Chromatin Reorganization During Epithelial-to-Mesenchymal Transition

A Role for Euchromatin Lamin B1 Domains

Laura Pascual-Reguant

TESI DOCTORAL UPF / 2018

DIRECTOR DE LA TESI

Dra. Sandra Peiró Sales

DEPARTAMENT DE CIÈNCIES EXPERIMENTALS I DE LA SALUT

A l'Adrià, per fer-me sentir sempre capaç.

A la meva familia, pel seu suport incondicional.

Abstract

The nuclear lamins A/C and B are major constituents of the nuclear lamina (NL). Structurally conserved lamina-associated domains (LADs) are formed by genomic regions that contact the NL. However, lamins are also found in the nucleoplasm, where their role is still unknown. In this thesis, we map the genome-wide localization of lamin B1 into an euchromatin-enriched fraction of the mouse genome and follow its dynamics during the epithelial-to-mesenchymal transition (EMT). Lamin B1 associates with actively expressed and open euchromatin, forming dynamic euchromatin laminB1-associated domains (eLADs) with an average size of 0.3 Mb. Hi-C data link eLADs to the three-dimensional (3D) organization of the mouse genome during EMT process and correlate lamin B1 enrichment at topologically associating domain (TAD) borders with increased border strength. Importantly, reduced levels of lamin B1 alter the EMT signature and compromise acquisition of mesenchymal traits. Therefore, the mouse genome requires dynamic and plastic domains (eLADs) that contain predetermined lamin B1 levels in order to undergo EMT.

Resum

Les làmines nuclears (A/C i B) són els majors constituents de la làmina nuclear (NL). Els estructuralment conservats dominis associats a les làmines (LADs) estan formats per regions genòmiques que contacten amb la NL. Tot i així, les làmines també s'han trobat dins del nucleoplasma però el paper de les làmines nuceloplasmàtiques encara està per definir. En aquesta tesis, hem mapat la localització de la Lamina B1 en tot el genoma de ratolí en una fracció enriquida en eucromatina i hem seguit el seu dinamisme durant la transició epiteli-mesènquima (EMT). La lamina B1 s'associa amb eucromatina activament expressada i oberta, formant dominis d'eucromatina associats a la lamina B1 (eLADs) dinàmics amb una mida mitjana de 0.3Mb. Les dades del Hi-C relacionen els eLADs amb la organització tri-dimensional del genoma de ratolí durant el procés de transició epiteli-mesènquima i correlaciona l'enriquiment de la lamina B1 a les fronteres de dominis d'associació topològics (TADs) amb un increment de la força de les fronteres. Nivells reduïts de lamina B1 alteren el la EMT i comprometen l'adquisició de característiques mesenquimals. Per tant, el genoma de ratolí necessita de dominis plàstics i dinàmics (eLADs) que continguin nivells predeterminats de lamina B1 per tal que es completi la EMT.

