Chapter 1

Introduction

The nucleus is a highly complex organelle responsible for the faithful replication and maintenance of the DNA template and regulation its transcriptional products. To achieve this, the contents of the nucleus are organised into compartments specialising in processes such as transcription and replication. Abnormalities in the organisation of the nucleus are often associated with diseases such as cancer. In this chapter I will discuss the current understanding of nuclear structure and organisation and how it may be involved in the initiation of oncogenic chromosomal translocations.

1.1 Chromatin

Chromatin is a general term used to describe DNA packaged around histone proteins to form nucleosomes, and the plethora of additional proteins that bind to them both. There are two principle forms of chromatin: euchromatin and heterochromatin. These were first described in the early twentieth century due to their differential staining within the nucleus (Heitz, 1928). Heterochromatin stains darkly because it remains highly condensed during interphase and often associated with the nuclear periphery. Constitutive heterochromatin consists of repetitive elements found in centromeres and telomeres, it plays a structural role within the genome and is highly compacted during interphase. Facultative heterochromatin is less compact and consists of inactive chromatin that can vary between cell types as they differentiate. Heterochromatin is well known as being a repressive environment for gene expression. An example of this is the position-effect variegation seen in *Drosophila melanogaster* - a chromosomal inversion in the X chromosome characterised in the early 1930s can place the *white* gene close to pericentric heterochromatin, leading to the spreading of heterochromatin marks which silences the *white* gene and results in a change in eye colour (Vogel et al., 2009). In contrast, euchromatin is the site of most genic transcription (Chesterton et al., 1974), its looser compaction allowing cellular machinery access to the DNA, enabling the binding of transcription factors and the initiation of transcription.

The differences between heterochromatin and euchromatin lie within the proteins that they contain. Chromatin acts as a platform for proteins to bind to. Differences in histone modifications, histone variants, nucleosome packing and DNA modifications affect the accessibility and binding profile of the chromatin, and therefore affect how the DNA is interpreted by the cell.

1.1.1 Histones

To package DNA, the double helix is wrapped around an octomer of core histones: two H2A, two H2B, two H3 and two H4. 146 base pairs of DNA interact with these positively charged proteins to form the nucleosome, which is then bound by histone H1 with linker DNA to make a total of 166 base pairs (Fig 1.1, Davey et al., 2002). This packing forms the 10 nm fibre, often known as 'beads on a string' due to its appearance in electron micrographs. At their most basic level, histones function to compact DNA by counteracting the negative charge of the phosphorylated back bone.

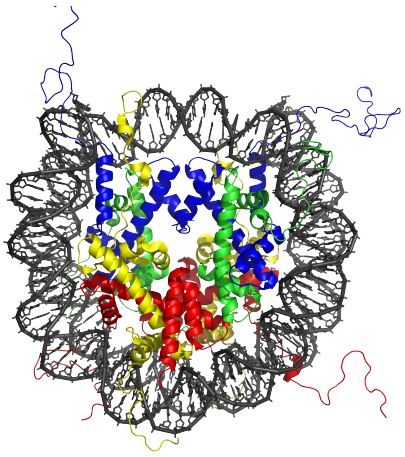

Figure 1.1 – Structure of the nucleosome. X-ray structure of a nucleosome core particle at a resolution of 1.9Å. DNA can be seen wrapped around the core histones, which have flexible tails extending into the nuclear matrix. PDB structure 1KX5 (Davey et al., 2002).

1.1.2 Histone modifications

Core histones have flexible amino-terminal tails which extend outside of the nucleosome, and are accessible to proteins within the nucleoplasm. These tails can be post-translationally modified at a large number of residues - lysine (methylation, acetylation, ubiquitination, sumoylation, ADP-Ribosylation), arginine (methylation) as well as serine and threonine (phosphorylation). These modifications can affect the packing of chromatin (Wolffe and Hayes, 1999) as well as which proteins can bind. The large number of combinatorial possibilities that result from these modifications have been dubbed the 'Histone Code' (Strahl and Allis, 2000). Different modifications are related to different chromatin states. For example, active promoters are typically enriched for di- and tri-methylation of histone 3 lysine 4 (H3K4) whilst inactive promoters are enriched for trimethylation at lysine residues 27 (H3K27me3) and 9 (H3K9me3) (See table 1.1 for a summary. For review, see Zhou et al., 2011).

As chromatin immunoprecipitation (ChIP) has become a common laboratory technique, combined with microarray techniques (ChIP on chip) and next generation sequencing (ChIP-Seq), our understanding of how histone modifications affect chromatin biology on a genome-wide scale has advanced dramatically. Profiling chromatin types using multiple datasets covering a large number of histone modifications is sufficient to predict the identity and function of regions

Modification	Histone	Residue	Effects on transcription
Acetylation	H2A	K5	Activation
	H2B	K5, K12, K15, K20	Activation
	H3	K4, K14, K18, K23, K27	Activation
		K9	Histone deposition
	H4	K5, K12	Histone deposition
		K8, K16	Activation
Methylation	H3	K4, K79	Euchromatin
		K9, K27	Silencing
		R17	Activation
		K36	Elongation
	H4	R3	Activation
		K20	Silencing
Phosphorylation	H2A	S1, T119	Mitosis
	H2AX	S139	DNA repair
	H3	T3, S10, T11, S28	Mitosis
	H4	S1	Mitosis
Ubiquitination	H2A	K119	Silencing
	H2B	K120	Activation

Table 1.1 – Summary of known mammalian histone modifications. H, histone; K, lysine; R, arginine; S, serine; T, threonine. Adapted from Sadri-Vakili and Cha, 2006.

in the genome with a high degree of accuracy, revealing previously unknown enhancers (Heintzman et al., 2007; Ernst and Kellis, 2010; Hon et al., 2009). Ernst *et al.* used the genome-wide profiles of nine histone modifications in nine different cell types to define fifteen chromatin states, including promoters, enhancers, insulators and transcribed regions (Ernst et al., 2011). They integrated data from genome-wide association studies (GWAS) and found numerous enhancer elements that coincide with disease associated mutations. Such genome-wide approaches can reveal the dynamics across different cell types and are powerful tools in understanding how the genome is interpreted in health and disease (see Section **??**).

1.1.3 Histone variants

In addition to histone tail modifications, chromatin can be modified by the incorporation of histone variants. Canonical core histone genes are found in clustered repeat arrays within the genome, are transcribed during replication and are highly conserved between species. Histone variants are found as single genes spread through the genome and are subject to far greater diversity (Talbert and Henikoff, 2010).

CENP-A is a human variant of histone H3 which replaces the canonical histone in centromeric heterochromatin. It is a key factor in the establishment of the centromeres and kinetochores required for mitosis. The histone variant is incorporated with the help of a number of chaperone proteins, including HJURP, after replication of DNA has finished (Dunleavy et al., 2009; Foltz et al., 2009). CENP-A is essential for the formation of centromeres.

Another frequent histone variant found in humans is H3.3, which differs from canonical H3 by just four amino

acids (Talbert and Henikoff, 2010). This histone variant is found within transcribed genes, promoters and regulatory elements, and is thought to be laid down during transcriptional elongation (Schwartz and Ahmad, 2005). Nucleosomes containing H3.3 appear to be less stable than canonical nucleosomes, with a high turnover (Schwartz and Ahmad, 2005). It is possible that this increased turnover of the nucleosomal components helps to keep the chromatin open and accessible to the transcriptional machinery (Talbert and Henikoff, 2010).

Other core histone proteins also have variants, such as H2A.Z, a histone variant located on either side of the nucleosome free regions found at the transcriptional start sites of active genes as well as insulator regions (Zlatanova and Thakar, 2008). H2A.Z is able to promote the recruitment of RNA polymerase II to certain regions, such as the yeast *GAL1-10* genes, by mediating C-terminal interactions with the transcriptional machinery (Adam et al., 2001).

1.2 Two-dimensional organisation

Since the inception of cytogenetics, it has been known that not all chromosomes are the same. The development of chromosome banding techniques in the 1970s allowed detailed human karyotypes to be determined, complete with differential staining of regions within each chromosome (Caspersson et al., 1970). Banding assays can show regions of heterochromatin and euchromatin, highlighting the variation in characteristics across regions of the genome (Trask, 2002). In prokaryotes, genes are often found in cistrons and can be coexpressed in single polycistronic mRNAs. This type of linear organisation is not present in most higher eukaryotes, with the notable exception of the nematode worm *Caenorhabditis elegans* (Blumenthal et al., 2002), though gene clusters resulting from tandem duplication are frequently found throughout mammalian genomes. Some specific examples of two-dimensional clustering have been shown: testes-specific genes in *Drosophila melanogaster* have been found in clusters more frequently than would be expected by chance (Boutanaev et al., 2002) and genes sharing transcription factors can be found in clusters in the yeast *Saccharomyces cerevisiae* (Janga et al., 2008).

The sequencing of the human genome allowed detailed analysis of GC content, gene density and repetitive sequence content (Lander et al., 2001). Versteeg *et al.* integrated a multitude of SAGE tag expression profiles from different cell types into the genomic map and built on earlier work defined regions of high transcriptional activity, called ridges (Versteeg et al., 2003; Caron et al., 2001). They found ridges to be gene-dense, highly transcribed, have a high GC content and low LINE repeat density. These features are based purely on the underlying sequence content and so do not vary amongst cell types. The different ridges and anti-ridges were found to contain different classes of genes, with weakly expressed genes clustering within anti-ridges and clusters of highly expressed housekeeping genes found predominantly in ridges (Versteeg et al., 2003).

Whilst the two-dimensional organisation of the genome cannot completely explain the degree of complexity found within the transcriptome, these studies demonstrate that the order of sequence within the genome is not entirely random

and can affect the transcriptional control of genes.

1.3 Chromosome Territories

The interphase nucleus is a highly structured organelle. As chromosomes decondense after metaphase they retain some degree of structure, forming "chromosome territories" (CTs) (Cremer and Cremer, 2001). Circumstantial evidence for interphase organisation of chromosomes has existed for a long time, first suggested by Carl Rabl in 1885 (Rabl, 1885). Oberservations by Stack *et al.* using microscopy with giemsa-band staining suggested that chromosomes retained some degree of organisation during interphase (Stack et al., 1977), and in 1982 Cremer *et al.,* showed that interphase chromosomes occupy territories by studying the pattern of DNA damage in metaphase chromosomes after spot irradiation during interphase (Cremer et al., 1982). The subsequent development of chromosome paints, a method to visualise entire or part-chromosomes with fluorescence *in-situ* hybridisation (FISH), confirmed these findings (Schardin et al., 1985; Manuelidis, 1985; Bolzer et al., 2005).

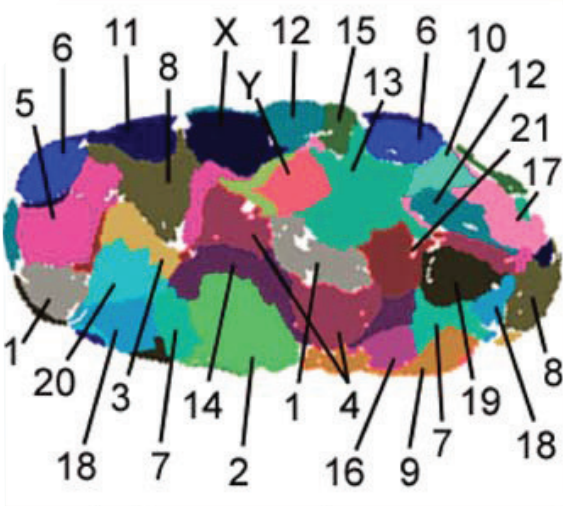

Figure 1.2 – Chromosome Territories. Image of a Human fibroblast cell in G0 with all chromosomes labelled using multi-colour FISH. Adapted from Bolzer *et al.* (Bolzer et al., 2005)

As FISH techniques have developed, so too has the detail with which CT organisation can be studied. Several groups have shown that chromosome territory position within the nucleus is not random and correlates with chromosome size (Sun et al., 2000; Cremer et al., 2001; Bolzer et al., 2005), gene-density (Cremer et al., 2001; Croft et al., 1999) and replication timing (Ferreira et al., 1997; Visser et al., 1998) - those near the centre of the nucleus tend to be gene-rich, early replicating and small. CT positioning appears to be conserved through evolution (Tanabe et al., 2002) and is cell type specific (Parada et al., 2004; Kuroda et al., 2004).

1.3.1 CTs and transcription

It is known that for some but not all genes, positioning at the nuclear periphery correlates with reduced gene expression (Kosak et al., 2002; Dietzel et al., 2004; Zink et al., 2004). To investigate whether nuclear positioning can cause changes in transcriptional activity or is simply a consequence, three groups published studies which artificially tethered genomic regions to the inner nuclear membrane using *lac* operators *(lacO)* (Finlan et al., 2008; Reddy et al., 2008; Kumaran and Spector, 2008). Finlan *et al.* and Reddy *et al.* both observed a decrease in the transcriptional activity of the

regions when tethered (Finlan et al., 2008; Reddy et al., 2008), an effect that was ablated when cells were treated with trichostatin A (TSA) to inhibit class I and class II histone deacetylases (Finlan et al., 2008). It should be noted that this effect does not appear to apply to all loci (Kumaran and Spector, 2008). Whilst these studies suggest that it is nuclear positioning that leads to transcriptional effects, a study by Croft *et al.* showed that inhibition of transcription causes a reversible change in CT position (Croft et al., 1999) and Bridger *et al.* showed a difference in CT positioning between proliferating and senescent human fibroblasts (Bridger et al., 2000). Large scale rearrangements of CTs have also been observed during cell differentiation (Stadler et al., 2004; Szczerbal et al., 2009), likely due to changes in transcriptional profiles and chromatin remodelling.

This evidence suggests that there is a dynamic interplay between CT positioning and transcription - gross transcriptional patterns may drive the position of chromosomes within the nucleus, and those positions may in turn affect the transcription of the genes they contain.

1.3.2 Chromosome territory dynamics

The movement of chromatin loci within the nucleus is known to be largely due to local Brownian motion, or "constrained diffusion" and seems to be limited by attachment to a nuclear cytoskeleton, nucleoli and the inner nuclear membrane (Marshall et al., 1997; Chubb et al., 2002), though some studies have shown a mixture of local diffusion and larger, active movements (Vazquez et al., 2001). Chuang *et al.* studied Chinese hamster ovary (CHO) cells after stimulation with a transcriptional activator and analysed the movement of a reporter locus from the nuclear periphery to the interior of the nucleus (Chuang et al., 2006). They found that chromosome movements happened in rapid unidirectional bursts, suggesting an active mechanism. This was supported by a later study in human fibroblast cells by Mehta *et al.* who demonstrated chromosome movement only 15 minutes after serum starvation, a process that was ablated by the inhibition of actin polymerisation or myosin activity (Mehta et al., 2010).

1.3.3 Chromatin decondensation and the inter-chromosomal space

In addition to the study of whole chromosome positioning, there has been a great deal of research into the positioning of individual sequences relative to their chromosome territory. Early studies suggested that transcribed genes were found at the surface of chromosome territories (Zirbel et al., 1993) leading to a model whereby transcriptionally inactive genes are buried within territories and expressed genes are able to contact transcriptional machinery in an inter-chromosome domain (ICD) (Cremer et al., 1993). This model gained support due to studies showing genes at the periphery of CTs (Kurz et al., 1996) and new techniques to visualise the ICD using microscopy (Bridger et al., 1998). A number of FISH studies showed genes moving away from their territories in large loops upon activation (Volpi et al., 2000; Mahy et al., 2002a; Chambeyron and Bickmore, 2004), suggesting that they may be recruited to the ICD for transcription.

Despite these case studies, looping is not a prerequisite for transcription; DNA-FISH studies have shown transcription within the volume of chromosome territories (Verschure et al., 1999; Mahy et al., 2002b). Osborne *et al.* showed that the actively transcribed *Uros* gene is more frequently outside the CT than the inactive gene *Fgfr2*, though this position was not necessary for transcription, suggesting that actively transcribed genes may preferentially locate to the surface of CTs, but that this alone is not sufficient to drive transcription (Osborne et al., 2004). As such, it has been suggested that the inter-chromosomal domain model should be renamed the inter-*chromatin* domain model, whereby chromosomes are invaginated with channels and subdivided into ~1 Mbp domains of chromatin (Cremer and Cremer, 2001).

1.3.4 Chromosome territory intermingling

A question that followed immediately from the discovery of loops extending from chromosome territories was that of chromosome intermingling. Mathematical modelling approaches capable of predicting intermingling volumes correlated with known rates irradiation induced DNA damage (Holley et al., 2002; Hlatky et al., 2002). This was later backed up with a study by Branco and Pombo, who studied thin cryosections of nuclei with high resolution light and electron microscopy; they found that there is significant intermingling of chromatin between chromosome territories, and that the extent of this intermingling correlated strongly with previously recorded irradiation induced translocation frequencies (Branco and Pombo, 2006). The degree of intermingling changed significantly for three chromosome pairs after transcription inhibition with α-amanitin, suggesting a role for specific transcription interactions in the organisation of the nucleus (Branco and Pombo, 2006).

In support of chromosome intermingling, the *HoxB* extra-chromosomal loops found to extend from the chromosome territory upon gene activation (Chambeyron and Bickmore, 2004) were found to make increased *trans* chromosomal interactions whilst looping out (Würtele and Chartrand, 2006), suggesting that the loops contact other chromosomes rather than occupying an empty inter-chromosomal space. A large number of inter-chromosomal contacts have been detected by recent genome-wide chromosome conformation studies, supporting the presence of chromosome intermingling (Lieberman-Aiden et al., 2009).

1.4 Nuclear compartmentalisation

Because the nuclear interior is devoid of membrane bound structures its organisation is defined by a dynamic equilibrium - a sum product of the many processes and requirements involved in nuclear biology. In addition to the non-random positioning of chromosomes, a number of proteins are also found in aggregates typically referred to as subnuclear compartments. The clustering of proteins into compartments increases the efficiency of biochemical processes and is predicted by the principle of molecular crowding (reviewed in Cook, 2002) and is a key feature in the organisation of the nucleus.

