. doi:10.1182/blood-2010-01-264432

Yuan W, Payton JE, Holt MS, Link DC, Watson MA, DiPersio JF, Ley TJ (2007) Commonly dysregulated genes in murine APL cells. Blood 109(3):961–970. doi:10.1182/blood-2006-07-036640, blood-2006-07-036640 [pii]

Yuen BT, Knoepfler PS (2013) Histone H3.3 mutations: a variant path to cancer. Cancer Cell 24 (5):567–574. doi:10.1016/j.ccr.2013.09.015

Zeisig BB, Kwok C, Zelent A, Shankaranarayanan P, Gronemeyer H, Dong S, So CW (2007) Recruitment of RXR by homotetrameric RARalpha fusion proteins is essential for transformation. Cancer Cell 12(1):36–51. doi:10.1016/j.ccr.2007.06.006, S1535-6108(07)00175-4 [pii]

Zhang H, He J, Li J, Tian D, Gu L, Zhou M (2013) Methylation of RASSF1A gene promoter is regulated by p53 and DAXX. FASEB J 27(1):232–242. doi:10.1096/fj.12-215491

Zhou J, Peres L, Honore N, Nasr R, Zhu J, de The H (2006) Dimerization-induced corepressor binding and relaxed DNA-binding specificity are critical for PML/RARA-induced immortalization. Proc Natl Acad Sci USA 103(24):9238–9243. doi:10.1073/pnas.0603324103, 0603324103 [pii]

Zhu J, Koken MH, Quignon F, Chelbi-Alix MK, Degos L, Wang ZY, Chen Z, de The H (1997) Arsenic-induced PML targeting onto nuclear bodies: implications for the treatment of acute promyelocytic leukemia. Proc Natl Acad Sci USA 94(8):3978–3983

Zhu J, Gianni M, Kopf E, Honore N, Chelbi-Alix M, Koken M, Quignon F, Rochette-Egly C, de The H (1999) Retinoic acid induces proteasome-dependent degradation of retinoic acid receptor alpha (RARalpha) and oncogenic RARalpha fusion proteins. Proc Natl Acad Sci USA 96(26):14807–14812

Zhu J, Zhou J, Peres L, Riaucoux F, Honore N, Kogan S, de The H (2005) A sumoylation site in PML/RARA is essential for leukemic transformation. Cancer Cell 7(2):143–153. doi:10.1016/j.ccr.2005.01.005, S1535-6108(05)00026-7 [pii]

Zhu J, Nasr R, Peres L, Riaucoux-Lormiere F, Honore N, Berthier C, Kamashev D, Zhou J, Vitoux D, Lavau C, de The H (2007) RXR is an essential component of the oncogenic PML/RARA complex in vivo. Cancer Cell 12(1):23–35. doi:10.1016/j.ccr.2007.06.004, S1535-6108(07)00173-0 [pii]

Zimonjic DB, Pollock JL, Westervelt P, Popescu NC, Ley TJ (2000) Acquired, nonrandom chromosomal abnormalities associated with the development of acute promyelocytic leukemia in transgenic mice. Proc Natl Acad Sci USA 97(24):13306–13311. doi:10.1073/pnas.97.24.13306

Part III
Chromosomes

Spatial Genome Organization and Disease

Karen J. Meaburn, Bharat Burman, and Tom Misteli

Abstract The nucleus is a complex organelle that performs a wide array of critical functions. Within the nucleus the genome is highly organized. Individual chromosomes form discrete chromosome territories. The organization of the genome is correlated with function, for example gene expression. Each chromosome and gene has a preferential spatial location, which can vary by cell type, differentiation stage and during disease. Active and inactive chromatin tends to be spatially separated both within the 3D nuclear space and within a chromosome territory. The molecular mechanisms that determine genome organization are currently poorly understood. However, it is known that the proximity of gene loci can contribute to translocation partner choice. The recent development of a plethora of new molecular techniques and imaging strategies, combined with fluorescent in situ hybridization, is being applied to both normal and diseased cells. Such studies will bring us closer to understanding the implications of genome organization and the molecules and mechanisms that determine it.

Keywords Genome organization • Chromosome territories • Nuclear architecture • Spatial positioning • Gene positioning • Disease • Cancer • Translocations • Chromatin structure

1 Introduction

The basic principles underlying how chromosomes behave and organize themselves during mitosis, compacting into tightly condensed X-shaped structures, aligning and then separating into daughter nuclei, have been appreciated since the late

K.J. Meaburn (✉) • T. Misteli (✉)
National Cancer Institute, NIH, Bethesda, MD 20892, USA
e-mail: meaburnk@mail.nih.gov; mistelit@mail.nih.gov

B. Burman
National Cancer Institute, NIH, Bethesda, MD 20892, USA

Program in Cell, Molecular and Developmental Biology, Sackler School of Biomedical Sciences, Tufts University, Boston, MA, USA
e-mail: burmanb@mail.nih.gov

D.P. Bazett-Jones, G. Dellaire (eds.), *The Functional Nucleus*,
DOI 10.1007/978-3-319-38882-3_5

1800s. The interphase nucleus, however, has historically been far more of a black box with regards to its compartmentalization and organization. Yet, cells, and consequently chromosomes, spend most of their time in interphase. It is during this time that they perform the majority of their functions.

Ground-breaking work over the last several decades has revealed that the nucleus is highly compartmentalized and that chromosomes exist as distinct nuclear subdomains, known as chromosome territories (CTs) (Fig. 1). These territories persist throughout the cell cycle, and remain largely separate from each other occupying reproducible positions within interphase nuclei (Cremer and Cremer 2001; Meaburn and Misteli 2007). The first evidence that chromosomes occupy discrete territories came in the 1970s and early 1980s from a series of UV-laser microbeam experiments (Cremer and Cremer 2001; Cremer et al. 2014). A microlaser was used to create DNA damage to a small volume of the interphase nucleus. The sites of damage detected in the subsequent metaphase were limited to only a few chromosomes per cell, indicating that chromosomes occupy discrete domains in the nucleus since damage to many more chromosomes would have been expected if the DNA from each chromosome was dispersed throughout the nucleus. The first direct visualization of CTs came in 1985, when total human genomic DNA was used as a probe to reveal the location of human DNA in interphase nuclei of

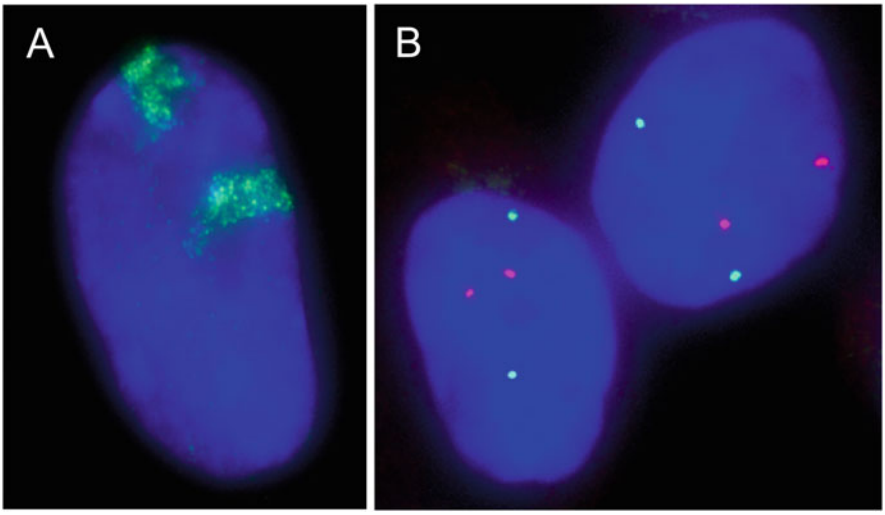

Fig. 1 Visualizing genomic loci. (**a**) Individual chromosomes form discrete chromosome territories (CTs) in interphase cells. CTs can be visualized using fluorescence in situ hybridization (FISH), in this case, HSA11 is shown in a MCF10A breast epithelial cell nucleus. (**b**) Individual genes assume preferred spatial locations. FISH reveals the location of *MMP1* and *TGFB3* gene loci in MCF10A nuclei. DAPI to stain DNA in *blue* (colour figure online)

human:rodent hybrid cells, which contained either one or a few human chromosomes (Cremer et al. 2014). However, it was not until the development and expanded use of DNA fluorescent in situ hybridization (FISH) in interphase cells in the 1980s (Cremer and Cremer 2001; Cremer et al. 2014) that it became possible to truly elucidate interphase chromosome topography, and to begin to probe the relationship between the spatial organization of the genome and nuclear function. In FISH, whole chromosomes or individual genomic regions are "painted" with fluorescent labels enabling the visualization of specific sub-regions of the genome. More recently, a host of complementary biochemical techniques have been devised to probe genome organization, such as chromatin conformation capture (3C)-based technologies and DNA adenine methyltransferase identification (DamID). These techniques are now confirming conclusions drawn from FISH experiments on a genome-wide scale and are yielding further insights into spatial genome organization (Bickmore and van Steensel 2013; van Steensel and Dekker 2010). The group of 3C-technologies measure the contact frequency of pairs of genomic loci in fixed nuclei, to determine regions that are in close 3D spatial proximity within a population of cells [reviewed in van Steensel and Dekker (2010)]. The different C-techniques are variations on the same basic principle, differing mainly by how much of the genome they cover. 3C detects the contact frequencies of a known locus to selected target loci of interest, 4C takes this one step further and asks where a single "bait" locus contacts other chromatin regions across the whole genome. In a more unbiased approach, Hi-C maps all contacts in the genome. The resolution of these methods is limited by the sequencing depth. For example, 5C is a similar "all against all" approach as Hi-C, but uses a small, defined region as a target, thus increasing resolution (Bickmore and van Steensel 2013; van Steensel and Dekker 2010). DamID is a complementary tool for detecting genomic interaction (proximity) with nuclear sub-compartments. To this end, DNA adenine methyltransferase is fused to a protein of interest, such as the nuclear lamina protein lamin B. This fusion protein methylates adenines in DNA that is in proximity with the fusion protein, thus mapping associated genomic regions [reviewed in van Steensel and Dekker (2010)].

While there is no question that the genome is highly organized within interphase nuclei, we are still a long way from fully understanding what this means functionally and what molecules are responsible for positioning and repositioning the genome. For the purpose of this review, we focus on mammalian genomes, but many properties of genome organization hold true for many other species, including plants, yeast, *C. elegans* and Drosophila (Bickmore and van Steensel 2013; Del Prete et al. 2014; Egecioglu and Brickner 2011; Sexton et al. 2012; Sharma and Meister 2013; Zimmer and Fabre 2011).

2 Non-random Genome Organization

2.1 Spatial Separation of Gene-Rich and Gene-Poor Genomic Regions

The genome is highly organized within interphase nuclei. Each chromosome and gene occupies a preferred nuclear position (Fig. 1) (Boyle et al. 2001; Cremer and Cremer 2010; Meaburn and Misteli 2007). The prototypical example of this is human chromosomes (HSA) 18 and 19 in proliferating cells. The gene-rich HSA19 is located in the center of the nucleus, while the gene-poor HSA18 locates to the nuclear periphery (Boyle et al. 2001; Croft et al. 1999). Indeed, in proliferating human cells the radial position (position relative to the edge and center of the nucleus) is highly correlated with gene density, whereby CTs generally follow the pattern of HSA18 and 19, with the gene-rich chromosomes in the center of the nucleus and gene-poor chromosome towards the periphery (Boyle et al. 2001). Chromosome positioning patterns can also be correlated with chromosome size (Bolzer et al. 2005; Bridger et al. 2000; Sun et al. 2000). In this case, the small chromosomes tend to position to the nuclear interior and the larger chromosomes prefer the periphery.

Remarkably, even within a CT there is spatial separation of gene-rich and gene-poor DNA (Fig. 2). Using FISH, it was demonstrated that gene-rich and gene-poor DNA within a 7 MB domain of Drosophila chromosome 2 (Boutanaev et al. 2005) or a 4.3 MB domain of mouse chromosome (MMU) 14 cluster separately from each other (Shopland et al. 2006). This also holds true at the whole CT level. The exome of the MMU2 chromosome is not found evenly throughout its CT, but instead, localizes to the part of the CT which faces the nuclear interior, and beyond the bulk of the CT (Boyle et al. 2011). In the case of the HSAX chromosome, genes, regardless of their activity, locate to the periphery of the CT, while the non-genic DNA localizes to the interior of the territory (Clemson et al. 2006). Genome-wide analysis has shown this spatial separation is not limited to a few isolated cases. Early Hi-C analysis, at the resolution of 1 Mb, revealed that gene-rich and transcriptionally active chromatin regions, and gene-rich chromosomes, tend to cluster together within the nuclear space, as do gene-poor, inactive chromatin regions. Moreover, gene-poor and gene-rich regions are far less likely to cluster with each other (Lieberman-Aiden et al. 2009). Hi-C and 4C have also shown that intra-chromosomal interactions are far more common than trans-chromosomal interactions, fitting with a genome organized into discrete CTs, and even the p and q arms of the same chromosome do not interact at high frequency (Bickmore and van Steensel 2013; Lieberman-Aiden et al. 2009). However, FISH data reveals that CTs are not completely isolated from each other. For example by cryo-FISH, ~40 % of a given CT volume intermingles with other CTs (Branco and Pombo 2006).

At a higher resolution, Hi-C and 5C can resolve internal CT domains further and has revealed the presence of the so-called topologically associated domains (TADs) (Bickmore and van Steensel 2013; Dixon et al. 2012). TADs are discrete kb-Mb-

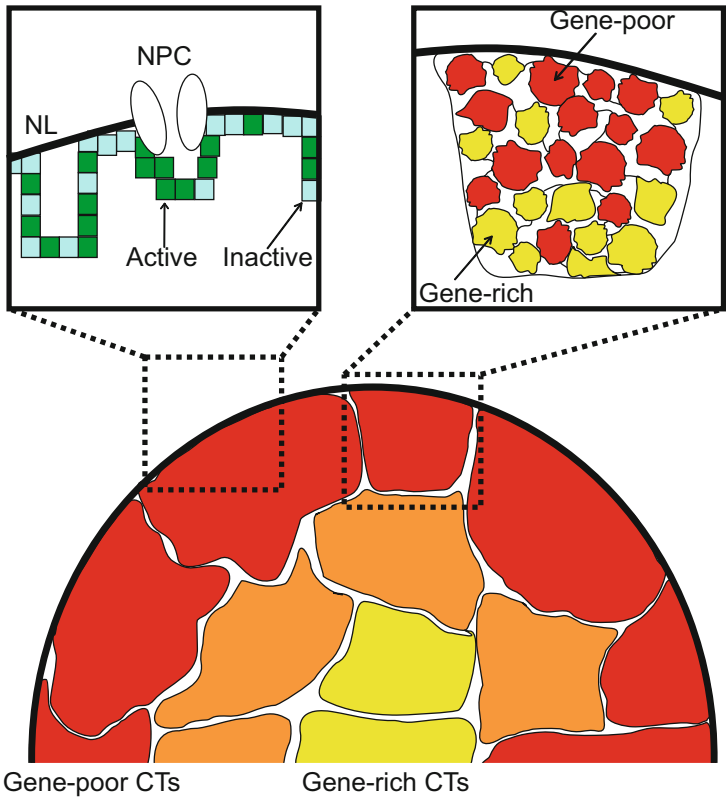

Fig. 2 The genome is non-randomly organized in interphase nucleus. CTs occupy preferred positions with the nucleus both relative to the periphery of the nucleus and to other chromosomes. Smaller chromosomes tend to be more internally located. Gene-poor chromosomes have a preference for the nuclear periphery and gene-rich chromosomes for the nuclear interior, and chromosomes with intermediate gene densities are intermediately positioned. The nuclear lamina (NL) is associated with inactive chromatin, whereas the nuclear pore complex (NPC) can associate with active and inactive chromatin. Even within a CT, gene-rich and gene poor-chromatin regions tend to be spatially separated, with genic regions tending to favor both the edge of the CT and the part of the CT which faces the nuclear interior

sized domains, which are defined as genomic regions that are more likely to be spatially close to other loci within the same TAD than another TAD (Bickmore and van Steensel 2013; Cavalli and Misteli 2013; Dixon et al. 2012). TADs are separated by sharp boundaries, and are conserved both between cell types and species (Dixon et al. 2012). TAD borders are characterized by enrichment for certain factors, such as CTCF, housekeeping genes and "active" histone modifications, and by depletion of heterochromatin marks such as H3K9me3 (Dixon et al. 2012). It has been suggested that the boundaries between TADs may function to prevent the spread of heterochromatin (Dixon et al. 2012). Despite the topological constraints imposed by such local domains, long range chromatin interactions

still occur within the nucleus, particularly between promoters and distal regulatory elements (Dekker and Misteli 2015).

Spatial positioning due to gene density and/or size cannot be the full story however, since the positioning patterns of chromosomes are not identical in all cells. For example, some, but not all, CTs are differentially positioned in proliferating and non-proliferating cells, as indicated by HSA18 and 13 that are peripheral in proliferating cells but internally positioned in G_0 cells (Bridger et al. 2000; Meaburn and Misteli 2008; Mehta et al. 2010). Similarly, although the organization of CTs is broadly similar across different tissues, some chromosomes occupy alternative positions in different cell types (Boyle et al. 2001; Foster et al. 2012; Parada et al. 2004; Zuleger et al. 2013), such as the small and gene-poor HSA21 which locates to the nuclear periphery in lymphoblasts but to the nuclear interior in fibroblasts (Boyle et al. 2001). Obviously, gene density and chromosome size are identical between these different cells, indicating other factors, such as differential gene expression or histone modifications, in CT positioning.

2.2 A Link Between Gene Activity and Positioning

Generally, frequently transcribed genes congregate in euchromatin, or "active chromatin", and localize in the nuclear interior, while repressed genes reside in heterochromatin, or "inactive chromatin", and localize in the nuclear periphery (Bickmore 2013). At the individual gene level there is also a correlation with gene expression and nuclear position (Ferrai et al. 2010a; Takizawa et al. 2008b). This is strikingly shown by the differential position of the active and inactive *Gfap* and *Il4* alleles, where the inactive allele is in a more peripheral position than the active allele, at least in the vast majority of cells in the population (Takizawa et al. 2008a). Another example of this is *CFTR*, which moves from the nuclear periphery upon activation, while its neighbouring, still silent, genes remain at the periphery (Zink et al. 2004a). In fact many loci, particularly developmentally regulated loci, reposition concomitantly with activation or silencing (Ferrai et al. 2010a; Takizawa et al. 2008b). For example, *Igh* and *Mash1* move from the edge of the nucleus to a more internal position as they become activated during lymphocyte and neuronal differentiation, respectively (Kosak et al. 2002; Williams et al. 2006). DamID, employing a lamin B fusion gene, found that the FISH studies on a limited number of genes coincided well at the genome-wide scale. In mammalian genomes there are 1100–1400 lamin associated domains (LADs), which are predominantly gene-poor and lowly transcribed/silent regions of the genome and include many developmentally regulated regions (Guelen et al. 2008; Peric-Hupkes et al. 2010; Pickersgill et al. 2006; Zullo et al. 2012). LADs vary between cell types and differentiation stages, and a subset of genomic regions change association with the lamina as they change expression (Peric-Hupkes et al. 2010). Interestingly, numerous genes move away from the periphery during differentiation before they change expression (Peric-Hupkes et al. 2010), suggesting that repositioning is not a mere consequence

of activity changes and involves separate mechanisms. While the majority of genes that associate with another lamina protein, lamin A/C, are also predominantly transcriptionally silent (Kubben et al. 2012), these observations should not be taken to mean that all silent genes locate to the nuclear periphery or that all peripheral genes are silent (Fig. 2). Many silent genes are located within the interior of the nucleus. As an example, even though *Gfap* moves to a more internal location upon activation, even it is inactive state it does not associate with the nuclear periphery (Takizawa et al. 2008a). Additionally, within each individual nucleus only ~30 % of the LAD domains are at the nuclear lamina (Kind et al. 2013). Conversely, numerous active genes locate to the nuclear periphery, often in association with nuclear pore complexes (Egecioglu and Brickner 2011), which may provide support for the "gene-gating" hypothesis put forward by Gunter Blobel nearly three decades ago (Blobel 1985).

Many gene loci change their association with other nuclear sub-compartments correlating with changes in transcription. These include nuclear bodies such as transcription factories, splicing speckles, Cajal bodies, polycomb bodies and blocks of heterochromatin [reviewed in Bickmore and van Steensel (2013) and Ferrai et al. (2010a)]. In at least some cases, this association is critical for gene expression (Khanna et al. 2014). Several examples have been reported of clustering of co-regulated genes, either in their active or inactive state, within the nuclear space, in many cases at nuclear bodies (Brown et al. 2008; Clowney et al. 2012; Fanucchi et al. 2013; Hakim et al. 2011; Rieder et al. 2014; Schoenfelder et al. 2010; Szczerbal and Bridger 2010; Takizawa et al. 2008a). Active genes may even loop away from the bulk of their CT in order to associate with co-regulated genes in "transcription factories" or splicing speckles (Branco and Pombo 2006; Brown et al. 2006; Ferrai et al. 2010b; Matarazzo et al. 2007; Morey et al. 2009). Splicing speckles are enriched in components of the splicing machinery and RNA polymerases. Transcription factories are sites of active transcription within the nucleus, which are enriched in RNA polymerases. Individual transcription factories have been suggested to be able to allow the simultaneous expression of several genes (Ferrai et al. 2010a).

2.3 Beyond Gene Expression

Changes in the spatial position of a gene may also occur in the absence of changes in gene expression (Harewood et al. 2010; Morey et al. 2009; Kubben et al. 2012; Kumaran and Spector 2008; Meaburn and Misteli 2008; Takizawa et al. 2008b; Williams et al. 2006). For example, when the *Mash1* locus moves to the nuclear interior upon activation during murine neuronal cell differentiation, neighbouring genes also become more internally positioned despite remaining silent (Williams et al. 2006). Equally, when the active *Hoxb* locus loops out of its CT during mouse ES cell differentiation, flanking genes also move away from the bulk of the CT without an accompanying change in expression (Morey et al. 2009). Moreover,

when loci are forced to change nuclear position, either by artificial tethering to the nuclear periphery or by chromosome translocation, some genes do alter expression, while others are unaffected (Finlan et al. 2008; Harewood et al. 2010; Kumaran and Spector 2008; Reddy et al. 2008). Conversely, another set of genes remain in the same nuclear position when their activity changes (Hakim et al. 2011; Hewitt et al. 2004; Meaburn and Misteli 2008), such as *Ifnγ*, which remains at the periphery in mouse T helper cells when activated (Hewitt et al. 2004). To complicate things further, in mouse erythroid differentiation, *β-globin* becomes active, localizes with active polymerase II and then, at a later time, moves from the nuclear periphery simultaneously with polymerase II foci concentrating away from the periphery (Ragoczy et al. 2006). In this case, the repositioning could be a consequence of changes in expression. However, gene expression increases further after the loci repositions, suggesting that positioning patterns may aid modulation of expression (Ragoczy et al. 2006).

One explanation for these varying observations on individual genes is that some loci require specific positioning for expression and others do not, and that the latter loci are passively moved along with the neighbouring regions for which positioning is important. Alternatively, other factors may also influence spatial positioning. It seems the genomic or regional neighbourhood a locus resides in is important for its behaviour (Takizawa et al. 2008b). In agreement with this, the same gene can behave differently between species when activated. For example, during erythroblast differentiation *α-globin* and *β-globin* loci co-localize, with *α-globin* looping out of its CT in mice when the gene activates, but neither of these events occur in humans (Brown et al. 2006, 2008). The genomic neighbourhood for these two genes (e.g. local gene density) is divergent between mouse and human (Brown et al. 2008). These data suggest that the repositioning and resulting proximity of these two genes is not a requirement per se for the regulation of particular genes, and that genomic context matters. Further evidence that sequence does not intrinsically dictate position but that the local environment plays a role comes from human:mouse hybrid cells. In these hybrid cells, many human chromosomes do not assume the same position in the mouse nucleus as they do in the human cells they originate from (Meaburn et al. 2008). Such comparisons across multiple species and assessing the effects of manipulating the nuclear environment may shed light on the factors that are important in defining positioning patterns.

Transcription is only one of many functions the nucleus carries out. While gene activity does not fully account for genome organization, it could be that other nuclear activities, either alone or in combination with gene expression, influence positioning patterns. For example, replication timing, which in itself is linked to gene expression, correlates with nuclear positioning patterns (Hiratani et al. 2008). Moreover, in yeast cells, broken chromosome ends move to the nuclear periphery (Nagai et al. 2008). Although in mammalian cells the position of broken DNA ends are generally positionally stable (Roukos and Misteli 2014), large scale DNA damage does seem to induce both decondensation of chromatin (Kruhlak et al. 2006; Dellaire et al. 2009) and a re-organization of the genome (Mehta et al. 2013). For example both UV-laser induced DNA damage (Kruhlak

et al. 2006) and etoposide-induced DNA DSBs (Dellaire et al. 2009) can trigger large scale decondensation of chromatin. In another example, five chromosomes were shown to reposition in fibroblasts cells upon treatment with global DNA damaging agents, including gene-rich HSA17, 19 and 20 relocating to the nuclear periphery (Mehta et al. 2013).

Chromosome organization may also play a structural role by contributing to the function of the organ within which the individual cells reside. For example, a rather surprising apparent role for genome organization with respect to cellular function has been identified in the retina. In rod cells of nocturnal mammals there is an inversion of the position of eu- and heterochromatin, with heterochromatin locating in the center of nuclei (Solovei et al. 2009). This evolutionary adaptation is important for filtering low levels of light through nuclei to the photoreceptors to optimize night vision (Solovei et al. 2009). Surprisingly, this massive reorganization of chromatin does not affect transcriptional output (Solovei et al. 2009).

Despite the ample evidence of the existence of chromosomal movement, it remains unclear what mechanisms move chromatin in the nucleus. Several mechanisms have been suggested. For example, passage though mitosis and particularly early G_1 has been associated with localizing some regions to the nuclear periphery, such as HSA18 as cells re-enter proliferation (Bridger et al. 2000). Similarly, the subset of LAD regions binding to the nuclear periphery are reset during early G_1, and are a different subset than the mother cell (Kind et al. 2013). Moreover, tethering of a reporter array to the periphery via lamin B also required passage through mitosis (Kumaran and Spector 2008). However, the observed repositioning of loci in quiescence (Bridger et al. 2000; Meaburn and Misteli 2008), by its very definition, cannot involve mitosis and early G_1. In fact, the repositioning of whole chromosomes in quiescent cells is remarkably quick, occurring within 15 min after cells are placed into low serum conditions, and requires ATP/GTP and nuclear myosin and actin (Mehta et al. 2010). Chromatin modifiers may also play a role is the positioning of the genome. In line with this, histone deacetylase 3 is important for the correct peripheral localization of certain gene loci (Demmerle et al. 2013; Zullo et al. 2012). Similarly, histone H3K9 di- and tri-methylation by Suv39H and G9a in mammalian cells (Bian et al. 2013; Kind et al. 2013) and by MET-2 and SET-25 in *C. elegans* (Towbin et al. 2012) are also important to anchor heterochromatin at the nuclear periphery. Non-coding RNA may also play a role in maintaining certain features of CT since it has recently been discovered that non-coding RNA from repetitive regions of the DNA coat CTs and appear to influence chromatin compaction (Hall et al. 2014). In addition, the long noncoding RNA Firre is necessary for the co-localization of specific trans-chromosomal loci (Hacisuleyman et al. 2014). The potential role of the nuclear envelope in positioning the genome will be discussed below (Sect. 4.1).

3 Genome Organization and Disease

The fact that genome organization is linked to correct cellular function is highlighted by the findings that the genome is often reorganized in disease states. Disease generally does not result in global nuclear re-organization and instead leads to repositioning of subsets of genome regions. Studying genome positioning patterns in disease is giving deeper insight into what regulates genome organization, may provide clues to disease mechanisms and is paving the way for spatial positioning to be used as a diagnostic tool.

3.1 The Nuclear Envelope, Laminopathies and Genome Organization

Laminopathies are a group of rare diseases caused by mutations in nuclear envelope (NE) proteins. The most prominent of these diseases result from mutations in lamin A/C and include Emery-Dreifuss muscular dystrophy (EDMD) and the premature aging disease Hutchison-Gilford progeria syndrome (HGPS) (Burke and Stewart 2013; Dittmer and Misteli 2011). The organization of CTs in proliferating fibroblasts from laminopathy patients resembles that of quiescent normal fibroblasts cells, with HSA13 and 18 no longer positioned at the nuclear periphery and HSA10 relocating to the periphery (Meaburn et al. 2007b; Mehta et al. 2011; Mewborn et al. 2010). These altered positioning patterns can be reversed. Treatment of HGPS fibroblasts with farnesyltransferase inhibitors (FTIs), which prevent the accumulation of mutant lamin A at the NE, rescues CT positioning back to normal (Mehta et al. 2011). Interestingly, the reorganization of CTs is similar for most laminopathy patients, irrespective of where the mutation maps to on the lamin A/C gene (*LMNA*), or what disease the patient suffers from (Meaburn et al. 2007a; Mewborn et al. 2010). The exceptions to this being *LMNA* mutations delta303 and D596N, where HSA13 is more tightly associated with the nuclear periphery (Mewborn et al. 2010). Not all chromosomes change their association with the NE in laminopathy fibroblasts, however. Similar to quiescent control cells, HSA4 and X remain at the nuclear periphery of laminopathy cells (Meaburn et al. 2007a; Mehta et al. 2011). That said, the conformation of each chromosome does appear to change, at least in late passage HGPS cells. Using Hi-C, there is a loss of spatial separation between active and inactive chromosomal domains in late passage, senescing HGPS fibroblasts (passage-19), but not in a proliferating culture of HGPS (passage-17) (McCord et al. 2013). Since senescing normal fibroblasts were not used as a control, it remains unclear to what extent these changes are associated with senescence and what changes are directly linked to HGPS. Interestingly, changes in H3K27me3 and lamin A/C binding patterns were noted to occur before the large scale changes in chromatin contact frequencies in these cells,

suggesting these are important determinants of chromatin organization (McCord et al. 2013).

The observations in laminopathies are in line with findings on the role of various NE proteins in genome organization. In mouse fibroblasts, the position of MMU18 to the nuclear periphery is dependent on lamin B1, while the central position of MMU19 is not (Malhas et al. 2007). As with lamins, several NE transmembrane proteins (NETs) influence the position of a specific subsets of chromosomes (Zuleger et al. 2013). NET29 and NET39 have a role in locating HSA5 and HSA13 to the nuclear periphery, whereas expression of NET5, NET45 and NET47 position HSA5 but not HSA13 to the periphery. Moreover, the presence of NET47 actually reduced HSA13's association with the nuclear periphery. On the other hand, the internal position of HSA19 and HSA17 are not influenced by presence of these NETs (Zuleger et al. 2013).

Given that the NE of different cell types is composed of distinct complements of NE proteins, the variations between different NE proteins and the chromosomes they influences may account, in part, for the tissue-specificity of genome organization (Solovei et al. 2013; Wong et al. 2014). For example, the expression of some NETs, including the five NETs mentioned above, are either restricted to only certain tissues or exhibit a wide range of expression levels between tissue types (Zuleger et al. 2013). In keeping with this, the positioning pattern of HSA5 in liver cells could be transformed into that of kidney cells by modifying the expression of NETs so that the NE better resembled the kidney NE (Zuleger et al. 2013). The differential influence of various NE proteins between different cell types is also highlighted in laminopathies. HSA13 and 18 remain peripheral in laminopathy patient lymphoblastiod cell lines, a cell type which does not require lamin A/C (Boyle et al. 2001; Meaburn et al. 2005). Consistent with this, in *C. elegans* an EDMD associated mutant lamin (Y59C) inhibited the release from the nuclear lamina, and full activation, of a muscle promoter array in muscle cells, but it did not interfere with an intestinal promoter containing array relocating to the nuclear interior during gut development (Mattout et al. 2011). Along with the tissue specific impairment of genome reorganization, the Y59C worms only had an aberrant muscle phenotype (Mattout et al. 2011). In the most detailed study to date of differences in NE composition between tissues and the consequences on spatial organization, NE proteins lamin B receptor (LBR) and lamin A/C were compared between rod nuclei from 39 different species and in 30 tissue types during mouse development (Solovei et al. 2013). Lack of both LBR and lamin A/C strongly correlated with a dramatic inversion of chromatin distribution, in which euchromatin shifts from an internal position to the periphery. In wild-type and LBR null mice inverted chromatin was only observed in cell types lacking both LBR and lamin A/C, suggesting that these proteins have roles in targeting heterochromatin to the nuclear periphery (Solovei et al. 2013). Indeed, Pelger–Huët anomaly, a disease in which LBR is mutated, is characterised by an altered heterochromatin distribution (Hoffmann et al. 2007). Developmentally regulated loss of LBR has also been implicated in the clustering of silenced olfactory receptor genes, away from the nuclear periphery, in olfactory neurons (Clowney et al. 2012).

The changes to genome organization in cells with either a mutation in a NE protein or with an altered complement of NE proteins may be due to a direct role of NE proteins in tethering chromatin. However, since these reorganizations occur in concert with misregulation of both gene expression and histone modifications (Malhas et al. 2007; McCord et al. 2013; Mewborn et al. 2010; Scaffidi and Misteli 2005; Shumaker et al. 2006), it cannot be ruled out that that the repositioning events are indirectly related to altered NE structure or function. Of course, direct roles via physical interaction and indirect roles of NE proteins in chromatin positioning are not mutually exclusive. It seems likely that genome repositioning events are a result of a release from tethering to the NE as well as due to changes in epigenetic modifications and gene expression that stem from alterations in NE interactions with other proteins. Several lines of evidence support a direct tethering role for NE proteins. These include the observation that NE proteins, such as lamins and LBR, can directly bind to DNA and chromatin (Dittmer and Misteli 2011; Makatsori et al. 2004). Moreover, HSA13 and 18 are mis-positioned in asymptomatic carriers of *LMNA* mutation R527H+/− (Meaburn et al. 2007a), where gene expression patterns and chromatin modifications are presumably similar to control cells. In *LMNA* E145K mutant fibroblasts centromeres and telomeres are mis-localized only in nuclei that are lobulated (Taimen et al. 2009). Live cell imaging revealed that the reorganization of the centromeres and telomeres and lobulation occurred together as the NE reformed after mitosis, suggesting an altered attachment of centromeres and telomeres to lamins (Taimen et al. 2009). In further support of a direct tethering role, in mouse cardiac myocytes, knock-down of lamin A/C caused a subset of genes to move away from the nuclear periphery, in the absence of a change of expression for that gene (Kubben et al. 2012). It may be that several NE proteins are required to work in concert to tether chromatin to the NE. At least in rod cells, LEM domain proteins, such as emerin, are required with lamin A/C to tether chromatin to the periphery (Solovei et al. 2013). Conversely, LBR does not seem to require a mediator to tether heterochromatin (Solovei et al. 2013).

3.2 Altered Genome Organization in Other Non-cancerous Diseases

Genome reorganization is not limited to diseases with altered NE. In one of the first studies of spatial genome organization in disease, Borden and Manuelidis established that the centromere of HSAX is relocated away from the nuclear periphery or the periphery of nucleoli in the epileptic focus of brain tissue from patients with chronic uncontrolled seizures. Conversely, loci mapping to 1q12, 9q12 and Yq12 remained proximal to the NE and/or nucleolus in most patients (Borden and Manuelidis 1988). Some diseases also have an increased clustering of genomic loci. In cheek cells from Alzheimer's patients there is an altered telomere aggregation and clustering, possibly as the result of shortened telomeres and

telomere dysfunction (Mathur et al. 2014) and in Down syndrome there is an increase clustering of HSA21 (Paz et al. 2013).

Furthering a link between spatial genome positioning and histone modifications, gene repositioning occurs in a disease associated with altered DNA methylation. Mutations in DNA methyltransferase 3B lead to immunodeficiency centromeric instability facial anomalies (ICF) syndrome, and several genomic loci are subjected to DNA hypomethylation and thus activation. One such site, *SYBL1,* which normally escapes X inactivation, loops out of the inactive X (X_i) in female ICF cells and the HSAY CT in male ICF cells, but it does not loop from the active X (Matarazzo et al. 2007). The normally methylated and silenced neighbouring gene *SPRY3* also loops out of the X_i CT, but not HSAY in ICF cells (Matarazzo et al. 2007). Both these differences point to the local genomic environment as an influential factor on positioning (see Sect. 3.3).

Altered spatial positioning patterns are also induced in cells infected with either viruses or parasites. For example, Epstein-Barr virus infection of lymphocytes results in HSA17 transiently moving closer to the periphery in the days following infection, while the position of HSA18 is unaffected (Li et al. 2010). Similarly, in snail Bge embryonic cells, infection with the parasite *Shistosoma manoni* results in temporal repositioning of gene loci (Knight et al. 2011). Again, the effect of infection varies between the loci studied. *Actin* shifted toward the nuclear periphery within 30 min of infection, in the absence of a change in gene expression. However, in these cells, gene expression was altered at both earlier (15 min) and later (2 h) times after infection. Conversely, *ferritin* was displaced from the nuclear periphery, peaking at 5 h and returning to a peripheral positioning by 24 h, matching the temporal changes in its expression. This repositioning is dependent on an active infection, and not simply the presence of a cellular breach or foreign entity within the nuclei since irradiated (non-functional) parasites did not induce repositioning (Knight et al. 2011). Interestingly, both HSA17 and *actin* exhibited cycling of positioning to the periphery after infection, with both returning to the periphery again after internal positioning had been restored (Knight et al. 2011; Li et al. 2010).

3.3 Altered Genome Organization in Cancer

Alterations in spatial genome organization have been prominently linked to cancer (Meaburn and Misteli 2007; Zink et al. 2004b). Indeed, distinctive alterations in chromatin staining patterns and nucleoli size and number, in addition to nuclear shape and tissue morphology, are used by pathologists to diagnose cancer. Yet these large scale changes at the chromatin level do not appear to reflect a global change in genomic spatial positioning patterns. For example, in pancreatic cancer tissues, HSA8 remains at the nuclear peripherally, although the CT shape changes (Timme et al. 2011; Wiech et al. 2005). Conversely, HSA18 and 19 change nuclear location in several cancers types, including cervical, colon and some thyroid cancers

(Cremer et al. 2003; Murata et al. 2007). Cancer related repositioning is not limited to whole chromosomes. For instance, in cervical squamous carcinoma, the *BCL2* locus repositions in a BCL2 positive tumor, but not in a BCL2 negative tumor (Wiech et al. 2009). The most extensively studied cancer to date, with respect to spatial genome positioning, is breast cancer. 4C data, and FISH validation, revealed *IGFBP3* changes long range interaction partners in breast cancer cell lines (Zeitz et al. 2013). Moreover, the centromere of HSA17 is more internally positioned in breast cancer (Wiech et al. 2005) and several genes reposition in breast cancer (Meaburn et al. 2009; Meaburn and Misteli 2008). In an in vitro mammary epithelial cell model of early breast cancer, 4 out of 11 tested genes (*AKT1*, *VEGF*, *BCL2* and *ERBB2*) significantly changed intranuclear position, but not their gene expression level, during carcinogenic transformation (Meaburn and Misteli 2008). Similarly, in breast cancer tissues, 8 of 20 tested genes (*HES5*, *HSP90AA1*, *TGFB3*, *MYC*, *ERBB2*, *FOSL2*, *CSF1R* and *AKT1*) occupied significantly different positions in breast cancer tissues compared to normal tissues (Meaburn et al. 2009). These differences were not due to inter-sample variance since these genes were positioned in 11–14 cancers and 6–9 normal tissues and even though a wide range of breast cancers specimens were used, these genes repositioned in 64.3–100 % of cancers (Meaburn et al. 2009). The differences in gene position were also not due to numerical chromosome abnormalities, and, for most genes, were not observed in the benign breast diseases hyperplasia and fibroadenoma or among normal tissues (Meaburn et al. 2009). These observations point to specific repositioning events of a subset of genes in breast cancer and they point to the possibility of using spatial positioning of the genome as a diagnostic biomarker for cancer detection (Meaburn et al. 2009; Meaburn and Misteli 2008).

4 The Role of Positioning and Chromatin in Translocation Formation

4.1 Spatial Proximity of Translocation Partners

While it is not entirely clear why the genome is positioned as it is within interphase cell and what functional role it has, the spatial positioning of the genome has emerged as an important factor in determining chromosome translocation partners (Fig. 3) (Meaburn et al. 2007b; Misteli 2010; Roukos and Misteli 2014). It was initially noticed in FISH experiments that frequent translocation partners appeared to be frequently found in close spatial proximity. For example, peripheral chromosomes HSA4, 9, 13 and 18 are more likely to translocate with each other than with internally located chromosomes (Bickmore and Teague 2002). Moreover, chromosomes that form preferred clusters within certain tissues, such as MMU12, 14 and 15 in splenocytes or MMU5 and 6 in hepatocytes, are more likely to translocate in cancers derived from those tissues than chromosomes that are not part of these

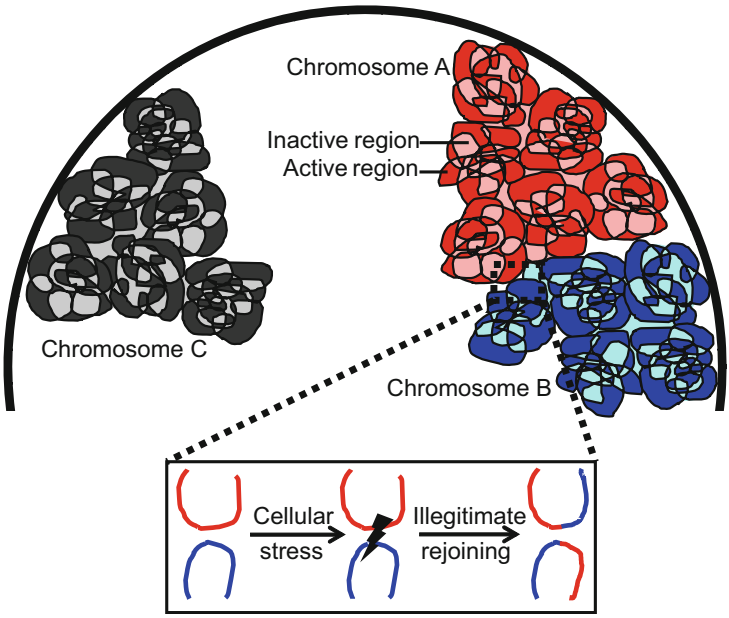

Fig. 3 Spatial positioning patterns influence translocation frequency. Translocations can arise from DSBs induced by cellular stress in the forms of genotoxic, oxidative, replicative, or transcriptional stress. Illegitimate joining of DSBs can result in the formation of translocations. Proximal chromosomes A and B translocate at higher frequency than distal chromosomes A and C, or B and C. Chromatin features may also predispose certain genomic regions to translocation

clusters (Parada et al. 2002, 2004). This data points to the role tissue specific genome organization may play in the prevalence of certain translocations in different tissues. In keeping with proximal chromosomes being more likely to be translocation partners, the amount of intermingling with neighbouring chromosomes in human lymphocytes correlates with translocation frequency (Branco and Pombo 2006).

Perhaps more importantly, beyond the level of entire chromosomes, genes that translocate are in close proximity, in the cell types where they translocate (Lukasova et al. 1997; Meaburn et al. 2007b; Misteli 2010; Neves et al. 1999; Roix et al. 2003; Roukos and Misteli 2014). For instance, in androgen deprived prostate cell nuclei, *TMPRSS2*, *ERG* and *ETV1,* which are common translocation partners in prostate cancers, are generally found in distant locations (Lin et al. 2009; Mani et al. 2009). Upon androgen stimulation these genes reposition to become proximal neighbours, predisposing them as translocation partners, as indicated by a dramatic increase in translocation frequency upon irradiation (Lin et al. 2009; Mani et al. 2009). Along the same lines, in anaplastic large-cell lymphoma cells the close proximity of *ALK* and *NPM1* facilities the formation of the *ALK-NPM1* gene fusion upon irradiation (Mathas et al. 2009).

The caveat of these correlation studies is they may not reflect translocation formation per se, since only translocations that carry a growth advantage, and thus expand to high levels within the cancer cell population, are analyzed. Moreover, most of these studies were limited to a few genes and control regions. To address this, several genome-wide studies have been performed in mouse B lymphocytes, which, crucially, were carried out before cellular selection skewed the translocation detection frequency. Translocation frequencies across the genome were measured and mapped onto linear chromosomes (Chiarle et al. 2011; Klein et al. 2011) or compared to Hi-C (Zhang et al. 2012) or 4C (Hakim et al. 2012; Rocha et al. 2012) data, to account for 3D genomic proximity. These studies have confirmed a correlation between spatial proximity and translocation frequency. Nevertheless, a low frequency of translocations from genes that were distally located was also detected (Zhang et al. 2012). In addition to spatial proximity, not surprisingly, the amount of DNA damage was also implicated in determining frequency of specific translocations (Hakim et al. 2012). These correlative studies were recently extended by direct observation of translocation formation (Roukos et al. 2013). For this, inducible DNA double strand break (DSB) sites, tagged with coloured fluorescently labels, were integrated into different mouse chromosomes. Upon DNA damage, the broken DNA ends where then tracked using live-cell imaging of individual cells. Consistent with both the FISH and genome-wide data, the vast majority of translocation forming breaks were proximal (within 2.5 μm) before pairing, however, ~10 % of translocations were formed from distant breaks (>4 μm apart) (Roukos et al. 2013).

Given the fact the translocation partners tend to be proximal neighbours before translocation, it is not surprising that many fusion chromosomes resulting from the translocation event position similarly to their intact counterparts (Cremer et al. 2003; Croft et al. 1999; Meaburn et al. 2007b; Parada et al. 2002). Interestingly, the orientations within the derivative CT tend to reflect the positioning patterns of the individual intact chromosomes (Cremer et al. 2003; Croft et al. 1999), for example, within the CT of the t(18;19) derivative chromosome, the HSA18 DNA is more peripherally located than the HSA19 DNA, similar to the intact chromosomes (Croft et al. 1999). This is not to say the positions of all translocation chromosomes are unaffected. In lymphoblastoid cells from multiple individuals with the t(11;22) (q23;q11) balanced translocation, the HSA11 DNA on the fusion chromosome is more centrally located that intact HSA11. Similarly, the HSA22 portion of the fusion chromosome was more peripherally located than the intact HSA22 for the derivative 11 chromosome, but was unaffected when it was part of the derivate 22 chromosome (Harewood et al. 2010). The repositioning of gene loci after translocation can also be variable depending on the translocation, with some translocated loci repositioning and others not (Meaburn et al. 2007b). In at least some case the repositioning reflects alterations to the local gene density around the fusion gene site (Harewood et al. 2010; Murmann et al. 2005) or aberrant expression at or around the fusion gene (Ballabio et al. 2009; Harewood et al. 2010).

4.2 Chromatin Organization and Translocations

While the spatial arrangement of the genome in vivo contributes to the formation of translocations, not all genomic loci are equally susceptible to translocation. Recently, evidence has emerged that points to higher-order chromatin structure as a key player in translocations formation, possibly by modulating DNA DSB susceptibility and repair (Misteli 2010). In support, genome-wide sequencing of translocation junctions after DSBs were experimentally induced found that most breakpoints localize within or near transcriptionally active regions (Fig. 3) (Chiarle et al. 2011; Klein et al. 2011). The break sites were enriched for histone modifications associated with active chromatin, such as H3K4me3, H3K36me3 and H3 acetylation (Klein et al. 2011). A link between transcriptional activity and translocation frequency has been demonstrated in prostate cancer, where the common translocation fusion-gene partners *TMPRSS2*, *ERG*, and *ETV1* contain binding sites for androgen receptor (AR), a potent transcriptional activator. Upon androgen treatment, AR was co-recruited with topoisomerase-IIβ to break sites, leading to a more open chromatin conformation and persistent DSBs (Haffner et al. 2010; Lin et al. 2009). Similarly, the regions near translocation breakpoints in anaplastic large cell lymphoma were found to be transcriptionally activated prior to translocation (Mathas et al. 2009). Taken together, these observations suggest that genomic regions with altered chromatin structure and transcription factor binding may be more susceptible to DSBs that lead to translocations.

Chromatin organization within the nucleus also appears to determine the efficiency of DNA repair. After a DSB occurs, damaged chromatin around the break is thought to rapidly decondense to facilitate access of repair machineries, and then recondense as the repair process progresses. These events are orchestrated by chromatin remodelers that reposition nucleosomes, histone chaperone proteins that exchange core histones for specific histone variants, and histone modifying enzymes (Groth et al. 2007). Higher-order chromatin structure can drastically influence the progression of repair, possibly by impeding recruitment of these proteins. For example, radiation-induced DSBs in heterochromatin were observed to repair more slowly than breaks in euchromatin (Goodarzi et al. 2008). In the case of translocations, several chromatin modifications have been implicated in the inaccurate repair of DSBs. In prostate cancer cells treated with liganded AR, H3K79me2, a modification associated with DNA recombination, was found to be enriched near *TMPRSS2* and *ERG* breakpoints. Overexpression of H3K79-specific methyltransferase DOT1L significantly increased translocation frequency (Lin et al. 2009). Along the same lines, genome-wide conversion to an H4K20 monomethylation state in mice led to defective DSB repair, Ig class-switch recombination, and *IgH* translocations (Schotta et al. 2008). Finally, in the absence of H2AX, a histone variant that is immediately phosphorylated after DSB formation, DSBs were shown to persist during Ig class-switch recombination, resulting in frequent translocations (Franco et al. 2006). These observations point to a potential, still poorly characterized, role of chromatin structure and histone modifications in determining translocation break points.

5 Summary

It has become increasingly apparent that the genome is non-randomly organized in interphase nuclei and that these positioning patterns generally correlate with nuclear function. Several recently developed technologies are driving forward our understanding of the extent and relevance of spatial genome organization. Combining genome-wide strategies and high-throughput siRNA screens with FISH will enable the identification of factors that directly regulate the position of genomic loci, the mechanisms of gene motion and will give a clearer understanding of the functional consequences of positioning patterns. In these studies it will be important to consider a given gene in the context of its neighbourhood. Comparisons of genome-wide data sets merging genome positioning information with gene expression, epigenetics, proteome, non-coding RNA expression and localisation, in multiple biological systems (different species, cell types, conditions, diseases) will be an important first step. Furthermore, studying the genome in live cells will continue to give important insights and allow the hypotheses generated from fixed cells to be tested in real time. New approaches, such as clustered regularly interspaced short palindromic repeats (CRISPR) (Chen et al. 2013), are adding to the arsenal of techniques for live cell imaging of specific regions of the genome, be it endogenous loci or engineered arrays. The time is ripe to integrate data from of the new and old techniques to further elucidate important properties of the spatial organization of genomes.

Acknowledgments Work in the Misteli laboratory is supported by the Intramural Research Program of the National Institutes of Health (NIH), NCI, Center for Cancer Research. KM and BB are supported by Department of Defense Idea Awards (W81XWH-12-1-0224 and W81XWH-12-1-0295).

References

Ballabio E, Cantarella CD, Federico C, Di Mare P, Hall G, Harbott J, Hughes J, Saccone S, Tosi S (2009) Ectopic expression of the HLXB9 gene is associated with an altered nuclear position in t (7;12) leukaemias. Leukemia 23(6):1179–1182. doi:10.1038/leu.2009.15

Bian Q, Khanna N, Alvikas J, Belmont AS (2013) beta-Globin cis-elements determine differential nuclear targeting through epigenetic modifications. J Cell Biol 203(5):767–783. doi:10.1083/jcb.201305027

Bickmore WA (2013) The spatial organization of the human genome. Annu Rev Genomics Hum Genet 14:67–84. doi:10.1146/annurev-genom-091212-153515

Bickmore WA, Teague P (2002) Influences of chromosome size, gene density and nuclear position on the frequency of constitutional translocations in the human population. Chromosome Res 10 (8):707–715

Bickmore WA, van Steensel B (2013) Genome architecture: domain organization of interphase chromosomes. Cell 152(6):1270–1284. doi:10.1016/j.cell.2013.02.001

Blobel G (1985) Gene gating: a hypothesis. Proc Natl Acad Sci USA 82(24):8527–8529

Bolzer A, Kreth G, Solovei I, Koehler D, Saracoglu K, Fauth C, Muller S, Eils R, Cremer C, Speicher MR, Cremer T (2005) Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. PLoS Biol 3(5):e157

Borden J, Manuelidis L (1988) Movement of the X chromosome in epilepsy. Science 242:1687–1691

Boutanaev AM, Mikhaylova LM, Nurminsky DI (2005) The pattern of chromosome folding in interphase is outlined by the linear gene density profile. Mol Cell Biol 25(18):8379–8386

Boyle S, Gilchrist S, Bridger JM, Mahy NL, Ellis JA, Bickmore WA (2001) The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. Hum Mol Genet 10(3):211–219

Boyle S, Rodesch MJ, Halvensleben HA, Jeddeloh JA, Bickmore WA (2011) Fluorescence in situ hybridization with high-complexity repeat-free oligonucleotide probes generated by massively parallel synthesis. Chromosome Res 19(7):901–909. doi:10.1007/s10577-011-9245-0

Branco MR, Pombo A (2006) Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. PLoS Biol 4(5), e138

Bridger JM, Boyle S, Kill IR, Bickmore WA (2000) Re-modelling of nuclear architecture in quiescent and senescent human fibroblasts. Curr Biol 10(3):149–152

Brown JM, Leach J, Reittie JE, Atzberger A, Lee-Prudhoe J, Wood WG, Higgs DR, Iborra FJ, Buckle VJ (2006) Coregulated human globin genes are frequently in spatial proximity when active. J Cell Biol 172(2):177–187

Brown JM, Green J, das Neves RP, Wallace HA, Smith AJ, Hughes J, Gray N, Taylor S, Wood WG, Higgs DR, Iborra FJ, Buckle VJ (2008) Association between active genes occurs at nuclear speckles and is modulated by chromatin environment. J Cell Biol 182(6):1083–1097. doi:10.1083/jcb.200803174

Burke B, Stewart CL (2013) The nuclear lamins: flexibility in function. Nat Rev Mol Cell Biol 14 (1):13–24. doi:10.1038/nrm3488

Cavalli G, Misteli T (2013) Functional implications of genome topology. Nat Struct Mol Biol 20 (3):290–299. doi:10.1038/nsmb.2474

Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li GW, Park J, Blackburn EH, Weissman JS, Qi LS, Huang B (2013) Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell 155(7):1479–1491. doi:10.1016/j.cell.2013.12.001

Chiarle R, Zhang Y, Frock RL, Lewis SM, Molinie B, Ho YJ, Myers DR, Choi VW, Compagno M, Malkin DJ, Neuberg D, Monti S, Giallourakis CC, Gostissa M, Alt FW (2011) Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells. Cell 147(1):107–119. doi:10.1016/j.cell.2011.07.049

Clemson CM, Hall LL, Byron M, McNeil J, Lawrence JB (2006) The X chromosome is organized into a gene-rich outer rim and an internal core containing silenced nongenic sequences. Proc Natl Acad Sci USA 103(20):7688–7693

Clowney EJ, LeGros MA, Mosley CP, Clowney FG, Markenskoff-Papadimitriou EC, Myllys M, Barnea G, Larabell CA, Lomvardas S (2012) Nuclear aggregation of olfactory receptor genes governs their monogenic expression. Cell 151(4):724–737. doi:10.1016/j.cell.2012.09.043

Cremer T, Cremer C (2001) Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nat Rev Genet 2(4):292–301

Cremer T, Cremer M (2010) Chromosome territories. Cold Spring Harb Perspect Biol 2(3): a003889. doi:10.1101/cshperspect.a003889

Cremer M, Kupper K, Wagler B, Wizelman L, von Hase J, Weiland Y, Kreja L, Diebold J, Speicher MR, Cremer T (2003) Inheritance of gene density-related higher order chromatin arrangements in normal and tumor cell nuclei. J Cell Biol 162(5):809–820

Cremer T, Cremer C, Lichter P (2014) Recollections of a scientific journey published in human genetics: from chromosome territories to interphase cytogenetics and comparative genome hybridization. Hum Genet 133(4):403–416. doi:10.1007/s00439-014-1425-5

Croft JA, Bridger JM, Boyle S, Perry P, Teague P, Bickmore WA (1999) Differences in the localization and morphology of chromosomes in the human nucleus. J Cell Biol 145:1119–1131

Dekker J, Misteli T (2015) Long-range genome interactions Epigenetics, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Del Prete S, Arpon J, Sakai K, Andrey P, Gaudin V (2014) Nuclear architecture and chromatin dynamics in interphase nuclei of Arabidopsis thaliana. Cytogenet Genome Res 143 (1-3):28–50. doi:10.1159/000363724

Dellaire G, Kepkay R, Bazett-Jones DP (2009) High resolution imaging of changes in the structure and spatial organization of chromatin, gamma-H2A.X and the MRN complex within etoposide-induced DNA repair foci. Cell Cycle 8(22):3750–3769

Demmerle J, Koch AJ, Holaska JM (2013) Emerin and histone deacetylase 3 (HDAC3) cooperatively regulate expression and nuclear positions of MyoD, Myf5, and Pax7 genes during myogenesis. Chromosome Res 21(8):765–779. doi:10.1007/s10577-013-9381-9

Dittmer TA, Misteli T (2011) The lamin protein family. Genome Biol 12(5):222. doi:10.1186/gb-2011-12-5-222

Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485 (7398):376–380. doi:10.1038/nature11082

Egecioglu D, Brickner JH (2011) Gene positioning and expression. Curr Opin Cell Biol 23 (3):338–345. doi:10.1016/j.ceb.2011.01.001

Fanucchi S, Shibayama Y, Burd S, Weinberg MS, Mhlanga MM (2013) Chromosomal contact permits transcription between coregulated genes. Cell 155(3):606–620. doi:10.1016/j.cell.2013.09.051

Ferrai C, de Castro IJ, Lavitas L, Chotalia M, Pombo A (2010a) Gene positioning. Cold Spring Harb Perspect Biol 2(6):a000588. doi:10.1101/cshperspect.a000588

Ferrai C, Xie SQ, Luraghi P, Munari D, Ramirez F, Branco MR, Pombo A, Crippa MP (2010b) Poised transcription factories prime silent uPA gene prior to activation. PLoS Biol 8(1), e1000270. doi:10.1371/journal.pbio.1000270

Finlan LE, Sproul D, Thomson I, Boyle S, Kerr E, Perry P, Ylstra B, Chubb JR, Bickmore WA (2008) Recruitment to the nuclear periphery can alter expression of genes in human cells. PLoS Genet 4(3):e1000039

Foster HA, Griffin DK, Bridger JM (2012) Interphase chromosome positioning in in vitro porcine cells and ex vivo porcine tissues. BMC Cell Biol 13:30. doi:10.1186/1471-2121-13-30

Franco S, Gostissa M, Zha S, Lombard DB, Murphy MM, Zarrin AA, Yan C, Tepsuporn S, Morales JC, Adams MM, Lou Z, Bassing CH, Manis JP, Chen J, Carpenter PB, Alt FW (2006) H2AX prevents DNA breaks from progressing to chromosome breaks and translocations. Mol Cell 21(2):201–214. doi:10.1016/j.molcel.2006.01.005

Goodarzi AA, Noon AT, Deckbar D, Ziv Y, Shiloh Y, Lobrich M, Jeggo PA (2008) ATM signaling facilitates repair of DNA double-strand breaks associated with heterochromatin. Mol Cell 31(2):167–177. doi:10.1016/j.molcel.2008.05.017

Groth A, Rocha W, Verreault A, Almouzni G (2007) Chromatin challenges during DNA replication and repair. Cell 128(4):721–733. doi:10.1016/j.cell.2007.01.030

Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, Eussen BH, de Klein A, Wessels L, de Laat W, van Steensel B (2008) Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature 453(7197):948–951. doi:10.1038/nature06947

Hacisuleyman E, Goff LA, Trapnell C, Williams A, Henao-Mejia J, Sun L, McClanahan P, Hendrickson DG, Sauvageau M, Kelley DR, Morse M, Engreitz J, Lander ES, Guttman M, Lodish HF, Flavell R, Raj A, Rinn JL (2014) Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre. Nat Struct Mol Biol 21(2):198–206. doi:10.1038/nsmb.2764

Haffner MC, Aryee MJ, Toubaji A, Esopi DM, Albadine R, Gurel B, Isaacs WB, Bova GS, Liu W, Xu J, Meeker AK, Netto G, De Marzo AM, Nelson WG, Yegnasubramanian S (2010) Androgen-induced TOP2B-mediated double-strand breaks and prostate cancer gene rearrangements. Nat Genet 42(8):668–675. doi:10.1038/ng.613

Hakim O, Sung MH, Voss TC, Splinter E, John S, Sabo PJ, Thurman RE, Stamatoyannopoulos JA, de Laat W, Hager GL (2011) Diverse gene reprogramming events occur in the same spatial clusters of distal regulatory elements. Genome Res 21(5):697–706. doi:10.1101/gr.111153.110

Hakim O, Resch W, Yamane A, Klein I, Kieffer-Kwon KR, Jankovic M, Oliveira T, Bothmer A, Voss TC, Ansarah-Sobrinho C, Mathe E, Liang G, Cobell J, Nakahashi H, Robbiani DF, Nussenzweig A, Hager GL, Nussenzweig MC, Casellas R (2012) DNA damage defines sites of recurrent chromosomal translocations in B lymphocytes. Nature 484(7392):69–74. doi:10.1038/nature10909

Hall LL, Carone DM, Gomez AV, Kolpa HJ, Byron M, Mehta N, Fackelmayer FO, Lawrence JB (2014) Stable C0T-1 repeat RNA is abundant and is associated with euchromatic interphase chromosomes. Cell 156(5):907–919. doi:10.1016/j.cell.2014.01.042

Harewood L, Schutz F, Boyle S, Perry P, Delorenzi M, Bickmore WA, Reymond A (2010) The effect of translocation-induced nuclear reorganization on gene expression. Genome Res 20 (5):554–564. doi:10.1101/gr.103622.109

Hewitt SL, High FA, Reiner SL, Fisher AG, Merkenschlager M (2004) Nuclear repositioning marks the selective exclusion of lineage-inappropriate transcription factor loci during T helper cell differentiation. Eur J Immunol 34(12):3604–3613

Hiratani I, Ryba T, Itoh M, Yokochi T, Schwaiger M, Chang CW, Lyou Y, Townes TM, Schübeler D, Gilbert DM (2008) Global reorganization of replication domains during embryonic stem cell differentiation. PLoS Biol 6, e245. doi:10.1371/journal.pbio.0060245

Hoffmann K, Sperling K, Olins AL, Olins DE (2007) The granulocyte nucleus and lamin B receptor: avoiding the ovoid. Chromosoma 116(3):227–235. doi:10.1007/s00412-007-0094-8

Khanna N, Hu Y, Belmont AS (2014) HSP70 transgene directed motion to nuclear speckles facilitates heat shock activation. Curr Biol 24(10):1138–1144. doi:10.1016/j.cub.2014.03.053

Kind J, Pagie L, Ortabozkoyun H, Boyle S, de Vries SS, Janssen H, Amendola M, Nolen LD, Bickmore WA, van Steensel B (2013) Single-cell dynamics of genome-nuclear lamina interactions. Cell 153(1):178–192. doi:10.1016/j.cell.2013.02.028

Klein IA, Resch W, Jankovic M, Oliveira T, Yamane A, Nakahashi H, Di Virgilio M, Bothmer A, Nussenzweig A, Robbiani DF, Casellas R, Nussenzweig MC (2011) Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. Cell 147(1):95–106. doi:10.1016/j.cell.2011.07.048

Knight M, Ittiprasert W, Odoemelam EC, Adema CM, Miller A, Raghavan N, Bridger JM (2011) Non-random organization of the Biomphalaria glabrata genome in interphase Bge cells and the spatial repositioning of activated genes in cells co-cultured with Schistosoma mansoni. Int J Parasitol 41(1):61–70. doi:10.1016/j.ijpara.2010.07.015

Kosak ST, Skok JA, Medina KL, Riblet R, Le Beau MM, Fisher AG, Singh H (2002) Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. Science 296 (5565):158–162

Kruhlak MJ, Celeste A, Dellaire G, Fernandez-Capetillo O, Muller WG, McNally JG, Bazett-Jones DP, Nussenzweig A (2006) Changes in chromatin structure and mobility in living cells at sites of DNA double-strand breaks. J Cell Biol 172(6):823–834. doi:10.1083/jcb.200510015

Kubben N, Adriaens M, Meuleman W, Voncken JW, van Steensel B, Misteli T (2012) Mapping of lamin A- and progerin-interacting genome regions. Chromosoma 121(5):447–464. doi:10.1007/s00412-012-0376-7

Kumaran RI, Spector DL (2008) A genetic locus targeted to the nuclear periphery in living cells maintains its transcriptional competence. J Cell Biol 180(1):51–65

Li C, Shi Z, Zhang L, Huang Y, Liu A, Jin Y, Yu Y, Bai J, Chen D, Gendron C, Liu X, Fu S (2010) Dynamic changes of territories 17 and 18 during EBV-infection of human lymphocytes. Mol Biol Rep 37(5):2347–2354. doi:10.1007/s11033-009-9740-y

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326(5950):289–293. doi:10.1126/science.1181369

Lin C, Yang L, Tanasa B, Hutt K, Ju BG, Ohgi K, Zhang J, Rose DW, Fu XD, Glass CK, Rosenfeld MG (2009) Nuclear receptor-induced chromosomal proximity and DNA breaks underlie specific translocations in cancer. Cell 139(6):1069–1083. doi:10.1016/j.cell.2009.11.030

Lukasova E, Kozubek S, Kozubek M, Kjeronska J, Ryznar L, Horakova J, Krahulcova E, Horneck G (1997) Localisation and distance between ABL and BCR genes in interphase nuclei of bone marrow cells of control donors and patients with chronic myeloid leukaemia. Hum Genet 100 (5-6):525–535

Makatsori D, Kourmouli N, Polioudaki H, Shultz LD, McLean K, Theodoropoulos PA, Singh PB, Georgatos SD (2004) The inner nuclear membrane protein lamin B receptor forms distinct microdomains and links epigenetically marked chromatin to the nuclear envelope. J Biol Chem 279(24):25567–25573. doi:10.1074/jbc.M313606200

Malhas A, Lee CF, Sanders R, Saunders NJ, Vaux DJ (2007) Defects in lamin B1 expression or processing affect interphase chromosome position and gene expression. J Cell Biol 176 (5):593–603. doi:10.1083/jcb.200607054

Mani RS, Tomlins SA, Callahan K, Ghosh A, Nyati MK, Varambally S, Palanisamy N, Chinnaiyan AM (2009) Induced chromosomal proximity and gene fusions in prostate cancer. Science 326(5957):1230. doi:10.1126/science.1178124

Matarazzo MR, Boyle S, D'Esposito M, Bickmore WA (2007) Chromosome territory reorganization in a human disease with altered DNA methylation. Proc Natl Acad Sci USA 104 (42):16546–16551. doi:10.1073/pnas.0702924104

Mathas S, Kreher S, Meaburn KJ, Johrens K, Lamprecht B, Assaf C, Sterry W, Kadin ME, Daibata M, Joos S, Hummel M, Stein H, Janz M, Anagnostopoulos I, Schrock E, Misteli T, Dorken B (2009) Gene deregulation and spatial genome reorganization near breakpoints prior to formation of translocations in anaplastic large cell lymphoma. Proc Natl Acad Sci USA 106 (14):5831–5836. doi:10.1073/pnas.0900912106

Mathur S, Glogowska A, McAvoy E, Righolt C, Rutherford J, Willing C, Banik U, Ruthirakuhan M, Mai S, Garcia A (2014) Three-dimensional quantitative imaging of telomeres in buccal cells identifies mild, moderate, and severe Alzheimer's disease patients. J Alzheimers Dis 39(1):35–48. doi:10.3233/JAD-130866

Mattout A, Pike BL, Towbin BD, Bank EM, Gonzalez-Sandoval A, Stadler MB, Meister P, Gruenbaum Y, Gasser SM (2011) An EDMD mutation in C. elegans lamin blocks muscle-specific gene relocation and compromises muscle integrity. Curr Biol 21(19):1603–1614. doi:10.1016/j.cub.2011.08.030

McCord RP, Nazario-Toole A, Zhang H, Chines PS, Zhan Y, Erdos MR, Collins FS, Dekker J, Cao K (2013) Correlated alterations in genome organization, histone methylation, and DNA-lamin A/C interactions in Hutchinson-Gilford progeria syndrome. Genome Res 23(2):260–269. doi:10.1101/gr.138032.112

Meaburn KJ, Misteli T (2007) Cell biology: chromosome territories. Nature 445(7126):379–781. doi:10.1038/445379a

Meaburn KJ, Misteli T (2008) Locus-specific and activity-independent gene repositioning during early tumorigenesis. J Cell Biol 180(1):39–50

Meaburn KJ, Levy N, Toniolo D, Bridger JM (2005) Chromosome positioning is largely unaffected in lymphoblastoid cell lines containing emerin or A-type lamin mutations. Biochem Soc Trans 33(Pt 6):1438–1440. doi:10.1042/BST20051438

Meaburn KJ, Cabuy E, Bonne G, Levy N, Morris GE, Novelli G, Kill IR, Bridger JM (2007a) Primary laminopathy fibroblasts display altered genome organization and apoptosis. Aging Cell 6(2):139–153. doi:10.1111/j.1474-9726.2007.00270.x

Meaburn KJ, Misteli T, Soutoglou E (2007b) Spatial genome organization in the formation of chromosomal translocations. Semin Cancer Biol 17(1):80–90. doi:10.1016/j.semcancer.2006.10.008

Meaburn KJ, Newbold RF, Bridger JM (2008) Positioning of human chromosomes in murine cell hybrids according to synteny. Chromosoma 117(6):579–591. doi:10.1007/s00412-008-0175-3

Meaburn KJ, Gudla PR, Khan S, Lockett SJ, Misteli T (2009) Disease-specific gene repositioning in breast cancer. J Cell Biol 187(6):801–812. doi:10.1083/jcb.200909127

Mehta IS, Amira M, Harvey AJ, Bridger JM (2010) Rapid chromosome territory relocation by nuclear motor activity in response to serum removal in primary human fibroblasts. Genome Biol 11(1):R5. doi:10.1186/gb-2010-11-1-r5

Mehta IS, Eskiw CH, Arican HD, Kill IR, Bridger JM (2011) Farnesyltransferase inhibitor treatment restores chromosome territory positions and active chromosome dynamics in Hutchinson-Gilford progeria syndrome cells. Genome Biol 12(8):R74. doi:10.1186/gb-2011-12-8-r74

Mehta IS, Kulashreshtha M, Chakraborty S, Kolthur-Seetharam U, Rao BJ (2013) Chromosome territories reposition during DNA damage-repair response. Genome Biol 14(12):R135. doi:10.1186/gb-2013-14-12-r135

Mewborn SK, Puckelwartz MJ, Abuisneineh F, Fahrenbach JP, Zhang Y, MacLeod H, Dellefave L, Pytel P, Selig S, Labno CM, Reddy K, Singh H, McNally E (2010) Altered chromosomal positioning, compaction, and gene expression with a lamin A/C gene mutation. PLoS One 5(12), e14342. doi:10.1371/journal.pone.0014342

Misteli T (2010) Higher-order genome organization in human disease. Cold Spring Harb Perspect Biol 2(8):a000794. doi:10.1101/cshperspect.a000794

Morey C, Kress C, Bickmore WA (2009) Lack of bystander activation shows that localization exterior to chromosome territories is not sufficient to up-regulate gene expression. Genome Res 19(7):1184–1194. doi:10.1101/gr.089045.108

Murata S, Nakazawa T, Ohno N, Terada N, Iwashina M, Mochizuki K, Kondo T, Nakamura N, Yamane T, Iwasa S, Ohno S, Katoh R (2007) Conservation and alteration of chromosome territory arrangements in thyroid carcinoma cell nuclei. Thyroid 17(6):489–496. doi:10.1089/thy.2006.0328

Murmann AE, Gao J, Encinosa M, Gautier M, Peter ME, Eils R, Lichter P, Rowley JD (2005) Local gene density predicts the spatial position of genetic loci in the interphase nucleus. Exp Cell Res 311(1):14–26

Nagai S, Dubrana K, Tsai-Pflugfelder M, Davidson MB, Roberts TM, Brown GW, Varela E, Hediger F, Gasser SM, Krogan NJ (2008) Functional targeting of DNA damage to a nuclear pore-associated SUMO-dependent ubiquitin ligase. Science 322(5901):597–602. doi:10.1126/science.1162790

Neves H, Ramos C, da Silva MG, Parreira A, Parreira L (1999) The nuclear topography of ABL, BCR, PML, and RARalpha genes: evidence for gene proximity in specific phases of the cell cycle and stages of hematopoietic differentiation. Blood 93(4):1197–1207

Parada LA, McQueen PG, Munson PJ, Misteli T (2002) Conservation of relative chromosome positioning in normal and cancer cells. Curr Biol 12(19):1692–1697

Parada L, McQueen P, Misteli T (2004) Tissue-specific spatial organization of genomes. Genome Biol 7:R44

Paz N, Zabala A, Royo F, Garcia-Orad A, Zugaza JL, Parada LA (2013) Combined fluorescent-chromogenic in situ hybridization for identification and laser microdissection of interphase chromosomes. PLoS One 8(4):e60238. doi:10.1371/journal.pone.0060238

Peric-Hupkes D, Meuleman W, Pagie L, Bruggeman SW, Solovei I, Brugman W, Graf S, Flicek P, Kerkhoven RM, van Lohuizen M, Reinders M, Wessels L, van Steensel B (2010) Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. Mol Cell 38(4):603–613. doi:10.1016/j.molcel.2010.03.016

Pickersgill H, Kalverda B, de Wit E, Talhout W, Fornerod M, van Steensel B (2006) Characterization of the Drosophila melanogaster genome at the nuclear lamina. Nat Genet 38(9):1005–1014. doi:10.1038/ng1852

Ragoczy T, Bender MA, Telling A, Byron R, Groudine M (2006) The locus control region is required for association of the murine beta-globin locus with engaged transcription factories during erythroid maturation. Genes Dev 20(11):1447–1457

Reddy KL, Zullo JM, Bertolino E, Singh H (2008) Transcriptional repression mediated by repositioning of genes to the nuclear lamina. Nature 452(7184):243–247

Rieder D, Ploner C, Krogsdam AM, Stocker G, Fischer M, Scheideler M, Dani C, Amri EZ, Muller WG, McNally JG, Trajanoski Z (2014) Co-expressed genes prepositioned in spatial neighborhoods stochastically associate with SC35 speckles and RNA polymerase II factories. Cell Mol Life Sci 71(9):1741–1759. doi:10.1007/s00018-013-1465-3

Rocha PP, Micsinai M, Kim JR, Hewitt SL, Souza PP, Trimarchi T, Strino F, Parisi F, Kluger Y, Skok JA (2012) Close proximity to Igh is a contributing factor to AID-mediated translocations. Mol Cell 47(6):873–885. doi:10.1016/j.molcel.2012.06.036

Roix JJ, McQueen PG, Munson PJ, Parada LA, Misteli T (2003) Spatial proximity of translocation-prone gene loci in human lymphomas. Nat Genet 34(3):287–291

Roukos V, Misteli T (2014) The biogenesis of chromosome translocations. Nat Cell Biol 16 (4):293–300. doi:10.1038/ncb2941

Roukos V, Voss TC, Schmidt CK, Lee S, Wangsa D, Misteli T (2013) Spatial dynamics of chromosome translocations in living cells. Science 341(6146):660–664. doi:10.1126/science.1237150

Scaffidi P, Misteli T (2005) Reversal of the cellular phenotype in the premature aging disease Hutchinson-Gilford progeria syndrome. Nat Med 11(4):440–445. doi:10.1038/nm1204

Schoenfelder S, Sexton T, Chakalova L, Cope NF, Horton A, Andrews S, Kurukuti S, Mitchell JA, Umlauf D, Dimitrova DS, Eskiw CH, Luo Y, Wei CL, Ruan Y, Bieker JJ, Fraser P (2010) Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. Nat Genet 42(1):53–61. doi:10.1038/ng.496

Schotta G, Sengupta R, Kubicek S, Malin S, Kauer M, Callen E, Celeste A, Pagani M, Opravil S, De La Rosa-Velazquez IA, Espejo A, Bedford MT, Nussenzweig A, Busslinger M, Jenuwein T (2008) A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. Genes Dev 22(15):2048–2061. doi:10.1101/gad.476008

Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, Parrinello H, Tanay A, Cavalli G (2012) Three-dimensional folding and functional organization principles of the Drosophila genome. Cell 148(3):458–472. doi:10.1016/j.cell.2012.01.010

Sharma R, Meister P (2013) Nuclear organization in the nematode C. elegans. Curr Opin Cell Biol 25(3):395–402. doi:10.1016/j.ceb.2013.02.002

Shopland LS, Lynch CR, Peterson KA, Thornton K, Kepper N, Hase J, Stein S, Vincent S, Molloy KR, Kreth G, Cremer C, Bult CJ, O'Brien TP (2006) Folding and organization of a contiguous chromosome region according to the gene distribution pattern in primary genomic sequence. J Cell Biol 174(1):27–38

Shumaker DK, Dechat T, Kohlmaier A, Adam SA, Bozovsky MR, Erdos MR, Eriksson M, Goldman AE, Khuon S, Collins FS, Jenuwein T, Goldman RD (2006) Mutant nuclear lamin A leads to progressive alterations of epigenetic control in premature aging. Proc Natl Acad Sci USA 103(23):8703–8708. doi:10.1073/pnas.0602569103

Solovei I, Kreysing M, Lanctot C, Kosem S, Peichl L, Cremer T, Guck J, Joffe B (2009) Nuclear architecture of rod photoreceptor cells adapts to vision in mammalian evolution. Cell 137 (2):356–368. doi:10.1016/j.cell.2009.01.052

Solovei I, Wang AS, Thanisch K, Schmidt CS, Krebs S, Zwerger M, Cohen TV, Devys D, Foisner R, Peichl L, Herrmann H, Blum H, Engelkamp D, Stewart CL, Leonhardt H, Joffe B (2013) LBR and lamin A/C sequentially tether peripheral heterochromatin and inversely regulate differentiation. Cell 152(3):584–598. doi:10.1016/j.cell.2013.01.009

Sun HB, Shen J, Yokota H (2000) Size-dependent positioning of human chromosomes in interphase nuclei. Biophys J 79(1):184–190

Szczerbal I, Bridger JM (2010) Association of adipogenic genes with SC-35 domains during porcine adipogenesis. Chromosome Res 18(8):887–895. doi:10.1007/s10577-010-9176-1

Taimen P, Pfleghaar K, Shimi T, Moller D, Ben-Harush K, Erdos MR, Adam SA, Herrmann H, Medalia O, Collins FS, Goldman AE, Goldman RD (2009) A progeria mutation reveals functions for lamin A in nuclear assembly, architecture, and chromosome organization. Proc Natl Acad Sci USA 106(49):20788–20793. doi:10.1073/pnas.0911895106

Takizawa T, Gudla PR, Guo L, Lockett S, Misteli T (2008a) Allele-specific nuclear positioning of the monoallelically expressed astrocyte marker GFAP. Genes Dev 22(4):489–498. doi:10.1101/gad.1634608

Takizawa T, Meaburn KJ, Misteli T (2008b) The meaning of gene positioning. Cell 135(1):9–13. doi:10.1016/j.cell.2008.09.026

Timme S, Schmitt E, Stein S, Schwarz-Finsterle J, Wagner J, Walch A, Werner M, Hausmann M, Wiech T (2011) Nuclear position and shape deformation of chromosome 8 territories in pancreatic ductal adenocarcinoma. Anal Cell Pathol 34(1–2):21–33. doi:10.3233/ACP-2011-0004

Towbin BD, Gonzalez-Aguilera C, Sack R, Gaidatzis D, Kalck V, Meister P, Askjaer P, Gasser SM (2012) Step-wise methylation of histone H3K9 positions heterochromatin at the nuclear periphery. Cell 150(5):934–947. doi:10.1016/j.cell.2012.06.051

van Steensel B, Dekker J (2010) Genomics tools for unraveling chromosome architecture. Nat Biotechnol 28(10):1089–1095. doi:10.1038/nbt.1680

Wiech T, Timme S, Riede F, Stein S, Schuricke M, Cremer C, Werner M, Hausmann M, Walch A (2005) Human archival tissues provide a valuable source for the analysis of spatial genome organization. Histochem Cell Biol 123(3):229–238. doi:10.1007/s00418-005-0768-3

Wiech T, Stein S, Lachenmaier V, Schmitt E, Schwarz-Finsterle J, Wiech E, Hildenbrand G, Werner M, Hausmann M (2009) Spatial allelic imbalance of BCL2 genes and chromosome 18 territories in nonneoplastic and neoplastic cervical squamous epithelium. Eur Biophys J 38 (6):793–806

Williams RR, Azuara V, Perry P, Sauer S, Dvorkina M, Jorgensen H, Roix J, McQueen P, Misteli T, Merkenschlager M, Fisher AG (2006) Neural induction promotes large-scale chromatin reorganisation of the Mash1 locus. J Cell Sci 119(Pt 1):132–140

Wong X, Luperchio TR, Reddy KL (2014) NET gains and losses: the role of changing nuclear envelope proteomes in genome regulation. Curr Opin Cell Biol 28:105–120. doi:10.1016/j.ceb.2014.04.005

Zeitz MJ, Ay F, Heidmann JD, Lerner PL, Noble WS, Steelman BN, Hoffman AR (2013) Genomic interaction profiles in breast cancer reveal altered chromatin architecture. PLoS One 8(9): e73974. doi:10.1371/journal.pone.0073974

Zhang Y, McCord RP, Ho YJ, Lajoie BR, Hildebrand DG, Simon AC, Becker MS, Alt FW, Dekker J (2012) Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. Cell 148(5):908–921. doi:10.1016/j.cell.2012.02.002

Zimmer C, Fabre E (2011) Principles of chromosomal organization: lessons from yeast. J Cell Biol 192(5):723–733. doi:10.1083/jcb.201010058

Zink D, Amaral MD, Englmann A, Land S, Clarke LA, Rudolph C, Alt F, Luther K, Braz C, Sadoni N, Rosenecker J, Schindelhauer D (2004a) Transcription-dependent spatial arrangement of CFTR and adjacent genes in human cell nuclei. J Cell Biol I166:815–825

Zink D, Fische AH, Nickerson JA (2004b) Nuclear structure in cancer cells. Nat Rev Cancer 4 (9):677–687

Zuleger N, Boyle S, Kelly DA, de Las Heras JI, Lazou V, Korfali N, Batrakou DG, Randles KN, Morris GE, Harrison DJ, Bickmore WA, Schirmer EC (2013) Specific nuclear envelope transmembrane proteins can promote the location of chromosomes to and from the nuclear periphery. Genome Biol 14(2):R14. doi:10.1186/gb-2013-14-2-r14

Zullo JM, Demarco IA, Pique-Regi R, Gaffney DJ, Epstein CB, Spooner CJ, Luperchio TR, Bernstein BE, Pritchard JK, Reddy KL, Singh H (2012) DNA sequence-dependent compartmentalization and silencing of chromatin at the nuclear lamina. Cell 149(7):1474–1487. doi:10.1016/j.cell.2012.04.035

Telomeres and Chromosome Stability

Tsz Wai Chu and Chantal Autexier

1 Discovery of Telomeres

The discovery of telomeres, originating from the Greek words "telos" (end) and "meros" (part), was the unexpected outcome of a simple experiment conducted by Hermann J. Müller in the 1930s (Müller 1938). This brilliant geneticist irradiated Drosophila chromosomes and made the astonishing observation that the ends of the DNA, unlike the rest of the genome, evade aberrant genomic modifications such as inversions and deletions, suggesting the presence of a protective mechanism. In 1941, the famous geneticist Barbara McClintock's work further strengthened the importance of telomeres in the maintenance of genome integrity (McClintock 1941). Using a special *Zea mays* strain characterized by high frequencies of chromosomal breakage and fusion, she readily saw the appearance of dicentric chromosomes. Surprisingly, she also noted that chromosome ends are protected from fatal "rupture-fusion-bridge" cycles by the addition of new DNA. In the 1980s, the combination of Blackburn, Greider and Szostak's experiments revealed that the telomere is composed of short tandem G-rich hexameric repeats that are evolutionarily and functionally conserved across different organisms (Szostak and

T.W. Chu
Division of Experimental Medicine, McGill University, Montréal, QC, Canada

Bloomfield Center for Aging, Lady Davis Institute, Jewish General Hospital, Montréal, QC, Canada

C. Autexier (✉)
Division of Experimental Medicine, McGill University, Montréal, QC, Canada

Bloomfield Center for Aging, Lady Davis Institute, Jewish General Hospital, Montréal, QC, Canada

Department of Anatomy and Cell Biology, McGill University, Montréal, QC, Canada
e-mail: chantal.autexier@mcgill.ca

D.P. Bazett-Jones, G. Dellaire (eds.), *The Functional Nucleus*,
DOI 10.1007/978-3-319-38882-3_6

Blackburn 1982). Their work also led to the discovery of telomerase, the enzyme responsible for telomere synthesis (Greider and Blackburn 1985).

2 Telomere: Structure and Function

Telomeres are nucleoprotein structures located at the extremities of the chromosomes, and are highly conserved from yeast to human. The sequence and the length of telomeres vary among different organisms. Depending on the tissue type and the age of the organism, telomere length differs and its homeostasis is under complex genetic control. In humans, telomeres are comprised of a 9–15 kb double stranded region of the tandem hexameric repeats (TTAGGG)n, and a 3′ single-stranded G-rich overhang (150–200 nucleotides). Due to the linear organization of the eukaryotic chromosomal ends, which resemble double-stranded breaks, protection of the telomeres requires a capping mechanism to avoid the detrimental activation and repair of the chromosome end by the DNA damage response (DDR) machinery. In 1999, this physical protective "cap", termed the "T-loop", was identified and visualized by electron microscopy of human and mouse chromosomal DNA (Griffith et al. 1999). The T-loop is a remarkably sophisticated lariat-like secondary structure that is dynamic throughout the cell replication cycle (Fig. 1a). It is generated by the backward folding of the single-stranded 3′ overhang followed by its invasion into the double-stranded region of the telomeres, resulting in the formation of a second loop known as the displacement D-loop. This higher order organization is formed and maintained by six telomeric proteins, known as the "shelterin" complex (discussed below).

During the normal cellular aging process, telomeres shorten after each round of cell division due to nucleolytic processing and failure of the semi-conservative nature of DNA polymerase to completely replicate the ends of linear chromosomes during lagging strand DNA synthesis ("end-replication problem"). Eventually, in the absence of a telomere maintenance mechanism (TMM), cells with telomeres that have reached a certain threshold length will irreversibly enter a state of permanent growth arrest known as the "Hayflick limit', or senescence. In human cells, the presence of five nonfusogenic uncapped, or dysfunctional, telomeres in G1 is sufficient to trigger p53-mediated senescence (Kaul et al. 2012).

Occasionally, a cell can bypass the activation of senescence via the inactivation of tumor suppressor pathways and continues to divide despite increased levels of DNA damage. The accumulation of genomic instability typically triggers apoptosis. However, ~one in a million cells will undergo a change that confers a growth advantage, ultimately resulting in uncontrolled cell division and tumorigenesis. Thus, telomere length can act as a "molecular clock", limiting the replicative potential of normal somatic cells to prevent unlimited cell division and carcinogenesis.

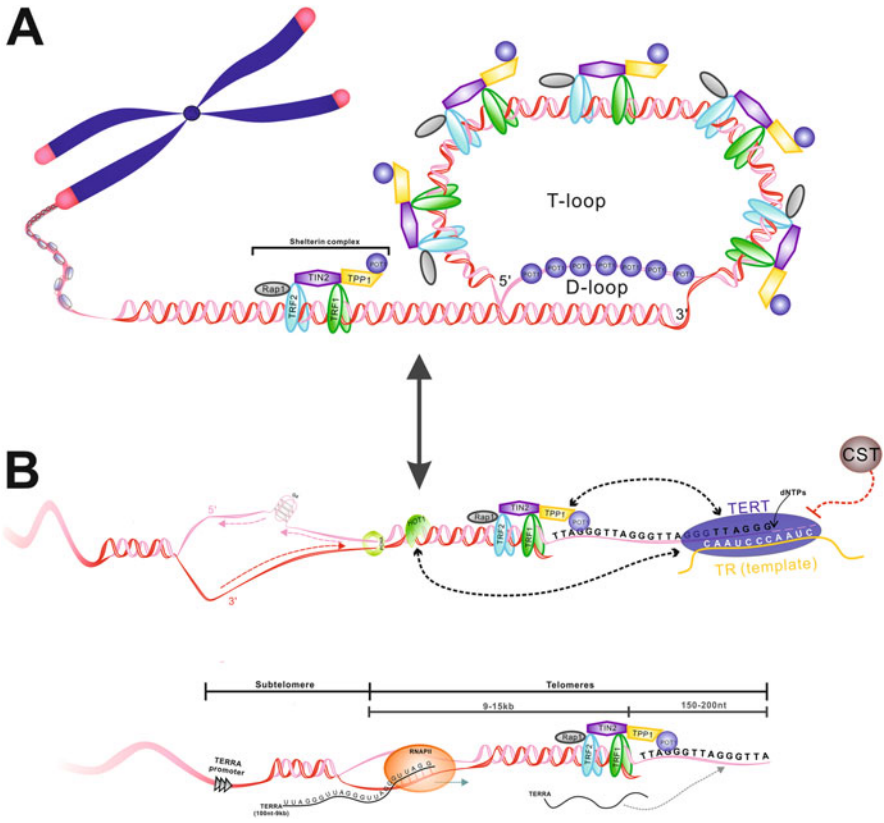

Fig. 1 Telomere structure. (**a**) Telomeres are nucleoprotein structures located at the end of linear chromosomes. The shelterin complex forms and maintains the protective T-loop cap. The T-loop is generated by the invasion of the 3′ G-rich overhang into the double-stranded region, resulting in the D-loop formation (*top*). During telomere replication, the T-loop structure is disrupted to allow recruitment of telomerase to the telomeres by TPP1 and HOT1 proteins. G4 structures and the CST complex negatively impact telomere length by inhibiting telomerase (*bottom*). (**b**) TERRA are transcribed by RNA polymerase II (RNAPII). TERRA transcription starts from the TERRA promoter located in the subtelomeres and continues into the telomeric regions. The complementarities of the TERRA transcripts with the telomeres allow their association and as such, negatively regulate telomere length

3 Telomere-Associated Proteins

Telomere length homeostasis is a crucial determinant of stem cell survival, cellular aging as well as carcinogenesis. Low levels of telomerase expression in human somatic stem cells and progenitor cells is necessary to sustain the self-renewal capacity and the pluripotency of these cells but is insufficient to completely halt telomere shortening (Weissbein et al. 2014). In mouse and human, the efficiency of cellular reprogramming and rate of cell proliferation directly correlates with

telomere length (Weissbein et al. 2014). Various premature aging diseases such as dyskeratosis congenita and aplastic anemia highlight the crucial role of TMM in the stem cell compartments. Telomere maintenance and protection rely on a complex network of telomere-associated proteins which inter-regulate each other via functional interactions.

3.1 Shelterin Complex

Early studies on the shelterin complex (Fig. 1), also termed the "telosome", mainly described its role as a "shelter" to protect the telomeres from being recognized as a double-stranded break. Later findings demonstrated the importance of this complex beyond simply forming and maintaining the T-loop structure to ensure end-protection. Each subunit of the shelterin complex also plays a crucial role in mediating telomere length homeostasis. Shelterin is composed of six members: the double-stranded DNA binding proteins TRF1 and TRF2 (telomeric repeat-binding factor 1 and 2), the single-stranded DNA binding protein, POT1 (protection of telomeres 1), the TRF2-binding protein Rap1 (repressor/activator protein 1), and TIN2 (TRF1-interacting protein 2), which bridges the TPP1 (formerly known as TINT1, PTOP and PIP1)-POT1 heterodimer to the rest of the telosome (de Lange 2010). The binding affinity of TRF1/2 and POT1 to the telomeric sequence confers the sequence specificity of the shelterin complex for the telomeres (de Lange 2010).

3.2 Roles of the Shelterin Complex at the Telomeres: T-Loop Maintenance, Telomere Length Regulation and Suppression of DDR

The shelterin complex is highly conserved in mammals and can associate with several factors involved in telomere replication, length regulation and end-protection (de Lange 2010). TRF1/2 are negative regulators of telomere length, containing a C-terminal Myb-type DNA binding domain which allows their association with the double-stranded telomeric tracts (de Lange 2010). Post-translational modifications (e.g. ubiquitinylation, sumoylation, phosphorylation) of these subunits are important for their proper functioning and regulation (Peuscher and Jacobs 2012). The poly(ADP-ribose) polymerase tankyrase 1 (Tank1) binds TRF1 and contributes to telomere length control and resolution of mitosis (Donigian and de Lange 2007). In humans specifically, Tank1 acts as a positive regulator of telomere length by binding and removing TRF1 from the telomeres, thus allowing telomerase access to the telomeres. PinX1, a negative regulator of telomerase activity involved in cell cycle-specific telomerase trafficking, interacts with the TRF homology (TRFH) domain of TRF1 to control telomere

length (Lu et al. 2013; Cheung et al. 2012). Structural and mutational studies have demonstrated the importance of the TRFH domain in protein homodimerization, high-affinity DNA binding and recruitment to telomeres (Fairall et al. 2001).

Extensive replicative telomere erosion causes telomere deprotection and the activation of the tumor suppressor p53. Phosphorylation of p53 by the ATM (ataxia telangiectasia mutated) and ATR (ataxia telangiectasia and Rad3 related) kinases results in cellular senescence or apoptosis (Karlseder et al. 1999). An elegant study by de Lange's group directly addressed the role of the telosome in telomere end-protection, by completely removing all the components of the shelterin complex from mouse telomeres (Sfeir and de Lange 2012). They reported the recruitment of general DDR factors by the shelterin complex and the activation of six distinct repair pathways: ATM, ATR, classical-NHEJ (non-homologous end joining), alternative-NHEJ, HR (homologous recombination) and resection (Sfeir and de Lange 2012).

The function of TRF2 and POT1 in protection and preventing telomere end-to-end fusion is well established (Lu et al. 2013). Removing TRF2 or the use of a dominant negative TRF2 leads to spontaneous telomere deprotection (uncapping) and cellular senescence. The association of TRF2 with Apollo (SMN1/Pso2-type nuclease) promotes telomere end-processing and the generation of the 3′ overhang by the resection of the 5′ strand, thus supporting the efficient formation and stabilization of the T-loop (Longhese et al. 2012). The use of TRF2-depleted mouse embryonic fibroblasts (MEFs) provided evidence for the function of TRF2 in the repression of the ATM pathway. In $TRF2^{-/-}$ MEFs, accumulation of the MRN (Mre11, Rad50 and Nbs1) complex at the leading telomeric strand and phosphorylation of H2AX in the subtelomeric and telomeric chromatin are readily observable (de Lange 2010). High levels of telomere-dysfunction-induced foci (TIFs) characterized by enrichment of DDR factors such as 53BP1 at the telomeres are also induced (de Lange 2010). A TRF2-ATM-p53 positive-feedback loop implicating the Siah1 E3 ubiquitin ligase was proposed to function in the amplification of the telomere uncapping-induced p53-mediated DDR (Horikawa et al. 2011). Their results suggest that the reduced level of TRF2 induces ATM pathway activation to phosphorylate p53. Activated p53 subsequently activates Siah1 to target TRF2 for degradation, leading to a further decrease in TRF2 levels at the telomeres. TRF2 is also important to prevent unwanted telomeric repair and end-to-end fusion by the NHEJ pathway and the generation of lethal dicentric chromosomes at mitosis. TRF2 hinders the association of the NHEJ proteins Ku70/Ku80 with the telomere by promoting the formation and stabilization of the T-loop structure (de Lange 2010). Thus, TRF2 plays a central role in telomere end-protection.

Protection of the single-stranded region of the telomeres is mediated by POT1, a single-stranded DNA binding protein with two oligosaccharide/oligonucleotide binding (OB) folds. The localization of POT1 relies on the interaction of its C-terminus with TPP1 to form a heterodimer linked to the rest of the shelterin complex via the TPP1-TIN2 interaction. While the telomeres are in an extendible state (open T-loop), POT1 coats the 3′ G-rich overhang. However, when the

telomeres are in a closed configuration, POT1 binds to the short single-stranded DNA segment of the D-loop (de Lange 2010; Cristofari and Lingner 2006) (Fig. 1a). POT1 regulates telomeric 5′-end resection by inhibiting nucleases such as Apollo, and therefore, is a determinant of the 5′ chromosomal end sequences (Hockemeyer et al. 2005; Latrick and Cech 2010). POT1 negatively controls telomere length by limiting the access of telomerase to the telomeres and inhibits telomerase activity *in vitro* (Lei et al. 2004). Interestingly, when in a complex with TPP1, POT1 acts synergistically to enhance telomerase activity and telomere elongation, playing dual opposing roles in telomerase and telomere regulation (Wang and Lei 2011). Studies of mouse POT1 demonstrated its importance in repressing NHEJ and HR at telomeres (He et al. 2006). POT1 inhibits the ATR pathway and its interaction with TIN2 and TPP1 allows it to effectively compete with the ATR activation protein, RPA (replication protein A) for binding to telomeres (de Lange 2010). RPA is a single-stranded DNA binding heterotrimer which participates in several pathways of DNA metabolism, including DNA replication and damage repair (Zou et al. 2006).

TIN2 is the core subunit of the shelterin complex and cooperates with TRF2 to inhibit the activation of the ATM pathway, facilitates TPP1-mediated recruitment of telomerase and contributes to telomere length homeostasis (Lu et al. 2013). Mammalian Rap1 does not bind to telomeric DNA and its association with the shelterin complex relies on its interaction partner TRF2. Furthermore, using mass spectrometry, Songyang's group identified several DDR factors that interact with the Rap1-TRF2 complex, such as Rad50, Mre11, PARP1 and Ku86/70 (O'Connor et al. 2004). Another study identified the DNA repair protein BTBD12 (human orthologue of yeast SLX4, a nuclease implicated in resolution of recombination intermediates) as an interactor of Rap1-TRF2 and suggests a role for SLX4 in the repression of HR (Svendsen et al. 2009).

In addition to its function in recruiting POT1 to the telomeres, TPP1 also directly interacts with the telomerase essential N-terminal (TEN)-domain of hTERT, via seven residues termed the TEL patch which are located within its OB fold, to recruit telomerase to the telomeres (Zhong et al. 2012; Nandakumar et al. 2012; Schmidt et al. 2014). Importantly, the TPP1-POT1 heterodimer acts to stimulate telomerase enzyme processivity (discussed below), an important determinant of telomere length maintenance (Cifuentes-Rojas and Shippen 2012). Furthermore, TPP1 also interacts with the single-stranded DNA binding protein Stn1, a component of the CST complex (discussed below) involved in 5′-end resection (Wan et al. 2009). These, and new data, highlight the role of TPP1 in regulating chromosomal end-protection, telomerase recruitment and activation, and telomere length homeostasis set point (Sexton et al. 2014).

Although individual telosome components regulate the telomere negatively or positively, the telosome complex is a negative regulator of telomere length. Proper control of telomere homeostasis sets the telomere length within a range of nucleotides. This is accomplished using a negative feed-back system dependent on the number of telomere-bound shelterin proteins, principally the negative regulator POT1 (Diotti and Loayza 2011). The number of shelterin complexes bound to the

telomeres is directly proportional to telomere length and thus, dictates telomere length regulation and homeostasis (Diotti and Loayza 2011). In the protein counting model (Marcand et al. 1997), as telomeres lengthen, the number of binding sites increases, and more shelterin proteins can bind, including POT1 that associates with the 3′ single-stranded telomeric DNA (Lei et al. 2004). Recruitment of POT1 enhances the inhibitory effect at the telomere by acting in *cis* to block the access of telomerase to the telomeric 3′ overhang (Diotti and Loayza 2011). In contrast, as telomeres shorten, the number of bound inhibitory proteins decreases and consequently, drives the telomeres to switch into an open state to allow telomere elongation by telomerase. TPP1 also controls stem cell telomere length homeostasis via a feedback regulation of telomerase by telomere length (Sexton et al. 2014).

3.3 The CST Complex

The CST complex, composed of Cdc13, Stn1 and Ten1, was first identified in yeast as a heterotrimer with structural similarity to the RPA protein complex (Chen and Lingner 2013). The CST complex binds to the 3′ telomeric overhang and contributes to its protection. Studies identified yeast homologs of Stn1 and Ten1 in humans in addition to a third peptide, CTC1, which possesses no conservation with Cdc13 but still trimerizes with the other two components to localize to the telomeric 3′ end. Human CTC1 and STN1 enhance the telomere binding affinity and the activity of DNA polymerase α-primase to facilitate telomere replication and stimulate C-strand fill-in synthesis (Chen and Lingner 2013). Also, the human CST complex competes with the heterodimer POT1-TPP1 for binding to the telomeres. The interaction of CST with telomeres peaks in late S/G2 phase, which corresponds to the timing of post-telomere replication by telomerase and therefore limits telomere over-elongation by terminating telomerase's action at telomeres (Fig. 1a).

4 Telomere Length Maintenance

Telomere homeostasis is a dynamic process and is the end result of two opposing forces: telomere maintenance or lengthening and telomere erosion. The majority of human stem, germ and cancer cells relies on the enzyme telomerase to counteract telomere attrition (Calado and Dumitriu 2013) but in some human cancer and immortalized cell lines, a homologous-recombination-based mechanism is used, termed the alternative lengthening of telomeres (ALT) (O'Sullivan and Almouzni 2014).

4.1 Enzyme-Based Telomere Maintenance: Telomerase

In the mid-1980s, Blackburn and Greider's work using *Tetrahymena themophila* extracts identified an enzyme, telomerase, with terminal transferase activity that adds tandem TTGGGG repeats to synthetic telomeric substrates (Greider and Blackburn 1985). Telomerase is a specialized reverse transcriptase that counteracts telomere shortening during cell replication. Following birth, expression of telomerase is down-regulated and is virtually undetectable in normal differentiated human somatic cells, with the exception of the highly proliferative stem cell compartments or the germ cells. Consequently progressive telomere shortening occurs throughout the lifetime of the organism. Ectopic expression of telomerase allows normal human cells to bypass senescence and extends their replicative lifespan (Bodnar et al. 1998), suggesting that telomerase expression is a limiting factor for cellular proliferation. Additionally, telomerase preferentially extends short telomeres in human cells (Britt-Compton et al. 2009) and a recent study demonstrated that the onset of replicative senescence in *Saccharomyces cerevisiae* is controlled by the length of the shortest telomere (Xu et al. 2013). Importantly, 85 % of human cancers express telomerase to overcome telomere shortening (Shay et al. 2012) through the addition of ~60 nucleotides at each round of extension (Zhao et al. 2011). During the last three decades, significant resources were invested to understand the structure and function of the enzyme in anticipation of developing specific and effective telomerase-based anti-cancer therapies.

4.2 Telomerase Structure

Human telomerase is minimally composed of the catalytic subunit hTERT (human telomerase reverse transcriptase) and its integral RNA component hTR or hTERC (human telomerase RNA; human telomerase RNA component), which serves as a template for the *de novo* synthesis of telomeres (Fig. 1a). *In vivo*, the presence of additional accessory proteins is essential for the proper biogenesis, cellular trafficking and catalytic activity of the telomerase holoenzyme. Telomerase associates with the dyskerin complex, pontin, reptin and TCAB1. DKC1 (dyskerin), GAR1, NHP2 and NOP10 are members of the dyskerin complex and are implicated in holoenzyme assembly and the maintenance of telomerase RNA stability (Mitchell et al. 1999). *In vivo* assembly of the active enzyme also requires the ATPases pontin and reptin (Venteicher et al. 2008). TCAB1 (WDR79), a WD repeat containing protein which binds to dyskerin and the telomerase RNA, is essential for the accumulation of hTR in the Cajal bodies (CB), for the trafficking of the holoenzyme to the telomeres (Venteicher and Artandi 2009) and for licensing of catalytically active telomerase ribonucleoprotein (Vogan and Collins 2015).

The crystal structures of the *Tribolium castaneum* TERT provided tremendous insights into the structural and functional organization of the catalytic subunit,

despite the absence of the N-terminal TEN domain in the beetle TERT (Mitchell et al. 2010). Based on this structure which highly resembles that of HIV-reverse transcriptase (RT), TERT forms a ring that can accommodate up to 8 bases of the DNA-RNA hybrid, whereby the 3′ end of the telomeric substrate is positioned within the active site to allow catalysis (Mitchell et al. 2010). Single particle electron microscopy enabled the 3D visualization of the active full length human telomerase as two molecules of hTERT associated with two molecules of hTR that can function as a dimer to bind to two telomeric DNA substrates (Sauerwald et al. 2013).

Unlike other conventional nucleic acid polymerases, telomerase possesses the unique ability to reiteratively add long stretches of telomeric DNA to the ends of the chromosomes using the short template (one and a half repeats in length for hTR) within its integral RNA component. The addition of multiple telomeric repeats is accomplished through repeated rounds of nucleotide addition, enzyme dissociation, translocation and realignment with the newly synthesized telomeres. This complex property, termed "repeat-addition processivity" (RAP), is tightly regulated and is an important determinant of telomere length homeostasis (D'Souza et al. 2013).

4.3 Telomerase Regulation

Telomerase is under very strict molecular and cellular control to ensure proper telomere length regulation, with the transcriptional regulation of the TERT component being the limiting component of enzyme activity (Cifuentes-Rojas and Shippen 2012). Furthermore, cellular compartmentalization, post-transcriptional and translational modifications, as well as interactions with associated proteins also contribute to the multi-layer control of telomerase function.

The most essential requirement for human telomere maintenance in telomerase-positive cells is the recruitment and localization of both hTERT and hTR to their site of action, the telomeres. Cellular trafficking of telomerase is cell cycle regulated (Tomlinson et al. 2006). TERT and TR are synthesized separately and accumulate in distinct subnuclear compartments in G1 (Tomlinson et al. 2006). hTR, through its interaction with TCAB1, accumulates in the CB for further processing and maturation whereas hTERT is enriched in nucleoplasmic foci termed TERT foci (Tomlinson et al. 2006). In early and mid-S phase, TERT foci are found adjacent to hTR-containing CB and colocalize with telomeres (Tomlinson et al. 2006). The mechanism by which hTERT associates with hTR and is shuttled to telomeres remains to be elucidated. A recent work by Collins's group demonstrated that hTERT is stably associated with hTR throughout the cell cycle but the association between hTR and TCAB1 is disrupted in M-phase cells, thus ensuring localization of telomerase to the telomeres in a cell cycle-dependent manner (Vogan and Collins 2015).

Numerous proteins are engaged in the recruitment of telomerase to the telomeres, including the shelterin protein TPP1, PinX1, and finally, the most recently

identified double-stranded telomeric DNA binding protein HOT1, a positive regulator of telomere length (Kappei et al. 2013). Telomere repeat-containing non-coding RNAs (TERRA) (discussed below) inhibit telomerase action in *cis* by binding to telomeres and interacting with telomerase through complementary base-pairing with the RNA template. Thus, TERRA negatively impacts telomerase function and telomere length (Redon et al. 2010). Another regulator of telomerase, the hPif1 DNA/RNA helicase, interacts with hTERT and preferentially binds to telomeric DNA, negatively influencing telomere length (Cifuentes-Rojas and Shippen 2012). A new study in yeast demonstrated that Pif1 can efficiently remove telomerase from the telomeres in a length dependent manner, thereby regulating telomerase activity and telomere length (Li et al. 2014).

4.4 Homologous Recombination Based: Alternative Lengthening of Telomeres (ALT)

Although the majority of human cancers expresses telomerase as a mean to confer unlimited replicative potential, at least 10 % of human cancers rely on a homologous recombination-based TMM, termed alternative lengthening of telomeres, or ALT (O'Sullivan and Almouzni 2014). The exact mechanisms controlling ALT-mediated telomere maintenance are unknown and therefore, targeting of ALT cancer cells is not yet possible.

4.5 Characteristics of ALT Cells

ALT cells are distinguished by their highly heterogeneous telomere length resulting from HR-based template copying from a sister chromatid or another chromosome. ALT cells display noticeably higher incidences of telomeric-sister-chromatid exchange events compared to telomerase-positive cells (O'Sullivan and Almouzni 2014). ALT cells also exhibit elevated levels of extrachromosomal telomeric repeat DNA including c-circles which are composed of a full-length telomeric C-rich strand hybridized to a partial G-rich strand (Henson et al. 2009). C-circles represent the most specific and quantifiable marker of ALT reported to date, are 750-fold more abundant in ALT-positive cells compared to ALT-negative cells and are detected in blood and tumor tissue from ALT cancer patients (Henson et al. 2009). The mechanism by which c-circles are generated is currently unknown.

ALT cells are also characterized by the presence of a special subtype of promyelocytic leukemia (PML) nuclear body, termed ALT-associated PML body (APB) (O'Sullivan and Almouzni 2014). PML bodies are usually involved in the regulation of normal cellular functions such as DNA replication, transcription, and DNA repair as well as tumor suppression. APBs contain the PML protein, telomeric

DNA, telosome components, and factors implicated in DNA repair, replication and recombination (O'Sullivan and Almouzni 2014). A recent study revealed a role for PML bodies in telomeric chromatin compaction and telomeric repeat clustering, leading to formation of APBs and depletion of TRF2 from the telomeres (Osterwald et al. 2015). This subsequently promotes ATM autophosphorylation in the APBs and activation of the DDR, whereby telomeres are replicated via the ALT pathway (Osterwald et al. 2015).

In normal cells, three variants of the telomeric sequence (TGAGGG, TCAGGG, TTGGGG) are detected at human telomeres and restricted to the telomeric proximal end (Conomos et al. 2012). In ALT cells however, variant sequences are detectable throughout the telomeres and are most likely the consequence of HR-based telomere replication (Conomos et al. 2012). The presence of non-cognate telomeric sequences was proposed to decrease the sequence-specific binding of the telosome and other telomere-associated proteins, and sequence-specifically recruit different proteins such as nuclear receptors, potentially altering the architecture of ALT telomeres to promote recombination (Conomos et al. 2012). For example, in ALT cells, the ratio of TRF2/total amount of telomeric DNA is reduced compared to ALT-negative cells (Conomos et al. 2013). Since TRF2 typically represses HR, telomeres in ALT cells are therefore prone to HR (Conomos et al. 2013). A decrease in the binding of the telosome in ALT cells could also explain the high levels of genomic instability at ALT telomeres (TIFs) but nevertheless, ALT cells are able to progress through the cell cycle, suggesting that the telomere uncapping is only a transient state and that dysfunctional telomeres are a requirement for ALT (Conomos et al. 2013).

4.6 Factors Regulating ALT

Numerous proteins associated with APBs, telomeres and HR are potentially implicated in the ALT pathway but only a few proteins or protein complexes are required for ALT-mediated telomere length maintenance (Bhattacharyya et al. 2010; Cesare and Reddel 2010). These include the MRN and SMC5/6 'structural maintenance of chromosome' recombination complexes, consistent with ALT being HR-dependent, the MUS81 and FEN1 endonucleases, and the FANCD2 and FANCA Fanconi anemia proteins.

No ALT inhibitors have been developed to date primarily because the molecular mechanisms governing the ALT pathway are incompletely characterized. A better understanding of the ALT mechanism will greatly contribute to the development of therapeutic approaches to specifically target ALT. One major challenge in the development of ALT specific inhibitors is that proteins with potential roles in telomeric recombination and the ALT pathway also regulate DNA transactions, including DDR proteins, recombination proteins, helicases and nucleases and are present in normal cells (Cesare and Reddel 2010). Nonetheless, depletion of MUS81, a DNA structure-specific recombination endonuclease induces decreased

telomeric sister chromatid exchanges and proliferation arrest in ALT cells, but not in telomerase-positive cells (Zeng et al. 2009). Interestingly, RNaseHI was recently found to associate specifically with ALT telomeres to regulate TERRA-telomeric hybrids and RPA at telomeres, and altering cellular RNaseH1 levels distinctively perturbed telomere homeostasis in ALT but not telomerase-positive cells (Arora et al. 2014). Indeed, ALT cancer cells were recently shown to display specific hypersensitivity to inhibition of ATR, a protein recruited by RPA and which localizes to and maintains APBs (Flynn et al. 2015). These results suggest that it may be possible to specifically target telomeres and viability of ALT cancer cells.

4.7 ALT in Human Cancer

In human cancer, 20–60 % of certain tumor types with a high unmet medical need, such as osteosarcomas, pancreatic neuroendocrine cancers (PanNETs), glioblastomas, and other tumors of the central nervous system (CNS), maintain telomere length by the ALT pathway (O'Sullivan and Almouzni 2014).

ATRX is part of the transcription/chromatin remodelling complex and functions to maintain a closed heterochromatin structure at pericentric heterochromatin, telomeres and several transcription factor binding sites by depositing the replication-independent histone 3 variant H3.3 (Schwartzentruber et al. 2012). Recently, mutation and loss of ATRX were found to correlate with features of ALT in pediatric glioblastomas, tumors of the central nervous system and PanNETs, consistent with ATRX being a repressor of the ALT pathway (Heaphy et al. 2011; Schwartzentruber et al. 2012; O'Sullivan and Almouzni 2014; Clynes et al. 2015). Loss of ATRX expression in PanNET patients is associated with increased genomic instability and shorter time of survival (Marinoni et al. 2014). The ALT phenotype is also associated with p53 deficiency (95 % of ALT cell lines, and ~78 % of ALT tumours have non-functional p53) (Cesare and Reddel 2008; Gocha et al. 2012). p53 mutations at codons 248 or 273 have been reported as hotspots in human tumours, including ALT tumours (Gocha et al. 2012). Many ALT tumors with these p53 mutations also harbor ATRX mutations (Gocha et al. 2012).

The mechanisms by which ATRX regulates telomeric recombination and the ALT pathway are unknown, but mutations or loss of ATRX could promote a more open chromatin conformation and favor telomeric recombination. Recent studies show that chromatin compaction is reduced at ALT telomeres and that ALT phenotypes are induced by depletion of the histone chaperone ASF1, supporting the hypothesis that ALT is a consequence of histone management dysfunction (O'Sullivan et al. 2014; Episkopou et al. 2014).

5 Telomeres and Epigenetics

Unlike the rest of the genome, mammalian telomeres do not contain genes and the subtelomeres are gene poor. The telomeric chromatin structure has the ability to reversibly silence subtelomeric genes, a repressive effect termed "telomere position effect" (TPE) (Gottschling et al. 1990). TPE was first discovered in *Drosophila melanogaster* and later extensively studied in yeast, and was shown to be dependent on the high binding affinity of the chromatin compaction protein heterochromatin protein 1 (HP1) to trimethylated H3K9me3 (Schoeftner and Blasco 2010). This association triggers a switch from the open euchromatin conformation into the closed heterochromatin state. In mammals, the existence of TPE remained controversial for many years. Studies using insertion of transgenes near human telomeres demonstrated that the NAD-dependent protein deacetylase SIRT6 acts as a key player in the maintenance of TPE by keeping the telomeres in a repressive hypoacetylated state (Tennen et al. 2011). Furthermore, telomere length determines the extent of TPE, with long telomeres exhibiting a stronger silencing effect due to a higher density of repressive histone marks and enrichment of chromatin compaction proteins (Baur et al. 2001).

The assumption that telomeres merely serve as a protective cap and are consistently in a heterochromatin state persisted for many years. It was only with the discovery of TERRA, transcribed by RNA polymerase II (RNAPII), that this notion was disproved (Azzalin et al. 2007; Schoeftner and Blasco 2008). TERRA is transcribed from subtelomeric promoters toward the telomeric regions (Fig. 1b). They are comprised of RNA sequences ranging in size from 100 bases to approximately 9 kb complementary to both subtelomeric and telomeric sequences (Schoeftner and Blasco 2010). In 2012, Decottignies's group assessed the physiological existence of TPE at native human chromosomal ends for the first time by monitoring the transcription of the subtelomeric regions into TERRA (Arnoult et al. 2012). Mammalian subtelomeric regions can adopt a more open conformation characterized by enrichment of H2BK5me1 and H3K4me3. Consistent with earlier yeast studies, Decottignies's group showed that long human telomeres negatively impact the level of TERRA transcription, with a corresponding enrichment of H3K9me3 and HP1α at the telomeres. Longer telomeres also result in a concomitant increase in TERRA length despite the lower number of TERRA molecules (Arnoult et al. 2012). Due to the complementary nature of the TERRA sequence $(UUAGGG)_N$ to the telomeric C-strand $(AATCCC)_N$ sequences, these RNAs can transiently bind to subsets of telomeres and, in turn, recruit heterochromatin-associated proteins such as HP1α, resulting in a negative feedback that shifts the telomeres from the open euchromatin conformation back to the heterochromatin state (Arnoult et al. 2012). Furthermore, a new study in telomerase-negative yeast cells (Type II survivors in yeast equivalent to ALT in mammalian cells) showed that the association of TERRA with telomeric DNA promotes telomeric recombination and senescence bypass (Yu et al. 2014).

TERRAs are strictly nuclear and are cell cycle regulated, with levels peaking in G1 and early S phase (Maicher et al. 2014). Furthermore, the shelterin protein TRF1 can interact with RNAPII and changes in TRF1 or RNAPII expression impact TERRA levels (Schoeftner and Blasco 2008). Another mechanism regulating TERRA was found by Azzalin's group, when they discovered that factors of the nonsense-mediated RNA decay pathway can physically interact with the telomeric chromatin and trigger TERRA degradation or displacement from the telomeres (Azzalin et al. 2007). A recent study demonstrated that the chromatin organizing factor CTCF and the multiprotein complex cohesin are fundamental components of the subtelomeres in human cells (Deng et al. 2012). CTCF and cohesin recruit RNAPII to telomeres, promote TERRA expression, while their depletion reduces TERRA expression and induces TIF formation, suggesting a potential role for TERRA in telomere end-protection (Deng et al. 2012).

Nucleosomes at mammalian telomeric and subtelomeric regions are spaced distinctly compared to their organization in the rest of the genome (Schoeftner and Blasco 2010). More recently, a study showed that *in vitro*, the shelterin component TRF1 can change the nucleosomal spacing, thereby allowing chromatin remodelling at telomeres (Galati et al. 2015). Increasing evidence supports a role for epigenetics in telomere length regulation. Several studies have shown that defective epigenetic control leads to compromised telomere length maintenance and integrity. SUV39H1/H2 are suppressors of variegation histone methyltransferases responsible for the trimethylation of H3K9 (Peters et al. 2001). Deficiency in SUV39H1/H2 results in the decrease of H3K9me3, altered telomeric heterochromatin state and aberrantly elongated telomeres (Garcia-Cao et al. 2004). Mammalian telomeres are not methylated due to the absence of CpG sequences, which are the substrates of DNA methyltransferases (DNMTs) (Gonzalo et al. 2006). Instead, association of DNMT1, 3a and 3b has been reported at the subtelomeric regions enriched in CpG sequences (Gonzalo et al. 2006). Subtelomeric DNA methylation enforces TPE and restricts the access of HR proteins to telomeres, thus suppressing ALT-mediated TMM (Gonzalo et al. 2006).

Emerging evidence suggests a link between the epigenetic status of human telomeric and subtelomeric chromatin and disease states. Such a connection has already been demonstrated by the altered telomeric methylation status in Alzheimer's disease (AD) (Guan et al. 2013). Also, TERRA downregulation has been observed in advanced stages of human larynx, colon and lymph node cancer, compared to normal tissue (Schoeftner and Blasco 2008). An interesting cohort study on astrocytic tumors revealed that promoter methylation-mediated epigenetic silencing leads to TERRA downregulation (Sampl et al. 2012). Elevated TERRA level inversely correlates with tumor grade and is associated with better patient survival (Sampl et al. 2012). Increasing efforts are currently directed towards studies that investigate the use of telomere epigenetic marks as diagnostic or prognostic tools in human disease.

6 Nuclear Organization and Telomeres

Initial studies in yeast showed that telomeres are distributed in a non-random fashion and are localized to specific subnuclear domains in a cell cycle-dependent manner. This specific telomere distribution implicating Sir3, Sir4 and Rap1 was first described in yeast, through the observation that telomeric clustering forms at the nuclear rim (Gotta et al. 1996). In human interphase cells, telomeres are also attached to the nuclear matrix (de Lange 1992). Based on live cell imaging data, telomeres are highly dynamic structures with a mobility that increases with telomere attrition and uncapping (Dimitrova et al. 2008; Wang et al. 2008).

Type A lamin, the main component of the lamina layer, is a member of the class V intermediate filaments generally involved in the maintenance of nuclear structural integrity and the regulation of transcription. Depletion of lamin A, which interacts with telomeres, causes telomere redistribution while mutation in this filament protein leads to lamina layer distortion and the formation of telomere aggregates (Novo and Londono-Vallejo 2013). Members of the shelterin complex are also implicated in telomere positioning, serving as intermediates that bridge the telomeres to the nuclear matrix. Time-lapse confocal microscopy showed that similarly to yeast, human Rap1 is required for the association of telomeres with the nuclear architecture (Crabbe et al. 2012). In *Caenorhabditis elegans*, POT1 is essential for nuclear peripheral tethering of the telomeres (Ferreira et al. 2013). Furthermore, TIN2L, an isoform of TIN2, binds strongly to the nuclear matrix and is required for telomere tethering (Kaminker et al. 2009).

To decipher the 3D telomere positioning in the nucleus, Mai's group developed a technique using 3D nuclear imaging in combination with TeloView (Klonisch et al. 2010). This technique led to the discovery that the nuclear positioning of telomeres differs between normal and cancer cells. In normal cells, telomeres are confined within an ellipsoid area and are positioned at the periphery of a central telomeric disk formed in G2 (Klonisch et al. 2010). In cancer cells however, telomeric fusions generate telomeric aggregates (TAs), hallmarks of tumor cells, in the interphase nucleus (Chuang et al. 2004). TAs promote genomic instability and nuclear architecture remodelling distinct from that in normal cells (Novo and Londono-Vallejo 2013). In recent years, Mai's group conducted studies profiling the telomere architecture of multiple myeloma, myelodysplastic syndrome and acute myeloid leukemia patients by measuring the telomere number, TAs, telomere signal intensity, nuclear volume as well as telomeric distribution (Gadji et al. 2012; Klewes et al. 2013). This 3D imaging approach allows the assessment of these various parameters with a potential application in disease stratification and the development of patient-specific treatments (Gadji et al. 2012; Klewes et al. 2013).

7 Telomere and Disease

The existence of a link between telomere length and longevity in humans remains an open question. However, convincing evidence suggests that deprotected telomeres contribute to the development of cancer and various premature aging diseases. This was first demonstrated using telomerase knockout mice characterized by critically short, uncapped telomeres. They also displayed increased incidence of tumorigenesis with several characteristics of premature aging (Artandi et al. 2000; Blasco et al. 1997; Sahin and Depinho 2010).

7.1 Telomeres and Aging

Telomere shortening results from the natural replication dependent-aging process. However, environmental factors, including pollution, smoking, diet, infection, inflammation, and DNA damaging agents such as UV light, enhance the rate of telomere loss (Calado and Dumitriu 2013). The heritability of telomere length has also been demonstrated in human leukocytes, where the paternal telomere length determines the ones in the offsprings (Eisenberg et al. 2012). Heritability of telomere length is particularly important in individuals with a defective telomere maintenance mechanism in their stem and germ cell compartments. Persons with degenerative premature aging disorders caused by defective telomere maintenance, termed "telomeropathies", exhibit disease anticipation; that is, the onset of the disease occurs at a younger age and symptoms are more severe with each subsequent generation (Holohan et al. 2014). These diseases arise from defects in the shelterin complex, telomerase or its accessory proteins such as dyskerin, in addition to other factors required for telomere replication, for example, the RTEL1 helicase.

The first disease identified to be directly associated with defective telomere biology is dyskeratosis congenita (DC) (Mitchell et al. 1999). DC patients are generally diagnosed by a triad of nail dystrophy, oral leukoplakia and reticulated hyperpigmentation of the skin (Armanios and Blackburn 2012). The lengths of germ line telomeres are very short in these individuals due to the inability to sustain telomere maintenance and, as such, they often suffer from aplastic anemia, bone marrow failure, pulmonary fibrosis and/or liver cirrhosis. Furthermore, high levels of genomic instability resulting from short telomeres predispose patients to cancer development (acute myeloid leukemia and myelodysplastic syndrome) (Armanios and Blackburn 2012). The average lifespan of DC patient is 30 years as a result of bone marrow failure (Holohan et al. 2014). DC is an extremely rare inherited genetic disorder that can be classified based on the pattern of inheritance: autosomal dominant, recessive or X-linked. To date, germline mutations causing DC and Hoyeraal-Hreidarsson syndrome (HHS), a more severe form of DC, have been identified in genes involved in telomere maintenance including hTERT, hTERC, DKC1, NOP10, NHP2, TCAB1, TINF2 (codes for TIN2), CTC1, RTEL1 and TPP1

(Holohan et al. 2014; Kocak et al. 2014; Guo et al. 2014). The manifestation of DC and HHS is highly heterogeneous due to differences in the level of penetrance (Holohan et al. 2014). Detrimental compound heterozygous RTEL1 mutations have also been identified in hereditary HHS (Walne et al. 2013; Holohan et al. 2014). RTEL1 was extensively studied in mice and evidence support its role in the resolution of G-quadruplex (G4)-DNA (see below) and D-loop HR intermediates during telomere replication (Vannier et al. 2012). More recently, an HHS mutation within the TRF2-interacting site of RTEL1, the C4C4 domain, was shown to disrupt its interaction with TRF2, thus preventing the recruitment of RTEL1 to the telomeres for T-loop unwinding during S-phase (Sarek et al. 2015). Loss of RTEL1 in human cells also results in hyper-recombination, suggesting a role for RTEL1 in the suppression of HR at telomeres (Holohan et al. 2014).

Other severe premature aging diseases resulting from aberrant telomere maintenance includes the Revesz syndrome (*TINF2*), and the Coats Plus syndrome (*CTC1*), which also presents features that overlap with DC. Idiophathic pulmonary fibrosis (IPF) is typically sporadic and is associated with mutations in hTERT and hTR, possibly as a result of haploinsufficiency (Armanios and Blackburn 2012). IPF is a progressive disease that usually develops around the age of 50 and patients usually do not survive more than 3 years post-diagnosis (Armanios and Blackburn 2012).

Tremendous efforts continue to be invested into deciphering the role of the different components to uncover new targets for regenerative medicine. The only option available at this time is tissue or organ transplant but the wait for a compatible donor and potential relapses limit the chances of survival. Research focusing on the development of small molecules that can specifically activate telomerase without triggering tumorigenesis is currently ongoing. Using a heterozygous TERT murine model of IPF, a novel molecule activator of telomerase, GRN510, showed encouraging results in suppressing lung damages (Le Saux et al. 2013). Telomerase gene therapy based on telomerase overexpression, tested in both adult and old mice, has been shown to delay aging and increase longevity without increasing cancer incidence (Bernardes de Jesus et al. 2012).

7.2 Telomeres and Cancer

The core components of the telomere maintenance machinery are potentially specific targets for the development of anti-cancer therapies. Telomerase is active in ~85 % of tumors, but only weakly active in primary cells, thus it is an attractive target for cancer cell-specific therapy. The telomerase inhibitor GRN163L (Imetelstat) is an hTR antagonist that has been tested in human cancer cell lines including pancreatic, esophageal, and leukemic cancer cells and in xenograft models. Results initially showed promising inhibitory effects on cell growth and increased sensitization of cells to DNA damaging agents (Harley 2008). The efficacy of GRN163L was investigated in various clinical trials for different types

of cancer, but to date it has had limited efficacy in the clinic, with beneficial effects noted in hematological cancers, in tumors with short telomeres and in combination therapy (Buseman et al. 2012; Williams 2013).

One drawback of anti-telomerase based therapy is the lag period between the time of treatment and the observable proliferative arrest resulting from telomere erosion, which requires several rounds of cell division and thus, would likely only be effective in cancer cells with short telomeres. Furthermore, 15 % of cancer cells employ the ALT pathway for telomere maintenance and are insensitive to telomerase inhibition (Shay et al. 2012). An ingenious method to overcome such limitations involves directly triggering telomere uncapping, the activation of the DNA damage checkpoint and ultimately cellular growth arrest (McEachern et al. 2000; Guiducci et al. 2001). In recent years, G4-DNA-stabilizing ligands and mutant hTR gene therapy-based approaches demonstrated their efficacy in triggering telomere uncapping (Harley 2008; Cerone et al. 2006; Brault and Autexier 2011; Tauchi et al. 2006; Goldkorn and Blackburn 2006). Recently, a small molecule telomerase substrate was reported to induce telomere uncapping in a telomerase-dependent fashion in cell and xenograft models (Mender et al. 2015).

Formation of the telomere cap is dependent on the telomere sequence-specific binding of the TRF1, TRF2 and POT1 shelterin proteins. Alteration in the telomeric sequences by mutant hTR can trigger telomere uncapping resulting from impaired shelterin-telomere associations. In eukaryotes, the single-stranded guanine-rich DNA sequences at the telomeres favor the formation of higher order intramolecular structures known as G4 (Harley 2008) (Fig. 1a). The observation that certain G4 ligands and mutant hTR expression lead to specific anti-proliferative effects in cancer cells implies that perhaps in these cells, the telomere cap is distinct from that in normal cells (Fakhoury et al. 2007; Riou 2004; Mahalingam et al. 2011). Importantly, though G4 structures are typically viewed as preventing the access of telomerase to the telomeres, a recent study demonstrated that *in vivo*, telomerase can localize to parallel G4 formed at telomeres and subsequently allow DNA extension (Moye et al. 2015). Their data suggest that specific G4 conformations and location have important biological implications.

The altered expression profile of telosome components has been reported in some human tumors (Oh et al. 2005; Poncet et al. 2008). TRF2 expression was shown to contribute to multidrug resistance in gastric cancer (Ning et al. 2006) while other studies found an increased expression of POT1 and TPP1 in multiple myeloma (Panero et al. 2014; Ferrandon et al. 2013). POT1 was also recently identified as the first shelterin protein mutated in human cancer (Ramsay et al. 2013). Altogether, these results support the hypothesis that the telomere cap in cancer cells is distinct from that in normal cells and highlight the potential of telomere uncapping strategies.

Nonetheless, the development of telomerase and telomere-based anti-cancer and regenerative therapies remains challenging due to the potential risk to induce tumorigenesis upon telomerase activation or the development of resistance to anti-cancer therapies. Both ALT and telomerase-based telomere maintenance can co-exist in human immortalized and cancer cells (Queisser et al. 2013).

Consequently, telomerase inhibition in these cells may elicit a switch from telomerase to ALT-dependent telomere maintenance, allowing the cells to resist anti-telomerase-based strategies. Indeed, telomerase inhibition or deletion of telomerase components in human cells or in mice leads to telomeric recombination, ALT or ALT-like activation and cell survival (Bechter et al. 2004; Chang et al. 2003; Hande et al. 1999; Morrish and Greider 2009; Niida et al. 2000; Laud et al. 2005; Hu et al. 2012; Queisser et al. 2013). Telomere uncapping strategies in yeast and mice also cause an increase in telomeric recombination and cell survival (Bechard et al. 2009; Grandin et al. 2001; Iyer et al. 2005; Celli et al. 2006; He et al. 2006; Wu et al. 2006). Recently, ALT activity was detected in normal mammalian somatic cells (Neumann et al. 2013) and telomere uncapping in telomerase-positive human cells activates an ALT-like phenotype characterized by telomeric HR (Brault and Autexier 2011; Conomos et al. 2012). These results expose a possible resistance mechanism against telomere uncapping-mediated anti-cancer therapies.

There is growing interest in combining anti-telomerase strategies with conventional therapies to maximize the efficacy and specificity of these methods. A study on Barrett's adenocarcinoma (BAC) demonstrated that targeting telomerase using GRN163L in combination with an HR inhibitor such as nilotinib or an shRNA against RAD51 decreases telomere length and increases BAC cell apoptosis (Lu et al. 2014). A phase II trial using the telomerase vaccine GV1001 (15 amino acid peptide) post-chemoradiotherapy treatment showed an elevated immune response rate of 80 % with low level of toxicity in non-small cell lung cancer (NSCLC) patients (Brunsvig et al. 2011). The increased survival observed in responders suggests that GV1001, in combination with chemoradiotherapy, may have a beneficial effect in NSCLC patients (Brunsvig et al. 2011).

8 Concluding Remarks

Despite the tremendous amount of knowledge we currently possess about telomere biology and regulation, there are still many pending questions regarding the underlying mechanisms of telomere homeostasis. Furthermore, an increasing number of studies aim to develop telomere-based strategies to target cancer cells more specifically and efficiently. Importantly, regenerative therapies are also under investigation, such as the generation of human induced pluripotent stem cells and telomere length resetting, in the hope to find a cure for individuals suffering from premature aging syndromes.

Acknowledgements The authors wish to thank Johanna Mancini and Deanna MacNeil for comments on the manuscript. Work in the lab of C. Autexier is supported by the Canadian Institutes of Health, the Canadian Cancer Society, the Cancer Research Society and the Natural Sciences and Engineering Research Council of Canada. T.W. Chu is a recipient of Le Fonds de recherche du Québec – Santé Doctoral Award.