Index

Figure Index

Introduction

Results

Discussion

Concluding remarks

Tables Index

Materials and methods

Annex

Abbreviations

3C: Chromosome Conformation Capture

ATAC: Assay for transposase-accessible chromatin

bp: DNA base pairs

BSA: Bovine serum albumin

C: Control

CDH1: E-cadherin gene

ChIP: Chromatin immunoprecipitation

cLAD: Constitutive lamin-associated domain

CTCF: CCCTC-binding factor

DamID: DNA adenine methyltransferase identification

DAPI : 4',6-diamidino-2-phenylindole

DNA: Deoxyribonucleic acid

DNMT: DNA methyltransferase

ECM: Extracellular matrix

EGF: Epidermal growth factor

eLAD: Euchromatin lamina-associated domain

EMT: Epithelial-to-mesenchymal transition

ESC: Embryonic stem cell

EU: 5-Ethynyl uridine

FGF: Fibroblast growth factor

FISH: Fluorescent in situ hybridization

FRAP: Fluorescence recovery after photobleaching

GFP: Green fluorescent protein

HDAC: Histone deacetylase

HP1: Heterochromatin protein 1

IF: Intermediate filament

IgG: Immunoglobulin

INM: Inner nuclear membrane

KD: Knockdown

LAD: Lamina-associated domain

LB1: Lamin B1

LOXL2: Lysyl oxidase–like 2

LSD1: Lysine-specific histone demethylase 1

MEFs: Mouse embryonic fibroblasts

MET: Mesenchymal-to-epithelial transition

MNase: Micrococcal nuclease

mRNA: Messenger ribonucleic acid

NE: Nuclear envelope

NLS: Nuclear localization signal

NMuMG: Normal murine mammary gland

NPC: Nuclear pore complex

PADs: Pericentromere-associated domains

PCA: Principal component analysis

PcG: Polycomb group

PRC: Polycomb repressive complex

PRE: Polycomb response element

PTM: Post-translational modification

RNA: Ribonucleic acid

SD: Standard deviation

SIM: Structured illumination microscopy

SNAI1: Snail family transcriptional factor 1

Suv39h: Suppressor of variegation 3-9 homolog

TADs: Topologically associating domains

TF: Transcription factor

TGF-β: Transforming growth factor beta

TSS: Transcription start site

INTRODUCTION

1. Chromatin structure and organization

Eukaryotic cells have approximately 2 m of DNA that must be correctly packaged to fit into the small volume of the nucleus. For that reason, different levels of organization are required.

1.1. Nucleosomes as chromatin unit

DNA was discovered in 1953 by Watson and Crick as a double helix able to store and copy all the individual genetic information. However, the double helix formation of DNA is only the first step of DNA structure. In the nuclei of eukaryotic cells, DNA is further re-assembled into chromatin structures that determine the activity and organization of genomic DNA.

The basic unit of chromatin is the nucleosome core particle, which consists in 147 bp of core DNA wrapped approximately 1.75 times around an octamer of proteins called histones, thus generating a 225 kDa nucleosome core structure[26] of about 11 nm in diameter, which resembles “beads on a string”[27,28] (Figure I.2). The histone octamer is formed by two copies each of the four core histones, namely, H2A, H2B, H3, and H4, that have a relative small molecular weight (11–15 kDa)[29] (Figure I.1). As the core DNA is in tightly associated with the histone octamer, it is protected from nuclease digestion and from binding of trans-acting factors[30]. Histones have a globular core domain that folds into the histone fold domain, representing 70%–75% of the protein, a highly conserved structure around which DNA wraps. Each histone also contains a histone tail, which comprises about 25%–30% of the

protein and is highly conserved, although structurally undefined. Tail domains are located in the N-terminal region of all four core histones, extending into the exterior of nucleosomes. Histone tails are highly basic, with a dominance of lysine and arginine, and contain residues that are targets of posttranslational modifications (PTMs) important for epigenetic signaling, such as acetylation, phosphorylation, and methylation[30,31].