1.4.1 Transcription factories

One class of nuclear subcompartment which has come to light within the past thirty years is the transcription factory, seen as foci of hyper-phosphorylated RNA polymerase II scattered throughout the nucleus. The majority of genic transcription appears to take place at transcription factories (Jackson et al., 1993; Osborne et al., 2004; Eskiw et al., 2008), challenging the classical model of transcription found in many text books (discussed below).

The first study to suggest that eukaryotic transcription does not take place with a processive RNA polymerase moving along a DNA template was by Jackson, McCreedy and Cook in 1981. They showed that nascent RNA transcripts labelled with [^{3}H] uridine remained within the nucleus when loops of DNA were removed using a nuclease (Jackson et al., 1981). They went on to show that RNA polymerase II and active genes were also resistant to elution after chromatin digestion (Jackson and Cook, 1985). The term 'transcription factories' was coined by Jackson *et al.* in 1993. Fluorescence microscopy was used to label the incorporation of bromouridine triphosphate (BrUTP) into nascent RNA; discrete foci of nascent transcription could then be seen within the nucleus which did not form in the presence of the RNA polymerase II inhibitor α-amanitin (Jackson et al., 1993). Further studies showed that these foci contained RNA polymerase II along with many other components required for transcription (Iborra et al., 1996; Grande et al., 1997). An ultrastructural study by Eskiw *et al.* used correlative microscopy with both electron spectroscopic imaging (ESI) and fluorescence microscopy to study nuclei sections. ESI can distinguish nitrogen and phosphorous atoms without labelling, and fluorescence light microscopy can visualise transcription through labelling BrUTP in nascent transcripts. The authors found that nascent RNA is almost always associated with the surface of large nitrogen-rich protein structures with a diameter of ~87 nm, comparable in size to that predicted for a transcription factory (Eskiw et al., 2008).

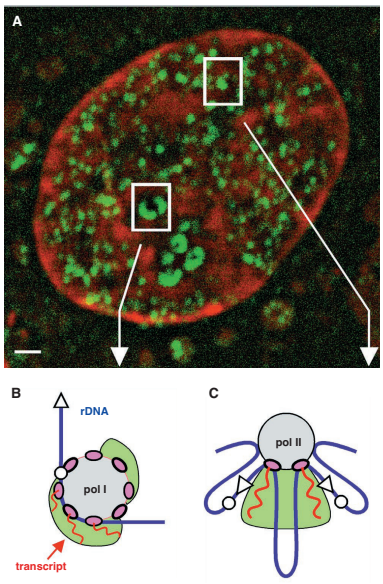

Figure 1.3 – Nascent RNA and transcription factories. (A) Transcription foci in HeLa cells, visualised with labelled Br-UTP in 100 nm cryosections. Nascent RNA (green) is concentrated in punctate foci. (B) Model for a nucleolar factory, showing a transcript with multiple polymerases generating a crescent shaped focus. (C) Model for a nucleoplasmic factory. Multiple transcribed regions each with a single polymerase generate a smaller cloud of nascent transcripts. Adapted from Cook *et al.*, Science (2009) (Cook, 1999).

The discovery of transcription factories has demanded a new model for the action of RNA polymerase II (Cook, 1999). The revised model proposes that instead of RNA polymerase II freely diffusing to active genes and tracking along the gene body, genes are recruited to transcription factories and are pulled through a stationary polymerase. Such a model provides a better explanation for the mechanics of transcription - clustering of transcriptional activity may

enable the cell to conduct transcription in a much more efficient manner; HeLa nuclei have a 1 μM concentration of active RNA polymerase II, whereas the local concentration within transcription factories is closer to 1 mM (Cook, 2002). Additionally, a polymerase enzyme moving along a gene would rotate with the helix of the DNA wrapping the nascent transcript around the template. Genes pulled through static transcription factories would extrude their RNA transcripts into the nucleoplasm (Iborra et al., 1996; Cook, 1999), creating topological loops within the template DNA which may be removed through the activity of topoisomerases.

Jackson *et al.* went on to work on a quantitative analysis of transcription factories in HeLa cells calculating the number of active RNA polymerases, the number of transcription sites and the number of polymerases associated with each transcriptional unit (Jackson et al., 1998). They showed that each HeLa cell nucleus contain approximately 2400 transcription factories, each with approximately 30 active RNA polymerase II complexes (Jackson et al., 1998). Importantly, this study showed that there are more transcribing units than there are foci of transcription, suggesting that genes must colocalise to transcribe. The number of transcription factories varies a great deal amongst cell types, but the observation that genes colocalise within transcription factories has been confirmed by a number of different techniques. Osborne *et al.* showed that genes situated within transcription factories are actively transcribed, whereas those outside are not (Osborne et al., 2004). They went on to demonstrate that transcription is a discontinuous process with the frequency of nascent RNA transcription foci related to primary transcript RNA concentrations, suggesting that transcription occurs in bursts. Multiple genes both in *cis* and in *trans* were seen to dynamically colocalise in transcription factories, supporting predictions that genes must share transcription factories (Osborne et al., 2004; Jackson et al., 1998). A later paper by Osborne *et al.* showed that the immediate-early genes *Myc* and *Fos* are dynamically recruited to existing transcription factories within five minutes of B-cell stimulation, suggesting that the recruitment of genes to pre-existing transcription factories may be a method of transcriptional control (Osborne et al., 2007).

It is worth noting that the concept of genes being recruited to immobile transcription factories is not universally accepted, with some doubt over the resolution achievable by FISH and 3C studies (reviewed in Sutherland and Bickmore, 2009).

1.4.1.1 Specialised transcription factories

After the demonstration that genes share transcription factories, a number of groups postulated that specific genes may colocalise at a subset of transcription factories. The self-organising principle applied to transcription factories would predict this; just as the local concentration of active RNA polymerase II is elevated by the clustering of transcribing units, the local concentration of transcription factors that bind to those transcribed units will also be elevated (Cook, 2002). It is thought that genes diffusing to a transcription factory already engaged with other genes that share the same factors are more likely to engage and be transcribed themselves (Bartlett et al., 2006). Indeed, genes regulated

by common transcription factors appear to cluster within specific chromosomes in yeast (Janga et al., 2008) and testis-specific genes are found clustered in *Drosophila* (Boutanaev et al., 2002) - genes clustered in linear sequence are more likely to associate in three dimensions.

Some evidence that such specialised transcription factories may exist within mammalian cells came from Osborne *et al.* (Osborne et al., 2007). *Myc* and *Igh* are commonly translocated in Burkitt's lymphoma and mouse plasmacytoma. They showed that approximately one quarter of actively transcribing *Myc* alleles shared a transcription factory with *Igh*, over double the rate of colocalisation found with the control genes tested. Using DNA-FISH they showed an overall reduction in spacing between *Myc* and *Igh* alleles upon B cell stimulation, suggesting that *Myc* alleles are being specifically recruited to transcription factories containing transcribing *Igh* (Osborne et al., 2007). Xu and Cook later demonstrated that the transcription of plasmids transfected into cells clustered together at a handful of transcription factories. The plasmids were generated with one of four promoter types, one of three genes and one of three 3' regions and were found to segregate according to their promoter and the presence of an intron (Xu and Cook, 2008).

Perhaps the best evidence for the existence of specialised transcription factories came in a publication by Schoenfelder *et al.* in 2010 (Schoenfelder et al., 2010). They used a variety of techniques to investigate the nuclear localisation of erythroid genes. A genome-wide screen of genes associating with *Hba* and *Hbb* at transcription factories showed enrichment for genes with CACC motifs capable of binding the erythroid-specific transcription factor Klf1. Using immunofluorescence they found that nuclear Klf1 foci overlap with active RNA polymerase II foci, suggesting that a subset of transcription factories are enriched for Klf1 in erythroid tissues (Fig 1.4). Co-localising Klf1 dependent genes associated with these Klf1 specific transcription factories at a high frequency, and a number of these gene associations were lost in $Klf^{-/-}$ knockout mice. These data suggest that a network of Klf1 specific transcription factories exist within mouse erythroid tissues, and that Klf1 specific genes are preferentially recruited to these sites (Schoenfelder et al., 2010). If such specialised transcription factories are a general feature in mammalian nuclei, they could be a key driving force in the organisation of the genome.

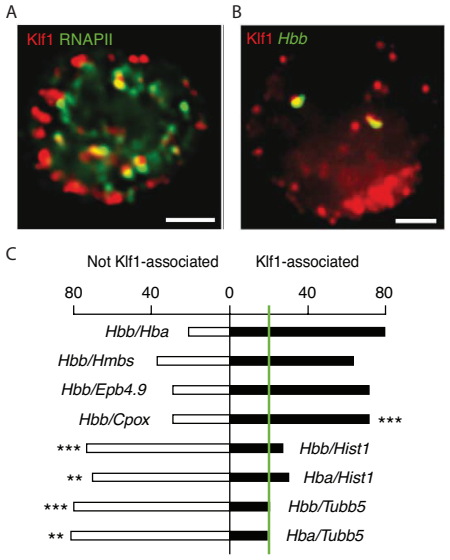

Figure 1.4 – Klf1 specialised transcription factories. (A) Immunofluorescence showing Klf1 (red) overlapping with a subset of RNA Polymerase II transcription factories (green). (B) RNA immuno-FISH showing transcribing *Hbb* alleles (green) co-localising with Klf1 (red). (C) Percentages of co-localising transcript pairs colocalisation with Klf1 associated transcription factories (right, black bars) and non-Klf1 associated factories (left, white bars). Expected level of association due to chance shown by green bar. Scale bars show in (A) and (B) represents 2 μm. Taken from Schoenfelder *et al.* (Schoenfelder et al., 2010).

It is important to note that the evidence above does not describe *exclusively* specific transcription factories, i.e. factories which are incapable of transcribing genes not controlled by

Klf1. Indeed, although statistically significant, the gene associations detected suggest a population bias. Specialised transcription factories are certainly not required for the transcription of the eukaryotic genome, but may aid efficiency and affect nuclear organisation and function.

1.5 Chromatin interactions in three dimensions

1.5.1 Promoter and enhancer interactions

The classical model of gene promoter regulation began with the characterisation of the *lac* operon in *Escherichia coli* in 1961 (Jacob and Monod, 1961). Three genes required for the digestion of lactose are controlled by the binding of a repressor protein which blocks transcription by RNA polymerase. This repressor can form a tetramer and bind two region simultaneously, requiring a topological loop in the chromatin (Savageau, 2011).

A number of eukaryotic genes were later found to require distal enhancer elements (Banerji et al., 1981), prompting speculation that the looping of chromatin allowing direct interaction between sequences may be a common mechanism (Dorsett, 1999). Direct evidence for *in-vivo* chromatin interactions was provided in 2002 with two papers studying the mouse beta globin *(Hbb)* locus (Carter et al., 2002; Tolhuis et al., 2002). This locus contains four beta-like genes arranged in order of their expression through development. Approximately 50 kilobases upstream is a locus control region (LCR), containing multiple DNase hypersensitive sites (HS1-6) (Bender et al., 2000). Carter *et al.* used a novel technique called RNA TRAP (tagging and recovery of associated proteins) which localises horseradish peroxidase (HRP) to the site of nascent RNA production of a specific gene. The HRP catalyses biotinylation of nearby chromatin which can be purified and quantified by quantitative real-time PCR (qPCR). They showed a 15-fold enrichment of biotin over the HS2 region of the LCR when the actively transcribing *Hbb-b1* transcript was used for the HRP localisation, demonstrating that this region of the LCR is in very close proximity to the *Hbb-b1* gene (Carter et al., 2002).

Tolhuis *et al.* (Tolhuis et al., 2002) used the chromosome conformation capture technique (3C) developed by Dekker *et al.* to study the organisation of chromatin within yeast (Dekker et al., 2002) (for discussion of this technique see Section **??**). They studied the same globin locus and also found evidence for interaction between the LCR hypersensitive regions and the *Hbb* genes. Importantly they showed that in mouse foetal brain tissue, where the *Hbb* locus is not expressed, the chromatin adopted a linear conformation without any looping (Tolhuis et al., 2002). Palstra *et al.* went on to show that each gene contacts the LCR as it is expressed, supporting a system where the developmentally controlled genes must contact an enhancer within the LCR in order to be expressed (Palstra et al., 2003). They showed that erythroid progenitor cells dedicated to the lineage but not yet expressing beta-like globin genes form a 'poised' structure, contacting the LCR but not engaging strongly with the HS2 enhancer (Palstra et al., 2003).

1.5.2 Chromatin hubs

The beta globin locus has become a model system for the active chromatin hub (ACH) model (de Laat and Grosveld, 2003), describing a system where loops of chromatin containing elements capable of controlling the expression of genes are held in close three-dimensional space to the genes that they control. Other gene clusters have also been shown to behave in a similar manner, notably the T_H2 locus and *Hox* clusters.

The T_H2 locus control region is involved in the transcriptional control of cytokine genes IL4, 5 and 13 (Lee et al., 2003). Spilianakis *et al.* showed that these genes cluster together in T cells, NK cells, B cells and fibroblasts, despite not being expressed in the last two cell types. In T cells and NK cells the genes additionally associate with the T_H2 LCR (Spilianakis and Flavell, 2004). The authors suggest that this mechanism allows the coordinate expression of the gene cluster in a controlled manner.

The *Hox* genes are master regulators of gene transcription and are responsible for the creation of vertebrate segments during development. *Hox* gene clusters *A* to *D* are transcribed in sequence as development of the embyro progresses and have been observed decondensing and looping out from their chromosome territories upon activation (Chambeyron and Bickmore, 2004). Noordermeer *et al.* studied the *Hox* clusters in three different mouse embryonic day 10.5 tissues: forebrain, anterior trunk and posterior trunk (Noordermeer et al., 2011). They found that the *Hoxd* cluster formed a discrete domain in forebrain, where it is inactive. In anterior and posterior trunk the *Hoxd* cluster is active, but different genes are transcribed. In both tissues they found the cluster to form two distinct compartments correlating with an inactive and active regions. Using circularised chromosome conformation capture (4C), a technique based on 3C, they showed that genes move from the inactive to active compartment as they are activated, correlating with active histone marks (Noordermeer et al., 2011).

1.5.3 Long range interactions

As the role for distal enhancers has become more established, evidence has been uncovered for increasingly distant interactions. For example, Sharpe *et al.* developed a mouse model for preaxial polydactyly by random insertion of a reporter cassette (Sharpe et al., 1999) which they found to affect a *cis* regulatory site over a megabase upstream of the gene *Shh*, known to be important in the condition (Lettice et al., 2002). Lettice *et al.* went on to characterise this enhancer, which lies within an intron of the gene *Lmbr1* unrelated to the condition, demonstrating that 7q36 abnormalities found in patients with preaxial polydactyly disrupt this enhancer (Lettice et al., 2003). In 2005, Velagaleti *et al.* characterised breakpoints found in two patients with the skeletal malformation syndrome campomelic dysplasia (Velagaleti et al., 2005). They found the breakpoints corresponded to two different regulatory elements, one 1.1 Mb upstream of the target gene *SOX9* and one 1.3 Mb downstream. Kleinjan *et al.* used mouse models carrying yeast artificial chromosomes (YACs) to characterise multiple distal enhancers of the developmental control gene *Pax6*

(Kleinjan et al., 2006). They found that as different enhancers were removed, expression of the gene was abolished in different tissues, suggesting a complex system of enhancer - promoter interactions driving the pattern of tissue-specific expression (Kleinjan et al., 2006).

1.5.4 Interactions in *trans*

Spilianakis *et al.* went on from characterising the intra-chromosomal interactions of the T_H2 LCR (described above) to show that the same locus forms inter-chromosomal interactions (Spilianakis et al., 2005). Depending on the stimulus received, naïve T cells can differentiate into either TH1 or TH2 cells, defined by the expression of either IFN-γ or IL-4. Spiliankis *et al.* showed that the *Ifng* gene on chromosome 10 can interact with the T_H2 LCR on chromosome 11 to stimulate *Ifng* expression whilst inhibiting *IL4* expression. This interaction is the first interchromosomal interaction known to regulate gene expression (Spilianakis et al., 2005).

Lomvardas *et al.* used 3C to demonstrate the association of an olfactory receptor gene enhancer made specific contacts to multiple other olfactory genes across the genome (Lomvardas et al., 2006). Mouse dendrites can express one of approximately 1300 odorant receptor genes and Lomvardas *et al.* suggested that this enhancer-gene interaction was the mechanism responsible for the expression of that gene. However, it should be noted that deletion of this enhancer had little effect on the usage of olfactory genes outside of it's cluster (Fuss et al., 2007).

A number of other studies have shown specific interchromosomal contacts involved in a number of processes ranging from X-inactivation to genomic imprinting, showing that these interactions may play an important role in chromatin biology (reviewed in Schneider and Grosschedl, 2007).

1.5.5 Global interaction maps

Our understanding of the three-dimensional organisation of the genome has advanced in leaps and bounds during the last decade largely because of the development of the 3C method and its derivatives (for review, see (Osborne et al., 2011)). There are a large number of 3C variants, but they can be grossly categorised into four classes based on how many loci can be interrogated in a single experiment: one-to-one (3C, quantitative 3C), one-to-all (4C, e4C, ACT), many-to-many (5C) and all-to-all (ChIA-PET, Hi-C, TCC).

The recent development of all-to-all methods has allowed the conformation of the entire genome to be probed in a single experiment. This approach has many advantages; such an unbiased approach allows the detection of unexpected interactions and associations can be probed in parallel allowing a far higher rate of data collection. Whilst these techniques have had great impact on the field, they are currently limited by the depth of sequencing that is achievable with today's technology. To address this Sanyal *et al.* recently published a paper as part of the ENCODE project describing the interaction profiles of 628 transcription start sites (TSS) and 4535 surrounding fragments, representing

approximately 1% of the genome (Sanyal et al., 2012). To achieve the resolution required for the robust detection of promoter-element interactions, Sanyal *et al.* used 5C, a many-to-many technique that uses a panel of oligonucleotides with common adapters to anneal to 3C products and create a library capable of being sequenced. They sequenced libraries from three ENCODE cell lines: K562, HeLa-S3 and GM12878. Only a small proportion of the looping interactions uncovered were shared between the three cell types, with approximately 60% of interactions being unique to a single cell line. The majority of TSS looping interactions could be classified as interacting with enhancer elements, promoters or regions bound by the structural protein CTCF. Looping interactions with enhancer elements were significantly enriched for actively expressed TSS, demonstrating the importance of three-dimensional chromatin contacts in the regulation of gene expression.