References

Armanios M, Blackburn EH (2012) The telomere syndromes. Nat Rev Genet 13(10):693–704. doi:10.1038/nrg3246

Arnoult N, Van Beneden A, Decottignies A (2012) Telomere length regulates TERRA levels through increased trimethylation of telomeric H3K9 and HP1alpha. Nat Struct Mol Biol 19 (9):948–956. doi:10.1038/nsmb.2364

Arora R, Lee Y, Wischnewski H, Brun CM, Schwarz T, Azzalin CM (2014) RNaseH1 regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells. Nat Commun 5:5220. doi:10.1038/ncomms6220

Artandi SE, Chang S, Lee SL, Alson S, Gottlieb GJ, Chin L, DePinho RA (2000) Telomere dysfunction promotes non-reciprocal translocations and epithelial cancers in mice. Nature 406 (6796):641–645. doi:10.1038/35020592

Azzalin CM, Reichenbach P, Khoriauli L, Giulotto E, Lingner J (2007) Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. Science 318(5851):798–801. doi:10.1126/science.1147182

Baur JA, Zou Y, Shay JW, Wright WE (2001) Telomere position effect in human cells. Science 292(5524):2075–2077. doi:10.1126/science.1062329

Bechard LH, Butuner BD, Peterson GJ, McRae W, Topcu Z, McEachern MJ (2009) Mutant telomeric repeats in yeast can disrupt the negative regulation of recombination-mediated telomere maintenance and create an alternative lengthening of telomeres-like phenotype. Mol Cell Biol 29(3):626–639. doi:10.1128/MCB.00423-08, MCB.00423-08 [pii]

Bechter OE, Zou Y, Walker W, Wright WE, Shay JW (2004) Telomeric recombination in mismatch repair deficient human colon cancer cells after telomerase inhibition. Cancer Res 64:3444–3451

Bernardes de Jesus B, Vera E, Schneeberger K, Tejera AM, Ayuso E, Bosch F, Blasco MA (2012) Telomerase gene therapy in adult and old mice delays aging and increases longevity without increasing cancer. EMBO Mol Med 4(8):691–704. doi:10.1002/emmm.201200245

Bhattacharyya S, Sandy A, Groden J (2010) Unwinding protein complexes in ALTernative telomere maintenance. J Cell Biochem 109(1):7–15. doi:10.1002/jcb.22388

Blasco MA, Lee HW, Hande MP, Samper E, Lansdorp PM, DePinho RA, Greider CW (1997) Telomere shortening and tumor formation by mouse cells lacking telomerase RNA. Cell 91 (1):25–34

Bodnar AG, Ouellette M, Frolkis M, Holt SE, Chiu CP, Morin GB, Harley CB, Shay JW, Lichtsteiner S, Wright WE (1998) Extension of life-span by introduction of telomerase into normal human cells. Science 279(5349):349–352

Brault ME, Autexier C (2011) Telomeric recombination induced by dysfunctional telomeres. Mol Biol Cell 22:179–188

Britt-Compton B, Capper R, Rowson J, Baird DM (2009) Short telomeres are preferentially elongated by telomerase in human cells. FEBS Lett 583(18):3076–3080. doi:10.1016/j.febslet.2009.08.029

Brunsvig PF, Kyte JA, Kersten C, Sundstrom S, Moller M, Nyakas M, Hansen GL, Gaudernack G, Aamdal S (2011) Telomerase peptide vaccination in NSCLC: a phase II trial in stage III patients vaccinated after chemoradiotherapy and an 8-year update on a phase I/II trial. Clin Cancer Res 17(21):6847–6857. doi:10.1158/1078-0432.CCR-11-1385

Buseman CM, Wright WE, Shay JW (2012) Is telomerase a viable target in cancer? Mutat Res 730 (1–2):90–97. doi:10.1016/j.mrfmmm.2011.07.006, S0027-5107(11)00182-5 [pii]

Calado RT, Dumitriu B (2013) Telomere dynamics in mice and humans. Semin Hematol 50 (2):165–174. doi:10.1053/j.seminhematol.2013.03.030

Celli GB, Denchi EL, de Lange T (2006) Ku70 stimulates fusion of dysfunctional telomeres yet protects chromosome ends from homologous recombination. Nat Cell Biol 8(8):885–890. doi:10.1038/ncb1444, doi:ncb1444 [pii]

Cerone MA, Londono-Vallejo JA, Autexier C (2006) Mutated telomeres sensitize tumor cells to anticancer drugs independently of telomere shortening and mechanisms of telomere maintenance. Oncogene 25:7411–7420

Cesare AJ, Reddel RR (2008) Telomere uncapping and alternative lengthening of telomeres. Mech Ageing Dev 129(1–2):99–108. doi:10.1016/j.mad.2007.11.006, S0047-6374(07)00182-0 [pii]

Cesare AJ, Reddel RR (2010) Alternative lengthening of telomeres: models, mechanisms and implications. Nat Rev Genet 11(5):319–330. doi:10.1038/nrg2763, nrg2763 [pii]

Chang S, Khoo CM, Naylor ML, Maser RS, DePinho RA (2003) Telomere-based crisis: functional differences between telomerase activation and ALT in tumor progression. Genes Dev 17 (1):88–100. doi:10.1101/gad.1029903

Chen LY, Lingner J (2013) CST for the grand finale of telomere replication. Nucleus 4 (4):277–282. doi:10.4161/nucl.25701, 25701 [pii]

Cheung DH, Kung HF, Huang JJ, Shaw PC (2012) PinX1 is involved in telomerase recruitment and regulates telomerase function by mediating its localization. FEBS Lett 586 (19):3166–3171. doi:10.1016/j.febslet.2012.06.028

Chuang TC, Moshir S, Garini Y, Chuang AY, Young IT, Vermolen B, van den Doel R, Mougey V, Perrin M, Braun M, Kerr PD, Fest T, Boukamp P, Mai S (2004) The three-dimensional organization of telomeres in the nucleus of mammalian cells. BMC Biol 2:12. doi:10.1186/1741-7007-2-12

Cifuentes-Rojas C, Shippen DE (2012) Telomerase regulation. Mutat Res 730(1–2):20–27. doi:10.1016/j.mrfmmm.2011.10.003

Clynes D, Jelinska C, Xella B, Ayyub H, Scott C, Mitson M, Taylor S, Higgs DR, Gibbons RJ (2015) Suppression of the alternative lengthening of telomere pathway by the chromatin remodelling factor ATRX. Nat Commun 6:7538. doi:10.1038/ncomms8538

Conomos D, Stutz MD, Hills M, Neumann AA, Bryan TM, Reddel RR, Pickett HA (2012) Variant repeats are interspersed throughout the telomeres and recruit nuclear receptors in ALT cells. J Cell Biol 199(6):893–906. doi:10.1083/jcb.201207189, jcb.201207189 [pii]

Conomos D, Pickett HA, Reddel RR (2013) Alternative lengthening of telomeres: remodeling the telomere architecture. Front Oncol 3:27. doi:10.3389/fonc.2013.00027

Crabbe L, Cesare AJ, Kasuboski JM, Fitzpatrick JA, Karlseder J (2012) Human telomeres are tethered to the nuclear envelope during postmitotic nuclear assembly. Cell Rep 2 (6):1521–1529. doi:10.1016/j.celrep.2012.11.019

Cristofari G, Lingner J (2006) Telomere length homeostasis requires that telomerase levels are limiting. EMBO J 25:565–574

D'Souza Y, Lauzon C, Chu TW, Autexier C (2013) Regulation of telomere length and homeostasis by telomerase enzyme processivity. J Cell Sci 126:676–687

de Lange T (1992) Human telomeres are attached to the nuclear matrix. EMBO J 11(2):717–724

de Lange T (2010) How shelterin solves the telomere end-protection problem. Cold Spring Harb Symp Quant Biol 75:167–177. doi:10.1101/sqb.2010.75.017

Deng Z, Wang Z, Stong N, Plasschaert R, Moczan A, Chen HS, Hu S, Wikramasinghe P, Davuluri RV, Bartolomei MS, Riethman H, Lieberman PM (2012) A role for CTCF and cohesin in subtelomere chromatin organization, TERRA transcription, and telomere end protection. EMBO J 31(21):4165–4178. doi:10.1038/emboj.2012.266

Dimitrova N, Chen YC, Spector DL, de Lange T (2008) 53BP1 promotes non-homologous end joining of telomeres by increasing chromatin mobility. Nature 456(7221):524–528. doi:10.1038/nature07433

Diotti R, Loayza D (2011) Shelterin complex and associated factors at human telomeres. Nucleus 2(2):119–135. doi:10.4161/nucl.2.2.15135, 1949-1034-2-2-8 [pii]

Donigian JR, de Lange T (2007) The role of the poly(ADP-ribose) polymerase tankyrase1 in telomere length control by the TRF1 component of the shelterin complex. J Biol Chem 282 (31):22662–22667. doi:10.1074/jbc.M702620200

Eisenberg DT, Hayes MG, Kuzawa CW (2012) Delayed paternal age of reproduction in humans is associated with longer telomeres across two generations of descendants. Proc Natl Acad Sci U S A 109(26):10251–10256. doi:10.1073/pnas.1202092109

Episkopou H, Draskovic I, Van Beneden A, Tilman G, Mattiussi M, Gobin M, Arnoult N, Londono-Vallejo A, Decottignies A (2014) Alternative Lengthening of Telomeres is characterized by reduced compaction of telomeric chromatin. Nucleic Acids Res 42(7):4391–4405. doi:10.1093/nar/gku114

Fairall L, Chapman L, Moss H, de Lange T, Rhodes D (2001) Structure of the TRFH dimerization domain of the human telomeric proteins TRF1 and TRF2. Mol Cell 8(2):351–361

Fakhoury J, Nimmo GAM, Autexier C (2007) Harnessing telomerase in cancer therapeutics. Anticancer Agents Med Chem 7:475–483

Ferrandon S, Saultier P, Carras J, Battiston-Montagne P, Alphonse G, Beuve M, Malleval C, Honnorat J, Slatter T, Hung N, Royds J, Rodriguez-Lafrasse C, Poncet D (2013) Telomere profiling: toward glioblastoma personalized medicine. Mol Neurobiol 47(1):64–76. doi:10.1007/s12035-012-8363-9

Ferreira HC, Towbin BD, Jegou T, Gasser SM (2013) The shelterin protein POT-1 anchors Caenorhabditis elegans telomeres through SUN-1 at the nuclear periphery. J Cell Biol 203 (5):727–735. doi:10.1083/jcb.201307181

Flynn RL, Cox KE, Jeitany M, Wakimoto H, Bryll AR, Ganem NJ, Bersani F, Pineda JR, Suva ML, Benes CH, Haber DA, Boussin FD, Zou L (2015) Alternative lengthening of telomeres renders cancer cells hypersensitive to ATR inhibitors. Science 347(6219):273–277. doi:10.1126/science.1257216

Gadji M, Adebayo Awe J, Rodrigues P, Kumar R, Houston DS, Klewes L, Dieye TN, Rego EM, Passetto RF, de Oliveira FM, Mai S (2012) Profiling three-dimensional nuclear telomeric architecture of myelodysplastic syndromes and acute myeloid leukemia defines patient subgroups. Clin Cancer Res 18(12):3293–3304. doi:10.1158/1078-0432.CCR-12-0087

Galati A, Micheli E, Alicata C, Ingegnere T, Cicconi A, Pusch MC, Giraud-Panis MJ, Gilson E, Cacchione S (2015) TRF1 and TRF2 binding to telomeres is modulated by nucleosomal organization. Nucleic Acids Res 43(12):5824–5837. doi:10.1093/nar/gkv507

Garcia-Cao M, O'Sullivan R, Peters AH, Jenuwein T, Blasco MA (2004) Epigenetic regulation of telomere length in mammalian cells by the Suv39h1 and Suv39h2 histone methyltransferases. Nat Genet 36(1):94–99. doi:10.1038/ng1278

Gocha AR, Harris J, Groden J (2012) Alternative mechanisms of telomere lengthening: permissive mutations, DNA repair proteins and tumorigenic progression. Mutat Res 743–744:142–150. doi:10.1016/j.mrfmmm.2012.11.006, S0027-5107(12)00222-9 [pii]

Goldkorn A, Blackburn EH (2006) Assembly of mutant-template telomerase RNA into catalytically active telomerase ribonucleoprotein that can act on telomeres is required for apoptosis and cell cycle arrest in human cancer cells. Cancer Res 66:5763–5771

Gonzalo S, Jaco I, Fraga MF, Chen T, Li E, Esteller M, Blasco MA (2006) DNA methyltransferases control telomere length and telomere recombination in mammalian cells. Nat Cell Biol 8(4):416–424. doi:10.1038/ncb1386

Gotta M, Laroche T, Formenton A, Maillet L, Scherthan H, Gasser SM (1996) The clustering of telomeres and colocalization with Rap1, Sir3, and Sir4 proteins in wild-type Saccharomyces cerevisiae. J Cell Biol 134(6):1349–1363

Gottschling DE, Aparicio OM, Billington BL, Zakian VA (1990) Position effect at S. cerevisiae telomeres: reversible repression of Pol II transcription. Cell 63(4):751–762

Grandin N, Damon C, Charbonneau M (2001) Cdc13 prevents telomere uncapping and Rad50-dependent homologous recombination. EMBO J 20(21):6127–6139. doi:10.1093/emboj/20.21.6127

Greider CW, Blackburn EH (1985) Identification of a specific telomere terminal transferase activity in *Tetrahymena* extracts. Cell 43:405–413

Griffith JD, Comeau L, Rosenfield S, Stansel RM, Bianchi A, Moss H, de Lange T (1999) Mammalian telomeres end in a large duplex loop. Cell 97(4):503–514

Guan JZ, Guan WP, Maeda T, Makino N (2013) Analysis of telomere length and subtelomeric methylation of circulating leukocytes in women with Alzheimer's disease. Aging Clin Exp Res 25(1):17–23. doi:10.1007/s40520-013-0006-0

Guiducci C, Cerone MA, Bacchetti S (2001) Expression of mutant telomerase in immortal telomerase-negative human cells results in cell cycle deregulation, nuclear and chromosomal abnormalities and rapid loss of viability. Oncogene 20:714–725

Guo Y, Kartawinata M, Li J, Pickett HA, Teo J, Kilo T, Barbaro PM, Keating B, Chen Y, Tian L, Al-Odaib A, Reddel RR, Christodoulou J, Xu X, Hakonarson H, Bryan TM (2014) Inherited bone marrow failure associated with germline mutation of ACD, the gene encoding telomere protein TPP1. Blood. doi:10.1182/blood-2014-08-596445

Hande MP, Samper E, Lansdorp P, Blasco MA (1999) Telomere length dynamics and chromosomal instability in cells derived from telomerase null mice. J Cell Biol 144:589–601

Harley CB (2008) Telomerase and cancer therapeutics. Nat Rev Cancer 8(3):167–179. doi:10.1038/nrc2275, nrc2275 [pii]

He H, Multani AS, Cosme-Blanco W, Tahara H, Ma J, Pathak S, Deng Y, Chang S (2006) POT1b protects telomeres from end-to-end chromosomal fusions and aberrant homologous recombination. EMBO J 25(21):5180–5190. doi:10.1038/sj.emboj.7601294

Heaphy CM, de Wilde RF, Jiao Y, Klein AP, Edil BH, Shi C, Bettegowda C, Rodriguez FJ, Eberhart CG, Hebbar S, Offerhaus GJ, McLendon R, Rasheed BA, He Y, Yan H, Bigner DD, Oba-Shinjo SM, Marie SK, Riggins GJ, Kinzler KW, Vogelstein B, Hruban RH, Maitra A, Papadopoulos N, Meeker AK (2011) Altered telomeres in tumors with ATRX and DAXX mutations. Science 333(6041):425. doi:10.1126/science.1207313, science.1207313 [pii]

Henson JD, Cao Y, Huschtscha LI, Chang AC, Au AY, Pickett HA, Reddel RR (2009) DNA C-circles are specific and quantifiable markers of alternative-lengthening-of-telomeres activity. Nat Biotechnol 27(12):1181–1185. doi:10.1038/nbt.1587

Hockemeyer D, Sfeir AJ, Shay JW, Wright WE, de Lange T (2005) POT1 protects telomeres from a transient DNA damage response and determines how human chromosome end. EMBO J 24:2667–2678

Holohan B, Wright WE, Shay JW (2014) Cell biology of disease. Telomeropathies: an emerging spectrum disorder. J Cell Biol 205(3):289–299. doi:10.1083/jcb.201401012

Horikawa I, Fujita K, Harris CC (2011) p53 governs telomere regulation feedback too, via TRF2. Aging 3(1):26–32

Hu J, Hwang SS, Liesa M, Gan B, Sahin E, Jaskelioff M, Ding Z, Ying H, Boutin AT, Zhang H, Johnson S, Ivanova E, Kost-Alimova M, Protopopov A, Wang YA, Shirihai OS, Chin L, DePinho RA (2012) Antitelomerase therapy provokes ALT and mitochondrial adaptive mechanisms in cancer. Cell 148(4):651–663. doi:10.1016/j.cell.2011.12.028, S0092-8674(12)00026-8 [pii]

Iyer S, Chadha AD, McEachern MJ (2005) A mutation in the STN1 gene triggers an alternative lengthening of telomere-like runaway recombinational telomere elongation and rapid deletion in yeast. Mol Cell Biol 25(18):8064–8073. doi:10.1128/MCB.25.18.8064-8073.2005, 25/18/8064 [pii]

Kaminker PG, Kim SH, Desprez PY, Campisi J (2009) A novel form of the telomere-associated protein TIN2 localizes to the nuclear matrix. Cell Cycle 8(6):931–939

Kappei D, Butter F, Benda C, Scheibe M, Draskovic I, Stevense M, Novo CL, Basquin C, Araki M, Araki K, Krastev DB, Kittler R, Jessberger R, Londono-Vallejo JA, Mann M, Buchholz F (2013) HOT1 is a mammalian direct telomere repeat-binding protein contributing to telomerase recruitment. EMBO J 32(12):1681–1701. doi:10.1038/emboj.2013.105, emboj2013105 [pii]

Karlseder J, Broccoli D, Dai Y, Hardy S, de Lange T (1999) p53- and ATM-dependent apoptosis induced by telomeres lacking TRF2. Science 283(5406):1321–1325

Kaul Z, Cesare AJ, Huschtscha LI, Neumann AA, Reddel RR (2012) Five dysfunctional telomeres predict onset of senescence in human cells. EMBO Rep 13(1):52–59. doi:10.1038/embor.2011.227, embor2011227 [pii]

Klewes L, Vallente R, Dupas E, Brand C, Grun D, Guffei A, Sathitruangsak C, Awe JA, Kuzyk A, Lichtensztejn D, Tammur P, Ilus T, Tamm A, Punab M, Rubinger M, Olujohungbe A, Mai S (2013) Three-dimensional nuclear telomere organization in multiple myeloma. Transl Oncol 6 (6):749–756

Klonisch T, Wark L, Hombach-Klonisch S, Mai S (2010) Nuclear imaging in three dimensions: a unique tool in cancer research. Ann Anat 192(5):292–301. doi:10.1016/j.aanat.2010.07.007

Kocak H, Ballew BJ, Bisht K, Eggebeen R, Hicks BD, Suman S, O'Neil A, Giri N, Maillard I, Alter BP, Keegan CE, Nandakumar J, Savage SA (2014) Hoyeraal-Hreidarsson syndrome caused by a germline mutation in the TEL patch of the telomere protein TPP1. Genes Dev. doi:10.1101/gad.248567.114

Latrick CM, Cech TR (2010) POT1-TPP1 enhances telomerase processivity by slowing primer dissociation and aiding translocation. EMBO J 29(5):924–933. doi:10.1038/emboj.2009.409, emboj2009409 [pii]

Laud PR, Multani AS, Bailey SM, Wu L, Ma J, Kingsley C, Lebel M, Pathak S, DePinho RA, Chang S (2005) Elevated telomere-telomere recombination in WRN-deficient, telomere dysfunctional cells promotes escape from senescence and engagement of the ALT pathway. Genes Dev 19(21):2560–2570. doi:10.1101/gad.1321305, 19/21/2560 [pii]

Le Saux CJ, Davy P, Brampton C, Ahuja SS, Fauce S, Shivshankar P, Nguyen H, Ramaseshan M, Tressler R, Pirot Z, Harley CB, Allsopp R (2013) A novel telomerase activator suppresses lung damage in a murine model of idiopathic pulmonary fibrosis. PLoS One 8(3), e58423. doi:10.1371/journal.pone.0058423

Lei M, Podell ER, Cech TR (2004) Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end-protection. Nat Struct Mol Biol 11 (12):1223–1229. doi:10.1038/nsmb867

Li JR, Yu TY, Chien IC, Lu CY, Lin JJ, Li HW (2014) Pif1 regulates telomere length by preferentially removing telomerase from long telomere ends. Nucleic Acids Res 42 (13):8527–8536. doi:10.1093/nar/gku541

Longhese MP, Anbalagan S, Martina M, Bonetti D (2012) The role of shelterin in maintaining telomere integrity. Front Biosci (Landmark Ed) 17:1715–1728

Lu W, Zhang Y, Liu D, Songyang Z, Wan M (2013) Telomeres-structure, function, and regulation. Exp Cell Res 319(2):133–141. doi:10.1016/j.yexcr.2012.09.005

Lu R, Pal J, Buon L, Nanjappa P, Shi J, Fulciniti M, Tai YT, Guo L, Yu M, Gryaznov S, Munshi NC, Shammas MA (2014) Targeting homologous recombination and telomerase in Barrett's adenocarcinoma: impact on telomere maintenance, genomic instability and tumor growth. Oncogene 33(12):1495–1505. doi:10.1038/onc.2013.103

Mahalingam D, Tay LL, Tan WH, Chai JH, Wang X (2011) Mutant telomerase RNAs induce DNA damage and apoptosis via the TRF2-ATM pathway in telomerase-overexpressing primary fibroblasts. FEBS J 278(19):3724–3738. doi:10.1111/j.1742-4658.2011.08290.x

Maicher A, Lockhart A, Luke B (2014) Breaking new ground: digging into TERRA function. Biochim Biophys Acta 1839(5):387–394. doi:10.1016/j.bbagrm.2014.03.012

Marcand S, Gilson E, Shore D (1997) A protein-counting mechanism for telomere length regulation in yeast. Science 275(5302):986–990

Marinoni I, Kurrer AS, Vassella E, Dettmer M, Rudolph T, Banz V, Hunger F, Pasquinelli S, Speel EJ, Perren A (2014) Loss of DAXX and ATRX are associated with chromosome instability and reduced survival of patients with pancreatic neuroendocrine tumors. Gastroenterology 146 (2):453–460e455. doi:10.1053/j.gastro.2013.10.020

McClintock B (1941) The stability of broken ends of chromosomes in Zea mays. Genetics 26 (2):234–282

McEachern MJ, Iyer S, Boswell Fulton T, Blackburn EH (2000) Telomere fusions caused by mutating the terminal region of telomeric DNA. Proc Natl Acad Sci U S A 97:11409–11414

Mender I, Gryaznov S, Dikmen ZG, Wright WE, Shay JW (2015) Induction of telomere dysfunction mediated by the telomerase substrate precursor 6-thio-2′-deoxyguanosine. Cancer Discov 5(1):82–95. doi:10.1158/2159-8290.CD-14-0609

Mitchell JR, Wood E, Collins K (1999) A telomerase component is defective in the human disease dyskeratosis congenita. Nature 402(6761):551–555. doi:10.1038/990141

Mitchell M, Gillis A, Futahashi M, Fujiwara H, Skordalakes E (2010) Structural basis for telomerase catalytic subunit TERT binding to RNA template and telomeric DNA. Nat Struct Mol Biol 17(4):513–518. doi:10.1038/nsmb.1777

Morrish TA, Greider CW (2009) Short telomeres initiate telomere recombination in primary and tumor cells. PLoS Genet 5(1), e1000357. doi:10.1371/journal.pgen.1000357

Moye AL, Porter KC, Cohen SB, Phan T, Zyner KG, Sasaki N, Lovrecz GO, Beck JL, Bryan TM (2015) Telomeric G-quadruplexes are a substrate and site of localization for human telomerase. Nat Commun 6:7643. doi:10.1038/ncomms8643

Müller HJ (1938) The remaking of chromosomes. Collect Net 13:181–198

Nandakumar J, Bell CF, Weidenfeld I, Zaug AJ, Leinwand LA, Cech TR (2012) The TEL patch of telomere protein TPP1 mediates telomerase recruitment and processivity. Nature 492:285–289. doi:10.1038/nature11648, nature11648 [pii]

Neumann AA, Watson CM, Noble JR, Pickett HA, Tam PP, Reddel RR (2013) Alternative lengthening of telomeres in normal mammalian somatic cells. Genes Dev 27(1):18–23. doi:10.1101/gad.205062.112

Niida H, Shinkai Y, Hande MP, Matsumoto T, Takehara S, Tachibana M, Oshimura M, Lansdorp PM, Furuichi Y (2000) Telomere maintenance in telomerase-deficient mouse embryonic stem cells: characterization of an amplified telomeric DNA. Mol Cell Biol 20:4115–4127

Ning H, Li T, Zhao L, Li J, Liu J, Liu Z, Fan D (2006) TRF2 promotes multidrug resistance in gastric cancer cells. Cancer Biol Ther 5(8):950–956

Novo CL, Londono-Vallejo JA (2013) Telomeres and the nucleus. Semin Cancer Biol 23 (2):116–124. doi:10.1016/j.semcancer.2012.02.001

O'onnor MS, Safari A, Liu D, Qin J, Songyang Z (2004) The human Rap1 protein complex and modulation of telomere length. J Biol Chem 279(27):28585–28591. doi:10.1074/jbc.M312913200

O'Sullivan RJ, Almouzni G (2014) Assembly of telomeric chromatin to create ALTernative endings. Trends Cell Biol. doi:10.1016/j.tcb.2014.07.007

O'Sullivan RJ, Arnoult N, Lackner DH, Oganesian L, Haggblom C, Corpet A, Almouzni G, Karlseder J (2014) Rapid induction of alternative lengthening of telomeres by depletion of the histone chaperone ASF1. Nat Struct Mol Biol 21(2):167–174. doi:10.1038/nsmb.2754, nsmb.2754 [pii]

Oh BK, Kim YJ, Park C, Park YN (2005) Up-regulation of telomere-binding proteins, TRF1, TRF2, and TIN2 is related to telomere shortening during human multistep hepatocarcinogenesis. Am J Pathol 166(1):73–80. doi:10.1016/S0002-9440(10)62233-X, S0002-9440(10)62233-X [pii]

Osterwald S, Deeg KI, Chung I, Parisotto D, Worz S, Rohr K, Erfle H, Rippe K (2015) PML induces compaction, TRF2 depletion and DNA damage signaling at telomeres and promotes their alternative lengthening. J Cell Sci 128(10):1887–1900. doi:10.1242/jcs.148296

Panero J, Stanganelli C, Arbelbide J, Fantl DB, Kohan D, Garcia Rivello H, Rabinovich GA, Slavutsky I (2014) Expression profile of shelterin components in plasma cell disorders. Clinical significance of POT1 overexpression. Blood Cells Mol Dis 52(2–3):134–139. doi:10.1016/j.bcmd.2013.10.002

Peters AH, O'Carroll D, Scherthan H, Mechtler K, Sauer S, Schofer C, Weipoltshammer K, Pagani M, Lachner M, Kohlmaier A, Opravil S, Doyle M, Sibilia M, Jenuwein T (2001) Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. Cell 107(3):323–337

Peuscher MH, Jacobs JJ (2012) Posttranslational control of telomere maintenance and the telomere damage response. Cell Cycle 11(8):1524–1534. doi:10.4161/cc.19847

Poncet D, Belleville A, t'kint de Roodenbeke C, Roborel de Climens A, Ben Simon E, Merle-Beral H, Callet-Bauchu E, Salles G, Sabatier L, Delic J, Gilson E (2008) Changes in the expression of telomere maintenance genes suggest global telomere dysfunction in B-chronic

lymphocytic leukemia. Blood 111(4):2388–2391. doi:10.1182/blood-2007-09-111245, blood-2007-09-111245 [pii]
Queisser A, Heeg S, Thaler M, von Werder A, Opitz OG (2013) Inhibition of telomerase induces alternative lengthening of telomeres during human esophageal carcinogenesis. Cancer Genet 206(11):374–386. doi:10.1016/j.cancergen.2013.10.001
Ramsay AJ, Quesada V, Foronda M, Conde L, Martinez-Trillos A, Villamor N, Rodriguez D, Kwarciak A, Garabaya C, Gallardo M, Lopez-Guerra M, Lopez-Guillermo A, Puente XS, Blasco MA, Campo E, Lopez-Otin C (2013) POT1 mutations cause telomere dysfunction in chronic lymphocytic leukemia. Nat Genet 45(5):526–530. doi:10.1038/ng.2584, ng.2584 [pii]
Redon S, Reichenbach P, Lingner J (2010) The non-coding RNA TERRA is a natural ligand and direct inhibitor of human telomerase. Nucleic Acids Res 38(17):5797–5806. doi:10.1093/nar/gkq296
Riou JF (2004) G-quadruplex interacting agents targeting the telomeric G-overhang are more than simple telomerase inhibitors. Curr Med Chem Anticancer Agents 4(5):439–443
Sahin E, Depinho RA (2010) Linking functional decline of telomeres, mitochondria and stem cells during ageing. Nature 464(7288):520–528. doi:10.1038/nature08982
Sampl S, Pramhas S, Stern C, Preusser M, Marosi C, Holzmann K (2012) Expression of telomeres in astrocytoma WHO grade 2 to 4: TERRA level correlates with telomere length, telomerase activity, and advanced clinical grade. Transl Oncol 5(1):56–65
Sarek G, Vannier JB, Panier S, Petrini JH, Boulton SJ (2015) TRF2 recruits RTEL1 to telomeres in S phase to promote t-loop unwinding. Mol Cell 57(4):622–635. doi:10.1016/j.molcel.2014.12.024
Sauerwald A, Sandin S, Cristofari G, Scheres SH, Lingner J, Rhodes D (2013) Structure of active dimeric human telomerase. Nat Struct Mol Biol 20(4):454–460. doi:10.1038/nsmb.2530
Schmidt JC, Dalby AB, Cech TR (2014) Identification of human TERT elements necessary for telomerase recruitment to telomeres. eLife 3. doi:10.7554/eLife.03563
Schoeftner S, Blasco MA (2008) Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. Nat Cell Biol 10(2):228–236. doi:10.1038/ncb1685
Schoeftner S, Blasco MA (2010) Chromatin regulation and non-coding RNAs at mammalian telomeres. Semin Cell Dev Biol 21(2):186–193. doi:10.1016/j.semcdb.2009.09.015
Schwartzentruber J, Korshunov A, Liu XY, Jones DT, Pfaff E, Jacob K, Sturm D, Fontebasso AM, Quang DA, Tonjes M, Hovestadt V, Albrecht S, Kool M, Nantel A, Konermann C, Lindroth A, Jager N, Rausch T, Ryzhova M, Korbel JO, Hielscher T, Hauser P, Garami M, Klekner A, Bognar L, Ebinger M, Schuhmann MU, Scheurlen W, Pekrun A, Fruhwald MC, Roggendorf W, Kramm C, Durken M, Atkinson J, Lepage P, Montpetit A, Zakrzewska M, Zakrzewski K, Liberski PP, Dong Z, Siegel P, Kulozik AE, Zapatka M, Guha A, Malkin D, Felsberg J, Reifenberger G, von Deimling A, Ichimura K, Collins VP, Witt H, Milde T, Witt O, Zhang C, Castelo-Branco P, Lichter P, Faury D, Tabori U, Plass C, Majewski J, Pfister SM, Jabado N (2012) Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. Nature 482(7384):226–231. doi:10.1038/nature10833, nature10833 [pii]
Sexton AN, Regalado SG, Lai CS, Cost GJ, O'Neil CM, Urnov FD, Gregory PD, Jaenisch R, Collins K, Hockemeyer D (2014) Genetic and molecular identification of three human TPP1 functions in telomerase action: recruitment, activation, and homeostasis set point regulation. Genes Dev 28(17):1885–1899. doi:10.1101/gad.246819.114
Sfeir A, de Lange T (2012) Removal of shelterin reveals the telomere end-protection problem. Science 336(6081):593–597. doi:10.1126/science.1218498
Shay JW, Reddel RR, Wright WE (2012) Cancer. Cancer and telomeres–an ALTernative to telomerase. Science 336(6087):1388–1390. doi:10.1126/science.1222394
Svendsen JM, Smogorzewska A, Sowa ME, O'Connell BC, Gygi SP, Elledge SJ, Harper JW (2009) Mammalian BTBD12/SLX4 assembles a Holliday junction resolvase and is required for DNA repair. Cell 138(1):63–77. doi:10.1016/j.cell.2009.06.030

Szostak JW, Blackburn EH (1982) Cloning yeast telomeres on linear plasmid vectors. Cell 29 (1):245–255

Tauchi T, Shin-ya K, Sashida G, Sumi M, Okabe S, Ohyashiki JH, Ohyashiki K (2006) Telomerase inhibition with a novel G-quadruplex-interactive agent, telomestatin: in vitro and in vivo studies in acute leukemia. Oncogene 25(42):5719–5725. doi:10.1038/sj.onc.1209577, 1209577 [pii]

Tennen RI, Bua DJ, Wright WE, Chua KF (2011) SIRT6 is required for maintenance of telomere position effect in human cells. Nat Commun 2:433. doi:10.1038/ncomms1443

Tomlinson RL, Ziegler TD, Supakorndej T, Terns RM, Terns MP (2006) Cell cycle-regulated trafficking of human telomerase to telomeres. Mol Biol Cell 17(2):955–965. doi:10.1091/mbc.E05-09-0903

Vannier JB, Pavicic-Kaltenbrunner V, Petalcorin MI, Ding H, Boulton SJ (2012) RTEL1 dismantles T loops and counteracts telomeric G4-DNA to maintain telomere integrity. Cell 149 (4):795–806. doi:10.1016/j.cell.2012.03.030, S0092-8674(12)00418-7 [pii]

Venteicher AS, Artandi SE (2009) TCAB1: driving telomerase to Cajal bodies. Cell Cycle 8 (9):1329–1331

Venteicher AS, Meng Z, Mason PJ, Veenstra TD, Artandi SE (2008) Identification of ATPases pontin and reptin as telomerase components essential for holoenzyme assembly. Cell 132 (6):945–957. doi:10.1016/j.cell.2008.01.019

Vogan JM, Collins K (2015) Dynamics of human telomerase holoenzyme assembly and subunit exchange across the cell cycle. J Biol Chem. doi:10.1074/jbc.M115.659359

Walne AJ, Vulliamy T, Kirwan M, Plagnol V, Dokal I (2013) Constitutional mutations in RTEL1 cause severe dyskeratosis congenita. Am J Hum Genet 92(3):448–453. doi:10.1016/j.ajhg.2013.02.001

Wan M, Qin J, Songyang Z, Liu D (2009) OB fold-containing protein 1 (OBFC1), a human homolog of yeast Stn1, associates with TPP1 and is implicated in telomere length regulation. J Biol Chem 284(39):26725–26731. doi:10.1074/jbc.M109.021105

Wang F, Lei M (2011) Human telomere POT1-TPP1 complex and its role in telomerase activity regulation. Methods Mol Biol 735:173–187. doi:10.1007/978-1-61779-092-8_17

Wang X, Kam Z, Carlton PM, Xu L, Sedat JW, Blackburn EH (2008) Rapid telomere motions in live human cells analyzed by highly time-resolved microscopy. Epigenetics Chromatin 1(1):4. doi:10.1186/1756-8935-1-4

Weissbein U, Benvenisty N, Ben-David U (2014) Quality control: genome maintenance in pluripotent stem cells. J Cell Biol 204(2):153–163. doi:10.1083/jcb.201310135

Williams SC (2013) No end in sight for telomerase-targeted cancer drugs. Nat Med 19(1):6. doi:10.1038/nm0113-6, nm0113-6 [pii]

Wu L, Multani AS, He H, Cosme-Blanco W, Deng Y, Deng JM, Bachilo O, Pathak S, Tahara H, Bailey SM, Behringer RR, Chang S (2006) Pot1 deficiency initiates DNA damage checkpoint activation and aberrant homologous recombination at telomeres. Cell 126(1):49–62. doi:10.1016/j.cell.2006.05.037, S0092-8674(06)00763-X [pii]

Xu Z, Duc KD, Holcman D, Teixeira MT (2013) The length of the shortest telomere as the major determinant of the onset of replicative senescence. Genetics 194(4):847–857. doi:10.1534/genetics.113.152322

Yu TY, Kao YW, Lin JJ (2014) Telomeric transcripts stimulate telomere recombination to suppress senescence in cells lacking telomerase. Proc Natl Acad Sci U S A 111 (9):3377–3382. doi:10.1073/pnas.1307415111

Zeng S, Xiang T, Pandita TK, Gonzalez-Suarez I, Gonzalo S, Harris CC, Yang Q (2009) Telomere recombination requires the MUS81 endonuclease. Nat Cell Biol 11(5):616–623. doi:10.1038/ncb1867, ncb1867 [pii]

Zhao Y, Abreu E, Kim J, Stadler G, Eskiocak U, Terns MP, Terns RM, Shay JW, Wright WE (2011) Processive and distributive extension of human telomeres by telomerase under homeostatic and nonequilibrium conditions. Mol Cell 42(3):297–307. doi:10.1016/j.molcel.2011.03.020, S1097-2765(11)00248-6 [pii]

Zhong FL, Batista LF, Freund A, Pech MF, Venteicher AS, Artandi SE (2012) TPP1 OB-fold domain controls telomere maintenance by recruiting telomerase to chromosome ends. Cell 150 (3):481–494. doi:10.1016/j.cell.2012.07.012, S0092-8674(12)00874-4 [pii]

Zou Y, Liu Y, Wu X, Shell SM (2006) Functions of human replication protein A (RPA): from DNA replication to DNA damage and stress responses. J Cell Physiol 208(2):267–273. doi:10.1002/jcp.20622

Part IV
Nuclear Domains and Development

Polycomb Bodies

Vincenzo Pirrotta

Abstract Polycomb bodies are concentrations of Polycomb Group proteins detectable in the nucleus at various intensities. The largest or most intense have been shown to contain genomic clusters of Polycomb Group targets such as the *Hox* gene clusters. Since, in general, the number of Polycomb bodies visible is an order of magnitude smaller than the known number of Polycomb target genes in the genome, they are often thought to involve the association of multiple genomic regions that are distant from one another in the genome. This chapter reviews the evidence for Polycomb bodies, their formation and their genomic content. While different lines of evidence indicate that genomically remote Polycomb target genes can associate, often enhancing the repressive effect, other evidence indicates that this is not usually a stable interaction, varies from one tissue to another, and is strongly dependent on the presence of insulator protein binding sites near Polycomb targets. The effects of transcriptional derepression and of post-transcriptional modifications of Polycomb proteins or of insulator proteins as factors modulating the association lf remote Polycomb target sites are also discussed.

1 Nuclear "Bodies"

Eukaryotic nuclei contain many regions where certain particular proteins aggregate forming foci, generally accompanied by genomic chromatin with which their function is associated. Perhaps the most extreme example of this tendency might be said to be the nucleolus, the most prominent feature of the nucleus, where ribosomal RNA genes from multiple genomic sites gather together with the transcription apparatus, the nascent rRNA and the proteins that bind to, process, and eventually transport the rRNA. Many varieties of smaller "bodies", "speckles", "paraspeckles", "dots", etc. have been described (for review, see Mao et al. 2011)

V. Pirrotta (✉)
Department of Molecular Biology and Biochemistry, Rutgers University, Nelso Laboratories, 604 Allison Road, Piscataway, NJ 08854, USA
e-mail: pirrotta@dls.rutgers.edu

D.P. Bazett-Jones, G. Dellaire (eds.), *The Functional Nucleus*,
DOI 10.1007/978-3-319-38882-3_7

but in this article, we will be concerned primarily with two: Polycomb bodies and transcription factories. An earlier treatment of some of the issues discussed here was presented in Pirrotta and Li (2012). Polycomb bodies have been visualized principally by fluorescence microscopy of fixed tissues stained with antibodies against one or more Polycomb Group (PcG) proteins or against histone H3 trimethylated at lysine27 (H3K27me3), a modification characteristic of repressed chromatin associated with PcG proteins. Live cell microscopy has also been used, with GFP-tagged PcG proteins. Basic questions are whether these Polycomb bodies represent multiple chromatin sites or storage accumulations of proteins. If they contain multiple genomic sites, what are the mechanisms that bring these sites together and what is their function. Transcription factories are structures that include more than one RNA polymerase II, whose existence is shown by the fact that although many thousands of genes are transcriptionally active in any one cell, nascent RNA transcripts and RNA polymerase II are primarily found in a much smaller number of foci (Osborne et al. 2004). Since PcG proteins and the complexes in which they assemble silence transcription of many hundreds of genes, such target PcG genes are expected to be found in Polycomb bodies when they are repressed but in transcription factories when they are active. The question is therefore not just what forms Polycomb bodies but what brings PcG target genes together and where they are congregating. To examine these questions we need first to summarize the basic known features of Polycomb complexes. The material in the following section is more comprehensively reviewed in recent articles (Simon and Kingston 2013; Schwartz and Pirrotta 2013), therefore specific references will not be given here.

2 Polycomb Complexes and Their Genomic Function

Polycomb mechanisms were first discovered in the genetic analysis of the regulation of homeotic genes in *Drosophila*. Inappropriate expression of homeotic genes caused by mutations in the Polycomb Group of genes cause lethality or dramatic phenotypes. Many more genes are subject to repression by PcG mechanisms but PcG role in the case of the major homeotic genes *Antennapedia*, *Ultrabithorax* and *Abdominal-B* has remained an important influence in the thinking about the mechanisms involved.

Homeotic genes must be expressed in the appropriate embryonic segments in *Drosophila* to specify their identity in the anterior/posterior axis. Thus, Antp and the Bithorax Complex genes must be expressed in the thoracic segments to specify thoracic development but must be repressed in more anterior segments to prevent formation of thoracic structures in the head region of the fly. The homeotic genes are activated in the embryonic segments by segmentation genes and are initially repressed by positional factors that ultimately depend on maternal determinants deposited in the anterior or posterior regions of the embryo. Thus, the *Ubx* gene is activated principally by the segmentation genes *ftz* and *eve* but is repressed by maternal and zygotic *hb* in the anterior half and posterior extremity of the embryo.

These initial cues are short-lived and disappear around the time of gastrulation. The PcG system, which comes into function shortly before gastrulation, is responsible for maintaining the repressed state in the cells where the gene was repressed in early embryonic stages. Although PcG proteins are present in all cells, PcG repression does not occur wherever the target genes were active. Once established, the PcG-dependent repressed state tends to be maintained from one cell cycle to the next. This is particularly true in the case of the *Drosophila* homeotic genes where the repressed state is maintained for the rest of development.

PcG repression is recruited at target genes by specific DNA elements, the Polycomb Response Elements (PREs), regions that usually span a few hundred nucleotides and contain binding motifs for several DNA-binding proteins that are thought to act cooperatively to effect the recruitment of PcG complexes. Several complexes in fact bind at PREs and are important in producing stable binding of the two key PcG complexes, the Polycomb Repressive Complexes 1 and 2. PRC1 and PRC2 are almost always found together at PRE regions of target genes. Although the details have become rather complex in recent years, in brief, the PRC2 complex is a histone H3 lysine 27 methyltransferase that produces H3K27 trimethylation in the chromatin region surrounding a PRE to which it is bound. This region may extend over many tens of kb although the PRC2 complex remains localized at the PRE recruitment site. The PRC1 complex contains a core of four proteins RING/Sce, Psc, Ph and Pc. In particular, Ring and Psc together are responsible for histone H2A lysine 118 ubiquitylation, which is now thought to help recruit PRC2, while Pc, the eponymous Polycomb protein, contains a chromodomain that recognizes the H3K27me3 deposited by PRC2. Thus in some way the two complexes help stabilize one another. It is important to note, however, that ChIP studies find both PRC1 and PRC2 bound at the PRE and even the Pc protein remains at the PRE and is not stably associated with the entire domain of H3K27me3. It is thought, however that the interaction of the Pc chromodomain with H3K27me3 helps the chromatin looping that causes the spread of the methylated region, which probably stabilizes the repressed state and helps to maintain it from one cell cycle to the next.

In recent years, Polycomb complexes and PRC1 in particular have turned out to include a range of variants, particularly in mammalian cells, where there are six known alternative Psc homologs called PCGF1-6, two RING proteins, three Ph homologs, five Pc homologs called CBX2,4,6,7,8, as well as other associated proteins. Importantly, some of these variant complexes do not include a chromodomain-containing CBX component and therefore do not recognize H3K27me3 but instead have a strong H2AK119 ubiquitylating activity that helps recruit PRC2. Mammalian PcG complexes are generally found at CpG island promoters, where their recruitment is initiated by a variant PRC1 complex containing KDM2B, provided that such regions are not DNA-methylated and are not transcriptionally active. KDM2B is a histone H3K36 demethylase but its role here is to bind to unmethylated PcG islands. The H2AK119 ubiquitylation produced by this variant PRC1 complex favors the recruitment of PRC2 and in turn the H3K27me3 produced by PRC2 facilitates the recruitment of CBX-containing PRC1 complexes.

Genomic ChIP-seq studies, supported by a large number of functional studies have shown that, in addition to homeotic genes, a large number of other genes are targets of PcG binding and regulation in both mammals and flies. The number and identity of target genes varies from one cell type to another and PcG protein binding varies in intensity at different sites. There is clearly a large number of binding sites in any one cell, ranging from hundreds in *Drosophila* cells to more than a thousand mammalian cells. In embryonic stem cells, a large number of genes bind PcG complexes and are maintained in a bivalent state where PcG complexes and H3K27me3 coexist with the binding of RNA polymerase and the "active" histone modification H3K4me3.

3 Evidence for PcG Foci from Nuclear Imaging Studies

Antibody staining for various PRC1 or PRC2 components reveals a distribution in diploid nuclei in the form of speckles of variable intensity and number. In general however, what is striking about the images produced is that the number of significant speckles is relatively low, compared to the number of known genomic binding sites. Much of this is due to the fact that PcG target genes tend to be clustered in the genome. In fact in salivary gland polytene chromosome spreads where any long-range contacts have deliberately been broken by squashing, only about 100 bands appear to stain with antibodies against PcG proteins. The majority of these sites contain more than one PcG target genes and, in both flies and mammals, the major clusters are the homeotic gene complexes. Nevertheless, it has been commonly remarked that in intact diploid nuclei fewer foci of PcG proteins are visible than would be expected from the known number of genomic binding sites in a given cell type. Possible explanations are that target genes are not continuously binding PcG proteins, that many individual signals are too weak to be consistently detected; or that multiple target genes are gathered together to form a single focus.

In *Drosophila*, as in other dipterans, pairing of homologous chromosomes during interphase becomes detectable during embryogenesis and remains highly highly efficient during the rest of development. Alleles of PcG target genes therefore tend to be associated during interphase. Nuclear images from *Drosophila* embryos expressing GFP-tagged PcG proteins revealed "100 or more" foci of widely varying intensities but provided no evidence of association of multiple PcG loci (Buchenau et al. 1998). More discrete foci were reported in embryos stained with an antibody against Pc protein (Bantignies et al. 2011). Furthermore, combinations of FISH (Fluorescence In Situ Hybridization) and antibody staining demonstrated that the *Antp* gene co-localizes with the *Abd-B* gene in about 20 % of the embryonic nuclei, 7 % in larval imaginal disc nuclei. although the two genes are distant from each other more than 10 Mb along chromosome 3 L (Grimaud et al. 2006; Bantignies et al. 2011). In agreement with the known binding pattern of PcG complexes, co-localization does not take place in cells in which the *Abd-B* gene is repressed but the *Antp* gene is active. However, another careful study of

embryos, larvae, brain and cultured cells found that co-localization is absent or very low in most tissues and, when detectable, it is both tissue- and strain-dependent (Fedorova et al. 2008).

Similar Polycomb bodies have been reported in mammalian cells where they vary considerably during the cell cycle and among different cell types (Gil et al. 2004; Hernández-Muñoz et al. 2005; Ren et al. 2008). Tumor cell lines and primary tumor cultures tend to have a few but very large Polycomb bodies associated with centromeric satellite sequences and localized adjacent to heterochromatin (Saurin et al. 1998; Voncken et al. 1999). During the cell cycle, PcG proteins staining at these locations weakens and disappears in late S phase, returning only at the next G_1. These large sub-heterochromatic Polycomb bodies are clearly abnormal but might nevertheless be informative about the processes that produce them. They might be due, for example, to abnormal SUMOylation (see below) or loss of glycosylation (see below) or to loss of DNA methylation of CpG-rich satellite sequences that would turn them into massive targets of PcG recruitment (Cooper et al. 2014).

Discrete foci of PcG proteins have been observed in mammalian embryonic stem (ES) cells, indicating that they are not just consequences of pathological states but normal manifestations. Foci of the mammalian CBX2, 4, 6, 7 and 8 proteins have been studied in ES cells using GFP fusions (Ren et al. 2008). Interestingly, they have different nuclear distributions from one another, with CBX2 found in both small and large foci, CBX6 in a granular pattern and CBX4, 7 and 8 in numerous small foci. Although the significance of these patterns remains unclear, the differences are perhaps better understood in view of the known differences in the relative affinities for H3K27me3 and H3K9me3 (Bernstein et al. 2006) and the dominant role of CBX7, which represses pro-differentiation CBX2, 4 and 8 in undifferentiated ES cells (Morey et al. 2012; Klauke et al. 2013). When differentiation is induced, all the CBX proteins become dispersed in a granular pattern and, at least in this study, stronger foci such as would correspond to PcG target clusters such as the *Hox* loci, were not in evidence (Ren et al. 2008).

4 Transgene-Dependent Evidence

The study of PcG function in *Drosophila* involved the generation of large numbers of transgenes containing PREs or fragments of PREs and reporter genes. A common feature of PRE-containing transgenes is that often, but not in all cases, they are much more strongly repressed when they are homozygous than when present in a single copy. Since homologous chromosomes are stably paired during interphase in *Drosophila*, this implies that physical proximity of two PREs enhances their repressive ability, a phenomenon called Pairing-Dependent Silencing. *Drosophila* P-element transgene constructs generally insert at random sites in the genome but in some cases they were found to interact functionally with the genomic PcG target from which they were derived or with other PRE-containing transgenes inserted at

other genomic sites, resulting in enhanced silencing. These surprising observations, reminiscent of Pairing-Dependent Silencing, suggested that such transgenes inserted at remote genomic sites, sometimes on different chromosomes, could interact at long distances enhancing the function of the PREs they contained. Such behavior was studied with transgenes containing the *Mcp* or the *Fab-7* elements from the *Abd-B* gene, elements known to contain a PRE adjacent to a sequence with enhancer blocker/boundary activity (Müller et al. 1999; Bantignies et al. 2003). This interaction involved also the endogenous element but, while the presence of the latter was required in the case of *Fab-7* (Bantignies et al. 2003), the interaction between two *Mcp* transgenes did not require the endogenous *Mcp* (Li et al. 2011). The actual physical co-localization of such remote transgenes in the nucleus was demonstrated by combinations of Fluorescent In Situ Hybridization (FISH) and antibody staining on fixed tissues (Bantignies et al. 2003; Fedorova et al. 2008) or in live cells by tagging the transgenes with a *LacO* array to which bound a *Lac* repressor fused to GFP (Vazquez et al. 2006; Li et al. 2011, 2013). The remarkable stability of the co-localization in the latter case is demonstrated by the fact that even when five distinct transgene copies are present in different genomic sites, they all co-localize in a single focus in the live nuclei. However, Fedorova et al. (2008) concluded that the association of *Fab-7* transgenes with the endogenous locus might not occur in all tissues or for all PREs and did not appear to constitute the general rule.

A clue to this behavior was given by earlier experiments using the *bxd* PRE, one of the most powerful PcG-repressive elements in *Drosophila*. Transgenes containing the *bxd* PRE together with the *gypsy* Su(Hw) insulator element, inserted in some genomic sites were able to interact with a second copy inserted at other sites, resulting in increased silencing (Sigrist and Pirrotta 1997). Remarkably, this interaction was lost in the presence of mutations in the *Su(Hw)* gene that abolished insulator function. Similarly, such trans-interactions had never been seen with transgenes containing the bxd PRE but no insulator element. The suggestion implied by these experiments was therefore that the insulator in some way mediated long-distance trans-interactions between the transgenic PREs. This was explicitly demonstrated by subdividing the *Mcp* or *Fab-7* region into an insulator fragment and a PRE fragment. Transgenes containing the PRE fragment alone failed to show long-distance trans-interactions while transgenes containing the insulator fragment alone were still able to trans-interact (Li et al. 2011). The conclusion therefore was that interaction between the insulator elements, not an interaction between PcG complexes, mediated the trans-interaction and co-localization of the transgenes. This was confirmed by showing that the *Mcp* and *Fab-7* insulators both contained binding sites for the CTCF insulator protein and that co-localization was lost entirely when CTCF was knocked out (Li et al. 2013). The role of the insulator activity of *Fab-7* in co-localization might help explain the finding of Fedorova et al. (2008) that *Fab-7* transgene co-localization depended on the tissue. In fact the *Fab-7* insulator function has been reported to vary in embryonic stages and tissues (Aoki et al. 2008).

There remained the question of the much lower co-localization frequency seen with the insulator alone compared to the transgenes used by Vazquez et al. (2006). A critical difference was the presence of the eye enhancer in the transgenes of Vazquez et al. In fact, in all the transgene-based studies, the transgene contained a marker gene, usually the *white* gene, to identify the transgenic flies. The transgene construct is therefore at least partially transcriptionally active, active part of the time or in some of the cells, raising the question whether the transgene would be localized in a repressed or an active compartment. Further analysis showed that, in fact, although the insulator element alone was necessary and sufficient to mediate a low co-localization frequency, the presence of an enhancer plus the full PRE yielded far higher frequencies and stability of co-localization (Li et al. 2013). This was apparently not due to PRC1 complexes binding to the PRE but to the fact that the PRE also recruits the binding and activity of Trithorax. Mutations in the *trx* gene, even when heterozygous, abolished the high level of co-localization and reduced it to the level seen with the insulator alone. This expands the discussion from that of Polycomb foci where such a PRE + insulator transgene would localize when repressed, to foci of transcriptional activity or transcription factories (Osborne et al. 2004), where the transgene would localize when it becomes derepressed by the presence of appropriate enhancers. Thus, the *Antp* gene would co-localize with *Abd-B* when both are PcG-repressed but not when one is repressed and the other active (Bantignies et al. 2011) but might be expected to co-localize again, this time in a transcriptionally active compartment such as a transcription factory, when both genes are active in a Trithorax-dependent way (Li et al. 2013).

5 The Involvement of the RNAi Machinery

The involvement of insulator function in co-localization probably links two disparate sets of observations concerning the effects of RNAi components on co-localization and its effects on PcG-mediated repression. While studying the trans-acting effects of a transgene containing the *Fab-7* element (PRE + insulator) from the *Abd-B* gene, Grimaud et al. (2006) observed that the interaction of the transgene inserted on the X chromosome with the endogenous *Fab-7* element on chromosome 3 L, causing an enhanced repression when both were present, were visibly weakened by mutations in RNAi components such as *dicer*, *piwi*, and *AGO1*. These mutations do not interfere with the recruitment of PcG proteins to the transgenic *Fab-7* element but they reduce the frequency of their co-localization and that of endogenous PcG targets. These three proteins also are found associated with Polycomb foci by antibody staining. This connection between PcG silencing and the RNAi machinery is intriguing but mechanistically not understood. The authors concluded that the RNAi machinery is implicated in the association of PcG target sites in the nucleus to form Polycomb foci.

A possible way to connect RNAi proteins with the results on the co-localization of PRE-containing transgenes is through the role of insulator proteins such as CTCF

or CP190. Insulator function has been reported to decrease in the presence of mutations in RNAi components and the organization of insulator proteins in the nucleus is disrupted (Lei and Corces 2006). Furthermore, AGO2 but not AGO1 or dicer was found at many binding sites of CTCF and CP190, including *Fab-7* and other PREs in the Bithorax Complex BX-C (Moshkovich et al. 2011). The role of AGO2 in promoting insulator organization and function does not require its catalytic activity and is therefore distinct from its role in RNAi mechanisms. This account differs in some details from that of Grimaud et al. but agrees in arguing that RNAi components play a role in the association of remote PcG targets in the nucleus and suggests that, while the mechanism remains unclear, it is related to the part played by insulator elements. A very interesting and very recent report shows that mammalian CTCF binds RNA transcripts from thousands of sites, often in close proximity to CTCF genomic binding sites (Kung et al. 2015). At least in one rather special case, that of the X chromosome inactivation center (Xic), this interaction with nearby transcripts is important in mediating long-range interactions such as the pairing of two Xics.

According to this picture, therefore, it is not the PcG complexes themselves that mediate long-distance association of their binding sites in the nucleus, or at least, not primarily the PcG complexes. Rather, chromatin looping mediated by architectural elements such as the insulators of the *Mcp* or *Fab-7* elements allows certain PcG binding sites to reach out of chromosomal territories and into the nuclear interior to contact at least subsets of similar sites, possibly through interactions between the PcG proteins, and enhance their repressive functions. Alternatively, when the transgene is active, the same architectural/insulator elements mediate the association of the site to a transcription factory. What, according to this model, would specify to which Polycomb body or transcrption factory a gene would be targeted remains unclear. Possibly. RNA molecules might play a part.

6 The Assembly and Behavior of Polycomb Bodies

To gain a better understanding on the formation of these nuclear foci, Cheutin and Cavalli (2012) studied their kinetics of assembly in the live nuclei of developing embryos expressing GFP-tagged Pc or Ph proteins. 3D confocal imaging showed that both Pc and Ph assemble in relatively few foci of different intensities. The homeotic gene clusters ANT-C and BX-C are generally found in the largest of these. In the early embryo, the intensity of the signal is well correlated with the linear genomic extent of the individual PcG target region and does not suggest that more than one PcG target occupies a given Polycomb body. However, as the embryo develops, the accumulation of Pc or Ph in the Polycomb foci increases progressively. This may suggest that, after the initial recruitment of PcG complexes, additional copies of the complexes accrete to each target site. There is no evidence from ChIP data for extension of complex recruitment along the chromatin

but ChIP would not detect the accumulation of additional complexes atop those originally recruited at a PRE.

As earlier observed by Buchenau et al. (1998), PcG protein foci do not correspond to DNA-dense region of the nucleus and are not therefore associated with very dense chromatin. These studies confirmed that PcG proteins exchange rapidly between bound and unbound states (Ficz et al. 2005; Ren et al. 2008) and showed that the Polycomb foci move by constrained diffusion within relatively small confines but together with surrounding chromatin and rarely give rise to contacts between different Polycomb bodies. Overall therefore, these studies suggest that the majority of Polycomb bodies represent individual genomic regions and that association between remote PcG target sites is a relatively rare event, probably limited to sites with special characteristics, e.g., insulator/boundary elements such as the *Mcp* or *Fab-7* elements. Note that they do not exclude the possibility that PcG target clusters such as the *Antennapedia* complex ANT-C and the bithorax complex BX-C could associate, as the same laboratory had shown earlier by FISH in fixed tissues (Bantignies et al 2003).

7 Chromosome Conformation Capture and Nuclear Architecture

An approach to evaluate the occurrence of contacts between two chromatin regions was developed by J. Dekker (Dekker et al. 2002) and called Chromosome Conformation Capture (3C). The methodology has evolved to permit identification of multiple contact sites (4C, 5C) and, taking advantage of the enormous power of massively parallel sequencing, recent versions allow detection of all possible contacts of all genomic sequences with all other genomic sequences (Hi-C). 3C methods have been used to identify looping contacts between PREs within the BX-C homeotic cluster (Lanzuolo et al. 2007). Using an optimized 4C approach, Tolhuis et al. (2011) asked if long-distance contacts might be observed also among PcG target sites other than the homeotic gene clusters. They found that the homeotic loci on chromosome 3R of *Drosophila* had significant and specific contacts with other chromatin domains, primarily on the same chromosome arm and that the majority of these contacts involved other PcG-binding domains. A similar picture emerged if the analysis used other, non-homeotic PcG target sites as a starting point: PcG domains interact preferentially with other PcG domains on the same chromosome arm, while non-PcG regions, even when transcriptionally inactive, have no such preferential contacts with PcG-binding domains but interact with other genomic regions. The analysis of chromosomal rearrangements demonstrated that the preference for interactions within a chromosome arm is due to a general architectural constraint. Interchromosomal contacts are not entirely absent but much less frequent. This suggests that a chromosome arm forms a nuclear territory out of which constituent regions, both transcriptionally active and repressed, may

loop out while the centromere (and pericentromeric heterochromatin) constitutes a considerable obstacle for contacts between the two arms. A similar preference for intrachromosomal interactions has been observed in mammalian cells (Simonis et al. 2006; Lieberman-Aiden et al. 2009). A chromatin contact map of the genome in *Drosophila* embryos was obtained by Sexton et al. (2012) using the more general Hi-C approach. This study confirmed that, although interchromosomal contacts are far more rare than intra-chromosome arm contacts, they occur also between arms and, to a lesser extent between chromosomes. This study confirmed the preferential interactions among PcG target sites on a given chromosome arm and found that the contact domains had a good correspondence with domains of epigenetic marks. In other words, chromatin regions with similar epigenetic marks are more likely to interact with one another. Interestingly, the borders of contact domains were frequently binding sites of insulator proteins, particularly CP190, BEAF-32 and CTCF. Another protein found frequently at the nodes between domains was Chromator, a chromatin protein that is specifically enriched in interband regions of *Drosophila* polytene chromosomes (Vatolina et al. 2011). These studies support the idea that chromatin domains binding PcG proteins can and do interact with one another preferentially and that domains are often demarcated by insulator protein binding sites. However, they do not explain what mediates these interactions and, if insulator proteins are involved, what is responsible for the specificity of PcG site interaction with other PcG sites. It should be noted also that the contacts revealed by the Hi-C studies are statistically significant but do not correspond to stable interactions. Rather, they appear as relatively rare contacts, either in terms of stability or in terms of the frequency with which they occur.

Hi-C studies of mammalian cells present a similar picture of the architectural genome organization. A very recent study in human cell lines used improved procedures and high depth of sequencing to achieve the unprecedented resolution of about 1 kb (Rao et al. 2014) and can represent earlier studies for present purposes. The increase in resolution shows that the chromatin is organized in a hierarchical fashion at the root of which are simple chromatin loops frequently juxtaposing enhancers and promoters. The bases of these loops are often formed by binding sites for CTCF arranged in opposing orientation along the DNA sequence so that, when paired at the base of the loop, they align CTCF molecules in the same orientation, suggesting that the two CTCF units form dimers. These loops are further organized in domains containing multiple loops which may interact with one another to different extents. Contacts are correlated with similar histone modifications, one of which corresponds to the PcG-repressed state.

8 Comparing Conformation Capture and Imaging Studies

The architectural organization of genomic chromatin revealed by the Hi-C studies provides enormous detail about the type and detail of the long-range interactions that can occur but does not yet integrate them into a global view that gives us a

space-filling view of what happens in the nucleus. FISH studies can give a larger scale picture. Chromosome painting probes have shown that chromatin belonging to each chromosome tends to occupy a specific nuclear domain or territory. They also demonstrate vividly that gene-rich chromosomes or chromosome regions reside in the interior of the nucleus while gene-poor chromosomes or chromosome regions and heterochromatin are found at the nuclear periphery, often attached to the nuclear lamina, where condensed chromatin can be visualized. Inactive genes that become transcriptionally activated move from the nuclear periphery to a more internal position (see review by Deniaud and Bickmore 2009).

A detailed examination of the *HoxA* locus using FISH shows a compact configuration in embryonic stem cells, where it is inactive. The compaction corresponds to the region bearing the H3K27me3 mark and is shown by gene knockout experiments to depend on the binding of PRC1. When the *Hox* genes become activated during development or differentiation, they appear to become progressively less compact and to loop out from the chromosome territory in the order 3′ to 5′, corresponding to their sequential activation (Morey et al. 2007). Despite the extended appearance, however, 4C and 5C studies indicate that the active *Hox* genes engage in extensive interactions among themselves and their regulatory regions (Noordermeer et al. 2011, 2014; Wang et al. 2011; Chambeyron et al. 2005). This is then the problem: if in fact the PcG-repressed *Hox* loci are packaged within the chromosome territory, it is difficult to see how they could be accessible for interaction with other PcG targets and form Polycomb bodies. On the other hand, it is perhaps not so clear that the inactive *HoxA* cluster is really highly condensed and truly embedded in the chromosome territory. In fact, different studies concur in finding that H3K27me3-associated chromatin is not particularly dense (Buchenau et al. 1998; Ren et al. 2008; Cheutin and Cavalli 2012). In addition, FISH studies may not distinguish between compacted chromatin and chromatin folded by looping interactions. Thus, while the *HoxA* cluster might become more extended during differentiation and certainly subdivided into active and inactive genes, it is not necessarily inaccessible to interactions with other PcG loci in the repressed state.

9 Role of PRC1 Components

If PRC1 components mediate the interaction of one PcG target region with another, presumably when they are juxtaposed by insulator proteins such as CTCF, what could be those components and those interactions? An old observation that has returned to generate much interest is that a conserved domain of the PRC1 component Polyhomeotic (Ph), called SAM (Sterile Alpha Motif) or SPM, is responsible for its polymerization in vitro, producing an open-ended, head to tail helical filament (Kim et al. 2002). SAM domains with similar properties are found in other transcriptional regulators, including another PcG protein, Scm, which is required for PcG repression and is loosely associated with the PRC1 complex. The potential

for polymerization or heteropolymerization of Ph and Scm was seized upon as a possible explanation for the action of PRC1 complexes over long distances and perhaps to mediate interactions between remote genomic sites (Isono et al. 2013). Mutational analysis shows that the ability to oligomerize in vitro is well correlated with the in vivo function of Ph and other SAM-containing factors (Robinson et al. 2012). However, PRC1 complexes purified from a baculovirus expression system do not display any degree of oligomerization and no evidence exists for polymerization occurring in vivo other than the importance of the SAM domain for the silencing function. Instead, fly and mammalian Ph proteins have been shown to be targets for O-GlcNAcylation by the glycosyltransferase OGT, a modification required for effective PcG repression (Gambetta et al. 2009; Sinclair et al. 2009; Gambetta and Müller 2014). O-GlcNAcylation of a cluster of serine/threonines in the central part of the insect or human Ph proteins largely blocks the formation of oligomers through the SAM domain. In vivo, such oligomers form when O-GlcNAcylation is absent and are nonfunctional. It is possible, however, that a residual degree of SAM domain-mediated association still occurs and is important for function. A likely possibility is that the SAM domain is essential for repression because it mediates the interaction between Ph and Scm, both necessary for PRC1 function although the precise role of Scm is also poorly understood. Mutations in the SAM domain would then prevent recruitment of effective PRC1 function not because of the failure to form oligomers but because they disable the interaction with an important PRC1 participant protein. Overall, then, it appears that SAM domain-mediated oligomerization is not involved in the formation of functional aggregates of PcG complexes in the nucleus.

10 SUMO and Polycomb Body Formation

The question of the role of the higher order organization of PcG proteins in the nucleus was tackled directly in a high throughput screen to identify genes whose loss of function altered the distribution of a GFP-tagged Pc protein (Gonzalez et al. 2014). This study used RNAi knockdown with a library of 24,000 distinct dsRNAs, coupled with automated confocal microscopy to detect any changes in the pattern of fluorescence in the nuclei of *Drosophila* cultured S2 cells. While the screen yielded 129 confirmed candidate genes including many known to affect chromatin organization, among the most prominent were the *smt3* gene encoding the ubiquitin-related small protein SUMO and *veloren* (*velo*) encoding a peptidase that reverses SUMO conjugation. Loss of SUMO function resulted in the formation of fewer but more intense Pc foci while loss of *velo* function had the opposite effect of producing a loss of foci in favor of diffuse Pc distribution. The change in distribution affects Ph as well as Pc, suggesting that it involves the entire PRC1 complex. Furthermore, the Pc distribution corresponds to that of H3K27me3, implying that it represents chromatin-bound PRC1 complexes. It has been shown in fact that the human Pc2 protein is SUMOylated and itself acts as a SUMO E3

ligase on other chromatin proteins such as CTCF (Kagey et al. 2003; MacPherson et al. 2009). ChIP-seq analysis in *Drosophila* embryos confirms that both SUMO and Velo are frequently found at PcG binding sites but immunostaining of nuclei showed that neither protein is strongly co-localized with Pc foci (Gonzalez et al. 2014), suggesting that clustering involves un-SUMOylated Pc, or, alternatively, that there is a high level of other SUMOylated proteins in the nucleus that masks a modest degree of Pc SUMOylation. The knockdown analyses indicated that SUMOylation antagonizes Pc binding to chromatin and slows down its kinetics while de-SUMOylation by Velo promotes binding. Consistent with this, loss of Velo function induces developmental defects indicative of loss of PcG repression. SUMO is therefore a negative regulator of Pc binding to chromatin. In contrast, mammalian CTCF was found to co-localize well with the Pc2 protein (now called CBX4) in HeLa cells (MacPherson et al. 2009).

11 A Synthesis?

Polycomb bodies, foci of PcG proteins in the nucleus, clearly can form under certain conditions. It is not clear that they frequently involve the coming together of multiple loci separated by large genomic distances. However, it is clear that certain PcG targets have the capacity to contact other, distant PcG targets and associate with them at least part of the time. Such association would be expected to enhance and stabilize silencing, as in fact is known to happen in *Drosophila* when two PRE-containing transgene inserted at homologous sites pair through homologous chromosome pairing (Pairing-Dependent Silencing). Association of non-homologous PcG target genes is probably promoted by insulator or architectural proteins such as CTCF. However, it seems to occur at significant frequency only between sites possessing binding sites for these proteins close to the PcG target. Such would be the *Drosophila Mcp* and *Fab-7* elements. Such associations cannot be stable or long-lived in most cases or the enhancement of silencing observed in the case of homologous pairing would not be noticeable. The role of CTCF is particularly interesting in view of its involvement in large-scale chromatin architecture. However, this role must be regulated and reversible since when a PcG target gene becomes transcriptionally active, it must relocate to an active compartment such as a transcription factory. That CTCF or similar architectural proteins play a role in this relocation is suggested by the fact that CTCF is implicated in the high-level co-localization dependent on Trithorax reported for *Mcp* transgenes in *Drosophila* (Li et al. 2013). The co-localization of active transgenes in this case is much more frequent and stable than that seen when the transgenes are in the repressed state.

Mechanisms to promote or hinder the association of PcG complexes and, presumably, PcG-bound chromatin regions are indicated by the role played by SUMOylation. The genetic data in *Drosophila* shows that SUMOylation of Pc results in excessive aggregation either at a few favored sites or at non-productive

sites in the nucleus at the expense of the genomic sites where Pc would otherwise be recruited. This effect of SUMOylation is reminiscent of the case of PML bodies, where the PML protein recognizes SUMOylation and must be itself SUMOylated in order to aggregate with itself and other components of PML bodies (Bernardi and Pandolfi 2007). In the case of Polycomb bodies, such aggregation is either not essential or must be counteracted or controlled by the Velo SUMO peptidase. At the same time, like PML, the Pc protein is itself a SUMO E3 ligase. The human Pc2/CBX4 SUMOylates CTCF and CTCF co-localizes with Polycomb bodies (MacPherson et al. 2009). Regulated SUMOylation may therefore be a key to the assembly of at least a subset of Polycomb bodies. That high levels of aggregation of PRC1 complexes are undesirable is shown also by the necessity to control oligomerization of the Ph SAM domain by counteracting it through glycosylation, presumably because it would also lead to nonproductive sequestration of the Ph protein or of the entire PRC1 complex (Gambetta et al. 2009; Gambetta and Müller 2014).

References

Aoki T, Schweinsberg S, Manasson J, Schedl P (2008) A stage-specific factor confers Fab-7 boundary activity during early embryogenesis in Drosophila. Mol Cell Biol 28(3):1047–1060

Bantignies F, Grimaud C, Lavrov S, Gabut M, Cavalli G (2003) Inheritance of Polycomb-dependent chromosomal interactions in Drosophila. Genes Dev 17:2406–2420

Bantignies F, Roure V, Comet I, Leblanc B, Schuettengruber B, Bonnet J, Tixier V, Mas A, Cavalli G (2011) Polycomb-dependent regulatory contacts between Distant Hox Loci in *Drosophila*. Cell 144(2):214–226

Bernardi R, Pandolfi PP (2007) Structure, dynamics and functions of promyelocytic leukaemia nuclear bodies. Nat Rev Mol Cell Biol 8(12):1006–1016

Bernstein E, Duncan EM, Masui O, Gil J, Heard E, Allis CD (2006) Mouse Polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. Mol Cell Biol 26(7):2560–2569

Buchenau P, Hodgson J, Strutt H, Arndt-Jovin DJ (1998) The distribution of polycomb-group proteins during cell division and development in *Drosophila* embryos: impact on models for silencing. J Cell Biol 141:469–481

Chambeyron S, Da Silva NR, Lawson KA, Bickmore WA (2005) Nuclear re-organisation of the Hoxb complex during mouse embryonic development. Development 132(9):2215–2223

Cheutin T, Cavalli G (2012) Progressive Polycomb assembly on H3K27me3 compartments generates Polycomb bodies with developmentally regulated motion. PLoS Genet 8(1), e1002465

Dekker J, Rippe K, Dekker M, Kleckner N (2002) Capturing chromosome conformation. Science 295:1306–1311

Deniaud E, Bickmore WA (2009) Transcription and the nuclear periphery: edge of darkness? Curr Opin Genet Dev 19(2):187–191

Fedorova E, Sadoni N, Dahlsveen I, Koch J, Kremmer E, Eick D, Paro R, Zink D (2008) The nuclear organization of Polycomb/Trithorax group response elements in larval tissues of *Drosophila* melanogaster. Chromosome Res 16(4):649–673

Ficz G, Heintzmann R, Arndt-Jovin DJ (2005) Polycomb group protein complexes exchange rapidly in living *Drosophila*. Development 132(17):3963–3976

Gambetta MC, Oktaba K, Muller J (2009) Essential role of the glycosyltransferase Sxc/Ogt in Polycomb repression. Science 325(5936):93–96
Gambetta MC, Müller J (2014) O-GlcNAcylation prevents aggregation of the Polycomb group repressor polyhomeotic. Dev Cell 31(5):629–639
Gil J, Bernard D, Martinez D, Beach D (2004) Polycomb CBX7 has a unifying role in cellular lifespan. Nat Cell Biol 6:67–72
Gonzalez I, Mateos-Langerak J, Thomas A, Cheutin T, Cavalli G (2014) Identification of regulators of the three-dimensional Polycomb organization by a microscopy-based genome-wide RNAi screen. Mol Cell 54(3):485–499
Grimaud C, Bantignies F, Pal-Bhadra M, Ghana P, Bhadra U, Cavalli G (2006) RNAi components are required for nuclear clustering of Polycomb group response elements. Cell 124:957–971
Hernandez-Munoz I, Taghavi P, Kuijl C, Neefjes J, van Lohuizen M (2005) Association of BMI1 with Polycomb bodies is dynamic and requires PRC2/EZH2 and the maintenance DNA methyltransferase DNMT1. Mol Cell Biol 25(24):11047–11058
Isono K, Endo TA, Ku M, Yamada D, Suzuki R, Sharif J, Ishikura T, Toyoda T, Bernstein BE, Koseki H (2013) SAM Domain polymerization links subnuclear clustering of PRC1 to gene silencing. Dev Cell 26(6):565–577
Kagey MH, Melhuish TA, Wotton D (2003) The Polycomb protein Pc2 is a SUMO E3. Cell 113:127–137
Kim CA, Gingery M, Pilpa RM, Bowie JU (2002) The SAM domain of polyhomeotic forms a helical polymer. Nat Struct Biol 9:453–456
Klauke K, Radulovic V, Broekhuis M, Weersing E, Zwart E, Olthof S, Ritsema M, Bruggeman S, Wu X, Helin K, Bystrykh L, de Haan G (2013) Polycomb Cbx family members mediate the balance between haematopoietic stem cell self-renewal and differentiation. Nat Cell Biol 15:353–362
Kung JT, Kesner B, An JY, Ahn JY, Cifuentes-Rojas C, Colognori D, Jeon Y, Szanto A, del Rosario BC, Pinter SF, Erwin JA, Lee JT (2015) Locus-specific targeting to the X chromosome revealed by the RNA interactome of CTCF. Mol Cell 57(2):361–375
Lanzuolo C, Roure V, Dekker J, Bantignies F, Orlando V (2007) Polycomb response elements mediate the formation of chromosome higher-order structures in the bithorax complex. Nat Cell Biol 9(10):1167–1174
Lei EP, Corces VG (2006) RNA interference machinery influences the nuclear organization of a chromatin insulator. Nat Genet 38:936–941
Li H-B, Muller M, Bahechar IA, Kyrchanova O, Ohno K, Georgiev P, Pirrotta V (2011) Insulators, not Polycomb response elements, are required for long-range interactions between Polycomb targets in *Drosophila* melanogaster. Mol Cell Biol 31(4):616–625
Li H-B, Ohno K, Gui H, Pirrotta V (2013) Insulators target active genes to transcription factories and Polycomb-repressed genes to Polycomb bodies. PLoS Genet 9(4), e1003436
Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326(5950):289–293
MacPherson MJ, Beatty LG, Zhou W, Du M, Sadowski PD (2009) The CTCF insulator protein is posttranslationally modified by SUMO. Mol Cell Biol 29(3):714–725
Mao YS, Zhang B, Spector DL (2011) Biogenesis and function of nuclear bodies. Trends Genet 27 (8):295–306
Morey C, Da Silva NR, Perry P, Bickmore WA (2007) Nuclear reorganisation and chromatin decondensation are conserved, but distinct, mechanisms linked to Hox gene activation. Development 134(5):909–919
Morey L et al (2012) Nonoverlapping functions of the Polycomb Group Cbx family of proteins in embryonic stem cells. Cell Stem Cell 10(1):47–62

Moshkovich N, Nisha P, Boyle PJ, Thompson BA, Dale RK, Lei EP (2011) RNAi-independent role for Argonaute2 in CTCF/CP190 chromatin insulator function. Genes Dev 25 (16):1686–1701

Muller M, Hagstrom K, Gyurkovics H, Pirrotta V, Schedl P (1999) The *Mcp* element from the *Drosophila* melanogaster bithorax complex mediates long-distance regulatory interactions. Genetics 153:1333–1356

Noordermeer D, Leleu M, Splinter E, Rougemont J, De Laat W, Duboule D (2011) The dynamic architecture of Hox gene clusters. Science 334(6053):222–225

Noordermeer D, Leleu M, Schorderet P, Joye E, Chabaud F, Duboule D, Krumlauf R (2014) Temporal dynamics and developmental memory of 3D chromatin architecture at Hox gene loci. eLife 3, e02557

Osborne CS, Chakalova L, Brown KE, Carter D, Horton A, Debrand E, Goyenechea B, Mitchell JA, Lopes S, Reik W, Fraser P (2004) Active genes dynamically colocalize to shared sites of ongoing transcription. Nat Genet 36:1065–1071

Pirrotta V, Li H-B (2012) A view of nuclear Polycomb bodies. Curr Opin Genet Dev 22 (2):101–109

Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, Aiden EL (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell 159(7):1665–1680

Ren X, Vincenz C, Kerppola TK (2008) Changes in the distributions and dynamics of Polycomb repressive complexes during embryonic stem cell differentiation. Mol Cell Biol 28 (9):2884–2895

Robinson AK, Leal BZ, Chadwell LV, Wang R, Ilangovan U, Kaur Y, Junco SE, Schirf V, Osmulski PA, Gaczynska M, Hinck AP, Demeler B, McEwen DG, Kim CA (2012) The growth-suppressive function of the Polycomb group protein polyhomeotic is mediated by polymerization of its Sterile Alpha Motif (SAM) domain. J Biol Chem 287(12):8702–8713

Saurin AJ, Shiels C, Williamson J, Satijn DPE, Otte AP, Sheer D, Freemont PS (1998) The human Polycomb Group complex associates with pericentromeric heterochromatin to form a novel nuclear domain. J Cell Biol 142:887–898

Schwartz YB, Pirrotta V (2013) A new world of Polycombs: unexpected partnerships and emerging functions. Nat Rev Genet 14(12):853–864

Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, Parrinello H, Tanay A, Cavalli G (2012) Three-dimensional folding and functional organization principles of the *Drosophila* genome. Cell 148(3):458–472

Sigrist CJA, Pirrotta V (1997) Chromatin insulator elements block the silencing of a target gene by the Drosophila Polycomb Response Element (PRE) but allow trans interactions between PREs on different chromosomes. Genetics 147:209–221

Simon JA, Kingston RE (2013) Occupying chromatin: Polycomb mechanisms for getting to genomic targets, stopping transcriptional traffic, and staying put. Mol Cell 49(5):808–824

Simonis M, Klous P, Splinter E, Moshkin Y, Willemsen R, de Wit E, van Steensel B, de Laat W (2006) Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nat Genet 38:1348–1354

Sinclair DAR, Syrzycka M, Macauley MS, Rastgardani T, Komljenovic I, Vocadlo DJ, Brock HW, Honda BM (2009) *Drosophila* O-GlcNAc transferase (OGT) is encoded by the Polycomb group (PcG) gene, super sex combs (sxc). Proc Natl Acad Sci U S A 106(32):13427–13432

Tolhuis B, Blom M, Kerkhoven RM, Pagie L, Teunissen H, Nieuwland M, Simonis M, de Laat W, van Lohuizen M, van Steensel B (2011) Interactions among Polycomb domains are guided by chromosome architecture. PLoS Genet 7(3), e1001343

Vatolina TY, Boldyreva LV, Demakova OV, Demakov SA, Kokoza EB, Semeshin VF, Babenko VN, Goncharov FP, Belyaeva ES, Zhimulev IF (2011) Identical functional organization of nonpolytene and polytene chromosomes in *Drosophila* melanogaster. PLoS One 6(10), e25960

Vazquez J, Müller M, Pirrotta V, Sedat JW (2006) The *Mcp* element mediates stable long-range chromosome-chromosome interactions in *Drosophila*. Mol Biol Cell 17:2158–2165

Voncken JW, Schweizer D, Aagaard L, Sattler L, Jantsch MF, van Lohuizen M (1999) Chromatin association of the Polycomb group protein BMI1 is cell cycle-regulated and correlates with it phosphorylation status. J Cell Sci 112:4627–4639

Wang KC, Yang YW, Liu B, Sanyal A, Corces-Zimmerman R, Chen Y, Lajoie BR, Protacio A, Flynn RA, Gupta RA, Wysocka J, Lei M, Dekker J, Helms JA, Chang HY (2011) A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature 472(7341):120–124

The Continuing Flight of Ikaros

Karen E. Brown

Ikaros is a key regulator of lymphocyte differentiation and its loss within the hematopoietic system and can result in leukemic transformation. Ikaros functions at discrete steps during lineage specification: initially in early hematopoietic progenitors, where it is necessary for lymphoid lineage differentiation potential and later, at the proliferative stages of B- and T-cell differentiation, when selection of B- and T-repertoires occurs. Ikaros is able to both activate and silence gene expression through interaction with a variety of tissue- and developmental stage-specific co-factors. Ikaros family proteins can recruit chromatin-remodelling machinery to loci to control transcription. The activity of Ikaros can be regulated at multiple levels: at the transcriptional level, by competing with other transcription factors for binding sites, by co-operating with or antagonizing a range of differentiation stage- or cell type-specific partners, and by post-translational modification including phosphorylation or SUMOylation.

Understanding how Ikaros selects and regulates key target genes is a major challenge in the post-genomics era. The realization that many factors can be found at multiple genomic sites, yet regulate of only a small subset of these potential targets poses a real challenge for our understanding of their function. The number of potential Ikaros binding sites in the genome far exceeds the number of genes affected by gain or loss of this factor. A better understanding of chromatin structure (using DamID (Cléard et al. 2014) and 'Hi-C' (Dostie and Bickmore 2012) techniques to probe chromatin accessibility and three-dimensional organization) has revealed that the concept of chromatin as either 'accessible' or 'inaccessible' is overly simplistic (van Steensel 2011). Ikaros is involved in a complex web of interactions with signalling pathways regulating feedback and feedforward loops that define cellular function at successive stages of differentiation as cells'

K.E. Brown (✉)
MRC Clinical Sciences Centre, Imperial College London, Hammersmith Hospital Campus, Du Cane Road, London NW12 0NN, UK
e-mail: Karen.brown@csc.mrc.ac.uk

D.P. Bazett-Jones, G. Dellaire (eds.), *The Functional Nucleus*,
DOI 10.1007/978-3-319-38882-3_8

signalling and niche environments change. The way in which Ikaros interfaces with lineage-specific co-factors in the evolving chromatin landscape in response to intrinsic and extrinsic cues needs to be understood for a better mechanistic insight into both hematopoiesis and leukemic transformation, where even subtle perturbation of Ikaros function may tip the balance towards disease.

1 The Diverse Functions of Ikaros

Ikaros function is not limited to lymphocyte commitment, survival and homeostasis. Ikaros regulates the transcriptional program of early hematopoietic progenitors, influences erythroid cell progenitor homeostasis, erythroid cell differentiation, as well as dendritic and neutrophil cell specification. Ikaros can bind to proximal and distal regulatory regions of diverse genes, many of which are regulated by Ikaros in a lineage-specific way. Through its ability to dimerize, Ikaros-containing complexes bind to both target genes and to mouse γ-satellite repeat sequences present at pericentromeric heterochromatin. In this way, Ikaros can sequester genes to new locations of the nucleus enriched for gene silencing factors. Through interaction with chromatin remodelling machinery, Ikaros can alter the epigenetic landscape and influence locus organization, as well as transcriptional activity. How Ikaros achieves its exceptional degree of lineage-specific function is of increasing interest in the post-genomics era. A broad picture is emerging of the remarkable spectrum of Ikaros's interactions with DNA, the chromatin landscape as it evolves during development, as well as its interactions with lineage specific and ubiquitous transcriptional regulators and chromatin remodelling machinery and the way in which its function can be controlled by it's regulation at the transcriptional and post-translational levels. Some of the ways in which the Ikaros family interacts with DNA, chromatin, signalling pathways and nuclear organization in order to achieve its diverse effects are described below. In 23 years since its discovery (Lo et al. 1991; Georgolpoulos et al. 1992; Hahm et al. 1994), the flight of Ikaros continues to soar. The mechanistic principles underlying its pleiotropic effects may act as a paradigm for understanding the ways in which lineage-commitment and specification operate.

2 Ikaros Family and Structure

Ikaros proteins contain a highly conserved N-terminal C2H2 zinc finger DNA-binding domain and a C-terminal zinc finger domain. These mediate its interaction with other proteins, or its dimerization with Ikaros isoforms derived by alternate splicing of exons 3–7 (Georgolpoulos et al. 1992; Hahm et al. 1994). The isoforms are shown schematically in Fig. 1. All stable isoforms share the two C-terminal zinc fingers but differ in the number of N-terminal zinc fingers, and so

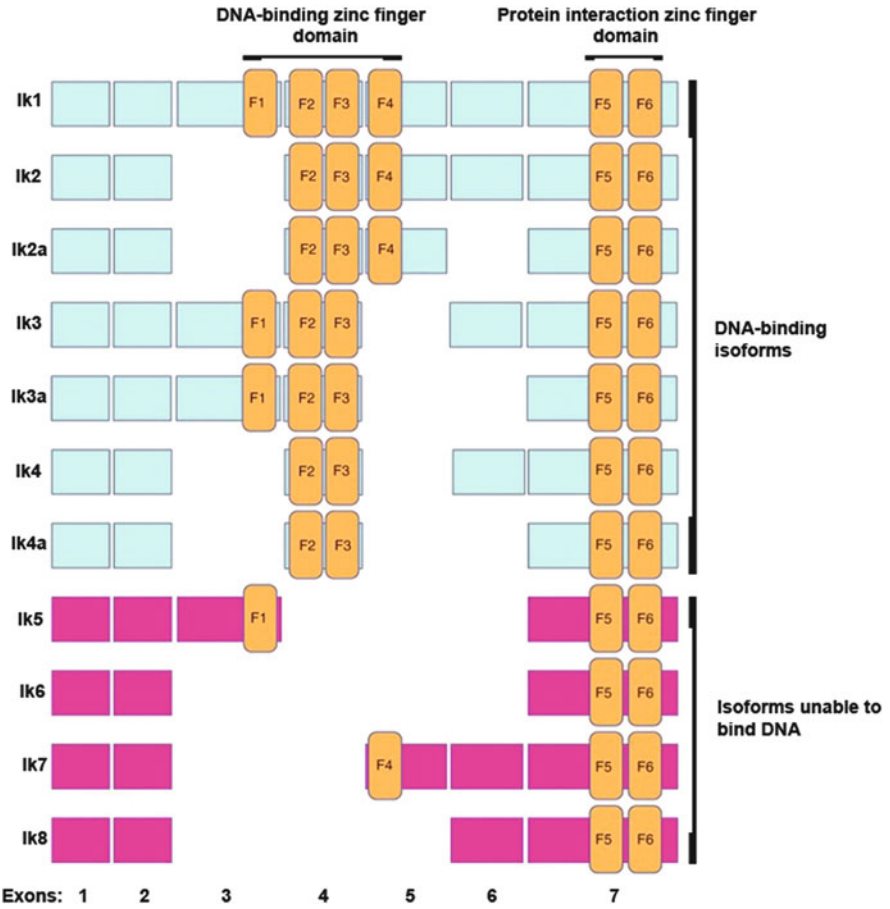

Fig. 1 Schematic representation of the isoforms of Ikaros generated by alternative splicing. Isoforms lacking zinc fingers 2 and 3 are unable to bind DNA and act as dominant negatives. After Davis (2011)

differ in their DNA-binding specificity. Ikaros isoforms lacking the DNA-binding domain can dimerize with full-length Ikaros and inhibit its function, so exerting a 'dominant negative' effect (Papathanasiou et al. 2003). Over-expression of dominant negative isoforms of Ikaros, such as Ik6 (Kastner et al. 2013) Ik10 (Wang et al. 2014) and Ik-11 (Capece et al. 2013), occur in transformed hematopoietic cell lines and primary leukemic cells, where they dimerize with endogenous Ik-1 or Ik-2 and interfere with DNA binding or compete for incorporation into chromatin-remodelling complexes. In transformed cells, dominant negative isoforms are often generated by internal deletions within rather than by splicing defects (Tokunaga et al. 2013).

In addition to full length Ikaros (Ik-1), most hematopoietic cells abundantly express isoform Ik-2, and the shorter isoforms (Ik-3 to Ik-8) are generally far less

abundant (Molnar and Gerorgopoulos 1994). Ikaros-x is an isoform that is selectively expressed in human myeloid cells (Payne et al. 2003). Differences in some functions of Ikaros between mice and humans may be reflected by the existence of a human-specific isoform of Ikaros (Ik-H) reported to be involved solely in gene activation (Li et al. 2011). As well as Ikaros-1, humans express the largest form of Ikaros, Ik-H, which includes a 30 amino acid region encoded by exon 3B, which restricts its DNA-binding to bipartite GGGAA Ikaros consensus sites. Ik-H can potentiate or inhibit the binding of Ik-1 to DNA, and is able to bind to the regulatory regions of genes that are upregulated by Ikaros, but not to repressed genes (Dovat 2011). Ik-H binds at pericentromeric heterochromatin as well as to euchromatic sites. The anti-silencing ability of satellite DNA (Kim et al. 2009) has been shown to be dependent on Ik-H binding (Dovat 2011).

There are five members of the Ikaros family in mouse and human: Ikaros (*Ikzf1*), Helios (*Ikzf2*) and Aiolos (*Ikzf3*) are largely restricted to hematopoietic cells, with Eos (*Ikzf4*) and Pegasus (*Ikzf5*) being more broadly expressed. Ikaros-deficient mice lack B-cells, NK cells, peripheral lymph nodes and show delayed T-cell development and perturbed myelopoiesis (Wang et al. 1996). Deletion of the Ikaros DNA-binding domain results in the absence of lymphocytes and their precursors (Georgopoulos et al. 1994). Ikaros hemizygous mice show abnormal B- and T-cell function and develop T-cell malignancies. Aiolos deficiency results in abnormal B-cell development and function (Morgan et al. 1997). Different Ikaros family members can exert different effects, despite having similar DNA-binding consensus sequences. For example, a Helios-deficient DT40 B-cell line shows increased response to B-cell receptor signalling, whereas the Ikaros-deficient equivalent shows a decreased response (Alinikula et al. 2010). The development of adult T-cells in Ikaros-deficient mice is thought to occur through the substitution of other Ikaros family members (Wang et al. 1996).

3 Ikaros DNA-Binding Specificity

Several mammalian transcription factors that are involved in diverse biological functions or are involved in multiple steps in a single pathway are modular in structure and contain more Zn fingers than the tandemly organized two or three required for stable DNA binding in vitro (Wolfe et al. 2000). Zinc fingers 2 and 3 of Ikaros's 4-finger N-terminal domain are sufficient for binding to the core DNA-binding consensus A/GGGAA (Molnar et al. 1996). It has recently been postulated that additional fingers or distinct combinations of fingers may modulate binding to (or select) specific target sites involved in different functions, increasing the functional repertoire of the transcription factor. Mice lacking either zinc-finger 1 or 4 were generated by germline deletion of either exon 4 or 6 of *Ikzf1* and phenotypes compared with Ikaros-null mice ($Ikzf1^{-/-}$). These mice exhibited distinct subsets of the phenotypes of the Ikaros null mice, with some interesting specific effects: large pre-BII cells were substantially decreased in $Ikzf1^{\Delta F1/\Delta F1}$ yet

were increased in *Ikzf1*$^{\Delta F4/\Delta F4}$ mice, and only *Ikzf1*$^{\Delta F4/\Delta F4}$ mice developed aggressive lymphomas (as are seen in *Ikzf1*$^{-/-}$ mice), indicating that zinc finger 4 is essential for tumour suppression whilst zinc finger 1 is dispensable. RNA-Seq data indicated that genes mis-regulated in *Ikzf1*$^{-/-}$ mice are a composite of those mis-regulated in *Ikzf1*$^{\Delta F1/\Delta F1}$ and *Ikzf1*$^{\Delta F4/\Delta F4}$ mice. Zinc fingers 1 and 4 could regulate different biological events by regulating distinct target genes, possibly only a subset of those bound by Ikaros according to ChIP-Seq data. Many genes were up-regulated from a low base, indicating that Ikaros often acts as a transcriptional repressor (Schjerven et al. 2013). A similar mode of action has been proposed for the multi-zinc-finger factor CTCF (Nakahashi et al. 2013).

4 Ikaros and the Chromatin Landscape

Ikaros is able to either activate or repress transcription of its target genes (Kim et al. 1999). One way in which it could influence transcriptional status is by modulating chromatin structure of the regulatory regions of its target genes. Biochemical purification experiments have shown that Ikaros can interact with chromatin remodelling machinery and recruit it to target genes.

Ikaros can associate with the NuRD complex in lymphoid and erythroid cells (Kim et al. 1999; Sridharan and Smale 2007). The NuRD complex contains histone deacatylases (HDACs) and Mi-2β, which can remodel chromatin by re-positioning nucleosomes (Becker and Horz 2002). The NuRD complex usually mediates gene repression. Ikaros is also able to associate with the gene-activating chromatin remodelling complexes, for example SWI/SNF binds together with Ikaros and NuRD in the PYR complex involved in globin switching in erythroid cells (O'Neill et al. 2009). Another way in which Ikaros regulates gene activation is through direct protein-protein interaction with GATA factors and the cyclin-dependent kinase Cdk9, which is a component of the positive transcriptional elongation factor P-TEFb, required for transcriptional avtivation of Ikaros target genes. Ikaros and GATA1 interact through their C-terminal zinc finger domains, whereas the N-terminal end of Ikaros is required for interaction with Cdk9 (Bottardi et al. 2013).

The developmental progression from multipotent progenitors to specialized cell types is guided by the iterative actions of transcription factors and chromatin modifiers over time. The availability of binding sites for particular factors depends on the existing landscape of histone modifications and nucleosome packaging, as well as the presence of other factors that are able to modify the chromatin landscape to potentiate the binding of factors that arrive later in lineage commitment. The hierarchical, directional nature of lineage commitment depends on the way in which present chromatin regulatory mechanisms set up opportunities for future gene expression, while removing the conditions needed for past gene expression. The binding of 'pioneer' factors (which modify chromatin to pave the way for future factor binding) at particular genomic sites depends on both the currently

bound co-factors and binding site competitors as well as on the prior regulatory state of the region. The 'availability' of a target site may differ between cell types. For example, Pax5 can bind different sets of potential targets in pre-B and pro-B cells compared with mature B-cells (Revilla-i-Domingo et al. 2012). A factor bound to a particular target may activate its expression only after being joined by a co-factor expressed at a later stage of differentiation. Such factors may also influence transcription by interacting with other factors located at distant regulatory regions, bringing them into physical contact, and enabling transcription factors act in a combinatorial way over time. Factor binding may depend on the presence of other interacting factors at a locus, rather than the possession of a core DNA consensus sequence for the factor in the promoter itself (reviewed by Rothenberg 2014). The epigenetic landscape can remain 'plastic' during lineage commitment. The $CD4^+$ specification factor Th-POK is active in early T-cell precursors, is silenced during commitment, when the locus acquires H3K27 trimethylation, then is re-activated a few stages later during positive selection (Zhang et al. 2012).

5 Lineage Specificity of Ikaros Activity

Ikaros has multiple lineage-specific effects and shows a remarkable degree of lineage specificity in target recognition. One way in which Ikaros can achieve lineage specificity is through interaction with factors expressed at defined stages of differentiation. The B lineage commitment factor Pax5 is absent in lymphoid progenitor cells but is expressed in pro-B cells, in which its binding showed enrichment at Ikaros ChIP-seq peaks on activated Ikaros targets. The observed minimal overlap of Ikaros-regulated genes in lymphoid progenitor cells and pro-B cells further supports the principal of stage-specific Ikaros function (Schwickert et al. 2014). Ikaros may interact with some lineage-specific factors functionally if not physically. Kruppel-associated box (KRAB)-associated protein 1 (KAP1) is a co-factor involved in repression. T-cell-specific deletion of *Kap1* resulted in expansion of immature thymocytes, imbalanced $CD4^+/CD8^+$ ratios and altered responses to TCR and TGFβ stimulation. KAP1 was found to bind to T-cell-specific *cis*-acting regulatory elements marked by H3K9me3 and enriched in Ikaros/NuRD complexes, and the frequent overlap of binding sites may indicate involvement of KAP1 in fine-tuning the expression of Ikaros targets (Santoni de Sio et al. 2012).

Chromatin enrichment studies of Ikaros binding sites in lymphocytes have led to the identification of several thousand (Ferreirós-Vidal et al. 2013), with strong selection for the core DNA motif AGAA in lymphocytes. Ikaros DNA-binding specificity is similar to that described for a variety of Ets factors on lymphoid-specific regulatory elements (Hollenhorst et al. 2007). The relative distributions of Ikaros binding sites in promoters and enhancers is cell-type dependent. At T-cell-specific enhancers, for example, the most frequent partners of Ikaros are the E-box E2A proteins and Runx1, implying possible functional interaction between these factors during T-cell differentiation (Zhang et al. 2011).

The lineage-specific effects of Ikaros are reflected by the limited overlap shown between targets identified by ChIP in different lineages: only a sub-set of putative targets was found to be common between megakaryocytes and thymocytes, for example.

6 Ikaros in Early Hematopoiesis

Ikaros is a critical self-renewal regulator in multipotent Long-Term Hematopoietic Stem Cells (LT-HSCs) (Papathanasiou et al. 2009). Here, the Polycomb factor Bmi-1 regulates Ikaros expression. Conditional knockout of Bmi-1 allowed extensive up-regulation of Ikaros and concomitant Ikaros-dependent lymphoid lineage priming prior to the subsequent Short-Term Hematopoietic Stem Cell (ST-HSC) stage. Bmi-1 thus appears to underlie the self-renewing and multilineage potential of HSCs and their immediate progeny by controlling Ikaros-dependent lineage specification (Arranz et al. 2012).

In Ikaros-deficient mice, the earliest defects occur in the lymphoid-primed multipotent progenitor (LMPP) stage, which is derived from the ST-HSC compartment. MPPs lacking Ikaros are unable to differentiate into Common Lymphoid Progenitors (CLPs). In LMPPs, specific erythro-megakaryocyte, myeloid and lymphoid genes are poised for activation, with some low-level 'priming' transcription.

As these cells differentiate to become LMPPs, the erythro-megakaryocyte genes are silenced and the myeloid and lymphoid genes remain poised. Loci that remain 'poised' for either future activation or silencing often contain histone modifications characteristic of both chromatin states, referred to as 'bivalent' (Bernstein et al. 2006). In the absence of Ikaros, the early lymphoid signature is mostly erased (the cells lack Flt3 and don't express *Rag1*), whereas the erythro-megakaryocyte signature remains intact, allowing the differentiation to an LMPP that lacks lymphoid potential and retains the ability for myeloid differentiation. Ikaros is important for establishing a lymphoid-specific gene expression program in ST-HSCs and so contributes to their multi-lineage plasticity. Promoting one hematopoietic lineage program while restricting others could be the balancing act played by Ikaros and its associated factors not only in ST-HSCs but also in the series of progressively restricted progenitors that constitute early hematopoiesis (Ng et al. 2007). Following lymphoid lineage specification, Ikaros expression is again increased at the small pre-BI stage in bone marrow, and in double positive (DP) T-cell precursors in the thymus. At these stages of lymphocyte development, precursor cells receive receptor signals and undergo a burst of proliferation followed by further antigen receptor rearrangements and differentiation. When Ikaros activity is reduced in human and mouse, B- and T-cell leukemias arise from the preceding proliferative stages (Yoshida and Georgopoulos 2014).

7 Ikaros and Early B-Cell Development

During B-cell differentiation from the LMPP, cells progress through the CLP stage, becoming pre-pro-B cells, pro-B cells, cycling pre-BI cells and then resting pre-BII cells. Subsequent immature B-cells express Ig heavy and light chains and undergo repertoire selection before becoming mature, recirculating B-cells. Specific gene expression patterns characterize these stages. In pro-B cells, *IgH* (heavy) chain rearrangement allows assembly of the pre-BCR (which consists of IgH and surrogate light chain components λ5 and Vpre-B1), and a burst of proliferation ensues. Ikaros and Aiolos are up-regulated, and the pre-BCR genes such as *Igll1* (which encodes λ5) are then permanently silenced and the cells exit the cell cycle on down-regulation of cell cycle control genes to form pre-BII (small) cells, in which the light chain loci are rearranged (reviewed by Melchers et al., 2000). Successful light chain rearrangement allows assembly of the BCR and subsequent progression to the immature and then mature B cell stages.

Ikaros stringently controls the transition from pro-B to pre-B cell by promoting pre-BCR signalling and cell migration, whilst suppressing cell adhesion. Genome-wide identification of Ikaros target genes by ChIP-Seq and RNA-Seq following conditional inactivation of *Ikzf1* in pro-B cells revealed a plethora of targets involved in cell adhesion and downstream signalling pathways of the pre-BCR in addition to targets involved in pre-BCR signalling. Until recently, the genome-wide integration of Ikaros binding and gene expression in B-cell progenitors had not been achieved owing to the high number of predicted Ikaros consensus binding sites in the genome, the severe developmental disturbances caused by Ikaros deletion and because Ikaros is an inhibitor of cell cycle progression. To overcome these limitations, Ferreirós-Vidal et al. combined Ikaros ChIP-seq with gene expression studies at enhanced temporal resolution to reveal that Ikaros target genes were highly enriched in pre-BCR signalling, cell cycle and V(D)J recombination. Integrating ChIP-seq and gene expression data with gene expression changes during hematopoietic development showed that Ikaros-regulated genes accounted for more than half of genes upregulated in vivo at the transition from CLPs to B-cell progenitors, implying that Ikaros regulates B lineage-associated genes from the earliest stages of B-cell differentiation. Remarkably, increased Ikaros expression was sufficient to drive the majority of gene expression changes associated with progression from cycling to resting pre-B cell stages in vivo. The expression of Ikaros and Aiolos increases incrementally during B-cell progenitor differentiation, and their dosage drives the progression of progenitors along a predetermined lineage by regulating multiple targets in key differentiation pathways (Ferreirós-Vidal et al. 2013).

Loss of Ikaros was shown to lead to a greater abundance of active histone marks (H3K4me3 and H3K9ac) at promoters and distal control elements of strongly repressed Ikaros target genes, whereas significantly fewer of these marks were present at genes activated by Ikaros. The enhanced substrate adhesion and diminished migration of Ikaros-deficient pro-B cells exemplifies a previously unknown function for Ikaros in early B-cell development (Schwickert et al. 2014).

The initial expansion of pre-B cells is driven by IL-7 and pre-BCR signalling. In cycling pre-BI cells, c-Myc is expressed and is subsequently down-regulated in small pre-BII cells, in which expression of Ikaros and Aiolos is elevated. In pre-BII cells, Ikaros and Aiolos directly down-regulate c-Myc expression, which precedes p27 de-repression (Ikaros can bind to the c-Myc promoter to repress expression and can also act to destabilize the p27 protein) and the slower down-regulation of cyclin D3, in an inhibitory network that terminates pre-B cell expansion (Ma et al. 2010). In pre-BI cells, Ikaros modulates λ5 expression by competing with the transcriptional activator EBF. On transition to the pre-BII stage, Aiolos is upregulated though pre-BCR signalling and the adaptor molecule SLP-65, and Aiolos is required for the permanent silencing of the *Igll1* gene. EBF is displaced through increased occupancy of the overlapping Ikaros/Aiolos binding sites on the *Igll1* promoter. Since expression of Aiolos itself is controlled by pre-BCR signalling, Aiolos can be considered to form part of a developmental stage-specific feedback loop linking pre-BRC signalling to the transcriptional silencing of *Igll1*, the gene encoding the pre-BCR component λ5 itself, so limiting the proliferative burst (Thompson et al. 2007). Hence, pre-B cells utilise both intrinsic (Ikaros and Aiolos induction) and extrinsic (attenuation of the IL-7 signal) as mechanisms to terminate their own expansion.

8 Ikaros in T-Cell Development and Function

Ikaros-deficient mice have a complete defect in fetal thymocyte development and T-cell development in adult mice is aberrant, with thymocyte development skewed towards $CD4^+$ T-cells (Wang et al. 1996). Ikaros also regulates T-cell receptor (TCR) signal transduction, and T-cells with reduced Ikaros show enhanced TCR signalling and responsiveness. Ikaros family members are regulated during T-cell activation and proliferation, and Ikaros co-localizes with DNA replication machinery during activation-induced proliferation (Avitahl et al. 1999). Helios is also upregulated at this stage (Akimova et al. 2012). In mature $CD4^+$ T-cells, Ikaros regulates many processes, including Th2 differentiation and expression of the cytokine IL-2 (Bandyopadhyay et al. 2007; Thomas et al. 2007). Ikaros family members Helios and Eos are active in regulatory T-cells (Tregs) and Aiolos is required in Th17 cells (Pan et al. 2009; Getnet et al. 2010). Helios is upregulated in $CD8^+$ T-cells during chronic infection (Doering et al. 2012).

9 T-Cell Progenitors

During αβT-cell development in the thymus, rearrangement of the T-cell receptor (*Tcr*) loci occurs in $CD4^-CD8^-$ double negative (DN) thymocytes, and surface expression of the pre-TCR (including the rearranged β chain) triggers a burst of

proliferation and maturation to the $CD4^+CD8^+$ double positive (DP) stage. Cells then exit the cell cycle and initiate rearrangement of the α-chain, eventually producing a population of $\alpha\beta TCR^+$ DP thymocytes. TCR gene rearrangement requires RAG1 and RAG2 enzymes together with a number of chromatin remodelling factors. Thymic development in mice lacking Ikaros is very inefficient, and mice lacking Ikaros develop thymomas. The Ikaros-null thymoma cell line JE131 has a DN phenotype similar to pre-T-cells, with a surface pre-TCR containing a single functional β-chain and many rearranged TCR α-loci (mainly out of frame). On re-introduction of Ikaros, the cells were able to differentiate into a DP-like population containing many $\alpha\beta TCR^+$ cells. Transcription from the *Tcr* locus rapidly increased, and new RAG-dependent in-frame rearrangements of the α-chain appeared. The process required SWI/SNF remodelling complex, suggesting that Ikaros functions to open the *Tcr* locus setting in motion processed that allow efficient recombination. Knockdown of Mi-2β (part of the NuRD chromatin remodelling complex) enhanced rearrangement, implying that the NuRD complex is antagonistic to the effects of Ikaros in this context (Collins et al. 2013).

Ikaros is able to regulate both the expression and the repression of the *Cd4* gene in a developmental stage-specific manner. A silencer located within the first intron is required for lineage-specific expression. CD4 is not expressed at the DN stage of thymocyte development but is rapidly upregulated at the DN to DP transition. At the $CD8^+$ single positive (SP) stage, *Cd4* expression is permanently silenced. Ikaros binds to the silencer element in a complex containing HDACs at the DN stage, but is able to recruit Mi-2β and histone acetyl transferases (HATs) at the DN to DP transition, resulting in a change in acetylation status at the *Cd4* locus correlating with its activity (Naito et al. 2007). A distinct epigenetic mechanism is used to ensure the permanent silencing of *Cd4* expression in $CD8^+$ SP cells (Zou et al. 2001).

10 Mature T-Cell Function

Mature $CD8^+$ T-cells in human and mouse express detectable levels of all five Ikaros family members. Helios expression is modestly increased in effector and effector memory cells relative to either naïve or memory $CD8^+$ T-cells. Following T-cell activation, Ikaros is involved in regulating cytokine responsiveness, regulating CD25 expression and the competitive ability to survive in pro-inflammatory conditions (e.g. high levels of IL-12), and so is involved in orchestrating $CD8^+$ T-cell immunity in an optimal way (Clamey et al. 2013).

Ikaros is a key regulator of IL-2 expression and sets the co-stimulatory requirement for CD28 in $CD4^+$ T cells. In naïve $CD4^+$ T-cells, the IL-2 locus is incorporated into an inaccessible chromatin structure. Upon activation by TCR and CD28, histones at the promoter and enhancer of the locus become acetylated, leading to opening of the chromatin and accessibility to activating transcription factors, allowing the cells to expand and differentiate. T-cells activated in the absence of

CD28 co-stimulatory signals do not remodel the IL-2 promoter, fail to produce IL-2, and so are rendered hyporesponsive to further stimuli, a state called anergy, which is required for peripheral tolerance. Autoimmune responses can develop where the induction of anergy is compromised.

$CD4^+$ T-cells with reduced Ikaros binding activity no longer require signals from the TCR or CD28 for histone acetylation of the IL-2 promoter, and no longer require CD28 co-stimulation for IL-2 expression. There are two Ikaros binding sites in the IL-2 promoter, which overlap with an NF-kB binding site. Ikaros can form a complex with the histone deacetylases HDAC1 and HCAD2 to maintain hypoacetylation of the IL-2 promoter in quiescent and anergic cells to maintain silencing at the epigenetic level. Since Ikaros can be found at the IL-2 promoter in resting $CD4^+$ T-cells, it has been postulated that Ikaros's ability to bind to the IL-2 promoter to interact with co-factors could be altered following its post-translational modification on co-stimulatory signals from CD28 or IL-2R (Thomas et al. 2014). Phosphorylation is a possibility (Gómez-del Arco et al. 2004).

In contrast to $CD4^+$ T-cells, signals from the TCR and CD28 are not sufficient to induce autocrine IL-2 production in naïve $CD8^+$ T-cells, which require cytokines provided by other cell types to drive their differentiation. Ikaros, which silences IL-2 in anergic $CD4^+$ T-cells, also restricts IL-2 production in $CD8^+$ T-cells. $CD8^+$ T-cell activation in vitro in the absence of exogenous cytokines and $CD4^+$ help results in increased Ikaros, which represses the IL-2 gene. Cells haploinsufficient for Ikaros fail to up-regulate Ikaros, produce IL-2 and differentiate in an IL-2-dependent manner into Interferon γ-producing Cytotoxic T-Cells (CTLs) in response to TCR/CD28 stimulation alone. IL-2 during the priming phase is required for effective T cell memory formation, as 'unhelped' $CD8^+$ T-cells fail to generate memory. High levels of IL-2, however, can promote terminal effector $CD8^+$ T-cell generation at the expense of memory formation, and a balance is required. By restricting autocrine IL-2 production by $CD8^+$ T-cells, Ikaros ensures that induction of an inflammatory and cyctotoxic program only occurs in cells that have been appropriately licensed by a third signal. The repressive activity of Ikaros renders $CD8^+$ T-cells dependent on environmental cues such as IL-12 from dendritic cells or IL-2 from activated $CD4^+$ T-cells to license their differentiation into effector CTLs. In this way, Ikaros normally functions to protect naïve $CD8^+$ T-cells from a loss of self-tolerance during infection or other inflammatory situations. There is a T-cell-specific enhancer region upstream of the Ikaros genes that could be involved in its inducible expression in $CD8^+$ T-cells (O'Brien et al. 2014).

Ikaros also directly regulates immunoglobulin class-switch recombination (CSR) in mature T-cells (Sellars et al. 2009).

11 Ikaros and the Immune Response

T regulatory cells (Tregs) maintain immune tolerance, limiting autoimmunity and chronic inflammation by inhibiting activation-induced proliferation and cytokine expression. Several sub-types of Tregs have been described based on their mechanisms of generation. Tregs express high levels of the IL-2α-receptor chain (CD25) and are strongly dependent on the cytokine IL-2. When T-cells are activated in the presence of Tregs, NFAT-dependent Ikaros expression is upregulated. Th1 cells transfected with a plasmid expressing dominant negative Ikaros mimicked NFAT-deficient T-cells and became more refractory to Tregs, showing that the activity of Ikaros in $CD4^+$ T-cells is required for efficient Treg-mediated suppression. Hence Ikaros is a regulator of suppression that is required for optimal inhibition of T-cell effector function (Shin et al. 2014).

Other Ikaros family members are involved in different aspects of Treg function. Whether the Treg lineage has the ability to be suppressive or not depends on up- or down-regulation of the Treg lineage-specific marker Forkhead box P3 transcription factor (FoxP3). Helios (which is expressed in both thymic and peripherally-induced Tregs) can bind to the *foxp3* promoter and is able to upregulate FoxP3 expression, whereas SATB1 is able to inhibit its expression. RNAi-depletion of Helios in Tregs resulted in downregulation of FoxP3, highlighting the importance of Helios in maintaining the function of Tregs as immune suppressors (Grzanka et al. 2013). Foxp3 associates with another Ikaros family member, Eos, which together with Tip60 and HDAC7 mediates the gene repression profile component of the Treg signature (Pan et al. 2009). A study involving a Foxp3-GFP reporter gene that was unable to interact with Eos revealed that in the absence of the interaction between Eos and Foxp3, there was an increase in IL-2 production in the cells, with increased acetylation of the IL-2 promoter, as well as upregulation of some non-Treg-associated genes (Bettini et al. 2012).

CD4+ IL-17 producing helper T-cells (T_H17 cells) are required for immune response to pathogens. Overactive T_H17 response results in tissue inflammation and autoimmunity, and so their development must be controlled. IL-2 suppresses their development, and so IL-2 production has to be suppressed to allow their expansion. In T_H17-polarizing conditions, transcription factors STAT3 and AhR upregulate the expression of the Ikaros family member Aiolos, which silences IL-2 production through promoting the formation of silent chromatin at the locus (Quintana et al. 2012).

12 Ikaros and Notch

Both Ikaros and Notch proteins were first identified for their essential functions in lymphocyte development as well as their mis-regulation in lymphoid malignancies. The Notch signalling pathway is active in many cell types, but its effects are

lineage- and stage-specific. Notch activity is required for the differentiation of T-cell progenitors, and activity is reduced in more mature thymocytes. Deregulation of Notch and Ikaros co-operate to promote T-cell leukemogenesis, and Ikaros is involved in the de-repression (rather than direct activation) of some Notch target genes. Ikaros has been reported to directly repress known Notch target genes such as Hairy Enhancer of Split 1 (*Hes1*), for example (Kathrein et al. 2008). Since Ikaros and the Notch transcriptional mediator RBP-Jk compete for two elements on the *Hes1* promoter, up-regulation of Ikaros in DN4 ($CD4^-/CD8^-/CD3^-$) thymocytes leads to increased occupancy of the *Hes1* promoter by Ikaros and subsequent repression, reducing thymocyte responsiveness to Notch signalling during the $CD4^-/CD8^-$ double negative (DN) to $CD4^+/CD8^+$ double positive (DP) transition (Kleinmann et al. 2008). More recent genome-wide analysis has shown that Ikaros is able to bind to and directly repress most genes that are activated by Notch, and loss of Ikaros leads to deregulated Notch signaling. Ikaros shapes the timing and repertoire of the Notch transcriptional response in T-cells through widespread targeting of elements adjacent to Notch regulatory sequences (Geimer Le Lay et al. 2014). Some Notch target genes (*Fjx1*, *Hey1* and *Dtx1* for example) that are over-expressed in Ikaros-deficient DP thymocytes (a population not known to receive strong Notch signalling) may be de-repressed through the generally reduced H3K27me3 associated with the loss of Ikaros, i.e. the more open chromatin structure permitted by lack of Ikaros could allow activation of these genes by other factors. Ikaros may require an appropriate chromatin environment to exert its repressive effects (Zhang et al. 2011).

Ikaros null ($Ikzf1^{-/-}$) mice show mild anaemia and reduced platelet count (Lopez et al 2002). Recent work has implicated Ikaros and Notch signaling in the regulation of myeloid, erythroid and megakaryocye differentiation. Ikaros expression was reported to be high in hematopoietic progenitors and low in progenitors committed to megakaryocyte or erythroid differentiation (Yoshida et al. 2006). Ikaros over-expression was found to reduce NOTCH-induced megakaryocytic specification (Malinge et al. 2013). Ikaros was found to have a broad impact on the megakaryocyte transcriptional network, including many Notch-independent effects. Together with other regulators of megakaryopoiesis, GATA1 expression (and consequently expression of key GATA1 target genes) was down-regulated by Ikaros and reciprocally, the abundance of Ikaros transcripts was reduced by GATA1 expression, and ChIP showed that GATA1 can bind to the *Ikzf1* locus. Hence, Ikaros represses selected Notch targets and megakaryocytic genes, and is itself turned off by GATA factors on terminal differentiation. Down-regulation of Ikaros on flicking this 'GATA switch' could then release repression of targets required for the terminal differentiation of megakaryocytes (Maillard 2013; Malinge et al. 2013). This mechanistic scenario is specific to the megakaryocytic lineage, since in the erythroid lineage, GATA1 and Ikaros co-operate.

Ikaros was found to be upregulated during erythroid differentiation of human $CD34^+$ cells (Dijon et al. 2008) and Ikaros can regulate the expression or activity of erythroid-specific genes, including *Gata1*, *EpoR* and *Nfe2*. It is also involved in the control of globin switching (described below). The Notch target gene *Hes1* is transiently repressed at the pro-erythroblast stage of erythroid differentiation in

fetal liver. RNAi-mediated reduction of *Hes1* was found to enhance erythroid differentiation, and decreased erythroid differentiation and *Hes1* up-regulation has been observed in Ikaros-deficient mice, suggesting that *Hes1* normally opposes terminal differentiation of erythroid cells. Ikaros facilitates the binding of GATA1 and its co-factor FOG1 to repress the *Hes1* locus at the pro-erythroblast stage when GATA1 expression is high. EZH2, a member of the polycomb repressive complex 2 (PRC2), is also required for this transient repression of *Hes1*. Repression correlates with increased H3K27me3 at the locus following recruitment of PRC2 by GATA1 and FOG1.

13 Ikaros and the Three-Dimensional Organization of Chromatin

The multi-genic human β-globin locus has been extensively studied as a paradigm for regulated developmental stage-specific gene expression. The locus contains five genes (ε, $^{G}\gamma$, $^{A}\gamma$, δ and β) and a distant upstream Locus Control Region (LCR), which regulates globin gene expression in erythroid cells. The LCR contains five DNAse1 hypersensitive sites (HSSs), which are rich in transcription factor binding sites. The locus undergoes two switches in gene expression during development: embryonic to fetal early in fetal life and then fetal to adult (γ- to β-) just after birth. The specificity of gene expression at each stage depends on long-range contacts between the LCR and the promoter of the expressed gene, such that the expression of each gene correlates with a distinct three-dimensional structure of the locus. Ikaros is involved in the regulation of globin switching. In Ikaros null mice (with a deletion in exon 8 that leads to lack Ikaros protein owing to protein instability), γ- to β-globin switching is delayed, and in Ikaros^{Plastic} mice (where Ikaros is unable to bind DNA owing to a mutation in zinc finger 3) embryonic erythroblast growth and differentiation fail (Papathanasiou et al. 2003). Ikaros is able to occupy sites in HSS3 of the LCR, as well as interacting with the promoters of the γ- and δ-globin genes, and so can alter long-range chromatin interactions and hence the three-dimensional structure of the locus. Ikaros recruits HDACs and the chromatin-remodelling factor Mi-2β to the γ-globin promoter region at the time of the γ- to β-globin switch, and in addition, the interaction between Ikaros and GATA1 alters the binding affinity of GATA1 at the γ-globin promoter, impairing contact between the γ-globin promoter and the LCR, resulting in down-regulation of γ-globin expression (Bottardi et al. 2009).

A further way in which Ikaros is able to regulate heritable silencing during development is through control of the global location of developmentally relevant genes within the nucleus. On detection of Ikaros protein together with pericentromeric γ-satellite DNA and locus-specific probes in cycling murine lymphocytes, a significant fraction of Ikaros was found to associate with clustered centromeric DNA (Brown et al. 1997). Ikaros was able to recruit silenced genes to

these nuclear compartments though dimerization, where one molecule bound to the target gene and the other bound specifically to the centromeric γ-satellite DNA (Cobb et al. 2000). For example, RAG (*Rag1*) and TdT (*Dntt*) loci were found to relocate to pericentromeric heterochromatin following heritable gene silencing in primary $CD4^+CD8^+$ thymocytes, but not in the phenotypically similar cell line VL3-3M2, where silencing occurred but was not heritable (Brown et al. 1999). The D element of the TdT promoter contains overlapping binding sites for both Ikaros and an Ets transcriptional activator, which compete for occupancy. Ikaros is required for repression of TdT, during which it recruits chromatin-modifying machinery to create silent chromatin at the locus. These changes are followed by repositioning of the locus to pericentromeric heterochromatin only when silencing is permanent, suggesting a model where pericentromeric regions, rich in repressive factors, enhance the silent state of loci repositioned there (Trinh et al. 2001; Su et al. 2005), as illustrated in Fig. 2. Other Ikaros family members are able to associate with Ikaros at pericentromeric heterochromatin. During B cell development, repositioning of the λ5 locus (*Igll1*) occurs following the pre-BI to pre-BII transition and correlates with its permanent inactivation. Similarly for the *Dntt* locus, the *Igll1* promoter contains overlapping binding sites for Ikaros and an activating factor EBF, and silencing ensues after the Ikaros family 'win' the competition for promoter binding on up-regulation of Aiolos (Thompson et al. 2007).

Since post-translational phosphorylation of Ikaros following signalling events is able to modify its DNA-binding affinity. Reversible phosphorylation of Ikaros by CK2 is able to alter its binding affinity for both the *Dntt* promoter and its ability to bind to pericentromeric heterochromatin, affecting its regulation of permanent gene silencing (Gurel et al. 2008).

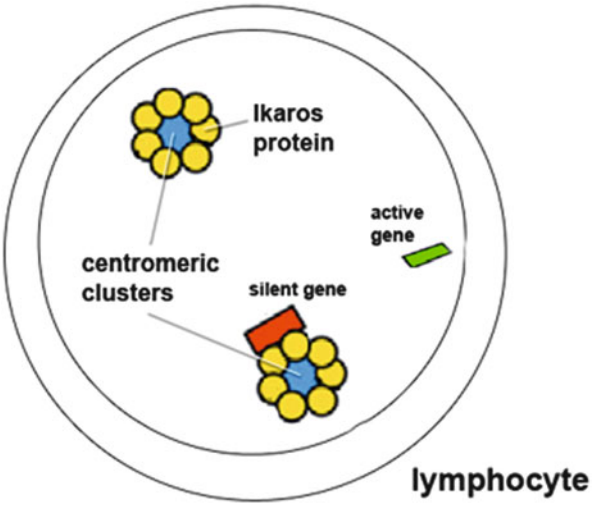

Fig. 2 Schematic illustration of Ikaros foci around clustered centromeres in mouse lymphocytes, where silent but not active genes can be found. After Merkenschlager (2010)

14 Role of Ikaros Family Members in Non-hematopoietc Tissues

The Ikaros family has been studied extensively in hematopoietic cells, but its roles in other tissues provide further insight into its mechanistic diversity (Ezzat and Asa 2008). In some examples, family members act through directly binding to Ikaros consensus sequences or by replacing it in complexes that usually contain Ikaros when protein levels of the family member are upregulated. Ikaros family members can regulate the transcriptional programs in tissues that don't express Ikaros. Eos (*Ikzf4*), for example, is expressed in skeletal muscle, heart, brain and liver. In muscle cells it regulates expression of the myosin gene Myh7b, which encodes a micro-RNA miR-499 involved in skeletal muscle biology. Eos bins to an upstream Ikaros consensus site and physical interaction between Eos and MyoD (a muscle-specific transcription factor) produces an activating complex leading to expression of the gene. In addition, Eos indirectly regulates additional myosin genes that depend on Sox6 expression, since miR-499 regulates Sox6 expression (Yeung et al. 2012).

Helios (*Ikzf2*) is expressed in the developing murine brain, in a dynamic spatio-temporal pattern between E14.5 and E18.5, though it is absent after birth. It takes part in the transcriptional hierarchy that drives striatal neuron development. Although Ikaros and Helios are co-expressed in a subpopulation of olfactory bulb cells, immunocytochemisty shows no co-localization, indicating that the Ikaros and Helios act on different target genes. As with other family members, Helios is involved in the regulation of lineage choice (Martín-Ibáñez et al. 2012).

Pegasus (*Ikzf5*) is the most divergent Ikaros family member. It has three N-terminal zinc fingers that bind to the 'atypical' consensus sequence GNNNGNNG. It can dimerize but doesn't interact with the other Ikaros family members. Pegasus is broadly expressed in diverse adult organs, including brain, heart, skeletal muscle, kidney and liver. Knockdown of the single Pegasus orthologue in Zebrafish using morpholino technology revealed a function in regulating left-right asymmetry of the embryo. Pegasus may directly regulate the gene *pitx2* (a bicoid-type homeodomain transcription factor), as well as other genes involved in left-right asymmetry including *lefty2* and *spaw* (John et al. 2013). Pegasus may retain the more archaic function by acting as a more global developmental regulator (John et al. 2009).

15 Regulation of Ikaros Activity

Ikaros has its zinc fingers in many pies, so the way in which its activity is regulated has a major impact on modulating its functions during lineage specification. Post-translational modification of Ikaros can influence its DNA-binding affinity, stability, nuclear localization and ability to interact with co-factors, perhaps through conformational change. It is tempting to speculate that such regulation of Ikaros function could translate different levels of receptor signalling into different levels of Ikaros activity.

Bruton's Tyrosine Kinase (BTK) is a partner and post-translational regulator of Ikaros. BTK phosphorylates Ikaros at S214 and S215 near zinc finger 4 to augment its nuclear localization and DNA-binding capacity. Spleen Tyrosine Kinase (SYK) phosphorylates Ikaros at S358 and S361 at the C-terminus, and is also required for optimal Ikaros function. SYK and BTK co-operate via the B-cell-specific adapter molecule SLP-65, which is controlled by pre-BCR signaling. In the presence of SLP-65, SYK directly activates BTK by phosphorylating its Y551 residue in the activation loop of the BTK catalytic domain (Ma et al. 2013). IN BTK-deficient cells, Ikaros is cytoplasmic rather than focussed at pericentromeric heterochromatin. Casein Kinase 2 (CK2) can phosphorylate Ikaros at eleven sites in a serine/threonine-rich region near the C-terminal end of the protein. Regulation of Ikaros by CK2 has been implicated in regulating Ikaros's ability to negatively regulate the G1-S transition in cycling cells. Ikaros is normally de-phosphorylated in late G1 phase of the cell cycle, and is phosphorylated again during S-phase, so de-phosphorylation of Ikaros could act as a checkpoint in transformation. Changes in the phosphorylation status of Ikaros can affect its DNA-binding capacity and hence affect its regulation of cell cycle progression-promoting genes (Martín-Ibáñez et al. 2010, for example). Over-expression of CK2 leads to T cell leukaemia (Gómez-del Arco et al. 2004). Phosphorylation of Ikaros in the linker region of the DNA-binding zinc fingers occurs during metaphase, reducing its DNA-binding affinity and so removing it from mitotic chromosomes during early mitosis (Dovat et al. 2002). The tyrosine phosphorylation of Aiolos following IL-2 stimulation leads to its movement from the cytoplasm to the nucleus, hence regulating its access to target genes (Romero et al. 1999).

The de-phosphorylation of Ikaros by Protein Phosphatase 1 (PP1) can also influence Ikaros's stability and nuclear localization. Ikaros interacts with PP1 using its C-terminal domain, and point mutations in the PP1 binding site reduce the stability of the Ikaros protein and reduce its ability to bind to γ-satellite DNA probes and to pericentromeric heterochromatin. In vivo, Ikaros can target a fraction of cytoplasmic PP1 to the nucleus and recruit it to pericentromeric heterochromatin, where it may affect genes that have been localized there (Popescu et al. 2009). Ikaros undergoes de-phosphorylation during the induction of thymocyte differentiation, at which time Ikaros's affinity for the *Dntt* promoter is increased. De-phosphorylation of Ikaros at amino acids 13 and 294 occurs on PMA and ionomycin treatment of the thymocyte cell line VL3-3 M2 correlating with an

increase in Ikaros's affinity for the D9 regulatory region of the *Dntt* promoter, displacement of the activating transcription factor Elf-1 and subsequent repression of TdT (Gurel et al. 2008). The balance between phosphorylation and de-phosphorylation of a significant proportion of Ikaros in a cell could permit external cues to regulate its function (Song et al. 2011).

Ikaros-mediated gene repression can also be modulated by Ikaros SUMOylation, which can disrupt its participation in both HDAC-dependent and HDAC-independent repression. SUMOylation doesn't affect Ikaros's ability to bind to pericentromeric heterochromatin. Ikaros can be SUMOylated on K58 and K240 in the N-terminal half of the protein, which may affect its conformation, and there are corresponding conserved sites on Aiolos. SUMOylation disrupts Ikaros's interaction with Sin3A, Sin3B, Mi-2β and CtBP, whereas interaction with SWI/SNF complexes is less severely affected. Since a small but significant proportion of Ikaros is SUMOylated, SUMOylation may be a mechanism to initiate disassembly of repressive complexes that contain Ikaros. In this way, post-translational modification is able to switch the function of Ikaros from repression while its activation function is maintained (Gómez-del Arco et al. 2005).

The way in which Ikaros is made to act in the right way at the right time and place is key to understanding how it influences the progression of lymphocyte development (Liberg et al. 2003). Understanding how the binding capacity of Ikaros to both DNA and co-factors is affected by post-translational modification is required in order to understand how Ikaros activity can be rapidly modulated as signalling environments change during differentiation. An understanding the regulatory networks that operate both up- and down-stream of Ikaros is also important for an appreciation of the functional influence of the Ikaros family in both normal and malignant hematopoiesis (Yoshida and Georgopoulos 2014).

16 Ikaros and Leukemogenesis

Ikaros functions as a tumour suppressor (reviewed by Kastner et al. 2013). Mice heterozygous for a dominant negative mutation of *Ikzf1*, or homozygous for a knock-down *Ikzf1* mutation develop T-cell lymphomas and leukemias with 100 % penetrance (Winandy et al. 1995). In humans, Ikaros mutations occur in only 5 % of T-cell acute lymphoblastic leukemias (T-ALL) but surprisingly, are much more frequent in pediatric and adult B-cell acute lymphoblastic leukemias (B-ALL). The human Ikaros locus (*IKZF1*) spans over 120 kb on chromosome 7p12.2, and comprises 8 exons. There are three types of mutation: (i) deletions resulting in haploinsufficiency and reduction in Ikaros protein level, (ii) genomic deletions or splicing mutations resulting in the production of dominant negative forms of Ikaros, which lack DNA binding capacity but can dimerize with endogenous Ikaros family members to sequester them away from their genomic target genes or interfere with co-factor interactions, and (iii) mutations that perturb Ikaros function. Of the more common genomic mutations, deletion of exons 4–7 leads to abundant production of

Ik6, which is localized in the cytoplasm because it lacks a nuclear localization sequence. Deletion of exons 2–7 is likely to result in an Ikaros null effect owing to lack of a translational start site (Wang et al. 2014).

Ikaros mutations are significantly associated with Ph^+ ALLs and Chronic Myeloid Leukemia (CML) involving the Philadelphia translocation t(9;22)(q34;q11) which produces a fusion protein between BCR and ABL1, a tyrosine kinase. Translocation removes the N-terminal end of ABL1 rendering it constitutively active and able to aberrantly phosphorylate many targets, including signalling pathway components. Ikaros is BCR-ABL1's 'most frequent partner in crime', and over 80 % of Ph^+ ALLs have somatic Ikaros mutations, and they correlate with poor clinical outcome and probability of relapse (reviewed by Bernt and Hunger 2014). Ikaros mutations or deletions have been associated with 66 % of CML blast crisis (where the presence of both myeloid and lymphoid cell types indicates a defect at an early stage of hematopoietic differentiation). Ikaros mutations occur in 33 % of Ph^- ALLs, and inherited genetic variations within the Ikaros locus are associated with high risk of pediatric ALL (Payne and Dovat 2011). Several primary tumour cells and animal leukemic cell lines show either Ikaros haploinsufficiency or the over-expression of dominant negative isoforms such as Ik6, Ik10 and Ik11. Ik11 can interact with and sequester Helios in the cytoplasm, for example (Kastner et al. 2013; Wang et al. 2014; Capece et al. 2013).

The way in which Ikaros facilitates the development of leukemias is controversial. Since it is involved in regulating hematopoiesis at early stages, mis-regulation of Ikaros may lead to the retention of stem cell-like gene expression patterns that favour self-renewal and drug resistance, favouring survival of malignant clones. Since Ikaros is also a key regulator of later stages of B- and T-cell development that involve proliferative bursts, blockage of differentiation at such stages by defective Ikaros could lead to continued proliferation of precursor cells susceptible to malignant transformation (reviewed by Georgopoulos 2009). A recent example describes loss of Ikaros at the pre-B cell stage arrested the pre-B cells in the integrin-dependent, adherent, self-renewing, pro-proliferative stage of development, promoting malignant transformation (Joshi et al. 2014).

In a human case of severe bone marrow pancytopenia, the phenotype of a newborn closely resembled the phenotype of Ikaros null mice, and so Ikaros was analysed and found to have a point mutation in the zinc finger 4 in the N-terminal DNA-binding region. Although still able to bind to some specific target genes (VPAC1, for example), the mutant form of Ikaros was found by immunocytochemistry to mis-localize within the nucleus away from pericentromeric heterochromatin (in contrast to the distribution seen in the patient's unaffected sister and in control cells), perhaps demonstrating the importance of nuclear compartmentalization of Ikaros for its function. It is also possible that the mutation could skew DNA-binding specificity and so compromise the regulation of other Ikaros targets (Goldman et al. 2012).

The over-expression of dominant negative Ik6 can lead to transformation of non-hematopoietc tissues. Ikaros interacts with C-terminal Binding Protein (CtBP) (Mandel and Grosschedl 2010) in the pituitary to modulate growth, and its

replacement by dominant negative Ik6 occurs in murine pituitary tumours, for example (Dorman et al. 2012). Aiolos is expressed in lung carcinoma, where one of its mechanisms of action is the repression of genes involved in cell-cell and cell-matrix adhesion (Frisch and Schaller 2014).

The question of whether Ikaros mutations actively drive malignant transformation or whether they are passengers resulting from increased genomic instability in malignant cells remains controversial. Many studies have not included quantitative measurement of Ikaros protein levels, and identification of target genes consistently and significantly affected by Ikaros mutation has proved challenging (Qazi and Uckun 2013; Uckun et al. 2013), and so its value as a prognostic marker has been questioned. Analysis is further complicated by the correlation of additional mutations affecting other developmentally important factors (e.g. Pax5 mutations in many B-ALLs) or signalling pathways (e.g. the Notch pathway in T-ALLs) such that a range of defects and co-operativity with Ikaros mutations could contribute to leukemogenesis.

In a serendipitous way it was found that thalidomide and related drugs lenalidomide and pomalidomide were effective in the treatment of Multiple Myeloma (MM). Initially used as a sedative, thalidomide's severe teratogenic effects led to its abandonment. Since it has anti-inflammatory properties, it was tried as a treatment for MM and in combination with dexamethasone, was found to be effective in 60 % of patients. Related lenalidomide was effective in 80 % of patients treated. The drugs directly affect the proliferation of MM cells in culture, through upregulation of p21 and induction of apoptosis, while in patients, secretion of IL-2 from T-cells and tumour cell killing by natural killer (NK) cells was enhanced (Licht et al. 2014). Using a thalidomide derivative to purify interacting factors, Cereblon (CRBN) and DNA Damage Binding Protein 1 (DDB1) were identified, which together with CUL4 and ROC1, are components of the E3 ubiquitin ligase complex CRL4. When the drugs bound CRBN, the target specificity of the complex was altered, resulting in increased ubiqiuitinylation and subsequent degradation of Ikaros and Aiolos proteins, which are essential for MM cell growth and repression of IL-2 transcription (Gandhi et al. 2013; Krönke et al. 2014; Lu et al. 2013). The abundance of Helios, Eos and Pegasus was not affected, since they lack the 30 amino acid domain used for CRBN interaction that is common to Ikaros and Aiolos. Mutation to create the domain in Eos allowed its degradation, showing that the 30 amino acid domain was necessary and sufficient for interaction with CRBN. The drugs were also found to be effective in mantle cell lymphoma and CLL, where Aiolos expression is high (Krönke et al. 2014).

17 Evolution of the Ikaros Family

Evolutionary analysis suggests that the Ikaros family is derived from a primordial gene related to the proteosome Hunchback family. Orthologues of the five mammalian Ikaros genes occur in other tetrapod lineages including amphibians (Hansen

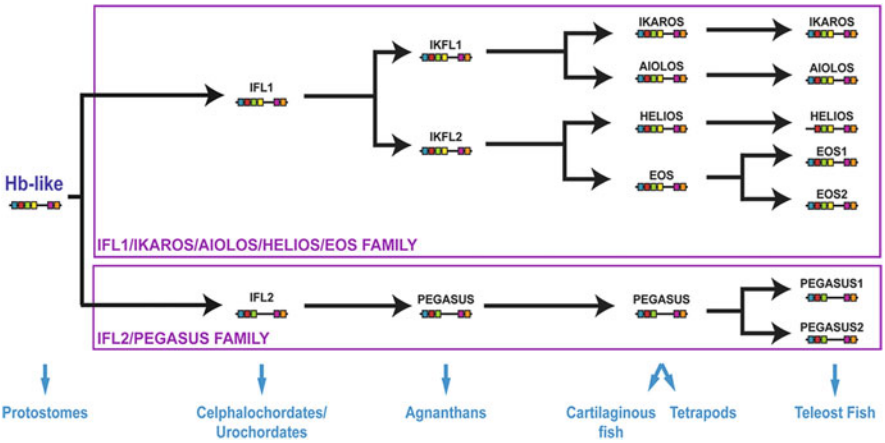

Fig. 3 The evolution of the Ikaros family. Duplication of a primordial Hunchback-like gene prior to cephalochordate and urochordate divergence produced two lineages, one related to IFL1, Ikaros, Aiolos, Helios and Eos, and another related to IFL2 and Pegasus (purple boxes). Further duplication of the IFL1 lineage before agnathan divergence produced precursors to the present day IKLF1 and IKLF2, followed by a second duplication that generated individual Ikaros members in cartilaginous fish and tetrapods: Ikaros and Aiolos from the IKFL1 lineage, and Helios and Eos from the IKFL2 lineage. The IFL2 lineage has produced only Pegasus (except in teleosts, which have two paralogues for Pegasus and Eos). After John and Ward (2011)

et al. 1997), birds (Kohonen et al. 2004) and cartilaginous fish (Haire et al. 2000), and the role of Ikaros in higher vertebrates appears to be conserved. Ikaros-factor-like (IFL or IKFL) genes have been identified in the more primitive urochordates and agnathans. IFL1 genes in urochordates and lower vertebrates cluster with Ikaros, Aiolos, Helios and Eos, while IFL2 genes cluster with Pegasus, a scheme shown in Fig. 3 (John and Ward 2011). It was further suggested that the ancestral Ikaros gene was duplicated prior to the divergence of lower and higher vertebrates (John et al. 2009). The generation of Ikaros, Aiolos, Helios and Eos appears to coincide with the development of adaptive immunity 685–891 million years ago, perhaps providing the additional regulatory control needed to generate the diverse cell populations that constitute the adaptive immune system (Liongue et al. 2011). Hunchback is able to both repress and activate gene expression in *Drosophila* (Zuo et al. 1991), suggesting a conserved mechanism of action and perhaps helping to explain some of the non-hematopoietic functions of the Ikaros family (John and Ward 2011).

18 Future Challenges

Much progress has been made in understanding the way in which transcriptional regulation by the Ikaros family of transcription factors impacts on mammalian hematopoiesis, lineage commitment, immune response and disease. Ikaros interacts

with signaling networks, co-factors, chromatin structure and nuclear organization in order to regulate important developmental and immunological events. Despite the wealth of data that has recently been generated using sophisticated genomics and proteomics methodologies, barriers to further progress remain. For example, many studies focus on narrowly defined events and genes that are directly regulated are difficult to distinguish from indirectly regulated targets. Functionally important protein-protein interactions may occur at low affinity and so be difficult to distinguish from functionally irrelevant interactions. Redundancy between related family members can mask biologically important functions. In leukemogenesis, additional mutations occur in combination with Ikaros family mutations, making mechanistic interpretation difficult. In addition, cause and consequence can be difficult to distinguish when changes in gene expression correlate with changes in chromatin structure and/or nuclear organization. The recent development of RNA-seq at population and single cell level will enhance our understanding of the transcriptional circuitry, and development of fast and easy genome editing techniques including the CRISPR/Cas9 system (Hsu et al. 2014) should further enhance our understanding of the way in which modular transcription factors such as Ikaros function. Improvements in high-resolution microscopy in single live cells will allow further insight into the role of nuclear organization in transcriptional regulation. Together with conventional biochemical and molecular approaches, these techniques will help to reveal the logic though which individual genes and groups of genes are regulated in key immunological settings, as well as the mechanisms by which immune responses are co-ordinated by interactions between individual cells and between populations and tissues (Smale 2014). Small changes in the levels of Ikaros family members and/or the factors with which they compete for binding at specific promoters could tip the balance towards or away from transcription of factors that profoundly influence downstream signalling pathways, for example, allowing small changes in the relative levels of Ikaros proteins to result in major changes in outcome, perhaps helping to partly explain the fact that mutations that do not massively reduce the levels of Ikaros in a cell can correlate with malignant transformation. Since Ikaros can be viewed as sitting at the centre of a web of interactions, the way in which evolutionary changes in one area of the web are felt at other areas of the web imply co-evolution of related sets of factors.

In conclusion, future identification of Ikaros family target genes and elucidation of the way in which Ikaros function is controlled at multiple levels should throw more light on how Ikaros drives developmental progression and how changes in its dosage can reconfigure transcriptional and signalling networks in lymphocytes and other cell types in which it is expressed. Identification of the mechanisms of Ikaros's diverse functions should lead to a greater understanding of the functional connectivity involved in the progression of development.

References

Akimova T, Beier UH, Wang L, Levine MH, Hancock WW (2012) Helios expression is a marker of T cell activation and proliferation. PLoS One 6, e24226

Alinikula J, Kohonen P, Nera K-P, Lassila O (2010) Concerted action of Helios and Ikaros controls the inositol 5-phosphatase SHIP. Eur J Immunol 40:2599–2607

Arranz L, Herrera-Merchan A, Ligos JM, de Molina A, Dominguez O, Gonzalez S (2012) Cell Cycle 11(1):65–78

Avitahl N, Winandy S, Friedrich C, Jones B, Ge Y, Georgopoulos K (1999) Ikaros sets thresholds for T cell activation and regulates chromosome propogation. Immunity 10:333–343

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Werning M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL, Lander ES (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125 (2):315–326

Bernt KM, Hunger SP (2014) Current concepts in pediatric Philadelphia chromosome-positive acute lymphoblastic leukemia. Front Oncol 4:1–21

Bettini ML, Pan F, Bettini M, Finkelstein D, Rehg JE, Floess S, Bell BD, Zeigler SF, Huehn J, Pardoll DM, Vignali DAA (2012) Loss of epigenetic modification driven by the Foxp3 transcription factor leads to regulatory T cell insufficiency. Immunity 36(5):717–730

Bottardi S, Ross J, Bourgoin V, Fotouhi-Ardakani N, Affar EB, Trudel M, Milot E (2009) Ikaros and GATA-1 combinatorial effect is required for silencing of human γ-globin genes. Mol Cell Biol 29(6):1526–1537

Bottardi S, Mavoungou L, Bourgoin V, Mashtalir N, el Affar B, Milot E (2013) Direct protein interactions are responsible for Ikaros-GATA and Ikaros-Cdk9 cooperativeness in hematopoietic cells. Mol Cell Biol 33(16):3064–3076

Brown KE, Guest SS, Smale ST, Hahm K, Merkenschlager M, Fisher AG (1997) Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. Cell 91:845–854

Brown KE, Baxter J, Graf D, Merkenschlager M, Fisher AG (1999) Dynamic repositioning of genes in the nucleus of lymphocytes preparing for cell division. Mol Cell 3(2):207–217

Capece D, Zazzeroni F, Manacarelli MM, Verzella D, Fischietti M, Di Thommaso A, Maccarone R, Plebani S, De Ianni M, Gulino A, Alesse E (2013) A novel, non-canonical splice variant of the Ikaros gene is aberrantly expressed in B-cell lymphoproliferative disorders. PLoS One 8(7), e68080

Clamey ET, Collins B, Young MH, Eberlein J, David A, Kappler JW, Marrak P (2013) The Ikaros transcription factor regulates responsiveness to IL-12 and expression of IL-2 receptor alpha in mature, activated CD8 T cells. PLoS One 8(2), e57435

Cléard F, Karch F, Maeda RK (2014) DamID as an approach to studying long-distance chromatin interactions. Methods Mol Biol 1196:279–289

Cobb BS, Morales-Alcelay S, Kleiger K, Brown KE, Fisher AG, Smale ST (2000) Targeting of Ikaros to pericentromeric heterochromatin by direct DNA binding. Genes Dev 14 (17):2146–2160

Davis KL (2011) Ikaros: master of hematopoiesis, agent of leukemia. Ther Adv Hematol 2 (6):359–368

Doering TA, Crawford A, Angelosanto JM, Paley MA, Ziegler CG, Wherry EJ (2012) Network analysis reveals centrally connected genes and pathways involved in CD8(+) T cell exhaustion and memory. Immunity 37:1130–1144

Dorman K, Shen Z, Yang C, Ezzat S, Asa SL (2012) CtBP1 interacts with Ikaros and modulates pituitary tumor cell survival and response to hypoxia. Mol Endocrinol 26(3):447–457

Dostie J, Bickmore WA (2012) Chromosome organization in the nucleus: Charting new territory across the Hi-Cs. Curr Opin Genet Dev 22(2):125–131
Dovat S (2011) Ikaros isoforms: the saga continues. World J Biol Chem 2(6):140–145
Dovat S, Ronni T, Russell R, Ferrini B, Cobb BS, Smale ST (2002) A common mechanism for mitotic inactivation of C2H2 zinc finger DNA-binding domains. Genes Dev 16:2985–2990
Ezzat S, Asa SL (2008) The emerging role of the Ikaros stem cell factor in the neuroendocrine system. J Molec Endocrinol 41:45–51
Ferreirós-Vidal I, CarrollT TB, Terry A, Liang Z, Bruno L, Dharmalingam G, Khadayate S, Cobb BS, Smale ST, Spivakov M, Srivastava P, Petretto E, Fisher AG, Merkenschlager M (2013) Genome-wide identification of Ikaros targets elucidates its contribution to mouse B-cell lineage specification and pre-B-cell differentiation. Blood 121(10):1769–1782
Frisch SM, Schaller MD (2014) The wind God promotes lung cancer. Cancer Cell 25:551–552
Gandhi AK, Kang J, Havens CG, Conklin T, Ning Y, Wu L, Ito T, Ando H, Waldman MF, Thakurta A, Kippel A, Handa H, Daniel TO, Schafer PH, Chopra R (2013) Immunomodulatory agents lenalidomide and pomalidomide co-stimulate T cells by inducing degradation of T cell repressors Ikaros and Aiolos via modulation of the E3 ubiquitin ligase complex CRL4 (CRBN). Br J Hematol 164(6):811–821
Geimer Le Lay A-S, Oravecz A, Mastio J, Jung C, Marchal P, Ebel C, Dembélé D, Jost B, Le Gras S, Thibault C, Borggrefe T, Kastner P, Chan S (2014) The tumour suppressor Ikaros shapes the repertoire of Notch target genes in T cells. Sci Signal 7(317):ra28.doi:10.1126/scisignal.2004545.
Georgolpoulos K, Moore DD, Derfler B (1992) Ikaros, an early lymphoid-specific transcription factor and a putative mediator for T cell commitment. Science 258:808–812
Georgopoulos K (2009) Acute lymphoblastic leukemia-on the wings of Ikaros. N Engl J Med 360 (5):524–526
Georgopoulos K, Bigby M, Wang JH, Molnar A, Wu P, Winandy S, Sharpe A (1994) The Ikaros gene is required for the development of all lymphoid lineages. Cell 79:143–156
Getnet D, Grosso JF, Goldberg MV, Harris TJ, Yen HR, Bruno TC, Durham NM, Hipkiss EL, Pyle KJ, Wada S, Pan F, Pardoll DM, Drake CG (2010) A role for the transcription factor Helios in human CD4(+) CD25(+) regulatory T cells. Mol Immunol 47:1595–1600
Goldman F, Gurel Z, Al-Zubeidi D, Freed A, Icardi M, Song C, Dovat S (2012) Congenital Pancytopenia and absence of B lymphocytes in a neonate with a mutation in the Ikaros gene. Pediatr Blood Cancer 58(4):591–597
Gómez-del Arco P, Maki K, Georgopoulos K (2004) Phosphorylation controls Ikaros's ability to negatively regulate the G1-S transition. Mol Cell Biol 24(7):2797–2807
Gómez-del Arco P, Koipally J, Georgopoulos K (2005) Ikaros SUMOylation: switching out of repression. Mol Cell Biol 25(7):2688–2697
Grzanka J, Leveson-Gower D, Golab K, Wang X-J, Marek-Trzonkowska N, Krzystyniak A, Wardowska A, Mills JM, Trzonkowsi P, Witkowski P (2013) FoxP3, Helios, and SATB1: roles and relationships in regulatory T cells. Int Immunopharmacol 16(3):343–347
Gurel Z, Ronni T, Ho S, Kuchar J, Payne KJ, Turk CW, Dovat S (2008) Recruitment of Ikaros to pericentromeric heterochromatin is regulated by phosphorylation. J Biol Chem 283 (13):8291–8300
Hahm K, Ernst P, Lo K, Turck GS, Smale ST (1994) The lymphoid transcription factor LyF-1 is encoded by specific, alternatively spliced mRNAs derived from the Ikaros gene. Mol Cell Biol 14:7111–7123
Hansen J, Strassburger P, Du Pasquier L (1997) Conservation of a master hematopoietic switch gene during vertebrate evolution: isolation and characterization of *Ikaros* from teleost and amphibian species. Eur J Immunol 27:3049–3058

Hollenhorst PC, Shah AA, Hopkins C, Graves BJ (2007) Genome-wide analyses reveal properties of redundant and specific promoter occupancy within the ETS gene family. Genes Dev 21:1882–1894
Hsu PD, Lander ES, Zhang F (2014) Development and applications of CRISPR-Cas9 for genome engineering. Cell 157(6):1262–1278
John LB, Ward AC (2011) The Ikaros gene family: transcriptional regulators of hematopoiesis and immunity. Mol Immunol 48:1272–1278
John LB, Yoong S, Ward AC (2009) Evolution of the Ikaros gene family: implications for the origins of adaptive immunity. J Immunol 182:4792–4799
John LB, Trengove MC, Fraser FW, Yoong SH, Ward AC (2013) Pegasus, the 'atypical' Ikaros family member, influences left-right asymmetry and regulates *pitx2* expression. Dev Biol 377:46–54
Joshi I, Yoshida T, Jena N, Qi X, Zhang J, Van Etten RA, Georgopoulos K (2014) Loss of Ikaros DNA-binding function confers integrin-dependent survival on pre-B cells and progression to acute lymphoblastic leukemia. Nat Immunol 15(3):294–304
Kastner P, Dupuis A, Gaub M-P, Herbrecht R, Lutz P, Chan S (2013) Function of Ikaros as a tumour suppressor in B cell acute lymphoblastic leukemia. Am J Blood Res 3(1):1–13
Kathrein KL, Chari S, Winandy S (2008) Ikaros directly represses the Notch target Hes1 in a leukemia cell line. J Biol Chem 283(16):10476–10484
Kim J, Sif S, Jones B, Jackson A, Koipally J, Heller E, Winandy S, Viel A, Sawyer A, Ikeda T et al (1999) Ikaros DNA-binding proteins direct formation of chromatin remodelling complexes in lymphocytes. Immunity 10:345–355
Kim J-H, Ebersole T, Kouprina N, Noskov VN, Ohzeki J-I, Masumoto H, Mravinac B, Sullivan BA, Pavlicek A, Dovat S, Pack SD, Kwon Y-W, Flanagan PT, Loukinov D, Lobanenkov V, Larinov V (2009) Human gamma-satellite DNA maintains open chromatin structure and protects a transgene from epigenetic silencing. Genome Res 19(4):533–544
Kleinmann E, A-s GLL, Sellars M, Kastner P, Chan S (2008) Ikaros represses the transcriptional response to Notch signalling in T-cell development. Mol Cell Biol 24:7465–7475
Kohonen P, Nera KP, Lassila O (2004) Avian Helios and evolution of the Ikaros family. Scand J Immunol 60:100–107
Krönke J, Udeshi ND, Narla A, Grauman P, Hurst SN, McConkey M, Svinkina T, Heckl D, Comer E, Li X, Carr SA, Ebert BL (2014) Lenalidomide causes selective degradation of IKZF1 and IKZF3 in multiple myeloma cells. Science 343:301–305
Li Z, Perez-casellas LA, Savic A, Song C, Dovat S (2011) Iakros isoforms: The saga continues. World J Biol Chem 2(6):140–145
Liberg D, Smale ST, Merkenschlager M (2003) Upstream of Ikaros. Trends Immunol 24 (11):567–570
Licht JD, Shortt J, Johnstone R (2014) From anecdote to targeted therapy: the curious case of thalidomide and multiple myeloma. Cancer Cell 25(1):9–11
Liongue C, John LB, Ward AC (2011) Origins of adaptive immunity. Crit Rev Immunol 31:53–63
Lo K, Landau NR, Smale ST (1991) LyF-1, a transcriptional regulator that interacts with a novel class of promoters for lymphocyte-specific genes. Mol Cell Biol 11(10):5229–5243
Lopez RA, Schoetz S, DeAngelis K, O'Neill D, Bank A (2002) Multiple hematopoietic defects and delayed globin switching in Ikaros null mice. Proc Natl Acad Sci U S A 99(2):602–607
Lu G, Middleton RE, Sun H, Naniong M, Ott CJ, Mitsiades CS, Wong KK, Bradner JE, Kaelin WH Jr (2013) The myeloma drug lenalidomide promotes the cereblon-dependent destruction of Ikaros proteins. Science 343:305–309
Ma S, Pathak S, Mandal M, Trinh L, Clark MR, Lu R (2010) Ikaros and Aiolos inhibit pre-B-cell proliferation by directly suppressing c-Myc expression. Mol Cell Biol 30(17):4149–4158
Ma H, Qazi S, Ozer Z, Zhang J, Ishkhanian R, Uckun FM (2013) Regulatory phosphorylation of Ikaros by Bruton's tyrosine kinase. PLoS One 8(8), e71302
Maillard I (2013) Ikaros, Notch, and GATA1 cross paths during megakaryopoiesis. Blood 121 (13):2376–2377

Malinge S, Thiollier C, Chlon TM, Doré LC, Diebold L, Bluteau O, Mabialah V, Vainchenker W, Dessen P, Winandy S, Mercher T, Crispino JD (2013) Ikaros inhibits megakaryopoiesis through functional interaction with GATA-1 and NOTCH signalling. Blood 121 (13):2440–2451

Mandel EM, Grosschedl R (2010) Transcription control of early B cell differentiation. Curr Opin Immunol 22:161–167

Martín-Ibáñez R, Crespo E, Urbán N, Sergent-Tanguy S, Herranz C, Jaumot M, Valiente M, Long JE, Pineda JR, Andreu C, Rubenstein JL, Marín O, Georgopoulos K, Mengod G, FariñasI BO, Alberch J, Canais JM (2010) Ikaros couples cell cycle arrest of late striatal precursors withneurogenesis of enkephalinergic neurons. J Comp Neurol 518(3):329–351

Martín-Ibáñez R, Crespo E, Esgleas M, Urban N, Wang B, Waclaw R, Georgopoulos K, Martínez S, Campbell K, Vicario-Abejón C, Alberch J, Chan S, Kastner P, Rubenstein JL, Canals JM (2012) *Helios* transcription factor expression depends on *Gsx2* and *Dlx1&2* function in developing striatal matrix neurons. Stem Cells Dev 21(12):2239–2251

Merkenschlager M (2010) Ikaros in immune receptor signalling, lymphocyte differentiation, and function. FEBS Lett 584:4910–4914

Molnar A, Gerorgopoulos K (1994) The Ikaros gene encodes a family of functionally diverse zinc finger DNA-binding proteins. Mol Cell Biol 14:8292–8303

Morgan B, Sun L, Avitahl N, Andrikopoulos K, Gonzales E, Wu P, Neben S, Georgopoulos K (1997) Aiolos, a lymphoid-restricted transcription factor that interacts with Ikaros to regulate lymphocyte differentiation. EMBO J 16(8):2004–2013

Naito T, Gómez-del Arco P, Williams CJ, Georgopoulos K (2007) Antagonistic interactions between Ikaros and the chromatin remodeler Mi-2β determine silencer activity and *Cd4* gene expression. Immunity 27:723–734

Nakahashi H et al (2013) A genome-wide map of CTCF multivalency redefines the CTCF code. Cell Rep 3:1678–1689

Ng SY-M, Yoshida T, Georgopoulos K (2007) Ikaros and chromatin regulation in early hematopoiesis. Curr Opin Immunol 19:116–122

O'Brien S, Thomas RM, Wertheim GB, Zhang F, Shen H, Wells AD (2014) Ikaros imposes a barrier to $CD8^+$ T cell differentiation by restricting autocrine IL-2 production. J Immunol 192:5118–5129

O'Neill DW, Schoetz SS, Lopez RA, Castle M, Rabinowitz L, Shor E, Krawchuk D, Goll MG, Renz M, Seelig HP, Han S, Seong RH, Park SD, Pan F, Yu H, Dang EV, Barbi J, Pan X, Grosso JF, Jinasena D, Sharma SM, McCadden EM, Getnet D, Drake CG, Liu JO, Ostrowski MC, Pardoll DM (2009) Eos mediates Foxp3-dependent gene silencing in CD4+ regulatory T cells. Science 325:1142–1146

Papathanasiou P, Perkins AC, Cobb BS, Ferrini R, Sridharan R, Hoyne GF, Nelms KA, Smale ST (2003) Widespread failure of hematolymphoid differentiation caused by a recessive niche-filling allele of the Ikaros transcription factor. Immunity 19:131–144

Papathanasiou P, Attema JL, Karsunky H, Hosen N, Sontani Y, Hoyne GF, Tunningley R, Smal ST, Weissman IL (2009) Self-renewal of long-term reconstituting subset of hematopoietic stem cells is regulated by Ikaros. Stem Cells 27(12):3082–3092

Payne KJ, Dovat S (2011) Ikaros and tumour suppression in acute lymphoblastic leukemia. Crit Rev Oncog 16(1–2):3–12

Payne KJ, Huang G, Sahakian E, Zhu JY, Barteneva NS, Barsky LW, Payne MA, Crooks GM (2003) Ikaros isoform x is selectively expressed in myeloid differentiation. J Immunol 170 (6):3091–3098

Popescu M, Gurel Z, Ronni T, Song C, Hung KY, Payne KJ, Dovat S (2009) J Biol Chem 284 (20):13869–13880

Qazi S, Uckun FM (2013) Incidence and biological significance of *IKZF1*/Ikaros gene deletions in pediatric Philadelphia chromosome positive B-cell precursor acute lymphoblastic leukemia. Hematologica 98, e151

Quintana FJ, Jin H, Burns EJ, Nadeau M, Yeste A, Kumar D, Rangachari M, Zhu C, Xiao S, Seavitt J, Georgopoulos K, Kuchroo VK (2012) Aiolos promotes T_H17 differentiation by directly silencing *Il2* expression. Nat Immunol 13(8):770–777

Revilla-i-Domingo R, Bilic I, Vilagos B, Tagoh H, Ebert A, Tamir IM, Smeenk L, Trupke J, Sommer A, Jaritz M, Busslinger M (2012) The B-cell identity factor Pax5 regulates distinct transcriptional programmes in early and late B lymphopoiesis. EMBO J 31(14):3130–3146

Romero F, Martinez-A C, Camonis J, Rebello A (1999) Aiolos transcription factor controls cell death by regulating Bcl-2 expression and its cellular localization. EMBO J 18:3419–3430

Rothenberg EV (2014) The chromatin landscape and transcription factors in T cell programming. Trends Immunol 35(5):195–204

Santoni de Sio FR, Barde I, Offner S, Kapopoulou A, Corsinotti A, Bojkowska K, Genolet R, Thomas JH, Luescher IF, Pinschewer D, Harris N, Trono D (2012) KAP1 regulates gene networks controlling T-cell development and responsiveness. FASEB J 26:4561–4575

Schjerven H, McLaughlin J, Teresita L, Frietze S, Cheng D, Wadsworth S, Lawson GW, Bensinger J, Farnham PJ, Witte ON, Smale ST (2013) Differential regulation of lymphopoiesis and leukemogenesis by individual zinc fingers of Ikaros. Nat Immunol 14(10):1073–1083

Schwickert TA, Tagoh H, Itekin SG, Dakic A, Axelsson E, Minnich M, Ebert A, Werner B, Roth M, Cimmino L, Dickins RA, Zuber J, Jaritz M, Busslinger M (2014) Stage-specific control of early B cell development by the transcription factor Ikaros. Nat Immunol 15 (3):283–293

Sellars M, Reina-San-Martin B, Kastner P, Chan S (2009) Ikaros controls isotype selection during immunoglobulin class switch recombination. J Exp Med 206:1073–1087

Shin DS, Jordan A, Basu S, Thomas RM, Bandyopadhyay, de Zotten EF, Wells AD, Macian F (2014) EMBO Rep 15(9):991–999

Smale ST (2014) Transcriptional regulation in the immune system: a status report. Trends Immunol 35(5):190–194

Song C, Li Z, Erbe AK, Savic A, Dovat S (2011) Regulation of Ikaros function by casein kinase 2 and protein phosphatase 1. World J Biol Chem 2(6):126–131

Sridharan R, Smale ST (2007) Predominant interaction of both Ikaros and Helios with the NuRD complex in immature thymocytes. J Biol Chem 282:30227–30238

Su RC, Sridharan R, Smale ST (2005) Assembly of silent chromatin during thymocyte development. Semmin Immunol 17(2):129–140

Thomas RM, Neelanjana C, Chen C, Umetsu SE, Winandy S, Wells AD (2007) Ikaros enforces the costimulatory requirement for *IL2* gene expression and is required for anergy induction in $CD4^+$ T lymphocytes. J Immunol 179:7305–7315

Thompson EC, Cobb BS, Sabbattini P, Meixlsperger S, Parelho V, Liberg D, Taylor B, Dillon N, Georgopoulos K, Jumaa H, Smale ST, Fisher AG, Merkenschlager M (2007) Ikaros DNA-binding proteins as integral components of B cell developmental-stage-specific regulatory circuits. Immunity 26:335–344

Tokunaga K, Yamaguchi S, Iwanaga E, Nanri T, Shimomura T, Suzushima H, Mitsuya H, Asou N (2013) Eur J Haematol 91(3):201–208

Trinh LA, ferrini R, Cobb BS, Weinmann AS, Hahm K, Ernst P, Garraway IP, Merkenschlager M, Smale ST (2001) Down-regulation of TDT transcription in CD4(+)CD8(+) thymocytes by Ikaros proteins in direct competition with an Ets activator. Genes Dev 15(14):1817–1832

Uckun FM, Ma H, Ishkhanian R, Arellano M, Shahidzadeh A, Termuhlen A, Gaynon PS, Qazi S (2013) Constitutive function of the Ikaros transcription factor in primary leukemic cells from pediatric newly diagnosed high-risk and relapsed B-precursor ALL patients. PloS One 8(11), e80732

van Steensel B (2011) Chromatin: constructing the big picture. EMBO J 30(10):1885–1895

Wang JH, Nichogiannopoulou A, Wu L, Sun L, Sharpe AH, Bigby M, Georgolpoulos K (1996) Selective defects in the development of the fetal and adult lymphoid system in mice with an Ikaros null mutation. Immunity 5:537–549

Wang L, Howarth A, Clark RE (2014) Ikaros transcripts Ik6/10 and levels of full-length transcript are critical for chromic myeloid leukemia blast crisis transformation. Leukemia 28 (8):1745–1747

Winandy S, Wu P, Georgopoulos K (1995) A dominant mutation in the Ikaros gene leads to rapid development of leukemia and lymphoma. Cell 83:289–299

Yeung F, Chung E, Guess MG, Bell ML, Leinwand LA (2012) Myh7b?miR-499 gene expression is transcriptionally regulated by MRFs and Eos. Nuclic Acid Res 40(15):7303–7318

Yoshida T, Georgopoulos K (2014) Ikaros fingers on lymphocyte differentiation. Int J Hematol 100(3):220–229

Yoshida T, Ng SY, Zuniga-Pflucker JC, Georgopoulos K (2006) Early hematopoietic lineage restrictions directed by Ikaros. Nat Immunol 7(4):382–391

Zhang J, Jackson AF, Natio T, Dose M, Seavitt J, Liu F, Heller EJ, Kashiwagi M, Yoshida T, Gounari F, Petrie HT, Georgopoulos K (2011) Hernessing of the nucleosome remodelling-deacetylase complex controls lymphocyte development and prevents leukemogenesis. Nat Immunol 13:86–94

Zhang JA, Mortazavi A, Williams BA, Wold BJ, Rothenberg EV (2012) Dynamic transformations of genome-wide epigenetic marking and transcriptional control establish T cell identity. Cell 149:467–482

Zou YR, Sunshine MJ, Taniuchi I, Hatam F, Killeen N, Littman DR (2001) Epigenetic silencing of CD4 in T cells committed to the cytotoxic lineage. Nat Genet 29:332–336

Zuo P, Stanojevic D, Colgan J, Han K, Levine M, Manley J (1991) Activation and repression of transcription by the gap proteins hunchback and Krüppel in cultured Drosophila cells. Genes Dev 5:254–264

Part V
Nuclear Domains and Cell Stress

Senescence Associated Heterochromatic Foci: SAHF

Tamir Chandra

Abstract The permanent cell cycle exit known as cellular senescence is a stress response that can be triggered by DNA damage, telomere erosion and activation of oncogenic signalling. Increasing evidence suggests that cellular senescence in cell culture mimics an in vivo situation, where a pathological cue (like the activation of an oncogene) triggers senescence as an intrinsic tumour suppressor defence. Oncogene induced senescence (OIS) is often accompanied by a global change in nuclear architecture, most dramatically exemplified by the formation of senescence associated heterochromatic foci (SAHF). Advances in imaging, chromatin profiling and mapping nuclear led to new insights into SAHF architecture, which will be summarised in this chapter.

1 Introduction

1.1 Cellular Senescence

Cellular senescence was first described as an irreversible cell cycle arrest in primary cells after long term culture (Hayflick and Moorhead 1961). Senescence triggers include DNA damage, oncogenic activity and telomere dysfunction (Di Leonardo et al. 1994; Serrano et al. 1997; Lin et al. 1998). Two key pathways have been established in oncogene induced senescence (OIS), namely the p53 and pRB/p16INK4a pathways with either or both pathways being mutated in the majority of cancers (Beauséjour et al. 2003). Especially p16 upregulation is prevalent in benign lesions and often lost during malignancy (Braig et al. 2005; Michaloglou et al. 2005; Haugstetter et al. 2010). Another key aspect of the

T. Chandra, Ph.D. (✉)
MRC Human Genetics Unit, The Institute of Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, Scotland
e-mail: tamir.chandra@igmm.ed.ac.uk

D.P. Bazett-Jones, G. Dellaire (eds.), *The Functional Nucleus*,
DOI 10.1007/978-3-319-38882-3_9

senescence response stems from the extensive secretion of cytokines, called senescence associated secretory phenotype (SASP), which plays a crucial role in recently reported immune clearance of cancer cells in vivo and paracrine induction of senescence in neighbouring cells (Xue et al. 2007; Acosta et al. 2008, 2013; Kuilman et al. 2008; Coppé et al. 2008). Moreover, two recent papers have implicated cellular senescence in normal development (Muñoz-Espín et al. 2013; Storer et al. 2013) and a role for senescence in the ageing process has been substantiated in mouse models of ageing where depletion of senescent cells led to a relieve of age-related symptoms (Baker et al. 2011, 2013; López-Otín et al. 2013; Martin et al. 2014). Overall, the functional roles of cellular senescence on an organismal level might be numerous and require further investigation.

