Characterization of an ectopic β-globin LCR

Good neighbors and a distant friend

Daan Noordermeer

ISBN/EAN: 978-90-79000-06-7
Deze publicatie is uitgegeven in samenwerking met Uitgeverij de Lawick Pers in Wageningen.

Cover design: Noël S. Keulen
Printed by: Wöhrmann Print Service, Zutphen

Characterization of an Ectopic β-Globin LCR

Good neighbors and a distant friend

Analyse van een Ectopisch Geïntegreerde β-Globine LCR

Goede buren en een verre vriend

Proefschrift

Ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de rector magnificus
Prof.dr. S.W.J. Lamberts
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
Woensdag 18 februari 2009 om 15.45

door

Daniël (Daan) Noordermeer

geboren te Bergen op Zoom

ERASMUS UNIVERSITEIT ROTTERDAM

Promotiecommissie

Promotor:	Prof.dr. F.G. Grosveld
Overige leden:	Dr. J. Gribnau Prof.dr. J.N.J. Philipsen Prof.dr. C. P. Verrijzer
Copromotor:	Prof.dr. W.L. de Laat

Dit proefschrift is tot stand gekomen binnen de vakgroep Celbiologie aan de faculteit der Geneeskunde en Gezondheidswetenschappen van het ErasmusMC te Rotterdam. De vakgroep maakt deel uit van het Medisch Genetisch Centrum Zuid-West Nederland.

Contents

	List of abbreviations	8
	Scope of the thesis	10
Chapter 1	Introduction: Nuclear and chromatin organisation	13
Chapter 2	Introduction: Joining the loops: β-globin gene regulation	71
Chapter 3	Transcription and chromatin organization of the housekeeping gene cluster 8C3/C4	95
Chapter 4	Transcription and chromatin organization of the housekeeping gene cluster 8C3/C4 containing an integrated β-globin Locus Control Region	117
Chapter 5	An ectopic β-globin LCR repositions its chromosomal integration site in the nucleus without searching for functionally related genes	149
Chapter 6	Trans-activation of an endogenous mouse β-globin gene by the human β-globin LCR and a β-globin gene inserted in another chromosome	179
Chapter 7	General discussion	207
	Summary	222
	Samenvatting	226
	Curriculum Vitae	230
	Publication list	231
	PhD portfolio	233
	Nawoord	234

List of abbreviations

3C	Chromosome conformation capture
4C	Chromosome conformation capture on chip *or* Circular chromosome conformation capture
5C	Chromosome conformation capture carbon copy
6C	Combined 3C – ChIP – cloning
3D	Deconvolution of DNA interactions by DSL
ACH	Active chromatin hub
ACT	Associated chromatin trap
bp	base pair
CH	Chromatin hub
CT	Chromosome territory
DMD / DMR	Differentially methylated domain / region
DNA	Deoxyribonucleic acid
ChIP	Chromatin immunoprecipitation
EDC	Epidermal differentiation complex
ES cell	Embryonic stem cell
FAIRE	Formaldehyde assisted identification of regulatory elements
FISH	Fluorescent *in situ* hybridization
FL	Fetal liver
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
HS	Hypersensitive site
kb	kilobase pair; one thousand base pairs
LAD	Lamina-associated domain
LCR	Locus control region
LTR	Long terminal repeat
MAR	Matrix attachment region
Mb	Megabase pair; one million base pairs
MBD	Methyl-CpG binding domain
MEL cell	Murine erythroid leukemia cell
MHC	Major Histocompatibility Complex
NuRD	Nucleosome remodeling and deacetylase

OR	Olfactory receptor
PcG	Polycomb group
PCR	Polymerase chain reaction
PML	Promyelocytic leukemia protein
qPCR	Quantitative polymerase chain reaction
qRT-PCR	Quantitative reverse transcriptase polymerase chain reaction
RIDGE	Region of increased gene expression
RNA	Ribonucleic acid
RNAP	RNA polymerase
SE	Standard error
WT	Wild-type

Scope of the thesis

The work discussed in this thesis is aimed at answering questions related to nuclear and chromatin organization, and to functioning of the β-globin LCR. In the introduction these topics are introduced and questions regarding their function are being posed. In the following four experimental chapters, the functional role of the LCR on nuclear and chromatin organization is addressed. In all chapters, the framework for the study is shortly introduced and the results are discussed within this framework. In **Chapter 7**, general conclusions are drawn from the total work in this thesis.

Introduction

In **Chapter 1** chromatin and nuclear organization are introduced. In **Paragraph 1.4** four different levels of organisation are discerned: epigenetic regulation, nuclear location of chromosomal loci, long-range chromatin interactions and interchromosomal transcriptional regulation. Each individual level of organisation has been studied intensively, though many important questions still remain. Therefore, one question that is addressed in this thesis is how the levels of chromatin and nuclear organisation influence each other.

In **Chapter 2** the β-globin Locus Control Region (LCR) and its relationship to chromatin and nuclear organisation are discussed. The LCR has been proposed to actively influence both chromatin and nuclear organisation, but thus far has mainly been studied as an integral part of the larger β-globin locus. The role of the LCR, independent from the linked β-globin genes and other regulatory elements, has therefore been difficult to discriminate. Therefore, the second topic that is treated in this thesis is how the LCR, independent from other elements in the β-globin locus, regulates gene-expression and how it influences nuclear and chromatin organisation.

Experimental chapters

In **Chapter 3**, "Transcription and chromatin organization of the housekeeping gene cluster 8C3/C4", the mouse gene-dense region 8C3/C4 is introduced. Analyses are presented at the level of gene expression, chromatin organisation and nuclear location of the region. The results increase our understanding of the relationship between the four different levels of organisation as introduced in **Paragraph 1.4**.

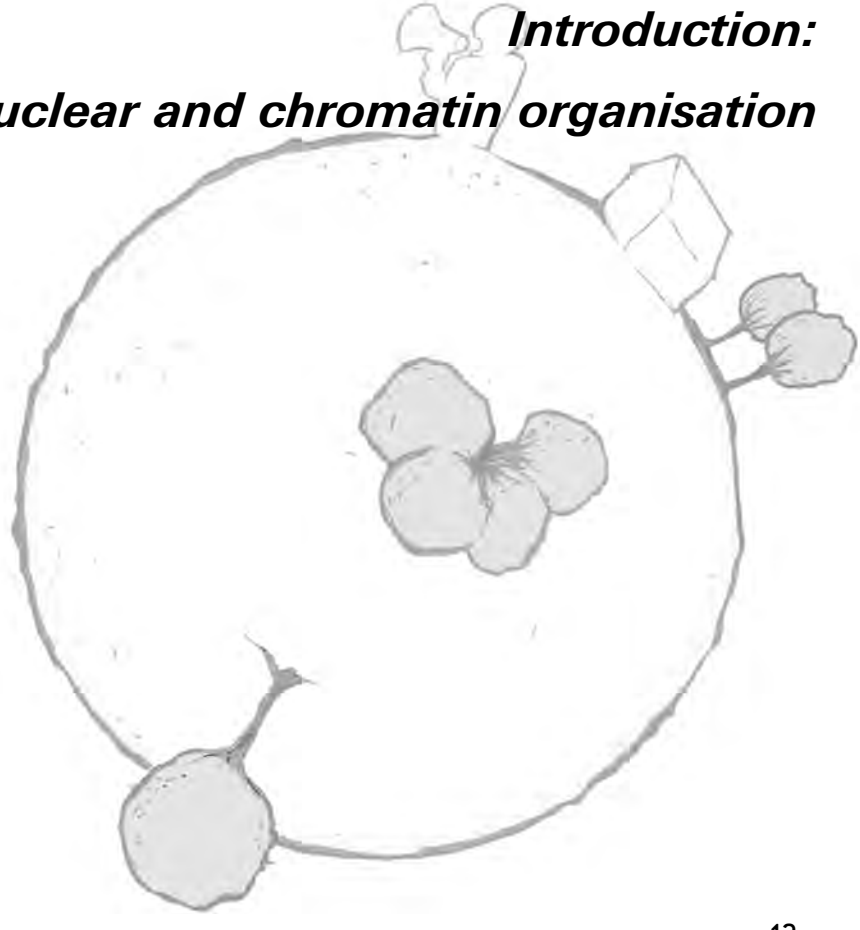
In **Chapter 4**, "Transcription and chromatin organization of the housekeeping gene cluster 8C3/C4 containing an integrated β-globin LCR", the effect of the integration of the human β-globin LCR in two orientations on 8C3/C4 is presented. Like in **Chapter 3**, gene expression, chromatin organisation and nuclear location of 8C3/C4 are examined. Integration of the LCR in two orientations does not only generate data on LCR function, but also allows the determination of hierarchy between these functions. Additionally, these results provide insight in the relationship between the different levels of organisation as introduced in **Paragraph 1.4**.

In **Chapter 5**, "An ectopic β-globin LCR repositions its chromosomal integration site in the nucleus without searching for functionally related genes", the intra- and inter-chromosomal associations of 8C3/C4 with and without the LCR are analyzed. The most important issue in this chapter is the question whether the LCR actively influences long-range and interchromosomal interactions. Furthermore, results in this chapter also shed light on how chromatin is spatially organized and to what extent this may be functionally relevant.

In **Chapter 6**, "Trans-activation of an endogenous mouse β-globin gene by the human β-globin LCR and a β-globin gene inserted in another chromosome", results are presented that indicate that both the human β-globin LCR and the human Aγ-globin gene at 8C3/C4 can functionally interact with one of the endogenous β-globin-like genes. Next to increasing our knowledge of interchromosomal transcriptional regulation, these results also have implications for LCR functioning and the establishment of long-range interactions.

Chapter 1

Introduction: Nuclear and chromatin organisation



1.1- The mammalian cell

The basic unit for building up organisms is the cell. 'Simple' organisms like bacteria and yeast usually are single cell organisms, but complex organisms like humans may consist of many trillions of cells. Cells in their basic form contain all machinery and information that is required for maintenance and proliferation [1]. An important difference in building plan of cells exists between eukaryotes and prokaryotes, where the former contain a nucleus and the latter don't. In mammals over 200 different cell types are found, all with specific functions and often very different morphologies (see Cellular Medical Subject Headings at http://en.wikipedia.org/wiki/Wikipedia:MeSH_A11). To perform all these functions, the cell contains many different specialised substructures (**Figure 1-1**).

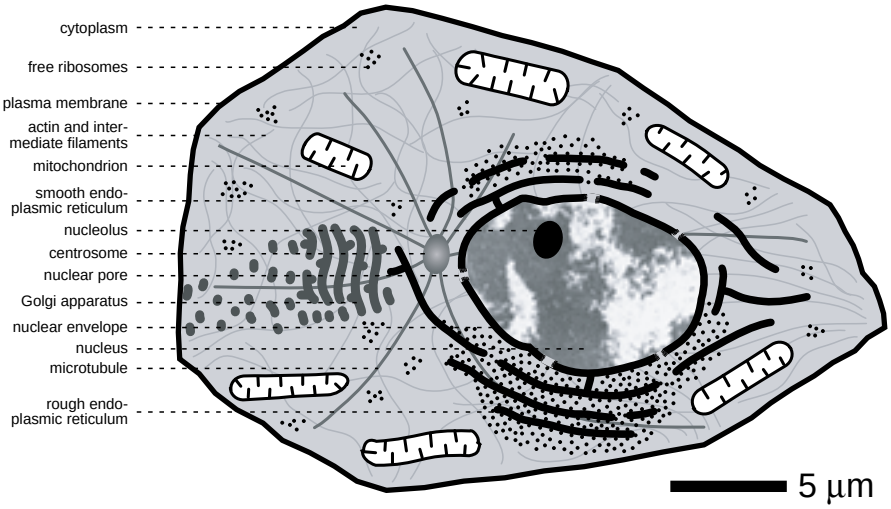


Figure 1-1. The mammalian cell. Example of a typical mammalian cell depicting a number, but not all, organelles and other specialised structures. A representative size bar is given below the cell.

Some of these substructures are physically separated from their surroundings by a membrane, like the cell itself is from its environment. Analogous to organs in the human body, these structures have been coined organelles. The most prominent organelle, and probably the most complex, is the nucleus containing the majority of DNA (**Figure 1-1**). The function and organisation of the nucleus will be discussed in further detail in the remainder of this introduction. Other important organelles and cellular substructures include (i) the ribosomes, that catalyse the synthesis of proteins from RNA and themselves consist of

both protein and RNA, (ii) the endoplasmic reticulum, which (a) is involved in folding, transport and sorting of proteins, (b) can be connected to the outer nuclear envelope and (c) can be covered with ribosomes at the outside (i.e. the rough endoplasmic reticulum), (iii) the mitochondria, that produce the large majority of ATP, the main intracellular energy carrier, (iv) the Golgi apparatus, involved in production, modification and sorting of many chemical compounds, (v) the centrosome, which is the microtubule organising centre and essential for the correct division of chromosomes during mitosis and meiosis, and (vi) actin filaments, intermediate filaments and microtubules, that are part of the cytoskeleton and have functions in intracellular transport and shaping of the cell ([1] and **Figure 1-1**). Next to these common structures, cells contain many other components that are not discussed. Some of these structures are found in many or all cell types (e.g. peroxisomes, lysosomes), while others are restricted to specific cell types with specialised functions [1].

1.2- The cell nucleus

The cell nucleus is the most prominent organelle in the cell and in mammals typically measures around 10 μm in diameter (**Figure 1-2**). The nucleus contains the large majority of the DNA, which is visible as a structure named chromatin [1]. Chromatin was initially discovered by Emil Heitz around 1928. Using conventional microscopy, Heitz found large dark and light areas in the interphase nucleus of both mosses and fruit flies (*Drosophila melanogaster*) [2-4]. The dark areas were originally termed heterochromatin and thought to be genetically inactive, while the lighter areas where named euchromatin and thought to consist of active chromatin [4, 5]. The molecular properties and organisation of chromatin will be further discussed in **Paragraphs 1.3** and **1.4** of this chapter.

The most distinct substructure in the nucleus is the nucleolus (**Figure 1-2**). It is clearly visible using conventional microscopy, even though it is not encapsulated by a membrane. The nucleolus is the site in the nucleus where all the steps required for ribosome assembly are performed. Structurally it is a large aggregate of active rRNA genes, RNA polymerases (RNAPs), the resulting transcribed rRNAs, ribosomal protein subunits and other proteins involved in rRNA-processing and ribosome assembly [6]. How the structural integrity of the nucleolus is maintained, or whether its integrity is actually required for proper function, is not clear [6].

The nucleus is physically separated from the cytoplasm by a double membrane, the nuclear envelope, that extents into the interior of the endoplasmic reticulum [7, 8] (**Figure**

1-2). To allow transport between the cytoplasm and the nucleus, the nuclear envelope is perforated by protein structures, called the nuclear pore complex (NPC) (**Figure 1-2**). The NPC and its associated proteins act as an aqueous channel that allows both passive diffusion of small molecules and active transport of large compounds including RNA and proteins between the nucleus and the cytoplasm [9]. The inside of the nuclear envelope is covered with a meshwork of proteins named the nuclear lamina (**Figure 1-2**). Lamins are intermediate filaments that are attached to NPCs and by interacting with each other form a dense network along the inside of the nuclear envelope. One established function of lamins is the positioning of heterochromatin at the nuclear envelope [10, 11]. The role of lamins in gene regulation will be further discussed in **Sections 1.4.1** and **1.4.2**.

1.2.1- Nuclear bodies

Next to the previously mentioned structures in the nucleus, labelling of individual proteins reveals a number of different aggregates, or so-called Nuclear Bodies ([12] and **Figure 1-2**). Spector, in [12], defined nuclear bodies as "*non-membrane-bound structures that can be visualized as independent domains by transmission electron microscopy without antibody labelling.*" According to this definition the nucleolus is also a nuclear body.
Nuclear bodies consist of aggregates of one or more proteins and sometimes specific RNA molecules and genomic regions. Their numbers in the cell can vary from 1 to about 50 and they can differ in size from about 0.1 μm to 3 μm ([12, 13] and references therein). In this section the general characteristics of nuclear bodies relevant to the research in this thesis will be introduced. Furthermore RNA polymerase (RNAP) factories will be introduced. RNAP factories are also distinct nuclear substructures, but they do depend on antibody labelling for their visualisation. Compared to other nuclear substructures, RNAP factories are far smaller (around 40–80 nm) and more abundant (up to 8000 RNAP II factories per cell) [14-18]. The association of nuclear bodies with genomic regions and how this is related to gene expression will be discussed in **Section 1.4.2**.
The first discovered nuclear body besides the nucleolus, was the Cajal body [19, 20], previously also known as the coiled body (**Figure 1-2**). Depending on the metabolic activity of the cell, the size of Cajal bodies varies between 0.1 - 2.0 μm and the number varies between 0 and 6 [12, 20]. Cajal bodies are local concentrations of snRNAs (small nuclear RNAs), snoRNAs (small nucleolar RNAs), snRNPs (small nuclear ribonucleoprotein particles, consisting of snRNAs and spliceosomal proteins) and other non-snRNP proteins, including coilin. The Cajal body seems mainly involved in the final

steps of spliceosome assembly, though certain other functions are also proposed [20]. Among the largest and most abundant nuclear bodies are the splicing speckles (**Figure 1-2**). Splicing speckles are known by a plethora of other names, including nuclear speckles, SC35 domains and interchromatin granule clusters [12, 13, 21]. Splicing speckles are irregularly shaped and can vary in size between 0.8 μm and several μms,

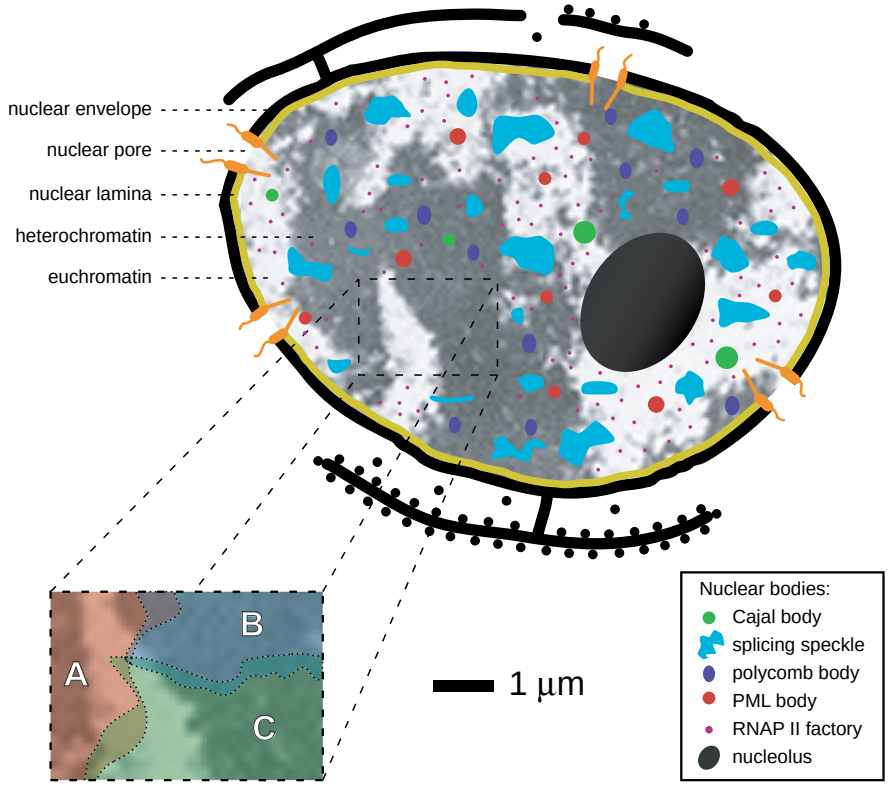


Figure 1-2. The mammalian cell nucleus. Example of a typical mammalian cell nucleus. The most important structural components are indicated on the left. Morphology and approximate location of relevant nuclear bodies are shown in the figure. Number of nuclear bodies indicate approximate total number of bodies in whole nuclei, except for Pol II factories that are 20-80 fold more abundant. A representative size bar is given below the nucleus. In the left lower part of the figure a schematic representation of three Chromosome Territories (A (red), B (blue) and C (green)) intermingling at their edges is given.

though actually they are composed of many 20-25 nm granules that are connected by thin fibrils. The number of splicing speckles varies between 15 – 50 per nucleus [12, 22].

Splicing speckles contain a large number of different proteins, of which the majority is involved in the splicing of pre-mRNAs into mature mRNA [22].
Polycomb (PcG) bodies are aggregates of repressing PcG proteins and are associated with heterochromatin ([12, 13, 23], **Figure 1-2**). Typically, in nuclei of tissue-culture cells, between 12 and 16 PcG bodies are present, that exhibit a fairly globular structure with a size between 0.3 - 1 μm [12, 23, 24]. PcG bodies specifically consist of the PRC1 Polycomb protein complex, which is required for maintenance, but not initiation, of PcG repression (see **Section 1.4.1**). Interestingly, PcG body binding to heterochromatin is maintained during mitosis. This probably allows the protein complex to stably repress genes after subsequent cell divisions, thereby maintaining specific gene expression patterns [23, 25].
Functionally the least well understood nuclear body is the PML body (Promyelocytic leukemia protein body, **Figure 1-2**). PML bodies, also known as ND10, Kr or PODs (PML oncogenic domain), are round structures of 0.3 – 1 μm and usually between 10 and 30 bodies are present in the nucleus [12, 13, 26]. The PML protein is required for the formation of the PML body, but many other functionally different proteins are also present in this subnuclear structure (e.g. p53, CBP, Sumo-1 and SP100) [26, 27]. The function of PML bodies is not clear and many different functions, like cell cycle regulation, apoptosis, tumour suppression, regulation of gene expression and viral infection have been proposed [26]. A different hypothesis for the role of PML bodies is that their main function is the sequestration of proteins whose functions may harm the cell [27]. Supporting this hypothesis are the findings that (i) many proteins present in PML bodies are not required under normal conditions, but need to be recruited rapidly in stress situations [27, 28] and (ii) that viral proteins and many ectopically expressed proteins localise at or close to PML bodies [27, 29-32].
RNA polymerase (RNAP) factories (also referred to as transcription factories and RNAP foci) are sites of RNA transcription and consist of accumulations of RNA polymerase and associated proteins [14, 18, 33, 34]. They can be seen by immunofluorescence, after antibody staining of fixed cells. RNAP factories were identified in the nuclei of HeLa cells, where around 90,000 transcripts are being elongated at only about 2400 sites [15, 33]. Additional studies revealed that a typical mammalian cell contains between 2000 and 8000 RNAP II factories that are between 40 and 80 nm in size (**Figure 1-2**) and contain at least 5, but probably up to 20 actively transcribing RNAP molecules [14-18]. Importantly, whether RNAP factories also exist *in vivo* is still under debate, since imaging

of GFP-tagged RNAP in live cells showed a uniform distribution, rather than the presence of discrete RNAP factories [35].
When considering the relationship between transcription and chromatin organisation, it is important to keep in mind that active RNAPs are known to generate considerable force [36], which may be sufficient to significantly move the chromatin fibre [37]. In a recent study RNAP factories were found to be surrounded by domains of decondensed chromatin, that would consist of recently transcribed genes, and it was suggested that the transcription machinery may have a considerable role in determining the nuclear location of actively transcribed genomic regions [38].

1.3- DNA and chromatin structure

DNA (Deoxyribonucleic acid) is the carrier of genetic information in mammals and all other living organisms. The base pairs (bp) that make up the DNA are organised into a complementary double helix, with Adenosine bases (A) pairing with Thymine bases (T) and Cytosine bases (C) pairing with Guanine bases (G) (**Figure 1-3**). Mammalian genomes contain several billions of base pairs. According to the latest numbers of the Ensembl consortium, the human genome contains 3,253,037,807 bp, and the mouse genome contains 3,420,842,930 bp (http://www.ensembl.org, NCBI 36 and NCBI m37). All cells in mammals contain the same genomic DNA, except for B- and T-cells, the gametes and the enucleated red blood cells. The majority of mammalian DNA is located in the nucleus, in the shape of linear chromosomes. The number of chromosomes between mammals varies, from 6 or 7 (*Muntiacus muntjak sp.,* including the Indian muntjac, whose cells are exploited in certain studies for this property [39, 40]) up to 102 chromosomes (Red viscacha rat, *Tympanoctomys barrerae*) (Gregory, T.R. (2008). Animal Genome Size Database. http://www.genomesize.com). Normal human somatic cells contain 46 chromosomes and mouse somatic cells contain 40 chromosomes.
Mammalian chromosomes typically add up to around 2 meters in length [41]. A typical mammalian nucleus on the other hand is only about 10 μm in diameter. Therefore chromosomes need to be intricately folded and condensed to allow them to fit in the nucleus without becoming entangled. Despite intensive and long-lasting studies, surprisingly little is known about the higher-order folding of chromatin. It has become clear that the folding of interphase chromosomes is achieved at different levels, but despite massive efforts to determine interphase chromatin structure, the actual organisation of most folding levels remains poorly understood [42-44]. The best characterized

level of genome condensation is the nucleosome, which is the basic building block of chromatin [42, 45]. The nucleosome consists of 147 bp of DNA wrapped around 1.7 times around a histone octamer [46]. This core histone octamer is assembled from two copies of the histones H2A, H2B, H3 and H4 (**Figure 1-3**), though a number of histone variants exist that generate variant histone octamers with specific functions. Next to the globular DNA-binding domain, the histones possess an N-terminal tail that emanates from the nucleosome and that can be post-translationally modified. A number of these modifications alter the higher order structure of chromatin (reviewed in [42] and [44]), which in turn may influence gene expression, a mechanism that is called epigenetic regulation (see **Section 1.4.1**).

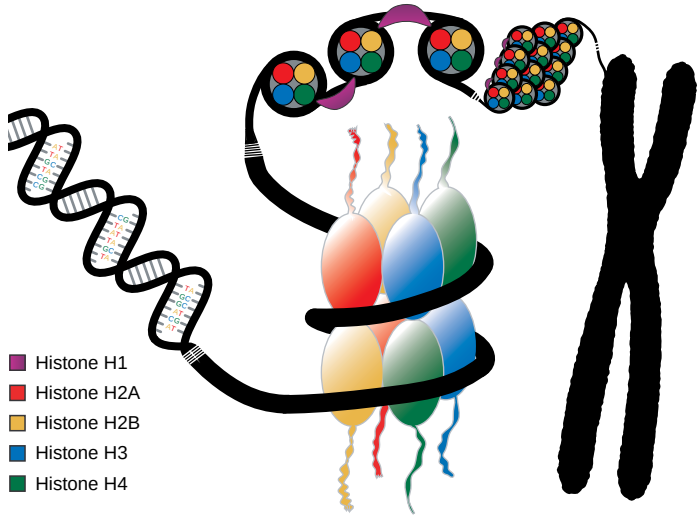


Figure 1-3. Chromatin organisation. Schematic representation of the different levels of chromatin organisation. *Left:* The DNA-double helix. A number of bps are highlighted as an example of complementary organisation of the helix. *Bottom centre:* Simplified nucleosome. Around 147 bps of DNA are wrapped around a nucleosome consisting of 2 copies of each of the 4 core histones. The N-terminal histone tails protruding from the nucleosome in reality adopt a more intricate structure, including many protein-DNA and protein-protein interactions. *Top centre:* 10 nm fibre or "beads on a string". Nucleosomes and linker DNA are bound by linker histone H1. *Top right:* 30 nm fibre. The 10 nm fibre and linker histones are further condensed in to a highly compacted structure. *Right:* Mitotic chromosome. Different compaction levels ultimately result in a highly compacted chromosome. Interphase chromosomes usually adopt a less well defined structure.

The individual nucleosomes are attached to each other by a small stretch of non-histone bound linker DNA. The linker DNA can be between 10 and 60 bp in length, and varies between different species, tissues and genomic regions, depending on the transcriptional state [47-50]. The nucleosomes together with the linker DNA can be visualised by electron microscopy as "beads on a string" structure [45], which is also referred to as the 10 nm fibre [42]. Additional compaction of chromatin is achieved by the binding of nucleosomes and linker DNA to linker histones ([42, 43, 51, 52] and **Figure 1-3**). In mammals there are 8 different linker histones, all belonging to the histone H1-family. The exact function of the linker histones has not been fully elucidated, but transcriptionally active chromatin seems depleted of histone H1 (see for instance [53, 54] and reviewed in [51]). On artificial constructs, histone H1 strongly inhibits transcription [55] and prevents the mobility of nucleosomes [56]. It therefore seems that the stabilisation of the nucleosome array inhibits the transcription machinery from accessing the DNA. In mice sufficient levels of linker histones are required for development and survival of the embryo [52, 57]. A 50% reduction of linker histones causes embryonal lethality before day E11.5. In ES cells that naturally already exhibit low levels of histone H1, the 50% reduction in histone H1 levels is tolerated but leads to a decreased chromatin compaction and interestingly a shortening of linker DNA size, together resulting in dramatic changes in chromatin structure [57, 58]. Surprisingly, in these ES cells the expression of only a very small number of genes is affected and all these genes share that they are regulated by DNA-methylation ([58] and **Section 1.4.1**).

Chromatin organisation at levels higher than the nucleosomal array is not well understood. *In-vitro* studies have revealed a 30 nm fibre, a highly ordered repressive chromatin structure which consists of condensed nucleosomes and linker histones (reviewed in [42] and [44]). *In-vivo* studies have not been able to readily detect this structure in mammalian cells. Possible explanations may be that in living cells chromatin always adopts an even higher order conformation, or that the 30 nm fibre in cells is less regularly structured [44]. Additional *in-vitro* experiments have revealed that the density of the 30 nm fibre is both dependent on interactions of the core-domain of the histones with specific proteins, notably the PcG proteins, and the acetylation of the histone tails. A single repressive PcG complex PCC compacts three nucleosomes, and this is independent of the histone tails [59]. In contrast, histone tail acetylation causes decondensation of chromatin, which increases the accessibility for RNAPs. [42, 60]. This further indicates that changes affecting the organisation of chromatin have a functional role *in vivo*.

Folding beyond the already disputed 30 nm fibre is even less clear. Extensive studies by the Belmont-group have revealed several packing stages of interphase chromatin beyond the 30 nm fibre. Using both electron microscopy and conventional light microscopy, large linear fibre-like elements are detected that consists of densely packed 30 nm fibres, so-called chromonema fibres [61, 62]. Depending on the stage between exit from mitosis to interphase, chromonema fibres are between 60 nm and 130 nm and can have lengths of up to 5 μm, while only very occasional stretches of native 30 nm fibres are observed [42, 61-63]. Within the 100 nm width of a chromonema fibre, the 30 nm fibres may contain over 10 kb of DNA, indicating that whole chromosomal domains may be contained in these fibres [42, 64].

The folding of chromosomes at the previously described levels results in highly compacted interphase chromosomes. Importantly, the compaction of chromosomes does not lead to randomly distributed chromatin throughout the nucleus, but rather results in the individual chromosomes occupying distinct areas, the so-called chromosome territories (CTs, **Figure 1-2**) [65-70]. Individual CTs are further subdivided in several stable and potentially functionally different subchromosomal foci of around 400-800 nm [71]. Initially it was thought that CTs were discrete structures, surrounded by an interchromatin compartment that contained the machinery for transcription, splicing, replication and DNA-repair [68]. Careful analysis with improved visualisation techniques though, has revealed that the borders of the CTs are not discrete and that in human lymphocytes on average 46% of each CT is intermingling with other CTs [72]. The presence of CTs implies that during the formation of interphase chromosome structure, intrachromosomal associations are highly favoured over interchromosomal interactions. The limited intermingling of CTs that is observed, is at least partially dependent on ongoing transcription and it is therefore hypothesised that this may be caused by clustering of genetic loci at nuclear substructures like RNAP factories [72]. Interestingly, recently it was found that integrity of RNAP factories and interchromosomal interactions are not dependent on ongoing transcription [73, 74], though implications for CT organisation were not addressed in these studies. The location of specific gene-loci versus the CT and other nuclear substructures, and the possible implications for gene-regulation are further discussed in **Sections 1.4.2** to **1.4.4**.

A plausible explanation for the presence of CTs is that a more chaotic organisation is not compatible with chromosome condensation and separation during mitosis. It seems likely that overall chromosome structure is mainly determined during mitosis and that

during the decondensation afterwards chromatin movement is highly restricted. In two studies the steady-state movements of loci within CTs has been determined. Movements of loci are generally small over longer periods in time. In the one study movement of loci was restricted to about 0.4 μm in one hour [75], while in the other study distances in the similar range were reported with incidental movements up to 1.5 μm [76]. In a number of recent studies the structure of chromosomes during interphase and different stages of mitoses has been compared. Location of genomic regions on interphase chromosomes is highly correlated to their location during mitosis [77-80]. Restructuring of chromosomal conformation was found in all studies, but they disagree when the reorganisation during cell-cycle takes place. The paper of Manders *et al.* reports that regions were most mobile during G2-phase and that late-replicating loci are mainly immobile [77]. In the study of Gerlich *et al.* reorganisation occurred during mitosis, though only moderate changes were observed [79]. In the studies of Walter *et al.* and Thomson *et al.* the majority of structural changes is reported to occur in the earliest stages of G1 [78, 80], which was further established in a more recent paper by Essers *et al.* [81]. Together these studies indicate that overall conformation of chromosomes is determined by strong condensation of chromosomes during the process of mitosis, and that the subsequent decondensation into CTs at the beginning of interphase does not allow large-scale restructuring of the chromosomes. As a result of the relatively fixed organisation of CTs during the last stages of mitosis and during interphase, CT organisation shows considerable similarities in the two daughter cells [78] and in daughter and mother cells [79].

1.4- Chromatin organisation and gene regulation

As explained in the previous paragraph, mammalian chromatin is not organised randomly. In addition to the regulatory information contained by the genomic code, there are many mechanisms that exert their regulatory functions by acting at the level of chromatin organisation. In this paragraph the relationship between the different levels of chromatin organisation and gene expression are discussed.

1.4.1- Epigenetic regulation

A typical mammalian genome carries around 25,000 genes, each having its own specific transcriptional program. It is therefore not surprising that regulation of mammalian transcription is an extremely complex and dynamic process. To allow the different transcription programs of individual genes, regulation is not only achieved by the binding

of transcription factors to specific sequences in the genome, but also by mechanisms that exert their effects by chemically modifying the DNA and nucleosomes at their target regions. These mechanisms, collectively referred to as epigenetic modifications, usually lead to changes in chromatin structure and DNA accessibility, thereby allowing alteration or maintainence of gene expression in these regions. Epigenetic regulation is achieved by two categories of chemical modifications: methylated cytosine residues in the primary DNA strand and a large array of modified histones, that make up the nucleosome [82]. The mechanisms of epigenetic regulation are diverse, and include both activating and repressive functions [83]. Certain epigenetic modifications are maintained during cell division and therefore are, like the genomic code, a means of propagating information through cell division or even passing it on to the next generation [82].

DNA methylation

DNA methylation in mammals occurs almost exclusively at cytosine residues of CpG dinucleotides. Methylation marks are deposited and maintained by Dnmt proteins. CpG dinucleotides are often found in clusters at promoters, so-called CpG islands. When methylated, these CpG islands have a role in repression of gene expression, though *in vivo* most CpG islands tend not to be methylated. Outside these islands, CpGs are underrepresented, due to the increased susceptibility to mutation of methylated CpGs that are not under evolutionary selection [84]. Methylation of CpG islands, and associated binding of protein factors, can interfere in the DNA-binding of regulatory proteins, but more commonly causes the recruitment of proteins that deposit repressing histone modifications [85, 86].

The most extensively characterized example of impaired binding of a transcription factor to a methylated template is found at the imprinted mouse Igf2/H19 locus. The *Igf2* gene, which encodes Insulin-like growth factor 2, is only expressed from the paternal allele, but is silenced at the maternal allele. About 90 kb downstream of the *Igf2* gene a differentially methylated region (DMR) is present. This DMR can only be bound by the mammalian insulator protein CTCF when it is not methylated [87, 88]. At the unmetylated maternal allele of the DMR, CTCF prevents the formation of a chromatin loop between the *Igf2* gene and the further downstream-located enhancer. Due to the lack of interaction between the gene and enhancer, the *Igf2* gene is not activated. Interestingly, the enhancer engages in an interaction with the nearby *H19* gene, leading to activation of this gene from the maternal allele [89, 90]. At the paternal allele the situation is the opposite: the DMR is

methylated, thereby preventing CTCF from DMR-binding and consequently allowing the formation of a chromatin loop between the enhancer and the *Igf2* gene, which is now exclusively expressed from this allele, while the *H19* gene is not activated anymore from this allele [90].

Repression by DNA methylation can be mediated by specific histone modifications. Most proteins binding to methylated CpGs contain the MBD-domain (Methyl binding domain) and recruit histone modifying enzymes to their site of binding. The MBD-containing MeCP2 protein targets the Sin3A histone deacetylase, while the MBD3 protein targets components of the nucleosome remodelling and histone deacetylating NuRD-complex [91-94]. Furthermore DNA methylation also influences H3K9 and H3K27 methylation [95, 96]. A complicating factor though, is that the opposite seems also true: changes in both histone acetylation and histone H3 methylation themselves can influence DNA methylation (see [95, 97]). The importance of DNA methylation in gene regulation is clearly evident from the strong enrichment of this modification in promoter regions. Even though only about 0.7% of the genome is considered to be part of a CpG-island, 60 – 70% of human promoters contain one or more CpG islands [82, 94, 98, 99]. Whether CpG islands in promoters are methylated, and to what extent they are methylated, is dependent on the tissue and development stage where the gene is analysed, on the function of the gene and on the organisation of the promoter itself [98, 100]. In [100] three categories of promoters are described. The first group of promoters are CpG-poor. Incidental CpGs that are present in these promoters are usually methylated, but this does not preclude the genes from being transcribed. The second group of promoters display a high CpG content, and these are usually not methylated, irrespective of their transcription status. The third group of promoters contain an intermediate number of CpGs, and these CpGs often become methylated during development. Among the genes that are most prone to *de novo* methylation in somatic tissues, germ-line specific genes are overrepresented, indicating that DNA methylation is applied as a general mechanism for silencing of these genes [100]. Two studies comparing DNA methylation between embryonic stem (ES) cells and differentiated cell lines revealed distinctly different methylation patterns. In the first study, promoter DNA methylation was higher in human ES cells than in several differentiated cell lines, with the exception of imprinted genes that were not statistically differently methylated [101]. In the second study, on an *in vitro* mouse differentiation system, changes in promoter DNA methylation were found to be only moderate, but considerable changes were found at long-range regulatory elements, though no strong

preference was observed for increased or decreased methylation [102]. Considering the results from these studies on changes in DNA methylation during development, it is surprising that in an elaborate study between normal tissue and tumours, only very small differences in CpG island methylation were observed [103]. A small number of promoters is hypermethylated in colon cancer cells, compared to healthy colon tissue, while *de novo* de-methylated promoters are not at all observed. Several large regions with overall hypomethylation are observed, but these are generally gene-poor [103]. What the significance of this hypomethylation could be, and whether this is a general tumour phenomenon needs to be determined. In several studies, decreased DNA methylation was found to lead to increased genomic instability, and therefore could be a cause rather than an effect of tumorigenesis [103-105].

Until recently, DNA methylation in somatic cells was thought to be a stable mark that constitutively represses genes. So far only proteins have been identified that establish and maintain methylation of CpGs, but recently an intriguing paper was published which showed that two of these proteins, Dnmt3a and Dnmt3b, also play a role in demethylation, leading to transcriptional activation [106]. Upon oestrogen stimulation, the human *pS2* gene is cyclically expressed, with a period length of about 40 minutes. Coinciding with the expression of the gene, certain CpGs in the promoter are methylated and demethylated. Addition of an oestrogen antagonist decreases both *pS2* expression and the presence of Dnmt3a and Dnmt3b at the promoter of the gene, indicating that these proteins may have a role in activation of gene expression. Interestingly, demethylation is not direct, but is achieved by deamination, which renders the methylated cytosine into a thymine. The T-G mismatches are then recognised by components of the Base Excision Repair pathway, that substitute the thymine for a cytosine. Together this mechanism would result in demethylation of a CpG [106].

Histone modifications

The second category of epigenetic regulation is modifications of histones. By now, over 100 modifications of the four core histones have been identified. Specific amino acids of histones can be subjected to different posttranslational modifications, including acetylation, methylation, phosphorylation, sumoylation and ubiquitylation (reviewed in [83]). Additional complexity is achieved by the methylation of lysine and arginine residues, that can be mono- or di- and trimethylated, each having their own function. Especially the N-terminal tails of histones H3 and H4 contain a large number of amino

acid residues that can be modified (**Figures 1-3** and **1-4**). Most histone modifications are involved in activation or repression of transcription, but some modifications have roles in DNA replication and repair [83, 107]. The large number of possible combinations of modifications present on a nucleosome, has led to the hypothesis of a "histone code", that would allow an enormous potential for fine-tuning of chromatin structure and transcriptional state of a locus [83, 107-109]. Epigenetic modifications exert their function through two different mechanisms, though the two mechanisms may be utilized simultaneously by one individual modification. Histone modifications can cause a change in higher-order structure of chromatin, by either modifying the interactions between adjacent nucleosomes or altering the interactions between histones and DNA. Next to this, the modifications can also induce the recruitment of non-histone proteins to their location on the chromatin template [83].

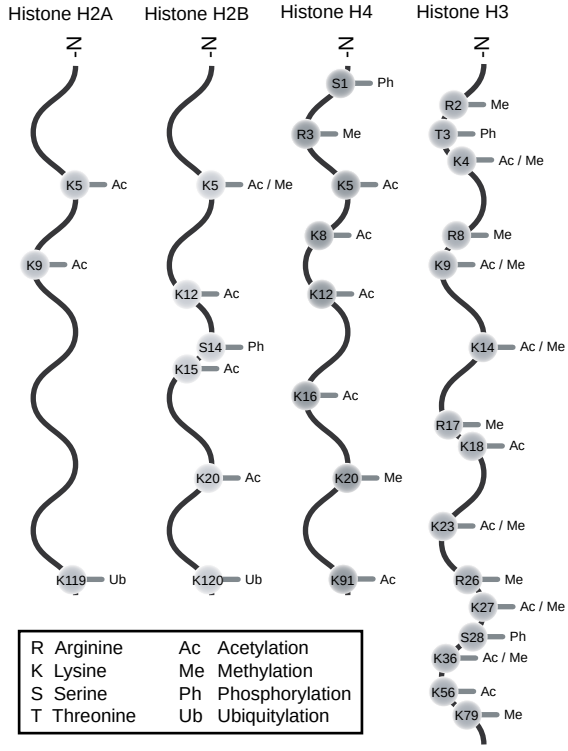


Figure 1-4. Histone modifications. Post-translational modifications of the 4 core histones in mammalian cells. Modifications are depicted from the N-terminal tail (top). Notice that the more C-terminal amino acid residues in histones H2A, H2B and H3 are not depicted according to their actual location on the primary protein strand.

The regulatory histone modifications that have been characterized most extensively, and which are most relevant to the scope of this thesis, are methylation and acetylation of lysine residues in the tail of histone H3.

Methylation of H3K9 and H3K27 (lysine 9 and 27 in the tail of histone H3) is usually associated with transcriptional repression; these are considered to be fairly stable modifications [83, 96, 110]. The function of H3K9 methylation is dependent on the degree of methylation. Trimethylation of H3K9 induces constitutive heterochromatin formation, by providing a binding site for the HP1 proteins [111-113]. Interestingly, HP1 associates with nuclear lamins that are present at the nuclear periphery (see **Figure 1-2**), providing a potential explanation for the relocation of heterochromatin to this repressive nuclear environment and linking nuclear organisation to histone modifications and heterochromatin formation [11, 113-115]. The targeted deposition of H3K9 trimethylation in mammals is not well understood. In yeast, heterochromatin formation is preceded by transcription of siRNA repeats (small interfering RNA) in these regions. Subsequently, the siRNAs bind the RNAi machinery, leading to recruitment of the Clr4 protein and subsequently HP1. The human homologues of Clr4 are the SUV39h1 and SUV39h2 proteins and these proteins are histone lysine methyltransferases that are specific for H3K9 [96, 116-119]. Besides establishing H3K9 methylation, human SUV39h1 has also been reported to target the repressing PRC1 Polycomb complex to chromatin loci [120], which may further inhibit the transcriptional potential of these heterochromatic regions through histone ubiquitylation. Interestingly, PRC1 is part of the previously discussed Polycomb bodies (see **Section 1.2.1**), thereby possibly also linking nuclear body location to histone modifications and heterochromatin formation.
In contrast to trimethylation, mono- and dimethylation marks at H3K9 are mainly enriched at promoters in euchromatic regions [96, 110]. The function of these modifications at promoters seems the silencing of individual genes in euchromatic regions. The two most important histone lysine methyltransferases involved in depositing H3K9 mono- and dimethylation are G9a and GLP, and a deficiency of G9a leads to upregulation of specific genes [121, 122]. Next to G9a and GLP, the previously mentioned SUV39h1 and SUV39h2 are also involved in gene-silencing at euchromatic regions, though this seems only the case in terminally differentiated cells [96, 123]. Like tri-methylated H3K9, mono- and dimethylated H3K9 also attract HP1 to its sites on the chromatin template [122, 124]. How HP1 locally silences gene expression is not well understood, though HP1 has also been found to interact with the DNA methyltransferase Dnmt3b, thereby possibly inducing promoter DNA methylation [125].
H3K27 trimethylation is involved in the repression and silencing of specific categories of loci, like HOX genes, imprinted loci and potentially the inactivated X-chromosome

[96]. Repression by H3K27 methylation is tightly associated with the PRC2 Polycomb protein complex. The EZH2 component of this human PcG complex deposits the H3K27 methylation mark at the promoters of silenced genes [126-129]. The H3K27 mark in turn is bound by PRC1, which induces silencing by preventing chromatin remodelling, potentially induced by the ubiquitylation of histone H2A [127, 130, 131]. Furthermore, EZH2 targets the DNA-methyltransferases Dnmt1, Dnmt3a and Dnmt3b to promoters, leading to simultaneous DNA methylation [132]. Very little is known about H3K27 di- and monomethylation. In a recent genome-wide study on human CD4+ T-cells, H3K27 trimethylation was mainly enriched at the promoters of silent genes. H3K27 dimethylation was also mainly enriched at promoters, though with less distinction between promoters of active and inactive genes. Contrary to expectations, H3K27 monomethylation was exclusively enriched at active promoters, [133].

The role of H3K4 methylation is very different to H3K9 methylation and H3K27 di- and trimethylation. Not only is the mark associated with transcriptional activation, but it is also deposited by the transcription machinery itself, thereby creating a positive feedback loop. In yeast, the SET1 methyltransferase associates with the elongating Ser5-phosphorylated isoform of RNAPII and deposits the H3K4 methylation mark during transcription [134]. Interestingly, this does not lead to a uniform distribution of H3K4 methylation through genes. Rather, H3K4 trimethylation is most abundant at promoters and the 5'-end of genes, H3K4 dimethylation is distributed fairly evenly through genes and H3K4 monomethylation is most enriched at the 3'-end of genes [135, 136]. Both in HeLa cells and human CD4+ T-cells similar patterns are observed, suggesting that in mammals a conserved mechanism is employed [133, 137, 138]. The function of H3K4 methylation still remains speculative, though it has been proposed that the mark functions as a memory of recent transcriptional activity [96, 134]. Recently several findings have been published that support this hypothesis. In human cells the non-transcription coupled MLL-methyltransferase complex binds dimethylated H3K4 and subsequently converts H3K4 to a trimethylated state [139], thereby possibly influencing the stability of the mark [96]. In the same human cells, a component of the ISWI chromatin remodelling NURF complex (Nuclear Remodelling Factor) binds trimethylated H3K4. The subsequent chromatin remodelling by the NURF complex may in turn lead to facilitated transcription [140]. This results in a positive feedback loop, with transcription itself positively influencing transcription, thereby indeed generating a transcriptional memory that precludes silencing of loci.

Like H3K4 methylation, histone H3 acetylation is a modification associated with activation of transcription. Functionally, histone acetylation neutralizes the positive charge of histone tails, which leads to a decreased affinity for DNA and thereby facilitates chromatin remodelling and transcription [141]. Acetylated lysine residues on histone H3 that are associated with transcriptional activation are K4, K9, K14, K18, K23, K27, K36 and K56 [83, 142]. Mammalian histone acetyltransferases (HATs) can be subdivided in three families: the GNAT family, the MYST family and the p300/CBP family [143]. They comprise a large number of HATs that interact with many other proteins, including members of the RNAPII complex, transcriptional co-activators that interact with transcription factors, and with other HATs [143, 144]. In contrast to methyltransferases, HATs usually modify more than one histone residue, and some do not only modify residues of histone H3, but also of the histones H2A, H2B and H4 [83]. Due to the interactions with other proteins, HATs are recruited to promoters, resulting in a large number of seemingly redundant modifications that facilitate the access of RNAP to the DNA [142, 144, 145]. In contrast to the quite stable methylation marks on H3K9 and H3K27, histone H3 acetylation marks are highly dynamic. In addition to HATs, also a large number of histone deacetylases (HDACs) exist and these usually have functions in transcriptional repression. Mammalian HDACs are also subdivided in three families: the RPD3-like HDACs, the HDA1-like HDACs and the SIR2-like HDACs and usually these are also present in large protein complexes [97]. The RPD3-like HDACs associate with co-repressor complexes like NuRD (Nucleosome remodelling and deacetylating complex) and the DNA-methyltransferase Dnmt1 [97, 146-149]. Intriguingly, the NuRD complex binds the MBD3 protein that recognizes methylated DNA. Furthermore, RPD3-like HDACs themselves interact with the MeCP2 protein that also recognizes methylated DNA. The importance of this cross-talk between HDACs and DNA methylation is shown by the finding that a loss of DNA methylation leads to hyperacetylation of histone H3, accompanied by hypomethylation of H3K9 in human cells, indicating that DNA methylation is required for maintenance of a repressive hypermethylated chromatin state [91-93, 95, 150].

All these histone modifications, including the ones not discussed in this section, potentially add up to a complex signalling mechanism, yet clear evidence for the existence of a histone code is still missing. Only recently several genome wide studies have been published, that combine the location of many histone modifications, RNAP location and transcriptional status of loci. These studies show distinct patterns of histone modifications, related to the location of promoters, genes and other regulatory elements

and to the expression status of loci, indeed suggesting the existence of a histone code [107, 133, 138, 142, 145, 151].

1.4.2- Genomic loci and their location in the nucleus

The location of genomic loci in the nucleus is not random and is often suggested to be related to gene expression (for several recent reviews see: [10, 68, 152-159]). In studies aimed at elucidating a potential relationship, the location of genomic loci classically is determined against several different nuclear landmarks (see **Figure 1-5A**). When comparing these results, two potential issues should be kept in mind. First, even though chromosomes are non-randomly organised, their location within the nucleus is usually not absolute but probabilistic. This issue has been extensively discussed by Parada *et al.*, who provided the following description: "*...chromosome positioning patterns are statistical representations of chromosome positions but do not provide information about the precise coordinates of a given chromosome in a single nucleus. It is important to realize that, although statistically significant non-random chromosome patterns can be described, they contain a significant degree of uncertainty and merely indicate a preferred, probabilistic position of a given chromosome in the cell nucleus.*" [160]. Not only is this true for whole chromosomes, but individual gene-loci are also positioned according to the same probabilistic rules. The location, and therefore nuclear environment of individual loci in specific cells may therefore vary greatly. These differences may be related to the observation that RNA levels of individual genes can greatly vary between "similar" cells in a population, showing that individual cells react probabilistic to similar environmental stimuli [161, 162]. The knowledge obtained from a population of cells is therefore merely indicative for the actual situation in individual cells and outliers may have a strong influence on the outcome for whole cell populations.

A second important issue when comparing reports on nuclear organisation is the lack of standardised methodology. The results in different reports are usually obtained with different Fluorescent In-Situ Hybridisation (FISH) techniques on different cell types. Subsequently, the results are analysed and visualised in different ways (see **Figure 1-5A, B**). Caution should therefore be maintained when comparing results from different studies [163].

When determining the location of genomic loci, it is first important to realise that CTs are non-randomly distributed in the nucleus. A very interesting example are human chromosomes 18 and 19. Both chromosomes have a comparable size (Chr. 18: 76 Mb,

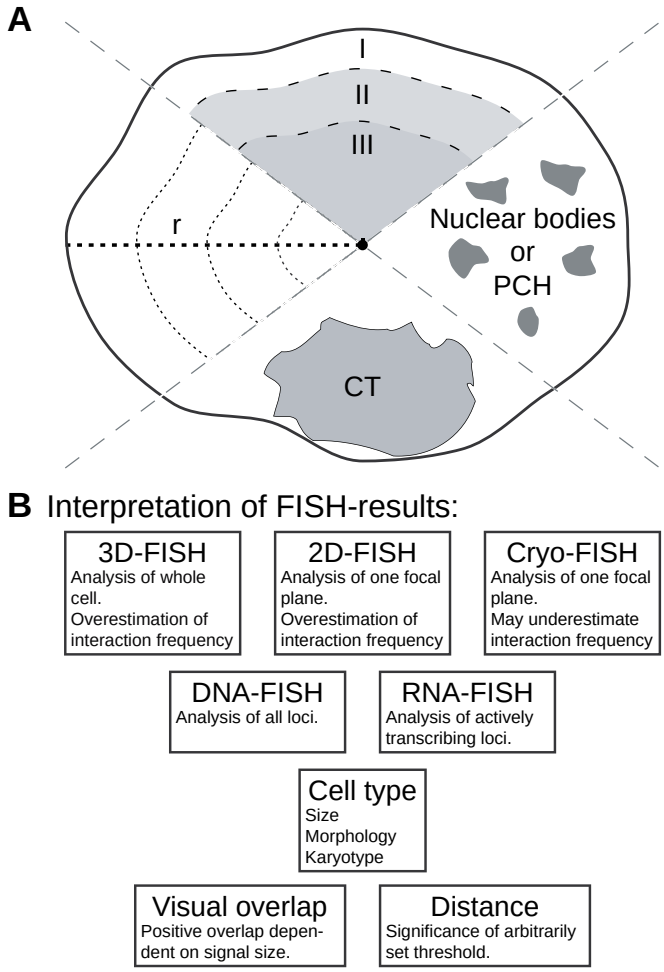


Figure 1-5. Location of genomic loci in the nucleus. (A) Examples of substructures and measurements that are used to determine the location of genomic loci. *Left:* Location of genomic loci versus the periphery. Here measured relative to the radius (r) of the nucleus. *Top:* Location of genomic loci versus the periphery. Here measured by location of the locus in one of three concentric shells, each with identical area. Notice that the shell located most to the centre (shell III) encompasses over 50% of the nuclear radius (left).*Right:* Location of genomic loci versus nuclear bodies or pericentric heterochromatin (PCH). Location is usually scored as visually overlapping or distance from the centre of the locus with the respective nuclear substructure. *Bottom:* Location of genomic loci versus the CT. Location can be scored as visual overlap, as distance from the edge of the CT or in categories, like inside, at the edge or outside the CT. **(B)** Scheme showing differences in FISH analysis, indicating the cause of potential differences.

Chr. 19: 64 Mb; http://www.ensembl.org, NCBI 36), but are structurally very different. Chromosome 18 is among the most gene-poor chromosomes (277 known protein coding genes, resulting in less than 4 genes/Mb), while chromosome 19 is the most gene-rich chromosome in the human genome (1363 known protein coding genes, resulting in over 21 genes/Mb). In an initial study on human lymphocytes and lymphoblasts, chromosome 19 is shown to be located closer to the nuclear centre than chromosome 18 [164]. In a follow-up study on an additional six unrelated cell types this preferential location was confirmed, showing that non-random radial distribution is not a lymphoid cell type specific phenomenon [165]. As a next step, the location of all human chromosomes was determined, again revealing a strong correlation between radial positioning and gene-density, but not chromosome size [166]. Two subsequent studies though, reported partially different results. In spherical lymphoid nuclei results are similar as observed previously, but in ellipsoid nuclei CT distribution is only dependent on chromosome size, with small CTs located at the nuclear interior [167]. Similarly, when all CTs were visualized simultaneously in two different flat-ellipsoidal human cell types, again CT distribution depends on chromosome size but not gene-density [168]. Whether CTs are organised according to gene-density or size therefore seems dependent on nuclear morphology.
Like whole CTs, genomic loci also adopt non-random radial positions in the nucleus, and this seems mainly determined by the expression status of the locus. In a first study, the IgH and Igκ loci were studied during mouse B-cell development. In pro-B cells both loci are inactive, but are prone for transcription in subsequent steps of development. In early stages of development, about 50% of alleles for both loci are located at the nuclear periphery. In further developed pro-B cells this percentage goes down to less than 10% [169]. Positioning of loci at the periphery has therefore been proposed to repress transcription and relocation towards the nuclear interior would allow the initiation of transcription [169]. Several recent studies have not reported this strict correlation between location at the nuclear interior and initiation of transcription. In a study where the location of three adjacent genes with different transcriptional programs was studied in six human cell types, genes again showed non-random nuclear organisation depending on their expression status. Inactive genes are most often located at the periphery and active genes are more often at the interior [170]. But more importantly, relocation of the CFTR gene to the nuclear interior does not lead to activation of the gene, but rather seems to be the result of transcription of surrounding genes [170]. In a different study, the β-globin genes were studied during erythroid differentiation (see also **Paragraph**

2.6). Again, the inactive locus is usually located at the periphery. When active, the β-globin locus is located at a more interior position in the nucleus, but activation of the β-globin locus actually precedes relocation to the nuclear interior [171]. Similar results are obtained in Chinese hamster cells, where activation of a transgene with an artificial viral transcriptional activator precedes nuclear relocation to the interior of the nucleus [172]. These findings are therefore suggested to indicate that relocation to the nuclear interior is not a prerequisite for transcription. Rather, this relocation may have a function in maintenance of transcriptional activity [171, 172], or be a consequence of transcription.

Interestingly, inactive loci in all these studies share that they are located at the nuclear periphery. The question therefore arises, whether this location causes active repression of genomic loci or whether it is a consequence of inactivity. The nuclear periphery is strongly associated with the nuclear lamina and these are thought to associate with heterochromatic regions containing repressive histone modifications (reviewed in [115] and **Section 1.4.1**). An analysis of these heterochromatic regions, or so-called Lamina-Associated Domains (LADs), revealed that these regions are not only gene-poorer than non-LADs, but also that genes in these regions generally are lower expressed [11]. Recently, three comparable studies in mouse and human cells have addressed the question whether relocation to the nuclear periphery results in silencing of genomic loci. In all three studies, genomic loci containing *lacO*-repeats were targeted to the nuclear periphery through their interaction with specific inner nuclear membrane proteins fused to the *E. coli* LacI repressor protein [173-175]. In the first study, a sophisticated reporter construct was directed to Lamin B1, a component of the nuclear lamina. The kinetics of transcriptional activation of this reporter gene is not altered when the locus is relocated to the nuclear periphery, as determined by MS2-YFP binding to RNA containing repeats of motifs recognized by MS2 [173]. The two other studies measured maintenance of transcription upon relocation to the periphery. In the second study, randomly integrated *lacO*-repeats are directed to the Emerin protein, which is part of the inner nuclear membrane. When the locus is relocated to the nuclear periphery, expression levels of the *hyg* selection marker in two clones and two genes surrounding the integration site in one clone are reduced by about 60% [174]. In the third study, randomly integrated *lacO*-repeat arrays at two different sites in the genome are recruited to the nuclear periphery through association with the inner nuclear membrane component Lap2β. In this study, relocation to the nuclear periphery also reduces the expression of the selection

marker, though only by 20-30%. Furthermore several, but not all, genes in the vicinity of the integration site are downregulated [175]. Interestingly, in the last two studies downregulation of genes is related to HDAC-activity [174, 175]. Together, it therefore seems that relocation to the nuclear periphery does not necessarily silence loci, but depending on the susceptibility can decrease the transcriptional activity of a gene by deacetylation of histones. Supporting this phenomenon and relating it to nuclear lamina function are the findings by Malhas *et al.* [176]. In this study, defects in the processing of Lamin B1, which decreases nuclear lamina stability, leads to upregulation of a large number of genes, indicating that this component of the nuclear lamina has a repressive function. Further support for the association of lamina with heterochromatic regions is provided by the observation that in these processing-deficient cells the previously discussed small gene-poor human chromosome 18 is not located at the nuclear periphery, but adopts a similar distribution pattern as the gene-dense chromosome 19 [164, 176].

A large body of work has also been devoted to determining the relationship between location of genomic loci versus the CT and their transcriptional activity (see **Figures 1-2** and **1-5A**). In a first study, three genes and two non-coding sequences were analysed in four different human cell types. Location at the edge of the CT was defined as being less than the spots diameter away from the territories edge. Interestingly, genes are preferentially located at the periphery of CTs, irrespective of whether they are active or not. The non-coding regions on the other hand are located randomly in their CT [177]. In a different study, the location of genes on the X-chromosome was determined. Location versus the CT in this study is quantitatively measured. The ANT3 gene, which escapes X-chromosome inactivation, is similarly located at the edge of both the active and inactive X-chromosome. In contrast, the ANT2 gene, which is silenced during X-chromosome inactivation, is significantly more internally located in the CT of the inactive X-chromosome than in the active X-chromosome [178]. A third initial study addressed the topology of the 3 Mb gene-dense MHC (Major Histocompatibility Complex) locus. The MHC cluster contains both many constitutively expressed and inducible genes. In this study, location at the edge of the CT is determined by visually scoring whether loci are touching the border of the CT. Under uninduced conditions, the human MHC locus is about 80% of the times located at the periphery and 12% of the times located outside the CT. Further activation of the MHC locus in fibroblasts increases looping out of the CT from around 10% to 26% [179]. From these initial studies the hypothesis was postulated that location of genomic

loci at the edge or outside the CT is a common theme for active loci and may even be required for efficient access to the transcriptional and RNA-processing machinery [68]. Later publications though, indicate that the situation is more complex and that active genes also reside inside CTs. In one study, the location of an evolutionary conserved genomic region containing ubiquitously expressed genes, tissue specific genes and three gene deserts of over 100 kb was studied in human and mouse cells and active genes were found also inside the CT [180]. Similarly, in a study on the human CFTR gene and adjacent genes in different tissues, the genes were mostly located inside the CT, irrespective of activity [170]. Finally, the HoxD locus, which is both active in the tailbud and limb bud in E9.5 mouse embryos, is located differently relative to the CT in these two tissues and adopts a position inside the CT in the limb bud [181]. Together, these results indicate that location inside the CT does not preclude transcriptional activity. A study on another large and conserved gene-dense region containing unrelated genes further revealed that location at the edge or outside the CT, besides not being absolutely required for transcription, also does not seem to cause a general stimulation of transcriptional activity. Like the previous studies, this region is mainly located at the edge or outside the CT. But more importantly, location outside the CT does not generate a collective positive response from the genes, but rather finds each gene maintaining its own transcriptional program [182]. An alternative explanation for location of active genes at the periphery of the CT could be that chromatin decondensation, either required for transcription or the consequence of transcription itself, leads to location at the edge or outside the CT (discussed in [179, 181, 183]). In a recent study no support for this hypothesis was found; the decondensation state of the mouse HoxD locus and location versus the CT were found not to be related [181]. Most likely though, the genomic context of a locus dictates its steady-state position in the CT and determines whether its decondensation leads to an appreciable relocation away from the CT.