1.6 What drives nuclear organisation?

As our understanding of the structure of the nucleus evolves, an increasing number of structural features and patterns are being uncovered. Teasing apart correlation and causation to find the driving forces behind nuclear organisation is not an easy task and remains a major challenge within the field.

1.6.1 Transcription

The discovery of transcription factories has changed our view of nuclear organisation substantially. If genetic templates are mobile and transcription factories are fixed, then it maybe be possible to use the process of transcription as a tool to fold the genome into specific conformations.

Kimura *et al.* quantified the amount of stable RNA polymerase II in HeLa cells (Kimura et al., 1999) adding to the work by Jackson *et al.* demonstrating the stability of nascent transcripts and Polymerase in the nucleus (Jackson et al., 1981; Jackson and Cook, 1985). Mitchell and Fraser later demonstrated that RNA polymerase II transcription factories remain in the absence of transcription, though gene association with factories is ablated if transcription initiation is inhibited (Mitchell and Fraser, 2008). These data, along with the observation that genes are recruited to pre-existing transcription factories upon activation (Osborne et al., 2004), support a model whereby transcription factories are attached to a relatively immobile nuclear substructure. This means that RNA polymerase II can act as a motor, dragging template chromatin through the nucleus as it is transcribed, powered by the removal of phosphate groups during RNA synthesis (Cook, 1999). Yin *et al.* measured the force produced by a single *E. coli* RNA polymerase using an immobilised enzyme transcribing a template bound to a polystyrene bead held by optical tweezers (Yin et al., 1995). RNA polymerase stalled when the force applied was greater than 14 piconewtons (pN), substantially more than kinesin or myosin, making RNA polymerase the most powerful biological motor known. Papantonis *et al.* demonstrated the potential of RNA polymerase to pull chromatin transcripts through the nucleus *in vivo* by using 3C to measure the

change in association between regions of DNA after activation of the TNFα gene (Papantonis et al., 2010). They found that as the gene was transcribed, downstream regions of chromatin progressively came into contact with other transcribing regions at the transcription factory.

The role for transcription in the organisation of the genome is further supported by the pervasive nature of transcription. As much as 93% of genomic bases in the human genome are though to be transcribed in at least one cell type (described in Clark et al., 2011), including many enhancers (Ling et al., 2005; Kim et al., 2010). In the recent ENCODE study of TSS interactions, Sanyal *et al.* found that enhancer elements looping to a TSS were significantly more likely to express enhancer RNAs (Sanyal et al., 2012), supporting a model that the transcription of a chromatin template could be responsible for the formation of chromatin loops.

An attractive model for larger scale genome organisation revolves around the transcription of housekeeping and tissue-specific genes. Lercher *et al.* have shown that genes with high expression in multiple tissue types have a propensity to be present in clusters within the genome (Lercher et al., 2002) and the mouse alpha globin locus has been shown to assemble at a transcription factory already transcribing a cluster of housekeeping genes in erythroid cells (Zhou et al., 2006). Gavrilov *et al.* showed that these housekeeping genes were bound stably to the nuclear matrix and resistant to high salt extraction, whereas the alpha globin genes were less stable and could be eluted (Gavrilov et al., 2010). These studies may point to a model of the nucleus where persistently transcribed housekeeping genes are responsible for forming transcription factories, with tissue specific genes being recruited to these sites upon activation (Gavrilov et al., 2010). Such a system could have far reaching consequences, with chromosome conformation being determined by housekeeping gene hubs.

1.6.2 CTCF

CTCF (CCCTC-binding factor) is a highly conserved DNA-binding protein with eleven zinc finger domains which binds the consensus sequence CCCTC as well as a range of variant sequences using combinations of different zinc fingers (Filippova et al., 1996) (Fig 1.5, reviewed in Ohlsson et al., 2001*)*. The core sequence has extremely high sequence conservation between mouse, chicken and human and mice homozygous for the gene knockout exhibit early embryonic lethality (Splinter et al., 2006). The protein is expressed ubiquitously and misregulation by over-expression or RNAi knockdown have a wide range of effects (Torrano et al., 2005).

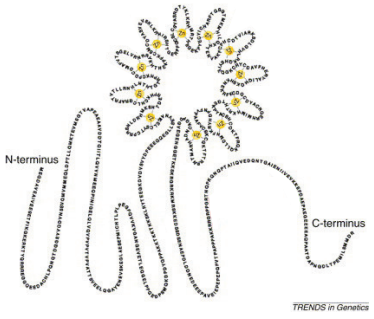

Figure 1.5 – Putative structure of CTCF. Adapted from Ohlsson *et al.* 2001 (Ohlsson et al., 2001).

When CTCF was first isolated it was thought to be a transcriptional repressor for the *c-Myc* gene (Lobanenkov et al., 1990). Since then its proposed functions have included transcriptional activator

(Vostrov and Quitschke, 1997), insulator and structural protein. Bell, West and Felsenfeld were the first to characterise the role of CTCF as an insulator in a study of the chicken beta-globin locus (Bell et al., 1999). They showed that CTCF binds the HS4 region of the β-globin LCR, a region previously shown to act as an insulator in a transgene enhancer-blocking assay (Chung et al., 1997). One of the best characterised examples of CTCF acting as an enhancer is at the imprinted *H19/Igf2* locus. Four CTCF binding sites were found in the imprinting control region (ICR) which exhibit methylation-sensitive binding (Bell and Felsenfeld, 2000; Hark et al., 2000; Szabó et al., 2000). On the unmethylated maternal allele, CTCF binds the ICR and the downstream enhancer element stimulates expression of the *H19* gene. On the methylated paternal allele, CTCF binding is abrogated and the enhancer contacts the distal *Igf2* gene instead, causing expression of *Igf2* and not *H19*. This model is supported by 3C data showing parent-of-origin specific interactions between the enhancer and the two genes (Murrell et al., 2004; Kurukuti et al., 2006).

Genome wide studies of CTCF binding using ChIP-on-chip and ChIP-Seq have greatly developed our understanding of this complex protein. The model of CTCF acting as a global transcriptional activator or repressor has been largely abandoned; binding by CTCF is not able to predict transcriptional activity or correlate with specific classes of genes (Chen et al., 2008). In support of CTCF in the role of a global insulator, Xie *et al.* determined a large number of conserved CTCF binding sites using computational analysis and derived a dataset containing divergent pairs of genes with an without a CTCF separating them (Xie et al., 2007). Divergent gene pairs facing away from each other frequently show correlated expression patterns (Trinklein et al., 2004; Li et al., 2006) - as expected, gene pairs not separated by a CTCF site showed a propensity to be coexpressed, yet those with a CTCF site had a close to background level chance of coexpression (Xie et al., 2007). Not all CTCF binding sites appear to act as enhancer-blocking insulators however, as demonstrated by looping over CTCF sites in the human beta-globin locus (Tolhuis et al., 2002). The recent 5C study by Sanyal *et al.* found that 79% of long range interactions skip one or more CTCF binding sites, prompting the authors to suggest that additional factors may be needed for the insulator function of CTCF (Sanyal et al., 2012).

The development of the ChIA-PET technique has allowed an in depth interrogation of all CTCF bound chromatin interactions in mouse embryonic stem (ES) cells (Handoko et al., 2011). Handoko *et al.* demonstrated with this technique that CTCF acts in multiple roles with different sizes of loops. Correlative analysis using data about surrounding histone modifications allowed the authors to determine five distinct classes of CTCF-derived loops: active domains enriched for H3K4me1, H3K4me2 and H3K36me3; repressive domains enriched for H3K9, H3K20 and H3K27 methylation; putative enhancer-promoter interactions; barrier insulator behaviour, separating active and inactive chromatin compartments; and finally, loops with no known correlation or function (Handoko et al., 2011). This study illustrates the complexity of CTCF binding within the genome and points to multiple roles for the protein involving chromatin organisation. This interaction has subsequently been shown to be dependent on protein-protein contacts between the SA2 subunit of cohesin and the C-terminal domain of CTCF (Xiao et al., 2011).

1.6.3 Cohesin

The structural maintenance of chromosome (SMC) proteins, also known as cohesin, form complexes required for sister chromatid cohesion and chromosome segregation in mitosis. Originally characterised in *Drosophila melanogaster* due to mutants having a defect in sister chromatid separation (Michaelis et al., 1997), cohesin is a multi subunit complex consisting of two large coiled-coil domain proteins (Smc1, Smc3) connected by a smaller dimer of Scc1 and Scc3. Cohesin complexes are thought to function in mitosis by holding sister chromatids together through topological looping (reviewed in Hudson et al., 2009).

A number of studies suggested that cohesin may play a role in the interphase nucleus as well as in mitosis, being implicated in gene regulation, recombination, repair and domain formation (reviewed in Hagstrom and Meyer, 2003). In 2008 three studies showed that the genome-wide binding profiles of cohesin subunits in mammalian cells correlates highly with that of CTCF (Parelho et al., 2008; Wendt et al., 2008; Rubio et al., 2008). Parelho *et al.* expressed a FLAG-tagged cohesin subunit of Rad21 in mouse lymphoid cell lines which was able to pull down the SMC1, SMC3 and SA1 cohesin subunits. They generated ChIP-chip libraries covering approximately 3% of the mouse genome and determined that cohesin binds DNAse hypersensitive regions enriched for a motif highly similar to the CTCF consensus binding sequence (Parelho et al., 2008). Wendt *et al.* followed a similar path, using the HeLa cell line to create ChIP-chip libraries covering approximately 1% of the human genome for both SMC3 and CTCF (Wendt et al., 2008). Rubio *et al.* studied the binding partners of CTCF by mass spectrometry and found that Scc3/SA1 was a key binding partner, leading them to create ChIP-chip libraries for CTCF and Scc3/SA1 in HBL100 cells (Rubio et al., 2008). All three studies came to the same conclusion - that the majority of CTCF and cohesin binding sites overlap in mammalian genomes, leading to a model where the sequence specific binding of CTCF is responsible for the targeting of the cohesin complex.

In these and investigations, cohesin has been implicated at many if not most of the loci described in the above sections. Stedman *et al.* and Wendt *et al.* both demonstrated that cohesin is enriched at the CTCF binding sites of the imprinted *H19/Igf2* locus in the same methylation- and parent of origin- sensitive manner and that this binding is ablated upon mutation of the CTCF sites (Stedman et al., 2008; Wendt et al., 2008). Enrichment of cohesin at this locus is also disrupted in mutants lacking the C-terminal domain of CTCF responsible for binding cohesin (Xiao et al., 2011) and the presence of cohesin is required for the three-dimensional conformation of the locus (Nativio et al., 2009). Cohesin has also been implicated in the control of the beta-globin locus (Wendt et al., 2008; Hou et al., 2010) and *Ifng* / T_H2 LCR (Parelho et al., 2008; Hadjur et al., 2009).

CTCF depletion does not affect the amount of cohesin bound to the genome, but rather the enrichment of cohesin at specific sites (Parelho et al., 2008; Wendt et al., 2008). Schmidt *et al.* showed a subset of cohesin bound regions independent of CTCF binding in MCF-7 cells which colocalise with ER-α binding (Schmidt et al., 2010). A similar

association of cohesin and mediator / Nipb1 at promoters and enhancers in ES cells has been described (Kagey et al., 2010). These studies suggest an attractive hypothesis that multiple tissue-specific proteins may be able to target cohesin binding to specific sites to affect transcriptional profiles.

1.6.4 Tethering

The nucleus is a structured organelle which contains a number of architectural features such as nuclear pores, the inner nuclear membrane (nuclear lamina) and the nucleolus. Chromatin can bind to these regions in a specific nature leading to changes in nuclear organisation and gene expression.

As microscopy studies have advanced our understanding of how the genomic positioning of genes can affect expression, it has became clear that association with the nuclear periphery generally correlates with gene silencing (Kosak et al. (2002); Dietzel et al. (2004); Zink et al. (2004); Finlan et al. (2008); Reddy et al. (2008); discussed in Section 1.3.1). In 2008, Guelen *et al.* used the DamID technique with lamin-B1 tethered to DNA adenine methyltransferase (Dam) to identify regions of the genome associated with the nuclear lamina in human lung fibroblasts (Guelen et al., 2008). They found that the fraction of lamina-associated chromatin on each chromosome correlated with known CT positioning preferences in fibroblasts. Their key finding was that chromatin-lamina associations existed as distinct regions, termed lamina associated domains (LADs). The domains range from 0.1 to 10 megabases in size and are enriched for chromatin marks associated with transcriptional repression: H3K27me3, H3K9me2, low H3K4me2, low RNA polymerase II, low gene expression and low gene density. LAD boundaries are enriched for CTCF binding and CpG island, suggesting a mechanism of association (Guelen et al., 2008). Shimi *et al.* showed in the same year that Lamins A and B form separate meshes on the inner nuclear membrane and relatively static structures within the nuclear matrix, hinting at the existence of highly complex micro-environments based on the binding of chromatin to lamin networks (Shimi et al., 2008). A subsequent study investigating cells with a mutation in the *LMNA* lamin gene has shown that the position, compaction and transcriptional activity of some lamin-associated regions are affected (Mewborn et al., 2010), suggesting that the many diverse conditions caused by lamin mutations may arise due to changes in chromatin structure and so gene expression.

Although LADs are associated with low gene expression, not all chromatin at the nuclear membrane is silenced. Early electron micrographs showed regions of less dense chromatin at nuclear pores (reviewed in Capelson and Hetzer (2009); Arib and Akhtar (2011)). This observation was recently validated with the observation of nuclear pore complexes (NPCs) contacting channels with heterochromatin using a new form of sub-diffraction limit light microscopy able to simultaneously image NPCs, lamins and chromatin (Schermelleh et al., 2008). Nuclear pore proteins have been found to be associated with active regions of chromatin and are present both at NPC and within the nucleoplasm (Vaquerizas et al., 2010; Kalverda et al., 2010).

1.6.5 Actin and myosin

The presence of nuclear actin has been debated for many years; its initial detection often labelled as artefacts due to the inability of phalloidin to stain actin fibrils within the nucleus (reviewed in Hofmann and de Lanerolle, 2006). Despite this skepticism, nuclear actin research has had a resurgence in recent years with a studies linking filamentous actin to processes such as transcription (Hofmann et al., 2004) and nuclear export (Hofmann et al., 2001). McDonald *et al.* used fluorescence recovery after photobleaching (FRAP) microscopy to study nuclear actin in HeLa cells (McDonald et al., 2006). Treatment with latrunculin which inhibits the polymerisation of actin lead to a loss of a slow moving population of actin, suggesting the existence of polymeric actin in the nucleus. Within the cytoplasm, force is generated through interaction between bundles of filamentous actin and polymerised myosin II. The nucleus does not contain any myosin II, though an isoform of myosin I incapable of forming filaments has been detected (Pestic-Dragovich et al., 2000). Our understanding of how nuclear actin and myosin are involved in nuclear organisation is still in its infancy, yet is rapidly gaining traction as techniques are developed which allow us to probe and manipulate their behaviour within the nucleus without disrupting cytoplasmic processes.

1.6.6 Replication

Proliferating cells must replicate their genomes once per cell cycle and do so in only a few hours, despite their size. This is accomplished by simultaneously replicating many regions of the genome at shared sites of replicated called replication factories (Jackson and Pombo, 1998; Ma et al., 1998). The genome is replicated in an organised manner; active genes are linked to early replication and inactive genes tend to replicate late (reviewed in Goren and Cedar, 2003). Clusters of replication foci that share replication factories continue to associate through multiple cycles of cell division (Jackson and Pombo, 1998). Many studies have shown correlation between behaviour in replication and genomic features or activity, though causative links are still lacking (reviewed in (Chakalova et al., 2005)).

1.6.7 Polycomb

Another nuclear subcompartment known to be involved in genomic organisation and looping is the polycomb body. The Polycomb protein complex (PcG) is involved in the directed silencing of regions at polycomb response elements (PREs), mediated by repressive histone modifications such as H3K27me3 (Cao et al., 2002), and are important for the maintenance of silencing of the *Hox* genes (reviewed in Pirrotta, 1998). Long range interactions and chromatin looping has been implicated in repression by PcG proteins Tiwari et al., 2008b;a.

1.6.8 The bigger picture

When seen as a whole, many of the models described above are not mutually exclusive. Many hold in common the presence of chromatin loops within the nucleus and describe domains of chromatin defined by epigenetic marks. It seems entirely plausible that the overall structure of the nucleus is determined by the combined result of many different processes directing specific interactions, driven by different processes and stabilised by different types of contact. Whilst this picture is still developing, it is clear that the nucleus is a highly dynamic and plastic organelle defined by millions of years of evolutionary pressures on every aspect of its content.

1.7 Chromosomal translocations

One of the major medical implications of genomic organisation is how it may relate to the formation of chromosomal translocations. Specific chromosomal rearrangements are frequently associated with certain cancer types, and can be highly predictive of patient prognosis. Furthermore, cancers involving different translocations can respond to treatments in different ways, paving the way for patient-specific treatment regimes (reviewed in Mitelman et al., 2007). Understanding the principles of chromosomal translocation formation is an important step in the development of novel treatments.

1.7.1 Formation of chromosomal translocations

Since the development of chromosome banding microscopy, it has been known that the genetic material within cancerous cells is frequently disrupted. Chromosomal translocations involve the rearrangement of genetic material between non-homologous chromosomes through the formation and aberrant repair of at least two double strand breaks (DSBs) that are situated on different chromosomes. The resulting product can be a straight swap (balanced translocations) or can result in deletions or even gain of material after malsegregation in mitosis (unbalanced translocations). Chromosomal translocations can involve multiple DSBs on different chromosomes resulting in hugely complex karyotypes, especially in cells predisposed to translocations due to defects in repair.

1.7.1.1 DSB repair

Double strand breaks can be caused by exogenous damage (ionising radiation, free radicals) and endogenous damage (physiological programmed DSBs). As many as 1 million DNA lesions are formed per cell per day (Alberts Lewis, Raff, Roberts, Walter, 2007) and can be highly deleterious to the cell, causing loss of genetic material, translocations and ultimately cell death if left unrepaired. Mammalian cells have an array of different DSB repair pathways that depend on the cellular context and type of lesion (Longhese et al., 2006).