1.2 Chromatin and Senescence

As opposed to epigenetic changes on single genes, cellular senescence results in a genome-wide alteration of its higher order chromatin structure (HOCS) resulting in SAHF (Senescence associated heterochromatic foci), which will be discussed in detail in a later paragraph (see below). One of the key effectors of the senescence response is the activation of the INK4a locus (Serrano et al. 1997). Given the importance of INK4a in cancer, the activation and repression of the INK4A locus has been extensively studied. Silencing of the INK4a locus is established through Polycomb group (PcG) proteins, such as the oncogene BMI1, a member of PRC1 (Polycomb Repressive Complex 1). BMI1 knock-out MEFs show upregulation of p16 and senesce after a few passages (Jacobs et al. 1999). CBX7 and CBX8, both PRC1 proteins, and the PRC1 interacting non-coding RNA ANRIL, contribute to repression of the INK4a locus (Gil et al. 2004; Bernard et al. 2005; Dietrich et al. 2007; Maertens et al. 2009; Yap et al. 2010). Mechanisms of activation of the INK4a locus are less described. Recently, the activation of the INK4a locus in senescence has been shown to be dependent on the H3K27me3 specific Histone demethylase JMJD3 (Agger et al. 2009; Barradas et al. 2009). Another key player in cellular senescence is the tumour suppressor protein p53 (Tumour Protein 53). The p53 pathway is inactivated in a variety of tumours and it was originally thought that apoptosis and checkpoint functions of p53 made for its tumour suppressive activity (Meek 2009; Junttila and Evan 2009). However, a recent series of papers suggest that other p53 activities such as senescence could lead to tumour suppression (Christophorou et al. 2006; Brady et al. 2011). Overall, the tumour suppressor function of p53 might be dependent on context.

2 Senescence Associated Heterochromatic Foci (SAHF)

2.1 A Model System for Chromatin Dynamics

SAHF are a chromosomal phenotype, formed from individual chromosomes. During this stress response, cells reorganize their chromatin into SAHF within days, probably hours, making SAHF a unique model system for chromatin dynamics beyond their important biological relevance.

HMGA proteins are a prime example of how studying SAHF formation can lead to the discovery of new tumour suppressor genes. Originally known to possess oncogenic activity, HMGA proteins are essential for SAHF formation, suggesting an involvement in tumour suppression (Narita et al. 2006). Indeed, HMGA1 knockout mice suggested an oncogenic and tumour suppressive role for HMGA1 (Fedele et al. 2006; Visone et al. 2008). In addition, new oncogenes can be discovered where SAHF formation is inhibited. One example is the p400 protein which inhibits SAHF formation and leads to transformation of human fibroblasts during overexpression of adenoviral E1A (Early Region 1A Gene of the Adenovirus Genome) (Chan et al. 2005).

Mapping the location of epigenetic marks genome-wide led to an effort to describe principal classes of chromatin (Ernst and Kellis 2010; Filion et al. 2010). Below I describe the segregation and clustering of chromatin in SAHF, which results in an expansion of coherent territories of certain chromatin types. The ease with which territories of certain chromatin types can be identified makes SAHF a good model system to interrogate structural and functional properties of chromatin subtypes.

2.2 SAHF Show Spatial Segregation and Clustering of Repressive Chromatin Marks

SAHF were initially described as distinct DAPI-dense foci, enriched for markers of constitutive heterochromatin (cHC), such as H3K9me3 and HP1 proteins (Narita et al. 2003). Staining SAHF with multiple markers for repressive chromatin, it was later shown that the cHC is restricted to the inner part of the SAHF, namely the core, which is surrounded by a ring, positive for H3K27me3, a marker of facultative heterochromatin (fHC) (Fig. 1) (Chandra et al. 2012). Staining for marks of active chromatin, such as H3K36me3, showed further chromatin segregation, as the active marks were found exclusively outside of the SAHF core and ring. Another anchor of SAHF architecture is the location of the pericentromeric and telomeric regions, both of which are found at the periphery of the SAHF core (Narita et al. 2003; Chandra et al. 2012). The SAHF architecture is summarised in Fig. 1.

The SAHF core and ring were initially described for their mutual exclusive enrichment in repressive histone marks, H3K9me3 and H3K27me3 respectively.

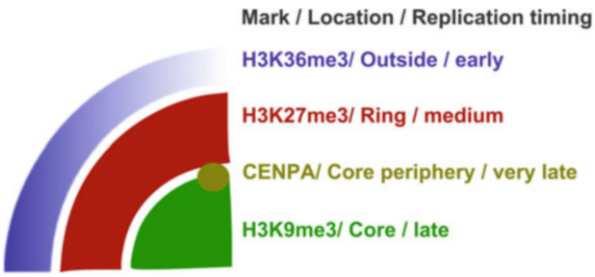

Fig. 1 The multi-layered architecture of a SAHF chromosome, where different heterochromatin types H3K9me3 and H3K27me3 seem to have segregated. Chromatin of the same type, H3K9me3 and H3K27me3, seems to cluster together forming a SAHF core and shell, respectively. Also shown are the locations of euchromatic regions (found outside of the SAHF shell) and pericentromeric regions (at the periphery of the SAHF core)

Although H3K9me3 and cHC had been implicated in SAHF function early on (a point I will revisit later), the role in SAHF formation remains unclear. There also is conflicting evidence on the total levels of H3K9me3 in senescence and SAHF formation, with levels reported to remain steady or increase (Di Micco et al. 2011; Chandra et al. 2012). There is evidence from imaging senescent cells overexpressing a H3K9me3 demethylase, that SAHF formation can occur when H3K9me3 levels are drastically reduced (Chandra et al. 2012). However, neither the quality of SAHF nor quantitative reductions thereof have been assessed. How and if cHC contributes to the segregation of chromatin types can only be speculated about. SAHF have been reproducibly accompanied by an upregulation of HP1 proteins on the chromatin, which could play a role in the clustering of the cHC regions. However, a role for HP1 proteins in SAHF formation has been weakened after overexpression of a dominant negative HP1, which delocalises 50–70 % of HP1 proteins from the chromatin, showed no effects (Zhang et al. 2007). However, qualitative changes in the SAHF structure may not have been sufficiently addressed.

H3K27me3 is a Xi-associated, fHC epigenetic mark (Plath et al. 2003). H3K9me3 and H3K27me3 are structurally (the sequences around the recognised lysines are largely identical) and functionally (both modifications have been associated with gene silencing and chromatin condensation) related. However, these histone marks often show a mutually exclusive localisation at the chromosomal level (Ringrose et al. 2004; Chadwick and Willard 2004). In SAHF, H3K27me3 exhibits a shell-like structure, surrounding the 'H3K9me3 core'. H3K27me3 plays a pivotal role in the genic regulation of the senescence response, such as in the regulation of the Ink/Arf locus, especially the expression of p16. However the role of global H3K27me3 levels in SAHF architecture is not understood. Despite a downregulation of EZH2 in senescence, steady levels for H3K27me3 have been reported (Chandra et al. 2012). Knock-down of an essential component of the PRC2 complex, SUZ12, did not prevent SAHF formation, although effects on the quantity and quality of SAHF have not been assessed. There is dynamic regulation of CBX

proteins, potential binders to the H3K27me3. Their role in the structural formation of SAHF remains to be investigated.

The dramatic change in repressive chromatin architecture in SAHF formation was speculated to be triggered by a spreading or redistribution of repressive histone marks as has been implicated in heterochromatin formation during the development of higher eukaryotes. In Drosophila PEV, the decisions in which genes are silenced are made during early development, leading to a variegated pattern (Ebert et al. 2006). Xi formation is developmentally regulated and the spreading of H3K9/27me3 repressive marks has been confirmed upon ESC differentiation and ESCs show lower levels of repressive histone modifications and more diffuse heterochromatin (Heard et al. 2001; Mermoud et al. 2002; Silva et al. 2003; Plath et al. 2003; Mikkelsen et al. 2007; Hawkins et al. 2010). Surprisingly, SAHF formation does not seem to be accompanied by a global redistribution or spreading of repressive marks (Chandra et al. 2012). This conclusion is largely based on ChIP-seq studies of repressive marks, which failed to find a global change during senescence but showed dramatic changes in ESC differentiation instead (Chandra et al. 2012). This observation suggests that SAHF architecture is the result of a spatial refolding, with no shifts of repressive chromatin blocks along the linear chromosome.

Cellular senescence can be induced by different pathological cues (Kuilman et al. 2010) and the extent of SAHF formation varies from a few percent in replicative senescent to over 90 % in c-raf induced senescence (Jeanblanc et al. 2012). The quality of SAHF, as assessed in consisting of a SAHF core and ring has been tested for a variety of cues, such as overexpression of oncogenes, ectopic p16 expression and DNA damage induction through etoposide. All tested conditions showed the same SAHF morphology, consisting of a core and a ring.

2.3 Replication Timing

Another approach to identify distinct regions within the SAHF architecture was made through labelling DNA replicating during distinct intervals of S-phase (Chandra et al. 2012). To monitor the re-localisation of genomic regions during SAHF formation, DNA labelling was performed using two pulses of thymidine analogue incorporation, which could be visualised individually. After the pulse labelling, oncogene induced senescence was induced and the cells were harvested 48 h later when a sizeable population of cells would have formed SAHF. In SAHF-positive cells there was a close correlation between RT and the layer structure of epigenetic marks (Fig. 1). Sequential comparison of different patterns observed in the SAHF positive cells revealed a spatiotemporal correlation between replication time and SAHF architecture. Importantly, late replicating regions were invariably located inside the early replicating regions, which indicated a directional association of RT with the radial distribution of chromatin in SAHF-forming chromosomes. Compared to the distinct boundaries observed for the repressive histone marks, the replication timing labelling seemed more continuous.

2.4 Histone Variants and Modifications

SAHF have been characterised by enrichment or absence of a variety of histones and histone marks. Histone variant macro-H2A (mH2A) and H3K9me2, both accumulate in the fHC of the Xi and are enriched in the entire SAHF (Zhang et al. 2005). Interestingly, mH2A has been shown to suppress tumour progression in malignant melanoma, a tumour suppressor response involving senescence (Kapoor et al. 2010). The linker histone H1 (H1) involved in condensation of chromatin is excluded from SAHF (Funayama et al. 2006). Moreover, histone modifications associated with active chromatin are also excluded from SAHF. Among those tested have been H3K9Ac, H3K4me3 and H3K36me3 (Chandra et al. 2012). The incorporation of the histone variant H3.3 into SAHF was suspected, because of the dependency of SAHF formation on the H3.3 specific chaperone histone regulator A (HIRA) (Zhang et al. 2005, 2007; Ye et al. 2007a, b; Rai et al. 2011). However, recent imaging studies found H3.3 excluded from SAHF (Corpet et al. 2014).

2.5 Architectural Proteins: HMGA1/2

In marked contrast to the debated change of H3K9/27me3 levels during SAHF formation, non-histone chromosomal architecture proteins HMGA1 and HMGA2 undoubtedly accumulate on SAHF positive chromatin (Narita et al. 2006). HMGA proteins probably accumulate on senescent chromatin in stoichiometric amounts to core histones as they were discovered by simply spotting a prominent band on a coomassie stained chromatin SDS-PAGE, running just above the core histones. Knock-down of HMGA1, and to a lesser degree HMGA2, completely prevents SAHF formation suggesting HMGA1 as a structural component of SAHF (Narita et al. 2006). HMGA knockdown in established SAHF cells leads to a reduction in SAHF, suggesting a role for HMGA proteins in the maintenance of SAHF (Narita et al. 2006; Chandra et al. 2012). HMGA proteins seem to cover the whole DAPI-dense SAHF without being restricted to the core or the ring. The role of HMGA2 in cancer (HMGA2 was originally described as an oncogene and is highly expressed in a subset of tumours) and the fundamental role of HMGA proteins in SAHF formation seem difficult to reconcile with the proposed tumour suppressive function of SAHF. However, recent studies suggested that HMGA1 can act as a tumour suppressor as well as an oncogene in a context dependent matter (Fedele et al. 2006; Narita et al. 2006; Visone et al. 2008). For example HMGA1−/− mice developed fewer skin papilloma in a two-stage chemical skin carcinogenesis protocol. However, the same study reported significantly less papilloma progression to carcinomas (Visone et al. 2008). These observations correlate well with the concept that senescence tumour suppressor activity acts in containing benign lesions, and thereby preventing malignancy.

2.6 Chromosomal Territories

The relocation of multiple types of chromatin (at least cHC and fHC) suggests that SAHF incorporate large regions of the chromosome. Does that mean that chromosomal territories change? The current model of interphase chromosomal territories suggests that chromosomes position their heterochromatic areas towards the nuclear periphery. With the heterochromatic areas inside, SAHF chromosomes may indeed show an inversion of chromosomal territories. It remains to be shown whether and to what extent chromosomal territories are preserved in SAHF.

2.7 Xi and SAHF

The inactive X chromosome (Xi) can easily be identified in DAPI stained interphase nuclei, as it is more condensed than autosomes and the active X chromosome (Xa) and usually located at the nuclear periphery. The Xi can also be identified by its enrichment in H3K27me3. A comparison between the Xi in growing (Xi/G) and SAHF positive cells (Xi/S) showed that the Xi/S retains its enrichment for H3K27me3 and remained distinguishable from the other chromosomes based on H3K27me3 staining, but not based on morphology by DAPI (Chandra et al. 2012). The part of the Xi/S that was not enriched in H3K27me3 showed a partial remodelling to a SAHF chromosome. It has been speculated that the Xi might be partially protected from SAHF remodelling, possibly through XIST RNA occupancy. This in turn could suggest the existence of an undiscovered, SAHF specific RNA that plays a role during remodelling and would repetition of suggest that the remodelled area might be the pseudoautosomal region. Another scenario, which might explain the observed differences between Xi/S and autosomal SAHF, could be the fundamentally different enrichment of repressive marks on the Xi/S, which may simply result in a different SAHF morphology, albeit the whole Xi/G being SAHF remodelled.

2.8 Histone Chaperones

The upregulation of histone chaperones HIRA and ASF1a in senescence led to speculations on the role of histone chaperones in SAHF formation. Indeed, it was shown that ectopic expression of HIRA and ASF1a alone is sufficient for SAHF formation in a fraction of cells (Zhang et al. 2005, 2007). HIRA accumulates in promyelocytic leukemia (PML) nuclear bodies upon senescence induction and it was shown that this relocalisation is required for SAHF formation (Zhang et al. 2007; Ye et al. 2007b). Additional members of the HIRA complex, UBN1

and Cabin1, were shown to localise to PML bodies and it seems that this mechanism is senescence specific (Banumathy et al. 2009; Rai et al. 2011).

2.9 Lamin Associated Domains in SAHF Formation

Lamin B1 (LMNB1) is lost in cellular senescence and this loss seems to be a prerequisite for SAHF formation (Shimi et al. 2011; Sadaie et al. 2013; Shah et al. 2013). It was for example possible to predict SAHF positive cells in a heterogeneous senescent population based on lower LMNB1 levels (Sadaie et al. 2013). This loss of LMNB1 leads to the dissociation of Lamin associated domains (LADs) from the nuclear periphery. It has also been shown that LMNB1 reduction and HMGA overexpression can work synergistically in promoting SAHF formation and showing a mild senescing effect on growing fibroblasts (Sadaie et al. 2013). In addition, ectopic LMNB1 expression led to a reduction of SAHF positive cells in senescence, without complete rescue from senescence. Profiling LADs led to the discovery of a small number of regions which seem to be gaining LMNB1, despite the genome-wide loss. These regions were suggested to being involved in gene silencing in senescence (Sadaie et al. 2013).

2.10 Ultrastructural Studies on SAHF Link Senescence and Progeria

A key conclusion of the studies on repressive histone marks in SAHF was that SAHF formation is fundamentally a three dimensional architectural change rather than a global redistribution of epigenetic marks (Chandra et al. 2012). A new study using Hi-C, a genome-wide proximity ligation approach, describes SAHF architecture on a three-dimensional level (Chandra et al. 2015). Surprisingly, the study revealed a loss of internal contacts in the cHC compartment, more specifically in a subset of LADs. These LADs were defined by their AT-richness. LADs have been classed by their GC content before in other contexts (Meuleman et al. 2013). The loss of internal connectivity is likely to represent a loss of chromatin compaction as was confirmed on one LAD by FISH studies. This observation seemed inconsistent with the idea of the SAHF core being a gene silencing compartment. Interestingly, the study also identified a subset of regions in the genome, which gained internal contacts and these regions overlap with the previously identified regions gaining LMNB1 in ChIP-seq experiments (Chandra et al. 2015; Sadaie et al. 2013).

Although a decompaction of cHC in senescence challenges a basic concept of SAHF function, it helps to integrate the controversial role of cHC in two of the most relevant cellular model systems of ageing. In contrast to the believed increase of heterochromatin in cellular senescence a loss of heterochromatin has been observed

in premature and healthy ageing (Scaffidi and Misteli 2006). The newly found loss of interactions in cHC in senescence allows the proposition that cHC decompaction is a common mechanism between progeroid and cellular senescence models. Indeed, a similar loss of connectivity in progeroid tissue was described in the same study (Chandra et al. 2015). In addition, the finding that loss of heterochromatin occurs in senescence was strengthened by a study showing decondensation and activation of satellite repeats, a mechanism named senescence associated distension of satellites (SADS) (Swanson et al. 2013).

Describing a common role for cHC in senescence and ageing leads to a new conundrum. Since progeroid cells do not show SAHF, there is no indication for the role of the unique structural reorganization that results in SAHF formation. In addition to the general loss of connectivity of heterochromatin, a spatial clustering of the decondensing regions on the same chromosome was observed. This observation was unique to SAHF positive cells and could not be reproduced in progeroid cells, which suggested that SAHF formation might be a two-step mechanism, with decompaction of heterochromatin as the first and spatial clustering of decompacted regions as the second step (Fig. 2). The same two-step mechanism was also proposed in the SADS study, where SADS, which has also been observed in progeria would represent the first step (Swanson et al. 2013).

While a decompaction of the cHC compartment within senescence may lead to the activation of satellite repeats, this may not be true for genic regions. A deeper understanding of the transcriptional consequences will have to be achieved before discarding the concept of SAHF as a silencing compartment. For example, it may be that the retention of repressive chromatin marks such as H3K9me3 and an increase in HP1 proteins is sufficient for gene silencing, despite chromatin

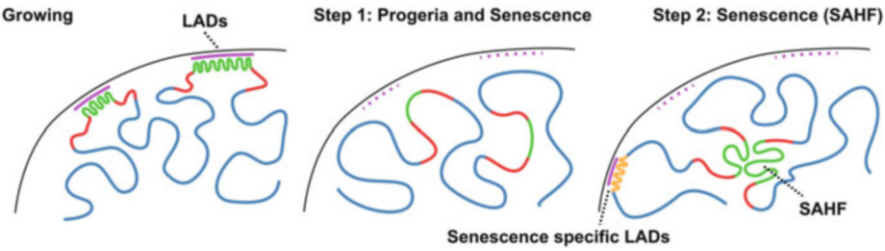

Fig. 2 A two-step model for SAHF formation. The schematics show the nuclear periphery of a cell on its way to senescence, starting with a proliferating cell to the *left*. Lamin associated domains (LADs), consist of compacted constitutive heterochromatin (cHC, *green*) and are attached to the nuclear lamina (*purple*). H3K27me3 regions (*red*) often flank the cHC in LADs. Euchromatic regions are shown in *blue*. As an initial stress response in acute to oncogene activation or chronic stress in progeria LADs detach from the lamina (*purple*) and it comes to a relaxation of the heterochromatin (*green*). In this model, Step 1 presents the endpoint for the chronic stress response in progeria, whereas it might be an intermediate step for SAHF forming cells. Step 2 shows the spatial clustering of cHC (*green* domains coming together) as suggested for SAHF positive cells. A set of unexplored regions in the genome gain LMNB1 and move towards the periphery in SAHF positive cells (*yellow*, LAD)

decondensation. Furthermore the function of the clustering observed in SAHF formation may involve re-silencing of areas activated through the decompaction.

2.11 SAHF Function

We are still oblivious as to whether the reorganization into SAHF chromatin is a functional process and if so what its function might be. Blocks of cHC have been shown to spread into adjacent active chromatin leading to the PEV in Drosophila. Could the clustering of cHC result in a heterochromatic entity (the SAHF core), which needs to be contained to prevent chromatin silencing at spatially proximal loci? In this scenario, the H3K27me3 shell could act as a buffer between transcribed and constitutive heterochromatic regions, establishing a spatial neighbourhood such as: Euchromatin (outside), gene-rich and transcriptionally active; fHC (ring): gene-rich and transcriptionally silent; cHC (core): gene poor and silent.

Another functional explanation for the multi-layered chromatin organization could be an efficient usage of the transcription machinery. SAHF formation may represent an end-point of continuous reorganization of the genome. Waiving plasticity, the cell may lock the genome in a senescence gene expression state. Of note, senescent cells become very efficient producers of a subset of proteins as suggested by the discovery of a compartment involved in protein turn-over (TASCC) and SAHF may represent the transcriptional counter-part to the TASCC compartment (Young et al. 2011).

Why the senescent genome may activate repeats as in SADS or indeed decompact its cHC domains is unknown. One plausible concept was proposed in a recent study showing the activation and transposition of LINE elements in replicative senescent. The authors suggest that the retrotransposition events may be a mechanism to induce or amplify a DNA damage signalling cascade inducing or reinforcing senescence (De Cecco et al. 2013a, b). The concept of an amplification of senescence signalling would also make sense in the light of cHC decompaction being an early event in senescence.

The response of regions gaining and losing compaction in senescence is so widespread that it could be described as a global response or even an inversion of global connectivity, with some of the most compacted regions opening up. Also, this is a fast response, showing effects in some cells after hours, with all cells being senescent within 2 days in, for example, Raf induced senescence (Jeanblanc et al. 2012). Furthermore the response can be activated due to a variety of triggers and cellular stresses. Taking together, the extent of reorganization, response time and the promiscuity to the inducing cue suggests the possibility that the chromatin reorganization in senescence may be hardwired in the genome, maybe even in the sequence. A strong correlation between changes in connectivity and isochores has been observed and leads to the questions of whether isochores have physical properties beyond sequence recognition that allow the genome to rearrange its architecture during a stress response. Such a concept would have far-reaching

implications for evolutionary biology as it would add constraints to genome evolution that have not been addressed so far.

References

Acosta JC, O'Loghlen A, Banito A et al (2008) Chemokine signaling via the CXCR2 receptor reinforces senescence. Cell 133:1006–1018. doi:10.1016/j.cell.2008.03.038

Acosta JC, Banito A, Wuestefeld T et al (2013) A complex secretory program orchestrated by the inflammasome controls paracrine senescence. Nat Cell Biol 15:978–990. doi:10.1038/ncb2784

Agger K, Cloos PAC, Rudkjaer L et al (2009) The H3K27me3 demethylase JMJD3 contributes to the activation of the INK4A-ARF locus in response to oncogene- and stress-induced senescence. Genes Dev 23:1171–1176. doi:10.1101/gad.510809

Baker DJ, Wijshake T, Tchkonia T et al (2011) Clearance of p16Ink4a-positive senescent cells delays ageing-associated disorders. Nature 479:232–236. doi:10.1038/nature10600

Baker DJ, Weaver RL, van Deursen JM (2013) p21 both attenuates and drives senescence and aging in BubR1 progeroid mice. Cell Rep 3:1164–1174. doi:10.1016/j.celrep.2013.03.028

Banumathy G, Somaiah N, Zhang R et al (2009) Human UBN1 is an ortholog of yeast Hpc2p and has an essential role in the HIRA/ASF1a chromatin-remodeling pathway in senescent cells. Mol Cell Biol 29:758–770. doi:10.1128/MCB.01047-08

Barradas M, Anderton E, Acosta JC et al (2009) Histone demethylase JMJD3 contributes to epigenetic control of INK4a/ARF by oncogenic RAS. Genes Dev 23:1177–1182. doi:10.1101/gad.511109

Beauséjour CM, Krtolica A, Galimi F et al (2003) Reversal of human cellular senescence: roles of the p53 and p16 pathways. EMBO J 22:4212–4222. doi:10.1093/emboj/cdg417

Bernard D, Martinez-Leal JF, Rizzo S et al (2005) CBX7 controls the growth of normal and tumor-derived prostate cells by repressing the Ink4a/Arf locus. Oncogene 24:5543–5551. doi:10.1038/sj.onc.1208735

Brady CA, Jiang D, Mello SS et al (2011) Distinct p53 transcriptional programs dictate acute DNA-damage responses and tumor suppression. Cell 145:571–583. doi:10.1016/j.cell.2011.03.035

Braig M, Lee S, Loddenkemper C et al (2005) Oncogene-induced senescence as an initial barrier in lymphoma development. Nature 436:660–665. doi:10.1038/nature03841

Chadwick BP, Willard HF (2004) Multiple spatially distinct types of facultative heterochromatin on the human inactive X chromosome. Proc Natl Acad Sci U S A 101:17450–17455. doi:10.1073/pnas.0408021101

Chan HM, Narita M, Lowe SW, Livingston DM (2005) The p400 E1A-associated protein is a novel component of the p53 → p21 senescence pathway. Genes Dev 19:196–201. doi:10.1101/gad.1280205

Chandra T, Kirschner K, Thuret J-Y et al (2012) Independence of repressive histone marks and chromatin compaction during senescent heterochromatic layer formation. Mol Cell 47:203–214. doi:10.1016/j.molcel.2012.06.010

Chandra T, Ewels P, Schoenfelder S et al (2015) Global reorganisation of the nuclear landscape in senescent cells. Cell Rep 10:471–483

Christophorou MA, Ringshausen I, Finch AJ et al (2006) The pathological response to DNA damage does not contribute to p53-mediated tumour suppression. Nature 443:214–217. doi:10.1038/nature05077

Coppé J-P, Patil CK, Rodier F et al (2008) Senescence-associated secretory phenotypes reveal cell-nonautonomous functions of oncogenic RAS and the p53 tumor suppressor. PLoS Biol 6:2853–2868. doi:10.1371/journal.pbio.0060301

Corpet A, Olbrich T, Gwerder M et al (2014) Dynamics of histone H3.3 deposition in proliferating and senescent cells reveals a DAXX-dependent targeting to PML-NBs important for pericentromeric heterochromatin organization. Cell Cycle 13:249–267. doi:10.4161/cc.26988

De Cecco M, Criscione SW, Peckham EJ et al (2013a) Genomes of replicatively senescent cells undergo global epigenetic changes leading to gene silencing and activation of transposable elements. Aging Cell 12:247–256. doi:10.1111/acel.12047

De Cecco M, Criscione SW, Peterson AL et al (2013b) Transposable elements become active and mobile in the genomes of aging mammalian somatic tissues. Aging (Albany NY) 5:867–883

Di Leonardo A, Linke SP, Clarkin K, Wahl GM (1994) DNA damage triggers a prolonged p53-dependent G1 arrest and long-term induction of Cip1 in normal human fibroblasts. Genes Dev 8:2540–2551

Di Micco R, Sulli G, Dobreva M et al (2011) Interplay between oncogene-induced DNA damage response and heterochromatin in senescence and cancer. Nat Cell Biol 13:292–302. doi:10.1038/ncb2170

Dietrich N, Bracken AP, Trinh E et al (2007) Bypass of senescence by the polycomb group protein CBX8 through direct binding to the INK4A-ARF locus. EMBO J 26:1637–1648. doi:10.1038/sj.emboj.7601632

Ebert A, Lein S, Schotta G, Reuter G (2006) Histone modification and the control of heterochromatic gene silencing in Drosophila. Chromosome Res 14:377–392. doi:10.1007/s10577-006-1066-1

Ernst J, Kellis M (2010) Discovery and characterization of chromatin states for systematic annotation of the human genome. Nat Biotechnol 28:817–825. doi:10.1038/nbt.1662

Fedele M, Fidanza V, Battista S et al (2006) Haploinsufficiency of the Hmga1 gene causes cardiac hypertrophy and myelo-lymphoproliferative disorders in mice. Cancer Res 66:2536–2543. doi:10.1158/0008-5472.CAN-05-1889

Filion GJ, van Bemmel JG, Braunschweig U et al (2010) Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. Cell 143:212–224. doi:10.1016/j.cell.2010.09.009

Funayama R, Saito M, Tanobe H, Ishikawa F (2006) Loss of linker histone H1 in cellular senescence. J Cell Biol 175:869–880. doi:10.1083/jcb.200604005

Gil J, Bernard D, Martínez D, Beach D (2004) Polycomb CBX7 has a unifying role in cellular lifespan. Nat Cell Biol 6:67–72. doi:10.1038/ncb1077

Haugstetter AM, Loddenkemper C, Lenze D et al (2010) Cellular senescence predicts treatment outcome in metastasised colorectal cancer. Br J Cancer 103:505–509. doi:10.1038/sj.bjc.6605784

Hawkins RD, Hon GC, Lee LK et al (2010) Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. Cell Stem Cell 6:479–491. doi:10.1016/j.stem.2010.03.018

Hayflick L, Moorhead PS (1961) The serial cultivation of human diploid cell strains. Exp Cell Res 25:585–621

Heard E, Rougeulle C, Arnaud D et al (2001) Methylation of histone H3 at Lys-9 is an early mark on the X chromosome during X inactivation. Cell 107:727–738

Jacobs JJ, Kieboom K, Marino S et al (1999) The oncogene and Polycomb-group gene bmi-1 regulates cell proliferation and senescence through the ink4a locus. Nature 397:164–168. doi:10.1038/16476

Jeanblanc M, Ragu S, Gey C et al (2012) Parallel pathways in RAF-induced senescence and conditions for its reversion. Oncogene 31:3072–3085. doi:10.1038/onc.2011.481

Junttila MR, Evan GI (2009) p53—a Jack of all trades but master of none. Nat Rev Cancer 9:821–829. doi:10.1038/nrc2728

Kapoor A, Goldberg MS, Cumberland LK et al (2010) The histone variant macroH2A suppresses melanoma progression through regulation of CDK8. Nature 468:1105–1109. doi:10.1038/nature09590

Kuilman T, Michaloglou C, Vredeveld LCW et al (2008) Oncogene-induced senescence relayed by an interleukin-dependent inflammatory network. Cell 133:1019–1031. doi:10.1016/j.cell.2008.03.039

Kuilman T, Michaloglou C, Mooi WJ, Peeper DS (2010) The essence of senescence. Genes Dev 24:2463–2479. doi:10.1101/gad.1971610

Lin AW, Barradas M, Stone JC et al (1998) Premature senescence involving p53 and p16 is activated in response to constitutive MEK/MAPK mitogenic signaling. Genes Dev 12:3008–3019

López-Otín C, Blasco MA, Partridge L et al (2013) The hallmarks of aging. Cell 153:1194–1217. doi:10.1016/j.cell.2013.05.039

Maertens GN, El Messaoudi-Aubert S, Racek T et al (2009) Several distinct polycomb complexes regulate and co-localize on the INK4a tumor suppressor locus. PLoS One 4, e6380. doi:10.1371/journal.pone.0006380

Martin N, Beach D, Gil J (2014) Ageing as developmental decay: insights from p16INK4a. Trends Mol Med. doi:10.1016/j.molmed.2014.09.008

Meek DW (2009) Tumour suppression by p53: a role for the DNA damage response? Nat Rev Cancer 9:714–723. doi:10.1038/nrc2716

Mermoud JE, Popova B, Peters AHFM et al (2002) Histone H3 lysine 9 methylation occurs rapidly at the onset of random X chromosome inactivation. Curr Biol 12:247–251

Meuleman W, Peric-Hupkes D, Kind J et al (2013) Constitutive nuclear lamina-genome interactions are highly conserved and associated with A/T-rich sequence. Genome Res 23:270–280. doi:10.1101/gr.141028.112

Michaloglou C, Vredeveld LCW, Soengas MS et al (2005) BRAFE600-associated senescence-like cell cycle arrest of human naevi. Nature 436:720–724. doi:10.1038/nature03890

Mikkelsen TS, Ku M, Jaffe DB et al (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448:553–560. doi:10.1038/nature06008

Muñoz-Espín D, Cañamero M, Maraver A et al (2013) Programmed cell senescence during mammalian embryonic development. Cell 155:1104–1118. doi:10.1016/j.cell.2013.10.019

Narita M, Nũnez S, Heard E et al (2003) Rb-mediated heterochromatin formation and silencing of E2F target genes during cellular senescence. Cell 113:703–716

Narita M, Narita M, Krizhanovsky V et al (2006) A novel role for high-mobility group a proteins in cellular senescence and heterochromatin formation. Cell 126:503–514. doi:10.1016/j.cell.2006.05.052

Plath K, Fang J, Mlynarczyk-Evans SK et al (2003) Role of histone H3 lysine 27 methylation in X inactivation. Science 300:131–135. doi:10.1126/science.1084274

Rai TS, Puri A, McBryan T et al (2011) Human CABIN1 is a functional member of the human HIRA/UBN1/ASF1a histone H3.3 chaperone complex. Mol Cell Biol 31:4107–4118. doi:10.1128/MCB.05546-11

Ringrose L, Ehret H, Paro R (2004) Distinct contributions of histone H3 lysine 9 and 27 methylation to locus-specific stability of polycomb complexes. Mol Cell 16:641–653. doi:10.1016/j.molcel.2004.10.015

Sadaie M, Salama R, Carroll T et al (2013) Redistribution of the Lamin B1 genomic binding profile affects rearrangement of heterochromatic domains and SAHF formation during senescence. Genes Dev 27:1800–1808. doi:10.1101/gad.217281.113

Scaffidi P, Misteli T (2006) Lamin A-dependent nuclear defects in human aging. Science 312:1059–1063. doi:10.1126/science.1127168

Serrano M, Lin AW, McCurrach ME et al (1997) Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. Cell 88:593–602

Shah PP, Donahue G, Otte GL et al (2013) Lamin B1 depletion in senescent cells triggers large-scale changes in gene expression and the chromatin landscape. Genes Dev 27:1787–1799. doi:10.1101/gad.223834.113

Shimi T, Butin-Israeli V, Adam SA et al (2011) The role of nuclear Lamin B1 in cell proliferation and senescence. Genes Dev 25:2579–2593. doi:10.1101/gad.179515.111

Silva J, Mak W, Zvetkova I et al (2003) Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes. Dev Cell 4:481–495

Storer M, Mas A, Robert-Moreno A et al (2013) Senescence is a developmental mechanism that contributes to embryonic growth and patterning. Cell 155:1119–1130. doi:10.1016/j.cell.2013.10.041

Swanson EC, Manning B, Zhang H, Lawrence JB (2013) Higher-order unfolding of satellite heterochromatin is a consistent and early event in cell senescence. J Cell Biol 203:929–942. doi:10.1083/jcb.201306073

Visone R, Iuliano R, Palmieri D et al (2008) Hmga1 null mice are less susceptible to chemically induced skin carcinogenesis. Eur J Cancer 44:318–325. doi:10.1016/j.ejca.2007.11.017

Xue W, Zender L, Miething C et al (2007) Senescence and tumour clearance is triggered by p53 restoration in murine liver carcinomas. Nature 445:656–660. doi:10.1038/nature05529

Yap KL, Li S, Muñoz-Cabello AM et al (2010) Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. Mol Cell 38:662–674. doi:10.1016/j.molcel.2010.03.021

Ye X, Zerlanko B, Kennedy A et al (2007a) Downregulation of Wnt signaling is a trigger for formation of facultative heterochromatin and onset of cell senescence in primary human cells. Mol Cell 27:183–196. doi:10.1016/j.molcel.2007.05.034

Ye X, Zerlanko B, Zhang R et al (2007b) Definition of pRB- and p53-dependent and -independent steps in HIRA/ASF1a-mediated formation of senescence-associated heterochromatin foci. Mol Cell Biol 27:2452–2465. doi:10.1128/MCB.01592-06

Young ARJ, Narita M, Narita M (2011) Spatio-temporal association between mTOR and autophagy during cellular senescence. Autophagy 7:1387–1388. doi:10.4161/auto.7.11.17348

Zhang R, Poustovoitov MV, Ye X et al (2005) Formation of Macroh2a-containing senescence-associated heterochromatin foci and senescence driven by ASF1a and HIRA. Dev Cell 8:19–30. doi:10.1016/j.devcel.2004.10.019

Zhang R, Chen W, Adams PD (2007) Molecular dissection of formation of senescence-associated heterochromatin foci. Mol Cell Biol 27:2343–2358. doi:10.1128/MCB.02019-06

DNA Repair Foci Formation and Function at DNA Double-Strand Breaks

Michael J. Hendzel and Hilmar Strickfaden

Abstract There are many examples of regions of the cell nucleus that accumulate a specific subset of nuclear proteins. We generally refer to these as nuclear compartments or nuclear bodies. Unlike those of the cytoplasm, nuclear compartments do not contain a surrounding membrane that physically separates the nucleoplasm from the interior of the compartment. Compartments that form in response to DNA double-strand breaks provide an interesting system to understand compartmentalization in the nucleus and its functional significance. Many of the protein modules responsible for targeting individual proteins to these compartments and the basic hierarchy of their assembly have been established. Most importantly, these compartments can be assembled on demand predictably in time and, using some model systems, predictably in space, thereby facilitating their study. In this chapter, we will examine the biophysical properties of the structure known as the DNA repair focus (DRF) that forms around a DNA double-strand break, reviewing experiments that examine the functional importance of the structure as a fundamental compartment of DNA repair.

Keywords DNA double-strand break • Compartmentalization • DNA damage checkpoint • Ionizing radiation induced focus • DNA double-strand break repair focus • Fluorescence microscopy • Transmission electron microscopy

1 Introduction

The cell nucleus contains no internal membrane system to segregate the volume into separate functional compartments and yet, when antibodies are used to stain for nuclear proteins, many of them show clear evidence for ordered association and segregation within the nucleoplasm (Sleeman and Trinkle-Mulcahy 2014). How does this occur and is it important? In many cases, we now know at least some of the molecular mechanisms required for compartment assembly as we will discuss in

M.J. Hendzel (✉) • H. Strickfaden
Department of Oncology, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Alberta, Canada
e-mail: mhendzel@ualberta.ca; Strickfaden@ualberta.ca

D.P. Bazett-Jones, G. Dellaire (eds.), *The Functional Nucleus*,
DOI 10.1007/978-3-319-38882-3_10

detail in the next section. These DNA repair compartments are initiated around DNA double-strand breaks (DSBs), and a single break is sufficient to nucleate the process. DSBs are among the most cytotoxic types of DNA damage. If left unrepaired, they can lead to loss of genetic material or chromosomal rearrangements, and promote the onset of diseases associated with genomic instability such as cancer. Consequently, the cell has developed machinery to repair DNA double-strand breaks. As part of this process, the cell assembles a compartment around the DSB that is enriched in proteins involved in the repair and in signalling from DNA DSBs. This makes DSB repair an attractive model to study the process of compartment formation in the cell nucleus because these compartments can essentially be assembled on demand without the requirement to artificially concentrate one or more proteins. Events that are associated with DNA repair are also easy to monitor, such as downstream protein phosphorylation, kinetics and fidelity of repair of the DNA break.

1.1 The Hierarchical Nature of DNA Double-Strand Break (DSB) Repair

When DNA DSBs arise in the genome, a signalling cascade is initiated by Ataxia telangiectasia mutated (ATM)-mediated phosphorylation of H2AX (γ-H2AX). Phosphorylation of H2AX by ATM is facilitated by an SQ amino acid motif in the C-terminus that is absent in the canonical H2A (Celeste et al. 2003b; Burma et al. 2001). This signalling pathway generates distinct nuclear compartments via the hierarchical assembly of DNA repair factors, which we term DSB Repair Foci (DRFs), around the DSBs (Fig. 1). The importance of the formation of these DRFs in the DSB repair process is not known but DRF formation is clearly not an absolute requirement for DSB repair because H2AX null mice are viable (Bassing et al. 2003; Celeste et al. 2002, 2003a). Cells lacking H2AX do not form stable DSB repair foci in response to DSBs (Celeste et al. 2003b). However, the contribution of compartmentalization to the repair of DSBs should be evident in H2AX−/− MEFs. H2AX−/− cells do have a defect in G2/M checkpoint activation after exposure to low doses (0.5 Gy) of ionizing radiation (Fernandez-Capetillo et al. 2002). Moreover, in a p53−/− background, H2AX has been shown to function as a dosage-dependent suppressor of oncogenic translocations (Bassing et al. 2003; Celeste et al. 2003a) and loss of one H2AX allele may contribute to the development of some human cancers (Dickey et al. 2009). Whether or not the absence of DRFs is responsible for this is not known. It is important to note that H2AX is acetylated differently than the major H2A subtypes and contains additional posttranslationally modified residues that respond to DNA damage (Chen et al. 2013; Gong and Miller 2013), which could indicate additional functions that are independent of repair or compartmentalization during repair.

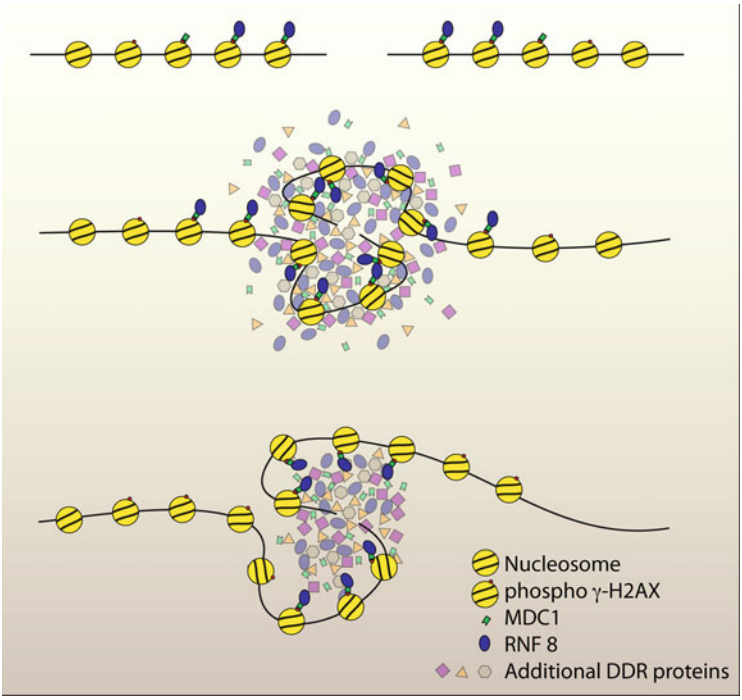

Fig. 1 Hierarchical assembly of DNA repair foci (DRF). *Top panel*: After the formation of a DNA double-strand break, the nucleosomes surrounding the DNA double-strand break are phosphorylated on histone H2AX and then recruit MDC1. *Middle panel*: The recruitment of proteins such as the E3 ubiquitin ligases RNF8 and RNF168 and the proteins that bind to phosphorylated, ubiquitylated, and sumoylated proteins that associate with the nucleosomes surrounding the DNA double-strand break initiate the focal organization of the DRF through interactions between the proteins and protein modification scaffolds such as K63-linked polyubiquitin chains, which bind RAP80 to recruit BRCA1. Bottom panel: When these foci persist for extended periods of time, the segregation of the protein network from the chromatin that initiated the formation of the focus becomes more apparent, with the chromatin largely being associated with the exterior of the focus. DDR = DNA damage response

Our understanding of hierarchy of DNA repair has been greatly aided by technical advances in live-cell microscopy, and one of the most revealing systems has been to introduce multiple copies of the DNA binding site or the bacterial repressor LacI, known as the Lac operator (LacO), into the genome (Dundr and Misteli 2010; Kaiser et al. 2008; Shevtsov and Dundr 2011; Soutoglou and Misteli 2008). In this system, candidate proteins for players in the compartmentalization of DNA repair can be forced to accumulate at a specific genomic locus within the nucleus, which contains multiple copies of the LacO sequence, via fusion of these proteins with the LacI repressor. These studies support both ordered and stochastic assembly models (Mao et al. 2011) and suggest that only a subset of proteins are required for the recruitment and retention of other members of the compartment. For example, when these experiments were done for proteins involved in

compartment formation around DNA double-strand breaks, the upstream kinase ATM, the upstream scaffolding protein MDC1, and the upstream sensor protein NBS1 were all able to nucleate the formation of a compartment that accumulated other DNA repair proteins (Soutoglou and Misteli 2008). In the case of MDC1, this required the BRCT domain. In contrast, downstream proteins CHK1 and CHK2 did not nucleate the formation of a compartment (Soutoglou and Misteli 2008).

Most remarkably, compared to constitutive nuclear domains such as the nucleolus, DRFs can be "assembled on demand" by treating cells with ionizing radiation or other agents that induce DSBs. As we discussed above, this assembly requires phosphorylation of histone H2AX (Celeste et al. 2003b). ChIP-seq experiments, where the chromatin containing phosphorylated H2AX is immunoprecipitated and the associated DNA sequenced, revealed that histone H2AX phosphorylation encompassed between 500 Kb and 2 Mb pairs of genomic sequence (Iacovoni et al. 2010), consistent with previous estimates (Rogakou et al. 1999). Measurements reveal that more than 1000 copies of individual DNA damage response (DDR) proteins can be found in these compartments (Mok and Henderson 2012a, b). In most cases, these proteins are rapidly exchanging between the DSB repair compartment and the surrounding nucleoplasm rather than accumulating statically at the DSB until it is repaired (Mok and Henderson 2012a, b). In the case of an enzyme like ATM, this means that the kinase is likely to access substrates outside of DRFs. In this chapter, we will examine the structure, dynamics, and functions of DRFs.

2 What Do We Know About the Structure of DNA Repair Foci (DRF)?

The remarkable thing about DNA double-strand breaks is that they initiate the formation of structures that are, despite the relatively small size of the DNA double-strand break itself, massive in size. This was initially observed with Rad51 (Haaf et al. 1995) and later additional proteins that respond to DSBs induced by ionizing radiation (Essers et al. 1997; Maser et al. 1997), leading to the term ionizing radiation induced foci (IRIF), but has subsequently been observed upon DNA replication stress and upon inducing a site-specific DNA double-strand break with a restriction endonuclease. With the discovery of a specific phosphorylated form of histone H2AX, phosphorylated on serine 139 (Rogakou et al. 1998), that was induced upon treatment with ionizing radiation and that phosphorylated histone H2AX forms IRIF (Rogakou et al. 1999), phosphorylated histone H2AX has become the most commonly employed probe to identify these foci in cells. Figure 2 shows a comparison of DRFs induced by ionizing radiation relative to sites of RNA transcription. These structures are comparatively large and irregular in shape. Fluorescence microscopy has revealed these structures to be relatively uniform in

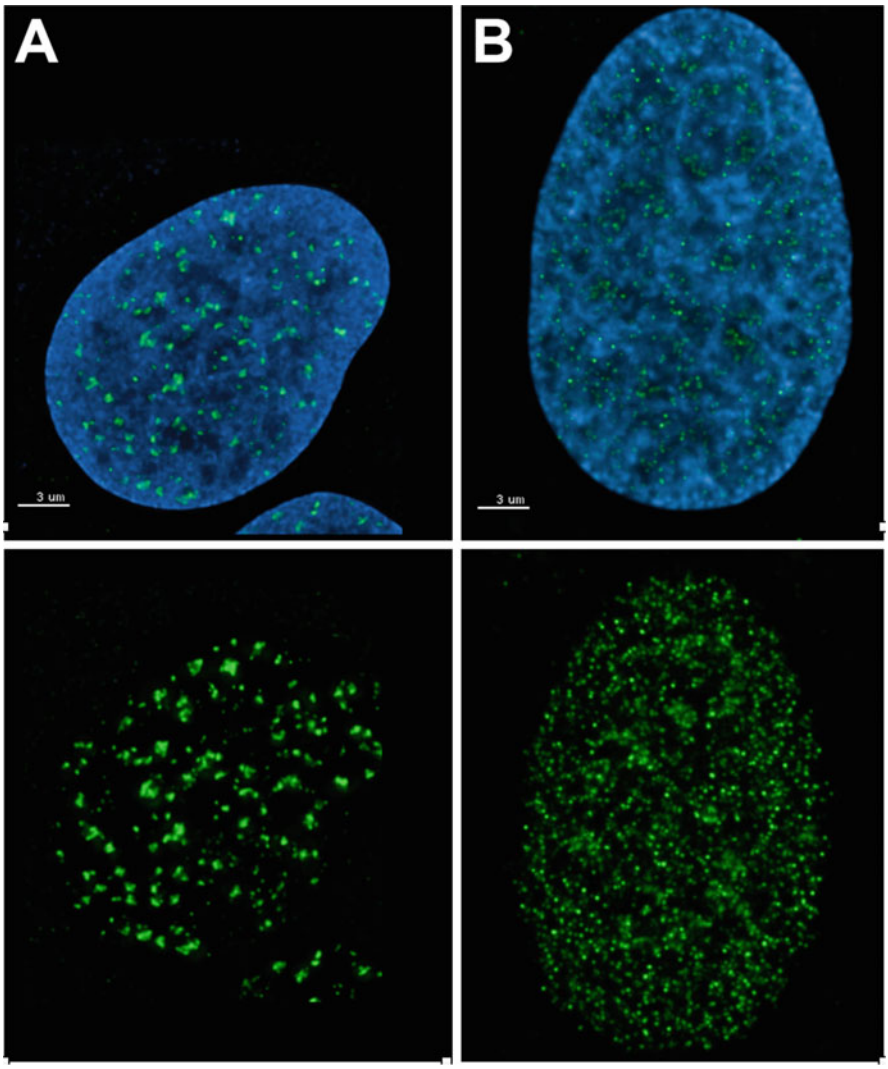

Fig. 2 A comparison of Double-strand break repair foci and sites of RNA transcription. (**a**) The left panels show DNA (*blue*) and phosphorylated histone H2AX (*green*). (**b**) The right panels show DNA (*blue*) and 5-fluorouridine (*green*), which is incorporated into newly synthesized RNAs. Scale bar represents 3 μm

composition. Thus, a common method to identify a new DSB signalling or repair protein is through colocalization with phosphorylated histone H2AX in DRFs.

The advent of advanced electron microscopy and super-resolution fluorescent light microscopy techniques has allowed a reinvestigation of the structure of DRFs. These experiments revealed that DRFs may subcompartmentalize proteins within the DRF (Chapman et al. 2012; Dellaire et al. 2009). Figure 3 shows an example of

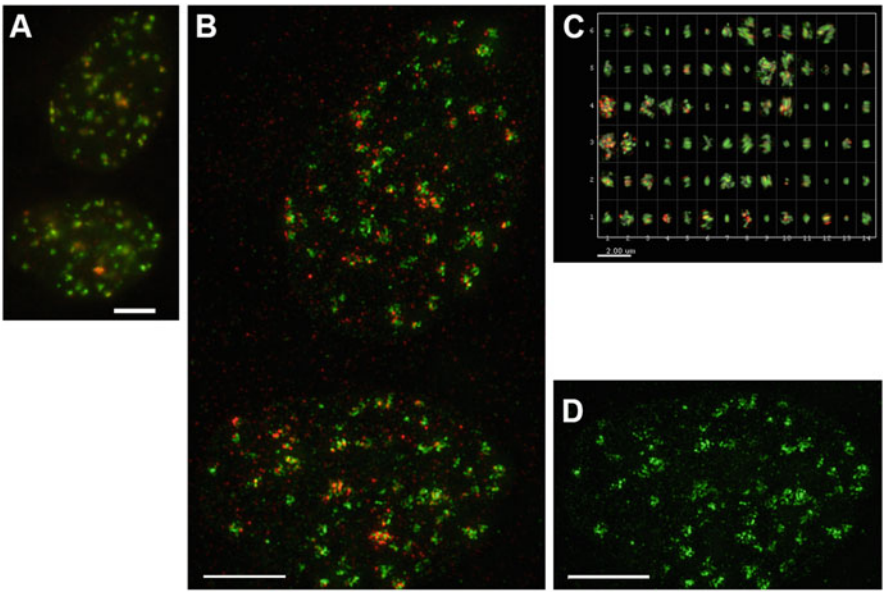

Fig. 3 Structured illumination microscopy of DSB repair foci. A pair of U2OS cells were stained with anti-phosphorylated H2AX after 2 h post-irradiation with 2 Gy. (**a**) Anti-phosphorylated H2AX (*red*) staining and anti-ubiquitin (FK2, *green*) are shown in a conventional fluorescence microscopy image. (**b**) The same pair of nuclei were imaged by structured illumination microscopy (**c**) Projections of the individual foci from A that have been segmented and sorted by size. (**d**) The staining of ubiquitin in the bottom nucleus from **b** is shown in isolation. In the structured illumination images, the foci are not homogenous but rather appear fibriillar. This could reflect higher-order chromatin organization. The scale bar represents 5 μm in **a**, **b**, and **d** and 2 μm in **c**

a super-resolution image of a DRF where the accumulation of ubiquitin (green) is partially separated from regions enriched in phosphorylated histone H2AX (red).

2.1 Ultrastructure of DRFs

Not surprisingly, much less is known about the ultrastructure of sites of DNA double-strand breaks. Conventional transmission electron microscopy coupled with immunogold labeling has revealed clustering of 53BP1, MDC1, and phosphorylated histone H2AX on electron dense material that is presumably chromatin (Lorat et al. 2012; Rube et al. 2011). Similarly, BRCA1 was shown to form clusters of about 100 nm in diameter and associated with electron dense material believed to be chromatin (Mok and Henderson 2012b). In contrast, Ku70 phosphorylated on serine 6 was found in pairs in euchromatin regions (defined as less electron dense regions in this study) but could be found in clusters of up to 6 gold particles in heterochromatin (Lorat et al. 2012). Interestingly, Ku70 only colocalized with

53BP1 and phosphorylated histone H2AX in heterochromatin domains, where these clusters were observed on the surface of heterochromatin (Lorat et al. 2012), consistent with fluorescence microscopy studies showing that DRFs rapidly move to the surface of heterochromatin (Jakob et al. 2011; Chiolo et al. 2011, 2013). While these studies showed evidence of clustering, they did not reveal any evidence of subcompartmentalization within the DRFs. In part, this may reflect the difficulty of distinguishing chromatin and nonchromatin structures when using electron dense stains in conventional transmission electron microscopy, particularly under conditions where the gold particles limit the range of contrast of the biological materials in the thin section.

In contrast to these conventional transmission electron microscopy experiments, electron spectroscopic imaging has revealed evidence for further differentiation of structure within the DRF (Dellaire et al. 2009). This method allows for the direct detection of phosphorus to delineate the chromatin and distinguish it from the protein, which is visualized with nitrogen. These experiments do reveal evidence of differentiation within the DRF compartment. The approach reveals separate chromatin and nonchromatin regions and allows the assignment of immunolabelled proteins to chromatin and nonchromatin structures within the DRF. Interestingly, in a comparison of MRE11 and phosphorylated histone H2AX localization, while the H2AX staining localized to chromatin, as expected, the MRE11 was most often found in nonchromatin structures that associated with the chromatin fibres (Dellaire et al. 2009). Moreover, while these antigens were dispersed throughout the focus at a 30 min time point, they adopted a more peripheral localization at a 3 h time point (Dellaire et al. 2009). This is consistent with specialization of subdomains within the DSB repair focus, at least at later time points.

2.2 How Do DRFs Assemble?

As noted above, histone H2AX is absolutely required for the assembly of most of the DDR proteins into foci surrounding DSBs. A clear exception is Rad51 foci, which form several hours after DNA damage and are still observed in H2AX−/− cells (Celeste et al. 2002). For most proteins, the initiation step in their recruitment is the phosphorylation of histone H2AX, typically through the activation of the ATM kinase (Misteli and Soutoglou 2009). An elegant approach to test the requirements for assembly of downstream proteins into DRFs is to make use of synthetic arrays of binding sites for prokaryotic DNA binding proteins. The lac repressor has been the most commonly used system. By creating a fusion protein with the lac repressor, it is possible to specifically accumulate any protein to this site in the nucleus containing the integrated DNA construct. This approach allows one to test for the dependencies for recruitment to these sites and to establish the minimal requirements to generate a DSB response signal without the presence of a double-strand break. When these experiments were performed, the catalytic domain of ATM, MRE11, NBS1, or MDC1 tethering were each sufficient to stimulate the

phosphorylation of histone H2AX and the generation of a DSB repair focus in the absence of DNA damage (Soutoglou and Misteli 2008). Doing so was sufficient, also, to activate the G2/M checkpoint. In contrast, tethering CHK1 or CHK2 to a single site was not sufficient to stimulate histone H2AX phosphorylation, the assembly of a focus, or activation of the G2/M checkpoint (Soutoglou and Misteli 2008). This established that locally concentrating any of these upstream components is sufficient to initiate the signalling cascade and that more downstream components, like MDC1, when concentrated, could initiate the cascade. MDC1 required the tandem BRCT domains that bind to the C-terminus of phosphorylated histone H2AX (Soutoglou and Misteli 2008). It would be interesting to determine if tethering other downstream components that localize to foci is sufficient to initiate the formation of a focus and the activation of the G2/M checkpoint. A previous study showing that histone H2AX is phosphorylated when BRCA1 is tethered to a lac operon array suggests that this may be the case (Ye et al. 2001). Nonetheless, these studies support the hypothesis that compartmentalization is important for signal amplification. Moreover, while downstream components may be capable of nucleating a DRF when tethered to a specific site in the nucleus at high concentrations, it is clear that the normal process of DRF assembly in vivo is ordered and initiated by the phosphorylation of histone H2AX.

2.3 DRF Protein Composition and Dynamics

Most proteins associated with the various nuclear compartments are dynamic, and individual molecules are constantly exchanging between the compartment and the nucleoplasm. Fluorescence recovery after photobleaching (FRAP) is a technique that has been used to quantify the binding of DDR proteins to foci or laser microirradiation-induced DNA damage. Because of the complexity of laser microirradiation-induced damage and because these results have been shown to differ from those obtained with foci (Mok and Henderson 2012a), we will only consider the results obtained with DRFs. In general, upon induction of a double-strand break, proteins that accumulate in DRFs show much slower rates of recovery in the DRF relative to the surrounding nucleoplasm in FRAP experiments (Mok and Henderson 2012a, b; Essers et al. 2002). This reflects the presence of binding events that reduce diffusion within the DRF. What is less certain, is whether or not there are populations that are bound more tightly and represent effectively permanent associations with DRFs over a time scale of minutes. Over the time scale of hours, DRFs typically resolve but it remains an open question whether any of the key signalling or repair molecules reside permanently in the DRF. One example of a protein that could be considered a more stable component of DRFs is Rad51 (Essers et al. 2002). For example, as shown in Table 1, 20 % of the pool of BRCA1 A complex subunits are apparently in stable association with the DRF. If these are genuinely stable interactions, then there would have to be two types of binding events in order to explain the data. However, there is always an uncertainty in

Table 1 Diffusion and mobility of selected DNA double-strand break response and repair proteins

Protein	T1/2 or Deff + DNA damage	T1/2 or Deff– DNA damage	Mobile Fraction + DNA damage	Mobile Fraction – DNA damage	Reference
MDC1	3 (19 %) 83 (81 %)	2 s (44 %); 32 s (56 %)	83 %	81 %	Mok and Henderson (2012a)
53BP1	3 s (17 %) 77 s (83 %)	2 s (24 %); 83 s (76 %)	55 %	36 %	Mok and Henderson (2012a)
Abraxas	10 s (34 %) 73 s (66 %)	ND	ND	80 %	Mok and Henderson (2012b)
Bard1	2 s (25 %) 101 s (75 %)	ND	ND	80 %	Mok and Henderson (2012b)
RAP80	0.4 s (40 %) 6 s (60 %)	ND	ND	80 %	Mok and Henderson (2012b)
BRCA1	4 s (23 %) 77 s (77 %)	ND	ND	80 %	Mok and Henderson (2012b)
Rad51	Not reached	~7 u2/s	ND	ND	Essers et al. (2002)
Rad52	26 s	~8 u2/s	ND	ND	Essers et al. (2002)
Rad54	0.5 s	~14u2/s	ND	ND	Essers et al. (2002)
Ku80	4.02 u2/s	14.42 u2/s	ND	ND	Abdisalaam et al. (2014)
DNAPKcs	3.35 u2/s	6.17 u2/s	ND	ND	Abdisalaam et al. (2014)

measuring the immobile fraction. This was highlighted by another study examining the mobility of the DNA repair protein RAD52, where FRAP experiments revealed ~80 % recovery but these results were inconsistent with combined FRAP and fluorescence loss in photobleaching (FLIP) experiments (Essers et al. 2002); where half the nucleus was photobleached and the recovery of the photobleached region and the loss from the region that was not photobleached were simultaneously monitored. When the stable population of RAD52 was assessed in this manner, the two fluorescence intensity curves converged to a single point, indicating that the entire population is mobile (Essers et al. 2002).

One of the interesting questions that can be answered by either FRAP or fluorescence correlation spectroscopy (FCS), a technique that is particularly good at resolving fast moving populations of molecules, is whether or not proteins are recruited as complexes or assembled de novo at the DNA DSB. For example, an

analysis of the behaviour of the BRCA1 A complex members Abraxas, Bard1, BRCA1, and RAP80 revealed that the RAP80 subunit was more dynamic than the remainder of the complex and was present in two-fold abundance (Mok and Henderson 2012b). This suggests that the interaction between RAP80 and BRCA1/BARD1 does not dictate the retention properties of BRCA1/BARD1 at DRF and suggests that the interaction between BRCA1 and RAP80 is more dynamic than anticipated. More compelling evidence for assembly at the DRF comes from an analysis of Rad51, Rad52, and Rad54. In this study, Rad51 was clearly more stably incorporated into DRFs than Rad52 and Rad54, despite interactions between these proteins (Essers et al. 2002). This likely reflects the assembly of Rad51 into polymeric filaments, thereby immobilizing the Rad51 molecules.

Using a viral particle expressing GFP as a standard, it has also been possible to estimate copy number of individual protein molecules inside the DRF. While this approach has the limitation of relying on the expression of a GFP-tagged version of a protein to determine its abundance, the size of the foci assembled in cells expressing tagged versions of DNA damage response proteins is not significantly altered, and hence such studies most likely provide a reasonable estimate of their abundance. Using this approach, it was estimated that there are over 1000 copies of MDC1, 53BP1, and RAP80 per focus as well as several hundred copies of BRCA1, BARD1, and Abraxas (Mok and Henderson 2012a, b).

2.4 Intranuclear Mobility of DRFs

One of the important consequences of misrepair is the formation of chromosome translocations by the ligation of DNA breaks arising on two different chromosomes. For example, the translocation of an immunoglobulin heavy gene enhancer to the c-myc gene drives the formation of Burkitt's lymphoma. Translocations are commonly observed in human cancers (Mitelman et al. 2007). Consequently, how much freedom of movement a double-strand break has will influence the probability of a translocation with a site more distant in nuclear space. Consistent with this expectation, it has been shown by several groups that the proximity of two genes in the normal interphase nucleus of a tissue is correlated with the probability of a translocation occurring between these two genes (Lin et al. 2009; Meaburn et al. 2007; Roix et al. 2003). Nonetheless, chromosome conformation capture assays and genome sequencing reveals that proximity is not the primary determinant of the probability of a translocation and that other factors such as transcriptional activity and the probability of a DSB being generated in the locus are more important considerations (Bunting and Nussenzweig 2013).

In order to appreciate the movement of broken DNA, it is important to know the mobility of undamaged loci within the living cell nucleus. This has been possible to study using the Lac operator (LacO)/LacI repressor system (Belmont and Straight 1998; Chubb et al. 2002) where a fluorescent LacI protein can be targeted to a specific region of a chromosome and the movement of the resulting spot can be

measured versus time. In mammalian cells, these studies show that individual loci exhibit diffusion by random walk confined to an area that is typically about 0.1–0.5 μm^2 depending on their subnuclear position relative to the nuclear lamina or nucleolus (Soutoglou and Misteli 2007) (Chubb et al. 2002). Although there is ATP-dependence to the motion that has been observed, the random walk taken by the individual loci indicate that an active transport process is not responsible for the movement. The ATP-dependence could reflect more local chromatin dynamics related to unwinding the chromatin fibre (Krawczyk et al. 2012; Soutoglou and Misteli 2007) although direct testing of a subset of chromatin remodelling factors suggest this may not be the case (Becker et al. 2014). It has also been possible to monitor the movement of large segments of chromatin by using photoactivation techniques to render specific regions of the genome fluorescent. Strickfaden et al. used photoactivatable histone H4 and, using a laser, photoactivated a grid system into the genome of living nuclei (Strickfaden et al. 2010). When this is done, a much larger contribution to relative motion is observed. These experiments revealed that, while the chromatin showed minimal diffusion over a period of hours, rotation of the nucleus could change the relative positioning of one region of the genome relative to another quite dramatically and can bring together sequences that are initially separated by several microns or separate by as much as the entire diameter of the nucleus sequences that initially sat in vertical alignment (Strickfaden et al. 2010). The latter observation reflects the tethering of the cellular chromatin to the lamina, enabling chromosomes attached to the upper lamina and those attached to the lower lamina to be physically separated through a treadmill-like rotation of the entire nucleus within the cytoplasm, at least in cells with a flattened phenotype.

The LacO/LacI array system, which has proven so valuable in studying chromatin dynamics and motion generally, has also been adapted to study the mobility of DSBs. Soutoglou and colleagues generated an NIH3T3 cell line containing a single integrated copy of an ISce1 restriction enzyme site flanked on one side by lac repressor binding sites and on the other side by Tet repressor (TetR) binding sites (Soutoglou et al. 2007). Transfection with cyan fluorescent protein (CFP)-LacI and yellow fluorescent protein (YFP)-TetR allowed them to identify either side of the ISceI site with a different fluorescent colour. In the absence of ISceI, the two sites remain together, as expected. In the presence of ISceI, the two sides show a slight increase in separation (average 0.2 μm versus 0.1 μm in untreated cells) but remain in close proximity, indicating positional stability of the two DSBs (Soutoglou et al. 2007). In contrast, knockdown of Ku80 increased the movement of DSBs, indicating that Ku proteins may be important in tethering DSBs (Soutoglou et al. 2007). Interestingly, Ku80 binds to polymeric actin and inhibition of actin polymerization inhibits the repair of DSBs (Andrin et al. 2012). It has been proposed that this limited mobility serves to increase the sampling of the local region of the break for homologous sequences for homologous recombination repair (Girst et al. 2013).

Using a variation of this system, where there are separate arrays on separate mouse chromosomes that are tagged with either the tetracycline repressor or the lac

operon, the question of how DSBs from separate chromosomes or distant regions of the same chromosome can drive translocations could be addressed. This allows the use of fusion proteins with fluorescent protein tags to identify each array. Interestingly, when DSBs are induced in both loci simultaneously, they often come in association but the co-association of both DSBs at a single site only resulted in translocations at a rate of one in three hundred (Roukos et al. 2013). The DSBs that were involved in translocations tended to be more mobile, surveying a larger portion of the nuclear area, but for 80 % of the loci, they were found within 1.5 μm of each other in the hour before translocation (Roukos et al. 2013). Several proteins were found to alter the likelihood of a translocation forming despite not altering the formation of pairing between the separate breaks. DNA PKcs inhibited translocations whereas MRE11 promoted translocations (Roukos et al. 2013). This indicates that there is an additional control on the specificity of the re-joining event that is independent of spatial proximity.

We have purposely not discussed the results with *S. cerevisiae* because there are important differences in the complexity of the signalling from DSBs in mammals compared to yeast, and the yeast nucleus is only 2 μm in diameter which is about the same size as the estimated freedom of movement of "immobile" loci in mammalian cells (Roukos et al. 2013). However, it is clear in yeast that individual breaks can coalesce to single "repair centers" (Lisby et al. 2003). This had been suggested with some of the initial analysis in mammalian cells, which involved the study of the evolution of tracks of DSBs produced by high linear energy transfer (LET) radiation.

In some experimental configurations, the source can be directed parallel to a coverslip, enabling the formation of a line of damage across a nucleus, analogous to what is achieved with laser microirradiation (Hauptner et al. 2006). Using this approach, foci were often fewer, larger and more widely spaced along the track, particularly in G1 cells, with increasing time post damage. This was interpreted as indicating that individual DSBs were being clustered into single DRFs, and that clustering increased with time and required energy (Aten et al. 2004); an interpretation that has been challenged and remains controversial in light of conflicting studies conducted by other groups (Jakob et al. 2009a, b). Nonetheless, it is clear that DSBs can move, that their movement on the time scale of an hour or less is quite limited (Roukos et al. 2013; Girst et al. 2013; Jakob et al. 2009a, b; Soutoglou et al. 2007; Aten et al. 2004; Krawczyk et al. 2012), and that this limited movement on shorter time scales is likely to minimize the misrepair of DSBs (Girst et al. 2013). Nonetheless, because translocations do occur, it is clear that more than one DSB can end up in the same nuclear location in space and time.

Most of the movement that has been described, energy-dependent or not, takes place as a random walk and, consequently, is not directional. There is also a type of DRF movement that can be considered non-random. That is the movement of DRFs from the interior of heterochromatin to the periphery of heterochromatin. It had been noted previously that heterochromatin regions were unexpectedly depleted in DRFs [reviewed in (Chiolo et al. 2013)]. This could indicate either that DSBs are poorly accessible in heterochromatin or that these breaks rapidly translocate to the

surface of heterochromatin. This was recently addressed directly by timelapse microscopy, where the translocation of DRFs to the periphery (surface) of heterochromatin occurs rapidly (first 20 min post DNA damage), involves decondensation of the heterochromatin domain, and, under normal conditions, was a requirement for the assembly of Rad51 into the foci (Chiolo et al. 2013; Jakob et al. 2011). These results support arguments against accessibility of DNA damage within heterochromatin domains being a barrier to responding to DSBs but provide a clear example of a type of directional movement of a DSB. Notably, though, the movement outside of the heterochromatin domain is mediated by a reorganization of the heterochromatin domain rather than a directed transport of individual DRF through heterochromatin to reach the surface.

3 What Is the Relationship Between DNA Damage Response and Repair Proteins and the DNA Sequences Found in DRFs?

Early estimates of the amount of DNA that is found in DSB repair foci was obtained by estimations based on stoichiometry relative to the number of DSBs and mitotic staining. This estimated 1–2 million base pairs of DNA could be found inside these foci (Rogakou et al. 1998, 1999). Chromatin immunoprecipitation (ChIP) allows the mapping of proteins or modified forms of proteins relative to the underlying DNA sequence and this approach has been used to map proteins relative to the double-strand break. To investigate the arrangement of DNA damage response proteins relative to the underlying DNA sequence using ChIP, it is necessary to introduce a DNA double-strand break within a defined DNA sequence. With the advent of targeted nuclease systems, this has become possible. The initial experiments were done using integrated reporter systems containing a recognition site for the endonuclease ISceI, which does not naturally occur in the genome of mammals. For example, mapping of Rad51 assembly showed a strong enrichment in the sequences immediately surrounding the DSB but dissipated as the distance increased (Rodrigue et al. 2006). This would be expected for a protein that is responsible for alignment of homologous sequences to repair the DSB through homologous recombination. Similarly, Ku80 was found in highest abundance near the break, but, in this case, predominantly in G1, when homologous recombination is inactive (Rodrigue et al. 2006). The first mammalian experiments involving ChIP mapping around endogenous genomic sites where DSBs have been generated using a transfected nuclease made use of an inducible PpoI enzyme whose nuclear entry was regulated by tamoxifen. This approach revealed that ATM and NBS1 are located proximal to the DSB and that phosphorylated histone H2AX is depleted near the DSB but highly enriched beginning 1–2 kbp distance from the DSB (Berkovich et al. 2007). The depletion near the DSB is associated with an apparent reduction in nucleosome density.

The most detailed analysis employed the nuclease AsiSI (Iacovoni et al. 2010). This restriction enzyme has an 8 nucleotide recognition sequence and cuts relatively infrequently throughout the human genome. Importantly, not all potential sites are cleaved by the restriction enzyme, which implies that chromatin structure regulates the access that the restriction enzyme has to the DNA sequence. Those that are not cut tend to be in regions of the genome that are highly methylated. Consequently, the sampling may be biased for euchromatic sequences and the results may not apply to heterochromatin domains. In this case, active genes, particularly promoters, were found to be depleted in phosphorylated histone H2AX despite being embedded in chromosomal domains enriched in phosphorylated histone H2AX. The phosphorylation of histone H2AX around the DSB is typically asymmetrical and uneven, as well as anti-correlated with histone H3 lysine 4 trimethylation, which is enriched in promoter sequences (Iacovoni et al. 2010; Seiler et al. 2011). These results indicate that aspects of chromatin structure regulate the spreading of histone H2AX phosphorylation away from the DSB and the irregularity of the H2AX phosphorylation might explain how proteins can enrich within DRFs but not colocalize with phosphorylated histone H2AX or each other, as has been seen by super-resolution fluorescence microscopy. Some of the biochemical mechanisms that regulate spreading have been identified and include histone acetylation and the recruitment of the acetylated histone binding protein Brd4 [reviewed in (Panier and Durocher 2013)].

3.1 Why Are So Many Copies of Proteins Assembled at Sites of DRFs?

The dominant hypothesis regarding the "massive" accumulation of DNA repair factors at DRF is that concentrating these molecules serves to rapidly amplify signalling such that a single DSB is sufficient to induce a DNA damage response and arrest the cell cycle (Misteli and Soutoglou 2009; Soutoglou and Misteli 2008). This hypothesis is driven by the observations that activation of the G2/M checkpoint at low doses (less than 2 Gy, typically 0.5 Gy) of ionizing radiation requires histone H2AX (Fernandez-Capetillo et al. 2002) and the mediator of damage checkpoint 1 (MDC1) protein (Lou et al. 2006). MDC1 functions as a protein scaffold for H2AX-mediated assembly of DSB repair compartments (Jungmichel and Stucki 2010). Thus, the formation of DSB repair compartments appears to be functionally critical only when there are small numbers of DSBs (Misteli and Soutoglou 2009; Soutoglou and Misteli 2008). This level of DNA damage, however, is more physiological than the high doses often employed for experimental characterization of the DDR. Interestingly, the low dose G2/M checkpoint defect in H2AX−/− cells mirrors that of ATM−/− cells (Fernandez-Capetillo et al. 2002). Thus, we favour the explanation that DRF formation is necessary for rapidly

amplifying the ATM kinase signalling cascade and that this plays an essential role in preserving genome fidelity when only small numbers of breaks are present.

4 Summary

DRFs are dynamic nuclear structures that form in response to DNA double-strand breaks. They form within the first few minutes post damage through a process that is driven by the activation of the ATM kinase and the phosphorylation of histone H2AX. The purpose of compartment formation has not been determined conclusively. However, there is good reason to believe, based on shared properties of knockout cell lines that fail to form DRFs, that DRFs form to amplify signalling and that this amplification is particularly important for the amplification of ATM signalling in order to establish a G2 arrest and prevent cells from entering mitosis with DSBs.

Much of the work done on the structure of these foci has taken place within the first few minutes up to the first hour post DNA damage. It is clear from transmission electron microscopy studies and from previous fluorescence microscopy studies that this structure may evolve over time (Dellaire et al. 2009; Maser et al. 1997; Rogakou et al. 1999; Bewersdorf et al. 2006).

Under conditions of normal repair, these foci eventually disassemble, presumably upon completion of the repair of the DSB. While many enzymes involved in the reversal of the DDR response have been identified [reviewed in (Panier and Durocher 2013)], how these are controlled locally to disassemble foci individually as the DSB is repaired and within a background of active DDR signalling remains to be determined. This is something that we need to better understand and is clearly relevant to compartmentalization because disassembly of individual DRFs can take place in a background where DDR signalling is maintained. In other words, DRF disassembly is not synchronous, as would be anticipated if these enzymes were being regulated nucleus-wide in the nucleoplasm.

A considerable challenge in connecting function with structures in the nucleoplasm is that the same interactions that are necessary for proteins to function in comparably small structures (protein complexes) are commonly necessary for proteins to remain associated with their resident compartments in the nucleus. With the vast numbers of known modified proteins generated in response to DSBs and the identification of alternative modules that could be used to target some of these signalling enzymes to DRFs in the absence of their normal partners, thereby allowing separation of targeting from assembly into a specific macromolecular complex, DRFs may be a model system that can be exploited to quantitatively define the role of compartmentalization in DDR signalling.

Acknowledgments Work in the Hendzel laboratory is supported by the Canadian Institutes of Health Research. Hilmar Strickfaden is funded by a post-doctoral Fellowship of the Alberta Cancer Foundation. We also want to thank the “Bayrische Forschungsallianz” for support.

References

Abdisalaam S, Davis AJ, Chen DJ, Alexandrakis G (2014) Scanning fluorescence correlation spectroscopy techniques to quantify the kinetics of DNA double strand break repair proteins after gamma-irradiation and bleomycin treatment. Nucleic Acids Res 42(1), e5. doi:10.1093/nar/gkt908

Andrin C, McDonald D, Attwood KM, Rodrigue A, Ghosh S, Mirzayans R, Masson JY, Dellaire G, Hendzel MJ (2012) A requirement for polymerized actin in DNA double-strand break repair. Nucleus 3(4):384–395. doi:10.4161/nucl.21055

Aten JA, Stap J, Krawczyk PM, van Oven CH, Hoebe RA, Essers J, Kanaar R (2004) Dynamics of DNA double-strand breaks revealed by clustering of damaged chromosome domains. Science (New York, NY) 303(5654):92–95. doi:10.1126/science.1088845

Bassing CH, Suh H, Ferguson DO, Chua KF, Manis J, Eckersdorff M, Gleason M, Bronson R, Lee C, Alt FW (2003) Histone H2AX: a dosage-dependent suppressor of oncogenic translocations and tumors. Cell 114(3):359–370

Becker A, Durante M, Taucher-Scholz G, Jakob B (2014) ATM alters the otherwise robust chromatin mobility at sites of DNA double-strand breaks (DSBs) in human cells. PLoS One 9(3), e92640. doi:10.1371/journal.pone.0092640

Belmont AS, Straight AF (1998) In vivo visualization of chromosomes using lac operator-repressor binding. Trends Cell Biol 8(3):121–124

Berkovich E, Monnat RJ, Kastan MB (2007) Roles of ATM and NBS1 in chromatin structure modulation and DNA double-strand break repair. Nat Cell Biol 9(6):683–690. doi:10.1038/ncb1599

Bewersdorf J, Bennett BT, Knight KL (2006) H2AX chromatin structures and their response to DNA damage revealed by 4Pi microscopy. Proc Natl Acad Sci U S A 103(48):18137–18142. doi:10.1073/pnas.0608709103

Bunting SF, Nussenzweig A (2013) End-joining, translocations and cancer. Nat Rev Cancer 13 (7):443–454. doi:10.1038/nrc3537

Burma S, Chen BP, Murphy M, Kurimasa A, Chen DJ (2001) ATM phosphorylates histone H2AX in response to DNA double-strand breaks. J Biol Chem 276(45):42462–42467. doi:10.1074/jbc.C100466200

Celeste A, Difilippantonio S, Difilippantonio MJ, Fernandez-Capetillo O, Pilch DR, Sedelnikova OA, Eckhaus M, Ried T, Bonner WM, Nussenzweig A (2003a) H2AX haploinsufficiency modifies genomic stability and tumor susceptibility. Cell 114(3):371–383

Celeste A, Fernandez-Capetillo O, Kruhlak MJ, Pilch DR, Staudt DW, Lee A, Bonner RF, Bonner WM, Nussenzweig A (2003b) Histone H2AX phosphorylation is dispensable for the initial recognition of DNA breaks. Nat Cell Biol 5(7):675–679. doi:10.1038/ncb1004

Celeste A, Petersen S, Romanienko PJ, Fernandez-Capetillo O, Chen HT, Sedelnikova OA, Reina-San-Martin B, Coppola V, Meffre E, Difilippantonio MJ, Redon C, Pilch DR, Olaru A, Eckhaus M, Camerini-Otero RD, Tessarollo L, Livak F, Manova K, Bonner WM, Nussenzweig MC, Nussenzweig A (2002) Genomic instability in mice lacking histone H2AX. Science 296 (5569):922–927. doi:10.1126/science.1069398

Chapman JR, Sossick AJ, Boulton SJ, Jackson SP (2012) BRCA1-associated exclusion of 53BP1 from DNA damage sites underlies temporal control of DNA repair. J Cell Sci 125 (Pt 15):3529–3534. doi:10.1242/jcs.105353

Chen WT, Alpert A, Leiter C, Gong F, Jackson SP, Miller KM (2013) Systematic identification of functional residues in mammalian histone H2AX. Mol Cell Biol 33(1):111–126. doi:10.1128/MCB.01024-12

Chiolo I, Minoda A, Colmenares SU, Polyzos A, Costes SV, Karpen GH (2011) Double-strand breaks in heterochromatin move outside of a dynamic HP1a domain to complete recombinational repair. Cell 144(5):732–744. doi:10.1016/j.cell.2011.02.012

Chiolo I, Tang J, Georgescu W, Costes SV (2013) Nuclear dynamics of radiation-induced foci in euchromatin and heterochromatin. Mutat Res 750(1-2):56–66. doi:10.1016/j.mrfmmm.2013.08.001

Chubb JR, Boyle S, Perry P, Bickmore WA (2002) Chromatin motion is constrained by association with nuclear compartments in human cells. Curr Biol 12(6):439–445

Dellaire G, Kepkay R, Bazett-Jones DP (2009) High resolution imaging of changes in the structure and spatial organization of chromatin, gamma-H2A.X and the MRN complex within etoposide-induced DNA repair foci. Cell Cycle 8(22):3750–3769

Dickey JS, Redon CE, Nakamura AJ, Baird BJ, Sedelnikova OA, Bonner WM (2009) H2AX: functional roles and potential applications. Chromosoma 118(6):683–692. doi:10.1007/s00412-009-0234-4

Dundr M, Misteli T (2010) Biogenesis of nuclear bodies. Cold Spring Harb Perspect Biol 2(12):a000711. doi:10.1101/cshperspect.a000711

Essers J, Hendriks RW, Swagemakers SM, Troelstra C, de Wit J, Bootsma D, Hoeijmakers JH, Kanaar R (1997) Disruption of mouse RAD54 reduces ionizing radiation resistance and homologous recombination. Cell 89(2):195–204

Essers J, Houtsmuller AB, van Veelen L, Paulusma C, Nigg AL, Pastink A, Vermeulen W, Hoeijmakers JH, Kanaar R (2002) Nuclear dynamics of RAD52 group homologous recombination proteins in response to DNA damage. EMBO J 21(8):2030–2037. doi:10.1093/emboj/21.8.2030

Fernandez-Capetillo O, Chen H-T, Celeste A, Ward I, Romanienko PJ, Morales JC, Naka K, Xia Z, Camerini-Otero RD, Motoyama N, Carpenter PB, Bonner WM, Chen J, Nussenzweig A (2002) DNA damage-induced G2-M checkpoint activation by histone H2AX and 53BP1. Nat Cell Biol 4(12):993–997. doi:10.1038/ncb884

Girst S, Hable V, Drexler GA, Greubel C, Siebenwirth C, Haum M, Friedl AA, Dollinger G (2013) Subdiffusion supports joining of correct ends during repair of DNA double-strand breaks. Sci Rep 3:2511. doi:10.1038/srep02511

Gong F, Miller KM (2013) Mammalian DNA repair: HATs and HDACs make their mark through histone acetylation. Mutat Res 750(1-2):23–30. doi:10.1016/j.mrfmmm.2013.07.002

Haaf T, Golub EI, Reddy G, Radding CM, Ward DC (1995) Nuclear foci of mammalian Rad51 recombination protein in somatic cells after DNA damage and its localization in synaptonemal complexes. Proc Natl Acad Sci U S A 92(6):2298–2302

Hauptner A, Krucken R, Greubel C, Hable V, Dollinger G, Drexler GA, Deutsch M, Lowe R, Friedl AA, Dietzel S, Strickfaden H, Cremer T (2006) DNA-repair protein distribution along the tracks of energetic ions. Radiat Prot Dosimetry 122(1-4):147–149. doi:10.1093/rpd/ncl420

Iacovoni JS, Caron P, Lassadi I, Nicolas E, Massip L, Trouche D, Legube G (2010) High-resolution profiling of gammaH2AX around DNA double strand breaks in the mammalian genome. EMBO J 29(8):1446–1457. doi:10.1038/emboj.2010.38

Jakob B, Splinter J, Conrad S, Voss KO, Zink D, Durante M, Lobrich M, Taucher-Scholz G (2011) DNA double-strand breaks in heterochromatin elicit fast repair protein recruitment, histone H2AX phosphorylation and relocation to euchromatin. Nucleic Acids Res 39(15):6489–6499. doi:10.1093/nar/gkr230

Jakob B, Splinter J, Durante M, Taucher-Scholz G (2009a) Live cell microscopy analysis of radiation-induced DNA double-strand break motion. Proc Natl Acad Sci U S A 106(9):3172–3177. doi:10.1073/pnas.0810987106

Jakob B, Splinter J, Taucher-Scholz G (2009b) Positional stability of damaged chromatin domains along radiation tracks in mammalian cells. Radiat Res 171(4):405–418. doi:10.1667/RR1520.1

Jungmichel S, Stucki M (2010) MDC1: The art of keeping things in focus. Chromosoma 119(4):337–349. doi:10.1007/s00412-010-0266-9

Kaiser TE, Intine RV, Dundr M (2008) De novo formation of a subnuclear body. Science 322(5908):1713–1717. doi:10.1126/science.1165216

Krawczyk PM, Borovski T, Stap J, Cijsouw T, ten Cate R, Medema JP, Kanaar R, Franken NAP, Aten JA (2012) Chromatin mobility is increased at sites of DNA double-strand breaks. J Cell Sci 125(Pt 9):2127–2133. doi:10.1242/jcs.089847

Lin C, Yang L, Tanasa B, Hutt K, Ju BG, Ohgi K, Zhang J, Rose DW, Fu XD, Glass CK, Rosenfeld MG (2009) Nuclear receptor-induced chromosomal proximity and DNA breaks underlie specific translocations in cancer. Cell 139(6):1069–1083. doi:10.1016/j.cell.2009.11.030

Lisby M, Mortensen UH, Rothstein R (2003) Colocalization of multiple DNA double-strand breaks at a single Rad52 repair centre. Nat Cell Biol 5(6):572–577. doi:10.1038/ncb997

Lorat Y, Schanz S, Schuler N, Wennemuth G, Rube C, Rube CE (2012) Beyond repair foci: DNA double-strand break repair in euchromatic and heterochromatic compartments analyzed by transmission electron microscopy. PLoS One 7(5), e38165. doi:10.1371/journal.pone.0038165

Lou Z, Minter-Dykhouse K, Franco S, Gostissa M, Rivera MA, Celeste A, Manis JP, van Deursen J, Nussenzweig A, Paull TT, Alt FW, Chen J (2006) MDC1 maintains genomic stability by participating in the amplification of ATM-dependent DNA damage signals. Mol Cell 21(2):187–200. doi:10.1016/j.molcel.2005.11.025

Mao YS, Zhang B, Spector DL (2011) Biogenesis and function of nuclear bodies. Trends Genet 27 (8):295–306. doi:10.1016/j.tig.2011.05.006

Maser RS, Monsen KJ, Nelms BE, Petrini JH (1997) hMre11 and hRad50 nuclear foci are induced during the normal cellular response to DNA double-strand breaks. Mol Cell Biol 17 (10):6087–6096

Meaburn KJ, Misteli T, Soutoglou E (2007) Spatial genome organization in the formation of chromosomal translocations. Semin Cancer Biol 17(1):80–90. doi:10.1016/j.semcancer.2006.10.008

Misteli T, Soutoglou E (2009) The emerging role of nuclear architecture in DNA repair and genome maintenance. Nat Rev Mol Cell Biol 10(4):243–254

Mitelman F, Johansson B, Mertens F (2007) The impact of translocations and gene fusions on cancer causation. Nat Rev Cancer 7(4):233–245. doi:10.1038/nrc2091

Mok MTS, Henderson BR (2012a) The in vivo dynamic interplay of MDC1 and 53BP1 at DNA damage-induced nuclear foci. Int J Biochem Cell Biol 44(9):1398–1409. doi:10.1016/j.biocel.2012.05.025

Mok MTS, Henderson BR (2012b) The in vivo dynamic organization of BRCA1-A complex proteins at DNA damage-induced nuclear foci. Traffic (Copenhagen, Denmark) 13 (6):800–814. doi:10.1111/j.1600-0854.2012.01355.x

Panier S, Durocher D (2013) Push back to respond better: regulatory inhibition of the DNA double-strand break response. Nat Rev Mol Cell Biol 14(10):661–672. doi:10.1038/nrm3659

Rodrigue A, Lafrance M, Gauthier MC, McDonald D, Hendzel M, West SC, Jasin M, Masson JY (2006) Interplay between human DNA repair proteins at a unique double-strand break in vivo. EMBO J 25(1):222–231. doi:10.1038/sj.emboj.7600914

Rogakou EP, Boon C, Redon C, Bonner WM (1999) Megabase chromatin domains involved in DNA double-strand breaks in vivo. J Cell Biol 146(5):905–916

Rogakou EP, Pilch DR, Orr AH, Ivanova VS, Bonner WM (1998) DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. J Biol Chem 273(10):5858–5868

Roix JJ, McQueen PG, Munson PJ, Parada LA, Misteli T (2003) Spatial proximity of translocation-prone gene loci in human lymphomas. Nat Genet 34(3):287–291. doi:10.1038/ng1177

Roukos V, Voss TC, Schmidt CK, Lee S, Wangsa D, Misteli T (2013) Spatial dynamics of chromosome translocations in living cells. Science 341(6146):660–664. doi:10.1126/science.1237150

Rube CE, Lorat Y, Schuler N, Schanz S, Wennemuth G, Rube C (2011) DNA repair in the context of chromatin: new molecular insights by the nanoscale detection of DNA repair complexes using transmission electron microscopy. DNA Repair 10(4):427–437. doi:10.1016/j.dnarep.2011.01.012

Seiler DM, Rouquette J, Schmid VJ, Strickfaden H, Ottmann C, Drexler GA, Mazurek B, Greubel C, Hable V, Dollinger G, Cremer T, Friedl AA (2011) Double-strand break-induced transcriptional silencing is associated with loss of tri-methylation at H3K4. Chromosome Res 19(7):883–899. doi:10.1007/s10577-011-9244-1

Shevtsov SP, Dundr M (2011) Nucleation of nuclear bodies by RNA. Nat Cell Biol 13(2):167–173. doi:10.1038/ncb2157

Sleeman JE, Trinkle-Mulcahy L (2014) Nuclear bodies: new insights into assembly/dynamics and disease relevance. Curr Opin Cell Biol 28:76–83. doi:10.1016/j.ceb.2014.03.004

Soutoglou E, Dorn JF, Sengupta K, Jasin M, Nussenzweig A, Ried T, Danuser G, Misteli T (2007) Positional stability of single double-strand breaks in mammalian cells. Nat Cell Biol 9 (6):675–682. doi:10.1038/ncb1591

Soutoglou E, Misteli T (2007) Mobility and immobility of chromatin in transcription and genome stability. Curr Opin Genet Dev 17(5):435–442. doi:10.1016/j.gde.2007.08.004

Soutoglou E, Misteli T (2008) Activation of the cellular DNA damage response in the absence of DNA lesions. Science (New York, NY) 320(5882):1507–1510. doi:10.1126/science.1159051

Strickfaden H, Zunhammer A, van Koningsbruggen S, Köhler D, Cremer T (2010) 4D chromatin dynamics in cycling cells: Theodor Boveri's hypotheses revisited. Nucleus (Austin, Tex) 1 (3):284–297. doi:10.4161/nucl.1.3.11969

Ye Q, Hu YF, Zhong H, Nye AC, Belmont AS, Li R (2001) BRCA1-induced large-scale chromatin unfolding and allele-specific effects of cancer-predisposing mutations. J Cell Biol 155 (6):911–921. doi:10.1083/jcb.200108049

Nuclear Domains and DNA Repair

Jordan Pinder, Alkmini Kalousi, Evi Soutoglou, and Graham Dellaire

Abstract The eukaryotic genome is being constantly subjected to DNA damage arising from environmental and endogenous mutagens. Rapid and faithful repair of damaged DNA is an essential process to life, and in multicellular organisms the presence of irreparable DNA damage induces cell cycle arrest to prevent propagation of a damaged genome, which can lead to cancer. DNA repair is mediated by the DNA damage response, a network of proteins that sense, signal, and repair through the ordered recruitment of DNA repair factors to sites of damage. Repair of DNA occurs in discrete regions of the nucleus called DNA repair foci, whose composition and localization to various nuclear subdomains depends on the stage of DNA repair and the severity of damage. In this chapter we will discuss two nuclear subdomains, the nuclear lamina, and promyelocytic leukemia nuclear bodies (PML-NBs), and how the DNA damage response is regulated at these domains.

Keywords PML nuclear body • Nuclear lamina • DNA recombination • DNA repair foci

J. Pinder
Department of Development and Stem Cells, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Laboratory of Cell Biology of Genome Integrity, Parc d'Innovation 1, Rue Laurent Fries, Illkirch 67404, France

A. Kalousi • E. Soutoglou (✉)
Department of Development and Stem Cells, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Laboratory of Cell Biology of Genome Integrity, Parc d'Innovation 1, Rue Laurent Fries, Illkirch 67404, France

Institut National de la Santé et de la Recherche Médicale, U964, Illkirch, France

Centre National de Recherche Scientifique, UMR7104, Illkirch, France

Université de Strasbourg, Illkirch, France
e-mail: evisou@igbmc.fr

G. Dellaire (✉)
Departments of Pathology and Biochemistry & Molecular Biology, Dalhousie University, P.O. Box 15000, Halifax, NS, Canada, B3H 4R2
e-mail: dellaire@dal.ca

D.P. Bazett-Jones, G. Dellaire (eds.), *The Functional Nucleus*,
DOI 10.1007/978-3-319-38882-3_11

1 Introduction

In addition to enclosing the dense chromatin fibers that package about two metres of DNA into a 6 μm-diameter sphere, the nucleus is also home to a number of distinct protein-rich subdomains. These structures are not membrane-bound, but are spatially constrained and have defined components and functions. At the nuclear periphery, the nuclear lamina is a fibrillar network of proteins that regulates cell division and chromatin organization. At the nuclear interior, numerous defined substructures have been identified. The largest of these, the nucleolus, is the site of ribosomal RNA synthesis. Others include promyelocytic leukemia nuclear bodies (PML-NBs), the perinucleolar compartment (PNC), Polycomb Group bodies, nuclear speckles, Cajal bodies, OPT (Oct1, PTF transcription) domains, gems, cleavage bodies, and Sam68 nuclear bodies (Dellaire and Bazett-Jones 2007; Dundr and Misteli 2001; Spector 2001). This chapter will discuss two nuclear structures, the nuclear lamina and PML-NBs, and will focus on how these subdomains regulate the DNA damage response.

The nuclear lamina is located under the inner nuclear membrane and is a filamentous structure that is composed of type A and B lamins and a growing number of lamin-associated proteins (Bridger et al. 2007; Burke and Stewart 2013). The integrity of the nuclear lamina is essential for the regulation of protein and RNA trafficking in and out of the nucleus via nucleopores. In addition to separating the cytoplasm from the nuclear contents and functions such as DNA transcription, replication and repair, the nuclear lamina also plays an important role in chromosome positioning in eukaryotes (Kind and van Steensel 2010; Padeken and Heun 2014) with evidence for both a general role of the lamina in silencing of genes as well as "gene-gating" of certain loci to facilitate cytoplasmic transport of specific messenger RNAs (Burns and Wente 2014). The nuclear lamina is also highly dynamic during both the cell cycle, where phosphorylation of the lamins regulates nuclear envelope breakdown during mitosis, and in response to genotoxic stress (Dechat et al. 2010). For example, induction of DNA damage with the topoisomerase inhibitor VP16 (etoposide) resulted in the formation of folds in the lamina and invaginations of the nuclear envelope within minutes of treating cells (Dellaire et al. 2009). However, the consequences or biological function of these invaginations remains to be elucidated. Perturbation of the nuclear lamina caused by mutations in lamin A also underly an aging syndrome known as Hutchinson-Guilford Progeria (Mounkes and Stewart 2004). This form of progeria is characterized by misshapen nuclei, defects in the nuclear lamina and increased genomic instability (Goldman et al. 2004; Liu et al. 2005; Scaffidi and Misteli 2006). These observations suggest a role for the nuclear lamina directly or indirectly in DNA repair and the maintenance of genomic instability, which will be explored in more detail later in this chapter.

PML-NBs were first identified by electron microscopy as interchromosomal structures, and are structurally organized by the PML protein (de The et al. 1960). PML is encoded by the promyelocytic leukemia (PML) gene, which

was identified at the site of a chromosomal translocation associated with the development of acute promyelocytic leukemia (APL). The translocation causes fusion of the *PML* gene with that of the retinoic acid receptor alpha (RARα). The resulting gene encodes the fusion oncoprotein, PML-RARα, which disrupts PML-NBs and affects promyelocyte survival and differentiation during hematopoiesis (de The et al. 1991; Kakizuka et al. 1991). APL can be effectively treated by arsenic trioxide and retinoic acid, which trigger degradation of PML-RARα and allow for reconstitution of functional PML-NBs, restoring promyelocyte differentiation (Zhu et al. 2001). Though numerous functions have been ascribed to PML-NBs, this chapter will concentrate on their roles in cell fate decisions and in DNA repair.

2 The Nuclear Lamina and DNA Repair

2.1 The Role of Inner Nuclear Membrane Proteins in DNA Repair

The first indications of a role for lamins in DNA repair came from studies of *zmpste24*$^{-/-}$ mice, which lack the metalloproteinase that processes prelamin A, and in cells from patients with laminopathies like Hutchinson-Gilford progeria syndrome (HGPS), caused by mutations in the *LMNA* gene that encodes lamin A. These cells exhibit elevated levels of endogenous DNA damage, increased genomic instability, persistent DNA lesions, and increased sensitivity to DNA damaging agents (Varela et al. 2005; Scaffidi and Misteli 2006; Liu et al. 2005). More extensive investigation using cells from patients from HGPS and other laminopathies revealed that enhanced endogenous DNA damage is due to increased levels of reactive oxygen species (ROS) produced in these patients (Richards et al. 2011). Treating these cells with ROS scavengers decreases the basal level of DNA damage (Richards et al. 2011). Moreover, genomic instability and inability to repair exogenously induced DNA damage was demonstrated in fibroblasts of patients with the laminopathy mandibuloacral dysplasia type A (MADA) that carry the homozygous mutation R527H in lamin A (di Masi et al. 2008). This may be due to an inability to activate the DDR, including p53 and its targets in these patient-derived cells (di Masi et al. 2008).

In another study, the contribution of disease-causing lamin A mutations to genomic instability was evaluated by expression of the mutant lamin A proteins (lamin A-Δ50, R471C, R527C and L530P) in HeLa cells (Manju et al. 2006). Formation of γ-H2AX and 53BP1 DNA repair foci in response to DNA damage was markedly reduced in these cells compared to cells expressing the wild-type lamin A, or the other disease-causing mutants H222P and R482L. Interestingly, lamin A mutants that impaired the formation of DNA repair foci also caused mislocalization of ataxia telangiectasia-mutated (ATM) and Rad3-related (ATR)

kinase, a key sensor in the response to DNA damage (Manju et al. 2006). It is still not clear why some of the disease-causing lamin A mutations do not appear to be sufficient to cause a defect in the DDR.

Other mechanisms have been proposed to explain the apparent genomic instability in progeria. In *zmpste24*$^{-/-}$ mouse embryonic fibrobasts (MEFs), the recruitment of 53BP1 and RAD51 is impaired, leading to a delay in activation of the DDR and defective DNA repair (Liu et al. 2005). In line with these observations, ectopic expression of unprocessed prelamin A is sufficient to induce defects in checkpoint activation and DNA repair, suggesting that unprocessed prelamin A acts in a dominant negative fashion to perturb DNA repair, leading to genomic instability (Liu et al. 2005). More extensive work in *zmpste24*$^{-/-}$ MEFs and HGPS patient cells revealed an abnormal accumulation of the xeroderma pigmentosum group A protein (XPA) that leads to aberrant activation of ATM and ATR signaling responsible for the apparent proliferation arrest. In these studies, depletion of XPA partially rescued the DNA repair defects (Liu et al. 2008). This differs from the mechanism that was proposed to explain the premature aging and senescence in induced pluripotent stem cells (iPSCs)-differentiated smooth muscle cells derived from HGPS (Liu et al. 2011). These cells as well as HGPS cells exhibit low levels of the DNAPKCs, which are important for DNA repair by non-homologous end joining (NHEJ) (Liu et al. 2011).

In addition to the cases where lamin A is mutated or not processed, recent studies have also examined the impact of total loss of lamin A in genome maintenance. MEFs that do not express lamin A (*lmna*$^{-/-}$) have increased genomic instability exemplified by chromosomal abnormalities and higher background levels of γ-H2AX (Gonzalez-Suarez et al. 2009a, b). More recent investigations on the mechanism involved in the progression of the genomic instability, showed that although the first steps of the DDR in the *lmna*$^{-/-}$ MEFs appear to function normally, there is a defect in the formation of 53BP1 foci upon irradiation (Gonzalez-Suarez et al. 2009b). The same study also showed that in *lmna*$^{-/-}$ cells, 53BP1 is degraded, and this phenotype can be suppressed if lamin A is ectopically expressed (Gonzalez-Suarez et al. 2009b). The decreased level of 53BP1 in these cells leads to failure of joining unprotected telomeres by NHEJ and delay of repairing ionizing radiation-induced DNA damage (Gonzalez-Suarez et al. 2009b). The degradation of 53BP1 in the absence of lamin A was not mediated by the proteasome but instead by the lysosomal protease cathepsin D (CTSL). Indeed the levels of CTSL in *lmna*$^{-/-}$ cells are elevated, and its depletion leads to restoration of 53BP1 levels and the delay in DNA repair (Gonzalez-Suarez et al. 2011).

DNA repair by homologous recombination (HR) is also impaired in mouse *lmna*$^{-/-}$ cells and in human cells depleted for lamin A. Depletion of lamin A results in transcriptional downregulation of BRCA1 and RAD51, two genes essential for HR (Das et al. 2013, 410–419). These studies suggest that lamin A is key in regulating both major DNA double-strand break (DSB) repair pathways (Fig. 1).

Although the majority of the above studies examined prolonged depletion of lamin A mainly in mouse cells, transient downregulation of Lamin A in human cells

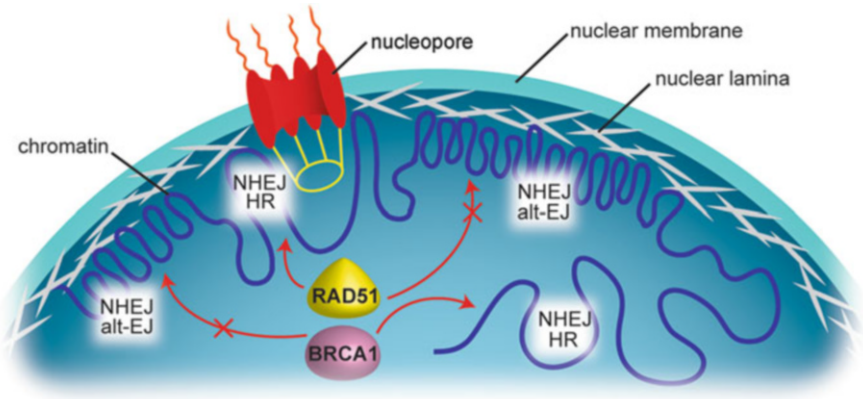

Fig. 1 DNA repair at the nuclear membrane. In mammalian cells DNA double-strand breaks in open chromatin are repaired by non-homologous end joining (NHEJ), or homologous recombination (HR). At the nuclear lamina, chromatin compaction impairs recruitment of homologous recombination repair factors such as BRCA1 and RAD51. Resected breaks that cannot complete repair by HR are instead repaired by alternative end-joining (alt-EJ). Chromatin associated with the nuclear pores is not compacted to the same degree as other regions at the nuclear membrane; therefore, DNA DSBs at the nuclear pores can be repaired by HR

had different phenotypes. It was shown that cells in which lamin A and lamin C were transiently depleted by siRNA have a normal response to ionizing radiation but are sensitive to agents that cause interstrand cross-links (ICLs) or replication stress. In response to treatment with ICL agents (cisplatin, camptothecin, and mitomycin), lamin A/C-deficient cells displayed normal γ-H2AX focus formation but had delayed γ-H2AX removal, decreased recruitment of the FANCD2 repair factor, and a higher frequency of chromosome aberrations (Singh et al. 2013). Similarly, following hydroxyurea-induced replication stress, lamin A/C-deficient cells exhibited defective repair factor recruitment (MRE11, CtIP, RAD51, RPA, and FANCD2) (Singh et al. 2013) and repair delay. These results suggest that lamin A is required for maintenance of genomic instability following replication stress (Singh et al. 2013).

Besides lamin A, other inner nuclear envelope components have been connected with genomic instability. It was reported that the inner nuclear envelope proteins SUN1 and SUN2 play a redundant role in DDR (Lei et al. 2012). MEFs from $sun1^{-/-}$ $sun2^{-/-}$ mice displayed premature proliferation arrest in S phase of cell cycle, increased apoptosis and DNA damage. $sun1^{-/-}$ $sun2^{-/-}$ fibroblasts exhibit impaired DDR activation through the ATM pathway. A biochemical screen identified interactions between SUN1 and SUN2 and DNA-dependent protein kinase (DNAPK) complex (Lei et al. 2012). Although the mechanism of action of SUN proteins in DDR and repair was not further explored, it has been proposed that the alterations of chromatin observed in $sun1^{-/-}$ $sun2^{-/-}$ cells leads to the observed defects in DNA repair (Lei et al. 2012) .

Another recent study shows that the outer nuclear envelope factor Nesprin-1 interacts with the mismatch repair proteins MSH2 and MSH6, and its downregulation leads to persistent DNA damage observed by elevated basal levels of γ-H2AX foci (Sur et al. 2014). In tumor cell lines Nesprin-1 levels are reduced, and this correlates with increased endogenous DNA damage (Sur et al. 2014) These findings indicate a role for Nesprin-1 in the DNA damage response pathway and propose Nesprin-1 as novel player in tumorigenesis and genome instability (Sur et al. 2014).

2.2 DDR and DNA Repair at the Nuclear Lamina

The nuclear periphery has recently emerged as a compartment that regulates DDR and DNA repair pathway choice. In yeast, persistent DSBs and telomeres relocate to the nuclear periphery, either at the inner nuclear membrane through interaction with the integral membrane protein Mps3 (Oza et al. 2009; Kalocsay et al. 2009; Schober et al. 2009) or at the nuclear pores through interaction with the nucleoporin Nup84 (Kalocsay et al. 2009; Khadaroo et al. 2009; Nagai et al. 2008; Therizols et al. 2006). Association with the two different compartments leads to distinct repair outcomes (Horigome et al. 2014). Discrete mechanisms mediate this relocalization of the irreparable DSBs to the nuclear membrane or to the pores, with the SWR-C chromatin remodeler being necessary for both pathways throughout the cell cycle, whereas INO80 is only required for the relocation to the nuclear membrane outside G1 (Horigome et al. 2014). Each nuclear compartment is linked to a different repair pathway. Breaks interacting with Mps3 are repaired by faithful HR pathway and excessive/inappropriate recombination is repressed, though at the pores non-canonical recombination can happen and the breaks can be repaired by error prone pathways like break-induced replication (BIR) (Horigome et al. 2014). Thus, according to the above data, the inner nuclear membrane can be a restrictive environment for recombination events in comparison with the permissive nuclear pores, possibly prompting breaks to relocalize to the pores in case their repair is inefficient at the nuclear membrane.

In *Schizosaccharomyces pombe*, DNA DSBs were also demonstrated to associate with the nuclear envelope proteins Sad1 and Kms1, two members of the LINC (linker of nucleoskeleton and cytoskeleton) complex in S/G2 phases of the cell cycle, connecting the DSBs to cytoplasmic microtubules (Swartz et al. 2014). The same study showed that the LINC component Kms1 and the cytoplasmic microtubule regulator Mto1 promote homology-directed DNA repair by gene conversion. Kms1 genetically interacts with a number of genes involved in HR thereby suggesting a role for the nuclear envelope and cytoskeleton in spatial segregation of DNA repair (Swartz et al. 2014).

Interestingly this spatial segregation of specific DNA repair pathways is conserved across species. Indeed, it was recently demonstrated that breaks induced at the nuclear pores in human cells were repaired by both NHEJ and HR pathways, whereas breaks induced at the nuclear lamina were unable to recruit HR factors and

were instead repaired by NHEJ or by the error-prone alternative end-joining (alt-EJ) mechanism (Lemaitre and Soutoglou 2014; Lemaitre et al. 2014). Although the exact mechanism by which HR is repressed at the nuclear lamina remains elusive, the highly compacted state of chromatin at the nuclear lamina is one of the factors involved in this inhibition, and both global and local induction of chromatin decompaction restored functional HR at the nuclear lamina. These observations suggest that gene positioning determines DNA repair pathway choice in mammals (Lemaitre and Soutoglou 2014; Lemaitre et al. 2014).

More detailed investigation on the way breaks are processed at the mammalian nuclear lamina showed that the resection can still occur, but the later steps of HR are not completed (Lemaitre et al. 2014). It is therefore likely that NHEJ cannot take place once resection has happened, forcing the use of alt-EJ (Lemaitre et al. 2014). Then again, alt-EJ is considered as a highly mutagenic repair pathway; hence its regular use would lead to significant genomic instability and high variability of the sequences associated with the nuclear lamina. In line with this hypothesis, lamina-associated domains (LADs) are conserved in size and genomic position but their overall sequence conservation is low (Meuleman et al. 2013). Moreover, since the alt-EJ pathway has been involved in the formation of chromosomal translocations in mouse (Meuleman et al. 2013), it would be interesting to investigate whether DSBs positioned at the nuclear periphery are at the origin of more frequent chromosomal rearrangements.

3 Promyelocytic Leukemia Nuclear Bodies (PML NBs) and DNA Repair

3.1 PML NB Composition and Functions

The PML-NB is a nuclear subdomain comprising of a protein-rich complex with a diameter of 0.3–1 μm. PML-NBs number from 5 to 30 per cell, and are also referred to as PML oncogenic domains (PODs), PML nuclear domains (PML-NDs) or nuclear domain 10 (ND-10). PML-NBs localize to interchromosomal space where they make extensive contacts with chromatin, and may be physically associated with specific chromosomal loci (Sun et al. 2003; Eskiw et al. 2004). PML-NBs are also highly dynamic and change in number and in composition during the cell cycle or following stress, such as DNA damage.

The main structural component of PML-NBs is the PML protein. PML homologues have been identified in a number of vertebrate species; however, thus far no homologues of PML have been identified in non-vertebrate or single-cell eukaryotes. PML is a member of the RBCC (RING, B-box, coiled coil) family of proteins, many members of which are E3 ubiquitin ligases that form macromolecular complexes with other proteins. PML may act as an E3 ligase specific to the small ubiquitin-like modifier protein SUMO, as it was shown to promote protein sumoylation in vivo and in vitro (Chu and Yang 2011).

Humans express seven canonical PML isoforms (PML-I to PML-VII) (Jensen et al. 2001). The amino terminus of PML (encoded by exons 1–4) is conserved between isoforms, while alternative splicing of the 3′ exons in the PML pre-mRNA produces distinct carboxy termini that mediate isoform-specific protein interactions. PML-I to PML-VI contain a nuclear localization signal (NLS) encoded within exon 6 and are highly enriched in the nucleus, while PML-VII lacks the NLS and is localized to the cytoplasm. Over-expression of individual nuclear-localized isoforms of PML generates morphologically distinct PML-NBs (Condemine et al. 2006). Individual PML-NBs are comprised of a heterogeneous mixture of PML isoforms, in addition to being constitutively or transiently associated with over 90 other proteins (Dellaire et al. 2003).

Assembly of PML-NBs is dependent on post-translational modification of PML with the small ubiquitin-like modifier SUMO. PML contains three lysine residues that can act as SUMO acceptor sites; substitution of all three with arginine prevents normal PML-NB assembly (Sternsdorf et al. 1997; Fu et al. 2005). PML also contains a SUMO-interaction motif that mediates interaction of PML with other sumoylated proteins, contributes to hetero-oligomerization of different isoforms of PML within a body, and is required for ubiquitin-mediated degradation of PML after treatment with arsenic trioxide (Shen et al. 2006; Maroui et al. 2012).

Given the dynamic association of PML-NBs with other proteins and with chromatin, it is not surprising that PML-NBs regulate a variety of cellular process, acting as scaffolds that mediate protein-protein interactions and/or the post-translational modification of a variety of proteins (Dellaire and Bazett-Jones 2004; Bernardi and Pandolfi 2007). PML-NBs are implicated in telomere maintenance, defense against viruses, gene expression and protein turnover. The discovery that mice lacking PML form tumors readily in response to carcinogens implicated PML in the DNA damage response (Wang et al. 1998). Indeed, as we will discuss below, PML-NBs have since been shown to promote cell cycle arrest following stress such as DNA damage, ultimately leading to apoptosis or cellular senescence. Additionally, PML-NBs have recently been demonstrated to be involved in DNA repair by homologous recombination (HR). Taken together, these findings highlight the importance of PML-NBs at both transient and sustained stages of the DNA damage response.

3.2 Multiple Roles of PML and PML NBs in the DNA Damage Response

3.2.1 Roles for PML in Regulating DNA Damage-Induced Cell Cycle Checkpoints

Failure to repair DNA damage causes sustained activation of cell cycle checkpoints that can lead to apoptosis or senescence, with both fates being regulated by PML-NBs. One way by which PML-NBs promote apoptosis is through stabilization

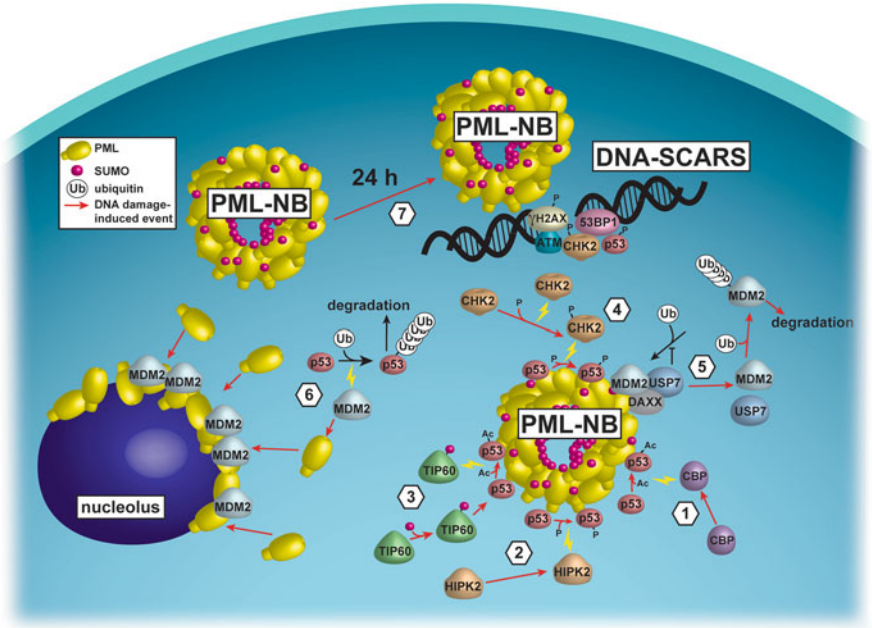

Fig. 2 PML-mediated p53 activation following DNA damage. *Red arrows* indicate events that occur following DNA damage. (1) The acetyltransferase CBP localizes to PML-NBs, where it activates p53 by acetylation of carboxy-terminal lysine residues. (2) HIPK2 kinase localizes to PML-NBs and activates p53 through phosphorylation at Ser-46. (3) The acetyltransferase TIP60 is sumoylated, activating it and promoting its localization to PML-NBs, where it acetylates p53 at Lys-120. (4) After damage PML-NBs mediate autophosphorylation and activation of CHK2, which stimulates phosphorylation of p53 at its carboxy terminus. (5) The MDM2-USP7-DAXX complex dissociates, relieving USP7-mediated inhibition of MDM2 polyubiquitylation. Proteasomal degradation of MDM2 leads to p53 stabilization. (6) PML sequesters MDM2 at the nucleolar periphery, inhibiting MDM2-mediated ubiquitylation of p53. (7) At later time points after damage, PML-NBs localize to sites of persistent DNA damage called DNA-SCARs

and activation of the tumor suppressor p53 (Ivanschitz et al. 2013) (Fig. 2). p53 is the key regulator of cell cycle checkpoints and cell fate decisions. In undamaged cells, p53 levels are kept in check by MDM2, an E3 ubiquitin ligase that targets p53 for proteasomal degradation. MDM2 forms a complex with the deubiquitinase USP7, which stabilizes MDM2, and the PML-NB protein DAXX, a modulator of apoptosis that stimulates the ubiquitin ligase activity of MDM2 toward p53 (Tang et al. 2006).

Following DNA damage, PML-NBs promote p53 stability in several ways. Co-recruitment of USP7 and p53 to PML-NBs stabilizes p53 (Everett et al. 1997; Li et al. 2002). The major DNA damage-activated kinase, ataxia telangiectasia mutated (ATM), phosphorylates DAXX, causing its dissociation from MDM2, resulting in degradation of MDM2 (Tang et al. 2006, 2013). Finally, PML

sequesters MDM2 to the nucleolus, restricting interaction with p53 (Kurki et al. 2003; Bernardi et al. 2004).

In addition to stabilizing p53, PML-NBs also contribute to p53 activation by acting as scaffolds that mediate post-translational modification of p53. Activity of p53 is regulated by a multitude of post-translational modifications, including ubiquitination, sumoylation, acetylation and phosphorylation. DNA damage promotes recruitment of p53 to PML-NBs through interaction with the carboxy-terminus of PML-IV (Guo et al. 2000; Fogal et al. 2000). Following ultraviolet (UV) radiation-induced damage, p53 is recruited to PML-NBs with the acetyltransferase CBP/p300 and the kinase HIPK2, which activate p53 through acetylation and phosphorylation, respectively (D'Orazi et al. 2002; Hofmann et al. 2002). PML-NBs also mediate the UV-induced acetylation of p53 by the acetyltransferase TIP60. UV damage upregulates sumoylation of TIP60, increasing its activity and promoting its association with PML-NBs where it acetylates p53 (Cheng et al. 2008). PML also mediates autophosphorylation and activation of the checkpoint kinase CHK2 (Yang et al. 2006) whose recruitment to PML-NBs promotes phosphorylation of p53 (Louria-Hayon et al. 2003).

In addition to promoting apoptosis through stabilization and activation of p53, PML-NBs also promote p53-independent apoptosis. When p53 is absent, PML promotes apoptosis after exposure to ionizing radiation (IR) via phosphorylation of the PML amino terminus by CHK2 (Yang et al. 2002). PML also contributes to arsenic trioxide-induced apoptosis by promoting autophosphorylation of CHK2 (Yang et al. 2006). PML is also implicated in p53-independent UV-induced apoptosis, which is mediated by the c-Jun-N-terminal kinase (JNK)/c-Jun (Bernardi et al. 2008). PML and c-Jun co-localize following UV damage (Salomoni et al. 2005), and though the exact role for PML-NBs in JNK-mediated apoptosis is not yet clear, it is proposed that PML may regulate JNK activity through DAXX (Salomoni et al. 2005; Khelifi et al. 2005).

For arrested cells bearing sustained DNA damage, an alternative to apoptosis is entry into cellular senescence, in which cells are metabolically active but no longer able to re-enter the cell cycle. Early studies demonstrated a role for PML-NBs in establishing cellular senescence induced by an oncogenic form of the Ras GTPase called Ras-V12 (Bischof et al. 2002). Disruption of PML-NBs by the adenovirus oncoprotein E1A prevents activation of p53 and attenuates Ras-V12-induced senescence. Additionally, over-expression of PML-IV, the only PML isoform that directly interacts with p53, can induce senescence even in the absence of DNA damage. Other PML isoforms are also likely required for this mechanism of establishing senescence, since PML-IV over-expression does not induce senescence in $PML^{-/-}$ cells (Bischof et al. 2002).

Another role for PML-NBs in senescence relates to establishing senescence-associated heterochromatin foci (SAHFs). SAHFs may be involved in repressing transcription of genes that promote entry into the cell cycle. Localization of histone chaperone HIRA to PML-NBs is required for the formation of SAHFs and for the establishing senescence (Zhang et al. 2005; Ye et al. 2007; Jiang et al. 2011).

PML-NBs also promote cellular senescence though the retinoblastoma protein (pRb)-E2F pathway. E2F transcription factors activate expression of multiple genes to promote progression through the cell cycle in response to growth factors. pRb is a tumor suppressor that binds to E2F to attenuate gene expression, which is essential in preventing DNA replication following oncogene expression (Chicas et al. 2010). During senescence in mouse embryonic fibroblasts (MEFs), PML is enriched at E2F promoters where it interacts with pRb and contributes to the heterochromatinization and silencing that occurs at these genes (Talluri and Dick 2014).

3.2.2 DNA Repair Factors and Persistent DNA Lesions Associate with PML NBs

In addition to regulating long-term consequences of DNA damage, PML-NBs are also implicated in early steps in the DNA damage response. PML-NBs respond to DNA damage by increasing in number through a fission mechanism, and remain at elevated numbers for a period of time that increases with dose (Dellaire et al. 2006). Besides acting as DNA damage sensors in this fashion, PML-NBs also play an active role in the sequestration and release of key proteins at early stages in the DNA damage response (Dellaire and Bazett-Jones 2004).

The MRN (MRE11-RAD50-NBS1) complex contributes to the initial processing of double-strand breaks. MRN localizes to PML-NBs in G2, possibly through interaction of NBS1 with the PML-NB protein SP100 (Naka et al. 2002). Upon DNA damage, MRN is released from PML-NBs and translocates to damage sites. MRN then re-associates with PML-NBs when repair is complete (Lombard and Guarente 2000; Mirzoeva and Petrini 2001). PML-NBs may stimulate MRN activity by promoting arginine methylation of the MRE11 endonuclease by the PML NB-localized methyltransferase PRMT1 (Boisvert et al. 2005). In cells that undergo Alternative Lengthening of Telomeres (ALT), a mechanism of maintaining telomeres that is often observed in cancer cells that do not express telomerase, the DNA repair proteins ATR (ATM and Rad3-related kinase) and RPA are sequestered at PML-NBs and released following DNA damage (Barr et al. 2003; Dellaire et al. 2006). However, this enrichment of RPA at PML-NBs is not observed in normal human fibroblasts, in which RPA is targeted to PML-NBs after damage (Dellaire et al. 2006). PML-NBs are also found at telomeres in mouse embryonic stem cells, and are required for normal telomere function (Chang et al. 2013).

PML-NBs also have a role in regulating S-phase localization of the BLM helicase, which has several roles in HR (Chu et al. 2010). Mutation of BLM causes Bloom syndrome, an autosomal recessive disease in which high levels of sister chromatid exchange lead to increased genomic instability (German et al. 1996). Sumoylation targets BLM to PML-NBs during G1/S (Zhong et al. 1999); failure to localize to PML-NBs causes the DNA damage, which localizes with DNA repair

protein to discrete nuclear sites called DNA repair foci. These foci contain BLM and other DNA repair factors such as γ-H2AX and BRCA1 (Eladad et al. 2005).

At late stages in repair, a subset of PML-NBs localize to DNA repair foci. The degree of DNA damage appears to determine whether DNA repair foci associate with PML-NBs, and affects the duration of association. For example, low-dose ionizing radiation leads to only slight, transient association of PML-NBs in DNA repair foci, while high dose radiation leads to more robust association that lasts for days (Munch et al. 2014). Even after high dose radiation, no more than about one third of DNA repair foci are observed at PML-NBs. Consistent with late association of DNA repair foci with PML-NBs, PML-NBs are not found to correlate spatially with γ-H2AX foci until later time points (18 h) following induction of DNA damage by the topoisomerase II inhibitor etoposide (Kepkay et al. 2011).

Repair foci that persist long after DNA damage may represent unrepaired DNA breaks that contribute to the amplification of the G1 cell cycle checkpoint following DNA damage (Yamauchi et al. 2008). Persistent double-strand breaks pose a threat to genomic integrity as they have the potential to recombine with other regions of the genome. Eukaryotes have evolved mechanisms to prevent these recombination events by sequestering the DNA breaks in specific nuclear domains. In yeast, which do not have a PML homologue, persistent DNA breaks are sequestered to the nuclear periphery to prevent recombination (Schober et al. 2009; Oza and Peterson 2010). In mammalian nuclei, persistent breaks are sequestered at both the nuclear periphery and PML NBs (Dellaire et al. 2009). Damaged senescent cells contain nuclear foci called DNA-SCARS (DNA Segments with Chromatin Alterations Reinforcing Senescence) that localize to the periphery of PML-NBs (Rodier et al. 2011) (Fig. 2). Association of DNA-SCARS with PML-NBs is observed 24–48 h following induction of DNA damage. These foci do not contain the repair proteins characteristic of active repair centers observed transiently after damage, but contain activated forms of proteins involved in establishing senescence, such as p53 and CHK2. PML-NBs do not appear to be required for formation of DNA-SCARS since disruption of PML-NBs does not interfere with DNA-SCARS formation (Rodier et al. 2011).

3.2.3 Role for PML in DNA Repair by Homologous Recombination

While early studies demonstrated co-localization of PML-NBs with DNA repair foci, a functional role for PML-NBs in DNA repair was not well defined until recent evidence demonstrated a role for PML-NBs in homologous recombination (HR). Clues that PML may be involved in HR came from observations that PML-NBs co-localize with repair foci that contain two HR pathway proteins, RAD51 and RPA (Dellaire et al. 2006; Boichuk et al. 2011; Munch et al. 2014). In addition, $PML^{-/-}$ cells exhibit high rates of sister chromatid exchange, which is often observed in cells defective for HR (Chu et al. 2010). To specifically examine DNA repair by HR, a direct repeat-green fluorescent protein (DR-GFP) assay was used. In this assay, a reporter cell line is used that bears two tandem chromosomally integrated inactive versions of the GFP coding sequence, with one GFP sequence containing a

recognition site for the I-SceI endonuclease. Following cleavage induced by ectopically expressed I-SceI, repair by HR leads to production of a functional GFP coding sequence, and the proportion of GFP-positive cells is determined by fluorescence automated cell sorting (FACS). Using this assay, depletion of PML led to a dramatic (~20-fold) decrease in HR (Boichuk et al. 2011; Yeung et al. 2012). Interestingly, over-expression of PML or the oncoprotein PML-RARα led to a 2- and 10-fold reduction in HR, respectively, suggesting that normal PML-NB structure and composition is important for efficient HR. Consistent with a role in DNA repair specific to HR, depletion of PML did not affect repair through non-homologous end joining (NHEJ), the major DNA repair pathway in mammals (Yeung et al. 2012). PML depletion did not affect formation of repair foci at early time points after damage but led to failure to establish MRE11 and RPA foci later, and has been suggested to be implicated in the formation of ssDNA during resection. It is not clear whether PML is required for formation of RAD51 foci, as this may be dependent on the cell line and degree of DNA damage (Boichuk et al. 2011; Yeung et al. 2012). A role for PML in HR supports the functional role for PML-NBs in maintaining telomeres in ALT cell lines, subset of PML-NBs co-localize with telomeres in ALT cell lines, which maintain their telomeres through HR (Chung et al. 2012).

4 Summary

Efficient and faithful repair of DNA damage is essential to life, ensuring that the daughter cell inherits an exact copy of the mother cell genome. Compartmentalization of chromatin in an organized nuclear architecture regulates the DNA damage response by controlling specific protein-protein and protein-DNA interactions. Nuclear structures can facilitate DNA repair by mediate critical interactions between repair pathway proteins, as has been shown for the various interactions of p53 at PML-NBs. Alternatively, DNA repair can be inhibited in other subnuclear domains, as in the case of the HR repair being inhibited at nuclear lamina due to the highly compacted state of chromatin. Studying the roles that various nuclear domains have in regulating the DNA damage response will help provide a better understanding of the pathology of diseases characterized by increased genomic instability, such as progeria and cancer.

Acknowledgments This work was supported by a Discovery Grant from the Natural Science and Engineering Research Council (NSERC) and an Operating Grant (MOP-84260) from the Canadian Institutes of Health Research (CIHR) to GD, and by grants from Agence Nationale de la Recherche (ANR), the Fondation Schlumberger pour l'Education et la Recherche (FSER), Fondation ARC (ARC fix), and La Ligue Contre le Cancer (Région) to ES. J.P. was supported by a trainee award from the Beatrice Hunter Cancer Research Institute with funds provided by the Harvey Graham Cancer Research Fund as part of The Terry Fox Foundation Strategic Health Research Training Program in Cancer Research at CIHR.

References

Barr SM, Leung CG, Chang EE, Cimprich KA (2003) ATR kinase activity regulates the intranuclear translocation of ATR and RPA following ionizing radiation. Curr Biol 13(12): 1047–1051

Bernardi R, Scaglioni PP, Bergmann S, Horn HF, Vousden KH, Pandolfi PP (2004) PML regulates p53 stability by sequestering Mdm2 to the nucleolus. Nat Cell Biol 6(7):665–672. doi:10.1038/ncb1147

Bernardi R, Papa A, Pandolfi PP (2008) Regulation of Apoptosis by PML and the PML-NBs. Oncogene 27(48):6299–6312. doi:10.1038/onc.2008.305

Bischof O, Kirsh O, Pearson M, Itahana K, Pelicci PG, Dejean A (2002) Deconstructing PML-induced premature senescence. EMBO J 21(13):3358–3369. doi:10.1093/emboj/cdf341

Boichuk S, Hu L, Makielski K, Pandolfi PP, Gjoerup OV (2011) Functional connection between Rad51 and PML in homology-directed repair. PloS One 6(10), e25814. doi:10.1371/journal.pone.0025814

Boisvert FM, Hendzel MJ, Masson JY, Richard S (2005) Methylation of MRE11 regulates its nuclear compartmentalization. Cell Cycle (Georgetown, TX) 4(7):981–989

Bridger JM, Foeger N, Kill IR, Herrmann H (2007) The nuclear lamina: both a structural framework and a platform for genome organization. FEBS J 274(6):1354–1361

Burke B, Stewart CL (2013) The nuclear lamins: flexibility in function. Nat Rev Mol Cell Biol 14(1):13–24. doi:10.1038/nrm3488

Burns LT, Wente SR (2014) From hypothesis to mechanism: uncovering nuclear pore complex links to gene expression. Mol Cell Biol 34(12):2114–2120. doi:10.1128/MCB.01730-13

Chang FT, McGhie JD, Chan FL, Tang MC, Anderson MA, Mann JR, Andy Choo KH, Wong LH (2013) PML bodies provide an important platform for the maintenance of telomeric chromatin integrity in embryonic stem cells. Nucleic Acids Res 41(8):4447–4458. doi:10.1093/nar/gkt114

Cheng Z, Ke Y, Ding X, Wang F, Wang H, Wang W, Ahmed K et al (2008) Functional characterization of TIP60 Sumoylation in UV-irradiated DNA damage response. Oncogene 27(7):931–941. doi:10.1038/sj.onc.1210710

Chicas A, Wang X, Zhang C, McCurrach M, Zhao Z, Mert O, Dickins RA, Narita M, Zhang M, Lowe SW (2010) Dissecting the unique role of the retinoblastoma tumor suppressor during cellular senescence. Cancer Cell 17(4):376–387. doi:10.1016/j.ccr.2010.01.023

Chu Y, Yang X (2011) SUMO E3 ligase activity of TRIM proteins. Oncogene 30(9):1108–1116. doi:10.1038/onc.2010.462

Chu WK, Hanada K, Kanaar R, Hickson ID (2010) BLM has early and late functions in homologous recombination repair in mouse embryonic stem cells. Oncogene 29(33): 4705–4714. doi:10.1038/onc.2010.214

Chung I, Osterwald S, Deeg KI, Rippe K (2012) PML body meets telomere: the beginning of an ALTernate ending? Nucleus (Austin, TX) 3(3):263–275. doi:10.4161/nucl.20326

Condemine W, Takahashi Y, Zhu J, Puvion-Dutilleul F, Guegan S, Janin A, de The H (2006) Characterization of endogenous human promyelocytic leukemia isoforms. Cancer Res 66(12): 6192–6198. doi:10.1158/0008-5472.CAN-05-3792

Das A, Grotsky DA, Neumann MA, Kreienkamp R, Gonzalez-Suarez I, Redwood AB, Kennedy BK, Stewart CL, Gonzalo S (2013) Lamin AΔexon9 mutation leads to telomere and chromatin defects but not genomic instability. Nucleus (Austin, TX) 4(5):410–419. doi:10.4161/nucl.26873

de The H, Riviere M, Bernhard W (1960) Examination by electron microscope of the VX2 tumor of the domestic rabbit derived from the Shope papilloma. Bulletin De L'Association Francaise Pour L'Etude Du Cancer 47:570–584

de The H, Lavau C, Marchio A, Chomienne C, Degos L, Dejean A (1991) The PML-RAR alpha fusion mRNA generated by the T(15;17) translocation in acute promyelocytic leukemia encodes a functionally altered RAR. Cell 66(4):675–684

Dechat T, Gesson K, Foisner R (2010) Lamina-independent lamins in the nuclear interior serve important functions. Cold Spring Harb Symp Quant Biol 75:533–543. doi:10.1101/sqb.2010.75.018

Dellaire G, Bazett-Jones DP (2004) PML nuclear bodies: dynamic sensors of DNA damage and cellular stress. BioEssays News Rev Mol Cell Dev Biol 26(9):963–977. doi:10.1002/bies.20089

Dellaire G, Bazett-Jones DP (2007) Beyond repair foci: subnuclear domains and the cellular response to DNA damage. Cell Cycle (Georgetown, TX) 6(15):1864–1872

Dellaire G, Farrall R, Bickmore WA (2003) The Nuclear Protein Database (NPD): sub-nuclear localisation and functional annotation of the nuclear proteome. Nucleic Acids Res 31(1):328–330

Dellaire G, Ching RW, Ahmed K, Jalali F, Tse KC, Bristow RG, Bazett-Jones DP (2006) Promyelocytic leukemia nuclear bodies behave as DNA damage sensors whose response to DNA double-strand breaks is regulated by NBS1 and the kinases ATM, Chk2, and ATR. J Cell Biol 175(1):55–66. doi:10.1083/jcb.200604009

Dellaire G, Kepkay R, Bazett-Jones DP (2009) High resolution imaging of changes in the structure and spatial organization of chromatin, gamma-00482A.X and the MRN complex within etoposide-induced DNA repair foci. Cell Cycle (Georgetown, TX) 8(22):3750–3769

di Masi A, D'Apice MR, Ricordy R, Tanzarella C, Novelli G (2008) The R527H mutation in LMNA gene causes an increased sensitivity to ionizing radiation. Cell Cycle (Georgetown, TX) 7(13):2030–2037

D'Orazi G, Cecchinelli B, Bruno T, Manni I, Higashimoto Y, Saito S, Gostissa M et al (2002) Homeodomain-interacting protein kinase-2 phosphorylates p53 at Ser 46 and mediates apoptosis. Nat Cell Biol 4(1):11–19. doi:10.1038/ncb714

Dundr M, Misteli T (2001) Functional architecture in the cell nucleus. Biochem J 356(Pt 2):297–310

Eladad S, Ye TZ, Hu P, Leversha M, Beresten S, Matunis MJ, Ellis NA (2005) Intra-nuclear trafficking of the BLM helicase to DNA damage-induced foci is regulated by SUMO modification. Hum Mol Genet 14(10):1351–1365. doi:10.1093/hmg/ddi145

Eskiw CH, Dellaire G, Bazett-Jones DP (2004) Chromatin contributes to structural integrity of promyelocytic leukemia bodies through a SUMO-1-independent mechanism. J Biol Chem 279(10):9577–9585. doi:10.1074/jbc.M312580200

Everett RD, Meredith M, Orr A, Cross A, Kathoria M, Parkinson J (1997) A novel ubiquitin-specific protease is dynamically associated with the PML nuclear domain and binds to a herpesvirus regulatory protein. EMBO J 16(7):1519–1530. doi:10.1093/emboj/16.7.1519

Fogal V, Gostissa M, Sandy P, Zacchi P, Sternsdorf T, Jensen K, Pandolfi PP, Will H, Schneider C, Del Sal G (2000) Regulation of p53 activity in nuclear bodies by a specific PML isoform. EMBO J 19(22):6185–6195. doi:10.1093/emboj/19.22.6185

Fu C, Ahmed K, Ding H, Ding X, Lan J, Yang Z, Miao Y et al (2005) Stabilization of PML nuclear localization by conjugation and oligomerization of SUMO-3. Oncogene 24(35):5401–5413. doi:10.1038/sj.onc.1208714

German J, Ellis NA, Proytcheva M (1996) Bloom's syndrome. XIX. Cytogenetic and population evidence for genetic heterogeneity. Clin Genet 49(5):223–231

Goldman RD, Shumaker DK, Erdos MR, Eriksson M, Goldman AE, Gordon LB, Gruenbaum Y et al (2004) Accumulation of mutant lamin A causes progressive changes in nuclear architecture in Hutchinson-Gilford Progeria syndrome. Proc Natl Acad Sci U S A 101(24):8963–8968. doi:10.1073/pnas.0402943101

Gonzalez-Suarez I, Redwood AB, Gonzalo S (2009a) Loss of A-type lamins and genomic instability. Cell Cycle (Georgetown, TX) 8(23):3860–3865

Gonzalez-Suarez I, Redwood AB, Perkins SM, Vermolen B, Lichtensztejin D, Grotsky DA, Morgado-Palacin L et al (2009b) Novel roles for A-type lamins in telomere biology and the DNA damage response pathway. EMBO J 28(16):2414–2427. doi:10.1038/emboj.2009.196

Gonzalez-Suarez I, Redwood AB, Grotsky DA, Neumann MA, Cheng EH, Stewart CL, Dusso A, Gonzalo S (2011) A new pathway that regulates 53BP1 stability implicates cathepsin L and vitamin D in DNA repair. EMBO J 30(16):3383–3396. doi:10.1038/emboj.2011.225

Guo A, Salomoni P, Luo J, Shih A, Zhong S, Gu W, Pandolfi PP (2000) The function of PML in p53-dependent apoptosis. Nat Cell Biol 2(10):730–736. doi:10.1038/35036365

Hofmann TG, Moller A, Sirma H, Zentgraf H, Taya Y, Droge W, Will H, Schmitz ML (2002) Regulation of p53 activity by its interaction with homeodomain-interacting protein kinase-2. Nat Cell Biol 4(1):1–10. doi:10.1038/ncb715

Horigome C, Oma Y, Konishi T, Schmid R, Marcomini I, Hauer MH, Dion V, Harata M, Gasser SM (2014) SWR1 and INO80 chromatin remodelers contribute to DNA double-strand break perinuclear anchorage site choice. Mol Cell 55(4):626–639. doi:10.1016/j.molcel.2014.06.027

Ivanschitz L, de The H, Le Bras M (2013) PML, SUMOylation, and senescence. Front Oncol 3: 171. doi:10.3389/fonc.2013.00171

Jensen K, Shiels C, Freemont PS (2001) PML protein isoforms and the RBCC/TRIM motif. Oncogene 20(49):7223–7233. doi:10.1038/sj.onc.1204765

Jiang WQ, Nguyen A, Cao Y, Chang AC, Reddel RR (2011) HP1-mediated formation of alternative lengthening of telomeres-associated PML bodies requires HIRA but not ASF1a. PloS One 6(2), e17036. doi:10.1371/journal.pone.0017036

Kakizuka A, Miller WH Jr, Umesono K, Warrell RP Jr, Frankel SR, Murty VV, Dmitrovsky E, Evans RM (1991) Chromosomal translocation T(15;17) in human acute promyelocytic leukemia fuses RAR alpha with a novel putative transcription factor, PML. Cell 66(4):663–674

Kalocsay M, Hiller NJ, Jentsch S (2009) Chromosome-wide Rad51 spreading and SUMO-H2A.Z-dependent chromosome fixation in response to a persistent DNA double-strand break. Mol Cell 33(3):335–343. doi:10.1016/j.molcel.2009.01.016

Kepkay R, Attwood KM, Ziv Y, Shiloh Y, Dellaire G (2011) KAP1 depletion increases PML nuclear body number in concert with ultrastructural changes in chromatin. Cell Cycle (Georgetown, TX) 10(2):308–322

Khadaroo B, Teixeira MT, Luciano P, Eckert-Boulet N, Germann SM, Simon MN, Gallina I et al (2009) The DNA damage response at eroded telomeres and tethering to the nuclear pore complex. Nat Cell Biol 11(8):980–987. doi:10.1038/ncb1910

Khelifi AF, D'Alcontres MS, Salomoni P (2005) Daxx is required for stress-induced cell death and JNK activation. Cell Death Differ 12(7):724–733. doi:10.1038/sj.cdd.4401559

Kind J, van Steensel B (2010) Genome-nuclear lamina interactions and gene regulation. Curr Opin Cell Biol 22(3):320–325. doi:10.1016/j.ceb.2010.04.002

Kurki S, Latonen L, Laiho M (2003) Cellular stress and DNA damage invoke temporally distinct Mdm2, p53 and PML complexes and damage-specific nuclear relocalization. J Cell Sci 116(Pt 19):3917–3925. doi:10.1242/jcs.00714

Lei K, Zhu X, Xu R, Shao C, Xu T, Zhuang Y, Han M (2012) Inner nuclear envelope proteins SUN1 and SUN2 play a prominent role in the DNA damage response. Curr Biol 22(17): 1609–1615. doi:10.1016/j.cub.2012.06.043

Lemaitre C, Soutoglou E (2014) Double Strand Break (DSB) repair in heterochromatin and heterochromatin proteins in DSB repair. DNA Repair 19:163–168. doi:10.1016/j.dnarep.2014.03.015

Lemaitre C, Grabarz A, Tsouroula K, Andronov L, Furst A, Pankotai T, Heyer V et al (2014) Nuclear position dictates DNA repair pathway choice. Genes Dev 28(22):2450–2463. doi:10.1101/gad.248369.114

Li M, Chen D, Shiloh A, Luo J, Nikolaev AY, Qin J, Gu W (2002) Deubiquitination of p53 by HAUSP is an important pathway for p53 stabilization. Nature 416(6881):648–653. doi:10.1038/nature737

Liu B, Wang J, Chan KM, Tjia WM, Deng W, Guan X, Huang JD et al (2005) Genomic instability in laminopathy-based premature aging. Nat Med 11(7):780–785. doi:10.1038/nm1266

Liu Y, Wang Y, Rusinol AE, Sinensky MS, Liu J, Shell SM, Zou Y (2008) Involvement of Xeroderma Pigmentosum Group A (XPA) in progeria arising from defective maturation of

prelamin A. FASEB J Off Publ Fed Am Soc Exp Biol 22(2):603–611. doi:10.1096/fj.07-8598com

Liu GH, Barkho BZ, Ruiz S, Diep D, Qu J, Yang SL, Panopoulos AD et al (2011) Recapitulation of premature ageing with iPSCs from Hutchinson-Gilford Progeria syndrome. Nature 472(7342): 221–225. doi:10.1038/nature09879

Lombard DB, Guarente L (2000) Nijmegen breakage syndrome disease protein and MRE11 at PML nuclear bodies and meiotic telomeres. Cancer Res 60(9):2331–2334

Louria-Hayon I, Grossman T, Sionov RV, Alsheich O, Pandolfi PP, Haupt Y (2003) The promyelocytic leukemia protein protects p53 from Mdm2-mediated inhibition and degradation. J Biol Chem 278(35):33134–33141. doi:10.1074/jbc.M301264200

Manju K, Muralikrishna B, Parnaik VK (2006) Expression of disease-causing lamin A mutants impairs the formation of DNA repair foci. J Cell Sci 119(Pt 13):2704–2714. doi:10.1242/jcs.03009

Maroui MA, Kheddache-Atmane S, El Asmi F, Dianoux L, Aubry M, Chelbi-Alix MK (2012) Requirement of PML SUMO interacting motif for RNF4- or arsenic trioxide-induced degradation of nuclear PML isoforms. PloS One 7(9), e44949. doi:10.1371/journal.pone.0044949

Meuleman W, Peric-Hupkes D, Kind J, Beaudry JB, Pagie L, Kellis M, Reinders M, Wessels L, van Steensel B (2013) Constitutive nuclear lamina-genome interactions are highly conserved and associated with A/T-rich sequence. Genome Res 23(2):270–280. doi:10.1101/gr.141028.112

Mirzoeva OK, Petrini JH (2001) DNA damage-dependent nuclear dynamics of the Mre11 complex. Mol Cell Biol 21(1):281–288. doi:10.1128/MCB.21.1.281-288.2001

Mounkes LC, Stewart CL (2004) Aging and nuclear organization: lamins and progeria. Curr Opin Cell Biol 16(3):322–327. doi:10.1016/j.ceb.2004.03.009

Munch S, Weidtkamp-Peters S, Klement K, Grigaravicius P, Monajembashi S, Salomoni P, Pandolfi PP, Weisshart K, Hemmerich P (2014) The tumor suppressor PML specifically accumulates at RPA/Rad51-containing DNA damage repair foci but is nonessential for DNA damage-induced fibroblast senescence. Mol Cell Biol 34(10):1733–1746. doi:10.1128/MCB.01345-13

Nagai S, Dubrana K, Tsai-Pflugfelder M, Davidson MB, Roberts TM, Brown GW, Varela E, Hediger F, Gasser SM, Krogan NJ (2008) Functional targeting of DNA damage to a nuclear pore-associated SUMO-dependent ubiquitin ligase. Science (New York, NY) 322(5901): 597–602. doi:10.1126/science.1162790

Naka K, Ikeda K, Motoyama N (2002) Recruitment of NBS1 into PML oncogenic domains via interaction with SP100 protein. Biochem Biophys Res Commun 299(5):863–871

Oza P, Peterson CL (2010) Opening the DNA repair toolbox: localization of DNA double strand breaks to the nuclear periphery. Cell Cycle (Georgetown, TX) 9(1):43–49

Oza P, Jaspersen SL, Miele A, Dekker J, Peterson CL (2009) Mechanisms that regulate localization of a DNA double-strand break to the nuclear periphery. Genes Dev 23(8):912–927. doi:10.1101/gad.1782209

Padeken J, Heun P (2014) Nucleolus and nuclear periphery: velcro for heterochromatin. Curr Opin Cell Biol 28:54–60. doi:10.1016/j.ceb.2014.03.001

Richards SA, Muter J, Ritchie P, Lattanzi G, Hutchison CJ (2011) The accumulation of un-repairable DNA damage in laminopathy progeria fibroblasts is caused by ROS generation and is prevented by treatment with N-acetyl cysteine. Hum Mol Genet 20(20):3997–4004. doi:10.1093/hmg/ddr327

Rodier F, Munoz DP, Teachenor R, Chu V, Le O, Bhaumik D, Coppe JP et al (2011) DNA-SCARS: distinct nuclear structures that sustain damage-induced senescence growth arrest and inflammatory cytokine secretion. J Cell Sci 124(Pt 1):68–81. doi:10.1242/jcs.071340

Salomoni P, Bernardi R, Bergmann S, Changou A, Tuttle S, Pandolfi PP (2005) The promyelocytic leukemia protein PML regulates C-Jun function in response to DNA damage. Blood 105(9): 3686–3690. doi:10.1182/blood-2004-09-3782

Scaffidi P, Misteli T (2006) Lamin A-dependent nuclear defects in human aging. Science (New York, NY) 312(5776):1059–1063. doi:10.1126/science.1127168

Schober H, Ferreira H, Kalck V, Gehlen LR, Gasser SM (2009) Yeast telomerase and the SUN domain protein Mps3 anchor telomeres and repress subtelomeric recombination. Genes Dev 23(8):928–938. doi:10.1101/gad.1787509

Shen TH, Lin HK, Scaglioni PP, Yung TM, Pandolfi PP (2006) The mechanisms of PML-nuclear body formation. Mol Cell 24(3):331–339. doi:10.1016/j.molcel.2006.09.013

Singh M, Hunt CR, Pandita RK, Kumar R, Yang CR, Horikoshi N, Bachoo R et al (2013) Lamin A/C depletion enhances DNA damage-induced stalled replication fork arrest. Mol Cell Biol 33(6):1210–1222. doi:10.1128/MCB.01676-12

Spector DL (2001) Nuclear domains. J Cell Sci 114(Pt 16):2891–2893

Sternsdorf T, Jensen K, Will H (1997) Evidence for covalent modification of the nuclear dot-associated proteins PML and Sp100 by PIC1/SUMO-1. J Cell Biol 139(7):1621–1634

Sun Y, Durrin LK, Krontiris TG (2003) Specific interaction of PML bodies with the TP53 locus in Jurkat interphase nuclei. Genomics 82(2):250–252

Sur I, Neumann S, Noegel AA (2014) Nesprin-1 role in DNA damage response. Nucleus (Austin, TX) 5(2):173–191. doi:10.4161/nucl.29023

Swartz RK, Rodriguez EC, King MC (2014) A role for nuclear envelope-bridging complexes in Homology-directed repair. Mol Biol Cell 25(16):2461–2471. doi:10.1091/mbc.E13-10-0569

Talluri S, Dick FA (2014) The retinoblastoma protein and PML collaborate to organize heterochromatin and silence E2F-responsive genes during senescence. Cell Cycle (Georgetown, TX) 13(4):641–651. doi:10.4161/cc.27527

Tang J, Qu LK, Zhang J, Wang W, Michaelson JS, Degenhardt YY, El-Deiry WS, Yang X (2006) Critical role for Daxx in regulating Mdm2. Nat Cell Biol 8(8):855–862. doi:10.1038/ncb1442

Tang J, Agrawal T, Cheng Q, Qu L, Brewer MD, Chen J, Yang X (2013) Phosphorylation of Daxx by ATM contributes to DNA damage-induced p53 activation. PloS One 8(2), e55813. doi:10.1371/journal.pone.0055813

Therizols P, Fairhead C, Cabal GG, Genovesio A, Olivo-Marin JC, Dujon B, Fabre E (2006) Telomere tethering at the nuclear periphery is essential for efficient DNA double strand break repair in subtelomeric region. J Cell Biol 172(2):189–199. doi:10.1083/jcb.200505159

Varela I, Cadinanos J, Pendas AM, Gutierrez-Fernandez A, Folgueras AR, Sanchez LM, Zhou Z et al (2005) Accelerated ageing in mice deficient in Zmpste24 protease is linked to p53 signalling activation. Nature 437(7058):564–568. doi:10.1038/nature04019

Wang ZG, Delva L, Gaboli M, Rivi R, Giorgio M, Cordon-Cardo C, Grosveld F, Pandolfi PP (1998) Role of PML in cell growth and the retinoic acid pathway. Science (New York, NY) 279(5356):1547–1551

Yamauchi M, Oka Y, Yamamoto M, Niimura K, Uchida M, Kodama S, Watanabe M, Sekine I, Yamashita S, Suzuki K (2008) Growth of persistent foci of DNA damage checkpoint factors is essential for amplification of G1 checkpoint signaling. DNA Repair 7(3):405–417. doi:10.1016/j.dnarep.2007.11.011

Yang S, Kuo C, Bisi JE, Kim MK (2002) PML-dependent apoptosis after DNA damage is regulated by the checkpoint kinase hCds1/Chk2. Nat Cell Biol 4(11):865–870. doi:10.1038/ncb869

Yang S, Jeong JH, Brown AL, Lee CH, Pandolfi PP, Chung JH, Kim MK (2006) Promyelocytic leukemia activates Chk2 by mediating Chk2 autophosphorylation. J Biol Chem 281(36):26645–26654. doi:10.1074/jbc.M604391200

Ye X, Zerlanko B, Zhang R, Somaiah N, Lipinski M, Salomoni P, Adams PD (2007) Definition of pRB- and p53-dependent and -independent steps in HIRA/ASF1a-mediated formation of senescence-associated heterochromatin foci. Mol Cell Biol 27(7):2452–2465. doi:10.1128/MCB.01592-06

Yeung PL, Denissova NG, Nasello C, Hakhverdyan Z, Chen JD, Brenneman MA (2012) Promyelocytic leukemia nuclear bodies support a late step in DNA double-strand break repair by homologous recombination. J Cell Biochem 113(5):1787–1799. doi:10.1002/jcb.24050

Zhang R, Poustovoitov MV, Ye X, Santos HA, Chen W, Daganzo SM, Erzberger JP et al (2005) Formation of MacroH2A-containing senescence-associated heterochromatin foci and senescence driven by ASF1a and HIRA. Dev Cell 8(1):19–30. doi:10.1016/j.devcel.2004.10.019

Zhong S, Hu P, Ye TZ, Stan R, Ellis NA, Pandolfi PP (1999) A role for PML and the nuclear body in genomic stability. Oncogene 18(56):7941–7947. doi:10.1038/sj.onc.1203367

Zhu J, Lallemand-Breitenbach V, de The H (2001) Pathways of retinoic acid- or arsenic trioxide-induced PML/RARalpha catabolism, role of oncogene degradation in disease remission. Oncogene 20(49):7257–7265. doi:10.1038/sj.onc.1204852

The Interplay Between Inflammatory Signaling and Nuclear Structure and Function

Sona Hubackova, Simona Moravcova, and Zdenek Hodny

Abstract Inflammation is an adaptive response that is triggered by noxious stimuli such as infection and tissue injury. Beside immune cells representing the main source of pro-inflammatory cytokines other, mainly damaged cells, contribute to production of these cytokines. For many of the pro-inflammatory cytokines were described that their prolonged exposure activates free radicals in normal and tumor cells which results in DNA damage, changes in nuclear structure and gene expression, activation of cell cycle checkpoints and cell cycle arrest. Failure of cell proliferation may help to stop the progression of premalignant and tumor cells, but changes in gene expression in these arrested cells (tumor and normal) are often accompanied with increased production of pro-inflammatory cytokines. If these arrested cells are not eliminated from organism they pose a threat to propagation of inflammation and development of degenerative diseases, cancer and aging.