The last category of non-random nuclear location of genomic loci that will be discussed in this chapter is location versus nuclear bodies (see **Figure 1-2**). Association of genomic loci with these substructures is thought to relate to the biological function of the nuclear body involved. In a considerable number of studies, specific genes have been shown to associate with splicing speckles, the nuclear substructure involved in processing of pre-mRNAs (see **Section 1.2.1**). It is important to realise though that speckles are transcriptionally inactive and that genomic loci localise at the periphery of splicing speckles with their associated RNA extending into the speckles [184, 185]. This

suggests that association is mediated through the binding of components of splicing speckles with pre-mRNAs that are still being transcribed and therefore also attached to the chromatin (illustrated in [21]). The highly active *COL1A1* and *COL1A2* genes, coding for intricate transcripts containing over 50 exons, associate 90% or more of the time with one of the 15-30 splicing speckles in human fibroblasts [184, 186, 187]. Similar observations have been made for the highly expressed *ACTB* gene (encoding β-actin; 6 exons, 89% associations), the *LMNA* gene (encoding Lamin A/C; between 8 and 13 exons, 70% association) [187]. In two different studies, co-localisation of genomic loci with splicing speckles was verified to be dependent on activity of loci. The *cMyHC* and *Myogenin* genes associate up to 90% with splicing speckles in myotube nuclei, where they are active. In contrast, only 10% of the alleles co-localise in cell types where they are inactive, which is similar to other inactive gene loci [188]. In an elegant study by Takizawa *et al.* the location of the mono-allelically expressed *GFAP* gene versus splicing speckles was determined. In 70% of the cells only the actively expressed *GFAP* allele is associated with splicing speckles, while in an additional 15% of cells both alleles are associated, which is comparable to random association [189]. Thus, at least some active loci co-localize with splicing speckles specifically, but this association seems not a prerequisite for transcription and may still merely be a consequence of RNA production. Association of genomic loci with RNAP factories has also been determined. In mouse fetal livers, three actively transcribed gene-loci on chromosome 7 overlap in around 90% of cases with RNAP factories [190]. A similar study in mouse B-cells showed comparable percentages for the actively transcribed Myc and Igh loci, indicating that this is not a tissue specific phenomenon [191]. In the previously mentioned study by Ragoczy *et al.* [171], relocation to the interior results in increased association with RNAP factories, but importantly, initial transcription precedes both relocation to the interior and association with RNAP factories. This last result raises the question whether the association with RNAP factories may actually influence location of genomic loci. Do genomic loci actively relocate to RNAP factories or could promoter bound "pre-RNAP factories" associate in the nuclear space, thereby concentrating their location at a discrete number of sites in the nucleus? So far, this question has not been answered, but in two recent studies RNAP factory integrity and association of genomic loci with these substructures is shown not to be dependent on ongoing transcription [73, 74]. From these results it becomes apparent that interactions of genomic loci with nuclear substructures may influence each others location in the nucleus, thereby potentially having a large effect

on overall nuclear organisation. The establishment of long-range and interchromosomal interactions mediated by the association with nuclear bodies will be further discussed in **Sections 1.4.3** and **1.4.4**.

1.4.3- Long-range chromatin interactions

Distant genomic loci can interact in the nucleus through the formation of chromatin loops. Until a few years ago, little was known about the existence of chromatin loops in mammals, since detection methods with sufficient resolution were lacking. The development of 3C (Chromosome Conformation Capture) and subsequent genome-wide derivatives did provide this resolution and has greatly increased our knowledge [192-195]. It is important to realise that these techniques need to be highly controlled to avoid over-interpretation of results (discussed in [194, 196] and **Paragraph 2.3**). Using 3C, the first chromatin loops in mammals were identified in the mouse and human β-globin loci [193, 197]. Thus far, most knowledge on chromatin looping has been obtained from these well-characterized loci, as is discussed in detail in **Chapter 2**. Depending on the developmental stage, the Locus Control Region (LCR), a strong enhancer element, loops towards one of the embryonic, fetal or adult β-globin genes, which is accompanied by high expression levels of these genes [197]. Nowadays, several similar regulatory mechanisms have been described at other gene loci. The mouse T_H2 LCR, which activates the cytokine genes *IL4*, *IL5* and *IL13* but itself is located in the unrelated *RAD50* gene, communicates with the individual genes through chromatin looping [198, 199]. At the mouse α-globin locus, the active genes are in physical contact with regulatory sequences surrounding the genes [200, 201]. Depending on the methylation state of the DMR in the mouse Igf2/H19 locus, activating chromatin loops are established either between the Igf2/H19 enhancer and the *Igf2* or *H19* gene ([90] and discussed in detail in **Section 1.4.1**). Furthermore, at the human MHC locus, the MHC class II genes *HLA-DRB1* and *HLA-DQA1* are simultaneously activated by spatial interactions with the *XL9* element [202]. Next to transcriptional activation, transcriptional repression involving looping has also been reported. Using a variation to the 3C-technique, MeCP2 and DNA methylation dependent chromatin loops between the *Dlx5* and *Dlx6* gene are observed in mouse brain cells. Ablation of these loops in MeCP2-null mice results in an about twofold upregulation of both genes, indicating that the chromatin loops have a repressive function [203]. These examples show that chromatin looping is involved in several different regulatory mechanisms, either activating or repressing one or multiple genes.

Interesting questions regarding these mechanisms are how chromatin loops are established, maintained, and how they exert their regulatory function. In [204] two different attachment platforms for genomic loops are proposed. First, transcription factors and other DNA-binding proteins can mediate chromatin looping through (hetero-) dimerisation. Secondly, association with and sharing of nuclear bodies, like splicing speckles and RNAP factories, may result in the formation of chromatin loops [204]. In this last example, the functionality of the long-range chromatin interactions themselves may be questioned. In several studies, deletion of specific DNA binding proteins is shown to result in disturbed chromatin looping and changes in transcriptional status of target genes ([202, 205-208] and **Paragraph 2.5**). One particularly interesting DNA-binding protein involved in chromatin looping is the CTCF protein (CCCTC-binding factor). CTCF is the best known insulator protein in mammalian cells and is required for insulator function [209]. The protein is thought to act by locating genes and enhancers on different chromatin loops, thereby preventing their interactions (e.g. in the Igf2/H19 locus) [90, 210]. Interestingly, a genome-wide analysis of CTCF binding sites in human fibroblasts revealed over 13,000 sites, and their location is strongly correlated with the location of genes. This large number of CTCF-sites located close to genes suggests that CTCF induces the formation of a very large number of regulatory chromatin loops, resulting in a major influence on overall chromatin organisation [211]. The recently discovered almost complete co-localisation of CTCF and the cohesin complex provides interesting insight in how CTCF may establish chromatin loops [212-214]. Originally, the cohesin complex was found to be responsible for sister chromatin cohesion in mitosis. During DNA replication, the cohesin complex has been proposed to form a ring around the replicated DNA-strands, thereby ensuring pairing and subsequent correct division of chromosomes during cell-division (reviewed in [215]). The observed co-localisation of CTCF and the cohesin complex opens up the intriguing possibility that CTCF-dependent chromatin loops may also be established by the formation of cohesin rings around the DNA strands. Further studies should reveal whether cohesin is indeed responsible for long range chromatin interactions and whether other DNA-binding proteins involved in chromatin loop formation employ a similar mechanism.

The second proposed platform in [204] to which chromatin loops are attached are nuclear bodies. In a limited number of studies this issue has been addressed. In the previously mentioned study by Shopland *et al.*, co-localisation at splicing speckles of the *Col1a2* and *ACTB* genes is addressed. Both genes are located on human chromosome

7, but are separated by over 88 Mb of DNA [187]. The two genes show a high, almost 90%, co-localisation frequency with the 15 – 30 nuclear speckles in the nucleus. Due to the high degree of association at this limited number of structures, a considerable percentage of co-localisation between the two loci can be expected. Interestingly, instead of the expected 11% of cells with co-localising loci a significantly higher 29% of cells are observed [187]. A possible explanation for this is the non-random organisation of chromatin into CTs. This results in a smaller distance in the nucleus than would be randomly expected, thereby increasing the chance of co-localisation at the same splicing speckle. This result suggests that splicing speckles are involved in establishing long-range chromatin interactions and thereby influence spatial chromatin organisation. Similar to splicing speckles, co-localisation of genomic loci at RNAP factories has been addressed in the previously mentioned study by Osborne *et al.* [190]. Importantly, RNAP factories are far more abundant than splicing speckles. Even though the 100 – 300 of RNAP factories per nucleus reported in [190] is lower than the several thousands mentioned in other studies [14-18], it is still more than the maximum 50 splicing speckles per nucleus [12]. Random co-localisation of two genomic loci at one RNAP factory will therefore be an infrequent event. Surprisingly, four active genes on mouse chromosome 7 co-localise with the β^{maj} gene at RNAP factories between 41% and 60%, even though they are located up to 40 Mb away [190]. A similar study on three genes located on different chromosomes showed maximum co-localisation frequencies of 20% at RNAP factories [191]. These results suggest that RNAP factories may exist that coordinate the transcription of specific genes. In a recent study this hypothesis was addressed by studying co-localisation at RNAP factories of artificial plasmids with different structural characteristics [216]. Shared characteristics were indeed found to promote clustering of certain plasmids in the nucleus at RNAP factories. Co-localisation of the several thousands of plasmids at about 20 RNAP factories is dependent on either identical promoters or the presence of intronic sequences. This first observation does suggest that specialised RNAP factories exist and that genes are recruited based on similar promoter properties. The second observation though questions the relevance of promoter-determined co-localisation, since identical plasmids with and without an intron are divided into different populations. Intron-containing plasmids often co-localise with the splicing factor SC-35, though usually not at the bright splicing speckles [216]. Still, the complete separation of plasmids depending on the presence of an intron makes it very attractive to speculate on the role of the splicing machinery in co-localisation. A

role for splicing factors in the co-localisation at RNAP factories in the report of Osborne *et al*. [190] has therefore also been hypothesised [217] and was recently shown [218]. In this last study, co-localisation of genes mentioned in the study of Osborne *et al.* is determined, though here in human erythroblasts. Association frequencies are observed of up to 30% between homologous alleles (α-*globin* and *SLCA1* genes) and between different erythroid specific genes (α-*globin* vs. *SLCA1* genes and α-*globin* vs. *ERAF* genes). Importantly, co-localisation is defined as less than 1 μm distance. In this study co-localisation is both addressed in respect to RNAP factories and splicing speckles. Even though co-localising genes often contact the same RNAP factory (up to 27%), almost 100% association with the same splicing speckle is observed. Noteworthy is that FISH images of these erythroid cells show the presence of one giant speckle in these cells and that association is highly enriched at this speckle. The association of these very active genes with this large splicing speckle may be a better explanation for their co-localisation within a region of 1 μm [218]. Spatial proximity induced by association with the same splicing speckle may increase the chance of interaction with the same RNAP factory, suggesting that association with RNAP factories is an effect rather than a cause of co-localisation.

Recently, more elaborate determinations of long-range chromatin looping and interactions have been used to study sub-chromosomal organisation. Two publications applying different techniques both found that individual genomic loci show a highly significant preference for association with other genomic loci [219, 220]. In both studies, co-localisation is mainly determined by transcriptional activity. In the first study, higher-order folding of a 4.3 Mb genomic region on mouse chromosome 14 was studied using FISH [219]. This region consists of four regions that contain actively transcribed genes and that are interspersed with three gene-desserts of minimum 500 kb. Intriguingly, the actively transcribed regions co-localise in over 60% of the cases, while the inactive regions show about 50% co-localisation [219]. In the other study, genome-wide interacting partners of two genomic loci were determined using 4C [220]. 4C (Chromosome Conformation Capture on Chip) is a modification to the 3C-technique, which allows unbiased micro-array-based screening of the entire genome for regions that contact a locus of choice. The ubiquitously expressed *Rad23a* gene, located on mouse chromosome 8, preferentially contacts other actively transcribed genomic regions. Similarly, the actively expressed mouse β-globin locus preferentially associates with transcriptionally active regions. In contrast, the inactive β-globin locus, in a different tissue, mainly contacts regions with

inactive genes and to a lesser extent gene-desserts [220]. Even though CT organisation is probabilistic rather than fixed (discussed in **Section 1.4.2**), these results strongly indicate that CTs are further organised in active and inactive domains.
A remaining issue that still is hardly understood, is how chromatin looping actually results in transcriptional regulation. At the well-characterized mouse β-globin locus, the presence of the LCR, that is located up to 60 kb away from its target genes, is able to upregulate gene expression 25 – 100 fold [221]. Similarly, the removal of several transcription factors known to be required for high levels of globin gene expression, also results in decreased chromatin looping at the β-globin locus [205, 206]. Spatial proximity of the LCR and its associated DNA-bound factors therefore results in an enormous increase in transcription. Thus, long-range chromatin interactions must allow RNAPII to increase initiation, elongation or re-initiation of transcription. Future research hopefully will be able to elucidate what mechanisms are responsible for long-range gene-activation and repression through chromatin looping.

1.4.4- Interchromosomal transcriptional regulation

In recent years, inter-chromosomal gene communication has become subject of intense research. In a large number of reports interacting partners on different chromosomes are identified and it is suggested that these interactions result in interchromosomal transcriptional regulation.
In insects and plants interchromosomal transcriptional regulation is well-studied and firmly established [222-224]. Transvection, mainly studied in Drosophila *ssp.*, is a mechanism where homologous chromosome pairing allows the regulation of a gene by a regulatory element on the other chromosome (see **Figure 1-6A**). Genetic proof of transvection is shown in a study on the Abdominal gene-locus. This locus contains the *Abd-A* and *Abd-B* genes, which are regulated by several infra-abdominal (*iab*) elements. Expression of the *Abd-B* gene, without the required *iab-7* regulatory element *in cis*, is restored when the *iab-7* element on the homologous allele is coupled to a defective *Abd-B* gene [225]. Homologous pairing, which is a common phenomenon in *Drosophila*, allows the two alleles to communicate and cause transcriptional activation *in trans*. Besides transcriptional activation, transvection effects leading to repression and silencing have also been observed (see **Figure 1-6A** (bottom)). A more general definition of transvection therefore is "*[a gene]... whose function can be altered by homologous pairing*" [223].

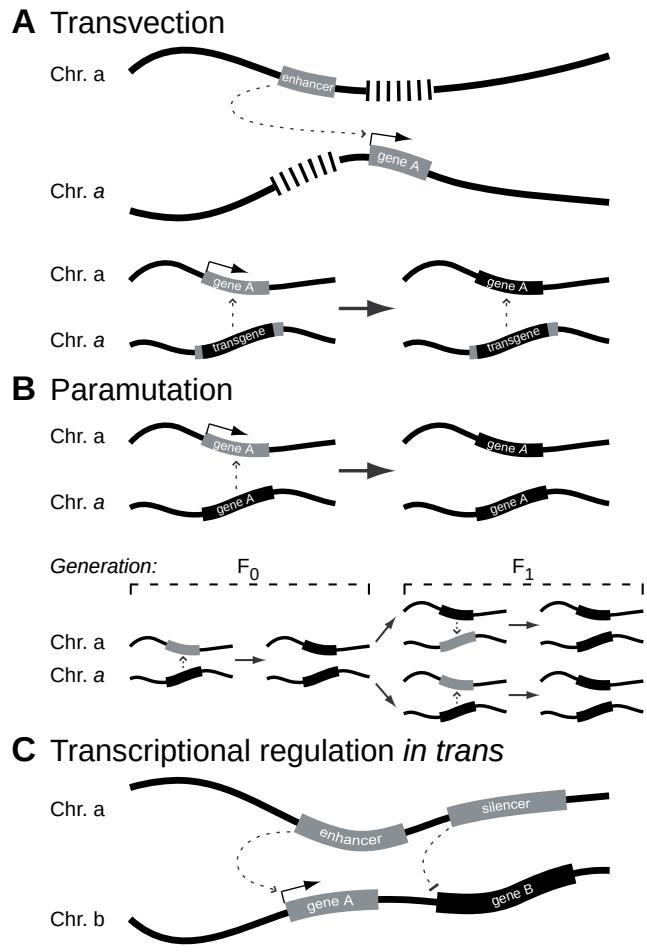


Figure 1-6. Interchromosomal chromatin interactions and transcriptional regulation. In all panels active genes are depicted by grey boxes and silenced gene are depicted by black boxes. **(A)** *(top)* Transvection as described is Drosophila melanogaster: an enhancer without its linked target gene on the one allele activates the target gene without its linked enhancer located on the homologous alleles. *(bottom)* Transvection as described in mammalian cells: a silenced transgene on one allele silences the endogenous locus on the homologous allele. **(B)** Paramutation is variation to transvection, where the active allele is hereditably silenced by the homologous silenced allele *(top)*. The newly silenced allele also possesses the potential to silence homologous alleles in subsequent generations *(bottom)*. **(C)** Trans-interaction is the association or interaction between two genomic loci located on different chromosomes. Trans-regulation involves activation *(left)* or repression *(right)* of a target gene by a regulatory element located at a different chromosome. Both transvection and paramutation are special examples of trans-regulation.

A second mechanism of interchromosomal transcriptional regulation is paramutation (see **Figure 1-6B**). Paramutation, which has been extensively studied in plants, may resemble transvection because it results in interchromosomal transcriptional regulation between homologous alleles. But, in contrast to transvection, paramutation involves the hereditary transfer of epigenetic information from the one allele to the other. Interestingly, this epigenetic state is not only heritably maintained for generations, but in these subsequent generations can even be transmitted from the newly modified allele to other unmodified alleles [226]. The best characterized example of paramutation involves the maize *b1* locus. Paramutation at *b1* involves two genetically identical alleles: (i) the highly expressed *B-I* allele and (ii) the weakly expressed *B'* allele. Most interestingly, crosses between homozygous *B-I* and *B'* plants lead to a heritable change of the *B-I* allele into the *B'* state. In subsequent crosses, the newly established B' allele is also able to convert B-I alleles into B' alleles [227, 228]. A tandem repeat that is located upstream from the locus plays an important role in paramutation. Its DNA methylation status differs between B-I and B' and the repeat is thought to induce paramutation either via an siRNA intermediate and/or through pairing of the homologs.

In recent years, also several examples of possible interchromosomal transcriptional regulation in mammals have been reported. Unfortunately, most reports depend on correlations rather than mechanistic determination. Therefore, the question still remains whether interchromosomal transcriptional regulation in mammals is a mechanism that is naturally occurring. A first issue, whether in mammalian cells transcriptional regulation between elements located on different DNA molecules is possible, has been unambiguously determined. In a first study, human cells were transfected with a plasmid containing an enhancer and a reporter gene. When the same plasmid is linearised, the distance between the promoter and enhancer precludes their efficient communication and transcription of the reporter gene dramatically drops. When communication between enhancer and gene is restored by a streptavidin bridge, efficient transcription is also largely restored, showing that gene and regulatory element don't need to be located on the same DNA molecule [229]. In a follow-up study, human cells were transfected with two different plasmids, one containing a reporter gene and the other containing an enhancer. Addition of *Drosophila* GAGA-binding sites to both plasmids results in a strong increase in transcriptional activation, when GAGA is expressed [230]. This result shows that proteins in mammalian cells can induce gene-expression by attaching two regulatory elements together that are not located on the same DNA molecule [229, 230].

Transvection and paramutation have been proposed as explanations for a number of transcriptionally regulatory effects observed in mammalian cells. In all these studies, changes in the expression of endogenous imprinted alleles are observed after genetic modification at the homologous allele or addition of additional transgenic copies and seem dictated by cellular compensation mechanisms that aim to prevent changes in a gene's expression level [231-233]. The active paternal *U2qf1-rs1* gene in mice testis becomes gradually methylated and deactivated after random integration of an additional copy of the gene. Introduction of the transgene initially results in increased overall transcription of the gene, which is partially abolished by methylation of the endogenous paternal allele [231]. Similarly, deletion of the active maternal *H19* gene in the murine Igf2/H19 locus results in a decrease in DMR methylation at the paternal allele. Deletion of the *H19* gene is proposed to increase transcription of the maternal *Igf2* allele, which is compensated for by decreasing the expression of the paternal allele [232]. The replacement of the paternal *Rasgrf1* DMD (differentially methylated domain) with an *Igf2r* methylated domain in mice decreases *Rasgfr1* expression on the paternal allele, but is compensated for by increased methylation and activation of the maternal allele [233]. In these three examples, transvection or paramutation effects that involve pairing of homologous or transgenic-homologous alleles are proposed (see **Figure 1-6A, B**). Unfortunately, these observations are based on expression analyses and methylation assays and not on co-localisation studies. Since these three studies all involve imprinted genes that are involved in cell-proliferation, another plausible explanation could be that cells expressing these genes at optimal levels have a growth advantage. Therefore, cells with spurious methylation at these alleles could be selected, providing an alternative explanation not involving homologous pairing (proposed in [232]). Recently, a case of transvection in human cancer cells was more thoroughly described [234]. In these cells, the *CCND1* locus is reactivated after translocation, and this is accompanied by demethylation on both alleles, proposed to be dependent on communication between alleles. Probably as a result of this reactivation, cells tend to lose the translocated chromosome containing the active *CCND1* locus. After loss of the translocated chromosome, remethylation at the unaffected chromosome is observed, which leads to silencing of the *CCND1* locus again. Subsequent cell-fusions between cells with the translocated chromosome and cells that had lost the translocated chromosome reveals a more than expected decrease in methylation at both *CCND1* loci again, indicating that the translocated allele indeed affects the homologous allele. Importantly, *CCND1* expression is never observed from

the non-translocated allele. In this study, in around 40% of cells alleles were found to co-localise (2 of 3 *CCND1* alleles / nucleus) at the edge of the nucleolus, which is proposed to allow the observed transvection effect [234]. Unfortunately, the relevance of this co-localisation and transvection is difficult to assess. Even without the translocated allele, the normal alleles frequently associate with the nucleolus. Combined with the polyploidy of the cells, the question therefore arises whether co-localisation could be due to random association with the nucleolus. Even though the observed changes in DNA methylation state are intriguing, the lack of effect on the transcriptional level indicates that the observed pairing has no regulatory function.

Two seemingly very promising examples of communication between alleles have recently been presented by one group [235, 236]. In the first study, transgenic mice carrying the Cre-recombinase and a LoxP reporter construct at the *Rosa26* or *Rxrα* locus, were found to lose recombination ability after the second generation [235]. In contrast to expectation, expression of the Cre-recombinase is maintained, but the LoxP sites become methylated, thereby precluding their activity. In a subsequent generation, the methylated region increases in size and this expansion progresses in somatic tissues. Studies on the *Rxrα* locus show that in mice carrying a methylated and an unmethylated allele, both alleles will be methylated in all offspring [235]. The authors propose that a paramutation-like effect must be involved, since the modification of the one allele by the other is hereditary maintained and the newly methylated alleles even seem to be able to induce methylation of previously unaffected alleles (initially this effect was coined transvection, but in [237] the authors acknowledge that the term paramutation better covers the effect). Unfortunately, in this study no co-localisation data is presented, which does not allow the assessment of homologous pairing in this process. The second example of communication *in trans* from Rassoulzadegan *et al.* involves a paramutation-like effect at the *Kit* locus in transgenic mice [236]. Over 90% of WT progeny from a heterozygous *Kit* knock-out parent and a WT parent carry a typical white spotted tail that is normally associated with heterozygous *Kit* animals. This high, non-mendelian, percentage shows that the effect is transmitted when the transgenic allele is not present. Interestingly, genotypical WT mice showing the heterozygous phenotype are able to transfer the phenotype to their progeny, though here only about 40% penetrance is observed [236]. The hereditary nature suggests a paramutation-like effect. Interestingly, there is no evidence that the effects are mediated by homologous pairing or that physical communication between the two alleles is involved. Rather, the

knock-out allele exhibits strongly increased aberrant transcription from the *Kit* locus, which induces the same transcriptional defect from the WT allele. Injection of these aberrant RNAs in one-cell embryos induces the same paramutation-like effect [236]. Importantly, this example of paramutation does not involve homologous pairing or physical interchromosomal communication. Therefore, physical contacts may not be necessary for alleles to communicate.

An intriguing, but solely observation-based, example of homologous pairing has been proposed to regulate X-chromosome inactivation [238, 239]. X-chromosome inactivation depends on stochastic mono-allelic activation of the non-coding *Xist*-RNA, which results in chromatin condensation of only the X-chromosome where *Xist* is expressed [240]. Counting of the number of X-chromosomes, and therefore the decision of the number of X-chromosomes that need to be inactivated, is proposed to depend on transient "kissing" of X-chromosomes prior to inactivation [238, 239]. The relevance of this "kissing" is unclear though. Even though the two X-chromosomes seem to become positioned more closely together prior to inactivation, the actual distance between the two chromosomes still tends to be quite large. Recently, as an alternative explanation, a chance model that includes positive selection of cells with the right number of active X-chromosomes was proposed [241]. In this model homologous X-chromosome pairing would not be required, and therefore alternative explanations for the observed repositioning of X-chromosomes may need to be considered.

Currently the holy grail in interchromosomal transcriptional regulation seems to be the regulation of genomic loci by regulatory sequences located on different chromosomes (transcriptional regulation *in trans*, **Figure 1-6C**). In recent years, 3C and derivative techniques have identified several of these possible interchromosomal regulatory interactions. The first transcriptional regulation *in trans* was reported between the *Ifng* gene and the T_H2 locus in mouse T-helper cells [242]. During development, naïve T-helper cells can commit to two cell-fates, which are distinguished by activation of either the *Ifng* or *IL4* gene. In T_H2 cells, the T_H2 LCR interacts with the *IL4*, *IL5* and *IL13* genes ([199] and **Section 1.4.3**). By applying the 3C-technique, the LCR is also found to contact the *Ifng* gene in naïve $CD4^+$ T-cells. In contrast to the situation in T_H2 cells, the interchromosomal interaction between the *Ifng* gene and T_H2 LCR seems to keep the two loci in a so-called poised conformation, which is a repressed state that allows rapid activation of one locus after cell-fate has been determined [242]. Maximum co-localisation frequencies, as determined by DNA-FISH, are about 40%. This implies that

the regulatory mechanism does not require permanent interchromosomal interaction, but leaves an epigenetic mark that can either be easily reversed or has a temporal character. A similar interchromosomal regulatory mechanism, where two loci are kept in a poised state, is presented in the previously discussed study on the mono-allelicly expressed *GFAP* gene [189]. The astrocyte marker genes *GFAP* and *S100*β show 20% interchromosomal co-localisation in neuronal precursor cells, where the two genes are inactive. Activation of the genes during development strongly decreases co-localisation, suggesting that these interchromosomal interactions could keep the genes in a poised state [189].

In two other studies, interchromosomal interactions involving the imprinted Igf2/H19 locus have been determined [243, 244]. Both studies, which use two different modified versions of the 3C-technique, find very different and contradictory interchromosomal interaction partners, thereby strongly complicating the interpretation of the results. In the one study, on mouse fibroblasts and ES cells, the ACT-technique (Associated Chromosome Trap) identifies two regions that interact with the DMR of the Igf12/H19 locus. Both these regions are intergenic regions located on other chromosomes. DNA-FISH studies reveal that the interaction with one of the regions exists in 32% to 40% of the cells. Co-localisation was also determined in cells where either the paternal or maternal DMR is deleted, showing that co-localisation only involves the maternal DMR [243]. Interestingly, this is the allele where the *Igf2* gene is silenced by preventing it to interact with its enhancer ([89, 90] and **Section 1.4.1**). An expression analysis of two genes surrounding one of the interchromosomal interaction partners shows that abolishment of interchromosomal interactions represses genes that are located on the paternal allele of this locus. This repressive interchromosomal regulatory mechanism therefore seems to be established between genomic loci of different parental origin [243]. The other study on interchromosomal interaction partner of the Igf2/H19 locus was done on neonatal mouse livers [244]. In this study, a different variation to the 3C-technique was applied, which is also abbreviated to 4C (Circular Chromosome Conformation Capture, not to be confused with the previously introduced 4C-technique in [220]). Here, a larger number of interchromosomal interaction partners is identified, some of which only interact with the maternal allele. Co-localisation, as determined by 3D DNA-FISH, is never detected in more than 17% of cells, indicating that interactions are relatively infrequent [244]. More interestingly though is that none of the interacting regions identified in the study of Ling *et al.* is found in this study. In the paper of Zhao *et al.*, it is acknowledged that they

likely only identified a subset of interaction partners, but this does not explain the very black-and-white result presented by Ling *et al.* [243, 244]. It could be that the Igf2/H19 locus engages in different interchromosomal interactions in different cell types, which would be an interesting topic for future research. Recently, the previously mentioned 4C-technique [244] has also been used to identify interchromosomal associations after virus infection [245]. After virus infection, the *IFN-β* gene should be mono-allelicly expressed to initiate antiviral response. In this report, low-level interchromosomal interactions are proposed to induce stochastic activation of the gene. Using 4C, two NF-κB binding sites *in trans* and one site *in cis* are identified that interact with the *IFN-β* gene. 3D-FISH experiments show that in maximal 27% of the cells one of the three loci is interacting with the *IFN-β* gene, while triple interactions are only observed in 4% of the cells. It is therefore proposed that this small population of interacting alleles could be an explanation for the initial antiviral response. This response includes a positive feedback loop that activates non-interacting alleles in the same and other cells [245]. Unfortunately this model is solely based on observations and no mechanistic studies have been performed, making it hard to assess the feasibility of this model.

Two studies, again generating highly contradictory data, have addressed the question whether interchromosomal interactions may be the mechanism behind olfactory receptor (OR) choice [246, 247]. Using 3C, the olfactory *H* enhancer element is found to contact OR genes on other chromosomes. Combined DNA and RNA-FISH reveals that the *H* enhancer contacts 85% of the actively transcribing *M50* OR alleles, while no inactive alleles are contacted. This suggests that contacts between the *H* enhancer and OR genes are the determinant for olfactory receptor choice [246]. Surprisingly, in a subsequent study where the *H* enhancer is deleted, only 7 OR genes are affected and these are all located in the direct vicinity of the element. Deletion of the *H* element does not affect any of the studied OR genes *in trans* [247]. A functional relevance for the interaction between the active OR gene and the *H* element therefore is lacking.

Another high-profile paper dealing with interchromosomal transcriptional regulation was recently retracted due to the presence of duplicated figures [248, 249]. Since the conclusions from this paper have explicitly not been retracted, this paper will still be discussed here. In this study, a number of estrogen induced genes located on different chromosomes are identified that engage in interchromosomal interactions. For the identification, again a different variation to the 3C-technique was used, named 3D (Deconvolution of DNA interaction by DSL). Interchromosomal interactions are established

by the directional movement to shared splicing speckles after transcriptional activation. This directional movement is dependent on the actin/myosin transport machinery and the LSD1 histone lysine demethylase. Surprisingly, loci in this study show absolute 100% co-localisation frequencies rather than the usual probabilistic numbers [248]. Further studies on estrogen-mediated co-localisation at splicing speckles will hopefully reveal whether this mechanism allows such discrete nuclear movements or whether an artefact was reported.

Most recently, a newly devised technique based on a combination of ChIP and 3C has been used to identify interchromosomal interactions through their association with specific proteins [250]. This technique, as the latest addition to 3C-based techniques, has been wittingly named 6C (Combined 3C-ChIP-cloning). The aim of this study was to identify interchromosomal interactions mediated by the EZH2 protein. EZH2 is part of the PRC2 and PRC3 PcG complexes and deposits the H3K27 methylation mark (see **Section 1.4.1**). Using the 6C-technique, several interacting regions *in trans* are identified, some of which may even be interacting with more than one region at the same time. Unfortunately, no FISH studies are shown and therefore it is difficult to determine whether these are frequent co-localisation event or whether these are random interactions that have accidentally been picked up [250]. Interestingly, the EZH2 protein is required for deposition of the H3K27 methylation mark, but not for maintenance. After deposition, maintenance of this mark is achieved by the PRC1 complex, which is found to cluster in PcG bodies. A more interesting question therefore seems to have been whether actively silenced regions may cluster at PcG bodies, like actively expressed regions may do at splicing speckles and RNAP factories.

Together, the results presented in this section show that reports on interchromosomal regulatory mechanisms in mammals are still very anecdotal. Novel 3C-based screening methods are often used to discover the interchromosomal interactions. It is confusing though that these techniques do not necessarily seem to generate similar results. For example, while 4C (3C-on-Chip) consistently finds abundant interactions with DNA segments proximal on the same chromosome and much less long-range and interchromosomal interactions, most other techniques appear to almost exclusively identify interchromosomal interactions. From a polymer physics point of view one would expect the former, but not the latter, result. The lack of genetic evidence supporting the existence of regulatory mechanisms acting *in trans*, combined with the usually far from high co-localisation frequencies, raises the question whether interchromosomal

regulatory mechanisms are a purposely employed mechanism in mammals. So far the only seemingly well-explainable interchromosomal interactions are at splicing speckles [218]. Since interaction frequencies reported for regulatory interchromosomal interactions are often in the same range as the interaction frequencies at splicing speckles, it may be that the proposed regulatory mechanisms are merely representing spurious association with splicing speckles or other nuclear substructures.

1.5- References

1. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002). Cells and genomes. In Molecular Biology of the Cell, Fourth Edition, B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter, eds. (New York, USA: Garland Science), pp. 3-46.
2. Heitz, E. (1928). Das Heterochromatin der Moose. Jahrb Wiss Bot 69, 762-818.
3. Heitz, E. (1932). Der Bau der somatischen Kerne von Drosophila melanogaster. Z indukt Abstammungs Vererbungsl 54, 248-249.
4. Zacharias, H. (1995). Emil Heitz (1892-1965): chloroplasts, heterochromatin, and polytene chromosomes. Genetics 141, 7-14.
5. Heitz, E. (1929). Heterochromatin, Chromocentren, Chromomeren (Vorläufige Mitteilung). Ber Deutsch Bot Ges 47, 274-284.
6. Scheer, U., and Hock, R. (1999). Structure and function of the nucleolus. Curr Opin Cell Biol 11, 385-390.
7. Watson, M.L. (1955). The nuclear envelope; its structure and relation to cytoplasmic membranes. J Biophys Biochem Cytol 1, 257-270.
8. Anderson, D.J., and Hetzer, M.W. (2008). The life cycle of the metazoan nuclear envelope. Curr Opin Cell Biol 20, 386-392.
9. Suntharalingam, M., and Wente, S.R. (2003). Peering through the pore: nuclear pore complex structure, assembly, and function. Dev Cell 4, 775-789.
10. Akhtar, A., and Gasser, S.M. (2007). The nuclear envelope and transcriptional control. Nat Rev Genet 8, 507-517.
11. Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W., and van Steensel, B. (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature 453, 948-951.
12. Spector, D.L. (2006). SnapShot: Cellular bodies. Cell 127, 1071.
13. Spector, D.L. (2001). Nuclear domains. J Cell Sci 114, 2891-2893.
14. Iborra, F.J., Pombo, A., Jackson, D.A., and Cook, P.R. (1996). Active RNA polymerases are localized within discrete transcription "factories' in human nuclei. J Cell Sci 109 (Pt 6), 1427-1436.
15. Jackson, D.A., Iborra, F.J., Manders, E.M., and Cook, P.R. (1998). Numbers and organization of RNA polymerases, nascent transcripts, and transcription units in HeLa nuclei. Mol Biol Cell 9, 1523-1536.
16. Pombo, A., Jackson, D.A., Hollinshead, M., Wang, Z., Roeder, R.G., and Cook, P.R. (1999). Regional specialization in human nuclei: visualization of discrete sites of transcription by RNA polymerase III. Embo J 18, 2241-2253.
17. Faro-Trindade, I., and Cook, P.R. (2006). A conserved organization of transcription during embryonic stem cell differentiation and in cells with high C value. Mol Biol Cell 17, 2910-2920.
18. Martin, S., and Pombo, A. (2003). Transcription factories: quantitative studies of nanostructures in the mammalian nucleus. Chromosome Res 11, 461-470.
19. Cajal, R. (1903). Un sencillo metodo de coloracion seletiva del reticulo protoplasmatico y sus efectos en los diversos organos nerviosos de vertebrados e invertebrados. Trab Lab Investig Biol Univ Madrid 2, 129-221.

20. Stanek, D., and Neugebauer, K.M. (2006). The Cajal body: a meeting place for spliceosomal snRNPs in the nuclear maze. Chromosoma 115, 343-354.
21. Lamond, A.I., and Spector, D.L. (2003). Nuclear speckles: a model for nuclear organelles. Nat Rev Mol Cell Biol 4, 605-612.
22. Saitoh, N., Spahr, C.S., Patterson, S.D., Bubulya, P., Neuwald, A.F., and Spector, D.L. (2004). Proteomic analysis of interchromatin granule clusters. Mol Biol Cell 15, 3876-3890.
23. Saurin, A.J., Shiels, C., Williamson, J., Satijn, D.P., Otte, A.P., Sheer, D., and Freemont, P.S. (1998). The human polycomb group complex associates with pericentromeric heterochromatin to form a novel nuclear domain. J Cell Biol 142, 887-898.
24. Voncken, J.W., Schweizer, D., Aagaard, L., Sattler, L., Jantsch, M.F., and van Lohuizen, M. (1999). Chromatin-association of the Polycomb group protein BMI1 is cell cycle-regulated and correlates with its phosphorylation status. J Cell Sci 112 (Pt 24), 4627-4639.
25. Lund, A.H., and van Lohuizen, M. (2004). Polycomb complexes and silencing mechanisms. Curr Opin Cell Biol 16, 239-246.
26. Zhong, S., Salomoni, P., and Pandolfi, P.P. (2000). The transcriptional role of PML and the nuclear body. Nat Cell Biol 2, E85-90.
27. Maul, G.G., Negorev, D., Bell, P., and Ishov, A.M. (2000). Review: properties and assembly mechanisms of ND10, PML bodies, or PODs. J Struct Biol 129, 278-287.
28. Maul, G.G., Yu, E., Ishov, A.M., and Epstein, A.L. (1995). Nuclear domain 10 (ND10) associated proteins are also present in nuclear bodies and redistribute to hundreds of nuclear sites after stress. J Cell Biochem 59, 498-513.
29. Desbois, C., Rousset, R., Bantignies, F., and Jalinot, P. (1996). Exclusion of Int-6 from PML nuclear bodies by binding to the HTLV-I Tax oncoprotein. Science 273, 951-953.
30. Maul, G.G. (1998). Nuclear domain 10, the site of DNA virus transcription and replication. Bioessays 20, 660-667.
31. Maul, G.G., Jensen, D.E., Ishov, A.M., Herlyn, M., and Rauscher, F.J., 3rd (1998). Nuclear redistribution of BRCA1 during viral infection. Cell Growth Differ 9, 743-755.
32. Anton, L.C., Schubert, U., Bacik, I., Princiotta, M.F., Wearsch, P.A., Gibbs, J., Day, P.M., Realini, C., Rechsteiner, M.C., Bennink, J.R., and Yewdell, J.W. (1999). Intracellular localization of proteasomal degradation of a viral antigen. J Cell Biol 146, 113-124.
33. Jackson, D.A., Hassan, A.B., Errington, R.J., and Cook, P.R. (1993). Visualization of focal sites of transcription within human nuclei. Embo J 12, 1059-1065.
34. Jackson, D.A. (2005). The amazing complexity of transcription factories. Brief Funct Genomic Proteomic 4, 143-157.
35. Kimura, H., Sugaya, K., and Cook, P.R. (2002). The transcription cycle of RNA polymerase II in living cells. J Cell Biol 159, 777-782.
36. Noy, A., Vezenov, D.V., Kayyem, J.F., Meade, T.J., and Lieber, C.M. (1997). Stretching and breaking duplex DNA by chemical force microscopy. Chem Biol 4, 519-527.
37. Cook, P.R. (1999). The organization of replication and transcription. Science 284, 1790-1795.
38. Muller, W.G., Rieder, D., Karpova, T.S., John, S., Trajanoski, Z., and McNally, J.G. (2007). Organization of chromatin and histone modifications at a transcription site. J Cell Biol 177, 957-967.
39. Scheuermann, M.O., Murmann, A.E., Richter, K., Gorisch, S.M., Herrmann, H., and Lichter, P. (2005).

Characterization of nuclear compartments identified by ectopic markers in mammalian cells with distinctly different karyotype. Chromosoma 114, 39-53.

40. Murmann, A.E., Gao, J., Encinosa, M., Gautier, M., Peter, M.E., Eils, R., Lichter, P., and Rowley, J.D. (2005). Local gene density predicts the spatial position of genetic loci in the interphase nucleus. Exp Cell Res 311, 14-26.
41. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002). DNA and chromosomes. In Molecular Biology of the Cell, Fourth Edition, B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter, eds. (New York, USA: Garland Science), pp. 191-234.
42. Horn, P.J., and Peterson, C.L. (2002). Chromatin higher order folding: wrapping up transcription. Science 297, 1824-1827.
43. Woodcock, C.L. (2006). Chromatin architecture. Curr Opin Struct Biol 16, 213-220.
44. Tremethick, D.J. (2007). Higher-order structures of chromatin: the elusive 30 nm fiber. Cell 128, 651-654.
45. Oudet, P., Gross-Bellard, M., and Chambon, P. (1975). Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. Cell 4, 281-300.
46. Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature 389, 251-260.
47. Thomas, G.J., Jr., and Prescott, B. (1977). Secondary structure of histones and DNA in chromatin. Science 197, 385-388.
48. Kornberg, R.D. (1977). Structure of chromatin. Annu Rev Biochem 46, 931-954.
49. Smith, R.D., Seale, R.L., and Yu, J. (1983). Transcribed chromatin exhibits an altered nucleosomal spacing. Proc Natl Acad Sci U S A 80, 5505-5509.
50. Benezra, R., Cantor, C.R., and Axel, R. (1986). Nucleosomes are phased along the mouse beta-major globin gene in erythroid and nonerythroid cells. Cell 44, 697-704.
51. Woodcock, C.L., Skoultchi, A.I., and Fan, Y. (2006). Role of linker histone in chromatin structure and function: H1 stoichiometry and nucleosome repeat length. Chromosome Res 14, 17-25.
52. Rundquist, I., and Lindner, H.H. (2006). Analyses of linker histone - chromatin interactions in situ. Biochem Cell Biol 84, 427-436.
53. Bresnick, E.H., Bustin, M., Marsaud, V., Richard-Foy, H., and Hager, G.L. (1992). The transcriptionally-active MMTV promoter is depleted of histone H1. Nucleic Acids Res 20, 273-278.
54. Kim, A., and Dean, A. (2003). A human globin enhancer causes both discrete and widespread alterations in chromatin structure. Mol Cell Biol 23, 8099-8109.
55. O'Neill, T.E., Meersseman, G., Pennings, S., and Bradbury, E.M. (1995). Deposition of histone H1 onto reconstituted nucleosome arrays inhibits both initiation and elongation of transcripts by T7 RNA polymerase. Nucleic Acids Res 23, 1075-1082.
56. Pennings, S., Meersseman, G., and Bradbury, E.M. (1994). Linker histones H1 and H5 prevent the mobility of positioned nucleosomes. Proc Natl Acad Sci U S A 91, 10275-10279.
57. Fan, Y., Nikitina, T., Morin-Kensicki, E.M., Zhao, J., Magnuson, T.R., Woodcock, C.L., and Skoultchi, A.I. (2003). H1 linker histones are essential for mouse development and affect nucleosome spacing in vivo. Mol Cell Biol 23, 4559-4572.
58. Fan, Y., Nikitina, T., Zhao, J., Fleury, T.J., Bhattacharyya, R., Bouhassira, E.E., Stein, A., Woodcock, C.L., and Skoultchi, A.I. (2005). Histone H1 depletion in mammals alters global chromatin structure

but causes specific changes in gene regulation. Cell 123, 1199-1212.
59. Francis, N.J., Kingston, R.E., and Woodcock, C.L. (2004). Chromatin compaction by a polycomb group protein complex. Science 306, 1574-1577.
60. Tse, C., Sera, T., Wolffe, A.P., and Hansen, J.C. (1998). Disruption of higher-order folding by core histone acetylation dramatically enhances transcription of nucleosomal arrays by RNA polymerase III. Mol Cell Biol 18, 4629-4638.
61. Belmont, A.S., Braunfeld, M.B., Sedat, J.W., and Agard, D.A. (1989). Large-scale chromatin structural domains within mitotic and interphase chromosomes in vivo and in vitro. Chromosoma 98, 129-143.
62. Belmont, A.S., and Bruce, K. (1994). Visualization of G1 chromosomes: a folded, twisted, supercoiled chromonema model of interphase chromatid structure. J Cell Biol 127, 287-302.
63. Robinett, C.C., Straight, A., Li, G., Willhelm, C., Sudlow, G., Murray, A., and Belmont, A.S. (1996). In vivo localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. J Cell Biol 135, 1685-1700.
64. Gerchman, S.E., and Ramakrishnan, V. (1987). Chromatin higher-order structure studied by neutron scattering and scanning transmission electron microscopy. Proc Natl Acad Sci U S A 84, 7802-7806.
65. Schardin, M., Cremer, T., Hager, H.D., and Lang, M. (1985). Specific staining of human chromosomes in Chinese hamster x man hybrid cell lines demonstrates interphase chromosome territories. Hum Genet 71, 281-287.
66. Manuelidis, L. (1985). Individual interphase chromosome domains revealed by in situ hybridization. Hum Genet 71, 288-293.
67. Lichter, P., Cremer, T., Borden, J., Manuelidis, L., and Ward, D.C. (1988). Delineation of individual human chromosomes in metaphase and interphase cells by in situ suppression hybridization using recombinant DNA libraries. Hum Genet 80, 224-234.
68. Cremer, T., and Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nat Rev Genet 2, 292-301.
69. Cremer, T., Cremer, M., Dietzel, S., Muller, S., Solovei, I., and Fakan, S. (2006). Chromosome territories - a functional nuclear landscape. Curr Opin Cell Biol 18, 307-316.
70. Meaburn, K.J., and Misteli, T. (2007). Chromosome territories. Nature 445, 379-781.
71. Zink, D., Cremer, T., Saffrich, R., Fischer, R., Trendelenburg, M.F., Ansorge, W., and Stelzer, E.H. (1998). Structure and dynamics of human interphase chromosome territories in vivo. Hum Genet 102, 241-251.
72. Branco, M.R., and Pombo, A. (2006). Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. PLoS Biol 4, e138.
73. Mitchell, J.A., and Fraser, P. (2008). Transcription factories are nuclear subcompartments that remain in the absence of transcription. Genes Dev 22, 20-25.
74. Palstra, R.J., Simonis, M., Klous, P., Brasset, E., Eijkelkamp, B., and de Laat, W. (2008). Maintenance of long-range DNA interactions after inhibition of ongoing RNA Polymerase II transcription. PLoS ONE 3, e1661.
75. Abney, J.R., Cutler, B., Fillbach, M.L., Axelrod, D., and Scalettar, B.A. (1997). Chromatin dynamics in interphase nuclei and its implications for nuclear structure. J Cell Biol 137, 1459-1468.

76. Chubb, J.R., Boyle, S., Perry, P., and Bickmore, W.A. (2002). Chromatin motion is constrained by association with nuclear compartments in human cells. Curr Biol 12, 439-445.
77. Manders, E.M., Kimura, H., and Cook, P.R. (1999). Direct imaging of DNA in living cells reveals the dynamics of chromosome formation. J Cell Biol 144, 813-821.
78. Walter, J., Schermelleh, L., Cremer, M., Tashiro, S., and Cremer, T. (2003). Chromosome order in HeLa cells changes during mitosis and early G1, but is stably maintained during subsequent interphase stages. J Cell Biol 160, 685-697.
79. Gerlich, D., Beaudouin, J., Kalbfuss, B., Daigle, N., Eils, R., and Ellenberg, J. (2003). Global chromosome positions are transmitted through mitosis in mammalian cells. Cell 112, 751-764.
80. Thomson, I., Gilchrist, S., Bickmore, W.A., and Chubb, J.R. (2004). The radial positioning of chromatin is not inherited through mitosis but is established de novo in early G1. Curr Biol 14, 166-172.
81. Essers, J., van Cappellen, W.A., Theil, A.F., van Drunen, E., Jaspers, N.G., Hoeijmakers, J.H., Wyman, C., Vermeulen, W., and Kanaar, R. (2005). Dynamics of relative chromosome position during the cell cycle. Mol Biol Cell 16, 769-775.
82. Bernstein, B.E., Meissner, A., and Lander, E.S. (2007). The mammalian epigenome. Cell 128, 669-681.
83. Kouzarides, T. (2007). Chromatin modifications and their function. Cell 128, 693-705.
84. Ioshikhes, I.P., and Zhang, M.Q. (2000). Large-scale human promoter mapping using CpG islands. Nat Genet 26, 61-63.
85. Jaenisch, R., and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet 33 Suppl, 245-254.
86. Klose, R.J., and Bird, A.P. (2006). Genomic DNA methylation: the mark and its mediators. Trends Biochem Sci 31, 89-97.
87. Bell, A.C., and Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature 405, 482-485.
88. Hark, A.T., Schoenherr, C.J., Katz, D.J., Ingram, R.S., Levorse, J.M., and Tilghman, S.M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. Nature 405, 486-489.
89. Leighton, P.A., Saam, J.R., Ingram, R.S., Stewart, C.L., and Tilghman, S.M. (1995). An enhancer deletion affects both H19 and Igf2 expression. Genes Dev 9, 2079-2089.
90. Murrell, A., Heeson, S., and Reik, W. (2004). Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops. Nat Genet 36, 889-893.
91. Nan, X., Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N., and Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature 393, 386-389.
92. Zhang, Y., Ng, H.H., Erdjument-Bromage, H., Tempst, P., Bird, A., and Reinberg, D. (1999). Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. Genes Dev 13, 1924-1935.
93. Bird, A.P., and Wolffe, A.P. (1999). Methylation-induced repression--belts, braces, and chromatin. Cell 99, 451-454.
94. Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes Dev 16, 6-21.

95. Fuks, F. (2005). DNA methylation and histone modifications: teaming up to silence genes. Curr Opin Genet Dev 15, 490-495.
96. Martin, C., and Zhang, Y. (2005). The diverse functions of histone lysine methylation. Nat Rev Mol Cell Biol 6, 838-849.
97. Gray, S.G., and Ekstrom, T.J. (2001). The human histone deacetylase family. Exp Cell Res 262, 75-83.
98. Fazzari, M.J., and Greally, J.M. (2004). Epigenomics: beyond CpG islands. Nat Rev Genet 5, 446-455.
99. Saxonov, S., Berg, P., and Brutlag, D.L. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proc Natl Acad Sci U S A 103, 1412-1417.
100. Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Paabo, S., Rebhan, M., and Schubeler, D. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet 39, 457-466.
101. Bibikova, M., Chudin, E., Wu, B., Zhou, L., Garcia, E.W., Liu, Y., Shin, S., Plaia, T.W., Auerbach, J.M., Arking, D.E., Gonzalez, R., Crook, J., Davidson, B., Schulz, T.C., Robins, A., Khanna, A., Sartipy, P., Hyllner, J., Vanguri, P., Savant-Bhonsale, S., Smith, A.K., Chakravarti, A., Maitra, A., Rao, M., Barker, D.L., Loring, J.F., and Fan, J.B. (2006). Human embryonic stem cells have a unique epigenetic signature. Genome Res 16, 1075-1083.
102. Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B.E., Nusbaum, C., Jaffe, D.B., Gnirke, A., Jaenisch, R., and Lander, E.S. (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature 454, 766-770.
103. Weber, M., Davies, J.J., Wittig, D., Oakeley, E.J., Haase, M., Lam, W.L., and Schubeler, D. (2005). Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat Genet 37, 853-862.
104. Chen, R.Z., Pettersson, U., Beard, C., Jackson-Grusby, L., and Jaenisch, R. (1998). DNA hypomethylation leads to elevated mutation rates. Nature 395, 89-93.
105. Gaudet, F., Hodgson, J.G., Eden, A., Jackson-Grusby, L., Dausman, J., Gray, J.W., Leonhardt, H., and Jaenisch, R. (2003). Induction of tumors in mice by genomic hypomethylation. Science 300, 489-492.
106. Metivier, R., Gallais, R., Tiffoche, C., Le Peron, C., Jurkowska, R.Z., Carmouche, R.P., Ibberson, D., Barath, P., Demay, F., Reid, G., Benes, V., Jeltsch, A., Gannon, F., and Salbert, G. (2008). Cyclical DNA methylation of a transcriptionally active promoter. Nature 452, 45-50.
107. Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science 293, 1074-1080.
108. Turner, B.M., Birley, A.J., and Lavender, J. (1992). Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in Drosophila polytene nuclei. Cell 69, 375-384.
109. Quina, A.S., Buschbeck, M., and Di Croce, L. (2006). Chromatin structure and epigenetics. Biochem Pharmacol 72, 1563-1569.
110. Lachner, M., and Jenuwein, T. (2002). The many faces of histone lysine methylation. Curr Opin Cell Biol 14, 286-298.
111. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3

lysine 9 creates a binding site for HP1 proteins. Nature 410, 116-120.
112. Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature 410, 120-124.
113. Kwon, S.H., and Workman, J. (2008). The Heterochromatin Protein 1 (HP1) Family: Put Away a bias toward HP1. Mol Cells 26.
114. Ye, Q., Callebaut, I., Pezhman, A., Courvalin, J.C., and Worman, H.J. (1997). Domain-specific interactions of human HP1-type chromodomain proteins and inner nuclear membrane protein LBR. J Biol Chem 272, 14983-14989.
115. Dechat, T., Pfleghaar, K., Sengupta, K., Shimi, T., Shumaker, D.K., Solimando, L., and Goldman, R.D. (2008). Nuclear lamins: major factors in the structural organization and function of the nucleus and chromatin. Genes Dev 22, 832-853.
116. Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., and Jenuwein, T. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature 406, 593-599.
117. Nakayama, J., Rice, J.C., Strahl, B.D., Allis, C.D., and Grewal, S.I. (2001). Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. Science 292, 110-113.
118. Volpe, T.A., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I., and Martienssen, R.A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science 297, 1833-1837.
119. Hall, I.M., Shankaranarayana, G.D., Noma, K., Ayoub, N., Cohen, A., and Grewal, S.I. (2002). Establishment and maintenance of a heterochromatin domain. Science 297, 2232-2237.
120. Sewalt, R.G., Lachner, M., Vargas, M., Hamer, K.M., den Blaauwen, J.L., Hendrix, T., Melcher, M., Schweizer, D., Jenuwein, T., and Otte, A.P. (2002). Selective interactions between vertebrate polycomb homologs and the SUV39H1 histone lysine methyltransferase suggest that histone H3-K9 methylation contributes to chromosomal targeting of Polycomb group proteins. Mol Cell Biol 22, 5539-5553.
121. Tachibana, M., Sugimoto, K., Nozaki, M., Ueda, J., Ohta, T., Ohki, M., Fukuda, M., Takeda, N., Niida, H., Kato, H., and Shinkai, Y. (2002). G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. Genes Dev 16, 1779-1791.
122. Tachibana, M., Ueda, J., Fukuda, M., Takeda, N., Ohta, T., Iwanari, H., Sakihama, T., Kodama, T., Hamakubo, T., and Shinkai, Y. (2005). Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. Genes Dev 19, 815-826.
123. Ait-Si-Ali, S., Guasconi, V., Fritsch, L., Yahi, H., Sekhri, R., Naguibneva, I., Robin, P., Cabon, F., Polesskaya, A., and Harel-Bellan, A. (2004). A Suv39h-dependent mechanism for silencing S-phase genes in differentiating but not in cycling cells. Embo J 23, 605-615.
124. Nielsen, S.J., Schneider, R., Bauer, U.M., Bannister, A.J., Morrison, A., O'Carroll, D., Firestein, R., Cleary, M., Jenuwein, T., Herrera, R.E., and Kouzarides, T. (2001). Rb targets histone H3 methylation and HP1 to promoters. Nature 412, 561-565.
125. Lehnertz, B., Ueda, Y., Derijck, A.A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T., and Peters, A.H. (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr Biol 13, 1192-

1200.

126. Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A., and Pirrotta, V. (2002). Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. Cell 111, 185-196.
127. Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science 298, 1039-1043.
128. Kirmizis, A., Bartley, S.M., Kuzmichev, A., Margueron, R., Reinberg, D., Green, R., and Farnham, P.J. (2004). Silencing of human polycomb target genes is associated with methylation of histone H3 Lys 27. Genes Dev 18, 1592-1605.
129. Pan, G., Tian, S., Nie, J., Yang, C., Ruotti, V., Wei, H., Jonsdottir, G.A., Stewart, R., and Thomson, J.A. (2007). Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. Cell Stem Cell 1, 299-312.
130. Shao, Z., Raible, F., Mollaaghababa, R., Guyon, J.R., Wu, C.T., Bender, W., and Kingston, R.E. (1999). Stabilization of chromatin structure by PRC1, a Polycomb complex. Cell 98, 37-46.
131. Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R.S., and Zhang, Y. (2004). Role of histone H2A ubiquitination in Polycomb silencing. Nature 431, 873-878.
132. Vire, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J.M., Bollen, M., Esteller, M., Di Croce, L., de Launoit, Y., and Fuks, F. (2006). The Polycomb group protein EZH2 directly controls DNA methylation. Nature 439, 871-874.
133. Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell 129, 823-837.
134. Ng, H.H., Robert, F., Young, R.A., and Struhl, K. (2003). Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. Mol Cell 11, 709-719.
135. Pokholok, D.K., Harbison, C.T., Levine, S., Cole, M., Hannett, N.M., Lee, T.I., Bell, G.W., Walker, K., Rolfe, P.A., Herbolsheimer, E., Zeitlinger, J., Lewitter, F., Gifford, D.K., and Young, R.A. (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. Cell 122, 517-527.
136. Li, B., Carey, M., and Workman, J.L. (2007). The role of chromatin during transcription. Cell 128, 707-719.
137. Schneider, R., Bannister, A.J., Myers, F.A., Thorne, A.W., Crane-Robinson, C., and Kouzarides, T. (2004). Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. Nat Cell Biol 6, 73-77.
138. Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., Wang, W., Weng, Z., Green, R.D., Crawford, G.E., and Ren, B. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet 39, 311-318.
139. Wysocka, J., Swigut, T., Milne, T.A., Dou, Y., Zhang, X., Burlingame, A.L., Roeder, R.G., Brivanlou, A.H., and Allis, C.D. (2005). WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. Cell 121, 859-872.
140. Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T.,

Badenhorst, P., Wu, C., and Allis, C.D. (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature 442, 86-90.

141. Hong, L., Schroth, G.P., Matthews, H.R., Yau, P., and Bradbury, E.M. (1993). Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA. J Biol Chem 268, 305-314.
142. Wang, Z., Zang, C., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Peng, W., Zhang, M.Q., and Zhao, K. (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. Nat Genet 40, 897-903.
143. Sterner, D.E., and Berger, S.L. (2000). Acetylation of histones and transcription-related factors. Microbiol Mol Biol Rev 64, 435-459.
144. Struhl, K. (1998). Histone acetylation and transcriptional regulatory mechanisms. Genes Dev 12, 599-606.
145. Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. Cell 130, 77-88.
146. Tong, J.K., Hassig, C.A., Schnitzler, G.R., Kingston, R.E., and Schreiber, S.L. (1998). Chromatin deacetylation by an ATP-dependent nucleosome remodelling complex. Nature 395, 917-921.
147. Zhang, Y., LeRoy, G., Seelig, H.P., Lane, W.S., and Reinberg, D. (1998). The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities. Cell 95, 279-289.
148. Rountree, M.R., Bachman, K.E., and Baylin, S.B. (2000). DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. Nat Genet 25, 269-277.
149. Robertson, K.D., Ait-Si-Ali, S., Yokochi, T., Wade, P.A., Jones, P.L., and Wolffe, A.P. (2000). DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. Nat Genet 25, 338-342.
150. Espada, J., Ballestar, E., Fraga, M.F., Villar-Garea, A., Juarranz, A., Stockert, J.C., Robertson, K.D., Fuks, F., and Esteller, M. (2004). Human DNA methyltransferase 1 is required for maintenance of the histone H3 modification pattern. J Biol Chem 279, 37175-37184.
151. Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P., Lee, W., Mendenhall, E., O'Donovan, A., Presser, A., Russ, C., Xie, X., Meissner, A., Wernig, M., Jaenisch, R., Nusbaum, C., Lander, E.S., and Bernstein, B.E. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448, 553-560.
152. Sproul, D., Gilbert, N., and Bickmore, W.A. (2005). The role of chromatin structure in regulating the expression of clustered genes. Nat Rev Genet 6, 775-781.
153. Lanctot, C., Cheutin, T., Cremer, M., Cavalli, G., and Cremer, T. (2007). Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. Nat Rev Genet 8, 104-115.
154. Misteli, T. (2007). Beyond the sequence: cellular organization of genome function. Cell 128, 787-800.
155. Fraser, P., and Bickmore, W. (2007). Nuclear organization of the genome and the potential for gene regulation. Nature 447, 413-417.
156. Sexton, T., Schober, H., Fraser, P., and Gasser, S.M. (2007). Gene regulation through nuclear organization. Nat Struct Mol Biol 14, 1049-1055.