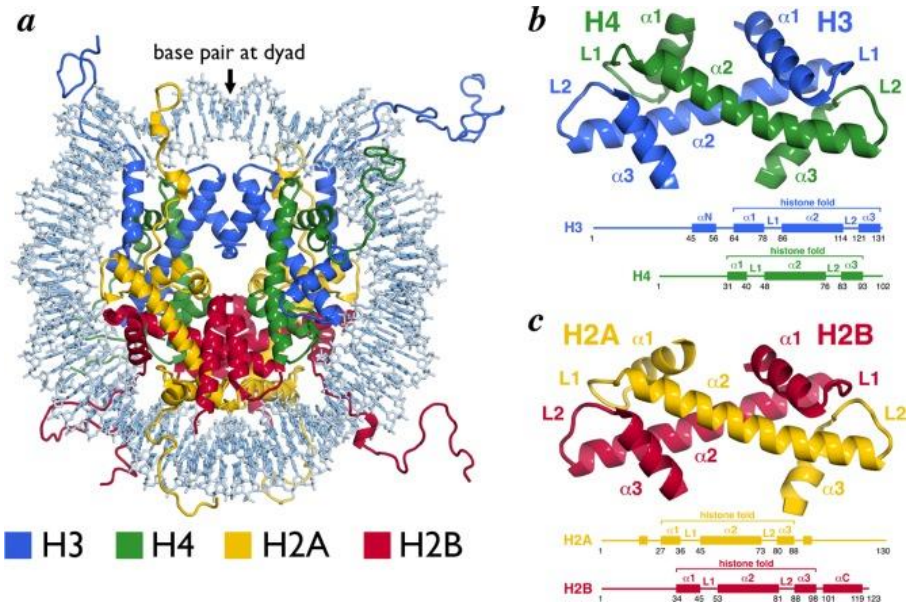


Figure I.1. Schematic representation of nucleosome and histone core structures. (A) Nucleosome core particle structure, with DNA and core histones colored as indicated. **(B)** H3/H4 histone-fold heterodimer. **(C)** H2A/H2B histone-fold heterodimer. Secondary structure elements are indicated in (B) and (C)[16].

Another component of the nucleosome subunit is the “linker” histone (H1 or H5), which is much less conserved between species. It functions to stabilize the wrapped DNA around the histone octamer by binding the sites where DNA enters into the nucleosome core particle[32,33].

Nucleosomes are joined to each other by the DNA that runs between them, which is called the “linker” DNA, which has an average length of 20 bp[27,34]. As linker DNA is not protected by histone proteins, it is preferentially cut by nucleases, such as micrococcal nuclease (MNase)[35].

1.2. Chromatin assembly

For a long time, chromatin assembly inside the nucleus was explained in most text books by the hierarchical folding model. As its name suggests, this model is based on successive coiling of DNA high-order structures, to give rise to even higher order chromatin structures that acquires the required condensation state in nuclei of eukaryotic cells[1,36,37]. Thus, different nucleosomes must fold into secondary 30 nm chromatin fibers. Subsequently, this fiber is assembled to give rise to 120 nm chromonema, which then fold to generate 300–700 nm chromatids and, finally, mitotic chromosomes (Figure I.2).

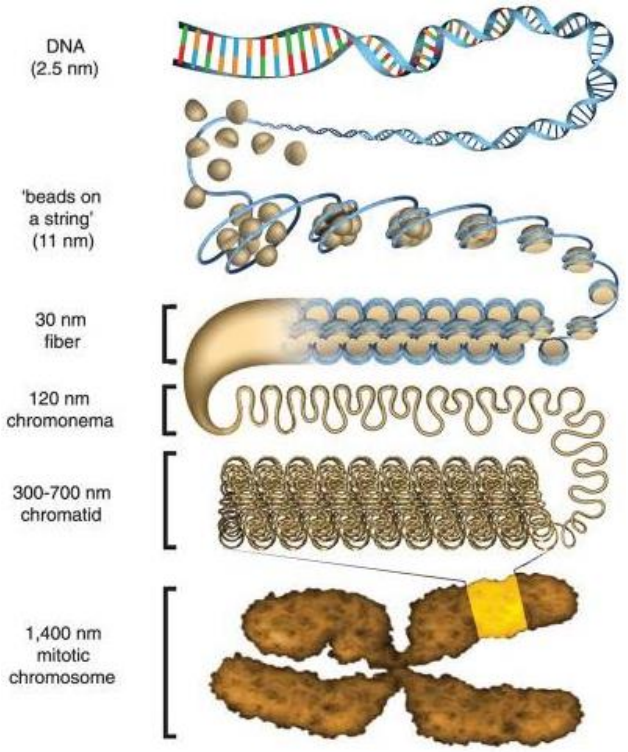


Figure I.2. Schematic representation of hierarchical chromatin-folding model. DNA is wrapped around nucleosomes resembling "beads on a string" that are folded to generate 30 nm fibers. At the same time, these fibers are folded to generate 120 nm chromonema to give rise to chromatids that form chromosomes[1].

However, over the past decade, several experimental studies coupled with computational models of the nucleus have questioned this hierarchical model. Indeed, although the presence of 30 nm fibers has been extensively studied *in vitro*, evidence for the existence of these structures *in vivo* are limited to some specialized systems, in which heterochromatin transcription repression and compaction prevail[1,36,37].