In this chapter we review the mutual interplay between DNA damage of non-immune cells and the inflammatory processes. The pro-inflammatory cytokines that are essential for the inflammatory signaling represent not only a potential consequence of the DNA damage response but these processes themselves can also induce DNA damage and trigger genome instability. The individual players involved in this interplay between cellular stress and inflammation are discussed in a detail with the intention to clearly characterize the cellular and nuclear processes and pathways that lead to release of the key pro-inflammatory cytokines. A cycle of events that subsequently affects the structural and functional features of the nucleus, including changes in the nuclear lamina and other subnuclear domains such as the promyelocytic leukemia nuclear body (PML NB) is discussed as well.

Keywords Genome organization • Chromosome territories • Nuclear architecture • Spatial positioning • Gene positioning • Disease • Cancer • Translocations • Chromatin structure • Inflammation • Cytokines • Senescence • ROS • DNA damage response • NFκB • Nuclear lamina • PML nuclear bodies • Cancer

S. Hubackova • S. Moravcova • Z. Hodny (✉)
Department of Genome Integrity, Institute of Molecular Genetics, v.v.i., Academy of Sciences of the Czech Republic, Videnska 1083, 142 20 Prague 4, Czech Republic
e-mail: hodny@img.cas.cz

D.P. Bazett-Jones, G. Dellaire (eds.), *The Functional Nucleus*,
DOI 10.1007/978-3-319-38882-3_12

1 Introduction

1.1 Inflammation

Inflammation represents a complex response to infection and to damaged cells, tissue injury, allergens and other harmful stimuli. The acute inflammatory response involves the recruitment of plasma proteins and leukocytes, including neutrophils, macrophages and/or mast cells, from the blood vessels to the site of the infection or injury. A successful outcome is considered to be the recovery from infection, repair of tissue damage and restoration of homeostasis, which is accompanied by the resolution phase when leukocytes undergo programmed cell death and apoptotic cells are cleared by macrophage-mediated phagocytosis (Savill et al. 1989).

The inflammatory response is orchestrated by various plasma- and cell-derived mediators. The plasma-mediators include among others the fragments of the complement components (Markiewski and Lambris 2007). Cell-derived mediators are secreted from the participating leukocytes—mainly the mast cells and macrophages—and they include vasoactive amines histamine and serotonin, lipid mediators such as eicosanoids, chemokines such as interleukin 8 (IL8) and pro-inflammatory cytokines (i.e., IL1, IL6, IL11, and TNFα). The transcription factor NFκB (nuclear factor of kappa light polypeptide gene enhancer in B-cells) is activated by crucial pro-inflammatory cytokines TNFα (tumor necrosis factor alpha) and IL1 while their expression, together with the expression of other pro-inflammatory genes, is in turn upregulated by this transcription factor (discussed in detail below). Activated neutrophils at the afflicted site are equipped to fight the invading pathogens by a wide range of different proteases that are released from their granules and by the generation of reactive oxygen species (ROS) and reactive nitrogen species (RNS).

In the case of a persistent infection, extended exposure to a toxic agent or as a consequence of an autoimmune disease the inflammation persists and is termed as the chronic inflammation. Major cells involved in the chronic inflammation response are macrophages and, in case of an infection, also lymphocytes. While some cytokines are involved in both acute and chronic inflammation signaling (such as IL1, IL6 and TNFα), some other cytokines come into play when the inflammatory response persists and these include, for example, TGFβ (transforming growth factor beta) or interferons (IFNs). In contrast to the acute response, the continuous low-grade chronic inflammation has a role in tissue destruction and is linked to a whole range of diseases including aging, Alzheimer's disease, atherosclerosis, cancer, cardiovascular diseases, dementia, neurological diseases, rheumatoid arthritis or type I diabetes mellitus. How acute and chronic inflammation contributes to structural changes in the nucleus, that in turn alter cellular homeostasis and contribute to disease processes, is only beginning to be explored and is the focus of this chapter.

1.2 Cancer-Related Inflammation

Although inflammation may contribute to the anti-tumor response by enhancing the adaptive anti-tumor immunity (Zhang et al. 2007), for example in some specific cases such as following therapy-induced inflammation and stimulation of antigen cross-presentation by dendritic cells, it is now widely accepted that inflammation underlies and accompanies tumor development and may even drive tumor progression [for a review, see (Mantovani et al. 2008)]. Subsequently, the inflammatory microenvironment including inflammatory cells and mediators such as cytokines and chemokines are an integral part of all tumors and represent therapeutic targets for cancer [reviewed in (Coussens and Werb 2002)].

Epidemiological data revealed that long-term users of nonsteroidal anti-inflammatory drugs have reduced risk of sporadic colorectal cancer and possibly of stomach and esophagus cancer as well (Garcia-Rodriguez and Huerta-Alvarez 2001). But the exact mechanism of this chemopreventive effect is still unclear. There is a complex interplay between inflammation and cancer which can be characterized by two pathways—termed the extrinsic and the intrinsic (Mantovani et al. 2008). In the extrinsic pathway, the chronic inflammation not only precedes cancer development but prominently contributes to it. Examples could be colon cancer in patients with inflammatory bowel disease, liver carcinoma in individuals with hepatitis C infection or stomach cancer in carriers of *Helicobacter pylori* (Ernst and Gold 2000).

The intrinsic pathway represents a process when oncogenic transformation and oncogenes activation promotes the generation of subsequent inflammatory microenvironment which in turn supports the tumor progression and neoangiogenesis. A key factor in this pathway is likely the transcription factor c-Myc, an oncogene overexpressed in many tumors, which upon activation induces expression and release of IL1β (Shchors et al. 2006). Another event that may trigger the intrinsic pathway is the activation of small GTPase Ras. Signal from Ras is mediated through cascade including Raf, MEK, ERK and MAP kinases and results in the induction of IL8 (Sparmann and Bar-Sagi 2004). Both IL1β and IL8 are prototypical pro-inflammatory cytokines that contribute to recruitment of leukocytes, vascularization and tumor growth progression. While the tumor microenvironment contains a variety of leukocyte populations including neutrophils or mast cells, tumor-associated macrophages are the most abundant immune cells infiltrating the neoplastic tissue. They support the tumor progression by releasing growth factors, cytokines or angiogenic factors, and their accumulation within the tumor microenvironment is linked to poor prognosis (Robinson-Smith et al. 2007).

The main players regulating cancer-related inflammation on the molecular level are transcription factors NFκB and STAT3 (signal transducer and activator of transcription 3). In addition to regulating the inflammatory cytokines, NFκB also plays a role in inducing the expression of anti-apoptotic genes of the Bcl2 (B cell lymphoma) family, iNOS (inducible nitric oxide synthase), COX2 (cyclooxygenase) or various growth factors (Surh et al. 2001). STAT3 regulates a wide range of

genes that contribute to the tumor growth and survival, including transcription factors c-Myc and c-Fos, antiapoptotic genes Bcl2 and Bcl-xL, regulators of the epithelial-mesenchymal transition Twist and Zeb1, Sox2 and NANOG, that are also linked to cell stemness, and many others [for a review, see (Carpenter and Lo 2014)].

The inflammatory response must be taken into account also as a consequence of cancer treatment using radiation therapy or DNA damage causing drugs as it will be elaborated in a detail below. Finally, as the cancer-related inflammation of either the intrinsic or extrinsic origin represents a crucial part of the tumor development, it opens a new perspective for new and effective therapeutic strategies.

2 DNA Damage Response and Pro-Inflammatory Signaling

DNA damage encompasses a number of lesions in DNA, including breaks and covalent changes in the structure of DNA, where the DNA backbone is broken on one or both strands and/or chemically modified, as well as non-covalent anomalous structures, including base-pair mismatches, loops, and "bubbles" arising from a string of mismatches. Independently of type of lesion, DNA damage activates cellular checkpoints which slow or arrest (reversibly or irreversibly) cell-cycle progression, thereby allowing time for appropriate repair mechanisms to correct genetic lesions before they are passed onto the next generation of daughter cells. These lesions to DNA can occur as a result of exogenous sources of DNA damage such as ionizing radiation or chemotherapy agents, or may arise as a result of endogenous processes, including inflammation, that generate reactive oxygen (ROS) and nitrogen species (RNS) that in turn damage the DNA backbone (Jena 2012).

DNA damage response pathway involves a cascade of sensor, transducer, mediator and effector proteins [reviewed in (Jackson and Bartek 2009)]. This signaling pathway ensures that DNA lesions are identified and cell cycle arrest is initiated in order to allow for efficient DNA repair and therefore for maintaining the genome integrity. The primary transducer proteins of the PI3K-related serine-threonine kinases family—the ATM, ATR and DNA-PK kinases are crucial for orchestrating the DNA damage response (Bartek and Lukas 2003). Another important feature of the DNA damage response pathway is a complex regulation of chromatin at the site of the DNA lesion as well as DNA repair proteins through posttranslational modifications including phosphorylation, ubiquitination, SUMOylation or poly (ADP-ribosyl)ation (PARylation) [reviewed in (Polo and Jackson 2011)]. Below we will discuss the relevant pathways and post-translational events that link inflammatory processes to DNA damage signaling and vice versa, whereby the DNA damage response can ultimately contribute to both cancer and age-related disease progression via the induction of pro-inflammatory cytokines via activation of the NFκB pathway.

2.1 Reactive Oxygen and Nitrogen Species

ROS are generated by the membrane-bound complex of NADPH (nicotinamide adenine dinucleotide phosphate) oxidase and by COX in mitochondria or exogenously, for example, during hypoxia, from tobacco smoke or by radiation. They include H_2O_2 (hydrogen peroxide), O_2^- (superoxide anion) or –OH (hydroxyl radical). RNS including NO (reactive nitric oxide) and $ONOO^-$ (peroxynitrite) are generated by iNOS and some NADPH oxidases. Pathologically increased levels of ROS were reported to cause epigenetic changes and can lead to changes in the DNA methylation status. As a direct consequence of ROS, the promoter of the tumor suppressor gene p16INK4a was shown to be methylated while hypomethylation of interspersed repeat sequences and satellites promoting instability were reported as well (Ziech et al. 2011). Finally, ROS activate a range of transcription factors such as NFκB or STAT3, key players in the subsequent inflammatory and oncogenic signaling pathways (Yoon et al. 2010; Baeuerle and Henkel 1994). Jaiswal and colleagues observed DNA damage caused in response to inflammatory cytokines that was dependent on activation of iNOS and excess production of NO (Jaiswal et al. 2000).

2.2 NFκB Pathway

Since the 1990s, there is increasing evidence suggesting that irradiated cells can affect the non-irradiated cells. This phenomenon was termed as the radiation-induced bystander effect. This effect was demonstrated in a study by Mothersill and Seymour in which they showed a decrease in cloning efficiency of non-irradiated cells exposed to medium from irradiated cells (Mothersill and Seymour 1997). In another study, the spread of damage from irradiated to non-irradiated region involved ROS and the induction of inflammatory cytokines TNFα and IL1. An anti-oxidant treatment reduced this out-of-field irradiation effects (Khan et al. 1998). Induction of inflammatory cytokines TNFα, IL1β, IL6 and IL8 was observed in vivo following irradiation (Linard et al. 2004). It was also noted that several different types of DNA damage including irradiation, UV light or camptothecin treatment were shown to be responsible for the induction of NFκB (Brach et al. 1991; Devary et al. 1993; Piret and Piette 1996).

NFκB refers to a family of transcription factors that includes NFκB1 (p50 processed from its precursor p105), NF-κB2 (p52 processed from p100), p65 (RelA), RelB and c-Rel. A variety of homo- and heterodimers of these proteins can be formed, while they activate their own specific sets of genes. For the inflammatory response, the subunit p65 is crucial. As an inactive form, NFκB resides in the cytoplasm associated with regulatory proteins called inhibitors of κB (IκB). To activate the prototypical p65/p50 dimer and to release it from its association with the IκBα inhibitor, the IKK (inhibitor of IκB kinase) signalosome

comes into play. IKK consists of two kinases IKKα and IKKβ and a regulatory subunit NEMO (NFκB essential modifier, also termed the IKKγ).

Early studies linked the NFκB activation in response to DNA damage to the ATM kinase activity although the signal transduction mechanism linking the DNA damage response pathway in the nucleus and NFκB activation in the cytoplasm was unclear and is still being elucidated in detail (Lee et al. 1998; Li et al. 2001). More insight into the NFκB activation brought studies that observed NEMO induction in response to DNA damage stimuli and the use of NEMO-deficient mouse cells and zinc finger deletion mutants of NEMO that both were defective in NFκB activation following the treatment with DNA damage agents (Huang et al. 2002, 2003). Finally, ATM and NEMO were shown to be the direct link between the nuclear response to DNA damage and the cytoplasmic events that lead to NFκB activation (Wu et al. 2006).

Upon genotoxic stress, NEMO undergoes a series of posttranslational modifications that are crucial for the subsequent NFκB activation. First, NEMO is SUMOylated at lysine residues 277 and 309 by the protein inhibitor of activated STAT Y (PIASy). While unmodified NEMO displays mainly cytoplasmic localization, its modification with SUMO-1 changes the cellular distribution of NEMO and promotes its nuclear residence and accumulation (Huang et al. 2003; Mabb et al. 2006). PARylation was recently identified to play a role upstream of this event. PAR chains of activated PARP1 were shown to serve as a platform that helps to assemble NEMO, PIASy and ATM in the nucleoplasm and that is required for the PIASy-mediated NEMO SUMOylation (Stilmann et al. 2009). NEMO is in the nucleus phosphorylated on serine 85 by ATM (Wu et al. 2006). This phosphorylation is an essential prerequisite for subsequent NEMO monoubiquitination on lysine residues 277 and 309 (Huang et al. 2003). It was shown that NEMO ubiquitination is mediated by cellular inhibitor of apoptosis-1 (cIAP1) (Jin et al. 2009). Monoubiquitination of NEMO is required for its nuclear export. ATM is sequestered to the cytoplasm with ubiquitinated NEMO and together they interact with the catalytic subunit of the IKK which allows for IκBα phosphorylation, subsequent proteasomal degradation and NFκB activation and its translocation to the nucleus (Wu et al. 2006).

The situation is even more complex, as ATM in the cytoplasm associates with ELKS (a protein rich in glutamate, leucine, lysine, and serine) that is also required for the IKK activation. NFκB activation following different stimuli including DNA damage also requires TAK1 (transforming growth factor beta-activated kinase 1) kinase and its activation was shown to be ATM dependent (Sato et al. 2005; Jin et al. 2009; Yang et al. 2011).

Different pathways leading to NFκB activation were demonstrated after cellular exposure to UV or after induction of ROS. Phosphorylation of IκBα followed by its degradation was shown to be dependent on casein kinase 2 (CK2) upon UV stimulation. CK2 activation was mediated by p38-MAPK (Kato et al. 2003). Oxidative stress associated with ROS production seems to include two different pathways of NFκB activation. The first one involves IκBα degradation after

tyrosine kinase-mediated phosphorylation. The second one depends on NFκB-inducing kinase (NIK) and its ability to phosphorylate IKKα (Nakano et al. 1998).

The main functional consequence of NFκB activation following DNA damage is the protection of the damaged cells from the apoptotic cell death. This was observed for the first time after the treatment with ionizing radiation or with the chemotherapeutics daunorubicin (Wang et al. 1996). In another study, it was shown that low levels of DNA damage caused by the etoposide treatment resulted in NFκB-dependent induction of TNFα and expression of prosurvival genes. NFκB stimulation also led to the establishment of autocrine feedforward signaling loop between TNFα and TNFR1 and the presence of extensive DNA damage correlated with secondary wave of NFκB activation, secretion of TNFα, IL8 and activation caspase 8 and apoptosis (Biton and Ashkenazi 2011). Finally, there has been some suggestion that NFκB functions to maintain cellular senescence and it was shown that signaling associated with persistent DNA damage triggers secretion of senescence-associated inflammatory cytokines (Rodier et al. 2009; Wang et al. 2009).

2.3 Role of p38-MAPK in Activation of Pro-Inflammatory Cytokines

Cytokines such as IL1 are mediators of the acute phase of inflammation by induction of local and systemic responses. One of the possible mechanisms involves the activation of stress p38-MAPK kinase in response to oxidative stress and DNA damage, but independently on p53 activation or Rb dephosphorylation. In turn, the p38-MAPK pathway induces activation of NFκB, a transcription activator of many pro-inflammatory cytokines including IL6 or IL8. Inhibition of p38-MAPK activation after DNA damage results in downregulation of pro-inflammatory cytokines expression and secretion. Activation of all of these cytokines results in expression of adhesion molecules on endothelial cells, which in turn allow infiltration of immunocompetent cells into stressed tissue and elimination of damaged or senescent cells. This p38-MAPK pathway provides insight into how damaged cells might be a source of the chronic inflammation, which might be an origin of some types of cancer or age-related diseases.

3 The Impact of Pro-Inflammatory Signaling Pathways on Nuclear Structure and Function

3.1 DNA Damage Induced by Pro-Inflammatory Cytokines

As was mentioned above, induction of DNA damage by different stimuli activates besides cell cycle checkpoints also a range of signaling pathways resulting in

Table 1 Role of cytokines in inflammation and cancer

Cytokine	Role in inflammation	Role in cancer
IL1	– Chemoattractant for neutrophils – Activation of NFκB signaling pathway – Expression of leukocyte adhesive molecules like VCAM-1 and ICAM-1 – Increased endothelial permeability – Mobilization of endothelial cells by regulation of VEGF	– Induction of DNA damage through generation of ROS – Support of angiogenesis – Induction of overexpression of metalloproteases (MMPs)
IL6	– Mediator of fever – Induction of PGE_2 synthesis – Stimulation of neutrophil production – Supports growth of B cells	– Activator of STAT3/STAT5 oncogenes – Induction of senescence
IL8	– Chemoattractant for neutrophils	– Induction of senescence
TGFβ	– Role in development of natural killer T cells (NKT), regulatory T cells (Treg) and CD8 positive T cells (CD8+)	– Induction of senescence – Induction of DNA damage through generation of ROS
TNFα	– Activator of anti-viral and anti-bacterial defense – Up-regulation of leukocyte adhesion molecules like E-selectin, ICAM-1 and VCAM-1 – Increase of endothelial permeability – Increase of vasodilatation	– Formation of 8-oxo-deoxyguanosine and induction of DNA damage – Destabilization of mitochondrial complex I and ATP synthesis – Increased expression of pro-coagulant proteins resulting in intravascular thrombosis

expression of number of pro-inflammatory cytokines. Damaged cells then may be a source of chronic inflammation, mainly by secretion of IL1, IL6, IL8, TGFβ and TNFα. As mentioned already, the inflammatory response is linked to a production of ROS and RNS, which induce DNA damage, cause mutations and therefore contribute to carcinogenesis (see Table 1). Chronic DNA damage signaling then results in activation of DNA damage response, cell cycle arrest and senescence [reviewed in (Bartek et al. 2008)]. This closes the positive-feedback loop where cytokine induced DNA damage leads to (and maintains) a range of nuclear and chromatin structural and functional changes.

3.1.1 IL1

Family of IL1 consists of IL1α, IL1β and IL1 receptor antagonist (IL1-Ra), the latter binds to IL1 receptor and inhibits IL1 signaling. The key signaling pathway activated by IL1β is the NFκB pathway (discussed above). Previous studies showed the role of NFκB in elevated expression of NOX4, which is a member of NADPH oxidase family known to regulate production of ROS. Elevated ROS production due to deregulated NADPH oxidases then induce DNA damage, genomic instability and premature cellular senescence (Lu et al. 2010).

3.1.2 IL6

IL6 was determined to have pleiotropic functions in a wide range of cells expressing receptor gp130 subunit and membrane bound IL6 receptor. This cytokine was found to have an important role in human malignancies as an activator of STAT3 oncogene (Heinrich et al. 2003). STAT3 is a member of STAT family that was found to control the regulation of apoptosis, cell differentiation, proliferation, metastasis and immune responses that candidate this protein as a target for cancer therapy. Most of the major human malignancies like leukemia, melanoma or breast and prostate carcinomas show elevated levels of constitutively activated STAT3 and this activation is associated with poor prognosis, thus inhibition of STAT3 signaling leads to a decrease of tumorigenicity (Frank 2007). The mechanism of constitutive activation of STAT3 depends on increased production of IL6 by autocrine/paracrine manner (Zhang et al. 2014) simultaneously with a decrease of its inhibitors like suppressor of cytokine signaling (SOCS), protein tyrosine phosphatases (PTP) or protein inhibitors of the activated STAT (PIAS) (Valentino and Pierre 2006).

Another member of STAT family activated by IL6, STAT5, was described to play a role in human malignances, especially in leukemia and breast cancer. Inhibition of JAK2/STAT5 signaling through depletion of IL6 from microenvironment increase a sensitivity of patients with acute myeloid leukemia (AML) to chemotherapy (Nishioka et al. 2009). JAK2/STAT5 was found to evoke positive feedback loop in activation of PI3K/mTOR signaling pathway of which hyperactivation is prevalent in malignances (Zhang et al. 2014). Combined inhibition of JAK2 and PI3K synergistically reduce breast tumor growth and decrease metastasis, and thus represents a potential therapy for the most resistant patients with triple-negative breast cancer (Britschgi et al. 2012).

Unexpectedly, IL6 has been found to be necessary for execution of oncogene-induced senescence (OIS) since its depletion abolished both senescence entry and maintenance. IL6 has also been found to cooperate with C/EBPβ (CCAAT-enhancer-binding protein beta) transcription factor to amplify the inflammatory network of senescent cells, including IL8. In addition to the changes of cytokine profile, downregulation of p15INK4b expression was observed after IL6 depletion in a model of OIS (Kuilman et al. 2008).

3.1.3 IL8

IL8, one of the most abundant chemokines involved in inflammation, was described as a main chemoattractant for neutrophils, as it induces their trafficking across the vascular wall and recruits them into the site of inflammation. Continued presence of IL8 and recruitment of neutrophils in response to inflammatory conditions may lead to tissue damage, and its role in rheumatoid arthritis, psoriasis or palmoplantar pustulosis has been demonstrated (Skov et al. 2008; Harada et al. 1994).

Together with IL6, IL8 is also one of the most secreted cytokines in senescent cells. Overexpression of the chemokine binding receptor CXCR2 (also known as IL8RB), which is regulated by the NFκB and C/EBPβ transcription factors, was found in oncogenic Ras-induced senescence. Knockdown of this receptor results in resistance to IL8 treatment, abolishes the cell cycle block and development of senescence indicating that senescent cells activate a self-amplifying secretory program, in which cytokines reinforce growth arrest (Acosta et al. 2008).

3.1.4 TGFβ

TGFβ is also the pleiotropic cytokine with potent regulatory and inflammatory activity. The diverse effects of TGFβ on numerous cells and immune functions depend on environmental context. Activation of TGFβ signaling results in SMAD2 and SMAD3 phosphorylation and their heterotrimerization with the SMAD4 coactivator. Relocalization of the SMAD2/3/4 complex from cytoplasm into nucleus triggers the expression of many genes.

Genes regulated by TGFβ include those linked to cell cycle arrest. It was found that TGFβ1-dependent growth arrest in G1 phase is accompanied by increased levels of cyclin dependent kinase inhibitors p15INK4B, p16INK4A and activation of p53/p21waf1 pathway. Depletion of TGFβ results in constitutive induction of CDK2 and CDK4 kinase activity and Rb phosphorylation (Vijayachandra et al. 2003). Importantly, ectopic expression or administration of TGFβ is capable of inducing premature senescence (Cipriano et al. 2011). Yoon et al. reported decreased activity of complex IV of mitochondrial respiratory chain and generation of ROS in dependence on TGFβ. TGFβ was found to elevate expression of NOX4 gene in consequence of NFκB activation (Yoon et al. 2005). TGFβ also plays a central role in fibroblast activation and fibroblast-to-myofibroblast differentiation and induces the expression of genes for extracellular matrix components including collagen 1 (Col1) (Leask 2010).

3.1.5 TNFα

Tumor necrosis factor alpha (TNFα) is a pro-inflammatory cytokine produced in response to tissue injury and viral or bacterial infection. ROS induced formation of 8-oxo-deoxyguanosine (8-oxodG) and induction of DNA damage were shown to be stimulated by the TNFα produced by tumor or inflammatory cells in the tumor microenvironment (Wheelhouse et al. 2003). Shoji et al. described TNFα induced DNA damage due to increased mitochondrial oxygen radical formation (Shoji et al. 1995). TNFα decreases function of mitochondrial complex I and depletes ATP synthesis which results in mitochondrial damage (Mariappan et al. 2009). Increased level of hydroxyl radicals generated from O_2^- and H_2O_2 produced by the mitochondria in a process catalysed by iron then contributes to DNA damage (Shoji et al. 1995).

3.2 *Structural Changes in Nucleus Induced by Pro-Inflammatory Cytokines*

Cellular senescence is a permanent state of cell cycle arrest and features one of the intrinsic barriers against tumorigenesis. More than half a century ago, Hayflick described a process that is limiting the proliferation of normal human cells in culture and termed it replicative senescence (Hayflick 1965). Later, cellular senescence was shown to be not only the consequence of "aging in vitro" but also a component of tumor suppressive mechanisms and can be induced in response to treatment by genotoxic compounds, physical stress (ionizing and UV irradiation) or oncogene activation [reviewed in (Campisi and d'Adda di Fagagna 2007)].

As mentioned above, cytokines produced during inflammation are involved in the activation of distinct pathways and regulate many cellular processes including cellular senescence, where several pro-inflammatory cytokines functioning in paracrine/autocrine manner have also been implicated in the decision mechanisms controlling the tumor suppressor pathways responsible for establishment of senescence-related growth arrest both in normal and tumor cells. Previously, Kuilman et al. described the role of IL6 in promotion and maintenance of oncogene-induced cellular senescence by paracrine manner (Kuilman et al. 2008). Moiseeva et al. reported that prolonged exposure of cells to interferon beta induces cell cycle arrest due to activation of p53 (Moiseeva et al. 2006). Furthermore, IL8 was found to play a role in regulation of p53 driven senescence (Acosta et al. 2008). Finally, it seems that TGFβ is one of the most potent cytokines capable to induce cell cycle arrest via activation of CDK inhibitors (Vijayachandra et al. 2003).

3.2.1 Chromatin Structure

The nucleus by virtue of its specific and dynamic structure houses, protects and controls the function of the genome and regulates many of cellular processes. A number of recent studies have demonstrated the importance of nuclear and genome organization for proper nuclear function [for a recent review, see e.g., (Ghosh and Zhou 2014)]. For example, misorganization and/or defects in nuclear structure, particularly alteration in telomeres and heterochromatin regions of the genome, which in turn coincide with gene expression changes, aberrant DNA damage signaling, and cell cycle arrest, can lead to cellular senescence during both normal aging and in progeroid syndromes and laminopathies (discussed in detail later in the chapter).

Telomeres consist of repetitive DNA elements at the end of linear chromosomes that protect the DNA ends from degradation and recombination. Reduction in telomere length, or damage of telomeres, activates the p53 tumor suppressor resulting in cell cycle arrest and telomere-initiated senescence. Disturbance of telomeres also propagates via losing an ability to protect the ends of chromosome

resulting in end-to-end fusions, which generate unstable dicentric chromosomes (d'Adda di Fagagna et al. 2003). Zimmermann et al. discovered the function of telomeric protein Rif1 in control of double-strand break repair and regulation of DNA damage response (Zimmermann et al. 2013). They identified Rif1 as the main factor used by 53BP1 to impair 5′end resection by limiting an accumulation of DNA repair proteins like CtIP, BLM, Exo1 and BRCA1/BARD1 complex at site of telomeric DNA damage which results in constitutive DNA damage signaling. Lanna et al. described that IFNα, secreted during viral infection, is able to inhibit telomerase activity and accelerate T cell differentiation in vivo (Lanna et al. 2013). Another example of cytokine regulation of telomerase activity is inhibition of c-Myc by TGFβ. This inhibition results in loss of the full length hTERT (human telomerase reverse transcriptase) transcript due to alternative splicing, decrease of telomerase level and induction of cell cycle arrest (Li et al. 2006).

Another structural change of chromatin detected in senescent cells is a formation of chromatin structures called senescence-associated heterochromatin foci (SAHF). SAHF can be detected cytologically as compacted foci of DNA stained by DAPI and each focus represents condensed, nuclease resistant compaction of chromatin from one chromosome. Recent evidence shows association between aging and heterochromatin formation in cells. Lezhava demonstrates higher percentage of condensed chromatin in interphase nuclei of cells derived from old people (80–93 years) in comparison with younger people (25–52 years) (Lezhava 2001). These structures may play a role in transcriptional repression of proliferation-associated genes. Presence of heterochromatin protein 1 (HP1), high mobility group A protein (HMGA) and transcription repressive variant of histone H2A (macro H2A) was observed in SAHF. Increased DNA methylation involved in gene silencing was detected at sites of SAHF, especially methylated histone H3 on Lys 9 (H3K9me3) (Narita et al. 2003). Formation of SAHF is also linked to relocalization of HIRA to PML nuclear bodies which initiates chromosome condensation. Ye et al. observed that this relocalization is dependent on downregulation of WNT signaling (Ye et al. 2007). They demonstrated that the loss of WNT signal activates GSK3β kinase, which in turn interacts with HIRA and phosphorylates it on serine 697 to allow its localization to PML nuclear bodies. Kuilman et al. showed that the C/EBPβ–IL6 axis is required for the formation of SAHF in p16INK4a positive cells (Kuilman et al. 2008).

Taken together, the above findings clearly identify role of inflammatory cytokines in alterations of cellular and molecular pathways involved in the maintenance of genomic integrity, chromatin dynamics and regulation of gene expression (Fig. 1). Several of these processes are accompanied by changes in nuclear structure including nuclear subcompartments (see below).

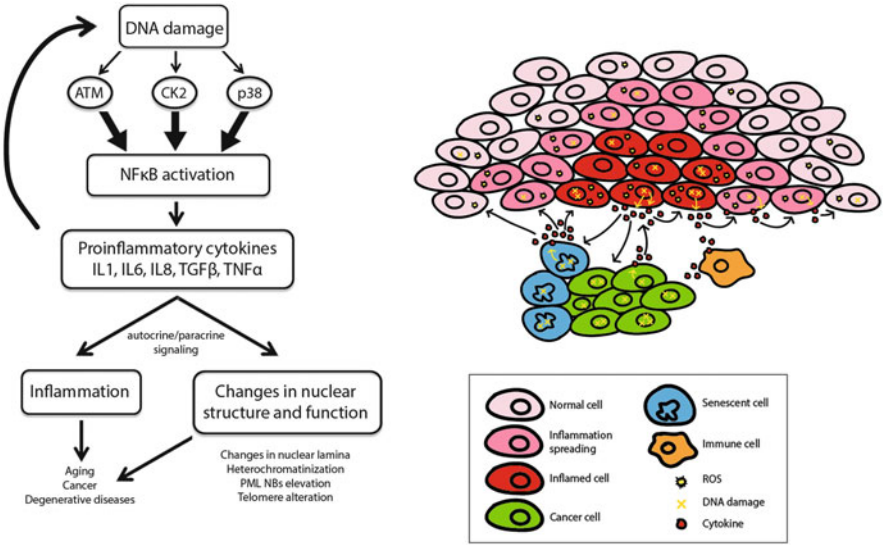

Fig. 1 Relationship of DNA damage and proinflammatory cytokines and their impact to nuclear structure and function, inflammation, cancer and aging. (*Left*) DNA damage leads to activation of NFκB via ATM, CK2 and p38-MAPK signaling pathways. NFκB triggers the expression of several proinflammatory cytokines. Secretion of these cytokines causes inflammatory response by immune system activation and changes in nuclear structure and function via autocrine/paracrine signaling. Protracted cytokine production contributes to development of cancer, degenerative diseases and aging. (*Right*) Immune, DNA damaged and senescent cells express and secrete inflammatory cytokines, which induce ROS production causing DNA damage in neighboring normal or tumor cells. Activated DNA damage response results in activation of inflammatory cytokine network in bystander cells thus closing the vicious circle. Provided the inflammatory cytokine producing cells are not eliminated, the focus of inflammation in tissue persists and further spreads

3.2.2 Nuclear Compartments

Nuclear Lamina

The nuclear lamina underlies the inner nuclear membrane and consists of intermediate filaments called lamins type A/C and B (encoded by genes LMNA and LMNB, respectively). Lamins provide a scaffold for a variety of nuclear proteins and maintain the architectural integrity of interphase nuclei. It was found that mutation in LMNA is associated with several human diseases, including the most studied Hutchinson-Gilford progeria, where the cells undergo premature senescence and patients prematurely die in young age on age-related diseases (Eriksson et al. 2003). Generally, senescence is associated with decrease of lamin B. This results in reorganization in the genome. It was found that lamin B binds preferentially chromatin regions enriched for methylated histone H3 on Lys 9 H3K9me3 and methylated histone H3 on Lys 27 H3K27me3. Decrease of lamin B facilitates

the spatial relocalization of perinuclear H3K9me3 heterochromatin and thus promotes SAHF formation (Sadaie et al. 2013). Disrupted interaction between lamin B and H3K27me3 is strongly associated with up-regulation of key pro-senescent and anti-proliferative genes including genes comprising senescence-associated secretory phenotype (SASP), as many of these genes are located near these methylated areas and thus unavailable for transcription under the normal condition, when the chromatin is attached closely to the nuclear envelope (Shah et al. 2013).

Several proteins involved in detection and repair of DNA damage by the nucleotide excision repair (NER) like proliferating cell nuclear antigen (PCNA), DNA polymerase eta (POLH) or damage-specific DNA binding protein 1 (DDB1) are downregulated after loss of LMNB, which results in protracted DNA damage signaling (persistent activation of γH2AX and 53BP1 foci) and induction of the cell cycle arrest in G1 and senescence (Butin-Israeli et al. 2013). Accumulation of non-functional lamin A has also been shown to trigger ATM/NEMO-dependent activation of NFκB in absence of DNA damage, which in turn can activate expression of pro-inflammatory cytokines and maintain senescence (Osorio et al. 2012).

Besides Hutchinson-Gilford progeria, lamin B1 downregulation and loss of nuclear envelope integrity in senescent cells result in generation of γH2AX and H3K27me3 strongly positive vesicles budding from nucleus. These vesicles containing histones are then transported to lysosomes and processed by proteolysis. In vivo, depletion of histones correlates with nevus maturation, which is associated with proliferation arrest and clinical benignancy (Ivanov et al. 2013). In contrast, Dvorakova et al. described the ability of these budding DNA/histones complexes to be released from the cells and to integrate into chromosomal DNA of recipient cells, in which persistent DNA damage signaling and senescence/apoptosis is induced. It is thought that in case of tumor cells, this process may be advantageous for elimination of tumor cells and thus suppression of tumor formation (Dvorakova et al. 2013).

PML Nuclear Bodies

Promyelocytic leukemia (PML) gene was originally discovered in patients with acute promyelocytic leukemia as a fusion partner of retinoic acid receptor alpha (RARα) gene. The PML protein is involved in many cellular processes like gene transcription, viral pathogenicity, DNA damage repair, cellular senescence, tumor suppression or apoptosis, although depletion of this protein is not lethal and PML knockout mice are viable but with increased sensitivity to infections and chemically induced carcinogenesis. PML is essential structural component of specific nuclear structure PML NBs (nuclear bodies) of which recognized function is to sequester (Negorev and Maul 2001), degrade (Qin et al. 2006) and posttranslationally modify proteins (Pearson et al. 2000).

PML is identified as a tumor suppressor and its levels in most types of human tumors are decreased. On the other hand, a dramatic increase in PML expression

was observed during inflammation in comparison with its low level in normal healthy tissues. Massive infiltration of immune cells is observed at sites of inflammation to target and remove infected, damaged or dead cells. Several of immune cells are strong producers of IFNγ. IFNγ activates JAK/STAT signaling pathway. Importance of activated STAT1 homodimer in regulation of PML expression and induction in infected as well as damaged and senescent cells was described as well as a protective role of STAT1/PML induction against tumorigenesis (Stadler et al. 1995).

In case of viral infection any stage in virus replication may be a target for inhibition by IFNγ induced PML, including viral entry, transcription, RNA stability, initiation of translation, maturation, assembly and release (Regad and Chelbi-Alix 2001). Since viral infection can be accompanied with production of pro-inflammatory cytokines, PML thus protect spreading of inflammation in affected tissue and organism. Overexpression of PML IV isoform in infected cells was found to enhance phosphorylation of transcription factor IRF3 resulting in IFNβ production in response to viral infection which demonstrate that, in addition to its intrinsic antiviral properties, PML is also implicated in an innate immune response (El Asmi et al. 2014).

As mentioned above, IL6 is a functionally pleiotropic cytokine produced by many cell types in response to injury, inflammation and infection. It was shown that IL6 regulates basal PML transcription via JAK/STAT3 and Akt/NFκB signaling (Hubackova et al. 2012b).

Chronic inflammation usually promotes angiogenesis. Besides VEGF, TNFα was found to be one of the important regulators of angiogenesis in local endothelium. Experiments revealed PML as a mediator for TNFα mediated inhibition of endothelial cell network formation and migration, where PML regulates expression of matrix-associated metalloproteinases (MMP) and integrin beta 1 (ITGβ1) (Cheng et al. 2012). The TNFα was also found to convert CD44 adhesion molecule present on surface of leukocytes from its inactive, nonbinding form, to its active form by inducing the sulfation of CD44. This posttranslational modification is required for CD44 mediated binding of leukocytes to vascular endothelial cells and to allow them to get into affected tissue (Maiti et al. 1998). Downregulation of PML resulted in loss of leukocytes adhesion due to inactivation of CD44. The data suggest that PML is the TNFα regulator required to finely control the TNF-α-mediated inflammatory response. Microarray analysis also revealed that PML is a putative regulator of class I HLA and that PML functions to reduce the susceptibility of damaged cells to escape the immune response (Cheng and Kao 2012).

PML also regulates inflammation through the inhibition of inflammasome. Inflammasome is a large cytoplasmic multiprotein complex that acts to regulate the activation of caspase-1, the enzyme that processes besides other things the pro-form of IL1β into its active form. During the inflammasome activation, ASC (the apoptosis-associated speck-like protein containing a caspase-activating recruitment domain) binds to this complex and recruits pro-caspase-1 leading to its autocleavage and activation. Interaction between elevated PML and ASC was found, which resulted in retention of ASC in the nucleus. These data indicates the

role of PML as a negative regulator of inflammasome-produced IL1β by regulation of ASC (Dowling et al. 2014).

PML null cells are resistant to many apoptotic stimuli including those induced by cytokines which participate in inflammation. It was described that PML null fibroblasts are resistant to TGFβ-induced apoptosis and senescence. PML has 7 isoforms, where isoform 1–6 are mainly nuclear, but isoform 7 is strictly cytoplasmic (cPML) due to the lack of nuclear localizing signal (NLS). cPML was found to be critical for SMAD2 phosphorylation by promoting its interaction with SARA (SMAD anchor for receptor activation) and accumulation of this complex in early endosome, where phosphorylation of SMAD2 takes place. In patients with acute promyelocytic leukemia (APL), PML-RARα fusion protein, apart from blocking formation of PML nuclear bodies and function of RARα, can also sequester cPML away from the TGFβ/SARA/SMAD2 complex and allow cancer cells to survive (Seo et al. 2006).

3.3 Secondary "Bystander" Senescence

Accumulating evidence shows that senescent cells can have deleterious effects on cells in their vicinity [see e.g., (Krtolica et al. 2001; Parrinello et al. 2005)]. In line with this, several reports indicate that organisms developed mechanisms to eliminate senescent cells (of some type at least) from tissues. It was shown that senescent cells are able to recruit immune cells, especially macrophages after their activation by CD4+ T-lymphocytes, by secreted pro-inflammatory cytokines to execute the clearance of senescent cells (Kang et al. 2011). The corroboration of aberrant effect of senescent cells on organism physiology emerges from findings that genetically engineered clearance of senescent cells from organism delays the onset of manifestations of tissue disturbances in progeroid mouse model (Baker et al. 2011). Nonetheless, clearance of a great number of senescent cells from tissue without compensation for cell loss, results in damage and dysfunction of affected organ (Krizhanovsky et al. 2008).

As mentioned above, cellular senescence is accompanied with changes in profile of secreted cytokines of which several have the pro-inflammatory function. It was shown that culture medium conditioned by senescent cells can induce DNA damage and senescence in surrounding normal and also tumor cells [secondary or bystander senescence, see (Nelson et al. 2012)]. The cytokines responsible for this effect, particularly IL1 and TGFβ, were identified. Their release from senescent cells activates ROS production and oxidative stress in neighboring cells. Given that oxidative stress accompanying primary replicative and oncogene-induced senescence is due to contribution of elevated expression of NADPH oxidases NOX4 and NOX1 (Ogrunc et al. 2014; Kodama et al. 2013) and that IL1 and TGFβ also induce NOX1 and NOX4, it is likely that similar effector mechanisms driven by cytokine paracrine signaling are involved both in primary and secondary senescence. Physiologically important for propagation of senescence and pro-inflammatory cytokine

activity in tissues might be the production of secondary wave of pro-inflammatory cytokines by bystander senescent cells (Fig. 1) (Hubackova et al. 2012a). In support of this notion, Acosta et al. observed that similarly to the control of production of pro-inflammatory cytokines by inflammasome-mediated IL1 signaling in primary senescent cells, the inflammasome/IL1 signaling is also activated in bystander senescent cells in vivo (Acosta et al. 2013). Similarly as senescent cells, also immune cells (especially T1-helper lymphocyte) were described to induce secondary senescence of mouse beta-cancer cells by production of TNFα and IFNγ, as demonstrated by Braumuller et al. (2013).

Besides the role of local production of pro-inflammatory cytokines in development of various pathological processes, pro-inflammatory cytokines like IL6 released into systemic circulation are also involved in pathogenesis of a variety of chronic degenerative diseases including cancer and immunosenescence [defined as functional decline of adaptive and innate immune system (Takahashi et al. 2003)]. Strikingly, Redon et al. described a role of the CCL2/MCP-1 chemokine in induction of double-strand breaks (DSB) and oxidatively induced non-DSB clustered DNA lesions through activation of ROS in tissues distant from the tumor site (Redon et al. 2010). Maturation and differentiation of megakaryocytes after receiving signal from thrombopoetin (TPO) to arrest cell cycle and enter the senescence-like state represents another example of distant secondary senescence (Besancenot et al. 2010).

Given accumulating evidence that the aging of organism has powerful systemic component mediated by chemokine and cytokine species including those of TGFβ family [reviewed in (Rando and Chang 2012)], the secretome of senescent cells can be an important contributor to mediate such a change of "young" into "aging-promoting" environment (for example, TGFβ family member MIC1/GDF15 produced also by senescent cells is supposed as a mortality predictor (Eggers et al. 2013; Wiklund et al. 2010)). Moreover, the reported capacity of senescent cells to convert surrounding cells to senescent ones by paracrine bystander effect can contribute to exponential onset of organismal aging. Thus elimination of senescent cells is emerging as a new therapeutic approach for treatment of age-associated diseases and ageing itself (Naylor et al. 2013).

4 Summary

Inflammation and inflammation-related processes represent a "double-edged" sword that is on one hand a crucial set of events for proper function of the immune system to combat infection or to promote healing following tissue injury, and on the other hand the imbalanced inflammatory response can promote and contribute to the development of a whole range of age-related diseases, including cancer. There is increasing evidence that protracted DNA damage response hosted in senescent cells and accompanied by the expression of pro-inflammatory cytokines, as a consequence of changes in nuclear and chromatin structure, can contribute to

formation of pro-inflammatory milieu in tissues. The released cytokines not only shape the physiological functions of cells, including expression of oncogenes and tumor suppressors in autocrine/paracrine manner, but can contribute to tissue damage by increasing the oxidative stress. Thus, a better understanding of both the sources and mechanisms underlying inflammation and its inter-relationship with cell senescence, that contribute to homeostasis by the removal of damaged cells from tissues, will pave a way for the design of new therapies for a range of disease associated with normal and premature aging.

Acknowledgments ZH and SM were supported by Grant Agency of the Czech Republic (Project No.: 13-17658S) and Institutional Grant RVO 68378050. SH was supported by Academy of Sciences of the Czech Republic (Project No.: L200521301).

References

Acosta JC, Banito A, Wuestefeld T, Georgilis A, Janich P, Morton JP, Athineos D, Kang TW, Lasitschka F, Andrulis M, Pascual G, Morris KJ, Khan S, Jin H, Dharmalingam G, Snijders AP, Carroll T, Capper D, Pritchard C, Inman GJ, Longerich T, Sansom OJ, Benitah SA, Zender L, Gil J (2013) A complex secretory program orchestrated by the inflammasome controls paracrine senescence. Nat Cell Biol 15(8):978–990. doi:10.1038/ncb2784

Acosta JC, O'Loghlen A, Banito A, Guijarro MV, Augert A, Raguz S, Fumagalli M, Da Costa M, Brown C, Popov N, Takatsu Y, Melamed J, d'Adda di Fagagna F, Bernard D, Hernando E, Gil J (2008) Chemokine signaling via the CXCR2 receptor reinforces senescence. Cell 133 (6):1006–1018. doi:10.1016/j.cell.2008.03.038, S0092-8674(08)00619-3 [pii]

Baeuerle PA, Henkel T (1994) Function and activation of NF-kappa B in the immune system. Annu Rev Immunol 12:141–179. doi:10.1146/annurev.iy.12.040194.001041

Baker DJ, Wijshake T, Tchkonia T, LeBrasseur NK, Childs BG, van de Sluis B, Kirkland JL, van Deursen JM (2011) Clearance of p16Ink4a-positive senescent cells delays ageing-associated disorders. Nature 479(7372):232–236. doi:10.1038/nature10600

Bartek J, Hodny Z, Lukas J (2008) Cytokine loops driving senescence. Nat Cell Biol 10 (8):887–889

Bartek J, Lukas J (2003) Chk1 and Chk2 kinases in checkpoint control and cancer. Cancer Cell 3 (5):421–429, S1535610803001107 [pii]

Besancenot R, Chaligne R, Tonetti C, Pasquier F, Marty C, Lecluse Y, Vainchenker W, Constantinescu SN, Giraudier S (2010) A senescence-like cell-cycle arrest occurs during megakaryocytic maturation: implications for physiological and pathological megakaryocytic proliferation. PLoS Biol 8(9). doi:10.1371/journal.pbio.1000476

Biton S, Ashkenazi A (2011) NEMO and RIP1 control cell fate in response to extensive DNA damage via TNF-alpha feedforward signaling. Cell 145(1):92–103. doi:10.1016/j.cell.2011.02.023

Brach MA, Hass R, Sherman ML, Gunji H, Weichselbaum R, Kufe D (1991) Ionizing radiation induces expression and binding activity of the nuclear factor kappa B. J Clin Invest 88 (2):691–695. doi:10.1172/JCI115354

Braumuller H, Wieder T, Brenner E, Assmann S, Hahn M, Alkhaled M, Schilbach K, Essmann F, Kneilling M, Griessinger C, Ranta F, Ullrich S, Mocikat R, Braungart K, Mehra T, Fehrenbacher B, Berdel J, Niessner H, Meier F, van den Broek M, Haring HU, Handgretinger R, Quintanilla-Martinez L, Fend F, Pesic M, Bauer J, Zender L, Schaller M, Schulze-Osthoff K, Rocken M (2013) T-helper-1-cell cytokines drive cancer into senescence. Nature 494(7437):361–365. doi:10.1038/nature11824

Britschgi A, Andraos R, Brinkhaus H, Klebba I, Romanet V, Muller U, Murakami M, Radimerski T, Bentires-Alj M (2012) JAK2/STAT5 inhibition circumvents resistance to PI3K/mTOR blockade: a rationale for cotargeting these pathways in metastatic breast cancer. Cancer Cell 22(6):796–811. doi:10.1016/j.ccr.2012.10.023

Butin-Israeli V, Adam SA, Goldman RD (2013) Regulation of nucleotide excision repair by nuclear lamin b1. PLoS One 8(7):e69169. doi:10.1371/journal.pone.0069169

Campisi J, d'Adda di Fagagna F (2007) Cellular senescence: when bad things happen to good cells. Nat Rev Mol Cell Biol 8(9):729–740

Carpenter RL, Lo HW (2014) STAT3 target genes relevant to human cancers. Cancers 6 (2):897–925. doi:10.3390/cancers6020897

Cheng X, Kao HY (2012) Microarray analysis revealing common and distinct functions of promyelocytic leukemia protein (PML) and tumor necrosis factor alpha (TNFalpha) signaling in endothelial cells. BMC Genomics 13:453. doi:10.1186/1471-2164-13-453

Cheng X, Liu Y, Chu H, Kao HY (2012) Promyelocytic leukemia protein (PML) regulates endothelial cell network formation and migration in response to tumor necrosis factor alpha (TNFalpha) and interferon alpha (IFNalpha). J Biol Chem 287(28):23356–23367. doi:10.1074/jbc.M112.340505

Cipriano R, Kan CE, Graham J, Danielpour D, Stampfer M, Jackson MW (2011) TGF-beta signaling engages an ATM-CHK2-p53-independent RAS-induced senescence and prevents malignant transformation in human mammary epithelial cells. Proc Natl Acad Sci USA 108 (21):8668–8673. doi:10.1073/pnas.1015022108

Coussens LM, Werb Z (2002) Inflammation and cancer. Nature 420(6917):860–867

d'Adda di Fagagna F, Reaper PM, Clay-Farrace L, Fiegler H, Carr P, Von Zglinicki T, Saretzki G, Carter NP, Jackson SP (2003) A DNA damage checkpoint response in telomere-initiated senescence. Nature 426(6963):194–198

Devary Y, Rosette C, DiDonato JA, Karin M (1993) NF-kappa B activation by ultraviolet light not dependent on a nuclear signal. Science 261(5127):1442–1445

Dowling JK, Becker CE, Bourke NM, Corr SC, Connolly DJ, Quinn SR, Pandolfi PP, Mansell A, O'Neill LA (2014) Promyelocytic leukemia protein interacts with the apoptosis-associated speck-like protein to limit inflammasome activation. J Biol Chem 289(10):6429–6437. doi:10.1074/jbc.M113.539692

Dvorakova M, Karafiat V, Pajer P, Kluzakova E, Jarkovska K, Pekova S, Krutilkova L, Dvorak M (2013) DNA released by leukemic cells contributes to the disruption of the bone marrow microenvironment. Oncogene 32(44):5201–5209. doi:10.1038/onc.2012.553

Eggers KM, Kempf T, Wallentin L, Wollert KC, Lind L (2013) Change in growth differentiation factor 15 concentrations over time independently predicts mortality in community-dwelling elderly individuals. Clin Chem 59(7):1091–1098. doi:10.1373/clinchem.2012.201210

El Asmi F, Maroui MA, Dutrieux J, Blondel D, Nisole S, Chelbi-Alix MK (2014) Implication of PMLIV in both intrinsic and innate immunity. PLoS Pathog 10(2):e1003975. doi:10.1371/journal.ppat.1003975

Eriksson M, Brown WT, Gordon LB, Glynn MW, Singer J, Scott L, Erdos MR, Robbins CM, Moses TY, Berglund P, Dutra A, Pak E, Durkin S, Csoka AB, Boehnke M, Glover TW, Collins FS (2003) Recurrent de novo point mutations in lamin A cause Hutchinson-Gilford progeria syndrome. Nature 423(6937):293–298

Ernst PB, Gold BD (2000) The disease spectrum of *Helicobacter pylori*: the immunopathogenesis of gastroduodenal ulcer and gastric cancer. Annu Rev Microbiol 54:615–640. doi:10.1146/annurev.micro.54.1.615

Frank DA (2007) STAT3 as a central mediator of neoplastic cellular transformation. Cancer Lett 251(2):199–210. doi:10.1016/j.canlet.2006.10.017

Garcia-Rodriguez LA, Huerta-Alvarez C (2001) Reduced risk of colorectal cancer among long-term users of aspirin and nonaspirin nonsteroidal antiinflammatory drugs. Epidemiology 12 (1):88–93

Ghosh S, Zhou Z (2014) Genetics of aging, progeria and lamin disorders. Curr Opin Genet Dev 26C:41–46. doi:10.1016/j.gde.2014.05.003

Harada A, Sekido N, Akahoshi T, Wada T, Mukaida N, Matsushima K (1994) Essential involvement of interleukin-8 (IL-8) in acute inflammation. J Leukoc Biol 56(5):559–564

Hayflick L (1965) The limited in vitro lifetime of human diploid cell strains. Exp Cell Res 37:614–636

Heinrich PC, Behrmann I, Haan S, Hermanns HM, Muller-Newen G, Schaper F (2003) Principles of interleukin (IL)-6-type cytokine signalling and its regulation. Biochem J 374(Pt 1):1–20. doi:10.1042/BJ20030407, BJ20030407 [pii]

Huang TT, Feinberg SL, Suryanarayanan S, Miyamoto S (2002) The zinc finger domain of NEMO is selectively required for NF-kappa B activation by UV radiation and topoisomerase inhibitors. Mol Cell Biol 22(16):5813–5825

Huang TT, Wuerzberger-Davis SM, Wu ZH, Miyamoto S (2003) Sequential modification of NEMO/IKKgamma by SUMO-1 and ubiquitin mediates NF-kappaB activation by genotoxic stress. Cell 115(5):565–576

Hubackova S, Krejcikova K, Bartek J, Hodny Z (2012a) IL1- and TGFbeta-Nox4 signaling, oxidative stress and DNA damage response are shared features of replicative, oncogene-induced, and drug-induced paracrine 'bystander senescence'. Aging (Albany NY) 4 (12):932–951

Hubackova S, Krejcikova K, Bartek J, Hodny Z (2012b) Interleukin 6 signaling regulates promyelocytic leukemia protein gene expression in human normal and cancer cells. J Biol Chem 287(32):26702–26714. doi:10.1074/jbc.M111.316869

Ivanov A, Pawlikowski J, Manoharan I, van Tuyn J, Nelson DM, Rai TS, Shah PP, Hewitt G, Korolchuk VI, Passos JF, Wu H, Berger SL, Adams PD (2013) Lysosome-mediated processing of chromatin in senescence. J Cell Biol 202(1):129–143. doi:10.1083/jcb.201212110

Jackson SP, Bartek J (2009) The DNA-damage response in human biology and disease. Nature 461 (7267):1071–1078

Jaiswal M, LaRusso NF, Burgart LJ, Gores GJ (2000) Inflammatory cytokines induce DNA damage and inhibit DNA repair in cholangiocarcinoma cells by a nitric oxide-dependent mechanism. Cancer Res 60(1):184–190

Jena NR (2012) DNA damage by reactive species: mechanisms, mutation and repair. J Biosci 37 (3):503–517

Jin HS, Lee DH, Kim DH, Chung JH, Lee SJ, Lee TH (2009) cIAP1, cIAP2, and XIAP act cooperatively via nonredundant pathways to regulate genotoxic stress-induced nuclear factor-kappaB activation. Cancer Res 69(5):1782–1791. doi:10.1158/0008-5472.CAN-08-2256

Kang TW, Yevsa T, Woller N, Hoenicke L, Wuestefeld T, Dauch D, Hohmeyer A, Gereke M, Rudalska R, Potapova A, Iken M, Vucur M, Weiss S, Heikenwalder M, Khan S, Gil J, Bruder D, Manns M, Schirmacher P, Tacke F, Ott M, Luedde T, Longerich T, Kubicka S, Zender L (2011) Senescence surveillance of pre-malignant hepatocytes limits liver cancer development. Nature 479(7374):547–551. doi:10.1038/nature10599

Kato T Jr, Delhase M, Hoffmann A, Karin M (2003) CK2 Is a C-terminal IkappaB kinase responsible for NF-kappaB activation during the UV response. Mol Cell 12(4):829–839

Khan MA, Hill RP, Van Dyk J (1998) Partial volume rat lung irradiation: an evaluation of early DNA damage. Int J Radiat Oncol Biol Phys 40(2):467–476

Kodama R, Kato M, Furuta S, Ueno S, Zhang Y, Matsuno K, Yabe-Nishimura C, Tanaka E, Kamata T (2013) ROS-generating oxidases Nox1 and Nox4 contribute to oncogenic Ras-induced premature senescence. Genes Cells 18(1):32–41. doi:10.1111/gtc.12015

Krizhanovsky V, Yon M, Dickins RA, Hearn S, Simon J, Miething C, Yee H, Zender L, Lowe SW (2008) Senescence of activated stellate cells limits liver fibrosis. Cell 134(4):657–667. doi:10.1016/j.cell.2008.06.049, S0092-8674(08)00836-2 [pii]

Krtolica A, Parrinello S, Lockett S, Desprez PY, Campisi J (2001) Senescent fibroblasts promote epithelial cell growth and tumorigenesis: a link between cancer and aging. Proc Natl Acad Sci USA 98(21):12072–12077

Kuilman T, Michaloglou C, Vredeveld LC, Douma S, van Doorn R, Desmet CJ, Aarden LA, Mooi WJ, Peeper DS (2008) Oncogene-induced senescence relayed by an interleukin-dependent inflammatory network. Cell 133(6):1019–1031. doi:10.1016/j.cell.2008.03.039, S0092-8674(08)00620-X [pii]

Lanna A, Coutavas E, Levati L, Seidel J, Rustin MH, Henson SM, Akbar AN, Franzese O (2013) IFN-alpha inhibits telomerase in human CD8(+) T cells by both hTERT downregulation and induction of p38 MAPK signaling. J Immunol 191(7):3744–3752. doi:10.4049/jimmunol.1301409

Leask A (2010) Potential therapeutic targets for cardiac fibrosis: TGFbeta, angiotensin, endothelin, CCN2, and PDGF, partners in fibroblast activation. Circ Res 106(11):1675–1680. doi:10.1161/CIRCRESAHA.110.217737

Lee SJ, Dimtchev A, Lavin MF, Dritschilo A, Jung M (1998) A novel ionizing radiation-induced signaling pathway that activates the transcription factor NF-kappaB. Oncogene 17(14):1821–1826. doi:10.1038/sj.onc.1202088

Lezhava T (2001) Chromosome and aging: genetic conception of aging. Biogerontology 2(4):253–260

Li H, Xu D, Li J, Berndt MC, Liu JP (2006) Transforming growth factor beta suppresses human telomerase reverse transcriptase (hTERT) by Smad3 interactions with c-Myc and the hTERT gene. J Biol Chem 281(35):25588–25600. doi:10.1074/jbc.M602381200, M602381200 [pii]

Li N, Banin S, Ouyang H, Li GC, Courtois G, Shiloh Y, Karin M, Rotman G (2001) ATM is required for IkappaB kinase (IKKk) activation in response to DNA double strand breaks. J Biol Chem 276(12):8898–8903. doi:10.1074/jbc.M009809200

Linard C, Marquette C, Mathieu J, Pennequin A, Clarencon D, Mathe D (2004) Acute induction of inflammatory cytokine expression after gamma-irradiation in the rat: effect of an NF-kappaB inhibitor. Int J Radiat Oncol Biol Phys 58(2):427–434

Lu X, Murphy TC, Nanes MS, Hart CM (2010) PPAR{gamma} regulates hypoxia-induced Nox4 expression in human pulmonary artery smooth muscle cells through NF-{kappa}B. Am J Physiol Lung Cell Mol Physiol 299(4):L559–L566. doi:10.1152/ajplung.00090.2010

Mabb AM, Wuerzberger-Davis SM, Miyamoto S (2006) PIASy mediates NEMO sumoylation and NF-kappaB activation in response to genotoxic stress. Nat Cell Biol 8(9):986–993. doi:10.1038/ncb1458, ncb1458 [pii]

Maiti A, Maki G, Johnson P (1998) TNF-alpha induction of CD44-mediated leukocyte adhesion by sulfation. Science 282(5390):941–943

Mantovani A, Allavena P, Sica A, Balkwill F (2008) Cancer-related inflammation. Nature 454(7203):436–444. doi:10.1038/nature07205, nature07205 [pii]

Mariappan N, Elks CM, Fink B, Francis J (2009) TNF-induced mitochondrial damage: a link between mitochondrial complex I activity and left ventricular dysfunction. Free Radic Biol Med 46(4):462–470. doi:10.1016/j.freeradbiomed.2008.10.049

Markiewski MM, Lambris JD (2007) The role of complement in inflammatory diseases from behind the scenes into the spotlight. Am J Pathol 171(3):715–727. doi:10.2353/ajpath.2007.070166

Moiseeva O, Mallette FA, Mukhopadhyay UK, Moores A, Ferbeyre G (2006) DNA damage signaling and p53-dependent senescence after prolonged beta-interferon stimulation. Mol Biol Cell 17(4):1583–1592. doi:10.1091/mbc.E05-09-0858

Mothersill C, Seymour C (1997) Medium from irradiated human epithelial cells but not human fibroblasts reduces the clonogenic survival of unirradiated cells. Int J Radiat Biol 71(4):421–427

Nakano H, Shindo M, Sakon S, Nishinaka S, Mihara M, Yagita H, Okumura K (1998) Differential regulation of IkappaB kinase alpha and beta by two upstream kinases, NF-kappaB-inducing kinase and mitogen-activated protein kinase/ERK kinase kinase-1. Proc Natl Acad Sci USA 95(7):3537–3542

Narita M, Nunez S, Heard E, Narita M, Lin AW, Hearn SA, Spector DL, Hannon GJ, Lowe SW (2003) Rb-mediated heterochromatin formation and silencing of E2F target genes during cellular senescence. Cell 113(6):703–716

Naylor RM, Baker DJ, van Deursen JM (2013) Senescent cells: a novel therapeutic target for aging and age-related diseases. Clin Pharmacol Ther 93(1):105–116. doi:10.1038/clpt.2012.193

Negorev D, Maul GG (2001) Cellular proteins localized at and interacting within ND10/PML nuclear bodies/PODs suggest functions of a nuclear depot. Oncogene 20(49):7234–7242. doi:10.1038/sj.onc.1204764

Nelson G, Wordsworth J, Wang C, Jurk D, Lawless C, Martin-Ruiz C, von Zglinicki T (2012) A senescent cell bystander effect: senescence-induced senescence. Aging Cell 11(2):345–349. doi:10.1111/j.1474-9726.2012.00795.x

Nishioka C, Ikezoe T, Yang J, Yokoyama A (2009) Multitargeted tyrosine kinase inhibitor stimulates expression of IL-6 and activates JAK2/STAT5 signaling in acute myelogenous leukemia cells. Leukemia 23(12):2304–2308. doi:10.1038/leu.2009.163

Ogrunc M, Di Micco R, Liontos M, Bombardelli L, Mione M, Fumagalli M, Gorgoulis VG, d'Adda di Fagagna F (2014) Oncogene-induced reactive oxygen species fuel hyperproliferation and DNA damage response activation. Cell Death Differ 21(6):998–1012. doi:10.1038/cdd.2014.16

Osorio FG, Barcena C, Soria-Valles C, Ramsay AJ, de Carlos F, Cobo J, Fueyo A, Freije JM, Lopez-Otin C (2012) Nuclear lamina defects cause ATM-dependent NF-kappaB activation and link accelerated aging to a systemic inflammatory response. Genes Dev 26(20):2311–2324. doi:10.1101/gad.197954.112

Parrinello S, Coppe JP, Krtolica A, Campisi J (2005) Stromal-epithelial interactions in aging and cancer: senescent fibroblasts alter epithelial cell differentiation. J Cell Sci 118(Pt 3):485–496

Pearson M, Carbone R, Sebastiani C, Cioce M, Fagioli M, Saito S, Higashimoto Y, Appella E, Minucci S, Pandolfi PP, Pelicci PG (2000) PML regulates p53 acetylation and premature senescence induced by oncogenic Ras. Nature 406(6792):207–210

Piret B, Piette J (1996) Topoisomerase poisons activate the transcription factor NF-kappaB in ACH-2 and CEM cells. Nucleic Acids Res 24(21):4242–4248

Polo SE, Jackson SP (2011) Dynamics of DNA damage response proteins at DNA breaks: a focus on protein modifications. Genes Dev 25(5):409–433. doi:10.1101/gad.2021311

Qin Q, Inatome R, Hotta A, Kojima M, Yamamura H, Hirai H, Yoshizawa T, Tanaka H, Fukami K, Yanagi S (2006) A novel GTPase, CRAG, mediates promyelocytic leukemia protein-associated nuclear body formation and degradation of expanded polyglutamine protein. J Cell Biol 172(4):497–504. doi:10.1083/jcb.200505079

Rando TA, Chang HY (2012) Aging, rejuvenation, and epigenetic reprogramming: resetting the aging clock. Cell 148(1–2):46–57. doi:10.1016/j.cell.2012.01.003

Redon CE, Dickey JS, Nakamura AJ, Kareva IG, Naf D, Nowsheen S, Kryston TB, Bonner WM, Georgakilas AG, Sedelnikova OA (2010) Tumors induce complex DNA damage in distant proliferative tissues in vivo. Proc Natl Acad Sci USA 107(42):17992–17997. doi:10.1073/pnas.1008260107, 1008260107 [pii]

Regad T, Chelbi-Alix MK (2001) Role and fate of PML nuclear bodies in response to interferon and viral infections. Oncogene 20(49):7274–7286

Robinson-Smith TM, Isaacsohn I, Mercer CA, Zhou M, Van Rooijen N, Husseinzadeh N, McFarland-Mancini MM, Drew AF (2007) Macrophages mediate inflammation-enhanced metastasis of ovarian tumors in mice. Cancer Res 67(12):5708–5716. doi:10.1158/0008-5472.CAN-06-4375

Rodier F, Coppe JP, Patil CK, Hoeijmakers WA, Munoz DP, Raza SR, Freund A, Campeau E, Davalos AR, Campisi J (2009) Persistent DNA damage signalling triggers senescence-associated inflammatory cytokine secretion. Nat Cell Biol 11(8):973–979. doi:10.1038/ncb1909, ncb1909 [pii]

Sadaie M, Salama R, Carroll T, Tomimatsu K, Chandra T, Young AR, Narita M, Perez-Mancera PA, Bennett DC, Chong H, Kimura H, Narita M (2013) Redistribution of the Lamin B1

genomic binding profile affects rearrangement of heterochromatic domains and SAHF formation during senescence. Genes Dev 27(16):1800–1808. doi:10.1101/gad.217281.113

Sato S, Sanjo H, Takeda K, Ninomiya-Tsuji J, Yamamoto M, Kawai T, Matsumoto K, Takeuchi O, Akira S (2005) Essential function for the kinase TAK1 in innate and adaptive immune responses. Nat Immunol 6(11):1087–1095. doi:10.1038/ni1255

Savill JS, Wyllie AH, Henson JE, Walport MJ, Henson PM, Haslett C (1989) Macrophage phagocytosis of aging neutrophils in inflammation. Programmed cell death in the neutrophil leads to its recognition by macrophages. J Clin Invest 83(3):865–875. doi:10.1172/JCI113970

Seo SR, Ferrand N, Faresse N, Prunier C, Abecassis L, Pessah M, Bourgeade MF, Atfi A (2006) Nuclear retention of the tumor suppressor cPML by the homeodomain protein TGIF restricts TGF-beta signaling. Mol Cell 23(4):547–559. doi:10.1016/j.molcel.2006.06.018, S1097-2765(06)00434-5 [pii]

Shah PP, Donahue G, Otte GL, Capell BC, Nelson DM, Cao K, Aggarwala V, Cruickshanks HA, Rai TS, McBryan T, Gregory BD, Adams PD, Berger SL (2013) Lamin B1 depletion in senescent cells triggers large-scale changes in gene expression and the chromatin landscape. Genes Dev 27(16):1787–1799. doi:10.1101/gad.223834.113

Shchors K, Shchors E, Rostker F, Lawlor ER, Brown-Swigart L, Evan GI (2006) The Myc-dependent angiogenic switch in tumors is mediated by interleukin 1beta. Genes Dev 20(18):2527–2538. doi:10.1101/gad.1455706

Shoji Y, Uedono Y, Ishikura H, Takeyama N, Tanaka T (1995) DNA damage induced by tumour necrosis factor-alpha in L929 cells is mediated by mitochondrial oxygen radical formation. Immunology 84(4):543–548

Skov L, Beurskens FJ, Zachariae CO, Reitamo S, Teeling J, Satijn D, Knudsen KM, Boot EP, Hudson D, Baadsgaard O, Parren PW, van de Winkel JG (2008) IL-8 as antibody therapeutic target in inflammatory diseases: reduction of clinical activity in palmoplantar pustulosis. J Immunol 181(1):669–679

Sparmann A, Bar-Sagi D (2004) Ras-induced interleukin-8 expression plays a critical role in tumor growth and angiogenesis. Cancer Cell 6(5):447–458. doi:10.1016/j.ccr.2004.09.028

Stadler M, Chelbi-Alix MK, Koken MH, Venturini L, Lee C, Saib A, Quignon F, Pelicano L, Guillemin MC, Schindler C et al (1995) Transcriptional induction of the PML growth suppressor gene by interferons is mediated through an ISRE and a GAS element. Oncogene 11(12):2565–2573

Stilmann M, Hinz M, Arslan SC, Zimmer A, Schreiber V, Scheidereit C (2009) A nuclear poly(ADP-ribose)-dependent signalosome confers DNA damage-induced IkappaB kinase activation. Mol Cell 36(3):365–378. doi:10.1016/j.molcel.2009.09.032

Surh YJ, Chun KS, Cha HH, Han SS, Keum YS, Park KK, Lee SS (2001) Molecular mechanisms underlying chemopreventive activities of anti-inflammatory phytochemicals: down-regulation of COX-2 and iNOS through suppression of NF-kappa B activation. Mutat Res 480–481:243–268

Takahashi S, Hakuta M, Aiba K, Ito Y, Horikoshi N, Miura M, Hatake K, Ogata E (2003) Elevation of circulating plasma cytokines in cancer patients with high plasma parathyroid hormone-related protein levels. Endocr Relat Cancer 10(3):403–407

Valentino L, Pierre J (2006) JAK/STAT signal transduction: regulators and implication in hematological malignancies. Biochem Pharmacol 71(6):713–721. doi:10.1016/j.bcp.2005.12.017

Vijayachandra K, Lee J, Glick AB (2003) Smad3 regulates senescence and malignant conversion in a mouse multistage skin carcinogenesis model. Cancer Res 63(13):3447–3452

Wang CY, Mayo MW, Baldwin AS Jr (1996) TNF- and cancer therapy-induced apoptosis: potentiation by inhibition of NF-kappaB. Science 274(5288):784–787

Wang J, Jacob NK, Ladner KJ, Beg A, Perko JD, Tanner SM, Liyanarachchi S, Fishel R, Guttridge DC (2009) RelA/p65 functions to maintain cellular senescence by regulating genomic stability and DNA repair. EMBO Rep 10(11):1272–1278. doi:10.1038/embor.2009.197, embor2009197 [pii]

Wheelhouse NM, Chan YS, Gillies SE, Caldwell H, Ross JA, Harrison DJ, Prost S (2003) TNF-alpha induced DNA damage in primary murine hepatocytes. Int J Mol Med 12 (6):889–894

Wiklund FE, Bennet AM, Magnusson PK, Eriksson UK, Lindmark F, Wu L, Yaghoutyfam N, Marquis CP, Stattin P, Pedersen NL, Adami HO, Gronberg H, Breit SN, Brown DA (2010) Macrophage inhibitory cytokine-1 (MIC-1/GDF15): a new marker of all-cause mortality. Aging Cell 9(6):1057–1064. doi:10.1111/j.1474-9726.2010.00629.x

Wu ZH, Shi Y, Tibbetts RS, Miyamoto S (2006) Molecular linkage between the kinase ATM and NF-kappaB signaling in response to genotoxic stimuli. Science 311(5764):1141–1146. doi:10.1126/science.1121513

Yang Y, Xia F, Hermance N, Mabb A, Simonson S, Morrissey S, Gandhi P, Munson M, Miyamoto S, Kelliher MA (2011) A cytosolic ATM/NEMO/RIP1 complex recruits TAK1 to mediate the NF-kappaB and p38 mitogen-activated protein kinase (MAPK)/MAPK-activated protein 2 responses to DNA damage. Mol Cell Biol 31(14):2774–2786. doi:10.1128/MCB.01139-10

Ye X, Zerlanko B, Kennedy A, Banumathy G, Zhang R, Adams PD (2007) Downregulation of Wnt signaling is a trigger for formation of facultative heterochromatin and onset of cell senescence in primary human cells. Mol Cell 27(2):183–196

Yoon S, Woo SU, Kang JH, Kim K, Kwon MH, Park S, Shin HJ, Gwak HS, Chwae YJ (2010) STAT3 transcriptional factor activated by reactive oxygen species induces IL6 in starvation-induced autophagy of cancer cells. Autophagy 6(8):1125–1138

Yoon YS, Lee JH, Hwang SC, Choi KS, Yoon G (2005) TGF beta1 induces prolonged mitochondrial ROS generation through decreased complex IV activity with senescent arrest in Mv1Lu cells. Oncogene 24(11):1895–1903. doi:10.1038/sj.onc.1208262

Zhang B, Bowerman NA, Salama JK, Schmidt H, Spiotto MT, Schietinger A, Yu P, Fu YX, Weichselbaum RR, Rowley DA, Kranz DM, Schreiber H (2007) Induced sensitization of tumor stroma leads to eradication of established cancer by T cells. J Exp Med 204(1):49–55. doi:10.1084/jem.20062056

Zhang X, Blaskovich MA, Forinash KD, Sebti SM (2014) Withacnistin inhibits recruitment of STAT3 and STAT5 to growth factor and cytokine receptors and induces regression of breast tumours. Br J Cancer 111(5):894–902. doi:10.1038/bjc.2014.349

Ziech D, Franco R, Pappa A, Panayiotidis MI (2011) Reactive oxygen species (ROS)-induced genetic and epigenetic alterations in human carcinogenesis. Mutat Res 711(1–2):167–173. doi:10.1016/j.mrfmmm.2011.02.015

Zimmermann M, Lottersberger F, Buonomo SB, Sfeir A, de Lange T (2013) 53BP1 regulates DSB repair using Rif1 to control 5′ end resection. Science 339(6120):700–704. doi:10.1126/science.1231573

Manipulation of PML Nuclear Bodies and DNA Damage Responses by DNA Viruses

Lori Frappier

Abstract To undergo a successful infection, viruses must overcome intrinsic host responses that would otherwise suppress their gene expression or replication. For nuclear viruses, these host responses include promyelocytic leukemia (PML) nuclear bodies (NBs) and DNA damage responses (DDR). DNA viruses have been found to be particularly adept at disabling PML NBs and DDRs in order to overcome their inhibitory effects on viral infection. In addition, DNA viruses often induce and use components of the DDR to augment their own replication. Studying the interplay between viruses and PML NBs and DDRs has been hugely informative in understanding the regulation of these cellular structures and responses and their associated functions. Here we discuss the multiple mechanisms by which DNA viruses manipulate PML NBs and DDRs.

Keywords Herpesvirus • Adenovirus • Polyomavirus • Papillomavirus • DNA damage • ATM activation • Promyelocytic leukemia nuclear bodies

1 Introduction

Many viruses undergo their infectious cycles in the cell nucleus. Successful infection by these nuclear viruses involves the modulation of nuclear structures and processes that would otherwise inhibit viral gene expression and replication and/or trigger apoptosis. DNA viruses are particularly well known for their ability to modulate nuclear processes as they encode a large number of proteins whose only functions appear to be in altering the cellular environment. The best studied example of a nuclear structure that is altered by DNA viruses are PML NBs. PML NBs contribute to a variety of functions that impact viral infection, including apoptosis and DNA repair, and are also known to be part of the innate immune response to suppress viral infection. Accordingly, numerous viruses have been shown to associate with, alter or completely disrupt PML NBs. In addition, several DNA viruses have been shown to alter DNA repair pathways in order to promote

L. Frappier (✉)
Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada
e-mail: lori.frappier@utoronto.ca

D.P. Bazett-Jones, G. Dellaire (eds.), *The Functional Nucleus*,
DOI 10.1007/978-3-319-38882-3_13

viral DNA replication. This chapter will focus on the effects of DNA viruses on PML NBs and DNA damage responses (DDR).

2 DNA Viruses and PML Nuclear Bodies

2.1 Antiviral Function of PML Nuclear Bodies

One of the first indications of a relationship between PML NBs (initially called Nuclear Domain 10s; ND10s) and viruses came from observations that viral genomes and proteins often localized to PML NBs, leading to the initial interpretation that PML NBs positively contribute to viral transcription and/or replication (Ishov and Maul 1996; Maul et al. 1996; Doucas et al. 1996). However, this interpretation seemed counter to the observation that specific viral proteins that disrupted PML NBs promoted viral infection (Everett and Maul 1994; Korioth et al. 1996; Everett et al. 1998; Doucas et al. 1996; Ahn and Hayward 1997). Indeed, current data indicates that PML NBs are part of the innate immune response that suppresses the gene expression and/or replication of many viruses, a topic that has been the subject of several excellent reviews (Everett and Chelbi-Alix 2007; Geoffroy and Chelbi-Alix 2011; Saffert and Kalejta 2008; Tavalai and Stamminger 2008). Evidence supporting the antiviral role of PML NBs is summarized as follows. First, PML expression is directly induced by interferons (Lavau et al. 1995; Chelbi-Alix et al. 1995, 1998). Since interferons are part of the innate immune system that suppresses viral infection, this suggests that PML proteins play a role in innate immunity. Second, lack of PML can increase viral infection. In part, this has been shown using $PML^{-/-}$ mice, which were found to be much more susceptible to infection by some viruses, while their fibroblasts show increased viral production (Bonilla et al. 2002; Mitchell et al. 2014; Djavani et al. 2001; El McHichi et al. 2010). In addition silencing of PML has been shown to increase lytic infection of several viruses, although for viruses that efficiently induce PML loss such as HSV-1, the inhibitory effect of PML is only evident in viral mutants that are unable to disrupt PML NBs (Everett et al. 2006, 2008; Ullman and Hearing 2008; Tavalai et al. 2006; Sivachandran et al. 2012b; Kyratsous and Silverstein 2009). Third, while it was initially assumed that incoming viral genomes moved to PML NBs, an elegant study by Everett and Murray (2005) indicated the opposite; namely, at least for HSV-1, PML and associated proteins move to the viral genomes as they enter the nucleus resulting in de novo formation of PML NBs. This indicates that PML proteins have some mechanism of sensing viral genomes and sequestering them. Note that, although PML NBs clearly suppress infection by several types of viruses, this may not be true for all viruses. For example, multiple studies on papillomaviruses and polyomaviruses suggest that associations of these viral genomes and proteins with PML NBs may be beneficial for infection and

accordingly, these viruses do not induce loss of PML NBs [reviewed in (Tavalai and Stamminger 2008)].

2.2 *Mechanisms of Disruption of PML NBs*

In keeping with the ability of PML NBs to suppress viral infection, many viruses encode proteins that promote viral infection by disrupting or altering PML NBs (summarized in Table 1). These can be grouped into four general mechanisms. First, some viral proteins induce the degradation of PML proteins resulting in the loss of PML NBs. This can be due to E3 ubiquitin ligase activity of the viral protein, as is the case for ICP0 of herpes simplex virus type 1 (HSV-1) (Everett et al. 1998; Boutell et al. 2002), or due to recruitment of cellular proteins, as is the case for the EBNA1 protein of Epstein-Barr virus (EBV) (Sivachandran et al. 2008, 2010). Second, some viral proteins induce the dispersal of PML proteins from the NB such that the number of NBs is decreased but the cellular level of PML protein is unaffected. These proteins, which include IE1 from human cytomegalovirus (CMV) (Lee et al. 2004; Muller and Dejean 1999) and BZLF1 from EBV (Adamson and Kenney 2001; Bowling and Adamson 2006), are thought to induce PML dispersal by interfering with the SUMO modification of PML, which contributes to NB formation. A third way that viral proteins can inactivate PML NBs is by altering their overall structure (without loss or dispersal of PML proteins). The adenovirus E4orf3 protein, which reorganizes PML NBs into track-like structures,

Table 1 Viral proteins that disrupt PML NBs by known mechanisms

Virus	Protein	Effect on PML NB
HSV-1	ICP0	Degradation of PML
VZV	ORF61	PML dispersal
CMV	pp71	Degradation of Daxx and release of ATRX
CMV	IE1	PML dispersal
CMV	UL35	Relocalization of PML, Sp100, Daxx
EBV	BZLF1	PML dispersal
EBV	BRLF1/Rta	PML dispersal
EBV	BGLF4	PML dispersal
EBV	BNRF1	Disrupts ATRX-Daxx interaction
EBV	EBNA1	PML degradation
KSHV	K-Rta	Degradation of SUMO2/3-modified PML
KSHV	ORF75	Loss of ATRX and Daxx from PML NBs
KSHV	LANA2	Increased PML SUMOylation and degradation
γHV-68	ORF61	Reorganization of PML into nuclear tracks
γHV-68	ORF75c	PML degradation
HVS	ORF3	Degradation of Sp100
Adeno	E4orf3	Reorganization of PML into nuclear tracks

is the best characterized example (Doucas et al. 1996; Carvalho et al. 1995), although alteration of PML NB structure has also been reported by some CMV and EBV proteins (Salsman et al. 2008). Finally, several viral proteins have been found to promote viral infection by altering or degrading specific components of PML NBs other than the PML proteins themselves. In particular the degradation of Daxx and Sp100, two major components of PML NBs with antiviral effects, is induced by several proteins including CMV pp71 and adenovirus E1B55K [which target Daxx (Hwang and Kalejta 2007; Schreiner et al. 2010)] and ORF3 of Herpesvirus Saimiri [which targets Sp100 (Full et al. 2012)]. The detailed mechanisms by which specific proteins from herpesviruses and adenoviruses interfere with PML NB functions are discussed below.

2.3 Herpes Simplex and Varicella-Zoster Viruses

Studies on HSV-1 provided the first evidence of a relationship between viral infection and PML NBs and have been instrumental in recognizing the antiviral role of PML NBs. Initial studies showed that HSV-1 lytic infection rapidly induced the loss of PML NBs but that this effect was abrogated when an HSV-1 mutant lacking ICP0 was used (Everett and Maul 1994; Maul and Everett 1994). In addition, both HSV-1 genomes and ICP0 were found to localize to PML NBs very early in infection prior to loss of these NBs, and localization of HSV-1 genomes to PML NBs increased in the absence of ICP0 (Ishov and Maul 1996; Maul et al. 1993; Everett and Maul 1994; Maul and Everett 1994). ICP0 expression was also found to be sufficient to induce PML NB loss in the absence of HSV-1 infection. Subsequent studies on multiple ICP0 mutants showed a good correlation between the ability of ICP0 to induce loss of PML NBs and its ability to promote viral infection, strongly suggesting that the effect of ICP0 on PML NBs is functionally significant and supporting the contention that PML NBs suppress lytic infection (Everett 2000; Everett et al. 2009). Experiments in which PML, Sp100 and Daxx proteins were silenced further indicated the contributions of all of these NB components in restricting HSV-1 infection in the absence of ICP0 (Everett et al. 2006; Glass and Everett 2013). Furthermore, studies in cells expressing single PML isoforms indicated that PML isoforms I and II are responsible for the PML-dependent restriction of HSV-1 infection (Cuchet et al. 2011).

The mechanisms of PML disruption by ICP0 involves the degradation of PML proteins and was shown to be due to intrinsic E3 ubiquitin ligase activity of ICP0 (Boutell et al. 2002, 2003; Everett et al. 1998; Van Sant et al. 2001). The loss of PML NBs requires both the catalytic RING domain of ICP0 and its C-terminal region, which mediates the interaction with PML NBs (Everett and Maul 1994; Maul and Everett 1994; Everett et al. 2009; Boutell et al. 2003). In keeping with these observations, ICP0 induces the polyubiquitylation of PML proteins in cells, although it remains unclear whether this is a direct effect, since ICP0 has not been found to ubiquitylate PML in vitro (Boutell et al. 2003). PML proteins are highly

modified by SUMO and ICP0 has been shown to preferentially induce the degradation SUMO-modified PML (Boutell et al. 2003, 2011). In keeping with these observations, ICP0 contains at least one SUMO interacting motif (SIM) which can target ICP0 to SUMOylated PML isoforms (Boutell et al. 2011). Therefore one mechanism by which ICP0 induces PML degradation is by acting as a SUMO-targeted ubiquitin ligase (STUbL). However, ICP0 has also been shown to induce the degradation of PML isoform I that is not SUMO-modified, supporting a dual mechanism by which ICP0 targets PML proteins (Boutell et al. 2011; Cuchet-Lourenco et al. 2012). SUMO-independent interactions of ICP0 with PML I occurs through C-terminal PML sequences encoded by exon 9 that are unique to isoform I and N-terminal ICP0 sequences outside of the RING domain (Cuchet-Lourenco et al. 2012). In addition, PML I depletion delays the degradation of other PML isoforms by ICP0, suggesting that the PML I interaction facilitates SUMO-dependent degradation of other PML isoforms in the NB. Note that, while ICP0 is the HSV-1 protein that most dramatically affects PML NBs, a genome-wide screen of herpesvirus proteins for PML effects identified UL14 and US10 as also inducing PML NB loss (Salsman et al. 2008). Interestingly, UL14 had been reported to have anti-apoptotic activity which would be an expected outcome of PML loss (Yamauchi et al. 2003).

All of the above studies were in the context of lytic infection, but HSV-1 also has a latent mode of infection in neurons, in which stable noncoding RNA molecules called LAT are expressed in the absence of any viral protein expression. A recent microscopy study showed that PML forms a shell around latent HSV-1 genomes and that, in cells from PML knockout mice, there are more latent genomes and increased LAT expression (Catez et al. 2012). The results suggest that PML can restrict HSV-1 expression in latency as it does in lytic infection.

Lytic infection with Varizella Zoster Virus (VZV), another α-herpesvirus, has also been found to disrupt PML NBs and this activity has been shown to be due to the ICP0 orthologue, ORF61 (Wang et al. 2011). Like ICP0, ORF61 has E3 ubiquitin ligase activity in vitro (Everett et al. 2010), however unlike ICP0, ORF61 does not degrade PML (Kyratsous et al. 2009; Wang et al. 2011). Instead ORF61 induces the dispersal of PML proteins and this effect was shown to require its three SIMs (Wang et al. 2011). Infection with VZV containing a ORF61 mutated in these SIMS showed reduced viral spread and failed to cause lesions typical of VZV infection, supporting the importance of PML disruption in viral infection (Wang et al. 2011). Interestingly, in addition to PML effects on early aspects of VZV infection, PML was also found to inhibit VZV infection by forming cage-like structures around both mature and immature viral capsids (Reichelt et al. 2011). This phenomenon was shown to be due to PML isoform IV, which interacted with the ORF23 capsid protein by a mechanism involving PML IV sequences encoded by exon 8b (Reichelt et al. 2011). Antiviral activity of PML IV was also shown to require exon 8b, supporting a role for PML-mediated capsid caging in inhibiting viral infection (Reichelt et al. 2011).

2.4 *Cytomegalovirus*

Like HSV-1, human CMV genomes are associated with PML NBs, as are the two CMV immediately early proteins (IE1 and IE2) and the pp71 tegument protein, initially suggesting that PML NBs positively contribute to CMV replication (Ahn and Hayward 1997; Ishov and Maul 1996; Maul et al. 1996; Ishov et al. 1997, 2002; Hofmann et al. 2002; Marshall et al. 2002). However studies in cells in which PML was silenced showed that PML depletion enhanced viral replication and increased expression of IE1 and IE2, supporting a role for PML in restricting CMV infection (Tavalai et al. 2006). In addition, similar to the findings for HSV-1 infection, silencing of either PML, Daxx or Sp100 was found to promote CMV lytic infection and reactivation upon differentiation of a monocyte cell line to macrophages, but did not affect CMV latency (Wagenknecht et al. 2015).

A variety of studies indicate pp71 and IE1 play important roles in overcoming suppression by PML NBs early in infection (Tavalai and Stamminger 2008, 2011). Pp71 localizes to PML NBs through a direct interaction with Daxx and this interaction is required for pp71 to activate the CMV major immediate early promoter (MIEP) (Cantrell and Bresnahan 2006; Hofmann et al. 2002; Ishov et al. 2002). While these results initially suggested a positive role for Daxx in activating this CMV promoter, it was later shown that Daxx represses the MIEP and that pp71 relieves this repression (Woodhall et al. 2006; Preston and Nicholl 2006; Cantrell and Bresnahan 2006; Saffert and Kalejta 2006). Furthermore Daxx overexpression was shown to inhibit CMV infection, while Daxx silencing increased CMV gene expression and replication (Woodhall et al. 2006; Saffert and Kalejta 2006, 2007; Preston and Nicholl 2006; Cantrell and Bresnahan 2006). Pp71 inhibits repression by Daxx by inducing its degradation by a mechanism that appears to be independent of ubiquitylation but requires the proteasome (Saffert and Kalejta 2006, 2007). This mechanism might also involve SUMOylation of Daxx, since pp71 was found to increase the relative amount of SUMO-modified Daxx (Hwang and Kalejta 2009). It was also shown that, prior to inducing Daxx degradation, pp71 causes the chromatin remodeling protein ATRX to be released from PML NBs and that ATRX silencing increases IE gene expression, suggesting that this is part of the mechanism by which pp71 relieves repression of MIEP (Lukashchuk et al. 2008).

The actions of pp71 allow for expression of IE1 which then plays a primary role in overcoming suppression by PML NBs. IE1 induces the dispersal of PML and associated proteins from PML NBs and this ability correlates with its role in activating viral transcription (Ahn and Hayward 1997; Korioth et al. 1996). IE physically interacts with PML proteins and reduces their degree of SUMOylation without affecting total PML levels, and mutational analyses showed that the IE1-PML interaction is required for the loss of PML SUMOylation and for PML dispersal (Lee et al. 2004; Muller and Dejean 1999; Xu et al. 2001). IE1 was also shown to be modified by SUMO-1 (Muller and Dejean 1999) raising the possibility that IE1 may inhibit PML SUMOylation by competing for SUMO. However IE1

was also found to induce the dispersal of PML aggregates lacking SUMO, suggesting that loss of SUMO-modified PML may be a consequence of dispersal rather than a requirement for dispersal (Kang et al. 2006). Therefore the mechanism of PML dispersal by IE1 remains unclear.

In addition to pp71 and IE1, a screen of CMV nuclear proteins for PML NB affects identified ten proteins that reduce the number of PML NBs and three that caused obvious alterations to the size and shaped of PML NBs (Salsman et al. 2008). This suggests that other CMV proteins might also manipulate PML-associated processes later in infection. The CMV UL35 tegument protein was one protein found to dramatically alter PML NBs, by recruiting PML, Sp100 and Daxx into ring-shaped structures formed by UL35 (Salsman et al. 2008, 2011). Interestingly UL35 was previously shown to interact with pp71 and act cooperatively with it in activating the MIEP (Schierling et al. 2004; Liu and Biegalke 2002). The results as a whole suggest that UL35 contributes to MIEP activation through its ability to associate with and relocalize PML NB components.

2.5 Epstein-Barr Virus

Like other herpesviruses, EBV genomes in lytic infection are associated with PML NBs and lytic infection results in decreased levels of PML NBs (Bell et al. 2000; Amon et al. 2006). In addition, PML silencing promotes EBV reactivation and lytic infection (Sivachandran et al. 2012b; Sides et al. 2011). However, a positive role for PML NBs in EBV lytic infection was also recently reported (Wang et al. 2015). The EBV capsid protein, BORF1, was found to associate with PML NBs and recruit other capsid proteins to the NBs. Furthermore, PML silencing inhibited capsid assembly, suggesting that PML NBs are sites of capsid assembly.

A few different EBV proteins have been reported to alter PML NBs. The two immediate early proteins, BZLF1 and BRLF1, as well as the BGLF4 kinase were found to disrupt PML NBs through PML dispersal (Adamson and Kenney 2001; Salsman et al. 2008; Li et al. 2012; Kuny et al. 2010). This PML effect was shown to require the SIM sequences in BGLF4 and a portion of the transactivation domain of BZLF1 (Li et al. 2012; Adamson and Kenney 2001). BZLF1 reduced the proportion of SUMO-modified PML and, since BZLF1 is also SUMO-modified, this effect on PML may be due to competition for SUMO. However, a BZLF1 mutant lacking the SUMO-modification sites also disrupted PML NBs indicating that the mechanism is more complex (Adamson and Kenney 2001). In addition, the EBV tegument protein BNRF1 and the EBV latency protein EBNA-LP have been shown to impact PML NB functions through interactions with Daxx and Sp100, respectively (Ling et al. 2005; Tsai et al. 2011).

Another EBV protein that has a well-characterized role in PML disruption is EBNA1, a protein expressed in both latent and lytic forms of EBV infection. Latent infection of nasopharyngeal carcinoma and gastric carcinoma cells by EBV decreases the number of PML NBs and the level of PML proteins, and these levels

are restored by silencing EBNA1 (Sivachandran et al. 2008). Conversely, EBNA1 expression in the absence of EBV results in pronounced loss of PML NBs and degradation of PML proteins (Sivachandran et al. 2008). Microscopy studies showed that EBNA1 is partly localized to PML NBs, while co-IP studies with endogenous PML showed preferential association of EBNA1 with PML isoform IV (Sivachandran et al. 2008). EBNA1 itself lacks enzymatic activities but its ability to induce PML degradation was shown to involve interactions with two host enzymes, CK2 kinase and ubiquitin specific protease 7 (USP7), both of which can negatively regulate PML levels (Sivachandran et al. 2008; Sarkari et al. 2011; Scaglioni et al. 2006). EBNA1 recruits these proteins to PML NBs, resulting in increased phosphorylation of PML by CK2, which is a trigger for PML degradation (Sivachandran et al. 2008, 2010; Scaglioni et al. 2006). The mechanism by which USP7 contributes to PML loss is unclear but it does not involve its catalytic activity (Sarkari et al. 2011). Intriguingly, USP7 also binds to ICP0 and can affect the ability of ICP0 to induce PML loss (Meredith et al. 1994; Everett et al. 1997).

The ability of EBNA1 to affect PML-associated functions was also examined. In keeping with the loss of PML NBs, EBNA1 expression in nasopharyngeal and gastric carcinoma cells was found to decrease apoptosis and DNA repair in response to DNA damaging agents and to increase cell survival (Sivachandran et al. 2008, 2012a). EBNA1 expression also resulted in decreased acetylation of p53 and subsequent induction of p21 in response to DNA damage, consistent with the previously reported requirement of PML NBs for p53 acetylation (Sivachandran et al. 2012a; Pearson et al. 2000). In addition, a comparison of PML staining in EBV-positive and EBV-negative gastric carcinoma biopsy samples showed that the presence of EBV greatly decreased PML levels in these tumours, confirming that the effects of EBNA1 on PML in cell lines also occurs in tumours (Sivachandran et al. 2012a). Therefore the results as a whole suggest that EBNA1 contributes to cell survival and EBV-induced nasopharyngeal and gastric carcinoma by disrupting PML NBs.

Since EBNA1 is also expressed during lytic EBV infection, the possibility that EBNA1 contributes to lytic infection by inducing the loss of PML NBs was examined. Consistent with this hypothesis, EBNA1 silencing during lytic infection was found to decrease EBV gene expression and viral replication, while PML silencing increased both EBV gene expression and viral replication and led to spontaneous EBV reactivation (Sivachandran et al. 2012b; Sides et al. 2011). Furthermore the requirement for EBNA1 for efficient EBV lytic gene expression was abrogated by silencing PML (Sivachandran et al. 2012b). Studies in cell lines expressing single PML isoforms showed that any PML isoform could suppress EBV reactivation, however suppression by PML IV was only seen upon EBNA1 silencing because PML IV NBs were most efficiently disrupted by EBNA1 (Sivachandran et al. 2012b). These results support the functional importance of the previously observed preferential association of EBNA1 with PML IV.

2.6 *Kaposi's Sarcoma-Associated Herpesvirus, Gammaherpesvirus 68 and Herpesvirus Saimiri*

Like EBV, Kaposi's Sarcoma-Associated Herpesvirus (KSHV) is a gammaherpesvirus and encodes an immediate early protein closely related to BZLF1, called K8 or K-bZIP. While K8 localizes to PML NBs, unlike BZLF1, it does not disrupt PML NBs (Wu et al. 2001). Instead PML NB disruption has been reported to involve two KSHV proteins, LANA2 and K-Rta. LANA2, which is required for survival of KSHV-infected primary effusion lymphoma (PEL) cells, was found to increase SUMOylation of PML and induce PML degradation (Marcos-Villar et al. 2009). LANA2 was also shown to relieve the PML-mediated repression of survivin expression, suggesting a mechanism by which LANA2 contributes to the malignant progression of PEL (Marcos-Villar et al. 2009). The K-Rta major transactivating protein has ubiquitin ligase activity and SIMs, and was recently shown to induce the degradation of SUMO2/3-modified PML (Izumiya et al. 2013). A K-Rta SIM mutant failed to degrade PML and was defective for transactivation, suggesting that PML loss contributes to the transactivation function of K-Rta (Izumiya et al. 2013). The EBV homologue of K-Rta (Rta or BRLF1) has also been found to induce loss of PML NBs in some circumstances, an effect that might be related to its SUMO modification (Chang et al. 2004; Salsman et al. 2008; Adamson and Kenney 2001).

Murine gammaherpesvirus 68 (γHV-68) and herpesvirus saimiri (HVS) are close relatives of KSHV (all belonging to the rhadinovirus genera) that infect mice and squirrel monkeys, respectively. γHV-68 induces pronounced loss of PML NBs due to the activity of the ORF75c tegument protein, a member of the FGARAT family of proteins present in all gammaherpesviruses (Ling et al. 2008). ORF75c induces the ubiquitylation and degradation of PML proteins and has been shown to have auto-ubiquitylation activity in vitro, suggesting that it might directly ubiquitylate PML proteins (Tavalai and Stamminger 2008; Sewatanon and Ling 2013). However, γHV-68 expressing a catalytically inactive ORF75c was not defective in replication but rather had PML NBs reorganized into track-like structures, suggesting the presence of another PML-altering γHV-68 protein. This protein was identified as the viral ribonucleotide reductase large subunit ORF61, which was shown to induce PML NBs to form track-like structures (Sewatanon and Ling 2014).

In contrast to γHV-68, HVS infection induces the degradation of Sp100 without affecting PML. Interestingly this activity has been attributed to the ORF75c homologue, ORF3 (Full et al. 2012). ORF3 induces the selective degradation of Sp100 and is needed to overcome suppression by PML NBs and Sp100. Recently it was shown that the KSHV ORF75 affects PML NB function in yet another way, by inducing loss of the PML component ATRX and dispersal of Daxx (Full et al. 2014). This is similar to the EBV homologue, BNRF1, which interacts with Daxx at PML NBs and disrupts the Daxx-ATRX complex (Tsai et al. 2011). It is

remarkable that this family of viral FGARAT proteins have adopted so many different mechanisms to disable suppression by PML NBs.

2.7 Adenovirus

Infection with human adenovirus type 5 (Ad5) has the unusual effect of reorganizing PML NBs into elongated tracks, and the Ad5 protein encoded by E4orf3 was shown to be both necessary and sufficient for this effect (Doucas et al. 1996; Carvalho et al. 1995; Puvion-Dutilleul et al. 1995). This activity of E4orf3 antagonizes the ability of PML NBs to suppress viral infection and, unlike PML-mediate suppression of herpesvirus infections which mainly inhibits viral gene expression, PML NB-mediated suppression of Ad5 was found to occur predominantly at the level of viral replication (Evans and Hearing 2003; Ullman and Hearing 2008; Ullman et al. 2007). In addition to promoting Ad5 infection in human cells, E4orf3 has been shown to contribute to the ability of Ad5 to transform primary rodent cells and this function was found to involve its interaction with PML NBs, suggesting that the E4orf3-induced tracks may be impaired for tumour suppressive functions (Nevels et al. 1999). The association of E4orf3 with PML NBs involves a direct interaction with a unique sequence in the C-terminus of PML isoform II (Leppard et al. 2009). In addition, the reorganization of PML NBs involves the ability of the E4orf3 dimers to multimerize into cable-like polymers (Patsalo et al. 2012; Ou et al. 2012), as was shown by the finding that E4orf3 mutants that are unable to multimerize do not reorganize PML NBs (Patsalo et al. 2012; Ou et al. 2012). However, a multimerization mutant of E4orf3 was shown to reorganize PML when multimerization was restored by fusing it to lamin A/C, further emphasizing the need for polymer formation for E4orf3-mediated PML reorganization (Ou et al. 2012).

Another adenovirus protein that associates with PML NBs is the E1B-55K oncoprotein, which is known to contribute to cell transformation by repressing p53 functions. EIB-55K is SUMO1-modified and interacts with PML isoforms IV and V (Wimmer et al. 2010). Binding to PML IV involves the SUMO1-modified sequence of E1B-55K, while PML V binding occurs through a distinct sequence (Wimmer et al. 2010, 2015). Mutations in EIB-55K that disrupt binding to both PML IV and V also disrupt the ability of E1B-55K to induce p53 SUMOylation, inhibit p53-mediated transactivation and transform rodent cells, indicating the importance of the PML interactions in oncogenesis (Wimmer et al. 2015).

3 DNA Viruses and DNA Damage Responses

DNA viruses have a complicated relationship with cellular DNA damage responses that can be summarized into three general features [reviewed in (Turnell and Grand 2012; Weitzman et al. 2010; McFadden and Luftig 2013; Hollingworth and Grand 2015)]. First, DNA viruses induce some aspects of the DDR. At first glance this might appear to be a response of the cell to the incoming viral genomes, especially for linear DNA genomes whose ends resemble broken DNA. However studies have not supported this scenario. Rather the actions of specific viral proteins and/or viral replication are responsible for activating the DDR, suggesting that the virus activates the DDR by design. Second, DNA viruses inactivate some aspects of the DDR, such that DNA damage signaling does not result in apoptosis. This is also a role of specific viral proteins (summarized in Fig. 1). Third, specific cellular proteins that are part of the DDR are recruited to sites of viral replication and are required for efficient viral DNA replication. As a whole the observations support the model that DNA viruses intentionally activate some components in the DDR because these proteins or modifications are needed for viral DNA replication, but

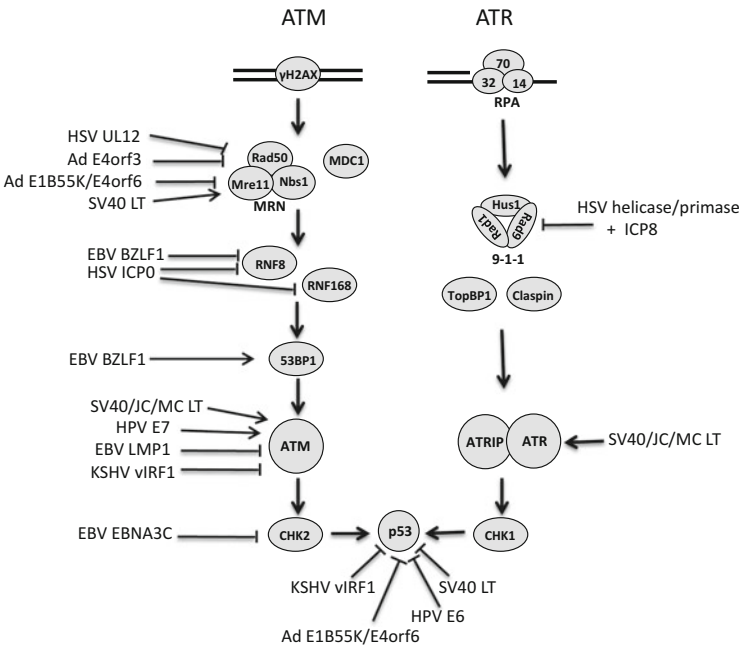

Fig. 1 Interactions of viral proteins with ATM and ATR signalling proteins. The key proteins in ATM and ATR signaling pathways are shown along with viral proteins that manipulate them. *Arrows* between viral and DDR proteins indicate cases where the viral protein uses or activates the DDR protein. *Blunted lines* indicate cases where the viral protein inactivates or degrades the DDR protein

inactivate other components so that cell survival is not compromised. Effects on the DDR are often closely linked to cell cycle alterations typical of specific viral infections and may also be related to PML effects, since PML NBs have important roles in mediating DDRs. The current state of knowledge on the relationship between specific DNA viruses and the DDRs is discussed below.

3.1 Herpes Simplex Virus

HSV-1 was first reported to induce a DDR by Wilkinson and Weller (2004), who showed that HSV-1 infections resulted in recruitment of the cellular homologous recombination (HR) proteins RPA, Nbs1 and Rad51 to sites of HSV-1 replication as well as hyper-phosphorylation of RPA and Nbs1, indicating activation of some aspects of the DDR. Both the recruitment and phosphorylation of these factors occurred after the loss of PML NBs and corresponded to recruitment of the viral polymerase and other replication factors to the HSV-1 genomes. These observations, in conjunction with previous work by the same authors (Wilkinson and Weller 2003), suggest that components of the HR pathway may contribute to HSV-1 DNA replication. In keeping with these observations, other studies found that HSV-1 activated ATM signaling and that activated ATM and the MRN complex components (Mre11, Rad50 and Nbs1) were recruited to sites of HSV-1 replication in an ATM-dependent manner (Lilley et al. 2005; Shirata et al. 2005). The recruitment of MRN was later shown to involve an interaction with the HSV-1 UL12 protein, which is important for recombination-dependent replication (Balasubramanian et al. 2010). The importance of ATM activation was further supported by the finding that HSV-1 infection is greatly decreased in the absence of ATM (Lilley et al. 2005). Furthermore, HSV-1 does not induce a DDR in neurons in which the virus does not replicate, suggesting that the inability to induce DDR may be a determinant in establishing latency (Lilley et al. 2005). Finally it has also been reported that p53 can positively contribute to HSV replication in promoting ICP22 expression (Maruzuru et al. 2013).

In addition to activating ATM, HSV-1 has been shown to inactivate ATR signaling. While several mechanisms of this inhibition have been proposed over the years, current data indicates that HSV-1 recruits ATR and its recruitment factor ATRIP to HSV-1 replication compartments but fails to recruit the 9-1-1 complex, an important step in ATR activation (Mohni et al. 2010). Interestingly, a complex formed by the HSV-1 helicase/primase complex and the ssDNA binding protein (ICP8) was shown to be necessary and sufficient to inactivate ATR signaling and to localize at sites of DNA damage with ATR, ATRIP and RPA. This suggests that these proteins disable ATR activation by binding DNA sites important for 9-1-1 recruitment and preventing 9-1-1 loading on the DNA (Mohni et al. 2013b). Not only are ATR-pathway proteins recruited to the sites of HSV-1 replication, but several (including ATR, ATRIP, RPA70, TopBP1, Claspin and CINP) have been shown to positively contribute to HSV-1 replication (Mohni et al. 2010, 2013a). In

addition, the mismatch repair proteins MSH2 and MLH1 have both been shown to be recruited to HSV-1 replication compartments and to be required for efficient HSV-1 replication (Mohni et al. 2011).

Finally another way that HSV-1 has been reported to affect DDRs is through the ICP0 ubiquitin ligase protein, which was found to induce the degradation of the cellular histone ubiquitin ligases RNF8 and RNF168. RNF8 and 168 are normally sequentially recruited to sites of double-stranded DNA breaks (DSB) in an Mdc1-dependent manner, where their further ubiquitylation of histone H2A and other targets leads to recruitment of downstream effectors of DSB repair (Bartocci and Denchi 2013). By inducing the degradation of RNF8 and 168, ICP0 was found to result in loss of ubiquitylated H2A and mobilization of DNA repair proteins (Lilley et al. 2010). Moreover, RNF8 was shown to suppress HSV-1 infection in the absence of ICP0 (Lilley et al. 2010). ICP0 targets RNF8 through a direct interaction between a phosphosite on ICP0 and the forkhead domain for RNF8, suggesting that ICP0 mimics a cellular phosphorylation mark (Chaurushiya et al. 2012). In addition, in the absence of ICP0, DDR proteins Mdc1 and 53BP1 are recruited to HSV-1 genomes as they enter the nucleus, whereas in the presence of ICP0 or in the absence of RNF8, recruitment of 53BP1 does not occur (Lilley et al. 2011). This suggests that RNF8 and DSB repair proteins downstream of it are part of an innate immune response to suppress HSV-1 replication, but that this suppression is prevented by ICP0-mediated degradation of RNF8 and RNF168.

3.2 Cytomegalovirus

Like HSV-1, CMV infection induces a DDR that involves ATM activation, and multiple proteins from the ATM-signalling pathway are recruited to viral replication sites (Luo et al. 2007b; E et al. 2011). The replication of CMV genomes was found to be compromised in cells with inactive or depleted ATM or with depleted H2AX (a target of ATM), indicating that ATM signaling is required for CMV replication (E et al. 2011). ATM signaling was shown to be activated by the viral IE1 protein and to be mediated by the cellular E2F1 protein (Xiaofei et al. 2011). Another CMV protein that affects the DDR is UL35. This tegument protein interacts with and relocalizes the CUL4A-DDB1-DCAF1 ubiquitin ligase complex that has functions in DNA repair (Salsman et al. 2012; Olma et al. 2009). This interaction involves the C-terminal region of DCAF1 and the N-terminal portion of UL35 since it did not occur with UL35a. Accordingly UL35 but not UL35a induced γH2AX/53BP1 DNA damage foci and G2 checkpoint activation (Salsman et al. 2012).

3.3 *Epstein-Barr Virus*

Like other herpesviruses, lytic infection by EBV induces ATM signaling without the downstream accumulation of p53, and also results in recruitment of multiple DDR proteins to the viral replication sites, including phosphorylated ATM, MRN, Rad51, hyper-phosphorylated RPA, γH2AX, MDC1 and RNF8 (Kudoh et al. 2005, 2009; Hau et al. 2015). Depletion of RPA32 and Rad51 indicated that these proteins are required for viral DNA synthesis (Kudoh et al. 2009). In addition, 53BP1 also appears to contribute to EBV lytic replication, as replication was reduced upon 53BP1 depletion (Bailey et al. 2009). 53BP1 interacts with the EBV BZLF1 protein, which is required for both transcriptional activation and lytic DNA replication, and BZLF1 mutants defective in 53BP1 binding were specifically disrupted for DNA replication, indicating the functional importance of the interaction (Bailey et al. 2009). BZLF1 has also been found to bind p53 and recruit it to sites of viral replication (Kudoh et al. 2005), and this could be a factor in the lack of p53 induction. In addition the EBNA1 protein, which is expressed in both latent and lytic modes of infection, may be an important contributor to the lack of p53 induction since EBNA1 has been shown to interfere with p53 induction by DDRs by blocking the ability of USP7 to bind and stabilize p53 (Saridakis et al. 2005; Holowaty et al. 2003). It has also been reported that several mismatch repair factors are recruited to lytically replicating EBV genomes but their functional importance is not known (Daikoku et al. 2006).

Although there have been mixed reports of the importance of ATM for EBV infection, current data favours an important role for ATM in lytic infection. Initially, Kudoh et al. (2005) reported that caffeine treatment did not affect expression of EBV lytic proteins in B95-8 B-cells, suggesting that ATM and ATR activation was not important for lytic infection. However, Hagemeier et al. (2012) later found that the ATM specific inhibitor ATM KU55933 inhibited the reactivation of EBV in both Akata Burkitt's lymphoma and gastric carcinoma cells. In addition, ATM or p53 silencing inhibited EBV reactivation in the carcinoma cells, while treatment with the p53 and ATM activator, nutlin, induced EBV reactivation (Hagemeier et al. 2012). Since BZLF1 overexpression over-rides the need for ATM, it was proposed that ATM contributes to the activation of the BZLF1 promoter. However the role of p53 in reactivation remains unknown. In addition, Hau et al. (2015) recently showed that EBV reactivation in nasopharyngeal carcinoma cells was inhibited by silencing or inhibiting ATM. Similarly ATM inhibition in Burkitt's lymphoma cell lines was recently reported by Wang'ondu et al. (2015) to greatly decrease EBV lytic protein expression in response to chemical induction. Hau et al. (2015) further showed that phosphorylation of the ATM target, Sp1, is critical for the recruitment of viral replication proteins to replication compartments and for subsequent viral replication, providing another rationale for the importance of ATM signaling in EBV lytic infection. The importance of DDR proteins in the EBV lytic cycle also fits with the fact that EBV

reactivation can be induced by treating with DNA damaging agents (Hagemeier et al. 2012; Westphal et al. 2000).

In addition to its induction by the DDR response, the BZLF1 (or ZEBRA) lytic protein itself affects the DDR in multiple ways. Expression of BZLF1 in the absence of other viral proteins has been shown to induce phospho-ATM foci and γH2AX, and this effect is abrogated by BZLF1 mutations that disrupt the DNA binding activity of BZLF1 (Wang'ondu et al. 2015). Conversely, BZLF1 has also been found to interfere with the formation of 53BP1 foci in response to DNA damage (Yang et al. 2015). This was shown to be due to failure to recruit RNF8 to DNA damage sites, apparently due to interference of BZLF1 with the Mdc-RNF8 interaction.

Like lytic infection, latent EBV infection can also induce ATM signaling during initial cell transformation and this has been shown to be attenuated by EBNA3C (Nikitin et al. 2010). In addition, cell transformation by EBV was increased by inhibiting ATM and Chk2 kinases, suggesting that growth transformation would be suppressed by the DDR if not attenuated by EBNA3C. This fits well with a previous study showing that EBNA3C is able to over-ride nocodazole-induced G2/M arrest through a direct interaction with Chk2 (Choudhuri et al. 2007). The DDR that occurs in latent EBV infection may involve LMP1 and EBNA1 as both have been found to induce DNA damage. LMP1 has been reported to induce DNA damage by inhibiting ATM (Gruhne et al. 2009), and recently co-expression of LMP1 and LMP2a was reported to interfere with γH2AX phosphorylation (Wasil et al. 2015). Several properties of EBNA1 likely contribute to its induction of DNA damage, including its ability to induce ROS (Gruhne et al. 2009; Cao et al. 2012), to induce the loss of PML NBs (Sivachandran et al. 2008) and to interfere with p53 stabilization by USP7 (Saridakis et al. 2005; Holowaty et al. 2003). Finally, some aspects of the DDR may positively contribute to EBV latent infection, as it was found that Mre11 and Nbs1 are recruited to the EBV latent origin of replication, oriP, and that their depletion inhibits oriP-dependent replication (Dheekollu et al. 2007).

3.4 KSHV

Like EBV, ATM signaling has been found to be activated during growth transformation as part of KSHV latent infection (Koopal et al. 2007; Singh et al. 2014). This DDR suppresses growth transformation by KSHV, and studies on the v-cyclin latency protein suggest that it is at least partly responsible for this growth suppression (Koopal et al. 2007). Effects of KSHV latent infection include induction of H2AX levels and phosphorylation, and recent studies indicate that γH2AX positively contributes to KSHV latent infection. Inhibition of ATM kinase or depletion of H2AX was shown to reduce the expression of LANA as well as its interactions with the KSHV terminal repeat sequences (Singh et al. 2014; Jha et al. 2013). This could have major effects on KSHV latent infection, since LANA has multiple roles

in latency including the replication and stable maintenance of the KSHV genomes through interaction with the terminal repeats (Jha et al. 2013).

Lytic infection by KSHV has been reported to result in increased DSBs, induction of γH2AX and activation of ATM and DNA-PK kinases without inducing 53BP1 foci (Xiao et al. 2013; Hollingworth et al. 2015). Inhibition of ATM was further shown to reduce viral replication, while inhibition of DNA-PK, which increases ATM activation, led to earlier viral release (Hollingworth et al. 2015). One protein that plays a major role in subverting the DDR is the viral interferon regulatory factor 1 (vIRF1). vIRF1 has been shown to bind and inhibit the activation of both p53 and ATM, the latter of which results in decreased γH2AX and Chk2 activation (Seo et al. 2001; Nakamura et al. 2001; Shin et al. 2006). Recently a screen for proteins inhibiting p53-mediated apoptosis identified several additional KSHV proteins that can also antagonize p53, suggesting that they may also participate in limiting DDRs (Chudasama et al. 2014).

3.5 Adenovirus

During adenovirus infection, multiple DDR proteins are recruited to the viral replication centers, including ATR, ATRIP, and RPA32, but little ATM or ATR activation is observed, due to the ability of adenovirus proteins to inactivate the MRN complex (Stracker et al. 2002; Carson et al. 2009). MRN inactivation involves E4orf3 as well as a complex of E1B55K and E4orf6. E4orf3 was shown to relocalize MRN into nuclear tracks as well as to cytoplasmic aggresomes (Stracker et al. 2005; Araujo et al. 2005; Liu et al. 2005). Like PML reorganization by E4orf3, the relocalization of MRN into nuclear tracks requires E4orf3 polymerization (Ou et al. 2012). Cytoplasmic aggresomes are induced by E1B-55K and contain an E3 ubiquitin ligase complex formed by E1B-55K, E4orf6 and the cellular proteins cullin 5, elongin B and C and RING-box 1 (Rbx1) (Harada et al. 2002; Liu et al. 2005). This complex ubiquitylates MRN, inducing its proteasomal degradation (Stracker et al. 2002; Carson et al. 2009). These MRN effects appear to promote adenovirus infection as an adenovirus mutant lacking the E4 region is greatly inhibited by MRN and ATM (Gautam and Bridge 2013).

A recent detailed study identified two temporally distinct ATM-mediated responses to adenovirus genomes (Shah and O'Shea 2015). In the first response, MRN binds early replicating viral genomes and recruits ATM, activating a localized signaling response that suppresses viral replication without affecting cellular DNA replication. This suppression is overcome by the actions of E1B-55K and E4-orf3. Later, the assembly of viral replication domains was shown to trigger a MRN-independent activation of ATM, leading to a more extensive DDR that does not inhibit viral replication. In both cases these adenovirus-associated DDRs differ from canonical cellular DDRs (Burgess and Misteli 2015).

A second important role of E1B-55K and E4orf6 in overcoming DDRs is in the degradation of p53. Infection by Ad 5 or Ad12 leads to rapid degradation of p53 and

this was found to be another function of the E1B-55K and E4orf6 proteins (Querido et al. 1997, 2001b; Steegenga et al. 1998). As for MRN degradation, p53 degradation involves polyubiquitylation at cellular aggresomes by a complex containing E1B55k/E4orf6 with cellular cullins (Cul5 for Ad5 and Cul2 for Ad12), elongin B and C and Rbx1 (Querido et al. 2001a; Harada et al. 2002; Luo et al. 2007a).

3.6 Polyomaviruses

Infection with SV40, JC, BK, Merkel cell (MC) or mouse polyomaviruses induces ATM and ATR signaling, and loss of ATM or ATR signaling reduces viral infection (albeit to various degrees in different viruses) (Justice et al. 2015; Dahl et al. 2005; Zhao et al. 2008; Shi et al. 2005; Orba et al. 2010; Jiang et al. 2012; Tsang et al. 2014). For SV40, JC and MC viruses, large T antigen (LT) alone is sufficient to activate ATM and ATR (Justice et al. 2015; Hein et al. 2009; Orba et al. 2010; Tsang et al. 2014). Activation of ATM and ATR by SV40 LT involves binding to Bub1, while this activation by JC LT involves DNA binding (Orba et al. 2010; Hein et al. 2009). In contrast ATM/ATR activation by BK polyomavirus is not triggered by LT alone but appears to involve viral DNA replication (Verhalen et al. 2015). In addition, multiple DDR proteins (including ATM, γH2AX, MRN, Rad51 and FANCD2) are recruited to sites of SV40 and MC polyomavirus replication, suggesting that some of them may directly contribute to DNA replication (Boichuk et al. 2010; Zhao et al. 2008; Hein et al. 2009; Tsang et al. 2014). Indeed, depletion of either Rad51 or FANCD2 decreased SV40 replication, confirming a positive role of these proteins (Boichuk et al. 2010). Interestingly ATM was found to phosphorylate large SV40 T antigen (LT) at Ser-120 at the onset of DNA replication, and mutation of Ser-120 impaired SV40 infection, indicating one reason that ATM signaling is important for SV40 replication (Shi et al. 2005). Similarly, LT from MC polyomavirus was recently shown to be phosphorylated by ATM, and mutagenesis of this site increased cell proliferation and decreased apoptosis (Li et al. 2015). Therefore regulation of LT activity is one reason that ATM activation is important for infection by at least some polyomaviruses.

In addition to inducing DDR proteins, SV40 infection was found to gradually induce the loss of MRN proteins and this also appears to be a function of LT (Zhao et al. 2008). Similar to adenovirus E1B55k/E4orf6, LT was found to form a E3 ubiquitin ligase complex with p185/Cul7, Rbx1, and the F box protein Fbw6 (Ali et al. 2004). A LT mutant disrupted in Cul7 binding was shown to have stable levels of MRN and reduce viral infection, indicating that MRN is inhibitory for viral infection and that the LT-Cul7 complex overcomes this inhibition by inducing MRN degradation (Zhao et al. 2008). Finally, the interaction of LT with p53 is known to be important for enabling viral infection. While p53 levels are elevated in polyomavirus infections, the direct interaction of LT with p53 inhibits p53 binding to DNA, thereby inactivating p53-mediated transactivation (Dey et al. 2002; Doherty and Freund 1997; Bargonetti et al. 1992; Jiang et al. 1993).

3.7 Papillomaviruses

Human papillomaviruses (HPV) initially infect undifferentiated epithelial cells, where they replicate to maintain a constant copy number, and, as cells differentiate and stop dividing, infection switches to a vegetative mode of replication involving DNA amplification (McKinney et al. 2015; Sakakibara et al. 2013). HPV infection has been found to induce ATM signaling to some degree in undifferentiated cells but to a greater degree in differentiated cells (Moody and Laimins 2009). In addition, ATM signaling only appears to contribute to infection in differentiated cells, as inhibition of the ATM kinase interferes with vegetative DNA replication in differentiated cells without affect HPV maintenance in undifferentiated cells (Moody and Laimins 2009). Examination of the localization of a variety of DDR factors in HPV infection showed that γH2AX, 53BP1, phospho-ATM and Chk2 all localized to sites of viral DNA replication in both differentiated and undifferentiated cells, although the foci are larger in differentiated cells (Gillespie et al. 2012). Moreover, ChIP assays indicated that γH2AX is associated with the viral origin of replication and that this association increases in vegetative DNA replication. In addition, the HR proteins Rad51 and BRCA1 were found to be induced in HPV-infected cells upon differentiation and to localize with the viral genomes (Gillespie et al. 2012). Therefore it appears that HR and ATM pathway proteins play active roles in HPV genome amplification in differentiated cells.

The induction of ATM signalling has been shown to involve HPV proteins, E1 and E7. E1 induces ATM signaling by a mechanism requiring its DNA binding and ATPase domains (Fradet-Turcotte et al. 2011; Sakakibara et al. 2011; Reinson et al. 2013; McKinney et al. 2015). E7 interacts with the ATM kinase and also induces replication stress that indirectly triggers ATM signaling (Moody and Laimins 2009; Bester et al. 2011). In addition, E7 activates STAT-5, which in turn activates ATM signaling (Hong and Laimins 2013). E7 also interacts with Nbs1 of the MRN complex and this interaction is required for viral replication (Anacker et al. 2014). In addition to the effects of E1 and E7, it was recently reported that E2 expressed on its own induces ATM signaling upon entry into mitosis in carcinoma cell lines (Xue et al. 2015).

Unlike other DNA viruses, to date identification of mechanisms to disable parts of the DDR in HPV infection have been lacking and are currently limited to p53 dysregulation. In particular, E6 can degrade and/or inactivate p53 and this is essential for long-term maintenance of the viral genomes (Park and Androphy 2002; Scheffner et al. 1990; Moody and Laimins 2010). In addition, inactivation or degradation of p53 by E6 is essential for genome amplification during vegetative infection, as p53 can directly inhibit vegetative replication independent of checkpoint activation (Kho et al. 2013; Lepik et al. 1998).

4 Summary

In summary, most DNA viruses have evolved multiple mechanisms to manipulate host PML NBs and DDRs in order to promote their gene expression and replication. PML NBs are repressive for most viral infections and uncovering the many mechanisms that viruses use to disable them has not only increased our understanding of viral infections, but also provided considerable insight into the relationship between the structure and function of PML NBs and mechanisms of regulation of PML NB components. DNA viruses also have a complicated relationship with the host DDR, in most cases both activating and inactivating components of the ATM signaling pathway. Although it is not surprising that some aspects of DDRs would inhibit viral infection and hence need to be neutralized by viral proteins, the finding that DNA viruses generally require some aspects of the DDRs, in particular ATM signaling, was unexpected. Understanding why particular DDR proteins are required for efficient viral replication will provide insight into the mechanisms of viral DNA replication, as well as a more complete understanding of viral-host interactions.

Acknowledgements Work in the Frappier laboratory is supported by grants from the Canadian Institutes of Health Research (CIHR) and the Canadian Cancer Society. L.F. is a tier 1 Canada Research Chair in Molecular Virology.

References

Adamson AL, Kenney S (2001) Epstein-Barr virus immediate-early protein BZLF1 is SUMO-1 modified and disrupts promyelocytic leukemia bodies. J Virol 75(5):2388–2399

Ahn JH, Hayward GS (1997) The major immediate-early proteins IE1 and IE2 of human cytomegalovirus colocalize with and disrupt PML-associated nuclear bodies at very early times in infected permissive cells. J Virol 71(6):4599–4613

Ali SH, Kasper JS, Arai T, DeCaprio JA (2004) Cul7/p185/p193 binding to simian virus 40 large T antigen has a role in cellular transformation. J Virol 78(6):2749–2757

Amon W, White RE, Farrell PJ (2006) Epstein-Barr virus origin of lytic replication mediates association of replicating episomes with promyelocytic leukaemia protein nuclear bodies and replication compartments. J Gen Virol 87(Pt 5):1133–1137. doi:87/5/1133 [pii] 10.1099/vir.0.81589-0

Anacker DC, Gautam D, Gillespie KA, Chappell WH, Moody CA (2014) Productive replication of human papillomavirus 31 requires DNA repair factor Nbs1. J Virol 88(15):8528–8544. doi:10.1128/JVI.00517-14

Araujo FD, Stracker TH, Carson CT, Lee DV, Weitzman MD (2005) Adenovirus type 5 E4orf3 protein targets the Mre11 complex to cytoplasmic aggresomes. J Virol 79(17):11382–11391. doi:79/17/11382 [pii] 10.1128/JVI.79.17.11382-11391.2005

Bailey SG, Verrall E, Schelcher C, Rhie A, Doherty AJ, Sinclair AJ (2009) Functional interaction between Epstein-Barr virus replication protein Zta and host DNA damage response protein 53BP1. J Virol 83(21):11116–11122. doi:JVI.00512-09 [pii] 10.1128/JVI.00512-09

Balasubramanian N, Bai P, Buchek G, Korza G, Weller SK (2010) Physical interaction between the herpes simplex virus type 1 exonuclease, UL12, and the DNA double-strand break-sensing MRN complex. J Virol 84(24):12504–12514. doi:10.1128/JVI.01506-10

Bargonetti J, Reynisdottir I, Friedman PN, Prives C (1992) Site-specific binding of wild-type p53 to cellular DNA is inhibited by SV40 T antigen and mutant p53. Genes Dev 6(10):1886–1898

Bartocci C, Denchi EL (2013) Put a RING on it: regulation and inhibition of RNF8 and RNF168 RING finger E3 ligases at DNA damage sites. Front Genet 4:128. doi:10.3389/fgene.2013.00128

Bell P, Lieberman PM, Maul GG (2000) Lytic but not latent replication of Epstein-Barr virus is associated with PML and induces sequential release of nuclear domain 10 proteins. J Virol 74(24):11800–11810

Bester AC, Roniger M, Oren YS, Im MM, Sarni D, Chaoat M, Bensimon A, Zamir G, Shewach DS, Kerem B (2011) Nucleotide deficiency promotes genomic instability in early stages of cancer development. Cell 145(3):435–446. doi:10.1016/j.cell.2011.03.044

Boichuk S, Hu L, Hein J, Gjoerup OV (2010) Multiple DNA damage signaling and repair pathways deregulated by simian virus 40 large T antigen. J Virol 84(16):8007–8020. doi:10.1128/JVI.00334-10

Bonilla WV, Pinschewer DD, Klenerman P, Rousson V, Gaboli M, Pandolfi PP, Zinkernagel RM, Salvato MS, Hengartner H (2002) Effects of promyelocytic leukemia protein on virus-host balance. J Virol 76(8):3810–3818

Boutell C, Sadis S, Everett RD (2002) Herpes simplex virus type 1 immediate-early protein ICP0 and is isolated RING finger domain act as ubiquitin E3 ligases in vitro. J Virol 76(2):841–850

Boutell C, Orr A, Everett RD (2003) PML residue lysine 160 is required for the degradation of PML induced by herpes simplex virus type 1 regulatory protein ICP0. J Virol 77(16):8686–8694

Boutell C, Cuchet-Lourenco D, Vanni E, Orr A, Glass M, McFarlane S, Everett RD (2011) A viral ubiquitin ligase has substrate preferential SUMO targeted ubiquitin ligase activity that counteracts intrinsic antiviral defence. PLoS Pathog 7(9), e1002245. doi:10.1371/journal.ppat.1002245, PPATHOGENS-D-11-00546 [pii]

Bowling BL, Adamson AL (2006) Functional interactions between the Epstein-Barr virus BZLF1 protein and the promyelocytic leukemia protein. Virus Res 117(2):244–253. doi:S0168-1702(05)00332-1 [pii] 10.1016/j.virusres.2005.10.018

Burgess RC, Misteli T (2015) Not all DDRs are created equal: non-canonical DNA damage responses. Cell 162(5):944–947. doi:10.1016/j.cell.2015.08.006

Cantrell SR, Bresnahan WA (2006) Human cytomegalovirus (HCMV) UL82 gene product (pp 71) relieves hDaxx-mediated repression of HCMV replication. J Virol 80(12):6188–6191

Cao JY, Mansouri S, Frappier L (2012) Changes in the nasopharyngeal carcinoma nuclear proteome induced by the EBNA1 protein of Epstein-Barr virus reveal potential roles for EBNA1 in metastasis and oxidative stress responses. J Virol 86(1):382–394. doi:10.1128/JVI.05648-11

Carson CT, Orazio NI, Lee DV, Suh J, Bekker-Jensen S, Araujo FD, Lakdawala SS, Lilley CE, Bartek J, Lukas J, Weitzman MD (2009) Mislocalization of the MRN complex prevents ATR signaling during adenovirus infection. Embo J 28(6):652–662. doi:10.1038/emboj.2009.15

Carvalho T, Seeler JS, Ohman K, Jordan P, Pettersson U, Akusjarvi G, Carmo-Fonseca M, Dejean A (1995) Targeting of adenovirus E1A and E4-ORF3 proteins to nuclear matrix-associated PML bodies. J Cell Biol 131(1):45–56

Catez F, Picard C, Held K, Gross S, Rousseau A, Theil D, Sawtell N, Labetoulle M, Lomonte P (2012) HSV-1 genome subnuclear positioning and associations with host-cell PML-NBs and centromeres regulate LAT locus transcription during latency in neurons. PLoS Pathog 8(8), e1002852. doi:10.1371/journal.ppat.1002852 PPATHOGENS-D-12-00913

Chang LK, Lee YH, Cheng TS, Hong YR, Lu PJ, Wang JJ, Wang WH, Kuo CW, Li SS, Liu ST (2004) Post-translational modification of Rta of Epstein-Barr virus by SUMO-1. J Biol Chem 279(37):38803–38812. doi:10.1074/jbc.M405470200

Chaurushiya MS, Lilley CE, Aslanian A, Meisenhelder J, Scott DC, Landry S, Ticau S, Boutell C, Yates JR 3rd, Schulman BA, Hunter T, Weitzman MD (2012) Viral E3 ubiquitin ligase-mediated degradation of a cellular E3: viral mimicry of a cellular phosphorylation mark targets the RNF8 FHA domain. Mol Cell 46(1):79–90. doi:10.1016/j.molcel.2012.02.004

Chelbi-Alix MK, Pelicano L, Quignon F, Koken MH, Venturini L, Stadler M, Pavlovic J, Degos L, de Thé H (1995) Induction of the PML protein by interferons in normal and APL cells. Leukemia 9(12):2027–2033

Chelbi-Alix MK, Quignon F, Pelicano L, Koken MH, de Thé H (1998) Resistance to virus infection conferred by the interferon-induced promyelocytic leukemia protein. J Virol 72(2): 1043–1051

Choudhuri T, Verma SC, Lan K, Murakami M, Robertson ES (2007) The ATM/ATR signaling effector Chk2 is targeted by Epstein-Barr virus nuclear antigen 3C to release the G2/M cell cycle block. J Virol 81(12):6718–6730. doi:JVI.00053-07

Chudasama P, Konrad A, Jochmann R, Lausen B, Holz P, Naschberger E, Neipel F, Britzen-Laurent N, Sturzl M (2014) Structural proteins of Kaposi's sarcoma-associated herpesvirus antagonize p53-mediated apoptosis. Oncogene. doi:10.1038/onc.2013.595

Cuchet D, Sykes A, Nicolas A, Orr A, Murray J, Sirma H, Heeren J, Bartelt A, Everett RD (2011) PML isoforms I and II participate in PML-dependent restriction of HSV-1 replication. J Cell Sci 124(Pt 2):280–291. doi:10.1242/jcs.075390

Cuchet-Lourenco D, Vanni E, Glass M, Orr A, Everett RD (2012) Herpes simplex virus 1 ubiquitin ligase ICP0 interacts with PML isoform I and induces its SUMO-independent degradation. J Virol 86(20):11209–11222. doi:JVI.01145-12

Dahl J, You J, Benjamin TL (2005) Induction and utilization of an ATM signaling pathway by polyomavirus. J Virol 79(20):13007–13017. doi:79/20/13007

Daikoku T, Kudoh A, Sugaya Y, Iwahori S, Shirata N, Isomura H, Tsurumi T (2006) Postreplicative mismatch repair factors are recruited to Epstein-Barr virus replication compartments. J Biol Chem 281(16):11422–11430

Dey D, Dahl J, Cho S, Benjamin TL (2002) Induction and bypass of p53 during productive infection by polyomavirus. J Virol 76(18):9526–9532

Dheekollu J, Deng Z, Wiedmer A, Weitzman MD, Lieberman PM (2007) A role for MRE11, NBS1, and recombination junctions in replication and stable maintenance of EBV episomes. PLoS ONE 2(12), e1257

Djavani M, Rodas J, Lukashevich IS, Horejsh D, Pandolfi PP, Borden KL, Salvato MS (2001) Role of the promyelocytic leukemia protein PML in the interferon sensitivity of lymphocytic choriomeningitis virus. J Virol 75(13):6204–6208. doi:10.1128/JVI.75.13.6204-6208.2001

Doherty J, Freund R (1997) Polyomavirus large T antigen overcomes p53 dependent growth arrest. Oncogene 14(16):1923–1931. doi:10.1038/sj.onc.1201025

Doucas V, Ishov AM, Romo A, Juguilon H, Weitzman MD, Evans RM, Maul GG (1996) Adenovirus replication is coupled with the dynamic properties of the PML nuclear structure. Genes Dev 10(2):196–207

Xiaofei E, Pickering MT, Debatis M, Castillo J, Lagadinos A, Wang S, Lu S, Kowalik TF (2011) An E2F1-mediated DNA damage response contributes to the replication of human cytomegalovirus. PLoS Pathog 7(5), e1001342. doi:10.1371/journal.ppat.1001342

El McHichi B, Regad T, Maroui MA, Rodriguez MS, Aminev A, Gerbaud S, Escriou N, Dianoux L, Chelbi-Alix MK (2010) SUMOylation promotes PML degradation during encephalomyocarditis virus infection. J Virol 84(22):11634–11645. doi:10.1128/JVI.01321-10

Evans JD, Hearing P (2003) Distinct roles of the Adenovirus E4 ORF3 protein in viral DNA replication and inhibition of genome concatenation. J Virol 77(9):5295–5304

Everett RD (2000) ICP0, a regulator of herpes simplex virus during lytic and latent infection. Bioessays 22(8):761–770

Everett RD, Chelbi-Alix MK (2007) PML and PML nuclear bodies: implications in antiviral defence. Biochimie 89(6–7):819–830

Everett RD, Maul GG (1994) HSV-1IE protein Vmw110 causes redistribution of PML. Embo J 13(21):5062–5069

Everett RD, Murray J (2005) ND10 components relocate to sites associated with herpes simplex virus type 1 nucleoprotein complexes during virus infection. J Virol 79(8):5078–5089

Everett R, Meredith M, Orr A, Cross A, Kathoria M, Parkinson J (1997) A novel ubiquitin-specific protease is dynamically associated with the PML nuclear domain and binds to a herpesvirus regulatory protein. EMBO J 16:1519–1530

Everett RD, Freemont P, Saitoh H, Dasso M, Orr A, Kathoria M, Parkinson J (1998) The disruption of ND10 during herpes simplex virus infection correlates with the Vmw110- and proteasome-dependent loss of several PML isoforms. J Virol 72(8):6581–6591

Everett RD, Rechter S, Papior P, Tavalai N, Stamminger T, Orr A (2006) PML contributes to a cellular mechanism of repression of herpes simplex virus type 1 infection that is inactivated by ICP0. J Virol 80(16):7995–8005

Everett RD, Parada C, Gripon P, Sirma H, Orr A (2008) Replication of ICP0-null mutant herpes simplex virus type 1 is restricted by both PML and Sp100. J Virol 82(6):2661–2672. doi: JVI.02308-07

Everett RD, Parsy ML, Orr A (2009) Analysis of the functions of herpes simplex virus type 1 regulatory protein ICP0 that are critical for lytic infection and derepression of quiescent viral genomes. J Virol 83(10):4963–4977. doi:10.1128/JVI.02593-08

Everett RD, Boutell C, McNair C, Grant L, Orr A (2010) Comparison of the biological and biochemical activities of several members of the alphaherpesvirus ICP0 family of proteins. J Virol 84(7):3476–3487. doi:10.1128/JVI.02544-09

Fradet-Turcotte A, Bergeron-Labrecque F, Moody CA, Lehoux M, Laimins LA, Archambault J (2011) Nuclear accumulation of the papillomavirus E1 helicase blocks S-phase progression and triggers an ATM-dependent DNA damage response. J Virol 85(17):8996–9012. doi:10.1128/JVI.00542-11

Full F, Reuter N, Zielke K, Stamminger T, Ensser A (2012) Herpesvirus saimiri antagonizes nuclear domain 10-instituted intrinsic immunity via an ORF3-mediated selective degradation of cellular protein Sp100. J Virol 86(7):3541–3553. doi:10.1128/JVI.06992-11

Full F, Jungnickl D, Reuter N, Bogner E, Brulois K, Scholz B, Sturzl M, Myoung J, Jung JU, Stamminger T, Ensser A (2014) Kaposi's sarcoma associated herpesvirus tegument protein ORF75 is essential for viral lytic replication and plays a critical role in the antagonization of ND10-instituted intrinsic immunity. PLoS Pathog 10(1), e1003863. doi:10.1371/journal.ppat.1003863 PPATHOGENS-D-13-01510

Gautam D, Bridge E (2013) The kinase activity of ataxia-telangiectasia mutated interferes with adenovirus E4 mutant DNA replication. J Virol 87(15):8687–8696. doi:10.1128/JVI.00376-13 JVI.00376-13 [pii]

Geoffroy MC, Chelbi-Alix MK (2011) Role of promyelocytic leukemia protein in host antiviral defense. J Interferon Cytokine Res 31(1):145–158. doi:10.1089/jir.2010.0111

Gillespie KA, Mehta KP, Laimins LA, Moody CA (2012) Human papillomaviruses recruit cellular DNA repair and homologous recombination factors to viral replication centers. J Virol 86 (17):9520–9526. doi:10.1128/JVI.00247-12 JVI.00247-12 [pii]

Glass M, Everett RD (2013) Components of promyelocytic leukemia nuclear bodies (ND10) act cooperatively to repress herpesvirus infection. J Virol 87(4):2174–2185. doi:10.1128/JVI.02950-12

Gruhne B, Sompallae R, Masucci MG (2009) Three Epstein-Barr virus latency proteins independently promote genomic instability by inducing DNA damage, inhibiting DNA repair and inactivating cell cycle checkpoints. Oncogene 28(45):3997–4008. doi:10.1038/onc.2009.258

Hagemeier SR, Barlow EA, Meng Q, Kenney SC (2012) The cellular ataxia telangiectasia-mutated kinase promotes Epstein-Barr virus lytic reactivation in response to multiple different types of lytic reactivation-inducing stimuli. J Virol 86(24):13360–13370. doi:10.1128/JVI.01850-12

Harada JN, Shevchenko A, Pallas DC, Berk AJ (2002) Analysis of the adenovirus E1B-55K-anchored proteome reveals its link to ubiquitination machinery. J Virol 76(18):9194–9206

Hau PM, Deng W, Jia L, Yang J, Tsurumi T, Chiang AK, Huen MS, Tsao SW (2015) Role of ATM in the formation of the replication compartment during lytic replication of Epstein-Barr virus in nasopharyngeal epithelial cells. J Virol 89(1):652–668. doi:10.1128/JVI.01437-14

Hein J, Boichuk S, Wu J, Cheng Y, Freire R, Jat PS, Roberts TM, Gjoerup OV (2009) Simian virus 40 large T antigen disrupts genome integrity and activates a DNA damage response via Bub1 binding. J Virol 83(1):117–127. doi:10.1128/JVI.01515-08

Hofmann H, Sindre H, Stamminger T (2002) Functional interaction between the pp 71 protein of human cytomegalovirus and the PML-interacting protein human Daxx. J Virol 76(11): 5769–5783

Hollingworth R, Grand RJ (2015) Modulation of DNA damage and repair pathways by human tumour viruses. Viruses 7(5):2542–2591. doi:10.3390/v7052542

Hollingworth R, Skalka GL, Stewart GS, Hislop AD, Blackbourn DJ, Grand RJ (2015) Activation of DNA damage response pathways during lytic replication of KSHV. Viruses 7(6): 2908–2927. doi:10.3390/v7062752

Holowaty MN, Sheng Y, Nguyen T, Arrowsmith C, Frappier L (2003) Protein interaction domains of the ubiquitin-specific protease, USP7/HAUSP. J Biol Chem 278(48):47753–47761. doi:10.1074/jbc.M307200200

Hong S, Laimins LA (2013) The JAK-STAT transcriptional regulator, STAT-5, activates the ATM DNA damage pathway to induce HPV 31 genome amplification upon epithelial differentiation. PLoS Pathog 9(4), e1003295. doi:10.1371/journal.ppat.1003295 PPATHOGENS-D-12-02723

Hwang J, Kalejta RF (2007) Proteasome-dependent, ubiquitin-independent degradation of Daxx by the viral pp 71 protein in human cytomegalovirus-infected cells. Virology 367(2):334–338

Hwang J, Kalejta RF (2009) Human cytomegalovirus protein pp 71 induces Daxx SUMOylation. J Virol 83(13):6591–6598. doi:10.1128/JVI.02639-08

Ishov AM, Maul GG (1996) The periphery of nuclear domain 10 (ND10) as site of DNA virus deposition. J Cell Biol 134(4):815–826

Ishov AM, Stenberg RM, Maul GG (1997) Human cytomegalovirus immediate early interaction with host nuclear structures: definition of an immediate transcript environment. J Cell Biol 138 (1):5–16

Ishov AM, Vladimirova OV, Maul GG (2002) Daxx-mediated accumulation of human cytomegalovirus tegument protein pp 71 at ND10 facilitates initiation of viral infection at these nuclear domains. J Virol 76(15):7705–7712

Izumiya Y, Kobayashi K, Kim KY, Pochampalli M, Izumiya C, Shevchenko B, Wang DH, Huerta SB, Martinez A, Campbell M, Kung HJ (2013) Kaposi's sarcoma-associated herpesvirus K-Rta exhibits SUMO-targeting ubiquitin ligase (STUbL) like activity and is essential for viral reactivation. PLoS Pathog 9(8), e1003506. doi:10.1371/journal.ppat.1003506 PPATHOGENS-D-12-01048

Jha HC, Upadhyay SK, Prasad AJ, Lu J, Cai Q, Saha A, Robertson ES (2013) H2AX phosphorylation is important for LANA-mediated Kaposi's sarcoma-associated herpesvirus episome persistence. J Virol 87(9):5255–5269. doi:10.1128/JVI.03575-12

Jiang D, Srinivasan A, Lozano G, Robbins PD (1993) SV40 T antigen abrogates p53-mediated transcriptional activity. Oncogene 8(10):2805–2812

Jiang M, Zhao L, Gamez M, Imperiale MJ (2012) Roles of ATM and ATR-mediated DNA damage responses during lytic BK polyomavirus infection. PLoS Pathog 8(8), e1002898. doi:10.1371/journal.ppat.1002898 PPATHOGENS-D-12-00806

Justice JL, Verhalen B, Jiang M (2015) Polyomavirus interaction with the DNA damage response. Virol Sin 30(2):122–129. doi:10.1007/s12250-015-3583-6

Kang H, Kim ET, Lee HR, Park JJ, Go YY, Choi CY, Ahn JH (2006) Inhibition of SUMO-independent PML oligomerization by the human cytomegalovirus IE1 protein. J Gen Virol 87(Pt 8):2181–2190. doi:87/8/2181 [pii]

Kho EY, Wang HK, Banerjee NS, Broker TR, Chow LT (2013) HPV-18 E6 mutants reveal p53 modulation of viral DNA amplification in organotypic cultures. Proc Natl Acad Sci U S A 110(19):7542–7549. doi:10.1073/pnas.1304855110

Koopal S, Furuhjelm JH, Jarviluoma A, Jaamaa S, Pyakurel P, Pussinen C, Wirzenius M, Biberfeld P, Alitalo K, Laiho M, Ojala PM (2007) Viral oncogene-induced DNA damage response is activated in Kaposi sarcoma tumorigenesis. PLoS Pathog 3(9):1348–1360. doi:07-PLPA-RA-0224 10.1371/journal.ppat.0030140

Korioth F, Maul GG, Plachter B, Stamminger T, Frey J (1996) The nuclear domain 10 (ND10) is disrupted by the human cytomegalovirus gene product IE1. Exp Cell Res 229(1):155–158. doi:S0014-4827(96)90353-5

Kudoh A, Fujita M, Zhang L, Shirata N, Daikoku T, Sugaya Y, Isomura H, Nishiyama Y, Tsurumi T (2005) Epstein-Barr virus lytic replication elicits ATM checkpoint signal transduction while providing an S-phase-like cellular environment. J Biol Chem 280(9):8156–8163

Kudoh A, Iwahori S, Sato Y, Nakayama S, Isomura H, Murata T, Tsurumi T (2009) Homologous recombinational repair factors are recruited and loaded onto the viral DNA genome in Epstein-Barr virus replication compartments. J Virol 83(13):6641–6651. doi:10.1128/JVI.00049-09JVI

Kuny CV, Chinchilla K, Culbertson MR, Kalejta RF (2010) Cyclin-dependent kinase-like function is shared by the beta- and gamma-subset of the conserved herpesvirus protein kinases. PLoS Pathog 6(9), e1001092. doi:10.1371/journal.ppat.1001092

Kyratsous CA, Silverstein SJ (2009) Components of nuclear domain 10 bodies regulate varicella-zoster virus replication. J Virol 83(9):4262–4274. doi:10.1128/JVI.00021-09

Kyratsous CA, Walters MS, Panagiotidis CA, Silverstein SJ (2009) Complementation of a herpes simplex virus ICP0 null mutant by varicella-zoster virus ORF61p. J Virol 83(20):10637–10643. doi:10.1128/JVI.01144-09

Lavau C, Marchio A, Fagioli M, Jansen J, Falini B, Lebon P, Grosveld F, Pandolfi PP, Pelicci PG, Dejean A (1995) The acute promyelocytic leukaemia-associated PML gene is induced by interferon. Oncogene 11(5):871–876

Lee HR, Kim DJ, Lee JM, Choi CY, Ahn BY, Hayward GS, Ahn JH (2004) Ability of the human cytomegalovirus IE1 protein to modulate sumoylation of PML correlates with its functional activities in transcriptional regulation and infectivity in cultured fibroblast cells. J Virol 78(12):6527–6542

Lepik D, Ilves I, Kristjuhan A, Maimets T, Ustav M (1998) p53 protein is a suppressor of papillomavirus DNA amplificational replication. J Virol 72(8):6822–6831

Leppard KN, Emmott E, Cortese MS, Rich T (2009) Adenovirus type 5 E4 Orf3 protein targets promyelocytic leukaemia (PML) protein nuclear domains for disruption via a sequence in PML isoform II that is predicted as a protein interaction site by bioinformatic analysis. J Gen Virol 90(Pt 1):95–104. doi:10.1099/vir.0.005512-0

Li R, Wang L, Liao G, Guzzo CM, Matunis MJ, Zhu H, Hayward SD (2012) SUMO binding by the Epstein-Barr virus protein kinase BGLF4 is crucial for BGLF4 function. J Virol 86(10):5412–5421. doi:10.1128/JVI.00314-12

Li J, Diaz J, Wang X, Tsang SH, You J (2015) Phosphorylation of Merkel cell polyomavirus large tumor antigen at serine 816 by ATM kinase induces apoptosis in host cells. J Biol Chem 290(3):1874–1884. doi:10.1074/jbc.M114.594895