157. Schneider, R., and Grosschedl, R. (2007). Dynamics and interplay of nuclear architecture, genome organization, and gene expression. Genes Dev 21, 3027-3043.
158. Kumaran, R.I., Thakar, R., and Spector, D.L. (2008). Chromatin dynamics and gene positioning. Cell 132, 929-934.
159. Dillon, N. (2008). The impact of gene location in the nucleus on transcriptional regulation. Dev Cell 15, 182-186.
160. Parada, L.A., Roix, J.J., and Misteli, T. (2003). An uncertainty principle in chromosome positioning. Trends Cell Biol 13, 393-396.
161. Levsky, J.M., and Singer, R.H. (2003). Gene expression and the myth of the average cell. Trends Cell Biol 13, 4-6.
162. Raj, A., Peskin, C.S., Tranchina, D., Vargas, D.Y., and Tyagi, S. (2006). Stochastic mRNA synthesis in mammalian cells. PLoS Biol 4, e309.
163. de Laat, W., and Grosveld, F. (2007). Inter-chromosomal gene regulation in the mammalian cell nucleus. Curr Opin Genet Dev.
164. Croft, J.A., Bridger, J.M., Boyle, S., Perry, P., Teague, P., and Bickmore, W.A. (1999). Differences in the localization and morphology of chromosomes in the human nucleus. J Cell Biol 145, 1119-1131.
165. Cremer, M., Kupper, K., Wagler, B., Wizelman, L., von Hase, J., Weiland, Y., Kreja, L., Diebold, J., Speicher, M.R., and Cremer, T. (2003). Inheritance of gene density-related higher order chromatin arrangements in normal and tumor cell nuclei. J Cell Biol 162, 809-820.
166. Boyle, S., Gilchrist, S., Bridger, J.M., Mahy, N.L., Ellis, J.A., and Bickmore, W.A. (2001). The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. Hum Mol Genet 10, 211-219.
167. Cremer, M., von Hase, J., Volm, T., Brero, A., Kreth, G., Walter, J., Fischer, C., Solovei, I., Cremer, C., and Cremer, T. (2001). Non-random radial higher-order chromatin arrangements in nuclei of diploid human cells. Chromosome Res 9, 541-567.
168. Bolzer, A., Kreth, G., Solovei, I., Koehler, D., Saracoglu, K., Fauth, C., Muller, S., Eils, R., Cremer, C., Speicher, M.R., and Cremer, T. (2005). Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. PLoS Biol 3, e157.
169. Kosak, S.T., Skok, J.A., Medina, K.L., Riblet, R., Le Beau, M.M., Fisher, A.G., and Singh, H. (2002). Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. Science 296, 158-162.
170. Zink, D., Amaral, M.D., Englmann, A., Lang, S., Clarke, L.A., Rudolph, C., Alt, F., Luther, K., Braz, C., Sadoni, N., Rosenecker, J., and Schindelhauer, D. (2004). Transcription-dependent spatial arrangements of CFTR and adjacent genes in human cell nuclei. J Cell Biol 166, 815-825.
171. Ragoczy, T., Bender, M.A., Telling, A., Byron, R., and Groudine, M. (2006). The locus control region is required for association of the murine beta-globin locus with engaged transcription factories during erythroid maturation. Genes Dev 20, 1447-1457.
172. Chuang, C.H., Carpenter, A.E., Fuchsova, B., Johnson, T., de Lanerolle, P., and Belmont, A.S. (2006). Long-range directional movement of an interphase chromosome site. Curr Biol 16, 825-831.
173. Kumaran, R.I., and Spector, D.L. (2008). A genetic locus targeted to the nuclear periphery in living cells maintains its transcriptional competence. J Cell Biol 180, 51-65.
174. Reddy, K.L., Zullo, J.M., Bertolino, E., and Singh, H. (2008). Transcriptional repression mediated by

repositioning of genes to the nuclear lamina. Nature 452, 243-247.
175. Finlan, L.E., Sproul, D., Thomson, I., Boyle, S., Kerr, E., Perry, P., Ylstra, B., Chubb, J.R., and Bickmore, W.A. (2008). Recruitment to the nuclear periphery can alter expression of genes in human cells. PLoS Genet 4, e1000039.
176. Malhas, A., Lee, C.F., Sanders, R., Saunders, N.J., and Vaux, D.J. (2007). Defects in lamin B1 expression or processing affect interphase chromosome position and gene expression. J Cell Biol 176, 593-603.
177. Kurz, A., Lampel, S., Nickolenko, J.E., Bradl, J., Benner, A., Zirbel, R.M., Cremer, T., and Lichter, P. (1996). Active and inactive genes localize preferentially in the periphery of chromosome territories. J Cell Biol 135, 1195-1205.
178. Dietzel, S., Schiebel, K., Little, G., Edelmann, P., Rappold, G.A., Eils, R., Cremer, C., and Cremer, T. (1999). The 3D positioning of ANT2 and ANT3 genes within female X chromosome territories correlates with gene activity. Exp Cell Res 252, 363-375.
179. Volpi, E.V., Chevret, E., Jones, T., Vatcheva, R., Williamson, J., Beck, S., Campbell, R.D., Goldsworthy, M., Powis, S.H., Ragoussis, J., Trowsdale, J., and Sheer, D. (2000). Large-scale chromatin organization of the major histocompatibility complex and other regions of human chromosome 6 and its response to interferon in interphase nuclei. J Cell Sci 113 (Pt 9), 1565-1576.
180. Kim, S.H., McQueen, P.G., Lichtman, M.K., Shevach, E.M., Parada, L.A., and Misteli, T. (2004). Spatial genome organization during T-cell differentiation. Cytogenet Genome Res 105, 292-301.
181. Morey, C., Da Silva, N.R., Perry, P., and Bickmore, W.A. (2007). Nuclear reorganisation and chromatin decondensation are conserved, but distinct, mechanisms linked to Hox gene activation. Development 134, 909-919.
182. Mahy, N.L., Perry, P.E., and Bickmore, W.A. (2002). Gene density and transcription influence the localization of chromatin outside of chromosome territories detectable by FISH. J Cell Biol 159, 753-763.
183. Heard, E., and Bickmore, W. (2007). The ins and outs of gene regulation and chromosome territory organisation. Curr Opin Cell Biol 19, 311-316.
184. Shopland, L.S., Johnson, C.V., and Lawrence, J.B. (2002). Evidence that all SC-35 domains contain mRNAs and that transcripts can be structurally constrained within these domains. J Struct Biol 140, 131-139.
185. Xie, S.Q., Martin, S., Guillot, P.V., Bentley, D.L., and Pombo, A. (2006). Splicing speckles are not reservoirs of RNA polymerase II, but contain an inactive form, phosphorylated on serine2 residues of the C-terminal domain. Mol Biol Cell 17, 1723-1733.
186. Johnson, C., Primorac, D., McKinstry, M., McNeil, J., Rowe, D., and Lawrence, J.B. (2000). Tracking COL1A1 RNA in osteogenesis imperfecta. splice-defective transcripts initiate transport from the gene but are retained within the SC35 domain. J Cell Biol 150, 417-432.
187. Shopland, L.S., Johnson, C.V., Byron, M., McNeil, J., and Lawrence, J.B. (2003). Clustering of multiple specific genes and gene-rich R-bands around SC-35 domains: evidence for local euchromatic neighborhoods. J Cell Biol 162, 981-990.
188. Moen, P.T., Jr., Johnson, C.V., Byron, M., Shopland, L.S., de la Serna, I.L., Imbalzano, A.N., and Lawrence, J.B. (2004). Repositioning of muscle-specific genes relative to the periphery of SC-35 domains during skeletal myogenesis. Mol Biol Cell 15, 197-206.

189. Takizawa, T., Gudla, P.R., Guo, L., Lockett, S., and Misteli, T. (2008). Allele-specific nuclear positioning of the monoallelically expressed astrocyte marker GFAP. Genes Dev 22, 489-498.
190. Osborne, C.S., Chakalova, L., Brown, K.E., Carter, D., Horton, A., Debrand, E., Goyenechea, B., Mitchell, J.A., Lopes, S., Reik, W., and Fraser, P. (2004). Active genes dynamically colocalize to shared sites of ongoing transcription. Nat Genet 36, 1065-1071.
191. Osborne, C.S., Chakalova, L., Mitchell, J.A., Horton, A., Wood, A.L., Bolland, D.J., Corcoran, A.E., and Fraser, P. (2007). Myc dynamically and preferentially relocates to a transcription factory occupied by Igh. PLoS Biol 5, e192.
192. Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. Science 295, 1306-1311.
193. Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F., and de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. Mol Cell 10, 1453-1465.
194. Simonis, M., Kooren, J., and de Laat, W. (2007). An evaluation of 3C-based methods to capture DNA interactions. Nat Methods 4, 895-901.
195. Simonis, M., and de Laat, W. (2008). FISH-eyed and genome-wide views on the spatial organisation of gene expression. Biochim Biophys Acta.
196. Dekker, J. (2006). The three 'C' s of chromosome conformation capture: controls, controls, controls. Nat Methods 3, 17-21.
197. Palstra, R.J., Tolhuis, B., Splinter, E., Nijmeijer, R., Grosveld, F., and de Laat, W. (2003). The beta-globin nuclear compartment in development and erythroid differentiation. Nat Genet 35, 190-194.
198. Lee, G.R., Fields, P.E., Griffin, T.J., and Flavell, R.A. (2003). Regulation of the Th2 cytokine locus by a locus control region. Immunity 19, 145-153.
199. Spilianakis, C.G., and Flavell, R.A. (2004). Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. Nat Immunol 5, 1017-1027.
200. Zhou, G.L., Xin, L., Song, W., Di, L.J., Liu, G., Wu, X.S., Liu, D.P., and Liang, C.C. (2006). Active chromatin hub of the mouse alpha-globin locus forms in a transcription factory of clustered housekeeping genes. Mol Cell Biol 26, 5096-5105.
201. Vernimmen, D., De Gobbi, M., Sloane-Stanley, J.A., Wood, W.G., and Higgs, D.R. (2007). Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. Embo J 26, 2041-2051.
202. Majumder, P., Gomez, J.A., Chadwick, B.P., and Boss, J.M. (2008). The insulator factor CTCF controls MHC class II gene expression and is required for the formation of long-distance chromatin interactions. J Exp Med 205, 785 - 798.
203. Horike, S., Cai, S., Miyano, M., Cheng, J.F., and Kohwi-Shigematsu, T. (2005). Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. Nat Genet 37, 31-40.
204. Marenduzzo, D., Faro-Trindade, I., and Cook, P.R. (2007). What are the molecular ties that maintain genomic loops? Trends Genet 23, 126-133.
205. Drissen, R., Palstra, R.J., Gillemans, N., Splinter, E., Grosveld, F., Philipsen, S., and de Laat, W. (2004). The active spatial organization of the {beta}-globin locus requires the transcription factor EKLF. Genes Dev 18, 2485-2490.
206. Vakoc, C.R., Letting, D.L., Gheldof, N., Sawado, T., Bender, M.A., Groudine, M., Weiss, M.J., Dekker, J., and Blobel, G.A. (2005). Proximity among distant regulatory elements at the beta-Globin locus

requires GATA-1 and FOG-1. Mol Cell 17, 453-462.
207. Splinter, E., Heath, H., Kooren, J., Palstra, R.J., Klous, P., Grosveld, F., Galjart, N., and de Laat, W. (2006). CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. Genes Dev 20, 2349-2354.
208. Cai, S., Lee, C.C., and Kohwi-Shigematsu, T. (2006). SATB1 packages densely looped, transcriptionally active chromatin for coordinated expression of cytokine genes. Nat Genet 38, 1278-1288.
209. Bell, A.C., West, A.G., and Felsenfeld, G. (1999). The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. Cell 98, 387-396.
210. Kurukuti, S., Tiwari, V.K., Tavoosidana, G., Pugacheva, E., Murrell, A., Zhao, Z., Lobanenkov, V., Reik, W., and Ohlsson, R. (2006). CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. Proc Natl Acad Sci U S A 103, 10684-10689.
211. Kim, T.H., Abdullaev, Z.K., Smith, A.D., Ching, K.A., Loukinov, D.I., Green, R.D., Zhang, M.Q., Lobanenkov, V.V., and Ren, B. (2007). Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. Cell 128, 1231-1245.
212. Wendt, K.S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T., Yahata, K., Imamoto, F., Aburatani, H., Nakao, M., Imamoto, N., Maeshima, K., Shirahige, K., and Peters, J.M. (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature 451, 796-801.
213. Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H.C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T., Cobb, B.S., Yokomori, K., Dillon, N., Aragon, L., Fisher, A.G., and Merkenschlager, M. (2008). Cohesins functionally associate with CTCF on mammalian chromosome arms. Cell 132, 422-433.
214. Rubio, E.D., Reiss, D.J., Welcsh, P.L., Disteche, C.M., Filippova, G.N., Baliga, N.S., Aebersold, R., Ranish, J.A., and Krumm, A. (2008). CTCF physically links cohesin to chromatin. Proc Natl Acad Sci U S A 105, 8309-8314.
215. Gause, M., Schaaf, C.A., and Dorsett, D. (2008). Cohesin and CTCF: cooperating to control chromosome conformation? Bioessays 30, 715-718.
216. Xu, M., and Cook, P.R. (2008). Similar active genes cluster in specialized transcription factories. J Cell Biol 181, 615-623.
217. Chuang, C.H., and Belmont, A.S. (2005). Close encounters between active genes in the nucleus. Genome Biol 6, 237.
218. Brown, J.M., Green, J., das Neves, R.P., Wallace, H.A., Smith, A.J., Hughes, J., Gray, N., Taylor, S., Wood, W.G., Higgs, D.R., Iborra, F.J., and Buckle, V.J. (2008). Association between active genes occurs at nuclear speckles and is modulated by chromatin environment. J Cell Biol 182, 1083-1097.
219. Shopland, L.S., Lynch, C.R., Peterson, K.A., Thornton, K., Kepper, N., Hase, J., Stein, S., Vincent, S., Molloy, K.R., Kreth, G., Cremer, C., Bult, C.J., and O'Brien, T.P. (2006). Folding and organization of a contiguous chromosome region according to the gene distribution pattern in primary genomic sequence. J Cell Biol 174, 27-38.
220. Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., and de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nat Genet 38, 1348-1354.

221. Epner, E., Reik, A., Cimbora, D., Telling, A., Bender, M.A., Fiering, S., Enver, T., Martin, D.I., Kennedy, M., Keller, G., and Groudine, M. (1998). The beta-globin LCR is not necessary for an open chromatin structure or developmentally regulated transcription of the native mouse beta-globin locus. Mol Cell 2, 447-455.
222. Pirrotta, V. (1999). Transvection and chromosomal trans-interaction effects. Biochim Biophys Acta 1424, M1-8.
223. Wu, C.T., and Morris, J.R. (1999). Transvection and other homology effects. Curr Opin Genet Dev 9, 237-246.
224. Chandler, V.L., and Stam, M. (2004). Chromatin conversations: mechanisms and implications of paramutation. Nat Rev Genet 5, 532-544.
225. Hendrickson, J.E., and Sakonju, S. (1995). Cis and trans interactions between the iab regulatory regions and abdominal-A and abdominal-B in Drosophila melanogaster. Genetics 139, 835-848.
226. Chandler, V.L. (2007). Paramutation: from maize to mice. Cell 128, 641-645.
227. Patterson, G.I., Thorpe, C.J., and Chandler, V.L. (1993). Paramutation, an allelic interaction, is associated with a stable and heritable reduction of transcription of the maize b regulatory gene. Genetics 135, 881-894.
228. Stam, M., Belele, C., Dorweiler, J.E., and Chandler, V.L. (2002). Differential chromatin structure within a tandem array 100 kb upstream of the maize b1 locus is associated with paramutation. Genes Dev 16, 1906-1918.
229. Mueller-Storm, H.P., Sogo, J.M., and Schaffner, W. (1989). An enhancer stimulates transcription in trans when attached to the promoter via a protein bridge. Cell 58, 767-777.
230. Mahmoudi, T., Katsani, K.R., and Verrijzer, C.P. (2002). GAGA can mediate enhancer function in trans by linking two separate DNA molecules. Embo J 21, 1775-1781.
231. Hatada, I., Nabetani, A., Arai, Y., Ohishi, S., Suzuki, M., Miyabara, S., Nishimune, Y., and Mukai, T. (1997). Aberrant methylation of an imprinted gene U2af1-rs1(SP2) caused by its own transgene. J Biol Chem 272, 9120-9122.
232. Forne, T., Oswald, J., Dean, W., Saam, J.R., Bailleul, B., Dandolo, L., Tilghman, S.M., Walter, J., and Reik, W. (1997). Loss of the maternal H19 gene induces changes in Igf2 methylation in both cis and trans. Proc Natl Acad Sci U S A 94, 10243-10248.
233. Herman, H., Lu, M., Anggraini, M., Sikora, A., Chang, Y., Yoon, B.J., and Soloway, P.D. (2003). Trans allele methylation and paramutation-like effects in mice. Nat Genet 34, 199-202.
234. Liu, H., Huang, J., Wang, J., Jiang, S., Bailey, A.S., Goldman, D.C., Welcker, M., Bedell, V., Slovak, M.L., Clurman, B., Thayer, M., Fleming, W.H., and Epner, E. (2008). Transvection mediated by the translocated cyclin D1 locus in mantle cell lymphoma. J Exp Med 205, 1843-1858.
235. Rassoulzadegan, M., Magliano, M., and Cuzin, F. (2002). Transvection effects involving DNA methylation during meiosis in the mouse. Embo J 21, 440-450.
236. Rassoulzadegan, M., Grandjean, V., Gounon, P., Vincent, S., Gillot, I., and Cuzin, F. (2006). RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. Nature 441, 469-474.
237. Louwers, M., Haring, M., and Stam, M. (2005). When alleles meet: Paramutation. In Plant epigenetics, P. Meyer, ed. (Oxford, UK: Blackwell Publishers), pp. 134-173.
238. Bacher, C.P., Guggiari, M., Brors, B., Augui, S., Clerc, P., Avner, P., Eils, R., and Heard, E. (2006). Transient colocalization of X-inactivation centres accompanies the initiation of X inactivation. Nat

Cell Biol 8, 293-299.

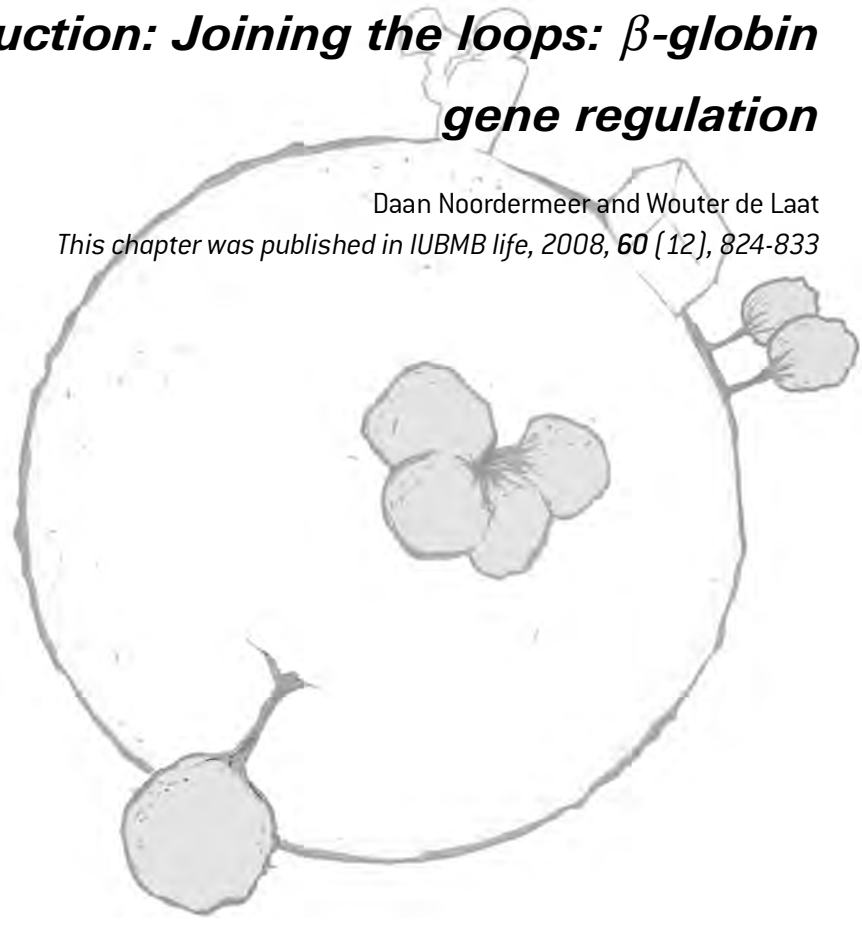
239. Xu, N., Tsai, C.L., and Lee, J.T. (2006). Transient homologous chromosome pairing marks the onset of X inactivation. Science 311, 1149-1152.
240. Penny, G.D., Kay, G.F., Sheardown, S.A., Rastan, S., and Brockdorff, N. (1996). Requirement for Xist in X chromosome inactivation. Nature 379, 131-137.
241. Monkhorst, K., Jonkers, I., Rentmeester, E., Grosveld, F., and Gribnau, J. (2008). X inactivation counting and choice is a stochastic process: evidence for involvement of an X-linked activator. Cell 132, 410-421.
242. Spilianakis, C.G., Lalioti, M.D., Town, T., Lee, G.R., and Flavell, R.A. (2005). Interchromosomal associations between alternatively expressed loci. Nature 453, 637-645.
243. Ling, J.Q., Li, T., Hu, J.F., Vu, T.H., Chen, H.L., Qiu, X.W., Cherry, A.M., and Hoffman, A.R. (2006). CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. Science 312, 269-272.
244. Zhao, Z., Tavoosidana, G., Sjolinder, M., Gondor, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Sandhu, K.S., Singh, U., Pant, V., Tiwari, V., Kurukuti, S., and Ohlsson, R. (2006). Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nat Genet 38, 1341-1347.
245. Apostolou, E., and Thanos, D. (2008). Virus infection induces NF-kappaB-dependent interchromosomal associations mediating monoallelic IFN-beta gene expression. Cell 134, 85-96.
246. Lomvardas, S., Barnea, G., Pisapia, D.J., Mendelsohn, M., Kirkland, J., and Axel, R. (2006). Interchromosomal interactions and olfactory receptor choice. Cell 126, 403-413.
247. Fuss, S.H., Omura, M., and Mombaerts, P. (2007). Local and cis effects of the H element on expression of odorant receptor genes in mouse. Cell 130, 373-384.
248. Nunez, E., Kwon, Y.S., Hutt, K.R., Hu, Q., Cardamone, M.D., Ohgi, K.A., Garcia-Bassets, I., Rose, D.W., Glass, C.K., Rosenfeld, M.G., and Fu, X.D. (2008). Nuclear receptor-enhanced transcription requires motor- and LSD1-dependent gene networking in interchromatin granules. Cell 132, 996-1010.
249. Nunez, E., Kwon, Y.S., Hutt, K.R., Hu, Q., Cardamone, M.D., Ohgi, K.A., Garcia-Bassets, I., Rose, D.W., Glass, C.K., Rosenfeld, M.G., and Fu, X.D. (2008). Nuclear receptor-enhanced transcription requires motor- and LSD1-dependent gene networking in interchromatin granules. Cell 134, 189.
250. Tiwari, V.K., Cope, L., McGarvey, K.M., Ohm, J.E., and Baylin, S.B. (2008). A novel 6C assay uncovers Polycomb-mediated higher order chromatin conformations. Genome Res 18, 1171-1179.

Chapter 2

Introduction: Joining the loops: β-globin gene regulation

Daan Noordermeer and Wouter de Laat

*This chapter was published in IUBMB life, 2008, **60** (12), 824-833*



2.1- Abstract

The mammalian β-globin locus is a multi-gene locus containing several globin genes and a number of regulatory elements. During development the expression of the genes changes in a process called “switching”. The most important regulatory element in the locus is the Locus Control Region (LCR) upstream of the globin genes, that is essential for high level expression of these genes. The discovery of the LCR initially raised the question how this element could exert its effect on the downstream globin genes. The question was solved by the finding that the LCR and activate globin genes are in physical contact, forming a chromatin structure named the Active Chromatin Hub (ACH). Here we discuss the significance of ACH formation, provide an overview of the proteins implicated in chromatin looping at the β-globin locus and evaluate the relationship between nuclear organization and β-globin gene expression.

2.2- The mammalian β-globin locus

The vertebrate hemoglobin gene loci have been intensively studied as model systems for developmentally regulated multi-gene loci. In mammals the α- and β- globin loci encode the proteins that form the heteromeric hemoglobin protein-complex involved in oxygen transport. Naturally occurring mutations in the loci show that co-regulation of both loci is required, since imbalance between the different proteins may lead to anemia. Interestingly, despite the need for tight co-regulation, the mammalian α- and β-locus are structurally very different and both loci are located in different genomic

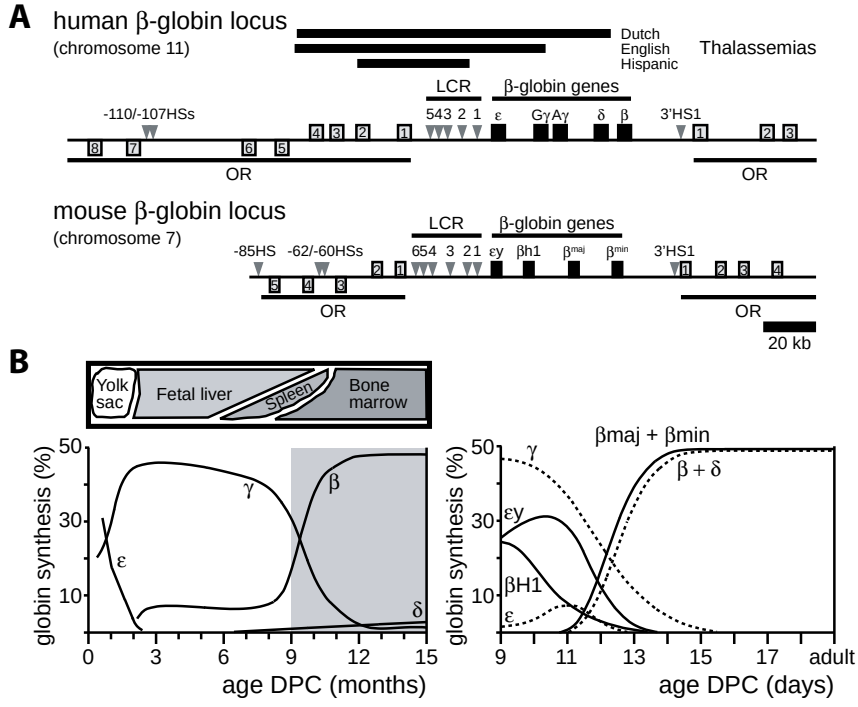


Figure 2-1. β-globin locus organisation and gene-expression. (A) The human and mouse β-globin locus. β-like genes are indicated by black boxes and hypersensitive sites by arrowheads. The genes in the surrounding olfactory receptor clusters (OR) are indicated by grey boxes and numbered according to their distance from the β-globin locus. Naturally occurring γδβ-Thalassemia deletions are indicated above the human locus. **(B)** β-like globin protein content during development in human *(left)* and mouse *(right)* as the percentage of the total globin content in the cells. Above the graph for the human locus the globin producing tissues at the stage of development is indicated. In the graph for the mouse locus both the globin content of the mouse globins (straight line) and human transgenically expressed globins (dashed line) are shown.

environments. The α-globin locus is located in a region containing mainly actively expressed housekeeping genes. In contrast, the β-globin locus is embedded in a large region of inactive olfactory receptor (OR) genes. During the later stages of erythroid differentiation, the genes in both loci are expressed at exceptionally high rates. This is necessary to fill the terminally differentiated erythrocyte with hemoglobin. How the β-globin genes achieve their extremely high expression rates, despite their location in a repressive chromatin environment has been the subject of intensive investigation.
The human and mouse β-globin loci are the most intensively studied mammalian globin loci. Like all other mammalian β-globin loci, they contain several globin genes, a large upstream regulatory element named the Locus Control Region (LCR) and a number of additional regulatory elements (**Figure 2-1A**). The genes are positioned on the chromosome in the order of their expression during development [1, 2]. Their expression is changed in a process called "switching". In the human locus, 5 expressed genes are present: ε, ^{G}γ, ^{A}γ, δ and β. The ε-gene is predominantly expressed in primitive erythroid cells in the embryo, both γ-genes are expressed in primitive cells and during the fetal stage in definitive cells and the β- and δ-genes are first activated in the fetal liver but mainly expressed perinatally (see **Figure 2-1B**). The mouse locus contains 4 highly expressed genes (εy, βh1, β^{maj} and β^{min}), of which the εy and βh1 genes are expressed during primitive erythropoiesis in embryonic tissues, and the β^{maj} and β^{min} genes are expressed during definitive erythropoiesis in the fetus and adult mouse (**Figure 2-1B**). Intriguingly, in embryonic tissue the βh1-gene initially appears more expressed than εy, suggesting that at defined stages of development gene expression may not strictly correlate with their order on the chromosome [3]. The LCR is the main regulatory element in the β-globin locus and is required for high levels of expression of all the genes. It was identified in patients that were suffering from anemia (γδβ-Thalassemia) due to β-chain imbalance, even though they were carrying a normal β-globin gene [4, 5]. Over the years, several other large deletions were identified that all comprised a region upstream of the ε-gene (see **Figure 2-1A**). In this region, a number of sites showed strong DNase I-hypersensitivity in erythroid cells [6]. The experiments by Grosveld *et. al.* [7], in which globin transgenes coupled to the LCR were found to be highly expressed, independent of the site of integration, showed the absolute importance of the LCR for high levels of globin gene expression and lead to a general definition of LCRs: elements that confer copy number dependent but position-of intergration independent expression to transgenes. In agreement with this, deletion of the mouse LCR leads to a 25 – 100 fold reduction of

globin gene expression [8]. The number of HSs that comprise the LCR varies between different species and is not always agreed on, but in both human and mouse the erythroid specific HS1 - HS4 exhibit the strongest enhancer function. The constitutive HS5 has a more structural role but has little effect on expression. Outside the LCR, both upstream and downstream of the globin genes, a number of additional HSs are found that seem to also have a structural role, comparable to HS5 of the LCR (**Figure 2-1A**).

2.3- Upregulation of the β-globin genes by the LCR

The discovery of enhancers and LCRs raised the question how a distant element could have such a large effect on the expression of its target genes. Many studies have been aimed at elucidating the mechanisms responsible for this long-range gene activation. Initially, a number of important discoveries were made that described the pattern of switching and demonstrated the relevance of gene order of the genes in the β-globin locus. In humans, the LCR upregulates only one gene at the time and the genes compete with each other for activation by the LCR [9]. In mice, the LCR seems to employ a similar mechanism, although nascent RNA signals from either εy and βh1 or β^{maj} and β^{min} are frequently detected simultaneously at one allele. A possible explanation for this could be LCR-mediated coregulation of the genes at the same allele, but after careful experiments the authors conclude that the most likely explanation is a rapid switching of the LCR between the genes on the same allele [10]. The main determinant for activation of a gene is the relative distance from the LCR. Introduction of a β-globin gene between the LCR and a γ-gene leads to premature activation of the β-gene and a strong reduction in expression of the γ-gene [11]. Introduction of a second β-gene in the locus causes the more proximal gene to be highest expressed, but the total output of the two β-genes together is not increased [2], similar to what has been found in patients that actively express the γ-globin genes in the adult stage [12]. In human adult tissues, the switch leading to the expression of the further downstream located β- and δ-genes is achieved by active silencing of the embryonic and fetal genes [13-15].

The observed competition between the β-like genes for upregulation by the LCR led to speculations that the LCR functioned by physically interacting with the genes. The simultaneous development of two new techniques, Chromosome Conformation Capture (3C) and RNA TRAP, allowed the determination of the spatial organization of the β-globin locus and showed that indeed the LCR and activated globin genes are in spatial proximity [16, 17]. The 3C technique proved very useful to study the changes during development

in the β-globin locus. Using 3C, the capture of *in vivo* interacting DNA fragments via formaldehyde crosslinking and subsequent ligation can be quantified by qPCR across ligation junctions. As a result this gives a measure for interaction frequencies of the fragments of interest [18]. If 3C experiments are well controlled they demonstrate the presence of chromatin loops [18-21]. It is important to realize though, that chromatin conformations as appreciated by 3C reflect steady-state averages measured across the population of cells. It may therefore well be that the structure of the β-globin locus at any given time is different from cell to cell [16].

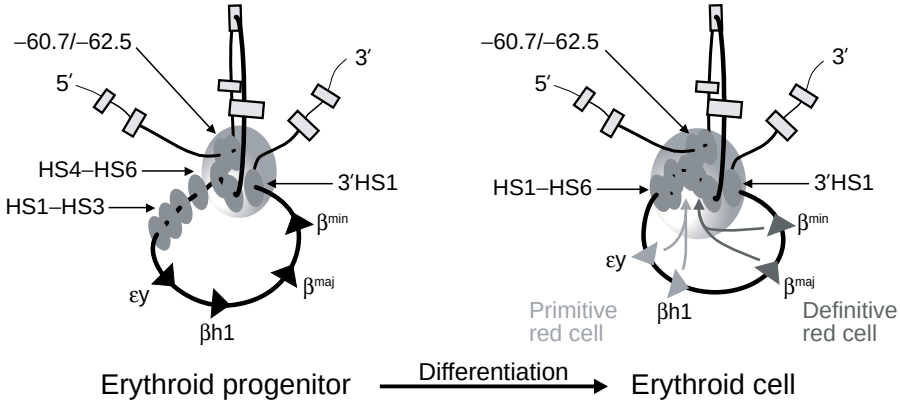


Figure 2-2. Chromatin organisation of the mouse β-globin locus. Structure of the mouse β-globin Chromatin Hub (CH, *left*) in erythroid progenitor cells and the Active Chromatin Hub (ACH, *right*) in β-globin expressing erythroid cells. Inactive β-globin genes are depicted by black arrowheads. Embryonic genes, participating in the ACH in embryonic tissues, are depicted by light grey arrowheads. Fetal and adult genes, participating in the ACH in fetal and adult tissues, are depicted by dark grey arrowheads. Hypersensitive sites are shown as grey ovals, olfactory receptor genes are depicted by grey boxes.

In globin expressing cells, depending on the tissue, the LCR co-localizes with the active genes, thereby forming a structure named the Active Chromatin Hub (ACH, see **Figure 2-2** for the mouse locus) [22]. In progenitor cells a substructure, the Chromatin Hub (CH), is present, consisting of the constitutive HS5 from the LCR and the upstream and downstream outer hypersensitive sites [22]. The appearance and disappearance of DNA interactions in the β-globin locus strikingly correlates with β-globin gene expression levels. Additional genetic experiments should reveal whether these loops are functional or not [21], but this does strongly suggest that ACH formation is important for the high expression levels of the β-globin genes.

Important questions now are how these chromatin loops are established and changed in the switching process, how frequent and dynamic these interactions are and which factors are involved. Furthermore it is interesting to consider whether the specific chromatin conformation of the β-globin locus may change the nuclear environment, thereby potentially facilitating the high expression of the genes. In the remainder of this review we will summarize the key players involved in chromatin looping at the β-globin locus and we will discuss the potential role of nuclear organization on the expression of the globin genes.

2.4- Factors involved in chromatin looping at the β-globin locus

Many proteins, both ubiquitously expressed and erythroid specific, are known to be involved in β-globin gene regulation and chromatin looping. Most of these factors bind the promoters of the globin genes and/or the HSs of the LCR and often are present in protein-complexes at these sites. Here we will discuss the current knowledge of factors involved in chromatin looping at the β-globin locus.

2.4.1- GATA-1 and FOG-1

The most intensively characterized regulator of globin gene expression is the erythroid specific GATA-1 transcription factor (previously also referred to as Gf-1 and NF-E1). GATA-1 can both activate and repress globin genes, but the study of this factor is complicated due to its involvement in activation and repression of many transcription factors involved in erythroid differentiation. In the mouse β-globin locus, GATA-1 binds the promoter of the β^{maj}-gene and HS1 - HS4 of the LCR [23-25]. Binding of GATA-1 to the promoter and HS2, but not to HS3 and HS4, depends on the GATA-1 interacting partner FOG-1 [23, 24]. The interaction with FOG-1 is also a determining factor for the involvement of GATA-1 in different protein-complexes that can have activating or repressing functions during development [26]. Therefore, FOG-1 interaction with GATA-1 may mediate recruitment of functionally different GATA-1 complexes with different regulatory functions to specific sites in the locus. Indeed site-specific epigenetic changes are observed in GATA-1 knockout cells expressing a GATA-1 mutant that is disrupted in its FOG-1 interaction. Acetylation of amino-acid residues in the tail of histone H3 are decreased at the β^{maj} promoter, where FOG-1 is required for GATA-1 binding. In contrast, no changes are observed at HS3 of the LCR, where GATA-1 binding is FOG-1 independent [23].

Next to targeting of regulatory protein-complexes, GATA-1 and FOG-1 binding is essential for loop formation between the mouse β^{maj} promoter and the LCR. 3C technology was applied to a cell line containing a GATA-1-fusion protein that is localized in the cytoplasm, but upon stimulation is rapidly recruited to the nucleus [27]. When GATA-1 is located in the cytoplasm, looping between the promoter of β^{maj} and HS2 and HS3 is nearly absent. Relocation of GATA-1 to the nucleus strongly increases looping between the promoter and the LCR. The direct involvement of GATA-1 in looping is further supported by the fact that loops are still being formed in the absence of protein synthesis, indicating that activation or repression of other genes by GATA-1 is not required for chromatin looping. Disruption of FOG-1 binding to the GATA-1-fusion protein strongly decreases chromatin looping between HS2 and the gene, which is accompanied by a severe decrease in β^{maj} expression [27].

Interestingly, GATA-1 has a role in gene-repression involving chromatin looping at the erythroid specific *Kit* locus. GATA-1 was found to be involved in chromatin loop switching, leading to silencing of the *Kit* locus in the cells previously used to examine GATA-1 inducible looping at the globin locus [27, 28]. When the *Kit* locus is active, a number of GATA-sites in the locus are occupied by GATA-2, a factor closely related to GATA-1. By applying 3C, a loop is observed between the 5'-end of the *Kit*-gene and a distant upstream enhancer. When GATA-1 is relocated into the nucleus, GATA-2 is replaced by GATA-1 at the *Kit* locus and *Kit* transcription is repressed. The replacement of GATA-2 by GATA-1 could be due to direct competition between the two factors, but more likely is caused by GATA-1-induced repression of the *GATA-2* locus. Switching of GATA factors in the locus causes a change in chromatin conformation of the locus. The initial loop between the start of the *Kit* gene and enhancer is abolished, and instead a new loop is formed between the 5'-end of the *Kit* gene and a region far more downstream in the gene [28]. Whether GATA-1 induced chromatin loop switches occur in the β-globin locus needs to be established, but it is an intriguing thought that GATA-1 may be a driving factor for globin switching. An interesting observation in this respect is the recently reported GATA-1 mediated silencing of the γ-genes in the human β-globin locus. GATA-1 binds a region upstream of both the Aγ- and Gγ-promoter in a FOG-1 dependent manner, leading to recruitment of a component of the repressive NuRD-complex [29]. Determination of whether binding of GATA-1 at these sites influences the structure of the ACH during development could potentially give further insight in how switching at the β-globin locus is achieved.

2.4.2- EKLF

A second well characterized erythroid specific factor involved in β-globin gene regulation and loop formation is the Krüppel-like zinc finger DNA binding protein ELKF (encoded by the *KLF1* gene). EKLF is not required for the expression of the embryonal ε- and fetal γ-genes, but strongly affects the expression of the adult β-globin gene [30]. In homozygous EKLF knockout mice containing a transgenic human β-globin locus, β-gene expression is absent and γ-globin expression is increased. In heterozygous EKLF mice, β-gene expression is delayed during differentiation, accompanied by increased γ-gene expression, but β-globin levels are unaffected in adult mice [30]. The exact mechanism of EKLF action in globin gene regulation is unclear due to several contradictory reports, most likely originating from different experimental designs. In mice, EKLF activates a human β-globin reporter construct containing HS3 of the LCR and is required for formation of this HS in the construct [31]. Similarly, HS3 is not formed at the full transgenic human locus in EKLF knockout mice [30]. Different results are obtained from the analysis of minichromosomes containing the β-globin gene coupled to HS2 and HS3 in human K562 genes. Here upregulation by HS3 is not dependent on the presence of EKLF, but rather HS2 function is EKLF dependent [32]. Interestingly, upregulation by HS3 in this study is dependent on the structural integrity of the EKLF binding sites at HS2. A construct containing HS3 coupled to HS2 without EKLF binding sites totally ablates HS3 enhancer function, suggesting an interaction between the HSs that is dependent on EKLF binding [32].

Application of 3C technology on the mouse β-globin locus in EKLF knockout mice revealed that interactions between the LCR and the β^{maj}-gene are lost when EKLF is not present [33]. To examine whether EKLF has a direct function in looping, 3C experiments were performed on EKLF knockout mice containing an EKLF-fusion protein that can be rapidly relocated to the nucleus. Recruitment of this EKLF-fusion protein to the nucleus in cells with blocked protein synthesis, showed reformation of the ACH and partially restored β^{maj} expression, indicating that EKLF is directly involved in chromatin looping between the LCR and the β^{maj} promoter [33].

Together, these results indicate that EKLF is an important factor in the switching process. EKLF allows the β-globin gene to compete efficiently with the γ-globin genes, possibly by initiating loops between the β-globin gene promoter and HS2 and HS3 of the LCR. The decreased levels of EKLF in heterozygous mice reduce the competing ability of the β-globin promoter, leading to upregulation of the γ-globin genes. Later during

development, the active silencing of the γ-globin genes relieves the β-globin gene from competition, allowing efficient loop formation despite the decreased EKLF levels, resulting in normal expression of the β-globin gene.

2.4.3- NF-E2

A third intensively studied factor involved in regulation of the globin genes is the NF-E2 heterodimer. The complex consists of two DNA binding subunits; the ubiquitously present MafK unit or p18 NF-E2 and its erythroid specific partner p45 NF-E2. NF-E2 most prominently binds HS2 of the LCR in mouse erythroid fetal liver tissue and in several mouse and human erythroleukemia cell lines, although a lower level of binding is also observed at the promoter of the adult β-genes. Depending on the cell system used, the necessity for NF-E2 in β-globin gene expression greatly varies [34]. In cell lines, p45 NF-E2 is required for β-globin gene expression, but in p45 NF-E2 knockout mice β^{maj} expression levels are only slightly reduced ([34, 35] and references therein). An explanation for this difference could be that in mice, in the absence of p45 NF-E2, redundant factors may associate with MafK, thereby compensating for the absence of this erythroid specific factor. Evidence for this hypothesis may be found in two recent reports showing very different requirements for the two components that make up NF-E2. In the first study, ACH formation was analyzed in p45 NF-E2 knockout mice [35]. In these mice, expression of the β^{maj} gene is slightly reduced (to a level similar as observed upon deletion of HS2 [36]) but the overall folding of the β-globin locus is not affected. Without p45 NF-E2, binding of the other NF-E2 subunit, MafK, is almost absent at the promoter and reduced to about 50% at HS2. MafK now seems associated with p45 NF-E2-related proteins like Nrf2 and to a lesser extent the repressing factor Bach1, which show strongly increased binding at HS2 [35]. These results are in contrast to a study on an established mouse erythroid leukemia cell line, DS19, in which MafK was knocked down using an siRNA approach [37]. MafK appears almost completely absent at the protein level, though binding of both MafK and p45 NF-E2 at HS2 and the promoter are only reduced to about one third of WT levels. Nevertheless, β^{maj} expression is about 80% lower, and this is reflected by a reduction of looping between HS2 and the β^{maj} promoter to levels close to non-globin expressing precursor cells. In contrast to the p45 NF-E2 knockout, in these cells no increase in Nrf2 binding is observed when MafK is knocked down [37]. The latter study may be interpreted to suggest that NF-E2 is involved in loop formation and that *in vivo* function of p45 NF-E2 may be compensated by other dimerization partners

like Nrf2. However, it should be noted that mice lacking both p45 NF-E2 and Nrf2, or p45 NF-E2 and Nrf3 show no erythroid phenotype beyond that seen with deletion of p45 NF-E2 alone [38-40], implying that compound knockout mice lacking all NF-E2 related factors need to be analyzed to unambiguously address this issue.

2.4.4- Other transcriptional regulators involved in looping

In addition to the previously discussed well characterized transcription factors, recently also a number of other proteins have been reported to have a potential role in chromatin looping at the β-globin locus. LDB1, a ubiquitous non-DNA binding protein, participates in many protein-complexes, which can both activate and repress the expression of target genes. Interestingly, the protein interacts with GATA-1 in murine erythroid cells [26, 41]. In a recent study, both LDB1 and GATA-1 were reported to be present at the promoter of the β^{maj} gene and HS1 – HS4 of the LCR in mouse fetal liver and MEL (Murine erythroid leukemia) cells and at HS1 – HS4 in human K562 cells [25]. Truncation of LDB1 reduces the expression of the ε-gene on a minichromosome containing HS2 in K562 cells, while a knockdown of LDB1 leads to decreased β^{maj} expression in differentiated MEL cells. In these knockdown cells, interaction between HS2 and the promoter of β^{maj} are reduced to background levels, showing that LDB1 is either directly or indirectly involved in long range chromatin interactions at the β-globin locus [25]. Since LDB1 is a non-DNA-binding protein, but interacts with many factors known to be involved in the regulation of erythroid genes, it may function by attaching several transcription factors to each other, thereby structuring the ACH. It would be interesting to know whether depletion of LDB1 also abrogates binding of its complexing partners GATA-1 and MafK to the β-globin locus.

Two other proteins have recently been reported to have a potential role in the switching process. SATB1, a protein known to bind specific AT-rich sequences in the genome, has been proposed to serve as spatial "genomic organizer" in certain cell types. In K562 cells, the protein binds MARs in the promoter of the ε-globin gene and HS2 of the LCR. Overexpression of SATB1 increases ε-gene expression, while simultaneously downregulating γ-globin [35]. Silencing of the ε-globin gene and activation of the γ-globin gene coincides with SATB1 repression, suggesting a potential relationship [42]. In T helper cells, SATB1 is involved in the formation of an intricate chromatin conformation at the T_H2 cytokine locus [43]. This raises the question whether SATB1 may change chromatin loops during the switching process.

The Ikaros protein is also thought to be involved in switching. Expression of a dominant negative version of Ikaros in primary human cells decreases the expression of both the fetal and adult globin genes, but more interestingly, strongly alters the ratio between γ- and β-gene expression in favor of the γ-globin genes [44]. In transgenic mice containing the human β-globin locus and expressing a DNA-binding impaired version of Ikaros, switching from γ- to β-gene expression is severely delayed, but β-globin gene expression is ultimately fully restored. Ikaros binds HS3 and to a lesser extent regions in the Gγ-gene and upstream of the δ-gene. In the mice containing the DNA-binding mutant of Ikaros, chromatin looping between a large fragment in the LCR and the β-gene is disturbed, and a new loop with the γ-globin genes is observed [45]. Therefore, like EKLF, Ikaros may have a role in switching from fetal to adult globin gene expression. Whether the binding of Ikaros at both the Gγ-gene and the δ-gene may be a key event in directing the switch in chromatin looping will be an interesting topic for future research.

2.4.5- CTCF and cohesin

CTCF (CCCTC-binding factor) and the cohesin complex are present at the β-globin locus and implied in chromatin looping, but don't seem to be involved in β-globin transcriptional regulation. CTCF is the most intensively characterized mammalian insulator protein. It blocks the effect of an enhancer on a target gene by binding in between [46, 47]. Enhancer blocking is proposed to be achieved by locating the gene and enhancer on different chromatin loops, which would preclude them from interacting with each other. CTCF binds HS5 of the LCR and the downstream 3'HS1 in the human and mouse β-globin loci. Additional binding sites have been reported at the upstream –62/-60 and –85 HSs in the mouse locus (see **Figure 1A**). All these HSs, except the –62/-60 sites and possibly the –85 site, can act as enhancer blockers [48, 49]. In non-globin expressing erythroid precursor cells, the CTCF-binding sites are in spatial proximity, thereby forming a chromatin structure called the chromatin hub (CH) ([22] and **Figure 2**). In primary mouse CTCF knockout cells, association between these sites is severely decreased, showing that CTCF is indeed involved in the shaping of the CH [50]. However, it is unclear what the function of these enhancer blocker sites in the β-globin locus is. It has been hypothesized that these sites either shield the LCR from activating the surrounding inactive OR-genes or that they block a potential signal from outside the locus to act on the globin genes. Deletion of the CTCF-binding site at 3'HS1 disturbs the participation of this site in the ACH. Surprisingly, looping between the LCR and the promoter of the β^{maj}

gene at later stages of erythroid differentiation is not affected and expression of the β^{maj} gene is not changed. The inactivity of the further downstream located OR-genes is also not altered [50]. Different studies supporting this finding showed that deletion of 3'HS1 and the upstream –62 HS does not affect εy and β^{maj} expression [51], that a human β-globin gene is upregulated when located upstream of the human LCR in transgenic mice [52] and that an integrated human β-globin LCR in an ectopic locus in mice upregulates genes upstream of HS5 from the LCR despite binding of CTCF [53]. Together, these results indicate that in erythroid progenitor cells CTCF has an important role in structuring the CH, but that the transition to the ACH and activation of the globin genes is not dependent on the presence of CTCF. Furthermore, CTCF-binding is not required to prevent deregulation of genes in and directly outside the β-globin locus, arguing against a role for CTCF in insulating the locus from its surrounding. The functional relevance of β-globin chromatin loop formation by CTCF early during erythropoiesis therefore remains enigmatic.

Recently a strong correlation between binding of CTCF and localization of cohesin to genomic loci was reported [54, 55]. During replication, the cohesin complex keeps the replicated sister chromatids together by forming a ring structure around the two strands. A role for cohesin in gene regulation during interphase has also been established, but until recently was not well understood. The colocalization with CTCF makes it attractive to speculate that CTCF recruits cohesin, and that cohesin in turn stabilizes interactions between CTCF binding sites, possibly by formation of rings around the chromatin strands. In HeLa cells, cohesin binds transgenic chicken HS4, the HS that is functionally related to HS5 in the mammalian β-globin locus. Depletion of both CTCF or cohesin in these cells results in impaired insulator function of this HS, indicating that cohesin may play a role together with CTCF in CH formation [54].

2.5- The Active Chromatin Hub and its nuclear environment

Gene expression and nuclear organization are tightly correlated mechanisms. Changes in the transcription state of a locus often coincide with changes in localization versus nuclear substructures. Several studies have reported nuclear relocation of the β-globin locus during erythroid differentiation, when the expression status of the locus changes from inactive or moderately active, to very highly active [56-58]. At the level of the β-globin locus this process coincides with the transition from CH to ACH [22]. An

interesting question is how these changes relate to each other.
When inactive, the β-globin locus is mainly located at the periphery of the nucleus and associates with repressive centromeric heterochromatin. When the locus is activated, it relocates to a position more to the interior of the nucleus and away from centromeric heterochromatin [58, 59]. Activation of the β-globin locus precedes the relocation to the nuclear interior, indicating that relocation is not a consequence of transcription at the locus. Rather, relocation may have a function in maintenance of transcriptional activity and may be driven by the association of the locus with nuclear entities known as transcription factories [58]. Deletion of the LCR strongly impairs this association with transcription factories and the relocation away from the nuclear periphery, but had previously been reported to have no effect on movement away from centromeric heterochromatin [58, 59].
A recent topic of interest is the location of the β-globin locus versus other genomic loci in the nucleus. In a number of studies the location of loci in the chromosome territory (CT) has been determined. Depending on the cell system used for examination, different results are obtained. In mouse MEL cells, where the globin genes are poised for transcription, the locus was reported to be significantly more often located outside the CT compared to non erythroid cells [56]. Furthermore, in this study MEL cells containing human chromosome 11 with or without the β-globin LCR were used to show that location of the β-globin locus outside the CT is strongly dependent on the presence of the LCR [56]. In contrast, another study showed that in mouse anemic spleen cells and human erythroblasts sorted at various differentiation stages the β-globin locus is almost exclusively located inside the CT [57]. To further explore the capacity of the LCR to relocate surrounding chromatin, we recently investigated an ectopically integrated human β-globin LCR integrated in a gene-dense region in the mouse genome. Localization outside the CT is strongly increased when the LCR is integrated in this region. The LCR upregulates many genes in the gene dense region, most likely by chromatin looping, but not all genes respond to the integration of the LCR and upregulation is dependent on the orientation of this element [53]. Together, these results show that looping out of the CT is not required for increased transcription, and that looping out does not automatically lead to increased transcription. It could be that the LCR locates the globin locus outside the CT to induce a poised state of the locus, as was suggested in the study of Ragoczy *et. al.* [56]. Whether this relocation is related to the spatial conformation of the locus will be an interesting topic for future studies.

The location of the β-globin locus versus other genomic loci has been characterized in two recent studies. In the first study, the location of the mouse β-globin locus was determined versus a selected number of other highly expressed loci on the same chromosome. These highly expressed genes are often in close spatial proximity when active. Colocalization was proposed to be at pre-assembled RNA polymerase II foci in the nucleus, where genes would need to migrate to for activation [60]. As two out the four loci that interacted with the β-globin locus also carried erythroid-specific genes, the data may be interpreted to suggest that functionally related genes preferentially come together in the nuclear space. In another study, genome-wide interacting partners of the β-globin locus were identified by using 4C, an unbiased micro-array-based method that allows screening the entire genome for regions that contact a locus of choice [61]. It was found that the active β-globin locus co-localizes with many regions containing active genes on the same chromosome, including those found in the study by Osborne *et. al.* (see reference [60]). In the context of all interacting loci, no preference was found for the β-globin locus to contact other erythroid-specific genes [61]. Interestingly, when 4C was applied to the inactive β-globin locus, it was found to associate with very different regions on the same chromosome which mainly contained inactive genes [61]. This suggests that the overall transcriptional status determines which chromosomal regions come together in the nucleus, with active chromatin separating from inactive chromatin. Whether clustering of the active β-globin locus with other active regions is only determined by transcription, or whether the LCR has a role in it is not known. Both studies have been repeated in the absence of transcription, and both studies reported no differences when transcription was blocked [62, 63]. Therefore it can be concluded that transcription itself is not necessary for the maintenance of these long-range associations, but this does not exclude that transcription is necessary for the initiation of these interactions.

Together these data show that in the process of development, and possibly also during erythroid differentiation, the β-globin locus is relocated to a very different nuclear environment. To understand the functional significance of the interactions formed by the active β-globin locus, it would be interesting to apply 4C technology to the locus with and without the LCR. Likewise, it would be interesting to apply 4C technology to the experimental system containing the ectopically integrated LCR and investigate the DNA contacts formed with and without the LCR. Question is: does the LCR search for erythroid-specific or functionally related genes elsewhere in the genome? If this were

true, it would strongly support the idea of dedicated transcription factories exclusively transcribing subsets of functionally related genes.

2.6- The Active Chromatin Hub, from steady state to dynamics

Since the elucidation of the structure of the ACH, considerable effort has been dedicated to determine the factors involved in structuring this 3D chromatin structure. In this review we have tried summarizing the current factors known or proposed to be involved in long-range chromatin interactions at the β-globin locus and how the structure of the

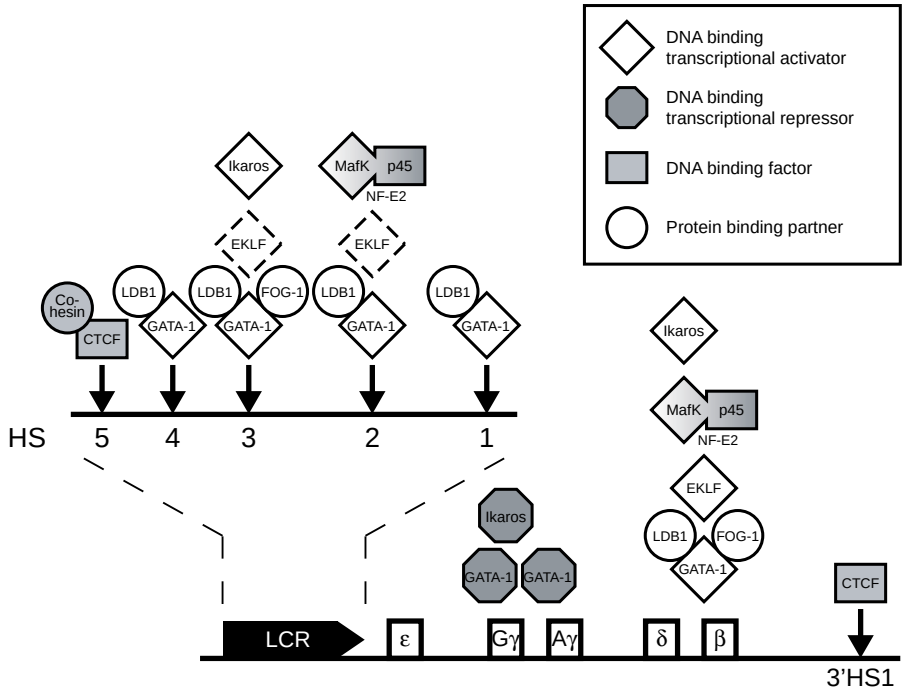


Figure 2-3. Transcription factor binding in the β-globin locus. Binding of factors known or proposed to be involved in chromatin looping to different sites in the β-globin locus. Binding sites for factors from studies on the adult human and mouse locus have been combined in this figure. Globin genes are depicted by boxes, hypersensitive sites are indicated by arrows. Reports on binding of EKLF to sites in the LCR are contradictory and therefore the factor is shown with a dashed line.

locus relates to its nuclear environment. In **Figure 2-3** a summary of factors present at sites in the β-globin locus is shown. Even though a considerable number of factors

have already been shown to be involved in looping at the globin locus, we have little doubt many more proteins play a role in chromatin looping at the locus. Probably the biggest challenge is to move from the current steady state studies that are based on cell populations to single cell studies and the visualization of changes in the locus during differentiation and the switching process. Indeed, it is to be expected that the ACH describes a structure that in each locus is dynamically formed and de-stabilized, with different HSs making and breaking contacts over time. At the molecular level it is important to investigate how these interactions affect the process of transcription initiation, re-initiation and possibly also elongation. At the level of nuclear organization the main challenge will be to determine where exactly the activated locus moves to and whether such movement is a prerequisite for proper expression of the β-globin genes or a mere consequence of gene activation. For this, the nuclear environment of the β-globin locus needs to be characterized at different stages of differentiation both in terms of protein content and in terms of the DNA loci present. 4C technology, in combination with immuno-FISH experiments will help addressing these issues.

2.7- Acknowledgements

We thank Robert-Jan Palstra for sharing pictures and Frank Grosveld and members of the lab for useful discussion. Wouter de Laat is funded by grants 016-006-026 and 912-04-082 from the Dutch Scientific Organization (NWO).

2.8- References

1. Grosveld, F., Dillon, N., and Higgs, D. (1993). The regulation of human globin gene expression. Baillieres Clin Haematol *6*, 31-55.
2. Dillon, N., Trimborn, T., Strouboulis, J., Fraser, P., and Grosveld, F. (1997). The effect of distance on long-range chromatin interactions. Mol Cell *1*, 131-139.
3. Kingsley, P.D., Malik, J., Emerson, R.L., Bushnell, T.P., McGrath, K.E., Bloedorn, L.A., Bulger, M., and Palis, J. (2006). "Maturational" globin switching in primary primitive erythroid cells. Blood *107*, 1665-1672.
4. Van der Ploeg, L.H., Konings, A., Oort, M., Roos, D., Bernini, L., and Flavell, R.A. (1980). gamma-beta-Thalassaemia studies showing that deletion of the gamma- and delta-genes influences beta-globin gene expression in man. Nature *283*, 637-642.
5. Kioussis, D., Vanin, E., deLange, T., Flavell, R.A., and Grosveld, F.G. (1983). Beta-globin gene inactivation by DNA translocation in gamma beta-thalassaemia. Nature *306*, 662-666.
6. Tuan, D., Solomon, W., Li, Q., and London, I.M. (1985). The "beta-like-globin" gene domain in human erythroid cells. Proc Natl Acad Sci U S A *82*, 6384-6388.
7. Grosveld, F., Blom van Assendelft, G., Greaves, D.R., and Kollias, G. (1987). Position-independent, high-level expression of the human beta-globin gene in transgenic mice. Cell *51*, 975-985.
8. Epner, E., Reik, A., Cimbora, D., Telling, A., Bender, M.A., Fiering, S., Enver, T., Martin, D.I., Kennedy, M., Keller, G., and Groudine, M. (1998). The beta-globin LCR is not necessary for an open chromatin structure or developmentally regulated transcription of the native mouse beta-globin locus. Mol Cell *2*, 447-455.
9. Wijgerde, M., Grosveld, F., and Fraser, P. (1995). Transcription complex stability and chromatin dynamics in vivo. Nature *377*, 209-213.
10. Trimborn, T., Gribnau, J., Grosveld, F., and Fraser, P. (1999). Mechanisms of developmental control of transcription in the murine alpha- and beta-globin loci. Genes Dev *13*, 112-124.
11. Hanscombe, O., Whyatt, D., Fraser, P., Yannoutsos, N., Greaves, D., Dillon, N., and Grosveld, F. (1991). Importance of globin gene order for correct developmental expression. Genes Dev *5*, 1387-1394.
12. Giglioni, B., Casini, C., Mantovani, R., Merli, S., Comi, P., Ottolenghi, S., Saglio, G., Camaschella, C., and Mazza, U. (1984). A molecular study of a family with Greek hereditary persistence of fetal hemoglobin and beta-thalassemia. Embo J *3*, 2641-2645.
13. Raich, N., Enver, T., Nakamoto, B., Josephson, B., Papayannopoulou, T., and Stamatoyannopoulos, G. (1990). Autonomous developmental control of human embryonic globin gene switching in transgenic mice. Science *250*, 1147-1149.
14. Dillon, N., and Grosveld, F. (1991). Human gamma-globin genes silenced independently of other genes in the beta-globin locus. Nature *350*, 252-254.
15. Sabatino, D.E., Cline, A.P., Gallagher, P.G., Garrett, L.J., Stamatoyannopoulos, G., Forget, B.G., and Bodine, D.M. (1998). Substitution of the human beta-spectrin promoter for the human agamma-globin promoter prevents silencing of a linked human beta-globin gene in transgenic mice. Mol Cell Biol *18*, 6634-6640.
16. Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F., and de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. Mol Cell *10*, 1453-1465.
17. Carter, D., Chakalova, L., Osborne, C.S., Dai, Y.F., and Fraser, P. (2002). Long-range chromatin

regulatory interactions in vivo. Nat Genet *32*, 623-626.
18. Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. Science *295*, 1306-1311.
19. Dekker, J. (2003). A closer look at long-range chromosomal interactions. Trends Biochem Sci *28*, 277-280.
20. Splinter, E., Grosveld, F., and de Laat, W. (2004). 3C technology: analyzing the spatial organization of genomic loci in vivo. Methods Enzymol *375*, 493-507.
21. Simonis, M., Kooren, J., and de Laat, W. (2007). An evaluation of 3C-based methods to capture DNA interactions. Nat Methods *4*, 895-901.
22. Palstra, R.J., Tolhuis, B., Splinter, E., Nijmeijer, R., Grosveld, F., and de Laat, W. (2003). The beta-globin nuclear compartment in development and erythroid differentiation. Nat Genet *35*, 190-194.
23. Letting, D.L., Chen, Y.Y., Rakowski, C., Reedy, S., and Blobel, G.A. (2004). Context-dependent regulation of GATA-1 by friend of GATA-1. Proc Natl Acad Sci U S A *101*, 476-481.
24. Im, H., Grass, J.A., Johnson, K.D., Kim, S.I., Boyer, M.E., Imbalzano, A.N., Bieker, J.J., and Bresnick, E.H. (2005). Chromatin domain activation via GATA-1 utilization of a small subset of dispersed GATA motifs within a broad chromosomal region. Proc Natl Acad Sci U S A *102*, 17065-17070.
25. Song, S.H., Hou, C., and Dean, A. (2007). A positive role for NLI/Ldb1 in long-range beta-globin locus control region function. Mol Cell *28*, 810-822.
26. Rodriguez, P., Bonte, E., Krijgsveld, J., Kolodziej, K.E., Guyot, B., Heck, A.J., Vyas, P., de Boer, E., Grosveld, F., and Strouboulis, J. (2005). GATA-1 forms distinct activating and repressive complexes in erythroid cells. Embo J *24*, 2354-2366.
27. Vakoc, C.R., Letting, D.L., Gheldof, N., Sawado, T., Bender, M.A., Groudine, M., Weiss, M.J., Dekker, J., and Blobel, G.A. (2005). Proximity among distant regulatory elements at the beta-Globin locus requires GATA-1 and FOG-1. Mol Cell *17*, 453-462.
28. Jing, H., Vakoc, C.R., Ying, L., Mandat, S., Wang, H., Zheng, X., and Blobel, G.A. (2008). Exchange of GATA factors mediates transitions in looped chromatin organization at a developmentally regulated gene locus. Mol Cell *29*, 232-242.
29. Harju-Baker, S., Costa, F.C., Fedosyuk, H., Neades, R., and Peterson, K.R. (2008). Silencing of Agamma-globin gene expression during adult definitive erythropoiesis mediated by GATA-1-FOG-1-Mi2 complex binding at the -566 GATA site. Mol Cell Biol *28*, 3101-3113.
30. Wijgerde, M., Gribnau, J., Trimborn, T., Nuez, B., Philipsen, S., Grosveld, F., and Fraser, P. (1996). The role of EKLF in human beta-globin gene competition. Genes Dev *10*, 2894-2902.
31. Gillemans, N., Tewari, R., Lindeboom, F., Rottier, R., de Wit, T., Wijgerde, M., Grosveld, F., and Philipsen, S. (1998). Altered DNA-binding specificity mutants of EKLF and Sp1 show that EKLF is an activator of the beta-globin locus control region in vivo. Genes Dev *12*, 2863-2873.
32. Jackson, D.A., McDowell, J.C., and Dean, A. (2003). Beta-globin locus control region HS2 and HS3 interact structurally and functionally. Nucleic Acids Res *31*, 1180-1190.
33. Drissen, R., Palstra, R.J., Gillemans, N., Splinter, E., Grosveld, F., Philipsen, S., and de Laat, W. (2004). The active spatial organization of the {beta}-globin locus requires the transcription factor EKLF. Genes Dev *18*, 2485-2490.
34. Forsberg, E.C., Downs, K.M., and Bresnick, E.H. (2000). Direct interaction of NF-E2 with hypersensitive site 2 of the beta-globin locus control region in living cells. Blood *96*, 334-339.