Recently, the “polymer melt” model has been proposed to explain the chromatin organization inside the nucleus of mammalian cells as an alternative to the 30 nm chromatin fiber model. This model suggests that chromosomes exist as a highly disordered and interdigitated state, similar to a polymer melt with dynamic movements. Notably, this model predicts that nucleosomes are not linear neighbors in the DNA strand but can interact within a chromatin region, thus favoring inter-fiber contacts rather than intra-fiber contacts (Figure I.3). This dynamic, irregular folding might provide a more flexible organization of the genome inside the nucleus, which would facilitate dynamic cellular processes, such as transcription, DNA replication, and enhancer-promoter interactions[21,38].

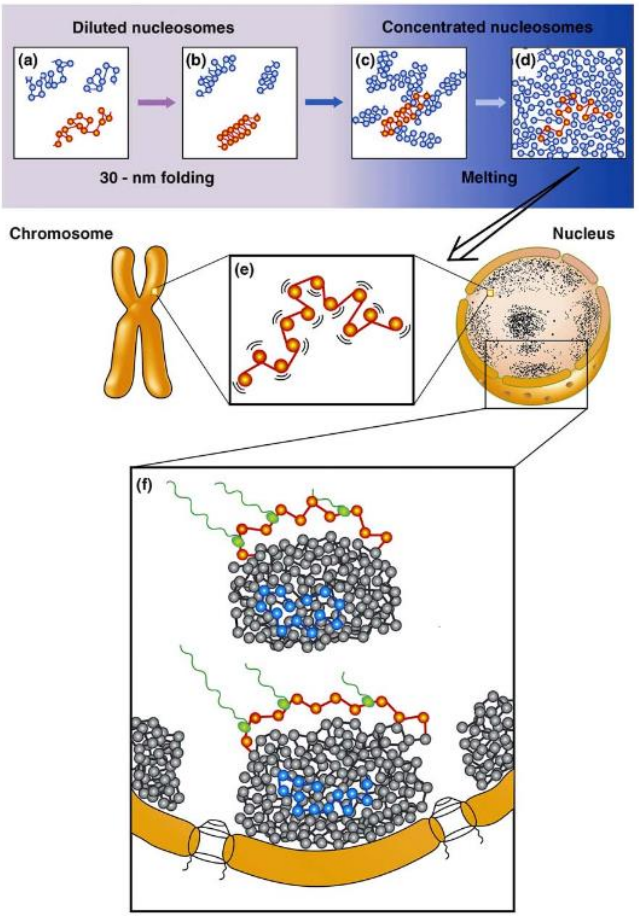


Figure I.3. "Polymer melt" model. (A, B) Under diluted conditions, nucleosomes may compact through close neighbor associations and rearrange to form 30 nm chromatin fibers. **(C)** An increase in nucleosome concentrations, leads to an increase in inter-fiber nucleosome interactions instead of intra-fiber contacts. **(D)** Nucleosomes of adjacent fibers interdigitate forming a "polymer melt" structure. **(E)** Chains in the polymer melt are nucleosomes fibers in constant movement, facilitating dynamic processes in the nucleus. **(F)** Nucleosome interactions generate higher order chromatin domains responsible for the compaction level of the chromatin and thus, responsible of the activation or repression of chromatin regions (active chromatin regions are shown in orange and inactive regions are shown in blue)[21].

1.3. Epigenetics

Epigenetics is the study of meiotically and mitotically heritable changes in gene function and expression that do not involve changes in the DNA sequence itself. Therefore, although all cells in an organism contain the same genetic information, epigenetic mechanisms mediate different gene expression patterns in different cell types[39].

Epigenetic regulation involves chromatin alterations that can be induced chemically (DNA methylation and histone modifications) or structurally (chromatin remodeling and DNA inter/intrachromosomal interactions), affecting gene transcription, and hence cellular function, without altering the primary DNA sequence[40]. Non-coding RNAs has recently emerged as an additional epigenetic mechanism[41].