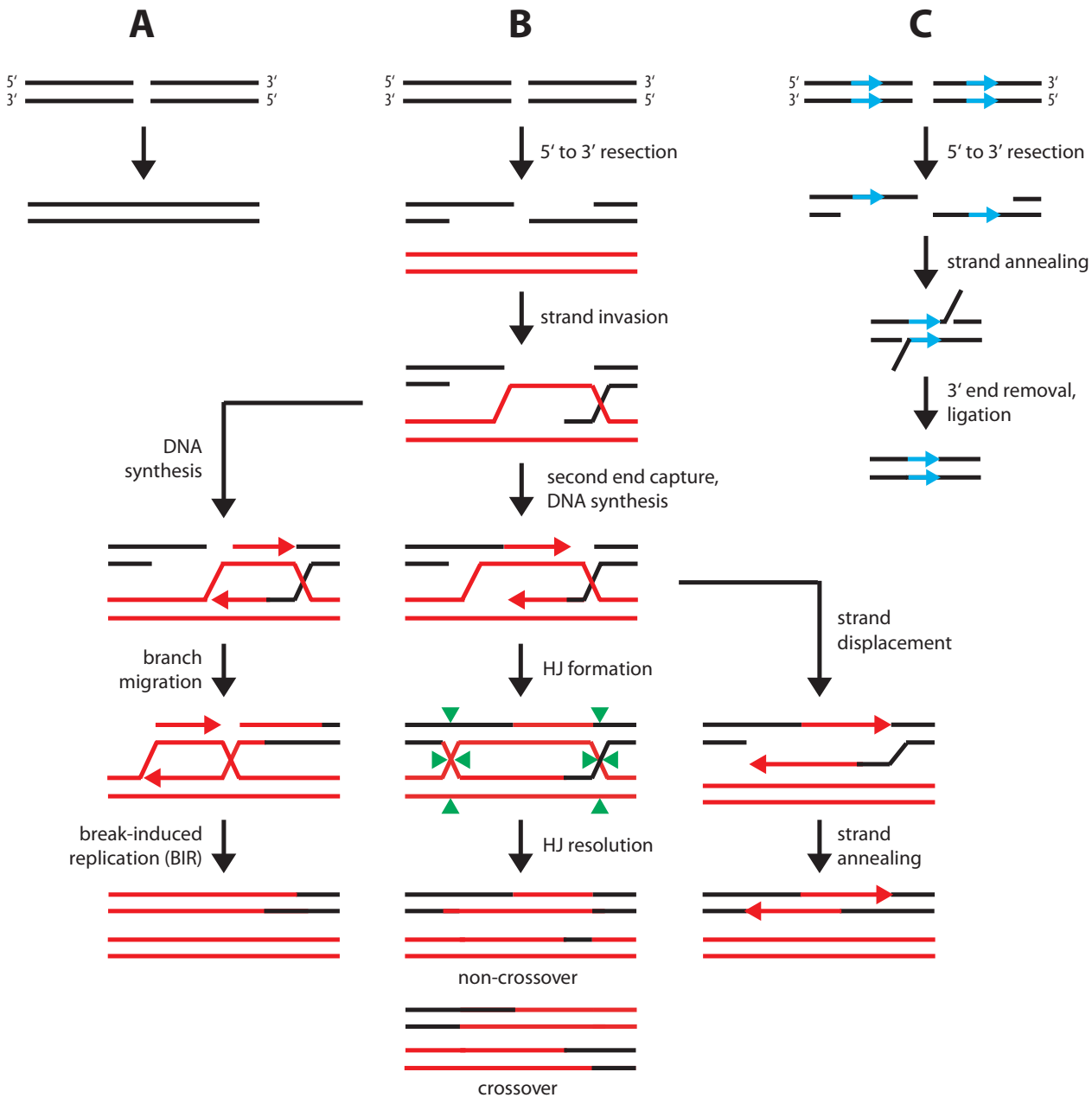

Figure 1.6 – **DSB repair.** Three double strand break repair pathways. **(A)** NHEJ (Non-Homologous End Joining) requires little or no sequence homology and joins any two nearby ends. **(B)** HR (Homologous Recombination) uses a strand of homologous DNA on a duplicated interphase chromosome or sister chromatid to reconstitute sequence before joining. Although it only requires a single double strand break to generate a crossover, the extensive region of homology needed makes errors rare. **(C)** SSA/NHEJ (Single Strand Annealing / Non-Homologous End Joining) searches for regions of homology, such as repeat sequences, and deletes any sequence between. HJ is Holliday junction. From Longhese *et al.*, 2006 (Longhese et al., 2006).

Homologous recombination (HR, Fig 1.6B) repairs double strand breaks by using long regions of homology on undamaged sister chromatids or homologous chromosomes. HR is primarily active during cell replication and rarely leads to serious chromosomal rearrangements or translocations, although 'crossing-over' can sometimes occur,

whereby two sections of homologous chromosomes switch to their partner chromosomes. This process is encouraged during meiosis with the formation of chiasmata, crossovers that are 10^4 to 10^5 times more frequent than in mitosis (Lee et al., 2009). During meiosis homologous chromosomes are preferentially used for HR to promote crossovers, but sister chromatids are more commonly used in mitosis to minimise variation (Schwacha and Kleckner, 1997; Haber, 2000). Non-allelic homologous recombination (NAHR) is a variation of HR characterised by the aberrant use of low-copy repeats during strand invasion. NAHR can result in duplications, inversions and translocations and is responsible for a number of spontaneous genetic disorders such as Potocki-Lupski syndrome (Potocki et al., 2007) and cancers involving recurring breakpoints (Darai-Ramqvist et al., 2008) (reviewed in Gu et al., 2008).

Single strand annealing (SSA, Fig 1.6C) creates a short region of single stranded DNA at the site of the DSB which is used to search for regions of microhomology, usually repeat sequences in the same orientation. These regions anneal and any excess single stranded tails are removed to allow the nicks to be ligated. SSA is prone to introducing deletions into the genome, and can form chromosomal translocations if two sequences with similar repeat regions are nearby, and both suffer simultaneous DSBs.

Non-homologous end joining (NHEJ, Fig 1.6A) is a last resort for the cell, whereby any two adjacent DSBs are ligated, needing little or no sequence homology. This pathway is extremely prone to errors as it is capable of joining any two strands of DNA irrespective of their identity.

A number of serious conditions exist due to mutations in genes key in the DSB repair pathways described above. Ataxia telangiectasia is an autosomal-recessive neurodegenerative disease caused by mutations in the *ATM* gene which is involved in NHEJ and HR (Beucher et al., 2009). Fanconi anaemia is a condition caused by mutations in one of a number of proteins involved in DSB recognition and repair which is associated with a high incidence of leukaemia and a number of congenital defects. Blooms syndrome is characterised by a excessive HR and genomic instability caused by mutations in the *BLM* gene. These, and other similar diseases, are indicative of the importance of DSB repair and the danger of genomic instability.

1.8 Leukaemia

Chromosomal translocations are frequently observed within cancer cells and can be one of the initiating events leading to oncogenesis. Leukaemias are one of the best studied cancer models due to the ease of accessibility of the affected cells; many of the best characterised cancer pathologies have been within leukaemias.

1.8.1 Chronic myeloid leukaemia

Chronic myeloid leukaemia (CML) is one of the best studied cancers, despite having a relatively low occurrence (Hehlmann et al., 2007). The term 'leukaemia' was based on CML patients in the mid 1800s (Geary, 2000), and it was

the first cancer for which a direct causative genotype was found (Nowell and Hungerford, 1960) (see below). A number of standardised therapies exist for CML and up to 87% of patients achieve complete cytogenetic remission (Hehlmann et al., 2007), making it one of the success stories of cancer research.

1.8.1.1 The Philadelphia chromosome

This genetic abnormality, primarily associated with CML, was first the first consistent chromosomal abnormality found with any cancer. Identified in 1960 by Nowell and Hungerford (Nowell and Hungerford, 1960) they noticed the presence of an abnormal chromosome and called it the Philadelphia chromosome, or Ph chromosome, after the city in which it was discovered. The chromosome was initially thought to be the result of a deletion until chromosome banding enabled its identification as a product of a translocation (Rowley, 1973). The breakpoints within chromosomes 9 (cytogenetic band q34) and 22 (band q22) were subsequently identified as being within the *c-ABL* and *BCR* genes, respectively. The *c-ABL* gene (also known as *ABL1)* was named as such because of its similarity to the Abelson murine leukaemia virus gene *v-ABL (Rowley, 2001).* Less is known about the *BCR* gene, so named because it was found at the Breakpoint Cluster Region of chromosome 22.

The *ABL1* gene encodes a tyrosine kinase which is ubiquitously expressed in mammalian cells. Involved in cell cycle regulation, it is thought to be involved in several cell signalling pathways (Deininger et al., 2000). The *BCR* protein contains a serine-threonine kinase and has GTPase activity, but its function is not known (Deininger et al., 2000). The t(9;22)(q34;q11) translocation creates a fusion protein lacking the SH2 domain of *ABL1* which normally regulates its activity. This results in constitutively active tyrosine-kinase activity which drives oncogenesis (Deininger et al., 2000). The mechanism of translocation formation is yet to be characterised, though ionising radiation is known to be a risk factor (Tanaka et al., 1989; Corso et al., 1995). Identification and characterisation of the *BCR-ABL* translocation and its fusion protein led to the development of a number of kinase inhibitors such as the drug *imatinib* (also known as *gleevec*), which achieve excellent success in the treatment of CML and a number of other cancers involving the t(9;22)(q34;q11) translocation (Hehlmann et al., 2007).

1.8.2 Mixed lineage leukaemia

Acute leukaemia is typically categorised as either Acute Lymphoblastic Leukaemia (ALL) or Acute Myeloid Leukaemia (AML) according to which lineage of blood cells are cancerous. ALL is defined by the uncontrolled proliferation of lymphoblasts, precursors to lymphocytes which differentiate into B cells, T cells and NK cells. ALL was estimated to account for approximately 12% of all leukaemia cases in the US in 2008 with around five and a half thousand cases occurring annually. It is the most common form of cancer in children aged under fourteen (Jemal et al., 2009).

AML is a cancer of the myeloid lineages of blood cells, responsible for approximately 30% of all cases of leukaemia in the US for 2008 with a higher mortality rate than ALL (66% and 27% mortality for AML and ALL, respectively)

(Jemal et al., 2009). The incidence of AML increases with age; the median age of AML diagnosis between 2002 and 2006 in the US was 67 years, and the median age of mortality was 72 (Jemal et al., 2009).

The biology of leukaemia is not always as discrete as this classification however, some patients present with expansion of both lymphoid and myeloid lineages (Matutes et al., 1997). Two genetic abnormalities sometimes found in these patients include the t(9;22)(q34;q11) Philadelphia chromosome and structural changes in 11q23 (Matutes et al., 1997). Chromosomal translocations within band q23 of chromosome 11 have been implicated in both AML and ALL, all involve the gene *MLL* (also known as *ALL-1*, *Htrx*, *HRX*) identified by Ziemin-van der Poel *et al.* in 1991 (Ziemin-van der Poel et al., 1991). Translocations involving the *MLL* gene are found in over 70% of all infant leukaemias (Biondi et al., 2000) and approximately 10% of adult AML cases (Krivtsov and Armstrong, 2007). The translocation correlates with poor patient prognosis and is of high clinical interest (Chen et al., 1993).

In 2002, Armstrong *et al.* showed that acute lymphoblastic leukaemias containing a translocation within the *MLL* gene have a unique expression profile that is different to ALL and AML, and suggests an origin within a less committed progenitor cell which can produce cells in both the myeloid and lymphoid cell lineages (Armstrong et al., 2002). They suggest that these leukaemias are substantially different from AML and ALL and deserve a new, distinct, class of leukaemia called Mixed Lineage Leukaemia (MLL). Further cytogenetic studies have supported this theory; leukaemic cells in MLL have been found to express cell surface antigens normally present on both myeloid and lymphoid cells such as CD14 and CD19 (Krivtsov and Armstrong, 2007).

The suggestion that MLL is initiated within a progenitor cell is in line with previous evidence showing that some cases of ALL and AML are initiated within undifferentiated haematopoietic stem cells (HSCs) (Sutherland et al., 1996; Bonnet and Dick, 1997). HSCs are present within the $CD34^+$ progenitor cell population, accounting for approximately 3% of normal human bone marrow and 0.3 to 0.5% of human cord blood mononuclear cells (Libura et al., 2008). They are long lived and capable of self-renewal, differentiating into lineage restricted progenitors and eventually mature terminally differentiated white blood cells. HSCs are necessary for the long term maintenance of the haematopoietic system and are commonly used to repopulate bone marrow after myeloablative therapy, as well as in the treatment of a number of other disorders such as autoimmune, cardiac and vascular diseases (Burt et al., 2008). Cancer stem cells are thought to be present in both leukaemias and solid tumours (Hamburger and Salmon, 1977) and HSCs are a probable founder population due to their ability to self-renew (Reya et al., 2001). The concept of a small pool of cancer stem-cells driving the large heterogeneous pool of cancer cells has a number of implications for treatment - these are the cells that must be targeted for the efficient and long lasting cure of cancer (Reya et al., 2001).

1.8.2.1 The MLL protein

The MLL protein is a H3K4 methyltransferase involved in the positive regulation of global gene regulation, including the maintenance of expression of the *Hox* genes (Yu et al., 1995). *MLL* is required for embryonic haematopoiesis

(Hess et al., 1997) and adult bone marrow maintenance (Jude et al., 2007). MLL is a mammalian homologue of the *Drosophila melanogaster* trithorax complex and is thought to bind DNA via an AT-hook domain (Zeleznik-Le et al., 1994) and a zinc finger domain (Birke et al., 2002). It is thought that the zinc finger domain targets MLL to unmethylated CpG island DNA (Birke et al., 2002) and ChIP studies have shown that MLL binds to a subset of transcribed genes (Milne et al., 2005). MLL binds promoters and gene bodies, associating tightly with RNA polymerase II (Milne et al., 2005).

All known MLL fusion proteins contain exons 8-13 of *MLL* and in-frame exons of a partner gene (Krivtsov and Armstrong, 2007). Fusion proteins always retain their AT-hook and zinc-finger CxxC motifs, which are essential for their transforming potential (Slany et al., 1998). The H3K4 methyltransferase domain of MLL is often lost in fusion proteins (Krivtsov and Armstrong, 2007), despite this the fusion proteins can drive constitutive expression of *HOXA9* and *MEIS1* which, if over-expressed together in the absence of a MLL fusion protein, give a similar phenotype (Zeisig et al., 2004).

1.8.2.2 The *MLL* gene

Translocations within the *MLL* gene are found in approximately 10% of all human leukaemias (Huret et al., 2001). There are 87 documented *MLL* translocation partners of which 51 have been characterised at the molecular level (Meyer et al., 2006); the five most frequent translocation partners, *AF4*, *AF9*, *ENL*, *AF10* and *AF6* account for approximately 80% of cases (Meyer et al., 2006). Translocations within the *MLL* gene usually occur within an 8.3 kb *BamHI* fragment known as the breakpoint cluster region (Gu et al., 1994). This region contains exons 5-11 as well as a number of repeat regions, notably eight direct *Alu* SINE repeats, five direct L1 and L2 LINE repeats and two MER elements, as well as a number of putative topoisomerase II binding sites and a SAR/MAR (Sung et al., 2006) (Fig 1.7). An internal promoter is present within the murine *Mll* breakpoint cluster region, correlating with etoposide-induced DSBs (Scharf et al., 2007), DNase I hypersensitive sites (Strissel et al., 1998) and histone modifications associated with transcription (Khobta et al., 2004).

1.8.2.3 Mechanisms of *MLL* translocation formation

A clue to how translocations form within the *MLL* gene comes from the observation that 11q23 translocations are especially prevalent in therapy-related leukaemias - secondary leukaemias that develop in patients after treatment for a primary cancer with topoisomerase II inhibitors (Krivtsov and Armstrong, 2007). Topoisomerases are found in all eukaryotic nuclei and are able to relieve supercoiling and promote chromosome disentanglement (Buck and Zechiedrich, 2004). They function by binding DNA, forming a transient double strand break and passing another strand of DNA through the gap, before ligating the DSB.

Topoisomerases are important in transcription, which creates supercoils as template DNA is processed through

static RNA polymerase II enzymes (Liu and Wang, 1987). DNA topoisomerase II associates with gene promoters (Collins et al., 2001) and is required for the transcription of genes longer than 3 Kbp in yeast (Joshi et al., 2012). Topoisomerase II induced DSBs have been implicated in the regulation of certain genes via the assembly of transcription complexes and changes in chromatin structure (Ju et al., 2006).

DNA topoisomerase II inhibitors are commonly used as chemotherapeutic agents and work by decreasing the ligation rate, disrupting the cleavage / ligation equilibrium. This leads to an increase in DNA cleavage and an accumulation of DSBs, triggering the cell DNA damage response and leading to cell death by apoptosis (Burden and Osheroff, 1998). Topoisomerase II inhibitors are widely used chemotherapeutic agents, effective against a range of malignancies including small-cell lung cancer and gonadal tumours (Arnold and Whitehouse, 1981).

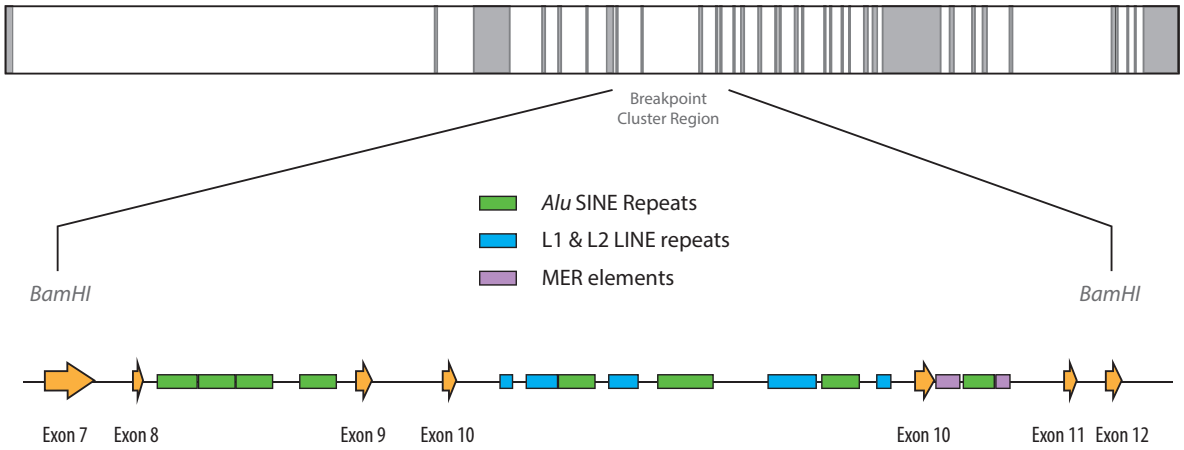

Figure 1.7 – Diagram of the *MLL* breakpoint cluster region.