35. Kooren, J., Palstra, R.J., Klous, P., Splinter, E., von Lindern, M., Grosveld, F., and de Laat, W. (2007). Beta-globin active chromatin Hub formation in differentiating erythroid cells and in p45 NF-E2 knock-out mice. J Biol Chem *282*, 16544-16552.
36. Fiering, S., Epner, E., Robinson, K., Zhuang, Y., Telling, A., Hu, M., Martin, D.I., Enver, T., Ley, T.J., and Groudine, M. (1995). Targeted deletion of 5'HS2 of the murine beta-globin LCR reveals that it is not essential for proper regulation of the beta-globin locus. Genes Dev *9*, 2203-2213.
37. Du, M.J., Lv, X., Hao, D.L., Zhao, G.W., Wu, X.S., Wu, F., Liu, D.P., and Liang, C.C. (2008). MafK/NF-E2 p18 is required for beta-globin genes activation by mediating the proximity of LCR and active beta-globin genes in MEL cell line. Int J Biochem Cell Biol *40*, 1481-1493.
38. Martin, F., van Deursen, J.M., Shivdasani, R.A., Jackson, C.W., Troutman, A.G., and Ney, P.A. (1998). Erythroid maturation and globin gene expression in mice with combined deficiency of NF-E2 and nrf-2. Blood *91*, 3459-3466.
39. Kuroha, T., Takahashi, S., Komeno, T., Itoh, K., Nagasawa, T., and Yamamoto, M. (1998). Ablation of Nrf2 function does not increase the erythroid or megakaryocytic cell lineage dysfunction caused by p45 NF-E2 gene disruption. J Biochem *123*, 376-379.
40. Derjuga, A., Gourley, T.S., Holm, T.M., Heng, H.H., Shivdasani, R.A., Ahmed, R., Andrews, N.C., and Blank, V. (2004). Complexity of CNC transcription factors as revealed by gene targeting of the Nrf3 locus. Mol Cell Biol *24*, 3286-3294.
41. Meier, N., Krpic, S., Rodriguez, P., Strouboulis, J., Monti, M., Krijgsveld, J., Gering, M., Patient, R., Hostert, A., and Grosveld, F. (2006). Novel binding partners of Ldb1 are required for haematopoietic development. Development *133*, 4913-4923.
42. Wen, J., Huang, S., Rogers, H., Dickinson, L.A., Kohwi-Shigematsu, T., and Noguchi, C.T. (2005). SATB1 family protein expressed during early erythroid differentiation modifies globin gene expression. Blood *105*, 3330-3339.
43. Cai, S., Lee, C.C., and Kohwi-Shigematsu, T. (2006). SATB1 packages densely looped, transcriptionally active chromatin for coordinated expression of cytokine genes. Nat Genet *38*, 1278-1288.
44. Dijon, M., Bardin, F., Murati, A., Batoz, M., Chabannon, C., and Tonnelle, C. (2008). The role of Ikaros in human erythroid differentiation. Blood *111*, 1138-1146.
45. Keys, J.R., Tallack, M.R., Zhan, Y., Papathanasiou, P., Goodnow, C.C., Gaensler, K.M., Crossley, M., Dekker, J., and Perkins, A.C. (2008). A mechanism for Ikaros regulation of human globin gene switching. Br J Haematol *141*, 398-406.
46. Bell, A.C., West, A.G., and Felsenfeld, G. (1999). The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. Cell *98*, 387-396.
47. Wallace, J.A., and Felsenfeld, G. (2007). We gather together: insulators and genome organization. Curr Opin Genet Dev *17*, 400-407.
48. Farrell, C.M., West, A.G., and Felsenfeld, G. (2002). Conserved CTCF insulator elements flank the mouse and human beta-globin loci. Mol Cell Biol *22*, 3820-3831.
49. Bulger, M., Schubeler, D., Bender, M.A., Hamilton, J., Farrell, C.M., Hardison, R.C., and Groudine, M. (2003). A complex chromatin landscape revealed by patterns of nuclease sensitivity and histone modification within the mouse beta-globin locus. Mol Cell Biol *23*, 5234-5244.
50. Splinter, E., Heath, H., Kooren, J., Palstra, R.J., Klous, P., Grosveld, F., Galjart, N., and de Laat, W. (2006). CTCF mediates long-range chromatin looping and local histone modification in the beta-

globin locus. Genes Dev *20*, 2349-2354.

51. Bender, M.A., Byron, R., Ragoczy, T., Telling, A., Bulger, M., and Groudine, M. (2006). Flanking HS-62.5 and 3' HS1, and regions upstream of the LCR, are not required for beta-globin transcription. Blood *108*, 1395-1401.
52. Wai, A.W., Gillemans, N., Raguz-Bolognesi, S., Pruzina, S., Zafarana, G., Meijer, D., Philipsen, S., and Grosveld, F. (2003). HS5 of the human beta-globin locus control region: a developmental stage-specific border in erythroid cells. Embo J *22*, 4489-4500.
53. Noordermeer, D., Branco, M.R., Splinter, E., Klous, P., van Ijcken, W., Swagemakers, S., Koutsourakis, M., van der Spek, P., Pombo, A., and de Laat, W. (2008). Transcription and chromatin organization of a housekeeping gene cluster containing an integrated beta-globin locus control region. PLoS Genet *4*, e1000016.
54. Wendt, K.S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T., Yahata, K., Imamoto, F., Aburatani, H., Nakao, M., Imamoto, N., Maeshima, K., Shirahige, K., and Peters, J.M. (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature *451*, 796-801.
55. Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H.C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T., Cobb, B.S., Yokomori, K., Dillon, N., Aragon, L., Fisher, A.G., and Merkenschlager, M. (2008). Cohesins functionally associate with CTCF on mammalian chromosome arms. Cell *132*, 422-433.
56. Ragoczy, T., Telling, A., Sawado, T., Groudine, M., and Kosak, S.T. (2003). A genetic analysis of chromosome territory looping: diverse roles for distal regulatory elements. Chromosome Res *11*, 513-525.
57. Brown, J.M., Leach, J., Reittie, J.E., Atzberger, A., Lee-Prudhoe, J., Wood, W.G., Higgs, D.R., Iborra, F.J., and Buckle, V.J. (2006). Coregulated human globin genes are frequently in spatial proximity when active. J Cell Biol *172*, 177-187.
58. Ragoczy, T., Bender, M.A., Telling, A., Byron, R., and Groudine, M. (2006). The locus control region is required for association of the murine beta-globin locus with engaged transcription factories during erythroid maturation. Genes Dev *20*, 1447-1457.
59. Schubeler, D., Francastel, C., Cimbora, D.M., Reik, A., Martin, D.I., and Groudine, M. (2000). Nuclear localization and histone acetylation: a pathway for chromatin opening and transcriptional activation of the human beta-globin locus. Genes Dev *14*, 940-950.
60. Osborne, C.S., Chakalova, L., Brown, K.E., Carter, D., Horton, A., Debrand, E., Goyenechea, B., Mitchell, J.A., Lopes, S., Reik, W., and Fraser, P. (2004). Active genes dynamically colocalize to shared sites of ongoing transcription. Nat Genet *36*, 1065-1071.
61. Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., and de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nat Genet *38*, 1348-1354.
62. Mitchell, J.A., and Fraser, P. (2008). Transcription factories are nuclear subcompartments that remain in the absence of transcription. Genes Dev *22*, 20-25.
63. Palstra, R.J., Simonis, M., Klous, P., Brasset, E., Eijkelkamp, B., and de Laat, W. (2008). Maintenance of long-range DNA interactions after inhibition of ongoing RNA Polymerase II transcription. PLoS ONE *3*, e1661.

Chapter 3

Transcription and chromatin organization of the housekeeping gene cluster 8C3/C4

Daan Noordermeer*, Miguel R. Branco*[,1], Erik Splinter, Petra Klous, Wilfred van IJcken[2], Sigrid Swagemakers[3], Manousos Koutsourakis, Peter van der Spek[3], Ana Pombo[1] and Wouter de Laat

[*] These authors contributed equally to this work

[1] MRC Clinical Sciences Centre, Faculty of Medicine, Imperial College London, Hammersmith Hospital Campus, Du Cane Road, London, W12 0NN, United Kingdom

[2] Erasmus Center for Biomics, Erasmus Medical Center, PO Box 2040, 3000 CA Rotterdam, The Netherlands

[3] Erasmus Center for Bioinformatics, Erasmus Medical Center, PO Box 2040, 3000 CA Rotterdam, The Netherlands

*Data in this chapter was published as part of PLOS Genetics, 2008, **4**, e1000016*

3.1- Abstract

Knowledge on the regulation of gene transcription is strongly biased towards studies of genes showing atypical, highly restricted expression patterns and remarkably little is known about transcription regulation of ubiquitously expressed genes. Here, we analysed in detail gene expression and chromatin organization of 8C3/C4, a gene-dense region in the mouse genome containing many functionally unrelated housekeeping genes as well as tissue-specific genes. We show that expression and chromatin modifications within a gene cluster are importantly controlled at the level of individual genes, and that the locus frequently associates with transcription factories and is positioned at the periphery of its chromosome territory (CT). We propose that aspects of nuclear organization such as position in relation to the CT or association with factories may be necessary but are not sufficient for determining expression levels of single genes within a gene cluster.

3.2- Introduction

One of the main aims in the post-genomic era has been to understand how genes are regulated at the level of transcription, giving rise to cell-type specific transcriptomes. Most of our knowledge on the regulation of gene transcription is biased towards studies of a small number of atypical genes showing highly restricted expression patterns.

Remarkably little is known about the regulation of more ubiquitously expressed genes, which comprise the major part of the coding genome. High-throughput expression studies revealed that housekeeping genes often cluster in large gene-dense regions on mouse and human chromosomes [1, 2]. Breakpoints of synteny were shown to be under-represented in these regions, suggesting that this organization is under natural selection [3]. Gene expression seems to benefit from clustering along the linear genomic sequence. When genes are integrated at random positions in the genome, their expression is often subject to position effect variegation [4]. Furthermore, reporter genes express at higher levels when integrated in active gene-dense regions, indicating the existence of domain-wide regulatory mechanisms [5].

It has been suggested that gene clustering promotes the maintenance of transcriptionally competent domains of open or decondensed chromatin across the gene-dense region, with expression of individual genes being dictated by the availability of specific transcription factors [5, 6]. Another emerging hypothesis is that clustered genes may collectively stabilize their position at nuclear zones of increased transcriptional competence, which in turn would positively affect the expression levels of genes within the cluster [2, 7]. The location of genes relative to several nuclear landmarks has been correlated with gene expression. For example, activation of MHCII cluster genes, epidermal differentiation complex (EDC) and Hox genes promotes a large-scale relocation of the subchromosomal regions that contain them, away from the respective chromosome territory (CT) [8-10]. In general however, there seems to exist little correlation between gene activity and position versus the CT [11-15].

Proximity of genomic regions to nuclear structures rich in the RNA processing machinery has also been found to correlate with increased expression. Gene-rich R-bands are more frequently associated with 'splicing speckles' than their gene-poor counterparts, G-bands [16], and transcriptional activity of the globin genes also correlates with increased association with speckles [17]. Whilst it is clear that transcription, and therefore co-transcriptional splicing, can occur both away and near splicing speckles, it appears that locus association with these structures may facilitate mRNA processing,

especially in the case of intron-rich transcripts, such as COL1A1 [18]. Lastly, proximity to clusters of multiple active RNA polymerases (RNAP), known as transcription factories, may facilitate expression of genes that lie adjacent on the linear DNA template [19-22]. Live-cell imaging of GFP-tagged RNAP II shows that RNAP II complexes are extremely mobile and can access the whole nucleoplasm [23, 24]. When active, RNAP II is immobile and found in transcription factories that contain on average 8 active RNAP II complexes [20, 25, 26]. Irrespective of cell type, the size of factories as determined by electron microscopy is 50-70 nm [20, 21, 25]. Differences in transcriptional activity amongst different cell types correlate with the number of factories present per nucleus, rather than with changes in factory size [21].

While it is clear that gene expression is controlled at various levels, it is difficult to assess the hierarchy and importance of each level of regulation, as many of the observations are made on different gene loci and in different types of cells, often cultured in vitro. In this chapter we investigated the hierarchy and importance of each level of regulation by analysing in detail gene expression, chromatin structure and nuclear positioning of 8C3/C4, an extremely gene-dense region in the mouse genome. We found that the locus was frequently located at the periphery of its CT and close to splicing speckles and transcription factories, consistent with the active expression state of most genes in the cluster. However, the expression levels of individual genes measured in different tissues varied greatly and in a manner that was independent of expression changes of neighbouring genes, suggesting that expression levels are largely regulated gene autonomously. Promoters of active genes within the locus are hyperacetylated, but the intervening sequences are not, arguing against a generalized chromatin conformation pattern. We propose that aspects of nuclear organization such as position in relation to CT or association with factories may be necessary but are not sufficient for determining expression levels of single genes within a gene cluster.

3.3- Materials and methods

3.3.1- Affymetrix gene expression analysis

Total RNA was isolated using the RNeasy Mini kit (Qiagen) from livers and brains of three independent embryos and mice. Biotinylated cRNA was generated using the One-cycle Target Labeling and Control Reagents Kit (Affymetrix). All previous procedures and hybridisation, washing and scanning of the Affymetrix Mouse Genome 430 2.0 Arrays were done according to manufacturers' instructions. Array-data was normalized using Bioconductor RMA ca-tools. For each probe set, the values of the three independent micro-arrays were averaged. When multiple probe sets represented the same gene, the highest value was chosen to represent the gene.

3.3.2- qPCR gene expression analysis

Total RNA was isolated as described previously ("Affymetrix gene expression analysis"). cDNA synthesis was performed on 3 independent DNAseI treated RNA samples using SuperScript II Reverse Transcriptase and Random Hexamer primers according to the manufacturers' instructions (Invitrogen). Products were quantified by qPCR, using Platinum Taq DNA polymerase (Invitrogen) and SYBR Green (Sigma) on an Opticon 2 Real-Time PCR Detection System (BioRad). Primer sequences in **Table 3-1**. Transcript levels were normalized to the *Hprt1* transcript, encoding a relatively high expressed housekeeping gene on an unrelated chromosome.

3.3.3- Chromatin immunoprecipitation

ChIP was performed according to the Upstate protocol (http://www.upstate.com), with two modifications: (1) E14.5 fetal livers were made single cell by applying a cell-strainer cap (BD Falcon #352340, BD Biosciences) and (2) cells were fixed for 5 minutes in a 2% formaldehyde solution at room temperature. Chromatin fragments were quantified by qPCR (sequences of primers in **Table 3-1**) using Platinum Taq DNA polymerase (Invitrogen) and SYBR Green (Sigma) on an Opticon 2 Real-Time PCR Detection System (Biorad). Enrichments were calculated relative to the endogenous β-globin promoter or amylase promoter and values were normalized to input measurements. Antibodies used: anti acetyl-Histone H3 (#06-599, Upstate); anti K4 trimethyl H3 (#07-473, Upstate); anti C-terminal-Histone H3 (#ab1791; Abcam).

3.3.4- Cell preparation and cryosectioning

For the preparation of cell blocks for cryosectioning, E14.5 fetal liver and brain tissues were fixed in 4 and then 8% paraformaldehyde in 250 mM HEPES pH 7.6 (10 min and 2 h respectively) [27]. Cell pellets were embedded in 2.1 M sucrose in phosphate-buffered saline (PBS) and frozen in liquid nitrogen as described previously [20]. Cryosections (140-180 nm in thickness, deduced from interference colour) were cut using an UltraCut UCT52 ultracryomicrotome (Leica), captured in sucrose drops, and transferred to glass coverslips.

Expression analysis

Primer set	Sequence
mRNA Hprt	AGCCTAAGATGAGCGCAAGT ATGGCCACAGGACTAGAACA
Primary transcripts CalR	CATAGAATGGAGGACATCTGG GTTCCCACTCTCCATCCA
Primary transcripts Nfix	CAGCCACATCACATTGGAG CTGAACAAATACCAGCAACTG
Primary transcripts Rad23a	GGTGTCTTGGTGTGTTAGTG GCTTCTATCTTCTCCTTCAGC
Primary transcripts Klf1	CAGTGCCTACCATTCAAGC AAGGGTCCTCCGATTTCAG
Primary transcripts Prdx2	TTTCCTGTCTCTACCCGTG ATAGAGGTCGTGATGAGGC

Chromatin immunoprecipitation

Primer set	Sequence
Promoter Lyl1	TTCAAAGCTAGACCAACCTCA CACAGCACTGGAAGACCC
Promoter Nfix	TCCAAACCACACTTCAGTAG GAAGGAGAAACACAGCGTT
Intergenic region Nfix - Dand5	ACAGAGACAGCCGAATACC GACCTCCCTCTGTCTGTTC
Promoter Dand5	AACTCTCAAGCTGCTCTCC CTGTGCAGTCGTTTGTCTG
Promoter Gadd45gip1	TCGGAGGGTAAAGGCATT AGTGTTGAAGTGTTGGTGAT
Promoter Rad23a	CACCAAGACAGCGGAATG GGCTGCACCTTACCTTAGA
Promoter CalR	CTGGGAAGCAATGGAAAG TTATATTCACCTACCTCTCACCC
Intergenic region CalR - Farsla	AGCACCTTCTGACTTCCAA GGAGAGATGGCTGAGAGTAA
Promoter Farsla	AACTAAAAGCCACTGGGGT TAAGTGTGGCAAATGAGCC
Promoter Syce2	CGCACCTACGCATTATGA TGCCTTTTGGGCTATGCT
3' Syce2, 3' Gcdh	ATTGCTCTCCCAAGGATCA TGAGCTGAAGATTCCAAACC
Promoter Gcdh	GGAACCATAACCTGGAAGGG AAGGAGGAACCAATGAGCAA
Promoter Klf1	CTTTGCCTGGGTCTTATCA TCCTCTCTCTCTTCTGAATC
Promoter Dnase2a	GGGTCACGAGATTCAGATG GCTTCGTCTCCACCCTCG
Intergenic region Mast1 - Rtbdn	ATGCTCAGCCAGTAGTAGTT TTTACTCTGTGGGTCCTGG
Promoter Prdx2	ATGCCCGGATTCCAACCG TCCACACGCTTTCACAAG
Promoter Amylase	CTCCTTGTACGGGTTGGT AATGATGTGCACAGCTGAA
Promoter Beta Major	GGGAGAAATATGCTTGTCATC CAACTGATCCTACCTCACCTT

Table 3-1 Primers

3.3.5- Cryo-FISH

Cryo-FISH was performed as described previously [15]. A probe for the 8C3/C4 locus was obtained by labeling a BAC (RP24-319P23) with biotin or rhodamine by nick-translation (Roche). The BAC probe was co-precipitated with mouse Cot1 DNA (Roche; 1.7 μg/μl final concentration) and resuspended in either hybridisation buffer (50% deionized formamide, 10% dextran sulfate, 2xSSC, 50 mM phosphate buffer pH 7.0) or a FITC-labelled mouse whole chromosome 8 paint (Applied Spectral Imaging). Probes were denatured at 70°C for 10 min, and re-annealed at 37°C for 30 min before hybridisation. Probe specificity was confirmed on mouse spleen metaphase spreads.

3.3.6- Immunolabeling

Immunolabeling of cryosections was performed as described previously [15]. The biotin-labeled BAC probe for the 8C3/C4 locus was detected using rhodamine-conjugated neutravidin (1/500; Molecular Probes), followed by a biotin-conjugated goat anti-avidin antibody (1/500; Vector) and rhodamine-conjugated neutravidin. Splicing speckles were detected with a human autoimmune serum against Sm antigen (1/2000; ANA-CDC), followed by a biotin-conjugated donkey anti-human antibody (1/100; Jackson ImmunoResearch) and an AlexaFluor488-conjugated neutravidin (1/100; Molecular Probes). Serine 2 phosphorylated RNAP II was immunolabeled with H5 (1/1000; Covance), followed by an IgM-specific biotin-conjugated donkey anti-mouse antibody (1/250; Jackson ImmunoResearch) and Alexa 488-conjugated neutravidin or Alexa 647-conjugated streptavidin (1/100; Molecular Probes). After immunolabeling and before cryo-FISH, antibodies were fixed with 8% paraformaldehyde in 250 mM HEPES pH 7.6 (1 h), or with 2 mM EGS in PBS (30 min, 37°C).

3.3.7- Microscopy and image processing

For confocal laser scanning microscopy, images were collected sequentially on a Leica TCS SP2 (100X PL APO 1.40 Oil objective) equipped with Argon (488 nm) and HeNe (543 nm; 633 nm) lasers. For wide-field light microscopy, images were collected sequentially on a Delta-Vision Spectris system (Applied Precision) equipped with an Olympus IX70 widefield microscope (100X UPlanFI 1.3 Oil objective), a charge-coupled device camera, and the following filters: DAPI, FITC, RD-TR-PE, CY-5, CFP, YFP. The use of ultrathin cryosections allows for the use of wide-field microscopy with no reduction in axial (*z*) resolution and only a small reduction in lateral resolution. No bleedthrough was detected in these conditions, and images were collected without saturation of intensities. The images presented were contrast stretched, without further thresholding or filtering. For the generation of randomly positioned loci, the experimental images collected for the analysis of 8C3/C4 position relative to transcription factories were used. We developed a macro on ImageJ (Wayne Rasband, NIH, USA) that thresholds the DAPI signal and, for each locus, generates random coordinates until they fall within the mask of the corresponding thresholded nuclear section. These coordinates were used to measure the distance to the nearest transcription factory.

3.4- Results

3.4.1- 8C3/C4: a gene-dense region on mouse chromosome 8

To study the expression, chromatin organization and nuclear position of a gene dense region, we focused on a 300 kb genomic region that contains 18 genes. This extremely gene dense area is located within a 3 Mb gene cluster with 79 genes that resides in the middle of mouse chromosome 8 at the transition of band C3 and C4 (NCBI assembly m34) (**Figure 3-1A, B**). The syntenic region in human was previously classified as a RIDGE (Region of Increased Gene Expression), indicating that the majority of its genes

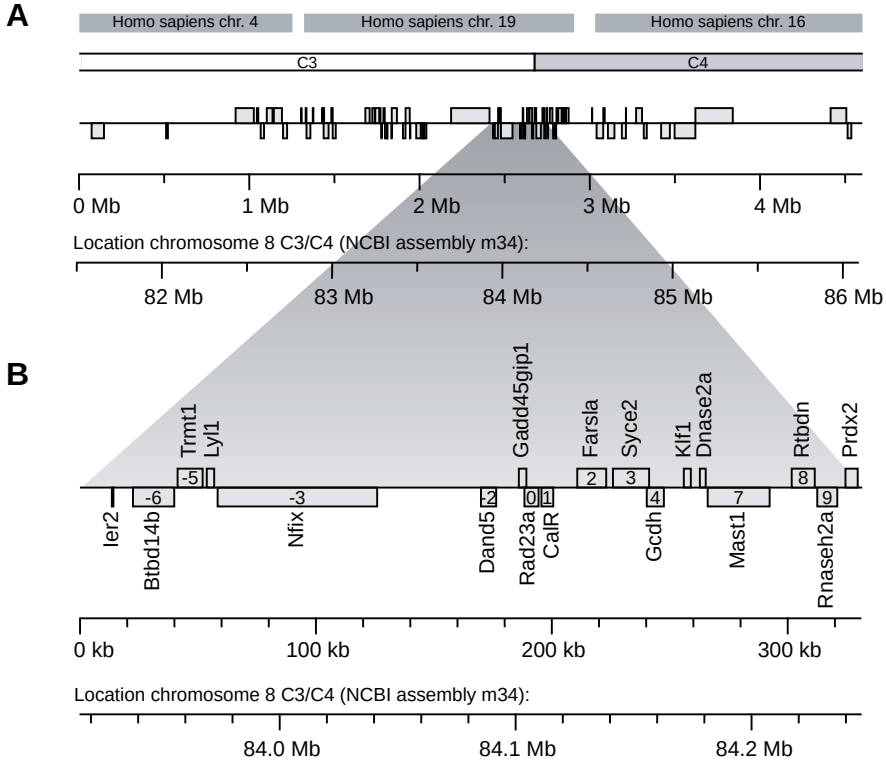


Figure 3-1. Murine gene dense region 8C3/C4. (A) 3 Mb region containing 79 genes (NCBI assembly m34). Location on murine chromosome 8 is depicted below the region. Human synteny and cytogenic chromosome bands are depicted on top. **(B)** Zoom in on 300 kb region containing 18 genes characterized in this study (NCBI assembly m34). Genes are numbered relative to the center gene *Rad23a* (gene 0). Location on murine chromosome 8 is depicted below the region.

are expressed in a large number of tissues and therefore are considered housekeeping genes [1]. The genes do not share sequence homology and encode for proteins that function in very diverse cellular processes, suggesting that gene clustering at this genomic location is not the result of sequence duplication during evolution. The 300 kb gene-dense region is hereafter referred to as 8C3/C4.

3.4.2- Gene expression at 8C3/C4

To gain insight in the behaviour of the genes in this gene-dense region, we analysed the expression of the individual genes clustered at 8C3/C4 by comparing their transcription profile in 61 different tissues. For this, we used a publicly available gene expression database (http://symatlas.gnf.org/SymAtlas/) that provides Affymetrix micro-array

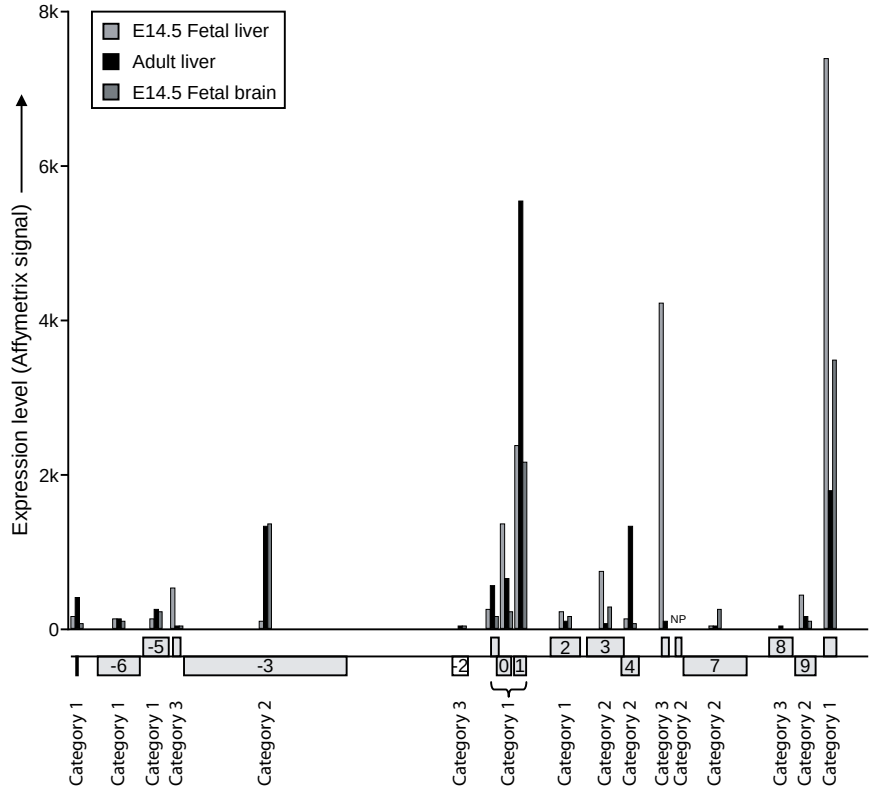


Figure 3-2. Expression levels in different tissues and developmental stages. Expression levels in fetal liver, fetal brain and adult liver of genes at 8C3/C4 represented on murine Affymetrix 430 2.0 Micro-array. Below each gene the expression category as discussed on page 102 is indicated. NP: not present on the micro-array.

measurements of steady-state mRNA levels in many tissues [28]. In addition, we performed a series of Affymetrix expression array experiments ourselves in order to understand expression in tissues relevant for this particular study, such as E14.5 fetal liver and brain (**Figure 3-2**). Three categories of genes were discerned: (1) genes that were expressed in most tissues and at levels comparable between the tissues, (2) genes that were expressed in most tissues but at levels that differed significantly between the tissues, and (3) tissue-specific genes that were expressed in only one

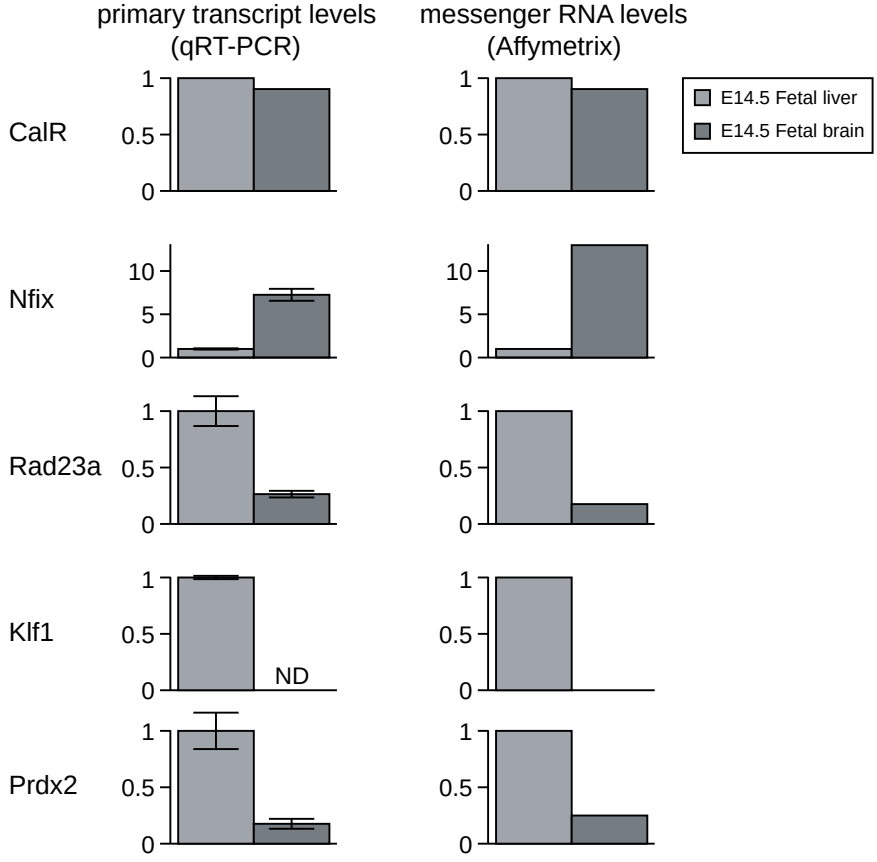


Figure 3-3. Primary transcript levels of genes at 8C3/C4. qRT-PCR analysis of primary transcript levels confirmed that the different mRNA levels between tissues, as measured by murine Affymetrix 430 2.0 Micro-array, represent differences in transcription efficiency of a given gene. Primary transcript levels were normalized to *CalR* primary transcript levels, taking into account the difference in *CalR* expression as measured by microarray analysis between fetal liver and brain. Error bars depict standard deviations. ND: not detected.

or a few tissues. Genes belonging to each category were found in this 300 kb region and appeared along the chromosome template in a seemingly random order. Eight out of the eighteen genes were expressed at similar levels in all tissues (category 1). Four genes were tissue-specific (category 3) and active in a different cell-type. Six genes belonged to category 2, each having its own favourite tissue for high expression levels (**Figure 3-2**). While it is clear that the absolute mRNA levels measured by micro-arrays are also influenced by post-transcriptional events and micro-array hybridisation efficiencies, analysis of primary transcript levels by quantitative reverse transcriptase PCR (qRT-PCR) indicated that the measured differences in messenger RNA levels of a given gene between tissues reflect mostly a difference in the transcriptional activity of that gene (**Figure 3-3**). Taken together, the expression data show that knowledge of expression breadth and level of a given gene has little, if any, predictive value for the transcription of nearby genes. For example, *Klf1* (gene -5) which is highly expressed in E14.5 liver and lowly in adult liver lies only 5 kb from *Gcdh* (gene -4) with an opposite expression pattern in these tissues. Similarly, *Syce2* (gene -3) and *Gcdh*, two neighbouring genes that have partially overlapping coding regions on the chromosome, showed very distinct levels of expression in different tissues. This apparent lack of co-regulation between the closely juxtaposed genes is in agreement with genome-wide analyses showing that expression of individual genes does not correlate over distances of more than two genes [29]. Thus, while clustering of genes will be beneficial to maintain active states and ensure pan-cellular expression, mechanisms working at the level of single genes determine the expression levels in the various tissues.

3.4.3- Histone H3 acetylation and lysine 4 trimethylation at 8C3/C4

Active chromatin is associated with specific histone modifications like acetylation of lysine residues 9, 14, 18 and 23 and di- and trimethylation of lysine residue 4 in the N-terminal tail of histone H3 [30, 31]. First, chromatin immunoprecipitation (ChIP) experiments were performed on E14.5 fetal liver and fetal brain to analyse whether histone H3 acetylation is homogeneously distributed across 8C3/C4 or whether it is more restricted to the individual genes. A region-wide analysis in both tissues revealed enrichment of hyperacetylated histone H3 at the promoters of the actively transcribed genes, but not at the inactive promoters or intergenic areas present in the region (**Figure 3-4A**). Such punctuated, rather than domain-wide, pattern of hyperacetylation was previously observed at gene-rich regions in a genome-wide mapping study [32]. Tissue-

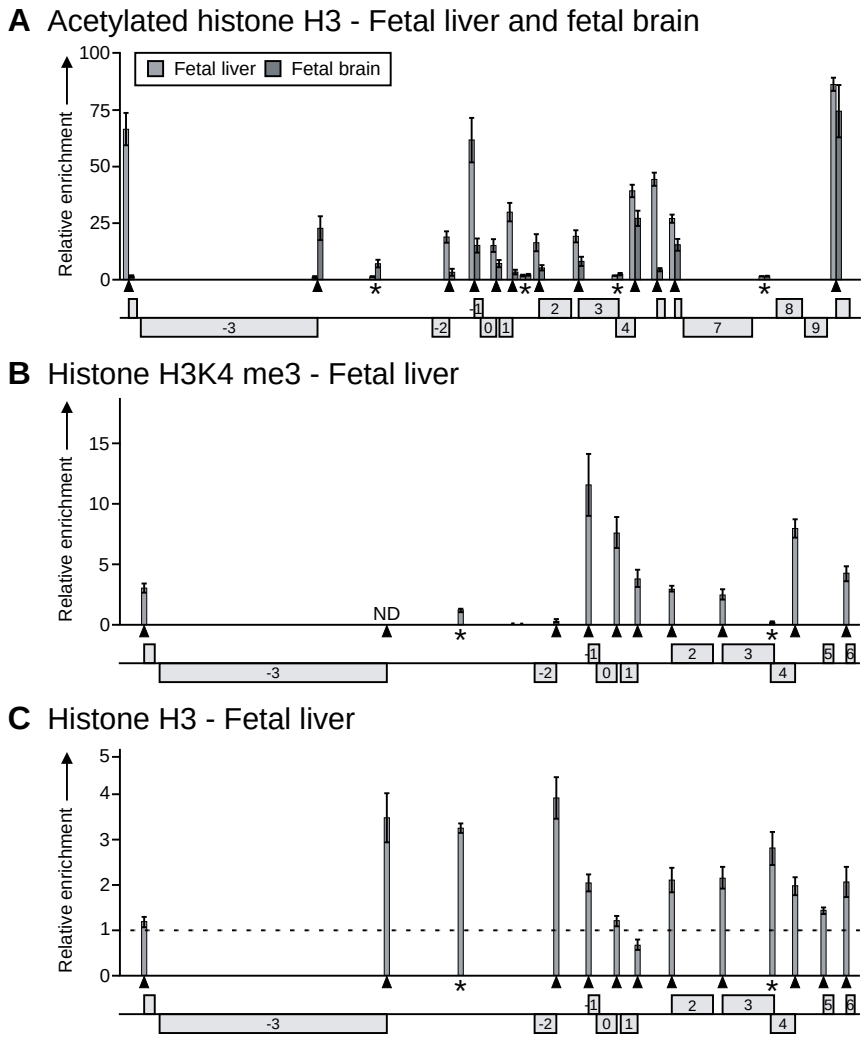


Figure 3-4. Chromatin organization at 8C3/C4. (A) Histone H3 acetylation at promoters (▲) and other regions (*) of 8C3/C4 in E14.5 fetal liver and brain as determined by ChIP-analysis. Enrichments were normalized to the Amylase promoter. **(B)** Histone H3 lysine 4 tri-methylation at 8C3/C4 in E14.5 fetal liver. Enrichments were normalized to the endogenous β-major promoter. **(C)** Histone H3 occupancy at 8C3/C4 in E14.5 fetal liver. Enrichments were normalized to the endogenous β-major promoter. Error bars in all graphs represent standard error (SE) of at least two independent experiments.

specific genes such as *Nfix* (-3) and *Klf1* (+5) switch promoter acetylation status in relation to their expression status in each tissue. Similar observations were made when we analysed H3K4 trimethylation across 8C3/C4 in E14.5 fetal liver (**Figure 3-4B**). To analyse whether our data were influenced by differences in histone H3 occupancy at these sites, we performed ChIP with an antibody against the C-terminus of histone H3 in E14.5 fetal liver (**Figure 3-4C**). Considerable differences in H3 enrichment were observed in a pattern generally opposite to that found with the antibody against acetylated and trimethylated H3. Thus, H3 was most abundantly present at intergenic regions and at inactive promoters and relatively depleted from active promoters, as seen before [33]. The punctuated pattern of H3 depletion, H3 acetylation and H3K4 trimethylation at 8C3/C4 reflects the expression status of the individual genes present in this region, suggesting that tissue-specific genes in gene-dense regions modulate these chromatin characteristics in a manner that is mainly independent of surrounding genes.

3.4.4- Nuclear positioning of 8C3/C4

We next investigated the nuclear position of the 8C3/C4 gene cluster in relation to other nuclear landmarks, including its own CT, splicing speckles and transcription factories. We first analysed the position of 8C3/C4 (using a BAC probe) relative to its own CT labeled with a whole chromosome 8 probe in E14.5 fetal liver. For this, we performed cryo-FISH (fluorescence in situ hybridization on thin cryosections), a novel technology that offers increased resolution in the z-axis (150-200 nm) compared to standard 2D- and 3D-FISH protocols (>500 nm) [15]. Location was scored in four categories: (1) inside the CT, (2) at the inner edge of the CT, (3) at the outer edge of the CT, and (4) looped out from the CT (see [21]). As observed previously for other actively transcribed gene clusters, 8C3/C4 was found mostly at the (inner and outer) edge (82%) or looped out (14%) from its CT (**Figure 3-5A**; n=232 loci).

We then performed a cryo-immuno-FISH experiment to determine association of 8C3/C4 with splicing speckles, labeled with the Sm antigen, a spliceosomal factor [34]. We found 8C3/C4 often adjacent to (63%) or overlapping with (21%) splicing speckles (**Figure 3-5B**), as would be expected for a region containing many active genes [16]. As splicing speckles are located outside CTs, the preferred association of 8C3/C4 with these domains is consistent with the favoured position of the locus towards the periphery of its CT.

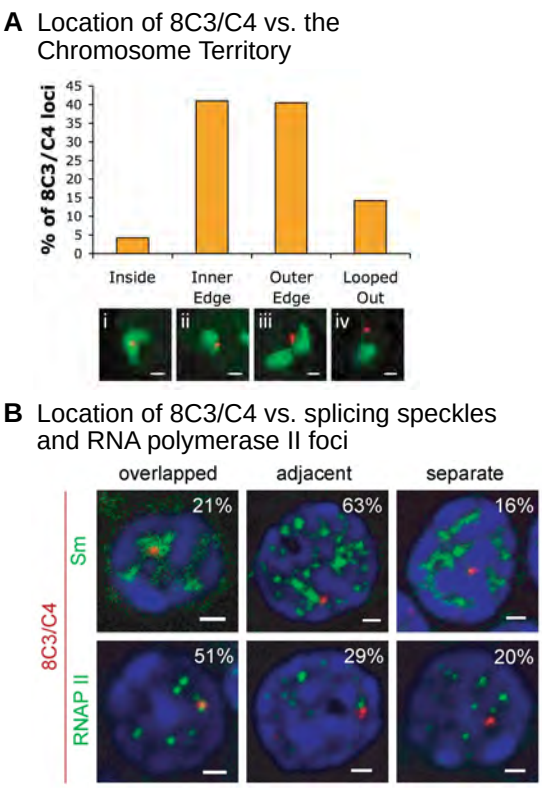


Figure 3-5. Nuclear positioning of 8C3/C4 in mouse fetal liver. (A) 8C3/C4 is often positioned at the edge of its own CT. Cryosections of fetal liver nuclei were hybridized to a BAC probe containing the 8C3/C4 locus (bottom, red) and a chromosome 8 specific paint (green); the distances from the centre of each locus to the nearest CT edge were measured (n=240 loci). **(B)** 8C3/C4 is frequently associated with splicing speckles and foci of active RNAP II. Nuclear cryosections (pseudocolored blue) were immunolabeled with antibodies against the Sm antigen (top, green) or the Ser2-phosphorylated form of RNAP II (bottom, green), hybridized to the 8C3/C4 BAC probe (red), and the locus position scored in relation to either nuclear structure (n=56 and 231 loci for Sm and RNAP II, respectively). Adjacent refers to the red and green signals touching each other.

Bars, 1 μm.

We next investigated the association of 8C3/C4 with the active form (serine2-phosphorylated) of RNAP II, which marks transcription factories as determined after co-localization with sites of Br-UTP incorporation [20, 35]. Using cryo-immuno-FISH to simultaneously visualize the elongating isoform of RNAP II and 8C3/C4 (**Figure 3-5B**), we found that 51% of 8C3/C4 alleles overlapped with RNAP II foci and 29% were adjacent to these ('adjacent' loci were $0.46 \pm 0.09\,\mu$m away from the centre of a RNAP II focus, whilst 'overlapped' loci were $0.23 \pm 0.10\,\mu$m away). Although a certain degree of spurious, non-functional association of 8C3/C4 with RNAP II foci is expected, this frequent association with the transcription machinery is in agreement with the high density of actively transcribed genes at 8C3/C4.

We then wondered whether the association of 8C3/C4 with RNAP II foci was dependent on its position relative to its CT, i.e. whether it preferentially occurred for 8C3/C4 loci found more externally in relation to their CT. To test this hypothesis, we simultaneously

visualized 8C3/C4, the territory of chromosome 8 and the active form of RNAP II in cryosections (**Figure 3-6**). Percentages of association with RNAP II foci scored for 8C3/C4 at the inner edge, at the outer edge and outside of the territory were 52%, 51% and 61%, respectively (n=173 loci; **Figure 3-6**). In addition, 3 out of 10 loci found inside the territory were also associated with RNAP II foci. There was no significant difference

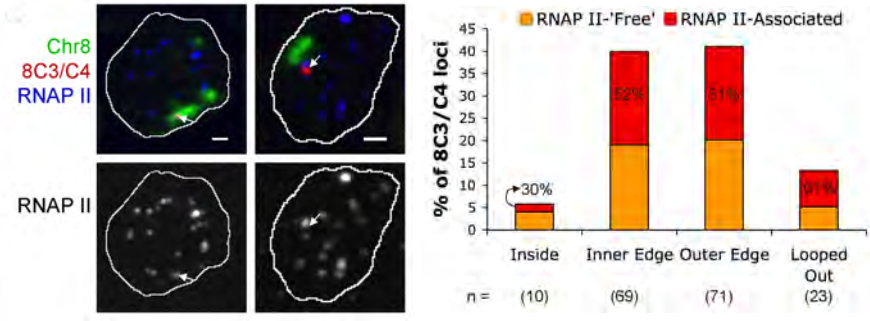


Figure 3-6. Association of 8C3/C4 with RNAP II foci is independent of position vs. the CT. The active form of RNAP II (blue or grayscale), 8C3/C4 (red; arrows) and chromosome 8 (green) were labelled as is Figure 3-5. Nuclei are outlined by a white line. Loci were classified as RNAP II-'free' (pool of 'separate' and 'adjacent' categories in B; n=89 loci) or RNAP II-associated ('overlapped' category in B; n=98 loci) and the respective distances to CT edge measured. No significant difference between the distributions was detected (p=0.73; K-S test).

Bars, 1 μm.

between the distributions of RNAP II-'free' and RNAP II-associated loci in relation to the CT (chi-squared test for independence). These data suggest that transcription of the genes at 8C3/C4 takes place regardless of the position of the locus relative to its CT. This is in agreement with the observation that transcription factories are present within CTs [18,19,21].

3.5- Discussion

Gene activation has been correlated with repositioning of a locus away from nuclear landmarks such as CTs, pericentromeric heterochromatin or the nuclear periphery. One interpretation of this phenomenon is that relocation critically drives transcription by positioning loci in nuclear zones of increased transcriptional competence. This idea may predict some degree of co-regulation between genes closely juxtaposed on the chromosome template [*7*]. However, here we demonstrate that the expression levels of functionally unrelated genes closely juxtaposed on the chromosome template are regulated in a largely autonomous fashion. Detailed expression analysis showed that the expression level of each gene in a gene-dense region varied between many different tissues in a manner that is independent of the expression changes of neighbouring genes, and is accompanied by gene-specific changes in histone H3 acetylation and occupancy. Thus, while clustering of genes will be beneficial to maintain active states and ensure pan-cellular expression, which may be caused by association of these clusters with active regions of the nucleus like splicing speckles and transcription factories, our results suggest that mechanisms working at the level of single genes determine the expression levels in the various tissues.

3.6- Accession numbers

Micro-array data used for **Figure 3-2**: Gene Expression Omnibus (GEO) GSE5891 (E14.5 fetal liver and E14.5 fetal brain) and ArrayExpress E-MEXP-839 (adult liver) [36, 37].

3.7- Acknowledgements

We thank Frank Grosveld for discussion, George Garinis and Marieke Simonis for sharing expression data and Robert-Jan Palstra and other members of the laboratories for assistance.

3.8- References

1. Caron, H., van Schaik, B., van der Mee, M., Baas, F., Riggins, G., van Sluis, P., Hermus, M.C., van Asperen, R., Boon, K., Voute, P.A., Heisterkamp, S., van Kampen, A., and Versteeg, R. (2001). The human transcriptome map: clustering of highly expressed genes in chromosomal domains. Science *291*, 1289-1292.
2. Lercher, M.J., Urrutia, A.O., and Hurst, L.D. (2002). Clustering of housekeeping genes provides a unified model of gene order in the human genome. Nat Genet *31*, 180-183.
3. Singer, G.A., Lloyd, A.T., Huminiecki, L.B., and Wolfe, K.H. (2005). Clusters of co-expressed genes in mammalian genomes are conserved by natural selection. Mol Biol Evol *22*, 767-775.
4. Grosveld, F., Blom van Assendelft, G., Greaves, D.R., and Kollias, G. (1987). Position-independent, high-level expression of the human beta-globin gene in transgenic mice. Cell *51*, 975-985.
5. Gierman, H.J., Indemans, M.H., Koster, J., Goetze, S., Seppen, J., Geerts, D., van Driel, R., and Versteeg, R. (2007). Domain-wide regulation of gene expression in the human genome. Genome Res *17*, 1286-1295.
6. Sproul, D., Gilbert, N., and Bickmore, W.A. (2005). The role of chromatin structure in regulating the expression of clustered genes. Nat Rev Genet *6*, 775-781.
7. Chakalova, L., Debrand, E., Mitchell, J.A., Osborne, C.S., and Fraser, P. (2005). Replication and transcription: Shaping the landscape of the genome. Nat Rev Genet *6*, 669 - 677.
8. Volpi, E.V., Chevret, E., Jones, T., Vatcheva, R., Williamson, J., Beck, S., Campbell, R.D., Goldsworthy, M., Powis, S.H., Ragoussis, J., Trowsdale, J., and Sheer, D. (2000). Large-scale chromatin organization of the major histocompatibility complex and other regions of human chromosome 6 and its response to interferon in interphase nuclei. J Cell Sci *113 (Pt 9)*, 1565-1576.
9. Williams, R.R., Broad, S., Sheer, D., and Ragoussis, J. (2002). Subchromosomal positioning of the epidermal differentiation complex (EDC) in keratinocyte and lymphoblast interphase nuclei. Exp Cell Res *272*, 163-175.
10. Chambeyron, S., and Bickmore, W.A. (2004). Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. Genes Dev *18*, 1119-1130.
11. Kurz, A., Lampel, S., Nickolenko, J.E., Bradl, J., Benner, A., Zirbel, R.M., Cremer, T., and Lichter, P. (1996). Active and inactive genes localize preferentially in the periphery of chromosome territories. J Cell Biol *135*, 1195-1205.
12. Abranches, R., Beven, A.F., Aragon-Alcaide, L., and Shaw, P.J. (1998). Transcription sites are not correlated with chromosome territories in wheat nuclei. J Cell Biol *143*, 5-12.
13. Verschure, P.J., van Der Kraan, I., Manders, E.M., and van Driel, R. (1999). Spatial relationship between transcription sites and chromosome territories. J Cell Biol *147*, 13-24.
14. Scheuermann, M.O., Tajbakhsh, J., Kurz, A., Saracoglu, K., Eils, R., and Lichter, P. (2004). Topology of genes and nontranscribed sequences in human interphase nuclei. Exp Cell Res *301*, 266-279.
15. Branco, M.R., and Pombo, A. (2006). Intermingling of Chromosome Territories in Interphase Suggests Role in Translocations and Transcription-Dependent Associations. PLoS Biol *4*, e138.
16. Shopland, L.S., Johnson, C.V., Byron, M., McNeil, J., and Lawrence, J.B. (2003). Clustering of multiple specific genes and gene-rich R-bands around SC-35 domains: evidence for local euchromatic neighborhoods. J Cell Biol *162*, 981-990.

17. Brown, J.M., Leach, J., Reittie, J.E., Atzberger, A., Lee-Prudhoe, J., Wood, W.G., Higgs, D.R., Iborra, F.J., and Buckle, V.J. (2006). Coregulated human globin genes are frequently in spatial proximity when active. J Cell Biol *172*, 177-187.
18. Johnson, C., Primorac, D., McKinstry, M., McNeil, J., Rowe, D., and Lawrence, J.B. (2000). Tracking COL1A1 RNA in osteogenesis imperfecta. splice-defective transcripts initiate transport from the gene but are retained within the SC35 domain. J Cell Biol *150*, 417-432.
19. Iborra, F.J., Pombo, A., Jackson, D.A., and Cook, P.R. (1996). Active RNA polymerases are localized within discrete transcription "factories' in human nuclei. J Cell Sci *109 (Pt 6)*, 1427-1436.
20. Pombo, A., Jackson, D.A., Hollinshead, M., Wang, Z., Roeder, R.G., and Cook, P.R. (1999). Regional specialization in human nuclei: visualization of discrete sites of transcription by RNA polymerase III. Embo J *18*, 2241-2253.
21. Faro-Trindade, I., and Cook, P.R. (2006). A conserved organization of transcription during embryonic stem cell differentiation and in cells with high C value. Mol Biol Cell *17*, 2910-2920.
22. Osborne, C.S., Chakalova, L., Brown, K.E., Carter, D., Horton, A., Debrand, E., Goyenechea, B., Mitchell, J.A., Lopes, S., Reik, W., and Fraser, P. (2004). Active genes dynamically colocalize to shared sites of ongoing transcription. Nat Genet *36*, 1065-1071.
23. Kimura, H., Sugaya, K., and Cook, P.R. (2002). The transcription cycle of RNA polymerase II in living cells. J Cell Biol *159*, 777-782.
24. Becker, M., Baumann, C., John, S., Walker, D.A., Vigneron, M., McNally, J.G., and Hager, G.L. (2002). Dynamic behavior of transcription factors on a natural promoter in living cells. EMBO Rep *3*, 1188-1194.
25. Jackson, D.A., Iborra, F.J., Manders, E.M., and Cook, P.R. (1998). Numbers and organization of RNA polymerases, nascent transcripts, and transcription units in HeLa nuclei. Mol Biol Cell *9*, 1523-1536.
26. Martin, S., and Pombo, A. (2003). Transcription factories: quantitative studies of nanostructures in the mammalian nucleus. Chromosome Res *11*, 461-470.
27. Guillot, P.V., Xie, S.Q., Hollinshead, M., and Pombo, A. (2004). Fixation-induced redistribution of hyperphosphorylated RNA polymerase II in the nucleus of human cells. Exp Cell Res *295*, 460-468.
28. Su, A.I., Cooke, M.P., Ching, K.A., Hakak, Y., Walker, J.R., Wiltshire, T., Orth, A.P., Vega, R.G., Sapinoso, L.M., Moqrich, A., Patapoutian, A., Hampton, G.M., Schultz, P.G., and Hogenesch, J.B. (2002). Large-scale analysis of the human and mouse transcriptomes. Proc Natl Acad Sci U S A *99*, 4465-4470.
29. Semon, M., and Duret, L. (2006). Evolutionary origin and maintenance of coexpressed gene clusters in mammals. Mol Biol Evol *23*, 1715-1723.
30. Bernstein, B.E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D.K., Huebert, D.J., McMahon, S., Karlsson, E.K., Kulbokas, E.J., 3rd, Gingeras, T.R., Schreiber, S.L., and Lander, E.S. (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. Cell *120*, 169-181.
31. Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693-705.
32. Roh, T.Y., Cuddapah, S., and Zhao, K. (2005). Active chromatin domains are defined by acetylation islands revealed by genome-wide mapping. Genes Dev *19*, 542-552.
33. Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., Wang, W., Weng, Z., Green, R.D., Crawford, G.E., and Ren, B. (2007). Distinct

and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet *39*, 311-318.

34. Xie, S.Q., Martin, S., Guillot, P.V., Bentley, D.L., and Pombo, A. (2006). Splicing speckles are not reservoirs of RNA polymerase II, but contain an inactive form, phosphorylated on serine2 residues of the C-terminal domain. Mol Biol Cell *17*, 1723-1733.
35. Grande, M.A., van der Kraan, I., de Jong, L., and van Driel, R. (1997). Nuclear distribution of transcription factors in relation to sites of transcription and RNA polymerase II. J Cell Sci *110 (Pt 15)*, 1781-1791.
36. Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., and de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nat Genet *38*, 1348-1354.
37. van der Pluijm, I., Garinis, G.A., Brandt, R.M., Gorgels, T.G., Wijnhoven, S.W., Diderich, K.E., de Wit, J., Mitchell, J.R., van Oostrom, C., Beems, R., Niedernhofer, L.J., Velasco, S., Friedberg, E.C., Tanaka, K., van Steeg, H., Hoeijmakers, J.H., and van der Horst, G.T. (2006). Impaired genome maintenance suppresses the growth hormone--insulin-like growth factor 1 axis in mice with Cockayne syndrome. PLoS Biol *5*, e2.

Chapter 4

Transcription and chromatin organization of the housekeeping gene cluster 8C3/C4 containing an integrated β-globin Locus Control Region

Daan Noordermeer*, Miguel R. Branco*,[1], Erik Splinter, Petra Klous, Wilfred van IJcken[2], Sigrid Swagemakers[3], Manousos Koutsourakis, Peter van der Spek[3], Ana Pombo[1] and Wouter de Laat

* These authors contributed equally to this work

[1] MRC Clinical Sciences Centre, Faculty of Medicine, Imperial College London, Hammersmith Hospital Campus, Du Cane Road, London, W12 0NN, United Kingdom

[2] Erasmus Center for Biomics, Erasmus Medical Center, PO Box 2040, 3000 CA Rotterdam, The Netherlands

[3] Erasmus Center for Bioinformatics, Erasmus Medical Center, PO Box 2040, 3000 CA Rotterdam, The Netherlands

*Data in this chapter was published as part of PLOS Genetics, 2008, **4**, e1000016*

4.1- Abstract

The activity of locus control regions (LCR) has been correlated with chromatin decondensation, spreading of active chromatin marks, locus repositioning away from its chromosome territory (CT), increased association with transcription factories, and long-range interactions via chromatin looping. To investigate the relative importance of these events in the regulation of gene expression, we targeted the human β-globin LCR in two opposite orientations to a gene-dense region in the mouse genome containing mostly housekeeping genes. We found that each oppositely oriented LCR influenced gene expression on both sides of the integration site and over a maximum distance of 150 kilobases. A subset of genes was transcriptionally enhanced, some of which in an LCR orientation-dependent manner. The locus resides mostly at the edge of its CT and integration of the LCR in either orientation caused a more frequent positioning of the locus away from its CT. Locus association with transcription factories increased moderately, both for loci at the edge and outside of the CT. These results show that nuclear repositioning is not sufficient to increase transcription of any given gene in this region. We identified long-range interactions between the LCR and two upregulated genes and propose that LCR-gene contacts via chromatin looping determine which genes are transcriptionally enhanced.

4.2- Introduction

One of the main aims in the post-genomic era has been to understand how genes are regulated at the level of transcription, giving rise to cell-type specific transcriptomes. Most of our knowledge on the regulation of gene transcription is biased towards studies of a small number of atypical genes showing highly restricted expression patterns. Expression of these tissue-specific genes is often controlled by distant transcription regulatory DNA elements. The β-globin locus control region (LCR) is a prototype of a strong mammalian regulatory DNA element. At its endogenous position, the LCR enhances the expression of the mouse β-globin-like genes 25-100 fold [1]. LCR-mediated transcriptional enhancement has been correlated with chromatin opening [2], histone acetylation of the locus (in case of the LCR of the human growth hormone cluster) [3], the initiation of intergenic transcripts [4, 5], the spreading of a histone methyltransferase [6] and intrachromosomal interactions with active genes via chromatin looping [7, 8]. At the level of nuclear organization, LCR activity has been correlated with repositioning of loci away from their chromosome territory (CT) and towards the nuclear interior, and increased association with transcription factories [9, 10]. However, the relative importance of these events for the enhancement of gene transcription is currently unclear. It is difficult to assess the hierarchy and importance of each level of regulation, as many of the observations are made on different gene loci and in different types of cells, often cultured in vitro.

Here, we generated transgenic mice containing the human β-globin LCR in opposite orientations integrated at the gene-dense region 8C3/C4, that we previously characterized extensively (see **Chapter 3**). We analysed in detail gene expression, chromatin structure and nuclear positioning of the region containing the integrated LCR and correlated our findings to the results found for the WT locus. We found that insertion of the LCR induced a relocation of the locus away from the edge of its CT and increased the association with RNA polymerase (RNAP) II factories. The latter occurred independently of the locus being positioned at the edge or outside the CT. Changes in gene expression occurred bi-directionally and as far as 150 kb from the site of LCR insertion, but were variable for individual genes, depending on the orientation of the integrated LCR. The fact that both orientations of the LCR caused a repositioning of the locus away from its CT, yet genes within the locus responded selectively to only one LCR orientation, shows that the observed repositioning is not the only factor determining the increase in transcription of the individual genes. The gene expression patterns seem incompatible with tracking or

spreading mechanisms, and we demonstrate that the LCR physically interacts with the two most upregulated genes. We propose that LCR-gene contacts via chromatin looping are key to the upregulation of at least a number of genes in this gene-dense region.

4.3- Materials and methods

4.3.1- Gene targeting and the generation of transgenic mice

For the insertion of the human β-globin LCR into the mouse *Rad23a* gene, the *ClaI* Neo resistance cassette of an existing *Rad23a* targeting construct that removes *Rad23a* exon II-VII [11] was replaced by a *ClaI* fragment containing a PGK-Puro resistance cassette coupled to a 21.5 kb *SalI-ClaI* fragment containing the human β-globin LCR. Constructs with the *ClaI* fragment in opposite orientations were obtained: LCR-S, with hypersensitive site 1 of the LCR at the 3'-end of the *Rad23a* gene, and LCR-AS with HS1 of the LCR at the 5'-end of the *Rad23a* gene. Targeting in Ola129-derived ES cells, blastocyst injection to generate chimeric mice and breeding to obtain homozygous transgenic animals in an FVB background was done as described [11]. Genotyping was performed by Southern blot. Animal experiments were carried out according to institutional and national guidelines (Committee on Experiments with Laboratory Animals (DEC-Consult); Ministry of Agriculture, Nature and Food Quality, The Hague, The Netherlands).

4.3.2- DNaseI hypersensitivity assays

DNaseI hypersensitivity assays were carried out on isolated nuclei from 6 E14.5 fetal livers. Nuclei were isolated in ice-cold lysis mix (10 mM Tris-HCl pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% NP40) by dounce homogenisation and subsequent slow spinning. Nuclei were incubated for 3 min at 37°C in lysis mix without NP40, substituted with 1 mM $CaCl_2$ and increasing amounts of DNaseI. Reactions were stopped by adding equal amounts of 2x stop-mix (0.6 M NaCl, 20 mM Tris-HCl pH 8.0, 10 mM EDTA, 1% SDS), treated overnight with Proteinase K and DNA purified with phenol/chlorophorm. Samples were digested overnight (HS1, HS2: *PstI*; HS3, HS4, HS5: *HinDIII*), run on 0.7% agarose gels and visualized by Southern blotting. Primer sequences for Southern probes are presented in **Table 4-1**.

4.3.3- FAIRE assay

Formaldehyde Assisted Identification of Regulatory Elements (FAIRE) was performed as described before [12], except that genomic sites were analysed by qPCR (sequences of primers in **Table 4-1**) using Platinum Taq DNA polymerase (Invitrogen) and SYBR Green (Sigma) on an Opticon 2 Real-Time PCR Detection System (Biorad). Enrichments were calculated relative to the amylase promoter and values were normalized to input measurements.