DNA methylation is a key epigenetic silencing mechanism fundamental for gene regulation and cell differentiation, catalyzed by a family of DNA methyltransferases (DNMTs) that covalently transfer a methyl group to the fifth carbon of a cytosine residue[39]. Besides DNA methylation, histone modifications are the main epigenetic mechanism involved in the regulation of chromatin. Many enzymes are able to post-translationally modify histones. The most common histone modifications are methylation, acetylation, phosphorylation, and ubiquitination[42].

On the other hand, chromatin remodeling mainly refers to ATP-dependent process in which different enzymes, such as the SWI/SNF complex, shift nucleosomes in order to change genomic

conformation. This mechanism is key for many cellular processes, such as the exposure of a promoter to its transcriptional machinery[43].

Finally, non-coding RNAs are functional transcripts that are not translated into proteins. Non-coding RNAs can have important roles in epigenetic gene regulation, mainly by directly interacting with histone modifying complexes or DNA methyltransferases[41].

We will extend the explanation of histone modifications for the further understanding of this thesis.

1.3.1. Histone modifications

Nucleosomes must be correctly packaged inside the nucleus in order to allow different cellular processes, such as transcription regulation, DNA repair and replication. For this, eukaryotes have evolved a mechanism of histone modifications that influence chromatin accessibility and, in turn, chromatin organization and regulation[2,44,45]. In a general context, genes can be either actively transcribed or repressed, depending basically on the chromatin condensation pattern. “Open” or decondensed chromatin has been widely associated with gene activation, and “closed” or condensed chromatin, with gene silencing.

Post-translational modifications (PTMs) of histones represent a well-studied mechanism of influencing DNA-based processes, such as nucleosome dynamics, chromatin compaction and transcription regulation[2]. Until few years ago, it was thought that histone modifications mainly occur in the N-terminal tail of histone proteins

projecting from nucleosomes. However, modifications in the core of histone proteins have been recently well characterized[45-47].

Over 100 different modifications have been described since epigenetic field appeared; these include histone acetylation, methylation and phosphorylation, which are the most well-characterized PTMs associated with activation or repression of gene transcription[48] (Figure I.4).

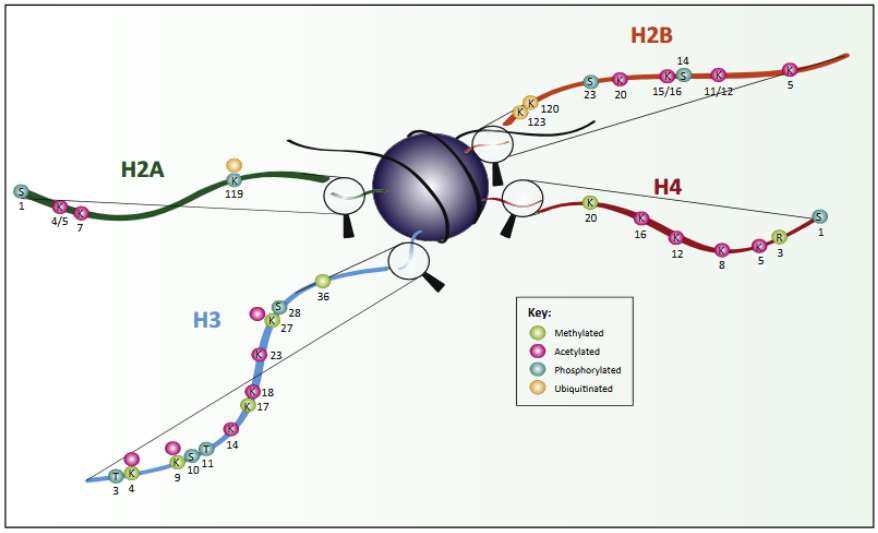


Figure I.4. Post-translational modifications of the N-terminal of histone tails. Numbers indicate the position of the amino acid modified and letters refers to the residue modification. Colors depict the type of PTMs (green; methylation, purple; acetylation, blue; phosphorylation and orange; ubiquitination)[2].