Lilley CE, Carson CT, Muotri AR, Gage FH, Weitzman MD (2005) DNA repair proteins affect the lifecycle of herpes simplex virus 1. Proc Natl Acad Sci U S A 102(16):5844–5849. doi:0501916102

Lilley CE, Chaurushiya MS, Boutell C, Landry S, Suh J, Panier S, Everett RD, Stewart GS, Durocher D, Weitzman MD (2010) A viral E3 ligase targets RNF8 and RNF168 to control histone ubiquitination and DNA damage responses. Embo J 29(5):943–955. doi:emboj2009400

Lilley CE, Chaurushiya MS, Boutell C, Everett RD, Weitzman MD (2011) The intrinsic antiviral defense to incoming HSV-1 genomes includes specific DNA repair proteins and is

counteracted by the viral protein ICP0. PLoS Pathog 7(6), e1002084. doi:10.1371/journal.ppat.1002084 PPATHOGENS-D-11-00124

Ling PD, Peng RS, Nakajima A, Yu JH, Tan J, Moses SM, Yang WH, Zhao B, Kieff E, Bloch KD, Bloch DB (2005) Mediation of Epstein-Barr virus EBNA-LP transcriptional coactivation by Sp100. Embo J 24(20):3565–3575. doi:7600820

Ling PD, Tan J, Sewatanon J, Peng R (2008) Murine gammaherpesvirus 68 open reading frame 75c tegument protein induces the degradation of PML and is essential for production of infectious virus. J Virol 82(16):8000–8012. doi:10.1128/JVI.02752-07

Liu Y, Biegalke BJ (2002) The human cytomegalovirus UL35 gene encodes two proteins with different functions. J Virol 76(5):2460–2468

Liu Y, Shevchenko A, Berk AJ (2005) Adenovirus exploits the cellular aggresome response to accelerate inactivation of the MRN complex. J Virol 79(22):14004–14016. doi:79/22/14004

Lukashchuk V, McFarlane S, Everett RD, Preston CM (2008) Human cytomegalovirus protein pp 71 displaces the chromatin-associated factor ATRX from nuclear domain 10 at early stages of infection. J Virol 82(24):12543–12554. doi:10.1128/JVI.01215-08

Luo K, Ehrlich E, Xiao Z, Zhang W, Ketner G, Yu XF (2007a) Adenovirus E4orf6 assembles with Cullin5-ElonginB-ElonginC E3 ubiquitin ligase through an HIV/SIV Vif-like BC-box to regulate p53. Faseb J 21(8):1742–1750. doi:fj.06-7241com

Luo MH, Rosenke K, Czornak K, Fortunato EA (2007b) Human cytomegalovirus disrupts both ataxia telangiectasia mutated protein (ATM)- and ATM-Rad3-related kinase-mediated DNA damage responses during lytic infection. J Virol 81(4):1934–1950. doi:JVI.01670-06

Marcos-Villar L, Lopitz-Otsoa F, Gallego P, Munoz-Fontela C, Gonzalez-Santamaria J, Campagna M, Shou-Jiang G, Rodriguez MS, Rivas C (2009) Kaposi's sarcoma-associated herpesvirus protein LANA2 disrupts PML oncogenic domains and inhibits PML-mediated transcriptional repression of the survivin gene. J Virol 83(17):8849–8858. doi:10.1128/JVI.00339-09

Marshall KR, Rowley KV, Rinaldi A, Nicholson IP, Ishov AM, Maul GG, Preston CM (2002) Activity and intracellular localization of the human cytomegalovirus protein pp71. J Gen Virol 83(Pt 7):1601–1612

Maruzuru Y, Fujii H, Oyama M, Kozuka-Hata H, Kato A, Kawaguchi Y (2013) Roles of p53 in herpes simplex virus 1 replication. J Virol 87(16):9323–9332. doi:10.1128/JVI.01581-13

Maul GG, Everett RD (1994) The nuclear location of PML, a cellular member of the C3HC4 zinc-binding domain protein family, is rearranged during herpes simplex virus infection by the C3HC4 viral protein ICP0. J Gen Virol 75(Pt 6):1223–1233

Maul GG, Guldner HH, Spivack JG (1993) Modification of discrete nuclear domains induced by herpes simplex virus type 1 immediate early gene 1 product (ICP0). J Gen Virol 74(Pt 12):2679–2690

Maul GG, Ishov AM, Everett RD (1996) Nuclear domain 10 as preexisting potential replication start sites of herpes simplex virus type-1. Virology 217(1):67–75. doi:S0042-6822(96)90094-X

McFadden K, Luftig MA (2013) Interplay between DNA tumor viruses and the host DNA damage response. Curr Top Microbiol Immunol 371:229–257. doi:10.1007/978-3-642-37765-5_9

McKinney CC, Hussmann KL, McBride AA (2015) The role of the DNA damage response throughout the Papillomavirus life cycle. Viruses 7(5):2450–2469. doi:10.3390/v7052450

Meredith M, Orr A, Everett R (1994) Herpes simplex virus type 1 immediate-early protein Vmw110 binds strongly and specifically to a 135-kDa cellular protein. Virology 200(2):457–469

Mitchell AM, Hirsch ML, Li C, Samulski RJ (2014) Promyelocytic leukemia protein is a cell-intrinsic factor inhibiting parvovirus DNA replication. J Virol 88(2):925–936. doi:10.1128/JVI.02922-13

Mohni KN, Livingston CM, Cortez D, Weller SK (2010) ATR and ATRIP are recruited to herpes simplex virus type 1 replication compartments even though ATR signaling is disabled. J Virol 84(23):12152–12164. doi:10.1128/JVI.01643-10

Mohni KN, Mastrocola AS, Bai P, Weller SK, Heinen CD (2011) DNA mismatch repair proteins are required for efficient herpes simplex virus 1 replication. J Virol 85(23):12241–12253. doi:10.1128/JVI.05487-11

Mohni KN, Dee AR, Smith S, Schumacher AJ, Weller SK (2013a) Efficient herpes simplex virus 1 replication requires cellular ATR pathway proteins. J Virol 87(1):531–542. doi:10.1128/JVI.02504-12

Mohni KN, Smith S, Dee AR, Schumacher AJ, Weller SK (2013b) Herpes simplex virus type 1 single strand DNA binding protein and helicase/primase complex disable cellular ATR signaling. PLoS Pathog 9 (10):e1003652. doi:10.1371/journal.ppat.1003652 PPATHOGENS-D-13-00621

Moody CA, Laimins LA (2009) Human papillomaviruses activate the ATM DNA damage pathway for viral genome amplification upon differentiation. PLoS Pathog 5(10), e1000605. doi:10.1371/journal.ppat.1000605

Moody CA, Laimins LA (2010) Human papillomavirus oncoproteins: pathways to transformation. Nat Rev Cancer 10(8):550–560. doi:10.1038/nrc2886

Muller S, Dejean A (1999) Viral immediate-early proteins abrogate the modification by SUMO-1 of PML and Sp100 proteins, correlating with nuclear body disruption. J Virol 73(6):5137–5143

Nakamura H, Li M, Zarycki J, Jung JU (2001) Inhibition of p53 tumor suppressor by viral interferon regulatory factor. J Virol 75(16):7572–7582. doi:10.1128/JVI.75.16.7572-7582.2001

Nevels M, Tauber B, Kremmer E, Spruss T, Wolf H, Dobner T (1999) Transforming potential of the adenovirus type 5 E4orf3 protein. J Virol 73(2):1591–1600

Nikitin PA, Yan CM, Forte E, Bocedi A, Tourigny JP, White RE, Allday MJ, Patel A, Dave SS, Kim W, Hu K, Guo J, Tainter D, Rusyn E, Luftig MA (2010) An ATM/Chk2-mediated DNA damage-responsive signaling pathway suppresses Epstein-Barr virus transformation of primary human B cells. Cell Host Microbe 8(6):510–522. doi:S1931-3128(10)00377-X

Olma MH, Roy M, Le Bihan T, Sumara I, Maerki S, Larsen B, Quadroni M, Peter M, Tyers M, Pintard L (2009) An interaction network of the mammalian COP9 signalosome identifies Dda1 as a core subunit of multiple Cul4-based E3 ligases. J Cell Sci 122(Pt 7):1035–1044. doi:10.1242/jcs.043539

Orba Y, Suzuki T, Makino Y, Kubota K, Tanaka S, Kimura T, Sawa H (2010) Large T antigen promotes JC virus replication in G2-arrested cells by inducing ATM- and ATR-mediated G2 checkpoint signaling. J Biol Chem 285(2):1544–1554. doi:10.1074/jbc.M109.064311

Ou HD, Kwiatkowski W, Deerinck TJ, Noske A, Blain KY, Land HS, Soria C, Powers CJ, May AP, Shu X, Tsien RY, Fitzpatrick JA, Long JA, Ellisman MH, Choe S, O'Shea CC (2012) A structural basis for the assembly and functions of a viral polymer that inactivates multiple tumor suppressors. Cell 151(2):304–319. doi:10.1016/j.cell.2012.08.035

Park RB, Androphy EJ (2002) Genetic analysis of high-risk e6 in episomal maintenance of human papillomavirus genomes in primary human keratinocytes. J Virol 76(22):11359–11364

Patsalo V, Yondola MA, Luan B, Shoshani I, Kisker C, Green DF, Raleigh DP, Hearing P (2012) Biophysical and functional analyses suggest that adenovirus E4-ORF3 protein requires higher-order multimerization to function against promyelocytic leukemia protein nuclear bodies. J Biol Chem 287(27):22573–22583. doi:10.1074/jbc.M112.344234

Pearson M, Carbone R, Sebastiani C, Cioce M, Fagioli M, Saito S, Higashimoto Y, Appella E, Minucci S, Pandolfi PP, Pelicci PG (2000) PML regulates p53 acetylation and premature senescence induced by oncogenic Ras. Nature 406(6792):207–210

Preston CM, Nicholl MJ (2006) Role of the cellular protein hDaxx in human cytomegalovirus immediate-early gene expression. J Gen Virol 87(Pt 5):1113–1121. doi:87/5/1113

Puvion-Dutilleul F, Chelbi-Alix MK, Koken M, Quignon F, Puvion E, de The H (1995) Adenovirus infection induces rearrangements in the intranuclear distribution of the nuclear body-associated PML protein. Exp Cell Res 218(1):9–16. doi:S0014-4827(85)71125-1

Querido E, Marcellus RC, Lai A, Charbonneau R, Teodoro JG, Ketner G, Branton PE (1997) Regulation of p53 levels by the E1B 55-kilodalton protein and E4orf6 in adenovirus-infected cells. J Virol 71(5):3788–3798

Querido E, Blanchette P, Yan Q, Kamura T, Morrison M, Boivin D, Kaelin WG, Conaway RC, Conaway JW, Branton PE (2001a) Degradation of p53 by adenovirus E4orf6 and E1B55K proteins occurs via a novel mechanism involving a Cullin-containing complex. Genes Dev 15 (23):3104–3117. doi:10.1101/gad.926401

Querido E, Morrison MR, Chu-Pham-Dang H, Thirlwell SW, Boivin D, Branton PE (2001b) Identification of three functions of the adenovirus e4orf6 protein that mediate p53 degradation by the E4orf6-E1B55K complex. J Virol 75(2):699–709. doi:10.1128/JVI.75.2.699-709.2001

Reichelt M, Wang L, Sommer M, Perrino J, Nour AM, Sen N, Baiker A, Zerboni L, Arvin AM (2011) Entrapment of viral capsids in nuclear PML cages is an intrinsic antiviral host defense against varicella-zoster virus. PLoS Pathog 7(2), e1001266. doi:10.1371/journal.ppat.1001266

Reinson T, Toots M, Kadaja M, Pipitch R, Allik M, Ustav E, Ustav M (2013) Engagement of the ATR-dependent DNA damage response at the human papillomavirus 18 replication centers during the initial amplification. J Virol 87(2):951–964. doi:10.1128/JVI.01943-12

Saffert RT, Kalejta RF (2006) Inactivating a cellular intrinsic immune defense mediated by Daxx is the mechanism through which the human cytomegalovirus pp 71 protein stimulates viral - immediate-early gene expression. J Virol 80(8):3863–3871

Saffert RT, Kalejta RF (2007) Human cytomegalovirus gene expression is silenced by Daxx-mediated intrinsic immune defense in model latent infections established in vitro. J Virol 81 (17):9109–9120

Saffert RT, Kalejta RF (2008) Promyelocytic leukemia-nuclear body proteins: herpesvirus enemies, accomplices, or both? Future Virol 3(3):265–277. doi:10.2217/17460794.3.3.265

Sakakibara N, Mitra R, McBride AA (2011) The papillomavirus E1 helicase activates a cellular DNA damage response in viral replication foci. J Virol 85(17):8981–8995. doi:10.1128/JVI.00541-11

Sakakibara N, Chen D, McBride AA (2013) Papillomaviruses use recombination-dependent replication to vegetatively amplify their genomes in differentiated cells. PLoS Pathog 9(7), e1003321. doi:10.1371/journal.ppat.1003321 PPATHOGENS-D-13-00367

Salsman J, Zimmerman N, Chen T, Domagala M, Frappier L (2008) Genome-wide screen of three herpesviruses for protein subcellular localization and alteration of PML nuclear bodies. PLoS Pathog 4(7), e1000100

Salsman J, Wang X, Frappier L (2011) Nuclear body formation and PML body remodeling by the human cytomegalovirus protein UL35. Virology 414(2):119–129. doi:10.1016/j.virol.2011.03.013

Salsman J, Jagannathan M, Paladino P, Chan PK, Dellaire G, Raught B, Frappier L (2012) Proteomic profiling of the human cytomegalovirus UL35 gene products reveals a role for UL35 in the DNA repair response. J Virol 86(2):806–820. doi:10.1128/JVI.05442-11

Saridakis V, Sheng Y, Sarkari F, Holowaty MN, Shire K, Nguyen T, Zhang RG, Liao J, Lee W, Edwards AM, Arrowsmith CH, Frappier L (2005) Structure of the p53 binding domain of HAUSP/USP7 bound to Epstein-Barr nuclear antigen 1 implications for EBV-mediated immortalization. Mol Cell 18(1):25–36. doi:S1097-2765(05)01145-7

Sarkari F, Wang X, Nguyen T, Frappier L (2011) The herpesvirus associated ubiquitin specific protease, USP7, is a negative regulator of PML proteins and PML nuclear bodies. PLoS ONE 6 (1), e16598. doi:10.1371/journal.pone.0016598

Scaglioni PP, Yung TM, Cai LF, Erdjument-Bromage H, Kaufman AJ, Singh B, Teruya-Feldstein-J, Tempst P, Pandolfi PP (2006) A CK2-dependent mechanism for degradation of the PML tumor suppressor. Cell 126(2):269–283

Scheffner M, Werness BA, Huibregtse JM, Levine AJ, Howley PM (1990) The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. Cell 63(6):1129–1136. doi:0092-8674(90)90409-8 [pii]

Schierling K, Stamminger T, Mertens T, Winkler M (2004) Human cytomegalovirus tegument proteins pp UL82 (pp71) and ppUL35 interact and cooperatively activate the major immediate-early enhancer. J Virol 78(17):9512–9523

Schreiner S, Wimmer P, Sirma H, Everett RD, Blanchette P, Groitl P, Dobner T (2010) Proteasome-dependent degradation of Daxx by the viral E1B-55K protein in human adenovirus-infected cells. J Virol 84(14):7029–7038. doi:10.1128/JVI.00074-10

Seo T, Park J, Lee D, Hwang SG, Choe J (2001) Viral interferon regulatory factor 1 of Kaposi's sarcoma-associated herpesvirus binds to p53 and represses p53-dependent transcription and apoptosis. J Virol 75(13):6193–6198. doi:10.1128/JVI.75.13.6193-6198.2001

Sewatanon J, Ling PD (2013) Murine gammaherpesvirus 68 ORF75c contains ubiquitin E3 ligase activity and requires PML SUMOylation but not other known cellular PML regulators, CK2 and E6AP, to mediate PML degradation. Virology 440(2):140–149. doi:10.1016/j.virol.2013.02.014

Sewatanon J, Ling PD (2014) Murine gammaherpesvirus 68 encodes a second PML-modifying protein. J Virol 88(6):3591–3597. doi:10.1128/JVI.03081-13

Shah GA, O'Shea CC (2015) Viral and cellular genomes activate distinct DNA damage responses. Cell 162(5):987–1002. doi:10.1016/j.cell.2015.07.058

Shi Y, Dodson GE, Shaikh S, Rundell K, Tibbetts RS (2005) Ataxia-telangiectasia-mutated (ATM) is a T-antigen kinase that controls SV40 viral replication in vivo. J Biol Chem 280(48):40195–40200. doi:C500400200

Shin YC, Nakamura H, Liang X, Feng P, Chang H, Kowalik TF, Jung JU (2006) Inhibition of the ATM/p53 signal transduction pathway by Kaposi's sarcoma-associated herpesvirus interferon regulatory factor 1. J Virol 80(5):2257–2266. doi:80/5/2257

Shirata N, Kudoh A, Daikoku T, Tatsumi Y, Fujita M, Kiyono T, Sugaya Y, Isomura H, Ishizaki K, Tsurumi T (2005) Activation of ataxia telangiectasia-mutated DNA damage checkpoint signal transduction elicited by herpes simplex virus infection. J Biol Chem 280(34):30336–30341

Sides MD, Block GJ, Shan B, Esteves KC, Lin Z, Flemington EK, Lasky JA (2011) Arsenic mediated disruption of promyelocytic leukemia protein nuclear bodies induces ganciclovir susceptibility in Epstein-Barr positive epithelial cells. Virology 416(1–2):86–97. doi:10.1016/j.virol.2011.04.005

Singh VV, Dutta D, Ansari MA, Dutta S, Chandran B (2014) Kaposi's sarcoma-associated herpesvirus induces the ATM and H2AX DNA damage response early during de novo infection of primary endothelial cells, which play roles in latency establishment. J Virol 88(5):2821–2834. doi:10.1128/JVI.03126-13

Sivachandran N, Sarkari F, Frappier L (2008) Epstein-Barr nuclear antigen 1 contributes to nasopharyngeal carcinoma through disruption of PML nuclear bodies. PLoS Pathog 4(10), e1000170. doi:10.1371/journal.ppat.1000170

Sivachandran N, Cao JY, Frappier L (2010) Epstein-Barr virus nuclear antigen 1 hijacks the host kinase CK2 to disrupt PML nuclear bodies. J Virol 84(21):11113–11123

Sivachandran N, Dawson CW, Young LS, Liu FF, Middeldorp J, Frappier L (2012a) Contributions of the Epstein-Barr virus EBNA1 protein to gastric carcinoma. J Virol 86(1):60–68. doi:JVI.05623-11

Sivachandran N, Wang X, Frappier L (2012b) Functions of the Epstein-Barr virus EBNA1 protein in viral reactivation and lytic infection. J Virol 86(11):6146–6158. doi:JVI.00013-12

Steegenga WT, Riteco N, Jochemsen AG, Fallaux FJ, Bos JL (1998) The large E1B protein together with the E4orf6 protein target p53 for active degradation in adenovirus infected cells. Oncogene 16(3):349–357. doi:10.1038/sj.onc.1201540

Stracker TH, Carson CT, Weitzman MD (2002) Adenovirus oncoproteins inactivate the Mre11-Rad50-NBS1 DNA repair complex. Nature 418(6895):348–352. doi:10.1038/nature00863

Stracker TH, Lee DV, Carson CT, Araujo FD, Ornelles DA, Weitzman MD (2005) Serotype-specific reorganization of the Mre11 complex by adenoviral E4orf3 proteins. J Virol 79(11):6664–6673. doi:79/11/6664

Tavalai N, Stamminger T (2008) New insights into the role of the subnuclear structure ND10 for viral infection. Biochim Biophys Acta 1783(11):2207–2221. doi:10.1016/j.bbamcr.2008.08.004

Tavalai N, Stamminger T (2011) Intrinsic cellular defense mechanisms targeting human cytomegalovirus. Virus Res 157(2):128–133. doi:S0168-1702(10)00361-8

Tavalai N, Papior P, Rechter S, Leis M, Stamminger T (2006) Evidence for a role of the cellular ND10 protein PML in mediating intrinsic immunity against human cytomegalovirus infections. J Virol 80(16):8006–8018

Tsai K, Thikmyanova N, Wojcechowskyj JA, Delecluse HJ, Lieberman PM (2011) EBV tegument protein BNRF1 disrupts DAXX-ATRX to activate viral early gene transcription. PLoS Pathog 7(11), e1002376. doi:10.1371/journal.ppat.1002376 PPATHOGENS-D-10-00589

Tsang SH, Wang X, Li J, Buck CB, You J (2014) Host DNA damage response factors localize to merkel cell polyomavirus DNA replication sites to support efficient viral DNA replication. J Virol 88(6):3285–3297. doi:10.1128/JVI.03656-13

Turnell AS, Grand RJ (2012) DNA viruses and the cellular DNA-damage response. J Gen Virol 93(Pt 10):2076–2097. doi:10.1099/vir.0.044412-0

Ullman AJ, Hearing P (2008) Cellular proteins PML and Daxx mediate an innate antiviral defense antagonized by the adenovirus E4 ORF3 protein. J Virol 82(15):7325–7335

Ullman AJ, Reich NC, Hearing P (2007) Adenovirus E4 ORF3 protein inhibits the interferon-mediated antiviral response. J Virol 81(9):4744–4752. doi:JVI.02385-06 [pii]

Van Sant C, Hagglund R, Lopez P, Roizman B (2001) The infected cell protein 0 of herpes simplex virus 1 dynamically interacts with proteasomes, binds and activates the cdc34 E2 ubiquitin-conjugating enzyme, and possesses in vitro E3 ubiquitin ligase activity. Proc Natl Acad Sci U S A 98(15):8815–8820. doi:10.1073/pnas.161283098

Verhalen B, Justice JL, Imperiale MJ, Jiang M (2015) Viral DNA replication-dependent DNA damage response activation during BK polyomavirus infection. J Virol 89(9):5032–5039. doi:10.1128/JVI.03650-14

Wagenknecht N, Reuter N, Scherer M, Reichel A, Muller R, Stamminger T (2015) Contribution of the major ND10 proteins PML, hDaxx and Sp100 to the regulation of human cytomegalovirus latency and lytic replication in the monocytic cell line THP-1. Viruses 7(6):2884–2907. doi:10.3390/v7062751

Wang L, Oliver SL, Sommer M, Rajamani J, Reichelt M, Arvin AM (2011) Disruption of PML nuclear bodies is mediated by ORF61 SUMO-interacting motifs and required for varicella-zoster virus pathogenesis in skin. PLoS Pathog 7(8), e1002157. doi:10.1371/journal.ppat.1002157 PPATHOGENS-D-11-00011

Wang WH, Kuo CW, Chang LK, Hung CC, Chang TH, Liu ST (2015) Assembly of Epstein-Barr virus capsid in promyelocytic leukemia nuclear bodies. J Virol 89(17):8922–8931. doi:JVI.01114-15

Wang'ondu R, Teal S, Park R, Heston L, Delecluse H, Miller G (2015) DNA damage signaling is induced in the absence of Epstein-Barr Virus (EBV) lytic DNA replication and in response to expression of ZEBRA. PLoS ONE 10(5), e0126088. doi:10.1371/journal.pone.0126088

Wasil LR, Wei L, Chang C, Lan L, Shair KH (2015) Regulation of DNA damage signaling and cell death responses by Epstein-Barr virus Latent Membrane Protein 1 (LMP1) and LMP2A in nasopharyngeal carcinoma cells. J Virol 89(15):7612–7624. doi:10.1128/JVI.00958-15

Weitzman MD, Lilley CE, Chaurushiya MS (2010) Genomes in conflict: maintaining genome integrity during virus infection. Annu Rev Microbiol 64:61–81. doi:10.1146/annurev.micro.112408.134016

Westphal EM, Blackstock W, Feng W, Israel B, Kenney SC (2000) Activation of lytic Epstein-Barr virus (EBV) infection by radiation and sodium butyrate in vitro and in vivo: a potential method for treating EBV-positive malignancies. Cancer Res 60(20):5781–5788

Wilkinson DE, Weller SK (2003) The role of DNA recombination in herpes simplex virus DNA replication. IUBMB Life 55(8):451–458. doi:10.1080/15216540310001612237

Wilkinson DE, Weller SK (2004) Recruitment of cellular recombination and repair proteins to sites of herpes simplex virus type 1 DNA replication is dependent on the composition of viral proteins within prereplicative sites and correlates with the induction of the DNA damage response. J Virol 78(9):4783–4796

Wimmer P, Schreiner S, Everett RD, Sirma H, Groitl P, Dobner T (2010) SUMO modification of E1B-55K oncoprotein regulates isoform-specific binding to the tumour suppressor protein PML. Oncogene 29(40):5511–5522. doi:10.1038/onc.2010.284

Wimmer P, Berscheminski J, Blanchette P, Groitl P, Branton PE, Hay RT, Dobner T, Schreiner S (2015) PML isoforms IV and V contribute to adenovirus-mediated oncogenic transformation by functionally inhibiting the tumor-suppressor p53. Oncogene. doi:10.1038/onc.2015.63

Woodhall DL, Groves IJ, Reeves MB, Wilkinson G, Sinclair JH (2006) Human Daxx-mediated repression of human cytomegalovirus gene expression correlates with a repressive chromatin structure around the major immediate early promoter. J Biol Chem 281(49):37652–37660. doi: M604273200

Wu FY, Ahn JH, Alcendor DJ, Jang WJ, Xiao J, Hayward SD, Hayward GS (2001) Origin-independent assembly of Kaposi's sarcoma-associated herpesvirus DNA replication compartments in transient cotransfection assays and association with the ORF-K8 protein and cellular PML. J Virol 75(3):1487–1506

Xiao Y, Chen J, Liao Q, Wu Y, Peng C, Chen X (2013) Lytic infection of Kaposi's sarcoma-associated herpesvirus induces DNA double-strand breaks and impairs non-homologous end joining. J Gen Virol 94(Pt 8):1870–1875. doi:10.1099/vir.0.053033-0

Xu Y, Ahn JH, Cheng M, Aprhys CM, Chiou CJ, Zong J, Matunis MJ, Hayward GS (2001) Proteasome-independent disruption of PML oncogenic domains (PODs), but not covalent modification by SUMO-1, is required for human cytomegalovirus immediate-early protein IE1 to inhibit PML-mediated transcriptional repression. J Virol 75(22):10683–10695

Xue Y, Toh SY, He P, Lim T, Lim D, Pang CL, Abastado JP, Thierry F (2015) HPV16-E2 induces prophase arrest and activates the cellular DNA damage response in vitro and in precursor lesions of cervical carcinoma. Oncotarget 6(33):34979–34991. doi:10.18632/oncotarget.5512

Yamauchi Y, Daikoku T, Goshima F, Nishiyama Y (2003) Herpes simplex virus UL14 protein blocks apoptosis. Microbiol Immunol 47(9):685–689

Yang J, Deng W, Hau PM, Liu J, Lau VM, Cheung AL, Huen MS, Tsao SW (2015) Epstein-Barr virus BZLF1 protein impairs accumulation of host DNA damage proteins at damage sites in response to DNA damage. Lab Invest 95(8):937–950. doi:10.1038/labinvest.2015.69

Zhao X, Madden-Fuentes RJ, Lou BX, Pipas JM, Gerhardt J, Rigell CJ, Fanning E (2008) Ataxia telangiectasia-mutated damage-signaling kinase- and proteasome-dependent destruction of Mre11-Rad50-Nbs1 subunits in Simian virus 40-infected primate cells. J Virol 82(11): 5316–5328. doi:10.1128/JVI.02677-07

Part VI
Macromolecular Dynamics Within the Nucleus

Energy-Dependent Intranuclear Movements: Role of Nuclear Actin and Myosins

Guillaume Huet and Maria K. Vartiainen

Abstract The actin cytoskeleton is the major driving force for motility of the cell and its constituents. The mechanisms of actin-based motility in the cytoplasm are well characterized, and actin produces the force harnessed for movement either through polymerization of actin monomers to polar filaments or together with motor proteins myosins. Actin, as well as several myosins, are also present in the cell nucleus, and have been linked to gene expression processes. Nuclear actin therefore harbours both the required connections and biochemical properties to power motility events also in this cellular compartment. Indeed, recent functional data have demonstrated a role for actin and myosin in energy-dependent movements of chromosomes and chromosomal loci. In this chapter, we discuss the current knowledge on the organization of nuclear actin, its connection to nuclear gene expression machineries and the present evidence for actin-based motility in the nucleus, which has the potential to play a critical role in the functional organization of the nucleus itself.

Keywords Nuclear actin • Myosins • Intranuclear movement • ATP-dependent • Nuclear organization

1 Introduction

The first step of gene expression, transcription, takes place in the 3D environment of the nucleus, which is non-randomly organized at the level of both chromatin and its associated processes, such as transcription and replication. In addition, nuclear bodies, the membraneless protein assemblies, create nuclear neighborhoods, which may influence gene expression by concentrating different transcription and RNA processing factors. The extent of chromatin dynamics in interphase cells is still not completely clear. Microscopy studies have suggested that chromatin undergoes only limited mobility in the form of constrained diffusion (Chubb et al. 2002) or large-scale coupled motions for several seconds (Zidovska et al. 2013). There are

G. Huet • M.K. Vartiainen (✉)
Institute of Biotechnology, University of Helsinki, Viikinkaari 9, 00014 Helsinki, Finland
e-mail: Maria.Vartiainen@helsinki.fi

D.P. Bazett-Jones, G. Dellaire (eds.), *The Functional Nucleus*,
DOI 10.1007/978-3-319-38882-3_14

also numerous reports of individual chromosomal site movements from one nuclear neighborhood to another (reviewed in Bridger 2011). In addition, genomic methods (such as chromosome conformation capture (3C) and its variants) are capturing interactions between chromosomal loci and even different chromosomes (Dostie and Bickmore 2012). Also whole chromosomes have been shown to change their nuclear localization after growth factor stimulation (Mehta et al. 2010). "Does a gene's location in the nucleus matter" was posed as one of the remaining mysteries in the cell in 2011 (Pennisi 2011). Experiments addressing the functional significance of gene positioning have yielded variable results, and the current view seems to be that gene positioning can influence, but not determine, gene activity (Hubner et al. 2013). Nevertheless, in the context of the whole organism, and for example during development of multicellular organisms, even small changes may matter. Also the mechanism underlying chromatin movements are still somewhat obscure. Genome organization is thought to be a self-organizing process, where the nuclear processes, such as transcription, determine the formation of chromatin domains, which in turn reinforces both the process and the structural features of the genome (Cavalli and Misteli 2013). Stochastic diffusion has long been seen as the default displacement mode in the nucleus (Austin and Bellini 2010). Yet, some observed movements are long-range (over several microns), fast (1 μm/min) and appear highly directional. These findings imply that active mechanisms may operate in the nucleus to power these movements. In the cytoplasm, many motility events are known to be driven by the dynamic actin cytoskeleton. During the past decade it has become apparent that actin, together with many of its accessory factors, is also present in the nucleus, and linked to many gene expression processes. In this chapter we discuss the current evidence for actin-based motility in the nucleus.

1.1 Actin and Myosin-Based Motility in the Cytoplasm

Actin is a 42 kDa globular protein that is the most abundant protein in the cell. Actin binds ATP or ADP, and one of its key properties is its ability to transition from actin monomers (G-actin) to a polymerized state known as filamentous actin (F-actin); filaments that exhibit a dynamic and polarized structure with a growing (plus end or barbed end) and a depolymerizing ends (minus or pointed end). In a process called "treadmilling", ATP-bound actin monomers are added to the filament at the barbed (plus) end, the ATP is hydrolyzed, the terminal phosphate leaves the filament and ADP-bound monomers are depolymerized from the pointed (minus) ends. The ability of actin to form filaments with rapid turnover rates is critical for actin-based movement and for remodeling the shape of the plasma membrane (Dominguez and Holmes 2011). In cells, numerous actin-binding proteins (ABPs) regulate actin filament formation in space and time, and arrange actin into higher order structures, such as bundles and networks (reviewed in Siripala and Welch 2007).

The first step in cell movement is the formation of plasma membrane protrusion through actin polymerization in the direction of locomotion (Fig. 1). Spontaneous nucleation of actin filaments is slow, and therefore the cell contains several proteins that aid in this process. At the leading edge of a motile cell, Arp2/3 complex nucleates new filaments from the side of existing filaments, creating a branched actin filament array that pushes the membrane forward. ABPs are under tight regulation of cellular signaling cascades, and for example WASp (Wiscott-Aldrich syndrome protein) family of proteins activate Arp2/3 complex in response to Rho GTPase signaling at the leading edge. Actin polymerization is promoted by the small actin-monomer binding protein profilin, which facilitates filament elongation at the barbed end. Profilin also participates in actin polymerization by another actin-nucleator family, formins, which nucleate straight actin filaments. Equally important to polymerization is the ability to limit filament growth and depolymerization of old filaments to allow recycling of the actin monomers. Capping protein caps the barbed ends to prevent further growth, and cofilin severs and depolymerizes actin filaments. In addition to these core proteins, numerous other actin regulators fine-tune the process, and make it spatially and temporally coordinated with other cellular events (reviewed in Bugyi and Carlier 2010).

Actin can also produce force and motility in conjunction with the motor proteins called myosins (Fig. 1), which are ATP-dependent motor proteins. Structurally, myosins contain a head, neck and tail domains. The myosin head, which binds actin filaments, changes its conformation in an ATP-dependent manner at the level of the neck. In this way, it can change the position of the head relatively to the cargo-bound tail domain. This generates the displacement of the molecular cargo associated with the myosin relative to the actin filament. To date, the myosin superfamily contains no less than eighteen classes. Class II myosins are the so-called conventional myosins, and the other myosins are described as unconventional. The superfamily members differ mainly by the structure and length of their tail domains, as well as the molecular processes by which they generate movement (Hartman and Spudich 2012). Almost all myosins move towards the barbed end of the actin filaments, with an exception of the myosin VI, which moves towards the pointed end (Wells et al. 1999).

The best characterized example of actin-myosin driven movement is muscle contraction. During contraction of muscle, overlapping actin and myosin II filaments slide past each other within the sarcomeres as a result of ATP-dependent movement of myosin heads on the actin filaments (Batters et al. 2014). Actin-myosin motors operate also in non-muscle cells. For example, myosin creates contraction between actin filaments to pull up the rear of the cell as it moves forward during migration (Cramer 2013). Cytokinesis, the final separation of daughter cells during cell division, is realized by the formation, followed by the contraction, of a ring of actin and myosin II (Pollard 2010). Transport of vesicles and some organelles is mediated by myosins tracking along actin filaments. For example, myosin V transports vesicles originating from the endoplasmic reticulum (Tabb et al. 1998), whereas myosin VI has been linked to transport of endocytic vesicles on actin filaments (Buss et al. 2001).

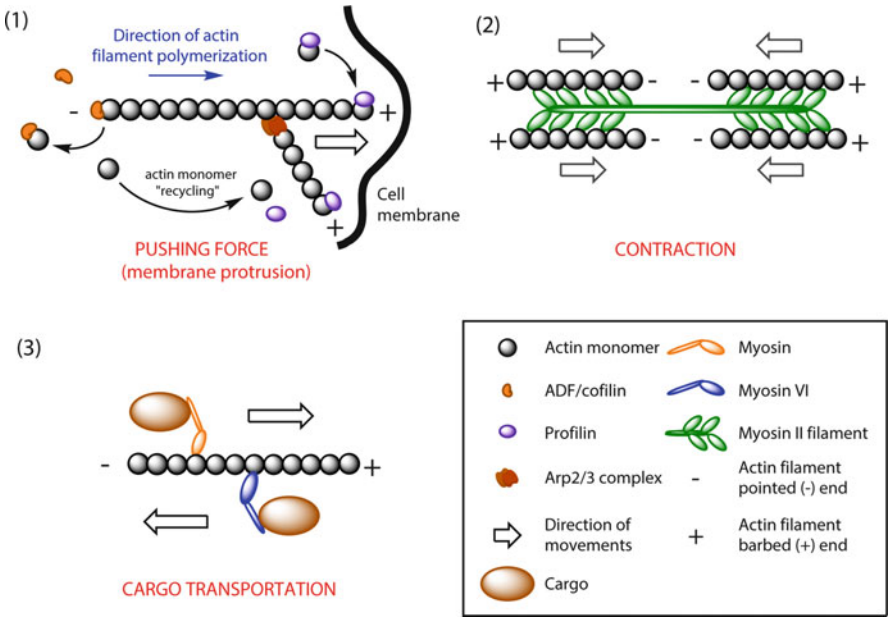

Fig. 1 Mechanisms of actin and myosin-generated movements. Actin can create cellular movement mainly by three mechanisms. (1) Actin polymerization, an oriented energy-dependent process, can push on cell structures, such as the cell membranes. (2) Interconnected actin and myosin II filaments can slide relatively to each other. This mechanism is found in muscle contraction and cytokinesis. Finally, (3) actin can be used as a path for cargo transportation by myosins, in both directions on the filament. Myosin VI is the only known class of myosins able to transport cargo towards the pointed (−) end of the actin filaments

2 Evidence Supporting the Existence of Nuclear Actin and Myosins

Considered as a contaminant for decades, mounting evidence has elevated nuclear actin from an artefact to an essential nuclear protein. The amount of actin in the nucleus appears to be tightly regulated. This was realized already in 2003, when the Görlich-lab identified Exportin-6 as the nuclear export receptor for actin, and suggested that active export was critical to prevent actin polymerization within the nucleus (Stuven et al. 2003). Subsequent studies have demonstrated that actin constantly and rapidly shuttles in and out of the nucleus. Actin monomer concentration limits the transport rate in both directions, and the export-competent actin monomers seem to determine the nuclear steady-state actin levels. Active nuclear import is required for the maintenance of nuclear actin levels, and is mediated by Importin-9 and a small ABP, actin depolymerizing factor (ADF)/cofilin. Significantly, this machinery is required for the maximal transcription activity of the cell (Dopie et al. 2012; Skarp et al. 2013). Altered nuclear actin levels have been reported in different cellular states. At steady state, the nuclear actin levels seem

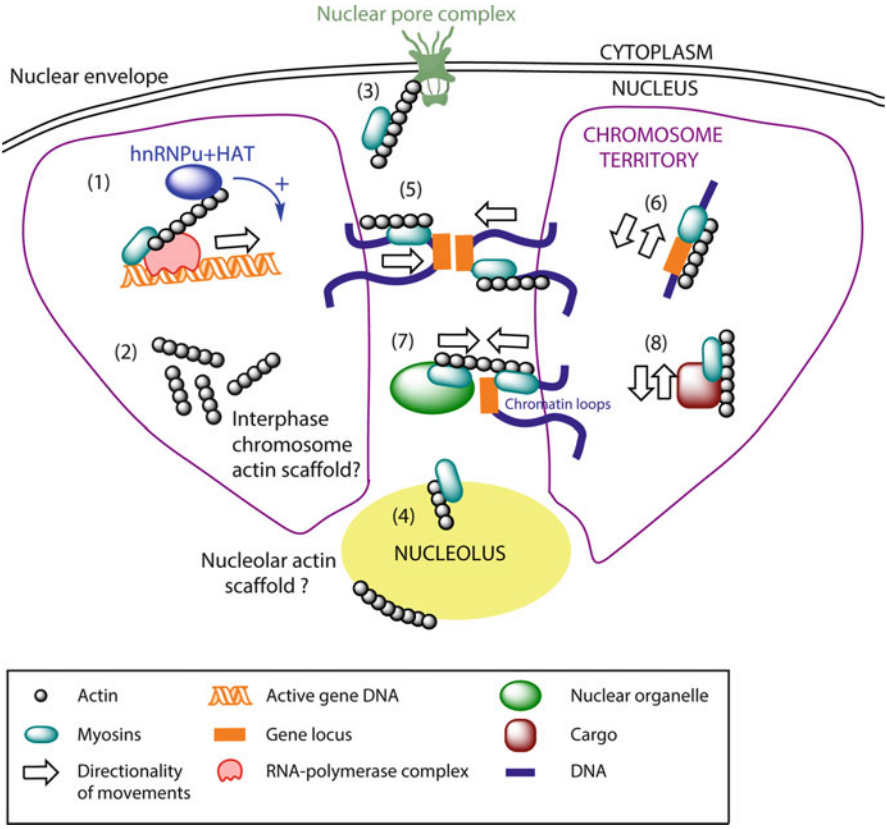

Fig. 2 Summary of putative roles of actin-myosin couples in movements inside the cell nucleus. Actin and myosins could elicit and regulate movements in the nucleus at different levels. (1) Polymerization of actin during transcription may help the transcribing RNA polymerase complex to move forward. First, actin polymers may serve as a recruitment platform for chromatin remodelers such as histone acetyl-transferases (HAT) that would open the chromatin in front of the transcribing complex. Myosins may also aid in the process of pushing forward the polymerase complex. At a wider level, a nuclear actin meshwork could be used to maintain chromosome territories intact (2) or to emanate from the nuclear pore as pore-linked filaments (3) or support the nucleolus (4). Actin, together with myosin, could then drive the co-localization of gene loci from different chromosomal territories (5) or mediate movement of gene loci within a same nuclear territory (6). Actin and myosin could also contribute to looping of gene loci out of the territory, and its recruitment to different nuclear bodies such as nuclear speckles or Cajal bodies (7). Also other cargo such as large protein complexes may be carried by myosins along actin polymers (8)

to vary from cell to cell (Skarp et al. 2013). For example, a decrease in nuclear actin has been linked to quiescence (Spencer et al. 2011), and increased nuclear actin has been observed upon both differentiation (Xu et al. 2010) and in response to certain stresses such as heat shock, DMSO treatment or ATP depletion (Fukui and Katsumaru 1979; Pendleton et al. 2003).

As mentioned, in the cytoplasm actin filaments provide a great part of the highly dynamic cytoskeleton, necessary for maintaining and remodeling the cell shape. It is therefore conceivable that nuclear actin could be polymeric and participate in movement events within the nucleus. Although there is abundant evidence for the existence of nuclear actin polymers, the exact form of nuclear actin is still somewhat unclear (reviewed in Grosse and Vartiainen 2013). The existence of a different conformation of nuclear actin has been hypothesized (Gonsior et al. 1999), and the theory consolidated at the time by the nuclear absence of phalloidin staining; the traditional method to visualize cytoplasmic actin filaments. Blocking nuclear actin export by Exportin-6 depletion accumulates actin in the nucleus, and causes the formation of phalloidin-stainable nuclear actin bars (Stuven et al. 2003), suggesting that the usually low amount of nuclear actin might, at least in part, explain the lack of phalloidin staining in the nuclei of most cells. One exception is the giant nuclei of *Xenopus* oocytes, which have a high abundance of nuclear actin that forms a network that can be stained by phalloidin due to posttranscriptional silencing of Exportin-6 (Bohnsack et al. 2006). In the same system, a network of actin-containing filaments, visible by electron microscopy, was reported to connect nuclear organelles with nuclear pores (Kiseleva et al. 2004). Concordantly, an F-actin scaffold was found to maintain ribonucleoproteic structures, nucleoli and histone locus bodies, in place against gravity (Feric and Brangwynne 2013). Despite the absence of phalloidin staining in most cell nuclei, fluorescence recovery after photobleaching (FRAP) experiments have hinted that a similar type of actin network might be present in other cells as well. FRAP profiles from HeLa cell nuclei revealed two populations of nuclear actin, with the slower recovering fraction responding to actin depolymerizing drugs, and thus likely representing actin in some type of polymeric form (McDonald et al. 2006). A subsequent FRAP study concluded that there are actually at least three pools of nuclear actin: monomeric actin, a polymeric form, and a third less mobile fraction that may correspond to actin stably bound to nuclear complexes (Dopie et al. 2012).

The recent development of new fluorescent probes based on actin-binding peptides has finally provided the first glimpses of nuclear actin polymers in "normal" cells. A shortened and nuclear localized version of the Utrophin calponin homology domain-probe, Utr230-EN, demonstrated submicron length structures that were excluded from chromatin-rich regions (Belin et al. 2013). Nuclear localization of the LifeAct probe reported a transient formation of nuclear actin bundles after serum stimulation, which is known to trigger the polymerization of actin via the activation of Rho GTPases. These same structures could be stained with phalloidin, and were linked to activation of the actin-responsive transcriptional coregulator megakaryoblastic leukemia 1 (MKL1) (Baarlink et al. 2013). Other indirect evidence supporting the existence of a functional nuclear actin network comes from the study of nuclei from different species. For example, disruption of actin filaments with actin binding drugs can impair nuclear export of RNAs in Neurospora *crassa* (Turian and Barja 1995) and leads to a disorganization of chromosome territories in human MCF7 cells (Ondrej et al. 2008), pointing towards

the existence of a network of nuclear polymeric actin essential for the proper chromosome organization.

The best described nuclear myosin is nuclear myosin I (NM1), an isoform of Myosin 1c with a modified N-terminal tail that localizes predominantly in the nucleus (Pestic-Dragovich et al. 2000). Also many other myosins are found in the nucleus, such as the conventional Myosin II (Li and Sarna 2009). Myosin Vb is found in nucleoli (Lindsay and McCaffrey 2009) and Myosin VI in transcription factories (Vreugde et al. 2006). Myosin XVIII, on the other hand, translocates to the muscle cell nucleus after their differentiation (Salamon et al. 2003). Hence, both actin, presumably forming some type of polymeric network, and several myosins are present in the cell nucleus, and could therefore serve as a basis for an active transport within the nucleus.

3 Actin and Myosin in Transcription: Movement on the Gene?

In order to facilitate motility events in the nucleus, actin needs to be connected to chromatin. Indeed, nuclear actin is found in several DNA-binding machineries linked with transcription, such as chromatin remodeling, transcription initiation, RNA polymerase (RNAP) elongation and pre-mRNA processing complexes. Actin also regulates the activity of specific transcription factors. The best known example is MKL1, a key coregulator of serum response factor (SRF), which is regulated by cellular actin monomer levels (Vartiainen et al. 2007).

The initiation of gene transcription usually requires local remodeling of chromatin. Actin is part of several key chromatin remodeling complexes such as the mammalian ATP-dependent BAF (Brahma-related gene 1-associated factor) complex (Rando et al. 2002) and the Ino80 complex in yeast (Kapoor et al. 2013). Actin is also a binding partner for all three RNA polymerase complexes (Hofmann et al. 2004; Hu et al. 2004; Philimonenko et al. 2004). Also NM1 has been linked to both pol I and pol II dependent transcription (Hofmann et al. 2006; Philimonenko et al. 2004; Ye et al. 2008). Actin is also part of the pre-initiation complex of RNA polymerase II (RNAP II) (Hofmann et al. 2004), and is also thought to participate in transcriptional elongation through its interaction with the transcription elongation factor P-TEFb (Qi et al. 2011). Actin also interacts with hnRNP-U (heterogeneous nuclear ribonucleoprotein-U) (Kukalev et al. 2005), and this interaction recruits the histone acetyltransferase PCAF to facilitate transcription elongation (Obrdlik et al. 2008). Also NM1 has been linked to recruitment of PCAF to ribosomal genes as part of the multiprotein complex B-WICH (Percipalle et al. 2006; Sarshad et al. 2013), and this activity seems to be regulated in a cell cycle-dependent manner by glycogen synthase kinase 3β (Sarshad and Percipalle 2014). Of interest, the mature ribosomal RNAs (rRNAs) associate both with nuclear pore complexes and

NM1, suggesting that NM1 could play a role both in rRNA transcription and their subsequent nuclear export (Obrdlik et al. 2010).

Interestingly, also many ABPs that regulate actin polymerization have been linked to RNA polymerase function. Actin filament disassembling protein cofilin-1 (Obrdlik and Percipalle 2011), and actin filament nucleator Arp2/3 complex (Yoo et al. 2007) together with its activator WASp (Wu et al. 2006) are all present in the RNAP II complex and linked to its activity. Another activator of Arp2/3 complex, Wave-1, is required for transcription reprograming in *Xenopus* oocytes (Miyamoto et al. 2013), which also requires actin polymerization regulated by the actin signaling protein Toca-1 (Miyamoto et al. 2011). Other myosins besides NM1 have also been linked to transcription. Nuclear myosin II contributes to RNAP II initiation complex assembly on the *Icam-1* gene (Li and Sarna 2009), while myosin VI binds to RNAP II and intragenic, but not intergenic, chromatin regions (Vreugde et al. 2006). Myosin Vb may also play a role in RNAP I transcription, as it physically interacts with its transcription complex (Lindsay and McCaffrey 2009).

Actin, myosin and other ABPs have therefore been linked to many different processes affecting gene expression. However, the exact mechanism by which these proteins co-operate for efficient transcription is still not completely clear. Nevertheless, it has been speculated that dynamic actin polymerization occurs during transcription. Different forms of actin could play a role during distinct steps of transcription (Percipalle 2013), and/or actin-myosin motors could be involved in facilitating RNAP II sliding on DNA (de Lanerolle et al. 2005). Regardless of the mechanism, it is obvious that actin is connected to numerous nuclear machineries, some of which will likely play a critical role in the actin-based chromatin movements.

4 Long Range Movements in the Nucleus Using Actin and Myosins

Engineered gene loci have played a significant part in studies and especially imaging of long-range chromosomal movements. The Belmont-lab has demonstrated that tethering of the transcription activation domain of VP16 to a peripheral chromosomal site leads to its repositioning towards the nuclear interior (Tumbar and Belmont 2001). An inducible version of this system showed that this repositioning took 1–2 h. The movement occurred in several periods, at velocities of 0.1–0.9 μm/min, cut by relatively long periods of immobility. The path of these movements, with distances of 1–5 μm, was described as curvilinear, approximately perpendicular to the nuclear envelope. Importantly, expression of either an actin mutant that does not polymerize, or a myosin mutant with presumably low actin-binding and treatment with myosin inhibitor butanedione monoxime (BDM), blocked or delayed the chromosomal site redistribution. Moreover, an actin mutant that favors actin polymerization accelerated the redistribution (Chuang et al. 2006).

These data support the active movement of chromosomes along filamentous structures. A later work from the same laboratory visualized the live movement of a Heat shock protein 70 (Hsp70) transgene towards nuclear speckles after heat shock. Again, the pattern followed by the transgene was curvilinear, with distances of 0.5–6 μm, and velocities of 1–2-μm/min. The nuclear speckle association was abolished either with drugs that prevented actin dynamics, or by expressing an actin mutant that does not polymerize. Moreover, chromatin was observed to stretch in the direction of the movement, further supporting the existence of a force-generating mechanism. Significantly, transcription of the transgene was almost exclusively observed only after the loci had come in contact with the nuclear speckle (Khanna et al. 2014). This demonstrates that the directed, actin-dependent motion of the HSP70 transgene has functional consequences on transcriptional activation, and thus likely playing a role in optimizing the cellular response to heat shock.

Actin may also play a wider role in directing motility events towards different nuclear bodies (Fig. 2). For example, rapid, directed and long-range movement of U2 small nuclear RNA (snRNA) genes towards relatively immobile Cajal bodies was observed by live-cell imaging by Spector and colleagues (Dundr et al. 2007). Once more, the repositioning was perturbed by overexpression of a non-polymerizing actin mutant (Dundr et al. 2007); a trend confirmed by additional studies. Actin is associated with nuclear speckles, and reorganization of these structures after transcription inhibition appeared to be sensitive to actin depolymerizing drugs (Wang et al. 2006). Actin also co-localizes with Pknox1, which is required for *Hox* gene expression, and is found together with elongating pol II and THOC in nuclear speckles (Naum-Ongania et al. 2013). Moreover, induction of *Hox* gene expression seems to require actin polymerization (Ferrai et al. 2009). In the same manner, two estrogen receptor target genes, situated in two different chromosomes, were shown to co-localize in an actin-dependent manner at nuclear speckles upon transcriptional activation in response to estrogen (Hu et al. 2008); although the co-localization of the genes in this case has been subsequently questioned (Kocanova et al. 2010). Interestingly, the actin-based movement may not always be directed towards nuclear bodies. Nuclear actin polymerization has been linked to displacement of HIV-1 proviruses from promyelocytic leukemia nuclear bodies (PML NBs). Association of the provirus with the PML NB inhibits HIV-1 gene expression, and coincides with repressive histone marks. Actin-dependent movements of the provirus therefore contribute to HIV-1 latency (Lusic et al. 2013). Thus, actin and myosins are clearly implicated in linking gene loci to different nuclear neighborhoods upon transcriptional changes.

At an even wider level, the arrangement and movements of whole chromosomes could rely on actin and myosin. In human primary fibroblasts a reorganization of chromosome territories was observed after serum starvation leading to quiescence (Mehta et al. 2010). This repositioning is rapid, occurring in less than 15 min, and dependent on both actin and NM1. Curiously, the return to the non-quiescent chromosome organization and myosin localization was a much longer process, taking more than 24 h (Mehta et al. 2010). One explanation could be that the

starvation provokes a rapid collapse of a complex actin-myosin structure supporting the chromosomes, and that its re-establishment is a slow process. The localization of chromosome territories may therefore be dependent on a nuclear actin network, as mentioned also above (Ondrej et al. 2008). Actin-based nuclear movements of chromosomes are not limited to interphase. During meiotic prophase in yeast, nuclear actin is involved in relocation of the synaptonemal complex to the nuclear periphery (Scherthan et al. 2007). A related observation was made in plant cells. Live cell imaging of meiotic prophase in maize allowed the observation of actin and microtubule-dependent fast chromosome movements towards the nuclear envelope (Sheehan and Pawlowski 2009). This process seems conserved, as it was also observed in yeast (Sonntag Brown et al. 2011). Altogether, actin-myosin complexes appear as keystones of the dynamic reorganization of chromatin, especially in response to significant changes in the gene expression status.

5 Summary

The existence of polymeric nuclear actin has been finally established, and equally clear is the presence of numerous ABPs and myosins in the nucleus. These same proteins also have extensive links to many gene expression machineries that connect them to chromatin. Hence, the nucleus harbors chromatin-associated machinery capable of eliciting active movement. Interphase chromatin dynamics, and its implications to nuclear functions, is one of the key unsolved aspects in modern cell biology. Visualization of chromatin dynamics has clearly been lagging behind, for instance, protein dynamics studies. Live-imaging of individual chromosomal loci and larger chromatin domains is absolutely essential to describe and quantitate the type of movements, and to separate active, energy-dependent movements from diffusion, which are likely to coexist in the nucleus. Development of new tools and technologies for imaging chromatin dynamics, perhaps based on the CRISPR/Cas system (Chen et al. 2013), will be vital for these studies.

The few studies so far, which have directly imaged and detected long-range movements of chromosomal loci (Chuang et al. 2006; Dundr et al. 2007; Khanna et al. 2014), have all implicated actin in the process. This clearly highlights the need for more mechanistic studies on nuclear actin and its binding partners. One key aspect will be to determine how actin is connected to the moving chromatin by identifying direct binding partners for nuclear actin. Equally important are studies aimed at gaining further insights into the polymeric forms of nuclear actin. Actin polymers are absolutely necessary for both polymerization-driven and myosin-based motility. However, the filamentous actin structures recognized by at least the Utr230-EN probe seem insufficient for long-range nuclear transport. It is therefore possible that actin could contribute to chromatin dynamics as part of the viscoelastic structure of the nucleoplasm (Belin et al. 2013), or as short filaments that are used as “landing pads” during intranuclear movement. In this context, it is intriguing to note that Myosin 1c, the best characterized nuclear myosin to date, and

the one that has been linked to intranuclear movements together with actin, can be classified as slow "strain sensor" myosin. This indicates that it would be better suited for tethering/anchoring of heavy cargo on actin filaments than to long-range processive movements (Bond et al. 2013). Clearly further studies are needed to clarify these issues, and to fully realize the potential of actin and its binding partners in functional organization of the nucleus.

References

Austin CM, Bellini M (2010) The dynamic landscape of the cell nucleus. Mol Reprod Dev 77:19–28

Baarlink C, Wang H, Grosse R (2013) Nuclear actin network assembly by formins regulates the SRF coactivator MAL. Science 340:864–867

Batters C, Veigel C, Homsher E, Sellers JR (2014) To understand muscle you must take it apart. Front Physiol 5:90

Belin BJ, Cimini BA, Blackburn EH, Mullins RD (2013) Visualization of actin filaments and monomers in somatic cell nuclei. Mol Biol Cell 24:982–994

Bohnsack MT, Stuven T, Kuhn C, Cordes VC, Gorlich D (2006) A selective block of nuclear actin export stabilizes the giant nuclei of Xenopus oocytes. Nat Cell Biol 8:257–263

Bond LM, Brandstaetter H, Kendrick-Jones J, Buss F (2013) Functional roles for myosin 1c in cellular signaling pathways. Cell Signal 25:229–235

Bridger JM (2011) Chromobility: the rapid movement of chromosomes in interphase nuclei. Biochem Soc Trans 39:1747–1751

Bugyi B, Carlier MF (2010) Control of actin filament treadmilling in cell motility. Annu Rev Biophys 39:449–470

Buss F, Arden SD, Lindsay M, Luzio JP, Kendrick-Jones J (2001) Myosin VI isoform localized to clathrin-coated vesicles with a role in clathrin-mediated endocytosis. EMBO J 20:3676–3684

Cavalli G, Misteli T (2013) Functional implications of genome topology. Nat Struct Mol Biol 20:290–299

Chen C, Fenk LA, de Bono M (2013) Efficient genome editing in Caenorhabditis elegans by CRISPR-targeted homologous recombination. Nucleic Acids Res 41, e193

Chuang CH, Carpenter AE, Fuchsova B, Johnson T, de Lanerolle P, Belmont AS (2006) Long-range directional movement of an interphase chromosome site. Curr Biol 16:825–831

Chubb JR, Boyle S, Perry P, Bickmore WA (2002) Chromatin motion is constrained by association with nuclear compartments in human cells. Curr Biol 12:439–445

Cramer LP (2013) Mechanism of cell rear retraction in migrating cells. Curr Opin Cell Biol 25:591–599

de Lanerolle P, Johnson T, Hofmann WA (2005) Actin and myosin I in the nucleus: what next? Nat Struct Mol Biol 12:742–746

Dominguez R, Holmes KC (2011) Actin structure and function. Annu Rev Biophys 40:169–186

Dopie J, Skarp KP, Rajakyla EK, Tanhuanpaa K, Vartiainen MK (2012) Active maintenance of nuclear actin by importin 9 supports transcription. Proc Natl Acad Sci U S A 109:E544–552

Dostie J, Bickmore WA (2012) Chromosome organization in the nucleus—charting new territory across the Hi-Cs. Curr Opin Genet Dev 22:125–131

Dundr M, Ospina JK, Sung MH, John S, Upender M, Ried T, Hager GL, Matera AG (2007) Actin-dependent intranuclear repositioning of an active gene locus in vivo. J Cell Biol 179:1095–1103

Feric M, Brangwynne CP (2013) A nuclear F-actin scaffold stabilizes ribonucleoprotein droplets against gravity in large cells. Nat Cell Biol 15:1253–1259

Ferrai C, Naum-Ongania G, Longobardi E, Palazzolo M, Disanza A, Diaz VM, Crippa MP, Scita G, Blasi F (2009) Induction of HoxB transcription by retinoic acid requires actin polymerization. Mol Biol Cell 20:3543–3551

Fukui Y, Katsumaru H (1979) Nuclear actin bundles in Amoeba, Dictyostelium and human HeLa cells induced by dimethyl sulfoxide. Exp Cell Res 120:451–455

Gonsior SM, Platz S, Buchmeier S, Scheer U, Jockusch BM, Hinssen H (1999) Conformational difference between nuclear and cytoplasmic actin as detected by a monoclonal antibody. J Cell Sci 112(Pt 6):797–809

Grosse R, Vartiainen MK (2013) To be or not to be assembled: progressing into nuclear actin filaments. Nat Rev Mol Cell Biol 14:693–697

Hartman MA, Spudich JA (2012) The myosin superfamily at a glance. J Cell Sci 125:1627–1632

Hofmann WA, Stojiljkovic L, Fuchsova B, Vargas GM, Mavrommatis E, Philimonenko V, Kysela K, Goodrich JA, Lessard JL, Hope TJ, Hozak P, de Lanerolle P (2004) Actin is part of pre-initiation complexes and is necessary for transcription by RNA polymerase II. Nat Cell Biol 6:1094–1101

Hofmann WA, Vargas GM, Ramchandran R, Stojiljkovic L, Goodrich JA, de Lanerolle P (2006) Nuclear myosin I is necessary for the formation of the first phosphodiester bond during transcription initiation by RNA polymerase II. J Cell Biochem 99:1001–1009

Hu P, Wu S, Hernandez N (2004) A role for beta-actin in RNA polymerase III transcription. Genes Dev 18:3010–3015

Hu Q, Kwon YS, Nunez E, Cardamone MD, Hutt KR, Ohgi KA, Garcia-Bassets I, Rose DW, Glass CK, Rosenfeld MG, Fu XD (2008) Enhancing nuclear receptor-induced transcription requires nuclear motor and LSD1-dependent gene networking in interchromatin granules. Proc Natl Acad Sci U S A 105:19199–19204

Hubner MR, Eckersley-Maslin MA, Spector DL (2013) Chromatin organization and transcriptional regulation. Curr Opin Genet Dev 23:89–95

Kapoor P, Chen M, Winkler DD, Luger K, Shen X (2013) Evidence for monomeric actin function in INO80 chromatin remodeling. Nat Struct Mol Biol 20:426–432

Khanna N, Hu Y, Belmont AS (2014) HSP70 transgene directed motion to nuclear speckles facilitates heat shock activation. Curr Biol 24:1138–1144

Kiseleva E, Drummond SP, Goldberg MW, Rutherford SA, Allen TD, Wilson KL (2004) Actin- and protein-4.1-containing filaments link nuclear pore complexes to subnuclear organelles in Xenopus oocyte nuclei. J Cell Sci 117:2481–2490

Kocanova S, Kerr EA, Rafique S, Boyle S, Katz E, Caze-Subra S, Bickmore WA, Bystricky K (2010) Activation of estrogen-responsive genes does not require their nuclear co-localization. PLoS Genet 6, e1000922

Kukalev A, Nord Y, Palmberg C, Bergman T, Percipalle P (2005) Actin and hnRNP U cooperate for productive transcription by RNA polymerase II. Nat Struct Mol Biol 12:238–244

Li Q, Sarna SK (2009) Nuclear myosin II regulates the assembly of preinitiation complex for ICAM-1 gene transcription. Gastroenterology 137:1051–1060, 1060 e1051–1053

Lindsay AJ, McCaffrey MW (2009) Myosin Vb localises to nucleoli and associates with the RNA polymerase I transcription complex. Cell Motil Cytoskeleton 66:1057–1072

Lusic M, Marini B, Ali H, Lucic B, Luzzati R, Giacca M (2013) Proximity to PML nuclear bodies regulates HIV-1 latency in CD4+ T cells. Cell Host Microbe 13:665–677

McDonald D, Carrero G, Andrin C, de Vries G, Hendzel MJ (2006) Nucleoplasmic beta-actin exists in a dynamic equilibrium between low-mobility polymeric species and rapidly diffusing populations. J Cell Biol 172:541–552

Mehta IS, Amira M, Harvey AJ, Bridger JM (2010) Rapid chromosome territory relocation by nuclear motor activity in response to serum removal in primary human fibroblasts. Genome Biol 11:R5

Miyamoto K, Pasque V, Jullien J, Gurdon JB (2011) Nuclear actin polymerization is required for transcriptional reprogramming of Oct4 by oocytes. Genes Dev 25:946–958

Miyamoto K, Teperek M, Yusa K, Allen GE, Bradshaw CR, Gurdon JB (2013) Nuclear Wave1 is required for reprogramming transcription in oocytes and for normal development. Science 341:1002–1005

Naum-Ongania G, Diaz VM, Blasi F, Rivera-Pomar R (2013) Nuclear actin polymerization from faster growing ends in the initial activation of Hox gene transcription are nuclear speckles involved? Transcription 4:260–272

Obrdlik A, Kukalev A, Louvet E, Farrants AK, Caputo L, Percipalle P (2008) The histone acetyltransferase PCAF associates with actin and hnRNP U for RNA polymerase II transcription. Mol Cell Biol 28:6342–6357

Obrdlik A, Louvet E, Kukalev A, Naschekin D, Kiseleva E, Fahrenkrog B, Percipalle P (2010) Nuclear myosin 1 is in complex with mature rRNA transcripts and associates with the nuclear pore basket. FASEB J: official publication of the Federation of American Societies for Experimental Biology 24:146–157

Obrdlik A, Percipalle P (2011) The F-actin severing protein cofilin-1 is required for RNA polymerase II transcription elongation. Nucleus 2:72–79

Ondrej V, Lukasova E, Krejci J, Matula P, Kozubek S (2008) Lamin A/C and polymeric actin in genome organization. Mol Cells 26:356–361

Pendleton A, Pope B, Weeds A, Koffer A (2003) Latrunculin B or ATP depletion induces cofilin-dependent translocation of actin into nuclei of mast cells. J Biol Chem 278:14394–14400

Pennisi E (2011) Mysteries of the cell. Does a gene's location in the nucleus matter? Science 334:1050–1051

Percipalle P (2013) Co-transcriptional nuclear actin dynamics. Nucleus 4:43–52

Percipalle P, Fomproix N, Cavellan E, Voit R, Reimer G, Kruger T, Thyberg J, Scheer U, Grummt I, Farrants AK (2006) The chromatin remodelling complex WSTF-SNF2h interacts with nuclear myosin 1 and has a role in RNA polymerase I transcription. EMBO Rep 7:525–530

Pestic-Dragovich L, Stojiljkovic L, Philimonenko AA, Nowak G, Ke Y, Settlage RE, Shabanowitz J, Hunt DF, Hozak P, de Lanerolle P (2000) A myosin I isoform in the nucleus. Science 290:337–341

Philimonenko VV, Zhao J, Iben S, Dingova H, Kysela K, Kahle M, Zentgraf H, Hofmann WA, de Lanerolle P, Hozak P, Grummt I (2004) Nuclear actin and myosin I are required for RNA polymerase I transcription. Nat Cell Biol 6:1165–1172

Pollard TD (2010) Mechanics of cytokinesis in eukaryotes. Curr Opin Cell Biol 22:50–56

Qi T, Tang W, Wang L, Zhai L, Guo L, Zeng X (2011) G-actin participates in RNA polymerase II-dependent transcription elongation by recruiting positive transcription elongation factor b (P-TEFb). J Biol Chem 286:15171–15181

Rando OJ, Zhao K, Janmey P, Crabtree GR (2002) Phosphatidylinositol-dependent actin filament binding by the SWI/SNF-like BAF chromatin remodeling complex. Proc Natl Acad Sci U S A 99:2824–2829

Salamon M, Millino C, Raffaello A, Mongillo M, Sandri C, Bean C, Negrisolo E, Pallavicini A, Valle G, Zaccolo M, Schiaffino S, Lanfranchi G (2003) Human MYO18B, a novel unconventional myosin heavy chain expressed in striated muscles moves into the myonuclei upon differentiation. J Mol Biol 326:137–149

Sarshad A, Sadeghifar F, Louvet E, Mori R, Bohm S, Al-Muzzaini B, Vintermist A, Fomproix N, Ostlund AK, Percipalle P (2013) Nuclear myosin 1c facilitates the chromatin modifications required to activate rRNA gene transcription and cell cycle progression. PLoS Genet 9, e1003397

Sarshad AA, Percipalle P (2014) New insight into role of myosin motors for activation of RNA polymerases. Int Rev Cell Mol Biol 311:183–230

Scherthan H, Wang H, Adelfalk C, White EJ, Cowan C, Cande WZ, Kaback DB (2007) Chromosome mobility during meiotic prophase in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A 104:16934–16939

Sheehan MJ, Pawlowski WP (2009) Live imaging of rapid chromosome movements in meiotic prophase I in maize. Proc Natl Acad Sci U S A 106:20989–20994
Siripala AD, Welch MD (2007) SnapShot: actin regulators I. Cell 128:626
Skarp KP, Huet G, Vartiainen MK (2013) Steady-state nuclear actin levels are determined by export competent actin pool. Cytoskeleton 70:623–634
Sonntag Brown M, Zanders S, Alani E (2011) Sustained and rapid chromosome movements are critical for chromosome pairing and meiotic progression in budding yeast. Genetics 188:21–32
Spencer VA, Costes S, Inman JL, Xu R, Chen J, Hendzel MJ, Bissell MJ (2011) Depletion of nuclear actin is a key mediator of quiescence in epithelial cells. J Cell Sci 124:123–132
Stuven T, Hartmann E, Gorlich D (2003) Exportin 6: a novel nuclear export receptor that is specific for profilin-actin complexes. EMBO J 22:5928–5940
Tabb JS, Molyneaux BJ, Cohen DL, Kuznetsov SA, Langford GM (1998) Transport of ER vesicles on actin filaments in neurons by myosin V. J Cell Sci 111(Pt 21):3221–3234
Tumbar T, Belmont AS (2001) Interphase movements of a DNA chromosome region modulated by VP16 transcriptional activator. Nat Cell Biol 3:134–139
Turian G, Barja F (1995) Nuclear actin and RNA export in conidial cells of Neurospora crassa. Cell Biol Int 19:77–78
Wang IF, Chang HY, Shen CK (2006) Actin-based modeling of a transcriptionally competent nuclear substructure induced by transcription inhibition. Exp Cell Res 312:3796–3807
Vartiainen MK, Guettler S, Larijani B, Treisman R (2007) Nuclear actin regulates dynamic subcellular localization and activity of the SRF cofactor MAL. Science 316:1749–1752
Wells AL, Lin AW, Chen LQ, Safer D, Cain SM, Hasson T, Carragher BO, Milligan RA, Sweeney HL (1999) Myosin VI is an actin-based motor that moves backwards. Nature 401:505–508
Vreugde S, Ferrai C, Miluzio A, Hauben E, Marchisio PC, Crippa MP, Bussi M, Biffo S (2006) Nuclear myosin VI enhances RNA polymerase II-dependent transcription. Mol Cell 23:749–755
Wu X, Yoo Y, Okuhama NN, Tucker PW, Liu G, Guan JL (2006) Regulation of RNA-polymerase-II-dependent transcription by N-WASP and its nuclear-binding partners. Nat Cell Biol 8:756–763
Xu YZ, Thuraisingam T, Morais DA, Rola-Pleszczynski M, Radzioch D (2010) Nuclear translocation of beta-actin is involved in transcriptional regulation during macrophage differentiation of HL-60 cells. Mol Biol Cell 21:811–820
Ye J, Zhao J, Hoffmann-Rohrer U, Grummt I (2008) Nuclear myosin I acts in concert with polymeric actin to drive RNA polymerase I transcription. Genes Dev 22:322–330
Yoo Y, Wu X, Guan JL (2007) A novel role of the actin-nucleating Arp2/3 complex in the regulation of RNA polymerase II-dependent transcription. J Biol Chem 282:7616–7623
Zidovska A, Weitz DA, Mitchison TJ (2013) Micron-scale coherence in interphase chromatin dynamics. Proc Natl Acad Sci U S A 110:15555–15560

Nucleosome Dynamics Studied by Förster Resonance Energy Transfer

Alexander Gansen and Jörg Langowski

Abstract Chromatin is a hierarchical structure that condenses the genetic material into a nucleus of less than 10 μm in diameter, while at the same time providing rapid on-demand access to specific DNA loci. This dual role requires a tightly regulated yet highly dynamic DNA packaging, with the nucleosome as the central repeating unit. The nucleosome is composed of multiple protein subunits with distinct dynamic properties and can undergo spontaneous conformational transitions.

Förster Resonance Energy Transfer (FRET) is a powerful method to analyze such conformational changes in nucleosomes. In this chapter we will review its concepts and describe different implementations of FRET, with a special emphasis on single molecule techniques. We will conclude with a brief overview of recent experiments, where FRET was successfully used to shed light on the dynamic properties of nucleosomes. These include the unwrapping dynamics in the entry/exit region, the pathway by which they disassemble and the role of posttranslational modifications on nucleosome architecture.

1 Introduction

In chromatin DNA is compacted with a roughly equal mass of protein into a hierarchical structure that fulfills two seemingly opposing purposes. On one hand, it serves as a scaffold to condense the genetic material into a nucleus of less than 10 μm in diameter. On the other hand, it has to provide rapid on-demand access to specific DNA loci when needed while maintaining overall chromatin integrity. This dynamic role requires a tightly regulated yet highly dynamic DNA packaging that can control genetic activity through changes in chromatin morphology.

The central component in this network is the nucleosome, the repeating unit of DNA compaction in eukaryotes (Olins and Olins 1974). The nucleosome core

A. Gansen (✉) • J. Langowski
Division Biophysics of Macromolecules, DKFZ (German Cancer Research Center), Im Neuenheimer Feld 580, 69120 Heidelberg, Germany
e-mail: alexander.gansen@gmail.com; jl@dkfz.de

D.P. Bazett-Jones, G. Dellaire (eds.), *The Functional Nucleus*,
DOI 10.1007/978-3-319-38882-3_15

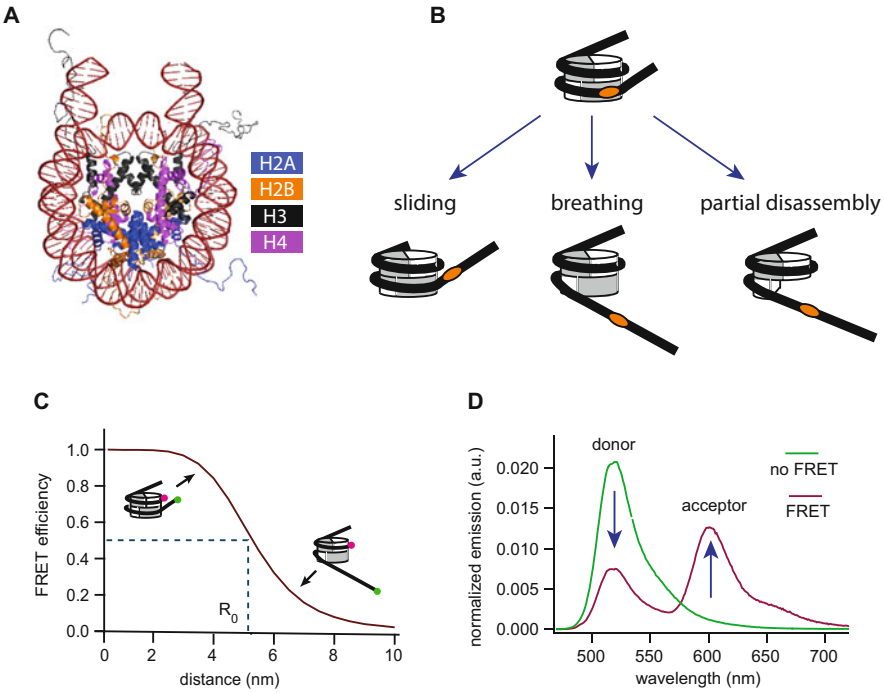

Fig. 1 (**a**) Crystal structure of the nucleosome (PDB entry 1kx5, with added linker DNA). DNA wraps around a protein core of 2 copies each of histones H2A (*blue*), H2B (*orange*), H3 (*black*) and H4 (*purple*). (**b**) Mechanisms to facilitate access to a DNA site that is buried inside the nucleosome (*orange patch*): In sliding, the octamer is permanently repositioned along the DNA by means of ATP-dependent remodelling factors. Spontaneous DNA unwrapping in the entry/exit region can provide transient access to DNA. Removal of one or both dimers prevents DNA rewrapping and renders DNA more accessible. (**c**) FRET efficiency as a function of interdye distance. R_0 is the distance, at which 50% of energy is transferred to the acceptor. When nucleosomal DNA is labeled near the dyad axis, the closed nucleosome structure will show large FRET, since both dyes are in close proximity. Upon DNA breathing, FRET efficiency is decreased when the labeled DNA arm detaches from the octamer. (**d**) FRET-induced change in fluorescence after selective donor excitation. In the absence of FRET only donor emission is observed (*green curve*). FRET between both dyes results in an increase in acceptor signal and reduced donor emission (*red curve*)

particle (NCP) consists of 147 bp DNA wrapped around a wedge-shaped protein octamer in 1.65 superhelical turns. In its canonical form the octamer is composed of two copies each of histone H2A, H2B, H3 and H4 (Fig. 1a). Nucleosomes are assembled in a precisely orchestrated fashion, where first two (H3-H4) dimers are deposited on DNA to organize the central ≈60 bp, followed by addition of two H2A-H2B dimers that bind an additional 30 bp on either side (Andrews and Luger 2011). In vivo, NCPs are connected by short stretches (10–90 bp) of linker DNA to form a bead-on-a-string structure, which subsequently condenses with the help of linker histones into higher-order structures, the exact conformation of which is still under discussion. Concepts such as the 30 nm fiber (Widom and Klug 1985) or the

"zig-zag" structure (Bednar et al. 1998) have been developed to explain in-vitro data, but whether these exist in the cell is still not proven (McDowall et al. 1986; Fussner et al. 2011). Changes in nucleosome structure affect all levels of chromatin organization—from local nucleosome-nucleosome interactions to interactions between global chromatin domains (Jin et al. 2013)—and thus have immense impact on genetic activity.

Through their sheer presence nucleosomes pose an important barrier to genetic processes by restricting access of DNA to nuclear proteins during transcription, replication and DNA repair. Eukaryotic cells employ various mechanisms that alter nucleosome structure and stability to facilitate temporary access to DNA (Fig. 1b). These include repositioning of nucleosomes along DNA by specific remodeling factors, chaperone-mediated re- and disassembly and transient site exposure through thermal fluctuations, commonly referred to as "breathing".

These processes need temporary breakage of DNA-protein interactions. The length of DNA that wraps around one histone octamer is a little less than 50 nm. This is comparable to its persistence length, the length scale at which unperturbed B-DNA behaves like a stiff elastic rod. Wrapping DNA of this length almost twice around a protein core of less than 7 nm diameter requires of the order of 100 k_BT of bending energy. This is balanced by many electrostatic interactions and hydrogen bonds between histones and DNA. Crystallographic studies (Luger et al. 1997; Davey et al. 2002) revealed that these interactions are not uniformly distributed along the DNA path, but clustered approximately every 10 bp. The average net binding energy per helix turn just exceeds 1 k_BT, so that thermal fluctuations suffice to frequently break interactions between DNA and protein, in particular in the DNA entry-exit region (Prinsen and Schiessel 2010). Recently, force experiments revealed the detailed distribution and strength of these interactions with near base-pair resolution. By monitoring the resilience of nucleosomal DNA to unzipping Hall et al. (Hall et al. 2009) generated an interaction map which showed highly non-uniform interaction strength along the DNA path. Strongest DNA-protein interactions were found near the dyad axis, with two side maxima of weaker strength approximately 40 bp away from the dyad axis. The highly uneven distribution of DNA-protein interaction strength has profound consequences on nucleosome dynamics, as we will discuss below.

Histones are among the most conserved proteins found in nature and can virtually bind to every DNA sequence—in human cells about 80 % of DNA is organized in nucleosomes. DNA sequence has strong impact on positioning and stability of nucleosomes, the binding strength between sequences can vary more than 1000-fold (Thastrom et al. 1999). Evidence is growing that in particular the physical properties of DNA, such as curvature or flexibility, determine relative binding strength to a great extent (Segal et al. 2006). Besides variation in DNA sequence nucleosomes can also vary in their protein composition—numerous variants of canonical histones are known to modulate nucleosome architecture and alter compaction into higher-order structures (reviewed in (Luger et al. 2012)). Finally, posttranslational modifications (PTMs) on DNA and protein add another dimension of structural variability, with strong impact on genetic activity

(Strahl and Allis 2000). PTMs can directly modulate electrostatic interactions, for example by neutralizing the charge of a residue, or serve as a docking point for specific nuclear factors.

From this perspective it is clear that nucleosomes—once thought to be a purely static component in DNA compaction—represent highly dynamic entities that have profound impact on gene regulation. Biochemical and crystallographic data have provided enormous insight in the overall structural properties of nucleosomes. It was shown that DNA sequence, posttranslational modifications and replacement of canonical histones with variants do not significantly alter nucleosome structure. Since these techniques provided only a static snapshot of nucleosome structure, dynamic properties associated with these changes evaded analysis. The change of a single amino acid, for example, is already sufficient to significantly alter nucleosome mobility (Flaus et al. 2004) and its dynamic properties (Neumann et al. 2009). Thus, to understand the complex mechanisms governing genetic activity, one needs to probe the dynamic behavior of nucleosomes.

Here, we will discuss the application of Förster Resonance Energy Transfer (FRET) to probe the dynamic properties of individual nucleosomes and small nucleosome arrays. We will first describe the fundamental principles behind FRET and the different implementations that are currently used. Particular emphasis will be laid on single-molecule FRET techniques, which provide information that is otherwise hidden within the ensemble; they can help to determine rates and lifetimes of dynamic processes and reveal the presence of short-lived intermediate states. We will then review recent key experiments that shed light on the dynamic properties of nucleosomes. These include site exposure by spontaneous fluctuations in the entry-exit region and the detailed sequence by which nucleosomes disassemble. We will conclude the chapter with a brief discussion of the impact of post-translational modifications on nucleosome architecture. Throughout the chapter we will focus on an experimentalist's perspective. For a modeling approach to nucleosome dynamics we refer the reader to (Biswas et al. 2013).

2 Förster Resonance Energy Transfer (FRET)

FRET is a spectroscopic technique to quantify changes in biomolecular structure and dynamics. It is based on a non-radiative energy transfer between a donor and an acceptor fluorophore as they approach each other within a few nanometers (Fig. 1c, d). Upon selective donor excitation part of its energy can be transferred to the nearby acceptor. This results in a decrease of donor fluorescence intensity and donor lifetime and usually an increase in fluorescence emission of the acceptor. This process strongly depends on the distance between both dyes and provides a "molecular ruler"; changes in molecular structure result in changes in distance and energy transfer between both dyes.

The first classical treatment by Perrin described the phenomenon as the interaction of two dipole oscillators (Perrin 1927). His theory captured the essence of the

process but could not reproduce the correct distance-dependence observed in experiments. An exact quantum mechanical theory was presented in 1946 by Theodor Förster (Förster 1946). He showed that FRET scales with the inverse 6th power of the interdye distance r:

$$E = \frac{1}{1 + (r/R_0)^6}.$$

Here, R_0 is the Förster radius, which is defined as the distance at which 50 % of the excitation energy is transferred to the acceptor. R_0 is a property of the dye pair that is used in the experiments, and typically ranges between 4 and 6 nm.

2.1 Experimental Determination of Energy Transfer and Interdye Distances

Energy transfer can be quantified by measuring donor and acceptor intensities or donor lifetime. For brevity, we will restrict our discussion to intensity-based experiments, since they are readily implemented on a many different instruments such as microscopes, spectrometers or multimode imaging systems. A discussion of lifetime-based FRET analysis can be found in (Clegg 1996).

FRET is usually calculated from the sensitized emission of the acceptor dye (Clegg 1992). To do so, the donor fluorophore is selectively excited near its absorption maximum, and fluorescence intensities are recorded from the donor and acceptor dye. These need to be corrected for background, spectral crosstalk between the dyes and residual direct excitation of the acceptor, see (Gansen et al. 2015). After correction, I_A only contains the sensitized emission due to FRET, and I_D represents the remaining donor emission in the presence of FRET. Energy transfer is then computed as:

$$E = \frac{I_A}{I_A + \gamma \cdot I_D}.$$

γ is the detection factor of the system and depends on the detection efficiencies and quantum yields of donor and acceptor. Usually, it can be determined with less than 10 % uncertainty. Often, researchers are interested in relative changes in FRET and a simplified expression of Eq. 2 can be used, where γ is taken as 1. This so-called proximity ratio suffices as a reporter of conformational changes. Previously, we pointed out that using the proximity ratio under conditions, where the detection factor is deliberately increased or reduced (deliberately detuned detection—D3), can increase the separation of subspecies within a given region in FRET (Gansen et al. 2009a).

If measured FRET efficiencies need to be converted into accurate interdye distances, the Förster radius has to be known precisely. R_0 is not a constant

and depends on several parameters, including the spectral overlap between donor and acceptor and their relative orientation. In particular the latter is a frequent source of uncertainty. R_0 values quoted in the literature often assume that the fluorophores can rotate freely in all directions on a time scale faster than the fluorescence lifetime. This approximation is justified, if the fluorophores are attached to DNA or protein by a long carbon linker. In cases, however, where the fluorophore mobility is restricted, for example when bifunctional linker chemistry is used or in the case of steric hindrance by DNA or protein, this assumption is no longer valid. Here, assuming freely rotating dyes can easily introduce an error of 20 % or more in distance (Ivanov et al. 2009).

3 Sample Preparation

Successful FRET experiments require well-defined nucleosome substrates, with accurate label positions on DNA and protein. They usually rely on strong nucleosome positioning sequences, such as the SELEX 601 sequence (Lowary and Widom 1998), but also a number of naturally occurring DNA sequences have been used in FRET experiments, such as the 5S rDNA or several promoter sequences, like GAL10, GUB or MMTV. Regardless of the DNA sequence used, the position of the reporter dyes on DNA or protein has to be chosen carefully to optimally follow the dynamic process of interest. The crystal structure of the nucleosome is a good starting point to optimize dye positions. Fluorophores can be virtually attached to any position of interest and computer simulations can predict the effective interdye distance as well as restrictions in dye mobility that could affect interpretation of FRET data (Kalinin et al. 2012).

A few typical labeling schemes are shown in Fig. 2. Dyes located in the entry-exit region of the nucleosome provide information on DNA unwrapping dynamics, while labels positioned in the internal region or on histone proteins can report on

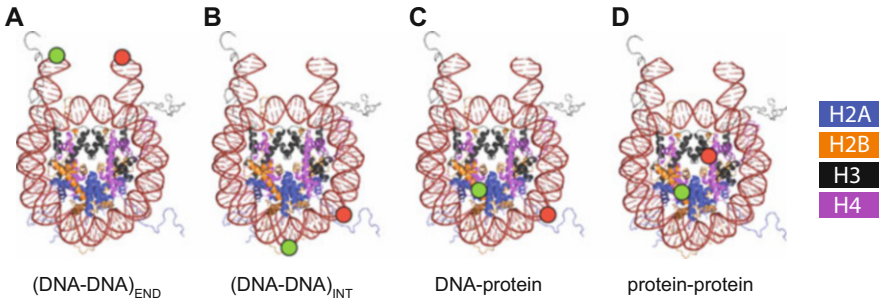

Fig. 2 Typical labelling schemes used for FRET experiments on nucleosomes. (**a**) Fluorophores in the entry/exit region report on DNA unwrapping. (**b**) Labels at the inner part of nucleosomal DNA are sensitive to nucleosome dissociation. (**c**, **d**) Labeling one or more histones provides information on the integrity of the octamer and the fate of individual histones during disassembly

nucleosome disassembly. Labeling of nucleosomal DNA is conveniently done by preparative PCR using fluorescently labeled primers. Protein labeling usually requires site-specific mutation to introduce a cysteine at a specific location to which a small organic fluorophore can be coupled via standard maleimide chemistry (Böhm et al. 2011).

Nucleosomes can be assembled by mixing recombinant histones and DNA at high NaCl concentration (typically 2 M) and slow dialysis against low salt buffer. The gradual decrease in ionic strength then promotes the step-wise assembly of nucleosomes; first, the $(H3\text{-}H4)_2$ tetramer associates with DNA around 1 M NaCl, followed by addition of the two H2A-H2B dimers at lower salt (600-700 mM) to organize the remaining DNA. An alternative approach to reconstitute nucleosomes at physiological salt is to use chaperones like NAP-1 to orchestrate correct nucleosome assembly (Choy et al. 2010).

4 FRET Methods

In this paragraph we will describe the implementation of FRET experiments in bulk and on the single molecule level. A schematic layout of the different techniques is shown in Fig. 3.

4.1 Ensemble FRET Experiments

In bulk experiments, billions of molecules are observed at the same time and the FRET information is averaged over the whole ensemble—molecular individuality and diversity is lost. Traditional bulk experiments are performed in cuvette-based fluorimeters with tens of microliters of sample at concentrations of several tens of nanomolar. Typically, whole fluorescence spectra are recorded; donor and acceptor intensities are inferred from a small region around their respective maxima and bulk FRET can be determined with good accuracy. A detailed description of the experimental procedures specifically adapted for nucleosome samples can be found in (Toth et al. 2001). To probe the response of the ensemble to an external perturbation, the emission at specific wavelengths can be continuously monitored over time, providing kinetic information in the seconds to minutes range.

Acquisition of full emission spectra is time-consuming and only one sample is processed at a time. Experiments probing nucleosome properties under various conditions, such as ionic strength and PTMs, would require long acquisition times, and care has to be taken to guarantee identical system performance throughout the experiment. To overcome these limitations Gansen et al. introduced a powerful array method to scale-up bulk FRET experiments at increased sensitivity (Gansen et al. 2013). In this "microplate-scanning FRET" (Fig. 3a) samples are distributed into 384-well microplates and fluorescence is imaged on a commercial multimode

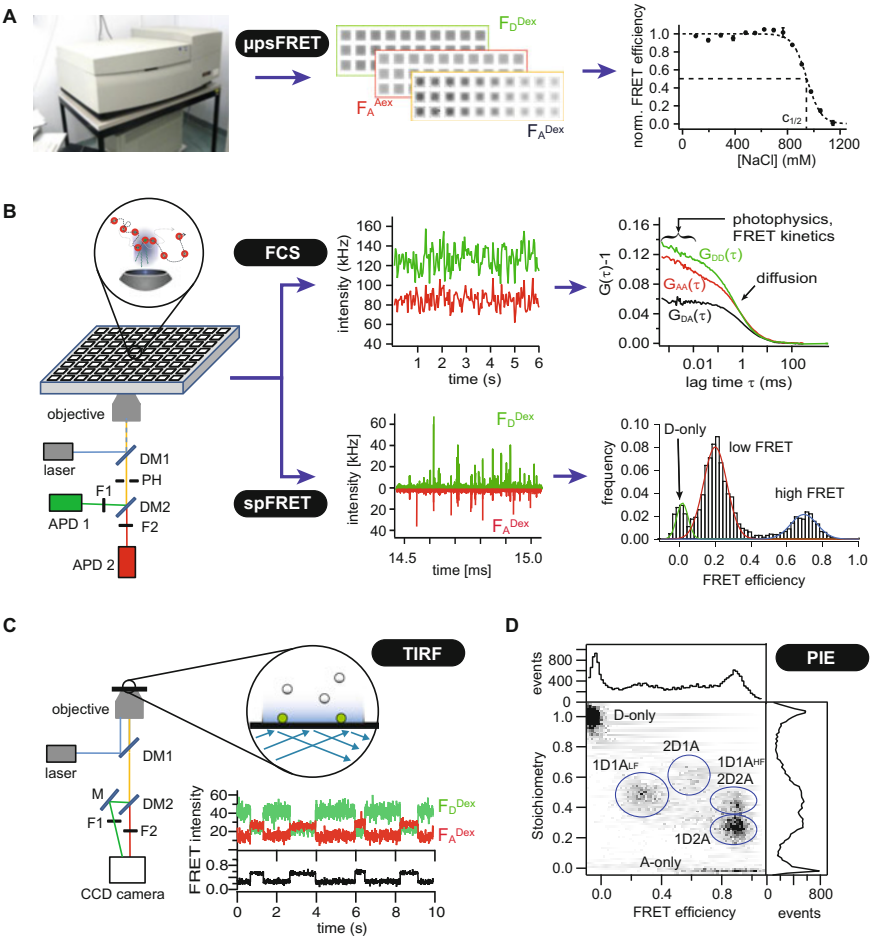

Fig. 3 Experimental strategies to measure FRET. (**a**) The principle of microplate-scanning FRET (μpsFRET). Samples are distributed into a multiwell plate and imaged on a multimode scanner. Three images are taken: donor emission upon donor excitation (donor signal, F_D^{Dex}), acceptor emission upon donor excitation (FRET signal, F_A^{Dex}) and acceptor emission upon acceptor excitation (acceptor signal, F_A^{Aex}). FRET is calculated per well and bulk FRET can be analyzed as a function of various parameters, such as ionic strength or selective modification of certain amino acids. For example, the NaCl concentration at which 50% of FRET is lost (denoted as $c_{1/2}$) can serve as a measure of overall nucleosome stability. (**b**) Schematic of a confocal microscope setup. A laser beam is focused into solution and fluorescence is detected by single photon counting avalanche photodiodes (APDs). The passage of fluorescent molecules through the focus generates intensity fluctuations that can be analyzed via their auto- or cross correlation function (Fluorescence Correlation Spectroscopy (FCS), *upper panel*). In spFRET, sample concentration is reduced such that the transit of individual molecules through the focus can be followed as separate bursts of fluorescence. FRET can be computed for each burst and the heterogeneity within the ensemble is reflected in the distribution of single molecule FRET efficiencies (*lower panel*). (**c**) Schematic of a TIRF microscope. Total reflection of the laser beam at the glass-solution interface generates an evanescent field at the surface. Only molecules within a few hundred nanometers from the surface

scanner. Spectral analysis is replaced by detection of photons within a predefined spectral range using bandpass filters and fast photomultiplier detectors. For each set of samples three images are taken: (a) donor emission upon donor excitation, (b) acceptor emission upon donor excitation, and (c) acceptor emission upon selective acceptor excitation. Intensities from (a) and (b) are used to determine FRET, while intensities from image (c) are used to correct for direct acceptor excitation. Energy transfer is then computed for each well and hundreds of samples can be analyzed in parallel. The sensitivity of this method approaches sample concentrations used in single molecule experiments, which allows to follow binding reactions with sub-nM binding constants.

4.2 Fluorescence Correlation Spectroscopy (FCS)

A common method to probe dynamic behavior on a faster time scale (microseconds to milliseconds) is Fluorescence Correlation Spectroscopy (FCS). Most FCS experiments are performed in a confocal microscope (Fig. 3b). A femtoliter-sized observation volume is generated by focusing a collimated laser beam into solution. Fluorescence from molecules passing in and out of the focus is imaged onto a pinhole that is located in the conjugated image plane. Only light emerging from the center of the focus is transmitted and spectrally separated onto two fast avalanche photodiodes. FCS is nestled between the ensemble and single molecule regime. Compared to bulk FRET, the number of observed molecules is vastly reduced—for a typical focus volume of 1 femtoliter and a sample concentration of 10 nM, on average only 6 molecules are observed at the same time. At such low concentrations random diffusion of molecules in and out of the focus generates significant fluctuations in the fluorescence signal I(t). These contain information about the diffusion properties of molecules, which depend on their shape and size, as well as photophysical properties and fluctuations in FRET due to changes in molecular structure. The time scales of the fluctuations are analyzed by computing the correlation function of the fluorescence signal:

Fig. 3 (continued) can be excited (*green circles*). No signal is observed from molecules that remain in the bulk solution (*white circles*). The fluorescence is spectrally separated and imaged onto a CCD camera. Time traces of donor and acceptor intensities can be extracted for each molecule and transitions in FRET can be followed over time. (**d**) Exemplary distribution of dye stoichiometry vs. FRET for nucleosomes with the donor on histone H2B and the acceptor attached to histone H4. Due to incomplete protein labeling and photophysical artifacts several subpopulations are observed that can be classified according to FRET efficiency and the relative number of active donors (D) and acceptors (A), called the stoichiometry S. Additional abbreviations: *PH* pinhole, *DM* dichroic mirrors, *M* reflective mirror, *F* emission filter

$$G_{X,Y}(\tau) = \frac{\langle I_X(t) \cdot I_Y(t+\tau) \rangle}{\langle I_X(t) \rangle \langle I_Y(t) \rangle},$$

where τ is the lag time and X and Y represent the different detection channels, for example donor and acceptor signal. For $X = Y$ the autocorrelation of the respective signal is computed, while $X \neq Y$ represents the cross-correlation of both signals. Either correlation function is dominated by the diffusion of the particles, but fluctuations on a faster time scale show up as additional components in $G(\tau)$. FRET fluctuations result in anti-correlated donor and acceptor signals and show up differently in the cross-correlation function than in the autocorrelation function. A global fit of all correlation functions can be used to extract the kinetic information from the FCS data (Torres and Levitus 2007).

4.3 Diffusion-Based spFRET

Diffusion-based single molecule experiments are a powerful tool to study the structural diversity within an ensemble and are usually performed in a confocal microscope. The concentration of fluorescent molecules is reduced to several tens of pM, and the average number of molecules in the focus is much smaller than one. Then, the observed signal consists of spurious bursts of fluorescence that reflect the transit of individual molecules through the laser focus. These bursts are discriminated against the background, and the detected photons in each burst allow to compute an energy transfer (and an apparent interdye distance) for each molecule. Single molecule events are then collected in a histogram where subspecies differing in their FRET values show up as individual peaks (Gansen et al. 2007). The events may be further classified by additional parameters such as fluorescence lifetime, polarization etc. Subspecies can be characterized by their individual FRET and selected for further analysis, for example a selective FCS analysis that determines the diffusion properties of bursts within a certain subspecies (Koopmans et al. 2009; Felekyan et al. 2012).

The observation time per molecule is limited by their diffusion time and of the order of a few milliseconds. In general, the properties of every detected molecule can be determined with an accuracy that is limited by the number of photons that were collected during this time. This number is fairly small—only a few tens to hundred photons per molecule—and shot noise of the detection process results in considerable broadening of individual peaks in the FRET distribution. Subpopulations with similar FRET can merge into a broad, seemingly undefined distribution, which renders the interpretation of structural heterogeneity difficult. The shot noise scales with the inverse square root of the number of detected photons and can be accounted for during data analysis to recover the actual distribution of FRET states from experimental data (Antonik et al. 2006; Widengren et al. 2006; Nir et al. 2006).

4.4 TIRF Microscopy

Often, researchers wish to observe the dynamic behavior of a single molecule over an extended period of time. TIRF microscopy of surface-immobilized molecules can provide kinetic traces of conformational dynamics of many seconds to minutes. In TIRF, a collimated laser beam emerges from the objective at an angle larger than the critical angle for total reflection. The laser beam is completely reflected off the solution-glass interface, generating an evanescent field that penetrates only a few hundred nanometers into solution. Only molecules that are directly located at the surface are efficiently excited. The signals from donor and acceptor are separated and imaged side-by-side on a CCD camera (Fig. 3c). Single molecule detection is ensured by keeping the surface immobilization density low enough, so that the images of each molecule are well separated. This way, several molecules can be observed in real time with a time resolution that depends on the speed of the camera, typically around 10 ms/frame. The total observation time for a molecule is only limited by photobleaching—under proper conditions time traces of several minutes can be obtained.

For analysis, donor and acceptor images in each frame need to be properly aligned to assign a donor and an acceptor spot to the same molecule. In general this requires software-based translation of one image relative to the other and, in some cases, additional rotation and rescaling. From the relative intensities of the matched image pair FRET can then be computed for each molecule in each frame. Extended observation of nucleosomes requires their immobilization on the surface. Over the years several surface functionalization schemes have been developed to specifically tether nucleosomes onto glass surfaces [reviewed in (Koopmans et al. 2008)].

5 Practical Aspects of spFRET on Nucleosomes

Single molecule detection is one of the most demanding fluorescence techniques currently in use. Its unique ability to resolve subpopulations in a heterogeneous sample requires careful optimization of several experimental parameters.

The signal from the fluorophore has to outmatch that of all other molecules in the buffer solution. Unwanted background from fluorescent impurities can be minimized by using ultrapure chemicals and ultrapure water, and most of the scattered laser light can be rejected by high-grade optical filters. However, some residual background due to inelastic Raman scattering, where the scattered laser light is shifted to longer wavelengths and superimposes on the emission spectrum of the fluorophore, remains.

Diffusion-based spFRET experiments rely on much higher laser intensities than bulk FRET or FCS for efficient single molecule detection. Higher intensities result in more photons being emitted, but the useable laser intensity is limited by the photophysical properties of the fluorophores. Background from scattering increases

linearly with laser intensity, but emission of the dye saturates due to triplet accumulation—resulting in a decrease in signal-to-noise ratio at too high laser intensities. A good compromise is the so called saturation intensity, where 50 % of maximum dye emission is achieved (Moerner and Fromm 2003).

Triplet accumulation is also a major pathway for photodestruction. The dwell time in the triplet state is long enough to allow for additional reactions with the local environment. Molecular oxygen can quench the triplet state by converting into singlet oxygen, which can then photooxidize the fluorophore. Thus, in spFRET it is mandatory to provide sufficient photostabilization; popular strategies include the use of oxygen scavenging systems (Roy et al. 2008), triplet quencher (Gansen et al. 2007) or continuous removal of oxygen from solution.

In aqueous solution, a typical fluorophore can emit about 1 million photons before it photobleaches. Due to incomplete photon collection in the optical setup and the limited quantum efficiency of the detector only a small fraction of these will be recorded. The rate at which photons are emitted depends on the laser intensity and eventually determines for how long the molecule can be observed. In other words the time scale of the dynamic process under study determines the experimental method of choice. Confocal experiments use high intensities and aim to collect as many photons as possible during the short transit time through the focus. In TIRF, molecules should be observed for an extended time period and the laser intensities are much lower than in confocal experiments. Available photons are then emitted over several seconds or minutes.

5.1 Multiparameter FRET Analysis

FRET experiments on nucleosomes are complicated by incomplete labeling of DNA or protein and insufficient photostabilization. If only the donor is excited it is almost impossible to distinguish an intact nucleosome that does not undergo FRET because of an inactive or missing acceptor from a disrupted nucleosome, where the dyes are too far apart for FRET. Incomplete histone labeling results in a mixture of nucleosomes with different numbers of donor and acceptor which renders interpretation of spFRET data difficult.

To overcome these issues, donor and acceptor can be sequentially excited during the nucleosome's diffusion through the focus. Alternating laser excitation (ALEX) (Kapanidis et al. 2004) or pulsed interleaved excitation (PIE) (Müller et al. 2005) not only monitor the presence and integrity of both fluorophores, but also provide information on the relative stoichiometry of both dyes—complexes that carry one donor and one acceptor can be discriminated from those with two donors and one acceptor or vice versa. FRET and dye stoichiometry are analyzed in a 2-dimensional histogram and appropriate selection of subspecies helps to remove artifacts from incomplete labeling or imperfect photostability.

A different approach was developed in the Seidel lab, which used the simultaneous acquisition of different fluorescence parameters, such as intensity, lifetime or

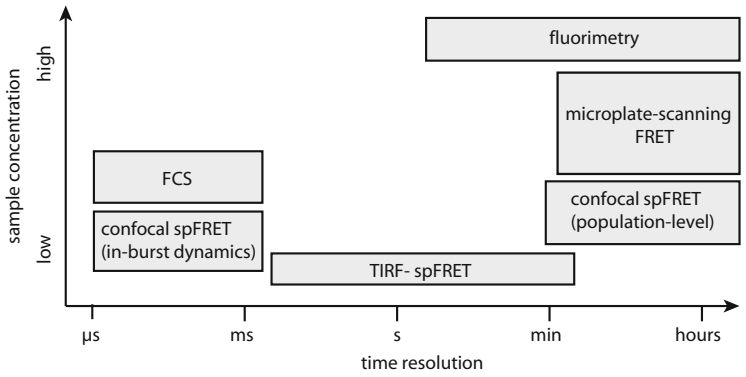

Fig. 4 Sample concentrations and dynamic time scales that can be analyzed with different intensity-based FRET methods. In FCS, fluctuations are best resolved if they occur on time scales lower than the diffusion time. In confocal spFRET, advanced data analysis tools allow for analysis of fast FRET fluctuations within a single molecule burst ("confocal in-burst"). Alternatively, the change in the number of molecules within a FRET subpopulation can be followed over time with much lower time resolution. A statistically significant number of events has to be collected to correctly interpret changes in histogram shape. The time resolution and maximum observation time in TIRF are limited by the camera speed and photobleaching of the molecule respectively. In microplate-scanning FRET, time resolution is limited to a few minutes by the image acquisition speed. Temporal resolution is compromised in favor of enhanced sample throughput. In fluorimetry, kinetic information on the seconds time scale can be obtained by following the emission at the peak wavelengths over time

anisotropy (Multiparameter Fluorescence Detection, MFD) (Widengren et al. 2006). A powerful tool emerges when MFD is combined with PIE to maximize information from both the donor and acceptor fluorophore (Kudryavtsev et al. 2012). The benefits of PIE for studies on protein-labeled nucleosomes is illustrated in Fig. 3f, where the nucleosomes were labeled on histones H2B and H4 and measured at 50 mM NaCl, Due to incomplete labeling multiple nucleosome species are observed, that can only be resolved by combining FRET efficiency and stoichiometry information. All subspecies are well resolved in the 2D representation, but not in the individual 1D histograms of E and S.

Figure 4 summarizes the sample concentration requirements and time scales for the different FRET-based approaches.

5.2 Sample Integrity and Stability

At single molecule conditions only few molecules are present in solution and their adsorption to the chamber walls has to be minimized. A straightforward strategy is the addition of excess inert protein in solution, which can adsorb to the chamber walls. The presence of about 0.2 mg/ml bovine serum albumine (BSA) in solution

prevented nucleosome degradation up to several hundred mM NaCl (Gansen et al. 2007), but proved less effective at higher ionic strength. Better passivation is achieved by small amounts of non-ionic surfactant added into solution. Popular detergents include Non-Idet P40 (Gansen et al. 2013) or CHAPS (Menshikova et al. 2011). For experiments on salt-induced disassembly we found that a combination of 0.01 % Non-Idet P40 and silanization of the bottom glass surface provided excellent passivation (Gansen et al. 2013).

In TIRF microscopy, the properties of the surface affect the integrity of immobilized nucleosomes and influence their dynamics. Unspecific adsorption has to be avoided to ensure that the observed signal stems from actual nucleosomes. Koopmans et al. compared various strategies and reported that a 6-arm branched polyethylene glycol (PEG) coating prevented unspecific binding and increased the fraction of intact nucleosomes more than two-fold compared to standard PEG-coated or BSA-coated surface (Koopmans et al. 2008). An intriguing alternative to surface tethering is encapsulation of single molecules into surface-attached liposomes (Cisse et al. 2007). The liposome acts as a hollow container, while the molecule is free to move within its aqueous compartment. This concept proved useful to study the dynamic landscape of small proteins and nucleic acids, but whether is it also applicable to a complex structure like the nucleosome awaits confirmation.

6 Advantage of smFRET Over Bulk

The essential advantage of spFRET over bulk FRET is that individual subspecies in a population can be characterized by their different FRET values. This is illustrated in Fig. 5, which compares a salt-titration experiment of nucleosomes in bulk and on the single molecule level (Gansen et al. 2013). In both cases 300 pM nucleosomes were incubated at different ionic strength for 1 h. In bulk, the average FRET efficiency was determined using microplate-scanning FRET. For spFRET 50 pM labeled nucleosomes were mixed with 250 mM unlabeled nucleosomes prior to incubation. The structural heterogeneity was then analyzed for each salt concentration. The fraction of intact nucleosomes showed a similar salt-dependence as the average FRET efficiency measured in bulk. Inspection of the spFRET histograms revealed that prior to dissociation a second population emerges at higher FRET values, suggesting a pronounced structural rearrangement or formation of an intermediate structure. This information is not accessible in bulk FRET experiments but is crucial for our understanding of the mechanism of nucleosome disassembly as described below.

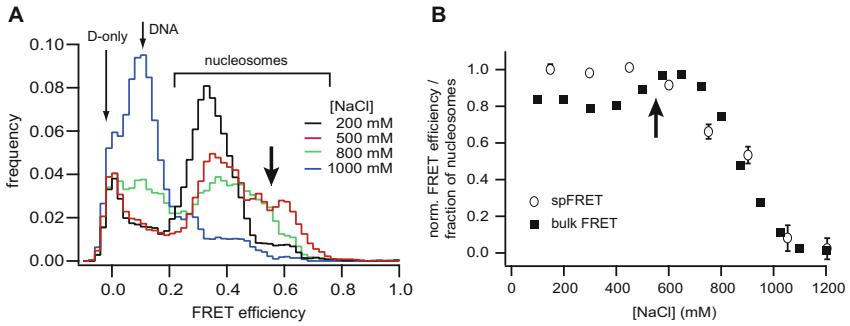

Fig. 5 The advantage of spFRET over bulk FRET. (**a**) spFRET distributions of nucleosomes at various NaCl concentrations. Going from 200 to 500 mM NaCl no increase in free DNA is observed, but some nucleosomes transit into a conformation with shorter interdye distance and higher FRET (marked by an *arrow*). Dissociation at higher ionic strength results in a decrease in overall nucleosome population and an increase in the fraction of free DNA. (**b**) Comparison of salt-induced dissociation measured in bulk and spFRET. Data are normalized to the respective maximum of each curve. In bulk FRET, formation of the high-FRET population results in an increase in ensemble FRET prior to dissociation (see *arrow*). In spFRET, the total fraction of nucleosomes is not affected by this transition. The subsequent decrease in nucleosome population correlates with the loss of FRET observed in bulk, confirming that both methods yield comparable results on nucleosome dissociation. The structural transition prior to disassembly is hidden within the ensemble FRET; spFRET, on the other hand, can quantify the fraction of nucleosomes undergoing the transition and the extent of the structural change involved (see example in Fig. 8)

7 Dynamic Properties of Canonical Nucleosomes

Research on the structure-function relationship of the nucleosome advanced explosively in the 1990s after the essential role of the DNA sequence for successful crystallization had been determined (Harp et al. 1996) and the first high-resolution X-ray structure had been solved in the Richmond lab (Luger et al. 1997); for a historical account, see (Vence 2015). The high-resolution X-ray structure, later refined to 1.9 Å (Davey et al. 2002) showed that the DNA ends are quite loosely bound while internal sites show the strongest contacts to the histone core. This finding is essential for interpreting the results on nucleosome breathing dynamics and disassembly in solution, outlined in the following.

7.1 Nucleosome Breathing

In general, there is a consensus that the terminal ends of nucleosomal DNA undergo spontaneous unwrapping and rewrapping, thereby transiently exposing binding sites for nuclear proteins. The first quantitative estimates of DNA opening in nucleosomes were given by Polach and Widom (Polach and Widom 1995) through

an interpretation of restriction enzyme digestion rates of sites at various positions on the nucleosomal DNA. Later, by analyzing the kinetics of transcription factor binding through a combination of FRET with stopped-flow kinetics and fluorescence correlation spectroscopy (FCS), the same lab gave evidence for spontaneous fluctuations of the outer parts of nucleosomal DNA on the millisecond time scale (Li et al. 2005). This time scale was further refined by direct observation of DNA unwrapping in immobilized nucleosomes (Tomschik et al. 2005, 2009; Koopmans et al. 2007). Using TIRF microscopy, transitions between open and closed nucleosome conformations were directly followed as step-wise fluctuations in FRET efficiency. These experiments agree on time scales of about 25–50 ms for the open and 250 ms for the closed state; in other words, nucleosomes are open for about 10–20 % of the time.

The rates of DNA unwrapping and rewrapping strongly depend on the position of DNA inside the nucleosome. Using diffusion-based spFRET, Koopmans et al. showed that DNA breathing at a location 27 bp inside the nucleosome was reduced by half compared to breathing at the DNA ends (Koopmans et al. 2007, 2009). Tims et al. (2011) compared spontaneous fluctuations and transcription factor binding at different positions inside the nucleosome. They also observed a decrease in equilibrium constant, which was dominated by a decrease in the rate of DNA unwrapping, while DNA rewrapping was less affected. This is consistent with a simple model for DNA breathing; successive DNA unwrapping from the ends requires more DNA-protein interactions to be broken. The total energy cost increases and extended DNA breathing becomes less likely.

North et al. have shown that the unwrapping rate, but not the rewrapping rate, is modulated by the DNA sequence in the entry/exit region (North et al. 2012). Changing the first 7 bp of a 601 DNA fragment into 5S DNA led to a 2.5-fold decrease in unwrapping rate. Variation of DNA sequence more than 28 bp inside the nucleosome, however, did not affect the kinetic rates at all.

DNA unwrapping also depends on the length of linker DNA and the presence of linker histone H1. Bulk FRET data from our own lab showed that the linker arms assume a more open conformation free in solution than in the crystal, and that addition of histone H1 decreased the distance between linker DNA ends (Toth et al. 2001). These data could be interpreted as a change in the ratio of open and closed conformation—the presence of linker DNA promotes DNA unwrapping due to an increase in electrostatic repulsion of the DNA arms. Similar findings were reported in a recent spFRET study that showed that the addition of 300 bp linker DNA can increase the fraction of nucleosomes in the open state by 30–50 % (Buning et al. 2015).

7.2 Nucleosome Disassembly

Unwrapping of terminal DNA not only provides transient access to binding sites buried within the nucleosome but can also initiate nucleosome disassembly.

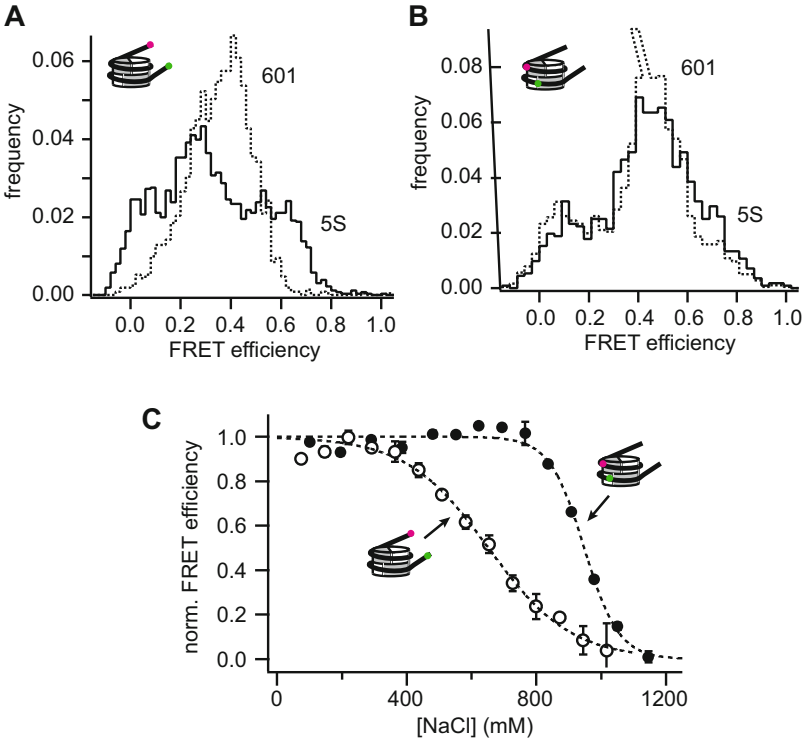

Fig. 6 Nucleosome disassembly proceeds from the terminal ends. (**a**, **b**) Sequence-dependent structural variability of nucleosomes in the linker DNA region (**a**) and at a DNA site buried inside the nucleosome (**b**) at 100 mM NaCl. Nucleosomes carrying the 5S sequence show increased heterogeneity of DNA end-to-end distance than nucleosomes reconstituted onto 601 DNA. No significant differences in internal structure were observed. (**c**) Salt-induced disassembly measured by microplate-scanning FRET. FRET between the ends of the DNA is lost at lower NaCl concentration than FRET between labels in the internal region. This demonstrates that dissociation of the nucleosome at higher ionic strength is facilitated by opening of the DNA ends

Comparing nucleosomes with end-labeled and internally labeled DNA of different sequence showed that the disassembly commences by increased dynamics of the linker DNA ends (Fig. 6a, b). At physiological ionic strength both internally labeled nucleosomes show very similar FRET histograms, while sequence-specific differences were observed in the entry/exit region of the DNA. The strongest positioning 601 sequence shows a much narrower distribution than the 5S rDNA, indicating that the latter sequence has more conformational freedom. End-labeled 5S nucleosomes frequently assume a more compact structure with higher FRET than the 601 samples, in agreement with reduced DNA unwrapping observed by North et al. At higher salt, nucleosomes progressively disassemble from the ends. FRET between the DNA ends is lost at lower ionic strength than FRET between internal sites on the DNA (Fig. 6c).

Nucleosome disassembly is essential to overcome the steric barrier that nucleosomes pose to DNA-processing enzymes. In vivo studies have shown that the speed of polymerase transcribing through the nucleosome is similar to that measured on free DNA. Therefore, an efficient mechanism has to be employed to rapidly disassemble nucleosomes before and reassemble them behind the elongating polymerase to maintain chromatin integrity (Clark and Felsenfeld 1992).

The stability of nucleosomes against disassembly correlates with their intrinsic binding affinity and reflects the number and strength of DNA-protein and protein-protein interactions. Early work from the Luger and Widom labs (Gottesfeld and Luger 2001; Thastrom et al. 2004) quantified nucleosome stability through the ratio of bound to unbound DNA in polyacrylamide gels after incubation at different ionic strength. Interpretation of the apparent binding constants is difficult, however, since nucleosomes are multicomponent complexes, and binding constants are only defined for interaction of the individual subcomponents in equilibrium (Andrews and Luger 2011).

The investigation of salt-induced nucleosome disassembly by FRET has become a well-established method to analyze nucleosome stability, and it is generally believed that intermediate species observed at higher ionic strength are similar to those that are transiently populated in the cellular context (Arimura et al. 2012). Often, only the overall stability of the complex needs to be known and bulk FRET measured in microplates containing many samples in parallel has an advantage over single molecule studies because of its speed, ease of use and sensitivity. Detailed investigation of the disassembly pathway requires analysis of individual nucleosomes, however, to capture transient intermediate structures that are otherwise hidden within the ensemble.

Earlier studies have shown that nucleosome disassembly proceeds by partial loss of H2A/H2B dimers (Yager et al. 1989; Claudet et al. 2005; Kelbauskas et al. 2008), but the exact sequence and mechanism of disassembly is not yet fully understood. Little is known about the existence and dynamic properties of intermediate structures that are populated during nucleosome disassembly and assembly. Even the spontaneous formation of such partially disrupted structures is currently under debate.

A thorough spFRET study of nucleosome disassembly was published by Gansen et al. (2009b). Nucleosomes were reconstituted on a 170 bp DNA fragment containing the Widom 601 positioning sequence, with both fluorophores placed in the central region of the nucleosome. Increase in salt concentration was used to disassemble the nucleosomes in a controlled way. For the first time multiparameter fluorescence detection (MFD) (Antonik et al. 2006) was used on individual nucleosomes and revealed the coexistence of three states with different FRET; intact, fully compacted nucleosomes, a partially opened intermediate species and fully open DNA.

A model for stepwise disassembly was proposed that was subsequently refined by collecting FRET information from different DNA and protein locations in the nucleosome. Böhm et al. (2011) analyzed nucleosome populations through a systematic variation of donor and acceptor label positions on H2B, H4 and DNA

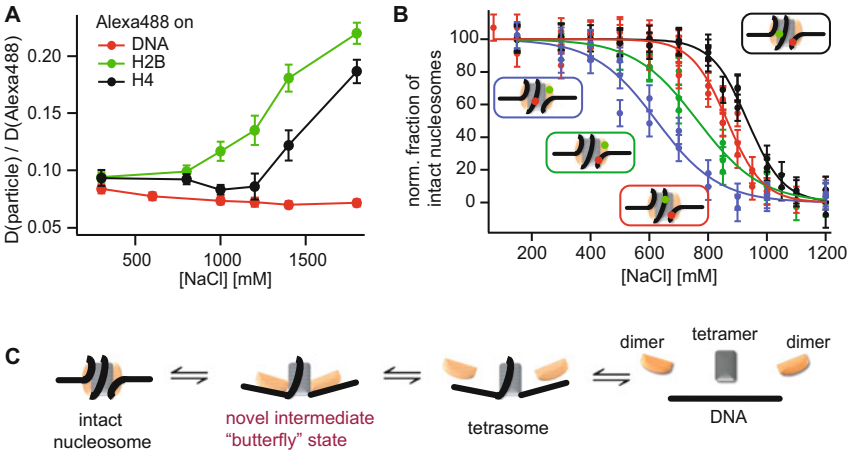

Fig. 7 Nucleosome disassembly proceeds through an intermediate structure with partially disrupted histone contacts. (**a**) Sequential release of subcomponents from the nucleosome as measured with FCS. Alexa488 was attached to either histone H2B, histone H4 or DNA. The increase in diffusion constant with salt indicates eviction of the different subcomponents from the complex. H2A/H2B dimers are released first, followed by the $(H3–H4)_2$ tetramer. (**b**) Identification of a novel intermediate species during salt-induced disassembly by systematic variation of FRET labels across the nucleosome. FRET between H2B and nearby DNA (*green*) is lost at higher salt than FRET between H2B and DNA near the dyad axis (*blue*). This suggests an intermediate conformation, where contacts between dimer and tetramer are broken, but the dimer remains bound to DNA ("butterfly state"). After dimer dissociation, DNA opens up and the tetramer is lost around 1 M NaCl. (**c**) A model for nucleosome disassembly through the intermediate "butterfly" state. $(H3–H4)_2$ tetramers are shown in *gray*, H2A–H2B dimers in *yellow* and DNA in *black*

(Fig. 7). From the variation of FRET efficiency with NaCl concentration a new intermediate state could be concluded where all histones were still associated with the DNA but the nucleosome core was opened up at the dimer:tetramer interface. In a follow-up study dissociation through this intermediate state was observed for different DNA sequences and histone protein origins; it thus appears to be a general hallmark of nucleosome disassembly (Tóth et al. 2013). Additional support for such a conformation with partially disrupted dimer-tetramer contacts was recently provided by time resolved small-angle X-ray scattering (SAXS) data (Chen et al. 2014).

Ngo et al. provided evidence for the existence of an alternative structure, where the two halves of the protein core transiently open in a direction perpendicular to the symmetry axis ("gaping") (Ngo and Ha 2015). Existence of this clam shell-like motion has been predicted theoretically (Mozziconacci and Victor 2003) but evaded experimental validation until now. The time scale of this gaping motion is several of minutes, much slower than DNA unwrapping in the entry/exit region.

In vivo, DNA topology also impacts nucleosome disassembly and genetic activity. Cellular chromatin can experience transient torsional stress, in particular

during transcription. Positive superhelical torsion is generated in front of and negative supercoiling behind the polymerase, as described by the "twin-supercoiled domain model" (Liu and Wang 1987). It has been observed that during transcription nucleosomes are displaced downstream and re-loaded upstream from the polymerase (Clark and Felsenfeld 1992), often accompanied by partial or complete histone loss (Kireeva et al. 2002; Hodges et al. 2009), a process that is promoted by supercoiling. Until recently, it was not clear whether this was directed by the topological change in DNA or by torsion-related effects on associated proteins. While earlier ultracentrifugation studies suggested that positive or negative supercoiling had no effect on DNA compaction by nucleosome formation (Clark et al. 1993), two more recent studies now conclusively proved that positive superhelical torsion can indeed destabilize nucleosomes. Elbel and Langowski studied the effect of superhelical density on the dissociation of nucleosomes assembled on plasmid DNA. Combining scanning force microscopy and FCS data they found direct proof that positive torsion promotes dimer loss (Elbel and Langowski 2015). Sheinin et al. arrived at the same conclusion through indirect observation of H2A/H2B dimer loss in mechanic force experiments (Sheinin et al. 2013).

The kinetics of nucleosome disassembly have only recently come into focus. Gansen et al. characterized the kinetics of salt-induced disassembly on a minute time scale (Gansen et al. 2009b). Intermediate structures over similar time scales have also been found in recent spFRET experiments by Plavner-Hazan et al. (Hazan et al. 2015). Using multiparameter single molecule fluorescence techniques (Antonik et al. 2006; Kalinin et al. 2010), rapid structural fluctuations in nucleosomes have become accessible; an analysis of nucleosomes labeled at the DNA with donor and acceptor fluorophores shows a rapid interconversion between partially disrupted conformations on the tens of microseconds time scale, where one H2A/H2B dimer is lost and the DNA labels are closer together (Gansen et al. manuscript in preparation).

8 Regulation of Nucleosome Dynamics Through Histone Tails and Posttranslational Modifications

Canonical nucleosomes whose DNA or protein content has not been modified provide a pure model system to study the dynamic landscape of nucleosomes. In the biological context, however, nucleosomes often carry chemical modifications on histones and DNA, which can affect their dynamic properties and have been recognized in the last 15–20 years as a key element in the regulation of genetic activity. They can affect chromatin morphology by altering electrostatic DNA-protein and protein-protein interactions, nucleosome stability and nucleosome-nucleosome interactions and constitute binding epitopes for chromatin-associated factors that can further modulate genetic activity. Thus, it is not surprising that

the effect of PTMs on nucleosome structure and dynamics has become a very active field of research. For a recent review, see (Lawrence et al. (2016)).

Many PTMs target the unstructured histone tails, which comprise about 30 % of the total histone mass. These regions feature a large number of basic residues, such as lysines or arginines, which can be acetylated or methylated. Histone tails and their modification directly alter the properties of nucleosomes; they can interact with nucleosomal DNA, restraining its position, and mediate nucleosome-nucleosome interactions in chromatin arrays. Partial removal of tails has been observed in promoter regions and simultaneous removal of the H3 and H4 tail increased site exposure to nuclear proteins (Polach et al. 2000). Ferreira et al. observed that removal of the H2B and H4 tails synergistically reduced thermal mobility, while clipping of H3 and H2A tails had the opposite effect. Removal of the H3 tail also facilitated H2A/H2B dimer eviction and opened up the nucleosome ends as measured by FRET (Ferreira et al. 2007).

The exact interplay of the long N-terminal tails of H3 and H4 is still a matter of debate. Crosstalk of PTMs between both tails has been observed (Fischle et al. 2003), where the presence of PTMs on one tail can affect the deposition of modifications on the other tail. Modeling studies correlated histone tail removal with increased DNA unwrapping (Voltz et al. 2012) and structural transitions in the nucleosome core (Biswas et al. 2011). These predictions were supported by a recent bulk FRET study by Nurse et al. (2013), who suggested that simultaneous removal of H3 and H4 tails increases DNA unwrapping compared to clipping either tail individually. Experiments on chemically acetylated histone tails, however, challenged the idea of a purely synergistic interplay between H3 and H4. Earlier bulk FRET studies from our lab analyzed the effect of selective histone acetylation (Toth et al. 2006). It was shown that H3 acetylation as well as acetylation of all histones led to an increase of DNA end-to-end distance but H4 acetylation to a decrease. The same tendency was later confirmed for trinucleosomes by FRET and AFM studies (Bussiek et al. 2006). Gansen et al. presented a detailed analysis of the role of histone acetylation on DNA unwrapping and disassembly. Using a combination of "microplate-scanning FRET" and spFRET the authors further established an antagonistic role of H3 and H4 acetylation in nucleosome stability (Gansen et al. 2015), see Fig. 8. Overall, the effect of histone tails on stability is more subtle than for example the effect of DNA sequence, but can provide regulatory fine tuning of nucleosome stability. The special role of histone H4 that emerged from these studies warrants further investigation into its role in nucleosome architecture.

Several recent studies provided evidence that individual modifications can directly affect nucleosome dynamics and stability. One of the best understood modifications is acetylation of lysine 56 in histone H3. H3K56 is located on the αN-helix of H3 and can interact with the nucleosomal DNA ends. Acetylation of H3K56 was shown to enhance DNA unwrapping significantly for the outermost 30 bp, rendering DNA more accessible in the entry/exit region (Neumann et al. 2009). A destabilizing role of acetylation of nearby lysine 64 in histone H3 has recently been observed by Di Cerbo et al. (2014), Simon et al. systematically investigated the role of specific residues in the histone-fold region of the

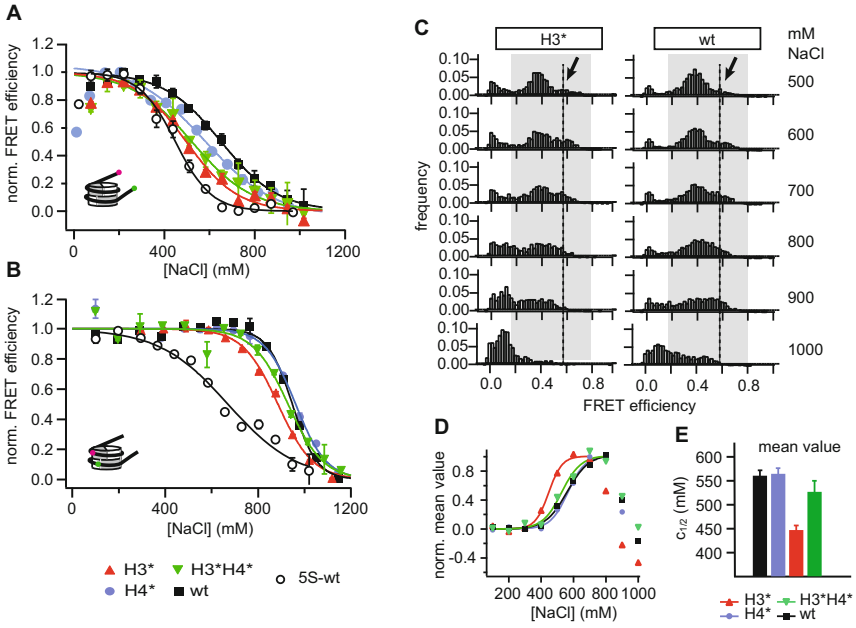

Fig. 8 Opposing roles of H3- and H4-acetylation in nucleosome stability. Salt-induced DNA unwrapping (**a**) and disassembly (**b**) of selectively acetylated nucleosomes measured with microplate-scanning FRET. DNA unwrapping and disassembly are dominated by H3-acetylation (H3*). Acetylation of histone H4 (H4*) increased unwrapping but does not affect nucleosome dissociation. Simultaneous acetylation of H3 and H4 (H3*H4*) induced less unwrapping and increased stability compared to H3-acetylation alone. For visualization, all curves were normalized between 1 at the maximum value and 0 at high ionic strength. Salt-titration curves of 5S non-acetylated nucleosomes are shown for comparison (open circles). (**c**) Exemplary single molecule distributions of H3-acetylated and non-acetylated nucleosomes at 500, 600, 700, 800, 900 and 1000 mM NaCl (*from top to bottom*). The same transition into a high-FRET intermediate prior to disassembly (*dashed line*) was observed for both samples. The salt concentration at which this transition is observed, however, depended on histone acetylation. (**d**) Normalized mean value of the nucleosome subpopulation (*gray region*) as a function of NaCl concentration. The increase in mean value is caused by formation of the high-FRET state. For quantification, data were approximated by a Sigmoid function. (**e**) Midpoint value ($c_{1/2}$) of the Sigmoidal fit as a function of histone acetylation. Formation of the high-FRET population correlates with the change in stability observed in bulk FRET; strongest effect was observed for H3-acetylation, while acetylation of H4 partially counteracts the effect of acetylated H3 in H3*H4* samples

nucleosome (Simon et al. 2011). Chemical ligation was used to introduce modifications at precise positions in the entry/exit region (H3K56ac, H4K77ac and H4K79ac) and near the dyad axis (H3K115ac and H3K122ac). A combination of magnetic force spectroscopy and FRET revealed that PTMs near the dyad axis selectively affected nucleosome disassembly without impacting unwrapping at the ends, while the opposite was observed for modifications in the entry/exit region. These studies imply that histone-DNA interactions are decoupled to selectively

modulate different dynamic processes within the nucleosome. Similar results were recently obtained in our lab, where a conservative mutation of two arginine residues in histone H2A located at the interface between dimer and tetramer dramatically reduced nucleosome stability and promoted formation of intermediate structures during disassembly (Tegeler et al., manuscript in preparation).

Besides directly impacting nucleosome stability many PTMs represent recognition sites for effector proteins. This way they can indirectly take action on genetic activity. A recent study by Musselman et al. combined NMR and FRET to show that binding of the protein PHF1 to trimethylated K36 in histone H3 increased the rate of DNA breathing and enhanced the binding of a second protein factor to its binding site inside the nucleosome (Musselman et al. 2013).

PTMs do not only occur on histone proteins: methylation of DNA has been associated with gene silencing. On the level of single nucleosomes, extensive methylation of 30 sites in a NCP leads to small changes in FRET distributions that indicate an increase in compaction and tighter DNA wrapping in the entry/exit region (Lee and Lee 2012). However, the effect was surprisingly small, given the large number of modifications present, suggesting that methylation has only minor influence on nucleosome structure, and an effect of less dense methylation on nucleosome structure is yet to be seen.

9 Towards Higher-Order Chromatin: FRET Studies on Di- and Trinucleosomes

In vivo, nucleosomes are part of a complex DNA-protein network, where the presence of neighboring nucleosomes and the crowded intranuclear environment could significantly modulate nucleosome properties. Earlier studies using bulk FRET on trinucleosomes reconstitutes on end-labeled 600 bp DNA suggested that the structural changes due to salt concentration, acetylation and linker histone binding are similar to that observed in mononucleosome samples (Bussiek et al. 2006). Again, single molecule experiments could further advance our understanding of oligonucleosome structure. Dinucleosomes have been used as templates for spFRET studies on the remodeling factor ACF (Hwang et al. 2014), to analyze the effect of H4-acetylation on internucleosome interactions (Lee et al. 2011) or the effect of flanking nucleosomes on DNA unwrapping in the linker DNA region (Buning et al. 2015). Several FRET studies probed the dynamic behavior of nucleosomes in small trinucleosome arrays. Poirier et al. (2009) showed by a combination of FRET and FCS that the rate of DNA unwrapping and protein binding to buried nucleosomal DNA is not significantly changed by the presence of flanking nucleosomes on either side. In a related study, it was then reported that flanking nucleosomes promote dissociation of the transcription factor, once bound to its binding site on the central nucleosome (Luo et al. 2014). However, a recent spFRET study by Buning et al. suggests that DNA unwrapping is enhanced by the

presence of a second nucleosome nearby (Buning et al. 2015). Further experiments will thus be needed to clarify, how the behavior of nucleosomes within an array compares to that observed for isolated complexes.

10 Conclusion

Over the last decades, FRET studies and in particular single molecule methods have provided enormous new insight into the dynamic properties of nucleosomes, and their consequences on chromatin organization. Single molecule mechanical studies converged with spFRET for a uniform picture of DNA wrapping and unwrapping, nucleosome disassembly and the effect of post-transcriptional modifications on these mechanisms.

References

Andrews AJ, Luger K (2011) Nucleosome structure(s) and stability: variations on a theme. Annu Rev Biophys 40:99–117

Antonik M, Felekyan S, Gaiduk A, Seidel CA (2006) Separating structural heterogeneities from stochastic variations in fluorescence resonance energy transfer distributions via photon distribution analysis. J Phys Chem B Condens Matter Mater Surf Interfaces Biophys 110 (13):6970–6978

Arimura Y, Tachiwana H, Oda T, Sato M, Kurumizaka H (2012) Structural analysis of the hexasome, lacking one histone H2A/H2B dimer from the conventional nucleosome. Biochemistry 51(15):3302–3309

Bednar J et al (1998) Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. Proc Natl Acad Sci U S A 95(24):14173–14178

Biswas M, Voltz K, Smith JC, Langowski J (2011) Role of histone tails in structural stability of the nucleosome. PLoS Comput Biol 7(12), e1002279

Biswas M, Langowski J, Bishop TC (2013) Atomistic simulations of nucleosomes. Wiley Interdiscip Rev Comput Mol Sci 3(4):378–392

Böhm V et al (2011) Nucleosome accessibility governed by the dimer/tetramer interface. Nucleic Acids Res 39(8):3093–3102

Buning R, Kropff W, Martens K, & van Noort J (2015) spFRET reveals changes in nucleosome breathing by neighboring nucleosomes. J Phys Condens Matter Inst Phys J 27(6):064103.

Bussiek M, Toth K, Schwarz N, Langowski J (2006) Trinucleosome compaction studied by fluorescence energy transfer and scanning force microscopy. Biochemistry 45 (36):10838–10846

Chen Y et al (2014) Revealing transient structures of nucleosomes as DNA unwinds. Nucleic Acids Res 42(13):8767–8776

Choy JS et al (2010) DNA methylation increases nucleosome compaction and rigidity. J Am Chem Soc 132(6):1782–1783

Cisse I, Okumus B, Joo C, Ha T (2007) Fueling protein DNA interactions inside porous nanocontainers. Proc Natl Acad Sci U S A 104(31):12646–12650

Clark DJ, Felsenfeld G (1992) A nucleosome core is transferred out of the path of a transcribing polymerase. Cell 71(1):11–22

Clark DJ, Ghirlando R, Felsenfeld G, Eisenberg H (1993) Effect of positive supercoiling on DNA compaction by nucleosome cores. J Mol Biol 234(2):297–301
Claudet C, Angelov D, Bouvet P, Dimitrov S, Bednar J (2005) Histone octamer instability under single molecule experiment conditions. J Biol Chem 280(20):19958–19965
Clegg RM (1992) Fluorescence resonance energy transfer and nucleic acids. Methods Enzymol 211:353–388
Clegg RM (1996) Fluorescence resonance energy transfer. In: Wang XF, Herman B (eds) Fluorescence imaging spectroscopy and microscopy. Wiley, New York, pp 179–252
Davey CA, Sargent DF, Luger K, Maeder AW, Richmond TJ (2002) Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. J Mol Biol 319 (5):1097–1113
Di Cerbo V et al (2014) Acetylation of histone H3 at lysine 64 regulates nucleosome dynamics and facilitates transcription. eLife 3, e01632
Elbel T, Langowski J (2015) The effect of DNA supercoiling on nucleosome structure and stability. J Phys Condens Matter Inst Phys J 27(6):064105
Felekyan S, Kalinin S, Sanabria H, Valeri A, Seidel CA (2012) Filtered FCS: species auto- and cross-correlation functions highlight binding and dynamics in biomolecules. Chemphyschem 13(4):1036–1053
Ferreira H, Somers J, Webster R, Flaus A, Owen-Hughes T (2007) Histone tails and the H3 alphaN helix regulate nucleosome mobility and stability. Mol Cell Biol 27(11):4037–4048
Fischle W, Wang Y, Allis CD (2003) Histone and chromatin cross-talk. Curr Opin Cell Biol 15 (2):172–183
Flaus A, Rencurel C, Ferreira H, Wiechens N, Owen-Hughes T (2004) Sin mutations alter inherent nucleosome mobility. EMBO J 23(2):343–353
Förster T (1946) Energiewanderung und Fluoreszenz. Naturwissenschaften 6:166–175
Fussner E, Ching RW, Bazett-Jones DP (2011) Living without 30 nm chromatin fibers. Trends Biochem Sci 36(1):1–6
Gansen A, Hauger F, Toth K, Langowski J (2007) Single-pair fluorescence resonance energy transfer of nucleosomes in free diffusion: optimizing stability and resolution of subpopulations. Anal Biochem 368(2):193–204
Gansen A, Toth K, Schwarz N, Langowski J (2009a) Structural variability of nucleosomes detected by single-pair forster resonance energy transfer: histone acetylation, sequence variation, and salt effects. J Phys Chem B 113(9):2604–2613
Gansen A et al (2009b) Nucleosome disassembly intermediates characterized by single-molecule FRET. Proc Natl Acad Sci U S A 106(36):15308–15313
Gansen A, Hieb AR, Bohm V, Toth K, Langowski J (2013) Closing the gap between single molecule and bulk FRET analysis of nucleosomes. PLoS One 8(4), e57018
Gansen A, Toth K, Schwarz N, Langowski J (2015) Opposing roles of H3- and H4-acetylation in the regulation of nucleosome structure—a FRET study. Nucleic Acids Res 43(3):1433–1443
Gottesfeld JM, Luger K (2001) Energetics and affinity of the histone octamer for defined DNA sequences. Biochemistry 40(37):10927–10933
Hall MA et al (2009) High-resolution dynamic mapping of histone-DNA interactions in a nucleosome. Nat Struct Mol Biol 16(2):124–129
Harp JM et al (1996) X-ray diffraction analysis of crystals containing twofold symmetric nucleosome core particles. Acta Crystallogr Sect D Biol Crystallogr 52(Part 2):283–288
Hazan NP et al (2015) Nucleosome core particle disassembly and assembly kinetics studied using single-molecule fluorescence. Biophys J 109(8):1676–1685
Hodges C, Bintu L, Lubkowska L, Kashlev M, Bustamante C (2009) Nucleosomal fluctuations govern the transcription dynamics of RNA polymerase II. Science 325(5940):626–628
Hwang WL, Deindl S, Harada BT, Zhuang X (2014) Histone H4 tail mediates allosteric regulation of nucleosome remodelling by linker DNA. Nature 512(7513):213–217
Ivanov V, Li M, Mizuuchi K (2009) Impact of emission anisotropy on fluorescence spectroscopy and FRET distance measurements. Biophys J 97(3):922–929

Jin F et al (2013) A high-resolution map of the three-dimensional chromatin interactome in human cells. Nature 503(7475):290–294
Kalinin S, Valeri A, Antonik M, Felekyan S, Seidel CA (2010) Detection of structural dynamics by FRET: a photon distribution and fluorescence lifetime analysis of systems with multiple states. J Phys Chem B 144:7983–7995
Kalinin S et al (2012) A toolkit and benchmark study for FRET-restrained high-precision structural modeling. Nat Methods 9(12):1218–1225
Kapanidis AN et al (2004) Fluorescence-aided molecule sorting: analysis of structure and interactions by alternating-laser excitation of single molecules. Proc Natl Acad Sci U S A 101 (24):8936–8941
Kelbauskas L et al (2008) Sequence-dependent variations associated with H2A/H2B depletion of nucleosomes. Biophys J 94(1):147–158
Kireeva ML et al (2002) Nucleosome remodeling induced by RNA polymerase II: loss of the H2A/H2B dimer during transcription. Mol Cell 9(3):541–552
Koopmans WJ, Brehm A, Logie C, Schmidt T, van Noort J (2007) Single-pair FRET microscopy reveals mononucleosome dynamics. J Fluoresc 17(6):785–795
Koopmans WJ, Schmidt T, van Noort J (2008) Nucleosome immobilization strategies for single-pair FRET microscopy. Chemphyschem 9(14):2002–2009
Koopmans WJ, Buning R, Schmidt T, van Noort J (2009) spFRET using alternating excitation and FCS reveals progressive DNA unwrapping in nucleosomes. Biophys J 97(1):195–204
Kudryavtsev V et al (2012) Combining MFD and PIE for accurate single-pair Forster resonance energy transfer measurements. Chemphyschem 13(4):1060–1078
Lawrence M, Daujat S, Schneider R (2016) Lateral thinking: how histone modifications regulate gene expression. Trends Genet 32(1):42–56
Lee JY, Lee TH (2012) Effects of DNA methylation on the structure of nucleosomes. J Am Chem Soc 134(1):173–175
Lee JY, Wei S, Lee TH (2011) Effects of histone acetylation by Piccolo NuA4 on the structure of a nucleosome and the interactions between two nucleosomes. J Biol Chem 286 (13):11099–11109
Li G, Levitus M, Bustamante C, Widom J (2005) Rapid spontaneous accessibility of nucleosomal DNA. Nat Struct Mol Biol 12(1):46–53
Liu LF, Wang JC (1987) Supercoiling of the DNA template during transcription. Proc Natl Acad Sci U S A 84:7024–7027
Lowary PT, Widom J (1998) New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. J Mol Biol 276(1):19–42
Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature 389:251–260
Luger K, Dechassa ML, Tremethick DJ (2012) New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? Nat Rev Mol Cell Biol 13(7):436–447
Luo Y, North JA, Poirier MG (2014) Single molecule fluorescence methodologies for investigating transcription factor binding kinetics to nucleosomes and DNA. Methods 70(2–3):108–118
McDowall AW, Smith JM, Dubochet J (1986) Cryo-electron microscopy of vitrified chromosomes in situ. EMBO J 5(6):1395–1402
Menshikova I, Menshikov E, Filenko N, Lyubchenko YL (2011) Nucleosomes structure and dynamics: effect of CHAPS. Int J Biochem Mol Biol 2(2):129–137
Moerner WE, Fromm DP (2003) Methods of single-molecule fluorescence spectroscopy and microscopy. Rev Sci Instrum 74(8):3597–3619
Mozziconacci J, Victor J-M (2003) Nucleosome gaping supports a functional structure for the 30 nm chromatin fiber. J Struct Biol 143(1):72–76
Müller BK, Zaychikov E, Brauchle C, Lamb DC (2005) Pulsed interleaved excitation. Biophys J 89(5):3508–3522
Musselman CA et al (2013) Binding of PHF1 Tudor to H3K36me3 enhances nucleosome accessibility. Nat Commun 4:2969

Neumann H et al (2009) A method for genetically installing site-specific acetylation in recombinant histones defines the effects of H3 K56 acetylation. Mol Cell 36(1):153–163
Ngo TT, Ha T (2015) Nucleosomes undergo slow spontaneous gaping. Nucleic Acids Res 43 (8):3964–3971
Nir E et al (2006) Shot-noise limited single-molecule FRET histograms: comparison between theory and experiments. J Phys Chem B 110(44):22103–22124
North JA et al (2012) Regulation of the nucleosome unwrapping rate controls DNA accessibility. Nucleic Acids Res 40(20):10215–10227
Nurse NP, Jimenez-Useche I, Smith IT, Yuan C (2013) Clipping of flexible tails of histones H3 and H4 affects the structure and dynamics of the nucleosome. Biophys J 104(5):1081–1088
Olins AL, Olins DE (1974) Spheroid chromatin units (v bodies). Science 183(122):330–332
Perrin J (1927) Fluorescence et induction moleculaire par resonance. C R Hebd Seances Acad Sci 184:1097–1100
Poirier MG, Oh E, Tims HS, Widom J (2009) Dynamics and function of compact nucleosome arrays. Nat Struct Mol Biol 16(9):938–944
Polach KJ, Widom J (1995) Mechanism of protein access to specific DNA sequences in chromatin: a dynamic equilibrium model for gene regulation. J Mol Biol 254(2):130–149
Polach KJ, Lowary PT, Widom J (2000) Effects of core histone tail domains on the equilibrium constants for dynamic DNA site accessibility in nucleosomes. J Mol Biol 298(2):211–223
Prinsen P, Schiessel H (2010) Nucleosome stability and accessibility of its DNA to proteins. Biochimie 92(12):1722–1728
Roy R, Hohng S, Ha T (2008) A practical guide to single-molecule FRET. Nat Methods 5 (6):507–516
Segal E et al (2006) A genomic code for nucleosome positioning. Nature 442(7104):772–778
Sheinin MY, Li M, Soltani M, Luger K, Wang MD (2013) Torque modulates nucleosome stability and facilitates H2A/H2B dimer loss. Nat Commun 4:2579
Simon M et al (2011) Histone fold modifications control nucleosome unwrapping and disassembly. Proc Natl Acad Sci U S A 108(31):12711–12716
Strahl BD, Allis CD (2000) The language of covalent histone modifications. Nature 403(6765): 41–45
Thastrom A et al (1999) Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences. J Mol Biol 288(2):213–229
Thastrom A, Gottesfeld JM, Luger K, Widom J (2004) Histone-DNA binding free energy cannot be measured in dilution-driven dissociation experiments. Biochemistry 43(3):736–741
Tims HS, Gurunathan K, Levitus M, Widom J (2011) Dynamics of nucleosome invasion by DNA binding proteins. J Mol Biol 411(2):430–448
Tomschik M, Zheng H, van Holde K, Zlatanova J, Leuba SH (2005) Fast, long-range, reversible conformational fluctuations in nucleosomes revealed by single-pair fluorescence resonance energy transfer. Proc Natl Acad Sci U S A 102(9):3278–3283
Tomschik M, van Holde K, Zlatanova J (2009) Nucleosome dynamics as studied by single-pair fluorescence resonance energy transfer: a reevaluation. J Fluoresc 19(1):53–62
Torres T, Levitus M (2007) Measuring conformational dynamics: a new FCS-FRET approach. J Phys Chem B 111(25):7392–7400
Toth K, Brun N, Langowski J (2001) Trajectory of nucleosomal linker DNA studied by fluorescence resonance energy transfer. Biochemistry 40(23):6921–6928
Toth K, Brun N, Langowski J (2006) Chromatin compaction at the mononucleosome level. Biochemistry 45(6):1591–1598
Tóth K et al (2013) Histone- and DNA sequence-dependent stability of nucleosomes studied by single-pair FRET. Cytometry A 83(9):839–846
Vence T (2015) Crystal Unclear. The Scientist (Oct 15). http://www.the-scientist.com/?articles.view/articleNo/44196/title/Crystal-Unclear/
Voltz K, Trylska J, Calimet N, Smith JC, Langowski J (2012) Unwrapping of nucleosomal DNA ends: a multiscale molecular dynamics study. Biophys J 102(4):849–858

Widengren J et al (2006) Single-molecule detection and identification of multiple species by multiparameter fluorescence detection. Anal Chem 78(6):2039–2050

Widom J, Klug A (1985) Structure of the 300A chromatin filament: X-ray diffraction from oriented samples. Cell 43(1):207–213

Yager TD, McMurray CT, van Holde KE (1989) Salt-induced release of DNA from nucleosome core particles. Biochemistry 28(5):2271–2281

Part VII
Chromatin Transactions and Epigenetics

Mapping and Visualizing Spatial Genome Organization

Christopher J.F. Cameron, James Fraser, Mathieu Blanchette, and Josée Dostie

Abstract Non-random genome organization and its correlation with various nuclear activities point to the existence of functional chromatin folding and positioning patterns. The importance of genome architecture is highlighted in disease states where altered chromatin organization can be found. These studies have provided insight into what regulates genome architecture and how this contributes to human disease. While much progress has been made towards defining chromosome organization in various cell types and states, the functional consequences of organization patterns remain poorly understood. The impact of chromatin folding and position within the nucleus appears to depend on nuclear context and neighbouring sequences. Thus, to understand the relevance of chromosome position and movement with respect to transcription and other activities, chromatin organization must be mapped across genomic scales in various nuclear contexts. Several techniques are available to map chromosome folding at high-resolution. These include the family of chromosome conformation capture (3C) technologies, which use the frequency of chromatin contacts as a measure of physical proximity. The 3C techniques differ in their genomic coverage with the genome-wide approaches relying more extensively on informatics to organize and visualize chromosome architecture. This book chapter provides an overview of the 3C technologies and explains how the data is used to infer genome organization.

Keywords Chromatin/genomic organization • 3D modelling • Deep sequencing • Genome architecture • Transcription • Long-range interactions

C.J.F. Cameron • M. Blanchette
McGill Center for Biochemistry and School of Computer Science, McGill University, Montréal, QC, Canada, H3B 0B1

J. Fraser • J. Dostie (✉)
Department of Biochemistry and Goodman Cancer Research Center, McGill University, McIntyre Medical Sciences Building, 3655 Promenade Sir-William-Osler, Room 815A, Montréal, QC, Canada, H3G 1Y6
e-mail: josee.dostie@mcgill.ca

D.P. Bazett-Jones, G. Dellaire (eds.), *The Functional Nucleus*,
DOI 10.1007/978-3-319-38882-3_16

1 Introduction

Even when strictly considered from a biomechanical perspective, chromosome folding is essential in eukaryotes because of the very large size difference between genomes and nuclei. The human genome, for instance, consists of 46 chromosomes with a combined linear length of over 2 m that must fit inside micron-scale nuclei. It thus stands to reason that genomic DNA must be packaged, at the very least to reduce chromosome length. As the first level of packaging, the DNA wraps around assemblies of histone octamers composed of two copies of each H2A, H2B, H3, and H4 core histone (Andrews and Luger 2011). This packaging forms the nucleosome, which represents the basic repeating unit in chromatin. There are two tight superhelical DNA turns (147 base pairs in length) to each nucleosome core particle, and this effectively shortens chromosomes by sevenfold.

Under physiological conditions, chromatin is mainly composed of DNA and histones, but also contains a very large repertoire of non-histone proteins. These can either be recruited directly by DNA binding or indirectly through other proteins. An important mechanism by which chromatin composition and activity are regulated involves post-translational modification (PTM) of histones. PTMs such as acetylation, methylation, phosphorylation, sumoylation, ADP-ribosylation, and ubiquitinylation have been found on multiple histones residues, particularly on their flexible amino- and carboxy-terminal tails. Histone PTMs can modify access to specific DNA sequences by changing the way nucleosomes interact with each other and/or with the DNA. Histone marks are also recognized by many different protein complexes that can modify the landscape and activity of chromatin.