DnaseI hypersensitivity assay

Primer set	Sequence
Human β-globin LCR HS1	GAAGCCTCTGGTCAGCAT
	CAGTGGTAGAAGCAGGAAGAT
Human βglobin LCR HS2	GACACATGCTCACATACGG
	AACCCTCACAGCTGCTAAC
Human β-globin LCR HS3	AGGGCTCCAGCATGTAGA
	TCAAGGTCGAAGGTAGGAAC
Human β-globin LCR HS4	AATTTAAGCTCCAGTCTCTGC
	TCCCATATTCTGAAGCATTC
Human β-globin LCR HS5	CAGCCCTGAGCACTTACA
	TTTCTCCTTTGTTGACCAAG

Gene expression analysis

Primer set	Sequence
mRNA Hprt	AGCCTAAGATGAGCGCAAGT
	ATGGCCACAGGACTAGAACA
mRNA Ier2	ATCACTACCGTCGCTCAA
	TACGCAAGAGGAAGTGCT
mRNA Btbd14b	TTGCCACCCAAAGCTCTA
	TTCATAAGCTGTGCCCTTG
mRNA Trmt1	TATAACCCGGTGCAGGAAT
	TTCTCTGAGTCCTTCTCTCC
mRNA Lyl1	GACCCTTCAGCATCTTCC
	CTGTTGGTGAACACTCGC
mRNA Nfix	GAACTGGACCTTTATCTGGC
	CGCAACTGGAGTCTGTGA
mRNA Dand5	CTTCTACATTCCCAGCTCG
	CTGGACCAATACCGTGGA
mRNA Gadd45gip1	AAAGCAGAAGCGAGAACG
	ATAGCAGCAATTCGTGCC
mRNA Calr	GACTTTCTGCCACCCAAG
	GTTCCCACTCTCCATCCA
mRNA Farsla	GAGTAGCCATGGCGGATA
	CCAGCACTTGGTAGAACGA
mRNA Syce2	TTGTACACCGTTTCCACAGT
	GACCATGCACTTATGACCAA
mRNA Gcdh	CTGCCGATGAGAAACTGATA
	TCGACCTGTAGCCACTGTC
mRNA Klf1	CATCAGTACACTCACCACCCT
	CGGAACCTGGAAAGTTTG
mRNA Dnase2a	TGCCAATCCTTGCAAACT
	CGACCAACCTCCTAAATCC
mRNA Mast1	TGGAAGGTGGTGACTGTG
	AATTGTGTAAGTACTCAAGGGC

Chromatin immunoprecipitation and FAIRE

Primer set	Sequence
Promoter Lyl1	TTCAAAGCTAGACCAACCTCA
	CACAGCACTGGAAGACCC
Promoter Nfix	TCCAAACCACACTTCAGTAG
	GAAGGAGAAACACAGCGTT
Intergenic region Nfix - Dand5	ACAGAGACAGCCGAATACC
	GACCTCCCTCTGTCTGTTC
Promoter Dand5	AACTCTCAAGCTGCTCTCC
	CTGTGCAGTCGTTTGTCTG
Promoter Gadd45gip1	TCGGAGGGTAAAGGCATT
	AGTGTTGAAGTGTTGGTGAT
Promoter Rad23a	CACCAAGACAGCGGAATG
	GGCTGCACCTTACCTTAGA
Promoter Calr	CTGGGAAGCAATGGAAAG
	TTATATTCACCTACCTCTCACCC
Promoter Farsla	AACTAAAAGCCACTGGGGT
	TAAGTGTGGCAAATGAGCC
Promoter Syce2	CGCACCTACGCATTATGA
	TGCCTTTTGGGCTATGCT
3' Syce2, 3' Gcdh	ATTGCTCTCCCAAGGATCA
	TGAGCTGAAGATTCCAAACC
Promoter Gcdh	GGAACCATAACCTGGAAGGG
	AAGGAGGAACCAATGAGCAA
Promoter Klf1	CTTTGCCTGGGTCTTATCA
	TCCTCTCTCTCTTCTGAATC
Promoter Dnase2a	GGGTCACGAGATTCAGATG
	GCTTCGTCTCCACCCTCG
Human HS5	CTGAAGCTGCTGTTATGACC
	ACAACTTCCTTGCTTGGG
Human HS4	TTTCTCTCTCCCACTCAGC
	TGCTATCAAAGCCCTGACA
Intergenic region human HS3 – human HS4	CACTTTATTCGTCTGTGTCCTG
	GGTGTAGTCTTTCCTGATGTC
Human HS3	GCTCAGATAGGTGGTTAGGT
	TGGTCTATCTCTCCTGGCT
Human HS2	CTCCATTAGTGACCTCCCA
	TTACACAGAACCAGAAGGC
Intergenic region human HS1 – human HS2	CCAACATAGTGAAACCTGGT
	CTCAGCCTCCCAAGTAGTT
Human HS1	TGCGGTTGTGGAAGTTTAC
	CACTAAGGGTGAGGATGCT
3' human HS1	GGGAAGAATGATGGGATTACC
	CAAGAGTTTGGAGAACAGGTT
Mouse HS5	TTCACCACTAGAGGGAAGG
	GAACGTGCAAGTGTCCTG
Mouse HS2	GTCATGCTGAGGCTTAGG
	TTCCCTGTGGACTTCCTC
Promoter Amylase	CTCCTTGTACGGGTTGGT
	AATGATGTGCACAGCTGAA
Promoter Beta Major	GGGAGAAATATGCTTGTCATC
	CAACTGATCCTACCTCACCTT

Chromosome Conformation Capture (3C)

Primer	Sequence
8C3/C4 fixed fragment	CCTTCCTCCACCATGATGA
8C3/C4 fragment I	GCATATTTGACTTTTACAAGCTGG
8C3/C4 fragment II	GTGGTAGCAGAAGTCTCAAG
8C3/C4 fragment III	ACTCTTAAACTGGCTGTGATG
8C3/C4 fragment IV	CGGGTCACGAGATTCAGA
Probe	Sequence
8C3/C4 double-dye probe	AAAGCTTAGGGCTCCAGCTTCCC
8C3/C4 fragment V	AGTTCCTAGCCGTTCCTTAG
8C3/C4 fragment VI	CAGAATGGGTCTATGGAAAGAG
Ercc3 fragment I	ATGGCCTGAAGAAACCGC
Ercc3 fragment II	CTTAGGCAACACACTCAAGC
Probe	Sequence
Ercc3 double-dye probe	CCATCTTCCACATCAATGGGACCC

Table 4-1 Primers and probes

4.3.4- Gene expression analysis

Total RNA was isolated using the RNeasy Mini kit (Qiagen) from livers of at least three independent embryos. cDNA synthesis was performed using SuperScript II Reverse Transcriptase and Oligo(dT)$_{12\text{-}18}$ primer according to the manufacturers' instructions (Invitrogen). Products were quantified by qPCR, using Platinum Taq DNA polymerase (Invitrogen) and SYBR Green (Sigma) on an Opticon 2 Real-Time PCR Detection System (BioRad). Primer sequences in **Table 4-1**. Transcript levels were normalized to the *Hprt1* transcript, encoding a relatively high expressed housekeeping gene on an unrelated chromosome, verified with Affymetrix gene expression analysis (see **Chapter 3**) not to be influenced by the integration of the human β-globin LCR (results not shown).

4.3.5- Chromatin immunoprecipitation

ChIP was performed according to the Upstate protocol (http://www.upstate.com), with two modifications: (1) E14.5 fetal livers were made single cell by applying a cell-strainer cap (BD Falcon #352340, BD Biosciences) and (2) cells were fixed for 5 minutes in a 2% formaldehyde solution at room temperature. Chromatin fragments were quantified by qPCR (sequences of primers in **Table 4-1**) using Platinum Taq DNA polymerase (Invitrogen) and SYBR Green (Sigma) on an Opticon 2 Real-Time PCR Detection System (Biorad). Enrichments were calculated relative to the endogenous β-globin promoter or amylase promoter and values were normalized to input measurements. Antibodies used: anti acetyl-Histone H3 (#06-599, Upstate); anti K4 trimethyl H3 (#07-473, Upstate); anti C-terminal-Histone H3 (#ab1791; Abcam); anti CTCF [13].

4.3.6- Cell preparation and cryosectioning

For the preparation of cell blocks for cryosectioning, E14.5 fetal liver and brain tissues were fixed in 4 and then 8% paraformaldehyde in 250 mM HEPES pH 7.6 (10 min and 2 h respectively) [14]. Cell pellets were embedded in 2.1 M sucrose in phosphate-buffered saline (PBS) and frozen in liquid nitrogen as described previously [15]. Cryosections (140-180 nm in thickness, deduced from interference colour) were cut using an UltraCut UCT52 ultracryomicrotome (Leica), captured in sucrose drops, and transferred to glass coverslips.

4.3.7- Cryo-FISH

Cryo-FISH was performed as described previously [16]. A probe for the 8C3/C4 locus was obtained by labeling a BAC (RP24-319P23) with biotin or rhodamine by nick-translation (Roche). The BAC probe was co-precipitated with mouse Cot1 DNA (Roche; 1.7 μg/μl final concentration) and resuspended in either hybridisation buffer (50% deionized formamide, 10% dextran sulfate, 2xSSC, 50 mM phosphate buffer pH 7.0) or a FITC-labeled mouse whole chromosome 8 paint (Applied Spectral Imaging). Probes were denatured at 70°C for 10 min, and re-annealed at 37°C for 30 min before hybridisation. Probe specificity was confirmed on mouse spleen metaphase spreads.

4.3.8- Immunolabeling

Immunolabeling of cryosections was performed as described previously [16]. The biotin-labeled BAC probe for the 8C3/C4 locus was detected using rhodamine-conjugated neutravidin (1/500; Molecular Probes), followed by a biotin-conjugated goat anti-avidin antibody (1/500; Vector) and rhodamine-conjugated neutravidin. Splicing speckles were detected with a human autoimmune serum against Sm antigen (1/2000; ANA-CDC), followed by a biotin-conjugated donkey anti-human antibody (1/100; Jackson ImmunoResearch) and an AlexaFluor488-conjugated neutravidin (1/100; Molecular Probes). Serine 2 phosphorylated RNAP II was immunolabeled with H5 (1/1000; Covance), followed by an IgM-specific biotin-conjugated donkey anti-mouse antibody (1/250; Jackson ImmunoResearch) and Alexa 488-conjugated neutravidin or Alexa 647-conjugated streptavidin (1/100; Molecular Probes). After immunolabeling and before cryo-FISH, antibodies were fixed with 8% paraformaldehyde in 250 mM HEPES pH 7.6 (1 h), or with 2 mM EGS in PBS (30 min, 37°C).

4.3.9- Microscopy and image processing

For confocal laser scanning microscopy, images were collected sequentially on a Leica TCS SP2 (100X PL APO 1.40 Oil objective) equipped with Argon (488 nm) and HeNe (543 nm; 633 nm) lasers. For wide-field light microscopy, images were collected sequentially on a Delta-Vision Spectris system (Applied Precision) equipped with an Olympus IX70 widefield microscope (100X UPlanFl 1.3 Oil objective), a charge-coupled device camera, and the following filters: DAPI, FITC, RD-TR-PE, CY-5, CFP, YFP. The use of ultrathin cryosections allows for the use of wide-field microscopy with no reduction in axial (z) resolution and only a small reduction in lateral resolution. No bleedthrough was detected

in these conditions, and images were collected without saturation of intensities. The images presented were contrast stretched, without further thresholding or filtering. For the generation of randomly positioned loci, the experimental images collected for the analysis of 8C3/C4 position relative to transcription factories were used. We developed a macro on ImageJ (Wayne Rasband, NIH, USA) that thresholds the DAPI signal and, for each locus, generates random coordinates until they fall within the mask of the corresponding thresholded nuclear section. These coordinates were used to measure the distance to the nearest transcription factory.

4.3.10- Chromosome Conformation Capture (3C) analysis

3C analysis was performed as described before, with slight modifications [17-19]. 3C material was digested using the restriction enzyme *BglII*. Ligation frequencies were quantified by qPCR (Opticon 2 Real-Time PCR Detection System, Biorad) using Platinum Taq DNA polymerase (Invitrogen) and double-dye oligonucleotides (5'FAM + 3'TAMRA) as probe (sequences of primers and probes in **Table 4-1**). To correct for differences in quality and quantity of templates, ligation frequencies between the fragments in the region on chromosome 8 were normalized to two fragments in the *Ercc3* locus, assumed to have a constant spatial organization independent of the presence of the human β-globin LCR. To correct for PCR amplification efficiency of different primer sets a *BglII* digested and re-ligated control template containing equimolar amounts of all possible ligation products was used. This control template was composed of two BAC clones containing all the analysed fragments in the region (RP24-136A15 and RP24-319P23, Ensemble Genome Browser), a construct containing the human β-globin LCR (see "Construction of targeting vectors") and a PAC clone containing the *Ercc3* locus (443-C18, MRC geneservice).

4.4- Results

4.4.1- Introduction of the β-globin LCR into the gene-dense region 8C3/C4

In order to study the function of the β-globin LCR and the hierarchy among changes in chromatin structure, nuclear organization and transcription regulation, we introduced the human β-globin LCR in two orientations into 8C3/C4, the gene-dense region on mouse chromosome 8 that was characterized in **Chapter 3** (see **Figure 3-1A, B**).

The β-globin LCR is a strong, erythroid-specific, regulatory DNA element that confers position-independent and copy-number dependent expression to transgenes in mice [2]. At its endogenous genomic location, the LCR is required for high expression levels of the β-globin-like genes [1]. Furthermore, the LCR has been implicated in the relocation of transgenes away from pericentromeric heterochromatin [8] and the nuclear periphery, and was suggested to be required for association of the β-globin locus with transcription factories [9].

We targeted the full, 21 kb, human LCR in sense (S) and anti-sense (AS) orientations into the *Rad23a* gene (gene 0 in **Figure 3-1B**) using homologous recombination in ES cells (**Figure 4-1A, B**). *Rad23a* is located centrally in 8C3/C4 and can be knocked out on both alleles without getting an apparent change in phenotype, due to the redundant presence of the homologous *Rad23b* gene in the mouse genome [11]. Mice homozygous for the integrated LCR in both orientations also did not show an abnormal phenotype. To test for functionality of the integrated LCR, DNaseI hypersensitivity assays and FAIRE assays were performed. Each of the 5 hypersensitive sites (HSs) characteristic for the human β-globin LCR were found by DNaseI hypersensitivity assays to be present in E14.5 liver cells of LCR-S +/+ and LCR-AS +/+ fetuses, indicating that the integrated LCR binds its normal repertoire of transcription factors (**Figure 4-1C**). A comparison between different regions of the LCR in E14.5 liver and brain cells of LCR-S +/+ fetuses using FAIRE assays (Formaldehyde Assisted Identification of Regulatory Elements) [12] showed strong nucleosomal depletion at HS2-4 in liver, but not in brain. HS5 is mildly enriched, both in liver and brain. HS1 in liver shows an enrichment comparable to non-hypersensitive sites in the LCR, but this enrichment is higher than for these sites in brain (**Figure 4-1D**), as was previously reported [1].

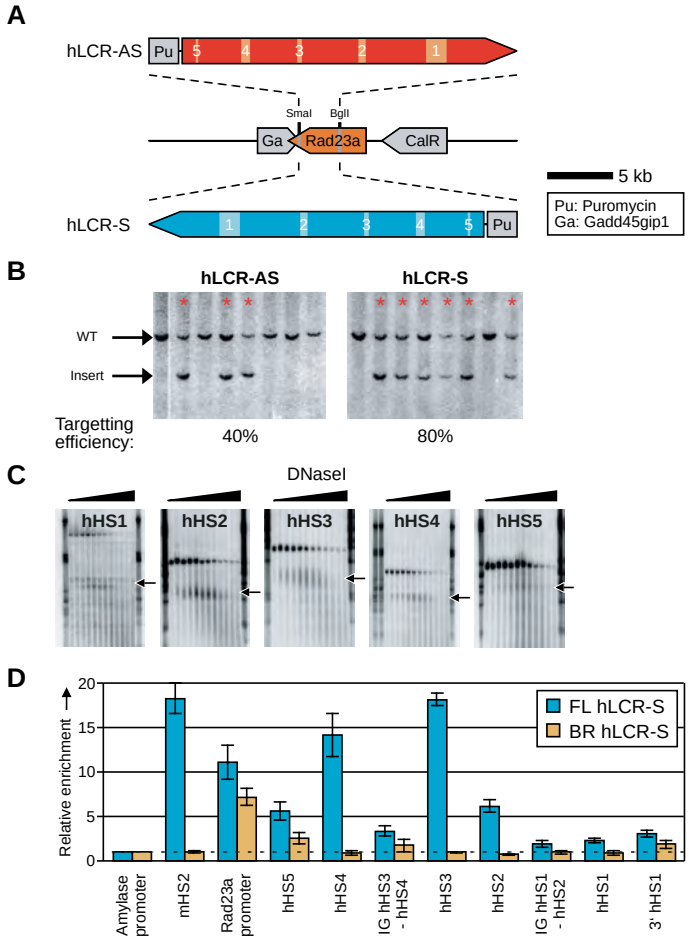


Figure 4-1. Integration of the human β-globin LCR in 8C3/C4. (A) The full human β-globin LCR, linked to a puromycin selection marker, was integrated in anti-sense orientation (LCR-AS, red) and sense orientation (LCR-S, blue) relative to Rad23a, removing exons II-VII from the *Rad23a* gene. Hypersensitive sites (HSs) in the LCR are numbered and represented by shaded boxes. **(B)** Southern blot showing targeting of the LCR into the *Rad23a* locus. Positive clones are indicated by red arrows. Note that targeting is extremely efficient. **(C)** DnaseI hypersensitivity of HSs in the integrated LCR in the sense orientation. E14.5 fetal liver DNA was digested with increasing amounts of DnaseI and DNA fragments containing the respective HSs were visualized by Southern blotting. Hypersensitive DNA fragments are indicated by arrows. LCR-AS HSs showed comparable patterns (not shown). **(D)** FAIRE analysis of regions representing LCR hypersensitive sites in E14.5 fetal liver and brain containing the LCR integrated in the sense orientation. Error bars represent standard error (SE) of at least two independent experiments. IG: intergenic region.

4.4.2- Gene expression at 8C3/C4 in the presence of an integrated LCR

We first studied the effect of the integration of the LCR in either orientation on the expression of the surrounding genes by quantitative reverse transcriptase PCR (qRT-PCR). A complex pattern of transcriptional upregulation around the integration site was observed in E14.5 livers (**Figure 4-2A**). To exclude that integration *per se*, or the disruption of the *Rad23a* gene, had an effect on gene expression at 8C3/C4, we also analysed transcription of several genes in E14.5 brain cells, which do not contain LCR activity. No upregulation of genes in the region was observed in these cells (**Figure 4-2B**), with the exception of one gene, which appears around 1.5 fold upregulated in one of the transgenic lines. Possibly, this is due to its altered genetic background. We conclude that the complex pattern of transcriptional upregulation in fetal liver cells is dependent on the erythroid-specific LCR activity.

In fetal liver, each oppositely oriented LCR enhanced the expression of at least seven genes surrounding the integration site, the most distal one (*Ier2*, gene –7) being over 150 kb away from the integrated LCR. Both LCRs activated genes on the plus and minus strands, as well as upstream and downstream of the integration site, demonstrating bi-directional activity of the LCR. The latter observation was somewhat surprising given that the LCR is thought to function in a unidirectional manner in the β-globin locus [20]. However, this finding was in agreement with results showing that a marked β-globin gene placed upstream of the LCR competes with downstream genes for activation by the LCR [21]. We found that upstream gene activation occurred despite the binding of the insulator protein CTCF to the outer hypersensitive site (HS5) of the LCR (**Figure 4-2C**), confirming that binding of CTCF does not necessarily lead to enhancer-blocking [22]. Genes present in 8C3/C4 reacted very differently to the integration of the two LCRs (**Figure 4-2A**). Genes at positions -1 (*Gadd45gip1*), +4 (*Gcdh*), and +6 (*Dnase2a*) relative to the integration site (defined as gene position 0) were upregulated to levels that were similar between both LCRs. Genes at positions -2 (*Dand5*) and +1 (*CalR*) also responded to both LCRs, but reached significantly different levels of mRNA. Interspersed, at positions -6 (*Btbd14b*), -5 (*Trmt1*), -3 (*Nfix*), +2 (*Farsla*) and +5 (*Klf1*) were genes not responding to the LCR in either orientation. We reasoned that structural constraints or a lack of certain *cis*-regulatory elements prevented these genes from communicating with the LCR, or that they were transcribed already at maximum rates in wild-type cells. The finding that gene +5 (*Klf1*) is not upregulated by the LCR is particularly interesting, since

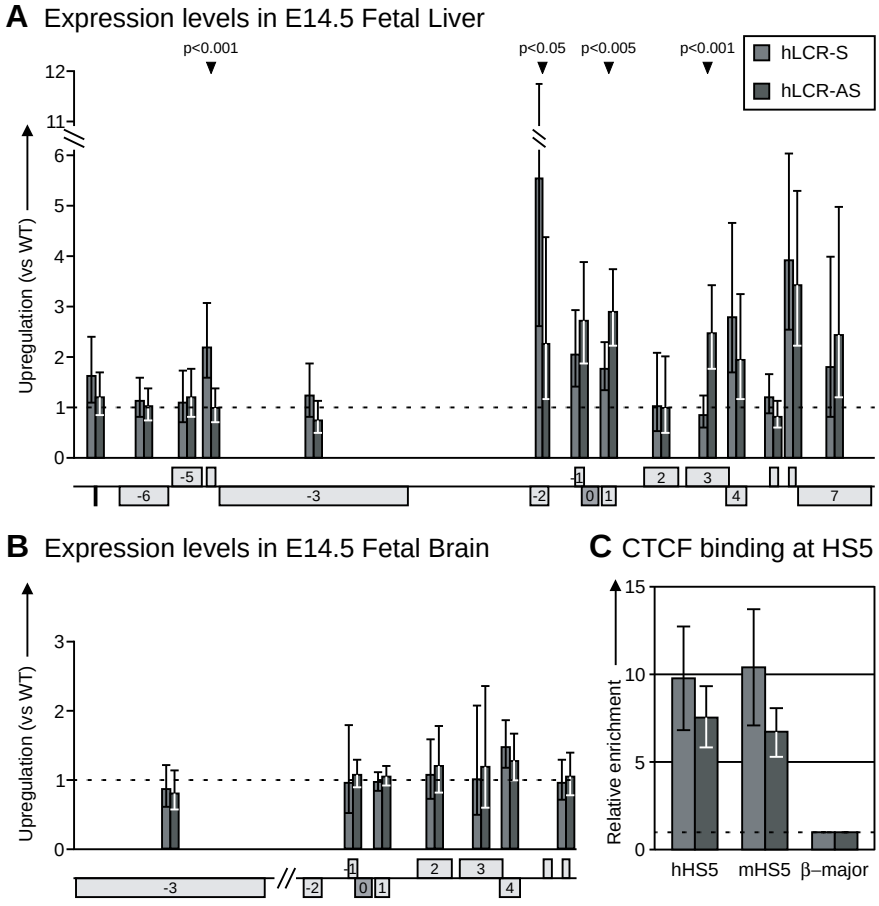


Figure 4-2. Effect of LCR insertion on expression of genes at 8C3/C4. (A) Upregulation of gene expression relative to WT levels (for each gene set at 1) in strains with the human β-globin LCR integrated at the *Rad23a* locus (orange). Expression levels were determined in E14.5 fetal liver using qRT-PCR. Error bars depict 95% confidence intervals obtained from a Student's t-test using the Welch-Satterthwaite approximation for the degrees of freedom. P-values of significant differences in expression levels measured between the two oppositely LCRs are shown above the relevant genes. **(B)** Upregulation of gene expression relative to WT levels (for each gene set at 1) in E14.5 fetal brain containing the human β-globin LCR. Expression levels were determined using qRT-PCR. Error bars depict 95% confidence intervals. **(C)** Binding of CTCF at HS5 of the human β-globin LCR (hHS5), as determined by ChIP-analysis in E14.5 fetal liver. Enrichments were normalized to the endogenous β-major promoter. Endogenous β-globin HS5 (mHS5) is shown as a control. Error bars represent SE of at least two independent experiments.

this is the only gene in the region that is specifically expressed in erythroid cells. This suggests that upregulation of genes by the LCR is not determined by a shared functional relationship. Finally, the gene at position -4 *(Lyl1)* was significantly upregulated only by the LCR integrated in the sense (S) orientation, whereas the gene at position +3 (*Syce2*) responded significantly only to LCR-AS.

4.4.3- Nuclear repositioning of 8C3/C4 by the integrated LCR

The LCR has been implicated in nuclear repositioning of loci and we therefore analysed whether insertion of the LCR had an effect on the position of 8C3/C4 relative to the CT, the nuclear periphery, transcription factories and splicing speckles. Cryo-FISH experiments, as described in **Chapter 3**, were performed to measure the position of the locus relative to its CT in LCR-S and LCR-AS E14.5 liver and brain cells. In E14.5 liver cells, where the LCR is active, the BAC signal had a similar dot-like appearance to wild-type cells, despite the insertion of the 21 kb transgenic fragment (**Figure 4-3A**). The distances between the centre of each locus and the nearest CT edge were then measured. To better assess for differences between WT and transgenic loci, a more quantitative measurement was performed than in **Chapter 3** (see **Figure 3-5A**). As observed previously for other actively transcribed gene clusters [23], WT 8C3/C4 in fetal liver was already found mostly near the edge of its CT (61% within 0.2 μm of the edge), and 7% was >0.4 μm away from the edge, looping out from the CT (**Figure 4-3B**). Insertion of the LCR in either orientation caused a highly significant shift of the distribution of distances between 8C3/C4 and the CT edge towards a more external position ($p<0.001$; K-S test). The percentage of loci found looped out at >0.4 μm away from the CT increased from 7% in wild-type to 19% in LCR-S and 17% in LCR-AS mice (**Figure 4-3B**). Thus, the LCR induces relocation of 8C3/C4 away from its CT, independently of its orientation in the locus. In wild-type brain cells, 8C3/C4 is also found at the periphery of its CT but does not relocate after LCR insertion (**Figure 4-3C**). As the LCR is inactive in brain cells, these results suggest that the repositioning of the locus in liver cells is dependent on LCR activity. The repositioning of 8C3/C4 relative to its CT does not reflect a change in radial nuclear position, as both wild-type and LCR-integrated loci occupy the same preferred radial position away from the nuclear periphery in fetal liver (data not shown). We next performed a cryo-immuno-FISH experiment to determine whether the LCR increased association of 8C3/C4 with splicing speckles, the nucleoplasmic regions enriched for splicing factors [24]. In **Chapter 3** we reported that the WT region was

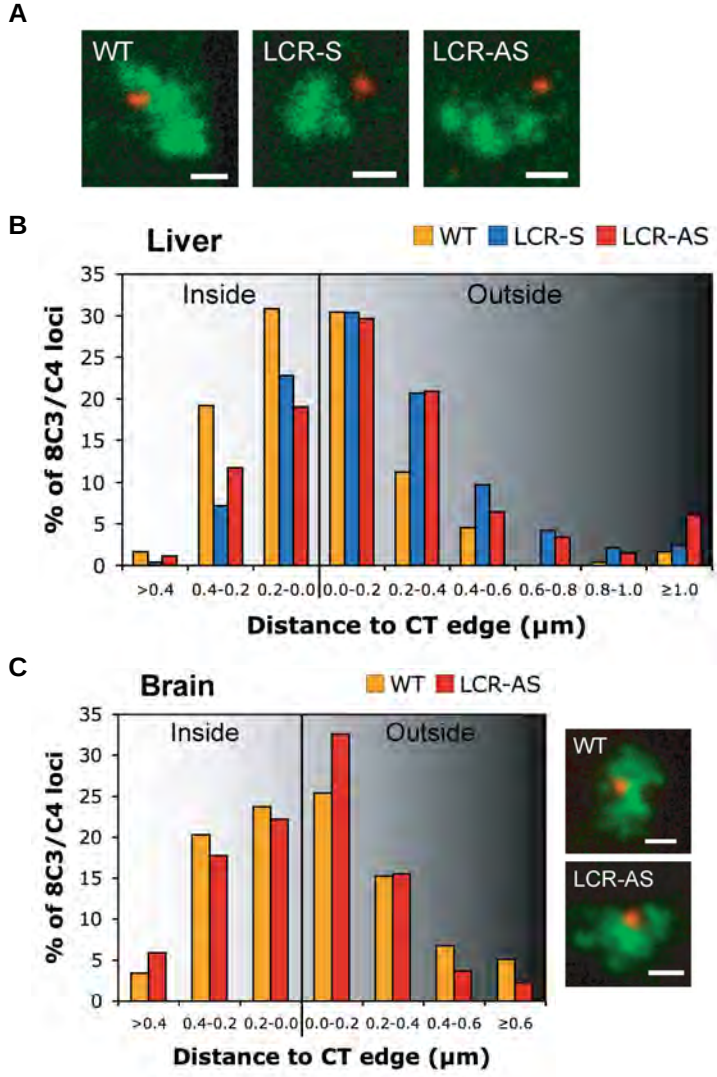


Figure 4-3. Effect of LCR insertion on the positioning of 8C3/C4 in relation to its CT. (A) Cryosections from WT, LCR-S and LCR-AS fetal liver nuclei were hybridized to a BAC probe containing the 8C3/C4 locus (red) and a chromosome 8 specific paint (green). Bars, 1 μm. **(B)** After cryo-FISH, the distances from the center of each locus to the nearest CT edge were measured (n>237 loci). Statistically significant differences were found between the wild-type locus and both transgenic loci containing the LCR ($p<0.001$, K-S test). **(C)** Distances between 8C3/C4 and the CT edge were measured for sections from WT (n=59 loci) and LCR-AS (n=135 loci) fetal brain, showing no differences between the two populations ($p=0.45$, K-S test). Bars, 1 μm.

often adjacent to (63%) or overlapping with (21%) splicing speckles labeled with the Sm antigen (**Figure 3.5B**, **Figure 4-4A, B**). Integration of the LCR had no effect on interaction frequencies with nuclear speckles, which may be due to the already high interaction frequency of the WT locus. As splicing speckles are located outside CTs, the preferred association of 8C3/C4 with these domains is consistent with the favoured position of the locus towards the periphery or outside of its CT. We then analysed whether the LCR influenced the association of 8C3/C4 with foci containing the transcriptionally active (serine2-phosphorylated) form of RNAP II, which marks transcription factories as determined after co-localization with sites of Br-UTP incorporation [15, 25]. Using cryo-immuno-FISH to simultaneously visualize the elongating isoform of RNAP II and 8C3/C4, we found that integration of the β-globin LCR had a small positive effect on the frequency of association of 8C3/C4 with RNAP II foci. Visual association, meaning that signals overlapped or touched (without background pixels in between; **Figure 4-5A**) increased from 51% to 61% in wild-type versus LCR-S +/+ fetal liver sections ($p = 0.03$, Fisher's test using pooled data from two independent hybridisations).

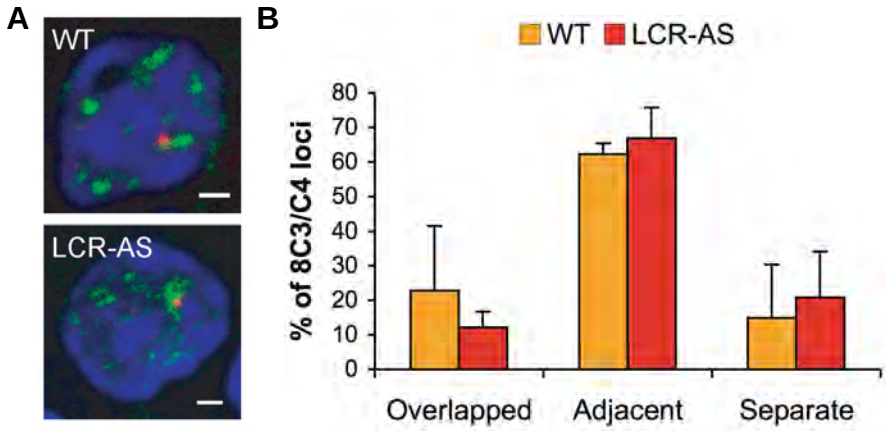


Figure 4-4. Insertion of the LCR has no effect on the association of 8C3/C4 with splicing speckles. (A) Nuclear sections (blue) from WT and LCR-AS livers were immunolabeled with an antibody against the Sm antigen (green), hybridized to the 8C3/C4 BAC probe (red). Bars, 1 μm. **(B)** After cryo-FISH, the locus position in relation to splicing speckles was scored (n>56 loci), showing there is no significant difference between the two populations (p=0.33, chi-squared test).

To more carefully determine the association of 8C3/C4 with factories in an unbiased fashion, we performed distance measurements between the centres of the fluorescent signals (**Figure 4-5B**). The two distance distributions differed significantly, with loci in LCR-S cells being closer to transcription factories than those in WT cells (p=0.003, K-S test), confirming the previous results. The largest difference observed in individual distance categories was for loci closest to a factory (<0.2 μm). In wild-type cells 15% of the loci were scored in this category, while this percentage doubled to 29% in LCR-S cells. To test whether these frequencies may also be explained by random, non-functional associations with RNAP II foci, we performed an in silico experiment. We used the experimental images of nuclei labeled with RNAP II and applied a computational algorithm to generate a randomly positioned locus within each nuclear section. We subsequently measured the distance to the nearest factory and found that the overall distribution of randomly positioned loci was very different from the distributions measured for the wild-type and the LCR-containing locus (**Figure 4-5B**; p<0.001, K-S test), which were more frequently close to the factories. These results suggest that the measured associations between 8C3/C4 loci and transcription factories reflect functionally significant interactions that are increased in the presence of the LCR. The increased factory association may reflect the increased overall transcriptional activity of the region and/or be a consequence of the capacity of the LCR to recruit RNAP II. Given the concomitant relocation of the locus away from the CT and its increased association with RNAP II in the presence of the LCR, we tested whether LCR-mediated looping out from the CT was driving the increased association with RNAP II foci. We simultaneously visualized 8C3/C4, the territory of chromosome 8 and the active form of RNAP II in cryosections (**Figure 4-5C**). Distances between 8C3/C4 and, respectively, the CT edge and RNAP II foci were measured, and RNAP II association frequencies (<0.2 μm) were scored for loci inside, outside and at the edge (±0.2 μm) of the CT (**Figure 4-5D**). The data showed that in the presence of the LCR, the association frequency of 8C3/C4 with RNAP II foci increases independent of its position relative to the CT (**Figure 4-5D**). Logistic regression analysis shows that the increased association of 8C3/C4 with RNAP II is not dependent on CT position (p=0.26), but solely on the presence of the LCR (p=0.0002). Logistic regression also indicates that the apparently smaller effect of the LCR insertion for loci inside the CT is not statistically different from effects at the edge or outside the CT, although we cannot exclude that a regional effect was not observed due to the small number of loci observed in this region. Additionally, we found no correlation

between distance to the CT edge and distance to transcription factories for both WT (R=0.01) and LCR-S (R=0.08) cells. The data suggest that LCR-induced looping out of 8C3/C4 from the CT is not a prerequisite for more frequent association with RNAP II foci. This is in agreement with the observation that active transcription factories are present within CTs, thus transcribing loci that are positioned inside CTs [16, 26-28].

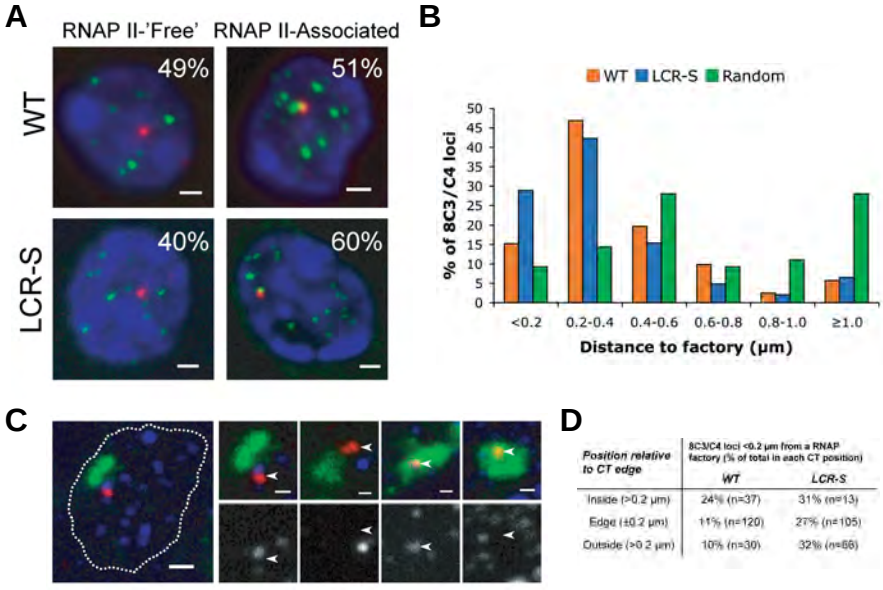


Figure 4-5. Effect of LCR insertion on the association of 8C3/C4 with transcription factories. **(A)** Nuclear sections from WT and LCR-S fetal livers (blue) were immunolabeled with an antibody against the Ser2-phosphorylated form of RNAP II (green), hybridized to the 8C3/C4 BAC probe (red), and the association of the locus with transciption factories was scored (n>231 loci). Bars, 1 μm. **(B)** To obtain an unbiased measurement of factory association, distances were measured between the centres of 8C3/C4 loci and the nearest RNAP II focus, showing significant differences in the distances distributions between WT and LCR-S loci (p=0.003, K-S test). Both loci are found in closer proximity to transcription factories than predicted by a random model (p<0.001, K-S test), in which loci were randomly placed inside experimental images of RNAP II-labeled nuclei and distances measured. **(C)** The active form of RNAP II (blue or grayscale), 8C3/C4 (red; arrows) and chromosome 8 (green) were labeled as before. The nucleus is outlined by a white line. Bars, 1 μm. **(D)** Distances of 8C3/C4 loci to the nearest CT edge and transcription factory were measured after performing the triple labeling described above (C) in sections from WT and LCR-S fetal livers. The frequency of loci <0.2 μm away from a transcription factory was calculated within each position relative to the CT.

4.4.4- Chromatin organization at 8C3/C4 in the presence of an integrated LCR

We next investigated whether LCR-mediated upregulation of gene expression was caused by the spreading of an epigenetic signal from the LCR to the neighbouring genes. In the human growth hormone cluster, the spreading of histone H3 acetylation from its LCR towards the target genes has been suggested to underlie the enhancement of their expression [3, 29] and to be responsible for the bystander activation of an unrelated gene present in the region [30], but the mechanism of such spreading is unknown. Recently it was also suggested that the β-globin LCR recruits an MLL2-containing protein complex, which after LCR-binding, would dissociate to allow for the spreading of the histone methyltransferase MLL2 and subsequent H3K4 methylation specifically at the active gene promoter [6].

Both acetylation of histone H3 and di- and trimethyl H3K4 are associated with active chromatin and on a genome-wide basis the levels of trimethyl H3K4 appear to correlate with transcriptional activity [31, 32]. Therefore we analysed whether the LCR had an effect on histone acetylation levels at 8C3/C4. We found a marked change in the level of acetylation at a site immediately downstream of the integrated LCR-S, and no or minor changes at 12 other positions in the locus (**Figure 4-6A**). Similar observations were made when we analysed H3K4 trimethylation across 8C3/C4, although here we noticed that increased levels of H3K4 tri-methylation corresponded with elevated expression levels at some genes (-4, +4, +6) (**Figure 4-6B**). In general though, the minor changes observed in the presence of the LCRs in acetylation and trimethylation levels of histone H3 appeared not to be strictly related to each other or to changes in transcriptional activity. Most likely, the changes in gene expression levels (2 to 6-fold) are too subtle to be reflected by changes in histone modification patterns. To analyse whether our data were influenced by differences in histone H3 occupancy, we performed ChIP with an antibody against the C-terminus of histone H3 against a subset of sites in WT and transgenic livers (**Figure 4-6C**), but only relatively small differences were observed between WT and transgenic livers.

Our data do not provide evidence for spreading of histone H3 acetylation or H3K4 trimethylation from the integrated LCRs. This is in agreement with the observation that deletion of the endogenous β-globin LCR has little effect on histone H3 acetylation patterns elsewhere in the β-globin locus [33]. We argue that the complex pattern of transcriptional up-regulation observed with the two LCRs is difficult to explain by a

mechanism involving the linear spreading of a signal from the LCR.

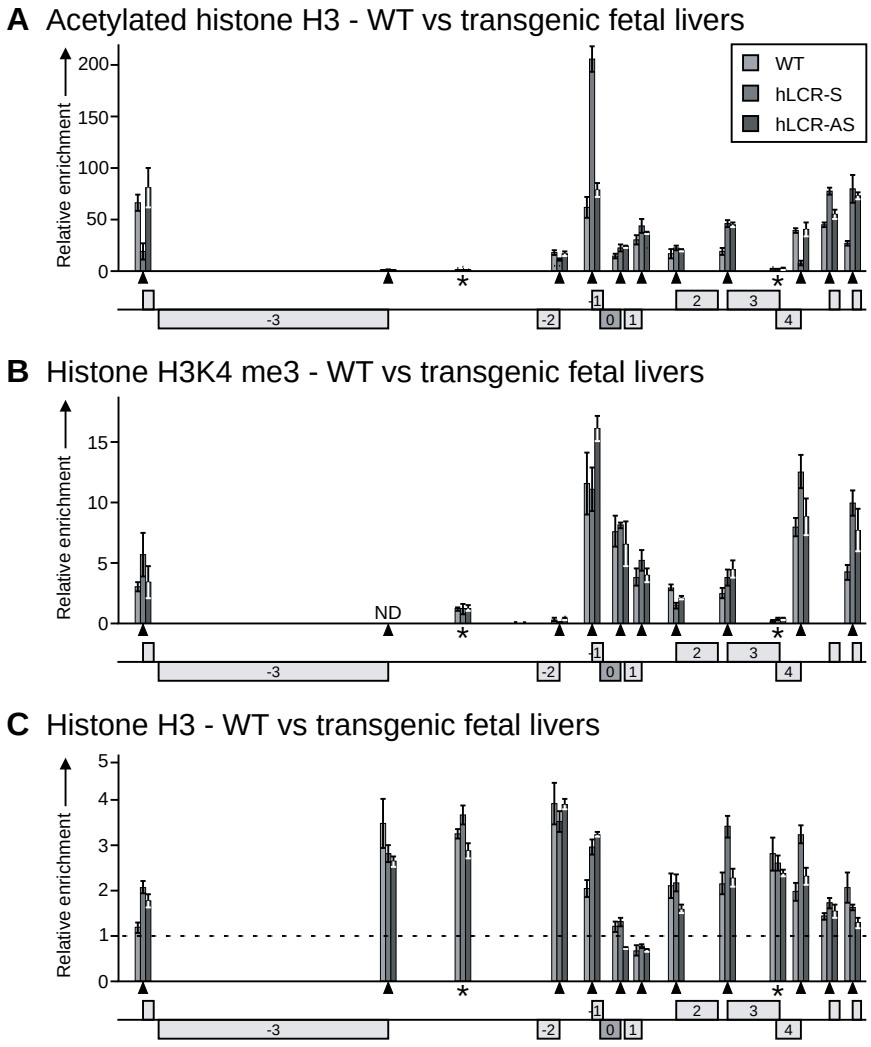


Figure 4-6. Chromatin organization at 8C3/C4 in the presence of the LCR. (A) Histone H3 acetylation at promoters (▲) and other regions (*) of 8C3/C4 in WT and transgenic strains in E14.5 fetal liver. Enrichments were normalized to the Amylase promoter. **(B)** Histone H3 lysine 4 tri-methylation at 8C3/C4 in WT and transgenic strains in E14.5 fetal liver. Enrichments were normalized to the endogenous β-major promoter. **(C)** Histone H3 occupancy at 8C3/C4 in WT and transgenic strains in E14.5 fetal liver. Enrichments were normalized to the endogenous β-major promoter. Error bars in all graphs represent standard error (SE) of at least two independent experiments.

Chromatin looping was previously observed at the human and mouse β-globin locus, where the LCR was found to specifically contact the actively transcribed β-globin genes [7, 8, 17]. To investigate whether increased expression levels were due to looping of the integrated LCR with specific genes at 8C3/C4, 3C technology (Chromosome Conformation Capture) was applied [18, 34], using a recently developed Taqman approach for a more accurate detection of crosslinked ligation products [19]. We focused on a *Rad23a*

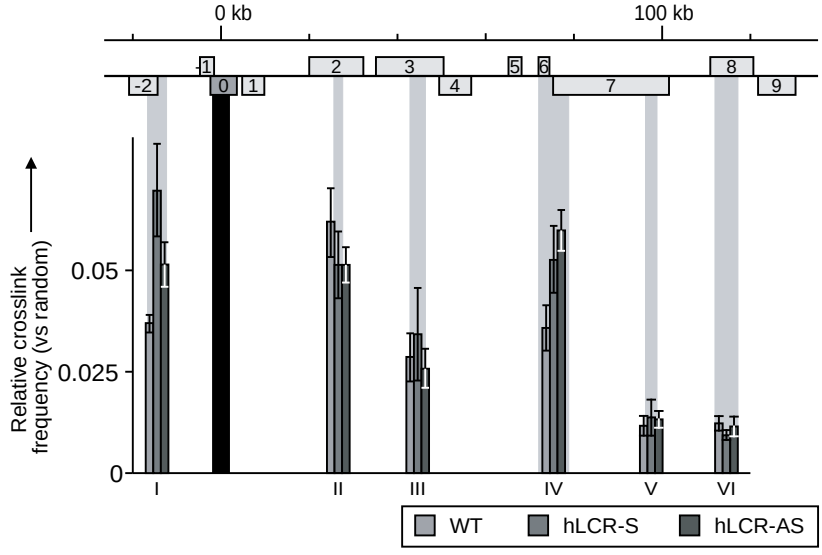


Figure 4-7. Chromatin looping at 8C3/C4 in the presence of the LCR. Relative crosslink frequencies (vs. random) in a part of the 300 kb region containing the two most upregulated genes (genes -2, *Dand5* and +6, *Dnase2a*). Crosslink frequencies between the fixed *BglII* fragment, depicted by a black bar, containing either the WT *Rad23a* gene or the LCR and other fragments, depicted by gray bars are visualized, as determined by 3C analysis and quantified by qPCR in WT and transgenic E14.5 fetal liver. Note that for the fixed fragment in each strain the same primer and Taqman probe combination was used. Error bars represent SE of at least three independent experiments.

restriction fragment that spanned the integration site and designed one primer/probe combination that could be used for the analysis of interactions with the *Rad23a* gene in wild-type fetal liver cells, as well as for the analysis of interactions with the integrated LCR in transgenic cells (**Figure 4-7**). We found that upon introduction of the LCR, interaction frequencies increased specifically between this site and two fragments that contained the promoters of the two genes that were most activated by the integrated

LCRs, suggesting that the LCR-dependent upregulation of these genes is mediated by chromatin looping (**Figure 4-7**). None of the other restriction fragments analysed across 8C3/C4 showed a significant difference in interaction frequency between the wild-type and transgenic loci, not even when the fragments analysed were at the promoters of up-regulated genes. This is not unexpected, as the LCR has been shown to contact only one gene at a time [35], such that increased contacts with the large number of target genes at 8C3/C4 will average out and be below the threshold of detection in the population of cells analysed. In combination with previous observations made in the β-globin locus [7, 8, 17], we conclude that physical interactions between the LCR and at least the genes at +6 (*DNaseII*) and -2 (*Dand5*) are likely to be key to their transcriptional enhancement.

4.5- Discussion

We have investigated the effect of the integration of the β-globin LCR into a gene-dense region containing many housekeeping genes. Two transgenic mouse lines were generated, each containing an oppositely oriented LCR inserted at the same genomic position. Gene activity, chromatin modifications and chromatin structure were analysed across a large genomic region containing the integration site. In addition, detailed studies were performed on the position of the locus relative to its chromosome territory, the nuclear periphery and nuclear entities such as speckles and transcription factories. To our knowledge, this is the first study that analyses in detail the impact of an integrated LCR on its surrounding genes and chromatin. As the wild-type region is already highly active, some of the changes induced by the LCR are subtle. More profound effects may be expected if the LCR is integrated at less active genomic positions. Nevertheless, our results demonstrate that the LCR can act on a relatively large number of genes (at least seven) spread over at least 150 kb, causes nuclear repositioning of the targeted locus and physically contacts the most upregulated genes via chromatin looping. Collectively, the data provide an integrated view on LCR functioning and on the various levels that control gene expression within a gene-dense region.

Gene activation has been correlated with repositioning of loci away from nuclear landmarks such as CTs, pericentromeric heterochromatin or the nuclear periphery. One interpretation of this phenomenon is that relocation critically drives transcription by positioning loci in nuclear zones of increased transcriptional competence. This idea may predict some degree of co-regulation between genes closely juxtaposed on the chromosome template [36]. We showed that introduction of the β-globin LCR into a gene-dense region of functionally unrelated genes caused (1) the more frequent positioning of this locus away from its CT, (2) a small but significant increase in association frequency with transcription factories, (3) no change in association frequency of the locus with splicing speckles and no change in the nuclear radial position. We also found an increase in expression levels of multiple genes surrounding the integration site. LCR-driven repositioning of the locus may facilitate increased transcriptional activity. Nevertheless, we consider it unlikely that the subset of cells in which 8C3/C4 is looped out from the CT can account for the overall increased levels of expression. Transcription also takes place at the interior of CTs [16, 23, 26, 27] and we show here that 8C3/C4 association with transcription factories increases both for loci at the edge and outside the CT. Moreover and as discussed below, our data demonstrate that at least for some genes repositioning

is not sufficient to drive increased transcriptional activity.
Transcriptional enhancement was maximally 4-6 fold, which is modest compared to the impact the mouse LCR has at its endogenous position, where it increases β-major gene expression 25-100 fold [2]. An explanation for this difference could be that more genes compete for interaction with the LCR at 8C3/C4. Furthermore, gene promoter intrinsic properties could preclude that the LCR increases their expression more efficiently. In agreement with this notion is the fact that both the mouse and human β-globin locus contain a second adult β-globin-like gene (called β-minor and δ-globin, respectively) that is activated much less dramatically than their prototypic counterparts (β-major and β-globin). For example, mouse β-minor expression is 10-fold less than β-major expression and the difference between human δ- and β-globin gene expression is even more pronounced.
When we integrated an oppositely oriented LCR at the same genomic position at 8C3C4, it had an identical impact on the positioning of the locus relative to these nuclear landmarks, but upregulated a partially overlapping yet different set of genes. Previously it has been shown that tissue-specific genes surrounded by housekeeping genes maintain their inactive status upon nuclear repositioning in unrelated tissues [37]. This observation may be explained by the absence of cell-type specific transcription factors in these tissues [38]. In another recent study, *Lnp*, a gene located near the *Hoxd* cluster, was shown to be active and not change its expression level upon looping away from its CT during ES cell differentiation [28]. We observed a similar phenomenon at 8C3/C4, where genes at position -6 (*Btbd14b*), -5 (*Trmt1*), -3 (*Nfix*), +2 (*Farsla*) and +5 (*Klf1*) did not change their expression. An explanation for this may be that gene or chromatin intrinsic properties preclude the more efficient transcription of these genes. Importantly, the genes at position -4 *(Lyl1)* and +3 (*Syce2*), and to a less significant extent -7 (*Ier2*) increased their transcription activity, but responded to only one orientation of the LCR (as confirmed independently by Affymetrix micro-array expression analysis (data not shown)). This shows that at least for these genes repositioning is not sufficient to drive their upregulation and that other mechanisms are involved.
Transcription regulation by the β-globin LCR has also been associated with the spreading of histone acetylation (in case of the LCR of the human growth hormone cluster) [3] and of the methyltransferase MLL2 [6]. We find the same punctuated pattern of acetylated and lysine-4 methylated histone H3 specifically at the active promoters of 8C3/C4 before and after LCR integration and find no evidence for spreading of these marks. The data do

not exclude that the LCR attracts MLL2 and subsequently spreads the methyltransferase across the locus. However, the finding that genes distant to the LCR are upregulated, while genes more proximal are not, is difficult to explain by a mechanism involving the linear spreading of any signal from the LCR. If spreading occurs, it also not clear why a gene like *Syce2* (+3) is upregulated by the LCR in one, but not the other, orientation.
Previously, we have shown that the β-globin LCR contacts the active genes in the β-globin locus [8, 17]. Here, we find such contacts being formed between the LCR and the most strongly upregulated genes, suggesting that the ectopic LCR also acts by looping. Looping is thought to result from random collisions between chromatin sites that are stabilized when proteins bound to these sites have affinity for each other [56]. Alternatively, looping could also be the outcome of the LCR tracking along the intervening chromatin fibre towards gene promoters. The latter seems less compatible with our observation that LCR-S fails to activate *Syce2* at position +3 (a gene that is upregulated by LCR-AS), while being able to upregulate the genes at position +4 and +6 further down the chromatin fibre. Spatial proximity of the LCR with active genes is thought to increase the local concentration of transcription factors and/or RNA polymerase II, which might allow for more efficient transcription [39, 40]. Productive loop formation is predicted to depend on affinities between trans-acting factors bound to the LCR and to the gene. Thus, at 8C3/C4, the genes at position -6 (*Btbd14b*), -5 (*Trmt1*), -3 (*Nfix*), +2 (*Farsla*) and +5 (*Klf1*) that are not upregulated in the presence of an LCR would lack proteins that can interact with LCR-associated factors. This is especially surprising for gene +5 (*Klf1*), since it encodes an erythroid specific transcription factor, and therefore could be expected to bind similar trans-acting factors to the promoter as the LCR. We notice, however, that an erythroid-specific enhancer less than 1 kb upstream of the promoter acts on this gene and may compete out the LCR. Why would the productive formation of a loop be dependent on the orientation of the LCR, as was found for some, but not all, genes at 8C3/C4? The LCR encompasses more than 20 kb of DNA and is asymmetric with regard to nucleotide sequence and transcription factor binding sites. One explanation may therefore be that physical constraints intrinsic to the chromatin fibre allow one, but not the other LCR to correctly juxtapose itself relative to a given gene. Interestingly, the genes that respond to only a single LCR are always located downstream of that particular LCR. We found CTCF bound to the outer hypersensitive site 5 and we propose that CTCF-mediated loops may interfere with contacts between the LCR and some of the upstream genes.

How to interpret the looping out of 8C3/C4 from the CT particularly when it contains an integrated LCR? One possibility is that this reflects increased local decondensation of chromatin after the insertion of 5 additional (erythroid-specific) DNase I hypersensitive sites, which may result in increased mobility [41, 42]. Regulatory DNA elements, such as the LCR, serve as binding platforms for trans-acting factors that locally disrupt the nucleosome fibre, as revealed by DNaseI hypersensitivity, and cause decondensation of the region. Indeed, live cell imaging studies that measured the compaction of a transgene array demonstrated that decondensation was not dependent on transcription but was dictated by the binding of transcriptional activator proteins [43]. A region-wide increase in accessibility may facilitate the regulation of individual genes and the simultaneous increase in mobility of the locus may promote the interaction with RNAPII foci. The collective stabilization of a decondensed chromatin state could explain why housekeeping genes tend to cluster in the genome [38].

4.6- Acknowledgements

We thank Frank Grosveld for discussion, Francisco Ramirez for the logistic regression statistical testing, Bert van der Horst for providing targeting constructs and Robert-Jan Palstra and other members of the laboratories for assistance.

4.7- References

1. Epner, E., Reik, A., Cimbora, D., Telling, A., Bender, M.A., Fiering, S., Enver, T., Martin, D.I., Kennedy, M., Keller, G., and Groudine, M. (1998). The beta-globin LCR is not necessary for an open chromatin structure or developmentally regulated transcription of the native mouse beta-globin locus. Mol Cell *2*, 447-455.
2. Grosveld, F., Blom van Assendelft, G., Greaves, D.R., and Kollias, G. (1987). Position-independent, high-level expression of the human beta-globin gene in transgenic mice. Cell *51*, 975-985.
3. Ho, Y., Elefant, F., Cooke, N., and Liebhaber, S. (2002). A defined locus control region determinant links chromatin domain acetylation with long-range gene activation. Mol Cell *9*, 291-302.
4. Routledge, S.J., and Proudfoot, N.J. (2002). Definition of transcriptional promoters in the human beta globin locus control region. J Mol Biol *323*, 601-611.
5. Ling, J., Ainol, L., Zhang, L., Yu, X., Pi, W., and Tuan, D. (2004). HS2 enhancer function is blocked by a transcriptional terminator inserted between the enhancer and the promoter. J Biol Chem *279*, 51704-51713.
6. Demers, C., Chaturvedi, C.P., Ranish, J.A., Juban, G., Lai, P., Morle, F., Aebersold, R., Dilworth, F.J., Groudine, M., and Brand, M. (2007). Activator-mediated recruitment of the MLL2 methyltransferase complex to the beta-globin locus. Mol Cell *27*, 573-584.
7. Carter, D., Chakalova, L., Osborne, C.S., Dai, Y.F., and Fraser, P. (2002). Long-range chromatin regulatory interactions in vivo. Nat Genet *32*, 623-626.
8. Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F., and de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. Mol Cell *10*, 1453-1465.
9. Ragoczy, T., Telling, A., Sawado, T., Groudine, M., and Kosak, S.T. (2003). A genetic analysis of chromosome territory looping: diverse roles for distal regulatory elements. Chromosome Res *11*, 513-525.
10. Francastel, C., Walters, M.C., Groudine, M., and Martin, D.I. (1999). A functional enhancer suppresses silencing of a transgene and prevents its localization close to centrometric heterochromatin. Cell *99*, 259-269.
11. Ng, J.M., Vermeulen, W., van der Horst, G.T., Bergink, S., Sugasawa, K., Vrieling, H., and Hoeijmakers, J.H. (2003). A novel regulation mechanism of DNA repair by damage-induced and RAD23-dependent stabilization of xeroderma pigmentosum group C protein. Genes Dev *17*, 1630-1645.
12. Giresi, P.G., Kim, J., McDaniell, R.M., Iyer, V.R., and Lieb, J.D. (2007). FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. Genome Res *17*, 877-885.
13. Heath, H., de Almeida, C.R., Sleutels, F., Dingjan, G., van de Nobelen, S., Jonkers, I., Ling, K.W., Gribnau, J., Renkawitz, R., Grosveld, F., Hendriks, R.W., and Galjart, N. (2008). CTCF regulates cell cycle progression of alphabeta T cells in the thymus. Embo J *27*, 2839-2850.
14. Guillot, P.V., Xie, S.Q., Hollinshead, M., and Pombo, A. (2004). Fixation-induced redistribution of hyperphosphorylated RNA polymerase II in the nucleus of human cells. Exp Cell Res *295*, 460-468.
15. Pombo, A., Jackson, D.A., Hollinshead, M., Wang, Z., Roeder, R.G., and Cook, P.R. (1999). Regional specialization in human nuclei: visualization of discrete sites of transcription by RNA polymerase

III. Embo J *18*, 2241-2253.

16. Branco, M.R., and Pombo, A. (2006). Intermingling of Chromosome Territories in Interphase Suggests Role in Translocations and Transcription-Dependent Associations. PLoS Biol *4*, e138.
17. Palstra, R.J., Tolhuis, B., Splinter, E., Nijmeijer, R., Grosveld, F., and de Laat, W. (2003). The beta-globin nuclear compartment in development and erythroid differentiation. Nat Genet *35*, 190-194.
18. Splinter, E., Grosveld, F., and de Laat, W. (2004). 3C technology: analyzing the spatial organization of genomic loci in vivo. Methods Enzymol *375*, 493-507.
19. Splinter, E., Heath, H., Kooren, J., Palstra, R.J., Klous, P., Grosveld, F., Galjart, N., and de Laat, W. (2006). CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. Genes Dev *20*, 2349-2354.
20. Tanimoto, K., Liu, Q., Bungert, J., and Engel, J.D. (1999). Effects of altered gene order or orientation of the locus control region on human beta-globin gene expression in mice. Nature *398*, 344-348.
21. Wai, A.W., Gillemans, N., Raguz-Bolognesi, S., Pruzina, S., Zafarana, G., Meijer, D., Philipsen, S., and Grosveld, F. (2003). HS5 of the human beta-globin locus control region: a developmental stage-specific border in erythroid cells. Embo J *22*, 4489-4500.
22. Defossez, P.A., Kelly, K.F., Filion, G., Perez-Torrado, R., Magdinier, F., Menoni, H., Nordgaard, C.L., Daniel, J.M., and Gilson, E. (2005). The human enhancer-blocker CTCF interacts with the transcription factor kaiso. J Biol Chem *280*, 43017 - 43023.
23. Scheuermann, M.O., Tajbakhsh, J., Kurz, A., Saracoglu, K., Eils, R., and Lichter, P. (2004). Topology of genes and nontranscribed sequences in human interphase nuclei. Exp Cell Res *301*, 266-279.
24. Xie, S.Q., Martin, S., Guillot, P.V., Bentley, D.L., and Pombo, A. (2006). Splicing speckles are not reservoirs of RNA polymerase II, but contain an inactive form, phosphorylated on serine2 residues of the C-terminal domain. Mol Biol Cell *17*, 1723-1733.
25. Grande, M.A., van der Kraan, I., de Jong, L., and van Driel, R. (1997). Nuclear distribution of transcription factors in relation to sites of transcription and RNA polymerase II. J Cell Sci *110 (Pt 15)*, 1781-1791.
26. Abranches, R., Beven, A.F., Aragon-Alcaide, L., and Shaw, P.J. (1998). Transcription sites are not correlated with chromosome territories in wheat nuclei. J Cell Biol *143*, 5-12.
27. Verschure, P.J., van Der Kraan, I., Manders, E.M., and van Driel, R. (1999). Spatial relationship between transcription sites and chromosome territories. J Cell Biol *147*, 13-24.
28. Morey, C., Da Silva, N.R., Perry, P., and Bickmore, W.A. (2007). Nuclear reorganisation and chromatin decondensation are conserved, but distinct, mechanisms linked to Hox gene activation. Development *134*, 909-919.
29. Elefant, F., Su, Y., Liebhaber, S.A., and Cooke, N.E. (2000). Patterns of histone acetylation suggest dual pathways for gene activation by a bifunctional locus control region. Embo J *19*, 6814-6822.
30. Cajiao, I., Zhang, A., Yoo, E.J., Cooke, N.E., and Liebhaber, S.A. (2004). Bystander gene activation by a locus control region. Embo J *23*, 3854-3863.
31. Bernstein, B.E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D.K., Huebert, D.J., McMahon, S., Karlsson, E.K., Kulbokas, E.J., 3rd, Gingeras, T.R., Schreiber, S.L., and Lander, E.S. (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. Cell *120*, 169-181.
32. Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. Cell *130*, 77-88.

33. Schubeler, D., Francastel, C., Cimbora, D.M., Reik, A., Martin, D.I., and Groudine, M. (2000). Nuclear localization and histone acetylation: a pathway for chromatin opening and transcriptional activation of the human beta-globin locus. Genes Dev *14*, 940-950.
34. Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. Science *295*, 1306-1311.
35. Wijgerde, M., Grosveld, F., and Fraser, P. (1995). Transcription complex stability and chromatin dynamics in vivo. Nature *377*, 209-213.
36. Chakalova, L., Debrand, E., Mitchell, J.A., Osborne, C.S., and Fraser, P. (2005). Replication and transcription: Shaping the landscape of the genome. Nat Rev Genet *6*, 669 - 677.
37. Brown, J.M., Leach, J., Reittie, J.E., Atzberger, A., Lee-Prudhoe, J., Wood, W.G., Higgs, D.R., Iborra, F.J., and Buckle, V.J. (2006). Coregulated human globin genes are frequently in spatial proximity when active. J Cell Biol *172*, 177-187.
38. Sproul, D., Gilbert, N., and Bickmore, W.A. (2005). The role of chromatin structure in regulating the expression of clustered genes. Nat Rev Genet *6*, 775-781.
39. Droge, P., and Muller-Hill, B. (2001). High local protein concentrations at promoters: strategies in prokaryotic and eukaryotic cells. Bioessays *23*, 179-183.
40. de Laat, W., and Grosveld, F. (2003). Spatial organization of gene expression: the active chromatin hub. Chromosome Res *11*, 447-459.
41. Volpi, E.V., Chevret, E., Jones, T., Vatcheva, R., Williamson, J., Beck, S., Campbell, R.D., Goldsworthy, M., Powis, S.H., Ragoussis, J., Trowsdale, J., and Sheer, D. (2000). Large-scale chromatin organization of the major histocompatibility complex and other regions of human chromosome 6 and its response to interferon in interphase nuclei. J Cell Sci *113 (Pt 9)*, 1565-1576.
42. Gilbert, N., Boyle, S., Fiegler, H., Woodfine, K., Carter, N.P., and Bickmore, W.A. (2004). Chromatin architecture of the human genome; gene-rich domains are enriched in open chromatin fibers. Cell *118*, 555-566.
43. Tumbar, T., Sudlow, G., and Belmont, A.S. (1999). Large-scale chromatin unfolding and remodeling induced by VP16 acidic activation domain. J Cell Biol *145*, 1341-1354.