The *MLL* breakpoint cluster region contains a number of putative topoisomerase II binding sites suggesting a mechanism of topoisomerase inhibitor related translocations (Broeker et al., 1996). Libura *et al.* showed that $CD34^{+}$ HSPCs exposed to etoposide formed 11q23 chromosomal aberrations typical of those seen in clinical samples (Libura et al., 2005). They went on to show that such exposure to etoposide increased the proliferative potential of the cells in a bone marrow graft assay using immune-deficient mice (Libura et al., 2008). This data supports a model whereby poisoned topoisomerase II may create double strand breaks within the *MLL* gene, allowing the NHEJ machinery to aberrantly repair the locus due to microhomology found in the nearby repetitive elements (Fig 1.7).

1.9 Effect of nuclear organisation on translocation formation

1.9.1 Breakage first and contact first models

For a chromosomal translocation to form, two double strand breaks must exist simultaneously and be adjacent in three-dimensional space. Two models have been described to explain how DSBs may meet in the nucleus – the breakage first model and the contact first model. The breakage first model states that DSBs can form anywhere and are able to freely diffuse in the nuclear space. They undergo large scale movement through the nucleus until they meet and are joined. The contact first model states that the two DSBs form in sequences already close to each other within the nuclear space, therefore large scale chromatin movements are not required for the two breaks to meet.

In support of the breakage first model, Aten *et al*. showed that DSBs formed simultaneously in HeLa nuclei clustered together (Aten et al., 2004). However, others have argued that this may be the result of a higher degree of chromatin mobility along the ion beam trajectory used to generate the DSBs (Jakob et al., 2009). A larger body of evidence supports a contact first model; DSBs created with ultra-soft X-rays by Nelms *et al.* remained in a fixed position for several hours after the damage was caused (Nelms et al., 1998) and Jakob *et al.* used live cell microscopy to visualise proteins involved in DNA damage signalling and repair. After accumulation at sites of DNA damage caused by heavy ion impacts, the foci exhibited only a small degree of movement (Jakob et al., 2009). Soutoglou *et al.* showed a high degree of positional stability of DSB ends created by endonuclease digestion, with a greater degree of local diffusion seen in the absence of the Ku80 DNA-end binding protein (Soutoglou et al., 2007).

The contact first model of translocation formation has large implications within the field of nuclear organisation, as it requires specific chromosomal contacts within cells prior to translocation.

1.9.2 Chromosome territories and translocations

The large scale organisation of chromosome territories has been implicated in translocation frequency by a number of studies. Kozubek *et al.* showed that chromosomes 9 and 22 were found in the centre of nuclei more frequently than would be expected by chance in lymphocytes, T- and B-cells, HL60 cells and bone marrow cells (Kozubek et al., 1999). They used neutron irradiation to show that transfer of genetic material was much higher than that found with chromosome 8, which was situated towards the nuclear periphery (Kozubek et al., 1999). Parada *et al.* investigated the positions of chromosomes 12, 14 and 15 in a mouse lymphoma cell line and mouse splenocytes (Parada et al., 2002). The found that two translocated chromosomes preferentially paired together in the nucleus of the cell line as well as in normal cells not containing the translocation. (Parada et al., 2002). Parada *et al.* went on in a further study to examine the positioning of a larger range of chromosomes in a number of different tissue types (Parada et al., 2004). They found that chromosome pairing was tissue specific and correlated with the occurrence of tissue-specific translocation

events (Parada et al., 2004). Kuroda *et al.* published a similar study in the same year, showing that association between chromomes 12 and 16 varies through adipocyte differentiation. These two chromosomes are involved a chromosomal translocation that can lead to liposarcomas, thought to be initiated within pre-adipocytes (Kuroda et al., 2004).

In 2006, Branco and Pombo published a study where they developed a new technique known as cryo-FISH to enhance the resolution of chromosome territory detection (Branco and Pombo, 2006). This technique had sufficient resolution to study the degree of intermingling between chromosome territories and they found that the degree of intermingling between chromosome pairs correlated strongly with their propensity to form translocations when subjected to radiation (Fig 1.8, Branco and Pombo, 2006).

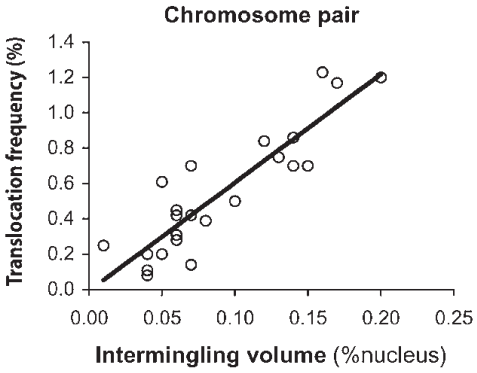

Figure 1.8 – Correlation of CT intermingling and radiation induced translocation frequencies. Adapted from Branco and Pombo, 2006 (Branco and Pombo, 2006).

Interestingly, derivative chromosomes that result from balanced translocations affect the organisation of CTs within the nucleus (Harewood et al., 2010), raising the possibility that the global changes in gene expression observed after oncogenic translocation formation could be in part due to changes in genome organisation (Harewood et al., 2010).

1.9.3 Transcription factories and translocations

Mounting evidence has shown that in addition to CT pairing, the position of specific genes involved in chromosomal translocations are frequently found in close proximity (Neves et al., 1999; Roix et al., 2003). As discussed in Section 1.6.1, localisation at transcription factories can drive organisation within the genome. Osborne *et al.* used DNA- and RNA-FISH to study the localisation of the *Myc* proto-oncogene upon activation of B cells (Osborne et al., 2007). They showed that *Myc,* on mouse chromosome 15, is dynamically recruited to an existing transcription factory and that this transcription factory was preferentially occupied by *Igh,* on mouse chromosome 12. These observations of preferential association at transcription factories support to a model whereby genes sharing transcription factories are predisposed to the formation of chromosomal translocations (Osborne et al., 2007). Similar transcriptional associations have recently been observed for *MLL* and its frequent translocation partners *AF4* and *AF9* (Cowell et al., 2012).

1.10 Thesis overview

Our understanding of nuclear architecture and organisation has developed rapidly in the last twenty years, in concert with the development of techniques allowing ever larger and less biased studies. Correlations have been found between the organisation of specific chromosomes and gene loci and their propensity to form chromosomal translocations which can lead to oncogenesis. A limitation of many of these studies is their scope: either they describe the gross organisation of chromosomes and large genomic features, or they describe the association of specific loci. Gross nuclear organisation is valuable in determining global rules and patterns of genomic organisation, but lacks the resolution needed to study specific loci. Studying the positions of individual genes is also useful, but limited to the study of anticipated associations, leading to candidate choice bias.

In this thesis I describe the further development of the e4C technique which allows the interrogation of the three-dimensional contacts made by a chosen locus in a relatively unbiased, genome-wide manner. I investigate the association profiles of three proto-oncogenes: *BCR, ABL1* and *MLL*. I do so in human $CD34^+$ haematopoietic stem cells, the tissue though to be the site of initial chromosomal translocations for these genes, and the human lymphoblastoid cell line GM12878.

The e4C protocol was previously used in mouse foetal liver cells for analysis with microarrays (Schoenfelder et al., 2010). In chapter three I describe my modification of the protocol for use with limited numbers of human $CD34^+$ cells and analysis with high-throughput sequencing technologies. This work demonstrates the potential of the technique to analyse the interaction profiles of any genomic locus, as well as a number of pitfalls which must be negotiated in its use.

Chapter four describes my subsequent development of analysis tools and techniques which I use to understand the sequencing data. At the time this work was carried out, such genome-wide interaction datasets were being produced for the first time and a great deal of exploration was required in order to find useful forms of analysis which could provide accurate and unbiased representations of the association data. In this chapter I describe a number of steps developed to normalise biases and analyse the data, several of which can be found in parallel studies published during the same period.

In chapter five I describe the initial analysis of association data from the three genes, which exhibit a common association with actively transcribed regions of the genome. This analysis is feasible due to the availability of a number of publicly accessible datasets describing the binding profiles of RNA polymerase, histone modifications and abundance of transcripts. The patterns I uncover support similar findings by other groups showing the separation of the genome into active and inactive compartments.

In chapter six I analyse the interaction of *BCR* with the telomeric region of chromosome 9, containing its recurrent translocation partner *ABL1*. I demonstrate that this interaction is the strongest in the genome and highly significant.

Interestingly, the maximal point of association with *BCR* is not at the *ABL1* locus, but rather in three loci several megabases away. I describe my use of DNA-FISH to validate these interactions in $CD34^+$ cells and GM12878 cells.

Unbiased yet specific studies of genomic interactions such as those described within this thesis help to demonstrate the importance of genomic organisation in the formation of chromosomal translocations and initiation of cancer. By furthering our understanding of how healthy cells are regulated, we may better understand how to prevent disease causing events.

Bibliography

Adam, M., F. Robert, M. Larochelle, and L. Gaudreau: 2001, 'H2A.Z is required for global chromatin integrity and for recruitment of RNA polymerase II under specific conditions.'. *Molecular and cellular biology* **21**(18), 6270–9.

Alberts Lewis, Raff, Roberts, Walter, J.: 2007, *Molecular Biology of the Cell*. Garland Science, 5 edition.

Arib, G. and A. Akhtar: 2011, 'Multiple facets of nuclear periphery in gene expression control.'. *Current opinion in cell biology* **23**(3), 346–53.

Armstrong, S. a., J. E. Staunton, L. B. Silverman, R. Pieters, M. L. den Boer, M. D. Minden, S. E. Sallan, E. S. Lander, T. R. Golub, and S. J. Korsmeyer: 2002, 'MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia.'. *Nature genetics* **30**(1), 41–7.

Arnold, A. M. and J. M. Whitehouse: 1981, 'Etoposide: a new anti-cancer agent'. *Lancet* **2**(8252), 912–915.

Aten, J. A., J. Stap, P. M. Krawczyk, C. H. van Oven, R. A. Hoebe, J. Essers, and R. Kanaar: 2004, 'Dynamics of DNA double-strand breaks revealed by clustering of damaged chromosome domains'. *Science* **303**(5654), 92–95.

Banerji, J., S. Rusconi, and W. Schaffner: 1981, 'Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences.'. *Cell* **27**(2 Pt 1), 299–308.

Bartlett, J., J. Blagojevic, D. Carter, C. Eskiw, M. Fromaget, C. Job, M. Shamsher, I. F. Trindade, M. Xu, and P. R. Cook: 2006, 'Specialized transcription factories.'. *Biochemical Society symposium* **181**(73), 67–75.

Bell, a. C. and G. Felsenfeld: 2000, 'Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene.'. *Nature* **405**(6785), 482–5.

Bell, a. C., a. G. West, and G. Felsenfeld: 1999, 'The protein CTCF is required for the enhancer blocking activity of vertebrate insulators.'. *Cell* **98**(3), 387–96.

Bender, M. a., M. Bulger, J. Close, and M. Groudine: 2000, 'Beta-globin gene switching and DNase I sensitivity of the endogenous beta-globin locus in mice do not require the locus control region.'. *Molecular cell* **5**(2), 387–93.

Beucher, A., J. Birraux, L. Tchouandong, O. Barton, A. Shibata, S. Conrad, A. a. Goodarzi, A. Krempler, P. a. Jeggo, and M. Löbrich: 2009, 'ATM and Artemis promote homologous recombination of radiation-induced DNA double-strand breaks in G2.'. *The EMBO journal* **28**(21), 3413–27.

Biondi, A., G. Cimino, R. Pieters, and C. H. Pui: 2000, 'Biological and therapeutic aspects of infant leukemia'. *Blood* **96**(1), 24–33.

Birke, M., S. Schreiner, M. P. Garcia-Cuellar, K. Mahr, F. Titgemeyer, and R. K. Slany: 2002, 'The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation'. *Nucleic Acids Res* **30**(4), 958–965.

Blumenthal, T., D. Evans, C. D. Link, A. Guffanti, D. Lawson, J. Thierry-Mieg, D. Thierry-Mieg, W. L. Chiu, K. Duke, M. Kiraly, and S. K. Kim: 2002, 'A global analysis of Caenorhabditis elegans operons.'. *Nature* **417**(6891), 851–4.

Bolzer, A., G. Kreth, I. Solovei, D. Koehler, K. Saracoglu, C. Fauth, S. Müller, R. Eils, C. Cremer, M. R. Speicher, and T. Cremer: 2005, 'Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes.'. *PLoS biology* **3**(5), e157.

Bonnet, D. and J. E. Dick: 1997, 'Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell'. *Nature Medicine* **3**(7), 730–737.

Boutanaev, A. M., A. I. Kalmykova, Y. Y. Shevelyov, and D. I. Nurminsky: 2002, 'Large clusters of co-expressed genes in the Drosophila genome.'. *Nature* **420**(6916), 666–9.

Branco, M. R. and A. Pombo: 2006, 'Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations.'. *PLoS biology* **4**(5), e138.

Bridger, J. M., S. Boyle, I. R. Kill, and W. a. Bickmore: 2000, 'Re-modelling of nuclear architecture in quiescent and senescent human fibroblasts.'. *Current biology : CB* **10**(3), 149–52.

Bridger, J. M., H. Herrmann, C. Münkel, and P. Lichter: 1998, 'Identification of an interchromosomal compartment by polymerization of nuclear-targeted vimentin.'. *Journal of cell science* **111 (Pt 9**, 1241–53.

Broeker, P. L., H. G. Super, M. J. Thirman, H. Pomykala, Y. Yonebayashi, S. Tanabe, N. Zeleznik-Le, and J. D. Rowley: 1996, 'Distribution of 11q23 breakpoints within the MLL breakpoint cluster region in de novo acute leukemia and in treatment-related acute myeloid leukemia: correlation with scaffold attachment regions and topoisomerase II consensus binding sites'. *Blood* **87**(5), 1912–1922.

Buck, G. R. and E. L. Zechiedrich: 2004, 'DNA disentangling by type-2 topoisomerases'. *Journal of Molecular Biology* **340**(5), 933–939.

Burden, D. A. and N. Osheroff: 1998, 'Mechanism of action of eukaryotic topoisomerase II and drugs targeted to the enzyme'. *Biochimica et Biophysica Acta* **1400**(1-3), 139–154.

Burt, R. K., Y. Loh, W. Pearce, N. Beohar, W. G. Barr, R. Craig, Y. Wen, J. a. Rapp, and J. Kessler: 2008, 'Clinical applications of blood-derived and marrow-derived stem cells for nonmalignant diseases.'. *JAMA : the journal of the American Medical Association* **299**(8), 925–36.

Cao, R., L. Wang, H. Wang, L. Xia, H. Erdjument-Bromage, P. Tempst, R. S. Jones, and Y. Zhang: 2002, 'Role of histone H3 lysine 27 methylation in Polycomb-group silencing.'. *Science (New York, N.Y.)* **298**(5595), 1039–43.

Capelson, M. and M. W. Hetzer: 2009, 'The role of nuclear pores in gene regulation, development and disease.'. *EMBO reports* **10**(7), 697–705.

Caron, H., B. van Schaik, M. van der Mee, F. Baas, G. Riggins, P. van Sluis, M. C. Hermus, R. van Asperen, K. Boon, P. a. Voûte, S. Heisterkamp, a. van Kampen, and R. Versteeg: 2001, 'The human transcriptome map: clustering of highly expressed genes in chromosomal domains.'. *Science (New York, N.Y.)* **291**(5507), 1289–92.

Carter, D., L. Chakalova, C. S. Osborne, Y.-f. Dai, and P. Fraser: 2002, 'Long-range chromatin regulatory interactions in vivo.'. *Nature genetics* **32**(4), 623–6.

Caspersson, T., L. Zech, and C. Johansson: 1970, 'Analysis of human metaphase chromosome set by aid of DNA-binding fluorescent agents.'. *Experimental cell research* **62**(2), 490–2.

Chakalova, L., E. Debrand, J. a. Mitchell, C. S. Osborne, and P. Fraser: 2005, 'Replication and transcription: shaping the landscape of the genome.'. *Nature reviews. Genetics* **6**(9), 669–77.

Chambeyron, S. and W. a. Bickmore: 2004, 'Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription.'. *Genes & development* **18**(10), 1119–30.

Chen, C. S., P. H. Sorensen, P. H. Domer, G. H. Reaman, S. J. Korsmeyer, N. A. Heerema, G. D. Hammond, and J. H. Kersey: 1993, 'Molecular rearrangements on chromosome 11q23 predominate in infant acute lymphoblastic leukemia and are associated with specific biologic variables and poor outcome'. *Blood* **81**(9), 2386–2393.

Chen, X., H. Xu, P. Yuan, F. Fang, M. Huss, V. B. Vega, E. Wong, Y. L. Orlov, W. Zhang, J. Jiang, Y.-H. Loh, H. C. Yeo, Z. X. Yeo, V. Narang, K. R. Govindarajan, B. Leong, A. Shahab, Y. Ruan, G. Bourque, W.-K. Sung, N. D. Clarke, C.-L. Wei, and H.-H. Ng: 2008, 'Integration of external signaling pathways with the core transcriptional network in embryonic stem cells.'. *Cell* **133**(6), 1106–17.

Chesterton, C. J., B. E. Coupar, and P. H. Butterworth: 1974, 'Transcription of fractionated mammalian chromatin by mammalian ribonucleic acid polymerase. Demonstration of temperature-dependent rifampicin-resistant initiation sites in euchromatin deoxyribonucleic acid.'. *The Biochemical journal* **143**(1), 73–81.