Several chromatin-binding proteins are known to interact with each other even over long distances and form so-called "chromatin loops" that represent a second level of packaging (Fig. 1). Long-range interactions have been found between proteins like CTCF, cohesin, Mediator, and transcription factors, and can either promote or restrict the physical proximity between control DNA elements (Sanyal et al. 2012). For example, the insulator-binding protein CTCF (CCCTC-binding factor) was shown to mediate chromatin loops that either prevent (Mishiro et al. 2009) or promote (Stadhouders et al. 2012) contacts between active enhancers and genes. Some interactions, on the other hand, like those between Mediator and cohesin, mostly enhance physical proximity between elements because they tend to bridge enhancers to promoters. These types of contacts can be highly tissue-specific in nature, while others such as those between CTCF and cohesin most often associate with invariant conformations (Phillips-Cremins et al. 2013).

The fact that control DNA elements can physically interact with distal genes explains how they can regulate them, and why the functional organization of genomes (i.e., the physical order of genes and control sequences) is not strictly linear along chromosomes. Although long-range interactions can occur over very large distances and even between chromosomes (Kleinjan and van Heyningen 2005; Ling et al. 2006; Lomvardas et al. 2006; West and Fraser 2005), most chromatin loops are likely to bridge regions spaced by <1 Mb because chromatin

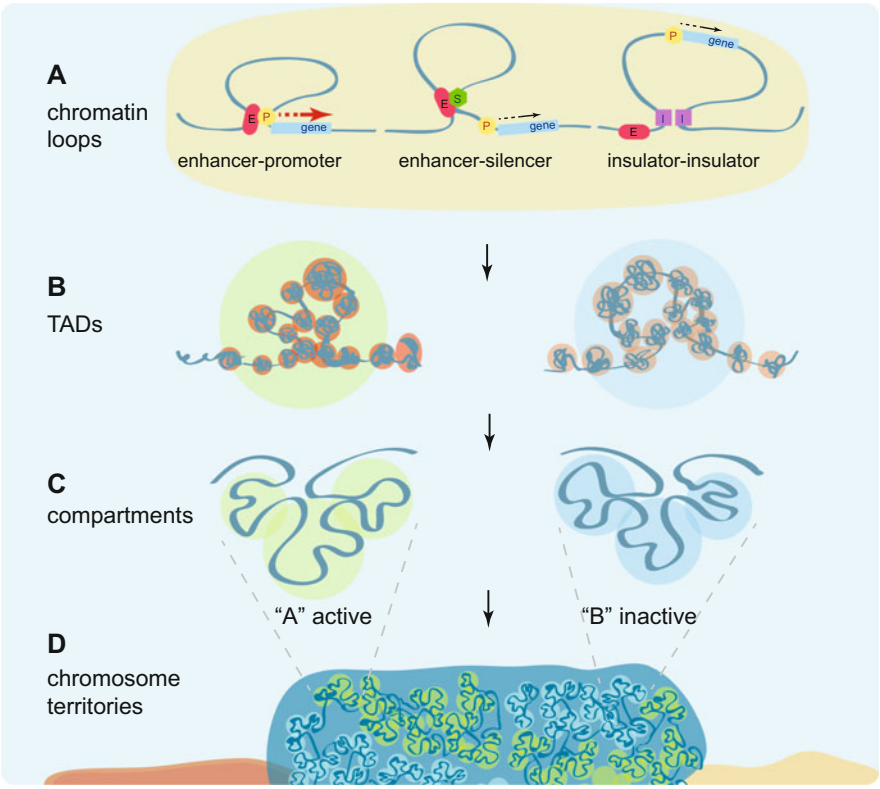

Fig. 1 Chromosome folding at different size scales. The four main levels of chromatin organization identified in mammals are shown from high (*top*) to low (*bottom*) resolution. (**a**) Three examples of chromatin looping structures are shown. Enhancer–promoter contacts (*left*) lead to active transcription represented by a *thick red arrow*. Chromatin loops can also lead to transcription inhibition as exemplified by enhancer–silencer contacts (*middle*) and insulator loops (*right*) that prevent enhancer–promoter interactions. Inactive transcription is shown as *thin black arrows*. *E* enhancer; *P* promoter; *S* silencer; *I* insulator. (**b**) Chromatin folds into topologically associating domains (TADs) at the megabase scale. TADs are contiguous blocks of chromatin along chromosomes that are enriched in chromatin contacts including long-range interactions that can form substructures (sub-TADs). (**c**) TADs are contained within compartments, which are multi-megabase structures that are either enriched in transcribed genes (A compartments) or heterochromatin (B compartments). (**d**) Compartments of the same type preferentially interact with each other to form chromosome territories

was shown to fold into higher-order structures named "topologically associating domains" (TADs) averaging 0.5–1 Mb in size (Fig. 1) (Dixon et al. 2012; Nora et al. 2012). TADs are regions characterized by locally enriched chromatin contacts that are observed throughout chromosomes. They have been found in many different mammalian cell types and species, and thus appear to represent an inherent feature of genome organization. Long-range interactions between control DNA elements within TADs can sometimes form substructures or "sub-TADs"

(Phillips-Cremins et al. 2013). TADs were suggested to compartmentalize chromosomes into functionally distinct domains. Accordingly, one study shows that disrupting a TAD boundary causes long-range transcriptional misregulation associated with the formation of ectopic chromatin contacts (Nora et al. 2012).

At the multi-megabase scale, chromatin is packaged further into "compartments" (Fig. 1) (Lieberman-Aiden et al. 2009). At least two types of compartments exist: Compartments "A" are associated with open chromatin (DNaseI hypersensitivity), are enriched in genes, highly transcribed, have a high GC content and more histone marks associated with active (H3K36me3) and poised (H3K27me3) chromatin. "B" compartments in contrast, are enriched in the silencing H3K9me3 histone modification, are more compact (higher interaction frequencies), and show a greater tendency towards self-association. The partitioning of chromosomes into compartments is pervasive across mammalian cell types and species, suggesting that it too represents a basic feature of genome organization, although the fact that compartments can vary also points to a regulatory role.

Beyond the compartment level, chromosomes are known to fold into nuclear substructures known as chromosome territories (CTs) (Fig. 1) (Bolzer et al. 2005) (detailed in Meaburn et al., this book). Although chromatin contacts are enriched within CTs, chromosomes can also interact with each other. For instance, a significant amount of cell-specific intermingling between different chromosomes was previously reported (Branco and Pombo 2006; Visser et al. 2000). Interchromosomal contacts are also enriched in highly transcribed regions (Kalhor et al. 2012; Yaffe and Tanay 2011), which correlates well with the prevalence of exons at the surface of CTs (Boyle et al. 2011). Accordingly, gene-rich chromosomes in human lymphocytes were found to interact more frequently with each other than would be predicted based on their size (Lieberman-Aiden et al. 2009).

The current model of genome architecture outlined above raises many questions about the nature, function, and regulation of chromatin organization. For instance, a unifying model explaining how chromatin conformations relate to each other across hierarchical length scales is lacking. What chromatin folding patterns are important for proper gene expression? Which ones contribute to human disease? What is the relationship between DNA sequence variants, genome organization and gene expression? These questions can only be answered through a better understanding of the chromatin structure-function relationship, which will require further mapping of chromatin organization under various conditions and at different size scales. Three-dimensional chromatin organization can be studied with several methods including the family of chromosome conformation capture (3C) technologies. These approaches measure contact frequency between chromatin regions, and this information can be used to infer chromatin architecture by assuming that contact frequency reflects physical distances *in vivo*.

2 Mapping Genome Organization: The 3C Technologies

Three-dimensional (3D) chromosome organization is typically characterized by measuring distances between chromatin segments. Distances can either be measured directly by microscopy (Dostie and Bickmore 2012) or inferred based on the frequency at which the chromatin fragments interact. Chromatin interactions can be quantified with molecular techniques such as the 3C technologies, which capture contacts by proximity-ligation in chemically cross-linked cell populations (Fig. 2). This type of data can then be used to predict the average architecture of chromatin by considering that regions interact proportionally to their distance in the nucleus. There are multiple 3C-type methods, and the ideal 3C method will depend on the scale and resolution desired, whether the question is hypothesis-based, the budget, and the amount of sample available.

2.1 Chromosome Conformation Capture (3C)

The chromosome conformation capture (3C) approach developed by Dekker et al. was the founding method of this class of "next generation" spatial genome mapping technologies (Dekker et al. 2002). During 3C and in all other derived methods, chromatin contacts are converted into unique ligation products, which are quantified with a method of choice (Fig. 2, *middle*). Cells are first fixed with formaldehyde to create covalent bonds, and the cross-linked chromatin is digested with a restriction enzyme. The digested DNA is then ligated to create new ligation junctions that can be measured specifically using different polymerase chain reaction (PCR) methods. 3C products are conventionally quantified individually by PCR amplification of the predicted ligation junctions and migration on agarose gel. Detection can also be achieved by other methods such as TaqMan quantitative PCR or melting curve analysis with SYBR green I (Abou El Hassan and Bremner 2009; Dekker et al. 2002; Hagege et al. 2007). Regardless of the detection method selected, 3C remains a low-throughput approach that requires a large amount of material because contacts are quantified individually or "point-by-point". 3C is therefore used mainly for small-scale analysis, to characterize the organization of short regions at high-resolution or the behaviour of a few contacts such as chromatin loops.

3C was first used to characterize the yeast chromosome III conformation in *Saccharomyces cerevisiae* (Dekker 2003; Dekker et al. 2002), and confirmed that it adopts a shape consistent with its known Rabl-like organization (Bystricky et al. 2005; Rabl 1885; Spector 2003; Taddei et al. 2010). 3C was then applied to examine the mouse beta-globin cluster during erythroid differentiation (Palstra et al. 2003; Tolhuis et al. 2002). These studies demonstrated how higher-order chromatin structure might impinge on gene regulation by revealing physical contacts between the locus control region and actively transcribed genes of the cluster.

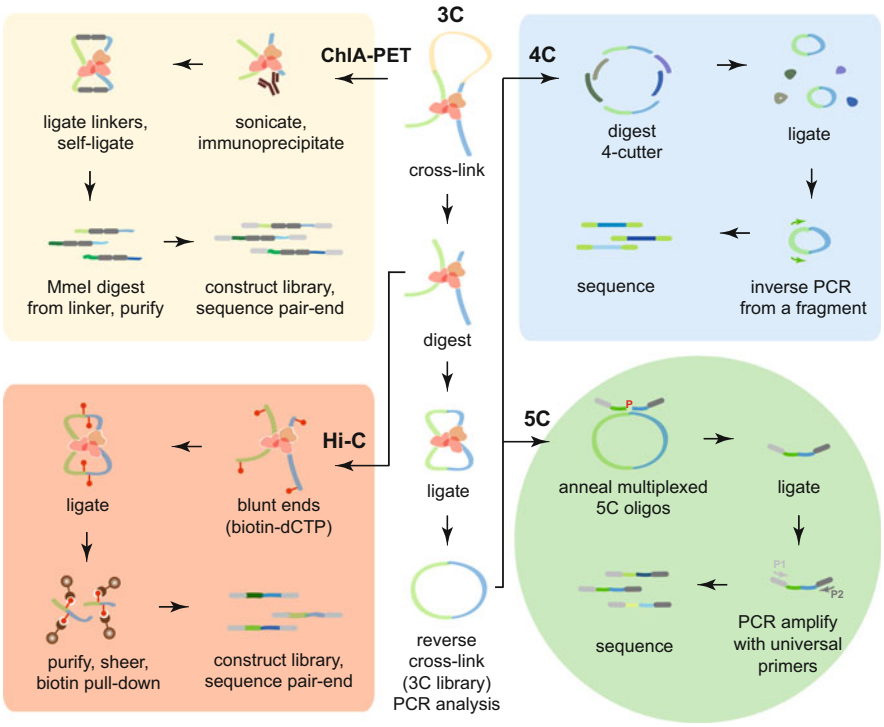

Fig. 2 Capturing chromatin contacts with 3C technologies and ChIA-PET. The founding 3C method is outlined from top to bottom in the *middle*. In all approaches, the chromatin is first chemically cross-linked to capture interactions between DNA regions (*blue* and *green lines*) stabilised by protein complexes (*coloured shapes*). The next step of 3C methods is the digestion of chromatin with a restriction fragment. This is in contrast to ChIA-PET, which uses sonication to solubilise chromatin prior to immunoprecipitation (*top, left*). The third step of 3C technologies is the ligation of free DNA ends to create 3C junctions between different restriction fragments. For 3C, the cross-links are then reversed, the DNA purified, and libraries are analyzed point-by-point with a preferred PCR method. The genome-scale 4C (*top, right*) and 5C (*bottom, right*) methods, each involve the transformation of 3C libraries into short PCR products that can be quantified by deep sequencing methods. These products are obtained by inverse PCR with primers against a bait region (*green arrows*) during 4C, and by ligating 5C primers at 3C junctions by ligation-mediated amplification (LMA) in 5C. The universal primer sequences at the ends of 5C primers are shown as grey lines. The red "P" is to indicate that reverse 5C primers are phosphorylated. The genome-wide ChIA-PET (*top, left*) and Hi-C (*bottom, left*) techniques each involve tagging and selection from the DNA fragment ends to enrich in contacts before deep sequencing. During ChIA-PET, the free DNA ends are ligated to linkers containing Mme I restriction sites, which are then used to release contacts from circularized products. Hi-C uses biotinylated nucleotides to mark contact junctions, which are enriched on Streptavidin beads. Biotin labels nucleotides are shown as *red dots* and Streptavidin beads are in *brown*. The Y-shaped molecule represents antibodies

The studies that followed revealed that long-range interactions regulate the activity of enhancers at promoters, insulator function, transcriptional silencing, imprinting, and X inactivation genome-wide [reviewed in (Ethier et al. 2012)]. It was also

found that chromatin contacts could be regulated by DNA methylation demonstrating a cross talk between genome architecture and epigenetic regulation (Murrell et al. 2004; Nativio et al. 2011). In fact, the epigenetic state of chromatin and its conformation are thought to mutually affect each other in self-enforcing structure-function feedback loops that are propagated from one generation to the next (Cavalli and Misteli 2013).

2.2 Chromosome Conformation Capture-On-Chip (4C)

The 4C techniques are a set of 3C-based methods that can be used to characterize the chromatin environment of given genomic regions without prior knowledge of what it might interact with. Although different groups independently developed four variations of 4C, the techniques all involve the transformation of 3C contacts from a specific "bait" region into very short ligation products that are either sequenced or quantified onto microarrays (Ling et al. 2006; Simonis et al. 2006; Wurtele and Chartrand 2006; Zhao et al. 2006). Of the four methods, the "Chromosome conformation capture-on-chip" version developed by the de Laat group was the only easily scalable approach, and is the one used today (Simonis et al. 2006; van de Werken et al. 2012a, b). During 4C, a 3C library is first produced and the ligation product are re-digested with a different restriction enzyme that cuts more frequently (Fig. 2, *top right*). The digested DNA is then re-ligated, and the resulting small circular products are used as template to simultaneously amplify all interacting fragments by inverse PCR from a given bait region. The amplicons, which were originally measured on custom microarrays, are now quantified by deep sequencing for greater sensitivity and scale. This 4C method generates high-resolution interaction profiles with the entire genome that remain unmatched by any other 3C method. However, because 4C only captures interactions from single viewpoints, the spatial organization of entire chromosomes or domains cannot be inferred from these data.

4C methods were first used to characterize the environment of the *HoxB1* gene in mouse ES cells (Wurtele and Chartrand 2006), identify regions contacting the mouse insulin-like growth factor (Igf2/H19) imprinting control region (Ling et al. 2006; Zhao et al. 2006), and study the active and inactive beta-globin locus conformation (Simonis et al. 2006). The 4C version that prevailed was then used in a very large number of studies, including those aimed at finding DNA elements controlling the expression of *Hox* genes during development (Andrey et al. 2013; Lonfat et al. 2014; Montavon et al. 2011; Noordermeer et al. 2011), interactions between Polycomb domains (Denholtz et al. 2013; Tolhuis et al. 2011), and the inactive X chromosome (Splinter et al. 2011).

2.3 Chromosome Conformation Capture Carbon Copy (5C)

5C is a 3C-based approach that can be used to map the three-dimensional organization of large chromatin domains, or the physical connectivity of numerous genomic regions simultaneously. This method combines the conventional 3C protocol with ligation-mediate amplification (LMA) to detect thousands of contacts simultaneously in one experiment (Dostie and Dekker 2007; Dostie et al. 2006). During 5C, a 3C library is first produced by the conventional protocol (Miele et al. 2006; Naumova et al. 2012), and the DNA is annealed to 5C primers in a multiplexed setting (Fig. 2, *bottom right*). 5C primers are designed to anneal at the 3′ end of restriction fragments such that when Taq DNA ligase is added to the mixture, only existing 3C junctions will be quantitatively represented in the 5C library. The 5C contacts are then amplified by PCR with primers against the universal 5C primer tails. The entire process results in the selective amplification of thousands of chromatin contacts simultaneously.

5C offers many advantages over 3C, including a much higher throughput, the need for less starting material given that contacts are amplified together in single tubes, better accuracy, and a greater dynamic detection range. Another significant advantage of 5C is that amplified 3C junctions are identified by sequence rather than expected size on gels or melting temperatures (see Section "Chromosome conformation capture (3C)" above). However, the fact that it uses primers to quantitatively detect contacts is limiting because it implies that target regions must first be selected before analysis.

5C has been used to probe different types of biological questions at various scales. It is an ideal approach to characterize the internal organization of TADs (sub-TADs) (Nora et al. 2012; Phillips-Cremins et al. 2013). Many studies have used 5C to study the organization of the human *HOX* clusters in various cell systems (Berlivet et al. 2013; Fraser et al. 2009; Wang et al. 2011), and to compare and contrast 5C data with information provided by fluorescence in situ hybridization (FISH) (Williamson et al. 2014). 5C was also used to characterize the inactive and active beta-globin (Dostie et al. 2006) and alpha-globin (Bau et al. 2011) clusters, the three-dimensional organization of mitotic chromosomes (Naumova et al. 2013), and the bacterial *Caulobacter crescentus* genome (Umbarger et al. 2011). The role of the CTCF and cohesin architectural proteins in sub-TAD structure maintenance was also probed with 5C as well as the spatial chromatin landscape of promoters (Sanyal et al. 2012). We have also used 5C data to model three-dimensional organization computationally (Fraser et al. 2010; Rousseau et al. 2011), and classify leukemia types based on chromatin conformation (Rousseau et al. 2014).

2.4 Genome-Wide Chromosome Conformation Capture (Hi-C)

The Hi-C technology can be used to probe the spatial organization of entire genomes. Like 3C, it captures *in vivo* chromatin contacts by proximity-ligation, but uses deep sequencing to quantify all ligation products present in libraries instead of individual contacts on gel (Fig. 2, *bottom left*) (Lieberman-Aiden et al. 2009). Generating Hi-C libraries is achieved in five general steps. The first two steps are identical between 3C and Hi-C and involve the chemical fixation of cell populations with formaldehyde and digestion of the fixed chromatin with a restriction enzyme. The third step is specific to Hi-C library production and consists in filling the ends of restriction fragments with Klenow and a mixture of dNTPs that include biotin-14-dCTP. This is followed with blunt end ligation with T4 DNA ligase, which joins DNA fragments that were close to each other in space. Cross-links are next reversed, yielding unprocessed Hi-C libraries that are sheared by sonication and enriched in ligation products on Streptavidin beads. Illumina Paired End adapters are finally added to the ends of Hi-C products and the sample is amplified by PCR before sequencing.

Unlike 3C, the amount of Hi-C data produced even from a single experiment is staggering. Depending on budget, hundreds of millions of sequencing reads are generated on average per Hi-C experiment with the current sequencing methods. Handling such large data sets therefore requires computational tools, some of which are already available (see Section "Processing, Organization, and Analysis of Chromatin Contacts"). The original Hi-C article presented conformational data for two human cell lines at one megabase resolution (Lieberman-Aiden et al. 2009). This study demonstrated the power of Hi-C by confirming the existence of key features like CTs and the proximity between gene-rich chromosomes, and uncovering the existence of the chromatin compartments described in the Introduction (Section "Introduction"). Hi-C (Dixon et al. 2012) and 5C (Nora et al. 2012) were also used to generate the chromosome interaction maps that revealed the presence of TADs. The complexity of Hi-C libraries from large genomes like those from mammals is very high and for this reason, chromatin organization is usually described at resolutions between 40 kb and 1 Mb. However, a recent study using a modified Hi-C protocol achieved kb-resolutions (Rao et al. 2014).

Hi-C data was used to correlate genome architecture with replication timing, and it was found that differential firing at origins could be explained by the compartmentalization of origins into units of three-dimensional chromatin architectures (Pope et al. 2014; Ryba et al. 2010). Hi-C was also used to show that the proximity of chromatin correlates well with the incidence of intra and inter-chromosomal translocations from double stranded DNA breaks in the mouse (Zhang et al. 2012).

2.5 Chromatin Interaction Analysis by Paired-End Tag Sequencing (ChIA-PET)

The ChIA-PET method is not a 3C-derived technique *per se*, but complements them well because it captures regions bound by specific proteins genome-wide (Fig. 2, *top left*) (Fullwood et al. 2009; Fullwood and Ruan 2009). This method therefore provides information about the nature of chromatin contacts, whereas Hi-C only identifies them. In ChIA-PET, the formaldehyde-fixed chromatin from cell populations is sonicated rather than digested with a restriction enzyme. Sonication is required to solubilize the fixed complexes prior to immunoprecipitation, although less stable contacts are likely lost at this step. DNA linkers containing an MmeI restriction site are then ligated at the ends of the co-immunoprecipitated DNA fragments, and the resulting products are ligated together intra-molecularly. The ChIA-PET junctions produced by this ligation step are then excised by MmeI digestion, which cuts ~20 bp away from the enzyme recognition sequence, purified, and processed for high-throughput DNA sequencing. ChIA-PET has been used to construct the physical interactome of estrogen receptor alpha (ERalpha), the insulator-binding protein CTCF, and RNA polymerase II (Fullwood et al. 2009; Handoko et al. 2011; Sandhu et al. 2012).

3 Processing, Organization, and Analysis of Chromatin Contacts

3C technologies provide unparalleled views of the chromatin folding patterns present *in vivo*. These analyses, particularly the high-throughput derivatives like Hi-C and ChIA-PET, generate very large sets of sequence reads that must be processed and organized computationally. Correct and efficient handling of such "big data" sets is in itself a challenge that will likely grow with increased sequencing depth and the development of higher resolution technologies. Given the size and complexity of the data, biologically relevant relationships are identified with specialized computational tools. Below we present informatics tools currently used to handle genome-wide Hi-C chromatin conformation data.

3.1 Generating Robust Interaction Matrices

Chromatin interactions captured with Hi-C are usually represented for each chromosome using the interaction frequency matrix format (Fig. 3), whose rows and columns correspond to a segmentation of chromosome in regions of a fixed size (often 1 Mb or 100 kb). Each cell in these matrices contains the frequency of contacts between the corresponding pair of regions. Interaction frequencies are

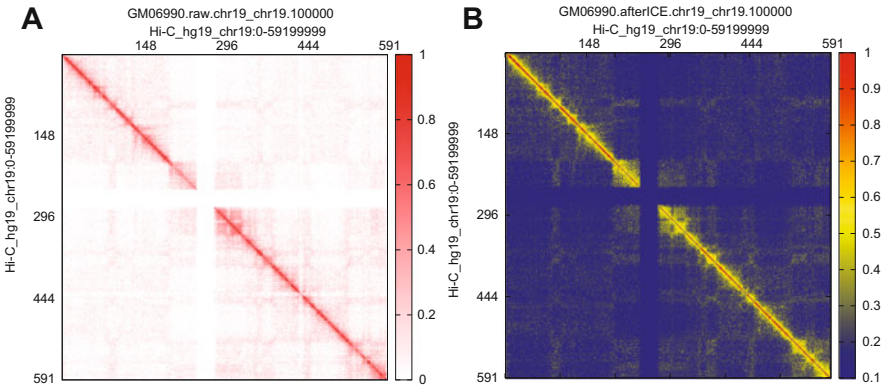

Fig. 3 Comparison of raw and ICE transformed interaction matrices. Log-transformed raw (**a**) and ICE normalized (**b**) interaction matrices for GM06990 lymphoblastoid cell line chromosome 19 (Lieberman-Aiden et al. 2009). Log transformation allows for a comparison between two different interaction matrices. The *white vertical* and *horizontal bands* starting approximately at the 250th bin correspond to centromeric regions where interactions cannot be mapped

obtained from paired-end sequencing of the Hi-C proximity-ligation products, and mapping the reads separately to a reference genome using standard alignment algorithms (Langmead and Salzberg 2012; Li et al. 2008). Hi-C read pairs that contain two uniquely mapped ends are then selected, and the data filtered for molecular by-products and other biases to obtain a collection of high-confidence Hi-C read pairs.

Systematic biases that are either biological or technical in nature are thought to alter the "raw" Hi-C data in ways that influence the set of observed contacts (Yaffe and Tanay 2011). At least four sources of Hi-C biases are thought to exist and interfere with the analysis: (1) the length of restriction fragments, (2) the GC content of interacting genomic regions, (3) the free DNA fragment sizes available during the ligation step, and (4) the reads' mappability (Cournac et al. 2012; Yaffe and Tanay 2011). Long and short fragments are known to have different ligation efficiencies and to compete differently for ligation with other fragment ends (Yaffe and Tanay 2011). For instance, the interaction probability of restriction fragments $<$800 base pairs in length was shown to decrease linearly with fragment size (Cournac et al. 2012).

Restriction fragments with a high GC content tend to be underrepresented in the raw Hi-C data because current sequencing techniques favour DNA with a GC content of approximately 45 %. Read count is consequently reduced at either extreme, particularly the GC-rich regions of genomes (Cournac et al. 2012). The length of DNA ends available for ligation can influence the frequency of Hi-C contacts because of the mechanical properties of the DNA polymer. Segments that are too short will be poorly ligated due to high bending persistence length, while longer segments will interfere with ligation by increasing entropy (Cournac et al. 2012). Sequence read mapping can also bias Hi-C analysis because reads

that originate from highly repetitive and low complexity regions are difficult to map unambiguously.

In 2011, Yaffe and Tanay described a probabilistic model to eliminate the systematic biases outlined above. This approach computes an expected interaction matrix corrected for fragment length, GC content and read mappability to normalize for known sources of biases in observed Hi-C matrices. While effective, this correction method is computationally expensive and only accounts for pre-defined bias sources. The Mirny group later described a less computationally demanding alternative, which does not rely on prior definition of bias source (Imakaev et al. 2012). The Iterative Correction and Eigenvalue decomposition (ICE) pipeline is available open-source and corrects the data based on DNA visibility with the understanding that the bias attributed to detecting contacts between fragment ends is factorizable. This assumption is justified since 99.99 % of the variance captured by Yaffe and Tanay could be accounted for with ICE (Imakaev et al. 2012).

During iterative correction, sequencing reads are first truncated to 25 bp before mapping, and uniquely aligned reads are pooled for downstream processing. The size of the remaining reads is then increased incrementally by 5 bp and re-mapped until full read sizes are reached, and any unaligned read are discarded. Sequence truncation is essential to account for the hybrid structure of Hi-C ligation products, and aggregating alignments over increasing lengths yields more double-sided mapped reads than if mapping is done at a fixed length (Imakaev et al. 2012). Hi-C read pairs containing two uniquely mapped ends are then filtered to eliminate uninformative/problematic reads by examining the position of mated reads relative to predicted cut sites: self-circularized products, reads that align too far from Hi-C junctions (typically more than 500 bp, defined by the library sonication settings), and those that map to fragments smaller than 100 bp or larger than 100 kb. Dangling ends, duplicate reads, and fragments in the top 0.2 % of read counts that likely originate from PCR biases, are also removed to minimize background.

The observed read count in a given cell of the interaction frequency matrix is finally corrected with normalization factors derived from the sum of columns and rows, to obtain a frequency matrix that offers equal representation to all restriction fragments (Fig. 3). Specifically, let $O_{i,j}$ represent the observed number of read pairs mapping to fragments i and j. This approach to bias correction attempts to identify two sets of values:

1. A vector of biases B, where B_i is the bias for ($B_i > 1$) or against ($B_i < 1$) observing a read from fragment i.
2. A matrix T, where T_{ij} is the true relative contact frequency matrix.

This normalization is achieved by solving the following system of equations using an iterative approach:

$$O_{ij} = B_i B_j T_{ij} \tag{1}$$

$$\sum_{i=1,\,|i-j|>1}^{n} T_{ij} = 1 \tag{2}$$

To solve for the true interaction frequency matrix, T_{ij}, a vector of biases is iteratively collected and applied to the observed frequencies.

Begin by summing the column of values O_{ij}:

$$S_j = \sum_i O_{ij} \tag{3}$$

Renormalize column sums (S_j) to have the unit mean and set as bias (ΔB_j) for current iteration:

$$\Delta B_j = \frac{S_j}{\overline{S}} \quad \text{where} \quad \overline{S} = \frac{\sum_j S_j}{n} \tag{4}$$

Repeat for row vector of the observed Hi-C matrix. Apply calculated biases to observed values (adjust region visibility):

$$W_{ij} = \frac{O_{ij}}{\Delta B_i \Delta B_j} \tag{5}$$

The algorithm is iterated, each time using the previous iteration's W_{ij} to replace O_{ij} (Eq. 5), until the variance of biases becomes negligible (W_{ij} has converged to T_{ij}).

3.2 Mining Chromatin Interaction Data

Extracting information from 3C and 4C-type data is relatively straightforward because one-dimensional profiles of chromatin interactions are typically obtained with these approaches. For instance, long-range looping interactions from given genomic regions are easily observed in these datasets. In contrast, the two-dimensionality of Hi-C interaction matrices significantly complicates data interpretation, particularly because of the very large amount of data contained within them. A popular approach to simplify analysis is to transform interaction frequencies into one-dimensional information. The resulting one-dimensional tracks can then be used to identify chromatin features and correlate conformational data with other genomics and epigenomic features. The two analysis methods employed for this purpose thus far are the principal component analysis (PCA) and the directionality index (DI) (Dixon et al. 2012; Lieberman-Aiden et al. 2009).

Lieberman-Aiden et al. used PCA to partition chromosomes into two sets of loci (compartments A and B, see Section "Introduction") each enriched in contacts and

depleted in interactions between them. The corresponding eigenvectors were then correlated with several metrics including gene density and expression, DNAseI hypersensitivity, chromatin composition, and histone modifications. This analysis identified the chromatin compartments as a new type of higher-order chromatin organization. Although PCA has proven useful to study chromosome organization, important structural information is clearly lost in the transformation process because regions are classified based on their overall contact behaviours. Specific contacts patterns are therefore not observable with this type of analysis. Dixon et al. introduced the notion of directionality index to map chromatin topology along chromosomes (Dixon et al. 2012). The DI groups chromatin interactions along chromosomes by measuring the difference in normalized read counts between upstream and downstream regions within a given window size (Fig. 4). DI analysis of Hi-C interaction matrices revealed that mammalian chromosomes are segmented into Mb-size TADs (Dixon et al. 2012; Filippova et al. 2014; Phillips-Cremins et al. 2013).

Since interactions at the periphery of TADs varied considerably in these Hi-C data sets, Dixon et al. also applied a Hidden Markov Model (HMM) based on the DI to delineate regions initiated by significant downstream interactions, and terminated by elevated upstream contacts. This analysis defined genome structure as a collection of spatial modules connected by relatively short chromatin segments or boundaries. A high level of TAD conservation was observed between human and mouse, suggesting that evolutionarily conserved elements and/or mechanisms are responsible for higher-order chromatin structures (Dixon et al. 2012). The fact that chromatin organization into Mb-sized TADs was found in multiple cell types and species suggests that it represents a general structural property of mammalian genomes (Dixon et al. 2012), and perhaps for all metazoan, since similar domain organizations have been found in Drosophila embryos (Sexton et al. 2012).

An alternative method to the use of a HMM for defining TADs is to encode the quality score of a domain (similar to a DI) in a dynamic programming algorithm (Filippova et al. 2013, 2014). The quality score of a domain is defined by the number of interactions between domain boundaries, scaled by the length of the domain. This alternative approach to TAD identification explores whether Mb-sized domains are the only topologically and functionally relevant collection of domains (Filippova et al. 2014). Whereas Dixon's HMM restricts TADs to a constant size at a particular resolution, the dynamic programming alternative allows for multiple TAD sizes along the diagonal to be evaluated to obtain the optimal set of domains for the given resolution. Filippova et al.'s (2013) approach is twofold. First, given a set of n levels resolutions, for each resolution determine the optimal set of domains based on interactivity measured by a Hi-C interaction matrix. Second, the consensus of domains across all n levels of resolutions is discovered. The result of the dynamic programming algorithm shows that Mb-size TADs are identified similarly to Dixon et al. (2012) yet there may exist a hierarchy of topological sub-domains that become apparent at higher resolutions.

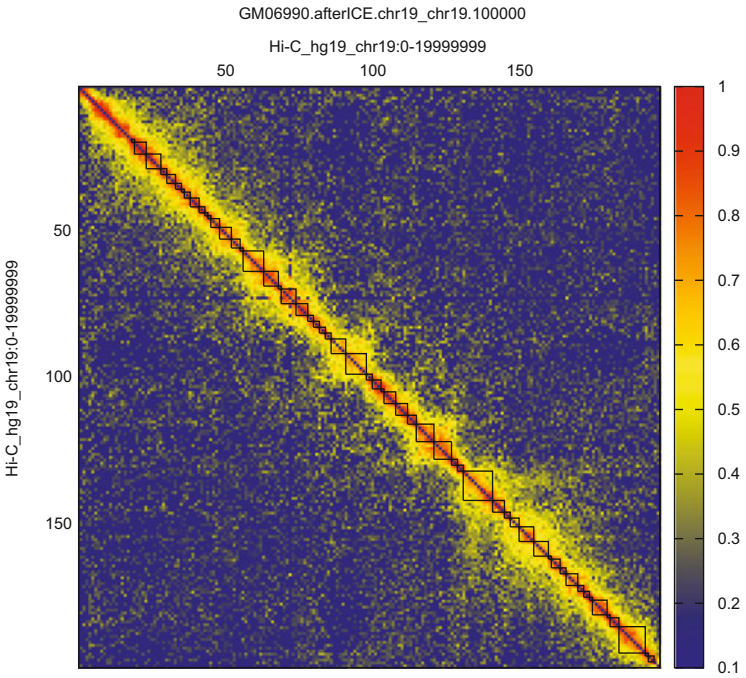

Fig. 4 Visualization of TADs. Topologically Associating Domains (TADs) identified within the first 20 Mb of chromosome 19 in the GM06990 lymphoblastoid cell line (Lieberman-Aiden et al. 2009). In this example, TADs were determined similarly to Dixon et al. (2012), but without the additional filtering of a HMM. A directionality index (DI) was first calculated by finding the difference in the sum of contacts 2 Mb bases downstream and upstream for each position across the chromosome. TAD boundaries were then identified as sites where the DI switches from negative to positive values

4 Visualizing Genome Organization

A key objective of measuring chromatin contacts with 3C methods is to reconstruct the 3D organization of chromatin, and to use these models to identify structural mechanisms that regulate genome function. All 3C-based techniques measure some form of interaction frequency, a value that captures the incidence of interactions between DNA fragments. Although the exact relation between interaction frequency and distance is complex, and likely varies based on chromatin composition and nuclear environment (Williamson et al. 2014), they generally appear to be inversely related (Fraser et al. 2009; Rousseau et al. 2011). Modeling one-dimensional data under this assumption is relatively straightforward and has led to the discovery of many regulatory links/elements simply by identifying chromatin loops. Modeling two-dimensional 5C or Hi-C data on the other hand is more complex because all interactions frequencies must fit within the resulting

spatial models. We discuss below some of the approaches used to model chromatin computationally from 3C-types interaction frequencies.

4.1 Computational Modeling of Chromatin in 3D

The goal of computational modeling approaches is to produce some sort of 3D model of chromatin from a set of interaction frequencies measured by 5C or Hi-C. A number of different modeling schemes now exist although they are relatively similar to each other for the most part, in that they treat modeling as an optimization problem. A majority of these computational approaches transform interaction frequency data to distance measures, and then infer one or more 3D models that satisfy the inferred pairwise distances, where the 3D structure is computed from a set of genomic loci constrained by local distances. Generally speaking, vertices representing the center of each restriction fragment are initialized randomly in space. These vertices are then iteratively moved within the model in order to optimize their placement with respect to all of the constraints that exist between pairs of restriction fragments. The best-fit model is obtained once the overall deviation between physical distances and corresponding constraints for each pair of fragments has been minimized (Fig. 5).

Both the *HOXA* cluster and the α-globin locus are examples of regions that have been extensively profiled using this type of approach (Bau and Marti-Renom 2011; Bau et al. 2011; Fraser et al. 2009; Rousseau et al. 2011; Umbarger et al. 2011). For example, (Rousseau et al. 2011) equate interaction frequencies found in the matrix to distance constraints with the following formula:

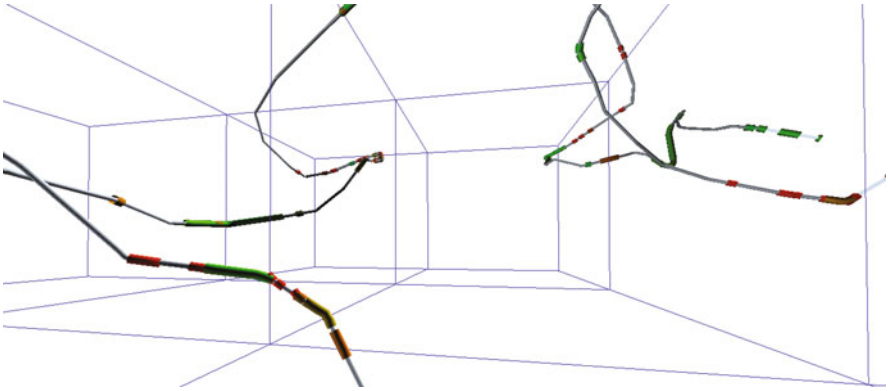

Fig. 5 3D structure prediction. A potential three-dimensional structure of GM06990 lymphoblastoid cell line chromosome five, derived from Hi-C data (Trieu and Cheng 2014) and visualized using a 3D genome browser [Butyaev et al. (2015), *In preparation*]. The DNA segments shown are portions of chr5: 50000000–74000000. Various coloured segments represent available experiment data (e.g., gene/transcription factor expression levels) that are cumulated by the browser

$$IF(i,j) \propto \frac{1}{D_s(ij)^{\alpha}} + noise \tag{6}$$

Here, $IF(i,j)$ is the interaction frequency for a pair of bins i and j of the interaction matrix, $D_s(ij)$ is the Euclidean distance between i and j, α is a parameter estimated from the observed data.

Generally, the structure inference problem is defined as an optimization task where one seeks to minimize the error between pairwise distances inferred from the IF data and those in the 3D model. There are, however, two problems that complicate optimization-based modeling of chromatin interactions from biological data. First, from a completely computational point of view, it is possible for the optimization of a given chromatin model to be caught in local minima, i.e., structures that cannot be improved by any local perturbation but do not correspond to the best-fitting structure. A simple approach to dealing with this is to provide a way for the model to extricate itself, for example by allowing the movement of vertices in a non-optimal direction (Rousseau et al. 2011). Second, 3C-based data generally originates from populations of cells, often diploid, for which a given region of chromatin may adopt more than one conformation. Due to the diploid nature of the cell, differences could also exist between copies of chromosomes within individual cells. Interaction matrices therefore represent an average of all these potentially different structures, for which no single chromatin model truly represent the actual biology.

These two considerations, inspired perhaps by the structural biology of proteins (Vendruscolo 2007), might explain why ensembles of chromatin models are also being considered. The approach involves sampling multiple representative structures from using a Markov chain Monte Carlo approach. Statistical analyses performed on a collection of chromatin structures might reveal trends in the data that a single model cannot describe. For example, ensembles of structures can be clustered to demonstrate the variability of chromatin organization within a population of cells and determine the most frequently occurring, and perhaps most interesting, chromatin model (Bau and Marti-Renom 2011). In another example, an ensemble of genome-scale chromatin models depicts the differential radial localization of small gene-rich chromosomes (Grosberg et al. 1988).

4.2 DNA as a Polymer and Other Constraints

In theory, knowing how far each fragment is from each other should be enough to place them in a three-dimensional space. Additional constraints like the shape and size of the nucleus may also be included for 3D structure prediction as some groups have already done (Bau and Marti-Renom 2012; Kalhor et al. 2012). In reality, multiple other factors should be accounted for, including biological and technical noise, and the innate behavior of the DNA polymer. The integration of polymer

physics with model optimization might be highly beneficial with regards to generating accurate chromatin structures. The first model was of the entire chromosome III of *Saccharomyces cerevisiae,* where 3C interaction frequencies were converted into physical measures of distance through the application of parameters derived from existing polymer models (Dekker et al. 2002). Two other examples used Hi-C-type data to model the entire *Saccharomyces cerevisiae* and *pombe* genomes at the kilobase level (Duan et al. 2010; Tanizawa et al. 2010). Converting interaction frequencies into distances were the primary constraints within the system. A number of other known features of the yeast nucleus were applied as constraints in the system, including the spherical nucleus, the location of the nucleolus and its relation to the centromeres, and other chromatin packing and polymer related restraints. Although this approach has the advantage of incorporating multiple sources of structural information within the model, care must be taken to properly deal with potentially contradicting information between sources.

4.3 Polymer Models of DNA Folding

Polymer models are interesting to study chromatin organization because they suggest potential physical mechanisms by which particular conformations might be formed instead of simply describing them. There are as many polymer model variants as there are potential mechanisms, and whether or not a given model should be used to explain chromatin folding depends on how well it recapitulates the actual interaction data. Questions on chromatin folding can be addressed at any genomic scale with polymer models. One such question, for example, aims to explain how chromosome territories are formed. The Dynamic Loop polymer model has been used to address this question and demonstrates that the formation of chromatin loops through diffusional motion can be a driving force for the formation of chromosome territories, even in the absence of other biological factors (Bohn and Heermann 2010; Bohn and Heermann 2011). Other work with chromatin polymers highlight how variations in polymer properties, such as looping interactions, fiber packing, and the presence of transcription factories, can each individually localize active gene-rich chromosome territories to the center of the nucleus and push inactive chromatin to the periphery (Cook and Marenduzzo 2009; Dorier and Stasiak 2010; Jerabek and Heermann 2012). These models together suggest a rather simple and elegant mechanism by which chromosome territories could be formed and positioned in the nuclear that is consistent with current biological data.

4.3.1 The Equilibrium and Fractal Globule Models

A pair of chromatin polymer models was brought into the spotlight with the development of Hi-C: the equilibrium and fractal globule models (Lieberman-Aiden et al. 2009). Power-law scaling exponents compatible with those derived

from the fractal globule model (Grosberg et al. 1988, 1993; Mirny 2011) were derived from the original Hi-C data at the scale of 0.5–7 Mb. With this model, the DNA polymer is crumpled into small globules, which then together form the basis for another self-similar polymer that undergoes the same type of crumpling over and over again as in a fractal. Regions that are close in the linear genome are thusly also proximal in the compacted structure. The major difference with the similar equilibrium model is the fractal globule does not generate knots when pulled tight from either end. The presence of knots in the polymer is not ideal for dynamic molecules like proteins or DNA, as they can greatly interfere with folding (Bolinger et al. 2010; Grosberg 2000; Vasilyev 2003).

In contrast to the knotted equilibrium model, the fractal globule might reveal very effective at describing folding and unfolding in response to various cellular stimuli. Biological support for this model, however, currently originates only from the power-law scaling of chromatin interactions observed in initial Hi-C experiments (Lieberman-Aiden et al. 2009). The original exponents derived from Hi-C data were from a genome-wide average across all chromosomes. When examined individually or in different cell types however, the exponents varied widely (Barbieri et al. 2012). While the fractal globule might still explain chromatin folding in some conditions, it cannot therefore represent all existing states.

4.3.2 The Strings and Binders Switch Model

It can be difficult for single models to explain trends observed in interaction data coming from different sources like those from FISH and 3C. The strings and binders switch (SBS) model addresses some of these issues (Barbieri et al. 2012; 2013a, b; Nicodemi et al. 2008; Nicodemi and Prisco 2009). Here, it is suggested that chromatin folding actually results from the binding of factors to the self-avoiding DNA polymer at pre-determined binding sites, forming chromatin "loops" that bring two distal regions of the polymer together. This system provides a large amount of flexibility as binding affinity, binder concentration, and the location of binding sites themselves can be greatly varied to suit different scenarios. This flexibility enables the reproduction of more diverse systems, including the variable exponents in Hi-C power-law scaling and the plateau observed in FISH experiments, as well as TADs and more complex structures (Barbieri et al. 2012). In the SBS model, changing either the concentration or the affinity of binders results in sharp transitions from open to compact states with brief transitions through intermediate phases. This mechanism might provide cells numerous ways to induce very sharp changes in chromatin conformation from small environmental changes.

5 Summary

This past decade has witnessed the development of techniques like Hi-C that capture chromatin organization genome-wide. These approaches offer unprecedented high-resolution views of genome organization that can now be used to assess how chromatin conformation may contribute to human disease. Defining which chromatin organization patterns and nuclear positions associate with genome dysfunction will require the integration of 3D data with other types of information like sequence variants, epigenomics and gene expression. For this, it will be essential to transform chromatin conformation data into formats accessible to a wide range of users. Various attempts have been made by the bioinformatics community to explore the application of chromatin organization data to genome browsers in the hope of creating a virtual exploration of the chromatin's 3D structures. These efforts have been met with difficulties, as designing back-end databases that efficiently handle the challenges of genomic "big data" set has not yet been accomplished. Virtual environments are still in their infancy and the resulting front-end for a browser of this type will require significant development to overcome many challenges. These challenges include designing visual elements to easily convey various experimental data, and of course instructing users on how to navigate such 3D environments.

Acknowledgments Work in the Dostie laboratory is supported by the Canadian Institutes of Health Research (CIHR MOP-115127), and the Canadian Cancer Society Research Institute (CCSRI # 702834). The Blanchette laboratory is funded by the National Sciences and Engineering Research Council (NSERC). JD is an FRSQ Research Scholar (Fonds de la Recherche en Santé du Québec Junior 2).

References

Abou El Hassan M, Bremner R (2009) A rapid simple approach to quantify chromosome conformation capture. Nucleic Acids Res 37:e35

Andrews AJ, Luger K (2011) Nucleosome structure(s) and stability: variations on a theme. Annu Rev Biophys 40:99–117

Andrey G, Montavon T, Mascrez B, Gonzalez F, Noordermeer D, Leleu M, Trono D, Spitz F, Duboule D (2013) A switch between topological domains underlies HoxD genes collinearity in mouse limbs. Science 340:1234167

Barbieri M, Chotalia M, Fraser J, Lavitas LM, Dostie J, Pombo A, Nicodemi M (2012) Complexity of chromatin folding is captured by the strings and binders switch model. Proc Natl Acad Sci USA 109:16173–16178

Barbieri M, Chotalia M, Fraser J, Lavitas LM, Dostie J, Pombo A, Nicodemi M (2013a) A model of the large-scale organization of chromatin. Biochem Soc Trans 41:508–512

Barbieri M, Fraser J, Lavitas LM, Chotalia M, Dostie J, Pombo A, Nicodemi M (2013b) A polymer model explains the complexity of large-scale chromatin folding. Nucleus 4:267–273

Bau D, Marti-Renom MA (2011) Structure determination of genomic domains by satisfaction of spatial restraints. Chromosome Res Int J Mol Supramol Evol Asp Chromosome Biol 19:25–35

Bau D, Marti-Renom MA (2012) Genome structure determination via 3C-based data integration by the integrative modeling platform. Methods 58:300–306

Bau D, Sanyal A, Lajoie BR, Capriotti E, Byron M, Lawrence JB, Dekker J, Marti-Renom MA (2011) The three-dimensional folding of the alpha-globin gene domain reveals formation of chromatin globules. Nat Struct Mol Biol 18:107–114

Berlivet S, Paquette D, Dumouchel A, Langlais D, Dostie J, Kmita M (2013) Clustering of tissue-specific sub-TADs accompanies the regulation of *HoxA* genes in developing limbs. PLoS Genet 9:e1004018

Bohn M, Heermann DW (2010) Diffusion-driven looping provides a consistent framework for chromatin organization. PLoS One 5:e12218

Bohn M, Heermann DW (2011) Repulsive forces between looping chromosomes induce entropy-driven segregation. PLoS One 6:e14428

Bolinger D, Sulkowska JI, Hsu HP, Mirny LA, Kardar M, Onuchic JN, Virnau P (2010) A Stevedore's protein knot. PLoS Comput Biol 6:e1000731

Bolzer A, Kreth G, Solovei I, Koehler D, Saracoglu K, Fauth C, Muller S, Eils R, Cremer C, Speicher MR, Cremer T (2005) Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. PLoS Biol 3:826–842

Boyle S, Rodesch MJ, Halvensleben HA, Jeddeloh JA, Bickmore WA (2011) Fluorescence in situ hybridization with high-complexity repeat-free oligonucleotide probes generated by massively parallel synthesis. Chromosome Res Int J Mol Supramol Evol Asp Chromosome Biol 19:901–909

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

Butyaev A, Mavlyutov R, Blanchette M, Cudré-Mauroux P, Waldispühl J (2015) A low-latency, big database system and browser for storage, querying and visualization of 3D genomic data. Nucleic Acids Res 43, e103

Bystricky K, Laroche T, van Houwe G, Blaszczyk M, Gasser SM (2005) Chromosome looping in yeast: telomere pairing and coordinated movement reflect anchoring efficiency and territorial organization. J Cell Biol 168:375–387

Cavalli G, Misteli T (2013) Functional implications of genome topology. Nat Struct Mol Biol 20:290–299

Cook PR, Marenduzzo D (2009) Entropic organization of interphase chromosomes. J Cell Biol 186:825–834

Cournac A, Marie-Nelly H, Marbouty M, Koszul R, Mozziconacci J (2012) Normalization of a chromosomal contact map. BMC Genomics 13:436

Dekker J (2003) A closer look at long-range chromosomal interactions. Trends Biochem Sci 28:277–280

Dekker J, Rippe K, Dekker M, Kleckner N (2002) Capturing chromosome conformation. Science 295:1306–1311

Denholtz M, Bonora G, Chronis C, Splinter E, de Laat W, Ernst J, Pellegrini M, Plath K (2013) Long-range chromatin contacts in embryonic stem cells reveal a role for pluripotency factors and polycomb proteins in genome organization. Cell Stem Cell 13:602–616

Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485:376–380

Dorier J, Stasiak A (2010) The role of transcription factories-mediated interchromosomal contacts in the organization of nuclear architecture. Nucleic Acids Res 38:7410–7421

Dostie J, Bickmore WA (2012) Chromosome organization in the nucleus - charting new territory across the Hi-Cs. Curr Opin Genet Dev 22:125–131

Dostie J, Dekker J (2007) Mapping networks of physical interactions between genomic elements using 5C technology. Nat Protoc 2:988–1002

Dostie J, Richmond TA, Arnaout RA, Selzer RR, Lee WL, Honan TA, Rubio ED, Krumm A, Lamb J, Nusbaum C, Green RD, Dekker J (2006) Chromosome conformation capture carbon

copy (5C): a massively parallel solution for mapping interactions between genomic elements. Genome Res 16:1299–1309

Duan Z, Andronescu M, Schutz K, McIlwain S, Kim YJ, Lee C, Shendure J, Fields S, Blau CA, Noble WS (2010) A three-dimensional model of the yeast genome. Nature 465:363–367

Ethier SD, Miura H, Dostie J (2012) Discovering genome regulation with 3C and 3C-related technologies. Biochim Biophys Acta 1819:401–410

Filippova D, Patro R, Duggal G, Kingsford C (2013) Multiscale identification of topological domains in chromatin. Proc 13th Workshop Algorithms Bioinf (WABI) 8126:300–3012

Filippova D, Patro R, Duggal G, Kingsford C (2014) Identification of alternative topological domains in chromatin. Algorithms Mol Biol 9:14

Fraser J, Rousseau M, Shenker S, Ferraiuolo MA, Hayashizaki Y, Blanchette M, Dostie J (2009) Chromatin conformation signatures of cellular differentiation. Genome Biol 10:R37

Fraser J, Rousseau M, Blanchette M, Dostie J (2010) Computing chromosome conformation. Methods Mol Biol 674:251–268

Fullwood MJ, Ruan Y (2009) ChIP-based methods for the identification of long-range chromatin interactions. J Cell Biochem 107:30–39

Fullwood MJ, Liu MH, Pan YF, Liu J, Xu H, Mohamed YB, Orlov YL, Velkov S, Ho A, Mei PH, Chew EG, Huang PY, Welboren WJ, Han Y, Ooi HS, Ariyaratne PN, Vega VB, Luo Y, Tan PY, Choy PY, Wansa KD, Zhao B, Lim KS, Leow SC, Yow JS, Joseph R, Li H, Desai KV, Thomsen JS, Lee YK, Karuturi RK, Herve T, Bourque G, Stunnenberg HG, Ruan X, Cacheux-Rataboul V, Sung WK, Liu ET, Wei CL, Cheung E, Ruan Y (2009) An oestrogen-receptor-alpha-bound human chromatin interactome. Nature 462:58–64

Grosberg AY (2000) Critical exponents for random knots. Phys Rev Lett 85:3858–3861

Grosberg A, Nechaev SK, Shakhnovich EI (1988) The role of topological limitations in the kinetics of homopolymer collapse and self-assembly of biopolymers. Biofizika 33:247–253

Grosberg A, Rabin I, Khavlin S, Nir A (1993) Self-similarity in the structure of DNA: why are introns needed? Biofizika 38:75–83

Hagege H, Klous P, Braem C, Splinter E, Dekker J, Cathala G, de Laat W, Forne T (2007) Quantitative analysis of chromosome conformation capture assays (3C-qPCR). Nat Protoc 2:1722–1733

Handoko L, Xu H, Li G, Ngan CY, Chew E, Schnapp M, Lee CW, Ye C, Ping JL, Mulawadi F, Wong E, Sheng J, Zhang Y, Poh T, Chan CS, Kunarso G, Shahab A, Bourque G, Cacheux-Rataboul V, Sung WK, Ruan Y, Wei CL (2011) CTCF-mediated functional chromatin interactome in pluripotent cells. Nat Genet 43:630–638

Imakaev M, Fudenberg G, McCord RP, Naumova N, Goloborodko A, Lajoie BR, Dekker J, Mirny LA (2012) Iterative correction of Hi-C data reveals hallmarks of chromosome organization. Nat Methods 9:999–1003

Jerabek H, Heermann DW (2012) Expression-dependent folding of interphase chromatin. PLoS One 7:e37525

Kalhor R, Tjong H, Jayathilaka N, Alber F, Chen L (2012) Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. Nat Biotechnol 30:90–98

Kleinjan DA, van Heyningen V (2005) Long-range control of gene expression: emerging mechanisms and disruption in disease. Am J Hum Genet 76:8–32

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359

Li H, Ruan J, Durbin R (2008) Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res 18:1851–1858

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326:289–293

Ling JQ, Li T, Hu JF, Vu TH, Chen HL, Qiu XW, Cherry AM, Hoffman AR (2006) CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. Science 312:269–272

Lomvardas S, Barnea G, Pisapia DJ, Mendelsohn M, Kirkland J, Axel R (2006) Interchromosomal interactions and olfactory receptor choice. Cell 126:403–413

Lonfat N, Montavon T, Darbellay F, Gitto S, Duboule D (2014) Convergent evolution of complex regulatory landscapes and pleiotropy at Hox loci. Science 346:1004–1006

Miele A, Gheldof N, Tabuchi TM, Dostie J, Dekker J (2006) Mapping chromatin interactions by chromosome conformation capture. Curr Protoc Mol Biol Chapter 21:Unit 21.11

Mirny L (2011) The fractal globule as a model of chromatin architecture in the cell. Chromosome Res Int J Mol Supramol Evol Asp Chromosome Biol 19:37–51

Mishiro T, Ishihara K, Hino S, Tsutsumi S, Aburatani H, Shirahige K, Kinoshita Y, Nakao M (2009) Architectural roles of multiple chromatin insulators at the human apolipoprotein gene cluster. EMBO J 28:1234–1245

Montavon T, Soshnikova N, Mascrez B, Joye E, Thevenet L, Splinter E, de Laat W, Spitz F, Duboule D (2011) A regulatory archipelago controls Hox genes transcription in digits. Cell 147:1132–1145

Murrell A, Heeson S, Reik W (2004) Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops. Nat Genet 36:889–893

Nativio R, Sparago A, Ito Y, Weksberg R, Riccio A, Murrell A (2011) Disruption of genomic neighbourhood at the imprinted IGF2-H19 locus in Beckwith–Wiedemann syndrome and Silver–Russell syndrome. Hum Mol Genet 20:1363–1374

Naumova N, Smith E, Zhan Y, Dekker J (2012) Analysis of long-range chromatin interactions using chromosome conformation capture. Methods (San Diego, Calif) 58:192–203

Naumova N, Imakaev M, Fudenberg G, Zhan Y, Lajoie BR, Mirny LA, Dekker J (2013) Organization of the mitotic chromosome. Science 342:948–953

Nicodemi M, Prisco A (2009) Thermodynamic pathways to genome spatial organization in the cell nucleus. Biophys J 96:2168–2177

Nicodemi M, Panning B, Prisco A (2008) A thermodynamic switch for chromosome colocalization. Genetics 179:717–721

Noordermeer D, Leleu M, Splinter E, Rougemont J, De Laat W, Duboule D (2011) The dynamic architecture of Hox gene clusters. Science 334:222–225

Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piolot T, van Berkum NL, Meisig J, Sedat J, Gribnau J, Barillot E, Bluthgen N, Dekker J, Heard E (2012) Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature 485:381–385

Palstra RJ, Tolhuis B, Splinter E, Nijmeijer R, Grosveld F, de Laat W (2003) The beta-globin nuclear compartment in development and erythroid differentiation. Nat Genet 35:190–194

Phillips-Cremins JE, Sauria ME, Sanyal A, Gerasimova TI, Lajoie BR, Bell JS, Ong CT, Hookway TA, Guo C, Sun Y, Bland MJ, Wagstaff W, Dalton S, McDevitt TC, Sen R, Dekker J, Taylor J, Corces VG (2013) Architectural protein subclasses shape 3D organization of genomes during lineage commitment. Cell 153:1281–1295

Pope BD, Ryba T, Dileep V, Yue F, Wu W, Denas O, Vera DL, Wang Y, Hansen RS, Canfield TK, Thurman RE, Cheng Y, Gulsoy G, Dennis JH, Snyder MP, Stamatoyannopoulos JA, Taylor J, Hardison RC, Kahveci T, Ren B, Gilbert DM (2014) Topologically associating domains are stable units of replication-timing regulation. Nature 515:402–405

Rabl C (1885) Uber Zelltheilung. Morphol Jahrbuch 10:214–330

Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, Aiden EL (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell 159:1665–1680

Rousseau M, Fraser J, Ferraiuolo MA, Dostie J, Blanchette M (2011) Three-dimensional modeling of chromatin structure from interaction frequency data using Markov chain Monte Carlo sampling. BMC Bioinf 12:414

Rousseau M, Ferraiuolo MA, Crutchley JL, Wang XQ, Miura H, Blanchette M, Dostie J (2014) Classifying leukemia types with chromatin conformation data. Genome Biol 15:R60

Ryba T, Hiratani I, Lu J, Itoh M, Kulik M, Zhang J, Schulz TC, Robins AJ, Dalton S, Gilbert DM (2010) Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. Genome Res 20:761–770

Sandhu K, Li G, Poh H, Quek Y, Sia Y, Peh S, Mulawadi F, Lim J, Sikic M, Menghi F, Thalamuthu A, Sung W, Ruan X, Fullwood M, Liu E, Csermely P, Ruan Y (2012) Large-scale functional organization of long-range chromatin interaction networks. Cell Rep 2:1207–1219

Sanyal A, Lajoie BR, Jain G, Dekker J (2012) The long-range interaction landscape of gene promoters. Nature 489:109–113

Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, Parrinello H, Tanay A, Cavalli G (2012) Three-dimensional folding and functional organization principles of the Drosophila genome. Cell 148:458–472

Simonis M, Klous P, Splinter E, Moshkin Y, Willemsen R, de Wit E, van Steensel B, de Laat W (2006) Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nat Genet 38:1348–1354

Spector DL (2003) The dynamics of chromosome organization and gene regulation. Annu Rev Biochem 72:573–608

Splinter E, de Wit E, Nora EP, Klous P, van de Werken HJ, Zhu Y, Kaaij LJ, van Ijcken W, Gribnau J, Heard E, de Laat W (2011) The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA. Genes Dev 25:1371–1383

Stadhouders R, Thongjuea S, Andrieu-Soler C, Palstra RJ, Bryne JC, van den Heuvel A, Stevens M, de Boer E, Kockx C, van der Sloot A, van den Hout M, van Ijcken W, Eick D, Lenhard B, Grosveld F, Soler E (2012) Dynamic long-range chromatin interactions control Myb proto-oncogene transcription during erythroid development. EMBO J 31:986–999

Taddei A, Schober H, Gasser SM (2010) The budding yeast nucleus. Cold Spring Harb Perspect Biol 2:a000612

Tanizawa H, Iwasaki O, Tanaka A, Capizzi JR, Wickramasinghe P, Lee M, Fu Z, Noma K (2010) Mapping of long-range associations throughout the fission yeast genome reveals global genome organization linked to transcriptional regulation. Nucleic Acids Res 38:8164–8177

Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W (2002) Looping and interaction between hypersensitive sites in the active beta-globin locus. Mol Cell 10:1453–1465

Tolhuis B, Blom M, Kerkhoven RM, Pagie L, Teunissen H, Nieuwland M, Simonis M, de Laat W, van Lohuizen M, van Steensel B (2011) Interactions among Polycomb domains are guided by chromosome architecture. PLoS Genet 7:e1001343

Trieu T, Cheng J (2014) Large-scale reconstruction of 3D structures of human chromosomes from chromosomal contact data. Nucleic Acids Res 42:e52

Umbarger MA, Toro E, Wright MA, Porreca GJ, Bau D, Hong SH, Fero MJ, Zhu LJ, Marti-Renom MA, McAdams HH, Shapiro L, Dekker J, Church GM (2011) The three-dimensional architecture of a bacterial genome and its alteration by genetic perturbation. Mol Cell 44:252–264

van de Werken HJ, de Vree PJ, Splinter E, Holwerda SJ, Klous P, de Wit E, de Laat W (2012a) 4C technology: protocols and data analysis. Methods Enzymol 513:89–112

van de Werken HJ, Landan G, Holwerda SJ, Hoichman M, Klous P, Chachik R, Splinter E, Valdes-Quezada C, Oz Y, Bouwman BA, Verstegen MJ, de Wit E, Tanay A, de Laat W (2012b) Robust 4C-seq data analysis to screen for regulatory DNA interactions. Nat Methods 9:969–972

Vasilyev OA (2003) Universality of the crossing probability for the Potts model for $q = 1, 2, 3, 4$. Phys Rev E Stat Nonlin Soft Matter Phys 68:026125

Vendruscolo M (2007) Determination of conformationally heterogeneous states of proteins. Curr Opin Struct Biol 17:15–20

Visser AE, Jaunin F, Fakan S, Aten JA (2000) High resolution analysis of interphase chromosome domains. J Cell Sci 113:2585–2593

Wang KC, Yang YW, Liu B, Sanyal A, Corces-Zimmerman R, Chen Y, Lajoie BR, Protacio A, Flynn RA, Gupta RA, Wysocka J, Lei M, Dekker J, Helms JA, Chang HY (2011) A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature 472:120–124

West AG, Fraser P (2005) Remote control of gene transcription. Hum Mol Genet 14:R101–R111

Williamson I, Berlivet S, Eskeland R, Boyle S, Illingworth RS, Paquette D, Dostie J, Bickmore WA (2014) Spatial genome organization: contrasting views from chromosome conformation capture and fluorescence in situ hybridization. Genes Dev 28:2778–2791

Wurtele H, Chartrand P (2006) Genome-wide scanning of HoxB1-associated loci in mouse ES cells using an open-ended chromosome conformation capture methodology. Chromosome Res Int J Mol Supramol Evol Asp Chromosome Biol 14:477–495

Yaffe E, Tanay A (2011) Probabilistic modeling of Hi-C contact maps eliminates systematic biases to characterize global chromosomal architecture. Nat Genet 43:1059–1065

Zhang Y, McCord RP, Ho YJ, Lajoie BR, Hildebrand DG, Simon AC, Becker MS, Alt FW, Dekker J (2012) Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. Cell 148:908–921

Zhao Z, Tavoosidana G, Sjolinder M, Gondor A, Mariano P, Wang S, Kanduri C, Lezcano M, Sandhu KS, Singh U, Pant V, Tiwari V, Kurukuti S, Ohlsson R (2006) Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nat Genet 38:1341–1347

Developmental Roles of Histone H3 Variants and Their Chaperones

Sebastian Müller, Dan Filipescu, and Geneviève Almouzni

Abstract Within the nucleus, chromatin dynamics with the interplay between lineage-specific transcription factors, histone dynamics, non-histone proteins and their modifications along with contribution of non coding RNA emerge as critical in defining a stable cellular identity. A fine tuning of these different parameters can impact the balance between stability and plasticity during the development and life span of a multicellular organism. At the chromatin level, dynamics of histone H3 variants through their regulated incorporation into the genome, orchestrated in time and location by dedicated histone chaperones, have gained importance in our understanding of these processes. The combination of their individual modifications and specific binding partners provide distinct features that together contribute to the establishment of genomic *loci*. Here, we describe the network of histone chaperones that governs their handling with a focus on deposition, to highlight how their distinct distribution impacts genome organization and function. Next, we integrate the importance of H3 variants in the context of current knowledge related to nuclear reprogramming and cell differentiation. Then, using the centromere as a paradigm, we describe a case where the identity of a given genomic *locus* is propagated across different cell types. Finally, to place H3 dynamics in a developmental context, we underline the role of chromatin changes in cell differentiation associated with gastrulation, myogenesis or neurogenesis.

Keywords H3 variants • Histone chaperone • Embryo • Gametogenesis • Development • Differentiation, oocyte

S. Müller • D. Filipescu • G. Almouzni (✉)
Institut Curie, PSL Research University, 75005 Paris, France

CNRS, UMR3664, 75005 Paris, France

Equipe Labellisée Ligue contre le Cancer, 75005 Paris, France

Sorbonne Universités, 75005 Paris, France

UPMC Univ Paris 06, 75005 Paris, France
e-mail: genevieve.almouzni@curie.fr

D.P. Bazett-Jones, G. Dellaire (eds.), *The Functional Nucleus*,
DOI 10.1007/978-3-319-38882-3_17

1 Introduction

Long before the discovery of the molecular nature and structure of the DNA double helix (Franklin and Gosling 1953; Watson and Crick 1953) and before the nucleosome was defined as the basic unit of chromatin (Kornberg 1974; Oudet et al. 1975), Heitz classified distinct chromatin regions into euchromatin and heterochromatin based on cytological properties (Heitz 1928). This classification is now further documented and enlarged thanks to studies on the distinct biochemical and structural properties characteristic of a given chromosomal region, which in particular includes nucleosome composition. In the nucleosome core particle DNA wraps in about two superhelical turns around a protein octamer composed of histones, to form a basic module that, together with linker DNA, represents the repeat unit of chromatin. A nucleosome octamer contains two molecules of each of the core histones H2A, H2B, H3 and H4 (Kornberg 1974; Oudet et al. 1975) and nucleosomes are connected via a linker DNA associated to linker histones (McGhee and Felsenfeld 1980). Negative charges of the DNA backbone are neutralized by the basic nature of histones (Luger et al. 1997), these units are thought to fold into higher order chromatin structures according to different models [for a reviews see (Fudenberg and Mirny 2012; Kalashnikova et al. 2013; Luger et al. 2012)]. Importantly, histones beyond a structural role to establish the building block of chromatin, also represent important modulators of many cellular functions operating on DNA, including replication, gene expression, repair and recombination. In metazoans, all histones exist as histone variants (Franklin and Zweidler 1977) (Table 1) [see (Talbert et al. 2012) for a list of the current nomenclature]. This variation in primary sequence contributes to enlarge the repertoire of chromatin features, along with posttranslational histone and DNA modifications and non-coding RNAs. *In vitro*, different histone variants can confer differential stability to nucleosomes (Andrews and Luger 2011; Marino-Ramirez et al. 2005), they can be differently modified, and their expression during development is regulated. All this offers means to control chromatin structure and function. For their molecular properties and functions, we direct readers to the following reviews (Mattiroli et al. 2015; Szenker et al. 2014) as we focus on the aspects related to development that underline H3 variants.

The number of H3 variants differs between species and are generally classified into two subgroups. Expression of replicative histones peaks during S-phase providing the main histone supply throughout DNA replication. These replicative histones are deposited in a DNA synthesis-coupled (DSC) manner, and in humans they are mainly represented by H3.1 and H3.2. Replacement variants, generally expressed independently of S-phase, are incorporated in a DNA synthesis-independent (DSI) manner. In humans, the replacement variants best characterized are H3.3 [for a review see (Szenker et al. 2011)] and the centromeric H3 variant $CenH3^{CENP\text{-}A}$ [for a review see (Müller and Almouzni 2014)]. Table 1 and Fig. 1 give a list and illustration including other variants identified to date. Importantly, outside chromatin, soluble histones are escorted by a network of histone chaperones. This can help to neutralize their positive charge and prevent promiscuous

Table 1 List of major histone variants and their chaperones

	Histone H3 (*Hs*)	Conservation	Main variant functions	Histone chaperones	Main chaperone functions	Key references
H3	H3.1	H3.2 (*Mm*, *Dm*, *Xl*) H3 (*Sc*, *Sp*)	Canonical variant *Sc* and *Sp* only have H3 and no identified variants.	CAF-1 complex: – p150 (*Hs, Mm*, *Xl*); Rif2/Cac1 (*Sc*); SPBC29A10.03C (*Sp*); Fas1 (*At*) – p60 (*Hs, Mm*, *Xl*); p150 (*Dm*) Cac2 (*Sc*); SPAC26H5.03 (*Sp*); Fas2 (*At*) – RbAp46/48 (*Hs, Mm*); Mis16 (*Sp*)	H3.1 deposition in a DNA synthesis dependent manner (i.e., replicative histone H3 variant) by CAF-1 In *Sc* and *Sp* H3 can be deposited in DNA synthesis-dependent or-independent manners	Kaufman et al. (1995), Smith and Stillman (1989), Tagami et al. (2004)
				ASF1a and ASF1b (*Hs, Mm*); Asf1 (*Sc*, *Xl*); Cia1 (*Sp*); Sga1 and Sga2 (*At*)	ASF1a and ASF1b are a histone donors for CAF-1	Le et al. (1997), Munakata et al. (2000), Tagami et al. (2004), Tyler et al. (1999)
	H3.2	H3.1 (*Mm*) H3 (*Sc*, *Sp*)	Canonical H3 histone variant	CAF-1 complex	H3.2 is deposited by CAF-1	Akiyama et al. (2011), Latreille et al. (2014)
				ASF1a and ASF1b	ASF1a and ASF1b are a histone donors for CAF-1	Le et al. (1997), Munakata et al. (2000), Tagami et al. (2004), Tyler et al. (1999)
	H3.3	H3.3 (*Mm*, *Dm*, *Xl*) H3 (*Sc*, *Sp*)	Replacement H3 histone variant	HIRA complex: – HIRA (*Hs, Mm*, *Xl*, *Dm*, *At*); HIR1/HIR2 (*Sc*); Hip1, Sim2 (*Sp*) – Cabin1 (*Hs, Mm*); Hir3 (*Sc*); Hip3 (*Sp*); AT4G32820# (*At*) – UBN1; Hpc4 (*Sc*); Hip4 (*Sp*); yemanuclein-α (*Dm*); T14N5.16# (*At*); mCG.1031012# (*Mm*)	H3.3 is deposited in a DNA-synthesis independent manner by the HIRA complex. The HIRA complex was the first complex identified specifically for H3.3	Banumathy et al. (2009), Rai et al. (2011), Ray-Gallet et al. (2002)

(continued)

Table 1 (continued)

	Histone H3 (*Hs*)	Conservation	Main variant functions	Histone chaperones	Main chaperone functions	Key references
				ASF1a and ASF1b	ASF1a and ASF1b are a histone donors for HIRA	Le et al. (1997), Munakata et al. (2000), Tagami et al. (2004), Tyler et al. (1999)
				DAXX/ATRX – DAXX (*Hs, Mm*) – ATRX (*Hs, Mm*); Xnp (*Dm*)	DAXX is a H3.3 deposition factor independent of DNA synthesis at telomeres, rDNA and pericentric heterochromatin	Drane et al. (2010), Goldberg et al. (2010)
				DEK (*Hs*, *Dm*)	Transcriptional co-activator	Sawatsubashi et al. (2010)
	CenH3 CENP-A	$CenH3^{CENP-A}$ (*Mm*, *Xl*) $CenH3^{CID}$ (*Dm*) $CenH3^{Cse4}$ (*Sc*) $CenH3^{Cnp1}$ (*Sp*) $CenH3^{HTR12}$ (*At*) $CenH3^{HCP-3}$ (*Ce*)	Replacement H3 histone variant	HJURP (*Hs*, *Mm*, *Xl*) Scm3 (*Sc*, *Sp*) Cal1 (*Dm*)[a]	CenH3 is the centromeric H3 variant enriched at centromeres and an epigenetic determinant of centromeres. HJURP/Scm3 or Cal1 (*Dm*)[a] are responsible for their centromere-specific deposition. HJURP interacts directly with the kinetochore factor CENP-C and has a DNA binding domain required for $CenH3^{CENP-A}$ deposition	Barnhart et al. (2011), Dunleavy et al. (2009), Foltz et al. (2009), Müller et al. (2014), Shuaib et al. (2010), Tachiwana et al. (2015)

				DAXX (*Hs*)	DAXX deposits CenH3 $^{CENP-A}$ at non-centromeric sites in CenH3^{CENP-A} overexpressing cells	Lacoste et al. (2014)
	H3.4	Also known as H3t (*Hs*)	Testis specific H3 replacement variant incorporated genome-wide at late meiosis or early spermiogenesis			Albig et al. (1996), Tachiwana et al. (2010)
	H3.5		Testis specific H3 replacement variant. Hominid-specific			Schenk et al. (2011)
	H3.Y.1	Also known as H3.Y (*Hs*)	Replacement H3 histone variant			
	H3.Y.2	Also known as H3.X (*Hs*)	Replacement H3 histone variant			
				BAF53 (*Hs*, *Mm*); BAP55 (*Dm*); Arp8 (*Sc*)	Assists in chromatin remodeling	Shen et al. (2003)
				Hif1 (*Sc*)	Histone Acetyltransferase assistance	Ai and Parthun (2004)
				tNASP/sNASP (*Hs*, *Mm*) N1/N2 (*Xl*)	NASP protects H3–H4 from degradation in human cells. N1/N2 regulates H3–H4 storage in *Xl* oocytes	Campos et al. (2010, Cook et al. (2011), Kleinschmidt et al. (1985)
				Rsf-1 (*Hs*, *Mm*)	Assists in chromatin remodeling	Loyola et al. (2001)
				Rtt106 (*Sc*); SPAC6G9.03c (*Sp*)	Silencing of heterochromatin	Huang et al. (2005)

(continued)

Table 1 (continued)

	Histone H3 (*Hs*)	Conservation	Main variant functions	Histone chaperones	Main chaperone functions	Key references
				Spt6 (*Hs, Mm*, *Xl*, *Dm*, *Sc*, *Sp*)	Transcription initiation and elongation	Bortvin and Winston (1996)
				SSRP1 (FACT subunit) (*Hs, Mm*, *Xl*, *Dm*); SSRP (*At*); Pob3 (*Sc*, *Sp*)	Transcription elongation; assists in chromatin remodeling. FACT also interacts with H2A-H2B	Belotserkovskaya et al. (2003)

This table shows multiple species conservation and their functions. Species conservation is indicated if seminal studies of histone variants or histone chaperones were conducted on these proteins

Dm Drosophila melanogaster; *Ce Caenorhabtidis elegans*; *Hs Homo sapiens*; *Mm Mus musculus*; *Sc Saccharomyces cerevisiae*; *Sp Schizosaccharomyces pombe*; *Xl Xenopus laevis*; *At Arabidopsis thaliana*

[a]Only found in *Dm* where it takes the role of HJURP/Scm3. Cal1 is not part of the HJURP/Scm3 protein family

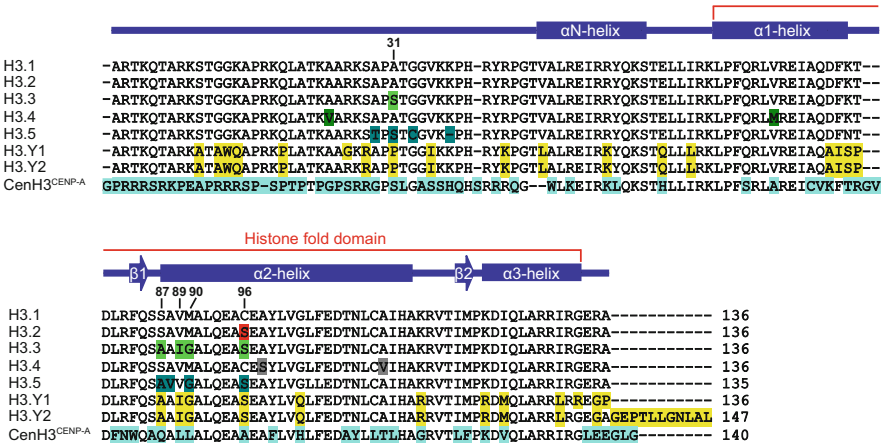

Fig. 1 Sequence alignment of human histone H3 variants. The domains of the canonical histone H3.1 are indicated. Sequence variations compared to H3.1 are indicated in distinct *colors*. Important residues of H3.3 distinct to H3.1 are indicated

interaction. By definition (De Koning et al. 2007), histone chaperones are defined as proteins that associate with histones that are involved in their transfer without being necessarily part of the final product. For further reading we address the reader to: (De Koning et al. 2007; Gurard-Levin et al. 2014; Mattiroli et al. 2015).

Histone chaperones (listed in Table 1) can either interact without particular variant specificity or selectivity, or rather dedicated to and preferentially associate with a particular histone variant. The isolation of H3.1 and H3.3 soluble protein complexes (Tagami et al. 2004) revealed particular association for these variants *in vivo*. Identified in both complexes, *N*uclear *a*utoantigenic *s*perm *p*rotein (NASP) (Richardson et al. 2000) and *A*nti *s*ilencing *f*actor *1* (ASF1) (Mello et al. 2002; Sharp et al. 2001) are examples of general H3–H4 chaperones. The *c*hromatin *a*ssembly *f*actor *1* (CAF-1) complex (Kaufman et al. 1995; Smith and Stillman 1989) only found with the replicative variants H3.1 (Tagami et al. 2004) and H3.2 (Akiyama et al. 2011; Latreille et al. 2014) shows more selectivity. Similarly, two dedicated H3.3 chaperone exist in humans: the *hi*stone cell cycle *r*egulator *A* (HIRA) complex (Banumathy et al. 2009; Daniel Ricketts et al. 2015; Rai et al. 2011; Ray-Gallet et al. 2002; Tagami et al. 2004) and the *D*eath-domain *a*ssociated protein 6 (DAXX)/*α-t*halassemia/mental *r*etardation syndrome *X*-linked (ATRX) complex (Goldberg et al. 2010; Lewis et al. 2010; Voon et al. 2015). Importantly, these H3.3 complexes differ in their capacity to promote enrichment of H3.3 at specific chromosomal regions (Fig. 2). Whereas the HIRA complex is required for H3.3 deposition at transient nucleosome-free regions, DAXX/ATRX are important for H3.3 enrichment at pericentric heterochromatin and telomeres (Fig. 2). *H*olliday *ju*nction *r*ecognition *p*rotein (HJURP) is the dedicated chaperone for the centromeric variant CenH3$^{CENP\text{-}A}$ (Dunleavy et al. 2009; Foltz et al. 2009), and targets CenH3$^{CENP\text{-}A}$ deposition at centromeres in a cell-cycle-dependent

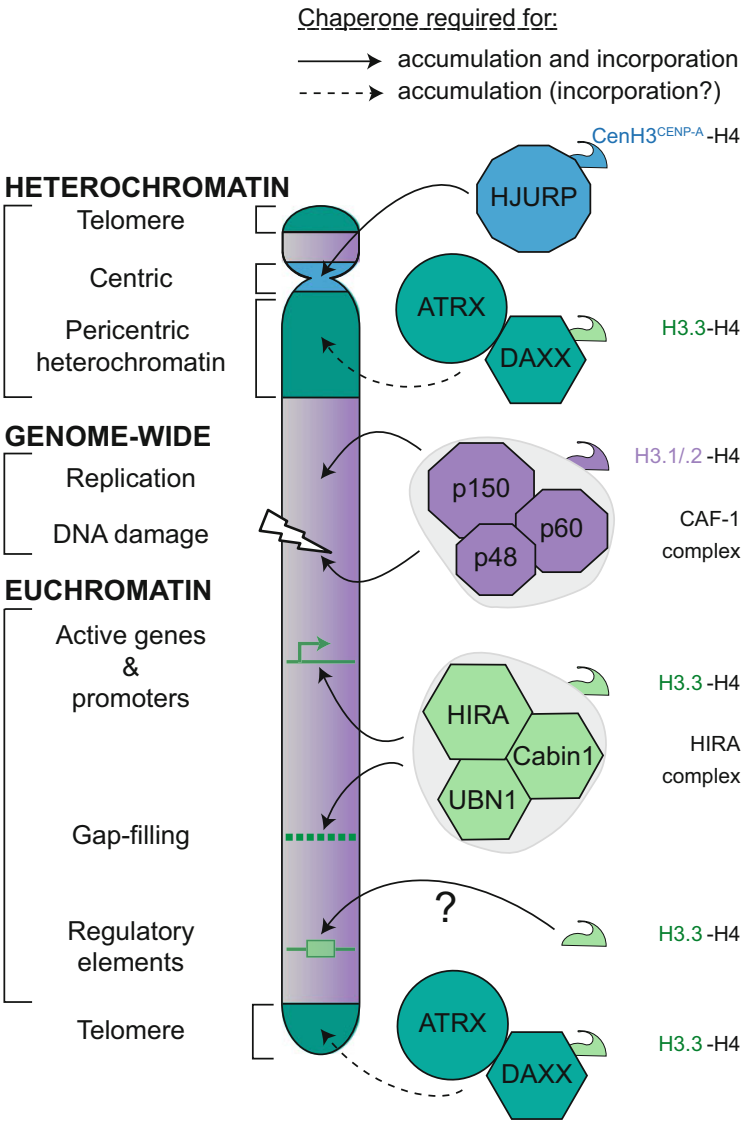

Fig. 2 Local enrichment of histone H3 variants and the role of their chaperones. Incorporation of replicative histones H3.1 and H3.2 (*purple*) occurs genome-wide during DNA replication and at sites of damage during DNA repair. These DNA synthesis coupled (DSC) events are promoted by the CAF-1 chaperone complex, which comprises three subunits referred to in humans as p150, p60 and p48. In mouse somatic and embryonic cells, the replacement variant H3.3 (*light green*) is enriched in coding regions and at specific chromatin landmarks. In heterochromatin, DAXX cooperates with the chromatin remodeler ATRX to give rise to the accumulation of H3.3 at pericentric satellite repeats and telomeres. In euchromatin, the HIRA complex (comprising the HIRA, Cabin1 and UBN1 proteins) is responsible for H3.3 enrichment in the body of transcribed genes and at promoters of transcribed or non-transcribed genes. In addition, a broad incorporation of H3.3 involves the HIRA complex by means thought to exploit the binding to transiently accessible non-nucleosomal DNA (gap-filling mechanism). DAXX has been suggested as a candidate for H3.3 enrichment at regulatory elements. The chaperone HJURP is involved in the

fashion. Importantly, HJURP does not merely escort the soluble histone variant, but is directly involved in the CenH3^{CENP-A} deposition step at the centromeric *locus* through its ability to bind DNA (Müller et al. 2014) and re-organize centromeric chromatin (Perpelescu et al. 2015; Tachiwana et al. 2015). Importantly, in cells overexpressing CenH3^{CENP-A}, DAXX can also interact with CenH3^{CENP-A}, and is able to deposit it in an unscheduled manner at non-centromeric sites (Lacoste et al. 2014). The specificity and selectivity of histone chaperones towards their histone cargo is complex *in vivo*, but can partly be attributed to differences in the amino acid sequence of the variants themselves [Fig. 1, reviewed in (Gurard-Levin et al. 2014; Szenker et al. 2014)]. These differences can be minimal, as illustrated by the case of DAXX, whose recognition of the residues 87 and 90 specific to H3.3 is sufficient to promote a selective interaction (DeNizio et al. 2014; Elsässer et al. 2012; Liu et al. 2012). Studies in various model organisms in the past two decades have explored how histone variants leverage their network of chaperones in the context of gene expression, replication and repair. These connections and selectivity may change according to physiology and cell type and while it has been traditionally studied through biochemical and cellular approaches, an emerging challenge is to understand how this network operates at the level of an organism.

Development is orchestrated at the level of genome function by the crosstalk between lineage and cell type-specific transcription factors with the modular nature of histones and their post-translational modifications (PTMs), noncoding RNAs, DNA methylation and other proteins that bind or modify these elements. The zygote gives rise to all cell types through cell division and differentiation, while exploiting the same DNA template during development. Therefore, the identity of each cellular lineage is defined by its unique gene expression profile and its capacity to adapt it to the environment. Faithful propagation to its daughter cells ensures lineage stability, whereas plasticity in response to developmental cues permits cell fate decisions as an integral part of differentiation. Thus, a constant balance between stability and plasticity is at work during the entire lifespan of an organism. In this context, the contribution of histone variants to developmental processes is interesting to highlight, as their discrete modes of incorporation influence chromatin structure, which offers direct control over various genomic functions that are critical during development [reviewed in (Filipescu et al. 2014; Gaume and Torres-Padilla 2015; Loyola and Almouzni 2007)]. Furthermore, the diversity of nucleosome composition, including the introduction of distinct variants and potentially histone modifications, provides a layer of information additional to that encoded by DNA. Thus, when contributing to propagate heritable changes of cellular phenotype or gene expression across mitosis or meiosis, without involving changes in the underlying DNA sequence, histone variants are versatile components of epigenetic nature [reviewed in (Russo et al. 1996)]. Alternately, when acting in a transient

Fig. 2 (continued) deposition of the centromere-specific CenH3 variant (*blue*). Adapted from (Szenker et al. 2011) with modifications

response, they form signaling modules recognized by specific readers. Importantly, a paradigm of epigenetic inheritance is illustrated by centromere identity, which so far has been, with the exception of budding yeast, not determined by a specific DNA sequence but by the presence of the H3 variant CenH3$^{CENP\text{-}A}$ in most eukaryotes (Black and Bassett 2008; Valente et al. 2012). However, one should also bear in mind that epigenetics was a concept that emerged in the context of development, when Waddington described it as the branch of biology which studies the causal interactions between genes and their products which bring the phenotype into being (Waddington 1942). He joined the terms genetics and epigenesis to explain the progressively irreversible nature of differentiation during development. Epigenesis is the process by which animals, plants and fungi develop from an egg, seed or spore through a sequence of steps (Aristotle 350 BC). While there was no molecular feature behind the concept at the time, this was before the discovery of the DNA double helix (Watson and Crick 1953), and this vision has stimulated much work and has regained interest today.

Much progress has been made at a cellular level concerning the mechanisms of deposition of histone H3 variants and their functions related to genome organization. However, their roles during development are only beginning to be uncovered. We review here recent contributions to the field, illustrating the importance of H3 variants and their chaperones for developmental processes that involve both major genome reorganization, as well as more discrete changes to sustain cell fate specification.

2 H3 Dynamics During Preimplantation and Early Development

2.1 Gametogenesis

Gametogenesis gives rise to one of two highly specialized cell types through meiosis and postmeiotic maturation: the sperm and the oocyte. During this overall process, major histone variant dynamics are involved to contribute to meiotic recombination, followed by chromatin remodeling to condense and reset the genome and inactivate transcription in mature gametes in a distinct manner when dealing with female or male gametogenesis. These processes of whole genome rearrangement and reprogramming are examples of major events at the chromatin level that are essential in order to prepare the next steps of development (summarized in Table 2).

Starting with Male Gametogenesis In *C. elegans*, H3.3 is present in mature sperm chromatin and associated with the female genome throughout meiosis, despite the fact that histones are generally removed during spermatogenesis (Ooi et al. 2006). H3.3 is required for fertility in *D. melanogaster*, as males and females lacking functional copies of H3.3A and H3.3B genes are viable but sterile (Hödl and Basler 2009; Sakai et al. 2009). The role of H3.3 for fertility is linked to meiosis

Table 2 Histone H3 variants and their chaperones during metazoan development

Process	Replicative H3 and chaperones	H3.3 and chaperones	CenH3 and chaperones
Spermatogenesis	*Xl*: retention of H3 variants/H4 (Katagiri and Ohsumi 1994) *Mm*: Retention of replicative H3 could propagate transcriptional repression between generations (Erkek et al. 2013)	*Dm*: An H3 variant expressed throughout the cell cycle is required for male fertility (Hödl and Basler 2009; Hödl and Basler 2012). H3.3 essential for chromosome segregation/remodeling during and after meiosis (Sakai et al. 2009) *Mm*: H3.3 required for spermatogenesis (Bush et al. 2013; Tang et al. 2013); H3.3, HIRA (van der Heijden et al. 2007) and DAXX (Rogers et al. 2004) accumulate at the XY body	*Ce*: $CenH3^{HCP-3}$ not propagated (Gassmann et al. 2012) *Dm*: CAL1-dependent loading before, during and after meiosis, $CenH3^{CID}$ retained in mature sperm (Dunleavy et al. 2012; Raychaudhuri et al. 2012) *Xl* and *Mm*: $CenH3^{CENP-A}$ present in sperm (Palmer et al. 1990; Zeitlin et al. 2005)
Oogenesis	*Mm*: H3.1 absent during oocyte maturation (Kawamura et al. 2012)	*Dm*: H3.3 required for female fertility (Hödl and Basler 2009; Sakai et al. 2009) *Mm*: H3.3 accumulates in oocyte euchromatin (Akiyama et al. 2011)	*Ce*: $CenH3^{HCP-3}$ removal and reloaded during meiotic prophase (Gassmann et al. 2012) *Dm*: gradual $CenH3^{CID}$ loading in meiotic prophase (Dunleavy et al. 2012)
Fertilization		*Dm*: global deposition of H3.3 on paternal genome (Loppin et al. 2005), dependent on HIRA (Bonnefoy et al. 2007; Loppin et al. 2005; Orsi et al. 2013), Yemanuclein (Orsi et al. 2013) and CHD1 (Konev et al. 2007) *Ca*: HIRA essential for the decondensation of sperm nucleus (Zhao et al. 2011) *Mm*: H3.3 and HIRA specifically enriched in paternal chromatin during pronucleus formation (van der Heijden et al. 2005); in female pronucleus, fertilization triggers removal of accumulated maternal H3.3 (Akiyama et al. 2011)	*Dm*: paternal $CenH3^{CID}$ essential for propagation of the paternal genome, maternal $CenH3^{CID}$ deposited on paternal centromeres soon after fertilization (Raychaudhuri et al. 2012)

(continued)

Table 2 (continued)

Process	Replicative H3 and chaperones	H3.3 and chaperones	CenH3 and chaperones
Early/Pre-implantation development	*Dm*: H3.2 is dispensable, DSC incorporation can be functionally replaced by H3.3 (Hödl and Basler 2012) *Xl*: H3.2 and p150 required for early gastrulation (Szenker et al. 2012) *Mm*: CAF-1 dependent H3.2 DSC incorporation (Akiyama et al. 2011), no H3.1 until morula stage (Kawamura et al. 2012); CAF-1 p150 null embryos arrest before 16-cell stage with decondensed pHC (Houlard et al. 2006)	*Dm*: H3.3 (Hödl and Basler 2009) and HIRA (Bonnefoy et al. 2007) are not required for development up to adult, role in transcription of H3.3 can be replaced by H3.2 (Hödl and Basler 2012; Sakai et al. 2009) *Mm*: H3.3 restricted to euchromatin from late preimplantation onwards (Akiyama et al. 2011) and H3.3K36 important to counterbalance genome compaction (Lin et al. 2013); H3.3K27 required for establishment of pericentric heterochromatin and developmental progression (Santenard et al. 2010)	*Ce*: CenH3^{HCP-3} incorporation into domains not transcribed during gametogenesis and early embryogenesis (Gassmann et al. 2012) *Dm*: CenH3^{CID} deposition in anaphase (Schuh et al. 2007)
Pluripotency and reprogramming to pluripotency	*Dm*: Retention of existing H3.2 during asymmetric division of germline stem cells (Tran et al. 2012) *Mm*: exchange with oocyte-derived H3.2/H3.1 independently of DNA synthesis in transplanted somatic nuclei (Nashun et al. 2011)	*Xl*: H3.3K4 required for the stability of gene expression patterns (Ng and Gurdon 2008); HIRA-mediated H3.3 deposition necessary for induced expression of pluripotency genes (Jullien et al. 2012) during reprogramming after nuclear transfer *Mm*: H3.3 of donor nuclei is rapidly exchanged with oocyte-derived protein upon nuclear transfer (Nashun et al. 2011); HIRA mediated H3.3 incorporation dispensable for ES cell self-renewal (Goldberg et al. 2010) but becomes essential upon differentiation for PRC2 distribution and cell fate restriction (Banaszynski et al. 2013); accumulation of HIRA and H3.3-associated PTMs in PGC nuclei that undergo reprogramming (Hajkova et al. 2008)	*Mm*: ES cells tolerate CenH3^{CENP-A} depletion to a larger extent than differentiated ones (Ambartsumyan et al. 2010)
Extraembryonic tissue derivation		Mm: ATRX required for formation of extraembryonic trophoblast (Garrick et al. 2006)	

Gastrulation and mesoderm induction		*Xl*: H3.3 expression peaks at gastrulation, HIRA mediated deposition required for late gastrulation and mesoderm differentiation (Lim et al. 2013; Szenker et al. 2012) *Mm*: HIRA-null embryos display abnormal gastrulation and die by E10–11 (Roberts et al. 2002) DAXX-null embryos don't differentiate and die before E9.5 (Michaelson et al. 1999)	*Mm*: $CenH3^{CENP-A}$ null embryos lose centromeric chromatin organization and die by E6.5 (Howman et al. 2000)
Mesodermal organogenesis		*Dr*: proper H3.3 incorporation required for ectomesenchyme development (Cox et al. 2012) *Mm*: HIRA and ASF1a mediate H3.3 incorporation to establish MyoD expression (Yang et al. 2011b) and control cell type-specific genes during myoblast conversion into osteoblasts or myotubes (Song et al. 2012); CHD2-dependent deposition of H3.3 at myogenic loci to facilitate gene expression (Harada et al. 2012); Cabin1 antagonizes HIRA and ASF1a-dependent activation of Mef2-target genes (Yang et al. 2011a)	*Mm*: $CenH3^{CENP-A}$ downregulated upon differentiation into myotubes (Yang et al. 2011b); cardiac precursors require $CenH3^{CENP-A}$ to sustain proliferation (McGregor et al. 2014)
Neurogenesis	*Ce*: CAF-1 mediated H3 deposition supports asymmetric cell division in the nervous system (Nakano et al. 2011)	*Mm*: H3.3 disappears from the bodies of pluripotency genes upon differentiation of ESCs to neural precursor cells (Goldberg et al. 2010); CHD1 depletion drives stem cells towards neurogenic differentiation (Gaspar-Maia et al. 2009); DAXX associated with regulatory regions of activity-regulated genes in cortical neural cells (Michod et al. 2012); ATRX critical for neuron survival during corticogenesis (Berube et al. 2005) contributes to maintain telomere and pericentric heterochromatin integrity in proliferating brain cells (Watson et al. 2013)	*Dm*: in asymmetric divisions of larval brain stem cells $CenH3^{CID}$ deposition occurs earlier in stem daughter than in differentiating daughter cell (Dunleavy et al. 2012)

A comparative view of existing studies across model organisms and major developmental events
Dr Danio rerio; *Ca Carassius auratus*

and postmeiotic genome remodeling, as both processes are impaired in mutant males (Sakai et al. 2009). Initially, experiments in *D. melanogaster* to rescue the phenotype with tagged H3.2 failed (Hödl and Basler 2009; Sakai et al. 2009). However, expression of untagged H3.2 from the *H3.3B* promoter can rescue the infertility of *H3.3A* and *B* double-mutants, suggesting that the specific role of H3.3 during gametogenesis in flies could be related to its availability outside of S-phase (Hödl and Basler 2012). This indicates that the availability of a histone H3 outside of S-phase is critical for gametogenesis in this model organism. In mice, H3.3 is highly expressed in ovaries and testes (Couldrey et al. 1999), and hypomorphic or null *H3.3A* alleles decrease male fertility in homozygous state (Couldrey et al. 1999; Tang et al. 2013). Given that loss of one *H3.3B* gene copies is sufficient to disrupt sperm maturation (Tang et al. 2013) and homozygous KO leads to complete infertility in males (Bush et al. 2013) (Fig. 3), this paralog could be of predominant importance for spermatogenesis. The H3.3 variant might be crucial here to orchestrate correct gene expression and repression of repetitive elements. This operates together with histone methylation (Siklenka et al. 2015) and DNA methylation (Kubo et al. 2015) to allow for correct orchestration of sperm maturation.

Although the exact role of H3.3 during mammalian spermatogenesis remains to be delineated in more detail, it is important in *D. melanogaster*, which shares similarities with the mammalian system. The presence of H3.3 and two of its dedicated chaperones, HIRA (van der Heijden et al. 2007) and DAXX (Rogers et al. 2004), in the transcriptionally inactive XY body, may point towards a potential function of H3.3 during mammalian spermatogenesis. Formation of the XY body, known as meiotic sex chromosome inactivation (MSCI), is required for normal meiosis to proceed, as aberrant MSCI leads to arrest in pachytene and apoptosis (Royo et al. 2010).

Repackaging of the male genome through the replacement of histones with protamines and protamine-like proteins is a challenge for the transgenerational propagation of chromatin marks. Although the majority of histones are lost in human and mouse sperm chromatin, particular sites in the genome, such as developmentally regulated promoters, certain noncoding RNAs, paternally expressed imprinted *loci* and regions in HOX clusters, retain histones and their associated marks (Brykczynska et al. 2010; Hammoud et al. 2009). In addition, H3K27me3 is significantly enriched at developmental promoters that are repressed in early embryos, including many promoters with bivalent (H3K4me3/H3K27me3) marks in ES cells. Later in embryogenesis, there is an accumulation of repressive histone modifications including H3K9me2/3, H3K27me3 and H3K20me3 in early embryos (Hatanaka et al. 2015; Voon et al. 2015), where the distribution of H3.3 may play a role (Elsässer et al. 2015). Retained H3 in mature sperm corresponds predominantly to H3.3 at unmethylated CpG dinucleotides at CpG island (CGI) promoters and H3.1/H3.2 at non-CGI due to high and low histone turnover, respectively (Erkek et al. 2013). A role of H3.1/H3.2 retention associated with Polycomb-mediated H3K27me3 may be implicated in propagating transcriptional repression between

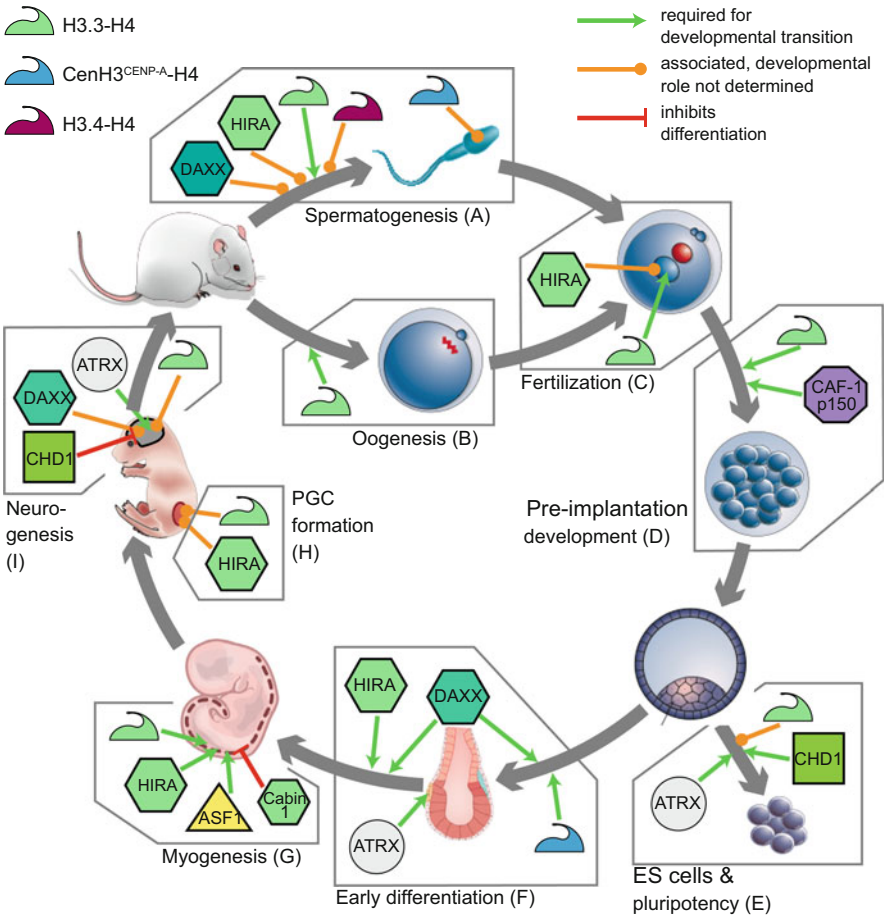

Fig. 3 The importance of H3 variants and their chaperones during various stages of mouse development. Mouse reproduction begins with the fusion of two highly differentiated gametes (**a** and **b**) into a zygote (**c**). This cell rapidly acquires totipotency and starts dividing (**d** and **e**), giving rise to daughter cells that will specialize progressively (**f**). The diverse cell lineages they establish will differentiate into the array of tissues in the adult organism (**g**, **i**). Among these lineages, primary germ cells (PGCs) undergo reprogramming to establish the germline of the adult (**h**), allowing it to produce either male or female gametes like those it originated from. H3 variants and their chaperones have been shown to contribute to the regulation of these processes (*green* or *red arrows*). For some, their presence or accumulation has been described at certain stages, but their contribution to developmental progression remains elusive (*orange*)

generations (Erkek et al. 2013). Further effort is needed to decipher further how variants and modifications are retained and how this can contribute to cellular fate and development of the whole organism.

Concerning Female Gametogenesis In female mice H3.3 is incorporated into the euchromatic regions of fully grown oocytes where it persists during meiosis

(Akiyama et al. 2011) and *H3.3B* KO results in almost uniform infertility (Bush et al. 2013) (Fig. 3). Surprisingly, H3.1 is lacking from chromatin during oocyte maturation and during the first part of preimplantation development (Kawamura et al. 2012). This shift from H3.1 to H3.3 might reflect an important feature of reprogramming or transcription in the absence of replication during oocyte maturation. Recent evidence suggests that continuous histone replacement by HIRA is essential for *de novo* DNA methylation and normal transcriptional regulation during mouse oogenesis (Nashun et al. 2015).

In mammals (Albig et al. 1996) and hominids (Schenk et al. 2011), the testis-specific H3 variants H3.4 and H3.5 are incorporated genome-wide at late meiosis or at early stages of spermiogenesis. H3.4 forms unstable nucleosomes (Tachiwana et al. 2010) and this could possibly contribute to its own removal during spermiogenesis to facilitate replacement with protamines. Moreover, these variants are associated with euchromatin when ectopically expressed in somatic cells (Schenk et al. 2011; Tachiwana et al. 2010) and this suggests a possible role in gene expression in particular for testis-specific genes in the germline. *In vitro* studies demonstrated that *h*uman *N*ucleosome *a*ssembly *p*rotein *2* (hNap2) promotes the deposition of H3.4-H4 (Tachiwana et al. 2008), but whether this is also the case *in vivo* remains to be determined.

The Centromere Paradigm CenH3 is though to be critical for centromere identity throughout metazoans [for reviews see (Boyarchuk et al. 2011; Müller and Almouzni 2014)] and retaining the identity of this *locus* poses a challenge in the male germline. *C. elegans* has holocentric chromosomes, which are marked by $CenH3^{HCP-3}$. Interestingly, $CenH3^{HCP-3}$ is not retained in sperm but is removed and subsequently reloaded during oogenic meiotic prophase (Gassmann et al. 2012). The lack of stable $CenH3^{HCP-3}$ transmission during gametogenesis in *C. elegans* could be explained by specific mechanisms to re-establish $CenH3^{HCP-3}$ distribution in early *C. elegans* embryos (Gassmann et al. 2012), as well as the capacity of meiotic kinetochores to form independently of $CenH3^{HCP-3}$ in oocytes (Monen et al. 2005). In *X. laevis*, where male gametes retain H3–H4 (Katagiri and Ohsumi 1994), $CenH3^{CENP-A}$ has been found in sperm (Zeitlin et al. 2005). This has also been shown in mammals (Palmer et al. 1990), where only a minor fraction of total nucleosomes is retained (Brykczynska et al. 2010; Hammoud et al. 2009). However, the mechanisms involved in $CenH3^{CENP-A}$ maintenance and deposition during gametogenesis remains to be analyzed in detail. In *D. melanogaster* $CenH3^{CID}$ survives the extensive chromatin remodeling during the compaction of the haploid genome into sperm heads (Dunleavy et al. 2012; Raychaudhuri et al. 2012). Its retention is essential for the propagation of the paternal genome in the next generation, as the absence of paternal $CenH3^{CID}$ prevents recruitment of maternal $CenH3^{CID}$ and formation of functional kinetochores during the first mitosis, which leads to the development of gynogenetic haploid embryos (Raychaudhuri et al. 2012). This demonstrates the importance of $CenH3^{CID}$ as a transgenerational mark in flies. *C*hromosome *al*ignment defect *1* (CAL1), the $CenH3^{CID}$ assembly factor in *D. melanogaster* (Chen et al. 2014), is required for the retention of the

synaptonemal complex at centromeres, to allow proper centromere clustering and homologous chromosome pairing (Unhavaithaya and Orr-Weaver 2013). Cell-cycle control is a major determinant of $CenH3^{CID}$ deposition. Loading occurs during G2 before the onset of male meiosis (Raychaudhuri et al. 2012) and over a period of days during male and female prophase (Dunleavy et al. 2012). In male prophase I, $CenH3^{CID}$ signal gradually increases above what is required to compensate for the twofold dilution in premeiotic S-phase, drops by more than half in the absence of DNA replication between the end of meiosis I and the beginning of meiosis II, and increases again after exit from meiosis II, in spermatids (Dunleavy et al. 2012).

Taken together, in several metazoan models H3 variants have proven essential for gametogenesis. A future challenge will be to characterize the mechanisms governing their eviction or retention during genome repackaging and dissect their exact contribution to meiotic recombination.

2.2 *Fertilization*

Following fertilization, the female genome having just completed meiosis and the male genome in a condensed, inactive state, undergo major differential reprogramming in the same cytoplasm, in order to attain totipotency. Following this, both genomes replicate and prepare for the first embryonic mitosis. Spatial separation of the two genomes, maintained as individual pronuclei, allows paternal chromatin to shed its protamines and massively incorporate histones of maternal origin in a DSI manner, while maternal chromatin can preserve its somatic-like organization (Probst and Almouzni 2011).

In *D. melanogaster*, *C. elegans* and mouse, H3.3 of maternal origin is globally deposited in the paternal chromatin upon fertilization, before the first round of replication and onset of detectable zygotic transcription (Loppin et al. 2005; Ooi et al. 2006; Torres-Padilla et al. 2006; van der Heijden et al. 2005). In *D. melanogaster*, this is the only developmental process involving H3.3 that has been shown to strictly require HIRA (Bonnefoy et al. 2007). Both yemanuclein-α and the remodeling factor chromodomain helicase DNA-binding protein 1 (CHD1), which associates with HIRA *in vivo* (Konev et al. 2007) have a contributing role (Orsi et al. 2013) (Table 1). HIRA is also present in the decondensing sperm nucleus in the crucian carp (Zhao et al. 2011) and in mouse (van der Heijden et al. 2005), where H3.3 is equally detected (Torres-Padilla et al. 2006), indicating a conserved role during fertilization in vertebrates. Interestingly, transcription of ribosomal RNA is required for the first cleavage after fertilization and depends on HIRA/H3.3 (Dutta et al. 2010). Contrary to this paradigm, *X. laevis* paternal chromatin retains H3 variants and H4 histones, possibly allowing sperm decondensation in the absence of global H3.3 DSI assembly. Using nucleosome-depleted paternal pronuclei by depleting maternal histone H3.3 or its chaperone

HIRA in mouse zygotes, recent findings show that a pronucleus devoid of nucleosomes forms a nuclear envelope devoid of nuclear pore complexes (Inoue and Zhang 2014). Thus, H3.3 nucleosomes also exhibit a crucial role for the overall structure of the nucleus, impacting development at this level.

In mouse, pericentric heterochromatin, which is essential for accurate chromosome segregation, assembles *de novo* to ensure competence for the first mitotic division. This requires that the H3.3 incorporated massively on the paternal genome display an intact lysine 27 (Santenard et al. 2010), a site that is methylated by the Polycomb group proteins associated with heterochromatin (Puschendorf et al. 2008). Mutation of H3.3 at this site, but not of replicative H3, results in dysfunctional chromosome segregation, developmental arrest aberrant accumulation of pericentric transcripts and heterochromatin protein 1 (HP1) mislocalization (Santenard et al. 2010) (Fig. 3). Importantly, a specific burst of transcription from pericentric repeats is required for early developmental progression (Casanova et al. 2013; Probst et al. 2010). It remains to be determined whether specific accumulation of H3.3 relates to this transcription and in particular, whether the differential H3.3 localization at pericentric domains (Santenard and Torres-Padilla 2009) depends on the distinct regulation of major satellite repeat transcription that operates on each parental genome. Given that the integrity of pericentric heterochromatin influences the deposition of $CenH3^{CENP\text{-}A}$ in cultured mouse cells (Boyarchuk et al. 2014), and the fact that H3.3 also forms an integral part of centromeric chromatin (Dunleavy et al. 2011), proper H3.3 dynamics at these sites may influence centromeric integrity during subsequent developmental stages. In the mouse female pronucleus, parthenogenetic activation and fertilization trigger the removal of accumulated maternal H3.3. This global histone deficit is reversed when H3.2 and H3.3 accumulate in the nuclei of embryos from the one-cell stage onwards (Akiyama et al. 2011). Loss of the H3.3 chromatin mark transmitted on the maternal genome could thus participate in the initiation of a new gene expression pattern in the totipotent zygote.

In *C. elegans*, the removal of $CenH3^{HCP\text{-}3}$ during spermatogenesis necessitates its subsequent *de novo* deposition in the embryo. The importance of gene expression in defining new $CenH3^{HCP\text{-}3}$ domains is indicated by the fact that its incorporation is prevented in regions transcribed during spermatogenesis and early development (Gassmann et al. 2012). In *D. melanogaster*, maternally derived $CenH3^{CID}$ is loaded onto paternal centromeres soon after fertilization. Interestingly, the amount of $CenH3^{CID}$ retained during spermatogenesis dictates how much $CenH3^{CID}$ will be incorporated after subsequent mitoses up to adult stage (Raychaudhuri et al. 2012). Thus, not only the site of centromere formation, but also the stoichiometry of its chromatin components might be determined transgenerationally.

Taken together, this data illustrates that H3 variants participate both to erase and reprogram existing chromatin marks, as well as to perpetuate the identity of a given genomic *locus* across generations.

2.3 Early Development

Mutation of the two H3.3 genes in *D. melanogaster* leads to partial lethality, but most animals deficient in H3.3 surviving to adulthood appear morphologically normal (Hödl and Basler 2009; Sakai et al. 2009). HIRA can be dispensable for H3.3 deposition in *D. melanogaster* embryos and adult cells (Bonnefoy et al. 2007), possibly because an overexpression of H3.2 or potential promiscuous use of other histone chaperones can compensate for this defect. Mutants lacking H3.2 entirely but harboring an S-phase-expressed H3.3 can survive into the first instar larval stage and show no obvious patterning defects (Hödl and Basler 2012). The proposed gap-filling mechanism for H3.3 deposition carried out by HIRA or by Xnp, the *D. melanogaster* ATRX homolog (Ray-Gallet et al. 2011; Schneiderman et al. 2012) could explain these observations. Thus, in *D. melanogaster*, replicative or replacement H3 variants can compensate for the lack of each other. In early mouse embryos Stella, initially identified through its role in maintaining the DNA methylation status (Nakamura et al. 2012), also controls the expression of DAXX, which directly influences chromatin reorganization (Arakawa et al. 2015). This dual role highlights how a careful balance between the histone H3 deposition network, histone modifications and DNA methylation orchestrate chromatin organization and gene expression in the early embryo.

Contrary to *D. melanogaster*, DSI and DSC modes of histone deposition are not interchangeable during development of vertebrate models studied so far. In *X. laevis*, inhibition of replicative H3.2 incorporation with a CAF-1 p150 dominant negative mutant (Quivy et al. 2001) or morpholinos that target both H3.2 and H3.3 (Szenker et al. 2012) leads to developmental arrest at the mid-blastula transition or early gastrulation stage, respectively. In mouse preimplantation embryos, distinct transcriptional (and potentially posttranscriptional) regulation leads to the absence of H3.1 incorporation until the morula stage (Kawamura et al. 2012). Therefore, DSC during the first cell cycles likely relies on H3.2. In mouse preimplantation embryos, depletion of CAF-1 p150 leads to the complete loss of H3.1 and H3.2 and an increase in the incorporation of H3.3. However, this compensatory mechanism is insufficient, as p150 depleted embryos fail to reach blastocyst stage, display nuclear enlargement and lack heterochromatin foci (Akiyama et al. 2011) (Fig. 3). Recent data showed that CAF-1 protects human embryonic stem cells (Yang et al. 2015) and preimplantation mouse embryos (Hatanaka et al. 2015) from endogenous retrotransposons by mediating histone H3.1/H3.2 deposition. This process of histone variant replacement is associated in the early embryo with accumulation of repressive histone modifications including H3K9me2/3, H3K27me3 and H3K20me3, illustrating the role of CAF-1 in retrotransposon silencing (Hatanaka et al. 2015). Interestingly, CAF-1 also plays a role in lineage differentiation in early *C. elegans* embryogenesis, by regulating Notch target gene expression and consequently the choice of cell fate (Du et al. 2015).

H3.3 knockdown induces arrest at the morula stage that can be rescued exclusively by H3.3 with an intact K36 residue, but not replicative H3. Furthermore, loss

of H3.3K36 decreases H4K16 acetylation, leading to H1-mediated overcondensation of the genome (Lin et al. 2013). Strikingly, null embryos for HIRA, DAXX and ATRX die at different stages post-implantation as late as E11 (Fig. 3) and display divergent phenotypes (Garrick et al. 2006; Michaelson et al. 1999; Roberts et al. 2002). This difference with respect to the phenotype of H3.3 loss might arise from a redundancy of DSI pathways, until specific functions of each chaperone become indispensable after implantation. Disruption of one of the two H3.3 isoforms leads to postnatal death in the case of *H3.3A* (Couldrey et al. 1999) and lethality in the second half of embryogenesis for *H3.3B* (Bush et al. 2013). However, both phenotypes have incomplete penetrance. Since conditional knock-out alleles for the two H3.3 genes are available now (Bush et al. 2013; Tang et al. 2013), recent advances have shed light on the role of H3.3 in mammalian differentiation and organogenesis (Tang et al. 2015). *H3.3A* null mutants are viable to adulthood, but whereas females were fertile, males are sub-fertile exhibiting dysmorphic spermatozoa. *H3.3B* null mutants are growth-deficient, dying at birth. *H3.3B* heterozygotes were also growth-deficient, with males being sterile because of arrest of round spermatids. Simultaneous ablation of H3f3a and H3f3b in folliculogenesis results in early primary oocyte death, demonstrating a crucial role for H3.3 in oogenesis (Tang et al. 2015). Taken together, the control of a balance for a fine-tuning of histone H3.1/H3.2 and H3.3 deposition using their dedicated chaperones CAF-1 or HIRA/ATRX/DAXX, respectively, represents a means to exert a direct control over differentiation programs in the early metazoan embryo.

In *D. melanogaster*, the timing of $CenH3^{CID}$ deposition is developmentally regulated. In early embryos, deposition occurs in anaphase (Schuh et al. 2007). However, later during development, in somatic mitoses, new $CenH3^{CID}$ is assembled in late telophase/early G1 (Dunleavy et al. 2012). Whether similar differences exist in other organisms is currently not known, due to the fact that most studies have been carried out only in cell culture. Mouse development can proceed partially in the absence of new $CenH3^{CENP\text{-}A}$ expression (Howman et al. 2000) or in the presence of a dysfunctional GFP fusion protein (Kalitsis et al. 2003). These mutant embryos succumb to a loss of centromeric chromatin organization, chromosome missegregation, aneuploidy and apoptosis, at E6.5 and E10.5 respectively. The delayed lethality suggests that $CenH3^{CENP\text{-}A}$ of zygotic origin might be required only from a certain point in development onwards, which might be due to specific mechanisms of centromere propagation in pluripotent cells (Ambartsumyan et al. 2010), or the availability of a maternally provided protein pool during preimplantation development. Following exhaustion of this stock, $CenH3^{CENP\text{-}A}$ is diluted twofold from the centromere at every passage through the cell cycle. Since centromere function and propagation can occur with severely reduced $CenH3^{CENP\text{-}A}$ levels (Fachinetti et al. 2013), a number of divisions may be required before mitotic failure eventually occurs. These data underline the importance of a developmental context and to compare how $CenH3^{CENP\text{-}A}$ incorporation occurs in embryos with respect to data derived from cell culture, both in terms of cell cycle timing and the function of HJURP, will be a major future avenue of research.

3 H3 Dynamics During Cell Fate Specification

3.1 Pluripotency and Reprogramming

Nuclear transfer (NT) of somatic nuclei into oocytes, which was initially developed by Briggs and King in frogs (Briggs and King 1952), has provided one of the most striking examples of reprogramming and continues to provide valuable insight into the regulation of cell identity.

NT into *X. laevis* oocytes revealed a role for H3.3 both in maintaining the expression of genes that do not undergo reprogramming, as well as the induction of pluripotency genes (Jullien et al. 2012; Ng and Gurdon 2008). In mouse, donor cell-derived H3.1, H3.2 and H3.3 of the transplanted nucleus are rapidly replaced with maternal protein in a DSI manner (Nashun et al. 2011). Maternal-derived H3.3 replaces H3 in the donor nucleus shortly after oocyte activation, with the amount of replacement directly related to the differentiation status of the donor nucleus in somatic cell nuclear transfer embryos (Wen et al. 2014b). Thus, this replacement could possibly occur at specific *loci* that are programmed for gene activation. H3.3 knockdown in mouse oocytes leads to compromised reprogramming and down-regulation of key pluripotency genes, which can be rescued by injecting exogenous H3.3 mRNA. Importantly, maternal H3.3, and not H3.3 in the donor nucleus, is essential for reprogramming the somatic cell nucleus (Wen et al. 2014a). H3.1 is incorporated in NT embryos (Nashun et al. 2011) but not in fertilized ones (Akiyama et al. 2011). The differential H3 dynamics outlined by NT experiments may represent a means to improve the success of reprogramming which could help in producing induced pluripotent stem (iPS) cells. Indeed, by studying transcription-factor-mediated reprogramming of mouse fibroblasts to iPS using an siRNA screen, CAF-1 emerged a regulator of somatic cell identity during transcription-factor-induced cell-fate transitions allowing to modulate cellular plasticity in a regenerative setting (Cheloufi et al. 2015). In this study, optimal modulation of both CAF-1 and transcription factor levels increased reprogramming efficiency and facilitated iPS cell formation. Cells that resemble 2-cell-stage blastomeres can arise in embryonic stem cell cultures, illustrating spontaneous return to a totipotent state. The occurrence of this rare event can be significantly amplified by down-regulating CAF-1 (Ishiuchi et al. 2015), suggesting that totipotent cells might display an atypical balance of histone variants.

H3.3 appears to be dispensable for mouse ES cell self-renewal (Banaszynski et al. 2013; Goldberg et al. 2010). However, H3.3 becomes critical upon differentiation, with a significant impact on the expression of developmentally regulated bivalent genes (Banaszynski et al. 2013). Similarly, $CenH3^{CENP\text{-}A}$ depletion is tolerated in mouse ES cells, but this tolerance is lost upon differentiation (Ambartsumyan et al. 2010), which reflects a difference in the dynamics or functional importance of H3 variants for chromatin organization related to developmental potential (Fig. 3).

HIRA-mediated H3.3 deposition is required to maintain elevated levels of H3K27me3 (but not H3K4me3) at bivalent promoters, by creating a favorable chromatin environment for recruiting PRC2. Genes encoding for transcription factors essential for trophectoderm establishment are upregulated upon H3.3 or HIRA depletion and H3.3-depleted ES cells have a propensity to form trophectoderm-like teratomas (Banaszynski et al. 2013). Therefore, HIRA-deposited H3.3 may act as a barrier to prevent differentiation of ES cells into otherwise restricted cell lineages. H3.3 marks the enhancer regions of all-trans retinoic acid (tRA)-induced genes in ES cells prior to their activation, to maintain an open chromatin state, which is necessary for the binding of the RAR/RXR transcription factor. H3.3 is displaced from the enhancer and begins to accumulate at the promoter of the respective genes (Chen et al. 2013) after induction with tRA. Changes in H3.3 distribution at developmentally regulated genes may not be mere consequences of transcription, but could also play an upstream role by decreasing H1 binding and increasing template accessibility, given the antagonistic effects of H3.3 and H1 on chromatin compaction in the early embryo (Lin et al. 2013). The correlation between increased expression of the histone *H3.3B* gene with increased levels of pluripotency markers distinguishes pluripotent human embryonic stem (ES) cells from early-stage differentiated hES cell populations (Laslett et al. 2007). In *D. melanogaster*, CHD1, which is involved in H3.3 incorporation at fertilization (Konev et al. 2007), is required in mouse ES cells to maintain their pluripotent state. CHD1-deficient ES cells have a high propensity for neural differentiation but lose the ability to give rise to primitive endoderm. CHD1 is also required for efficient reprogramming of fibroblasts to the pluripotent stem cell state (Gaspar-Maia et al. 2009). In mouse $Hira^{-/-}$ ES cells, the accumulation of H3.3 is perturbed at promoters and in the body of active genes but not at telomeres. The latter specifically require ATRX for H3.3 enrichment, in order to repress the *te*lomeric *r*epeat-containing *RNA* (TERRA) (Goldberg et al. 2010). ATRX also localizes to silenced imprinted alleles in mouse ES cells, and its knock-out impairs H3.3 deposition at these sites, leading to loss of H3K9me3 and thus loss of repression and aberrant allelic expression (Voon et al. 2015). The DAXX/ATRX complex also associates to endogenous retroviral elements (ERV) containing long terminal repeats, where it is involved in depositing H3.3. H3K9me3 at these sites, which normally represses ERVs, is reduced upon deletion of H3.3 (Elsässer et al. 2015), suggesting that DAXX/ATRX contributes to histone variant deposition at ERVs in ES cells, and thus influencing local histone methylation. However, this mechanism has to be dissected further, considering that an independent study showed an involvement of CAF-1 and H3.1 dynamics controlling local H3K9me3 in order to silence endogenous retrotransposons (Hatanaka et al. 2015).

A natural setting to observe dynamic changes of the stem cell state is the developing gonad, as primordial germ cells (PGCs) reprogram and sequentially differentiate into gametogenic precursors. Blastomere mitoses that give rise to the PGC are accompanied by a gradual loss of H3.3 in *C. elegans*, which is potentially linked to transcriptional quiescence (Ooi et al. 2006). In *D. melanogaster*, asymmetric division of germline stem cells is reflected by an asymmetric distribution of

canonical H3.2, but not of H3.3. Whereas newly synthesized histones are enriched in the differentiating daughter cell, the stem daughter cell selectively retains pre-existing H3.2, (Tran et al. 2012). This asymmetric histone distribution is lost when progenitor cells are manipulated to divide symmetrically. This result suggests that retention of pre-existing canonical histones during asymmetric cell divisions contributes to the maintenance of stem cell identity. In mouse, PGCs formed in the proximal epiblast undergo large-scale replacement of histone variants and PTMs and erasure of DNA methylation before populating the primordial gonads (Surani et al. 2007) (Fig. 3). HIRA accumulates in PGC nuclei that undergo reprogramming and is accompanied by enrichment of histone PTMs preferentially associated with H3.3 (Hajkova et al. 2008). Therefore, it will be interesting to study the dynamics of H3.3 in PGCs and its contribution to the changes in chromatin that allow reacquisition of a pluripotent state, potentially using developed culture systems that enable monitoring of PGC reprogramming *in vitro* (Leitch et al. 2013).

3.2 Specific Differentiation Programs

During chicken development, a gradual change from H3.2 to H3.3 occurs (Urban and Zweidler 1983), which continues after hatching until H3.3 becomes predominant over H3.2 in tissues that have stopped dividing. Although initially interpreted as a mere consequence of decreasing cell proliferation, further studies summarized below have provided additional insights into discrete roles for H3.3 during organogenesis.

3.2.1 Mesoderm Induction

Depletion of H3.3 leads to defects in late gastrulation in *X. laevis*, including an open blastopore, *spina bifida*, and a shortened anteroposterior axis. The lack of rescue of H3.3 depletion upon overexpression of H3.2 indicates a differential role for the two H3 variants in this context (Szenker et al. 2012). However, whether it is H3.3 *per se*, or DSI deposition via HIRA that is critically required at this developmental stage remains an open question. Embryos may be more dependent on a DSI pathway to maintain chromatin organization from this point onwards, given the lengthening of the cell cycle at gastrulation. Although H3.3 depletion does not affect early mesoderm inducing signals, H3.3 is critical for the proper expression of later mesodermal differentiation markers (Szenker et al. 2012). Potential crosstalk with the deposition of H2A.Z, also required for *X. laevis* gastrulation (Ridgway et al. 2004), could also be relevant. This suggests that HIRA-dependent H3.3 incorporation could help in maintaining mesoderm signals after their induction. Since a similar role for HIRA has been described in gibel carp, where HIRA morpholinos delay gastrulation, leading to aberrant somitogenesis (Wang

et al. 2014), this role during gastrulation could represent a common phenomenon across vertebrates.

3.2.2 Development of Mesodermal Derivatives

Several steps in the pathway of muscle differentiation from mesodermal precursors require deposition of H3.3. In the murine C2C12 system, in order to enable its transcriptional activation in myotubes, H3.3 is incorporated in regulatory regions of the MyoD gene, which is a master regulator of myogenesis. The mechanism of activation likely involves the acquisition of a permissive chromatin state to recruit polymerase II to the promoter (Yang et al. 2011b). Similarly, prior to their activation induced by differentiation, promoters of MyoD target genes also show H3.3 deposition (Harada et al. 2012). This involves *M*yocyte *e*nhancer *f*actor *2* (Mef2), a transcriptional coactivator of MyoD targets, which favors transcription during muscle differentiation by interacting with HIRA in cooperation with ASF1. In addition, Cabin1, a component of the HIRA complex (Rai et al. 2011), is downregulated upon myoblast differentiation. It is possible that its exit from the complex could allow HIRA and ASF1 to activate expression of Mef2 target genes (Yang et al. 2011a). Further data supports the important role of HIRA and ASF1a for differentiation of myoblasts. Whereas HIRA and ASF1a levels are maintained during conversion into myotubes, their levels diminish during conversion into osteoblasts (Song et al. 2012). Therefore, they may play a role in the choice of cell fate followed by the myogenic precursors. In vertebrates, cranial neural crest cells (CNCs) express a gene repertoire characteristic of mesoderm derivatives but originate in the ectoderm, and they give rise to the skeleton and muscles of the head. Expression of a H3.3 mutant leads to perturbed expression of CNC-specific markers and impaired ectomesenchyme formation in zebrafish (Cox et al. 2012). Taken together, in these model organisms the developmental potential of CNCs depends on proper H3.3 incorporation.

Mutations in the *ATRX locus* are responsible for ATR-X syndrome (Gibbons et al. 1995), a developmental disorder whose clinical presentation also includes skeletal, facial and urogenital anomalies. Consistent with symptoms of ATR-X syndrome, mice with targeted KO of *Atrx* in the mesenchyme display brachydactyly, attributed to both apoptosis and DNA damage accumulation in the developing limb bud (Solomon et al. 2013). Since chondrocytes lacking ATRX support normal skeletal development (Solomon et al. 2009) and do not accumulate DNA damage (Solomon et al. 2013), the role of ATRX on cell survival has to be considered in a cell-type specific manner. It will be interesting to explore further the links between ATRX and H3.3 deposition, and how they contribute together or separately to developmental phenotypes.

In the context of organogenesis, little is known about CenH3 dynamics, but consistent with a role for chromosome segregation in mitotic cells, an increased level usually correlates with undifferentiated state and proliferative capacity. $CenH3^{CENP\text{-}A}$ is downregulated in postmitotic cells, which no longer require a

functional centromere as found upon differentiation of myoblasts into myotubes (Yang et al. 2011b). Along the same lines, while cardiac precursor cells (CPCs) need CenH3^{CENP-A} in order to sustain their proliferation, CenH3^{CENP-A} is depleted from centromeres as these cells age and lose their regenerative potential (Goldberg et al. 2010). This latter case suggests that centromere integrity could be an important parameter to consider in the context of stem cell therapies.

3.2.3 Neuronal Differentiation

In the developing *C. elegans* nervous system CAF-1 mediated deposition of replicative H3 is required for the asymmetric cell division that forms the basis for bilateral asymmetry (Nakano et al. 2011). It is unknown whether asymmetric histone distribution occurs in this model of differentiation. In mice, indirect evidence links H3.3 to neurogenesis (Fig. 3), where its enrichment at the transcription start site (TSS) of bivalent genes in ES cells is retained at the TSS and spreads along the gene body only in the subset of genes that become expressed upon differentiation into *n*euronal *p*recursor *c*ells (NPCs) (Goldberg et al. 2010). In order to maintain adequate levels of H3K27me3, promoters that regulate neural tube mesenchyme and ear development in mouse rely on H3.3 incorporation (Banaszynski et al. 2013). In murine cortical neural cells DAXX is associated with the regulatory regions of a subset of activity-regulated genes, where calcium-dependent dephosphorylation regulates H3.3 loading and transcription upon neuronal activation (Michod et al. 2012). During corticogenesis ATRX is critical for neuron survival (Berube et al. 2005) and its absence from NPCs leads to DNA damage at telomeres and pericentric heterochromatin and replicative stress. These defects arise in mitotically active structures of the developing brain, but disappear after proliferation has ceased postnatally (Watson et al. 2013). The exact role of these factors, their differential recruitment and the importance of H3.3 deposition remain to be deciphered. The daughter cells generated by the asymmetric division of neuroblast stem cells are characterized by a different timing of CenH3^{CID} incorporation in the developing *D. melanogaster* brain. In the daughter cell which will go on to self-renew as a neuroblast, CenH3^{CID} assembly at mitotic exit occurs sooner than in the other daughter cell, which will differentiate into a neuron (Dunleavy et al. 2012). Notably, H3.3 dynamics are essential for the plasticity of both embryonic and adult neurons (Maze et al. 2015), and their regulation is only just beginning to be explored. Understanding this regulation is also crucial for the understanding of the development of cancers associated with the nervous system. Whole-genome sequencing studies have identified mutations in both genes encoding for H3.3 in several cancers. Pediatric glioblastomas (GBM) harbor frequent H3.3K27M, and G34R/G34V somatic mutations in *H3F3A* (Schwartzentruber et al. 2012; Wu et al. 2012), each residue characterizing tumors formed in a different anatomic compartment (Sturm et al. 2012) [for a review see (Fontebasso et al. 2013)]. The K27M substitution has also been detected in pediatric glioma of thalamic (Aihara

et al. 2014) or pontine origin, correlated with poorer survival rates in the latter (Khuong-Quang et al. 2012).

3.2.4 Lymphocyte Differentiation

Two examples concerning lymphocyte differentiation also illustrate situations where H3 variants may play specific roles. During lineage commitment of CD4+ T cells to helper T2 cells (TH2, as opposed to TH1), H3K9me3, a repressive mark usually underrepresented on H3.3, is present at the promoters of silenced TH1 genes. The SUV39H1-H3K9me3-HP1α pathway participates in maintaining the silencing of TH1 *loci* and functions to ensure stable TH2 identity (Allan et al. 2012). Given the associations between the H3K9me3 mark and replicative H3 variants, a role for H3.1/H3.2 in T cell differentiation could be the case. H3.3 is present at V(D)J recombining *loci* where it is heavily acetylated and carries the H3.3-specific S31 phosphorylation. This modified H3.3 becomes ubiquitylated by the recombination activating gene 1 (RAG1) recombinase (Jones et al. 2011) as a potential mechanism to recruit the DNA repair machinery. The distribution of H3.3 and the histone chaperone complex FACT across the *Igh locus* correlate with the efficiency of somatic hypermutation (Aida et al. 2013), which might be due to an increased histone turnover as a potential mark for successful sites of activation-induced cytidine deaminase (AID) binding. These findings underline that histone variants may play a role in the development of the immune system, opening exciting avenues for future studies.

4 Conclusion

The examples from the literature presented herein illustrate how major developmental transitions exploit replacement of H3 variants and remodeling. Genome-wide histone turnover and H3.3 incorporation appear to be critical for gametogenesis and fertilization, potentially providing a mechanism to reset the chromatin landscape of highly specialized gametes to a totipotent state in the zygote. The roles of H3 variants during mammalian oogenesis are only beginning to emerge. A range of studies has emphasized the role of H3.3 during early development and in pluripotency, but more remains to be learned about it in later developmental stages. In this context, it will be important to consider how H3.3 cooperates with other histone variants. Our current view of how different chaperones control H3 variant distribution across the genome during development will certainly evolve. This will rely on a refined mapping and assessment of the respective contributions of the deposition pathways, their dynamics and crosstalk. The key for such advances lies in bridging the gap between the *in vitro* methodology that has been used to characterize the dynamics of H3 variants and the *in vivo* approaches of developmental biology.

Acknowledgements We apologize for not having been able to acknowledge all the colleagues who contributed to this work. This work was supported by la Ligue Nationale contre le Cancer (Equipe labellisée Ligue), the European Commission Network of Excellence EpiGeneSys (HEALTH-F4-2010-257082), ERC Advanced Grant 2009-AdG_20090506 "Eccentric", the European Commission large-scale integrating project FP7_HEALTH-2010-259743 "MODHEP", ANR-11-LABX-0044_DEEP and ANR-10-IDEX-0001-02 PSL, ANR "CHAPINHIB" ANR-12-BSV5-0022-02, ANR "Epicure" ANR-14-CE16-0009, ANR "CELLECTCHIP" ANR-14-CE10-0013 and Aviesan-ITMO cancer project "Epigenomics of breast cancer".