Chapter 5

An ectopic β-globin LCR repositions its chromosomal integration site in the nucleus without searching for functionally related genes

Daan Noordermeer, Marieke Simonis, Petra Klous, Bert Eusen[1], Annelies de Klein[1], Wouter de Laat

[1] Department of Clinical Genetics, Erasmus Medical Center, PO Box 2040, 3000 CA Rotterdam, The Netherlands

Work in progress

5.1- Abstract

Transcriptional regulatory elements can change the nuclear location of their target loci, but the functional relevance of this relocation is poorly understood. The β-globin LCR, a prototype regulatory element, relocates genomic regions versus several nuclear substructures upon activation. In this study, we used transgenic mice containing an ectopically integrated β-globin LCR to study whether relocation versus its Chromosome Territory (CT) is a consequence of the LCR searching for preferred partners in the nuclear space. We first setup a transvection-like system, but we found no evidence that the LCR can actively establish interactions with one of its natural target genes present on the homologous allele. Subsequently, we performed 4C to investigate whether the LCR searches for other preferred partners located elsewhere in the genome. LCR-induced positioning outside the CT results in increased encounters with genomic regions located on different chromosomes, but the specific interacting partners *in cis* and *in trans* are largely identical, irrespective of whether the LCR is integrated or not. The LCR therefore does not scan the nucleus for preferred interaction partners. Two existing interaction partners *in trans* containing erythroid specific genes did significantly increase their co-localization frequency, while the 4C-signal of the α-globin locus increases. These results may indicate that the LCR stabilizes some existing interactions with functionally related genes. Future experiments to further address this issue are being discussed.

5.2- Introduction

Nuclear organization is an emerging contributor to genomic function and gene regulation. Genomic loci occupy non-random, probabilistic locations versus nuclear substructures like the periphery, their chromosome territory (CT), pericentromeric heterochromatin and versus other genomic loci [1-3]. Transcriptional regulatory elements are involved in nuclear location of their target loci [4, 5]. The β-globin Locus Control Region (LCR), a strong erythroid-specific regulatory element that enhances expression of the β-globin genes 25-100 fold, has been shown to reposition the endogenous mouse β-globin locus away from the nuclear periphery [6]. Furthermore, it can relocate linked transgenes away from centromeric heterochromatin [7] and, when placed at an ectopic position in the genome, it moves its integration site more frequently outside its chromosome territory (CT) [8]. In one study the LCR was reported to position the endogenous mouse β-globin locus outside the CT prior to β-globin gene activation [9], but later during erythroid differentiation when the locus is active, it is rarely found located away from its CT [10].
The functional relevance of the observed nuclear relocation of genomic loci by regulatory elements is poorly understood. One possibility is that there is no functional relevance but that repositioning is just the consequence of the different chromatin condensation states adopted by active and inactive loci. Another often suggested possibility is that regulatory elements function by migrating their target genes to nuclear zones or bodies that support gene expression, such as transcription factories [11]. These are nuclear bodies that contain increased concentrations of RNA polymerase II. Transcription factories can be visualized by fluorescent microscopy [12] and have been proposed to promote clustering of active genes in the nuclear space. A number of selected erythroid specific genes cluster at high frequencies at transcription factories when active [11, 13], while transiently transfected plasmids with similar transcription units cluster in high numbers at selected transcription factories when actively transcribed [14]. These data suggest the existence of dedicated transcription factories where functionally related genes come together for their transcription. In contrast, based on large-scale mapping of DNA interactions by 4C technology we found no evidence for preferential clustering of erythroid specific genes [15]. We and others suggested that clustering may be dependent on similar activity state of genomic regions, which could actually be mediated by the formation of a nuclear substructure named splicing speckles rather than transcription factories [16, 17].

We recently generated transgenic mice containing the human β-globin LCR site-specifically integrated into a gene-dense region named 8C3/C4 [8]. Previously, we reported that the LCR caused upregulation of genes up to 150 kb away from the integration site. Interestingly, we also found that the LCR induced a highly significant repositioning of the region away from its CT and a mildly increased co-localization with transcription factories [8]. Here, we used these transgenic mice to investigate whether the observed repositioning is a consequence of the LCR searching for functionally related partners in the nuclear space.

5.3- Materials and methods

5.3.1- Gene targeting and generation of transgenic mice

Targeting of the human β-globin LCR to the mouse *Rad23a* gene has been described [8]. The human γ-globin gene was targeted to the mouse *Rad23a* gene by substituting the *Cla*I Neo-resistance cassette from a construct that removes *Rad23a* exon II-VII [18] for a 7.6 kb *Cla*I fragment containing a TK-Neo resistance cassette coupled to a *Cla*I-*Sma*I fragment containing the human *Aγ-globin* gene with a *GFP*-gene at the translational start. Constructs with the *Cla*I fragment in two orientations were obtained: γ-globin-S (5'-γ-globin at the 5'-end of the *Rad23a* gene) and γ-globin-AS (5'- γ-globin gene at the 3'-end of the *Rad23a* gene). Targeting in Ola129-derived ES cells, blastocyst injection and breeding to obtain homozygous transgenic animals in an FvB background was done as described [8, 18]. Genotyping was performed by Southern blot. Animal experiments were carried out according to institutional and national guidelines (Committee on Experiments with Laboratory Animals (DEC-Consult); Ministry of Agriculture, Nature and Food Quality, The Hague, The Netherlands).

5.3.2- Gene expression analysis

Total RNA was isolated using Trizol reagent (Invitrogen) from E10.5 fetal blood or E12.5 livers of two independent embryos of each genotype. cDNA synthesis was performed using SuperScript II Reverse Transcriptase and Oligo(dT)$_{12\text{-}18}$ primer according to the manufacturers' instructions (Invitrogen). Products were quantified by qPCR, using Platinum Taq DNA polymerase (Invitrogen) and SYBR Green (Sigma) on an Opticon 2 Real-Time PCR Detection System (BioRad). Transcript levels were normalized to the *Hprt1* transcript, encoding a housekeeping gene on an unrelated chromosome (primer sequences: human Aγ-globin forward AGGTGCTGACTTCCTTGGG, human Aγ-globin reverse GGGTGAATTCTTTGCCGAA, Hprt1 forward AGCCTAAGATGAGCGCAAGT, Hprt1 reverse ATGGCCACAGGACTAGAACA).

5.3.3- FISH

DNA-FISH and Cryo-FISH was performed as described before [8, 15, 19]. BAC clones (BACPAC Resources Centre) used to visualize genomic regions are listed in **Table 5-1**. BAC probes for DNA–FISH were labeled with SpectrumGreen-dUTP (Green, Vysis) or ChromaTide Texas Red-12-dUTP (Red, Invitrogen). BAC probes for Cryo-FISH were

labeled with SpectrumGreen-dUTP (Green) or ChromaTide Alexa Fluor 594-5-dUTP (Red; Invitrogen). Probe specificity was confirmed on mouse spleen metaphase spreads. 500 ng labeled probe was co-precipitated with 5 μg mouse Cot1 DNA (Invitrogen).

BAC clone	Locus (Cytological band)	Chromosome	Position (Mb) (NCBI assembly m34)
RP23-317H16	7E1, negative control	7	80
RP23-370E12	7F1, β-globin locus	7	98
RP23-32C19	7F3, negative control	7	118
RP23-265I23	7F4	7	121
RP23-27B18	7F5	7	136
RP24-136A15*	8C3/C4	8	84
RP24-319P23**	8C3/C4	8	84
RP23-87K3	10B2	10	41
RP23-127D15	10B2	10	44
RP24-130O14	10B5.3, negative control	10	74
RP23-375D18	11A5, α-globin locus	11	32
RP24-306K19	11B4	11	70
RP24-236L11	11B5	11	78
RP23-311P1	11D	11	102
RP23-258M10	14C2	14	50
RP23-450E9	14D1	14	64
RP24-255K10	15C	15	37

* RP24-136A15 used for DNA-FISH and Cryo-FISH

** RP24-319P23 used for DNA-FISH

Table 5-1 BAC Clones for DNA-FISH and Cryo-FISH

5.3.4- Microscopy

Images were collected with a Zeiss Axio Imager Z1 epifluorescence microscope (100x plan apochromat, 1.4 oil objective) equipped with a charge-coupled device (CCD) camera. DNA-FISH images were analyzed with Zeiss AxioVision software (Zeiss). Cryo-FISH images were analyzed with Isis FISH Imaging System software (Metasystems). Filters used for DNA-FISH: DAPI, (Zeiss) FITC, AF594 (Chroma). Filters used for Cryo-FISH: DAPI, FITC, RD-TR-PE (Zeiss). No bleedthrough was detected and images were collected without saturation of intensities.

5.3.5- 4C analysis

4C analysis was performed as described [15], with *Hin*dIII as a primary restriction enzyme, *Dpn*II (WT) or *Nla*III (LCR-AS) as a secondary frequent cutter and *Pst*I (WT) or *Eco*NI (LCR-AS) as a tertiary restriction enzyme for linearization. Different primer sets were used to amplify fragments ligated to the LCR-AS region (primer sequences: human LCR forward ACACTTTCAGTCCGGTCC, human LCR reverse AGATTTCCTGTTCACTCACTG) and to WT *Rad23a* [15]. Data were highly reproducible between replicates ($\rho = 0.76$; Spearman's rank correlation coefficient) and the different genotypes ($\rho = 0.76$ or more; Spearman's

DNA-FISH

Allele 1	Allele 2	n	co-localized	co-localization %	G-score*
8C3/C4 WT	8C3/C4 WT	159	7	4.40%	-
hLCR-S 8C3/C4	8C3/C4 WT	162	9	5.56%	0.47
hLCR-S 8C3/C4	Aγ-globin-S 8C3/C4	151	9	5.96%	0.79

G-score	significance
> 3.84	P < 0.05
> 6.64	P < 0.01
> 10.83	P < 0.001

* vs 8C3/C4 WT - 8C3/C4 WT

Cryo-FISH

4C Negative regions:

	WT 8C3/C4				8C3/C4 hLCR-AS				
Locus	n	co-localized	co-localization %	G-score*	n	co-localized	co-localization %	G-score*	G-score**
7E1	506	5	0.99%	0.08	253	1	0.40%	1.46	1.18
7F3	503	4	0.80%	0.53	528	5	0.95%	0.09	0.15
10B5.3	508	8	1.57%	0.84	519	8	1.54%	0.93	0.00

4C Positive regions:

	WT 8C3/C4				8C3/C4 hLCR-AS				
Locus	n	co-localized	co-localization %	G-score*	n	co-localized	co-localization %	G-score*	G-score**
7F4	258	11	4.26%	13.44	509	16	3.14%	13.46	1.72
7F5	251	4	1.59%	0.45	254	5	1.97%	1.52	0.21
10B2	252	5	1.98%	1.38	250	6	2.40%	3.05	0.21
10B2	513	12	2.34%	5.24	508	17	3.35%	15.76	1.99
11A5 (α-globin)	1023	54	5.28%	84.12	1019	50	4.91%	75.13	0.29
11B4	509	21	4.13%	24.62	510	16	3.14%	13.42	1.37
11B5	482	22	4.56%	29.18	504	16	3.17%	13.68	2.49
11D	785	35	4.46%	45.15	724	50	6.91%	103.96	8.78
14C2	255	11	4.31%	13.64	256	12	4.69%	17.15	0.08
14D1	509	19	3.73%	19.49	717	51	7.11%	108.70	18.15
15C	251	2	0.80%	0.26	251	2	0.80%	0.20	0.00

Other regions:

	WT 8C3/C4				8C3/C4 hLCR-AS				
Locus	n	co-localized	co-localization %	G-score*	n	co-localized	co-localization %	G-score*	G-score**
7F1 (β-globin)	262	6	2.29%	2.48	534	12	2.25%	5.23	0.00

* G-test versus combined negative controls per genotype

** G-test of 8C3/C4 hLCR-AS vs WT 8C3/C4

Table 5-2 Co-localization frequencies for DNA-FISH and Cryo-FISH

rank correlation coefficient). The arrays used for this study (Nimblegen Systems) cover seven mouse chromosomes [15]. Arrays and analyses were based on NCBI build m34. Interacting regions in cis were identified by applying a running mean approach with a window size of 29 probes and a False Discovery Rate of 5% and positively scored when regions were present in both independent replicates. Interacting regions in trans were identified by applying a running median approach with a window size of 29 probes and a False Discovery Rate of 0% and positively scored when regions were present in both independent replicates.

5.3.6- Statistical analysis

Significance of co-localization between homologous 8C3/C4 alleles (DNA-FISH) and of 8C3/C4 with interchromosomal regions (Cryo-FISH) was determined by applying a replicated goodness-of-fit test (G-statistic) [20]. The null hypothesis in the DNA-FISH experiments was that integration of the LCR or LCR and Aγ-globin gene resulted in similar co-localization frequencies as the WT 8C3/C4 alleles. The null hypothesis in the Cryo-FISH experiments was that associating regions identified in the 4C analysis had colocalization frequencies comparable to the negative regions. The number of scored regions (by a person not knowing the probe combination applied to the sections) and co-localization frequencies are indicated in **Table 5-2**.

5.4- Results

5.4.1- An ectopic human LCR at 8C3/C4 does not interact with a globin target gene in a transvection-like setting

To investigate the capacity of the β-globin LCR to search the nuclear interior for preferred target genes, we first created a classical transvection system. Transvection, which has mainly been studied in *Drosophila*, depends on pairing of homologous chromosomes, which allows an enhancer to upregulate a target gene on the homologous allele [21]. In *Drosophila*, homologous pairing is a common phenomenon most prominently seen at polytene chromosomes. In mammals, one report claims substantial proximity in erythroid cells between homologous chromosomes rich in erythroid-specific genes [22], while others have reported that pairing of homologous X chromosomes plays an important role in X inactivation [23, 24]. The current idea though is, that generally there is no tendency for homologous chromosomes to pair in the mammalian interphase nucleus [25, 26]. While this may be true at the level of entire chromosomes, less is known about the relative proximity between homologous gene loci. Extensive chromosome intermingling has been reported, suggesting that individual gene loci can invade inside the territories of other chromosomes [15, 27].

At the endogenous mouse β-globin locus, the LCR loops the chromatin to form protein-mediated contacts with the active β-globin genes ([28] and **Chapter 2**). In gene competition experiments, the LCR prefers β-globin genes over other genes for their activation. We therefore determined whether the human LCR at 8C3/C4 may be able to establish a transvection-like system with one of its natural target genes located at the homologous allele. For this purpose we generated a new series of transgenic mice using ES cell targeting. A genomic fragment containing the human Aγ-globin gene, carrying a GFP reporter gene inserted at its translational start site and with a few kilobase of surrounding sequences, was introduced in two orientations into the *Rad23a* gene (**Figure 5-1A**). The human LCR has previously been reported to confer high levels of transcription to a linked Aγ-globin gene in mouse transgenic erythroid cells until embryonic day E12.5 [29]. Like the mice carrying the LCR [8], mice homozygous for the Aγ-globin-GFP gene did not show an abnormal phenotype. Homozygous mice containing the Aγ-globin-GFP gene were crossed with mice carrying the LCR or WT (wildtype) mice. Offspring had one 8C3/C4 allele carrying the reporter gene in the one or the other orientation, while the other 8C3/C4 allele was either WT or contained one of two LCRs. Quantitative reverse

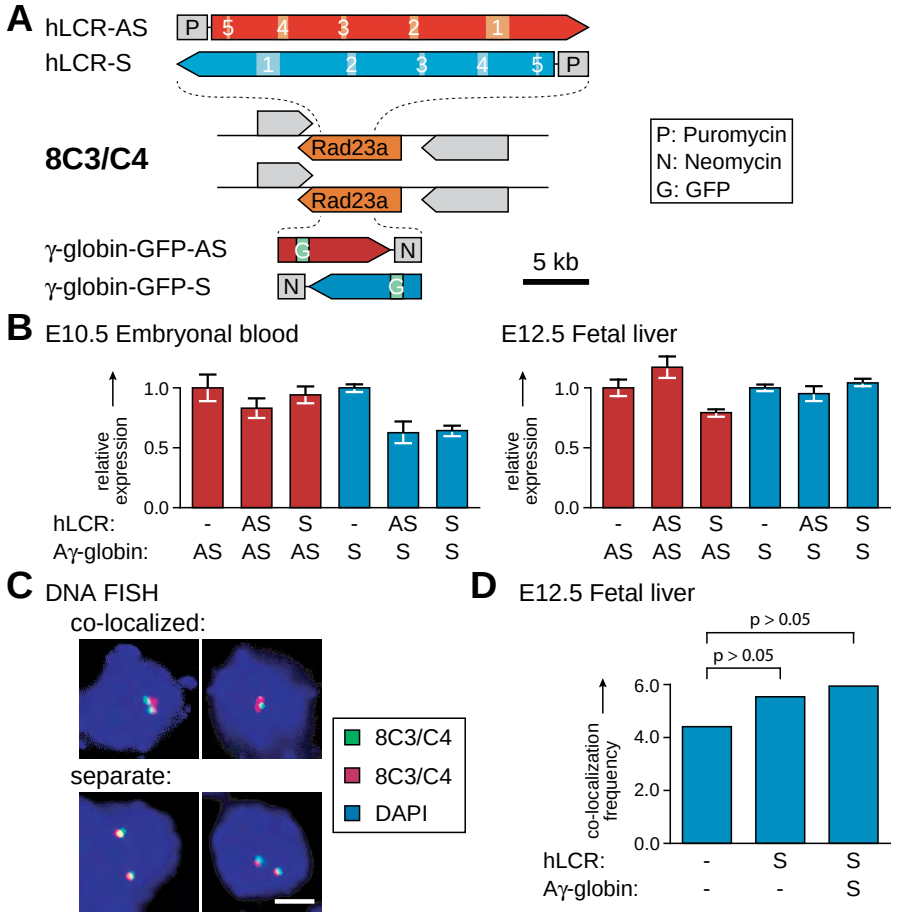


Figure 5-1. The human β-globin LCR at 8C3/C4 does not induce transvection. (A) Integration of the full human β-globin LCR in both orientations to one allele of the Rad23a gene (antisense (red) and sense (blue), relative to the orientation of the Rad23a gene) and a human Aγ-globin-GFP fusion-gene in both orientations to the other allele of the Rad23a gene (antisense and sense). Hypersensitive sites (HSs) in the LCR are numbered and indicated by shaded boxes. **(B)** Relative Aγ-globin transcript levels, normalized to Hprt1 transcript levels, in E10.5 embryonal blood and E12.5 fetal livers as determined by qRT-PCR. Genotypes are indicated below the x-axis. Aγ-globin transcript levels in samples not containing the human LCR were normalized to 1. Error bars represent standard error of at least two independent samples. **(C)** Examples of co-localized and separate 8C3/C4 alleles. Images obtained by DNA-FISH. Scale bar: 2 μm. (D) Co-localization frequencies of 8C3/C4 alleles. Presence of the human LCR and / or the Aγ-globin gene are indicated below the x-axis. Significance levels are indicated above the graph.

transcriptase PCR (qRT-PCR) analysis in two different erythroid cell populations, primitive E10.5 embryonal blood cells and definitive E12.5 FL cells, revealed that steady-state expression levels of the Aγ-globin-GFP gene are not increased when the LCR is located at the homologous allele (**Figure 5-1B**). We conclude that the Aγ-globin reporter gene is not upregulated by the LCR at the homologous allele. To further investigate if the LCR searches the nuclear interior for its natural target gene we performed 3D DNA-FISH and compared co-localization frequencies of 8C3/C4 alleles in E12.5 fetal liver cells (**Figure 5-1C, D** and **Table 5-2**). No significant difference in interaction frequencies was observed between combinations of 8C3/C4 alleles that were either WT/WT, WT/LCR or LCR/Aγ-globin (n > 150 alleles per genotype). We therefore found no evidence that 8C3/C4 is involved in high frequency homologous pairing, nor that the location of the LCR and a β-globin-like gene at 8C3/C4 leads to the formation of stable transvection-like interactions. We conclude that the human LCR, when located at 8C3/C4, does not actively search for its natural target gene at the homologous allele.

5.4.2- 8C3/C4 contacts essentially the same regions in cis *and* in trans *with and without the LCR*

To further investigate if the LCR searches for preferred interaction partners we performed 4C (Chromosome Conformation Capture on Chip) on mice homozygous for the LCR at 8C3/C4. 4C technology allows for an unbiased genome-wide screen for DNA segments that interact with a locus of choice. The strategy combines 3C technology [30] with dedicated micro-arrays [15, 31]. It involves the selective amplification by PCR of ligation products formed between DNA fragments that were formaldehyde crosslinked to a fragment of choice. This material is then labeled and hybridized to a micro-array which in this case contained probes for all possible unique restriction events on 7 mouse chromosomes including chromosome 8 where the LCR is integrated. The identified interaction partners of the LCR were compared to the previously identified interacting partners of the WT *Rad23a* locus in E14.5 FL [15]. When separated on gel prior to micro-array hybridization, the PCR products obtained with LCR-specific primers were highly reproducible between replicate experiments on transgenic animals but clearly distinct from products obtained with *Rad23a* primers on WT material (**Figure 5-2A, B**).

After micro-array hybridization, we first asked whether 4C confirmed our microscopy observation that the LCR positions 8C3/C4 more often outside its own CT. For this, we determined whether the ratio of interchromosomal over intrachromosomal ligation

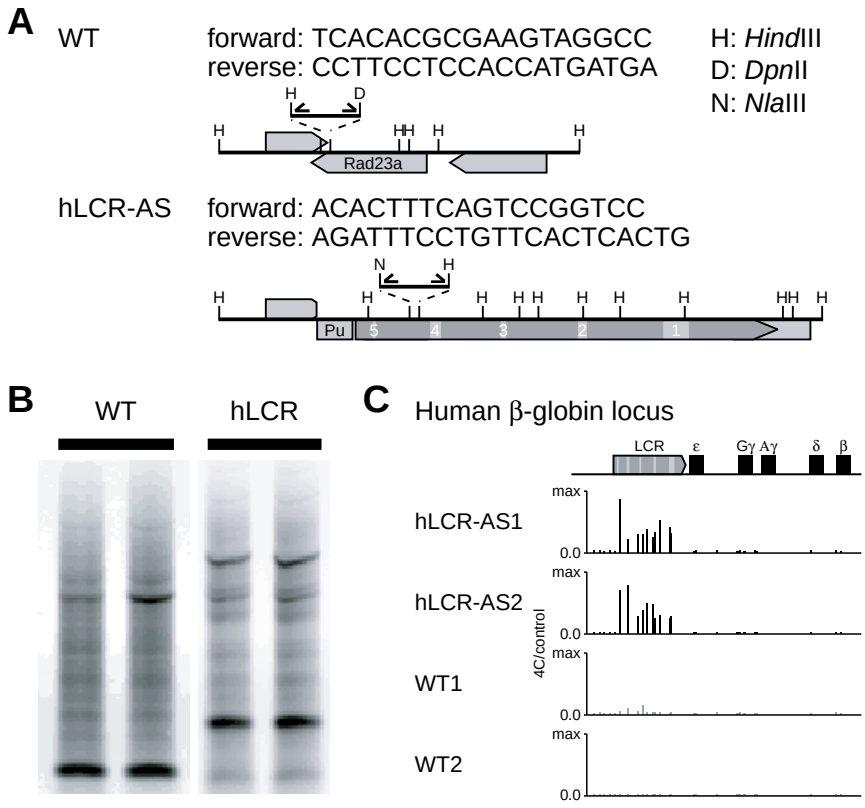


Figure 5-2. 4C analysis of 8C3/C4 with and without the integrated human β-globin LCR.(A) Location of restriction sites and primer sets used to amplify the WT Rad23a fragment and the LCR containing fragment. For reasons of clarity only relevant restriction sites of frequent cutters have been indicated. Completely different restriction sites and primer sets have been used to analyze the WT sample and the sample with the integrated LCR. **(B)** PCR amplified material separated by gel electrophoresis from the two independent samples of WT 8C3/C4 and the two independent samples of 8C3/C4 with human LCR-AS shows the reproducibility between replicates and the difference between the amplified 4C material generated with different primer sets. **(C)** Unprocessed 4C over genomic signal intensities for probes representing the human β-globin locus.

events changed after integration of the LCR. Probe signals from individual experiments were binned based on their intensities and subsequently for each bin the content of *cis* and *trans* probes was determined (**Figure 5-3**). Statistical analysis revealed a highly significant LCR-induced shift of *trans* probes to bins with the highest signal intensities

(K-S test; $P < 0.001$). This confirms that integration of the LCR results in increased encounters with genomic regions located on different chromosomes, as expected when a locus loops away from its CT more frequently. Thus, in agreement with a previous report that used a similar but less elaborate strategy to map DNA interactions with the *HoxB* locus [32], 4C technology can determine whether a locus changes its positioning relative to the CT.

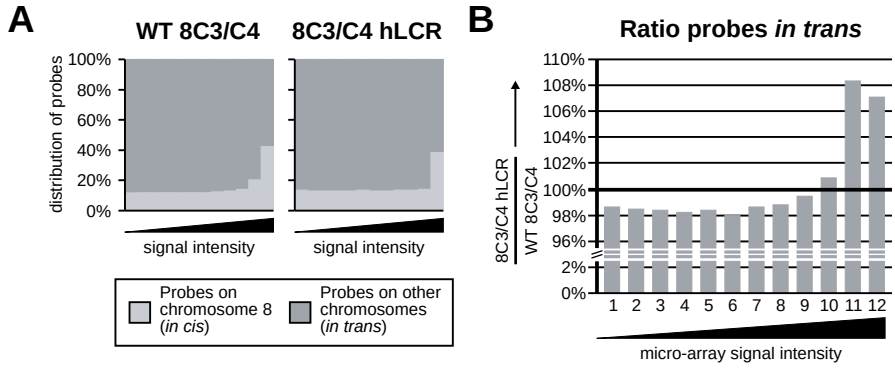


Figure 5-3. Location outside the CT increases interchromosomal interactions. (A) Bin plots showing unprocessed signal intensity of 4C probes. Probe sets from the 4C experiment have been distributed according to signal intensities in 12 equal sized bins and the relative number of probes on chromosome 8 and on the other chromosomes has been indicated for each bin. **(B)** Graph showing the ratio for each bin of the samples with the LCR versus the samples without the LCR. Probe sets in trans tend to be more enriched in the bins containing higher signal intensities when the LCR is integrated in 8C3/C4, which goes on the expense of probe sets in cis ($P < 0.001$; K-S test with null hypothesis that probes are similary binned independent of the presence of the LCR). Together, this shows that the restriction fragment with the LCR is ligated more often to fragments on other chromosomes than the WT restriction fragment, confirming that location away from the CT promotes clustering with genomic regions in trans.

As an additional control, probes representing the entire human β-globin locus were present on the micro-array. High signals were observed exclusively at probes that corresponded to the LCR and not to other parts of the human β-globin locus, and these signals were seen only if material from LCR-containing mice was hybridized. This further verified the specificity of the different primer sets (**Figure 5-2C**).

To investigate the genomic interactions in detail, 4C signal intensities were analyzed as described before [15, 31]. In brief, the analysis involves normalization of signal intensities to control genomic DNA signals, the application of a running median algorithm

to select clusters of positive signals and thresholding based on signal randomization. When 8C3/C4 interactions in fetal liver cells were plotted along the chromosomes they appeared highly similar between WT and LCR-transgenics (**Figure 5-4A**). Importantly, these signals were obtained from different animals and with entirely different primer sets (mouse *Rad23a* primers for WT and human β-globin LCR primers for transgenic mice). This shows that long-range interactions identified by 4C reflect contacts between regions, rather than specific restriction fragments, of the genome. Statistical analysis

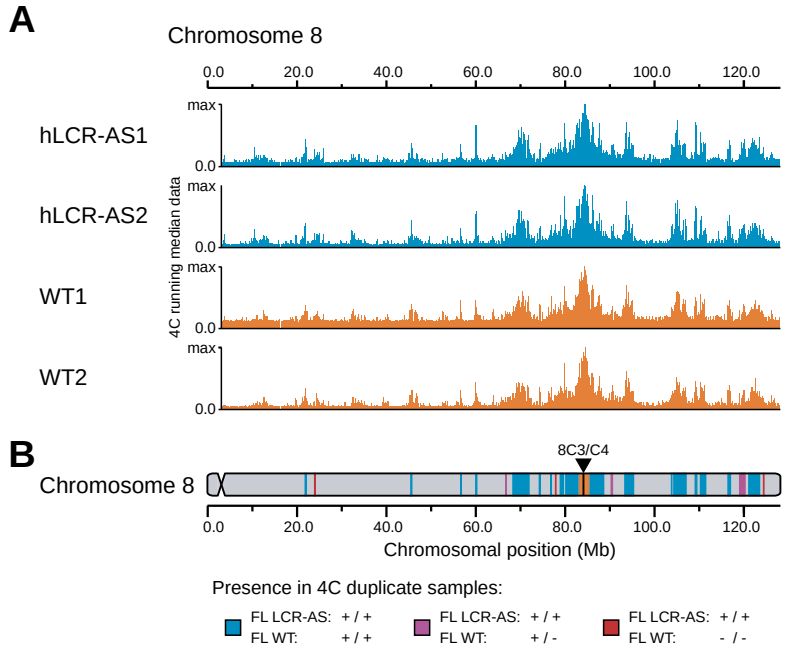


Figure 5-4. Integration of the human β-globin LCR essentially does not change intra-chromosomal interactions of 8C3/C4. (A) 4C running mean data of chromosome 8 of samples with and without the LCR integrated at 8C3/C4. Even though samples have been amplified with different primer sets, the overall chromosome-wide pattern is generally comparable. **(B)** 4C analysis of long-range interactions of 8C3/C4 with and without the human β-globin LCR. Regions identified in the 4C analysis to significantly associate with 8C3/C4 containing the LCR are color-coded for their association with WT 8C3/C4 (see below). The orange region is the region of around Mb around the integration site where probes on the micro-array are saturated.

confirmed the high correlation rate, both between LCR replicates ($\rho = 0.76$; Spearman's rank correlation coefficient) and between samples with and without the LCR ($\rho = 0.76$

or higher). 23 intrachromosomal interacting regions were identified in both replicates containing the LCR (**Figure 5-4B** and **Table 5-3**). Of these regions, 17 (>70%; blue) were also identified to interact with WT 8C3/C4. Additionally, 3 interacting regions of 8C3/C4 with the LCR were positive in one of the two WT replicates (13%; purple). Previously we

Interacting regions in cis

Chromosome	Locus (Cytological band)	Start (Mb)	End (Mb)	Presence in WT samples WT1	WT2
8	8A3	21.8	22.0	+	+
8	8A3	23.9	24.0	-	-
8	8B2	45.4	45.6	+	+
8	8B3.2	56.5	56.6	+	+
8	8B3.3	59.9	60.1	+	+
8	8C1	66.6	66.7	-	+
8	8C1/C2	68.2	71.8	+	+
8	8C2	74.1	74.3	+	+
8	8C3	76.7	76.8	+	+
8	8C3	77.7	77.8	-	-
8	8C3	78.8	79.0	+	+
8	8C3	79.5	82.4	+	+
Location 8C3/C4					
8	8C4	85.5	88.5	+	+
8	8C5	90.2	90.4	-	+
8	8C5	93.2	95.2	+	+
8	8D1	103.9	103.9	+	+
8	8D2	104.5	106.9	+	+
8	8D3	108.9	109.3	+	+
8	8D3	110.1	111.3	+	+
8	8E1	116.3	116.8	+	+
8	8E1	118.9	119.9	-	+
8	8E2	120.9	123.2	+	+
8	8E2	124.2	124.3	-	-

Interacting regions in trans

Chromosome	Locus (Cytological band)	Start (Mb)	End (Mb)	Presence in WT samples WT1	WT2
7	7D2	74.0	74.1	+	+
7	7F1	94.5	94.6	+	+
7	7F4	121.1	121.3	+	+
7	7F5	135.7	135.7	-	+
10	10B2	41.3	41.4	-	-
10	10B2	43.7	43.8	+	+
10	10C1	80.1	81.6	-	+
11	11A5	32.1	32.2	-	+
11	11B3	68.8	68.8	+	+
11	11B4	69.5	70.0	+	+
11	11B5	77.7	77.8	+	+
11	11D	97.2	97.3	-	+
11	11D	102.0	102.2	+	+
12	12D3	77.6	77.7	-	+
12	12F2	106.5	106.5	-	+
14	14C2	50.1	50.2	+	+
14	14D1	63.8	63.9	+	+
15	15C	36.6	36.7	-	-
15	15E1	75.8	76.4	-	+

Table 5-3 Interacting regions identified by 4C

showed by FISH that almost invariably these regions do interact under both conditions, but that they are not scored by 4C as such in one condition due to stringent thresholding [33]. Interactions with the 3 remaining regions appear to uniquely depend on the LCR, but this needs to be confirmed by FISH experiments. Inspection of these regions reveals that they do not contain other erythroid-specific genes. Together, our preliminary results based on 4C analysis indicate that interactions of 8C3/C4 *in cis* are largely identical, irrespective of whether the LCR is integrated or not. The LCR therefore does not seem to be the main element determining the interactions *in cis* of 8C3/C4, nor does it seem to induce drastic changes in chromosome 8 chromatin conformation. We therefore conclude that the LCR does not scan the chromosome it is integrated in for preferred interaction partners.

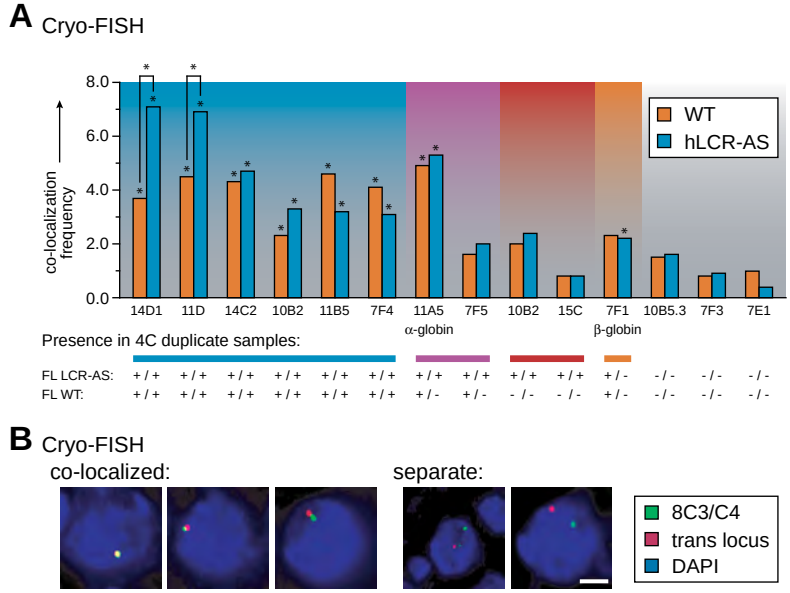


Figure 5-5. The human β-globin LCR does not promote appreciatably different interchromosomal interactions. (A) Co-localization frequencies of interchromosomal regions and 8C3/C4 with and without the human β-globin LCR. Regions are grouped according to their interaction status of the 4C replicates in both WT and transgenic replicates. Within groups, regions are sorted according to their co-localization frequency with 8C3/C4 containing the LCR. Location within genomic bands of interchromosomal regions are indicated below the graph. Significant co-localization versus the negative controls is indicated by asterisks above each bar. Significant co-localization frequencies between 8C3/C4 with and without the LCR is indicated by asterisks between bars. *: $p<0.05$. **(B)** Examples of Cryo-FISH sections showing co-localized and separate signals. Scale bar: 2 μm.

We next focused on the interacting regions *in trans* of 8C3/C4 with the LCR and compared these to the previously identified interchromosomal interactions of WT 8C3/C4 (**Table 5-3**). A first inspection of the interchromosomal interactions identified by 4C suggested that the LCR had impact on the contacts of 8C3/C4 with other chromosomes. Out of 19 regions identified to interact with 8C3/C4 containing the LCR, only 10 (>50%) were found to also interact with WT 8C3/C4. However, an additional 7 regions were positive in one of the two WT replicate 4C experiments. To understand the meaning of these results, we systematically analyzed co-localization frequencies of 8C3/C4 with 10 of the 19 interaction partners by Cryo-FISH. Cryo-FISH is a FISH technique that uses ultrathin (150–200nm) cryosections which compared to standard 3D-FISH, provides increased resolution in the Z-axis [34]. To get reliable co-localization frequencies, a minimum of 250 alleles was analyzed per combination of probes and scoring was done by a person not aware of the probe combination used. Moreover, in each counting session always at least one slide was included that analyzed a negative interaction (between two chromosomal regions that scored negative for interaction by 4C), such that it was impossible for the investigator to have any expectations with regard to contact frequencies (**Figure 5-5A, B** and **Table 5-2**). Results of the 4C analysis were largely confirmed by Cryo-FISH, since regions identified to interact *in trans* with both 8C3/C4 with and without the LCR were always co-localizing significantly more often than negative control regions. The maximum co-localization frequency that we scored for allelic interaction of 8C3/C4 with a region *in trans* was 7.1%. When allelic interaction frequencies dropped below 2%, a correlation between the 4C data and Cryo-FISH measured interaction frequencies was found to be lost. This probably reflects limitations to the resolution of 4C, Cryo-FISH or both. The most striking observation was though that no novel interchromosomal contacts were made when 8C3/C4 carried the β-globin LCR. We therefore conclude that LCR-mediated relocation outside the CT is not caused by the regulatory element searching for new preferred interaction partners located on other chromosomes.

Cryo-FISH also demonstrated that 8C3/C4 did not significantly change its co-localization frequency with most of the interacting regions *in trans* (**Figure 5-5A** and **Table 5-2**). However, two regions did significantly change their interaction frequency with 8C3/C4. The regions 11D (Chromosome 11; around 102 Mb) and 14D1 (Chromosome 14; around 64 Mb) co-localized about two-fold more often with 8C3/C4 when the LCR was integrated, compared to the WT region. Both regions contain highly expressed erythroid specific genes (**Figure 5-6** and http://symatlas.gnf.org/SymAtlas/). This could be an indication

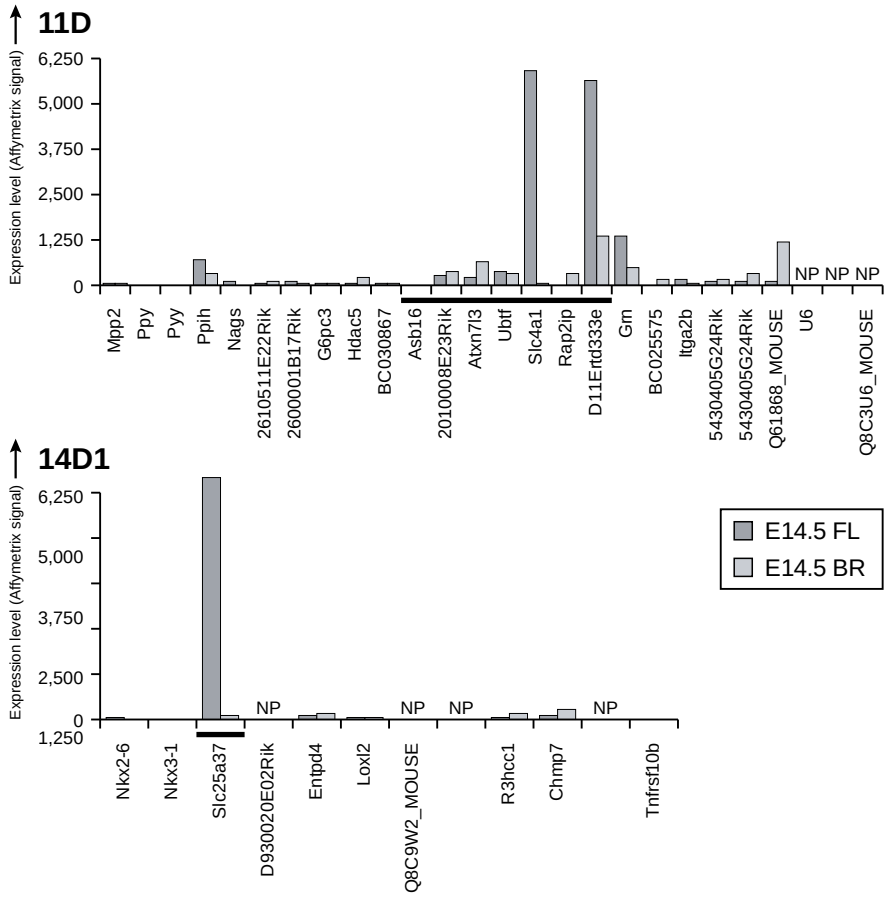
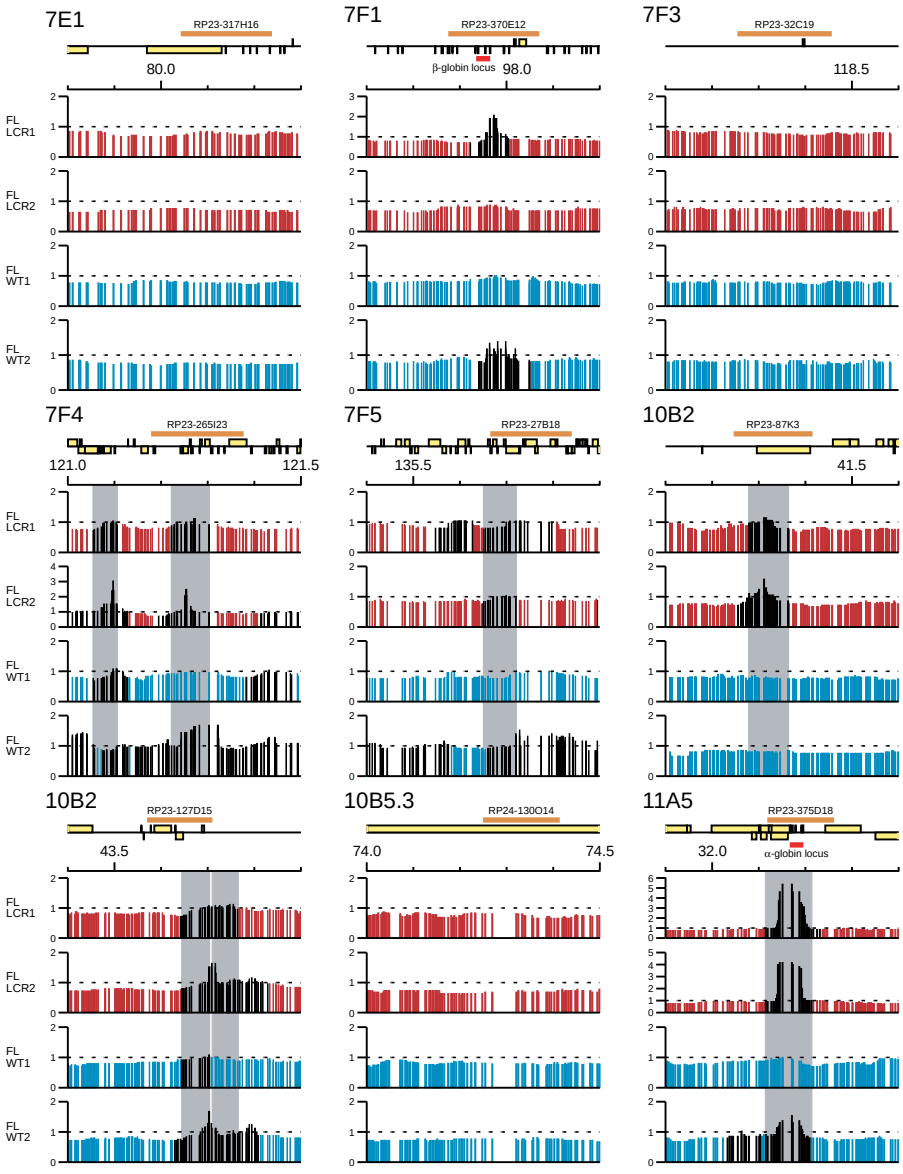


Figure 5-6. Expression of genes in differentially associating regions. Expression levels from murine Affymetrix 430 2.0 Micro-array in WT E14.5 fetal liver and fetal brain of genes located in regions differentially co-localizing with 8C3/C4 depending on the presence of the LCR. Expression levels are shown for genes located within 0.5 Mb from the middle of the interacting region in samples containing the LCR, as determined by Cryo-FISH. Black bar below the x-axis indicates genes present in the interacting region identified by 4C on E14.5 fetal liver. NP: not present on the micro-array.

Microarray data from Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., and de Laat, W. 2006. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nat Genet 38(11): 1348-1354.

that the integrated LCR can stabilize existing interactions with functionally related genes. When interpreting this data though it is important to realize that one of these regions, 11D, also interacts with 8C3/C4 in WT fetal brain [15]. Thus, in WT tissue this particular interaction is not dependent on an erythroid-specific component, but the combination of an LCR and an erythroid-specific gene may stabilize this contact in transgenic fetal livers. Interestingly, 4C technology also identified interchromosomal interactions between WT 8C3/C4 and the α-globin locus on chromosome 11 and between WT 8C3/C4 and the β-globin locus on chromosome 7. In both cases, this interaction was identified in only one of the two replicate 4C experiments. The interaction with α-globin, but not with β-globin, was also identified in fetal brain, again showing that its proximity to WT 8C3/C4 did not depend on an erythroid-specific factor. Cryo-FISH measurements could not discern the interaction frequencies with β-globin from background (around 2%) but did demonstrate a specific interaction of 8C3/C4 with the α-globin locus (around 5%). While integration of the LCR at 8C3/C4 had no impact on 4C signals at the endogenous β-globin locus, it resulted in strongly increased 4C signals at the α-globin locus (**Figure 5-7**). Interestingly though, preliminary Cryo-FISH data suggested that this was not accompanied by increased interaction frequencies, as they stayed around 5% in the presence of the LCR. Together, these results indicate that integration of the LCR in 8C3/C4 does not result in the formation of new interactions *in trans*, nor does it lead to large-scale changes in interaction frequencies of existing interactions *in trans*. Whether or not the LCR stabilizes existing interactions with erythroid-specific genes requires further investigation.


7E1
RP23-317H16
80.0
7F1
RP23-370E12
β-globin locus
98.0
7F3
RP23-32C19
118.5
FL LCR1
FL LCR2
FL WT1
FL WT2
7F4
RP23-265I23
121.0
121.5
7F5
RP23-27B18
135.5
10B2
RP23-87K3
41.5
10B2
RP23-127D15
43.5
10B5.3
RP24-130O14
74.0
74.5
11A5
RP23-375D18
32.0
α-globin locus

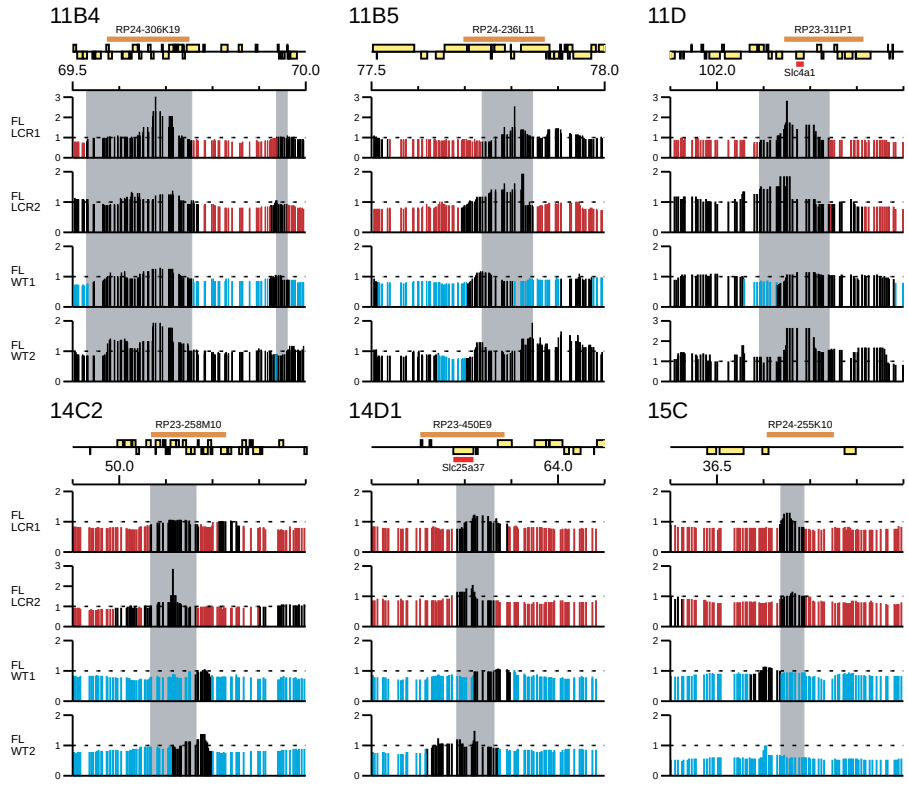


Figure 5-7. 4C / control ratios of regions analyzed by Cryo-FISH. Processed running median 4C signal for regions analyzed by Cryo-FISH. For each region a total of 0.5 Mb around the BAC probe used to analyze the region in the Cryo-FISH analysis is shown. 4C signals versus 8C3/C4 with the LCR are visualized by red bars, 4C signals versus WT 8C3/C4 are visualized by blue bars. Regions scored positive in the running median analysis are visualized by black bars. Double positive regions in the 8C3/C4 samples with the LCR are highlighted by the large grey bar. For each region, the name of the genomic region, the genomic coordinates, the location of the BAC used to visualize the region in the Cryo-FISH analysis and the location of genes in the region are indicated above the 4C graphs. Erythroid specific genes are indicated by a red bar and their name below the specific gene.

5.5- Discussion

In this study we have determined whether an ectopically integrated human β-globin LCR can actively influence the nuclear organization of a gene-dense region it is integrated in. Previously, we reported that introduction of the LCR in this region results in significant relocation versus the CT and a mild increase in association with transcription factories [8]. We hypothesized that repositioning could be a consequence of the LCR searching for functionally related partners in the nuclear space, for instance by promoting interactions with transcription factories dedicated to erythroid transcription.

Initially, we wondered whether the human LCR could actively search a human Aγ-globin gene in a transvection-like setup. Plasmids actively transcribing the Aγ-globin gene have been reported to co-localize with the endogenous β-globin locus in human HeLa cells, thereby inducing a significant activation of the endogenous β-globin locus [35]. Furthermore, transvection has been proposed as the mechanism regulating X-chromosome inactivation [23, 24]. We therefore systematically tested whether the LCR could initiate a transvection-like process. In contrast to the previous study showing clustering of plasmids at the human β-globin locus, we found no increased clustering of the homologous alleles or upregulation of the globin gene by the LCR on the homologous allele. An explanation for the observed differences with the previous study could be the different mobility of the genomic elements involved. The position of chromosomal loci in general is highly constraint [36, 37], while plasmids sharing similar characteristics cluster in only a few spots, implying that they are allowed to move relatively freely through the nucleus [14]. Together though, our experiments fail to provide evidence for transvection in mice cells, or the ability of the human LCR to initiate such a process.

As a next step we determined whether the LCR promotes clustering with preferred, non-homologous, interaction partners. Using 4C, we identified genomic interacting partners *in cis* and *in trans* of 8C3/C4 with and without the integrated LCR. Interestingly, we found that the large majority of interaction partners, both on the same chromosome and on different chromosomes, were identical regardless of the presence of the LCR. Subsequent high resolution Cryo-FISH further supported the conclusion that introduction of the LCR in 8C3/C4 for the large majority of interaction partners *in trans* does not influence their co-localization frequency. Currently, we are setting up a similar FISH analysis for interacting partners *in cis*. In general though, we found no evidence for the LCR actively repositioning its genomic integration site to nuclear zones with specialized functions, which is in contrast to previous reports that the β-globin LCR influences the nuclear location of the

locus it is integrated in [6, 7, 9]. Similarly, these data also do not support a previously hypothesized role in the promotion of clustering with functionally related genes [11, 13]. Rather, the LCR seems to more often explore the same nuclear space outside the CT that is also visited by WT 8C3/C4. Previously, we proposed that clustering of genomic loci may be dependent on overall activity state and associated chromatin characteristics [15]. Similarly, we also proposed that LCR-induced repositioning versus the CT and transcription factories could be caused by increased chromatin decondensation and mobility [8]. The lack of LCR-mediated changes in genomic environment observed in this study, provides further evidence that the LCR itself is not actively involved in determining the nuclear location of the target locus it is involved in. Rather, the LCR seems to exert its influence by changing chromatin characteristics, which may potentially and partially be mediated by increased transcription in the region around its integration site.

An issue that remains is the observed increase in co-localization of 8C3/C4 with two interaction partners containing erythroid specific genes. Importantly, both these regions also significantly co-localize with WT 8C3/C4, while one of the regions (11D, chromosome 11) was previously also identified as an interacting partner *in trans* of 8C3/C4 in E14.5 fetal brain [15]. Furthermore, the highly erythroid specific α-globin and β-globin loci do not increase their interaction frequency with 8C3/C4 when the LCR is integrated. Increased co-localization therefore is not a general phenomenon for erythroid specific loci. A recent study on the human *Slc4a1* gene, which is the erythroid specific gene located in 11D and that is located in a very large region synthenic to the mouse 11D locus, may provide insight into the mechanism behind the increased LCR-mediated co-localization [16]. This gene is located in a highly decondensed genomic region, which may suggest it also loops out of the CT frequently. In contrast, the active mouse α-globin and β-globin loci have previously been shown to be located outside the CT in only a very small percentage of cells [10]. It may therefore be that the LCR-induced location of 8C3/C4 outside the CT specifically promotes increased co-localization with other decondensed genomic loci that loop out of the CT with high frequency. Additional studies will be necessary to verify this hypothesis.

Another intriguing observation is the highly increased 4C signals, exclusively observed for the interaction between 8C3/C4 containing the LCR and the α-globin locus. This very strong increase seems not accompanied by changes in co-localization frequencies as determined by preliminary Cryo-FISH experiments. Given this apparent discrepancy, it will be important to repeat the Cryo-FISH experiments with an independent probe set.

If results are confirmed (i.e. more ligation products captured by 4C without increased interaction frequencies measured by FISH), an explanation may be that once in proximity, the LCR stabilizes interactions between 8C3/C4 and the α-globin locus. In this scenario the LCR would only stabilize an interaction when 8C3/C4 is already located at a very short distance from the α-globin locus, which is different from the LCR searching for preferred interaction partners in the nuclear space.
In conclusion, current results suggest that the LCR does not search the nuclear interior for interactions with functionally related genes, but the data leave open the possibility that it stabilizes existing interactions with (some) erythroid-specific genes. We plan to perform an analysis of 4C-results based on p-values that should allow a more quantitative assessment of differences measured by 4C in WT and LCR-containing transgenics. Among the regions that are captured more efficiently by 4C when the LCR is present we will ask whether they are enriched in erythroid-specific genes and/or EKLF target genes and/or GATA-1 target genes in order to more confidently determine whether indeed the LCR stabilizes interactions with selected target genes in the nuclear space.

5.6- Acknowledgements

We thank Frank Grosveld and members of the lab for discussion, Harmen van de Werken for assistance with the statistical analysis and Manousos Koutsourakis for assistance with the generation of transgenic mice. This work was supported by grants from the Dutch Scientific Organization (NWO) (016-006-026) and (912-04-082) and the Netherlands Genomics Initiative (050-71-324) to W.d.L.

5.7- References

1. Misteli, T. (2007). Beyond the sequence: cellular organization of genome function. Cell *128*, 787-800.
2. Lanctot, C., Cheutin, T., Cremer, M., Cavalli, G., and Cremer, T. (2007). Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. Nat Rev Genet *8*, 104-115.
3. Schneider, R., and Grosschedl, R. (2007). Dynamics and interplay of nuclear architecture, genome organization, and gene expression. Genes Dev *21*, 3027-3043.
4. Lundgren, M., Chow, C.M., Sabbattini, P., Georgiou, A., Minaee, S., and Dillon, N. (2000). Transcription factor dosage affects changes in higher order chromatin structure associated with activation of a heterochromatic gene. Cell *103*, 733-743.
5. Hewitt, S.L., Farmer, D., Marszalek, K., Cadera, E., Liang, H.E., Xu, Y., Schlissel, M.S., and Skok, J.A. (2008). Association between the Igk and Igh immunoglobulin loci mediated by the 3' Igk enhancer induces 'decontraction' of the Igh locus in pre-B cells. Nat Immunol *9*, 396-404.
6. Ragoczy, T., Bender, M.A., Telling, A., Byron, R., and Groudine, M. (2006). The locus control region is required for association of the murine beta-globin locus with engaged transcription factories during erythroid maturation. Genes Dev *20*, 1447-1457.
7. Francastel, C., Walters, M.C., Groudine, M., and Martin, D.I. (1999). A functional enhancer suppresses silencing of a transgene and prevents its localization close to centrometric heterochromatin. Cell *99*, 259-269.
8. Noordermeer, D., Branco, M.R., Splinter, E., Klous, P., van Ijcken, W., Swagemakers, S., Koutsourakis, M., van der Spek, P., Pombo, A., and de Laat, W. (2008). Transcription and chromatin organization of a housekeeping gene cluster containing an integrated beta-globin locus control region. PLoS Genet *4*, e1000016.
9. Ragoczy, T., Telling, A., Sawado, T., Groudine, M., and Kosak, S.T. (2003). A genetic analysis of chromosome territory looping: diverse roles for distal regulatory elements. Chromosome Res *11*, 513-525.
10. Brown, J.M., Leach, J., Reittie, J.E., Atzberger, A., Lee-Prudhoe, J., Wood, W.G., Higgs, D.R., Iborra, F.J., and Buckle, V.J. (2006). Coregulated human globin genes are frequently in spatial proximity when active. J Cell Biol *172*, 177-187.
11. Osborne, C.S., Chakalova, L., Brown, K.E., Carter, D., Horton, A., Debrand, E., Goyenechea, B., Mitchell, J.A., Lopes, S., Reik, W., and Fraser, P. (2004). Active genes dynamically colocalize to shared sites of ongoing transcription. Nat Genet *36*, 1065-1071.
12. Iborra, F.J., Pombo, A., Jackson, D.A., and Cook, P.R. (1996). Active RNA polymerases are localized within discrete transcription "factories' in human nuclei. J Cell Sci *109 (Pt 6)*, 1427-1436.
13. Chakalova, L., Debrand, E., Mitchell, J.A., Osborne, C.S., and Fraser, P. (2005). Replication and transcription: Shaping the landscape of the genome. Nat Rev Genet *6*, 669 - 677.
14. Xu, M., and Cook, P.R. (2008). Similar active genes cluster in specialized transcription factories. J Cell Biol *181*, 615-623.
15. Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., and de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by

chromosome conformation capture-on-chip (4C). Nat Genet *38*, 1348-1354.

16. Brown, J.M., Green, J., das Neves, R.P., Wallace, H.A., Smith, A.J., Hughes, J., Gray, N., Taylor, S., Wood, W.G., Higgs, D.R., Iborra, F.J., and Buckle, V.J. (2008). Association between active genes occurs at nuclear speckles and is modulated by chromatin environment. J Cell Biol *182*, 1083-1097.
17. de Laat, W. (2007). Long-range DNA contacts: romance in the nucleus? Curr Opin Cell Biol *19*, 317-320.
18. Ng, J.M., Vermeulen, W., van der Horst, G.T., Bergink, S., Sugasawa, K., Vrieling, H., and Hoeijmakers, J.H. (2003). A novel regulation mechanism of DNA repair by damage-induced and RAD23-dependent stabilization of xeroderma pigmentosum group C protein. Genes Dev *17*, 1630-1645.
19. Solovei, I., Cavallo, A., Schermelleh, L., Jaunin, F., Scasselati, C., Cmarko, D., Cremer, C., Fakan, S., and Cremer, T. (2002). Spatial preservation of nuclear chromatin architecture during three-dimensional fluorescence in situ hybridization (3D-FISH). Exp Cell Res *276*, 10-23.
20. Sokal, R.R., and Rohlf, F.J. (1995). Biometry: the Principles and Practice of Statistics in Biological Research 3rd edn. (New York: W.H. Freeman).
21. Pirrotta, V. (1999). Transvection and chromosomal trans-interaction effects. Biochim Biophys Acta *1424*, M1-8.
22. Kosak, S.T., Scalzo, D., Alworth, S.V., Li, F., Palmer, S., Enver, T., Lee, J.S., and Groudine, M. (2007). Coordinate gene regulation during hematopoiesis is related to genomic organization. PLoS Biol *5*, e309.
23. Xu, N., Donohoe, M.E., Silva, S.S., and Lee, J.T. (2007). Evidence that homologous X-chromosome pairing requires transcription and Ctcf protein. Nat Genet *39*, 1390-1396.
24. Bacher, C.P., Guggiari, M., Brors, B., Augui, S., Clerc, P., Avner, P., Eils, R., and Heard, E. (2006). Transient colocalization of X-inactivation centres accompanies the initiation of X inactivation. Nat Cell Biol *8*, 293-299.
25. Cremer, T., and Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nat Rev Genet *2*, 292-301.
26. Bolzer, A., Kreth, G., Solovei, I., Koehler, D., Saracoglu, K., Fauth, C., Muller, S., Eils, R., Cremer, C., Speicher, M.R., and Cremer, T. (2005). Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. PLoS Biol *3*, e157.
27. Branco, M.R., and Pombo, A. (2006). Intermingling of Chromosome Territories in Interphase Suggests Role in Translocations and Transcription-Dependent Associations. PLoS Biol *4*, e138.
28. Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F., and de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. Mol Cell *10*, 1453-1465.
29. Strouboulis, J., Dillon, N., and Grosveld, F. (1992). Developmental regulation of a complete 70-kb human beta-globin locus in transgenic mice. Genes Dev *6*, 1857-1864.
30. Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. Science *295*, 1306-1311.
31. Simonis, M., Kooren, J., and de Laat, W. (2007). An evaluation of 3C-based methods to capture DNA interactions. Nat Methods *4*, 895-901.
32. Wurtele, H., and Chartrand, P. (2006). Genome-wide scanning of HoxB1-associated loci in mouse ES cells using an open-ended Chromosome Conformation Capture methodology. Chromosome Res *14*, 477-495.

33. Palstra, R.J., Simonis, M., Klous, P., Brasset, E., Eijkelkamp, B., and de Laat, W. (2008). Maintenance of long-range DNA interactions after inhibition of ongoing RNA Polymerase II transcription. PLoS ONE *3*, e1661.
34. Pombo, A. (2007). Advances in imaging the interphase nucleus using thin cryosections. Histochem Cell Biol *128*, 97-104.
35. Ashe, H.L., Monks, J., Wijgerde, M., Fraser, P., and Proudfoot, N.J. (1997). Intergenic transcription and transinduction of the human beta-globin locus. Genes Dev *11*, 2494-2509.
36. Abney, J.R., Cutler, B., Fillbach, M.L., Axelrod, D., and Scalettar, B.A. (1997). Chromatin dynamics in interphase nuclei and its implications for nuclear structure. J Cell Biol *137*, 1459-1468.
37. Chubb, J.R., Boyle, S., Perry, P., and Bickmore, W.A. (2002). Chromatin motion is constrained by association with nuclear compartments in human cells. Curr Biol *12*, 439-445.