Histone PTMs can modulate chromatin regulation through two different mechanisms. First, PTMs can directly modulate the packaging of chromatin by neutralizing the positive charge of some residues, thus weakening interactions between histones and DNA allowing the chromatin becomes more accessible for transcriptional

machinery. Interestingly, histone acetylation was the first discovered PTM able to directly modulate chromatin accessibility[49].

On the other hand, PTMs can also regulate chromatin structure by recruiting specific proteins, named “histone readers”. Histone readers are able to recognize modified histone residues through specific recognition domains, such as bromodomain and chromodomain, which bind acetylated and methylated residues, respectively[50]. These histone readers reorganize chromatin fibers to allow correct regulation of gene expression (Figure I.5).

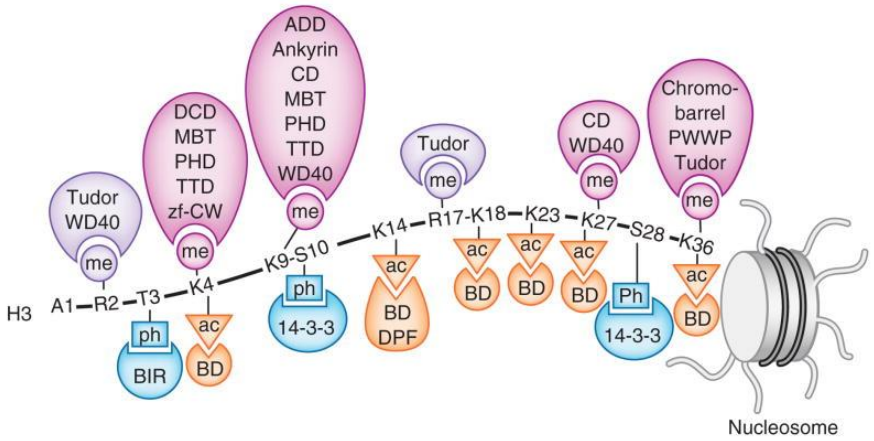


Figure I.5. Readers of histone post-translational modifications. Post-translational modifications of histone H3 N-terminal tail including methylation (me), phosphorylation (ph) and acetylation (ac) are recognized by different domains of histone readers[18].

2. Euchromatin and heterochromatin

In 1928, Heitz categorized eukaryotic genomes into two major structurally and functionally distinguishable territories, depending on their differential chromosome staining using DNA dyes: heterochromatin and euchromatin[51]. These domains have different patterns of histone modifications, are associated with different modes of nucleosome packaging and have different nuclear organization packaging[52]. This functional distinction has crucial implications for epigenetic control of gene expression programs[53]. In general, euchromatin characteristics are antagonist to those of heterochromatin, and the other way around. Therefore, in contrast to heterochromatin, euchromatin is typically less condensed, more accessible and enriched in genes, and thus more easily transcribed. It is generally located at chromosome arms, replicates in S-phase and undergoes recombination during meiosis, whereas heterochromatin remains condensed throughout the cell cycle[54,55].

Nucleosome modifications are also important distinguishing marks between these two major types of territories. Euchromatin presents hyperacetylated histone H3 and H4 tails, tri-methylation on lysine 4 and 36 of histone H3 (H3K4me3 and H3K36me3, respectively)[56,57] and (to a lesser extent) lysine 79 methylation of histone H3 (H3K79me)[58,59]. In contrast, heterochromatin is mainly characterized by hypoacetylated histones and methylation of histone 3 lysine 9 and 27 (H3K9me and H3K27me)[60,61] (Figure I.6 and Figure I.7).