Chuang, C.-H., A. E. Carpenter, B. Fuchsova, T. Johnson, P. de Lanerolle, and A. S. Belmont: 2006, 'Long-range directional movement of an interphase chromosome site.'. *Current biology : CB* **16**(8), 825–31.

Chubb, J. R., S. Boyle, P. Perry, and W. a. Bickmore: 2002, 'Chromatin motion is constrained by association with nuclear compartments in human cells.'. *Current biology : CB* **12**(6), 439–45.

Chung, J. H., a. C. Bell, and G. Felsenfeld: 1997, 'Characterization of the chicken beta-globin insulator.'. *Proceedings of the National Academy of Sciences of the United States of America* **94**(2), 575–80.

Clark, M. B., P. P. Amaral, F. J. Schlesinger, M. E. Dinger, R. J. Taft, J. L. Rinn, C. P. Ponting, P. F. Stadler, K. V. Morris, A. Morillon, J. S. Rozowsky, M. B. Gerstein, C. Wahlestedt, Y. Hayashizaki, P. Carninci, T. R. Gingeras, and J. S. Mattick: 2011, 'The reality of pervasive transcription.'. *PLoS biology* **9**(7), e1000625; discussion e1001102.

Collins, I., A. Weber, and D. Levens: 2001, 'Transcriptional consequences of topoisomerase inhibition.'. *Molecular and cellular biology* **21**(24), 8437–51.

Cook, P. R.: 1999, 'The Organization of Replication and Transcription'. *Science* **284**(5421), 1790–1795.

Cook, P. R.: 2002, 'Predicting three-dimensional genome structure from transcriptional activity.'. *Nature genetics* **32**(3), 347–52.

Corso, A., M. Lazzarino, E. Morra, S. Merante, C. Astori, P. Bernasconi, M. Boni, and C. Bernasconi: 1995, 'Chronic myelogenous leukemia and exposure to ionizing radiation–a retrospective study of 443 patients.'. *Annals of hematology* **70**(2), 79–82.

Cowell, I. G., Z. Sondka, K. Smith, K. C. Lee, C. M. Manville, M. Sidorczuk-Lesthuruge, H. A. Rance, K. Padget, G. H. Jackson, N. Adachi, and C. a. Austin: 2012, 'Model for MLL translocations in therapy-related leukemia involving topoisomerase IIβ-mediated DNA strand breaks and gene proximity.'. *Proceedings of the National Academy of Sciences of the United States of America* **109**(23), 8989–94.

Cremer, M., J. von Hase, T. Volm, A. Brero, G. Kreth, J. Walter, C. Fischer, I. Solovei, C. Cremer, and T. Cremer: 2001, 'Non-random radial higher-order chromatin arrangements in nuclei of diploid human cells.'. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* **9**(7), 541–67.

Cremer, T. and C. Cremer: 2001, 'Chromosome territories, nuclear architecture and gene regulation in mammalian cells.'. *Nature reviews. Genetics* **2**(4), 292–301.

Cremer, T., C. Cremer, H. Baumann, E. K. Luedtke, K. Sperling, V. Teuber, and C. Zorn: 1982, 'Rabl's model of the interphase chromosome arrangement tested in Chinese hamster cells by premature chromosome condensation and laser-UV-microbeam experiments.'. *Human genetics* **60**(1), 46–56.

Cremer, T., A. Kurz, R. Zirbel, S. Dietzel, B. Rinke, E. Schrock, M. Speicher, U. Mathieu, A. Jauch, P. Emmerich, H. Scherthan, T. Ried, C. Cremer, and P. Lichter: 1993, 'Role of Chromosome Territories in the Functional Compartmentalization of the Cell Nucleus'. *Cold Spring Harbor Symposia on Quantitative Biology* **58**(0), 777–792.

Croft, J. A., J. M. Bridger, S. Boyle, P. Perry, P. Teague, and W. A. Bickmore: 1999, 'Differences in the localization and morphology of chromosomes in the human nucleus'. *Journal of Cell Biology* **145**(6), 1119–1131.

Darai-Ramqvist, E., A. Sandlund, S. Müller, G. Klein, S. Imreh, and M. Kost-Alimova: 2008, 'Segmental duplications and evolutionary plasticity at tumor chromosome break-prone regions.'. *Genome research* **18**(3), 370–9.

Davey, C. a., D. F. Sargent, K. Luger, A. W. Maeder, and T. J. Richmond: 2002, 'Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution.'. *Journal of molecular biology* **319**(5), 1097–113.

de Laat, W. and F. Grosveld: 2003, 'Spatial organization of gene expression: the active chromatin hub.'. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* **11**(5), 447–59.

Deininger, M. W., J. M. Goldman, and J. V. Melo: 2000, 'The molecular biology of chronic myeloid leukemia.'. *Blood* **96**(10), 3343–56.

Dekker, J., K. Rippe, M. Dekker, and N. Kleckner: 2002, 'Capturing chromosome conformation'. *Science* **295**(5558), 1306–1311.

Dietzel, S., K. Zolghadr, C. Hepperger, and A. S. Belmont: 2004, 'Differential large-scale chromatin compaction and intranuclear positioning of transcribed versus non-transcribed transgene arrays containing beta-globin regulatory sequences.'. *Journal of cell science* **117**(Pt 19), 4603–14.

Dorsett, D.: 1999, 'Distant liaisons: long-range enhancer-promoter interactions in Drosophila.'. *Current opinion in genetics & development* **9**(5), 505–14.

Dunleavy, E. M., D. Roche, H. Tagami, N. Lacoste, D. Ray-Gallet, Y. Nakamura, Y. Daigo, Y. Nakatani, and G. Almouzni-Pettinotti: 2009, 'HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres.'. *Cell* **137**(3), 485–97.

Ernst, J. and M. Kellis: 2010, 'Discovery and characterization of chromatin states for systematic annotation of the human genome.'. *Nature biotechnology* **28**(8), 817–25.

Ernst, J., P. Kheradpour, T. S. Mikkelsen, N. Shoresh, L. D. Ward, C. B. Epstein, X. Zhang, L. Wang, R. Issner, M. Coyne, M. Ku, T. Durham, M. Kellis, and B. E. Bernstein: 2011, 'Mapping and analysis of chromatin state dynamics in nine human cell types.'. *Nature* **473**(7345), 43–9.

Eskiw, C. H., A. Rapp, D. R. F. Carter, and P. R. Cook: 2008, 'RNA polymerase II activity is located on the surface of protein-rich transcription factories.'. *Journal of cell science* **121**(Pt 12), 1999–2007.

Ferreira, J. a., G. Paolella, C. Ramos, and A. I. Lamond: 1997, 'Spatial organization of large-scale chromatin domains in the nucleus: a magnified view of single chromosome territories.'. *The Journal of cell biology* **139**(7), 1597–610.

Filippova, G. N., S. Fagerlie, E. M. Klenova, C. Myers, Y. Dehner, G. Goodwin, P. E. Neiman, S. J. Collins, and V. V. Lobanenkov: 1996, 'An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes.'. *Molecular and cellular biology* **16**(6), 2802–13.

Finlan, L. E., D. Sproul, I. Thomson, S. Boyle, E. Kerr, P. Perry, B. Ylstra, J. R. Chubb, and W. A. Bickmore: 2008, 'Recruitment to the nuclear periphery can alter expression of genes in human cells'. *PLoS Genet* **4**(3), e1000039.

Foltz, D. R., L. E. T. Jansen, A. O. Bailey, J. R. Yates, E. a. Bassett, S. Wood, B. E. Black, and D. W. Cleveland: 2009, 'Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP.'. *Cell* **137**(3), 472–84.

Fuss, S. H., M. Omura, and P. Mombaerts: 2007, 'Local and cis effects of the H element on expression of odorant receptor genes in mouse.'. *Cell* **130**(2), 373–84.

Gavrilov, A. a., I. S. Zukher, E. S. Philonenko, S. V. Razin, and O. V. Iarovaia: 2010, 'Mapping of the nuclear matrix-bound chromatin hubs by a new M3C experimental procedure.'. *Nucleic acids research* **38**(22), 8051–60.

Geary, C. G.: 2000, 'The story of chronic myeloid leukaemia.'. *British journal of haematology* **110**(1), 2–11.

Goren, A. and H. Cedar: 2003, 'Replicating by the clock.'. *Nature reviews. Molecular cell biology* **4**(1), 25–32.

Grande, M. a., I. van der Kraan, L. de Jong, and R. van Driel: 1997, 'Nuclear distribution of transcription factors in relation to sites of transcription and RNA polymerase II.'. *Journal of cell science* **110 (Pt 1**, 1781–91.

Gu, W., F. Zhang, and J. R. Lupski: 2008, 'Mechanisms for human genomic rearrangements.'. *PathoGenetics* **1**(1), 4.

Gu, Y., H. Alder, T. Nakamura, S. A. Schichman, R. Prasad, O. Canaani, H. Saito, C. M. Croce, and E. Canaani: 1994, 'Sequence analysis of the breakpoint cluster region in the ALL-1 gene involved in acute leukemia'. *Cancer Research* **54**(9), 2327–2330.

Guelen, L., L. Pagie, E. Brasset, W. Meuleman, M. B. Faza, W. Talhout, B. H. Eussen, A. de Klein, L. Wessels, W. de Laat, and B. van Steensel: 2008, 'Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions.'. *Nature* **453**(7197), 948–51.

Haber, J. E.: 2000, 'Partners and pathwaysrepairing a double-strand break.'. *Trends in genetics : TIG* **16**(6), 259–64.

Hadjur, S., L. M. Williams, N. K. Ryan, B. S. Cobb, T. Sexton, P. Fraser, A. G. Fisher, and M. Merkenschlager: 2009, 'Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus.'. *Nature* **460**(7253), 410–3.

Hagstrom, K. a. and B. J. Meyer: 2003, 'Condensin and cohesin: more than chromosome compactor and glue.'. *Nature reviews. Genetics* **4**(7), 520–34.

Hamburger, A. W. and S. E. Salmon: 1977, 'Primary bioassay of human tumor stem cells'. *Science* **197**(4302), 461–463.

Handoko, L., H. Xu, G. Li, C. Y. Ngan, E. Chew, M. Schnapp, C. W. H. Lee, C. Ye, J. L. H. Ping, F. Mulawadi, E. Wong, J. Sheng, Y. Zhang, T. Poh, C. S. Chan, G. Kunarso, A. Shahab, G. Bourque, V. Cacheux-Rataboul, W.-K. Sung, Y. Ruan, and C.-L. Wei: 2011, 'CTCF-mediated functional chromatin interactome in pluripotent cells.'. *Nature genetics* **43**(7), 630–8.

Harewood, L., F. Schütz, S. Boyle, P. Perry, M. Delorenzi, W. a. Bickmore, and A. Reymond: 2010, 'The effect of translocation-induced nuclear reorganization on gene expression.'. *Genome research* **20**(5), 554–64.

Hark, a. T., C. J. Schoenherr, D. J. Katz, R. S. Ingram, J. M. Levorse, and S. M. Tilghman: 2000, 'CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus.'. *Nature* **405**(6785), 486–9.

Hehlmann, R., A. Hochhaus, and M. Baccarani: 2007, 'Chronic myeloid leukaemia.'. *Lancet* **370**(9584), 342–50.

Heintzman, N. D., R. K. Stuart, G. Hon, Y. Fu, C. W. Ching, R. D. Hawkins, L. O. Barrera, S. Van Calcar, C. Qu, K. a. Ching, W. Wang, Z. Weng, R. D. Green, G. E. Crawford, and B. Ren: 2007, 'Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome.'. *Nature genetics* **39**(3), 311–8.

Heitz, E.: 1928, 'Das heterochromatin der moose.'. *I Jahrb Wiss Botanik* **69**, 762–818.

Hess, J. L., B. D. Yu, B. Li, R. Hanson, and S. J. Korsmeyer: 1997, 'Defects in yolk sac hematopoiesis in Mll-null embryos'. *Blood* **90**(5), 1799–1806.

Hlatky, L., R. K. Sachs, M. Vazquez, and M. N. Cornforth: 2002, 'Radiation-induced chromosome aberrations: insights gained from biophysical modeling.'. *BioEssays : news and reviews in molecular, cellular and developmental biology* **24**(8), 714–23.

Hofmann, W., B. Reichart, a. Ewald, E. Müller, I. Schmitt, R. H. Stauber, F. Lottspeich, B. M. Jockusch, U. Scheer, J. Hauber, and M. C. Dabauvalle: 2001, 'Cofactor requirements for nuclear export of Rev response element (RRE)- and constitutive transport element (CTE)-containing retroviral RNAs. An unexpected role for actin.'. *The Journal of cell biology* **152**(5), 895–910.

Hofmann, W. A. and P. de Lanerolle: 2006, 'Nuclear actin: to polymerize or not to polymerize.'. *The Journal of cell biology* **172**(4), 495–6.

Hofmann, W. a., L. Stojiljkovic, B. Fuchsova, G. M. Vargas, E. Mavrommatis, V. Philimonenko, K. Kysela, J. a. Goodrich, J. L. Lessard, T. J. Hope, P. Hozak, and P. de Lanerolle: 2004, 'Actin is part of pre-initiation complexes and is necessary for transcription by RNA polymerase II.'. *Nature cell biology* **6**(11), 1094–101.

Holley, W. R., I. S. Mian, S. J. Park, B. Rydberg, and A. Chatterjee: 2002, 'A model for interphase chromosomes and evaluation of radiation-induced aberrations.'. *Radiation research* **158**(5), 568–80.

Hon, G., W. Wang, and B. Ren: 2009, 'Discovery and annotation of functional chromatin signatures in the human genome.'. *PLoS computational biology* **5**(11), e1000566.

Hou, C., R. Dale, and A. Dean: 2010, 'Cell type specificity of chromatin organization mediated by CTCF and cohesin.'. *Proceedings of the National Academy of Sciences of the United States of America* **107**(8), 3651–6.

Hudson, D. F., K. M. Marshall, and W. C. Earnshaw: 2009, 'Condensin: Architect of mitotic chromosomes.'. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* **17**(2), 131–44.

Huret, J. L., P. Dessen, and A. Bernheim: 2001, 'An atlas of chromosomes in hematological malignancies. Example: 11q23 and MLL partners.'. *Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, U.K* **15**(6), 987–9.

Iborra, F. J., a. Pombo, D. a. Jackson, and P. R. Cook: 1996, 'Active RNA polymerases are localized within discrete transcription "factories' in human nuclei.'. *Journal of cell science* **109 (Pt 6**, 1427–36.

Jackson, D. and P. Cook: 1985, 'Transcription occurs at a nucleoskeleton.'. *The EMBO Journal* **4**(4), 919.

Jackson, D., A. Hassan, R. Errington, and P. Cook: 1993, 'Visualization of focal sites of transcription within human nuclei.'. *The EMBO Journal* **12**(3), 1059.

Jackson, D. A., F. J. Iborra, E. M. Manders, and P. R. Cook: 1998, 'Numbers and organization of RNA polymerases, nascent transcripts, and transcription units in HeLa nuclei.'. *Molecular biology of the cell* **9**(6), 1523–36.

Jackson, D. A., S. J. McCready, and P. R. Cook: 1981, 'RNA is synthesized at the nuclear cage'. *Nature* **292**(5823), 552–555.

Jackson, D. a. and a. Pombo: 1998, 'Replicon clusters are stable units of chromosome structure: evidence that nuclear organization contributes to the efficient activation and propagation of S phase in human cells.'. *The Journal of cell biology* **140**(6), 1285–95.

Jacob, F. and J. Monod: 1961, 'Genetic regulatory mechanisms in the synthesis of proteins'. *Journal of Molecular Biology* **3**(3), 318–356.

Jakob, B., J. Splinter, M. Durante, and G. Taucher-Scholz: 2009, 'Live cell microscopy analysis of radiation-induced DNA double-strand break motion.'. *Proceedings of the National Academy of Sciences of the United States of America* **106**(9), 3172–7.

Janga, S. C., J. Collado-Vides, and M. M. Babu: 2008, 'Transcriptional regulation constrains the organization of genes on eukaryotic chromosomes.'. *Proceedings of the National Academy of Sciences of the United States of America* **105**(41), 15761–6.

Jemal, A., R. Siegel, E. Ward, Y. Hao, J. Xu, T. Murray, and M. J. Thun: 2009, 'Cancer statistics, 2008.'. *CA: a cancer journal for clinicians* **58**(2), 71–96.

Joshi, R. S., B. Piña, and J. Roca: 2012, 'Topoisomerase II is required for the production of long Pol II gene transcripts in yeast.'. *Nucleic acids research* **40**(16), 1–9.

Ju, B.-G., V. V. Lunyak, V. Perissi, I. Garcia-Bassets, D. W. Rose, C. K. Glass, and M. G. Rosenfeld: 2006, 'A topoisomerase IIbeta-mediated dsDNA break required for regulated transcription.'. *Science (New York, N.Y.)* **312**(5781), 1798–802.

Jude, C. D., L. Climer, D. Xu, E. Artinger, J. K. Fisher, and P. Ernst: 2007, 'Unique and independent roles for MLL in adult hematopoietic stem cells and progenitors.'. *Cell stem cell* **1**(3), 324–37.

Kagey, M. H., J. J. Newman, S. Bilodeau, Y. Zhan, D. a. Orlando, N. L. van Berkum, C. C. Ebmeier, J. Goossens, P. B. Rahl, S. S. Levine, D. J. Taatjes, J. Dekker, and R. a. Young: 2010, 'Mediator and cohesin connect gene expression and chromatin architecture.'. *Nature* **467**(7314), 430–5.

Kalverda, B., H. Pickersgill, V. V. Shloma, and M. Fornerod: 2010, 'Nucleoporins directly stimulate expression of developmental and cell-cycle genes inside the nucleoplasm.'. *Cell* **140**(3), 360–71.

Khobta, A., C. Carlo-Stella, and G. Capranico: 2004, 'Specific histone patterns and acetylase/deacetylase activity at the breakpoint-cluster region of the human MLL gene'. *Cancer Research* **64**(8), 2656–2662.