Chapter 6

Trans-activation of an endogenous mouse β-globin gene by the human β-globin LCR and a β-globin gene inserted in anotherchromosome

Daan Noordermeer, Petra Klous and Wouter de Laat

Work in progress

6.1- Abstract

In recent years, interchromosomal interactions between functionally related mammalian genes have been observed, which were proposed to be important for regulation of their expression. However, genetic evidence showing that the loss of DNA elements on the one chromosome alters gene expression in *trans* is mostly lacking. In this study, we determined whether an ectopic human β-globin LCR integrated on mouse chromosome 8 activates genes elsewhere in the genome. Using a micro-array expression analysis, we identified a single gene located on a different chromosome that is upregulated over twofold. Interestingly, this βh1 gene on chromosome 7 is a natural target gene of the endogenous mouse β-globin LCR. Even more intriguing though, we found this same gene to be upregulated when the ectopic LCR was replaced by the human fetal Aγ–globin gene. Extensive analysis showed that interchromosomal upregulation of βh1 gene expression was driven by the insertion of β-globin sequences into chromosome 8. This study therefore provides unique genetic evidence for interchromosomal transcriptional regulation in mammals. The change in expression was not accompanied by increased interaction frequencies when both loci contained β-globin sequences since interchromosomal contacts invariably remained rare. We nevertheless hypothesize that the communication between loci on different chromosomes relies on their physical interactions and we propose future experiments to test whether interactions that occur at any given time in only a small percentage of cells account for a population-wide 2-3 fold increased expression level.

6.2- Introduction

The mouse β-globin locus consists of four developmentally regulated globin genes and an upstream regulatory element named the Locus Control Region (LCR). The LCR enhances the expression of the globin genes 25-100 fold through the formation of long-range chromatin loops [1-3]. Even though the existence of these loops is nowadays well established, the mechanisms behind the establishment of chromatin loops is still under debate [4-6]. An important difference between the models that explain loop formation is the question whether genomic loci need to be physically linked on the chromosome template to allow loop formation. In insects, this requirement does not exist, as regulatory elements can contact and upregulate genes on the homologous chromosome through a mechanism called transvection [7]. Transvection requires homologous pairing, which is a common phenomenon in *Drosophila*, where it is most prominently seen at polytene chromosomes. In mammals though, evidence for homologous pairing is generally lacking [8, 9]. Still, transvection has been proposed as an explanation for certain regulatory mechanisms in mammals [10-12]. We previously found the human β-globin LCR not to be involved in transvection when ectopically integrated at a specific position in the mouse genome (see **Chapter 5**). In recent years, several studies have been published where mammalian regulatory elements are proposed to regulate genes *in trans* (on other chromosomes) [13-17]. Generally, these studies identify interchromosomal interactions between functionally related genes or sequences, which are then interpreted as being evidence for functional communication between these regions. Unfortunately though, several of these studies have generated contradictory or controversial results ([14] vs. [16]; [15] vs. [20]) or had to be retracted ([18, 19]), complicating the interpretation of these conclusions. Moreover, genetic evidence, where it is demonstrated that deleting a region on the one chromosome affects gene expression on another chromosome, is mostly lacking. Perhaps an exception is the report which shows that deleting a transcription regulatory element at the T_H2 locus on mouse chromosome 11 delays the expression of the *Ifng* gene on chromosome 10 in an *ex vivo* differentiation experiment [13], although it also seems possible that this is an indirect effect, reflecting a differentiation problem of the genetically modified cells. In all these studies interaction frequencies measured across the cell population at a given time are considerably lower than 100%. This implies a long term memory of the regulatory effect (e.g. a stable messenger RNA) and considerable chromatin mobility, if the mechanism would control a pan-cellular process. Alternatively, the stochastic nature

of *trans*-interactions may activate a process only in a sub-population of cells, allowing them to selectively respond to a trigger [15, 17].
Recently, we reported the generation of transgenic mice containing the human β-globin LCR integrated in a gene-dense region we named 8C3/C4 [21]. We demonstrated that integration of the β-globin LCR in 8C3/C4 results in an orientation dependent upregulation of genes up to 150 kb away from the integration site [21]. Here, we investigated if this prototype of a strong transcription regulatory DNA element can activate genes elsewhere in the genome. Interestingly, the single gene we identify *in trans* that is upregulated more than twofold is the βh1 gene, an early embryonic endogenous mouse β-globin-like gene that at the investigated developmental stage is not detected by RNA-FISH to be active [22]. Even more intriguingly, we find this same gene to be upregulated if the LCR at 8C3/C4 is replaced by the human fetal Aγ–globin gene. The micro-array data are confirmed by quantitative real-time PCR (qRT-PCR) analysis of messenger RNA levels. We exclude that the upregulation of the βh1 gene is caused by differences in genetic background between wildtype (WT) and transgenic mice, and also show that it is not an indirect consequence of altered gene expression at 8C3/C4. *Trans*-activation is not found upon the insertion of a selection cassette at 8C3/C4, showing that it depends on these globin sequences. This study therefore provides unique genetic evidence for interchromosomal gene activation. We hypothesize that this communication *in trans* relies on a physical interaction between the loci, meaning that the low percentage of cells with this interchromosomal contact accounts for the overall 2-3 fold increased level of βh1 messenger RNA measured across the cell population. We aim to investigate this using combined RNA-DNA FISH. If true, it would show that functional compatibility, more than interaction frequency, determines whether contacts between genomic sites result in gene expression changes.

6.3- Materials and methods

6.3.1- Gene targeting and the generation of transgenic mice

Insertion of the human β-globin LCR into the mouse *Rad23a* gene and generation of transgenic mice was described in **Section 4.3.1**. Insertion of the human Aγ-globin gene into the mouse *Rad23a* gene was described in **Section 5.3.1**.

6.3.2- Affymetrix gene expression analysis

Total RNA was isolated using the RNeasy Mini kit (Qiagen) from three independent male E14.5 fetal livers for each genotype. Biotinylated cRNA was generated using the One-cycle Target Labeling and Control Reagents Kit (Affymetrix). All previous procedures and hybridisation, washing and scanning of the Affymetrix Mouse Genome 430 2.0 Arrays were done according to manufacturers' instructions. Array-data was normalized using Bioconductor RMA ca-tools. For each probe set, the values of the three independent micro-arrays were averaged.

6.3.3- qRT-PCR gene expression analysis

Total RNA was isolated using the RNeasy Mini kit (Qiagen) from tissues of at least two independent embryos or adult animals. cDNA synthesis was performed using SuperScript II Reverse Transcriptase and Oligo(dT)$_{12\text{-}18}$ primer according to the manufacturers' instructions (Invitrogen). Products were quantified by qPCR, using Platinum Taq DNA polymerase (Invitrogen) and SYBR Green (Sigma) on an Opticon 2 Real-Time PCR Detection System (BioRad). Primer sequences in **Table 6-1**. Transcript levels of normally expressed genes were normalized to the *Hprt1* transcript, encoding a relatively high expressed housekeeping gene on an unrelated chromosome, verified with Affymetrix gene expression analysis not to be influenced by the integration of the human β-globin LCR (results not shown). Transcript levels of extremely high expressed genes (β^{maj}, β^{min}, α-globin) were normalized to the *Gapdh* transcript.

6.3.4- Cell preparation

E14.5 fetal livers and E10.5 embryonic blood were disrupted in PBS to obtain single-cell suspensions. Cells were fixed onto poly-L-lysine coated slides in 3.7% formaldehyde / 5% acetic acid for 20 min at room temperature and subsequently washed 3 times in PBS for 5 min. Cells were stored at -20°C in 70% ethanol.

6.3.5- RNA-FISH

RNA-FISH was performed as described previously [23], with some modifications. Slides were pretreated for hybridization by a 0.01% pepsin digestion in 0.01 M HCl for 4 min at 37°C. Formaldehyde fixation was done using a 3% formaldehyde / PBS mixture for 5 min at RT. The hybridization mixture contained 1.0 ng/μl of each DNP or digoxygenin labelled oligonucleotide (**Table 6-1**).

Gene expression analysis

Primer set	Sequence
mRNA Hprt	AGCCTAAGATGAGCGCAAGT ATGGCCACAGGACTAGAACA
mRNA Scyl3	CTTACCATCTGGACTTGCTG GGGTGACGGAGTGTCTTTA
mRNA Cdc20	TGCTCCATCCTCTGGTCT CGTGCTGTGTGTCCTTTG
mRNA Lsm14a	CCACCCAAACCACAATGT GGACTGAACTGACTGTATGC
mRNA βh1	TGGACAACCTCAAGGAGAC AGTAGAAAGGACAATCACCAAC
mRNA Lactb	CGTGGTTGGAGTTTCTGTAG TGCTGATGCTTGCGATTC
mRNA εy	GAACTTGTCCTCTGCCTCT ATCACCAGCACATTACCCA
mRNA Gapdh	TTCACCACCATGGAGAAGGC GGCATGGACTGTGGTCATGA
mRNA βmaj	ATGCCAAAGTGAAGGCCCAT CCCAGCACAATCACGATCAT
mRNA βmin	ATCCCAAGGTGAAGGCCCAT CCCAGCACAATCACGATCGC

Primer set	Sequence
mRNA ζ-globin	GAGAGAGCTATCATCATGTCC AAGTAGGTCTTCGTCTGGG
mRNA α-globin	TGGCCATGGTGCTGAATATG TCTTGCCGTGACCCTTGAC
mRNA Dand5	CTTCTACATTCCCAGCTCG CTGGACCAATACCGTGGA
mRNA Gadd45gip1	AAAGCAGAAGCGAGAACG ATAGCAGCAATTCGTGCC
mRNA CaIR	GACTTTCTGCCACCCAAG GTTCCCACTCTCCATCCA
mRNA Gcdh	CTGCCGATGAGAAACTGATA TCGACCTGTAGCCACTGTC
mRNA Dnase2a	TGCCAATCCTTGCAAACT CGACCAACCTCCTAAATCC
mRNA Mast1	TGGAAGGTGGTGACTGTG AATTGTGTAAGTACTCAAGGGC
Primary transcripts βh1	TCTGGGAGTTGAGACTGTGA TGGACCCATGGACTCTAACA

RNA-FISH

Probe	Sequence
βh1 mRNA	GGCATCATAGACACATGGGATTGCCAGTGTACT AACAAAATTTGTGCTCTCAATGCCTAATCCAGT
α-globin mRNA	TTCCCCCAGGCAGCCTTGATGTTGCTTTTTG TGGAGGTCAGCACGGTGCTCACAGAAGGCAA
βh1 primary transcript	AAAACCCTATAGAAAACCCTGGAAATTTCTGCCATGCATAAGGATAATT GGACCCATGGACTCTAACATCTGACAAGGCATTGCCAATCACAGTCTC AATGCTGGGCGCTCACTCAAATCTGCACCCAAATCATTGTTGCCCACA CATAGATGTATTAATTTATAAAAACATACTCCTTTTTAAAAAAGATCC
βmaj primary transcript	AAAGGAGGAGGGGAAGCTGATATCAGGATGGGAAGTAAATAACCAGCT AAACAGGGACATATCTTCCTTGTCCTCTGAGCAAGTTACAAGGCAATA TCAAACAGAATTTATATGTAAAATATATTCTTCCCCTGTCACCCTGGCA CATAACTGTAGAGCAAAAATACCAGATACTGCAGGCTTATTTACAAAG

Table 6-1 Primers and probes

6.3.6- Microscopy

Images were collected with a Zeiss Axio Imager Z1 epifluorescence microscope (100x plan apochromat, 1.4 oil objective) equipped with a charge-coupled device (CCD) camera. RNA-FISH images were analyzed with Zeiss AxioVision software (Zeiss). Filters used: DAPI, (Zeiss) FITC, AF594 (Chroma). No bleedthrough was detected and images were collected without saturation of intensities.

6.4- Results

6.4.1- The human β-globin LCR at 8C3/C4 trans-activates an endogenous β-globin gene

To address whether the human β-globin LCR at 8C3/C4 can upregulate genes beyond the region directly surrounding its integration site, we performed an Affymetrix gene-expression analysis. Expression data from E14.5 fetal liver (FL) samples containing the LCR integrated at 8C3/C4 in the anti-sense orientation (see [21]) were compared to WT littermate samples, using a cut-off rate of twofold upregulation. Both datasets were highly similar, revealing only 11 probe-sets (representing 8 genes) being 2 fold upregulated in the LCR samples (**Figure 6-1A, B** and **Table 6-2**). Previously, we reported that ectopic integration of the β-globin LCR at 8C3/C4 results in upregulation of genes up to 150 kb away from the integration site [21]. Reassuringly, three of these genes were among the top 8 genes identified on the micro-array (**Figure 6-1A**). Other genes previously found upregulated at 8C3/C4 either lacked probes on the micro-array (*Dnase2a*), scored below threshold (*Mast1*, *Dand5*) or were scored just below 2-fold upreglulation (*Gcdh*).

The five remaining genes that were identified were all located on a different chromosome than chromosome 8 that contained the LCR (**Figure 6-1B** and **Table 6-2**). Upregulation of their expression was verified by quantitative reverse transcriptase PCR (qRT-PCR) (**Figure 6-1C**). Four out of five genes were not confirmed to be upregulated. Retrospectively, this can be explained by their low micro-array signals, which are known to often give false positive signals (**Table 6-2**). The one remaining gene was the βh1 gene, which is an endogenous mouse β-globin-like gene located on chromosome 7. This gene is known to express at high levels in embryonic blood earlier during development, but activity of the gene drops dramatically in FL later during differentiation [22]. However, expression of the βh1 gene in E14.5 FL was readily detectable on the micro-array. The gene is represented by three different probesets and each of them showed counts significantly above background in WT E14.5 FL. Sequence alignment confirmed these probes to be specific for βh1. All three probesets measured βh1 mRNA levels consistently higher (around 1.7-2.3 fold) in transgenics homozygous for LCR-AS (LCR-AS +/+) as compared to their WT littermates. An independent primerset specific for βh1 messenger RNA was used to verify this observation by qRT-PCR. While in E14.5 FL the βh1 gene showed only marginal activity compared to the highly active β^{maj} gene [22], βh1 messenger was readily detected at this stage also by qRT-PCR. In strong agreement with the micro-array results,

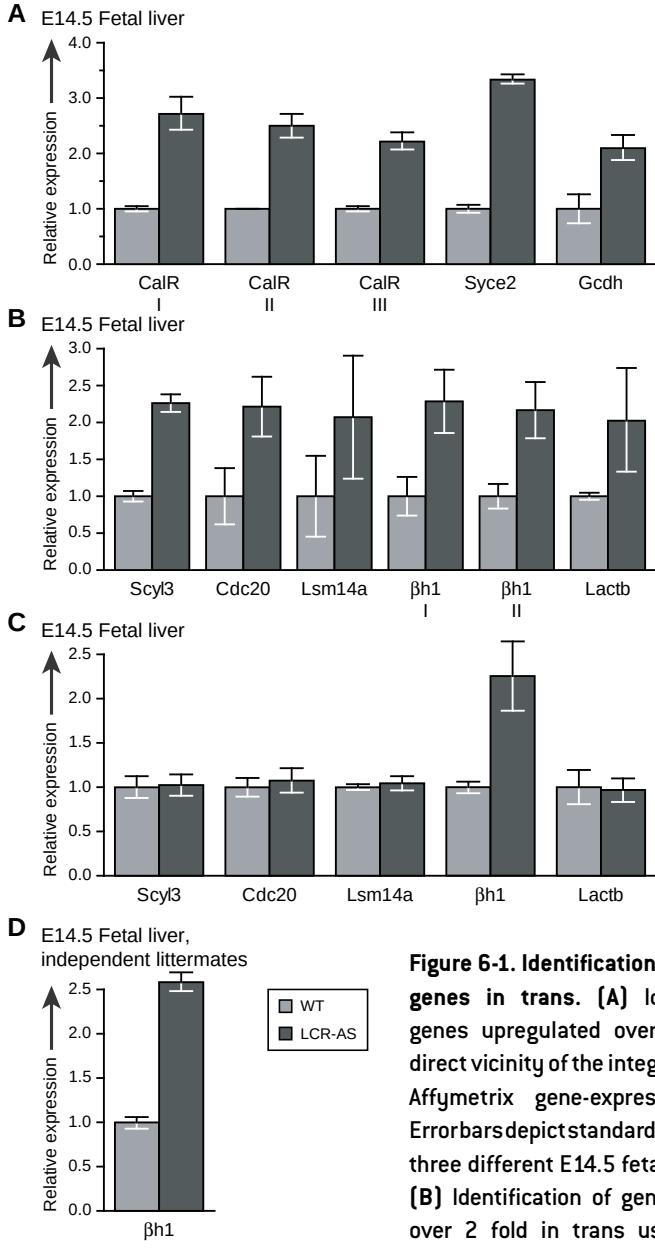


Figure 6-1. Identification of upregulated genes in trans. (A) Identification of genes upregulated over 2 fold in the direct vicinity of the integrated LCR using Affymetrix gene-expression analysis. Error bars depict standard errors (SE) from three different E14.5 fetal liver samples. **(B)** Identification of genes upregulated over 2 fold in trans using Affymetrix gene-expression analysis. Error bars depict SE from three different E14.5 fetal liver samples. **(C)** Determination of mRNA steady state levels in E14.5 FL using qRT-PCR. Error bars depict SE from at least three different samples. **(D)** Determination of mRNA steady state levels in E14.5 FL using qRT-PCR. Error bars depict SE from at least three different, indepentdent samples.

qRT-PCR revealed a 2-3 fold upregulation of βh1 gene expression in each transgenic over WT FL (**Figure 6-1C**). Thus, four different sets of oligonucleotides designed specifically for βh1 (three on the micro-array, one for qRT-PCR) independently showed the same effect, minimizing the chance that results are influenced due to cross-hybridisation with, for example, other globin sequences or human LCR read-through transcripts. To verify that we were not studying an artefact caused by accidental differences in genetic backgrounds between littermates we performed another round of back-crossing between homozygous transgenics and WT animals. The obtained heterozygotes were mated and littermate offspring was again analysed for gene expression. Once more βh1 expression levels, as measured by qRT-PCR, were found to be 2-3 fold higher consistently in LCR-AS +/+ embryos compared to WT littermates (**Figure 6-1D**).

>2 fold upregulated Affymetrix probes on chromosome 8:

Gene	Location		Affymetrix ID	WT		LCR-AS	
	Chromosome	Chromosomal band		Average probe intensity	Standard deviation	Average probe intensity	Standard deviation
CalR I	8	8C4	1456170_x_at	2225.7	145.9	6048.5	1139.8
CalR II	8	8C4	1433806_x_at	1917.9	11.5	4794.0	722.7
CalR III	8	8C4	1417606_a_at	2359.7	221.7	5252.3	622.7
Syce2	8	8C4	1429270_a_at	722.0	81.6	2414.6	96.3
Gcdh	8	8C4	1448717_at	121.1	53.2	254.9	45.8

>2 fold upregulated Affymetrix probes on other chromosomes:

Gene	Location		Affymetrix ID	WT		LCR-AS	
	Chromosome	Chromosomal band		Average probe intensity	Standard deviation	Average probe intensity	Standard deviation
Scyl3	1	1H1	1434365_a_at	72.6	7.5	164.0	14.7
Cdc20	4	4D1	1439394_x_at	66.8	45.3	147.3	46.9
Lsm14a	7	7B1	1428437_at	104.4	99.4	217.1	150.3
βh1 I	7	7E3	1437810_a_at	837.2	378.6	1911.7	623.1
βh1 II	7	7E3	1437990_x_at	786.0	212.9	1695.5	516.1
Lactb	9	9D	1449014_at	51.5	4.0	104.9	62.1

Affymetrix probes representing globin genes:

Gene	Affymetrix ID	WT		LCR-AS		LCR-S		Aγ-AS	
		Average probe intensity	Standard deviation	Average probe intensity	Standard deviation	Average probe intensity	Standard deviation	Average probe intensity	Standard deviation
εy I*	1450621_a_at	7351.1	956.5	6502.5	1496.1	6664.7	1290.0	8098.4	1454.4
εy II*	1436823_x_at	9031.2	494.8	8544.0	759.3	8603.2	586.5	9401.9	226.6
εy III*	1436717_x_at	9256.4	506.8	8760.0	745.8	8826.3	770.0	9549.3	244.9
βh1 I	1437810_a_at	837.2	378.6	1911.7	623.1	1988.2	272.5	2792.6	881.5
βh1 II	1437990_x_at	786.0	212.9	1695.5	516.1	2121.7	97.4	2787.6	912.1
βh1 III	1450736_a_at	714.9	298.6	1223.6	438.7	1085.5	372.3	1680.1	529.8
βmaj*	1417184_s_at	8494.6	272.6	8381.6	796.1	9106.1	160.6	8533.3	543.7
ζ-globin	1448716_at	2384.3	385.0	1698.7	479.4	2428.6	598.5	4148.4	1742.3
α-globin I*	1428361_x_at	10319.1	670.6	10068.4	656.9	10136.9	707.1	9582.9	970.5
α-globin II*	1452757_s_at	9738.7	239.0	9639.9	568.1	9625.8	297.1	9493.1	292.3

* βmaj, α-globin and potentially εy probes are expected to be saturated. Probe intensities should therefore not be considered informative for the detection of upregulation.

Table 6-2 Affymetrix expression levels

We therefore conclude that integration of the human LCR at 8C3/C4 on chromosome 8 causes upregulation of βh1 gene expression on chromosome 7. The data also show that, besides a number of functionally unrelated genes which directly surround the integrated LCR, βh1 is the only other gene in the genome that appreciably responds to the integration of the LCR. Since βh1 is a natural target gene of the mouse LCR, its upregulation is highly suggestive to be a consequence of the ectopic LCR acting directly on the gene.

6.4.2- Both the LCR and an ectopically integrated Aγ-globin gene upregulate the endogenous βh1 gene

We previously also generated mice carrying the human LCR in the opposite orientation (LCR-S; see [21]) at 8C3-C4, as well as mice carrying the human Aγ-globin gene in

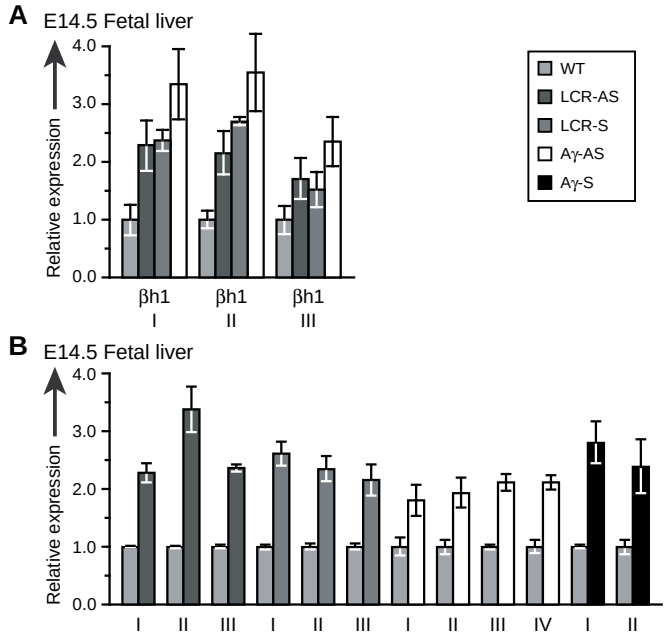


Figure 6-2. Both the human β-globin LCR and the Aγ-globin gene can upregulate the endogenous mouse βh1 gene. (A) Affymetrix gene-expression analysis of mouse βh1 probesets present on the micro-array. Error bars depict SE from three different E14.5 fetal liver samples. WT 8C3/C4 and 8C3/C4 LCR-AS samples are from littermates. 8C3/C4 LCR-S and 8C3/C4 Aγ-AS samples are unrelated samples, which show an increased overall experimental noise (not shown). **(B)** Determination of βh1 mRNA steady state levels in individual E14.5 fetal livers of littermates with different genotypes using qRT-PCR. Error bars depict standard deviations from at least two independent experiments.

either orientation at 8C3/C4 (gamma-S and gamma-AS; see **Section 5.4.1**). To determine whether the orientation of the LCR influenced βh1 upregulation and whether the effect was specific for the LCR, we analysed micro-array data from LCR-S +/+ and gamma-S +/+ E14.5 FLs. The data revealed that the βh1 gene was also upregulated by the LCR in the sense orientation, showing that LCR-orientation is not important for upregulation *in trans* of the βh1 gene. Highly intriguingly though, the data also showed upregulation of the βh1 gene in mice carrying the human Aγ-globin gene on chromosome 8 (**Figure 6-2A**). Again, to rule out the possibility that accidental changes in genetic background caused the effect, each homozygous transgenic line was back-crossed to WT and the obtained heterozygotes were subsequently mated to collect embryos. As was found for LCR-AS transgenics, LCR-S +/+, gamma-S +/+ and gamma-AS +/+ embryos each showed elevated βh1 gene expression in E14.5 FL compared to their WT littermates (**Figure 6-2B**). Sex of the embryos was determined and was found not to influence the results. We next considered the option that integration per se at 8C3/C4, or the knockout of *Rad23a*, was underlying the observed effect on globin gene expression. All four transgenic lines carried globin sequences in the *Rad23a* gene at 8C3/C4 and were generated based on a targeting construct made to knock-out this gene, thereby introducing along a constitutively expressed selection marker [24]. *Rad23a* encodes mHR23A, a DNA repair factor that is ubiquitously expressed. The gene can be knocked out in mice without phenotypic consequences [24]. To exclude the unlikely possibility

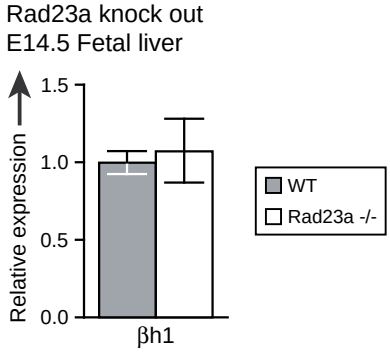


Figure 6-3. Upregulation of βh1 is not caused by *Rad23a* knockout or the introduction of a constitutively expressed selection marker at 8C3/C4. Determination of mRNA steady state levels in E14.5 FL using qRT-PCR. Error bars depict SE from at least two samples of WT and Rad23a knock out littermates.

that the *Rad23a* gene product, or the introduction of the selection marker, plays a role in βh1 gene expression, and to rule out that we had stumbled upon the first phenotype observed in *Rad23a* knockout mice, we re-examined the original knockout mice that just

carried a NEO selectable marker inserted at the same position into the *Rad23a* gene. When compared to their WT littermates (or any other WT embryo) no change in βh1 gene expression was observed (**Figure 6-3**). We conclude that deficient mHR23A function and the presence of a constitutively expressed selection marker at 8C3/C4 is not the cause of the upregulation of βh1 *in trans*. Instead, upregulation of the βh1 gene on chromosome 7 relies on the insertion of β-globin sequences on chromosome 8.

6.4.3- βh1 is the only globin gene that is upregulated in trans, as determined by micro-array and qRT-PCR analysis

After identifying one of the murine globin genes as a regulatory target of the ectopic human LCR and the Aγ-globin gene, we considered the possibility that the LCR also upregulates other globin genes in the mouse genome. Therefore, we first carefully re-analysed all globin genes represented on the Affymetrix micro-array (**Figure 6-4A**). None of the other β- or α-globin genes were significantly upregulated with the possible exception of the ζ-globin gene by the Aγ-globin gene (see below for additional analysis). Results of the Affymetrix analysis for the highly expressed globin genes is non-informative though, since probes were probably saturated (see **Table 6-2**). To further analyse these and the other globin genes we repeated the analysis on transgenics and their WT littermates by using qRT-PCR. The ζ-globin gene was found not upregulated regardless of the presence of the LCR or the Aγ-globin gene at 8C3/C4 (**Figure 6-4B**). It is therefore likely that the observed mild increase in ζ-globin gene expression measured on the micro-array was due to genetic variation between non-littermates. Similarly, analysis of the remaining globin genes showed they were not influenced by the presence of the human LCR or Aγ-globin gene at 8C3/C4 either (**Figure 6-4C**). Together, these results show that upregulation *in trans* in these transgenic animals is restricted to the βh1 gene, as determined by methods that measure steady-state mRNA levels across the cell population. This does not exclude the possibility though that in a small percentage of cells other (globin) genes also respond to the ectopic globin sequences at 8C3/C4, without appreciably affecting the overall mRNA levels measured across the cell population.

Integration of the LCR at 8C3/C4 results in an LCR-orientation dependent upregulation of genes up to 150 kb away from the integration site [21]. Six genes at 8C3/C4 (*Mast1*, *Dnase2a*, *Gcdh*, *CalR*, *Gadd45gip1* and *Dand5*) are upregulated by the LCR in both orientations. None of the corresponding gene products have been associated with β-globin gene expression and their overexpression is therefore very unlikely to result

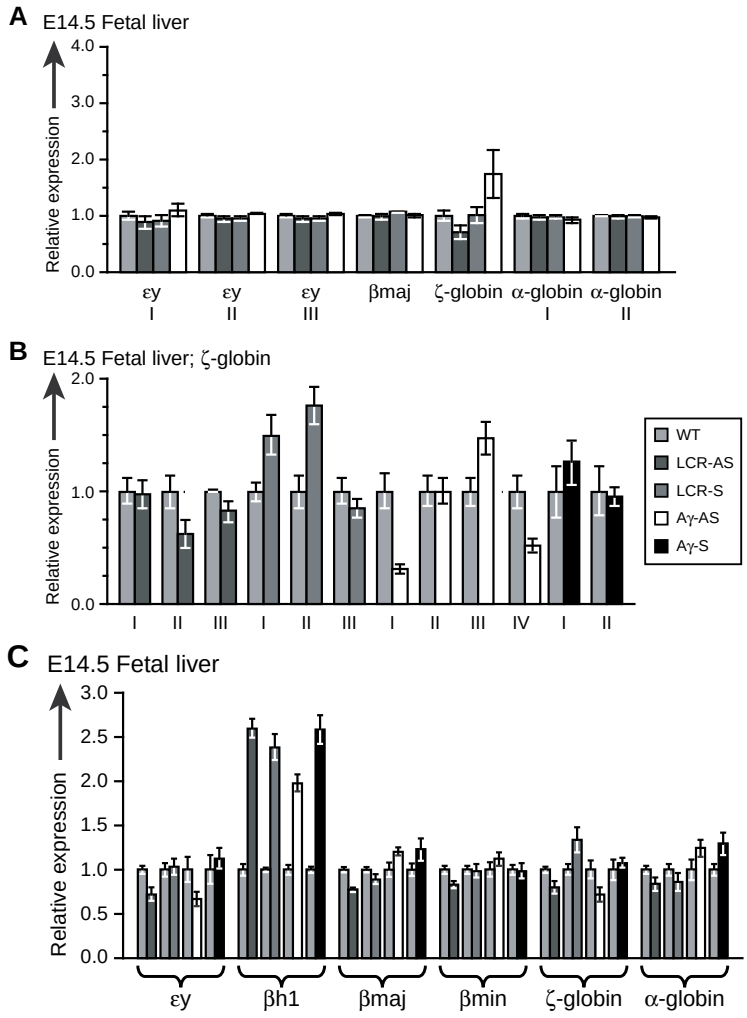


Figure 6-4. Other globin genes in the mouse genome do not profit from the human β-globin LCR and Aγ-globin gene. (A) Affymetrix gene-expression analysis of mouse β-globin and α-globin genes present on the micro-array. Error bars depict SE from three different E14.5 fetal liver samples. WT 8C3/C4 and 8C3/C4 LCR-AS samples are from littermates. 8C3/C4 LCR-S and 8C3/C4 Aγ-AS samples are unrelated samples, which show an increased overall experimental noise (not shown). **(B)** Determination of ζ-globin mRNA steady state levels in individual E14.5 fetal livers of littermates with different genotypes using qRT-PCR. Error bars depict standard deviations from at least two independent experiments. **(C)** Determination of globin mRNA steady state levels in E14.5 FL using qRT-PCR. For comparison, βh1 mRNA steady state levels have also been included. Error bars depict SE from at least two WT or transgenic littermates.

in an exclusive upregulation of βh1. In fact, the one erythroid-specific gene, encoding the transcription factor Klf1, that is located at 8C3/C4 does not respond to any of the LCRs and expresses at WT levels [21]. To further investigate if the upregulation of any of these genes may account for the observed βh1 gene activation we assessed their expression level in transgenic mice containing the human Aγ-globin gene in the anti-sense orientation (**Figure 6-5**). qRT-PCR revealed that these six genes were not

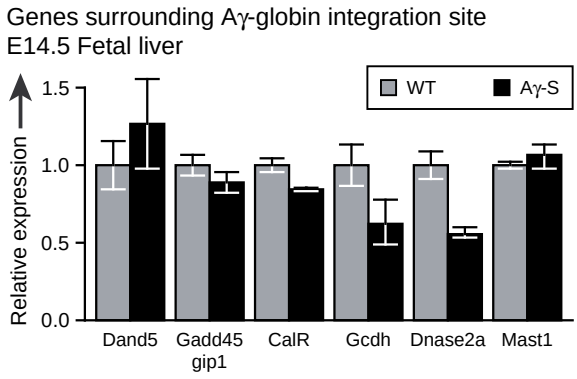


Figure 6-5. Upregulation of βh1 is not caused by the upregulation of genes surrounding the integration site at 8C3/C4. Determination of mRNA steady state levels of genes in 8C3/C4 surrounding the integration site of the Aγ-globin gene and previously reported to be upregulated by the integrated LCR. Analysis was done in E14.5 FL using qRT-PCR. Error bars depict SE from at least two samples

upregulated, showing that the insertion of the Aγ-globin gene, unlike the LCR, had no effect on the expression of surrounding genes. This ruled out that upregulated genes at 8C3/C4 accounted for activation of the βh1 gene. Together with the notion that the only upregulated gene found in the entire genome is a natural target of the β-globin LCR and the mouse equivalent of the human Aγ-globin gene, upregulation of βh1 *in trans* most likely is a direct effect of the integration of human β-globin sequences at 8C3/C4.

6.4.4- Upregulation of βh1 originates in later stages of embryonic development and is maintained in adult tissue

To get additional insight into how upregulation *in trans* may function, we analysed βh1 expression at several time points during mouse development. During normal development, the βh1 gene is highly active during primitive erythropoiesis in yolk-sac derived embryonic blood [22]. To assess whether βh1 is consistently upregulated throughout development, including at this early stage where the gene is highly active, messenger RNA levels of the gene were determined in WT and transgenic E10.5

embryonic blood (**Figure 6-6A**). At this early stage, upregulation *in trans* of the βh1 gene was not detected. To exclude that the effect measured at E14.5 was reflecting an excess of remaining primitive erythroid cells in the transgenic animals, we measured ongoing transcription in E14.5 FL by determining primary transcript levels of the βh1 gene by qRT-PCR (**Figure 6-6B**). Consistently, when the LCR was present at 8C3/C4, ongoing transcription of the βh1 gene was increased at E14.5. Thus, yet another primerset that hybridised to intronic sequences confirmed the upregulation of βh1 and showed that the βh1 gene is indeed more active in transgenic animals at this developmental time point.

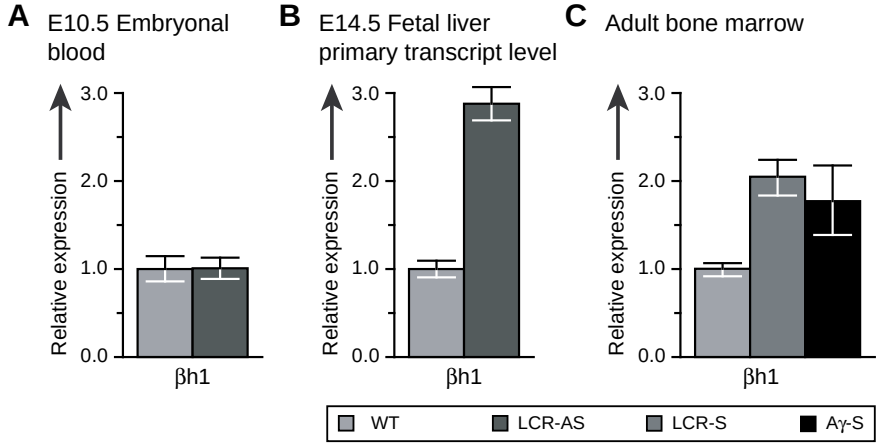


Figure 6-6. Upregulation of βh1 is established in later stages of development and is maintained in adult tissues. (A) Determination of mRNA steady state levels in E10.5 embryonic blood using qRT-PCR. Error bars depict SE from at least two independent samples. **(B)** Determination of primary transcript levels in E14.5 FL using qRT-PCR. Error bars depict SE from at least four independent samples. **(C)** Determination of mRNA steady state levels in adult bone marrow using qRT-PCR. Error bars depict SE from at least two independent samples.

Next we wondered whether upregulation of βh1 is maintained during mouse development, or whether it is a representation of LCR and Aγ-globin regulatory capacity at E14.5. To test this, we determined expression of the βh1 gene in bone marrow of adult transgenic and WT animals of similar age (**Figure 6-6C**). βh1 expression in bone marrow was reliably detected by qRT-PCR, although the observed CT-values were considerably lower than in E14.5 FL, which suggests that absolute expression of the gene is further decreased, but not absent, in adult erythroid tissue (not shown). Importantly though, the ectopic

human LCR and Aγ-globin gene still boost the expression level of the βh1 gene in bone marrow of transgenic mice.

6.4.5- Single cell analysis of βh1 gene expression

In the endogenous locus the β-globin LCR upregulates β-globin genes 25-100 fold through physical contact [1-3]. To understand the mechanism behind interchromosomal activation of the βh1 gene by ectopic β-globin sequences, we designed experiments to study their interaction in single cells under the microscope. Previously, we determined co-localization frequencies of the mouse β-globin locus with 8C3/C4 using high-resolution Cryo-FISH (see **Figure 5-5A** and **Table 5-2**). Regardless of the presence of the LCR at 8C3/C4, co-localization frequencies of around 2.5% between the two alleles were observed (i.e. in about 5% of the cells). Indeed, the LCR was not found to search for the

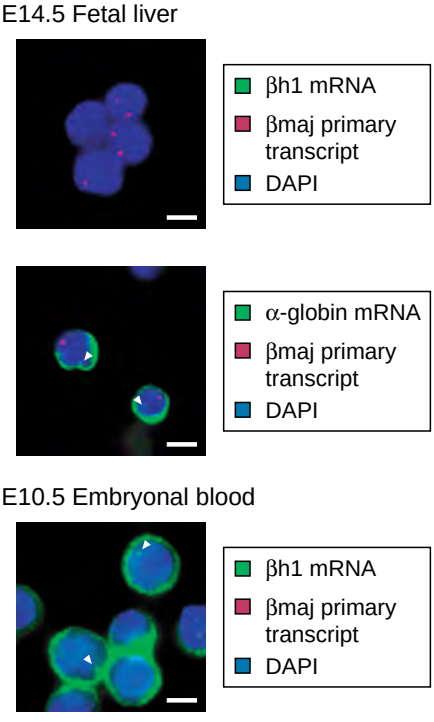


Figure 6-7. Lack of detection of βh1 at E14.5 is not a technical problem. Examples of globin expressing cells. Probes used for visualization of globin transcripts and messengers are indicated next to each picture. The βh1 messenger is not detected in E14.5FL (top). The same antibody detection scheme is able to visualize the β-globin messenger (middle), while the same βh1 probe does detect the βh1 messenger in E10.5 embryonal blood (bottom). Images obtained by RNA-FISH. Arrowheads indicate primary transcript signals detected by mRNA probe. Scale bars: 5 μm.

endogenous β-globin locus, nor did it stabilize this interaction. We hypothesized though that this small percentage of interacting loci may account for the overall increase in βh1 expression levels measured across the cell population. If true, high βh1 expression levels are expected specifically in cells showing contacts between the two interchromosomal

regions. To address this hypothesis, we first performed RNA-FISH on E14.5 FL cells and looked for the presence of βh1 expressing "jackpot" cells (**Figure 6-7**). Visual inspection of a large number of E14.5 FL cells, both with and without the integrated LCR, failed to identify cells with βh1 mRNA signals accumulated around the nucleus. βh1 mRNA signals were easily detected in E10.5 embryonic blood, a tissue where the gene is highly active, ruling out probe-dependent detection problems. From these preliminary experiments two conclusions may be drawn: (i) upregulation of the βh1 gene is a very rare event and/or (ii) upregulation of βh1 does not result in comparably high expression levels as detected in E10.5 embryonic blood, thereby failing to reach detectable levels. FISH experiments with increased sensitivity and with probes designed to detect primary βh1 transcripts will now be performed to identify potential βh1 expressing jackpot cells, and/or to compare βh1 gene activity between alleles that are separate or interacting with 8C3/C4 in transgenic FLs.

6.5- Discussion

In this study, we determined the potential of an ectopically integrated β-globin LCR to upregulate genes *in trans*. An initial micro-array expression analysis identified the endogenous β-globin-like βh1 gene as the most prominent upregulated gene *in trans*. Previously we studied the effect of the ectopic LCR on an Aγ-globin gene integrated into the homologous allele at developmental stages where it is highly active, and found no effect of the LCR on transcription (see **Chapter 5**). While the latter is in agreement with the idea that there is no tendency for homologous alleles to interact, it is interesting to mention that 4C identified interactions *in trans* between 8C3/C4, with and without the LCR, and the endogenous β-globin locus, though in both cases only in one of two replicate experiments (see **Figure 5-7**).

In E14.5 FL the βh1 gene is lowly expressed and supposed to not interact with the endogenous mouse β-globin LCR [2, 22, 25]. LCR function has been proposed to depend on promoter-competition and distance from the target gene ([22, 26-28]; see also **Chapter 2**). In WT E14.5 FL, interactions between the endogenous LCR and the adult β^{maj} and β^{min} genes are highly favoured over interactions with the embryonic βh1 and εy genes [2, 25]. In this study we show that the ectopic human LCR is able to override this system, thereby allowing the upregulation of the βh1 gene, but not the εy gene. We envision two possible models allowing this effect. In the first model, the presence of the Active Chromatin Hub (ACH; [2]) is accompanied by a loop protruding from this chromatin structure that contains the βh1 gene. Spurious co-localization of the ectopic LCR close to the loop containing the βh1 gene allows the formation of a functional interchromosomal interaction and a strong increase in transcriptional activity of the βh1 gene. Importantly, the εy gene in this model should be inaccessible to the human LCR, which may be due to its location near to the endogenous LCR. In the second model for the transcriptional regulation *in trans*, the βh1 gene and the human LCR bind protein factors that allow functional interactions, but these factors are lacking on the endogenous mouse LCR. These factors would allow the βh1 gene to interact with the human LCR, but not the endogenous mouse LCR, which allows it to compete with the adult globin genes for upregulation by the human LCR. In both models, the ectopic human LCR could also interact with the adult mouse β-globin gene at moments when the gene is not interacting with the endogenous LCR. If these are rare events, the high expression levels of adult β-globin genes in E14.5 FL may not allow appreciation of this effect on steady-state mRNA levels. It would allow co-expression of both adult β-globin genes from one

allele though, which is proposed not to occur in the WT β-globin locus [22, 23]. FISH experiments should provide insight on whether co-expression occurs when the ectopic LCR is nearby. If so, one may expect this to not happen with the human Aγ-globin gene integrated at 8C3/C4.

The ability of the human LCR to upregulate the βh1 gene raises the question how the ectopic LCR selects genes for their upregulation. In **Chapter 5** we reported that 4C identifies interactions between 8C3/C4 and the β-globin locus in one of the two replicate 4C experiments, though co-localization frequencies of 8C3/C4 with the mouse β-globin locus, measured by Cryo-FISH, are only marginally above background levels. Furthermore, 4C-interactions and Cryo-FISH co-localization frequencies are not dependent on the presence of the LCR at 8C3/C4. These results suggest that the LCR does not actively search for the βh1 gene, which is similar to the conclusion of the transvection experiments described in **Section 5.4.1**. Assuming that the observed upregulation of βh1 depends on interchromosomal contacts, which we still need to experimentally address, then our data are in agreement with studies showing that the compatibility between regulatory elements and gene promoters importantly influences the outcome of an interaction [28, 29]. Functional transcriptional regulation *in trans* therefore depends on an LCR dependent component (promoter affinity for the LCR) and an LCR independent component (interaction frequency, most likely determined by chromatin characteristics of both interaction partners ([30] and **Paragraph 5.5**)).

The finding that the LCR on one chromosome is able to upregulate a gene on a different chromosome provides important insight in LCR functioning. Several models for LCR functioning exist in which signals are either transported along the chromatin template towards the β-globin genes (spreading model) or where the intervening chromatin would be required for chromatin loop formation (facilitated tracking model) [4-6]. The results in this chapter show that physical linking of the LCR to target genes is not required for transcriptional activation, thereby essentially ruling out these models. Functional interactions rather seem to be spuriously established between high affinity genomic loci that are in each others close vicinity. Due to the small genomic distance and mutual affinity between the LCR and the endogenous β-globin genes, such interactions will occur frequently at the endogenous β-globin locus. The small genomic distance between the ectopically integrated LCR and the genes at 8C3/C4 similarly allows their upregulation [21], but their non-optimal affinity results in relatively mild levels of upregulation and allows for a more promiscuous behaviour of the LCR. Increasing distance from the

integration site decreases the chance to collide, thereby restricting the observable effect to only a limited genomic region.

How to relate these results to the previously reported examples of transcriptional regulation *in trans* [13-17]? So far in all examples of transcriptional regulation, with the exception of [17], a pan-cellular effect is assumed, even though interaction frequencies are considerably lower than 100%. In the introduction, transcriptional regulation *in trans* was therefore determined to depend on transient interactions, a long term memory of the regulatory effect and considerable chromatin mobility. The observation that the integrated LCR does not increase co-localization between 8C3/C4 and the endogenous β-globin locus, indicates that transcriptional regulation *in trans* is dependent on spurious association between loci. Combined with previous reports that showed that the mobility of chromatin is highly constrained during interphase [31, 32] and that the position in the nucleus of individual loci is mainly determined during the process of mitosis [33-37], it seems unlikely that interchromosomal communication could be applied as a faithful pan-cellular regulatory mechanism. Rather, directional movement of the involved loci would be required, which has previously been proposed to be driven by recruitment to sites dedicated to transcription of functionally related genes [38-40]. But again, such a mechanism is not in agreement with our observations that introduction of the LCR in 8C3/C4 does not change intrachromosomal and interchromosomal interaction partners and has no apparent effect on clustering with functionally related genes (see **Chapter 5**). It seems therefore unlikely that transcriptional regulation *in trans* would be mediated by active recruitment to such nuclear substructures.

The finding that the human Aγ-globin gene can upregulate the βh1 gene opens up the intriguing possibility that genes and their promoters can act as enhancers *in trans*. It has previously been noted that genomic sites defined as enhancers, promoters, silencers or insulators may act differently depending on their genomic context ([41] and references therein). Previous observations were published that may resemble our observations [42]. In this study, plasmids containing actively transcribed human globin genes, including the Aγ-globin gene, induce intergenic transcription from the human β-globin locus in normally non-expressing human cells. Genic transcription is not observed in this study though, which is an important difference from the results in this study. The experimental setup in the two studies is highly different, which may explain the observed differences. The question now is how to explain the enhancer function *in trans* of the Aγ-globin gene. A shared property of enhancers and promoters is that they

both contain high concentrations of transcription factor binding sites, with the difference being that promoters contain a transcriptional start site and are located immediately upstream from the coding region of genes [41]. Transcription from the ectopic Aγ-globin gene is still detected at E14.5, though it is severely decreased compared to earlier time points (not shown). We hypothesize that when the βh1 gene and 8C3/C4 are co-localizing, transcription of βh1 benefits from the transcription factors brought in by the Aγ-globin gene. Future studies will hopefully gain insight in this mechanism, and may also reveal whether simultaneous activity of the βh1 and Aγ-globin gene in possible. It would be interesting to see what the outcome will be if both the human LCR and the Aγ-globin or one of the human adult globin genes are inserted at 8C3/C4 together. Gene competition experiments would suggest that βh1 no longer benefits from the presence of these globin sequences *in trans*, as the ectopic LCR is kept busy by the nearby globin genes. However, one may also hypothesize that the presence of both elements near the endogenous β-globin locus would cause a further increase in local concentration of relevant transcription factors [43, 44] from which βh1 could benefit even more. The results open up the intriguing possibility that promoters not only regulate their downstream located gene, but may also influence the expression of other, most likely related genes ([41] and references therein). Future research should reveal whether such regulatory networks exist, or whether the effect we observed is an anomaly, induced by the integration of human sequences in the mouse genome.

6.6- Accession numbers

Micro-array data used for **Figures 6-1**, **6-2** and **6-4**: Gene Expression Omnibus (GEO) GSE5891 (E14.5 WT fetal liver). E14.5 fetal liver LCR-AS, LCR-S and Aγ-AS data sets will be submitted upon publication.

6.7- Acknowledgements

We thank Mariette van de Corput for assistance with RNA-FISH and Frank Grosveld for discussion.

6.8- References

1. Epner, E., Reik, A., Cimbora, D., Telling, A., Bender, M.A., Fiering, S., Enver, T., Martin, D.I., Kennedy, M., Keller, G., and Groudine, M. (1998). The beta-globin LCR is not necessary for an open chromatin structure or developmentally regulated transcription of the native mouse beta-globin locus. Mol Cell *2*, 447-455.
2. Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F., and de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. Mol Cell *10*, 1453-1465.
3. Carter, D., Chakalova, L., Osborne, C.S., Dai, Y.F., and Fraser, P. (2002). Long-range chromatin regulatory interactions in vivo. Nat Genet *32*, 623-626.
4. Blackwood, E.M., and Kadonaga, J.T. (1998). Going the distance: a current view of enhancer action. Science *281*, 60-63.
5. Bulger, M., and Groudine, M. (1999). Looping versus linking: toward a model for long-distance gene activation. Genes Dev *13*, 2465-2477.
6. Palstra, R.J., de Laat, W., and Grosveld, F. (2008). Beta-globin regulation and long-range interactions. Adv Genet *61*, 107-142.
7. Pirrotta, V. (1999). Transvection and chromosomal trans-interaction effects. Biochim Biophys Acta *1424*, M1-8.
8. Cremer, T., and Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nat Rev Genet *2*, 292-301.
9. Bolzer, A., Kreth, G., Solovei, I., Koehler, D., Saracoglu, K., Fauth, C., Muller, S., Eils, R., Cremer, C., Speicher, M.R., and Cremer, T. (2005). Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. PLoS Biol *3*, e157.
10. Xu, N., Donohoe, M.E., Silva, S.S., and Lee, J.T. (2007). Evidence that homologous X-chromosome pairing requires transcription and Ctcf protein. Nat Genet *39*, 1390-1396.
11. Bacher, C.P., Guggiari, M., Brors, B., Augui, S., Clerc, P., Avner, P., Eils, R., and Heard, E. (2006). Transient colocalization of X-inactivation centres accompanies the initiation of X inactivation. Nat Cell Biol *8*, 293-299.
12. Liu, H., Huang, J., Wang, J., Jiang, S., Bailey, A.S., Goldman, D.C., Welcker, M., Bedell, V., Slovak, M.L., Clurman, B., Thayer, M., Fleming, W.H., and Epner, E. (2008). Transvection mediated by the translocated cyclin D1 locus in mantle cell lymphoma. J Exp Med *205*, 1843-1858.
13. Spilianakis, C.G., Lalioti, M.D., Town, T., Lee, G.R., and Flavell, R.A. (2005). Interchromosomal associations between alternatively expressed loci. Nature.
14. Ling, J.Q., Li, T., Hu, J.F., Vu, T.H., Chen, H.L., Qiu, X.W., Cherry, A.M., and Hoffman, A.R. (2006). CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. Science *312*, 269-272.
15. Lomvardas, S., Barnea, G., Pisapia, D.J., Mendelsohn, M., Kirkland, J., and Axel, R. (2006). Interchromosomal interactions and olfactory receptor choice. Cell *126*, 403-413.
16. Zhao, Z., Tavoosidana, G., Sjolinder, M., Gondor, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Sandhu, K.S., Singh, U., Pant, V., Tiwari, V., Kurukuti, S., and Ohlsson, R. (2006). Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nat Genet *38*, 1341-1347.

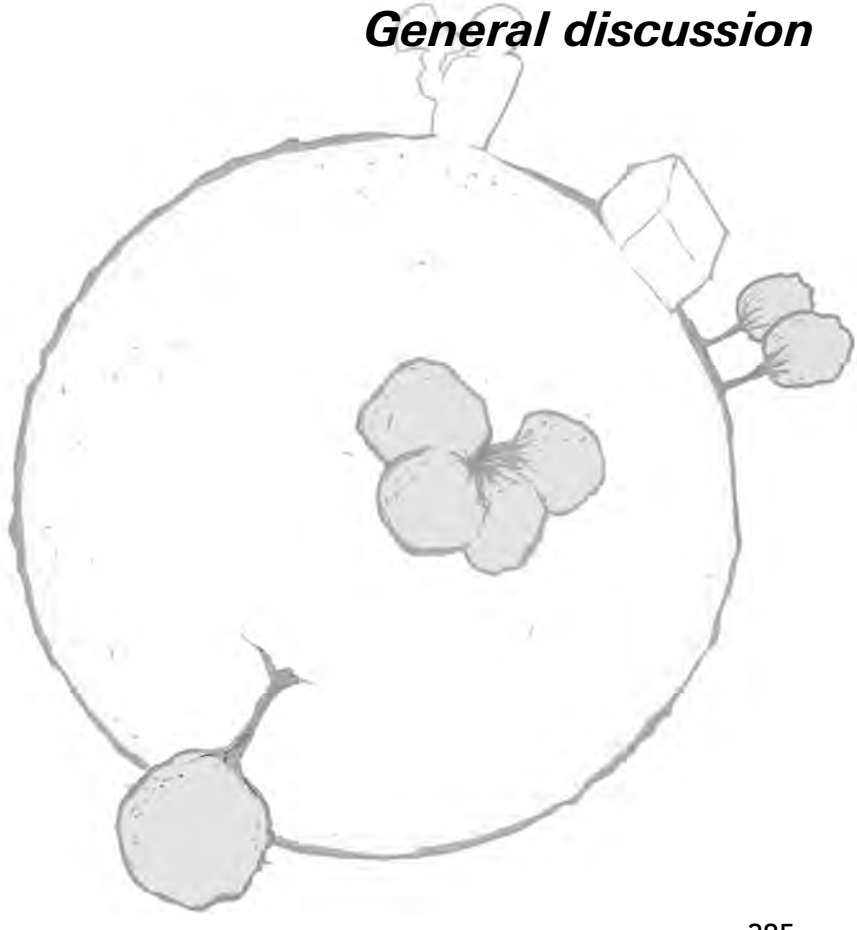
17. Apostolou, E., and Thanos, D. (2008). Virus infection induces NF-kappaB-dependent interchromosomal associations mediating monoallelic IFN-beta gene expression. Cell *134*, 85-96.
18. Nunez, E., Kwon, Y.S., Hutt, K.R., Hu, Q., Cardamone, M.D., Ohgi, K.A., Garcia-Bassets, I., Rose, D.W., Glass, C.K., Rosenfeld, M.G., and Fu, X.D. (2008). Nuclear receptor-enhanced transcription requires motor- and LSD1-dependent gene networking in interchromatin granules. Cell *132*, 996-1010.
19. Nunez, E., Kwon, Y.S., Hutt, K.R., Hu, Q., Cardamone, M.D., Ohgi, K.A., Garcia-Bassets, I., Rose, D.W., Glass, C.K., Rosenfeld, M.G., and Fu, X.D. (2008). Nuclear receptor-enhanced transcription requires motor- and LSD1-dependent gene networking in interchromatin granules. Cell *134*, 189.
20. Fuss, S.H., Omura, M., and Mombaerts, P. (2007). Local and cis effects of the H element on expression of odorant receptor genes in mouse. Cell *130*, 373-384.
21. Noordermeer, D., Branco, M.R., Splinter, E., Klous, P., van Ijcken, W., Swagemakers, S., Koutsourakis, M., van der Spek, P., Pombo, A., and de Laat, W. (2008). Transcription and chromatin organization of a housekeeping gene cluster containing an integrated beta-globin locus control region. PLoS Genet *4*, e1000016.
22. Trimborn, T., Gribnau, J., Grosveld, F., and Fraser, P. (1999). Mechanisms of developmental control of transcription in the murine alpha- and beta-globin loci. Genes Dev *13*, 112-124.
23. Wijgerde, M., Grosveld, F., and Fraser, P. (1995). Transcription complex stability and chromatin dynamics in vivo. Nature *377*, 209-213.
24. Ng, J.M., Vermeulen, W., van der Horst, G.T., Bergink, S., Sugasawa, K., Vrieling, H., and Hoeijmakers, J.H. (2003). A novel regulation mechanism of DNA repair by damage-induced and RAD23-dependent stabilization of xeroderma pigmentosum group C protein. Genes Dev *17*, 1630-1645.
25. Palstra, R.J., Tolhuis, B., Splinter, E., Nijmeijer, R., Grosveld, F., and de Laat, W. (2003). The beta-globin nuclear compartment in development and erythroid differentiation. Nat Genet *35*, 190-194.
26. Raich, N., Enver, T., Nakamoto, B., Josephson, B., Papayannopoulou, T., and Stamatoyannopoulos, G. (1990). Autonomous developmental control of human embryonic globin gene switching in transgenic mice. Science *250*, 1147-1149.
27. Hanscombe, O., Whyatt, D., Fraser, P., Yannoutsos, N., Greaves, D., Dillon, N., and Grosveld, F. (1991). Importance of globin gene order for correct developmental expression. Genes Dev *5*, 1387-1394.
28. Sabatino, D.E., Cline, A.P., Gallagher, P.G., Garrett, L.J., Stamatoyannopoulos, G., Forget, B.G., and Bodine, D.M. (1998). Substitution of the human beta-spectrin promoter for the human agamma-globin promoter prevents silencing of a linked human beta-globin gene in transgenic mice. Mol Cell Biol *18*, 6634-6640.
29. Ohtsuki, S., Levine, M., and Cai, H.N. (1998). Different core promoters possess distinct regulatory activities in the Drosophila embryo. Genes Dev *12*, 547-556.
30. Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., and de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nat Genet *38*, 1348-1354.
31. Abney, J.R., Cutler, B., Fillbach, M.L., Axelrod, D., and Scalettar, B.A. (1997). Chromatin dynamics in interphase nuclei and its implications for nuclear structure. J Cell Biol *137*, 1459-1468.
32. Chubb, J.R., Boyle, S., Perry, P., and Bickmore, W.A. (2002). Chromatin motion is constrained by association with nuclear compartments in human cells. Curr Biol *12*, 439-445.
33. Essers, J., van Cappellen, W.A., Theil, A.F., van Drunen, E., Jaspers, N.G., Hoeijmakers, J.H., Wyman,

C., Vermeulen, W., and Kanaar, R. (2005). Dynamics of relative chromosome position during the cell cycle. Mol Biol Cell *16*, 769-775.

34. Manders, E.M., Kimura, H., and Cook, P.R. (1999). Direct imaging of DNA in living cells reveals the dynamics of chromosome formation. J Cell Biol *144*, 813-821.
35. Walter, J., Schermelleh, L., Cremer, M., Tashiro, S., and Cremer, T. (2003). Chromosome order in HeLa cells changes during mitosis and early G1, but is stably maintained during subsequent interphase stages. J Cell Biol *160*, 685-697.
36. Gerlich, D., Beaudouin, J., Kalbfuss, B., Daigle, N., Eils, R., and Ellenberg, J. (2003). Global chromosome positions are transmitted through mitosis in mammalian cells. Cell *112*, 751-764.
37. Thomson, I., Gilchrist, S., Bickmore, W.A., and Chubb, J.R. (2004). The radial positioning of chromatin is not inherited through mitosis but is established de novo in early G1. Curr Biol *14*, 166-172.
38. Osborne, C.S., Chakalova, L., Brown, K.E., Carter, D., Horton, A., Debrand, E., Goyenechea, B., Mitchell, J.A., Lopes, S., Reik, W., and Fraser, P. (2004). Active genes dynamically colocalize to shared sites of ongoing transcription. Nat Genet *36*, 1065-1071.
39. Chakalova, L., Debrand, E., Mitchell, J.A., Osborne, C.S., and Fraser, P. (2005). Replication and transcription: Shaping the landscape of the genome. Nat Rev Genet *6*, 669 - 677.
40. Osborne, C.S., Chakalova, L., Mitchell, J.A., Horton, A., Wood, A.L., Bolland, D.J., Corcoran, A.E., and Fraser, P. (2007). Myc dynamically and preferentially relocates to a transcription factory occupied by Igh. PLoS Biol *5*, e192.
41. Dillon, N., and Sabbattini, P. (2000). Functional gene expression domains: defining the functional unit of eukaryotic gene regulation. Bioessays *22*, 657-665.
42. Ashe, H.L., Monks, J., Wijgerde, M., Fraser, P., and Proudfoot, N.J. (1997). Intergenic transcription and transinduction of the human beta-globin locus. Genes Dev *11*, 2494-2509.
43. Droge, P., and Muller-Hill, B. (2001). High local protein concentrations at promoters: strategies in prokaryotic and eukaryotic cells. Bioessays *23*, 179-183.
44. de Laat, W., and Grosveld, F. (2003). Spatial organization of gene expression: the active chromatin hub. Chromosome Res 11, 447-459.

Chapter 7

General discussion



7.1- Aim of the thesis

The aim of this thesis was to increase our understanding of how mammalian transcription regulatory DNA elements control the expression of genes located away on the same chromosome and how they influence chromatin organisation inside the living cell nucleus. In addition, we have explored the capacity of transcription regulatory DNA elements to activate genes on other chromosomes. We have focussed on the prototype of a strong mammalian transcription regulatory DNA element, the β-globin locus control region (LCR) and generated unique transgenic mouse models that carried an ectopic human LCR in one or the other orientation at a defined position in chromosome 8.

In **Chapter 3** we analysed gene expression and chromatin organization of WT 8C3/C4, the gene-dense region selected for LCR integration. We showed that expression and chromatin modifications within the gene cluster are importantly controlled at the level of individual genes, and that the locus frequently associates with transcription factories and is positioned at the periphery of its chromosome territory (CT).

In **Chapter 4** we showed that integration of the LCR in each orientation influenced gene expression in a partially directional matter on both sides of the integration site and over a maximum distance of 150 kilobases (kb). Using 3C, we identified long-range interactions between the LCR and two upregulated genes and proposed that LCR-gene contacts via chromatin looping determined which genes were transcriptionally enhanced. At the level of nuclear organisation, integration of the LCR causes a highly significant shift of the locus away from the CT. Furthermore, association with transcription factories increases, but this is independent of whether loci are located at the edge or outside of the CT.

In **Chapter 5** we addressed whether the LCR can actively search for preferred partners in the nuclear space. Using a transvection-like system that involved the generation of a new series of transgenic mice, we found that the LCR is not able to actively establish interactions with one of its natural target genes present on the homologous allele. Subsequently, we performed 4C to investigate whether the LCR searches for other preferred partners located elsewhere in the genome. LCR-induced positioning outside the CT results in increased encounters with genomic regions located on different chromosomes, but the specific interacting partners *in cis* and *in trans* are largely identical, irrespective of whether the LCR is integrated or not. The LCR therefore does not scan the nucleus for preferred interaction partners. Cryo-FISH showed that two existing interaction partners *in trans* that contain erythroid specific genes significantly increase their co-localization frequency, and we also observed that interaction frequencies with

the α-globin locus increase, as measured by 4C technology. Whether this indicates that the LCR stabilizes some existing interactions with functionally related genes needs to be further investigated.