References

Ai X, Parthun MR (2004) The nuclear Hat1p/Hat2p complex: a molecular link between type B histone acetyltransferases and chromatin assembly. Mol Cell 14:195–205

Aida M, Hamad N, Stanlie A, Begum NA, Honjo T (2013) Accumulation of the FACT complex, as well as histone H3.3, serves as a target marker for somatic hypermutation. Proc Natl Acad Sci USA 110:7784–7789

Aihara K, Mukasa A, Gotoh K, Saito K, Nagae G et al (2014) H3F3A K27M mutations in thalamic gliomas from young adult patients. Neuro Oncol 16:140–146

Akiyama T, Suzuki O, Matsuda J, Aoki F (2011) Dynamic replacement of histone H3 variants reprograms epigenetic marks in early mouse embryos. PLoS Genet 7:e1002279

Albig W, Ebentheuer J, Klobeck G, Kunz J, Doenecke D (1996) A solitary human H3 histone gene on chromosome 1. Hum Genet 97:486–491

Allan RS, Zueva E, Cammas F, Schreiber HA, Masson V et al (2012) An epigenetic silencing pathway controlling T helper 2 cell lineage commitment. Nature 487:249–253

Ambartsumyan G, Gill RK, Perez SD, Conway D, Vincent J et al (2010) Centromere protein A dynamics in human pluripotent stem cell self-renewal, differentiation and DNA damage. Hum Mol Genet 19:3970–3982

Andrews AJ, Luger K (2011) Nucleosome structure(s) and stability: variations on a theme. Annu Rev Biophys 40:99–117

Arakawa T, Nakatani T, Oda M, Kimura Y, Sekita Y et al (2015) Stella controls chromocenter formation through regulation of Daxx expression in 2-cell embryos. Biochem Biophys Res Commun 466:60–65

Aristotle Stagiritis son of Nicomachus. 350 BC. On the generation of animals

Banaszynski LA, Wen D, Dewell S, Whitcomb SJ, Lin M et al (2013) Hira-dependent histone H3.3 deposition facilitates PRC2 recruitment at developmental loci in ES cells. Cell 155:107–120

Banumathy G, Somaiah N, Zhang R, Tang Y, Hoffmann J et al (2009) Human UBN1 is an ortholog of yeast Hpc2p and has an essential role in the HIRA/ASF1a chromatin-remodeling pathway in senescent cells. Mol Cell Biol 29:758–770

Barnhart MC, Kuich PH, Stellfox ME, Ward JA, Bassett EA et al (2011) HJURP is a CENP-A chromatin assembly factor sufficient to form a functional de novo kinetochore. J Cell Biol 194:229–243

Belotserkovskaya R, Oh S, Bondarenko VA, Orphanides G, Studitsky VM, Reinberg D (2003) FACT facilitates transcription-dependent nucleosome alteration. Science 301:1090–1093

Berube NG, Mangelsdorf M, Jagla M, Vanderluit J, Garrick D et al (2005) The chromatin-remodeling protein ATRX is critical for neuronal survival during corticogenesis. J Clin Invest 115:258–267

Black BE, Bassett EA (2008) The histone variant CENP-A and centromere specification. Curr Opin Cell Biol 20:91–100

Bonnefoy E, Orsi GA, Couble P, Loppin B (2007) The essential role of Drosophila HIRA for de novo assembly of paternal chromatin at fertilization. PLoS Genet 3:1991–2006

Bortvin A, Winston F (1996) Evidence that Spt6p controls chromatin structure by a direct interaction with histones. Science 272:1473–1476

Boyarchuk E, Montes de Oca R, Almouzni G (2011) Cell cycle dynamics of histone variants at the centromere, a model for chromosomal landmarks. Curr Opin Cell Biol 23:266–276

Boyarchuk E, Filipescu D, Vassias I, Cantaloube S, Almouzni G (2014) The histone variant composition of centromeres is controlled by the pericentric heterochromatin state during the cell cycle. J Cell Sci 127:3347–3359

Briggs R, King TJ (1952) Transplantation of living nuclei from blastula cells into enucleated frogs' eggs. Proc Natl Acad Sci USA 38:455–463

Brykczynska U, Hisano M, Erkek S, Ramos L, Oakeley EJ et al (2010) Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. Nat Struct Mol Biol 17:679–687

Bush KM, Yuen BT, Barrilleaux BL, Riggs JW, O'Geen H et al (2013) Endogenous mammalian histone H3.3 exhibits chromatin-related functions during development. Epigenetics Chromatin 6:7

Campos EI, Fillingham J, Li G, Zheng H, Voigt P et al (2010) The program for processing newly synthesized histones H3.1 and H4. Nat Struct Mol Biol 17:1343–1351

Casanova M, Pasternak M, El Marjou F, Le Baccon P, Probst AV, Almouzni G (2013) Heterochromatin reorganization during early mouse development requires a single-stranded noncoding transcript. Cell Rep 4:1156–1167

Cheloufi S, Elling U, Hopfgartner B, Jung YL, Murn J et al (2015) The histone chaperone CAF-1 safeguards somatic cell identity. Nature 528:218–224

Chen P, Zhao J, Wang Y, Wang M, Long H et al (2013) H3.3 actively marks enhancers and primes gene transcription via opening higher-ordered chromatin. Genes Dev 27:2109–2124

Chen CC, Dechassa ML, Bettini E, Ledoux MB, Belisario C et al (2014) CAL1 is the Drosophila CENP-A assembly factor. J Cell Biol 204:313–329

Cook AJ, Gurard-Levin ZA, Vassias I, Almouzni G (2011) A specific function for the histone chaperone NASP to fine-tune a reservoir of soluble H3–H4 in the histone supply chain. Mol Cell 44:918–927

Couldrey C, Carlton MB, Nolan PM, Colledge WH, Evans MJ (1999) A retroviral gene trap insertion into the histone 3.3A gene causes partial neonatal lethality, stunted growth, neuromuscular deficits and male sub-fertility in transgenic mice. Hum Mol Genet 8:2489–2495

Cox SG, Kim H, Garnett AT, Medeiros DM, An W, Crump JG (2012) An essential role of variant histone H3.3 for ectomesenchyme potential of the cranial neural crest. PLoS Genet 8:e1002938

Daniel Ricketts M, Frederick B, Hoff H, Tang Y, Schultz DC et al (2015) Ubinuclein-1 confers histone H3.3-specific-binding by the HIRA histone chaperone complex. Nat Commun 6:7711

De Koning L, Corpet A, Haber JE, Almouzni G (2007) Histone chaperones: an escort network regulating histone traffic. Nat Struct Mol Biol 14:997–1007

DeNizio JE, Elsässer SJ, Black BE (2014) DAXX co-folds with H3.3/H4 using high local stability conferred by the H3.3 variant recognition residues. Nucleic Acids Res 42:4318–4331

Drane P, Ouararhni K, Depaux A, Shuaib M, Hamiche A (2010) The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. Genes Dev 24:1253–1265

Du Z, Santella A, He F, Shah PK, Kamikawa Y, Bao Z (2015) The regulatory landscape of lineage differentiation in a metazoan embryo. Dev Cell 34:592–607

Dunleavy EM, Roche D, Tagami H, Lacoste N, Ray-Gallet D et al (2009) HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. Cell 137:485–497

Dunleavy EM, Almouzni G, Karpen GH (2011) H3.3 is deposited at centromeres in S phase as a placeholder for newly assembled CENP-A in G(1) phase. Nucleus 2:146–157

Dunleavy EM, Beier NL, Gorgescu W, Tang J, Costes SV, Karpen GH (2012) The cell cycle timing of centromeric chromatin assembly in Drosophila meiosis is distinct from mitosis yet requires CAL1 and CENP-C. PLoS Biol 10:e1001460

Dutta D, Ray S, Home P, Saha B, Wang S et al (2010) Regulation of angiogenesis by histone chaperone HIRA-mediated incorporation of lysine 56-acetylated histone H3.3 at chromatin domains of endothelial genes. J Biol Chem 285:41567–41577
Elsässer SJ, Huang H, Lewis PW, Chin JW, Allis CD, Patel DJ (2012) DAXX envelops a histone H3.3-H4 dimer for H3.3-specific recognition. Nature 491:560–565
Elsässer SJ, Noh KM, Diaz N, Allis CD, Banaszynski LA (2015) Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells. Nature 522:240–244
Erkek S, Hisano M, Liang CY, Gill M, Murr R et al (2013) Molecular determinants of nucleosome retention at CpG-rich sequences in mouse spermatozoa. Nat Struct Mol Biol 20:868–875
Fachinetti D, Folco HD, Nechemia-Arbely Y, Valente LP, Nguyen K et al (2013) A two-step mechanism for epigenetic specification of centromere identity and function. Nat Cell Biol 15:1056–1066
Filipescu D, Müller S, Almouzni G (2014) Histone H3 variants and their chaperones during development and disease: contributing to epigenetic control. Annu Rev Cell Dev Biol 30:615–646
Foltz DR, Jansen LE, Bailey AO, Yates JR 3rd, Bassett EA et al (2009) Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. Cell 137:472–484
Fontebasso AM, Liu XY, Sturm D, Jabado N (2013) Chromatin remodeling defects in pediatric and young adult glioblastoma: a tale of a variant histone 3 tail. Brain Pathol 23:210–216
Franklin RE, Gosling RG (1953) Molecular configuration in sodium thymonucleate. Nature 171:740–741
Franklin SG, Zweidler A (1977) Non-allelic variants of histones 2a, 2b and 3 in mammals. Nature 266:273–275
Fudenberg G, Mirny LA (2012) Higher-order chromatin structure: bridging physics and biology. Curr Opin Genet Dev 22:115–124
Garrick D, Sharpe JA, Arkell R, Dobbie L, Smith AJ et al (2006) Loss of Atrx affects trophoblast development and the pattern of X-inactivation in extraembryonic tissues. PLoS Genet 2:e58
Gaspar-Maia A, Alajem A, Polesso F, Sridharan R, Mason MJ et al (2009) Chd1 regulates open chromatin and pluripotency of embryonic stem cells. Nature 460:863–868
Gassmann R, Rechtsteiner A, Yuen KW, Muroyama A, Egelhofer T et al (2012) An inverse relationship to germline transcription defines centromeric chromatin in *C. elegans*. Nature 484:534–537
Gaume X, Torres-Padilla ME (2015) Regulation of reprogramming and cellular plasticity through histone exchange and histone variant incorporation. Cold Spring Harb Symp Quant Biol pi:027458
Gibbons RJ, Picketts DJ, Villard L, Higgs DR (1995) Mutations in a putative global transcriptional regulator cause X-linked mental retardation with alpha-thalassemia (ATR-X syndrome). Cell 80:837–845
Goldberg AD, Banaszynski LA, Noh KM, Lewis PW, Elsaesser SJ et al (2010) Distinct factors control histone variant H3.3 localization at specific genomic regions. Cell 140:678–691
Gurard-Levin ZA, Quivy JP, Almouzni G (2014) Histone chaperones: assisting histone traffic and nucleosome dynamics. Annu Rev Biochem 83:487–517
Hajkova P, Ancelin K, Waldmann T, Lacoste N, Lange UC et al (2008) Chromatin dynamics during epigenetic reprogramming in the mouse germ line. Nature 452:877–881
Hammoud SS, Nix DA, Zhang H, Purwar J, Carrell DT, Cairns BR (2009) Distinctive chromatin in human sperm packages genes for embryo development. Nature 460:473–478
Harada A, Okada S, Konno D, Odawara J, Yoshimi T et al (2012) Chd2 interacts with H3.3 to determine myogenic cell fate. EMBO J 31:2994–3007
Hatanaka Y, Inoue K, Oikawa M, Kamimura S, Ogonuki N et al (2015) Histone chaperone CAF-1 mediates repressive histone modifications to protect preimplantation mouse embryos from endogenous retrotransposons. Proc Natl Acad Sci USA 112:14641–14646
Heitz E (1928) Das Heterochromatin der Moose. I Jahrb Wiss Botan 69:762–818
Hödl M, Basler K (2009) Transcription in the absence of histone H3.3. Curr Biol 19:1221–1226

Hödl M, Basler K (2012) Transcription in the absence of histone H3.2 and H3K4 methylation. Curr Biol 22:2253–2257

Houlard M, Berlivet S, Probst AV, Quivy JP, Hery P et al (2006) CAF-1 is essential for heterochromatin organization in pluripotent embryonic cells. PLoS Genet 2:e181

Howman EV, Fowler KJ, Newson AJ, Redward S, MacDonald AC et al (2000) Early disruption of centromeric chromatin organization in centromere protein A (Cenpa) null mice. Proc Natl Acad Sci USA 97:1148–1153

Huang S, Zhou H, Katzmann D, Hochstrasser M, Atanasova E, Zhang Z (2005) Rtt106p is a histone chaperone involved in heterochromatin-mediated silencing. Proc Natl Acad Sci USA 102:13410–13415

Inoue A, Zhang Y (2014) Nucleosome assembly is required for nuclear pore complex assembly in mouse zygotes. Nat Struct Mol Biol 21:609–616

Ishiuchi T, Enriquez-Gasca R, Mizutani E, Boskovic A, Ziegler-Birling C et al (2015) Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly. Nat Struct Mol Biol 22:662–671

Jones JM, Bhattacharyya A, Simkus C, Vallieres B, Veenstra TD, Zhou M (2011) The RAG1 V(D)J recombinase/ubiquitin ligase promotes ubiquitylation of acetylated, phosphorylated histone 3.3. Immunol Lett 136:156–162

Jullien J, Astrand C, Szenker E, Garrett N, Almouzni G, Gurdon JB (2012) HIRA dependent H3.3 deposition is required for transcriptional reprogramming following nuclear transfer to Xenopus oocytes. Epigenetics Chromatin 5:17

Kalashnikova AA, Porter-Goff ME, Muthurajan UM, Luger K, Hansen JC (2013) The role of the nucleosome acidic patch in modulating higher order chromatin structure. J R Soc Interface 10:20121022

Kalitsis P, Fowler KJ, Earle E, Griffiths B, Howman E et al (2003) Partially functional Cenpa-GFP fusion protein causes increased chromosome missegregation and apoptosis during mouse embryogenesis. Chromosome Res 11:345–357

Katagiri C, Ohsumi K (1994) Remodeling of sperm chromatin induced in egg extracts of amphibians. Int J Dev Biol 38:209–216

Kaufman PD, Kobayashi R, Kessler N, Stillman B (1995) The p150 and p60 subunits of chromatin assembly factor I: a molecular link between newly synthesized histones and DNA replication. Cell 81:1105–1114

Kawamura M, Akiyama T, Tsukamoto S, Suzuki MG, Aoki F (2012) The expression and nuclear deposition of histone H3.1 in murine oocytes and preimplantation embryos. J Reprod Dev 58:557–562

Khuong-Quang DA, Buczkowicz P, Rakopoulos P, Liu XY, Fontebasso AM et al (2012) K27M mutation in histone H3.3 defines clinically and biologically distinct subgroups of pediatric diffuse intrinsic pontine gliomas. Acta Neuropathol 124:439–447

Kleinschmidt JA, Fortkamp E, Krohne G, Zentgraf H, Franke WW (1985) Co-existence of two different types of soluble histone complexes in nuclei of *Xenopus laevis* oocytes. J Biol Chem 260:1166–1176

Konev AY, Tribus M, Park SY, Podhraski V, Lim CY et al (2007) CHD1 motor protein is required for deposition of histone variant H3.3 into chromatin in vivo. Science 317:1087–1090

Kornberg RD (1974) Chromatin structure: a repeating unit of histones and DNA. Science 184:868–871

Kubo N, Toh H, Shirane K, Shirakawa T, Kobayashi H et al (2015) DNA methylation and gene expression dynamics during spermatogonial stem cell differentiation in the early postnatal mouse testis. BMC Genomics 16:624

Lacoste N, Woolfe A, Tachiwana H, Garea AV, Barth T et al (2014) Mislocalization of the centromeric histone variant CenH3/CENP-A in human cells depends on the chaperone DAXX. Mol Cell 53:631–644

Laslett AL, Grimmond S, Gardiner B, Stamp L, Lin A et al (2007) Transcriptional analysis of early lineage commitment in human embryonic stem cells. BMC Dev Biol 7:12

Latreille D, Bluy L, Benkirane M, Kiernan RE (2014) Identification of histone 3 variant 2 interacting factors. Nucleic Acids Res 42:3542–3550

Le S, Davis C, Konopka JB, Sternglanz R (1997) Two new S-phase-specific genes from *Saccharomyces cerevisiae*. Yeast 13:1029–1042

Leitch HG, McEwen KR, Turp A, Encheva V, Carroll T et al (2013) Naive pluripotency is associated with global DNA hypomethylation. Nat Struct Mol Biol 20:311–316

Lewis PW, Elsaesser SJ, Noh KM, Stadler SC, Allis CD (2010) Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. Proc Natl Acad Sci USA 107:14075–14080

Lim CY, Reversade B, Knowles BB, Solter D (2013) Optimal histone H3 to linker histone H1 chromatin ratio is vital for mesodermal competence in Xenopus. Development 140:853–860

Lin CJ, Conti M, Ramalho-Santos M (2013) Histone variant H3.3 maintains a decondensed chromatin state essential for mouse preimplantation development. Development 140:3624–3634

Liu CP, Xiong C, Wang M, Yu Z, Yang N et al (2012) Structure of the variant histone H3.3-H4 heterodimer in complex with its chaperone DAXX. Nat Struct Mol Biol 19:1287–1292

Loppin B, Bonnefoy E, Anselme C, Laurencon A, Karr TL, Couble P (2005) The histone H3.3 chaperone HIRA is essential for chromatin assembly in the male pronucleus. Nature 437:1386–1390

Loyola A, Almouzni G (2007) Marking histone H3 variants: how, when and why? Trends Biochem Sci 32:425–433

Loyola A, LeRoy G, Wang YH, Reinberg D (2001) Reconstitution of recombinant chromatin establishes a requirement for histone-tail modifications during chromatin assembly and transcription. Genes Dev 15:2837–2851

Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature 389:251–260

Luger K, Dechassa ML, Tremethick DJ (2012) New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? Nat Rev Mol Cell Biol 13:436–447

Marino-Ramirez L, Kann MG, Shoemaker BA, Landsman D (2005) Histone structure and nucleosome stability. Expert Rev Proteomics 2:719–729

Mattiroli F, D'Arcy S, Luger K (2015) The right place at the right time: chaperoning core histone variants. EMBO Rep 16:1454–1466

Maze I, Wenderski W, Noh KM, Bagot RC, Tzavaras N et al (2015) Critical role of histone turnover in neuronal transcription and plasticity. Neuron 87:77–94

McGhee JD, Felsenfeld G (1980) Nucleosome structure. Annu Rev Biochem 49:1115–1156

McGregor M, Hariharan N, Joyo AY, Margolis RL, Sussman MA (2014) CENP-A is essential for cardiac progenitor cell proliferation. Cell Cycle 13:739–748

Mello JA, Sillje HH, Roche DM, Kirschner DB, Nigg EA, Almouzni G (2002) Human Asf1 and CAF-1 interact and synergize in a repair-coupled nucleosome assembly pathway. EMBO Rep 3:329–334

Michaelson JS, Bader D, Kuo F, Kozak C, Leder P (1999) Loss of Daxx, a promiscuously interacting protein, results in extensive apoptosis in early mouse development. Genes Dev 13:1918–1923

Michod D, Bartesaghi S, Khelifi A, Bellodi C, Berliocchi L et al (2012) Calcium-dependent dephosphorylation of the histone chaperone DAXX regulates H3.3 loading and transcription upon neuronal activation. Neuron 74:122–135

Monen J, Maddox PS, Hyndman F, Oegema K, Desai A (2005) Differential role of CENP-A in the segregation of holocentric *C. elegans* chromosomes during meiosis and mitosis. Nat Cell Biol 7:1248–1255

Müller S, Almouzni G (2014) A network of players in H3 histone variant deposition and maintenance at centromeres. Biochim Biophys Acta 1839:241–250

Müller S, Montes de Oca R, Lacoste N, Dingli F, Loew D, Almouzni G (2014) Phosphorylation and DNA binding of HJURP determine its centromeric recruitment and function in CenH3 (CENP-A) loading. Cell Rep 8:190–203

Munakata T, Adachi N, Yokoyama N, Kuzuhara T, Horikoshi M (2000) A human homologue of yeast anti-silencing factor has histone chaperone activity. Genes Cells 5:221–233

Nakamura T, Liu YJ, Nakashima H, Umehara H, Inoue K et al (2012) PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. Nature 486:415–419

Nakano S, Stillman B, Horvitz HR (2011) Replication-coupled chromatin assembly generates a neuronal bilateral asymmetry in *C. elegans*. Cell 147:1525–1536

Nashun B, Akiyama T, Suzuki MG, Aoki F (2011) Dramatic replacement of histone variants during genome remodeling in nuclear-transferred embryos. Epigenetics 6:1489–1497

Nashun B, Hill PW, Smallwood SA, Dharmalingam G, Amouroux R et al (2015) Continuous histone replacement by Hira is essential for normal transcriptional regulation and de novo DNA methylation during mouse oogenesis. Mol Cell 60:611–625

Ng RK, Gurdon JB (2008) Epigenetic memory of an active gene state depends on histone H3.3 incorporation into chromatin in the absence of transcription. Nat Cell Biol 10:102–109

Ooi SL, Priess JR, Henikoff S (2006) Histone H3.3 variant dynamics in the germline of *Caenorhabditis elegans*. PLoS Genet 2:e97

Orsi GA, Algazeery A, Meyer RE, Capri M, Sapey-Triomphe LM et al (2013) Drosophila Yemanuclein and HIRA cooperate for de novo assembly of H3.3-containing nucleosomes in the male pronucleus. PLoS Genet 9:e1003285

Oudet P, Gross-Bellard M, Chambon P (1975) Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. Cell 4:281–300

Palmer DK, O'Day K, Margolis RL (1990) The centromere specific histone CENP-A is selectively retained in discrete foci in mammalian sperm nuclei. Chromosoma 100:32–36

Perpelescu M, Hori T, Toyoda A, Misu S, Monma N et al (2015) HJURP is involved in the expansion of centromeric chromatin. Mol Biol Cell 26:2742–2754

Probst AV, Almouzni G (2011) Heterochromatin establishment in the context of genome-wide epigenetic reprogramming. Trends Genet 27:177–185

Probst AV, Okamoto I, Casanova M, El Marjou F, Le Baccon P, Almouzni G (2010) A strand-specific burst in transcription of pericentric satellites is required for chromocenter formation and early mouse development. Dev Cell 19:625–638

Puschendorf M, Terranova R, Boutsma E, Mao X, Isono K et al (2008) PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos. Nat Genet 40:411–420

Quivy JP, Grandi P, Almouzni G (2001) Dimerization of the largest subunit of chromatin assembly factor 1: importance in vitro and during Xenopus early development. EMBO J 20:2015–2027

Rai TS, Puri A, McBryan T, Hoffman J, Tang Y et al (2011) Human CABIN1 is a functional member of the human HIRA/UBN1/ASF1a histone H3.3 chaperone complex. Mol Cell Biol 31:4107–4118

Raychaudhuri N, Dubruille R, Orsi GA, Bagheri HC, Loppin B, Lehner CF (2012) Transgenerational propagation and quantitative maintenance of paternal centromeres depends on Cid/Cenp-A presence in Drosophila sperm. PLoS Biol 10:e1001434

Ray-Gallet D, Quivy JP, Scamps C, Martini EM, Lipinski M, Almouzni G (2002) HIRA is critical for a nucleosome assembly pathway independent of DNA synthesis. Mol Cell 9:1091–1100

Ray-Gallet D, Woolfe A, Vassias I, Pellentz C, Lacoste N et al (2011) Dynamics of histone H3 deposition in vivo reveal a nucleosome gap-filling mechanism for H3.3 to maintain chromatin integrity. Mol Cell 44:928–941

Richardson RT, Batova IN, Widgren EE, Zheng LX, Whitfield M et al (2000) Characterization of the histone H1-binding protein, NASP, as a cell cycle-regulated somatic protein. J Biol Chem 275:30378–30386

Ridgway P, Brown KD, Rangasamy D, Svensson U, Tremethick DJ (2004) Unique residues on the H2A.Z containing nucleosome surface are important for *Xenopus laevis* development. J Biol Chem 279:43815–43820

Roberts C, Sutherland HF, Farmer H, Kimber W, Halford S et al (2002) Targeted mutagenesis of the Hira gene results in gastrulation defects and patterning abnormalities of mesoendodermal derivatives prior to early embryonic lethality. Mol Cell Biol 22:2318–2328

Rogers RS, Inselman A, Handel MA, Matunis MJ (2004) SUMO modified proteins localize to the XY body of pachytene spermatocytes. Chromosoma 113:233–243

Royo H, Polikiewicz G, Mahadevaiah SK, Prosser H, Mitchell M et al (2010) Evidence that meiotic sex chromosome inactivation is essential for male fertility. Curr Biol 20:2117–2123

Russo V, Martienssen RA, Riggs AD (1996) Epigenetic mechanisms of gene regulation. Cold Spring Harbor Laboratory Press, Plainview, NY

Sakai A, Schwartz BE, Goldstein S, Ahmad K (2009) Transcriptional and developmental functions of the H3.3 histone variant in Drosophila. Curr Biol 19:1816–1820

Santenard A, Torres-Padilla ME (2009) Epigenetic reprogramming in mammalian reproduction: contribution from histone variants. Epigenetics 4:80–84

Santenard A, Ziegler-Birling C, Koch M, Tora L, Bannister AJ, Torres-Padilla ME (2010) Heterochromatin formation in the mouse embryo requires critical residues of the histone variant H3.3. Nat Cell Biol 12:853–862

Sawatsubashi S, Murata T, Lim J, Fujiki R, Ito S et al (2010) A histone chaperone, DEK, transcriptionally coactivates a nuclear receptor. Genes Dev 24:159–170

Schenk R, Jenke A, Zilbauer M, Wirth S, Postberg J (2011) H3.5 is a novel hominid-specific histone H3 variant that is specifically expressed in the seminiferous tubules of human testes. Chromosoma 120:275–285

Schneiderman JI, Orsi GA, Hughes KT, Loppin B, Ahmad K (2012) Nucleosome-depleted chromatin gaps recruit assembly factors for the H3.3 histone variant. Proc Natl Acad Sci USA 109:19721–19726

Schuh M, Lehner CF, Heidmann S (2007) Incorporation of Drosophila CID/CENP-A and CENP-C into centromeres during early embryonic anaphase. Curr Biol 17:237–243

Schwartzentruber J, Korshunov A, Liu XY, Jones DT, Pfaff E et al (2012) Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. Nature 482:226–231

Sharp JA, Fouts ET, Krawitz DC, Kaufman PD (2001) Yeast histone deposition protein Asf1p requires Hir proteins and PCNA for heterochromatic silencing. Curr Biol 11:463–473

Shen X, Ranallo R, Choi E, Wu C (2003) Involvement of actin-related proteins in ATP-dependent chromatin remodeling. Mol Cell 12:147–155

Shuaib M, Ouararhni K, Dimitrov S, Hamiche A (2010) HJURP binds CENP-A via a highly conserved N-terminal domain and mediates its deposition at centromeres. Proc Natl Acad Sci USA 107:1349–1354

Siklenka K, Erkek S, Godmann M, Lambrot R, McGraw S et al (2015) Disruption of histone methylation in developing sperm impairs offspring health transgenerationally. Science 350: aab2006

Smith S, Stillman B (1989) Purification and characterization of CAF-I, a human cell factor required for chromatin assembly during DNA replication in vitro. Cell 58:15–25

Solomon LA, Li JR, Berube NG, Beier F (2009) Loss of ATRX in chondrocytes has minimal effects on skeletal development. PLoS One 4:e7106

Solomon LA, Russell BA, Watson LA, Beier F, Berube NG (2013) Targeted loss of the ATR-X syndrome protein in the limb mesenchyme of mice causes brachydactyly. Hum Mol Genet 22:5015–5025

Song TY, Yang JH, Park JY, Song Y, Han JW et al (2012) The role of histone chaperones in osteoblastic differentiation of C2C12 myoblasts. Biochem Biophys Res Commun 423:726–732

Sturm D, Witt H, Hovestadt V, Khuong-Quang DA, Jones DT et al (2012) Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. Cancer Cell 22:425–437

Surani MA, Hayashi K, Hajkova P (2007) Genetic and epigenetic regulators of pluripotency. Cell 128:747–762

Szenker E, Ray-Gallet D, Almouzni G (2011) The double face of the histone variant H3.3. Cell Res 21:421–434

Szenker E, Lacoste N, Almouzni G (2012) A developmental requirement for HIRA-dependent H3.3 deposition revealed at gastrulation in Xenopus. Cell Rep 1:730–740

Szenker E, Boyarchuk E, Almouzni G (2014) Properties and functions of histone variants. In: Workman J (ed) Fundamentals of chromatin. Springer, Berlin

Tachiwana H, Osakabe A, Kimura H, Kurumizaka H (2008) Nucleosome formation with the testis-specific histone H3 variant, H3t, by human nucleosome assembly proteins in vitro. Nucleic Acids Res 36:2208–2218

Tachiwana H, Kagawa W, Osakabe A, Kawaguchi K, Shiga T et al (2010) Structural basis of instability of the nucleosome containing a testis-specific histone variant, human H3T. Proc Natl Acad Sci USA 107:10454–10459

Tachiwana H, Müller S, Blumer J, Klare K, Musacchio A, Almouzni G (2015) HJURP involvement in de novo CenH3(CENP-A) and CENP-C recruitment. Cell Rep 11:22–32

Tagami H, Ray-Gallet D, Almouzni G, Nakatani Y (2004) Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. Cell 116:51–61

Talbert PB, Ahmad K, Almouzni G, Ausio J, Berger F et al (2012) A unified phylogeny-based nomenclature for histone variants. Epigenetics Chromatin 5:7

Tang MC, Jacobs SA, Wong LH, Mann JR (2013) Conditional allelic replacement applied to genes encoding the histone variant H3.3 in the mouse. Genesis 51:142–146

Tang MC, Jacobs SA, Mattiske DM, Soh YM, Graham AN et al (2015) Contribution of the two genes encoding histone variant h3.3 to viability and fertility in mice. PLoS Genet 11:e1004964

Torres-Padilla ME, Bannister AJ, Hurd PJ, Kouzarides T, Zernicka-Goetz M (2006) Dynamic distribution of the replacement histone variant H3.3 in the mouse oocyte and preimplantation embryos. Int J Dev Biol 50:455–461

Tran V, Lim C, Xie J, Chen X (2012) Asymmetric division of Drosophila male germline stem cell shows asymmetric histone distribution. Science 338:679–682

Tyler JK, Adams CR, Chen SR, Kobayashi R, Kamakaka RT, Kadonaga JT (1999) The RCAF complex mediates chromatin assembly during DNA replication and repair. Nature 402:555–560

Unhavaithaya Y, Orr-Weaver TL (2013) Centromere proteins CENP-C and CAL1 functionally interact in meiosis for centromere clustering, pairing, and chromosome segregation. Proc Natl Acad Sci USA 110:19878–19883

Urban MK, Zweidler A (1983) Changes in nucleosomal core histone variants during chicken development and maturation. Dev Biol 95:421–428

Valente LP, Silva MC, Jansen LE (2012) Temporal control of epigenetic centromere specification. Chromosome Res 20:481–492

van der Heijden GW, Dieker JW, Derijck AA, Muller S, Berden JH et al (2005) Asymmetry in histone H3 variants and lysine methylation between paternal and maternal chromatin of the early mouse zygote. Mech Dev 122:1008–1022

van der Heijden GW, Derijck AA, Posfai E, Giele M, Pelczar P et al (2007) Chromosome-wide nucleosome replacement and H3.3 incorporation during mammalian meiotic sex chromosome inactivation. Nat Genet 39:251–258

Voon HP, Hughes JR, Rode C, De La Rosa-Velazquez IA, Jenuwein T et al (2015) ATRX plays a key role in maintaining silencing at interstitial heterochromatic loci and imprinted genes. Cell Rep 11:405–418

Waddington CH (1942) The epigenotype. Endeavour 1:18–20

Wang MY, Guo QH, Du XZ, Zhou L, Luo Q et al (2014) HIRA is essential for the development of gibel carp. Fish Physiol Biochem 40:235–244

Watson JD, Crick FH (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature 171:737–738

Watson LA, Solomon LA, Li JR, Jiang Y, Edwards M et al (2013) Atrx deficiency induces telomere dysfunction, endocrine defects, and reduced life span. J Clin Invest 123:2049–2063

Wen D, Banaszynski LA, Liu Y, Geng F, Noh KM et al (2014a) Histone variant H3.3 is an essential maternal factor for oocyte reprogramming. Proc Natl Acad Sci USA 111:7325–7330

Wen D, Banaszynski LA, Rosenwaks Z, Allis CD, Rafii S (2014b) H3.3 replacement facilitates epigenetic reprogramming of donor nuclei in somatic cell nuclear transfer embryos. Nucleus 5:369–375

Wu Q, Qian YM, Zhao XL, Wang SM, Feng XJ et al (2012) Expression and prognostic significance of centromere protein A in human lung adenocarcinoma. Lung Cancer 77:407–414

Yang JH, Choi JH, Jang H, Park JY, Han JW et al (2011a) Histone chaperones cooperate to mediate Mef2-targeted transcriptional regulation during skeletal myogenesis. Biochem Biophys Res Commun 407:541–547

Yang JH, Song Y, Seol JH, Park JY, Yang YJ et al (2011b) Myogenic transcriptional activation of MyoD mediated by replication-independent histone deposition. Proc Natl Acad Sci USA 108:85–90

Yang BX, El Farran CA, Guo HC, Yu T, Fang HT et al (2015) Systematic identification of factors for provirus silencing in embryonic stem cells. Cell 163:230–245

Zeitlin SG, Patel S, Kavli B, Slupphaug G (2005) Xenopus CENP-A assembly into chromatin requires base excision repair proteins. DNA Repair (Amst) 4:760–772

Zhao ZK, Li W, Wang MY, Zhou L, Wang JL, Wang YF (2011) The role of HIRA and maternal histones in sperm nucleus decondensation in the gibel carp and color crucian carp. Mol Reprod Dev 78:139–147

Epigenetics in Development, Differentiation and Reprogramming

Nuphar Salts and Eran Meshorer

1 Introduction

The creation of a complex multicellular organism begins with two haploid cells, the sperm and the egg, merging into one diploid cell. The two genomes are first kept separated as two pronuclei (Cantone and Fisher 2013), but soon thereafter join to form the zygote. This totipotent cell will then go through a series of mitotic cleavage divisions, first creating a small mass of cells called the blastula, and then (by day 3.5 in mouse or day 5.5 in human) form the blastocyst, comprised of the outer trophoblast cells and the pluripotent inner cell mass (ICM). The latter will subsequently differentiate into the three germ layers, creating the new developing organism. During differentiation, each cell expresses a different set of genes, according to its location, function, signaling cascades, etc. Since all cells share the exact same genome, epigenetic processes will dictate the silencing/activation of desired genes, and the maintenance of cellular states.

In early development, DNA demethylation plays an important role (Fig. 1). Several waves of global demethylation occur, permitting to reestablish the formation of patterns afresh. Later, as patterns and tissues are forming, the epigenetic landscape is being created, controlling and regulating differentiation. In order to explore development, researchers may follow the process in vivo at different organisms with somewhat large ethical limitations, especially when it comes to exploring higher organisms in general and humans in particular. For this reason, a methodology which uses pluripotent stem cells (PSCs) as a model for development has evolved, providing an endless resource of pluripotent cells for further characterization and

N. Salts • E. Meshorer (✉)
Department of Genetics, The Institute of Life Sciences, and The Edmond and Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem, Edmond J. Safra Campus, Jerusalem 91904, Israel
e-mail: meshorer@huji.ac.il

D.P. Bazett-Jones, G. Dellaire (eds.), *The Functional Nucleus*,
DOI 10.1007/978-3-319-38882-3_18

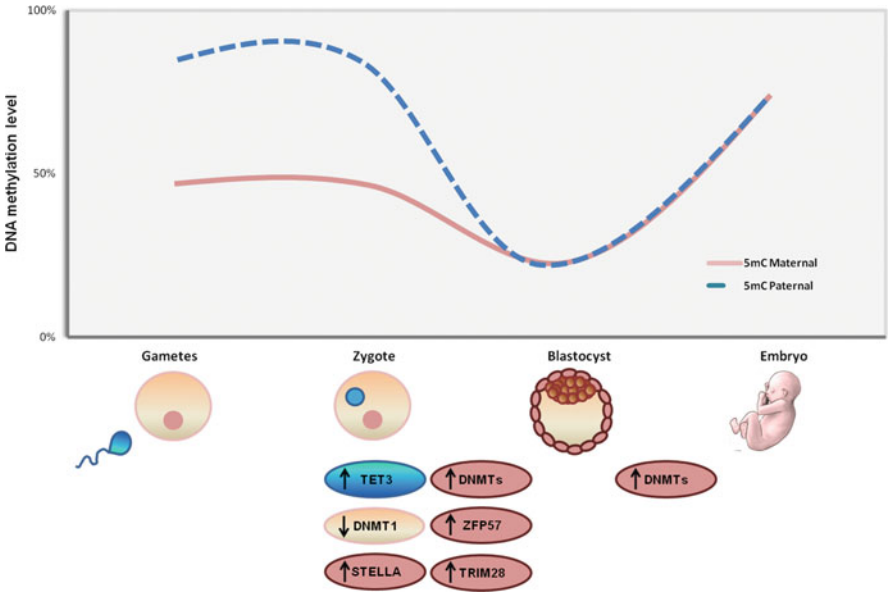

Fig. 1 DNA methylation levels at early developmental stages. Shown are the maternal (*solid line*) and paternal (*dashed line*) DNA methylation levels during early development. The corresponding developmental stages are shown below. TET3 facilitates the active demethylation of the paternal genome. DNMT1 down-regulation enhances the maternal passive demethylation. STELLA protects the maternal and several paternal genes from the active TET3 demethylation. DNMT1 and/or DNMT3A sustain ICRs methylation by targeting ZFP57 and TRIM28 which mark ICRs. DNMTs catalyze DNA methylation at further developmental stages

differentiation. It also enables in vitro examination of the process by altering and affecting the cells themselves, as well as their surroundings and conditions. Consequently, the field concentrates mainly on exploring PSCs, including their formation and differentiation, mimicking development. In this chapter we discuss the epigenetic processes that shape early developmental decisions, concentrating on the first embryonic days post fertilization and germ cell differentiation, as well as in embryonic stem cells (ESC), stem cell differentiation and somatic cell reprogramming to pluripotency.

2 Epigenetics in Early Development

Following fertilization, the highly methylated (80–90 %) paternal genome undergoes global demethylation (Fig. 1) (Kobayashi et al. 2012). This first wave of demethylation, which occurs before the first replication, is thought to rely on the active conversion of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) by TET3, followed by a passive replication-dependent demethylation until it reaches ~20 % methylation at the blastocyst stage (Kobayashi et al. 2012;

Messerschmidt et al. 2014; Gu et al. 2011; Hackett et al. 2013; Wu and Zhang 2014; Smith and Meissner 2013; Boland et al. 2014; Cantone and Fisher 2013). The maternal genome, on the other hand, begins with a moderate level of methylation (40–50 %) (Kobayashi et al. 2012) and undergoes passive erasure in a replication dependent manner, assisted by the down regulation of the maintenance DNA methyltransferase 1 (DNMT1) (Cantone and Fisher 2013) (Fig. 1).

Comparing the DNA methylation landscape of the paternal versus the maternal genome reveals over 60 % common methylation pattern regions. The shared hypermethylated regions are highly enriched for transposable elements and introns, whereas the common hypomethylated regions are mostly enriched with high-density CpG promoters, exons and CpG islands (CGIs) (Guo et al. 2014). In contrast to these common features between the human paternal and the maternal genomes, there are over 17,000 sperm-specific differentially methylated regions (DMRs) enriched in intergenic regions harboring tissue-specific enhancers with incomplete methylation of repetitive elements (Smith et al. 2014) and above 5000 (Smith et al. 2014; Guo et al. 2014) oocyte-specific DMRs enriched with intragenic regions with high localization to CGIs (Guo et al. 2014). Mouse gametes show similar results, but with a different amount of DMRs (Smith et al. 2012, 2014). Interestingly, while the methylation reduction trend is overall similar in both mouse and human, in mouse it occurs at the one cell stage whereas in human it is spread between fertilization and the two cell stage.

In order to sustain methylation of imprinted genes, DNMT1 and/or DNMT3A as well as histone modifiers, including the H3K9me3 methyltransferase Setdb1, are targeted to imprinted gene loci. These loci are marked by a hexanucleotide motif via the zinc finger transcription factor 57 (ZFP57) and tripartite motif-containing 28 (TRIM28, also known as KAP1) (Messerschmidt et al. 2014; Wu and Zhang 2014; Smith and Meissner 2013; Cantone and Fisher 2013; Guo et al. 2014; Lorthongpanich et al. 2013). The maternal genome, as well as a few paternally methylated imprinted genes, retrotransposons and centromeric heterochromatin, all appear to be protected from the active TET3 mediated demethylation by several means. One partially uncovered mechanism is based on the essential developmental protein called STELLA (also known as PGC7 or DPPA3), which alters the chromatin configuration by its interaction with H3K9me2. The H3K9me2 mark is specifically enriched on the maternal genome and on STELLA protected paternal genes (Messerschmidt et al. 2014; Wu and Zhang 2014; Smith and Meissner 2013; Cantone and Fisher 2013; Lorthongpanich et al. 2013).

Demethylation and methylation of pre-implantation embryos is a critical step in development and pluripotency. Mice deficient for DNMTs are embryonically lethal and do not develop past implantation. Dnmt3a and Dnmt3b are specifically responsible for *de novo* methylation and their inactivation causes global genome demethylation together with deficiency in methylating new retroviral DNA and in vitro aberrant differentiation (Okano et al. 1999; Jackson et al. 2004). Dnmt3a and Dnmt3b have partial functional redundancy since the deletion of both is embryonically lethal, whereas the deletion of Dnmt3a or Dnmt3b alone results in a variety of defects or complete embryonic lethality, respectively (Okano et al. 1999). Dnmt3l

was found to be an essential cofactor since its knockout produces infertile male mice and embryonic lethality in the case of maternal null-derived embryos. In this case global genome methylation levels are not affected and the lethality is a consequence of biallelic expression of imprinted genes (Bourc'his et al. 2001; Messerschmidt et al. 2014). Deletion of the maintenance Dnmt1 causes DNA hypomethylation, abnormal differentiation, aberrant development and embryonic lethality (Li et al. 1992; Jackson et al. 2004; Howell et al. 2001).

Correspondingly, ESCs deficient for all three Tet enzymes, which mediate the conversion from 5mC to 5hmC, exhibit low contribution to developing embryos and inadequate support for development (Dawlaty et al. 2014). Moreover, Tet1/2/3 triple knockout cells display aberrant differentiation as well as reprogramming ability (Dawlaty et al. 2014; Hu et al. 2014). Upon implantation, global DNA methylation level increases, reaching almost 80 % in human embryos (Guo et al. 2014). This includes the remethylation of the oocyte and sperm specific DMRs. Taken together, these studies demonstrate the crucial role that both methylation and demethylation play during early embryonic development (Fig. 1).

An interesting epigenetic program occurs during the differentiation of Primordial germ cells (PGCs). PGCs, derived from the embryonic ectoderm, are set aside early during development following implantation around day E6.5 in the mouse (Messerschmidt et al. 2014; Wu and Zhang 2014). At this stage, PGCs still carry the somatic, embryonic DNA methylation pattern (∼70 % CpG methylation) (Messerschmidt et al. 2014; Wu and Zhang 2014; Smith and Meissner 2013). Upon their specification, PGCs undergo global demethylation with the exception of several retroviral elements that are only partially demethylated (Messerschmidt et al. 2014; Smith and Meissner 2013; Cantone and Fisher 2013). It is believed that demethylation begins passively by 5mC dilution through cell proliferation at the time of migration from the proximal epiblast to the genital ridges through the hindgut. At the gonads themselves, erasure ends with an active demethylation process presumably by Tet1, which is essential for demethylating promoters of meiotic genes and imprinting control regions (ICRs) of imprinted genes (Yamaguchi et al. 2012, 2013; Messerschmidt et al. 2014; Wu and Zhang 2014; Cantone and Fisher 2013).

Several other global epigenetic features observed during the erasure process include H3K9 demethylation and H3K27 methylation, which might suggest a poised state allowing efficient de-repression (Smith and Meissner 2013; Cantone and Fisher 2013). After this erasure, remethylation occurs in a different manner between female and male gametes. Female gametes reach full methylation only at sexual maturation with one major, although incomplete, methylation wave at birth during the oocyte growth phase at prophase of meiosis I (Messerschmidt et al. 2014; Smith and Meissner 2013). In oocytes, Non-CpG methylation is also present around gene bodies and is correlated with the expression level of the regulated genes (Tomizawa et al. 2011; Shirane et al. 2013). Conversely, male gametes achieve complete remethylation before birth, followed by a large number of mitotic divisions before they enter meiosis (Messerschmidt et al. 2014; Smith and Meissner 2013). The effect of remethylation establishment also varies between the sexes.

Failure of this process is harmless for fertilization-competent oocytes while detrimental in spermatocytes, leading to severe spermatogenesis defects (Messerschmidt et al. 2014; Smith and Meissner 2013).

The erasure of DNA methylation followed by its reestablishment and maintenance is extremely important for PGC formation and proper development. The hallmark of this process is the imprinting of several hundred genes (Xie et al. 2012; DeVeale et al. 2012). Through the methylation and silencing of one of the two copies of the gene, the cells ensure a single parent of origin expression from the non-methylated copy. Thus, methylation regulation in early development in general and in primordial germ cells in particular, plays a pivotal role during embryogenesis. Any minute failure to properly execute the methylation/demethylation program results in dire consequences for development.

3 Embryonic Stem Cells and Differentiation

As noted above, the blastocyst stage embryo at E3.5 contains cells of the ICM and extra-embryonic trophoectoderm (Tee and Reinberg 2014). When ICM cells are successfully cultured in vitro as embryonic stem cells (ESCs), they maintain their potential to differentiate along the three germ layers, but unlike their in vivo counterparts, ESCs are also self-renewing when grown in vitro, proliferating essentially indefinitely. ESCs were first isolated from mouse in the early 1980s (Evans and Kaufman 1981), and almost 20 years later from human (Thomson et al. 1998). One of the reasons for this prolonged delay, apart from the technical difficulties and ethical concerns associated with working with human, rather than mouse embryos is that human ESCs (hESCs) required somewhat different conditions and signaling cascades to maintain their pluripotent, self-renewing state in vitro (Thomson et al. 1998; Reubinoff et al. 2000; Xu et al. 2001).

The first successful culture conditions applied to mESCs relied on factors secreted from irradiated mouse embryonic fibroblasts (MEFs) together with fetal bovine serum supplemented with the cytokine leukemia inhibitory factor (LIF) which predominantly affects signaling of the Stat3 pathway (Ying et al. 2008). More recently, the understanding that ESCs, while differentiating, use fibroblast growth factor-4 (FGF4) signaling to facilitate stimulation of mitogen-activated protein kinase (MAP/ERK kinase, MEK) pathway eventually yielded the 2i medium (Ying et al. 2008), named after the two kinase inhibitors PD0325901 and CHIR99021, which inhibit the MEK pathway and the glycogen synthase kinase (GSK3) pathway, respectively. 2i-grown ESCs, termed ‘naïve’, show a hypomethylated genome resembling that of the ICM (Smith et al. 2014; Wu and Zhang 2014).

In contrast to mouse, 2i conditions are not sufficient to maintain hESCs undifferentiated (Gafni et al. 2013; Chan et al. 2013; Theunissen et al. 2014). Traditionally, hESCs were grown in serum conditions supplemented with FGF. These colonies appeared flat and displayed considerably slower proliferation rates

than that of mouse ESCs. It was therefore argued that hESCs represent a more developed state resembling epiblast stem cells (epiSCs), derived from the post-implantation embryos, and that mESCs represent the pre-implantation embryo (Tesar et al. 2007; Najm et al. 2011; Brons et al. 2007). Very recently, human stem cell medium containing 2i/LIF synergistically with a combination of several additional cytokines and small molecule kinase inhibitors at specific concentrations, was shown to support a naïve-like state in hESCs. Naïve hESCs appeared mouse-like with round, highly proliferating colonies (Gafni et al. 2013; Chan et al. 2013; Theunissen et al. 2014).

But regardless of their origin (human or mouse), ESCs express pluripotency factors which act as transcription factors harboring the ability to silence lineage specific TFs or activate transcription of ESC specific genes in a direct or indirect manner. They regulate both protein coding genes as well as microRNAs (miRNAs), largely through occupying corresponding regulatory regions and enhancers. The core pluripotency factors include Nanog, Oct4 (a.k.a. Pou5f1) and Sox2, the latter two of which also function as a heterodimer in ESCs. Their regulation is facilitated by their own positive-feedback loops on each other (Young 2011), as well as by super-enhancers, which are defined as large genomic domains containing clusters of active enhancers (Whyte et al. 2013; Hnisz et al. 2013; Boland et al. 2014).

The pluripotency factors influence the global epigenetic state. All three may affect RNA polymerase II (RNAPII) recruitment by interacting with relevant co-activators. In order to maintain the ESC state, the repression of lineage specific regulators is required. This is accomplished by orchestrating pluripotency factors together with different epigenetic machineries, such as Polycomb Group (PcG) protein complexes, the histone methyltransferase SetDB1 and the histone acetyltransferase (HAT) Tip60-p400. One example includes the histone H3K9me3 repressive modification, catalyzed by sumoylated SetDB1, which is bound by Oct4. Subsequently, Polycomb Repressive Complexes (PRC) 1 and 2 associate with these modified nucleosomes and enhance the repression at these regions (Young 2011). PcG proteins can further add ubiquitin to H2A lysine 119 (H2AK119ub) as well as catalyze dimethylation and trimethylation of H3K27 (Boland et al. 2014; Young 2011; Riising et al. 2014). They possess the ability to silence developmental genes through H3K27me3 in an active poised state, thus providing them with quick activation ability during differentiation (Boland et al. 2014; Cantone and Fisher 2013; Tee and Reinberg 2014; Mikkelsen et al. 2007). These domains, which are characterized by regulatory regions carrying both active H3K4me3 marks alongside repressive H3K27me3 marks, are referred to as ‘bivalent’ (Bernstein et al. 2006; Pan et al. 2007).

Bivalent domains are evolutionarily conserved and are found abundantly, although not exclusively, in ESCs. Bivalent promoters are thought to poise developmental genes for lineage specific expression (Boland et al. 2014; Cantone and Fisher 2013; Tee and Reinberg 2014; Mikkelsen et al. 2007; Li et al. 2012). The histone methyltransferase Mixed lineage leukemia 2 (MLL2, a.k.a. KMT2D; ALR; MLL4) is responsible for regulating H3K4me3 deposition on bivalent promoters, although its loss did not cause changes in the expression pattern upon

differentiation (Tee and Reinberg 2014; Li et al. 2012). Supporting their role in repressing developmental genes, when PcG proteins are depleted, ESCs become less stable, more differentiation prone with significant differentiation defects (Boland et al. 2014; Cantone and Fisher 2013; Young 2011; Riising et al. 2014). While PcG act to maintain silencing mostly by H3K27 methylation, trithorax group (TrxG) proteins act to promote activation by catalyzing H3K4 methylation and antagonizing PcG proteins.

Akin to H3K4 methylation, histone acetylation also stimulates gene transcription, and when HATs such as p300, Gcn5 or Tip60-p400 are depleted in ESCs, these display severe aberrant differentiation phenotypes (Young 2011; Li et al. 2012). The HAT Mof (a.k.a. MYST1; KAT8) also has a role in ESC self-renewal as its depletion causes proliferation and self-renewal aberrations alongside defects in ESC gene expression including pluripotency genes and differentiation markers (Li et al. 2012). Taken together, these data illustrate that ESCs' epigenetic state supports a bivalent poised state and that tampering with the subtle balance between activators and suppressors pushes the cells away from their pluripotent state.

One hallmark of ESCs is their open chromatin conformation and hyperdynamic association of chromatin proteins with chromatin (Boland et al. 2014; Tee and Reinberg 2014; Meshorer et al. 2006; Gaspar-Maia et al. 2011; Efroni et al. 2008). ESCs also seem to harbor less heterochromatin foci (Efroni et al. 2008) and less chromatin associated with the nuclear lamina and nucleoli than differentiated cells (Boland et al. 2014), likely supporting a more permissive transcriptional state (Efroni et al. 2008). Supporting this view, ESCs are largely depleted of chromatin modifications associated with repressed chromatin such as H3K9me3 and H3K27me3, and are relatively enriched with chromatin modifications associated with active chromatin including histone acetylations (Boland et al. 2014) (Fig. 2).

Citrullination, the arginine conversion to citrulline by peptidylarginine deiminases (PADIs), is another modification involved in transcriptional activation (Christophorou et al. 2014; Zhang et al. 2012). Interestingly, citrullination was recently found to be important for the establishment and maintenance of pluripotency (Christophorou et al. 2014). One of the PADI family members, Padi4 (a.k.a Pad4 or Padv), is a member of the pluripotency network and its inhibition impairs reprogramming efficiency and reduces the number of pluripotent cells in the early mouse embryo. A proteomic screen identified linker histones as Padi4 targets. Padi4 citrullinates the arginine residue within the DNA-binding site of H1, resulting in its displacement from chromatin and to chromatin decondensation (Christophorou et al. 2014) (Fig. 2).

As discussed above, DNA methylation and demethylation plays crucial roles during preimplantation development but curiously, while ICM cells show a low, but measurable, methylation level of ~20 % (Wu and Zhang 2014) (Figs. 1 and 2), triple-knockout mouse ESCs deficient for all DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b can self-renew normally, although they exhibit strong differentiation defects (Tsumura et al. 2006). In addition, the methylation level of cultured ESCs is almost as high as somatic cells (Melcer et al. 2012; Pannetier

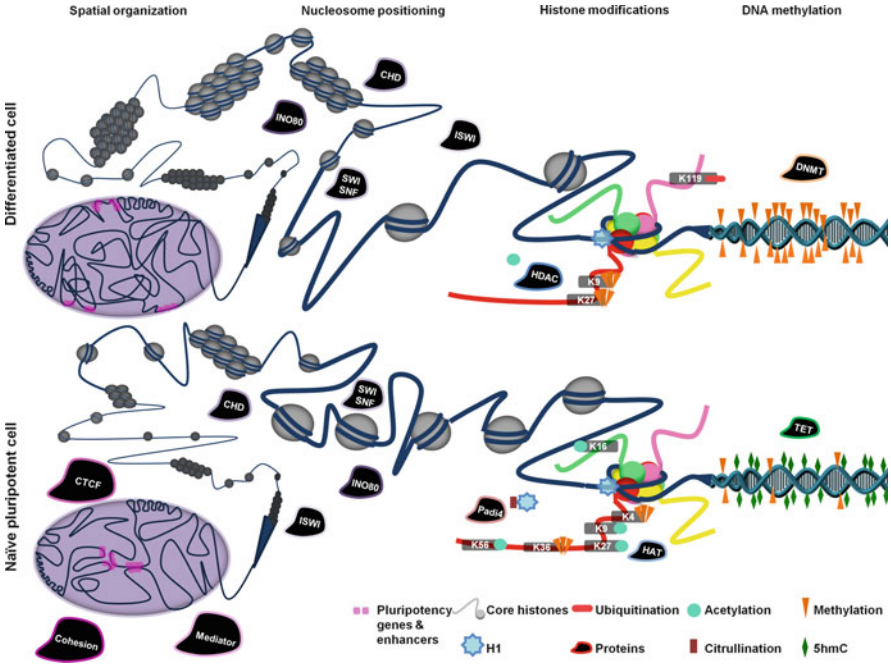

Fig. 2 Layers of epigenetic regulation, along with key enzymes and factors, distinguishing naïve pluripotent cells and differentiated cells. DNA methylation is lower, whereas 5hmC is higher, in naïve pluripotent cells. Naïve pluripotent cells also exhibit higher levels of open chromatin modifications than differentiated cells. The nucleosome repeat length (NRL) is shorter in naïve pluripotent cells although the overall chromatin compaction is higher in differentiated cells. The 3D chromatin conformation in naïve pluripotent cells enables distant contacts between different pluripotency genes and enhancers. These contacts are abolished in differentiated cells

and Feil 2007; Laurent et al. 2010), suggesting that they acquire spontaneous DNA methylation during culture. Finally, although ESCs are enriched with 5hmC (Fig. 2), which is negatively correlated with methylated regulatory elements (Williams et al. 2011; Wu et al. 2011), they are capable of maintaining their self-renewal even in the absence of the three Tet enzymes (Tet1,2,3) (Dawlaty et al. 2014), which as we have seen above, mediate the conversion from 5mC to 5hmC, thus effectively demethylating DNA. These results suggest that unlike their embryonic counterparts, ESCs can tolerate considerable changes in DNA methylation without affecting their self-renewal abilities. It is likely that once ESCs are injected into blastocysts for the generation of chimeric animals, their methylation level is reprogrammed in vivo to a methylation state that resembles that of the ICM. This, however, remains to be shown.

Following differentiation, ESCs acquire methylation changes as they shift from high to low DNA methylation turnover rates (Shipony et al. 2014) and gain or loss of methylation at specific sites (Meissner et al. 2008; Laurent et al. 2010; Lister et al. 2009) (Fig. 2). Curiously, non-CpG methylation, 5mC followed by A, T or C,

is enriched at neurons and pluripotent cells with a marked depletion at specific pluripotent enhancers (Wu and Zhang 2014; Boland et al. 2014; Laurent et al. 2010; Lister et al. 2009, 2013). It seems to be catalyzed by Dnmt3a and Dnmt3b, although its function is still an enigma (Wu and Zhang 2014; Boland et al. 2014; Laurent et al. 2010; Lister et al. 2009).

DNA methylation is also involved in the silencing of one of two X chromosomes in female cells during development. One hallmark of mouse female pluripotent cells are two active X chromosomes. This is a good example for an epigenetic change involving several machineries integrating and working together. The inactivation process involves the cis-acting non coding RNA, Xist, which is expressed only from the inactivated X and coats the X chromosome. In Parallel, Xist interacts with the protein YY1 and PRC2 complexes thus facilitating the accumulation of histone repressive marks and DNA methylation. The antagonizing Tsix RNA, the Xist antisense partner, can repress Xist on one allele thus aiding the allelic choice of silencing (Lee and Bartolomei 2013). While mouse female ESCs display two active X chromosomes, in most female hESCs one X chromosome is inactive (Boland et al. 2014; Tee and Reinberg 2014; Mekhoubad et al. 2012). Reactivation of the silenced X chromosome might occur in cultured hESCs, but this is often accompanied by the loss of its silencing ability upon differentiation (Mekhoubad et al. 2012). This reactivation procedure is escorted by the activation of the inactive X genes, Xist repression and H3K27me3 foci loss (Mekhoubad et al. 2012), a process involving the recruitment of the pluripotency factors as Oct4, Sox2, Nanog and Rex1 to the X inactivation machinery (Lee and Bartolomei 2013).

Nucleosome spacing is another important epigenetic layer influencing chromatin regulation and gene expression. To a large extent, DNA positioning is correlated with DNA sequence motifs (Segal et al. 2006; Brogaard et al. 2012). There is a marked depletion of nucleosomes alongside DNase I hypersensitivity sites at promoters, upstream to transcription start sites (TSS) and at the peak of RNAPII (Carone et al. 2014; Lee et al. 2004; Schwabish and Struhl 2004). In general, a reverse correlation is seen between the presence of a nucleosome at a given promoter and the expression level of the regulated gene, although many exceptions exist (Carone et al. 2014). In ESCs, the spacing between nucleosomes or the nucleosome repeat length (NRL) is shorter than in either neuronal progenitor cells (NPC) derived from ESCs or MEFs (Teif et al. 2012), suggesting, once again, that ESCs have a global chromatin conformation distinct from that of differentiated cells (Fig. 2). The fact that the open chromatin conformation of ESCs is associated with shorter, rather than longer, NRL, likely reflects the fact that ESCs are relatively depleted for linker histones, which act both as suppressors and as nucleosome separators (Fan et al. 2005; Woodcock et al. 2006). Accordingly, when three H1 variants are depleted in ESCs, the NRL shortens (Fan et al. 2005; Woodcock et al. 2006).

One of the protein complexes regulating nucleosome placement and displacement are the ATP-dependent chromatin remodeling proteins. According to the sequence and structure of their ATPase subunit, these proteins are classified to four different families: SWI/SNF, ISWI, CHD and INO80 (Ho and Crabtree 2010;

Gaspar-Maia et al. 2011) (Fig. 2). Several chromatin remodeling complexes were shown to be crucial for ESC pluripotency and maintenance. For instance, Chd1, a member of the CHD family, has a role in keeping a hyperdynamic and euchromatic state in ESCs, and when Chd1 is knocked down in ESCs, these exhibit impaired pluripotency (Gaspar-Maia et al. 2009). The NURD complex, also a member of the CHD family, includes at its core the Chd3, Chd4 and Mbd3 subunits. Mbd3 elimination causes developmental arrest after implantation, at the transition from ICM cells into mature epiblast (Kaji et al. 2007). In culture, Mbd3 deficient ESCs exhibit differentiation impairments, failing to commit to specific developmental lineages (Kaji et al. 2006). Lately, Mbd3 was also shown to play a key role in reprogramming (Luo et al. 2013; Rais et al. 2013) although the nature of its affect is debated and is yet to be clarified (dos Santos et al. 2014).

In the INO80 family, p400 is present in a complex together with the histone acetyltransferase TIP60. Both of these proteins alongside other components of the TIP60–p400 complex were found to be important for ESCs self-renewal and pluripotency (Fazzio et al. 2008). NURF, a member of the ISWI family, includes a large subunit named Bptf. Bptf mutant embryos show a decline in post-implantation proliferation, arresting at embryonic day (E) 8.5. ESCs lacking Bptf display differentiation defects into all three lineages, failing to activate Smad regulated genes. NURF might regulate development through its interaction with Smad transcription factors, followed by its recruitment to TGFβ responsive genes (Landry et al. 2008).

Finally, ESCs contain a unique composition of SWI/SNF remodelers termed esBAF, which regulates the core pluripotency transcriptional network (Ho et al. 2009, 2011; Yan et al. 2008; Fazzio et al. 2008). One mechanism by which esBAF contributes to pluripotency is through promoting chromatin accessibility at Stat3 binding targets, a downstream effector of the cytokine LIF signaling pathway (Ho et al. 2011). Overall, representative ATP-dependent chromatin remodeling proteins from all families demonstrate the essential function of mobilizing nucleosomes and altering chromatin structure in maintaining the pluripotent state, as well as in differentiation and reprogramming.

If DNA methylation signifies the primary, fundamental level of epigenetic regulation, the 3-dimentional configuration of the genome represents the higher order level of epigenetic control (Fig. 2). Spatial organization of DNA reflects a transcriptional related arrangement with long range contacts (Dixon et al. 2012; Hou et al. 2012; Sexton et al. 2012). In ESCs, the core pluripotency factors (Oct4, Sox2 and Nanog) seem to have high density of binding sites at interacting regions. This interaction network is specific to pluripotent cells and was found to be affected by chromatin organizers such as CTCF, cohesion subunits and protein of the mediator family. This organization is likely important to gather the different pluripotency affected regions, thereby enhancing transcriptional regulation and cell state (Boland et al. 2014; de Wit et al. 2013; Apostolou et al. 2013; Phillips-Cremins et al. 2013). Spatial rearrangements accompanies differentiation as several pluripotent genes move towards the nuclear periphery and their enhancer interactions are weakened (Apostolou and Hochedlinger 2013) (Fig. 2).

As noted above, apart from their potential for regenerative medicine, ESCs allow studying early development in vitro. By following in vitro differentiation to the three germ layers, it was shown that a massive transcriptional change occurs at the first 5 days of the process and that endodermal and ectodermal lineages are more related to each other than the more distant mesoderm (Gifford et al. 2013). A somewhat surprising finding was that while almost 40 % of the total expressed long non-coding RNAs (lncRNAs) are lineage restricted, <20 % of the protein coding genes are restricted to a specific lineage (Xie et al. 2013). Furthermore, lineage specific gene promoters are mostly GC poor with a relative paucity of CpG islands, compared with genes expressed at early embryonic lineages or housekeeping genes (Xie et al. 2013). In addition to expression changes, alternative splicing and alternative promoter usage also significantly contribute to changes in transcript diversity during early ESC differentiation. For example DNMT3B isoform 1 is expressed in hESCs whereas the alternative DNMT3B isoform 3 is expressed in differentiated cells (Gifford et al. 2013).

Recently, an intriguing connection between epigenetics and metabolomics has emerged. ESCs are naturally localized at a hypoxic environment, and they use anaerobic glycolysis in order to produce energy with minimizing oxidative stress (Weissbein et al. 2014; Yanes et al. 2010; Ochocki and Simon 2013). During differentiation, ESCs transform their metabolic properties to an oxidative phenotype, accompanied by the increased production of ATP. Furthermore, the cells change their lipid profile as they exhibit lower amounts of unsaturated lipids and fatty acids upon differentiation (Weissbein et al. 2014; Yanes et al. 2010; Ochocki and Simon 2013). Manipulating specific lipid pathways in ESCs, such as the eicosanoid pathway, has proven to affect their pluripotent state as well as their differentiation capacities (Shyh-Chang et al. 2013; Weissbein et al. 2014; Yanes et al. 2010; Ochocki and Simon 2013). An example of a mechanism behind such an effect involves threonine (Thr) and threonine dehydrogenase (Tdh) which can catalyze Thr conversion to glycine and acetyl-coenzyme A (CoA). These products subsequently control the generation of S-adenosylmethionine (SAM) and the SAM-to-S-adenosylhomocysteine (SAH) ratio, which influences the methylation reactions in the cell. In ESCs, the Thr-SAM pathway is activated, thus maintaining a high SAM/SAH ratio which apparently affects H3K4me3 and the overall epigenetic and pluripotent cell state (Shyh-Chang et al. 2013). These findings suggest an exciting link between chromatin regulation and the energetic homeostasis in ESCs and will no doubt lead to additional research into this newly discovered avenue.

Finally, the cell cycle also changes with differentiation. ESCs possess shortened G1 and G2 phases of the cell cycle and they are precluded of G1 checkpoint regulation (Pauklin and Vallier 2013). Upon differentiation, the G1 cell cycle phase is particularly extended. Interestingly, the cell cycle phase during the onset of differentiation affects the lineage outcome. hESCs at early G1 are more prone to differentiate into endoderm; hESCs at late G1 tend to differentiate into neuroectoderm, and hESCs at G2/S/M do not respond properly to differentiation signals. This cell cycle affect is thought to be a consequence of changes of the Activin/Nodal-Smad2/3 pathway and cyclin-D expression (Pauklin and Vallier

2013). If cells in vivo act similarly to their in vitro counterparts, it implies that an unsynchronized population of blastocyst ICM cells is necessary to generate different lineages during development. It would be interesting to see whether cell cycle-related differentiation propensities could be exploited for improved differentiation methods.

Together, the above studies demonstrate the multi-level regulation of ESCs, from the DNA methylation pattern, the positioning of nucleosomes and the action of chromatin remodelers, to histone modifications and the three-dimensional conformation of chromatin itself. All act in a concerted manner to regulate pluripotency and differentiation.

4 Reprogramming and iPS Cells

The 2012 Nobel Prize in Physiology or Medicine was awarded jointly to Sir John B. Gurdon and Shinya Yamanaka "for the discovery that mature cells can be reprogrammed to become pluripotent". John Gurdon was the first to show, in 1958, that a somatic nucleus combined with the fitting environment, is sufficient to create a whole embryo. He demonstrated this idea on a frog when he replaced the nucleus of a fertilized egg with a mature cell nucleus and succeeded in generating a new complete tadpole. These somatic cell nuclear transfer (SCNT) cells were characterized as pluripotent (Gurdon et al. 1958). It later became apparent that the fusion of two different cells triggered the activation of genes which are not meant to be expressed at the specific specialized cell, pointing towards the cytoplasm as the inducer of activation, and hinting on the ability to change cell identity via ectopic means (Blau et al. 1983; Tada et al. 1997). It was only in 2006 when Yamanaka discovered the specific genes required for the complete reprogramming of mice fibroblasts into pluripotent cells in vitro (Fig. 3), creating the so-called induced pluripotent stem cells (iPSC) (Takahashi and Yamanaka 2006). He was the first to accomplish direct reprogramming in vitro and pinpointed the necessary and sufficient factors to induce this conversion, first in mouse and later in human (Yu et al. 2007; Carey et al. 2009; Takahashi and Yamanaka 2006; Takahashi et al. 2007).

Reprogramming is achieved by the ectopic expression of Oct4, Sox2, Klf4 and c-Myc, collectively referred to as OSKM. Their expression is required transiently during the process until they are endogenously activated. From this point forward the cells are independently pluripotent. The OSKM transcription factors (TFs) induce the activation or repression of their downstream target genes. The oncogene c-Myc is responsible for reprogramming efficiency. Its early expression in the process prompts proliferation and a metabolic switch, coinciding with an overall transcriptional amplification (Apostolou and Hochedlinger 2013; Polo et al. 2012; Lin et al. 2012; Schmidt and Plath 2012; Buganim et al. 2013; Sridharan et al. 2009), although reprogramming is possible without c-Myc as well (Nakagawa et al. 2008; Wernig et al. 2008). While c-Myc is dispensable, Oct4, Sox2 and Klf4

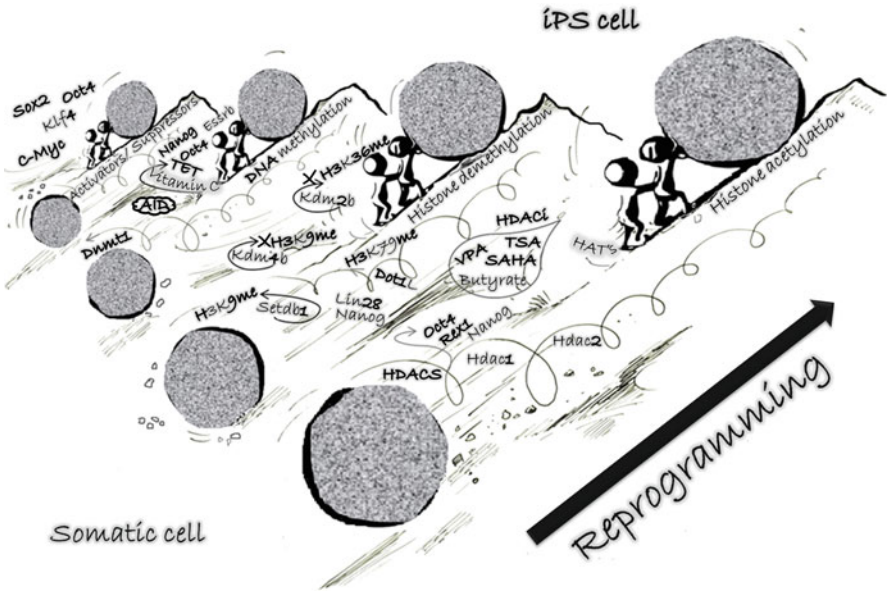

Fig. 3 Factors that alter reprogramming efficiency by chromatin decondensation. To achieve reprogramming, cells undergo transcriptional activation and repression by the OSKM TFs; DNA demethylation; histone demethylation; histone acetylation. For each of these processes, ways to push the cells towards or away from reprogramming are depicted

are essential, driving the cell toward an ESC state as they act to repress lineage specific genes, and activate a network of pluripotency genes including Nanog (Apostolou and Hochedlinger 2013; Polo et al. 2012; Schmidt and Plath 2012; Buganim et al. 2013; Sridharan et al. 2009). OSK factors share the majority of their binding sites, although Klf4 and the Sox2-Oct4 heterodimer also bind many regions selectively (Apostolou and Hochedlinger 2013; Polo et al. 2012; Schmidt and Plath 2012; Buganim et al. 2013; Sridharan et al. 2009).

Following this surprising realization that merely 3 TFs are sufficient for complete reprogramming, researchers worked out additional TF cocktails which are able to convert one cell type into another (Takeuchi and Bruneau 2009; Zhou et al. 2008; Zhou and Melton 2008), as well as other TFs, miRNAs or small molecules which are capable of replacing the OSKM cocktail and often improving reprogramming efficiency (Anokye-Danso et al. 2011; Schmidt and Plath 2012; Buganim et al. 2012; Ichida et al. 2009). This can be achieved by increasing the permissive chromatin state of the somatic reprogrammed cell (Biran and Meshorer 2012; Cantone and Fisher 2013; Rais et al. 2013) (Fig. 3). For example, TET proteins are thought to prime selected regions for DNA demethylation by converting 5mC to 5hmC and thus promote reprogramming (Wu and Zhang 2014; Apostolou and Hochedlinger 2013). This notion is based on the observation that TET1/2 interact with key pluripotency factors such as Nanog and Esrrb. Nanog recruits TET2 to chromatin and TET1 can facilitate Oct4 expression as well as

replace Oct4 in the OSKM reprogramming cocktail. In addition, vitamin C (ascorbic acid), which increases reprogramming efficiency, is thought to act as a direct regulator of TET and DNA methylation (Stadtfeld et al. 2012; Blaschke et al. 2013; Chen et al. 2013). Another crucial reprogramming regulator is the activation-induced cytidine deaminase (AID/AICDA) enzyme which mediates DNA demethylation at the first steps of reprogramming and is necessary for the induction of pluripotent genes (Bhutani et al. 2010). Reprogramming can also be promoted or inhibited by modifying the levels of DNA methyltransferases (Fig. 3). For example, reprogramming is enhanced when Dnmt1 is inhibited. A transient Dnmt1 inhibition is enough to aid the transition of partially reprogrammed to fully reprogrammed pluripotent cells (Mikkelsen et al. 2008).

Histone methylation also plays important roles (Fig. 3). The H3K36me2 demethylase Kdm2b, for example, acts early during the reprogramming process and increases reprogramming efficiency by demethylating histones on target gene promoters, thereby activating their expression (Liang et al. 2012). In contrast, depletion of the H3K9 methyltransferase Setdb1, enhances reprogramming as well as the transition from partially reprogrammed cells to fully reprogrammed iPSCs. Respectively, ectopic expression of the H3K9 demethylase Kdm4b has an opposite effect, promoting partial iPSCs into fully reprogrammed cells. Hence it may be argued that H3K9 methylation acts as a barrier for the reprogramming process (Chen et al. 2013). Along the same lines, inhibition of the H3K79 methyltransferase Dot1L, accelerates reprogramming as it facilitates H3K79me2 loss from fibroblast-specific genes and enhances Nanog and Lin28 expression in an indirect manner. It can also substitute for Klf4 and c-Myc during the reprogramming process (Onder et al. 2012).

If suppressive chromatin inhibits reprogramming, histone acetylation promotes reprogramming and supports the pluripotent state (Hezroni et al. 2011b; Huangfu et al. 2008a; Ware et al. 2009) (Fig. 3). Acetyl modifications are negatively charged thus reducing the histones' positive charge which interferes with their electrostatic interactions with the DNA (Haberland et al. 2009; Bannister and Kouzarides 2011). In addition, these modifications can alter the binding of chromatin remodeling complexes or transcriptional factors/activators (Bannister and Kouzarides 2011; Barth and Imhof 2010; Smith and Shilatifard 2010). Acetylation is a fairly transient modification with a relatively high turnover rate (Barth and Imhof 2010; Smith and Shilatifard 2010). The acetyl 'writers' are the family of histone acetyltransferases (HATs) and the 'erasers' are the histone deacetylases (HDACs) (Apostolou and Hochedlinger 2013; Bannister and Kouzarides 2011). HDACs are divided into five classes (class I, IIa, IIb, III and IV) and each enzyme has the ability to deacetylate several sites (Haberland et al. 2009; Bannister and Kouzarides 2011; Livyatan and Meshorer 2013). Upon deacetylation, the positive charge of the histones rises once again, increasing their interaction with the DNA, thus eventually repressing transcription (Bannister and Kouzarides 2011). Histone deacetylase inhibitors (HDACi) increase histone acetylation as well as chromatin plasticity and dynamics (Hezroni et al. 2011a; Melcer et al. 2012). The small-molecule HDACi, suberoylanilide hydroxamic acid (SAHA), trichostatin A (TSA), butyrate and

valproic acid (VPA), can enhance and increase the yield of reprogrammed cells (Huangfu et al. 2008a, b; Mali et al. 2010; Liang et al. 2010; Ware et al. 2009). VPA was also shown to facilitate reprogramming when combined with merely two additional TFs, Sox2 and Nanog (Huangfu et al. 2008b). HDACi can either block the activity of the majority of HDACs (e.g., TSA, SAHA) or selectively inhibit a class of HDACs (e.g., VPA). VPA, which is so far the most prominent HDACi at enhancing reprogramming, inhibits class I HDACs and specifically acts on Hdac2 (Kramer et al. 2003). Accordingly, suppression or globally low levels of Hdac2 can substitute for the use of VPA in reprogramming using miR302/367, which target Hdac2 (Anokye-Danso et al. 2011). Taken together, these studies show that chromatin conformation and histone modifications influences the reprogramming process, and that by understanding chromatin regulation, we may control reprogramming and increase reprogramming efficiency.

Several chromatin remodeling complexes also include HDACs. A prominent example is the Mbd3/NuRD complex, which, as we have seen above, regulates pluripotency. Apart from its core subunits, the NuRD complex also includes HDAC1 and HDAC2 (Xue et al. 1998), facilitating the repression activity of the complex. Knockout of Hdac1, but not Hdac2, leads to mouse embryonic lethality (Guan et al. 2009; Montgomery et al. 2007; Zimmermann et al. 2007; Trivedi et al. 2007) and aberrant ESC differentiation with elevated mesodermal and neuroectodermal progenies and decreased endodermal lineages (Dovey et al. 2010). Double Hdac1 and Hdac2 knockout ESCs lose their viability after 4 days (Jamaladdin et al. 2014). During this period, core pluripotency factors including Oct4, Nanog, and Rex1, all of which are bound by Hdac1 (Kidder and Palmer 2012), are down-regulated (Jamaladdin et al. 2014). This suggest that Hdac1 may also act as an activator in ESCs (Wang et al. 2009; Kurdistani et al. 2002), in addition to its function in several repressor complexes such as Sin3A (Cowley et al. 2005; Laherty et al. 1997), CoREST (Humphrey et al. 2001; You et al. 2001) and NuRD (Xue et al. 1998). Since different HDACs control different sets of genes and processes (Haberland et al. 2009), it seems that pinpointing HDACi to specific HDACs should further enhance reprogramming. The overall success in improving reprogramming through chromatin remodeling suggests that it functions as a rate limiting step. Furthermore, all the above exemplifies ways to induce reprogramming by small molecules in addition to transcription factors (Fig. 3). Achieving reprogramming by chemical means alone (Hou et al. 2013) will ease the procedure and raise its practical applications.

Recently, a more comprehensive RNAi screen in human cells systematically identified barriers for reprogramming. In addition to the chromatin modifiers and transcription factors, the screen identified proteins of various cell components including cell adhesion, vesicular transport, ubiquitination and dephosphorylation (Qin et al. 2014). These works combined, suggest that drivers of reprogramming are largely acting to induce open chromatin and a greater accessibility at pluripotency gene regions, while simultaneously silencing lineage restricted genes. This occurs mostly at the early phase of the reprogramming process thus largely preceding the expression changes of the target genes. Accordingly, lineage restricted genes gain

H3K27me3 (Koche et al. 2011) together with the loss of H3K4me2 (Koche et al. 2011) and H3K79me2 (Onder et al. 2012), while pluripotency-related and early developmental genes display elevated levels of H3K4me2 (Koche et al. 2011) and H3K36me2 (Liang et al. 2012).

Reprogramming at the cell level begins with the initial binding of the reprogramming factors to enhancers and other distal elements and at later steps to promoter regions (Papp and Plath 2013). This causes structural chromatin changes together with DNA methylation loss, followed by the expression of pluripotency genes (Cantone and Fisher 2013). Accordingly, the chromatin three-dimensional conformation resets from somatic to pluripotent following reprogramming (de Wit et al. 2013; Apostolou et al. 2013; Phillips-Cremins et al. 2013), possibly by the aid of pluripotency factors as they mediate long range interactions mainly through Oct4 and Klf4 (Wei et al. 2013; de Wit et al. 2013; Apostolou et al. 2013; Phillips-Cremins et al. 2013; Apostolou and Hochedlinger 2013). Not all the cells exposed to OSKM manage to transform uniformly into iPSCs. One way to explain this is by viewing the process as a stochastic chain of events beginning with TF expression and ending with a small minority of iPSCs (Buganim et al. 2012; Hanna et al. 2009). However, some consider reprogramming to follow an organized kinetic path. At the transcription level, cells appear to respond relatively homogeneously to the OSKM expression, starting with the silencing of the somatic state followed by pluripotency activation in a biphasic transcriptional wave according to the OSKM pattern described above (Cantone and Fisher 2013; Apostolou and Hochedlinger 2013; Polo et al. 2012; Sridharan et al. 2009; Mikkelsen et al. 2008).

One model divides and defines three different phases: initiation, maturation, and stabilization (Samavarchi-Tehrani et al. 2010). When reprogramming MEFs, the first wave/initiation step contains the critical phase called mesenchymal-to-epithelial transition (MET), essentially the opposite of the epithelial-to-mesenchymal transition (EMT) observed during development. MET includes the silencing of mesenchymal genes and the activation of epithelial markers, accompanied by morphological changes (Samavarchi-Tehrani et al. 2010; Polo et al. 2012; Mikkelsen et al. 2008; Li et al. 2010). The second wave consists of the pluripotency genes as they are gradually activated (Apostolou and Hochedlinger 2013; Samavarchi-Tehrani et al. 2010; Polo et al. 2012; Buganim et al. 2012), perhaps at a hierarchical manner (Samavarchi-Tehrani et al. 2010; Buganim et al. 2012). This biphasic pattern is true for mRNA and for miRNA expression level as well as for the bivalent histone marks H3K4me3 and H3K27me3, although the establishment of the bivalent promoters is gradual (Polo et al. 2012).

Overall, iPSCs and ESCs share the same epigenetic and expression profile alongside their differentiation capabilities (Boland et al. 2014; Sridharan et al. 2009), although upon closer inspection some differences between them may arise. For example, copy number variations (CNVs) appear abundantly in early passage iPSCs, as compared with their starting fibroblast population or ESCs. During passaging, the CNVs gradually decrease. This phenomenon indicates that the reprogramming process might spark the introduction of novel CNVs, especially at common fragile sites. These CNVs seem to cause a selective disadvantage,

therefore disappearing following iPSCs expansion (Boland et al. 2014; Weissbein et al. 2014; Hussein et al. 2011). When examining hiPSCs genomic mutations, researchers could not identify specific recurrent mutations, leading to the assumption that there is no selective advantageous reprogramming genetic alteration (Weissbein et al. 2014; Ruiz et al. 2013).

Scanning the DNA methylation pattern at single-base resolution generated whole-genome methylome profiles in which the iPSCs were found to largely match ESCs, although some consistent differences and several incomplete methylation motifs were detected (Lister et al. 2011). Several DMRs are highly stable through differentiation thus transmitting aberrant methylation onwards. Additional regions found at iPSCs which are not completely reprogrammed, are mega base scale non-CpG DMRs, which also exhibit altered H3K9me3 together with aberrant transcriptional activity, and are found in proximity to centromeres and telomeres (Boland et al. 2014; Lister et al. 2011). One significant limitation when exploring iPSCs versus ESC, is that the evaluation is assessed with respect to different cells. It's not a comparison of two different conditions of the same system, such as ESCs during differentiation or MEFs during reprogramming, but rather it is a comparison of two different systems. Each cell line holds stable epigenetic differences at specific genes. Depending on the role of these genes, processes such as differentiation or reprogramming might be affected. Various pluripotent cell lines exhibit different dispositions at differentiation, in addition to lineage specific differences in differentiation propensities for each cell line (Bock et al. 2011).

iPSCs opened a new world to science and medicine. It manifested a means for exploring chromatin changes and their interactions with different cell machineries and regulators. The simplicity of inducing pluripotent cells by expressing several core pluripotency factors implanted the field with the will to find and extend this concept. Researchers are exploring new ways to alter the cell identity including the discovery of new reprogramming factors, changing the environment by different small molecules or in vivo conversions (Abad et al. 2013) as well as finding new techniques that will allow the most efficient and non-harmful methods to introduce these changes. Moreover, as noted above, researchers now seek different cocktail factors which will enable direct conversion of one somatic cell type to another through trans-differentiation. Simultaneously it also provides us with the ability to derive patient-specific stem cells which can be beneficial for disease modeling as well as for regenerative medicine by generating autologous transplantable cells. iPSCs undoubtedly will continue to be a major topic in science and medicine, and have every potential to yield novel insights, tools and medical techniques.

5 Future Perspectives

Exploring epigenetics in development is a difficult task particularly when it comes to complex organisms such as mammals. The transitions and changes which occur along the creation of an organism are not yet fully captured and understood; nor are

the factors that control them, the mechanisms and even the motivations behind all processes. The derivation and maintenance of pluripotent cells opened up a window for exploring development and differentiated cells. The research of development as it proceeds in vivo established the chain of events that occur, beginning in the pluripotent cell and ending with a specific terminally differentiated cell. ESCs differentiation derivatives are characterized by loss of self-renewal and the eventual gain of senescence. By implementing this knowledge, researchers can now pass the pluripotent cells through pre-arranged scenarios thus mimicking the desired developmental path and creating their cell of choice. This process is achieved by different growth factors and small molecules applied at precise timing, thus affecting cell signaling and environment, ultimately changing their transcriptome through chromatin and epigenetic regulation.

The differentiation process can be accomplished by two broad approaches. One is by simulating embryonic organization through the creation of embryoid bodies from stem cell aggregates and the other is by directing the cells as they are kept in a monolayer. There are protocols for generating a variety of cell types from pluripotent cells. By using this amenable in vitro system it is also possible to track different parameters along the way to better understand developmental decisions as well as the cellular and the molecular mechanisms underlying each process. Deciphering and fully understanding epigenetics will pave the way for further manipulation of stem cells and accordingly, new therapeutic avenues.

References

Abad M, Mosteiro L, Pantoja C, Canamero M, Rayon T, Ors I, Grana O, Megias D, Dominguez O, Martinez D, Manzanares M, Ortega S, Serrano M (2013) Reprogramming in vivo produces teratomas and iPS cells with totipotency features. Nature 502(7471):340–345. doi:10.1038/nature12586

Anokye-Danso F, Trivedi CM, Juhr D, Gupta M, Cui Z, Tian Y, Zhang Y, Yang W, Gruber PJ, Epstein JA, Morrisey EE (2011) Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. Cell Stem Cell 8(4):376–388. doi:10.1016/j.stem.2011.03.001

Apostolou E, Hochedlinger K (2013) Chromatin dynamics during cellular reprogramming. Nature 502(7472):462–471. doi:10.1038/nature12749

Apostolou E, Ferrari F, Walsh RM, Bar-Nur O, Stadtfeld M, Cheloufi S, Stuart HT, Polo JM, Ohsumi TK, Borowsky ML, Kharchenko PV, Park PJ, Hochedlinger K (2013) Genome-wide chromatin interactions of the Nanog locus in pluripotency, differentiation, and reprogramming. Cell Stem Cell 12(6):699–712. doi:10.1016/j.stem.2013.04.013

Bannister AJ, Kouzarides T (2011) Regulation of chromatin by histone modifications. Cell Res 21(3):381–395. doi:10.1038/cr.2011.22

Barth TK, Imhof A (2010) Fast signals and slow marks: the dynamics of histone modifications. Trends Biochem Sci 35(11):618–626. doi:10.1016/j.tibs.2010.05.006

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL, Lander ES (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125(2):315–326. doi:10.1016/j.cell.2006.02.041

Bhutani N, Brady JJ, Damian M, Sacco A, Corbel SY, Blau HM (2010) Reprogramming towards pluripotency requires AID-dependent DNA demethylation. Nature 463(7284):1042–1047. doi:10.1038/nature08752, nature08752 [pii]

Biran A, Meshorer E (2012) Concise review: chromatin and genome organization in reprogramming. Stem Cells 30(9):1793–1799. doi:10.1002/stem.1169

Blaschke K, Ebata KT, Karimi MM, Zepeda-Martinez JA, Goyal P, Mahapatra S, Tam A, Laird DJ, Hirst M, Rao A, Lorincz MC, Ramalho-Santos M (2013) Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. Nature 500(7461):222–226. doi:10.1038/nature12362

Blau HM, Chiu CP, Webster C (1983) Cytoplasmic activation of human nuclear genes in stable heterocaryons. Cell 32(4):1171–1180

Bock C, Kiskinis E, Verstappen G, Gu H, Boulting G, Smith ZD, Ziller M, Croft GF, Amoroso MW, Oakley DH, Gnirke A, Eggan K, Meissner A (2011) Reference Maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. Cell 144 (3):439–452. doi:10.1016/j.cell.2010.12.032

Boland MJ, Nazor KL, Loring JF (2014) Epigenetic regulation of pluripotency and differentiation. Circ Res 115(2):311–324. doi:10.1161/CIRCRESAHA.115.301517

Bourc'his D, Xu GL, Lin CS, Bollman B, Bestor TH (2001) Dnmt3L and the establishment of maternal genomic imprints. Science 294(5551):2536–2539. doi:10.1126/science.1065848

Brogaard K, Xi L, Wang JP, Widom J (2012) A map of nucleosome positions in yeast at base-pair resolution. Nature 486(7404):496–501. doi:10.1038/nature11142

Brons IG, Smithers LE, Trotter MW, Rugg-Gunn P, Sun B, Chuva de Sousa Lopes SM, Howlett SK, Clarkson A, Ahrlund-Richter L, Pedersen RA, Vallier L (2007) Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature 448(7150):191–195. doi:10.1038/nature05950

Buganim Y, Faddah DA, Cheng AW, Itskovich E, Markoulaki S, Ganz K, Klemm SL, van Oudenaarden A, Jaenisch R (2012) Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. Cell 150 (6):1209–1222. doi:10.1016/j.cell.2012.08.023

Buganim Y, Faddah DA, Jaenisch R (2013) Mechanisms and models of somatic cell reprogramming. Nat Rev Genet 14(6):427–439. doi:10.1038/nrg3473

Cantone I, Fisher AG (2013) Epigenetic programming and reprogramming during development. Nat Struct Mol Biol 20(3):282–289. doi:10.1038/nsmb.2489

Carey BW, Markoulaki S, Hanna J, Saha K, Gao Q, Mitalipova M, Jaenisch R (2009) Reprogramming of murine and human somatic cells using a single polycistronic vector. Proc Natl Acad Sci USA 106(1):157–162. doi:10.1073/pnas.0811426106

Carone BR, Hung JH, Hainer SJ, Chou MT, Carone DM, Weng Z, Fazzio TG, Rando OJ (2014) High-resolution mapping of chromatin packaging in mouse embryonic stem cells and sperm. Dev Cell 30(1):11–22. doi:10.1016/j.devcel.2014.05.024

Chan YS, Goke J, Ng JH, Lu X, Gonzales KA, Tan CP, Tng WQ, Hong ZZ, Lim YS, Ng HH (2013) Induction of a human pluripotent state with distinct regulatory circuitry that resembles preimplantation epiblast. Cell Stem Cell 13(6):663–675. doi:10.1016/j.stem.2013.11.015

Chen J, Liu H, Liu J, Qi J, Wei B, Yang J, Liang H, Chen Y, Chen J, Wu Y, Guo L, Zhu J, Zhao X, Peng T, Zhang Y, Chen S, Li X, Li D, Wang T, Pei D (2013) H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. Nat Genet 45(1):34–42. doi:10.1038/ng.2491

Christophorou MA, Castelo-Branco G, Halley-Stott RP, Oliveira CS, Loos R, Radzisheuskaya A, Mowen KA, Bertone P, Silva JC, Zernicka-Goetz M, Nielsen ML, Gurdon JB, Kouzarides T (2014) Citrullination regulates pluripotency and histone H1 binding to chromatin. Nature 507 (7490):104–108. doi:10.1038/nature12942

Cowley SM, Iritani BM, Mendrysa SM, Xu T, Cheng PF, Yada J, Liggitt HD, Eisenman RN (2005) The mSin3A chromatin-modifying complex is essential for embryogenesis and T-cell development. Mol Cell Biol 25(16):6990–7004. doi:10.1128/MCB.25.16.6990-7004.2005

Dawlaty MM, Breiling A, Le T, Barrasa MI, Raddatz G, Gao Q, Powell BE, Cheng AW, Faull KF, Lyko F, Jaenisch R (2014) Loss of Tet enzymes compromises proper differentiation of embryonic stem cells. Dev Cell 29(1):102–111. doi:10.1016/j.devcel.2014.03.003

de Wit E, Bouwman BA, Zhu Y, Klous P, Splinter E, Verstegen MJ, Krijger PH, Festuccia N, Nora EP, Welling M, Heard E, Geijsen N, Poot RA, Chambers I, de Laat W (2013) The pluripotent genome in three dimensions is shaped around pluripotency factors. Nature 501(7466):227–231. doi:10.1038/nature12420

DeVeale B, van der Kooy D, Babak T (2012) Critical evaluation of imprinted gene expression by RNA-Seq: a new perspective. PLoS Genet 8(3):e1002600. doi:10.1371/journal.pgen.1002600

Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485 (7398):376–380. doi:10.1038/nature11082

dos Santos RL, Tosti L, Radzisheuskaya A, Caballero IM, Kaji K, Hendrich B, Silva JC (2014) MBD3/NuRD facilitates induction of pluripotency in a context-dependent manner. Cell Stem Cell 15(1):102–110. doi:10.1016/j.stem.2014.04.019

Dovey OM, Foster CT, Cowley SM (2010) Histone deacetylase 1 (HDAC1), but not HDAC2, controls embryonic stem cell differentiation. Proc Natl Acad Sci USA 107(18):8242–8247. doi:10.1073/pnas.1000478107

Efroni S, Duttagupta R, Cheng J, Dehghani H, Hoeppner DJ, Dash C, Bazett-Jones DP, Le Grice S, McKay RD, Buetow KH, Gingeras TR, Misteli T, Meshorer E (2008) Global transcription in pluripotent embryonic stem cells. Cell Stem Cell 2(5):437–447. doi:10.1016/j.stem.2008.03.021, S1934-5909(08)00161-6 [pii]

Evans MJ, Kaufman MH (1981) Establishment in culture of pluripotential cells from mouse embryos. Nature 292(5819):154–156

Fan Y, Nikitina T, Zhao J, Fleury TJ, Bhattacharyya R, Bouhassira EE, Stein A, Woodcock CL, Skoultchi AI (2005) Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation. Cell 123(7):1199–1212

Fazzio TG, Huff JT, Panning B (2008) An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. Cell 134(1):162–174. doi:10.1016/j.cell.2008.05.031, S0092-8674(08)00692-2 [pii]

Gafni O, Weinberger L, Mansour AA, Manor YS, Chomsky E, Ben-Yosef D, Kalma Y, Viukov S, Maza I, Zviran A, Rais Y, Shipony Z, Mukamel Z, Krupalnik V, Zerbib M, Geula S, Caspi I, Schneir D, Shwartz T, Gilad S, Amann-Zalcenstein D, Benjamin S, Amit I, Tanay A, Massarwa R, Novershtern N, Hanna JH (2013) Derivation of novel human ground state naive pluripotent stem cells. Nature 504(7479):282–286. doi:10.1038/nature12745

Gaspar-Maia A, Alajem A, Polesso F, Sridharan R, Mason MJ, Heidersbach A, Ramalho-Santos J, McManus MT, Plath K, Meshorer E, Ramalho-Santos M (2009) Chd1 regulates open chromatin and pluripotency of embryonic stem cells. Nature 460(7257):863–868. doi:10.1038/nature08212, nature08212 [pii]

Gaspar-Maia A, Alajem A, Meshorer E, Ramalho-Santos M (2011) Open chromatin in pluripotency and reprogramming. Nat Rev Mol Cell Biol 12(1):36–47. doi:10.1038/nrm3036, nrm3036 [pii]

Gifford CA, Ziller MJ, Gu H, Trapnell C, Donaghey J, Tsankov A, Shalek AK, Kelley DR, Shishkin AA, Issner R, Zhang X, Coyne M, Fostel JL, Holmes L, Meldrim J, Guttman M, Epstein C, Park H, Kohlbacher O, Rinn J, Gnirke A, Lander ES, Bernstein BE, Meissner A (2013) Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. Cell 153(5):1149–1163. doi:10.1016/j.cell.2013.04.037

Gu TP, Guo F, Yang H, Wu HP, Xu GF, Liu W, Xie ZG, Shi L, He X, Jin SG, Iqbal K, Shi YG, Deng Z, Szabo PE, Pfeifer GP, Li J, Xu GL (2011) The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. Nature 477(7366):606–610. doi:10.1038/nature10443

Guan JS, Haggarty SJ, Giacometti E, Dannenberg JH, Joseph N, Gao J, Nieland TJ, Zhou Y, Wang X, Mazitschek R, Bradner JE, DePinho RA, Jaenisch R, Tsai LH (2009) HDAC2

negatively regulates memory formation and synaptic plasticity. Nature 459(7243):55–60. doi:10.1038/nature07925

Guo H, Zhu P, Yan L, Li R, Hu B, Lian Y, Yan J, Ren X, Lin S, Li J, Jin X, Shi X, Liu P, Wang X, Wang W, Wei Y, Li X, Guo F, Wu X, Fan X, Yong J, Wen L, Xie SX, Tang F, Qiao J (2014) The DNA methylation landscape of human early embryos. Nature 511(7511):606–610. doi:10.1038/nature13544

Gurdon JB, Elsdale TR, Fischberg M (1958) Sexually mature individuals of *Xenopus laevis* from the transplantation of single somatic nuclei. Nature 182(4627):64–65

Haberland M, Montgomery RL, Olson EN (2009) The many roles of histone deacetylases in development and physiology: implications for disease and therapy. Nat Rev Genet 10 (1):32–42. doi:10.1038/nrg2485

Hackett JA, Sengupta R, Zylicz JJ, Murakami K, Lee C, Down TA, Surani MA (2013) Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. Science 339(6118):448–452. doi:10.1126/science.1229277

Hanna J, Saha K, Pando B, van Zon J, Lengner CJ, Creyghton MP, van Oudenaarden A, Jaenisch R (2009) Direct cell reprogramming is a stochastic process amenable to acceleration. Nature 462 (7273):595–601. doi:10.1038/nature08592

Hezroni H, Sailaja BS, Meshorer E (2011a) Pluripotency-related, valproic acid (VPA)-induced genome-wide histone H3 lysine 9 (H3K9) acetylation patterns in embryonic stem cells. J Biol Chem 286(41):35977–35988. doi:10.1074/jbc.M111.266254

Hezroni H, Tzchori I, Davidi A, Mattout A, Biran A, Nissim-Rafinia M, Westphal H, Meshorer E (2011b) H3K9 histone acetylation predicts pluripotency and reprogramming capacity of ES cells. Nucleus 2(4):300–311

Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-Andre V, Sigova AA, Hoke HA, Young RA (2013) Super-enhancers in the control of cell identity and disease. Cell 155(4):934–947. doi:10.1016/j.cell.2013.09.053

Ho L, Crabtree GR (2010) Chromatin remodelling during development. Nature 463 (7280):474–484. doi:10.1038/nature08911

Ho L, Ronan JL, Wu J, Staahl BT, Chen L, Kuo A, Lessard J, Nesvizhskii AI, Ranish J, Crabtree GR (2009) An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. Proc Natl Acad Sci USA 106 (13):5181–5186. doi:10.1073/pnas.0812889106

Ho L, Miller EL, Ronan JL, Ho WQ, Jothi R, Crabtree GR (2011) esBAF facilitates pluripotency by conditioning the genome for LIF/STAT3 signalling and by regulating polycomb function. Nat Cell Biol 13(8):903–913. doi:10.1038/ncb2285

Hou C, Li L, Qin ZS, Corces VG (2012) Gene density, transcription, and insulators contribute to the partition of the Drosophila genome into physical domains. Mol Cell 48(3):471–484. doi:10.1016/j.molcel.2012.08.031

Hou P, Li Y, Zhang X, Liu C, Guan J, Li H, Zhao T, Ye J, Yang W, Liu K, Ge J, Xu J, Zhang Q, Zhao Y, Deng H (2013) Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. Science 341(6146):651–654. doi:10.1126/science.1239278

Howell CY, Bestor TH, Ding F, Latham KE, Mertineit C, Trasler JM, Chaillet JR (2001) Genomic imprinting disrupted by a maternal effect mutation in the Dnmt1 gene. Cell 104(6):829–838

Hu X, Zhang L, Mao SQ, Li Z, Chen J, Zhang RR, Wu HP, Gao J, Guo F, Liu W, Xu GF, Dai HQ, Shi YG, Li X, Hu B, Tang F, Pei D, Xu GL (2014) Tet and TDG mediate DNA demethylation essential for mesenchymal-to-epithelial transition in somatic cell reprogramming. Cell Stem Cell 14(4):512–522. doi:10.1016/j.stem.2014.01.001

Huangfu D, Maehr R, Guo W, Eijkelenboom A, Snitow M, Chen AE, Melton DA (2008a) Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. Nat Biotechnol 26(7):795–797. doi:10.1038/nbt1418, nbt1418 [pii]

Huangfu D, Osafune K, Maehr R, Guo W, Eijkelenboom A, Chen S, Muhlestein W, Melton DA (2008b) Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. Nat Biotechnol 26(11):1269–1275. doi:10.1038/nbt.1502

Humphrey GW, Wang Y, Russanova VR, Hirai T, Qin J, Nakatani Y, Howard BH (2001) Stable histone deacetylase complexes distinguished by the presence of SANT domain proteins CoREST/kiaa0071 and Mta-L1. J Biol Chem 276(9):6817–6824. doi:10.1074/jbc.M007372200

Hussein SM, Batada NN, Vuoristo S, Ching RW, Autio R, Narva E, Ng S, Sourour M, Hamalainen R, Olsson C, Lundin K, Mikkola M, Trokovic R, Peitz M, Brustle O, Bazett-Jones DP, Alitalo K, Lahesmaa R, Nagy A, Otonkoski T (2011) Copy number variation and selection during reprogramming to pluripotency. Nature 471(7336):58–62. doi:10.1038/nature09871

Ichida JK, Blanchard J, Lam K, Son EY, Chung JE, Egli D, Loh KM, Carter AC, Di Giorgio FP, Koszka K, Huangfu D, Akutsu H, Liu DR, Rubin LL, Eggan K (2009) A small-molecule inhibitor of tgf-Beta signaling replaces sox2 in reprogramming by inducing nanog. Cell Stem Cell 5(5):491–503. doi:10.1016/j.stem.2009.09.012

Jackson M, Krassowska A, Gilbert N, Chevassut T, Forrester L, Ansell J, Ramsahoye B (2004) Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells. Mol Cell Biol 24(20):8862–8871

Jamaladdin S, Kelly RD, O'Regan L, Dovey OM, Hodson GE, Millard CJ, Portolano N, Fry AM, Schwabe JW, Cowley SM (2014) Histone deacetylase (HDAC) 1 and 2 are essential for accurate cell division and the pluripotency of embryonic stem cells. Proc Natl Acad Sci USA 111(27):9840–9845. doi:10.1073/pnas.1321330111

Kaji K, Caballero IM, MacLeod R, Nichols J, Wilson VA, Hendrich B (2006) The NuRD component Mbd3 is required for pluripotency of embryonic stem cells. Nat Cell Biol 8 (3):285–292

Kaji K, Nichols J, Hendrich B (2007) Mbd3, a component of the NuRD co-repressor complex, is required for development of pluripotent cells. Development 134(6):1123–1132. doi:10.1242/dev.02802

Kidder BL, Palmer S (2012) HDAC1 regulates pluripotency and lineage specific transcriptional networks in embryonic and trophoblast stem cells. Nucleic Acids Res 40(7):2925–2939. doi:10.1093/nar/gkr1151

Kobayashi H, Sakurai T, Imai M, Takahashi N, Fukuda A, Yayoi O, Sato S, Nakabayashi K, Hata K, Sotomaru Y, Suzuki Y, Kono T (2012) Contribution of intragenic DNA methylation in mouse gametic DNA methylomes to establish oocyte-specific heritable marks. PLoS Genet 8 (1):e1002440. doi:10.1371/journal.pgen.1002440

Koche RP, Smith ZD, Adli M, Gu H, Ku M, Gnirke A, Bernstein BE, Meissner A (2011) Reprogramming factor expression initiates widespread targeted chromatin remodeling. Cell Stem Cell 8(1):96–105. doi:10.1016/j.stem.2010.12.001

Kramer OH, Zhu P, Ostendorff HP, Golebiewski M, Tiefenbach J, Peters MA, Brill B, Groner B, Bach I, Heinzel T, Gottlicher M (2003) The histone deacetylase inhibitor valproic acid selectively induces proteasomal degradation of HDAC2. EMBO J 22(13):3411–3420. doi:10.1093/emboj/cdg315

Kurdistani SK, Robyr D, Tavazoie S, Grunstein M (2002) Genome-wide binding map of the histone deacetylase Rpd3 in yeast. Nat Genet 31(3):248–254. doi:10.1038/ng907

Laherty CD, Yang WM, Sun JM, Davie JR, Seto E, Eisenman RN (1997) Histone deacetylases associated with the mSin3 corepressor mediate mad transcriptional repression. Cell 89 (3):349–356

Landry J, Sharov AA, Piao Y, Sharova LV, Xiao H, Southon E, Matta J, Tessarollo L, Zhang YE, Ko MS, Kuehn MR, Yamaguchi TP, Wu C (2008) Essential role of chromatin remodeling protein Bptf in early mouse embryos and embryonic stem cells. PLoS Genet 4(10):e1000241. doi:10.1371/journal.pgen.1000241

Laurent L, Wong E, Li G, Huynh T, Tsirigos A, Ong CT, Low HM, Kin Sung KW, Rigoutsos I, Loring J, Wei CL (2010) Dynamic changes in the human methylome during differentiation. Genome Res 20(3):320–331. doi:10.1101/gr.101907.109

Lee JT, Bartolomei MS (2013) X-inactivation, imprinting, and long noncoding RNAs in health and disease. Cell 152(6):1308–1323. doi:10.1016/j.cell.2013.02.016

Lee CK, Shibata Y, Rao B, Strahl BD, Lieb JD (2004) Evidence for nucleosome depletion at active regulatory regions genome-wide. Nat Genet 36(8):900–905. doi:10.1038/ng1400

Li E, Bestor TH, Jaenisch R (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell 69(6):915–926

Li R, Liang J, Ni S, Zhou T, Qing X, Li H, He W, Chen J, Li F, Zhuang Q, Qin B, Xu J, Li W, Yang J, Gan Y, Qin D, Feng S, Song H, Yang D, Zhang B, Zeng L, Lai L, Esteban MA, Pei D (2010) A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. Cell Stem Cell 7(1):51–63. doi:10.1016/j.stem.2010.04.014

Li X, Li L, Pandey R, Byun JS, Gardner K, Qin Z, Dou Y (2012) The histone acetyltransferase MOF is a key regulator of the embryonic stem cell core transcriptional network. Cell Stem Cell 11(2):163–178. doi:10.1016/j.stem.2012.04.023

Liang G, Taranova O, Xia K, Zhang Y (2010) Butyrate promotes induced pluripotent stem cell generation. J Biol Chem 285(33):25516–25521. doi:10.1074/jbc.M110.142059

Liang G, He J, Zhang Y (2012) Kdm2b promotes induced pluripotent stem cell generation by facilitating gene activation early in reprogramming. Nat Cell Biol 14(5):457–466. doi:10.1038/ncb2483

Lin CY, Loven J, Rahl PB, Paranal RM, Burge CB, Bradner JE, Lee TI, Young RA (2012) Transcriptional amplification in tumor cells with elevated c-Myc. Cell 151(1):56–67. doi:10.1016/j.cell.2012.08.026

Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, Edsall L, Antosiewicz-Bourget J, Stewart R, Ruotti V, Millar AH, Thomson JA, Ren B, Ecker JR (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 462(7271):315–322. doi:10.1038/nature08514, nature08514 [pii]

Lister R, Pelizzola M, Kida YS, Hawkins RD, Nery JR, Hon G, Antosiewicz-Bourget J, O'Malley R, Castanon R, Klugman S, Downes M, Yu R, Stewart R, Ren B, Thomson JA, Evans RM, Ecker JR (2011) Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. Nature 471(7336):68–73. doi:10.1038/nature09798

Lister R, Mukamel EA, Nery JR, Urich M, Puddifoot CA, Johnson ND, Lucero J, Huang Y, Dwork AJ, Schultz MD, Yu M, Tonti-Filippini J, Heyn H, Hu S, Wu JC, Rao A, Esteller M, He C, Haghighi FG, Sejnowski TJ, Behrens MM, Ecker JR (2013) Global epigenomic reconfiguration during mammalian brain development. Science 341(6146):1237905. doi:10.1126/science.1237905

Livyatan I, Meshorer E (2013) The HDAC interaction network. Mol Syst Biol 9:671. doi:10.1038/msb.2013.33

Lorthongpanich C, Cheow LF, Balu S, Quake SR, Knowles BB, Burkholder WF, Solter D, Messerschmidt DM (2013) Single-cell DNA-methylation analysis reveals epigenetic chimerism in preimplantation embryos. Science 341(6150):1110–1112. doi:10.1126/science.1240617

Luo M, Ling T, Xie W, Sun H, Zhou Y, Zhu Q, Shen M, Zong L, Lyu G, Zhao Y, Ye T, Gu J, Tao W, Lu Z, Grummt I (2013) NuRD blocks reprogramming of mouse somatic cells into pluripotent stem cells. Stem Cells 31(7):1278–1286. doi:10.1002/stem.1374

Mali P, Chou BK, Yen J, Ye Z, Zou J, Dowey S, Brodsky RA, Ohm JE, Yu W, Baylin SB, Yusa K, Bradley A, Meyers DJ, Mukherjee C, Cole PA, Cheng L (2010) Butyrate greatly enhances derivation of human induced pluripotent stem cells by promoting epigenetic remodeling and the expression of pluripotency-associated genes. Stem Cells 28(4):713–720. doi:10.1002/stem.402

Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB, Gnirke A, Jaenisch R, Lander ES (2008) Genome-scale DNA

methylation maps of pluripotent and differentiated cells. Nature 454(7205):766–770. doi:10.1038/nature07107

Mekhoubad S, Bock C, de Boer AS, Kiskinis E, Meissner A, Eggan K (2012) Erosion of dosage compensation impacts human iPSC disease modeling. Cell Stem Cell 10(5):595–609. doi:10.1016/j.stem.2012.02.014

Melcer S, Hezroni H, Rand E, Nissim-Rafinia M, Skoultchi A, Stewart CL, Bustin M, Meshorer E (2012) Histone modifications and lamin A regulate chromatin protein dynamics in early embryonic stem cell differentiation. Nat Commun 3:910. doi:10.1038/ncomms1915

Meshorer E, Yellajoshula D, George E, Scambler PJ, Brown DT, Misteli T (2006) Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. Dev Cell 10(1):105–116

Messerschmidt DM, Knowles BB, Solter D (2014) DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. Genes Dev 28(8):812–828. doi:10.1101/gad.234294.113

Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, Lee W, Mendenhall E, O'Donovan A, Presser A, Russ C, Xie X, Meissner A, Wernig M, Jaenisch R, Nusbaum C, Lander ES, Bernstein BE (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448 (7153):553–560. doi:10.1038/nature06008, nature06008 [pii]

Mikkelsen TS, Hanna J, Zhang X, Ku M, Wernig M, Schorderet P, Bernstein BE, Jaenisch R, Lander ES, Meissner A (2008) Dissecting direct reprogramming through integrative genomic analysis. Nature 454(7200):49–55. doi:10.1038/nature07056

Montgomery RL, Davis CA, Potthoff MJ, Haberland M, Fielitz J, Qi X, Hill JA, Richardson JA, Olson EN (2007) Histone deacetylases 1 and 2 redundantly regulate cardiac morphogenesis, growth, and contractility. Genes Dev 21(14):1790–1802. doi:10.1101/gad.1563807

Najm FJ, Chenoweth JG, Anderson PD, Nadeau JH, Redline RW, McKay RD, Tesar PJ (2011) Isolation of epiblast stem cells from preimplantation mouse embryos. Cell Stem Cell 8 (3):318–325. doi:10.1016/j.stem.2011.01.016

Nakagawa M, Koyanagi M, Tanabe K, Takahashi K, Ichisaka T, Aoi T, Okita K, Mochiduki Y, Takizawa N, Yamanaka S (2008) Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nat Biotechnol 26(1):101–106. doi:10.1038/nbt1374, nbt1374 [pii]

Ochocki JD, Simon MC (2013) Nutrient-sensing pathways and metabolic regulation in stem cells. J Cell Biol 203(1):23–33. doi:10.1083/jcb.201303110

Okano M, Bell DW, Haber DA, Li E (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 99(3):247–257, S0092-8674(00)81656-6 [pii]

Onder TT, Kara N, Cherry A, Sinha AU, Zhu N, Bernt KM, Cahan P, Marcarci BO, Unternaehrer J, Gupta PB, Lander ES, Armstrong SA, Daley GQ (2012) Chromatin-modifying enzymes as modulators of reprogramming. Nature 483(7391):598–602. doi:10.1038/nature10953

Pan G, Tian S, Nie J, Yang C, Ruotti V, Wei H, Jonsdottir GA, Stewart R, Thomson JA (2007) Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. Cell Stem Cell 1(3):299–312. doi:10.1016/j.stem.2007.08.003, S1934-5909(07) 00122-1 [pii]

Pannetier M, Feil R (2007) Epigenetic stability of embryonic stem cells and developmental potential. Trends Biotechnol 25(12):556–562. doi:10.1016/j.tibtech.2007.09.003, S0167-7799(07)00261-2 [pii]

Papp B, Plath K (2013) Epigenetics of reprogramming to induced pluripotency. Cell 152 (6):1324–1343. doi:10.1016/j.cell.2013.02.043

Pauklin S, Vallier L (2013) The cell-cycle state of stem cells determines cell fate propensity. Cell 155(1):135–147. doi:10.1016/j.cell.2013.08.031

Phillips-Cremins JE, Sauria ME, Sanyal A, Gerasimova TI, Lajoie BR, Bell JS, Ong CT, Hookway TA, Guo C, Sun Y, Bland MJ, Wagstaff W, Dalton S, McDevitt TC, Sen R, Dekker J, Taylor J,

Corces VG (2013) Architectural protein subclasses shape 3D organization of genomes during lineage commitment. Cell 153(6):1281–1295. doi:10.1016/j.cell.2013.04.053

Polo JM, Anderssen E, Walsh RM, Schwarz BA, Nefzger CM, Lim SM, Borkent M, Apostolou E, Alaei S, Cloutier J, Bar-Nur O, Cheloufi S, Stadtfeld M, Figueroa ME, Robinton D, Natesan S, Melnick A, Zhu J, Ramaswamy S, Hochedlinger K (2012) A molecular roadmap of reprogramming somatic cells into iPS cells. Cell 151(7):1617–1632. doi:10.1016/j.cell.2012.11.039

Qin H, Diaz A, Blouin L, Lebbink RJ, Patena W, Tanbun P, LeProust EM, McManus MT, Song JS, Ramalho-Santos M (2014) Systematic identification of barriers to human iPSC generation. Cell 158(2):449–461. doi:10.1016/j.cell.2014.05.040

Rais Y, Zviran A, Geula S, Gafni O, Chomsky E, Viukov S, Mansour AA, Caspi I, Krupalnik V, Zerbib M, Maza I, Mor N, Baran D, Weinberger L, Jaitin DA, Lara-Astiaso D, Blecher-Gonen R, Shipony Z, Mukamel Z, Hagai T, Gilad S, Amann-Zalcenstein D, Tanay A, Amit I, Novershtern N, Hanna JH (2013) Deterministic direct reprogramming of somatic cells to pluripotency. Nature 502(7469):65–70. doi:10.1038/nature12587

Reubinoff BE, Pera MF, Fong CY, Trounson A, Bongso A (2000) Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. Nat Biotechnol 18(4):399–404. doi:10.1038/74447

Riising EM, Comet I, Leblanc B, Wu X, Johansen JV, Helin K (2014) Gene silencing triggers polycomb repressive complex 2 recruitment to CpG islands genome wide. Mol Cell 55(3):347–360. doi:10.1016/j.molcel.2014.06.005

Ruiz S, Gore A, Li Z, Panopoulos AD, Montserrat N, Fung HL, Giorgetti A, Bilic J, Batchelder EM, Zaehres H, Scholer HR, Zhang K, Izpisua Belmonte JC (2013) Analysis of protein-coding mutations in hiPSCs and their possible role during somatic cell reprogramming. Nat Commun 4:1382. doi:10.1038/ncomms2381

Samavarchi-Tehrani P, Golipour A, David L, Sung HK, Beyer TA, Datti A, Woltjen K, Nagy A, Wrana JL (2010) Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. Cell Stem Cell 7(1):64–77. doi:10.1016/j.stem.2010.04.015

Schmidt R, Plath K (2012) The roles of the reprogramming factors Oct4, Sox2 and Klf4 in resetting the somatic cell epigenome during induced pluripotent stem cell generation. Genome Biol 13(10):251. doi:10.1186/gb-2012-13-10-251

Schwabish MA, Struhl K (2004) Evidence for eviction and rapid deposition of histones upon transcriptional elongation by RNA polymerase II. Mol Cell Biol 24(23):10111–10117. doi:10.1128/MCB.24.23.10111-10117.2004

Segal E, Fondufe-Mittendorf Y, Chen L, Thastrom A, Field Y, Moore IK, Wang JP, Widom J (2006) A genomic code for nucleosome positioning. Nature 442(7104):772–778. doi:10.1038/nature04979

Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, Parrinello H, Tanay A, Cavalli G (2012) Three-dimensional folding and functional organization principles of the Drosophila genome. Cell 148(3):458–472. doi:10.1016/j.cell.2012.01.010

Shipony Z, Mukamel Z, Cohen NM, Landan G, Chomsky E, Zeliger SR, Fried YC, Ainbinder E, Friedman N, Tanay A (2014) Dynamic and static maintenance of epigenetic memory in pluripotent and somatic cells. Nature. doi:10.1038/nature13458

Shirane K, Toh H, Kobayashi H, Miura F, Chiba H, Ito T, Kono T, Sasaki H (2013) Mouse oocyte methylomes at base resolution reveal genome-wide accumulation of non-CpG methylation and role of DNA methyltransferases. PLoS Genet 9(4):e1003439. doi:10.1371/journal.pgen.1003439

Shyh-Chang N, Locasale JW, Lyssiotis CA, Zheng Y, Teo RY, Ratanasirintrawoot S, Zhang J, Onder T, Unternaehrer JJ, Zhu H, Asara JM, Daley GQ, Cantley LC (2013) Influence of threonine metabolism on S-adenosylmethionine and histone methylation. Science 339(6116):222–226. doi:10.1126/science.1226603

Smith ZD, Meissner A (2013) DNA methylation: roles in mammalian development. Nat Rev Genet 14(3):204–220. doi:10.1038/nrg3354

Smith E, Shilatifard A (2010) The chromatin signaling pathway: diverse mechanisms of recruitment of histone-modifying enzymes and varied biological outcomes. Mol Cell 40(5):689–701. doi:10.1016/j.molcel.2010.11.031

Smith ZD, Chan MM, Mikkelsen TS, Gu H, Gnirke A, Regev A, Meissner A (2012) A unique regulatory phase of DNA methylation in the early mammalian embryo. Nature 484 (7394):339–344. doi:10.1038/nature10960

Smith ZD, Chan MM, Humm KC, Karnik R, Mekhoubad S, Regev A, Eggan K, Meissner A (2014) DNA methylation dynamics of the human preimplantation embryo. Nature 511 (7511):611–615. doi:10.1038/nature13581

Sridharan R, Tchieu J, Mason MJ, Yachechko R, Kuoy E, Horvath S, Zhou Q, Plath K (2009) Role of the murine reprogramming factors in the induction of pluripotency. Cell 136(2):364–377. doi:10.1016/j.cell.2009.01.001

Stadtfeld M, Apostolou E, Ferrari F, Choi J, Walsh RM, Chen T, Ooi SS, Kim SY, Bestor TH, Shioda T, Park PJ, Hochedlinger K (2012) Ascorbic acid prevents loss of Dlk1-Dio3 imprinting and facilitates generation of all-iPS cell mice from terminally differentiated B cells. Nat Genet 44(4):398–405, S391–392. doi:10.1038/ng.1110

Tada M, Tada T, Lefebvre L, Barton SC, Surani MA (1997) Embryonic germ cells induce epigenetic reprogramming of somatic nucleus in hybrid cells. EMBO J 16(21):6510–6520. doi:10.1093/emboj/16.21.6510

Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126(4):663–676

Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 131 (5):861–872. doi:10.1016/j.cell.2007.11.019, S0092-8674(07)01471-7 [pii]

Takeuchi JK, Bruneau BG (2009) Directed transdifferentiation of mouse mesoderm to heart tissue by defined factors. Nature 459(7247):708–711. doi:10.1038/nature08039

Tee WW, Reinberg D (2014) Chromatin features and the epigenetic regulation of pluripotency states in ESCs. Development 141(12):2376–2390. doi:10.1242/dev.096982

Teif VB, Vainshtein Y, Caudron-Herger M, Mallm JP, Marth C, Hofer T, Rippe K (2012) Genome-wide nucleosome positioning during embryonic stem cell development. Nat Struct Mol Biol 19(11):1185–1192. doi:10.1038/nsmb.2419

Tesar PJ, Chenoweth JG, Brook FA, Davies TJ, Evans EP, Mack DL, Gardner RL, McKay RD (2007) New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature 448(7150):196–199. doi:10.1038/nature05972

Theunissen TW, Powell BE, Wang H, Mitalipova M, Faddah DA, Reddy J, Fan ZP, Maetzel D, Ganz K, Shi L, Lungjangwa T, Imsoonthornruksa S, Stelzer Y, Rangarajan S, D'Alessio A, Zhang J, Gao Q, Dawlaty MM, Young RA, Gray NS, Jaenisch R (2014) Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. Cell Stem Cell. doi:10.1016/j.stem.2014.07.002

Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM (1998) Embryonic stem cell lines derived from human blastocysts. Science 282 (5391):1145–1147

Tomizawa S, Kobayashi H, Watanabe T, Andrews S, Hata K, Kelsey G, Sasaki H (2011) Dynamic stage-specific changes in imprinted differentially methylated regions during early mammalian development and prevalence of non-CpG methylation in oocytes. Development 138 (5):811–820. doi:10.1242/dev.061416

Trivedi CM, Luo Y, Yin Z, Zhang M, Zhu W, Wang T, Floss T, Goettlicher M, Noppinger PR, Wurst W, Ferrari VA, Abrams CS, Gruber PJ, Epstein JA (2007) Hdac2 regulates the cardiac hypertrophic response by modulating Gsk3 beta activity. Nat Med 13(3):324–331. doi:10.1038/nm1552

Tsumura A, Hayakawa T, Kumaki Y, Takebayashi S, Sakaue M, Matsuoka C, Shimotohno K, Ishikawa F, Li E, Ueda HR, Nakayama J, Okano M (2006) Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. Genes Cells 11(7):805–814. doi:10.1111/j.1365-2443.2006.00984.x, GTC984 [pii]

Wang Z, Zang C, Cui K, Schones DE, Barski A, Peng W, Zhao K (2009) Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. Cell 138 (5):1019–1031. doi:10.1016/j.cell.2009.06.049

Ware CB, Wang L, Mecham BH, Shen L, Nelson AM, Bar M, Lamba DA, Dauphin DS, Buckingham B, Askari B, Lim R, Tewari M, Gartler SM, Issa JP, Pavlidis P, Duan Z, Blau CA (2009) Histone deacetylase inhibition elicits an evolutionarily conserved self-renewal program in embryonic stem cells. Cell Stem Cell 4(4):359–369. doi:10.1016/j.stem.2009.03.001, S1934-5909(09)00102-7 [pii]

Wei Z, Gao F, Kim S, Yang H, Lyu J, An W, Wang K, Lu W (2013) Klf4 organizes long-range chromosomal interactions with the oct4 locus in reprogramming and pluripotency. Cell Stem Cell 13(1):36–47. doi:10.1016/j.stem.2013.05.010

Weissbein U, Benvenisty N, Ben-David U (2014) Quality control: genome maintenance in pluripotent stem cells. J Cell Biol 204(2):153–163. doi:10.1083/jcb.201310135

Wernig M, Meissner A, Cassady JP, Jaenisch R (2008) c-Myc is dispensable for direct reprogramming of mouse fibroblasts. Cell Stem Cell 2(1):10–12. doi:10.1016/j.stem.2007.12.001

Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, Young RA (2013) Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell 153(2):307–319. doi:10.1016/j.cell.2013.03.035

Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PA, Rappsilber J, Helin K (2011) TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature 473 (7347):343–348. doi:10.1038/nature10066

Woodcock CL, Skoultchi AI, Fan Y (2006) Role of linker histone in chromatin structure and function: H1 stoichiometry and nucleosome repeat length. Chromosome Res 14(1):17–25

Wu H, Zhang Y (2014) Reversing DNA methylation: mechanisms, genomics, and biological functions. Cell 156(1–2):45–68. doi:10.1016/j.cell.2013.12.019

Wu H, D'Alessio AC, Ito S, Wang Z, Cui K, Zhao K, Sun YE, Zhang Y (2011) Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. Genes Dev 25(7):679–684. doi:10.1101/gad.2036011

Xie W, Barr CL, Kim A, Yue F, Lee AY, Eubanks J, Dempster EL, Ren B (2012) Base-resolution analyses of sequence and parent-of-origin dependent DNA methylation in the mouse genome. Cell 148(4):816–831. doi:10.1016/j.cell.2011.12.035

Xie W, Schultz MD, Lister R, Hou Z, Rajagopal N, Ray P, Whitaker JW, Tian S, Hawkins RD, Leung D, Yang H, Wang T, Lee AY, Swanson SA, Zhang J, Zhu Y, Kim A, Nery JR, Urich MA, Kuan S, Yen CA, Klugman S, Yu P, Suknuntha K, Propson NE, Chen H, Edsall LE, Wagner U, Li Y, Ye Z, Kulkarni A, Xuan Z, Chung WY, Chi NC, Antosiewicz-Bourget JE, Slukvin I, Stewart R, Zhang MQ, Wang W, Thomson JA, Ecker JR, Ren B (2013) Epigenomic analysis of multilineage differentiation of human embryonic stem cells. Cell 153 (5):1134–1148. doi:10.1016/j.cell.2013.04.022

Xu C, Inokuma MS, Denham J, Golds K, Kundu P, Gold JD, Carpenter MK (2001) Feeder-free growth of undifferentiated human embryonic stem cells. Nat Biotechnol 19(10):971–974. doi:10.1038/nbt1001-971

Xue Y, Wong J, Moreno GT, Young MK, Cote J, Wang W (1998) NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities. Mol Cell 2 (6):851–861

Yamaguchi S, Hong K, Liu R, Shen L, Inoue A, Diep D, Zhang K, Zhang Y (2012) Tet1 controls meiosis by regulating meiotic gene expression. Nature 492(7429):443–447. doi:10.1038/nature11709

Yamaguchi S, Shen L, Liu Y, Sendler D, Zhang Y (2013) Role of Tet1 in erasure of genomic imprinting. Nature 504(7480):460–464. doi:10.1038/nature12805

Yan Z, Wang Z, Sharova L, Sharov AA, Ling C, Piao Y, Aiba K, Matoba R, Wang W, Ko MS (2008) BAF250B-associated SWI/SNF chromatin-remodeling complex is required to maintain undifferentiated mouse embryonic stem cells. Stem Cells 26(5):1155–1165. doi:10.1634/stemcells.2007-0846, 2007–0846 [pii]

Yanes O, Clark J, Wong DM, Patti GJ, Sanchez-Ruiz A, Benton HP, Trauger SA, Desponts C, Ding S, Siuzdak G (2010) Metabolic oxidation regulates embryonic stem cell differentiation. Nat Chem Biol 6(6):411–417. doi:10.1038/nchembio.364

Ying QL, Wray J, Nichols J, Batlle-Morera L, Doble B, Woodgett J, Cohen P, Smith A (2008) The ground state of embryonic stem cell self-renewal. Nature 453(7194):519–523. doi:10.1038/nature06968, nature06968 [pii]

You A, Tong JK, Grozinger CM, Schreiber SL (2001) CoREST is an integral component of the CoREST-human histone deacetylase complex. Proc Natl Acad Sci USA 98(4):1454–1458. doi:10.1073/pnas.98.4.1454

Young RA (2011) Control of the embryonic stem cell state. Cell 144(6):940–954. doi:10.1016/j.cell.2011.01.032

Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, Nie J, Jonsdottir GA, Ruotti V, Stewart R, Slukvin II, Thomson JA (2007) Induced pluripotent stem cell lines derived from human somatic cells. Science 318(5858):1917–1920. doi:10.1126/science.1151526, 1151526 [pii]

Zhang X, Bolt M, Guertin MJ, Chen W, Zhang S, Cherrington BD, Slade DJ, Dreyton CJ, Subramanian V, Bicker KL, Thompson PR, Mancini MA, Lis JT, Coonrod SA (2012) Peptidylarginine deiminase 2-catalyzed histone H3 arginine 26 citrullination facilitates estrogen receptor alpha target gene activation. Proc Natl Acad Sci USA 109(33):13331–13336. doi:10.1073/pnas.1203280109

Zhou Q, Melton DA (2008) Extreme makeover: converting one cell into another. Cell Stem Cell 3(4):382–388. doi:10.1016/j.stem.2008.09.015

Zhou Q, Brown J, Kanarek A, Rajagopal J, Melton DA (2008) In vivo reprogramming of adult pancreatic exocrine cells to beta-cells. Nature 455(7213):627–632. doi:10.1038/nature07314

Zimmermann S, Kiefer F, Prudenziati M, Spiller C, Hansen J, Floss T, Wurst W, Minucci S, Gottlicher M (2007) Reduced body size and decreased intestinal tumor rates in HDAC2-mutant mice. Cancer Res 67(19):9047–9054. doi:10.1158/0008-5472.CAN-07-0312

Genomic Imprinting

Sanaa Choufani and Rosanna Weksberg

Abstract Genomic imprinting is an epigenetic mechanism that leads to parent of origin-specific differential expression of a subset of genes. The identification of the array of mechanisms regulating monoallelic parent-of-origin specific expression has been instrumental in elucidating a wide range of functional epigenetic mechanisms. Imprinted genes are crucial for normal development. Alterations in the relative expression of these genes, sometimes due to perturbations in gene dosage, cause imprinting disorders. Imprinting disorders are rare and often characterized by clinical features that highlight problems in growth and neurobehavior. Beckwith–Wiedemann (BWS) and Russell–Silver syndromes (RSS), two imprinting disorders with opposite clinical manifestations (overgrowth and undergrowth respectively), can be caused by opposite molecular alterations affecting imprinted gene expression on human chromosome 11p15.5. Studies on the orthologous BWS–RSS genomic region in the mouse have shown that several imprinted genes demonstrate dosage-sensitive functions with important roles in the regulation of placental function and embryonic growth.

Keywords Genomic imprinting • Imprinted genes • DNA methylation • Imprinting disorders • Beckwith–Wiedemann syndrome • Russell–Silver syndrome

S. Choufani
Genetics and Genome Biology Program, The Hospital for Sick Children, Toronto, ON, Canada

R. Weksberg (✉)
Genetics and Genome Biology Program, The Hospital for Sick Children, Toronto, ON, Canada

Division of Clinical and Metabolic Genetics, The Hospital for Sick Children, Toronto, ON M5G 1X8, Canada

Departments of Paediatrics and Molecular Genetics, University of Toronto, Toronto, ON, Canada

Institute of Medical Science, University of Toronto, Toronto, ON, Canada
e-mail: rweksb@sickkids.ca

D.P. Bazett-Jones, G. Dellaire (eds.), *The Functional Nucleus*,
DOI 10.1007/978-3-319-38882-3_19

1 Introduction

Normal imprinted gene function is critical for typical mammalian fetal growth and development (Verona et al. 2003). The human chromosome 11p15.5 region encompasses two imprinted domains, each regulated by their own differentially methylated region (DMR), also called an Imprinting Control Region (ICR). ICR1 at the H19/IGF2 domain (paternally methylated), and ICR2 (also known as KvDMR1) at the *KCNQ1*/*CDKN1C* domain (maternally methylated) (Smith et al. 2007). Loss of imprinting in at least one of these two domains can cause one of two clinically opposite growth disorders. A loss of DNA methylation (LOM) at ICR1 is identified in over 50 % of patients with Russell–Silver syndrome (RSS), which presents with intrauterine and postnatal growth retardation, body asymmetry and feeding difficulties (Gucev et al. 2013). In contrast, a gain of methylation at ICR1 is found in 10 % of patients with Beckwith–Wiedemann syndrome (BS), an overgrowth syndrome presenting with overgrowth of the body and tongue, body asymmetry, congenital malformations and an increased risk for childhood tumors (Choufani et al. 2013). The study of genomic imprinting and its specific role in disease has contributed tremendously to our understanding of the intricate regulation of parent-of-origin expression and normal embryonic development.

2 Genomic Imprinting

2.1 Imprinted Genes

Genomic imprinting is an epigenetically regulated process that causes genes to be expressed in a parent of origin-specific manner rather than from both chromosome homologues (Ferguson-Smith 2011). That is, imprinted genes are expressed exclusively, or preferentially, from either the paternally or maternally inherited allele. This unique transcription profile is driven by epigenetic regulatory mechanisms including DNA methylation (Girardot et al. 2013). DNA methylation is the best studied epigenetic modification, especially in the interrogation of imprinted gene regulation in both mouse and humans (Feinberg et al. 2002). Asymmetric DNA methylation often characterizes the two homologues of imprinted genes. That is, DNA methylation is present on one parental chromosome but not on the other (Hu et al. 1997). Such distinctive differentially methylated regions (DMRs) reflect the opposite transcriptional states of imprinted genes on the two parental chromosomes.

Estimates of imprinted gene prevalence in the human genome is modestly increasing and currently represents <1 % of the human genome. Presently, there are approximately 100 imprinted genes identified in humans (Morison et al. 2005; Choufani et al. 2011; Da Silva-Santiago et al. 2014) and many more predicted (Luedi et al. 2007; Sharp et al. 2010; Choufani et al. 2011; Yuen et al. 2011;

Das et al. 2013; Perez et al. 2015). For a complete list of known and predicted imprinted genes, see http://www.har.mrc.ac.uk/research/genomic_imprinting/ or http://www.geneimprint.com/ or http://igc.otago.ac.nz/home.html/. The majority of imprinted genes were first discovered in mice and then in humans. Although some genes are imprinted in both species, there is substantial species specificity. For example, *Igf2r*, *Ascl2, Tspan32*, *Cd81*, *Tssc4, Nap1l4, Gatm, Dcn*, and Impact are imprinted in mouse but not human (Morison et al. 2005; Monk et al. 2006). Conversely, the homeobox gene *DLX5* is imprinted in human (Okita et al. 2003) but not in mouse although a subtle maternal transcriptional preference was reported in mouse brain (Horike et al. 2005). This discordance means that direct interrogation of the mouse and the human genomes are required to establish species-specific imprinted genes.

In humans, few imprinted genes exist in isolation (Williamson et al. 1998; Evans et al. 2005). They are usually organised in clusters, forming chromosomal domains of tens of kilobases to several megabases in size (Verona et al. 2003; Kelsey and Feil 2013). Clusters of imprinted genes usually contain both maternally and paternally expressed genes. These genes are regulated *in cis* by multiple epigenetic mechanisms such as long noncoding RNAs and by imprinting control centers. ICRs are essential regulatory DNA elements, which are rich in CpG dinucleotides, and are asymmetrically marked by germline-derived DNA methylation on one of the two parental alleles. ICRs are also referred to as differentially methylated regions or DMRs (Kota and Feil 2010; Smallwood and Kelsey 2012). These ICRs regulate transcription of imprinted genes *in cis* resulting in preferential or exclusive transcription of the maternal or paternal allele (Thorvaldsen et al. 1998; Fitzpatrick et al. 2002; Mancini-DiNardo et al. 2003). Genomic imprints are regulated not only by 5-methyl-cytosine (5mC) methylation marks, but also various protein factors, covalent histone modifications and chromatin conformations (Kacem and Feil 2009; Kelsey and Feil 2013). Although most ICRs are methylated on the maternal allele, a few are subject to methylation on the paternal allele, e.g., ICR1 in both humans (Kerjean et al. 2000) and mice (Weaver et al. 2009).

2.2 Genomic Imprinting and Human Development

In mammals, genomic imprinting is essential for viability and normal embryonic development. This absolute requirement for both a maternal and a paternal genome in mammalian development became evident from experiments done in the early 1980's in which attempts to reconstitute a viable mouse embryo entirely from either the maternal germline (gynogenetic conceptus derived from the fusion of two female pronuclei) or the paternal germline (androgenetic conceptus from two male pronuclei) were unsuccessful (Barton et al. 1984; McGrath and Solter 1984; Surani et al. 1984). In humans uniparental inheritance is associated with very rare conditions. Conceptuses with two maternal genomes develop into a teratoma with features of embryonic tissues whereas conceptuses with two paternal genomes

develop into hydatidiform moles with features of placental tissue. In humans, dysregulation of genomic imprinting can lead to a variety of growth and neurodevelopmental disorders (Weksberg 2010; Weksberg et al. 2010; Choufani et al. 2013). Such disorders may result from chromosomal regions of uniparental disomy (UPD) or from targeted epigenetic alterations.

2.3 Genomic Imprints and Epigenomic Reprogramming

Studies of genomic imprinting have only begun to establish the genomic features and epigenetic mechanisms responsible for parent of origin-specific monoallelic gene expression. These studies demonstrate that normal mammalian development depends on a precise developmental epigenetic program with essential roles for differential epigenetic marks on the two parental chromosomes. Genomic imprints are erased and re-established in each generation (Arnaud and Feil 2005). The first phase of epigenetic reprogramming at imprinted genes occurs in the primordial germline and involves complete erasure of the methylation marks to remove the existing parental methylation patterns followed by establishment of a new sex-specific DNA methylation pattern in the gametes (Constancia et al. 1998; Lee et al. 2002). These processes take place during oocyte maturation in females and at birth in males (Hajkova 2011). A second phase of genome-wide epigenetic reprogramming occurs after fertilization but before implantation of the embryo. In this phase, there is global demethylation of non-imprinted genes followed by remethylation of such sites after implantation of the embryo. Notably, imprinted genes are normally protected from this second phase of epigenetic reprogramming and remain intact throughout embryonic development (Howell et al. 2001; Dean et al. 2003).

During fetal development imprinted genes are epigenetically programmed to be expressed in a tissue-specific and developmental time-specific manner (Ohlsson et al. 1993; Hetts et al. 1997). The mono-allelic DNA methylation marks and chromatin features that control imprinted gene expression are somatically maintained throughout development. Yet, for many imprinted genes, the mono-allelic expression is lineage or tissue-specific, presumably because imprinting marks are interpreted by cell type/tissue-specific factors.

2.4 Establishment and Maintenance of DNA Methylation at Imprinting Centers

DNA methylation marks at imprinting control regions can act as an epigenetic memory that is inherited through the germline and is stable throughout development. In order for imprinted marks to resist extensive epigenetic changes during

epigenomic reprogramming, specific factors need to be in place to confer such flexibility and resistance to a varying epigenomic environment. This is mediated through targeted epigenetic machinery to the ICRs. DNA methylation marks are established at ICRs by the de novo DNA methyltransferases, DNMT3A and DNMT3B, and their co-factor DNMT3L (Bourc'his et al. 2001). DNMT3L lacks DNA methyltransferase activity, but forms complexes with DNMT3A and DNMT3B, modulating their activity and attenuating their intrinsic sequence preferences (Wienholz et al. 2010). As for the maintenance of DNA methylation marks at ICRs, some of the factors important for this process have been identified. These include the KRAB zinc finger protein ZFP57 (Li et al. 2008) and developmental pluripotency-associated protein 3 (DPPA3; also known as PGC7 or Stella) (Nakamura et al. 2007). Dnmt1 is also required for faithful maintenance of methylation at DMRs not only in the embryo after fertilisation, but also during the later stages of spermatogenesis. Dnmt1 also has a subsidiary role in completing *de novo* methylation in oocytes by filling in gaps (hemimethylated CpG sites) left by Dnmt3a (Shirane et al. 2013).

This exceptional epigenetic maintenance of ICRs is also linked to differential patterns of histone lysine and arginine methylation between the parental chromosomes (Kelsey and Feil 2013). Chromatin associated with the DNA-methylated alleles of ICRs is consistently marked by histone H3 lysine-9 trimethylation (H3K9me3), H3 lysine-64 tri-methylation (H3K64me3) and H4 lysine-20 trimethylation (H4K20me3). Such ICRs are also bound by the heterochromatin protein-1 gamma (HP1γ) (Delaval et al. 2007; Pannetier et al. 2008; Girardot et al. 2014). Somatic maintenance of the unmethylated allele is conferred by a different set of nuclear proteins such as the zinc-finger proteins ZFP42 (Kim et al. 2011a, b) and 'CCCTC-binding factor' (CTCF). CTCF is an essential architectural protein, an insulator, involved in the formation of long-range interactions and chromatin loops (Lewis and Murrell 2004; Franco et al. 2014). It binds to the unmethylated allele of the IGF2/H19 ICR on chromosome 11p15.5. Interestingly, this ICR also carries binding sites for the transcription factors OCT4 and SOX2 (Abi Habib et al. 2014). Mutations at these sites lead to aberrant gain of DNA methylation and are associated with the foetal overgrowth disorder, Beckwith–Wiedemann syndrome (Abi Habib et al. 2014). This finding indicates that pluripotency-linked transcriptional regulators protect the unmethylated allele of this ICR against *de novo* DNA methylation in the early embryo.

Chromatin associated with the unmethylated alleles of ICRs is characterized by histone H3 and H4 acetylation at lysine residues and by enrichment of H3 lysine-4 di- and/or trimethylation (H3K4me2/H3K4me3) (Delaval et al. 2007). Biochemical studies have shown that H3K4-methylation is repressive as it prevents the DNMT3A/DNMT3L complex from binding and hence, protects against de novo DNA methylation (Ooi et al. 2007).

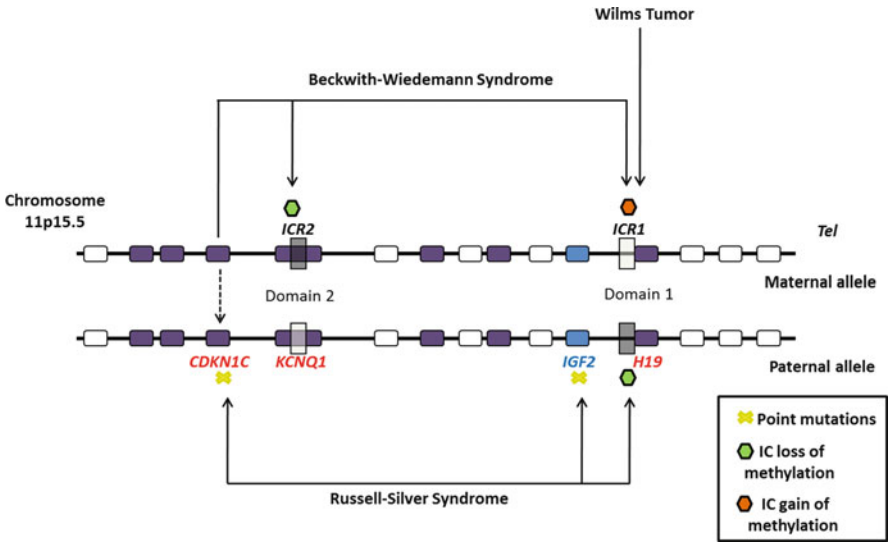

Fig. 1 Schematic representation of imprinted gene clusters on human chromosome 11p15.5. Imprinted genes are indicated as *filled boxes* and non-imprinted genes as *empty boxes*. Paternally expressed genes are indicated in *blue* and maternally expressed genes in *purple colour*. *Hollow rectangles* show the location on normally unmethylated imprinting center (ICR) and *filled rectangles* indicated the ICR is normally methylated. Methylation alterations, such as loss of methylation (*green hexagon*) and gain of methylation (*brown hexagon*), show the locations of these changes in each of the two syndromes: BWS and RSS spectrum. In the telomeric domain are two imprinted genes, *H19* and insulin-like growth factor 2 (*IGF2*). *IGF2* is a paternally expressed fetal growth factor and *H19* is a noncoding RNA. ICR1 is usually methylated on the paternal chromosome and unmethylated on the maternal chromosome. Normally, the H19 gene is expressed from the maternal allele and *IGF2* from the paternal allele. Loss of methylation (LOM) at ICR1 leads to bi-allelic expression of *H19* and no expression of *IGF2*, resulting in RSS. Conversely, gain of methylation (GOM) at ICR1 leads to bi-allelic expression of *IGF2* and no expression of *H19* resulting in BWS. The centromeric domain contains several imprinted genes, including *KCNQ1, KCNQ1OT1* (long non-coding RNA within the *KCNQ1* gene, not shown in this figure), and *CDKN1C*. ICR2 at the promoter for *KCNQ1OT1* regulates expression of *KCNQ1OT1*, which is a paternally expressed noncoding transcript that further regulates *in cis* the expression of the maternally expressed imprinted genes in the centromeric domain. LOM at ICR2 leading to bi-allelic expression of *KCNQ1OT1* is found in 50 % of BWS patients. This epigenetic alteration leads to reduced expression of the growth-regulating gene, *CDKN1C*

2.5 Chromosome 11p15.5 Imprinted Cluster

One of the best studied imprinted clusters is on human chromosome 11p15.5 (Fig. 1). Cis-acting regulatory elements and trans-acting factors are both known to be involved in the regulation of 11p15 imprinting, highlighting new potential mechanisms for many diseases associated with this region such as RSS, BWS and Wilms tumor (Netchine et al. 2012).

The chromosome 11p15.5 imprinted cluster is approximately 1 Mb in size. It contains 2 imprinted domains ICR1 and ICR2 (Fig. 1). ICR1 overlaps a 5 kb region

that contains 7 binding sites for CTCF. In the absence of DNA methylation on the maternal allele CTCF binding acts as an insulator, preventing the activation of *IGF2* and supporting *H19* expression via downstream enhancers. In contrast, on the paternal allele DNA methylation does not permit CTCF to bind. This enhancer-blocking activity results in *IGF2* activation and *H19* silencing (Singh and Srivastava 2008).

The *IGF2*/*H19* domain plays a dual role in development and tumorigenesis. It is regulated by ICR1 which is located 2 kb up- stream of the *H19* gene. Shared endodermal and mesodermal enhancers downstream of *H19* (Leighton et al. 1995) modulate the transcription of *IGF2* and *H19* genes in an allele-specific manner. *IGF2*, an important growth factor during embryogenesis, is transcribed from the methylated paternal allele, whereas the *H19* gene, which encodes an untranslated RNA, is transcribed from the unmethylated maternal allele. Several studies have illustrated the complex regulation of the *IGF2*/*H19* domain both in humans and in animal models. Chromatin conformation studies show that the domain undergoes complex allele-specific long-range chromatin interactions that prevent *H19* transcription from the paternal allele and *IGF2* transcription from the maternal allele (Qiu et al. 2008; Ribeiro de Almeida et al. 2012). Recent studies with human samples confirm that loss of imprinting (LOI) at the *IGF2*/*H19* domain disturbs these long-range chromatin interactions (Zhang et al. 2011).

Domain 2 spans 750-kb and is controlled by a maternally methylated, intragenic ICR located within intron 10 of the *KCNQ1* gene (Fitzpatrick et al. 2002). The *CDKN1C* gene is a negative growth regulator, expressed from the maternal allele, and the untranslated transcript *KCNQ1OT1* is expressed from the paternal allele. *KCNQ1OT1* stabilizes long-distance chromatin interactions between the ICR and the *KCNQ1* gene promoter and is thought to regulate *in cis* the expression of several imprinted genes in the domain 2 cluster [for review, see (Choufani et al. 2013)]. Imprint regulation in this domain is mediated by chromatin looping dependent on imprinted long non-coding RNA, *KCNQ1OT1* (Franco et al. 2014).

3 Imprinting Disorders

Imprinting disorders are rare. They are characterized by defects in growth and neurobehavioral regulation. Beckwith–Wiedemann syndrome (BWS) and Russell–Silver syndrome (RSS) are two imprinting disorders with opposite clinical manifestations caused by opposite imbalances on chromosome 11p15.5 imprinted gene expression from the ICR1 and ICR2 domains (Table 1).

Table 1 Molecular findings in Beckwith–Wiedemann Russell–Silver syndromes

Imprinting disorder	OMIM	Frequency	Affected chromosome	Type of mutation/ epimutation	Frequency of mutation/ epimutation
Russell–Silver syndrome	180860	1/75,000–1/1000,000	Chromosome 7	Maternal UPD7	Approximately 10 %
			Chromosome 11p15	Maternal UPD11p15	Single cases
				Maternal duplication	<1 %
				ICR1 loss of methylation	50 %
				Gain of function point mutations in *CDKN1C*	One family reported
				Paternally inherited mutations in *IGF2*	One family reported
Beckwith–Wiedemann syndrome	130650	1/15,000	Chromosome 11p15	Paternal UPD11p15	25 %
				Chromosomal aberrations	2–4 %
				ICR1 gain of methylation including maternally inherited deletions and OTC	
				Binding sites mutations	5–10 %
				ICR2 loss of methylation	50 %
				CDKN1C mutations	5 % (sporadic); 40 % (familial)
			Genome-wide	Multilocus imprinting defects	Approximately 20 %

3.1 Beckwith–Wiedemann Syndrome

Beckwith–Wiedemann syndrome (BWS; OMIM#130650) is a congenital overgrowth disorder associated with tumor predisposition. This phenotypically heterogeneous disorder can be characterized by one or more of the following features: macrosomia, macroglossia, ear abnormalities, neonatal hypoglycemia, visceromegaly, adrenocortical cytomegaly, hemihyperplasia, abdominal wall defects and perhaps most significantly, an increased tumor risk (Elliott and Maher 1994).

3.1.1 Epigenetic Alterations in BWS

BWS is associated with epigenetic defects affecting the two imprinted domains that regulate the gene expression of several imprinted genes on chromosome 11p15.5. While domain 1 (*H19*/*IGF2*) epigenetic abnormalities are found in only 30 % of BWS patients including cases with paternal uniparental disomy for 11p15, this type of epigenetic defect is associated with the highest risk (~30 %) of embryonal tumors especially Wilms tumor and hepatoblastoma. In domain 2, the epigenetic error, loss of methylation at ICR2, is found in ~50 % of BWS patients (Weksberg et al. 2001).

3.1.2 Genetic Alterations in BWS

Epigenetic abnormalities in BWS are often associated with genomic alterations especially in domain 1 (Baskin et al. 2014). Recently, OCT4 and SOX2 have been shown to be involved in controlling the methylation of the IC1 domain. Mutations or small deletions affecting these binding sites lead to complete gain of methylation at ICR1 by causing chromatin conformational changes (Kim et al. 2011a, b). Recent molecular studies have shown that OCT4/SOX2 mutations and deletions within ICR1 are common in patients with BWS and account for ~20 % of BWS patients with isolated IC1 gain of methylation (Abi Habib et al. 2014). BWS can be associated with mutations in the *CDKN1C* gene in 5 % of cases. The incidence of such mutations increases to 40 % in familial BWS cases. Furthermore, maternal gametic transmission of translocations and inversions involving imprinted domain 2 on chromosome 11p15.5, alter regional DNA methylation patterns and downregulate *CDKN1C* expression *in cis*, suggesting that these epigenetic alterations are generated by an alteration in "chromatin context" (Smith et al. 2012).

3.1.3 BWS and Multilocus Imprinting Disorder

Emerging literature on another imprinting disorder, transient neonatal diabetes (TND) led a new concept for the study of imprinting disorders, i.e., the study of multiple imprinted loci. Mackay et al. found that a subset of TND patients had maternal loss of methylation affecting multiple imprinted loci in addition to the primary TND locus on chromosome 6q24 (Mackay et al. 2006). This loss of methylation at multiple maternal imprinted loci was found to be associated with germline mutation in a zinc finger protein ZFP57 (Mackay et al. 2008). Recent studies using targeted methylation assays on individuals with BWS have identified that ~20 % of BWS patients with ICR2 loss of methylation also have multilocus loss of methylation (MLOM) at other imprinted loci, mostly at the TND locus (*PLAGL1*) on chromosome 6 and the pseudohypoparathyroidism locus (GNAS on chromosome 20) (Lim et al. 2009; Bliek et al. 2009a, b; Mackay et al. 2015).

Multilocus DNA methylation changes were not associated with either ICR1 gain of methylation or with UPD11 in BWS patients (Azzi et al. 2009).

A recent case report on two affected BWS sibling identified a homozygous frameshift mutation in *NLRP2* in the mother of the two children with BWS implicating *NLRP2* in the establishment and/or maintenance of genomic imprinting (Meyer et al. 2009). These data suggest the presence of trans-mechanism that might be affecting the epigenetic expression of imprinted loci (i.e., ICR2) in BWS. Whether epigenetic multilocus abnormalities in BWS patients affect phenotypic expression of the disorder is still not well defined.

3.1.4 BWS and Assisted Reproduction

In recent years, there has been an increase in the incidence of various imprinting disorders, especially among children conceived using assisted reproductive technologies, such as in vitro fertilization (Maher et al. 2003; Doornbos et al. 2007). Because imprinting occurs during gametogenesis, there are growing concerns that various elements of reproductive assistance procedures prevent genes from being properly imprinted or stably maintained during early embryonic development (Clayton-Smith 2003; Dolinoy et al. 2007).

The underlying mechanisms associating ART with aberrant imprinting are unclear. It is thought that the procedures involved in ART or subfertility itself may alter the primary acquisition of germ cell imprints or their maintenance (Weksberg et al. 2002; Ludwig et al. 2005; Fauque et al. 2007). Recently, there has been emerging studies highlighting the complexity of the domain 2 epigenetic abnormalities (ICR2 LOM). BWS patients born following conception involving assisted reproductive technologies have a 4–9 fold increase of ICR2 LOM (DeBaun et al. 2003; Gicquel et al. 2003; Maher et al. 2003; Halliday et al. 2004; Sutcliffe et al. 2006). Thus, ICR2 methylation could be easily altered by environmental insults. Understanding the process of imprinting and identifying conditions that interfere with normal imprinting may thus help reduce the incidence of these conditions.

3.2 Russell–Silver Syndrome

Russell–Silver syndrome (RSS; OMIM #180860) is characterized by pre- and postnatal growth deficiency, variable dysmorphic facial features, feeding difficulties and body asymmetry without major malformations.

3.2.1 Epigenetic Alterations in RSS

Loss of methylation of the paternal 11p15.5 imprinting control region 1 (ICR1), leading to loss of expression of *IGF2*, and maternal uniparental disomy of chromosome 7 is found in 50 % and in 10 % of RSS patients, respectively.

Children with the chromosome 11p15.5 epimutation had significantly lower weight and height at birth, but a significantly larger head circumference than children without this epigenetic aberration. Children with RSS present impaired somatic development compared to children with intrauterine growth restriction or IUGR, and those with a genetic aberration develop worse (Sienko et al. 2014).

RSS Patients with duplication of 11p15 show a more variable occipitofrontal head circumference at birth, a higher frequency of intellectual disability, and additional anomalies not reported in RSS (Fokstuen and Kotzot 2014). A follow up study of three RSS patients with ICR1 LOM, showed that these patients developed adult-onset diseases such as obesity, hypertension and diabetes mellitus (de Boo and Harding 2006; Takenouchi et al. 2015).

Imprinted domains other than on chromosome 11 and 7 have also been implicated in RSS as well. One recent study has reported several patients with an RSS-compatible clinical phenotype, but with unaltered DNA methylation at the *IGF2*/*H19* locus. In these patients, there was a complete loss of DNA methylation at the ICR controlling the imprinted *DLK1*/*DIO3* domain on human chromosome 14q32.2 (Kagami et al. 2015). Earlier studies had linked maternal uniparental disomy (UPD) of chromosome 14q32.2 to a clinically overlapping imprinting disorder, called Temple Syndrome (TS) (Ioannides et al. 2014). The etiology of observed methylation changes at the two different ICRs has not been delivered.

3.2.2 Genetic Mutations in RSS

CDKN1C mutations affecting the PCNA-binding domain, cause dominant maternally transmitted RSS (Brioude et al. 2013), underscoring again the opposite molecular pictures in BWS and RSS. Recently paternally inherited mutations in *IGF2* have been associated with RSS in a multigenerational pedigree (Begemann et al. 2015). Family members carrying *IGF2* mutations presented with severe growth restriction and clinical features of RSS highlighting the fact that *IGF2* affects postnatal growth in addition to prenatal growth.

3.3 Manifestations of Either BWS or RSS in Multigenerational Families

BWS and RSS can be manifested in multigenerational families depending on the parental transmission of the altered allele. For instance, duplications involving both

ICRs can cause either RSS or BWS, depending on the parental origin of the aberration. Recently, a 1.3-Mb duplication overlapping both ICRs on chromosome 11p15.5p15.4 was found in three generations of family causing an RSS or BWS phenotype depending on parent of origin of transmission (Vals et al. 2015). Another report of 4-generation family, where 1.9 Mb duplication of the chromosome 11p15.5 segment, led to a RSS phenotype when maternally inherited and a Beckwith–Wiedemann phenotype when paternally transmitted (Brown et al. 2014). These studies provide evidence that opposite phenotypes arise based on parental mode of transmission, a finding of particular relevance to genetic counseling of these patients and their families.

4 Summary

Genomic imprinting will benefit tremendously from the emerging whole genome sequencing technologies and their potential to unravel as yet unidentified imprinted regions either germline or somatic (i.e., tissue specific). In addition, the availability of single cell genomic and epigenomic maps currently being generated by large consortia such as Encode and the Epigenome Road Map will bring us a step closer to clarifying cell–type specific imprinting. Studies on genomic imprinting will offer valuable insights into our understanding of the basic mechanisms that underlie normal growth and development. They will also allow an in depth understanding of the intricate nature of the role of genomic imprinting in normal fetal development, epigenetic reprogramming, parent of origin-specific expression and tissue specific imprinting. Characterizing the human imprintome by exploring the roles of different types of epigenetic marks, e.g., DNA methylation and histone modifications in different tissues across development will offer new research avenues to compare findings in human to other model organisms such as mice where one can dissect the basic mechanisms underlying genomic imprinting in greater depth.

Acknowledgments Work in the Weksberg laboratory is supported by the Canadian Institutes of Health Research (CIHR, MOP-126054), Natural Sciences and Engineering Research Council of Canada (NSERC; 418734–2012).