Kim, T.-K., M. Hemberg, J. M. Gray, A. M. Costa, D. M. Bear, J. Wu, D. a. Harmin, M. Laptewicz, K. Barbara-Haley, S. Kuersten, E. Markenscoff-Papadimitriou, D. Kuhl, H. Bito, P. F. Worley, G. Kreiman, and M. E. Greenberg: 2010, 'Widespread transcription at neuronal activity-regulated enhancers.'. *Nature* **465**(7295), 182–7.

Kimura, H., Y. Tao, R. G. Roeder, and P. R. Cook: 1999, 'Quantitation of RNA polymerase II and its transcription factors in an HeLa cell: little soluble holoenzyme but significant amounts of polymerases attached to the nuclear substructure.'. *Molecular and cellular biology* **19**(8), 5383–92.

Kleinjan, D. a., A. Seawright, S. Mella, C. B. Carr, D. a. Tyas, T. I. Simpson, J. O. Mason, D. J. Price, and V. van Heyningen: 2006, 'Long-range downstream enhancers are essential for Pax6 expression.'. *Developmental biology* **299**(2), 563–81.

Kosak, S. T., J. A. Skok, K. L. Medina, R. Riblet, M. M. Le Beau, A. G. Fisher, and H. Singh: 2002, 'Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development'. *Science* **296**(5565), 158–162.

Kozubek, S., E. Lukásová, A. Marecková, M. Skalníková, M. Kozubek, E. Bártová, V. Kroha, E. Krahulcová, and J. Slotová: 1999, 'The topological organization of chromosomes 9 and 22 in cell nuclei has a determinative role in the induction of t(9,22) translocations and in the pathogenesis of t(9,22) leukemias.'. *Chromosoma* **108**(7), 426–35.

Krivtsov, A. V. and S. a. Armstrong: 2007, 'MLL translocations, histone modifications and leukaemia stem-cell development.'. *Nature reviews. Cancer* **7**(11), 823–33.

Kumaran, R. I. and D. L. Spector: 2008, 'A genetic locus targeted to the nuclear periphery in living cells maintains its transcriptional competence.'. *The Journal of cell biology* **180**(1), 51–65.

Kuroda, M., H. Tanabe, K. Yoshida, K. Oikawa, A. Saito, T. Kiyuna, H. Mizusawa, and K. Mukai: 2004, 'Alteration of chromosome positioning during adipocyte differentiation.'. *Journal of cell science* **117**(Pt 24), 5897–903.

Kurukuti, S., V. K. Tiwari, G. Tavoosidana, E. Pugacheva, A. Murrell, Z. Zhao, V. Lobanenkov, W. Reik, and R. Ohlsson: 2006, 'CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2.'. *Proceedings of the National Academy of Sciences of the United States of America* **103**(28), 10684–9.

Kurz, A., S. Lampel, J. E. Nickolenko, J. Bradl, A. Benner, R. M. Zirbel, T. Cremer, and P. Lichter: 1996, 'Active and inactive genes localize preferentially in the periphery of chromosome territories.'. *The Journal of cell biology* **135**(5), 1195–205.

Lander, E. S., L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, a. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J. P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, a. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, a. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, a. Coulson, R. Deadman, P. Deloukas, a. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, a. Hunt, M. Jones, C. Lloyd, a. McMurray, L. Matthews, S. Mercer, S. Milne, J. C. Mullikin, a. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R. H. Waterston, R. K. Wilson, L. W. Hillier, J. D. McPherson, M. a. Marra, E. R. Mardis, L. a. Fulton, a. T. Chinwalla, K. H. Pepin, W. R. Gish, S. L. Chissoe, M. C. Wendl, K. D. Delehaunty, T. L. Miner, a. Delehaunty, J. B. Kramer, L. L. Cook, R. S. Fulton, D. L. Johnson, P. J. Minx, S. W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J. F. Cheng, a. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, R. a. Gibbs, D. M. Muzny, S. E. Scherer, J. B. Bouck, E. J. Sodergren, K. C. Worley, C. M. Rives, J. H. Gorrell, M. L. Metzker, S. L. Naylor, R. S. Kucherlapati, D. L. Nelson, G. M. Weinstock, Y. Sakaki, a. Fujiyama, M. Hattori, T. Yada, a. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, D. R. Smith, L. Doucette-Stamm,

M. Rubenfield, K. Weinstock, H. M. Lee, J. Dubois, a. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, a. Rump, H. Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, a. Madan, S. Qin, R. W. Davis, N. a. Federspiel, a. P. Abola, M. J. Proctor, R. M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D. R. Cox, M. V. Olson, R. Kaul, N. Shimizu, K. Kawasaki, S. Minoshima, G. a. Evans, M. Athanasiou, R. Schultz, B. a. Roe, F. Chen, H. Pan, J. Ramser, H. Lehrach, R. Reinhardt, W. R. McCombie, M. de la Bastide, N. Dedhia, H. Blöcker, K. Hornischer, G. Nordsiek, R. Agarwala, L. Aravind, J. a. Bailey, a. Bateman, S. Batzoglou, E. Birney, P. Bork, D. G. Brown, C. B. Burge, L. Cerutti, H. C. Chen, D. Church, M. Clamp, R. R. Copley, T. Doerks, S. R. Eddy, E. E. Eichler, T. S. Furey, J. Galagan, J. G. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L. S. Johnson, T. a. Jones, S. Kasif, a. Kaspryzk, S. Kennedy, W. J. Kent, P. Kitts, E. V. Koonin, I. Korf, D. Kulp, D. Lancet, T. M. Lowe, a. McLysaght, T. Mikkelsen, J. V. Moran, N. Mulder, V. J. Pollara, C. P. Ponting, G. Schuler, J. Schultz, G. Slater, a. F. Smit, E. Stupka, J. Szustakowski, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, a. Williams, Y. I. Wolf, K. H. Wolfe, S. P. Yang, R. F. Yeh, F. Collins, M. S. Guyer, J. Peterson, a. Felsenfeld, K. a. Wetterstrand, a. Patrinos, M. J. Morgan, P. de Jong, J. J. Catanese, K. Osoegawa, H. Shizuya, S. Choi, Y. J. Chen, and J. Szustakowki: 2001, 'Initial sequencing and analysis of the human genome.'. *Nature* **409**(6822), 860–921.

Lee, G. R., P. E. Fields, T. J. Griffin, and R. a. Flavell: 2003, 'Regulation of the Th2 cytokine locus by a locus control region.'. *Immunity* **19**(1), 145–53.

Lee, P. S., P. W. Greenwell, M. Dominska, M. Gawel, M. Hamilton, and T. D. Petes: 2009, 'A fine-structure map of spontaneous mitotic crossovers in the yeast Saccharomyces cerevisiae.'. *PLoS genetics* **5**(3), e1000410.

Lercher, M. J., A. O. Urrutia, and L. D. Hurst: 2002, 'Clustering of housekeeping genes provides a unified model of gene order in the human genome.'. *Nature genetics* **31**(2), 180–3.

Lettice, L. A., S. J. H. Heaney, L. A. Purdie, L. Li, P. de Beer, B. A. Oostra, D. Goode, G. Elgar, R. E. Hill, and E. de Graaff: 2003, 'A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly.'. *Human molecular genetics* **12**(14), 1725–35.

Lettice, L. a., T. Horikoshi, S. J. H. Heaney, M. J. van Baren, H. C. van der Linde, G. J. Breedveld, M. Joosse, N. Akarsu, B. a. Oostra, N. Endo, M. Shibata, M. Suzuki, E. Takahashi, T. Shinka, Y. Nakahori, D. Ayusawa, K. Nakabayashi, S. W. Scherer, P. Heutink, R. E. Hill, and S. Noji: 2002, 'Disruption of a long-range cis-acting regulator for Shh causes preaxial polydactyly.'. *Proceedings of the National Academy of Sciences of the United States of America* **99**(11), 7548–53.

Li, Y.-Y., H. Yu, Z.-M. Guo, T.-Q. Guo, K. Tu, and Y.-X. Li: 2006, 'Systematic analysis of head-to-head gene organization: evolutionary conservation and potential biological relevance.'. *PLoS computational biology* **2**(7), e74.

Libura, J., D. J. Slater, C. a. Felix, and C. Richardson: 2005, 'Therapy-related acute myeloid leukemia-like MLL rearrangements are induced by etoposide in primary human CD34+ cells and remain stable after clonal expansion.'. *Blood* **105**(5), 2124–31.

Libura, J., M. Ward, J. Solecka, and C. Richardson: 2008, 'Etoposide-initiated MLL rearrangements detected at high frequency in human primitive hematopoietic stem cells with in vitro and in vivo long-term repopulating potential.'. *European journal of haematology* **81**(3), 185–95.

Lieberman-Aiden, E., N. L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, R. Sandstrom, B. Bernstein, M. a. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L. a. Mirny, E. S. Lander, and J. Dekker: 2009, 'Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome'. *Science* **326**(5950), 289–293.

Ling, J., B. Baibakov, W. Pi, B. M. Emerson, and D. Tuan: 2005, 'The HS2 enhancer of the beta-globin locus control region initiates synthesis of non-coding, polyadenylated RNAs independent of a cis-linked globin promoter.'. *Journal of molecular biology* **350**(5), 883–96.

Liu, L. and J. Wang: 1987, 'Supercoiling of the DNA template during transcription'. *Proceedings of the National Academy of Sciences* **84**(20), 7024.

Lobanenkov, V. V., R. H. Nicolas, V. V. Adler, H. Paterson, E. M. Klenova, A. V. Polotskaja, and G. H. Goodwin: 1990, 'A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-flanking sequence of the chicken c-myc gene.'. *Oncogene* **5**(12), 1743–53.

Lomvardas, S., G. Barnea, D. J. Pisapia, M. Mendelsohn, J. Kirkland, and R. Axel: 2006, 'Interchromosomal interactions and olfactory receptor choice.'. *Cell* **126**(2), 403–13.

Longhese, M. P., D. Mantiero, and M. Clerici: 2006, 'The cellular response to chromosome breakage.'. *Molecular microbiology* **60**(5), 1099–108.

Ma, H., J. Samarabandu, R. S. Devdhar, R. Acharya, P. C. Cheng, C. Meng, and R. Berezney: 1998, 'Spatial and temporal dynamics of DNA replication sites in mammalian cells.'. *The Journal of cell biology* **143**(6), 1415–25.

Mahy, N. L., P. E. Perry, and W. a. Bickmore: 2002a, 'Gene density and transcription influence the localization of chromatin outside of chromosome territories detectable by FISH.'. *The Journal of cell biology* **159**(5), 753–63.

Mahy, N. L., P. E. Perry, S. Gilchrist, R. a. Baldock, and W. a. Bickmore: 2002b, 'Spatial organization of active and inactive genes and noncoding DNA within chromosome territories.'. *The Journal of cell biology* **157**(4), 579–89.

Manuelidis, L.: 1985, 'Individual interphase chromosome domains revealed by in situ hybridization'. *Hum Genet* **71**(4), 288–293.

Marshall, W. F., A. Straight, J. F. Marko, J. Swedlow, A. Dernburg, A. Belmont, A. W. Murray, D. A. Agard, and J. W. Sedat: 1997, 'Interphase chromosomes undergo constrained diffusional motion in living cells.'. *Current biology : CB* **7**(12), 930–9.

Matutes, E., R. Morilla, N. Farahat, F. Carbonell, J. Swansbury, M. Dyer, and D. Catovsky: 1997, 'Definition of acute biphenotypic leukemia.'. *Haematologica* **82**(1), 64–6.

McDonald, D., G. Carrero, C. Andrin, G. de Vries, and M. J. Hendzel: 2006, 'Nucleoplasmic beta-actin exists in a dynamic equilibrium between low-mobility polymeric species and rapidly diffusing populations.'. *The Journal of cell biology* **172**(4), 541–52.

Mehta, I. S., M. Amira, A. J. Harvey, and J. M. Bridger: 2010, 'Rapid chromosome territory relocation by nuclear motor activity in response to serum removal in primary human fibroblasts'. *Genome Biol* **11**(1), R5.

Mewborn, S. K., M. J. Puckelwartz, F. Abuisneineh, J. P. Fahrenbach, Y. Zhang, H. MacLeod, L. Dellefave, P. Pytel, S. Selig, C. M. Labno, K. Reddy, H. Singh, and E. McNally: 2010, 'Altered chromosomal positioning, compaction, and gene expression with a lamin A/C gene mutation.'. *PloS one* **5**(12), e14342.

Meyer, C., B. Schneider, S. Jakob, S. Strehl, A. Attarbaschi, S. Schnittger, C. Schoch, M. W. J. C. Jansen, J. J. M. van Dongen, M. L. den Boer, R. Pieters, M.-G. Ennas, E. Angelucci, U. Koehl, J. Greil, F. Griesinger, U. Zur Stadt, C. Eckert, T. Szczepanski, F. K. Niggli, B. W. Schäfer, H. Kempski, H. J. M. Brady, J. Zuna, J. Trka, L. L. Nigro, A. Biondi, E. Delabesse, E. Macintyre, M. Stanulla, M. Schrappe, O. a. Haas, T. Burmeister, T. Dingermann, T. Klingebiel, and R. Marschalek: 2006, 'The MLL recombinome of acute leukemias.'. *Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, U.K* **20**(5), 777–84.

Michaelis, C., R. Ciosk, and K. Nasmyth: 1997, 'Cohesins: chromosomal proteins that prevent premature separation of sister chromatids.'. *Cell* **91**(1), 35–45.

Milne, T. A., Y. Dou, M. E. Martin, H. W. Brock, R. G. Roeder, and J. L. Hess: 2005, 'MLL associates specifically with a subset of transcriptionally active target genes'. *Proceedings of the National Academy of Sciences of the United States of America* **102**(41), 14765–14770.

Mitchell, J. a. and P. Fraser: 2008, 'Transcription factories are nuclear subcompartments that remain in the absence of transcription.'. *Genes & development* **22**(1), 20–5.

Mitelman, F., B. Johansson, and F. Mertens: 2007, 'The impact of translocations and gene fusions on cancer causation.'. *Nature reviews. Cancer* **7**(4), 233–45.

Murrell, A., S. Heeson, and W. Reik: 2004, 'Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops.'. *Nature genetics* **36**(8), 889–93.

Nativio, R., K. S. Wendt, Y. Ito, J. E. Huddleston, S. Uribe-Lewis, K. Woodfine, C. Krueger, W. Reik, J.-M. Peters, and A. Murrell: 2009, 'Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus.'. *PLoS genetics* **5**(11), e1000739.

Nelms, B. E., R. S. Maser, J. F. MacKay, M. G. Lagally, and J. H. Petrini: 1998, 'In situ visualization of DNA double-strand break repair in human fibroblasts'. *Science* **280**(5363), 590–592.

Neves, H., C. Ramos, M. G. Da Silva, A. Parreira, and L. Parreira: 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.

Noordermeer, D., M. Leleu, E. Splinter, J. Rougemont, W. De Laat, and D. Duboule: 2011, 'The Dynamic Architecture of Hox Gene Clusters'. *Science* **334**(6053), 222–225.

Nowell, P. and D. Hungerford: 1960, 'A minute chromosome in human chronic granulocytic leukemia'. *Science* **132**, 1497–501.

Ohlsson, R., R. Renkawitz, and V. Lobanenkov: 2001, 'CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease.'. *Trends in genetics : TIG* **17**(9), 520–7.

Osborne, C. S., L. Chakalova, K. E. Brown, D. Carter, A. Horton, E. Debrand, B. Goyenechea, J. a. Mitchell, S. Lopes, W. Reik, and P. Fraser: 2004, 'Active genes dynamically colocalize to shared sites of ongoing transcription.'. *Nature genetics* **36**(10), 1065–71.

Osborne, C. S., L. Chakalova, J. a. Mitchell, A. Horton, A. L. Wood, D. J. Bolland, A. E. Corcoran, and P. Fraser: 2007, 'Myc dynamically and preferentially relocates to a transcription factory occupied by Igh.'. *PLoS biology* **5**(8), e192.

Osborne, C. S., P. A. Ewels, and A. N. C. Young: 2011, 'Meet the neighbours: tools to dissect nuclear structure and function.'. *Briefings in functional genomics* **10**(1), 11–7.

Palstra, R.-J., B. Tolhuis, E. Splinter, R. Nijmeijer, F. Grosveld, and W. de Laat: 2003, 'The beta-globin nuclear compartment in development and erythroid differentiation.'. *Nature genetics* **35**(2), 190–4.

Papantonis, A., J. D. Larkin, Y. Wada, Y. Ohta, S. Ihara, T. Kodama, and P. R. Cook: 2010, 'Active RNA polymerases: mobile or immobile molecular machines?'. *PLoS biology* **8**(7), e1000419.

Parada, L. a., P. G. McQueen, and T. Misteli: 2004, 'Tissue-specific spatial organization of genomes.'. *Genome biology* **5**(7), R44.

Parada, L. A., P. G. McQueen, P. J. Munson, and T. Misteli: 2002, 'Conservation of relative chromosome positioning in normal and cancer cells.'. *Current biology : CB* **12**(19), 1692–7.

Parelho, V., S. Hadjur, M. Spivakov, M. Leleu, S. Sauer, H. C. Gregson, A. Jarmuz, C. Canzonetta, Z. Webster, T. Nesterova, B. S. Cobb, K. Yokomori, N. Dillon, L. Aragon, A. G. Fisher, and M. Merkenschlager: 2008, 'Cohesins functionally associate with CTCF on mammalian chromosome arms.'. *Cell* **132**(3), 422–33.

Pestic-Dragovich, L., L. Stojiljkovic, A. A. Philimonenko, G. Nowak, Y. Ke, R. E. Settlage, J. Shabanowitz, D. F. Hunt, P. Hozak, and P. de Lanerolle: 2000, 'A myosin I isoform in the nucleus.'. *Science (New York, N.Y.)* **290**(5490), 337–41.

Pirrotta, V.: 1998, 'Polycombing the genome: PcG, trxG, and chromatin silencing.'. *Cell* **93**(3), 333–6.