In **Chapter 6** we determined whether the LCR activates genes elsewhere in the genome. In this analysis, we identified the βh1 gene as the only gene located on a different chromosome that is upregulated over twofold. Even more intriguing though, we found this same gene to be upregulated when the ectopic LCR is replaced by the human fetal Aγ–globin gene. Extensive control experiments verified that interchromosomal upregulation of the βh1 gene is indeed driven by the insertion of β-globin sequences into chromosome 8. This study therefore provides unique genetic evidence for interchromosomal gene activation in mammals. The change in expression is not accompanied by increased interaction frequencies when 8C3/C4 contains β-globin sequences since interchromosomal contacts invariably remain rare. We nevertheless hypothesize that the communication between loci on different chromosomes relies on their physical interactions and we propose combined RNA-DNA FISH experiments to address this issue.

7.2- Long-range gene activation: finding each other in the nuclear space

Results presented in **Chapter 4** and **Chapter 6** of this thesis provide important insights in the process of enhancer – promoter interactions, chromatin loop formation and long-range gene activation. The existence of chromatin loops, and their involvement in gene regulation, is well established but the mechanism behind their establishment is still under debate [1-3]. Generally, there are two models that dominate the discussion on how chromatin loops are established. The "facilitated tracking" model proposes that proteins binding distant regulatory elements scan the chromosome for target genes, thereby dragging the bound regulatory DNA element along the intervening chromatin. The "random collision" model proposes that chromatin mobility allows genomic regions to randomly contact each other. Appropriate protein factors at regulatory elements may then be allowed to form high-affinity interactions and subsequent chromatin loop formation [3, 4]. In recent years, evidence for both models has been reported. In a recent report, the binding profile of the MLL2 protein in the mouse β-globin locus was determined. Binding of the protein showed a pattern that was interpreted to reflect the protein spreading through the locus. MLL2 was therefore proposed to have a role in

chromatin loop formation via a facilitated tracking mechanism [5]. A recent study on the conformation of the mouse IgH locus generated different conclusions. Efficient long-range interactions, necessary for ordered rearrangements, are facilitated by a specific chromatin structure that strongly decreases distance between elements involved [6]. These data are highly supportive for a random collision model, where interaction frequencies are negatively correlated with distance. Similarly, both experiments in insects and plasmid-based experiments in mammalian cells have shown that a physical DNA connection between genomic elements is not necessary for their interaction. In *Drosophila*, the process of transvection allows upregulation of genes by regulatory elements located on their homologous allele [7]. Importantly, homologous pairing is a common phenomenon in *Drosophila*, but in mammals evidence for such a mechanism is generally lacking [8, 9]. In mammals, transcriptional regulation between regulatory elements located on different DNA molecules is possible though. In two studies facilitation of interactions between two plasmids, one containing an enhancer and the other a reporter gene, strongly increased transcriptional activity. One study used a streptavidin bridge to bring together two DNA molecules in human cell extracts, while the other took advantage of a *Drosophila* bridging factor to couple two plasmids transfected into human cells [10, 11]. These results show that regulatory DNA elements and genes do not need to be together on the same DNA molecule to allow communication and transcriptional activation.

Data presented in this thesis is also highly supportive of the random collision model. In **Chapter 4**, upregulation of multiple genes in 8C3/C4 is found in a region of 150 kb on either side of the integrated LCR. It is interesting to consider the effect of the LCR on the *Syce2* gene (**Figure 4-2A**). The LCR in the sense orientation fails to activate *Syce2*, while the gene is upregulated by LCR-AS. Both LCRs upregulate the genes *Gcdh* and *Dnase2a* further down the chromatin fibre. If the LCR at 8C3/C4 would act through a facilitated tracking mechanism, or another mechanism that requires propagation of a signal along the linear chromatin fibre, it is not clear how depending on the orientation of the LCR certain genes can be ignored, while genes located further downstream can be upregulated. Rather, the punctuated pattern of upregulation, combined with the restricted area of upregulation, suggests that efficient promoter – enhancer interactions are determined by a random collision model based on affinities and interactions frequencies. Data in **Chapter 6** provides further support for this model, since the observed upregulation of the βh1 gene, that is located on a different chromosome, essentially excludes the

necessity for the LCR to be linked to its target genes on the same chromosome. The βh1 gene is a natural target gene of the endogenous β-globin LCR and its bound transcription factors, and may therefore have a very strong affinity for the human β-globin LCR. We hypothesize that the human LCR may also interact with the adult mouse β-globin genes, but the high expression levels of adult β-globin genes in E14.5 fetal liver (FL) may not allow appreciation of this effect on steady-state mRNA levels. Together, our data provide strong additional evidence that establishment of interactions by the LCR is dependent on random collisions and affinity mediated enhancer – promoter interactions, rather than on signals travelling from the LCR over intervening chromatin.

Results in **Chapters 4**, **5** and **6** have provided important insight in how enhancers influence the nuclear location of their target loci and to what extent they promote clustering with functionally related genes. The β-globin Locus Control Region (LCR) can reposition the genomic region it is located in versus the nuclear periphery, CT and centromeric heterochromatin [12-14]. Similar behaviour was reported for the IgH locus, that is also repositioned versus the nuclear periphery during development [15]. A direct relationship between transcriptional regulatory elements, transcription factor binding and nuclear location has been shown, indicating that transcriptional regulatory elements may indeed have functions in nuclear repositioning [16, 17]. Relocation of genomic loci has been proposed to actually be a representation of migration to specific nuclear bodies, like transcription factories that have been proposed to support transcription of functionally related genes [18-20]. In recent years, a number of examples have been published of transcriptional regulatory elements co-localizing with functionally related genes [21-25]. These results have spawned the idea that transcriptional regulatory elements actively search the nucleus for preferred interaction partners to initiate functional interactions. In **Chapter 4** we reported that the LCR significantly repositions 8C3/C4 versus its CTs. In **Chapter 5** of this thesis we have systematically tested whether this repositioning reflects the ability of the β-globin LCR at 8C3/C4 to actively scan the nuclear environment for functionally related interaction partners. In contrast to popular belief though, we found no evidence that the β-globin LCR could actively search for a natural target gene in a transvection-like situation or for functionally related genes located on the same or different chromosomes. In **Chapter 6** further support for this finding is provided, as the only interchromosomal interaction resulting in a measurable upregulation of gene expression induced by the integrated LCR is not accompanied by changes in co-localization frequencies between the two loci involved. The data therefore

shows that productive interactions are dependent on spurious interactions between the two genomic loci that are not influenced by the presence of the LCR. Rather than a locus, gene or regulatory element individually determining its specific location in the nucleus, we favour the idea that genomic context strongly impacts on positioning. In this model, chromosomes are probabilistically folded according to chromatin and expression characteristics of all the segments they are composed of [26-28], with increased probability of clustering of genomic regions with similar characteristics. The LCR itself is not actively involved in determining the nuclear location of the target locus it is involved in, but rather exerts its influence by changing chromatin characteristics and increasing transcription in the region around its integration site.

A final issue that arises from the data in **Chapters 4**, **5** and **6** is the question how dynamic the observed enhancer – promoter contacts and intra- and interchromosomal interactions are. The LCR at the endogenous β-globin locus has been proposed to contact only one gene at a time [29, 30]. The combination of local interactions observed in **Chapter 4** and the interaction with the βh1 gene reported in **Chapter 6** are not expected to occur in all cells at the same time. Still, the question remains how dynamic in particular interchromosomal interactions are (i.e. whether during one cell-cycle in a single cell the LCR is able to contact all interaction partners or whether the formation of another interchromosomal contact requires passage through mitosis. Previous studies have shown that chromatin mobility is highly restricted during interphase [31, 32] and that the overall conformation of chromosomes is mainly determined at a certain stage in G1 [33-36]. The data in **Chapter 5** and **Chapter 6** also seem to support that interaction frequencies measured between loci *in trans* by Cryo-FISH reflect cell-to-cell differences rather than genomic mobility inside the interphase cell nucleus. After all, if globin sequences at 8C3/C4 stabilize interactions with the endogenous β-globin locus, the number of interacting alleles is expected to increase in a dynamic situation. This is not what we observe and we therefore propose interaction percentages reflect cell-to-cell differences and that interactions between the LCR and the βh1 gene may only be stabilized when the two loci are already located at a very short distance from each other. An interesting strategy aimed at answering this question more definitely could be the life-imaging of these interactions by adding *LacO*- and/or *TetR*-repeat arrays to the interacting regions and study their dynamics during interphase and over cell division.

7.3- Altered gene expression around genomic integration sites

LCRs and other potent transcription regulatory DNA elements are used in gene therapy vectors to ensure full expression of rescue genes. In a series of initial experiments aimed at reverting severe combined immunodeficiency (SCID), such elements caused the inadvertent enhancement of host genes surrounding the viral integration site in at least two treated patients. These patients both developed leukemia after receiving gene therapy as a consequence of proviral activation of the *LMO2* oncogene [37-39]. Related integration effects have been observed in murine gene-targeting experiments, where transcription from constitutively expressed selection markers influences ongoing transcription from the gene they are located in or that is located immediately downstream from the integration site [40-42]. Gene therapy approaches have also been proposed for the treatment of sickle cell disease and Thalassemia disorders [43]. To ensure high expression levels of globin genes in such constructs, addition of the LCR or parts thereof seems compulsory [44, 45]. Results from **Chapter 4** and a recent paper [46] indicate that integration of the LCR at random positions in the genome can result in undesired upregulation of genes surrounding the integration site. Our results show altered expression of many genes upon integration of the β-globin LCR. Similarly, the data in the paper of Hargrove *et al.* [46], showed that the LCR can affect multiple genes surrounding its integration site. In this study, lentiviruses containing a human γ-globin coupled to HS2-4 of the human LCR were randomly inserted in *in vitro* cultured mouse primary erythroid tissue. Changes in gene expression were analysed in 18 clones with a total of 19 vector insertions. Analysis was restricted to about 300 kb up- and downstream of the integration site, and revealed altered expression in 5 of 18 clones (28%). In these 5 clones, 6 of 66 genes (11%) close to the integration sites were upregulated, with genes up to a maximum distance of 321 kb [46]. These data therefore show that the LCR is a potent element able to upregulate genes up to a considerable distance from its integration site, even when it is accompanied by a natural target gene like in the study of Hargrove *et al*. The data presented here should therefore be considered when designing safe gene therapy vectors for hematopoietic diseases. In **Chapter 4**, CTCF was unable to block the activation of multiple genes even though it was bound at a site in between the LCR and the surrounding genes. This was surprising as CTCF is often considered to be a factor that can insulate the integration cassette from surrounding host chromatin [47]. An interesting experiment addressing this issue may be to see if insulation by the LCR

integrated at 8C3/C4 works better with two CTCF-binding sites flanking the LCR. Results from a recent study suggest this would happen, as targeted integration of a retroviral cassette next to the *LMO2* oncogene strongly upregulated its expression, while this effect was up to 15 fold reduced when the viral cassette was flanked on both sites by CTCF binding sites [48].

7.4- Interchromosomal gene activation in mammals

In **Chapter 6** we determined the potential of the LCR at 8C3/C4 to upregulate genes *in trans* (located on different chromosomes) and identified the endogenous β-globin-like βh1 gene as the most prominent upregulated gene *in trans*. Initially, this finding surprised us highly, because this gene supposedly is inactive at the investigated stage of development. We wondered whether any experimental or technical artefact could provide an alternative explanation for these observations. Extensive control experiments essentially ruled out that these results could be explained by primer / probe cross reactions (**Figures 6-2** and **6-6**), developmental delays (**Figure 6-6**) and genetic background or various integration effects (**Figures 6-3** and **6-5**). Experimental data from E14.5 fetal liver cells generally shows some variation in globin gene expression levels (see for instance **Figure 6-4**), but the effect of the interchromosomal activation of the βh1 gene is so consistent that it even could be applied as a reliable method of genotyping (**Figure 6-2** and all other samples tested). These experiments ruled out that we are looking at random variation or an experimental artefact. We therefore strongly believe that upregulation by the βh1 gene must be a direct effect of LCR – gene interactions, most likely only occurring in a small but significant subpopulation of cells.

In recent years, a considerable number of high-profile publications have been issued claiming the identification of transvection / paramutation like effects [49-54] or interchromosomal transcriptional regulation [21-25] in mammals. Like discussed in **Section 1.4.4** and **Paragraph 6.2**, all these studies are based on observations suggesting that functionally related genes, loci or chromosomes interact or co-localize with a frequency higher than expected from random interactions. However, they almost invariably lack highly controlled genetic experiments, where deletion of the region on the one chromosome results in altered gene expression on another chromosome. In one instance, such a control was performed and it was shown that deletion of an important regulatory DNA element near IL4,5 and 13 on chromosome 11 caused a delay in expression of *Ifng* on chromosome 10 in an *ex vivo* differentiation model.

No overall changes in expression were observed though, and it seems possible that this difference in expression kinetics reflects a difference in differentiation capacity between wildtype and mutant T helper cells [21]. If we can confirm that communication between the globin sequences at 8C3/C4 and βh1 on chromosome *7* indeed relies on their physical interactions we think our study provides unique genetic *in vivo* evidence for interchromosomal gene activation in mammals.

Probably even more intriguing than the observed upregulation of βh1 by the human β-globin LCR, is the finding that the human Aγ-globin gene, including its promoter, integrated at 8C3/C4 is selectively able to upregulate the βh1 gene too (see **Figure 6-2**). Previously several reports in *Drosophila*, avian and mammalian systems have been published where genes, promoters and enhancers were proposed to function as insulators ([55-59] and discussed in [60]). Furthermore, elements of the β-globin LCR themselves possess promoter like properties [61, 62] and plasmids containing constitutively expressed human globin genes can initiate intergenic transcripts from the endogenous mouse β-globin locus [63]. Classification of such elements may therefore sometimes be more of a semantic discussion than being of genetic relevance. Still we are highly intrigued by the interchromosomal upregulation of βh1 gene by the Aγ-globin gene, especially since it occurs at a rate similar to the ectopic human β-globin LCR. Since the ectopic LCR has no impact on the interaction frequency of 8C3/C4 with the endogenous β-globin locus on chromosome *7*, we expect the same for the Aγ-globin gene. Possibly, when DNA elements cannot influence interaction frequencies anymore and the latter are 'fixed' by overall genomic folding patterns, as seen for the interchromosomal contacts between 8C3/C4 and βh1, both the Aγ-globin gene and LCR have the same capacity to influence the expression of interacting genes. When DNA elements can influence interaction frequencies, as expected for contacts constantly formed and terminated with genes nearby *in cis*, the LCR may have unique properties that allow productive interactions with genes. This may explain why the LCR and not the Aγ-globin gene activates the genes that surround the integration site. An alternative explanation for the fact that the two elements each have a similar impact on βh1 gene expression is that the observed levels reflect maximum βh1 expression at that stage of development.

Currently, we are planning FISH experiments with sensitive probes designed to detect primary βh1 transcripts to compare βh1 gene activity between alleles that are interacting with or separate from 8C3/C4 in transgenic FLs. Probes that have been used so far

detected βh1 mRNA accumulation (see **Figure 6-7**). They are also able to sometimes detect foci of βh1 primary transcripts, but other probes are required to more robustly detect βh1 primary transcripts (compare **Figure 6-7** vs. [30]). If we can indeed prove that interchromosomal transcriptional regulation of the βh1 gene by the Aγ-globin gene takes place, this would open up the intriguing possibility that promoters not only regulate their downstream located gene, but may also influence the expression of other, most likely related genes. Experiments that could shed further light on such mechanisms could be (i) the insertion of both the human LCR and the Aγ-globin gene at 8C3/C4 together and (ii) the integration of the human β-globin gene rather than the Aγ-globin gene at 8C3/C4. In the first experiment, gene competition experiments would suggest that βh1 would no longer benefit from the presence of these globin sequences at 8C3/C4, as the ectopic LCR is now interacting with the nearby Aγ-globin gene. However, accidental co-localization of these two elements near the endogenous β-globin locus could also cause a further increase in local concentration of relevant transcription factors [4, 64] from which βh1 could benefit even more. Thus, this experiment may give important information on the mechanism of long-range gene activation by transcriptional regulatory DNA elements. In the second experiment, the adult β-globin gene may be expected to have less impact on the expression of the embryonic βh1 gene, since it likely binds a different repertoire of transcription factors. Here, it would be interesting to see if it positively affects the expression of the endogenous adult β-globin genes (β^{maj} and β^{min}). For this, one would need to use DNA-RNA FISH to determine the percentage of actively transcribing β^{maj} and β^{min} alleles when the locus is or is not in contact with globin-containing 8C3/C4.

A last question that remains is whether mechanisms of regulation *in trans*, or *in cis*, of genes by other, related, genes are used in nature. Currently, we are unaware of such examples in mammals, but transvection studies in *Drosophila* and paramutation studies in plants suggest such mechanisms could exist in these organisms. The seemingly spurious nature of interchromosomal interactions combined with the reports of highly constrained chromatin mobility [31, 32] make it unlikely that interchromosomal communication could be applied as a faithful pan-cellular regulatory mechanism. This is different though for more local interactions *in cis*, where interactions may be expected to be far more frequent. Future research should reveal whether such additional regulatory complexity exists, or whether the effect that we have observed is an anomaly induced by the integration of specific human sequences in the mouse genome.

7.5- References

1. Blackwood, E.M., and Kadonaga, J.T. (1998). Going the distance: a current view of enhancer action. Science *281*, 60-63.
2. Bulger, M., and Groudine, M. (1999). Looping versus linking: toward a model for long-distance gene activation. Genes Dev *13*, 2465-2477.
3. Palstra, R.J., de Laat, W., and Grosveld, F. (2008). Beta-globin regulation and long-range interactions. Adv Genet *61*, 107-142.
4. de Laat, W., and Grosveld, F. (2003). Spatial organization of gene expression: the active chromatin hub. Chromosome Res *11*, 447-459.
5. Demers, C., Chaturvedi, C.P., Ranish, J.A., Juban, G., Lai, P., Morle, F., Aebersold, R., Dilworth, F.J., Groudine, M., and Brand, M. (2007). Activator-mediated recruitment of the MLL2 methyltransferase complex to the beta-globin locus. Mol Cell *27*, 573-584.
6. Jhunjhunwala, S., van Zelm, M.C., Peak, M.M., Cutchin, S., Riblet, R., van Dongen, J.J., Grosveld, F.G., Knoch, T.A., and Murre, C. (2008). The 3D structure of the immunoglobulin heavy-chain locus: implications for long-range genomic interactions. Cell *133*, 265-279.
7. Pirrotta, V. (1999). Transvection and chromosomal trans-interaction effects. Biochim Biophys Acta *1424*, M1-8.
8. Cremer, T., and Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nat Rev Genet *2*, 292-301.
9. Bolzer, A., Kreth, G., Solovei, I., Koehler, D., Saracoglu, K., Fauth, C., Muller, S., Eils, R., Cremer, C., Speicher, M.R., and Cremer, T. (2005). Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. PLoS Biol *3*, e157.
10. Mueller-Storm, H.P., Sogo, J.M., and Schaffner, W. (1989). An enhancer stimulates transcription in trans when attached to the promoter via a protein bridge. Cell *58*, 767-777.
11. Mahmoudi, T., Katsani, K.R., and Verrijzer, C.P. (2002). GAGA can mediate enhancer function in trans by linking two separate DNA molecules. Embo J *21*, 1775-1781.
12. Francastel, C., Walters, M.C., Groudine, M., and Martin, D.I. (1999). A functional enhancer suppresses silencing of a transgene and prevents its localization close to centrometric heterochromatin. Cell *99*, 259-269.
13. Ragoczy, T., Telling, A., Sawado, T., Groudine, M., and Kosak, S.T. (2003). A genetic analysis of chromosome territory looping: diverse roles for distal regulatory elements. Chromosome Res *11*, 513-525.
14. Ragoczy, T., Bender, M.A., Telling, A., Byron, R., and Groudine, M. (2006). The locus control region is required for association of the murine beta-globin locus with engaged transcription factories during erythroid maturation. Genes Dev *20*, 1447-1457.
15. Kosak, S.T., Skok, J.A., Medina, K.L., Riblet, R., Le Beau, M.M., Fisher, A.G., and Singh, H. (2002). Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. Science *296*, 158-162.
16. Lundgren, M., Chow, C.M., Sabbattini, P., Georgiou, A., Minaee, S., and Dillon, N. (2000). Transcription factor dosage affects changes in higher order chromatin structure associated with activation of a heterochromatic gene. Cell *103*, 733-743.
17. Hewitt, S.L., Farmer, D., Marszalek, K., Cadera, E., Liang, H.E., Xu, Y., Schlissel, M.S., and Skok, J.A.

(2008). Association between the Igk and Igh immunoglobulin loci mediated by the 3' Igk enhancer induces 'decontraction' of the Igh locus in pre-B cells. Nat Immunol *9*, 396-404.

18. Osborne, C.S., Chakalova, L., Brown, K.E., Carter, D., Horton, A., Debrand, E., Goyenechea, B., Mitchell, J.A., Lopes, S., Reik, W., and Fraser, P. (2004). Active genes dynamically colocalize to shared sites of ongoing transcription. Nat Genet *36*, 1065-1071.
19. Chakalova, L., Debrand, E., Mitchell, J.A., Osborne, C.S., and Fraser, P. (2005). Replication and transcription: Shaping the landscape of the genome. Nat Rev Genet *6*, 669 - 677.
20. Osborne, C.S., Chakalova, L., Mitchell, J.A., Horton, A., Wood, A.L., Bolland, D.J., Corcoran, A.E., and Fraser, P. (2007). Myc dynamically and preferentially relocates to a transcription factory occupied by Igh. PLoS Biol *5*, e192.
21. Spilianakis, C.G., Lalioti, M.D., Town, T., Lee, G.R., and Flavell, R.A. (2005). Interchromosomal associations between alternatively expressed loci. Nature.
22. Ling, J.Q., Li, T., Hu, J.F., Vu, T.H., Chen, H.L., Qiu, X.W., Cherry, A.M., and Hoffman, A.R. (2006). CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. Science *312*, 269-272.
23. Lomvardas, S., Barnea, G., Pisapia, D.J., Mendelsohn, M., Kirkland, J., and Axel, R. (2006). Interchromosomal interactions and olfactory receptor choice. Cell *126*, 403-413.
24. Zhao, Z., Tavoosidana, G., Sjolinder, M., Gondor, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Sandhu, K.S., Singh, U., Pant, V., Tiwari, V., Kurukuti, S., and Ohlsson, R. (2006). Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nat Genet *38*, 1341-1347.
25. Apostolou, E., and Thanos, D. (2008). Virus infection induces NF-kappaB-dependent interchromosomal associations mediating monoallelic IFN-beta gene expression. Cell *134*, 85-96.
26. Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., and de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nat Genet *38*, 1348-1354.
27. de Laat, W. (2007). Long-range DNA contacts: romance in the nucleus? Curr Opin Cell Biol *19*, 317-320.
28. Misteli, T. (2007). Beyond the sequence: cellular organization of genome function. Cell *128*, 787-800.
29. Wijgerde, M., Grosveld, F., and Fraser, P. (1995). Transcription complex stability and chromatin dynamics in vivo. Nature *377*, 209-213.
30. Trimborn, T., Gribnau, J., Grosveld, F., and Fraser, P. (1999). Mechanisms of developmental control of transcription in the murine alpha- and beta-globin loci. Genes Dev *13*, 112-124.
31. Abney, J.R., Cutler, B., Fillbach, M.L., Axelrod, D., and Scalettar, B.A. (1997). Chromatin dynamics in interphase nuclei and its implications for nuclear structure. J Cell Biol *137*, 1459-1468.
32. Chubb, J.R., Boyle, S., Perry, P., and Bickmore, W.A. (2002). Chromatin motion is constrained by association with nuclear compartments in human cells. Curr Biol *12*, 439-445.
33. Manders, E.M., Kimura, H., and Cook, P.R. (1999). Direct imaging of DNA in living cells reveals the dynamics of chromosome formation. J Cell Biol *144*, 813-821.
34. Walter, J., Schermelleh, L., Cremer, M., Tashiro, S., and Cremer, T. (2003). Chromosome order in HeLa cells changes during mitosis and early G1, but is stably maintained during subsequent interphase

stages. J Cell Biol *160*, 685-697.

35. Gerlich, D., Beaudouin, J., Kalbfuss, B., Daigle, N., Eils, R., and Ellenberg, J. (2003). Global chromosome positions are transmitted through mitosis in mammalian cells. Cell *112*, 751-764.
36. Thomson, I., Gilchrist, S., Bickmore, W.A., and Chubb, J.R. (2004). The radial positioning of chromatin is not inherited through mitosis but is established de novo in early G1. Curr Biol *14*, 166-172.
37. Aiuti, A., Slavin, S., Aker, M., Ficara, F., Deola, S., Mortellaro, A., Morecki, S., Andolfi, G., Tabucchi, A., Carlucci, F., Marinello, E., Cattaneo, F., Vai, S., Servida, P., Miniero, R., Roncarolo, M.G., and Bordignon, C. (2002). Correction of ADA-SCID by stem cell gene therapy combined with nonmyeloablative conditioning. Science *296*, 2410-2413.
38. Hacein-Bey-Abina, S., von Kalle, C., Schmidt, M., Le Deist, F., Wulffraat, N., McIntyre, E., Radford, I., Villeval, J.L., Fraser, C.C., Cavazzana-Calvo, M., and Fischer, A. (2003). A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. N Engl J Med *348*, 255-256.
39. Howe, S.J., Mansour, M.R., Schwarzwaelder, K., Bartholomae, C., Hubank, M., Kempski, H., Brugman, M.H., Pike-Overzet, K., Chatters, S.J., de Ridder, D., Gilmour, K.C., Adams, S., Thornhill, S.I., Parsley, K.L., Staal, F.J., Gale, R.E., Linch, D.C., Bayford, J., Brown, L., Quaye, M., Kinnon, C., Ancliff, P., Webb, D.K., Schmidt, M., von Kalle, C., Gaspar, H.B., and Thrasher, A.J. (2008). Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. J Clin Invest *118*, 3143-3150.
40. Taveau, M., Stockholm, D., Marchand, S., Roudaut, C., Le Bert, M., and Richard, I. (2004). Bidirectional transcriptional activity of the Pgk1 promoter and transmission ratio distortion in Capn3-deficient mice. Genomics *84*, 592-595.
41. Strathdee, D., Ibbotson, H., and Grant, S.G. (2006). Expression of Transgenes Targeted to the Gt(ROSA)26Sor Locus Is Orientation Dependent. PLoS ONE *1*, e4.
42. Hanada, K., Vermeij, M., Garinis, G.A., de Waard, M.C., Kunen, M.G., Myers, L., Maas, A., Duncker, D.J., Meijers, C., Dietz, H.C., Kanaar, R., and Essers, J. (2007). Perturbations of vascular homeostasis and aortic valve abnormalities in fibulin-4 deficient mice. Circ Res *100*, 738-746.
43. Nienhuis, A.W., Hanawa, H., Sawai, N., Sorrentino, B.P., and Persons, D.A. (2003). Development of gene therapy for hemoglobin disorders. Ann N Y Acad Sci *996*, 101-111.
44. May, C., Rivella, S., Callegari, J., Heller, G., Gaensler, K.M., Luzzatto, L., and Sadelain, M. (2000). Therapeutic haemoglobin synthesis in beta-thalassaemic mice expressing lentivirus-encoded human beta-globin. Nature *406*, 82-86.
45. Pawliuk, R., Westerman, K.A., Fabry, M.E., Payen, E., Tighe, R., Bouhassira, E.E., Acharya, S.A., Ellis, J., London, I.M., Eaves, C.J., Humphries, R.K., Beuzard, Y., Nagel, R.L., and Leboulch, P. (2001). Correction of sickle cell disease in transgenic mouse models by gene therapy. Science *294*, 2368-2371.
46. Hargrove, P.W., Kepes, S., Hanawa, H., Obenauer, J.C., Pei, D., Cheng, C., Gray, J.T., Neale, G., and Persons, D.A. (2008). Globin lentiviral vector insertions can perturb the expression of endogenous genes in beta-thalassemic hematopoietic cells. Mol Ther *16*, 525-533.
47. Recillas-Targa, F., Valadez-Graham, V., and Farrell, C.M. (2004). Prospects and implications of using chromatin insulators in gene therapy and transgenesis. Bioessays *26*, 796-807.
48. Ryu, B.Y., Evans-Galea, M.V., Gray, J.T., Bodine, D.M., Persons, D.A., and Nienhuis, A.W. (2008). An

experimental system for the evaluation of retroviral vector design to diminish the risk for proto-oncogene activation. Blood *111*, 1866-1875.

49. Hatada, I., Nabetani, A., Arai, Y., Ohishi, S., Suzuki, M., Miyabara, S., Nishimune, Y., and Mukai, T. (1997). Aberrant methylation of an imprinted gene U2af1-rs1(SP2) caused by its own transgene. J Biol Chem *272*, 9120-9122.
50. Forne, T., Oswald, J., Dean, W., Saam, J.R., Bailleul, B., Dandolo, L., Tilghman, S.M., Walter, J., and Reik, W. (1997). Loss of the maternal H19 gene induces changes in Igf2 methylation in both cis and trans. Proc Natl Acad Sci U S A *94*, 10243-10248.
51. Herman, H., Lu, M., Anggraini, M., Sikora, A., Chang, Y., Yoon, B.J., and Soloway, P.D. (2003). Trans allele methylation and paramutation-like effects in mice. Nat Genet *34*, 199-202.
52. Bacher, C.P., Guggiari, M., Brors, B., Augui, S., Clerc, P., Avner, P., Eils, R., and Heard, E. (2006). Transient colocalization of X-inactivation centres accompanies the initiation of X inactivation. Nat Cell Biol *8*, 293-299.
53. Xu, N., Tsai, C.L., and Lee, J.T. (2006). Transient homologous chromosome pairing marks the onset of X inactivation. Science *311*, 1149-1152.
54. Liu, H., Huang, J., Wang, J., Jiang, S., Bailey, A.S., Goldman, D.C., Welcker, M., Bedell, V., Slovak, M.L., Clurman, B., Thayer, M., Fleming, W.H., and Epner, E. (2008). Transvection mediated by the translocated cyclin D1 locus in mantle cell lymphoma. J Exp Med *205*, 1843-1858.
55. Kellum, R., and Schedl, P. (1992). A group of scs elements function as domain boundaries in an enhancer-blocking assay. Mol Cell Biol *12*, 2424-2431.
56. Dillon, N., Trimborn, T., Strouboulis, J., Fraser, P., and Grosveld, F. (1997). The effect of distance on long-range chromatin interactions. Mol Cell *1*, 131-139.
57. Ohtsuki, S., and Levine, M. (1998). GAGA mediates the enhancer blocking activity of the eve promoter in the Drosophila embryo. Genes Dev *12*, 3325-3330.
58. Bell, A.C., West, A.G., and Felsenfeld, G. (1999). The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. Cell *98*, 387-396.
59. Farrell, C.M., West, A.G., and Felsenfeld, G. (2002). Conserved CTCF insulator elements flank the mouse and human beta-globin loci. Mol Cell Biol *22*, 3820-3831.
60. Dillon, N., and Sabbattini, P. (2000). Functional gene expression domains: defining the functional unit of eukaryotic gene regulation. Bioessays *22*, 657-665.
61. Collis, P., Antoniou, M., and Grosveld, F. (1990). Definition of the minimal requirements within the human beta-globin gene and the dominant control region for high level expression. Embo J *9*, 233-240.
62. Tuan, D., Kong, S., and Hu, K. (1992). Transcription of the hypersensitive site HS2 enhancer in erythroid cells. Proc Natl Acad Sci U S A *89*, 11219-11223.
63. Ashe, H.L., Monks, J., Wijgerde, M., Fraser, P., and Proudfoot, N.J. (1997). Intergenic transcription and transinduction of the human beta-globin locus. Genes Dev *11*, 2494-2509.
64. Droge, P., and Muller-Hill, B. (2001). High local protein concentrations at promoters: strategies in prokaryotic and eukaryotic cells. Bioessays *23*, 179-183.

Summary

The basic unit for building up organisms is the cell. Cells come in a large variety of forms and with very different functions. Still, in its basic form, each individual cell contains all machinery for maintenance, functioning and proliferation. The information that encodes the cellular machinery is contained in the genome, which in mammals contains an estimated 25,000 genes. Structurally, genomic information is present in the cell nucleus as linear chromosomes. Chromosomes may add up to about 2 meters in length, but are contained in the cell nucleus with a diameter of around 10 μm. The large majority of cells contain the same genetic material, but in different cell types different subsets of genes are expressed. To allow these different gene expression patterns, mammalian genes are regulated by different regulatory mechanisms acting at different organizational levels. In **Chapter 1**, organization of chromatin, i.e. the chromosomes and their associated proteins, and nuclear organization are discussed. In this chapter, four different levels of organization are discussed. At a first level, transcriptional regulation may be achieved by local changes in chromatin characteristics, so-called epigenetic modifications, and the recruitment of *trans*-acting protein factors, i.e. transcription factors and associated enzymes. Regulation at other levels includes the three-dimensional recruitment of transcriptional regulatory elements through the formation of long-range chromatin loops, the non-random positioning of genomic loci versus nuclear substructures and possibly the association with genomic loci located on other chromosomes (*in trans*). Little is known though how these different levels of organisation influence each other. The first question that has been addressed in this thesis therefore is how chromatin and nuclear organisation influence each other.

Transcription factors in mammals do not only mediate their actions through genes and their promoters, but may also act through additional regulatory elements like enhancers. In **Chapter 2** the role of the β-globin Locus Control Region (LCR), a prototype strong regulatory element, in chromatin and nuclear organization is discussed. The β-globin LCR is required for the extremely high expression of the endogenous downstream located β-globin genes. In recent years, transcription factor mediated chromatin looping between the LCR and its target genes has been identified as an important mechanism in transcriptional enhancement. Otherwise, the LCR has been proposed to change the location during erythroid differentiation of its target genes versus the nuclear periphery, centromeric heterochromatin, transcription factories, the chromosome territory (CT) and

certain specific genomic regions. Thus far the LCR has mainly been studied as an integral part of the larger β-globin locus, which makes it difficult to discern specific LCR-function, independent from the linked β-globin genes and other regulatory elements. The second question that has been addressed in this thesis therefore is how the LCR regulates gene expression and how it influences chromatin and nuclear organization independent from other elements in the β-globin locus.

To answer these two main questions a mouse model was studied that contains the human β-globin LCR, without any of the linked globin genes, integrated in a gene-dense region containing mainly housekeeping genes.

In **Chapter 3** chromatin and nuclear organization of the wild-type gene-dense region, which we named 8C3/C4, was characterized. Gene expression and chromatin modifications in the region are importantly controlled at the level of individual genes. Furthermore we found that 8C3/C4 frequently associates with transcription factories and is positioned at the periphery of its CT. We therefore propose that aspects of nuclear organization such as position in relation to the CT or association with transcription factories may be important for transcription, but are not sufficient for determining expression levels of individual genes within this gene-dense region.

In **Chapter 4** we reported how integration of the human β-globin LCR in 8C3/C4 influences gene expression and chromatin and nuclear organization of this gene-dense region. The LCR at 8C3/C4 can increase gene expression on both sides of the integration site over a maximum distance of 150 kilobases. The integration of an oppositely oriented LCR at the same position allowed identifying genes that respond to both, the one or the other, or none of the LCRs. Genes that only responded to one of the two LCR did so in an orientation dependent manner, as they were always located downstream of the integrated LCR. We identified long-range interactions between the LCR and two upregulated genes and propose that LCR-gene contacts via chromatin looping determine which genes are transcriptionally enhanced. The presence of the LCR also changes the nuclear position of 8C3/C4. The LCR in either orientation at 8C3/C4 results in a highly significant shift of the locus away from its CT. Furthermore, locus association with transcription factories increases moderately, both for loci at the edge and outside of the CT. These results show that nuclear repositioning by the LCR is not sufficient to increase transcription of all genes in this region.

Transcriptional regulatory elements may change the nuclear location of their target loci, but the functional relevance of this relocation is poorly understood. Our finding that

the LCR induces a repositioning of 8C3/C4 versus its CT raised the possibility that this was a consequence of the LCR actively changing the nuclear environment by searching for preferred partners in the nuclear space. In **Chapter 5** we addressed this issue by determining whether LCR-induced relocation represents repositioning to a different nuclear environment. Using a transvection-like system we found no evidence that the LCR could actively establish interactions with one of its natural target genes present on the homologous 8C3/C4 allele. Next, we performed 4C (Chromosome Conformation Capture on Chip) to investigate whether the LCR searches for other preferred partners located elsewhere in the genome. LCR-induced positioning outside the CT results in increased encounters with genomic regions located on different chromosomes, but the specific interacting partners *in cis* and *in trans* are largely identical, irrespective of whether the LCR is present or not. We therefore conclude that the LCR does not scan the nucleus for preferred interaction partners. Interestingly, two existing interaction partners *in trans* containing erythroid specific genes did significantly increase their co-localization frequency as determined by FISH, while the 4C-signal of the α-globin locus shows a strong increase. Future investigations should confirm whether these results show that the LCR stabilizes some existing interactions with functionally related genes.
Even though we found no evidence for LCR-induced initiation of interactions *in trans*, we still wondered whether the LCR could activate genes in regions that were contacted by 8C3/C4. Previously, interchromosomal interactions between functionally related mammalian genes have been observed, but genetic evidence showing that the loss of DNA elements on the one chromosome alters gene expression in *trans* is mostly lacking. In **Chapter 6** we used our transgenic mice to address the issue whether transcriptional regulatory elements on one chromosome can upregulate genes on other chromosomes. Using a micro-array expression analysis, we identified only the βh1 gene on chromosome 7 to be upregulated over twofold by the LCR on chromosome 8. Interestingly, βh1 is a natural target gene of the endogenous mouse β-globin LCR. More intriguing though, we found that the βh1 gene is not only upregulated when the LCR is present at 8C3/C4 but also when the human fetal Aγ–globin gene is integrated. The change in expression is not accompanied by an increase in interaction frequencies between the endogenous β-globin locus and 8C3/C4 when the latter contains β-globin sequences: these interchromosomal contacts invariably remain rare. We nevertheless hypothesize that the communication between loci on different chromosomes relies on their physical interactions and we propose combined RNA-DNA FISH experiments to address this issue.

Together the results in this thesis provide important insight in β-globin LCR functioning and how such an element influences nuclear organisation of the locus it is present in. Secondly, these results show how different regulatory mechanisms acting at the level of chromatin and nuclear organisation influence the activity of genes present in a gene-dense region like 8C3/C4. Finally, the transgenic model system we have generated may give the first unambiguous example of interchromosomal gene regulation in mammals, an exciting observation that warrants further investigation.

Samenvatting

De bouwstenen waaruit organismen opgebouwd worden zijn cellen. De variatie tussen cellen in zowel vorm als functie is enorm. Ondanks deze grote verschillen bezitten toch vrijwel alle cellen de componenten die benodigd zijn voor onderhoud, functioneren en proliferatie. De informatie die hiervoor benodigd is, ligt opgeslagen in het genoom. In zoogdieren wordt het aantal genen, die ieder de informatie bevatten voor een component van de cellulaire machinerie, geschat op ongeveer 25.000. De genomische informatie ligt opgeslagen in de celkern in de vorm van lineaire chromosomen. Hoewel de celkern maar een diameter van ongeveer 10 μm heeft, loopt de totale lengte van de chromosomen op tot ongeveer 2 meter. De meeste cellen bevatten hetzelfde genen en genetische materiaal, maar afhankelijk van het celtype worden verschillende genen geactiveerd. In zoogdieren worden deze verschillende genexpressie patronen mogelijk gemaakt door een combinatie van regulerende mechanismen die functioneren op verschillende niveaus van organisatie. In **Hoofdstuk 1** wordt de organisatie van de celkern en de structuur van chromatine, dat wil zeggen de chromosomen en hun geassocieerde eiwitten, behandeld. In het hoofdstuk worden vier verschillende niveaus van genregulatie onderscheiden. Op het eerste niveau worden genen gereguleerd door lokale veranderingen van de chromatine structuur, zogenaamde epigenetische veranderingen, en door het rekruteren van regulerende eiwitten, zoals transcriptiefactoren en geassocieerde eiwitten. Het volgende niveau van regulatie is de vorming van driedimensionale chromatine loops, waardoor regulerende element over lange afstand in de celkern bij elkaar gebracht worden. De laatste twee bediscussieerde niveaus van genregulatie zijn niet-willekeurige lokalisatie van genomische regio's ten opzichte van bepaalde nucleaire substructuren en mogelijk de interactie tussen genomische regio's die op verschillende chromosomen liggen (*in trans*). Hoewel deze verschillende regulerende niveaus uitvoerig bestudeerd zijn, is er weinig bekend over hoe regulerende mechanismen op de verschillende niveaus elkaar beïnvloeden. De eerste vraag die behandeld is in dit proefschrift is daarom welk effect chromatine en nucleaire organisatie op elkaar hebben.

In zoogdieren oefenen transcriptie factoren hun regulerende functie niet alleen uit door binding aan genen en promotoren maar ook door de binding aan verder weg gelegen elementen zoals 'enhancers'. In **Hoofdstuk 2** is een van de prototype regulerende elementen, de β-globine Locus Control Region (LCR) geïntroduceerd en is haar rol in de organisatie van chromatine en de celkern besproken. De LCR is noodzakelijk voor de

extreem hoge activiteit van de verderop op het chromosoom gelegen β-globine genen. In de afgelopen jaren is naar voren gekomen dat chromatine loops tussen de LCR en de globine genen een belangrijk aspect zijn van de transcriptie versterkende functie van de LCR. Daarnaast wordt er gedacht dat de LCR tijdens de ontwikkeling van rode bloedcellen de locatie van de β-globine genen verandert ten opzichte van de nucleaire periferie, centromerisch heterochromatine, transcription factories (clusters in de nucleus waar meerdere genen tegelijkertijd actief worden afgeschreven) en het chromosoom territorium (CT). Tot nu toe is de LCR echter vooral bestudeerd als integraal onderdeel van de overkoepelende β-globine regio. Hierdoor kan moeilijk onderscheid gemaakt kan worden tussen LCR-specifieke functies en mechanismen veroorzaakt door de β-globine regio als geheel. De tweede vraag die daarom behandeld is in dit proefschrift is hoe de LCR de activiteit van genen reguleert en hoe zij chromatine en nucleaire organisatie beïnvloedt, onafhankelijk van andere elementen in the β-globine regio.
Om deze vragen te beantwoorden is een transgeen muis model bestudeerd, waarin de humane β-globine LCR, zonder andere elementen uit de β-globine regio, is geïntegreerd in een genrijk gebied met vooral genen betrokken bij de dagelijkse gang van zaken in de cel.
In **Hoofdstuk 3** is de chromatine en nucleaire organisatie van dit genrijke gebied dat wij 8C3/C4 hebben genoemd beschreven. Gen activiteit en de bijbehorende epigenetische veranderingen worden voor een belangrijk deel gecontroleerd op het niveau van het individuele gen. Daarnaast is 8C3/C4 veelvuldig geassocieerd met transcription factories en lokaliseert het aan de periferie van het CT. Onze conclusie is daarom dat hoewel componenten van nucleaire organisatie, zoals positie ten opzichte van het CT of associatie met transcription factories, belangrijk kunnen zijn voor gen expressie deze componenten niet voldoende zijn om de volledige expressie niveaus van genen in 8C3/C4 te bepalen.
In **Hoofdstuk 4** hebben we laten zien hoe genexpressie en chromatine en nucleaire organisatie in 8C3/C4 beïnvloed wordt door de integratie van de β-globine LCR. De LCR in 8C3/C4 verhoogt de activiteit van genen aan beide kanten van haar integratie site en die tot 150 kilobasen verder weggelegen zijn. Door integratie van de LCR op dezelfde plaats maar in de tegenovergestelde oriëntatie zijn genen geïdentificeerd die reageren op de LCR in beide oriëntaties, op de LCR in slechts één oriëntatie of helemaal niet op de LCR. Genen die reageren op de LCR in slechts één oriëntatie doen dit op een oriëntatie-afhankelijke manier, want deze genen liggenen altijd in het verlengde van de LCR. Vervolgens vonden

wij dat chromatine loops gevormd worden tussen de LCR en twee opgereguleerde genen. Daarom stellen wij voor dat contact tussen genen en de LCR de bepalende factor is voor welke genen opgereguleerd worden. De aanwezigheid van de LCR in 8C3/C4 zorgt ook voor een verandering in de nucleaire locatie van dit genrijke gebied. De LCR, onafhankelijk van haar oriëntatie, herpositioneert 8C3/C4 verder weg van het CT. Bovendien neemt de associatie met transcription factories enigszins toe, maar dit geldt zowel voor 8C3/C4 wanneer het aan de rand als buiten het CT gelokaliseerd is. Deze resultaten laten zien dat hoewel de LCR de positie van 8C3/C4 kan veranderen, dit niet voldoende is om de activiteit van alle genen in het genrijke gebied te verhogen.

Een van de voorgestelde functies van transcriptie regulerende elementen is het veranderen van de nucleaire locatie van de gebieden die zij beïnvloeden. De functie van deze herpositionering is echter niet duidelijk. Onze vinding dat 8C3/C4 door de LCR verplaatst wordt ten opzichte van het CT zou een reflectie kunnen zijn van een door de LCR geïnduceerde zoektocht in de nucleus naar functioneel gerelateerde genen. In **Hoofdstuk 5** hebben wij daarom de vraag behandeld of de verplaatsing van 8C3/C4 door de LCR resulteert in een herpositionering in een andere nucleaire omgeving. Door gebruik te maken van een transvectie-gerelateerd systeem hebben wij geen aanwijzing gevonden dat de LCR actief zoekt naar een natuurlijke interactie partner wanneer dit gen aanwezig is in het 8C3/C4 allel op het homologe chromosoom. Vervolgens hebben wij een 4C studie (Chromosome Conformation Capture on Chip) gedaan om te bepalen of de LCR actief interacteert met andere favoriete partners op andere plaatsen in het genoom. Herpositionering buiten het CT door de LCR leidt tot frequentere interacties met gebieden op andere chromosomen. Specifieke gebieden waarmee interacties aangegaan worden, zowel op hetzelfde als op andere chromosomen, worden echter hoofdzakelijk niet bepaald door de aanwezigheid van de LCR. Onze conclusie is daarom dat de LCR de nucleus niet doorzoekt naar specifieke gebieden om interacties mee aan te gaan. FISH-experimenten hebben daarentegen laten zien dat twee van de bestaande interacterende gebieden *in trans* die allebei rode bloedcel-specifieke genen bevatten significant vaker een interactie met 8C3/C4 aangaan wanneer de LCR aanwezig is. Bovendien laat het 4C-signaal van 8C3/C4 met de LCR een sterk verhoogd signaal zien met de α-globine regio. Vervolgstudies zullen moeten aantonen of de LCR inderdaad bestaande interacties kan stabiliseren met gebieden die rode bloedcel-specifieke genen bevatten.

Hoewel wij geen aanwijzingen hebben gevonden dat de LCR nieuwe interacties aangaat met genen *in trans*, vroegen wij ons toch af of de LCR de activiteit van genen kan

verhogen in gebieden die interacteren met 8C3/C4. In eerdere studies zijn interacties met een regulerende functie tussen gebieden op verschillende chromosomen gevonden. De meeste van deze studies missen echter essentiële genetische studies waarmee aangetoond wordt dat de afwezigheid van de regulerende DNA elementen op het ene chromosoom resulteert in veranderingen in de activiteit van genen op het andere chromosoom. In **Hoofdstuk 6** hebben wij onze transgene muizen gebruikt om te onderzoeken of een transcriptie regulerend element op het ene chromosoom de transcriptionele activiteit van genen op andere chromosomen kan veranderen. Door gebruik te maken van micro-arrays hebben we gevonden dat de LCR op chromosoom 8 van alle genen op andere chromosomen alleen de activiteit van het βh1 gen op chromosoom 7 kan verdubbelen. Deze vinding is erg interessant want het βh1 gen is een natuurlijke Partner van de endogene muizen β-globine LCR. Waarschijnlijk nog interessanter is dat dit gen niet alleen opgereguleerd wordt wanneer de LCR aanwezig is in 8C3/C4, maar ook wanneer een humaan foetaal Aγ-globine gen op dezelfde plek geïntegreerd is. De verandering in transcriptionele activiteit van βh1 leidt niet tot een verhoogde interactie frequentie tussen de endogene β-globine regio en 8C3/C4 wanneer hierin humane β-globine sequenties zijn geïntegreerd, want interacties blijven in alle gevallen van een vergelijkbare lage frequentie. Ondanks deze weinig frequente interacties, speculeren wij dat de communicatie tussen deze gebieden *in trans* afhankelijk is van fysieke interacties. Wij stellen daarom in dit hoofdstuk gecombineerde RNA-DNA-FISH experimenten voor om deze hypothese te testen.

De resultaten zoals gepresenteerd in dit proefschrift leveren belangrijke nieuwe inzichten in hoe de β-globine LCR functioneert en hoe dit soort elementen de nucleaire organisatie beïnvloeden van gebieden waarin zij geïntegreerd zijn. Bovendien laten deze resultaten zien hoe de verschillende regulatie mechanismen die functioneren op het niveau van chromatine en nucleaire organisatie de activiteit beïnvloeden van genen zoals die aanwezig zijn in 8C3/C4. Tenslotte kan het transgene muis model dat wij gecreëerd hebben het eerste goed karakteriseerbare voorbeeld van interchromosomale genregulatie in zoogdieren opleveren. Dit baanbrekende laatste resultaat zal nog uitgebreid verder onderzocht gaan worden.

Curriculum Vitae

Personal details

Name Daniël (Daan) Noordermeer
Birth date November 24th, 1978
Birth place Bergen op Zoom, The Netherlands

Education

1997 – 2004 BSc and MSc
Bioprocess Technology, Wageningen University, Wageningen, The Netherlands.

1991 – 1997 VWO
RKSG Serviam Lyceum, Sittard, The Netherlands.

Scientific research

From 2009 Post doctoral fellow
Lab of Developmental Genomics, École Polytechnique Fédéral de Lausanne, Lausanne, Switzerland.

2009 Post doctoral fellow
Genome Structure and Function Group, Hubrecht Intitute, Utrecht, The Netherlands.

2004 – 2009 PhD student
Department of Cell Biology and Genetics, Erasmus Medical Center, Rotterdam, The Netherlands.
Prof. Dr. F. Grosveld (promoter) & Prof. Dr. W. de Laat (co-promoter)

2003 Visiting Scientist
Department of Plant Pathology, University of Wisconsin – Madison, Madison, USA

2002 Internship
Department of Communication and Innovation Studies, Wageningen University, Wageningen, The Netherlands

2001 – 2002 Internship
Laboratory of Microbiology, Wageningen University, Wageningen, The Netherlands

Publication list

Noordermeer, D. and De Laat, W. **2008** Joining the loops: β-globin gene regulation. IUBMB Life. ***60***, 824-833

Noordermeer, D.*, Branco, M.R.*, Splinter, E., Klous, P., Van IJcken, W., Swagemakers, S., Koutsourakis, M., Van der Spek, P., Pombo, A. and De Laat, W. **2008** Transcription and chromatin organization of a housekeeping gene cluster containing an integrated β-globin Locus Control Region. PLoS Genetics. ***4***, e1000016

De Laat, W., Klous, P., Kooren, J., Noordermeer, D., Palstra, R.-J., Simonis, M., Splinter, E. and Grosveld, F. **2008** Three-dimensional organization of gene expression in erythroid cells. Current Topics in Developmental Biology ***82***, 117-139

Bok, J. W., Noordermeer, D., Kale, S. P. and Keller, N. P. **2006** Secondary metabolic gene cluster silencing in *Aspergillus nidulans*. Molecular Microbiology ***61***, 1636-1645

McDonald, T., Hammond, T., Noordermeer, D., Zhang, Y.-Q. and Keller, N. **2005** The sterigmatocystin cluster revisited: Lessons from a genetic model. pp. 117-136. In: Aflatoxin and Food Safety. Taylor & Francis Publishers.

* equal author contribution

PhD portfolio

Name PhD student: Daan Noordermeer Erasmus MC Department: Cell Biology and Genetics Research School: Medisch-Genetisch Centrum Zuid-West Nederland	PhD period: May 2004 – February 2009 Promotor(s): Prof. F. Grosveld Supervisor: Dr. W. de Laat

1. PhD training

	Year
General academic skills	
- Biomedical English Writing and Communication	2006
Research skills	
- Laboratory animal science	2004
- Safe handling of radioactive materials	2001
In-depth courses (e.g. Research school, Medical Training)	
- Epigenetics and Chromatin	2008
- From Development to Disease	2005
- Analysis of Microarray gene expression data	2005
International conferences	
- Oral presentation, Sixteenth Conference on Hemoglobin Switching. Monterey, USA.	2008
- Poster presentation, EMBO Conference Series on Nuclear Structure and Dynamics. Montpellier, France.	2007
- Oral presentation, Fifth and Fourth Winter School of the International Graduiertenkolleg "Transcriptional control in developmental processes". Kleinwalstertal, Austria.	2007, 2006
- Oral presentation, Seventh Joint MGC-Cancer Research UK Graduate Student Conference. Oxford, UK.	2006
- Oral presentation, Third Dutch Chromatin Community Meeting. Kerkrade, The Netherlands.	2005

Nawoord

Het is alweer bijna vijf jaar geleden dat ik aan mijn promotie bij het ErasmusMC ben begonnen. Dat je dit nawoord nu kunt lezen betekent dat ik de resultaten van het project in een proefschrift heb weten te gieten. Het voelt alsof de jaren in Rotterdam als een flits voorbij zijn gegaan. Terugkijkend heb ik wel heel veel dingen gedaan, zowel op het lab als daarbuiten. In dit onderdeel van mijn proefschrift wil ik iedereen bedanken die mij geholpen heeft tijdens mijn promotie of aan wie ik bijzondere herinneringen koester.

Als eerste wil ik mijn promotor Prof. Frank Grosveld bedanken. Ik ben je erg dankbaar dat je me de mogelijkheid hebt gegeven om binnen Celbiologie in Rotterdam mijn promotieonderzoek te doen. Ook al hebben we elkaar misschien minder gesproken dan de bedoeling was, je was altijd op de hoogte van mijn werk en kon uit een paar worden genoeg informatie halen om tot zinnig advies te komen.
Mijn meeste dank gaat uit naar mijn co-promotor Prof. Wouter de Laat. Je email kwam indertijd als geroepen en ik ben erg blij dat je me aangenomen hebt om aan dit uitdagende project te werken. Daarnaast ben ik je natuurlijk dankbaar voor je steun en dat je, ook toen het project minder makkelijk vooruit kwam dan de bedoeling was, in me bent blijven geloven. Ik hoop dat ik in Utrecht het laatste puntje op de 'ı' kan zetten en dat we zo de laatste hoofdstukken tot een mooie publicatie om kunnen smeden.
Ook wil ik graag de leden van mijn kleine promotiecommissie bedanken. Joost, Peter en Sjaak, bedankt voor het doornemen van mijn boekje en jullie constructieve commentaar. Dies, bedankt voor het zitting nemen in mijn grote promotiecommissie, ik zie uit naar de verdediging.
Dear Ana, thank you for joining my thesis committee and coming from England for my defense. And also thank you very much for your and Miguel's invaluable help during our collaboration. Our joint paper took a while to get published, but I think it ended up being a very nice story that combines our different expertise very well.
Om het hele promotiegebeuren compleet te maken, wil ik ook graag mijn paranimfen bedanken. Iris, behalve dat je in deze drukke tijd mijn paranimf wil zijn, wil ik je vooral ook graag bedanken voor je hulp aan het begin van mijn promotie. Zonder jouw tussenkomst was ik nooit in Rotterdam terecht gekomen. Dank je wel voor het laten vallen van mijn naam bij Wouter en voor de weken erna toen ik in je appartement heb gebivakkeerd. Ik ben blij dat we elkaar tijdens studie en promotie altijd hebben kunnen helpen. Veel succes

in de VS na jouw promotie. Wouter (van der Goot), jij ook bedankt dat je mijn paranimf wil zijn. Wij gaan natuurlijk nog veel langer terug en het is erg fijn dat we nog altijd zo goed werk- en wetenschapsonderwerpen weten te bepraten.

Ook wil ik graag al mijn (oud-)labgenoten bedanken. Het was niet alleen erg fijn om samen met jullie te werken, maar toch ook vooral de spelletjesavonden, bowlingpartijen, de welverdiende biertjes op z'n tijd en de andere activiteiten maakten lab 702 een fijne lab om onderdeel van uit te maken.
Petra, ik denk dat er bijna geen experimenten in dit proefschrift staan waar je niet op een of andere manier bij betrokken bent geweest. Zonder jouw hulp had dit boekje er dan ook nooit gelegen. Ik kijk er nu al naar uit om nog een paar maanden met je op het Hubrecht verder te werken.
Marieke, bedankt voor de hulp met de 4C experimenten. Ik hoop dat we er nog een mooie publicatie uit gaan halen. Veel succes in het lab van Edwin Cuppen, hoewel ik er niet aan twijfel dat dat wel goed gaat komen.
Erik, ook jij bedankt voor alle hulp bij het labwerk. Ik hoop dat ik je niet al te veel afgeleid heb met het "Erik, mag ik je even wat vragen". Ik ga je in ieder geval nog een paar maanden zien op het Hubrecht lab. Veel succes met het afmaken van je PhD.
Robert-Jan, ook voor jou als nestor van het lab (haha), bedankt voor de hulp en antwoorden op mijn vragen. Nu ook ik op het punt van vertrekken sta, blijf je binnenkort als laatste der "De Laat-mohikanen" in Rotterdam over. Gelukkig heb je je eigen plek alweer aardig gevonden.
Jurgen, het voelt alsof je alweer een hele tijd geleden gepromoveerd bent. Het was goed om iemand in het lab te hebben die een beetje voorliep, want zo kon ik mooi de kunst van het promoveren een beetje afkijken. Veel succes met je carrière als klinisch chemicus.
Emilie, it was great working with you during the year you where here. You have highly impressed me by the amazing amount of work you did in that short period. It was great that I could visit you this summer, and I hope I will soon find the time to stay in contact better.
Harmen, je bent een van de nieuwelingen, maar ik denk dat je je waarde al dubbel en dwars hebt bewezen. Alvast bedankt voor de analyses die je tot nu toe uitgevoerd hebt, en ik denk dat we elkaar nog veel zullen spreken over de analyses die nog komen gaan. Veel succes met het opzetten van de bioinformaticaclub in Utrecht!
Sjoerd and Yun, we only worked together for a little while, but I'm looking forward seeing

you a bit more in Utrecht. I wish both of you good luck in finishing your PhDs.
Maureen, het is maar iets meer dan een jaar geleden dat je opeens het lab in kwam lopen, maar het voelt alsof we elkaar al door en door kennen. Zonder onze evaluaties tijdens de lunch en koffie kom ik de dag toch moeilijk door. Ik hoop dat er nog vele WII-avonden en andere activiteiten mogen volgen, in Nederland of op andere exotische locaties.
Sanja, I enjoyed working with you very much. I'm amazed by the efficiency in the writing of your thesis, and I have no doubt that you'll produce a very nice book. Thanks for your great time in the lab and your enthusiasm for organizing parties and drinks. Good luck in graduating soon and I wish you the best of luck in Israel.
Athina, it was also very nice working with you in the last four years. I know you've been troubled by changes in supervision, but luckily your defense is now coming close. Good luck in finishing up and your work afterwards, whether that will be in The Netherlands or elsewhere.

Behalve mensen in het lab zijn er ook veel mensen binnen celbiologie, Cluster 15 en de rest van het ErasmusMC die ik graag wil bedanken. Sommigen omdat jullie hulp van onschatbare waarde was, anderen omdat het heel fijn was met jullie samen te werken of juist dingen buiten werktijd om te doen, en in veel gevallen zelfs een combinatie van deze drie dingen. Petros, Farzin, Eric, Charlotte, Dorotha, Boet, Ali, Ernie, Umut, Erik, Debbie, Tiago, Katy, Raymond, Alireza, Miyata, Kris, Rick, Phebe, Alex, Mariette, Dubi, Sylvia, Laura, Frank, Jeffrey, Marja, Bert, Annelies, Esther, thank you very much for being around and helping me out when necessary. Een extra bedankje aan lab 902 (Eskeww, Akiko, Annegien, Ruben), het was altijd beregezellig om met jullie op stap te gaan.
Karine, thanks for our time together. I hope Poland will bring you everything you're hoping for. Judith, dank je wel voor onze fijne gesprekken. Ik hoop dat je een fijne tijd in Parijs gaat hebben. Marianne, jij ook bedankt voor onze gesprekken. Ik hoop dat binnenkort die perfecte baan en zekerheid toch eens gaan komen. Mehlika, we've only met a little while ago, but I enjoy our evenings out and conversation very much. I hope we can stay in touch.
Melle, Marieke, Jasperina, Corianna en de computermannen: bedankt dat jullie altijd klaarstaan om binnen de kortste keren alle problemen op te lossen. AIOs zouden het een stuk moeilijker hebben zonder jullie hulp!

Gelukkig was er naast de promotie ook nog tijd voor de hoognodige ontspanning. Ik wil daarom ook nog een aantal belangrijke mensen bedanken voor hun indirecte steun tijdens mijn promotie. Sjoerd and Ekim, thanks a lot for the weekly Tuesday evening band rehearsal. I really enjoyed playing with you together and I feel very proud about the demos we made. Rock on! Ilse, Jeffrey, Jasper, Jeroen en Sanne: dank jullie wel voor het kroegen, wandelen, concerten, logeerpartijen en wat al niet meer. Met jullie op stap was altijd de perfecte afleiding. Ik hoop dat ik jullie niet al te veel met mijn PhD-verhalen aan jullie kop gezeurd heb. Sanne, jij nog extra bedankt voor het uitgeven van mijn proefschrift. Mannen van de voetbalpoule: nu die academische titel bijna binnen is, is er weer tijd om serieus aan de belangrijkere titels te gaan werken. Vanaf volgend jaar wordt mijn jaarlijks doel de driedubbelslag, dus ik zou jullie borst maar natmaken! Bedankt voor de gein, ongein en het opzoeken van papers op zijn tijd, zo blijf ik mooi op de hoogte. Neef Jeroen, dank je wel voor onze kroegbezoeken in Rotterdam. Je introductie van leuke plekken komt nog altijd van pas. Ik zal onze vakantie in de VS nooit vergeten. Noël en Cathrien, dank je wel dat jullie me gastvrij onderdak geboden hebben op de Maaskade. Ik geloof dat het uitzicht daar me pas echt van Rotterdam heeft leren houden. Noël, enorm bedankt voor de hulp met de lay-out en het ontwerpen van de kaft van dit proefschrift. Ik hoop dat we ook als ik verder weg zit nog heel veel gaan zien.

Als allerlaatste wil ik heel graag mijn familie bedanken. Pap en mam, dank jullie wel voor alle steun tijdens mijn promotie. Ik prijs me erg gelukkig dat jullie me altijd gestimuleerd en actief geholpen hebben met de keuzes die ik maak, ook al komen die misschien wel eens ruw op jullie dak vallen. Ik hoop dat het resultaat jullie toch bevalt. Minke, Edo, Sylvie, Jan Willem en Joanna, ook jullie bedankt voor de hulp en stimulerende discussies. Sylvie, schiet je wel een beetje op met jouw promotie, want vanaf nu val je wel een beetje uit de toon bij de rest van je broers en zussen. Zullen we het bij het volgende kerstdiner gewoon weer over Scheikunde hebben, dat bevalt me altijd wel!