References

Abi Habib W, Azzi S, Brioude F, Steunou V, Thibaud N, Das Neves C, Le Jule M, Chantot-Bastaraud S, Keren B, Lyonnet S, Michot C, Rossi M, Pasquier L, Gicquel C, Rossignol S, Le Bouc Y, Netchine I (2014) Extensive investigation of the IGF2/H19 imprinting control region reveals novel OCT4/SOX2 binding site defects associated with specific methylation patterns in Beckwith–Wiedemann syndrome. Hum Mol Genet 23(21):5763–5773

Arnaud P, Feil R (2005) Epigenetic deregulation of genomic imprinting in human disorders and following assisted reproduction. Birth Defects Res C Embryo Today 75(2):81–97

Azzi S, Rossignol S, Steunou V, Sas T, Thibaud N, Danton F, Le Jule M, Heinrichs C, Cabrol S, Gicquel C, Le Bouc Y, Netchine I (2009) Multilocus methylation analysis in a large cohort of 11p15-related foetal growth disorders (Russell Silver and Beckwith Wiedemann syndromes) reveals simultaneous loss of methylation at paternal and maternal imprinted loci. Hum Mol Genet 18(24):4724–4733

Barton SC, Surani MA, Norris ML (1984) Role of paternal and maternal genomes in mouse development. Nature 311(5984):374–376

Baskin B, Choufani S, Chen YA, Shuman C, Parkinson N, Lemyre E, Micheil Innes A, Stavropoulos DJ, Ray PN, Weksberg R (2014) High frequency of copy number variations (CNVs) in the chromosome 11p15 region in patients with Beckwith–Wiedemann syndrome. Hum Genet 133(3):321–330

Begemann M, Zirn B, Santen G, Wirthgen E, Soellner L, Buttel HM, Schweizer R, van Workum W, Binder G, Eggermann T (2015) Paternally inherited IGF2 mutation and growth restriction. N Engl J Med 373(4):349–356

Bliek J, Alders M, Maas SM, Oostra RJ, Mackay DM, van der Lip K, Callaway JL, Brooks A, van't Padje S, Westerveld A, Leschot NJ, Mannens MM (2009a) Lessons from BWS twins: complex maternal and paternal hypomethylation and a common source of haematopoietic stem cells. Eur J Hum Genet 17(12):1625–1634

Bliek J, Verde G, Callaway J, Maas SM, De Crescenzo A, Sparago A, Cerrato F, Russo S, Ferraiuolo S, Rinaldi MM, Fischetto R, Lalatta F, Giordano L, Ferrari P, Cubellis MV, Larizza L, Temple IK, Mannens MM, Mackay DJ, Riccio A (2009b) Hypomethylation at multiple maternally methylated imprinted regions including PLAGL1 and GNAS loci in Beckwith–Wiedemann syndrome. Eur J Hum Genet 17(5):611–619

Bourc'his D, Xu GL, Lin CS, Bollman B, Bestor TH (2001) Dnmt3L and the establishment of maternal genomic imprints. Science 294(5551):2536–2539

Brioude F, Oliver-Petit I, Blaise A, Praz F, Rossignol S, Le Jule M, Thibaud N, Faussat AM, Tauber M, Le Bouc Y, Netchine I (2013) CDKN1C mutation affecting the PCNA-binding domain as a cause of familial Russell Silver syndrome. J Med Genet 50(12):823–830

Brown LA, Rupps R, Penaherrera MS, Robinson WP, Patel MS, Eydoux P, Boerkoel CF (2014) A cryptic familial rearrangement of 11p15.5, involving both imprinting centers, in a family with a history of short stature. Am J Med Genet A 164A(6):1587–1594

Choufani S, Shapiro JS, Susiarjo M, Butcher DT, Grafodatskaya D, Lou Y, Ferreira JC, Pinto D, Scherer SW, Schaffer LG, Coullin P, Caniggia I, Beyene J, Slim R, Bartolomei M, Weksberg R (2011) A novel approach identifies new differentially methylated regions (DMRs) associated with imprinted genes. Genome Res 21(3):465–476

Choufani S, Shuman C, Weksberg R (2013) Molecular findings in Beckwith–Wiedemann syndrome. Am J Med Genet C Semin Med Genet 163C(2):131–140

Clayton-Smith J (2003) Genomic imprinting as a cause of disease. BMJ 327(7424):1121–1122

Constancia M, Pickard B, Kelsey G, Reik W (1998) Imprinting mechanisms. Genome Res 8 (9):881–900

Da Silva-Santiago SC, Pacheco C, Rocha TC, Brasil SM, Pacheco AC, Silva MM, Araujo FF, De Vasconcelos EJ, De Oliveira DM (2014) The linked human imprintome v1.0: over 120 genes confirmed as imprinted impose a major review on previous censuses. Int J Data Min Bioinform 10(3):329–356

Das R, Lee YK, Strogantsev R, Jin S, Lim YC, Ng PY, Lin XM, Chng K, Yeo G, Ferguson-Smith AC, Ding C (2013) DNMT1 and AIM1 imprinting in human placenta revealed through a genome-wide screen for allele-specific DNA methylation. BMC Genomics 14:685

de Boo HA, Harding JE (2006) The developmental origins of adult disease (Barker) hypothesis. Aust N Z J Obstet Gynaecol 46(1):4–14

Dean W, Santos F, Reik W (2003) Epigenetic reprogramming in early mammalian development and following somatic nuclear transfer. Semin Cell Dev Biol 14(1):93–100

DeBaun MR, Niemitz EL, Feinberg AP (2003) Association of in vitro fertilization with Beckwith–Wiedemann syndrome and epigenetic alterations of LIT1 and H19. Am J Hum Genet 72 (1):156–160

Delaval K, Govin J, Cerqueira F, Rousseaux S, Khochbin S, Feil R (2007) Differential histone modifications mark mouse imprinting control regions during spermatogenesis. EMBO J 26 (3):720–729

Dolinoy DC, Weidman JR, Jirtle RL (2007) Epigenetic gene regulation: linking early developmental environment to adult disease. Reprod Toxicol 23(3):297–307

Doornbos ME, Maas SM, McDonnell J, Vermeiden JP, Hennekam RC (2007) Infertility, assisted reproduction technologies and imprinting disturbances: a Dutch study. Hum Reprod 22 (9):2476–2480

Elliott M, Maher ER (1994) Beckwith–Wiedemann syndrome. J Med Genet 31(7):560–564

Evans HK, Weidman JR, Cowley DO, Jirtle RL (2005) Comparative phylogenetic analysis of blcap/nnat reveals eutherian-specific imprinted gene. Mol Biol Evol 22(8):1740–1748

Fauque P, Jouannet P, Lesaffre C, Ripoche MA, Dandolo L, Vaiman D, Jammes H (2007) Assisted reproductive technology affects developmental kinetics, H19 imprinting control region methylation and H19 gene expression in individual mouse embryos. BMC Dev Biol 7:116

Feinberg AP, Cui H, Ohlsson R (2002) DNA methylation and genomic imprinting: insights from cancer into epigenetic mechanisms. Semin Cancer Biol 12(5):389–398

Ferguson-Smith AC (2011) Genomic imprinting: the emergence of an epigenetic paradigm. Nat Rev Genet 12(8):565–575

Fitzpatrick GV, Soloway PD, Higgins MJ (2002) Regional loss of imprinting and growth deficiency in mice with a targeted deletion of KvDMR1. Nat Genet 32:426–431

Fokstuen S, Kotzot D (2014) Chromosomal rearrangements in patients with clinical features of Silver–Russell syndrome. Am J Med Genet A 164A(6):1595–1605

Franco MM, Prickett AR, Oakey RJ (2014) The role of CCCTC-binding factor (CTCF) in genomic imprinting, development, and reproduction. Biol Reprod 91(5):125

Gicquel C, Gaston V, Mandelbaum J, Siffroi JP, Flahault A, Le Bouc Y (2003) In vitro fertilization may increase the risk of Beckwith–Wiedemann syndrome related to the abnormal imprinting of the KCN1OT gene. Am J Hum Genet 72(5):1338–1341

Girardot M, Feil R, Lleres D (2013) Epigenetic deregulation of genomic imprinting in humans: causal mechanisms and clinical implications. Epigenomics 5(6):715–728

Girardot M, Hirasawa R, Kacem S, Fritsch L, Pontis J, Kota SK, Filipponi D, Fabbrizio E, Sardet C, Lohmann F, Kadam S, Ait-Si-Ali S, Feil R (2014) PRMT5-mediated histone H4 arginine-3 symmetrical dimethylation marks chromatin at G + C-rich regions of the mouse genome. Nucleic Acids Res 42(1):235–248

Gucev ZS, Saranac L, Jancevska A, Tasic V (2013) The degree of H19 hypomethylation in children with Silver–Russel syndrome (SRS) is not associated with the severity of phenotype and the clinical severity score (CSS). Prilozi 34(2):79–83

Hajkova P (2011) Epigenetic reprogramming in the germline: towards the ground state of the epigenome. Philos Trans R Soc Lond B Biol Sci 366(1575):2266–2273

Halliday J, Oke K, Breheny S, Algar E, Amor JD (2004) Beckwith–Wiedemann syndrome and IVF: a case–control study. Am J Hum Genet 75(3):526–528

Hetts SW, Rosen KM, Dikkes P, Villa-Komaroff L, Mozell RL (1997) Expression and imprinting of the insulin-like growth factor II gene in neonatal mouse cerebellum. J Neurosci Res 50 (6):958–966

Horike S, Cai S, Miyano M, Cheng JF, Kohwi-Shigematsu T (2005) Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. Nat Genet 37(1):31–40

Howell CY, Bestor TH, Ding F, Latham KE, Mertineit C, Trasler JM, Chaillet JR (2001) Genomic imprinting disrupted by a maternal effect mutation in the Dnmt1 gene. Cell 104(6):829–838

Hu JF, Nguyen PH, Pham NV, Vu TH, Hoffman AR (1997) Modulation of Igf2 genomic imprinting in mice induced by 5-azacytidine, an inhibitor of DNA methylation. Mol Endocrinol 11(13):1891–1898

Ioannides Y, Lokulo-Sodipe K, Mackay DJ, Davies JH, Temple IK (2014) Temple syndrome: improving the recognition of an underdiagnosed chromosome 14 imprinting disorder: an analysis of 51 published cases. J Med Genet 51(8):495–501

Kacem S, Feil R (2009) Chromatin mechanisms in genomic imprinting. Mamm Genome 20 (9–10):544–556

Kagami M, Mizuno S, Matsubara K, Nakabayashi K, Sano S, Fuke T, Fukami M, Ogata T (2015) Epimutations of the IG-DMR and the MEG3-DMR at the 14q32.2 imprinted region in two patients with Silver–Russell syndrome-compatible phenotype. Eur J Hum Genet 23 (8):1062–1067

Kelsey G, Feil R (2013) New insights into establishment and maintenance of DNA methylation imprints in mammals. Philos Trans R Soc Lond B Biol Sci 368(1609):20110336

Kerjean A, Dupont JM, Vasseur C, Le Tessier D, Cuisset L, Paldi A, Jouannet P, Jeanpierre M (2000) Establishment of the paternal methylation imprint of the human H19 and MEST/PEG1 genes during spermatogenesis. Hum Mol Genet 9(14):2183–2187

Kim JD, Kim H, Ekram MB, Yu S, Faulk C, Kim J (2011a) Rex1/Zfp42 as an epigenetic regulator for genomic imprinting. Hum Mol Genet 20(7):1353–1362

Kim YJ, Cecchini KR, Kim TH (2011b) Conserved, developmentally regulated mechanism couples chromosomal looping and heterochromatin barrier activity at the homeobox gene A locus. Proc Natl Acad Sci USA 108(18):7391–7396

Kota SK, Feil R (2010) Epigenetic transitions in germ cell development and meiosis. Dev Cell 19 (5):675–686

Lee J, Inoue K, Ono R, Ogonuki N, Kohda T, Kaneko-Ishino T, Ogura A, Ishino F (2002) Erasing genomic imprinting memory in mouse clone embryos produced from day 11.5 primordial germ cells. Development 129(8):1807–1817

Leighton PA, Saam JR, Ingram RS, Stewart CL, Tilghman SM (1995) An enhancer deletion affects both H19 and Igf2 expression. Genes Dev 9(17):2079–2089

Lewis A, Murrell A (2004) Genomic imprinting: CTCF protects the boundaries. Curr Biol 14(7): R284–R286

Li X, Ito M, Zhou F, Youngson N, Zuo X, Leder P, Ferguson-Smith AC (2008) A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. Dev Cell 15(4):547–557

Lim D, Bowdin SC, Tee L, Kirby GA, Blair E, Fryer A, Lam W, Oley C, Cole T, Brueton LA, Reik W, Macdonald F, Maher ER (2009). Clinical and molecular genetic features of Beckwith–Wiedemann syndrome associated with assisted reproductive technologies. Hum Reprod 24(3):741–747

Ludwig M, Katalinic A, Gross S, Sutcliffe A, Varon R, Horsthemke B (2005) Increased prevalence of imprinting defects in patients with Angelman syndrome born to subfertile couples. J Med Genet 42(4):289–291

Luedi PP, Dietrich FS, Weidman JR, Bosko JM, Jirtle RL, Hartemink AJ (2007) Computational and experimental identification of novel human imprinted genes. Genome Res 17 (12):1723–1730

Mackay D, Boonen S, Clayton-Smith J, Goodship J, Hahnemann J, Kant S, Njølstad P, Robin N, Robinson D, Siebert R, Shield J, White H, Temple I (2006) A maternal hypomethylation syndrome presenting as transient neonatal diabetes mellitus. Hum Genet 120(2):262–269

Mackay DJ, Callaway JL, Marks SM, White HE, Acerini CL, Boonen SE, Dayanikli P, Firth HV, Goodship JA, Haemers AP, Hahnemann JM, Kordonouri O, Masoud AF, Oestergaard E, Storr J, Ellard S, Hattersley AT, Robinson DO, Temple IK (2008) Hypomethylation of multiple imprinted loci in individuals with transient neonatal diabetes is associated with mutations in ZFP57. Nat Genet 40(8):949–951

Mackay DJ, Eggermann T, Buiting K, Garin I, Netchine I, Linglart A, de Nanclares GP (2015) Multilocus methylation defects in imprinting disorders. Biomol Concepts 6(1):47–57

Maher ER, Brueton LA, Bowdin SC, Luharia A, Cooper W, Cole TR, Macdonald F, Sampson JR, Barratt CL, Reik W, Hawkins MM (2003) Beckwith–Wiedemann syndrome and assisted reproduction technology (ART). J Med Genet 40(1):62–64

Mancini-DiNardo D, Steele SJ, Ingram RS, Tilghman SM (2003) A differentially methylated region within the gene Kcnq1 functions as an imprinted promoter and silencer. Hum Mol Genet 12(3):283–294

McGrath J, Solter D (1984) Completion of mouse embryogenesis requires both the maternal and paternal genomes. Cell 37(1):179–183

Meyer E, Lim D, Pasha S, Tee LJ, Rahman F, Yates JR, Woods CG, Reik W, Maher ER (2009) Germline mutation in NLRP2 (NALP2) in a familial imprinting disorder (Beckwith–Wiedemann Syndrome). PLoS Genet 5(3):e1000423

Monk D, Arnaud P, Apostolidou S, Hills FA, Kelsey G, Stanier P, Feil R, Moore GE (2006) Limited evolutionary conservation of imprinting in the human placenta. Proc Natl Acad Sci USA 103(17):6623–6628

Morison IM, Ramsay JP, Spencer HG (2005) A census of mammalian imprinting. Trends Genet 21:457–465

Nakamura T, Arai Y, Umehara H, Masuhara M, Kimura T, Taniguchi H, Sekimoto T, Ikawa M, Yoneda Y, Okabe M, Tanaka S, Shiota K, Nakano T (2007) PGC7/Stella protects against DNA demethylation in early embryogenesis. Nat Cell Biol 9(1):64–71

Netchine I, Rossignol S, Azzi S, Brioude F, Le Bouc Y (2012) Imprinted anomalies in fetal and childhood growth disorders: the model of Russell–Silver and Beckwith–Wiedemann syndromes. Endocr Dev 23:60–70

Ohlsson R, Nystrom A, Pfeifer-Ohlsson S, Tohonen V, Hedborg F, Schofield P, Flam F, Ekstrom TJ (1993) IGF2 is parentally imprinted during human embryogenesis and in the Beckwith–Wiedemann syndrome. Nat Genet 4(1):94–97

Okita C, Meguro M, Hoshiya H, Haruta M, Sakamoto YK, Oshimura M (2003) A new imprinted cluster on the human chromosome 7q21-q31, identified by human-mouse monochromosomal hybrids. Genomics 81(6):556–559

Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin SP, Allis CD, Cheng X, Bestor TH (2007) DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature 448(7154):714–717

Pannetier M, Julien E, Schotta G, Tardat M, Sardet C, Jenuwein T, Feil R (2008) PR-SET7 and SUV4-20H regulate H4 lysine-20 methylation at imprinting control regions in the mouse. EMBO Rep 9(10):998–1005

Perez JD, Rubinstein ND, Fernandez DE, Santoro SW, Needleman LA, Ho-Shing O, Choi JJ, Zirlinger M, Chen SK, Liu JS, Dulac C (2015) Quantitative and functional interrogation of parent-of-origin allelic expression biases in the brain. Elife 4:e07860

Qiu X, Vu TH, Lu Q, Ling JQ, Li T, Hou A, Wang SK, Chen HL, Hu JF, Hoffman AR (2008) A complex deoxyribonucleic acid looping configuration associated with the silencing of the maternal Igf2 allele. Mol Endocrinol 22(6):1476–1488

Ribeiro de Almeida C, Stadhouders R, Thongjuea S, Soler E, Hendriks RW (2012) DNA-binding factor CTCF and long-range gene interactions in V(D)J recombination and oncogene activation. Blood 119(26):6209–6218

Sharp AJ, Migliavacca E, Dupre Y, Stathaki E, Sailani MR, Baumer A, Schinzel A, Mackay DJ, Robinson DO, Cobellis G, Cobellis L, Brunner HG, Steiner B, Antonarakis SE (2010) Methylation profiling in individuals with uniparental disomy identifies novel differentially methylated regions on chromosome 15. Genome Res 20(9):1271–1278

Shirane K, Toh H, Kobayashi H, Miura F, Chiba H, Ito T, Kono T, Sasaki H (2013) Mouse oocyte methylomes at base resolution reveal genome-wide accumulation of non-CpG methylation and role of DNA methyltransferases. PLoS Genet 9(4):e1003439

Sienko M, Petriczko E, Zajaczek S, Zygmunt-Gorska A, Starzyk J, Korpysz A, Petriczko J, Walczak A, Walczak M (2014) A ten-year observation of somatic development of a first group of Polish children with Silver–Russell syndrome. Neuro Endocrinol Lett 35(4):306–313

Singh V, Srivastava M (2008) Enhancer blocking activity of the insulator at H19-ICR is independent of chromatin barrier establishment. Mol Cell Biol 28(11):3767–3775

Smallwood SA, Kelsey G (2012) De novo DNA methylation: a germ cell perspective. Trends Genet 28(1):33–42

Smith AC, Choufani S, Ferreira JC, Weksberg R (2007) Growth regulation, imprinted genes, and chromosome 11p15.5. Pediatr Res 61(5 Pt 2):43R–47R

Smith AC, Suzuki M, Thompson R, Choufani S, Higgins MJ, Chiu IW, Squire JA, Greally JM, Weksberg R (2012) Maternal gametic transmission of translocations or inversions of human chromosome 11p15.5 results in regional DNA hypermethylation and downregulation of CDKN1C expression. Genomics 99(1):25–35

Surani MA, Barton SC, Norris ML (1984) Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. Nature 308(5959):548–550

Sutcliffe AG, Peters CJ, Bowdin S, Temple K, Reardon W, Wilson L, Clayton-Smith J, Brueton LA, Bannister W, Maher ER (2006) Assisted reproductive therapies and imprinting disorders—a preliminary British survey. Hum Reprod 21(4):1009–1011

Takenouchi T, Awazu M, Eggermann T, Kosaki K (2015) Adult phenotype of Russell–Silver syndrome: a molecular support for Barker–Brenner's theory. Congenit Anom (Kyoto) 55 (3):167–169

Thorvaldsen JL, Duran KL, Bartolomei MS (1998) Deletion of the H19 differentially methylated domain results in loss of imprinted expression of H19 and Igf2. Genes Dev 12(23):3693–3702

Vals MA, Yakoreva M, Kahre T, Mee P, Muru K, Joost K, Teek R, Soellner L, Eggermann T, Ounap K (2015) The frequency of methylation abnormalities among Estonian patients selected by clinical diagnostic scoring systems for Silver–Russell syndrome and Beckwith–Wiedemann syndrome. Genet Test Mol Biomarkers 19:684–691

Verona RI, Mann MR, Bartolomei MS (2003) Genomic imprinting: intricacies of epigenetic regulation in clusters. Annu Rev Cell Dev Biol 19:237–259

Weaver JR, Susiarjo M, Bartolomei MS (2009) Imprinting and epigenetic changes in the early embryo. Mamm Genome 20(9–10):532–543

Weksberg R (2010) Imprinted genes and human disease. Am J Med Genet C Semin Med Genet 154C(3):317–320

Weksberg R, Nishikawa J, Caluseriu O, Fei YL, Shuman C, Wei C, Steele L, Cameron J, Smith A, Ambus I, Li M, Ray PN, Sadowski P, Squire J (2001) Tumor development in the Beckwith–Wiedemann syndrome is associated with a variety of constitutional molecular 11p15 alterations including imprinting defects of KCNQ1OT1. Hum Mol Genet 10(26):2989–3000

Weksberg R, Shuman C, Caluseriu O, Smith AC, Fei YL, Nishikawa J, Stockley TL, Best L, Chitayat D, Olney A, Ives E, Schneider A, Bestor TH, Li M, Sadowski P, Squire J (2002) Discordant KCNQ1OT1 imprinting in sets of monozygotic twins discordant for Beckwith–Wiedemann syndrome. Hum Mol Genet 11(11):1317–1325

Weksberg R, Shuman C, Beckwith JB (2010) Beckwith–Wiedemann syndrome. Eur J Hum Genet 18(1):8–14

Wienholz BL, Kareta MS, Moarefi AH, Gordon CA, Ginno PA, Chedin F (2010) DNMT3L modulates significant and distinct flanking sequence preference for DNA methylation by DNMT3A and DNMT3B in vivo. PLoS Genet 6(9):e1001106

Williamson CM, Beechey CV, Ball ST, Dutton ER, Cattanach BM, Tease C, Ishino F, Peters J (1998) Localisation of the imprinted gene neuronatin, Nnat, confirms and refines the location of a second imprinting region on mouse chromosome 2. Cytogenet Cell Genet 81(1):73–78

Yuen RK, Jiang R, Penaherrera MS, McFadden DE, Robinson WP (2011) Genome-wide mapping of imprinted differentially methylated regions by DNA methylation profiling of human placentas from triploidies. Epigenetics Chromatin 4(1):10

Zhang H, Niu B, Hu JF, Ge S, Wang H, Li T, Ling J, Steelman BN, Qian G, Hoffman AR (2011) Interruption of intrachromosomal looping by CCCTC binding factor decoy proteins abrogates genomic imprinting of human insulin-like growth factor II. J Cell Biol 193(3):475–487

Part VIII
Transcription and RNA Metabolism

Transcription Factories

Christopher Eskiw and Jenifer Mitchell

Abstract Transcription has classically been described as soluble complexes of polymerase and accessory molecules 'scanning' the genome, probing chromatin for accessible promoters in which to bind and initiate RNA synthesis. A number of studies have demonstrated that transcription occurs at focal accumulations of polymerase called transcription factories. Although the concept of a "factory" for transcription remains controversial, these factories may not only coordinate transcription of multiple genes but recent studies also implicate transcription factories in driving genome organization. Here we will discuss evidence for and against the "Factory" model, and we will identify some of the questions that still need to be addressed.

Keywords Transcription factory • Transcription • Transcription factor • Genome organization • Gene regulation

1 Introduction

Transcription is an intricate ensemble of functions, both spatially and temporally coordinating a diverse set of proteins to produce a single transcript. The proteins required facilitate the selection of the appropriate genes to transcribe, initiate the polymerization reaction and elongate the growing transcript as well as terminate the process. Furthermore, the growing transcript is processed by capping and splicing while still engaged with the elongating polymerase complex. Classic text book portrayals of transcription have RNA polymerase (RNAP) 'scanning' the genome (million to billions of base pairs of genetic material depending on the organism),

C. Eskiw (✉)
Department of Food and Bioproduct Sciences, College of Agriculture and Bioresources, University of Saskatchewan, Room 3E20, Agriculture Building, 51 Campus Drive, Saskatoon, SK S7N 5A8, Canada
e-mail: c.eskiw@usask.ca

J. Mitchell (✉)
Department of Cell and Systems Biology, University of Toronto, 25 Harbord Street, Rm 533a, Toronto, ON M5S 3G5, Canada
e-mail: ja.mitchell@utoronto.ca

D.P. Bazett-Jones, G. Dellaire (eds.), *The Functional Nucleus*,
DOI 10.1007/978-3-319-38882-3_20

testing sequences and binding 'open' available promoters in which to then initiate transcription. It is hard to image how, with such a large catalogue of proteins associated with this process as well as the large amount of genetic material that acts as a template, transcription occurs with components randomly diffusing throughout the nuclear volume.

It is obvious that the nucleus is highly compartmentalized with specific nuclear process occurring within domains such as Promyelocytic Leukemia (PML) bodies [reviewed in (Bernardi and Pandolfi 2014)], Cajal bodies [reviewed in (Sleeman and Trinkle-Mulcahy 2014)] and splicing speckles [reviewed in (Sleeman and Trinkle-Mulcahy 2014)]. Compartmentalization acts to cluster proteins and molecules with common function together to facilitate efficiency. A growing pool of data indicates that transcription is also compartmentalized in structures called transcription factories, allowing for transcription to occur within a specialized structure containing not only the proteins required for all stages of transcription and processing but also able to recruit genes and play a role in genome organization.

1.1 Origins of the "Factory" Model

The idea that nucleic acid synthesis could occur in small specialized compartments originated in the 1970's with the observation that DNA polymerases clustered within the nuclear volume and are attached to an extraction-resistant nuclear sub-architecture termed the nuclear matrix (Hozak et al. 1993, 1994; Berezney and Coffey 1975). Although controversial at the time, an extensive collection of biochemical and imaging data supports the hypothesis that insoluble, DNA polymerase containing, replication factories indeed exist.

It was not until the 1990's that similar observations were made for RNA polymerase (RNAP) II/III indicating that RNA synthesis could also occur within factories in eukaryotic cells. Jackson and colleagues observed that newly synthesized transcripts clustered within HeLa cell nuclei (Jackson et al. 1988, 1998; Wansink et al. 1993). This observation was made by first embedding live cells in agarose microbeads, permeablizing with a physiological buffer, and replacing the endogenous pool of nucleotides with a mixture of ribonucleotides (NTPs) containing brominated uridine triphosphate (BrUTP). In this system the physiological buffer maintains nuclear structure and BrUTP is incorporated [as brominated uridine (BrU)] into the elongating transcript. Subsequent immuno-labelling of BrU and visualization by light microscopy indicated the presence of much fewer foci of labeled transcript sites than what would be expected if all actively transcribing RNAPs were acting independently of one another within the nuclear volume (Jackson et al. 1998). Visualization of nascent transcripts labelled with either Biotin-CTP or Br-UTP using electron microscopy further supported this hypothesis and indicated that transcripts were localized into clusters ~71 nm in diameter (Iborra et al. 1996). Furthermore removal of ~90 % of the chromatin from cells using restriction enzyme digests demonstrated that transcription sites were still

present within the remaining nuclear material, similar to previously characterized DNA polymerase containing replication factories. Transcription factories therefore appear to be part of the insoluble nuclear sub-architecture, producing multiple transcripts each of which is associated with transcribing complexes as well as accessory molecules. Estimates comparing the number of transcripts being produced and the number of active polymerizing complexes to the number of factories present indicated there could be 7–14 transcription units/active polymerase complex at any factory (Jackson et al. 1998). Furthermore, since these complexes are part of an insoluble nuclear sub architecture it has been hypothesized that the DNA template needs to be recruited to a factory rather than polymerases seeking out the template.

The conventional hypothesis of transcription states that lone polymerases bind templates and track long its length. This tracking would create a problem; the extruded RNA molecules could become entangled with the template as the polymerase gyrates around the helix (Cook 2010). No evidence from biology indicates a problem with entanglement throughout evolution nor is there evidence for the existence of nuclear machinery/mechanisms to resolve this issue. However, there is no problem of entanglement if one considers the polymerase to be anchored with the template moving through the enzyme complex (Cook 2010). Although this would generate a torsional strain resulting in DNA supercoils (which are presumably alleviated by topoisomerases), the extruded transcript would not become entwined with the template. Thus a model where polymerizing RNAP complexes are anchored within factories resolves the problem of template/transcript entanglement.

2 Nucleolus: A Paradigm for the Transcription Factory Model

The nucleolus is the most visually distinct and prominent structure within the nucleus and is a paradigm for the transcription factory model. The nucleolus is the specialized site of transcription for ribosomal RNA genes and is comprised of three primary regions: (1) fibrilar component (FC) where RNA polymerase I (RNAPI) is concentrated, (2) dense fibrilar centers (DFC) containing nascent transcripts and, (3) granular component (GC) representing the soluble processed rRNA and ribosomal subunits [reviewed in (Sato et al. 2005)]. The ribosomal genes are organized in a head-to-tail fashion, and in humans, are located at the ends of the acrocentric chromosomes 13, 14, 15, 21 and 22. These regions of the chromosomes are known as nucleolar organizing regions (NORs) and are bound by the RNAPI transcription factor Upstream Binding Factor (UBF) (Dousset et al. 2000). NORs not bound by UBF are not recruited to newly forming nucleoli after mitosis, indicating a potential role for transcription factor binding in the recruitment of genes to sites of transcription. Once the nucleolus is assembled, UBF-associated

ribosomal DNA (rDNA) is pulled towards the surface of the FC where the ribosomal genes are transcribed to produce nascent transcripts. Electron micrographs of isolated transcribing ribosomal genes demonstrate the classic 'Christmas tree' indicating that many polymerases are bound to each gene to produce multiple transcripts concurrently (Miller and Beatty 1969; Koberna et al. 2002). *In vivo*, this region is the DFC and represents the site of nascent RNA synthesis and co-transcriptional splicing of the transcripts into their separate components. Regions of active transcription are further visible upon BrU incorporation followed by immuno-labeling and fluorescent microscopy. This clustering of ribosomal genes for active transcription demonstrates the central principles of the transcription factory model; polymerases and other transcriptional machinery recruit genes to their location, clustering requires the presence of a transcription factor with the transcription of specific subsets of genes occurring within a small portion of the nuclear volume.

3 Visualization of Transcription Factories

Visualization of transcription factories presents unique challenges; there are numerous foci within the nuclear volume with dimensions below the lower resolution limit for light microscopy (~250 nm in the XY plane). Initially, transcription factories were identified by the incorporation of BrUTP or other nucleotide analogues to localize sites of nascent transcript production, however this only provides the location of factories without information on the proteins or genes within each factory (Iborra et al. 1996). There is a large excess of polymerase components that are not engaged in transcript production that continually diffuse throughout the nucleoplasm; as a consequence, immuno-labeling of general polymerase components fails to reveal focal accumulations. Alternatively, antibody labeling against active forms of RNAPII, such as those against phosphorylated serine 5 of the carboxy terminal domain (CTD) repeat of the large subunit of RNAPII (which indicates the initiating form of the complex) labels foci distributed throughout the nucleus. These foci of active RNAPII colocalize with a subset of nascent transcript foci in the nucleoplasm, the number of which is fewer than the estimates of the total number of active genes being transcribed (Pombo et al. 1999). In whole cells it is difficult to visualize individual factories; however, embedding cells labeled for initiating RNAPII followed by ultra-thin (epoxy resin: 80–100 nm) (Fig. 1a) (Eskiw and Fraser 2011) or cryo (vitreous ice: 120–140 nm) (Fig. 1b) (Ferrai et al. 2010) sectioning and subsequent imaging by fluorescence microscopy reveals a clear focal pattern. Very few foci are seen in heterochromatic regions of the nucleus; a logical finding considering that factories are responsible for transcribing active genes located outside condensed inactive heterochromatin (Eskiw and Fraser 2011; Pombo et al. 2000). Imaging physical sections of cells dramatically improves the possible Z-resolution using conventional light microscopy techniques (typically

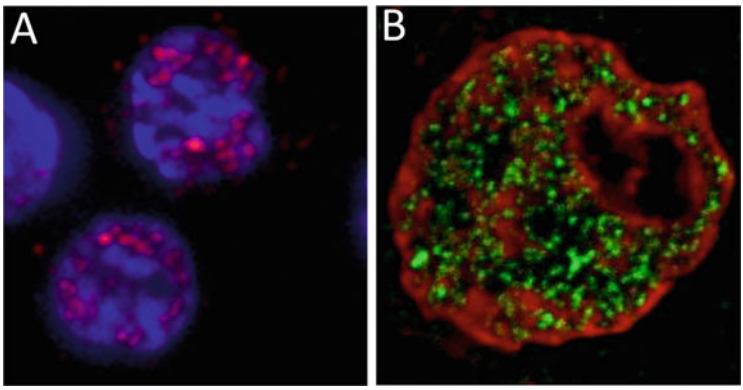

Fig. 1 Transcription factory visualization in thin sections. *Panel* (**a**) Day 14.5 mouse erythroblasts were indirectly labeled for phospho-serine 5 RPB1 (*red*), embedded in epoxy resin and sectioned to ~100 nm. Sections were collected and counterstained with Hoesch 33342 to mark chromatin (*blue*) [adapted from (Eskiw and Fraser 2011)]. *Panel* (**b**) HepG2 cells were indirectly labeled for phospho-serine 5 RPB1 (*green*) and cryosectioned (~140 nm). Chromatin was counterstained with TOTO-3 (*red*) [adapted from (Ferrai et al. 2010)]. Both methods clearly demonstrate the focal accumulations of initiating RNA polymerase II enzymes within euchromatic regions of the nucleus

700 nm or greater) and removes background fluorescence; however, transcription factories are still too small for definitive analysis by fluorescent microscopy.

Electron microscopy (EM) provides the capacity to resolve subcellular domains that is far greater than that of light microscopy, and therefore is a valuable tool for studying the structural organization of transcription factories. Initial EM analysis of nascent transcripts immune-labeled with gold particles revealed clusters of approximately ~71 nm (Chakalova and Fraser 2008). However, one should note that this distribution of gold particles marks the spatial distribution of the transcripts and not the RNAPII distribution. Therefore, these analyses may not reflect the size of the protein compartment. Conventional bright field EM employing sections stained with heavy metal such lead or uranium can only identify electron-dense RNP particles on a grey background, and ultimately fails to reveal structural information on factory organization. The use of correlative microscopy in combination with electron spectroscopic imaging/energy filtering electron microscopy (ESI/EFTEM) has been extremely useful for studying structure and organization of nuclear compartments (Bazett-Jones and Hendzel 1999; Bazett-Jones et al. 1999, 2008). ESI/EFTEM takes advantage of inelastic collisions of electrons with the specimen, with electrons losing specific energies corresponding to the atoms they interacted with. These electrons are collected and 'mapped' to reveal phosphorus rich (chromatin) and nitrogen rich (protein) structures. This technique also provides quantitative information on elemental composition allowing for the mass estimation of structures (Bazett-Jones and Hendzel 1999; Eskiw and Fraser 2011; Eskiw et al. 2008). Analysis of transcription factories in both HeLa cells and developing mouse embryonic erythroblasts (day e14.5) demonstrates that these transcription

factories are protein based with no nucleic acid within the core of the structure. Chromatin fibers make contact with the surface of the factories. However, due to the with the lack of nucleic acid (phosphorus signal) within the factory core, it is likely that transcripts are extruded towards the surface of the factory, similar to ribosomal RNA accumulation in DFCs within nucleoli. In HeLa cells, transcription factories are roughly spherical shaped structures with a mean diameter of 87 nm and an average mass of 10 MDa (Eskiw et al. 2008) while in developing mouse erythroblasts factories measured an average 130 nm in diameter and 26 MDa of mass (Eskiw and Fraser 2011). Both of these measurements demonstrate a Gaussian distribution of both diameter and mass from each cell type. Furthermore, the largest factories in mouse erythroblasts are associated with KLF1, one of the major transcription factors responsible for erythroid specific gene expression (Eskiw and Fraser 2011). These data support the hypothesis that not all transcription factories are equal, and demonstrate factory specialization, where larger factories associate with heavier transcription load than their counterparts.

3.1 Transcription Factories in Live Cells

Given the challenges of imaging transcription factories, it has been difficult to characterize these structures or even demonstrate their existence in live cells. In an attempt to address this issue, Kimura and colleagues developed a system for tracking two populations of RNAPII complexes, by employing a temperature resistant green fluorescent protein (GFP) tagged RPB1 (the major catalytic subunit of the RNAPII complex) in Chinese hamster ovary (CHO) cells containing a temperature sensitive RPB1 (Kimura et al. 2002). Using this system to perform fluorescence recovery after photo-bleaching (FRAP) experiments, they found that ~75 % of the GFP-RPB1 population was rapidly diffusing while the other ~25 % has slow recovery rates ($T_{1/2} = 20$ min). These data indicated that two distinct pools of RNAPII enzyme exist in mammalian cells, and the slow moving population was hypothesized to represent the engaged polymerase, presumably within transcription factories (Kimura et al. 2002). Treatment of cells with Actinomycin D, which inhibits transcription elongation but not initiation, increased the proportion of slow moving polymerase, possibly with more polymerase initiating and stalling on templates. Although this data did demonstrate two distinct populations of polymerase and correlated well with the population of polymerase molecules associated with transcription factories, it did not directly demonstrate that the immobile population is clustered in factories.

Super resolution light microscopy techniques, such as PALM (photoactivated localization microscopy) and STORM (stochastic optical reconstruction microscopy) provide localization information with increased spatial resolution (10–20 nm in the XY plane) to that available by conventional light microscopy techniques. Super resolution microscopy would therefore appear to be a tool ideally suited to study the localization of RNA polymerase molecules/complexes in nuclei. For

example, foci of mCherry-labeled CDK9 (an enzyme that phosphorylates RPB1 on serine 5 and 7 of the CTD to stimulate transcription initiation) in mouse primary cells overlap with foci containing RNAPII-pS5, but not SC35 domains/nuclear speckles, and therefore mark transcription factories (Ghamari et al. 2013). Live cell imaging of these foci demonstrated that although there was some local movement of the factories, possibly through changes in cell shape or constrained diffusion (Chubb and Bickmore 2002), transcription factories are stable entities over time. Furthermore inhibition of transcription using various chemical inhibitors such as 5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole (DRB) demonstrated that foci of CDK9-mCherry and overlapping transcription factories were stable in the absence of transcription. This observation mirrors finding that transcription factories remain present throughout the nucleoplasm during inhibition of transcription following heat shock (Mitchell and Fraser 2008). FRAP experiments for CDK9-mCherry demonstrate that although the foci themselves are stable over time there is a rapid exchange of protein from the nucleoplasmic pool; with 50 % of the fluorescence recovering within 2 s of the bleaching event. This exchange of protein between transcription factories and the highly dynamic soluble pool is similar to what is seen with other nuclear bodies such as PML bodies (Dellaire et al. 2006). These data further demonstrate that factories are likely composed of not only RNAP complexes but also accessory molecules required polymerase function.

3.2 Transcription Factory Distribution Within the Nuclear Volume

As described above, imaging studies have demonstrated that sties of nascent transcript production and initiating/elongating RNAPII can form discrete foci within the nuclear volume. RNAPI factories are sequestered to the nucleolus but what about RNAPII and RNAPIII factories? Comparisons between studies indicate that the number of transcription sites can vary dramatically from cell type to cell type as well as between species. For example, the number of activated RNAPII molecules and transcription factories is much higher in red-spotted salamander cells, which contain ~11 fold larger genome, than in mouse F9 and embryonic stem (ES) cells, (Faro-Trindade and Cook 2006). However, estimates indicate that the density of factories throughout the nucleoplasm is similar. Consistent with this, as mouse ES cells differentiate, there is a decrease in the number of factories, while the density of transcription sites remains relatively constant (Faro-Trindade and Cook 2006). An examination of developing mouse erythroblasts (e14.5) demonstrated that the density transcription factories per unit volume was relatively consistent. However, it is clear from imaging factories in ultra-thin sections that the factories are restricted to regions of euchromatin and excluded from densely packed heterochromatin (Eskiw and Fraser 2011). These observations indicate that the amount of DNA and the nuclear volume influence the number of transcription

factories, and that this is conserved throughout evolution (Faro-Trindade and Cook 2006).

4 Transcription Factory Specialization

Through detailed examination of structure and function, it is clear that the majority of nuclear processes are compartmentalized and thus specialized. Discrete nuclear structures have been described to be associated with specific genes. For example, PML bodies associate with the major histocompatibility locus (Wang et al. 2004) while OPT (Oct1/PTF/transcription) domains are associated with a specific region of Chromosome 6 (Pombo et al. 1998) and Cajal bodies associate with histone genes as well as the U1, U2 and U4 gene clusters [reviewed in (Nizami et al. 2010)]. Although the functional consequence of gene associations with nuclear bodies is unclear, it is likely that these associations facilitate transcriptional regulation. Transcription is compartmentalized into discrete structures, and subsets of factories appear to further specialize to facilitate the transcription of different sub-classes of genes. The most obvious and striking piece of evidence demonstrating transcription factory specialization is the nucleolus which contains only RNAPI and specifically recruits ribosomal genes for transcription within nucleoli. This specialization requires specific transcription factors, such as UBF, as well as processing machinery to generate ribosomal components.

Transcription factory specialization can also occur at nucleoplasmic transcription factories. Firstly, RNAP II and RNAP III are not found together in the same factory. Inhibition of transcription with the adenosine analogue 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB) prevents the elongation of RNAPII transcripts without impacting RNAPIII (Pombo et al. 1999). If factories shared RNAPII and RNAPIII the number of sites would remain constant, or at least relatively high, with the amount of transcript being generated from each factory being decreased. EM observations indicate that the total number of actively transcribing sites dramatically decreases, to one-fifth of the total number of sites, while those that remain still appear to have a high transcriptional output. This decrease mirrors the ratio of predicted RNAPII to RNAPIII factories and demonstrates that RNAPII and RNAPIII factories can specialize to meet the transcriptional requirements of each class of genes (Pombo et al. 1999).

Beyond specialized factories for RNAPI, II and III transcribed genes, transcription factories can also be specialized for transcription of specific types of RNAPII genes. For example, when two types of ectopic vectors, each carrying different RNAPII promoters, are transfected into COS7 African green monkey cells, specific factories will only accumulate the DNA from one type of vector and not the other. This also holds true if one vector contains an intron and the other does not, indicating that only specific factories are component to perform co-transcriptional splicing. Thus, only specific types of genes will be recruited to a given factory, either via promoter specific factors or as a result of alternative splicing (Xu and

Cook 2008) with intron-containing genes localizing to "splicing competent" factories; data which is supported by the RNAPII mediated association of intron-less genes *in vivo* (Li et al. 2012). Further evidence of specialization comes from the developing e14.5 mouse erythroblasts, where out of the several hundred transcription factories present in these erythroid precursor cells only ~60 factories contain the transcription factor Krupple-like factor-1 (KLF1). KLF1 is one of the major transcription factors required for erythroid specific development and many of the genes required for globin synthesis are regulated by this factor. There are >1300 KLF1 dependent genes in the mouse genome. Although not all genes will be transcribed at the same time, the numbers strongly indicate that KLF1-dependent genes must share the limited number of the 60 KLF1-containing factories, with between 15 and 25 genes sharing each site (Eskiw and Fraser 2011). However, it is unclear whether the transcription factors associate with the factory structure and then recruit genes, or if transcription factors first bind DNA elements and are then recruited to a nearby factory capable of supporting the transcription of those genes. Regardless of the model, increasing the local concentration of a specific transcription factor (which, like KLF1, can be the limiting resource for efficient transcriptional initiation) at a select number of sites would facilitate efficient transcription while also increasing the probability of genes regulated by similar mechanisms being with the same local environment (Fig. 2). Other transcription factors, such as TFIID and E2F-1, appear to lack localization at specific transcription factories. This apparent lack of specific localization may be due to the high amount of these factors within the nucleoplasm, the majority of which is unengaged and freely diffusing and therefore are behaving much like RNAP itself (Li et al. 2012; Chang and Hughes-Fulford 2009; Grande et al. 1997). It may very well be that only a subset of factories specialize as a result of a meaningful interaction with these factors.

Data from several studies has demonstrated that stimuli applied to cells results in transcriptional activation of a specific sub-classes of genes and that these genes co-associate (have a higher probability of interacting within the nuclear volume above that of genes from different pathways) at specialized transcription factories. Induction of human umbilical vein epithelial cells (HUVECs) with tumor necrosis factor α (TNF α) results in the phosphorylation and subsequent nuclear localization of the transcription factor NFkB. Prior to stimulation, genes such as *SAMD4A* are not bound by polymerase and are not found to associate with other NFkB regulated genes (Papantonis and Cook 2010; Papantonis et al. 2012). Within 10 min of TNF α induction, *SAMD4A* is bound by RNAPII and associates with other NFkB regulated genes. RNA FISH analysis in combination with super resolution microscopy demonstrated that these transcribing genes are localized to a structure ~90 nm in diameter (Papantonis and Cook 2010). Parallel observations have been made for MCF7 cells induced with estrogen, demonstrating that estrogen receptor α (ERα) genes associated after induction at specialized ERα factories (Li et al. 2012). These data reveal the existence of specialized transcription factories transcribing specific classes of genes regulated by the binding a specific transcription factor or set of shared factors. This specialization allows for the rapid and efficient transcription of

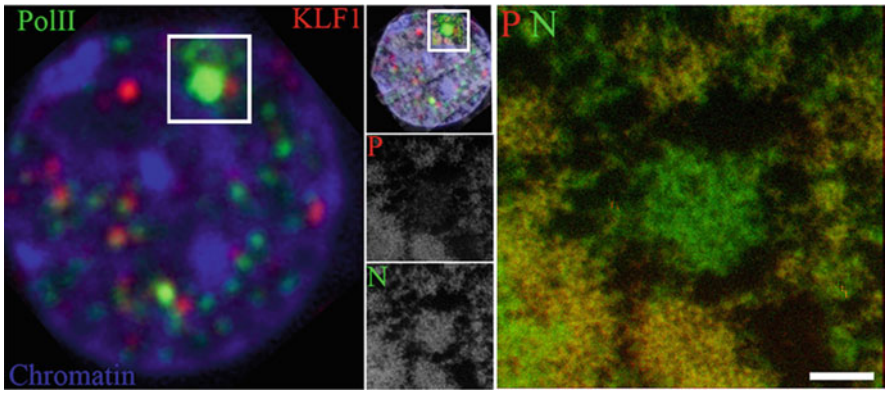

Fig. 2 KLF1 associated factories are larger than other erythroblast factories. Images (*left panel*) of an a day 14 mouse erythroblasts in an epoxy embedded thin section indirectly labeled for phospho-serine 5 RPB1 (PolII, *Green*) and KLF1 (*red*). Chromatin is counterstained with Hoescht 33342 (chromatin, *blue*). *Inset* indicates a large focus of polymerase colocalizing with KILF1. A low magnification ESI/EFTEM was overlayed with the fluorescence image to identify the inset factory (*upper center panel*) to identify the factory of interest for further imaging. Phosphorus (*middle center panel*) and nitrogen (*lower center panel*) maps indicating the location of each element were obtained and false colored *red* and *green* respectively and merged (*right panel*). Structures *yellow* in color contain both nucleic acids (phosphorus) and protein (nitrogen). The factory enriched for KLF1 is *green* indicating a protein based structure that makes contact with chromatin. This factory has a diameter of 194 nm and an estimated mass of ~38 MDa against the population average of 130 nm and a mass of 26 MDa. *Scale bar* = 200 nm

genes by recruiting all the prerequisite machinery for synthesis and processing of transcripts to discrete regions of the nucleus.

5 Transcription Factory Composition

What is a transcription factory composed of? Is there an underlying component that seeds the formation of a factory at a specific location? Since transcription initiation, elongation termination and co-transcriptional splicing is actively occurring we could assume that all the requisite components are localized at factories. Although these components must be at factories it is unclear if they are structural components required for the structure to maintained or if they are only transiently associated and not required for maintaining the structure. Live cell imaging and heat shock studies demonstrate that the factories themselves are metastable, however all components tested thus far show dynamic turnover rates with the surrounding nucleoplasm. nucleoplasm. Although ESI/EFTEM demonstrated the dimensions of transcription factories with extremely high resolution, it cannot elucidate which specific components form the underlying structure. For example, ESI/EFTEM identified that factories are primarily protein based structures (Fig. 2) but it is unclear how

much of this protein was a specific component, such as transcription factors, processing machinery or RNAP complexes themselves.

One method to address the composition of factories is to perform biochemical purification of the structures and use proteomics approaches to identify associated proteins. This is problematic for the study of factories as they are large mega-Dalton insoluble complexes attached to a nuclear sub-architecture. Using caspases (proteases normally activated in apoptotic cells) Melnik and colleagues partially digest the nuclear sub-architecture allowing the isolation transcription factories for proteomic analysis (Melnik et al. 2011). Three major complexes were identified each one containing RNAPI, II or III. Each one of the complexes had masses >8 MDa, agreeing with earlier mass estimates provided through imaging techniques (Eskiw and Fraser 2011; Eskiw et al. 2008) and all three complexes share common proteins including helicases and processing/splicing factors. However each class of complex also contains unique subsets of proteins; complex I represented the nucleolar factories and exhibited markers of the nucleolar proteome while class II complexes had proteins only associated with RNAPII activity. Interestingly, chromatin modifiers were identified to be associated with RNAPII factories indicating that histone modifications may be added to chromatin via the association of genes with specific transcription factories. Although these complexes underwent a battery of treatments during isolation, nascent RNAs remained associated. Examination of these RNAs further demonstrated factory specialization with RNAPI factories having significantly more ribosomal RNAs associated (Melnik et al. 2011). Isolated RNAPII factories also demonstrated a significant enrichment for protein coding genes. More impressively, specific genes, such as those responding to TNFα, which have been shown to interact *in vivo* by chromosome conformation capture (3C), were still associated in isolated complexes. These findings strongly indicate that factory composition is not only multiple polymerizing complexes, but also the associated genes elements and chromatin modifying proteins as well as processing components (Melnik et al. 2011).

Interestingly, all three complexes contained members of the spectrin family of actin binding proteins. Actin is present within nuclei however the role this essential structural protein plays in nuclear processes is still very much under debate. Actin filaments, such as those seen in the cytoplasm are not apparent in the nucleus however FRAP and fluorescence correlation spectroscopy (FCS) reveals that ~20 % of the nuclear actin pool behaves as polymers (McDonald et al. 2006). Actin and actin-associated proteins such as nuclear myosin 1 have been found to be essential for transcription and have further been found to associate with a diverse set of complexes including chromatin remodelers. Chemical inhibition that blocks actin polymerization or nuclear myosin function can inhibit transcription, further demonstrating the importance of these proteins in nuclear function (Ye et al. 2008). The presence of actin binding proteins in transcription factories indicates that there may be a significant role for nuclear actin in recruiting genes and transcription factors to the factory. Further experimentation will need to be performed to address the role of actin and actin binding proteins in transcription factory function and/or possibly formation.

6 Transcription Factories and Genome Organization

As alluded to above transcription and genome organization appear to be intimately linked. On a large scale, entire chromosomes are differentially positioned in the nucleus based on their gene density and state of cellular differentiation. The more gene dense chromosomes are positioned towards the nuclear interior whereas the gene poor chromosomes are more often found at the nuclear periphery (Croft et al. 1999; Boyle et al. 2001; Bolzer et al. 2005). In addition, translocations between differentially positioned chromosomes cause changes in position of the fused chromosome segments and widespread changes in gene expression, some of which may be due to altered nuclear position. Gene density and the underlying sequence is not the only modifier of chromosome position. For example, the mammalian inactive X chromosome in female nuclei occupies a distinct RNAPII free compartment, while the transcriptionally active X chromosome remains associated with RNAPII factories (Chaumeil et al. 2006). Furthermore, entire chromosomes can rapidly change their position in the interphase nucleus within 15 min of serum withdrawal demonstrating the dynamic properties of genome organization (Mehta et al. 2010). Individual genes can also alter their relative position in the nucleus upon receiving a transcriptional stimulus. This was first observed at the MHC locus, which is more often found outside of its chromosome territory 10 min after treatment with interferon-gamma (Volpi et al. 2000). Similarly, induction of *Myc* in B lymphocytes occurs within five minutes, concurrently with a relocation of *Myc* alleles to RNAPII factories occupied by the highly transcribed immunoglobulin heavy chain (*Igh*) gene located on a separate chromosome (Osborne et al. 2007). This relocation of genes has been observed on a larger scale in primary human endothelial cells stimulated with TNFα (Papantonis et al. 2012) in which TNFα responsive genes congregate at specialized transcription factories within 30 min. The rapid movement of chromosome territories during interphase does appear to require both the polymerization of nuclear actin and functional nuclear myosin, two proteins also required for transcription both *in vitro* and *in vivo* (Mehta et al. 2010; Dundr et al. 2007; Chuang et al. 2006). Together these studies highlight the role of transcription factories in driving changes in genome organization as well as revealing the dynamic nature of chromatin in the interphase nucleus.

Perhaps the most striking evidence that transcription factories and transcription influences genome organization comes from genome-wide studies of chromatin contacts. These studies use techniques based on the chromosome conformation capture (3C) technique, which monitors relative cross-linking frequency between chromatin regions in the nucleus (Dekker et al. 2002). Briefly, formaldehyde cross-linking fixes regions of the genome that interact *in vivo*. A restriction enzyme then cleaves DNA into soluble chromatin complexes. These are then ligated under dilute conditions to ensure that ligation events occur between DNA contained in the same large protein/DNA complex. After cross-link reversal and DNA purification, ligation events are identified by PCR using primers directed against different regions of the genome. The 3C technique, originally used to determine yeast chromosome

conformation, was adapted to identify chromatin loops formed between the β-globin gene, distal enhancers of the locus control region, and between genes frequently co-associated with β-globin in the same transcription factory (Tolhuis et al. 2002; Palstra et al. 2003; Osborne et al. 2004; Schoenfelder et al. 2010). The combination of massively parallel sequencing with 3C resulted in the development of Hi-C, which can be used to characterize genome-wide chromatin-chromatin interactions within a population of cells (Lieberman-Aiden et al. 2009). In Hi-C a biotinylated nucleotide is incorporated into the digested ends prior to the ligation step common to all 3C-based techniques. After ligation, sonication, and purification of biotinylated fragments, high-throughput sequencing allows detection of chromatin-chromatin interactions genome-wide. The resolution of Hi-C is limited by the number of ligation fragments sequenced and initially, interactions in the human genome were analyzed by Hi-C at a resolution of 1 Mb revealing interactions in both *cis* and *trans* genome-wide (Lieberman-Aiden et al. 2009). This analysis confirmed the proximity of gene rich chromosomes as observed in individual cells by DNA FISH. Furthermore, comparing two cell lines from the same species with different transcriptional programs shows a high correlation between the overall compartmentalization of the genome; however, specific loci are differentially compartmentalized in the two cell types (Lieberman-Aiden et al. 2009). Together, these data indicate a genome folding pattern common across cell types within a species, with specific local domains organized based on cell type transcriptional profiles. This differentiation in folding between cell types is consistent with the transcription factory model, where the specific cohort of transcription factors expressed in a cell may organize the genome via the formation of specialized transcription factories that in turn coordinate the transcription of gene loci in a particular cell type.

7 The Controversy Surrounding the "Factory" Model

Although there is a large pool of data in support of the "Factory" model, controversy still surrounds the hypothesis that RNA polymerases cluster into protein based-transcription factories. For example, some *in vitro* studies call in to question the requirement of factories. Biologists often take advantage of simple or reductionist systems to understand how complex biological processes work. One such system is the use of *in vitro* transcription to generate RNA molecules. It is quite apparent transcription factories are not required to generate these molecules *in vitro*. Examination of one of the most well characterized polymerase, the T7 bacteriophage RNAP, reveals this system does not require transcription factory formation and is capable of functioning to generate transcripts as monomers either in solution or when anchored to a surface substrate. Furthermore, examination of two actively transcribing DNA templates, one bound to beads via biotin-streptavidin and the other free in solution, showed no interaction, indicating that elongating complexes of T7 RNAP do not significantly interact *in vitro* (Finan

et al. 2012; Bjedov et al. 2010). Of note, T7 RNAP is a very simple enzyme compared to other polymerase (e.g., mammalian or plant) and does not require the assembly of a large protein complex, and the T7 RNAP is even much less complex than the 'simple' polymerizing complexes found in the host bacterium. Furthermore in these simple systems, where pure enzymes and components are used, it is not likely that all aspects of the *in vivo* transcriptional process will be recapitulated, for example the crowding effects of chromatin on the nuclear distribution and diffusion of polymerases and their gene templates in eukaryotic cells.

Another argument often used to challenge the factory model involves the work of Miller and colleagues (Scheer 1987a). Miller developed a method of spreading transcribing ribosomal genes from water beetle oocytes *in vitro* to form an iconic 'Christmas tree' pattern of multiple nascent transcripts being generated from a single template (Scheer 1987a). At face value, these images seem to contradict the factory model (Fig. 3). However, it must be noted that these images are the result of purification and conditions which were designed to disrupt chromatin and protein complexes in hypotonic media in way that no longer preserves their three dimensional organization as found *in vivo* (Scheer 1987a, b; Miller and Beatty 1969). The appearance of these Miller spreads suggests some polymerase complexes are bound more tightly to the DNA than to the DFC structure. In nucleoli the DFC structure contains many polymerase complexes but once the chromatin is stretched out on a flat surface the transcribing complexes most likely have broken away from the structure to generate this tree-like conformation. Although the base of each branch of the tree appears anchored to the template it is unclear where the

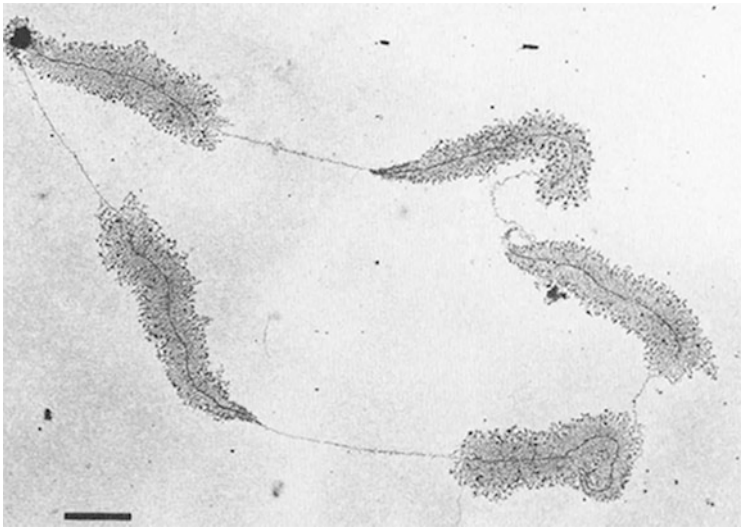

Fig. 3 'Christmas Tree' pattern of a Miller spread. Electron micrograph of nucleolar DNA prepared by Miller spread from an oocyte of the water beetle, *Dytiscus marginalis*. 5 repeating units of ribosomal genes are present with transcripts protruding from the DNA axis *scale bar* = 1 μm modified from Scheer (1987a)

polymerase is localized prior to the spreading. Although ground breaking and informative for the early stages of chromatin research it is unlikely that the *ex vivo* conditions used to generate Miller spreads give a true representation of the organization of transcription *in vivo* from ribosomal genes.

One of the main arguments against the factory model is that there are no transcription factories observable in *Drosophila melanogaster* salivary gland cells containing polytene chromosomes. Polytene chromosomes represent a unique nuclear environment and non-canonical and genome organization, consisting of several hundred copies of the same DNA fibers due to endoreplication of DNA in these cells. Imaging data of live salivary cells expressing GFP-tagged RNAPII indicates that loci of heat shock genes (87A and 87C) do not move after heat shock stimulation, and that heat shock causes the rapid recruitment of RNAPII molecules to this location. There is rapid turnover RNAPII molecules at these sites with FRAP recovery rates within 100 s of the bleaching event indicating a highly dynamic exchange with the nucleoplasmic pool of polymerases (Yao et al. 2007). Based on these data it has been suggested that upon heat shock, a transcription 'compartment' is formed and that dynamic 'local recycling' of RNAPII and other transcription components occurs (Yao et al. 2007). This interpretation has been viewed to be inconsistent with the transcription factory model; however, this could also be interpreted as being very consistent with the factory model as described above. Prior to the heat shock, large clusters of RNAPII molecules can be visualized at these loci which could be a preexisting transcription factory. The existing compartment would facilitate the transcription of a specific subset of like regulated genes (in this case a single gene but with multiple copies), with multiple transcripts being generated from multiple templates within the same structure. Furthermore, the size of the 'compartment' increases following heat shock, presumably to deal with the increased transcriptional load. This could be seen as analogous to the larger KLF1 positive factories observed in developing mouse erythroblasts, which are hypothesized to be larger structures than other factories within the nuclear volume, to accommodate the higher transcriptional load from KLF1 dependent genes (Eskiw and Fraser 2011). These images/movies may have captured the refurbishment of the factory, analogous to turning a small cottage operation into a super factory for the large scale assembly of transcripts.

Two recent studies have used super resolution microscopy techniques to examine the localization of RNAPII within mammalian nuclei. Both studies utilized U2OS osteosarcoma cell lines transfected with an exogenous modified RPB1 subunit and indicate that in both live (Cisse et al. 2013) and fixed (Zhao et al. 2014) cells RNAPII does not form factories within the nuclear volume. Cisse and colleagues observed clustering of polymerase in live cells, however their data indicates that clustering only exist transiently with an average lifetime of 5.1 s and are highly inducible following cell stimulation. This is contradictory to previous studies that indicated polymerase is rapidly exchanging from factories but the factories themselves are metastable (Kimura et al. 2002; Ghamari et al. 2013) and exist even under conditions that inhibit transcriptional initiation (Mitchell and Fraser 2008). Inhibition of transcription elongation in U2OS cells by flavopiridol

promoted the accumulation of RNAPII clustering (Cisse et al. 2013), similar to previous observations in the literature (Mitchell and Fraser 2008; Ghamari et al. 2013); however, despite this observation the authors conclude that factories are not present. Using light sheet super resolution microscopy of fixed nuclei, Zhoa and colleagues were unable to observe clustering of two differentially labeled RNAPII populations, and concluded that clustering of RNAPII does not occur. Despite the authors conclusions from these technically impressive studies that transcription factories do not exist, there are several issues with these studies which may call this conclusion into doubt: (1) there was no distinction between active and inactive polymerase population in these studies despite previous observation that these population have dramatically different mobile properties (Kimura et al. 2002); (2) the estimate of polymerase molecules within the nucleoplasm in these studies is much lower than previous studies (both soluble and insoluble) and is not supported by western blot or other confirmation that their measurements were physiologically relevant in regard to polymerase levels; (3) there was a lack of supporting evidence to ensure that the labeling protocol for polymerases was efficient and accurately represented the total population of RNAPII molecules; and (4) the authors of these studies used different fixation methods (methanol versus formaldehyde) for the positive control and the experimental samples, which are known to have dramatically different impacts on nuclear structure (Guillot et al. 2004). Further experiments using tagged endogenous RNA polymerase subunits and super resolution imaging will allow for a more physiologically relevant examination of the degree of polymerase clustering in live cells. Furthermore, the characterization of the phosphorylation status of the polymerases and their production of nascent transcripts (in both live and fixed cells) would provide the needed context for these observations and would help reconcile these results with previous data regarding the ability of RNAPII to cluster into factories (Ghamari et al. 2013).

8 Summary

Transcription has classically been thought of as soluble complexes of polymerase and accessory molecules being recruited to a gene to facilitate transcription. Experimental approaches that monitor transcription *in situ* have revealed levels of complexity in eukaryotic nuclei beyond the mere assembly of protein complexes at gene promoters. Specialized compartments in the nucleus transcribe co-regulated genes and exclude silenced genes from transcription compartments. These compartments contain numerous proteins directly involved in transcription as well as proteins indirectly involved in transcription, such as chromatin modifiers and RNA processing enzymes. Although here we have concentrated on the evidence for and against transcription factory function under normal growth conditions in eukaryotic cells, there also evidence for the existence and of transcription factories and clustering of polymerase both in bacteria and in response to viral infection. For

example, RNA dependent RNA polymerase from poliovirus exhibits co-operatively with high concentrations of polymerase molecules interacting to reach maximum efficiency of template utilization (Pata et al. 1995). Furthermore, in bacteria, RNA polymerase also appears to cluster within cells. The binding of polymerase to active genes creates 'barriers' or ties preventing DNA loop excision. Extraction of *E. coli*, with high salt and detergent to remove the soluble material and the majority of DNA, resulting in fragments of DNA that are actively being transcribed as well as the majority of the nascent RNA (Jackson et al. 1981). Transcription has also been demonstrated to reorganize supercoiled domains within bacterial genomes, parallel to observation indicating that transcription/transcription factory binding is driving eukaryotic genome organization (Deng et al. 2005). Inhibition of transcription by nutrient starvation, or via treatment with rifamycin, inhibits nucleiod condensation resulting in the dispersal of DNA throughout the cell as well the loss of polymerase clustering in the bacteria (Jin and Cabrera 2006).

A large pool of data now indicates that transcription occurs at focal accumulations of activated RNA polymerases called transcription factories, and that these nuclear domains recruit genes to their surface to be transcribed. Although controversial, the factory model provides a possible structural and conceptual framework from which to investigate the specialized and diverse functions of clustered genes in specific cell types that facilitates both cell type specific genome organization and transcriptional profiles (a summary model of transcription factories presented in Fig. 4). Nonetheless, a number of questions remain to be addressed, including: (1) how transcription factories nucleate following the reformation of the nuclear membrane after mitosis? (2) do transcription factors bind genes before being recruited to factories or do these factors accumulate in factories resulting in the recruitment of specific genes? (3) how are factories specialized for the transcription of a specific subclass of genes? (4) how does the recruitment of specific gene classes to specialized factories alter genome wide chromatin interactions and genome folding? Improvements and adaptations in various techniques and microscopy will no doubt help resolve some of these questions. For example, improved imaging techniques, such as super resolution microscopy and ESI/EFTEM, will provide further structural information on the composition of factories as well as shed light on the structures that make contact with them from the surrounding nucleoplasm. Improved resolution in live cell microscopy will help determine if transcription factories first recruit transcription factors before the genes, or if these factors bind genes and then are recruited to the nearest transcriptionally competent factory. Clues as to the structural basis for the formation of specialized factories may come from further biochemical purification. Adaptations in molecular techniques and sequencing technologies, such as those based on 3C will provide additional insight into the role transcription factories play in sculpting genome organization and folding. Ultimately, the better we are able to observe the process of transcription within cells *in vivo* without perturbation, the better we will understand fully the complexities of transcriptional gene regulation.

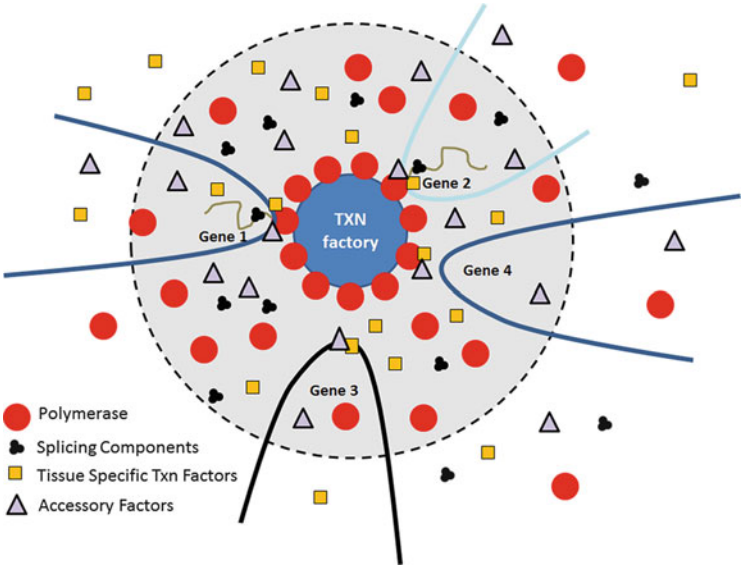

Fig. 4 Transcript factory organization model: RNAPII complexes (*red circles*) cluster into the factory structure (*blue circle* labelled TXN factory) however it is unclear what 'seeds' the formation, whether it be RNAP itself or other structural proteins. RNAP will exchange dynamically with the local environment although the factory remains as a metastable entity. Genes will be recruited to the surface of the factory. In this model, Genes 1 and 2 are actively transcribing interacting with tissue specific transcription factors (Tissue Specific TXN Factor, *orange triangles*) as well as other accessory factors (*grey triangle*) such as chromatin modifying complexes. The factory transcribing these genes is extruding nascent transcripts onto its and these transcripts are then bound by splicing components. It is currently unclear if genes are bound by transcription factors (Gene 3) and then recruited to factories or if factors are bound at transcription factories and then recruit genes (Gene 4). The accumulation of genes and factors at specific transcription factories will build up the local concentration (*grey circle* with *dashed line*) resulting in a higher probability of specific genes and their requisite factors localizing within the same region of the nucleus, causing transcription factory specialization and driving genome organization

References

Bazett-Jones DP, Hendzel MJ (1999) Electron spectroscopic imaging of chromatin. Methods 17: 188–200

Bazett-Jones DP, Hendzel MJ, Kruhlak MJ (1999) Stoichiometric analysis of protein- and nucleic acid-based structures in the cell nucleus. Micron 30:151–157

Bazett-Jones DP, Li R, Fussner E, Nisman R, Dehghani H (2008) Elucidating chromatin and nuclear domain architecture with electron spectroscopic imaging. Chromosome Res 16: 397–412

Berezney R, Coffey DS (1975) Nuclear protein matrix: association with newly synthesized DNA. Science 189:291–293

Bernardi R, Pandolfi PP (2014) A dialog on the first 20 years of PML research and the next 20 ahead. Front Oncol 4:23

Bjedov I, Toivonen JM, Kerr F, Slack C, Jacobson J, Foley A, Partridge L (2010) Mechanisms of life span extension by rapamycin in the fruit fly *Drosophila melanogaster*. Cell Metab 11: 35–46

Bolzer A, Kreth G, Solovei I, Koehler D, Saracoglu K, Fauth C, Muller S, Eils R, Cremer C, Speicher MR, Cremer T (2005) Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. PLoS Biol 3:e157

Boyle S, Gilchrist S, Bridger J, Mahy NL, Ellis JA, Bickmore WA (2001) The spatial organization of human chromosome within the nuclei of normal and emerin-mutant cells. Hum Mol Genet 10(3):211–219

Chakalova L, Fraser P (2008) Brushed aside and silenced. Dev Cell 14:461–462

Chang TT, Hughes-Fulford M (2009) Monolayer and spheroid culture of human liver hepatocellular carcinoma cell line cells demonstrate distinct global gene expression patterns and functional phenotypes. Tissue Eng Part A 15:559–567

Chaumeil J, Le Baccon P, Wutz A, Heard E (2006) A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. Genes Dev 20: 2223–2237

Chuang CH, Carpenter AE, Fuchsova B, Johnson T, de Lanerolle P, Belmont AS (2006) Long-range directional movement of an interphase chromosome site. Curr Biol 16:825–831

Chubb JR, Bickmore W (2002) Chromatin motion is constrained by association with nuclear compartments in human cells. Curr Biol 12:439–445

Cisse II, Izeddin I, Causse SZ, Boudarene L, Senecal A, Muresan L, Dugast-Darzacq C, Hajj B, Dahan M, Darzacq X (2013) Real-time dynamics of RNA polymerase II clustering in live human cells. Science 341:664–667

Cook PR (2010) A model for all genomes: the role of transcription factories. J Mol Biol 395:1–10

Croft JA, Bridger JM, Boyle S, Perry P, Teague P, Bickmore WA (1999) Differences in the localization and morphology of chromosomes in the human nucleus. J Cell Biol 145(6): 1119–1131

Dekker J, Rippe K, Dekker M, Kleckner N (2002) Capturing chromosome conformation. Science 295:1306–1311

Dellaire G, Eskiw CH, Dehghani H, Ching RW, Bazett-Jones DP (2006) Mitotic accumulations of PML protein contribute to the re-establishment of PML nuclear bodies in G1. J Cell Sci 119: 1034–1042

Deng S, Stein RA, Higgins NP (2005) Organization of supercoil domains and their reorganization by transcription. Mol Microbiol 57:1511–1521

Dousset T, Wang C, Verheggen C, Chen D, Hernandez-Verdun D, Huang S (2000) Initiation of nucleolar assembly is independent of RNA polymerase I transcription. Mol Biol Cell 11: 2705–2717

Dundr M, Ospina JK, Sung MH, John S, Upender M, Ried T, Hager GL, Matera AG (2007) Actin-dependent intranuclear repositioning of an active gene locus in vivo. J Cell Biol 179: 1095–1103

Eskiw CH, Fraser P (2011) Ultrastructural study of transcription factories in mouse erythroblasts. J Cell Sci 124:3676–3683

Eskiw CH, Rapp A, Carter DR, Cook PR (2008) RNA polymerase II activity is located on the surface of protein-rich transcription factories. J Cell Sci 121:1999–2007

Faro-Trindade I, Cook PR (2006) A conserved organization of transcription during embryonic stem cell differentiation and in cells with high C value. Mol Biol Cell 17:2910–2920

Ferrai C, Xie SQ, Luraghi P, Munari D, Ramirez F, Branco MR, Pombo A, Crippa MP (2010) Poised transcription factories prime silent uPA gene prior to activation. PLoS Biol 8:e1000270

Finan K, Torella JP, Kapanidis AN, Cook PR (2012) T7 RNA polymerase functions in vitro without clustering. PLoS One 7:e40207

Ghamari A, van de Corput MP, Thongjuea S, van Cappellen WA, van Ijcken W, van Haren J, Soler E, Eick D, Lenhard B, Grosveld FG (2013) In vivo live imaging of RNA polymerase II transcription factories in primary cells. Genes Dev 27:767–777

Grande MA, van der Kraan I, de Jong L, van Driel R (1997) Nuclear distribution of transcription factors in relation to sites of transcription and RNA polymerase II. J Cell Sci 110:1787–1791

Guillot PV, Xie SQ, Hollinshead M, Pombo A (2004) Fixation-induced redistribution of hyperphosphorylated RNA polymerase II in the nucleus of human cells. Exp Cell Res 295: 460–468

Hozak P, Hassan AB, Jackson DA, Cook PR (1993) Visualization of replication factories attached to nucleoskeleton. Cell 73:361–373

Hozak P, Jackson DA, Cook PR (1994) Replication factories and nuclear bodies: the ultrastructural characterization of replication sites during the cell cycle. J Cell Sci 107(Pt 8):2191–2202

Iborra FJ, Pombo A, Jackson DA, Cook PR (1996) Active RNA polymerases are localized within discrete transcription "factories" in human nuclei. J Cell Sci 109(Pt 6):1427–1436

Jackson DA, Mccready SJ, Cook PR (1981) RNA synthesis in a Cage. Science 292:552–556

Jackson DA, Yuan J, Cook PR (1988) A gentle method for preparing cyto- and nucleo-skeletons and associated chromatin. J Cell Sci 90(Pt 3):365–378

Jackson DA, Iborra FJ, Manders EM, Cook PR (1998) Numbers and organization of RNA polymerases, nascent transcripts, and transcription units in HeLa nuclei. Mol Biol Cell 9: 1523–1536

Jin DJ, Cabrera JE (2006) Coupling the distribution of RNA polymerase to global gene regulation and the dynamic structure of the bacterial nucleoid in *Escherichia coli*. J Struct Biol 156: 284–291

Kimura H, Sugaya K, Cook PR (2002) The transcription cycle of RNA polymerase II in living cells. J Cell Biol 159:777–782

Koberna K, Malinsky J, Pliss A, Masata M, Vecerova J, Fialova M, Bednar J, Raska I (2002) Ribosomal genes in focus: new transcripts label the dense fibrillar components and form clusters indicative of "Christmas trees" in situ. J Cell Biol 157:743–748

Li G, Ruan X, Auerbach RK, Sandhu KS, Zheng M, Wang P, Poh HM, Goh Y, Lim J, Zhang J, Sim HS, Peh SQ, Mulawadi FH, Ong CT, Orlov YL, Hong S, Zhang Z, Landt S, Raha D, Euskirchen G, Wei CL, Ge W, Wang H, Davis C, Fisher-Aylor KI, Mortazavi A, Gerstein M, Gingeras T, Wold B, Sun Y, Fullwood MJ, Cheung E, Liu E, Sung WK, Snyder M, Ruan Y (2012) Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. Cell 148:84–98

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326:289–293

Mcdonald D, Carrero G, Andrin C, de Vries G, Hendzel MJ (2006) Nucleoplasmic beta-actin exists in a dynamic equilibrium between low-mobility polymeric species and rapidly diffusing populations. J Cell Biol 172:541–552

Mehta IS, Amira M, Harvey AJ, Bridger JM (2010) Rapid chromosome territory relocation by nuclear motor activity in response to serum removal in primary human fibroblasts. Genome Biol 11:R5

Melnik S, Deng B, Papantonis A, Baboo S, Carr IM, Cook PR (2011) The proteomes of transcription factories containing RNA polymerases I, II or III. Nat Methods 8:963–968

Miller OL Jr, Beatty BR (1969) Visualization of nucleolar genes. Science 164:955–957

Mitchell JA, Fraser P (2008) Transcription factories are nuclear subcompartments that remain in the absence of transcription. Genes Dev 22:20–25

Nizami Z, Deryusheva S, Gall JG (2010) The Cajal body and histone locus body. Cold Spring Harb Perspect Biol 2:a000653

Osborne CS, Chakalova L, Brown KE, Carter D, Horton A, Debrand E, Goyenechea B, Mitchell JA, Lopes S, Reik W, Fraser P (2004) Active genes dynamically colocalize to shared sites of ongoing transcription. Nat Genet 36:1065–1071

Osborne CS, Chakalova L, Mitchell JA, Horton A, Wood AL, Bolland DJ, Corcoran AE, Fraser P (2007) Myc dynamically and preferentially relocates to a transcription factory occupied by Igh. PLoS Biol 5:e192

Palstra RJ, Tolhuis B, Splinter E, Nijmeijer R, Grosveld F, de Laat W (2003) The beta-globin nuclear compartment in development and erythroid differentiation. Nat Genet 35:190–194

Papantonis A, Cook PR (2010) Genome architecture and the role of transcription. Curr Opin Cell Biol 22:271–276

Papantonis A, Kohro T, Baboo S, Larkin JD, Deng B, Short P, Tsutsumi S, Taylor S, Kanki Y, Kobayashi M, Li G, Poh HM, Ruan X, Aburatani H, Ruan Y, Kodama T, Wada Y, Cook PR (2012) TNFalpha signals through specialized factories where responsive coding and miRNA genes are transcribed. EMBO J 31:4404–4414

Pata JD, Schultz SC, Kirkegaard K (1995) Functional oligomerization of poliovirus RNA-dependent RNA polymerase. RNA 1:466–477

Pombo A, Cuello P, Schul W, Yoon JB, Roeder RG, Cook PR, 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 DA, Hollinshead M, Wang Z, Roeder RG, Cook PR (1999) Regional specialization in human nuclei: visualization of discrete sites of transcription by RNA polymerase III. EMBO J 18:2241–2253

Pombo A, Jones E, Iborra FJ, Kimura H, Sugaya K, Cook PR, Jackson DA (2000) Specialized transcription factories within mammalian nuclei. Crit Rev Eukaryot Gene Expr 10:21–29

Sato S, Yano H, Makimoto Y, Kaneta T, Sato Y (2005) Nucleolonema as a fundamental substructure of the nucleolus. J Plant Res 118:71–81

Scheer U (1987a) Contributions of electron microscopic spreading preparations ("Miller spreads") to the analysis of chromosome structure. Results Probl Cell Differ 14:147–171

Scheer U (1987b) Structure of lampbrush chromosome loops during different states of transcriptional activity as visualized in the presence of physiological salt concentrations. Biol Cell 59: 33–42

Schoenfelder S, Sexton T, Chakalova L, Cope NF, Horton A, Andrews S, Kurukuti S, Mitchell JA, Umlauf D, Dimitrova DS, Eskiw CH, Luo Y, Wei CL, Ruan Y, Bieker JJ, Fraser P (2010) Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. Nat Genet 42:53–61

Sleeman JE, Trinkle-Mulcahy L (2014) Nuclear bodies: new insights into assembly/dynamics and disease relevance. Curr Opin Cell Biol 28C:76–83

Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W (2002) Looping and interaction between hypersensitive sites in the active beta-globin locus. Mol Cell 10:1453–1465

Volpi EV, Chevret E, Jones T, Vatcheva R, Williamson J, Beck S, Campbell RD, Goldsworthy M, Powis SH, Ragoussis J, Trowsdale J, Sheer D (2000) Large-scale chromatin organization of the major histocompatibility complex and other regions of human chromosome 6 and its response to interferon in interphase nuclei. J Cell Sci 113(Pt 9):1565–1576

Wang J, Shiels C, Sasieni P, Wu PJ, Islam SA, Freemont PS, Sheer D (2004) Promyelocytic leukemia nuclear bodies associate with transcriptionally active genomic regions. J Cell Biol 164:515–526

Wansink DG, Schul W, van der Kraan I, van Steensel B, van Driel R, 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

Xu M, Cook PR (2008) Similar active genes cluster in specialized transcription factories. J Cell Biol 181:615–623

Yao J, Ardehali MB, Fecko CJ, Webb WW, Lis JT (2007) Intranuclear distribution and local dynamics of RNA polymerase II during transcription activation. Mol Cell 28:978–990

Ye J, Zhao J, Hoffmann-Rohrer U, Grummt I (2008) Nuclear myosin I acts in concert with polymeric actin to drive RNA polymerase I transcription. Genes Dev 22:322–330

Zhao ZW, Roy R, Gebhardt JC, Suter DM, Chapman AR, Xie XS (2014) Spatial organization of RNA polymerase II inside a mammalian cell nucleus revealed by reflected light-sheet superresolution microscopy. Proc Natl Acad Sci USA 111:681–686

Dynamics and Transport of Nuclear RNA

Jonathan Sheinberger and Yaron Shav-Tal

Abstract The messenger RNA (mRNA) molecule passes the genetic information from the genome to the protein synthesis machinery. Decades of study of the spatial characteristics of mRNA distribution in fixed cells and tissues particularly by electron microscopy and *in situ* hybridization approaches, have revealed the sites of synthesis in relation to the nuclear DNA, and the paths taken en route to the nuclear pore. These studies are now complemented by experiments performed in living cells using fluorescent tags that specifically target mRNA transcripts. The use of high-end microscopy equipment improving the detection of mRNA molecules, together with the advent of new fluorescent tags and original means by which to label the mRNAs, allow us to spy on the mRNA within its natural context of the living cell. High-resolution time-lapse imaging has brought to light the dynamics of single molecules of mRNA during RNA polymerase II transcription, nucleoplasmic transport of mRNA-protein complexes (mRNPs), and the final nuclear event of mRNA export through the nuclear pore complex.

1 Introduction

In 1952 James Watson taped a small note on the wall of his room. It was his hypothesis for how protein synthesis takes place in eukaryotic cells. The note said: ↻DNA → RNA → Protein, putting forward the propositions that: (a) RNA is synthesized in the nucleus from the DNA molecule, and (b) that RNA must move into the cytoplasm where it serves as a template for protein synthesis. As Watson describes, his hypothesis was based on two findings (Watson 2011). First, was the experiment that showed that the nucleus (DNA) had no role in protein synthesis. Brachet and Chantrenne used the giant alga Acetabularia for these experiments; after cutting the cells in half they found that the half lacking the nucleus could

J. Sheinberger • Y. Shav-Tal (✉)
The Mina and Everard Goodman Faculty of Life Sciences and Institute of Nanotechnology, Bar-Ilan University, Ramat Gan 52900, Israel
e-mail: Yaron.Shav-Tal@biu.ac.il

D.P. Bazett-Jones, G. Dellaire (eds.), *The Functional Nucleus*,
DOI 10.1007/978-3-319-38882-3_21

sustain protein synthesis for 2 months (Brachet and Chantrenne 1951). The second finding was made by Beadle and Tatum, showing that the genetic information required for protein synthesis is harbored in the cell nucleus, also known as the "one gene–one enzyme hypothesis" (Beadle and Tatum 1941). Only in 1961 did the term messenger RNA (mRNA) emerge following the publication of two back-to-back papers in *Nature*. Meselson, Jacob and Brenner were searching for an RNA associated with the ribosomes that was not ribosomal RNA, and discovered the short-lived messenger RNA (Brenner et al. 1961). Independently around the same time, Watson and colleagues came to the same conclusion (Gros et al. 1961). Although Jacob et al. were ready for publication, Meselson tells that Watson requested that they wait with the submission, and so they did. Thus mRNA was discovered, and the rest is history. In the wide perspective of cell biology, it is safe to say that the mRNA molecule has become on its own right, one of the pillars of modern biology.

The study of mRNA biology has branched out in many directions focusing on the highly regulated processes of nuclear transcription and cytoplasmic translation. This chapter will describe the life cycle of the mRNA molecule in the nucleus, from the time it leaves the transcribing gene and travels toward the nuclear periphery, ending with nucleo-cytoplasmic export through the nuclear pore complex (NPC). We will emphasize the spatial considerations of mRNA dynamics in the nucleus with regard to the temporal information obtained from live-cell studies, and will use the technical developments in microscopy and imaging as stepping stones in describing key discoveries and the progression in our understanding of mRNA dynamics in eukaryotic cells. We will travel from the days where electron microscopy provided the first high resolution glimpses of mRNAs in fixed cells, through the appearance of the fluorescent molecules that lighted up mRNAs in cells, to bring us to the current era of live-cell imaging which provides real-time measurements of mRNA dynamics.

2 The mRNP

Electron microscopy was one of the first commonly used tools to study the appearance, location and even structure of mRNA molecules in the context of the cell. These large complexes, biochemically purified from cells or observed in EM specimens and identified as containing mRNAs, were considered mRNPs, meaning an mRNA molecule coated with a substantial protein component.

It is well recognized by now that mRNA molecules are not simple linear single stranded nucleic acids as depicted in many schemes, and are rather molecules rich in secondary structures that serve as scaffolds for the binding of tens or even hundreds of different proteins (Muller-McNicoll and Neugebauer 2013; Dreyfuss et al. 2002). These proteins have a variety of roles, one being a protective coat against the many mRNA decay factors roaming the cell in search for unprotected mRNA ends to latch onto. Some other functions might be the regulation of mRNA

processing and maturation, mRNP packaging, and determination of export properties. Many of the mRNP proteins assemble co-transcriptionally (Neugebauer and Roth 1997) and accompany the mRNA as it travels to the cytoplasm (Pinol-Roma and Dreyfuss 1992; Visa et al. 1996; Le Hir et al. 2000b). The mRNP is remodeled as some proteins are removed during or following export in a process of mRNP re-modelling, such as the nuclear polyA-binding protein (PABPN1) that is exchanged with the cytoplasmic polyA binding protein (PABPC1), or the nuclear cap binding complex (CBC) that is exchanged for eIF4E (Lejeune et al. 2002; Hosoda et al. 2006). Some of the proteins that assemble on the mRNA in the nucleus can determine the cytoplasmic fate of the mRNA in terms of RNA localization and translation efficiency (Ross et al. 1997) or even function in cytoplasmic mRNA decay (Kataoka et al. 2000; Le Hir et al. 2000a, 2001; Haimovich et al. 2013a, b).

What does an mRNP look like—is it a round granule or perhaps elongated or irregular? How many proteins does it contain? How is mRNA folded within the complex—is it hidden inside or perhaps is it wrapped around? Most of these questions remained unanswered. The extensive biochemical studies on mRNP composition (Sperling and Sperling 1990) have established that many of the RNA-binding proteins coating the mRNA belong to the hnRNP family (Dreyfuss et al. 2002; Dreyfuss 1986), which rapidly and co-transcriptionally bind to the nascent transcripts (Fakan et al. 1986), but their exact role in mRNP biogenesis and transport remains enigmatic for now. The heterogeneous ribonucleoprotein particles (hnRNPs) were characterized using purified hnRNP complexes which contained approximately a 4:1 stoichomoteric ratio of protein to RNA. The studied hnRNP particles were 20–25 nm in diameter but may have been part of larger complexes since they were obtained by nuclease digestion. Detection of similar sized particles on actively transcribing genes in *Drosophila* embryo cells was obtained by Beyer and Osheim who implemented the Miller chromatin spreading technique for the detection of RNA polymerase II (Pol II) transcribed mRNAs (Beyer and Osheim 1988; Osheim et al. 1985). They could not identify which active genes were being detected, but could clearly observe multiple nascent mRNAs of increasing length associated with the gene body. Each transcript had 2 RNP particles associated with it, one at the 5′ and one usually at a similarly spaced position downstream. However, subsequent studies revealed that these were splicing related particles and were not the mature mRNPs. Performing the Miller spreads at less stringent conditions revealed 50–60 nm granules close to the chromatin, which could be mRNPs (Sperling and Sperling 1990).

EM studies performed *in situ*, without the chromatin dispersal inherent to Miller spreads, examined the site of formation of newly synthesized mRNAs. The nascent mRNAs were coined ‘peri-chromatin fibrils’ since they were found in close proximity to chromatin (Fakan 1994; Bachellerie et al. 1975). These fibrils could sometimes be seen forming into single granules of 35–55 nm in diameter (Puvion-Dutilleul and Puvion 1981), and were structurally similar to the best documented example of single mRNP detection by EM, the Balbiani ring (BR) mRNP produced from the salivary glands of the dipteran *Chironomus tentans*.

The BR mRNPs contain extremely long mRNAs (35–40 kb) that are transcribed from the BR puff genes and therefore are easily detected as 50 nm granules that form co-transcriptionally and travel through the nucleoplasm until they reach the NPC (Fig. 1) (Stevens and Swift 1966; Skoglund et al. 1983; Daneholt 1999). The BR granules were spherical and uniformly dense after uranyl staining, and distinguishable from ribosomes or chromatin. These granules provided much insight to the process of mRNA export on account of the large size of the mRNPs, to be discussed later on. Studying the formation of the BR pre-mRNPs has shown that they are structurally diverse particles, mostly due to size and structure of the transcript, and the mRNA processing events that occur on the pre-mRNA (Bjork and Wieslander 2011). Structural analysis of the mature nucleoplasmic BR granules by EM tomography showed that they had similar structure, namely a sphere with a central hole. The particles contained a thick RNP ribbon folded into four domains, with the 5′ end at the top of the first domain and the 3′ end in the fourth domain, suggesting close proximity between the two ends of the mRNA (Skoglund et al. 1986). A study that purified nuclear mRNPs from yeast found similar configuration, exhibiting a ribbon-like elongated structure with lateral constrictions, and length which was dependent on mRNA size (Batisse et al. 2009). The authors suggest that these 25–30 nm long mRNPs contain mRNA in a condensed manner, since a 1-kb linear mRNA would be 340 nm in length, tenfold longer that

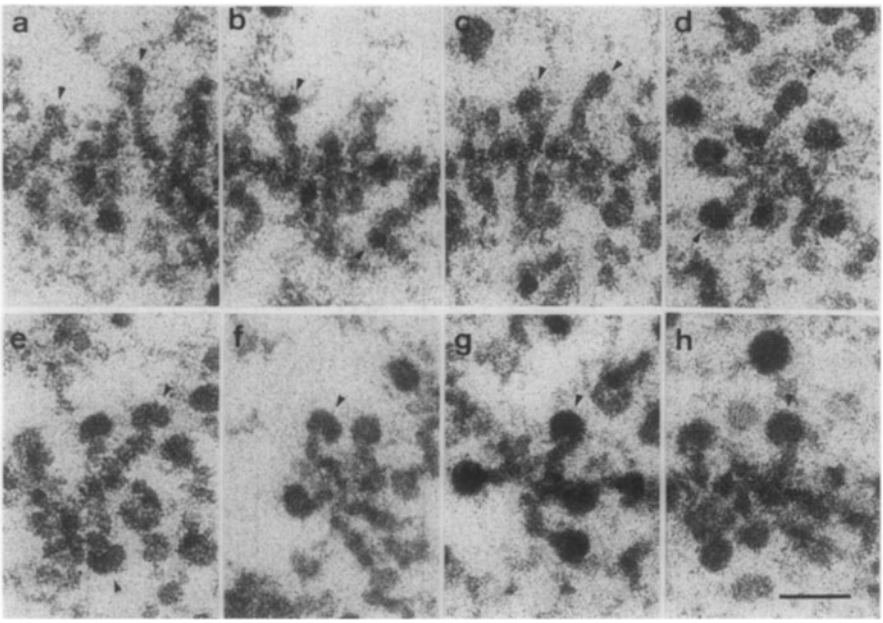

Fig. 1 Electron micrographs of growing Balbiani Ring (BR) RNP particles during various stages of maturation in *C. tentans* cells. Examples are shown from the (**a**) proximal, (**b**, **c**) transitional, and (**d**–**h**) distal portions of the gene. *Bar* = 100 nm. Reprinted by permission from *Cell Press* (Skoglund et al. 1983)

the mRNP particles. Very few studies of this sort document RNP structure, for example, purification of 30–120 nm influenza virus vRNPs (Wu et al. 2007), or *in vitro* generated RNPs with average lengths of ~130 nm (Matsumoto et al. 2003; Skabkin et al. 2004). Altogether, it is difficult to conclude whether there is common architecture for all mRNPs, and more study is required in this field.

3 Detecting Specific mRNAs

It was not always clear that DNA and RNA coexisted in the same cells. In fact, at the beginning of the twentieth century it was "common knowledge" that thymonucleic acid (DNA) was found only in animal cells whereas zymonucleic acid (RNA) was found in plants. At the time, Brachet was working on sea urchin eggs and found that cells producing high levels of protein also contained high concentrations of RNA (Thieffry and Burian 1996). In his studies, DNA and RNA nucleic acids were observed using cytochemical approaches (e.g., Feulgen, Unna and toluidine blue stains) until he finally developed the widely used methyl-green pyronin RNA stain. These staining approaches provided information as to the presence of RNA in all types of cells and within the different compartments of the cells.

A dramatic step forward in RNA observation came with the development of a method that could detect specific RNA or DNA sequences within cells. The *in situ* hybridization (ISH) method developed by Joe Gall used radioactive nucleic acid sequences that were complementary to DNA (or RNA) sequences. Gall and Pardue were the first to detect specific chromosomal regions such as satellite DNA using DNA probes, and the ribosomal DNA genes using radioactive ribosomal RNA as probe (Gall and Pardue 1969; Pardue and Gall 1969, 1970). This method was rapidly applied to many biological systems thus enabling the detection of endogenous genes and RNAs in fixed cells and tissues. The ISH protocol was refined and the hybridizing oligonucleotide sequences were labeled with enzymes that produced a colored stain in place of the radioactive labels. An additional improvement came with the appearance of fluorescence microscopy. Direct labeling of the oligonucleotides with fluorescent fluorophores established the fluorescence *in situ* hybridization (FISH) technique, popularly used in both basic sciences and diagnostics (Levsky and Singer 2003).

The ability to detect bulk RNA within cells as well as specific RNA targets led to important observations regarding the location of mRNA within the cell. In electron microscopy studies the nascent mRNA transcripts were observed as fibers protruding from the DNA (peri-chromatin fibrils), and transcription was found to take place in the peripheral areas of chromosomal regions that were in contact with the nucleoplasmic surroundings, whereas more internal regions of a chromosome were seen to be less transcriptionally active (Zirbel et al. 1993; Verschure et al. 1999). mRNA molecules that had left the site of transcription seemed to be randomly dispersed within the nucleoplasm, usually in between chromatin dense

regions (Singh et al. 1999; Pante et al. 1997). With the detection of nuclear bodies containing pre-mRNA splicing factors [termed inter-chormatin granules or nuclear speckles (Spector 1993)] it was thought that these might serve as the sites of splicing activity during the pathway an mRNA takes from the gene to the nuclear pore (Huang and Spector 1992). However, it turned out that bulk mRNA or specific transcripts were distributed throughout the nucleoplasm (Zachar et al. 1993; Dirks et al. 1995; Snaar et al. 2002) and no particular accumulations of mRNA could be detected. Even though poly(A) FISH detected considerable RNA signal in nuclear speckles, the actual accumulation was probably not more than twofold higher than the nucleoplasm (Fay et al. 1997). In the light of our current knowledge that many long non-coding RNAs (lncRNA) are transcribed by RNA polymerase II, contain poly(A) tails, and are nuclear retained [it was known early on that much of the nuclear RNA never left the nucleus (Perry et al. 1974; Herman et al. 1976)], it is possible that ncRNA is a substantial component of nuclear speckles as indicated by the detection of the MALAT/NEAT2 ncRNA in speckles (Hutchinson et al. 2007). In any case, the exact function of these nuclear speckle bodies remains controversial and they do not specifically accumulate mRNAs within (Lamond and Spector 2003; Hall et al. 2006). Intriguingly, some mRNA labeling studies demonstrated fiber-like tracks in the nucleus (Lawrence et al. 1989) suggesting that mRNAs could transport on a filamentous nuclear network, reminiscent of the actin or tubulin cytoskeletal highways observed in the cytoplasm. To date, such a nuclear transport system has not been detected.

The popularity of RNA FISH combined with conventional fluorescence microscopy provided much qualitative information on mRNA distribution in cells and tissues, but lacked a quantitative angle. Moreover, many of the protocols suffered from high background issues that did not enable the detection of single molecules. This required modification of the technique such as signal enhancement that will easily differentiate between the real molecules and the background. To overcome these issues, Singer and colleagues developed a single molecule RNA FISH approach in which each mRNA transcript of interest was targeted by a series of five short complementary DNA oligonucleotides (aprox. 50 nucleotides long), and each oligo (or probe) was labeled with several (3–5) fluorophores. Thereby, each transcript was labeled with many fluorophores, providing a strong point of fluorescence detectable by fluorescence imaging as well as high signal versus the background fluorescence of the specimen and non-specifically bound fluorescent probes. This approach enabled the detection of actively transcribing genes in mammalian cells and the counting of nascent and cellular mRNAs (Femino et al. 2003; Shav-Tal et al. 2004b). For instance, quantification of single molecules of β-actin mRNA showed that a quiescent population of cells contained 500 ± 200 β-actin mRNAs per cell, whereas a proliferating population had up to ~1500 copies per cell (Femino et al. 1998). At high activation levels the transcribing β-actin alleles harbored ~30 nascent transcripts suggesting the presence of numerous RNA polymerase II enzymes associated with the DNA along the β-*actin* gene body. Using differently labeled probe sets to the 5′-untranslated region (UTR) and the 3′UTR they could determine a rate of 1.1–1.4 kb/min for RNA Pol II transcription.

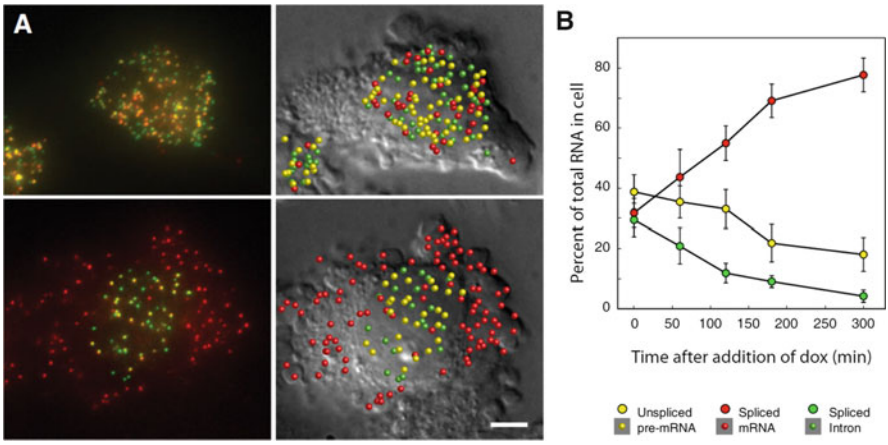

Fig. 2 Demonstration that pre-mRNA molecules dispersed in the nucleus are capable of being spliced. (**a**) *Upper panels* show composite RNA FISH images of cells in which a gene containing array with an intron was transcriptionally induced for a brief period (2 h). Many pre-mRNA molecules are seen scattered within the nucleoplasm with little accumulation of spliced mRNA molecules in the cytoplasm. *Lower panels* show images from the same batch of cells as above, but in which induction was followed by a period of suppression (2 h). There was a decrease in the proportion of pre-mRNA molecules in cells fixed after the chase period, with a concomitant increase in spliced mRNA molecules in the cytoplasm. Raw images are shown on the *left* and overlays with colored balls identifying the RNA species are presented on the *right*. (**b**) Percentage of the three different RNA species in individual cells as a function of time after the addition of doxycycline (dox). Doxycycline turns off new RNA synthesis within several minutes. Even though the proportion of spliced mRNAs continues to increase after 3 h, the overall number of RNAs declines due to degradation. *Error bars* represent 95 % CI. Reprinted by permission from *Cell Press* (Vargas et al. 2011)

The single molecule RNA FISH approach underwent another level of simplification by Raj, van Oudenaarden and Tyagi making it easily applicable in many laboratories (Raj et al. 2008). In place of the unique fluorophore conjugation procedures required for labeling the probes in 3–5 different positions within one probe, the DNA probes were typically labeled only at one end, and signal amplification was obtained by the use of between 40 and 100 short probes to the known mRNA sequence (compared to five probes in the previous approach). This approach and others are now commercially available [reviewed in (Pitchiaya et al. 2014)]. Subsequent studies have used these single molecule techniques to quantify mRNA expression levels in different types of cells and tissues generating a broad picture of stochastic behavior in gene expression patterns (Fig. 2), for instance see (Yunger et al. 2010; Raj et al. 2006; Vargas et al. 2011; Levsky et al. 2002; Zenklusen et al. 2008; Itzkovitz et al. 2012; Hansen and van Oudenaarden 2013; Waks et al. 2011; Battich et al. 2013; Chou et al. 2013; Hoyle and Ish-Horowicz 2013; Lee et al. 2014).

4 Bringing mRNAs to Life

Do drive RNA detection from fixed cell imaging to real-time imaging, Pederson and Politz applied the FISH method to living cells. Much of the initial detections of RNA by FISH in fixed cells were performed using an oligo(dT) fluorescent probe that hybridized with the poly(A) tails of all mRNAs, thus detecting the bulk poly (A)-containing populations of nuclear RNAs. This approach in living cells and the detection of several sub-populations based on their nuclear mobility [one of the first applications of fluorescence correlation spectroscopy (FCS) in the study of molecule mobility in living cells], suggested the existence of different mRNA populations that may vary in size (Politz et al. 1995, 1998). This study brought upon a whole new set of scientific questions and motivated the generation of new approaches for RNA labeling in living cells. It is important to note two of the major obstacles that had to be addressed in future development of studies in living cells. First, since excess oligo(dT) probe roamed the nucleus and could not hybridize with mRNA, it was difficult to distinguish between the mRNA-probe fraction and the unbound probe, hence the required use of FCS that could help differentiate between the populations. But the latter did not provide a solution for examining where in the nucleus do the mRNAs actually travel. Second, as with FISH in fixed cells, it became important to be able to examine specific mRNA transcripts rather than only the bulk poly(A) population.

An elegant approach helped solve the first issue of detection. Instead of labeling bulk mRNA with a fluorescent oligo(dT) probe, a caged fluorophore was attached to the probe, and only by specific activation of the caged fluorophore could the probe become detectable (Politz et al. 1999). In this manner, Politz and colleagues activated only a small portion of the probe in one area of the nucleus, and subsequently could follow the fluorescently tagged mRNA molecules over time. If mRNA were a non-mobile molecule one would expect the uncaged fluorescent signal bound to the mRNA to remain in one spot. This was in fact not the case at all, and the movement of the hybridized uncaged signal could be tracked over time. Importantly, this study included labeling of the DNA using the Hoescht 33342 dye that can be applied to living cells, and unequivocally demonstrated that mRNA traveled throughout all the nucleoplasmic space that was not occupied by chromatin (Fig. 3). In light of the abovementioned accumulation of poly(A) signal in nuclear speckles, it was later on shown that mRNA passed through nuclear speckles with the same mobility as within the rest of the nucleoplasm, not showing any "rest-stops" at which it might pause for further processing (Politz et al. 2006; Molenaar et al. 2004).

The appearance of green fluorescent protein (GFP) at the doorstep of cell biology expanded the toolbox for mRNA tagging in living cells. For instance, instead of using oligo(dT) probes, the group of Carmo-Fonseca used the natural nuclear poly(A)-binding protein (PABPN1) fused to GFP, to bind to the poly (A) region of mRNAs (Calapez et al. 2002). This study used fluorescence recovery after photobleaching (FRAP) to measure the nuclear mobility of the different

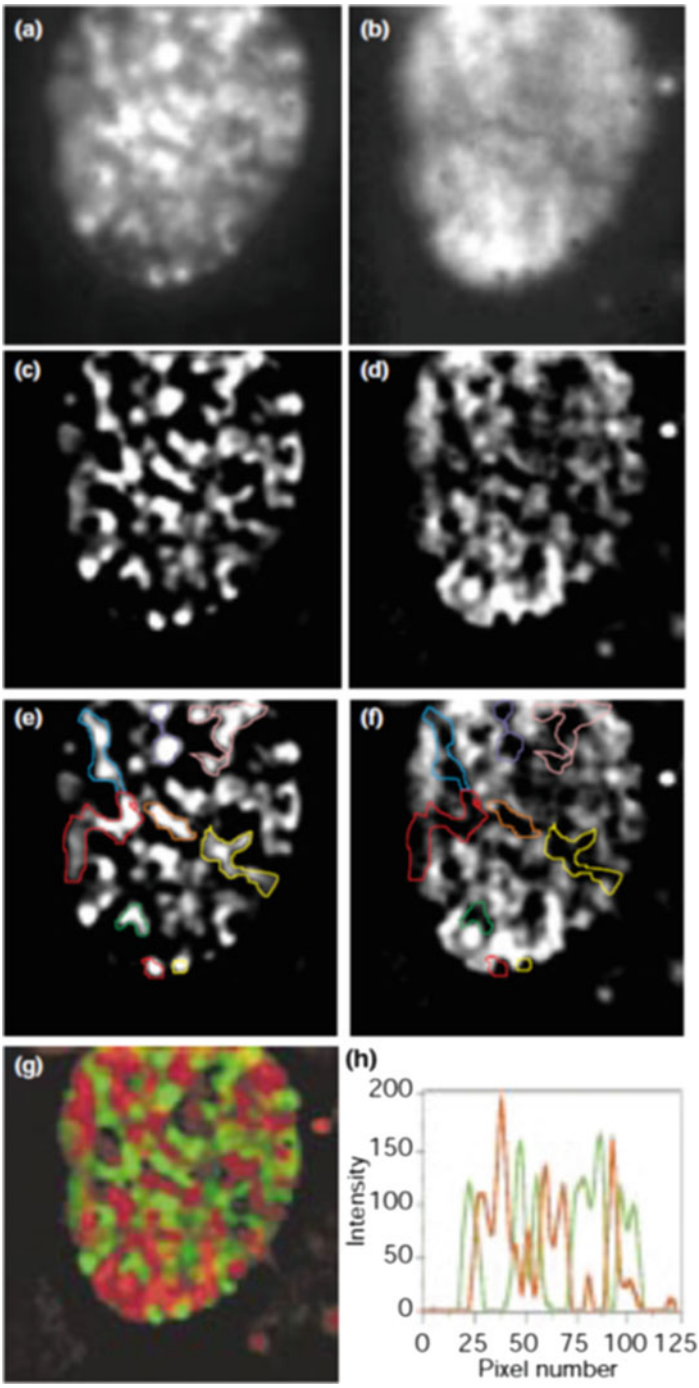

Fig. 3 Intranuclear localization of uncaged fluorescein labeled (FL) FL–oligo(dT) compared to chromatin distribution. Cells were incubated sequentially with caged FL–oligo(dT) and Hoechst 33342 and three-dimensional stacks in both (**a**, **c**, **e**) *blue* (Hoechst-labeled chromatin), and (**b**, **d**, **f**)

populations of moving molecules and could distinguish between free versus mRNA-bound GFP-PABPN1. Still, it was not possible to detect specific mRNAs. To this end, the laboratory of Singer generated a unique tagging sequence that could be inserted into a gene of interest, and subsequently the mRNA molecule would contain the tagging sequence within, that would be bound by a specific RNA-binding protein (RBP). In order for this tag not to interact with eukaryotic RBPs, the chosen sequence was taken from the MS2 bacteriophage, which contains an MS2-coat protein (MCP) that binds to a unique stem-loop structure in the phage MS2 RNA (Bertrand et al. 1998). The MS2 sequence was introduced as a series of 24 sequence repeats into the 3′UTR of a mammalian gene, thus forming 24 stem-loops in the mRNA transcribed from the gene, to be bound by GFP-MS2-CPs (Fig. 4).

The binding of the many GFP-CP RBPs to this specific mRNA immediately as this region was transcribed allowed the detection of the mRNA during transcription (Janicki et al. 2004; Darzacq et al. 2007; Boireau et al. 2007; Brody et al. 2011), co-transcriptional mRNA splicing (Martin et al. 2013; Coulon et al. 2014), release from the gene and nucleoplasmic travels (Shav-Tal et al. 2004a), and the final nuclear point of mRNA export (Mor et al. 2010b; Grunwald and Singer 2010). It is notable that this technique has been successfully implemented in prokaryotes as well as in almost every eukaryotic model organism used in experimental biology (Fig. 5) (Lionnet et al. 2011; Bertrand et al. 1998; Chubb et al. 2006; Muramoto et al. 2012; Golding and Cox 2004; Golding et al. 2005; Bothma et al. 2014). Additional RNA tagging platforms based on similar repeated sequences, known as PP7 and λ_N (Coulon et al. 2014; Martin et al. 2013; Schonberger et al. 2012; Daigle and Ellenberg 2007), have since emerged thus expanding the possibilities for simultaneous mRNA tagging in living cells (Hocine et al. 2013). The dynamics of single mRNPs could then be followed in living human cells showing that mRNPs travel by diffusion at rates that are between 10 and 100 fold slower than single proteins or small complexes (Shav-Tal et al. 2004a). Movement and diffusion rates of mRNPs have since been measured by a variety of additional mRNA tagging techniques (Vargas et al. 2005; Shav-Tal and Gruenbaum 2009; Siebrasse et al. 2008; Ishihama and Funatsu 2009; Thompson et al. 2010; Tyagi 2009; Bratu et al. 2003; Kubota et al. 2010; Santangelo et al. 2009; Gohring et al. 2014), altogether highlighting the bulkiness of the large mRNP particle as it

Fig. 3 (continued) green (uncaged FL–oligo(dT)), channels were captured and restored. (**a**, **b**) Raw and (**c**, **d**) restored midsections show the distribution of Hoechst signal and uncaged FL–oligo (dT) signal in the same nucleus. (**e**, **f**) The same images as in (**c**, **d**) but high intensity regions of Hoechst signal were (**e**) outlined and (**f**) the outlines superimposed on the oligo(dT) image. (**g**) A color encoded overlay in which the Hoechst signal is *green* and the oligo(dT) signal is *red*. (**h**) A plot (linescan) of the intensity (arbitrary units) versus pixel number for the Hoechst (*green*) and oligo(dT) (*red*) signals as they vary along a line across the middle of (**g**). For (**a**–**g**), each image is ~19 × 19 mm. Reprinted by permission from *Cell Press* (Politz et al. 1999)

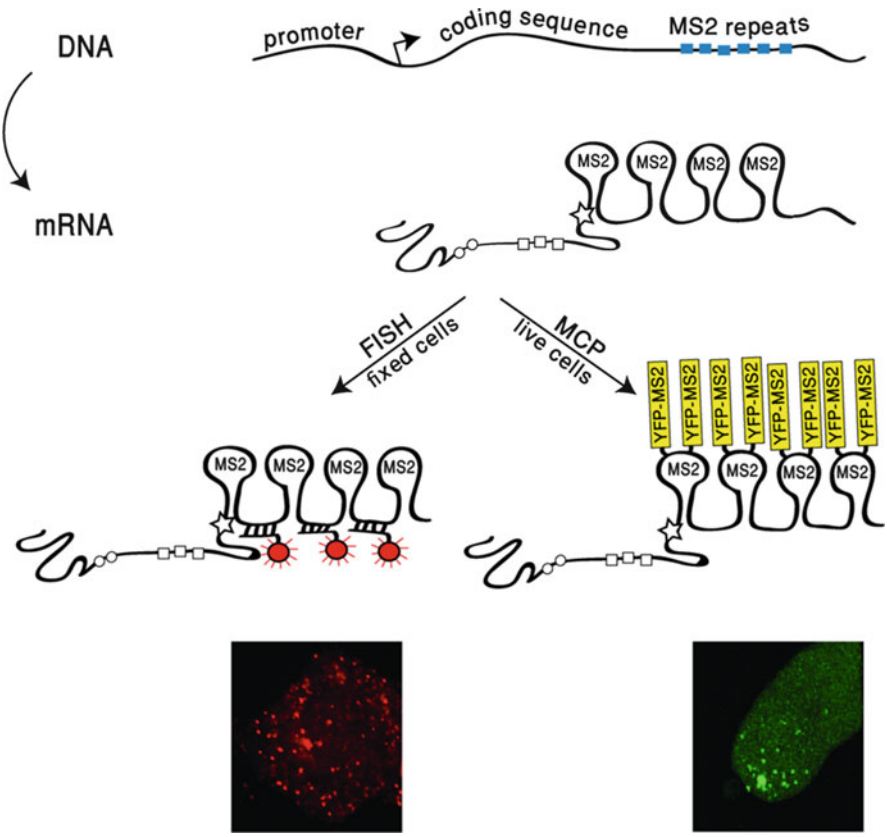

Fig. 4 The MS2 mRNA tagging system. The MS2 sequence, when transcribed, forms repeated stem-loop secondary structures in the mRNA molecule. To observe transcription in living cells, the YFP-MCP protein (can also be GFP/CFP/mCherry etc.) is transfected, and then binds the MS2 loops. For use by RNA FISH in fixed cells, fluorescently tagged DNA oligonucleotides complementary to MS2 repeated sequences, hybridize to the target mRNA. *Empty circle*, *square* and *star* indicate RBPs. At the *bottom*, images corresponding to each of the methods: single mRNPs are detected in the nucleus of 293 T cells using GFP-MCP (*right*), or by RNA FISH (*left*)

travels through the nucleoplasm, randomly moving to end up at the exit point at the nuclear pore.

Finally, these studies following bulk RNA and specific mRNA movement in cells showed that the nuclear mRNA movement was diffusion-based and that there was no energy-requiring process utilized by the cell to drive mRNA nucleocytoplasmic transport (Politz et al. 1998; Shav-Tal et al. 2004a; Politz et al. 2006). Although some studies could find an effect of ATP depletion on the mobility of mRNA in living cells (Molenaar et al. 2004; Calapez et al. 2002), it seems that was an indirect effect, while the primary site of energy depletion was on global chromatin structure thereby affecting the structure of the inner nuclear space and confining the movement of mRNPs within the nucleoplasm (Shav-Tal

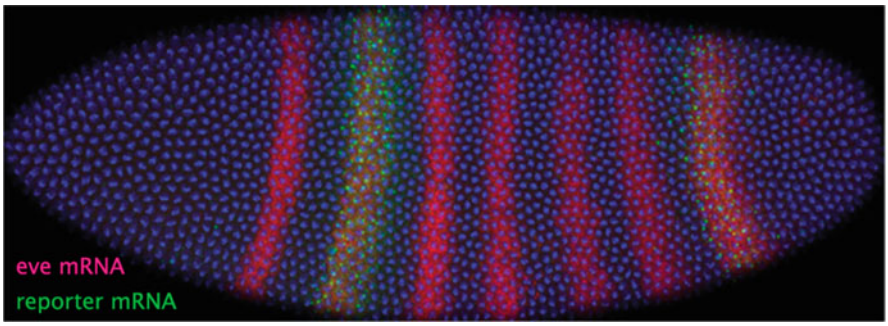

Fig. 5 Confocal image of a transgenic *Drosophila* embryo carrying an eve > MS2 transgene. The mRNA is labeled via *in situ* hybridization with probes for the reporter gene (*green*) and endogenous eve mRNAs (*red*) in the same embryo during nuclear cycle 14. Eve > MS2 transcripts identify authentic stripe 2 and stripe 7 expression patterns. Reprinted by permission from *PNAS* (Bothma et al. 2014)

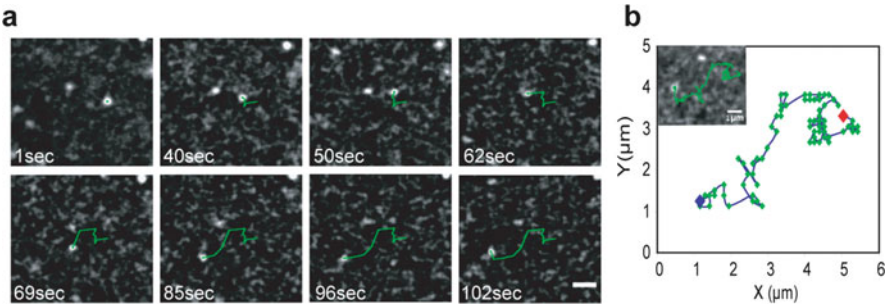

Fig. 6 Analysis of mRNP kinetics in the cell nucleoplasm. (**a**) Frames of a diffusive nucleoplasmic mRNP labeled with the YFP-MS2-CP tracked for 102 s (*green* track). (**b**) The full tracked movement from (**a**). *Red*, start of track; *blue*, end of track. *Bar* = 1 μm. Reprinted by permission from *Nature Publishing Group* (Mor et al. 2010b)

et al. 2004a; Mor et al. 2010b). To this end we can conclude that the bulk of genetic information moving from the nucleus into the cytoplasm in the form of messenger RNA molecules reaches it cytoplasmic destination by diffusion, and that the cell does not require energy investment for this step of message transport, in contrast to the cytoplasm where some of the mRNAs must be transported by molecular motors (Shav-Tal and Singer 2005). We could track the movement of single mRNPs in the nucleoplasm of mammalian cells (Fig. 6) and showed that the timescale of nucleocytoplsamic transport ranged from 5 to 40 min on average, with longer mRNAs prone to longer transport times (Ben-Ari et al. 2010; Mor et al. 2010b). Therefore, the timeframe of this path is determined by the particle size of the mRNP that would be influenced by mRNA length and the number of proteins coating the mRNA, as well as the biophysical properties of the nucleoplasmic space in which the mRNPs

travel and the hindrance of the chromatin structure (Mor and Shav-Tal 2010; Mor et al. 2010a; Roussel and Tang 2012).

5 Exit to the Cytoplasm

Export of mRNAs through the NPC and into the cytoplasm, where they are to reach the translation machinery, is an irreversible step, thus making this a key point of regulation in gene expression. One of the main systems used to visualize mRNA export are the abovementioned exceptionally large BR mRNPs easily observed by EM in the salivary glands cells of *C. tentans*, and their detection on the nucleoplasmic and cytoplasmic sides of the NPC allowed the examination of mRNPs during export. In an early study, Stevens and Swift showed images of the BR granule re-structuring to form a rod shape during translocation through the pore (Stevens and Swift 1966). Immunoelectron microscopy images demonstrating co-localization of the RNA helicase Dbp5 with the BR mRNP as it changes shape, implied the involvement of a helicase in this transformation (Zhao et al. 2002). Indeed, the remodeling of the mRNP on the cytoplasmic side is thought to prevent its return back into the nucleus, suggestive of a molecular ratchet mechanism imposing directionality on nucleo-cytoplasmic mRNA export (Stewart 2007). Using TEM and SEM (transmission and scanning EM, respectively), Kiseleva et al. were able to provide remarkable pictures of the BR mRNPs during the act of transport through the pore. Based on the analysis of these images, a model was proposed for mRNA export, in which the nuclear basket ring exhibits dynamic restructuring in order to allow the passage of the mRNP through it (Kiseleva et al. 1996, 1998; Daneholt 1997). The passage of the mRNP seems to be directional since Mehlin, Daneholt and Skoglund showed by applying 3D technology on high resolution images of BR mRNPs, that the 5′ region of the mRNA is first to

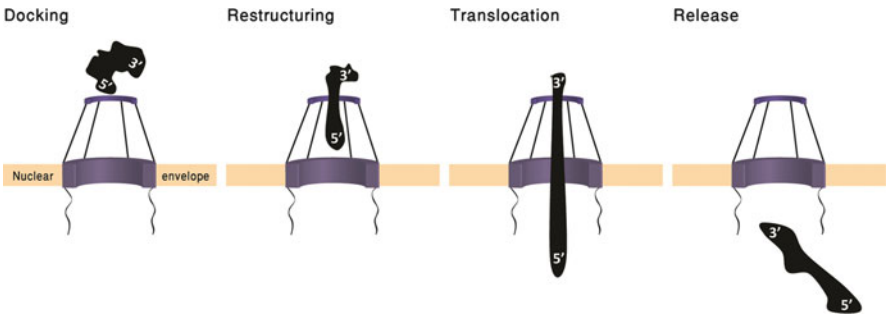

Fig. 7 A schematic overview of BR mRNP export, based on (Daneholt 1997). *Left to right*: the mRNP granule docks at the ring of the nuclear basket. Then, the mRNP changes shape to a rod-like structure and begins to enter the central channel of the pore, with the 5′ region inserted first. Next, the mRNP translocates through the pore, crossing from the nucleus to the cytoplasm, and finally released from the NPC

enter the pore (Fig. 7), suggestive of functional interactions between mRNP proteins situated at the 5′-end of the mRNA and the pore proteins, prior to translocation (Mehlin et al. 1992). Supporting information that this was a general phenomenon in mammalian cells as well, was provided by the group of Robin Reed who used RNA immunopercipitation experiments in human cells to show that the transcription-export protein complex (TREX) is situated on the first exon near the 5′-end of the mRNA (Cheng et al. 2006).

Other groups were motivated to search for proteins that may act as porters in nucleo-cytoplasmic transport. For instance, in a study conducted by the Dreyfuss laboratory during their studies of hnRNP proteins, the hnRNP A1 mRNA-binding protein was shown to shuttle between the nucleus and the cytoplasm (Michael et al. 1995). This seemed a suitable quality for a carrier of mRNAs, and indeed specific amino acid sequences were identified as crucial for export activity, and termed nuclear export signals (NES). In accordance, the titration of NES-receptors with NES-conjugated peptides in *Xenopus* oocytes cells resulted in mRNA export inhibition (Pasquinelli et al. 1997), suggesting that the NES-mediated mRNA export pathway is limited by NES-receptor availability. Meanwhile, the RBP Crm1 was discovered. This protein also possessed nucleo-cytoplasmic shuttling properties and was suspected as a carrier of mRNAs. Use of leptomycin B which specifically inhibits CRM1, caused poly(A) RNAs to accumulate in the nucleoplasm thus strengthening the notion of RBP-facilitated transport in mammalian cells (Watanabe et al. 1999). Using inhibition of LMB in heterokaryons of HeLa cells and *Xenopus* A6 cells demonstrated that hnRNP A1 can still translocate from the HeLa cells into the *Xenopus* cells, thus implying that the Crm1 export pathway and hnRNP A1 export are separable (Lichtenstein et al. 2001). The Crm1 pathway is currently considered important mostly in protein transport.

Accumulating data indicated separate export pathways for ncRNAs and mRNAs, the latter involving the mammalian protein TAP/NXF1 (or yeast Mex67). When constitutive transport element (CTE) containing mRNAs taken from viral RNAs, were microinjected by the Izaurralde group concomitantly with recombinant TAP into nuclei of *Xenopus* oocytes, an increase in mRNA export was registered. Indeed, the C-terminal domain of TAP interacts directly with the FG-repeat domains of different nucleoporins (Nups) both on the nuclear and cytoplasmic sides of the NPC (Bachi et al. 2000). The export machinery also interacts with upstream events of gene expression (Luna et al. 2008). For instance, coupling between pre-mRNA splicing and export was shown after the microinjection of ^{32}P-labeled pre-mRNA and mRNA into nuclei of *Xenopus* oocytes, and the observed increased export of spliced mRNAs whereas only 5 % of the unspliced mRNA underwent export (Luo and Reed 1999). This issue was re-examined by looking at β-globin mRNA distribution by RNA FISH. Quantifying single mRNA cellular localization demonstrated higher cytoplasm to nucleus ratio of the spliced mRNAs compared to unspliced transcripts (Valencia et al. 2008), demonstrating the enhancing effect of splicing on mRNA export.

Quite surprisingly, one RNA FISH study has shown regarding the signal sequence coding region (SSCR) amino acid sequence used to localize secreted

proteins to the ER, that this same sequence but on the mRNA nucleotide level will allow the mRNAs of these proteins to export independently of the TREX proteins but in a TAP mediated process (Palazzo et al. 2007). RNA FISH localization assays also helped in sorting out export pathways. For instance, influenza A vRNAs in MDCK cells showed co-localization with GFP-TAP thus implying that the TAP host cellular export mechanism is exploited for the packaging of influenza A virus (Wang et al. 2008). Currently, it is realized that many more RBPs and post-translational modifications are involved in defining the mRNA export process (Tutucci and Stutz 2011).

A further visual demonstration of the importance of Nups in mRNA export was obtained by the injection of antibodies to Nup153 into *C. tentans* salivary gland cells, and as a result, the export of BR mRNPs and rRNA was blocked (Soop et al. 2005). This study suggested that mRNP entry into the nuclear basket is a two-step process; first the mRNP binds to the tip of the basket fibrils and only then is it transferred through the basket by a Nup153-dependent process. Later on, live-cell studies following the behavior of single mRNPs in the human nucleus during blockage of mRNA export (using a dominant negative form of Dbp5) showed that indeed mRNP binding to the NPC occurred independently of export itself (Hodge et al. 2011). To provide compelling evidence as to the role of the already suspected DEAD-box ATPase Dbp5 in mRNA export, Lund and Guthrie employed oligo (dT) cellulose chromatography to extract mRNPs from *Saccharomyces cerevisiae* and quantified the bound fraction of Mex67-GFP on Dbp5 mutants compared to wild-type Dbp5. This resulted in an increase of the bound protein on the Dbp5 mutant, implying that Dbp5 is an active participant in the removal of Mex67 and as the terminator of the mRNA export process (Lund and Guthrie 2005).

It is almost dogmatically accepted that all import and export to the nucleus can only follow through the NPCs. Therefore, the field was overwhelmed by the demonstration of an alternative export pathway independent of NPCs. This process resembles herpes virus budding. The studied large mRNP granules in *Drosophila* synapses were found to exit the nucleus via budding through the inner and the outer nuclear membranes (Speese et al. 2012). However, to date, this is the exception rather than the rule, and in fact even the exact path taken by the mRNP inside the pore is not clearly defined. EM examination of various cargoes moving through the pore has distinguished between two pathways, central and peripheral. Use of nanometer sized RNA-gold conjugates offered the opportunity to examine different sub-classes of regular sized RNAs (mRNA, rRNA, tRNA) by EM to test questions regarding the exact pathway of passage within the pore and the competitive nature of RNA export, rather than using radiolabeled RNAs (Jarmolowski et al. 1994). The Mattaj group conjugated DHFR mRNA, tRNA and U1 snRNA to gold particles and microinjected them into *Xenopus* oocyte nuclei and found that only RNA species of the same type could inhibit export by competition (Pante et al. 1997). Analyzing these gold-mRNA conjugates at NPCs showed that mRNA passes through the center of the NPC, in accordance with an earlier study (Dworetzky and Feldherr 1988). In contrast, when Cook and colleagues examined mRNA localization at the pore using indirect immunogold labeling in HL-60 cells, the labeled transcripts

localized at the side of the pore channels and not in the center (Iborra et al. 2000). Huang and Spector presented similar findings using electron microscope pre-embedding *in situ* hybridization with eosin photo-oxidation to monitor poly (A) RNA in HeLa cells, to reveal a stronger staining in the periphery of the pores indicative for transport through the side of the NPC channel (Huang et al. 1994).

The well-known hypothesis termed "gene gating" proposed by Gunter Blobel (Blobel 1985), argued that NPCs play an active part in nuclear organization through interactions between NPC constituents and DNA sequences. It was proposed that the proximity of a gene to the nuclear envelope would facilitate export. Although some studies in yeast strengthen the latter (Casolari et al. 2004; Cabal et al. 2006), most studies in living mammalian cells demonstrate mRNAs slowly diffusing through the whole nucleoplasm on their way to the nuclear pores, as discussed above (Sheinberger and Shav-Tal 2013). mRNA export however, was always considered a rapid event, since not much mRNA was detected within the pores by the different staining approaches used. With the improvement of imaging techniques and rapid live-cell imaging, these type of studies could focus on the detection of mRNP dynamics at the pore. Large MS2-tagged mRNPs were visualized exiting the nucleus at an estimated time frame of 500 ms or less, and were seen to approach the NPC in a compact form to then emerge in the cytoplasm as a disorganized open structure, implying remodeling of the large mRNP during passage through the pore (Mor et al. 2010a, b). In a study performed on endogenous MS2-tagged β-actin mRNAs together with labeled NPCs, export kinetics were measured using super-registration which employs high-sensitive cameras and provide a time resolution of 20 ms. This revealed that a total of 180 ms was required for passage through the NPC, and that the longer dwelling times occurred at the nuclear and cytoplasmic sides (~80 ms) while the movement through the central channel was significantly faster (5–20 ms) (Grunwald and Singer 2010). Another rapid imaging approach could show that export times were in the range of ~12 ms (Ma et al. 2013). This study could also reconstruct a 3D pathway of mRNPs traveling through the NPC in live cells, and found that mRNPs moved along the periphery of NPC and not through the central axial channel that was used by small passively diffusing molecules. Even BR mRNPs were followed in live salivary gland cells They were indirectly tagged by a fluorescently labeled hrp36 protein (hnRNP A1 homologue) and export times ranging from 65 ms to 6 s were measured (Siebrasse et al. 2012). They could also detect significant binding times at the pore before export, pointing to a rate-limiting step occurring at the nuclear basket.

Future studies will enable direct examination of both mRNPs and NPCs in living cells to better understand the structural changes both undergo as the large mRNP complex travels through the channel in the NPC. It seems probable that both structures must transform to some extent but exactly what happens is not well understood. For instance, using wavelength anomalous dispersion on NPC crystals derived from *Rattus norvegicus*, Melcak et al. proposed that circumferential sliding of Nup58/45 affects the pore diameter and allows transport of macromolecules, potentially explaining how mRNPs translocate through the pores (Melcak et al. 2007). Another study proposed that the nucleoplasmic basket filaments are

connected at their distal ends, and only when an mRNP engages with this structure, does the nuclear basket ring form and continue to dilate as the mRNP passes through (Kiseleva et al. 1998). As a consequence, the filaments shorten, potentially assisting the mRNP approach to the central channel. It will be important to better understand the exact areas of interactions and molecular events of binding and remodeling of mRNA-associated RBPs moving through the NPC. These studies will take us further to examining the connections of nuclear trafficking in human disease and as potential targets for pharmaceutical intervention (Mor et al. 2014).

Acknowledgements The work in the Shav-Tal laboratory is supported by the European Research Council (ERC), the Israel Science Foundation (ISF), the Gassner Medical Research Fund and the United States-Israel Binational Science Foundation (BSF).

References

Bachellerie JP, Puvion E, Zalta JP (1975) Ultrastructural organization and biochemical characterization of chromatin—RNA—protein complexes isolated from mammalian cell nuclei. Eur J Biochem 58(2):327–337

Bachi A, Braun IC, Rodrigues JP, Pante N, Ribbeck K, von Kobbe C, Kutay U, Wilm M, Gorlich D, Carmo-Fonseca M, Izaurralde E (2000) The C-terminal domain of TAP interacts with the nuclear pore complex and promotes export of specific CTE-bearing RNA substrates. RNA 6(1):136–158

Batisse J, Batisse C, Budd A, Bottcher B, Hurt E (2009) Purification of nuclear poly(A)-binding protein Nab2 reveals association with the yeast transcriptome and a messenger ribonucleoprotein core structure. J Biol Chem 284(50):34911–34917

Battich N, Stoeger T, Pelkmans L (2013) Image-based transcriptomics in thousands of single human cells at single-molecule resolution. Nat Methods 10(11):1127–1133

Beadle GW, Tatum EL (1941) Genetic control of biochemical reactions in neurospora. Proc Natl Acad Sci USA 27(11):499–506

Ben-Ari Y, Brody Y, Kinor N, Mor A, Tsukamoto T, Spector DL, Singer RH, Shav-Tal Y (2010) The life of an mRNA in space and time. J Cell Sci 123(Pt 10):1761–1774

Bertrand E, Chartrand P, Schaefer M, Shenoy SM, Singer RH, Long RM (1998) Localization of ASH1 mRNA particles in living yeast. Mol Cell 2(4):437–445

Beyer AL, Osheim YN (1988) Splice site selection, rate of splicing, and alternative splicing on nascent transcripts. Genes Dev 2(6):754–765

Bjork P, Wieslander L (2011) Nucleocytoplasmic mRNP export is an integral part of mRNP biogenesis. Chromosoma 120(1):23–38

Blobel G (1985) Gene gating: a hypothesis. Proc Natl Acad Sci USA 82(24):8527–8529

Boireau S, Maiuri P, Basyuk E, de la Mata M, Knezevich A, Pradet-Balade B, Backer V, Kornblihtt A, Marcello A, Bertrand E (2007) The transcriptional cycle of HIV-1 in real-time and live cells. J Cell Biol 179(2):291–304

Bothma JP, Garcia HG, Esposito E, Schlissel G, Gregor T, Levine M (2014) Dynamic regulation of eve stripe 2 expression reveals transcriptional bursts in living *Drosophila embryos*. Proc Natl Acad Sci USA 111(29):10598–10603

Brachet J, Chantrenne H (1951) Protein synthesis in nucleated and non-nucleated halves of *Acetabularia mediterranea* studied with carbon-14 dioxide. Nature 168(4283):950

Bratu DP, Cha BJ, Mhlanga MM, Kramer FR, Tyagi S (2003) Visualizing the distribution and transport of mRNAs in living cells. Proc Natl Acad Sci USA 100(23):13308–13313

Brenner S, Jacob F, Meselson M (1961) An unstable intermediate carrying information from genes to ribosomes for protein synthesis. Nature 190:576–581

Brody Y, Neufeld N, Bieberstein N, Causse SZ, Bohnlein EM, Neugebauer KM, Darzacq X, Shav-Tal Y (2011) The in vivo kinetics of RNA polymerase II elongation during co-transcriptional splicing. PLoS Biol 9(1):e1000573

Cabal GG, Genovesio A, Rodriguez-Navarro S, Zimmer C, Gadal O, Lesne A, Buc H, Feuerbach-Fournier F, Olivo-Marin JC, Hurt EC, Nehrbass U (2006) SAGA interacting factors confine sub-diffusion of transcribed genes to the nuclear envelope. Nature 441(7094):770–773

Calapez A, Pereira HM, Calado A, Braga J, Rino J, Carvalho C, Tavanez JP, Wahle E, Rosa AC, Carmo-Fonseca M (2002) The intranuclear mobility of messenger RNA binding proteins is ATP dependent and temperature sensitive. J Cell Biol 159(5):795–805

Casolari JM, Brown CR, Komili S, West J, Hieronymus H, Silver PA (2004) Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization. Cell 117(4):427–439

Cheng H, Dufu K, Lee CS, Hsu JL, Dias A, Reed R (2006) Human mRNA export machinery recruited to the 5′ end of mRNA. Cell 127(7):1389–1400

Chou YY, Heaton NS, Gao Q, Palese P, Singer RH, Lionnet T (2013) Colocalization of different influenza viral RNA segments in the cytoplasm before viral budding as shown by single-molecule sensitivity FISH analysis. PLoS Pathog 9(5):e1003358

Chubb JR, Trcek T, Shenoy SM, Singer RH (2006) Transcriptional pulsing of a developmental gene. Curr Biol 16(10):1018–1025

Coulon A, Ferguson ML, de Turris V, Palangat M, Chow CC, Larson DR (2014) Kinetic competition during the transcription cycle results in stochastic RNA processing. Elife 3

Daigle N, Ellenberg J (2007) LambdaN-GFP: an RNA reporter system for live-cell imaging. Nat Methods 4(8):633–636

Daneholt B (1997) A look at messenger RNP moving through the nuclear pore. Cell 88(5): 585–588

Daneholt B (1999) Pre-mRNP particles: from gene to nuclear pore. Curr Biol 9(11):R412–R415

Darzacq X, Shav-Tal Y, de Turris V, Brody Y, Shenoy SM, Phair RD, Singer RH (2007) In vivo dynamics of RNA polymerase II transcription. Nat Struct Mol Biol 14(9):796–806

Dirks RW, Daniel KC, Raap AK (1995) RNAs radiate from gene to cytoplasm as revealed by fluorescence in situ hybridization. J Cell Sci 108(Pt 7):2565–2572

Dreyfuss G (1986) Structure and function of nuclear and cytoplasmic ribonucleoprotein particles. Annu Rev Cell Biol 2:459–498

Dreyfuss G, Kim VN, Kataoka N (2002) Messenger-RNA-binding proteins and the messages they carry. Nat Rev Mol Cell Biol 3(3):195–205

Dworetzky SI, Feldherr CM (1988) Translocation of RNA-coated gold particles through the nuclear pores of oocytes. J Cell Biol 106(3):575–584

Fakan S (1994) Perichromatin fibrils are in situ forms of nascent transcripts. Trends Cell Biol 4(3): 86–90

Fakan S, Leser G, Martin TE (1986) Immunoelectron microscope visualization of nuclear ribonucleoprotein antigens within spread transcription complexes. J Cell Biol 103(4):1153–1157

Fay FS, Taneja KL, Shenoy S, Lifshitz L, Singer RH (1997) Quantitative digital analysis of diffuse and concentrated nuclear distributions of nascent transcripts, SC35 and poly(A). Exp Cell Res 231(1):27–37

Femino AM, Fay FS, Fogarty K, Singer RH (1998) Visualization of single RNA transcripts in situ. Science 280(5363):585–590

Femino AM, Fogarty K, Lifshitz LM, Carrington W, Singer RH (2003) Visualization of single molecules of mRNA in situ. Methods Enzymol 361:245–304

Gall JG, Pardue ML (1969) Formation and detection of RNA–DNA hybrid molecules in cytological preparations. Proc Natl Acad Sci USA 63(2):378–383

Gohring J, Jacak J, Barta A (2014) Imaging of endogenous messenger RNA splice variants in living cells reveals nuclear retention of transcripts inaccessible to nonsense-mediated decay in Arabidopsis. Plant Cell 26(2):754–764

Golding I, Cox EC (2004) RNA dynamics in live *Escherichia coli* cells. Proc Natl Acad Sci USA 101(31):11310–11315

Golding I, Paulsson J, Zawilski SM, Cox EC (2005) Real-time kinetics of gene activity in individual bacteria. Cell 123(6):1025–1036

Gros F, Hiatt H, Gilbert W, Kurland CG, Risebrough RW, Watson JD (1961) Unstable ribonucleic acid revealed by pulse labelling of *Escherichia coli*. Nature 190:581–585

Grunwald D, Singer RH (2010) In vivo imaging of labelled endogenous beta-actin mRNA during nucleocytoplasmic transport. Nature 467(7315):604–607

Haimovich G, Choder M, Singer RH, Trcek T (2013a) The fate of the messenger is pre-determined: a new model for regulation of gene expression. Biochim Biophys Acta 1829(6-7):643–653

Haimovich G, Medina DA, Causse SZ, Garber M, Millan-Zambrano G, Barkai O, Chavez S, Perez-Ortin JE, Darzacq X, Choder M (2013b) Gene expression is circular: factors for mRNA degradation also foster mRNA synthesis. Cell 153(5):1000–1011

Hall LL, Smith KP, Byron M, Lawrence JB (2006) Molecular anatomy of a speckle. Anat Rec A Discov Mol Cell Evol Biol 288(7):664–675

Hansen CH, van Oudenaarden A (2013) Allele-specific detection of single mRNA molecules in situ. Nat Methods 10(9):869–871

Herman RC, Williams JG, Penman S (1976) Message and non-message sequences adjacent to poly (A) in steady state heterogeneous nuclear RNA of HeLa cells. Cell 7(3):429–437

Hocine S, Raymond P, Zenklusen D, Chao JA, Singer RH (2013) Single-molecule analysis of gene expression using two-color RNA labeling in live yeast. Nat Methods 10(2):119–121

Hodge CA, Tran EJ, Noble KN, Alcazar-Roman AR, Ben-Yishay R, Scarcelli JJ, Folkmann AW, Shav-Tal Y, Wente SR, Cole CN (2011) The Dbp5 cycle at the nuclear pore complex during mRNA export I: dbp5 mutants with defects in RNA binding and ATP hydrolysis define key steps for Nup159 and Gle1. Genes Dev 25(10):1052–1064

Hosoda N, Lejeune F, Maquat LE (2006) Evidence that poly(A) binding protein C1 binds nuclear - pre-mRNA poly(A) tails. Mol Cell Biol 26(8):3085–3097

Hoyle NP, Ish-Horowicz D (2013) Transcript processing and export kinetics are rate-limiting steps in expressing vertebrate segmentation clock genes. Proc Natl Acad Sci USA 110(46): E4316–E4324

Huang S, Spector DL (1992) Will the real splicing sites please light up? Curr Biol 2(4):188–190

Huang S, Deerinck TJ, Ellisman MH, Spector DL (1994) In vivo analysis of the stability and transport of nuclear poly(A) + RNA. J Cell Biol 126(4):877–899

Hutchinson JN, Ensminger AW, Clemson CM, Lynch CR, Lawrence JB, Chess A (2007) A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. BMC Genomics 8:39

Iborra FJ, Jackson DA, Cook PR (2000) The path of RNA through nuclear pores: apparent entry from the sides into specialized pores. J Cell Sci 113(Pt 2):291–302

Ishihama Y, Funatsu T (2009) Single molecule tracking of quantum dot-labeled mRNAs in a cell nucleus. Biochem Biophys Res Commun 381(1):33–38

Itzkovitz S, Blat IC, Jacks T, Clevers H, van Oudenaarden A (2012) Optimality in the development of intestinal crypts. Cell 148(3):608–619

Janicki SM, Tsukamoto T, Salghetti SE, Tansey WP, Sachidanandam R, Prasanth KV, Ried T, Shav-Tal Y, Bertrand E, Singer RH, Spector DL (2004) From silencing to gene expression; real-time analysis in single cells. Cell 116(5):683–698

Jarmolowski A, Boelens WC, Izaurralde E, Mattaj IW (1994) Nuclear export of different classes of RNA is mediated by specific factors. J Cell Biol 124(5):627–635

Kataoka N, Yong J, Kim VN, Velazquez F, Perkinson RA, Wang F, Dreyfuss G (2000) Pre-mRNA splicing imprints mRNA in the nucleus with a novel RNA-binding protein that persists in the cytoplasm. Mol Cell 6(3):673–682

Kiseleva E, Goldberg MW, Daneholt B, Allen TD (1996) RNP export is mediated by structural reorganization of the nuclear pore basket. J Mol Biol 260(3):304–311

Kiseleva E, Goldberg MW, Allen TD, Akey CW (1998) Active nuclear pore complexes in Chironomus: visualization of transporter configurations related to mRNP export. J Cell Sci 111(Pt 2):223–236

Kubota T, Ikeda S, Yanagisawa H, Yuki M, Okamoto A (2010) Sets of RNA repeated tags and hybridization-sensitive fluorescent probes for distinct images of RNA in a living cell. PLoS One 5(9):e13003

Lamond AI, Spector DL (2003) Nuclear speckles: a model for nuclear organelles. Nat Rev Mol Cell Biol 4(8):605–612

Lawrence JB, Singer RH, Marselle LM (1989) Highly localized tracks of specific transcripts within interphase nuclei visualized by in situ hybridization. Cell 57(3):493–502

Le Hir H, Izaurralde E, Maquat LE, Moore MJ (2000a) The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon–exon junctions. EMBO J 19(24):6860–6869

Le Hir H, Moore MJ, Maquat LE (2000b) Pre-mRNA splicing alters mRNP composition: evidence for stable association of proteins at exon–exon junctions. Genes Dev 14(9):1098–1108

Le Hir H, Gatfield D, Izaurralde E, Moore MJ (2001) The exon–exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. EMBO J 20(17):4987–4997

Lee JH, Daugharthy ER, Scheiman J, Kalhor R, Yang JL, Ferrante TC, Terry R, Jeanty SS, Li C, Amamoto R, Peters DT, Turczyk BM, Marblestone AH, Inverso SA, Bernard A, Mali P, Rios X, Aach J, Church GM (2014) Highly multiplexed subcellular RNA sequencing in situ. Science 343(6177):1360–1363

Lejeune F, Ishigaki Y, Li X, Maquat LE (2002) The exon junction complex is detected on CBP80-bound but not eIF4E-bound mRNA in mammalian cells: dynamics of mRNP remodeling. EMBO J 21(13):3536–3545

Levsky JM, Singer RH (2003) Fluorescence in situ hybridization: past, present and future. J Cell Sci 116(Pt 14):2833–2838

Levsky JM, Shenoy SM, Pezo RC, Singer RH (2002) Single-cell gene expression profiling. Science 297(5582):836–840

Lichtenstein M, Guo W, Tartakoff AM (2001) Control of nuclear export of hnRNP A1. Traffic 2(4):261–267

Lionnet T, Czaplinski K, Darzacq X, Shav-Tal Y, Wells AL, Chao JA, Park HY, de Turris V, Lopez-Jones M, Singer RH (2011) A transgenic mouse for in vivo detection of endogenous labeled mRNA. Nat Methods 8(2):165–170

Luna R, Gaillard H, Gonzalez-Aguilera C, Aguilera A (2008) Biogenesis of mRNPs: integrating different processes in the eukaryotic nucleus. Chromosoma 117(4):319–331

Lund MK, Guthrie C (2005) The DEAD-box protein Dbp5p is required to dissociate Mex67p from exported mRNPs at the nuclear rim. Mol Cell 20(4):645–651

Luo MJ, Reed R (1999) Splicing is required for rapid and efficient mRNA export in metazoans. Proc Natl Acad Sci USA 96(26):14937–14942

Ma J, Liu Z, Michelotti N, Pitchiaya S, Veerapaneni R, Androsavich JR, Walter NG, Yang W (2013) High-resolution three-dimensional mapping of mRNA export through the nuclear pore. Nat Commun 4:2414

Martin RM, Rino J, Carvalho C, Kirchhausen T, Carmo-Fonseca M (2013) Live-cell visualization of pre-mRNA splicing with single-molecule sensitivity. Cell Rep 4(6):1144–1155

Matsumoto K, Tanaka KJ, Aoki K, Sameshima M, Tsujimoto M (2003) Visualization of the reconstituted FRGY2-mRNA complexes by electron microscopy. Biochem Biophys Res Commun 306(1):53–58

Mehlin H, Daneholt B, Skoglund U (1992) Translocation of a specific premessenger ribonucleoprotein particle through the nuclear pore studied with electron microscope tomography. Cell 69(4):605–613

Melcak I, Hoelz A, Blobel G (2007) Structure of Nup58/45 suggests flexible nuclear pore diameter by intermolecular sliding. Science 315(5819):1729–1732

Michael WM, Choi M, Dreyfuss G (1995) A nuclear export signal in hnRNP A1: a signal-mediated, temperature-dependent nuclear protein export pathway. Cell 83(3):415–422

Molenaar C, Abdulle A, Gena A, Tanke HJ, Dirks RW (2004) Poly(A) + RNAs roam the cell nucleus and pass through speckle domains in transcriptionally active and inactive cells. J Cell Biol 165(2):191–202

Mor A, Shav-Tal Y (2010) Dynamics and kinetics of nucleo-cytoplasmic mRNA export. Wiley Interdiscip Rev RNA 1(3):388–401

Mor A, Ben-Yishay R, Shav-Tal Y (2010a) On the right track: following the nucleo-cytoplasmic path of an mRNA. Nucleus 1(6):492–498

Mor A, Suliman S, Ben-Yishay R, Yunger S, Brody Y, Shav-Tal Y (2010b) Dynamics of single mRNP nucleocytoplasmic transport and export through the nuclear pore in living cells. Nat Cell Biol 12(6):543–552

Mor A, White MA, Fontoura BM (2014) Nuclear trafficking in health and disease. Curr Opin Cell Biol 28:28–35

Muller-McNicoll M, Neugebauer KM (2013) How cells get the message: dynamic assembly and function of mRNA-protein complexes. Nat Rev Genet 14(4):275–287

Muramoto T, Cannon D, Gierlinski M, Corrigan A, Barton GJ, Chubb JR (2012) Live imaging of nascent RNA dynamics reveals distinct types of transcriptional pulse regulation. Proc Natl Acad Sci USA 109(19):7350–7355

Neugebauer KM, Roth MB (1997) Distribution of pre-mRNA splicing factors at sites of RNA polymerase II transcription. Genes Dev 11(9):1148–1159

Osheim YN, Miller OL Jr, Beyer AL (1985) RNP particles at splice junction sequences on *Drosophila chorion* transcripts. Cell 43(1):143–151

Palazzo AF, Springer M, Shibata Y, Lee CS, Dias AP, Rapoport TA (2007) The signal sequence coding region promotes nuclear export of mRNA. PLoS Biol 5(12):e322

Pante N, Jarmolowski A, Izaurralde E, Sauder U, Baschong W, Mattaj IW (1997) Visualizing nuclear export of different classes of RNA by electron microscopy. RNA 3(5):498–513

Pardue ML, Gall JG (1969) Molecular hybridization of radioactive DNA to the DNA of cytological preparations. Proc Natl Acad Sci USA 64(2):600–604

Pardue ML, Gall JG (1970) Chromosomal localization of mouse satellite DNA. Science 168(3937):1356–1358

Pasquinelli AE, Powers MA, Lund E, Forbes D, Dahlberg JE (1997) Inhibition of mRNA export in vertebrate cells by nuclear export signal conjugates. Proc Natl Acad Sci USA 94(26): 14394–14399

Perry RP, Kelley DE, LaTorre J (1974) Synthesis and turnover of nuclear and cytoplasmic polyadenylic acid in mouse L cells. J Mol Biol 82(3):315–331

Pinol-Roma S, Dreyfuss G (1992) Shuttling of pre-mRNA binding proteins between nucleus and cytoplasm. Nature 355(6362):730–732

Pitchiaya S, Heinicke LA, Custer TC, Walter NG (2014) Single molecule fluorescence approaches shed light on intracellular RNAs. Chem Rev 114(6):3224–3265

Politz JC, Taneja KL, Singer RH (1995) Characterization of hybridization between synthetic oligodeoxynucleotides and RNA in living cells. Nucleic Acids Res 23(24):4946–4953

Politz JC, Browne ES, Wolf DE, Pederson T (1998) Intranuclear diffusion and hybridization state of oligonucleotides measured by fluorescence correlation spectroscopy in living cells. Proc Natl Acad Sci USA 95(11):6043–6048

Politz JC, Tuft RA, Pederson T, Singer RH (1999) Movement of nuclear poly(A) RNA throughout the interchromatin space in living cells. Curr Biol 9(6):285–291

Politz JC, Tuft RA, Prasanth KV, Baudendistel N, Fogarty KE, Lifshitz LM, Langowski J, Spector DL, Pederson T (2006) Rapid, diffusional shuttling of poly(A) RNA between nuclear speckles and the nucleoplasm. Mol Biol Cell 17(3):1239–1249

Puvion-Dutilleul F, Puvion E (1981) Relationship between chromatin and perichromatin granules in cadmium-treated isolated hepatocytes. J Ultrastruct Res 74(3):341–350

Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S (2006) Stochastic mRNA synthesis in mammalian cells. PLoS Biol 4(10):e309

Raj A, van den Bogaard P, Rifkin SA, van Oudenaarden A, Tyagi S (2008) Imaging individual mRNA molecules using multiple singly labeled probes. Nat Methods 5(10):877–879

Ross AF, Oleynikov Y, Kislauskis EH, Taneja KL, Singer RH (1997) Characterization of a beta-actin mRNA zipcode-binding protein. Mol Cell Biol 17(4):2158–2165

Roussel MR, Tang T (2012) Simulation of mRNA diffusion in the nuclear environment. IET Syst Biol 6(4):125–133

Santangelo PJ, Lifland AW, Curt P, Sasaki Y, Bassell GJ, Lindquist ME, Crowe JE Jr (2009) Single molecule-sensitive probes for imaging RNA in live cells. Nat Methods 6(5):347–349

Schonberger J, Hammes UZ, Dresselhaus T (2012) In vivo visualization of RNA in plants cells using the lambdaN(2)(2) system and a GATEWAY-compatible vector series for candidate RNAs. Plant J 71(1):173–181. doi:10.1111/j.1365-313X.2012.04923.x

Shav-Tal Y, Gruenbaum Y (2009) Single-molecule dynamics of nuclear mRNA. F1000 Biol Rep 1:29–32

Shav-Tal Y, Singer RH (2005) RNA localization. J Cell Sci 118(Pt 18):4077–4081

Shav-Tal Y, Darzacq X, Shenoy SM, Fusco D, Janicki SM, Spector DL, Singer RH (2004a) Dynamics of single mRNPs in nuclei of living cells. Science 304(5678):1797–1800

Shav-Tal Y, Shenoy SM, Singer RH (eds) (2004b) Visualization and quantification of single RNA molecules in living cells. Live cell imaging: a laboratory manual. Cold Spring Harbor Laboratory Press, New York, NY

Sheinberger J, Shav-Tal Y (2013) The dynamic pathway of nuclear RNA in eukaryotes. Nucleus 4(3):195–205

Siebrasse JP, Veith R, Dobay A, Leonhardt H, Daneholt B, Kubitscheck U (2008) Discontinuous movement of mRNP particles in nucleoplasmic regions devoid of chromatin. Proc Natl Acad Sci USA 105(51):20291–20296

Siebrasse JP, Kaminski T, Kubitscheck U (2012) Nuclear export of single native mRNA molecules observed by light sheet fluorescence microscopy. Proc Natl Acad Sci USA 109(24): 9426–9431. doi:10.1073/pnas.1201781109

Singh OP, Bjorkroth B, Masich S, Wieslander L, Daneholt B (1999) The intranuclear movement of Balbiani ring premessenger ribonucleoprotein particles. Exp Cell Res 251(1):135–146

Skabkin MA, Kiselyova OI, Chernov KG, Sorokin AV, Dubrovin EV, Yaminsky IV, Vasiliev VD, Ovchinnikov LP (2004) Structural organization of mRNA complexes with major core mRNP protein YB-1. Nucleic Acids Res 32(18):5621–5635

Skoglund U, Andersson K, Bjorkroth B, Lamb MM, Daneholt B (1983) Visualization of the formation and transport of a specific hnRNP particle. Cell 34(3):847–855

Skoglund U, Andersson K, Strandberg B, Daneholt B (1986) Three-dimensional structure of a specific pre-messenger RNP particle established by electron microscope tomography. Nature 319(6054):560–564

Snaar SP, Verdijk P, Tanke HJ, Dirks RW (2002) Kinetics of HCMV immediate early mRNA expression in stably transfected fibroblasts. J Cell Sci 115(Pt 2):321–328

Soop T, Ivarsson B, Bjorkroth B, Fomproix N, Masich S, Cordes VC, Daneholt B (2005) Nup153 affects entry of messenger and ribosomal ribonucleoproteins into the nuclear basket during export. Mol Biol Cell 16(12):5610–5620

Spector DL (1993) Macromolecular domains within the cell nucleus. Annu Rev Cell Biol 9: 265–315

Speese SD, Ashley J, Jokhi V, Nunnari J, Barria R, Li Y, Ataman B, Koon A, Chang YT, Li Q, Moore MJ, Budnik V (2012) Nuclear envelope budding enables large ribonucleoprotein particle export during synaptic Wnt signaling. Cell 149(4):832–846

Sperling R, Sperling J (1990) Large nuclear ribonucleoprotein particles of specific RNA polymerase II transcripts. In: Strauss PR, Wilson SH (eds) The eukaryotic nucleus: molecular biochemistry and macromolecular assemblies, vol 2. The Telford Press, New Jersey, pp 453–476

Stevens BJ, Swift H (1966) RNA transport from nucleus to cytoplasm in Chironomus salivary glands. J Cell Biol 31(1):55–77

Stewart M (2007) Ratcheting mRNA out of the nucleus. Mol Cell 25(3):327–330

Thieffry D, Burian RM (1996) Jean Brachet's alternative scheme for protein synthesis. Trends Biochem Sci 21(3):114–117

Thompson MA, Casolari JM, Badieirostami M, Brown PO, Moerner WE (2010) Three-dimensional tracking of single mRNA particles in *Saccharomyces cerevisiae* using a double-helix point spread function. Proc Natl Acad Sci USA 107(42):17864–17871

Tutucci E, Stutz F (2011) Keeping mRNPs in check during assembly and nuclear export. Nat Rev Mol Cell Biol 12(6):377–384

Tyagi S (2009) Imaging intracellular RNA distribution and dynamics in living cells. Nat Methods 6(5):331–338

Valencia P, Dias AP, Reed R (2008) Splicing promotes rapid and efficient mRNA export in mammalian cells. Proc Natl Acad Sci USA 105(9):3386–3391

Vargas DY, Raj A, Marras SA, Kramer FR, Tyagi S (2005) Mechanism of mRNA transport in the nucleus. Proc Natl Acad Sci USA 102(47):17008–17013

Vargas DY, Shah K, Batish M, Levandoski M, Sinha S, Marras SA, Schedl P, Tyagi S (2011) Single-molecule imaging of transcriptionally coupled and uncoupled splicing. Cell 147(5): 1054–1065

Verschure PJ, van Der Kraan I, Manders EM, van Driel R (1999) Spatial relationship between transcription sites and chromosome territories. J Cell Biol 147(1):13–24

Visa N, Alzhanova-Ericsson AT, Sun X, Kiseleva E, Bjorkroth B, Wurtz T, Daneholt B (1996) A pre-mRNA-binding protein accompanies the RNA from the gene through the nuclear pores and into polysomes. Cell 84(2):253–264

Waks Z, Klein AM, Silver PA (2011) Cell-to-cell variability of alternative RNA splicing. Mol Syst Biol 7:506

Wang W, Cui ZQ, Han H, Zhang ZP, Wei HP, Zhou YF, Chen Z, Zhang XE (2008) Imaging and characterizing influenza A virus mRNA transport in living cells. Nucleic Acids Res 36(15): 4913–4928

Watanabe M, Fukuda M, Yoshida M, Yanagida M, Nishida E (1999) Involvement of CRM1, a nuclear export receptor, in mRNA export in mammalian cells and fission yeast. Genes Cells 4(5):291–297

Watson JD (2011) Prologue to the first edition of The RNA World. In: Atkins JF, Gesteland R, Cech T (eds) RNA worlds. Cold Spring Harbor Laboratory Press, New York, NY, p xi

Wu WW, Weaver LL, Pante N (2007) Ultrastructural analysis of the nuclear localization sequences on influenza A ribonucleoprotein complexes. J Mol Biol 374(4):910–916

Yunger S, Rosenfeld L, Garini Y, Shav-Tal Y (2010) Single-allele analysis of transcription kinetics in living mammalian cells. Nat Methods 7(8):631–633

Zachar Z, Kramer J, Mims IP, Bingham PM (1993) Evidence for channeled diffusion of pre-mRNAs during nuclear RNA transport in metazoans. J Cell Biol 121(4):729–742

Zenklusen D, Larson DR, Singer RH (2008) Single-RNA counting reveals alternative modes of gene expression in yeast. Nat Struct Mol Biol 15(12):1263–1271

Zhao J, Jin SB, Bjorkroth B, Wieslander L, Daneholt B (2002) The mRNA export factor Dbp5 is associated with Balbiani ring mRNP from gene to cytoplasm. EMBO J 21(5):1177–1187

Zirbel RM, Mathieu UR, Kurz A, Cremer T, Lichter P (1993) Evidence for a nuclear compartment of transcription and splicing located at chromosome domain boundaries. Chromosome Res 1(2):93–106