Potocki, L., W. Bi, D. Treadwell-Deering, C. M. B. Carvalho, A. Eifert, E. M. Friedman, D. Glaze, K. Krull, J. a. Lee, R. A. Lewis, R. Mendoza-Londono, P. Robbins-Furman, C. Shaw, X. Shi, G. Weissenberger, M. Withers, S. a. Yatsenko, E. H. Zackai, P. Stankiewicz, and J. R. Lupski: 2007, 'Characterization of Potocki-Lupski syndrome (dup(17)(p11.2p11.2)) and delineation of a dosage-sensitive critical interval that can convey an autism phenotype.'. *American journal of human genetics* **80**(4), 633–49.

Rabl, C.: 1885, 'Uber Zelltheilung'. *Morphologisches Jahrbuch* **10**, 214–330.

Reddy, K. L., J. M. Zullo, E. Bertolino, and H. Singh: 2008, 'Transcriptional repression mediated by repositioning of genes to the nuclear lamina.'. *Nature* **452**(7184), 243–7.

Reya, T., S. J. Morrison, M. F. Clarke, and I. L. Weissman: 2001, 'Stem cells, cancer, and cancer stem cells.'. *Nature* **414**(6859), 105–11.

Roix, J. J., P. G. McQueen, P. J. Munson, L. a. Parada, and T. Misteli: 2003, 'Spatial proximity of translocation-prone gene loci in human lymphomas.'. *Nature genetics* **34**(3), 287–91.

Rowley, J. D.: 1973, 'Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining.'. *Nature* **243**(5405), 290–3.

Rowley, J. D.: 2001, 'Chromosome translocations: dangerous liaisons revisited.'. *Nature reviews. Cancer* **1**(3), 245–50.

Rubio, E. D., D. J. Reiss, P. L. Welcsh, C. M. Disteche, G. N. Filippova, N. S. Baliga, R. Aebersold, J. a. Ranish, and A. Krumm: 2008, 'CTCF physically links cohesin to chromatin.'. *Proceedings of the National Academy of Sciences of the United States of America* **105**(24), 8309–14.

Sadri-Vakili, G. and J.-H. J. Cha: 2006, 'Mechanisms of disease: Histone modifications in Huntington's disease.'. *Nature clinical practice. Neurology* **2**(6), 330–8.

Sanyal, A., B. R. Lajoie, G. Jain, and J. Dekker: 2012, 'The long-range interaction landscape of gene promoters'. *Nature* **489**(7414), 109–113.

Savageau, M. a.: 2011, 'Design of the lac gene circuit revisited.'. *Mathematical biosciences* **231**(1), 19–38.

Schardin, M., T. Cremer, H. D. Hager, and M. Lang: 1985, 'Specific staining of human chromosomes in Chinese hamster x man hybrid cell lines demonstrates interphase chromosome territories.'. *Human genetics* **71**(4), 281–7.

Scharf, S., J. Zech, A. Bursen, D. Schraets, P. L. Oliver, S. Kliem, E. Pfitzner, E. Gillert, T. Dingermann, and R. Marschalek: 2007, 'Transcription linked to recombination: a gene-internal promoter coincides with the recombination hot spot II of the human MLL gene'. *Oncogene* **26**(10), 1361–1371.

Schermelleh, L., P. M. Carlton, S. Haase, L. Shao, L. Winoto, P. Kner, B. Burke, M. C. Cardoso, D. a. Agard, M. G. L. Gustafsson, H. Leonhardt, and J. W. Sedat: 2008, 'Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy.'. *Science (New York, N.Y.)* **320**(5881), 1332–6.

Schmidt, D., P. C. Schwalie, C. S. Ross-Innes, A. Hurtado, G. D. Brown, J. S. Carroll, P. Flicek, and D. T. Odom: 2010, 'A CTCF-independent role for cohesin in tissue-specific transcription.'. *Genome research* **20**(5), 578–88.

Schneider, R. and R. Grosschedl: 2007, 'Dynamics and interplay of nuclear architecture, genome organization, and gene expression.'. *Genes & development* **21**(23), 3027–43.

Schoenfelder, S., T. Sexton, L. Chakalova, N. F. Cope, A. Horton, S. Andrews, S. Kurukuti, J. a. Mitchell, D. Umlauf, D. S. Dimitrova, C. H. Eskiw, Y. Luo, C.-L. L. Wei, Y. Ruan, J. J. Bieker, and P. Fraser: 2010, 'Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells'. *Nat Genet* **42**(1), 53–61.

Schwacha, a. and N. Kleckner: 1997, 'Interhomolog bias during meiotic recombination: meiotic functions promote a highly differentiated interhomolog-only pathway.'. *Cell* **90**(6), 1123–35.

Schwartz, B. E. and K. Ahmad: 2005, 'Transcriptional activation triggers deposition and removal of the histone variant H3.3.'. *Genes & development* **19**(7), 804–14.

Sharpe, J., L. Lettice, J. Hecksher-Sorensen, M. Fox, R. Hill, and R. Krumlauf: 1999, 'Identification of sonic hedgehog as a candidate gene responsible for the polydactylous mouse mutant Sasquatch.'. *Current biology : CB* **9**(2), 97–100.

Shimi, T., K. Pfleghaar, S.-i. Kojima, C.-G. Pack, I. Solovei, A. E. Goldman, S. a. Adam, D. K. Shumaker, M. Kinjo, T. Cremer, and R. D. Goldman: 2008, 'The A- and B-type nuclear lamin networks: microdomains involved in chromatin organization and transcription.'. *Genes & development* **22**(24), 3409–21.

Slany, R. K., C. Lavau, and M. L. Cleary: 1998, 'The oncogenic capacity of HRX-ENL requires the transcriptional transactivation activity of ENL and the DNA binding motifs of HRX.'. *Molecular and cellular biology* **18**(1), 122–9.

Soutoglou, E., J. F. Dorn, K. Sengupta, M. Jasin, A. Nussenzweig, T. Ried, G. Danuser, and T. Misteli: 2007, 'Positional stability of single double-strand breaks in mammalian cells.'. *Nature cell biology* **9**(6), 675–82.

Spilianakis, C. G. and R. a. Flavell: 2004, 'Long-range intrachromosomal interactions in the T helper type 2 cytokine locus.'. *Nature immunology* **5**(10), 1017–27.

Spilianakis, C. G., M. D. Lalioti, T. Town, G. R. Lee, and R. a. Flavell: 2005, 'Interchromosomal associations between alternatively expressed loci.'. *Nature* **435**(7042), 637–45.

Splinter, E., H. Heath, J. Kooren, R.-J. Palstra, P. Klous, F. Grosveld, N. Galjart, and W. de Laat: 2006, 'CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus.'. *Genes & development* **20**(17), 2349–54.

Stack, S. M., D. B. Brown, and W. C. Dewey: 1977, 'Visualization of interphase chromosomes.'. *Journal of cell science* **26**, 281–99.

Stadler, S., V. Schnapp, R. Mayer, S. Stein, C. Cremer, C. Bonifer, T. Cremer, and S. Dietzel: 2004, 'The architecture of chicken chromosome territories changes during differentiation.'. *BMC cell biology* **5**(1), 44.

Stedman, W., H. Kang, S. Lin, J. L. Kissil, M. S. Bartolomei, and P. M. Lieberman: 2008, 'Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators.'. *The EMBO journal* **27**(4), 654–66.

Strahl, B. D. and C. D. Allis: 2000, 'The language of covalent histone modifications.'. *Nature* **403**(6765), 41–5.

Strissel, P. L., R. Strick, J. D. Rowley, and N. J. Zeleznik-Le: 1998, 'An in vivo topoisomerase II cleavage site and a DNase I hypersensitive site colocalize near exon 9 in the MLL breakpoint cluster region'. *Blood* **92**(10), 3793–3803.

Sun, H. B., J. Shen, and H. Yokota: 2000, 'Size-dependent positioning of human chromosomes in interphase nuclei'. *Biophysical Journal* **79**(1), 184–190.

Sung, P. a., J. Libura, and C. Richardson: 2006, 'Etoposide and illegitimate DNA double-strand break repair in the generation of MLL translocations: new insights and new questions.'. *DNA repair* **5**(9-10), 1109–18.

Sutherland, H. and W. a. Bickmore: 2009, 'Transcription factories: gene expression in unions?'. *Nature reviews. Genetics* **10**(7), 457–66.

Sutherland, H. J., A. Blair, and R. W. Zapf: 1996, 'Characterization of a hierarchy in human acute myeloid leukemia progenitor cells'. *Blood* **87**(11), 4754–4761.

Szabó, P., S. H. Tang, a. Rentsendorj, G. P. Pfeifer, and J. R. Mann: 2000, 'Maternal-specific footprints at putative CTCF sites in the H19 imprinting control region give evidence for insulator function.'. *Current biology : CB* **10**(10), 607–10.

Szczerbal, I., H. A. Foster, and J. M. Bridger: 2009, 'The spatial repositioning of adipogenesis genes is correlated with their expression status in a porcine mesenchymal stem cell adipogenesis model system'. *Chromosoma* **118**(5), 647–663.

Talbert, P. B. and S. Henikoff: 2010, 'Histone variants–ancient wrap artists of the epigenome.'. *Nature reviews. Molecular cell biology* **11**(4), 264–75.

Tanabe, H., F. A. Habermann, I. Solovei, M. Cremer, and T. Cremer: 2002, 'Non-random radial arrangements of interphase chromosome territories: evolutionary considerations and functional implications'. *Mutation Research* **504**(1-2), 37–45.

Tanaka, K., M. Takechi, J. Hong, C. Shigeta, N. Oguma, N. Kamada, Y. Takimoto, A. Kuramoto, H. Dohy, and T. Kyo: 1989, '9;22 translocation and bcr rearrangements in chronic myelocytic leukemia patients among atomic bomb survivors.'. *Journal of radiation research* **30**(4), 352–8.

Tiwari, V. K., L. Cope, K. M. McGarvey, J. E. Ohm, and S. B. Baylin: 2008a, 'A novel 6C assay uncovers Polycomb-mediated higher order chromatin conformations.'. *Genome research* **18**(7), 1171–9.

Tiwari, V. K., K. M. McGarvey, J. D. F. Licchesi, J. E. Ohm, J. G. Herman, D. Schübeler, and S. B. Baylin: 2008b, 'PcG proteins, DNA methylation, and gene repression by chromatin looping.'. *PLoS biology* **6**(12), 2911–27.

Tolhuis, B., R. J. Palstra, E. Splinter, F. Grosveld, and W. de Laat: 2002, 'Looping and interaction between hypersensitive sites in the active beta-globin locus'. *Mol Cell* **10**(6), 1453–1465.

Torrano, V., I. Chernukhin, F. Docquier, V. D'Arcy, J. León, E. Klenova, and M. D. Delgado: 2005, 'CTCF regulates growth and erythroid differentiation of human myeloid leukemia cells.'. *The Journal of biological chemistry* **280**(30), 28152–61.

Trask, B. J.: 2002, 'Human cytogenetics: 46 chromosomes, 46 years and counting.'. *Nature reviews. Genetics* **3**(10), 769–78.

Trinklein, N. D., S. F. Aldred, S. J. Hartman, D. I. Schroeder, R. P. Otillar, and R. M. Myers: 2004, 'An abundance of bidirectional promoters in the human genome.'. *Genome research* **14**(1), 62–6.

Vaquerizas, J. M., R. Suyama, J. Kind, K. Miura, N. M. Luscombe, and A. Akhtar: 2010, 'Nuclear pore proteins nup153 and megator define transcriptionally active regions in the Drosophila genome.'. *PLoS genetics* **6**(2), e1000846.

Vazquez, J., A. S. Belmont, and J. W. Sedat: 2001, 'Multiple regimes of constrained chromosome motion are regulated in the interphase Drosophila nucleus.'. *Current biology : CB* **11**(16), 1227–39.

Velagaleti, G. V. N., G. a. Bien-Willner, J. K. Northup, L. H. Lockhart, J. C. Hawkins, S. M. Jalal, M. Withers, J. R. Lupski, and P. Stankiewicz: 2005, 'Position effects due to chromosome breakpoints that map approximately 900 Kb upstream and approximately 1.3 Mb downstream of SOX9 in two patients with campomelic dysplasia.'. *American journal of human genetics* **76**(4), 652–62.

Verschure, P. J., I. van Der Kraan, E. M. Manders, and R. van Driel: 1999, 'Spatial relationship between transcription sites and chromosome territories.'. *The Journal of cell biology* **147**(1), 13–24.

Versteeg, R., B. D. C. van Schaik, M. F. van Batenburg, M. Roos, R. Monajemi, H. Caron, H. J. Bussemaker, and A. H. C. van Kampen: 2003, 'The human transcriptome map reveals extremes in gene density, intron length, GC content, and repeat pattern for domains of highly and weakly expressed genes.'. *Genome research* **13**(9), 1998–2004.

Visser, a. E., R. Eils, A. Jauch, G. Little, P. J. Bakker, T. Cremer, and J. a. Aten: 1998, 'Spatial distributions of early and late replicating chromatin in interphase chromosome territories.'. *Experimental cell research* **243**(2), 398–407.

Vogel, M. J., L. Pagie, W. Talhout, M. Nieuwland, R. M. Kerkhoven, and B. van Steensel: 2009, 'High-resolution mapping of heterochromatin redistribution in a Drosophila position-effect variegation model.'. *Epigenetics & chromatin* **2**(1), 1.

Volpi, E. V., E. Chevret, T. Jones, R. Vatcheva, J. Williamson, S. Beck, R. D. Campbell, M. Goldsworthy, S. H. Powis, J. Ragoussis, J. Trowsdale, and D. Sheer: 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.'. *Journal of cell science* **113 (Pt 9**, 1565–76.

Vostrov, A. A. and W. W. Quitschke: 1997, 'The zinc finger protein CTCF binds to the APBbeta domain of the amyloid beta-protein precursor promoter. Evidence for a role in transcriptional activation.'. *The Journal of biological chemistry* **272**(52), 33353–9.

Wendt, K. S., K. Yoshida, T. Itoh, M. Bando, B. Koch, E. Schirghuber, S. Tsutsumi, G. Nagae, K. Ishihara, T. Mishiro, K. Yahata, F. Imamoto, H. Aburatani, M. Nakao, N. Imamoto, K. Maeshima, K. Shirahige, and J.-M. Peters: 2008, 'Cohesin mediates transcriptional insulation by CCCTC-binding factor.'. *Nature* **451**(7180), 796–801.

Wolffe, a. P. and J. J. Hayes: 1999, 'Chromatin disruption and modification.'. *Nucleic acids research* **27**(3), 711–20.

Würtele, H. and P. Chartrand: 2006, 'Genome-wide scanning of HoxB1-associated loci in mouse ES cells using an open-ended Chromosome Conformation Capture methodology.'. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* **14**(5), 477–95.

Xiao, T., J. Wallace, and G. Felsenfeld: 2011, 'Specific sites in the C terminus of CTCF interact with the SA2 subunit of the cohesin complex and are required for cohesin-dependent insulation activity.'. *Molecular and cellular biology* **31**(11), 2174–83.

Xie, X., T. S. Mikkelsen, A. Gnirke, K. Lindblad-Toh, M. Kellis, and E. S. Lander: 2007, 'Systematic discovery of regulatory motifs in conserved regions of the human genome, including thousands of CTCF insulator sites.'. *Proceedings of the National Academy of Sciences of the United States of America* **104**(17), 7145–50.

Xu, M. and P. R. Cook: 2008, 'Similar active genes cluster in specialized transcription factories.'. *The Journal of cell biology* **181**(4), 615–23.

Yin, H., M. D. Wang, K. Svoboda, R. Landick, S. M. Block, and J. Gelles: 1995, 'Transcription against an applied force.'. *Science (New York, N.Y.)* **270**(5242), 1653–7.

Yu, B. D., J. L. Hess, S. E. Horning, G. A. Brown, and S. J. Korsmeyer: 1995, 'Altered Hox expression and segmental identity in Mll-mutant mice'. *Nature* **378**(6556), 505–508.

Zeisig, B. B., T. Milne, M.-P. García-Cuéllar, S. Schreiner, M.-e. Martin, U. Fuchs, A. Borkhardt, S. K. Chanda, J. Walker, R. Soden, J. L. Hess, and R. K. Slany: 2004, 'Hoxa9 and Meis1 are key targets for MLL-ENL-mediated cellular immortalization.'. *Molecular and cellular biology* **24**(2), 617–28.

Zeleznik-Le, N. J., A. M. Harden, and J. D. Rowley: 1994, '11q23 translocations split the "AT-hook" cruciform DNA-binding region and the transcriptional repression domain from the activation domain of the mixed-lineage leukemia (MLL) gene'. *Proceedings of the National Academy of Sciences of the United States of America* **91**(22), 10610–10614.

Zhou, G.-L., L. Xin, W. Song, L.-J. Di, G. Liu, X.-S. Wu, D.-P. Liu, and C.-C. Liang: 2006, 'Active chromatin hub of the mouse alpha-globin locus forms in a transcription factory of clustered housekeeping genes.'. *Molecular and cellular biology* **26**(13), 5096–105.

Zhou, V. W., A. Goren, and B. E. Bernstein: 2011, 'Charting histone modifications and the functional organization of mammalian genomes.'. *Nature reviews. Genetics* **12**(1), 7–18.

Ziemin-van der Poel, S., N. R. McCabe, R. Espinosa 3rd, H. J. Gill, Y. Patel, A. Harden, P. Rubinelli, S. D. Smith, M. M. LeBeau, J. D. Rowley, and E. al.: 1991, 'Identification of a gene, MLL, that spans the breakpoint in 11q23 translocations associated with human leukemias'. *Proceedings of the National Academy of Sciences of the United States of America* **88**(23), 10735–10739.

Zink, D., M. D. Amaral, A. Englmann, S. Lang, L. a. Clarke, C. Rudolph, F. Alt, K. Luther, C. Braz, N. Sadoni, J. Rosenecker, and D. Schindelhauer: 2004, 'Transcription-dependent spatial arrangements of CFTR and adjacent genes in human cell nuclei.'. *The Journal of cell biology* **166**(6), 815–25.

Zirbel, R. M., U. R. Mathieu, A. Kurz, T. Cremer, and P. Lichter: 1993, 'Evidence for a nuclear compartment of transcription and splicing located at chromosome domain boundaries.'. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* **1**(2), 93–106.

Zlatanova, J. and A. Thakar: 2008, 'H2A.Z: view from the top.'. *Structure (London, England : 1993)* **16**(2), 166–79.