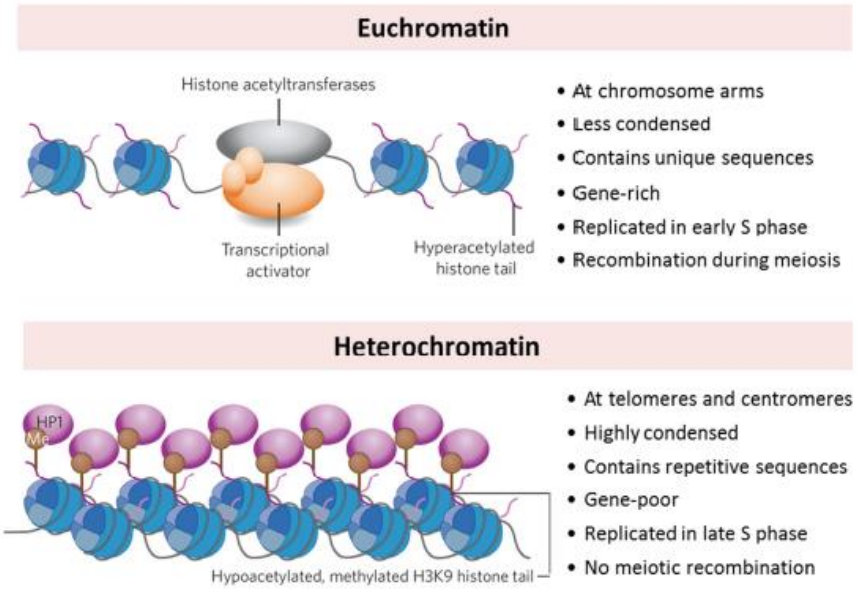


Figure I.6. Properties of euchromatin and heterochromatin domains. General characteristics of euchromatin and heterochromatin are listed. However, since include all specific properties is difficult, there are some exceptions in each case. Therefore characteristics of heterochromatin listed are representative of pericentromeric heterochromatin, located flanking centromeres in most eukaryotic genomes[17].

Heterochromatin has been further categorized into constitutive and facultative heterochromatin, depending on its genomic position[62]. Constitutive heterochromatin is mainly located in gene-poor regions that contain repeat-rich sequences, named satellites, that vary in size from 5 bp to a few hundred bp. In most organisms, satellite sequences are found in structural domains important for the maintenance of chromosomes and the transmission of genetic information, such as telomeres and pericentromeric regions, which are regions that flank centromers[63]. As these regions are repetitive

sequences, they represent a potential risk for genome instability due to their potential for self-duplication and illicit recombination. Therefore, the main function of constitutive heterochromatin is maintaining repeat-rich sequences in a transcriptionally silent state in the same genomic regions in every cell type[62,64,65].

Similar to constitutive heterochromatin, facultative heterochromatin is located in regions transcriptionally inactive, but it has the potential to interconvert heterochromatin into euchromatin, depending on whether the specific genomic regions need to be open or closed in a concrete spatial and temporal context[66]. Thus, facultative heterochromatin refers to genomic regions with developmental dynamic compaction and silencing, such as cell-type specific genes and enhancers[67].

Histone modifications also differ between both types of heterochromatin. Constitutive heterochromatin is mainly marked by histone methyltransferases that di- and tri-methylate histone 3 lysine 9 (H3K9me2 and H3K9me3)[68]. Suv39H1 and Suv39H2 are responsible for catalyzing H3K9me2 and H3K9me3[69], whereas GLP and G9a (also termed EHMT1 and EHMT2, respectively) catalyze H3K9me1 and H3K9me2[70]. However, these histone modifications *per se* are not able to compact and silence a specific region but require the contribution of other partners. Thus, heterochromatin protein 1 (HP1) binds to H3K9me2/me3 through its chromodomain and recruits other repressive histone modifiers to compact the specific chromatin region[71]. In contrast, repression of cell-type specific genes requires tri-methylation of histone 3 lysine 27 (H3K27me3), which is catalyzed by the Polycomb repressive complex 2 (PRC2)[72] (Figure I.7). Although both histone

modifications are related with gene repression, their mechanistics are completely different. Therefore, promoters marked with H3K27me3 are transcriptionally inactive yet remain accessible to binding of transcription factors and paused RNA polymerase, whereas chromatin regions marked with H3K9me2/me3 have impaired binding of transcription factors through recruitment of different DNA-binding proteins[73,74].

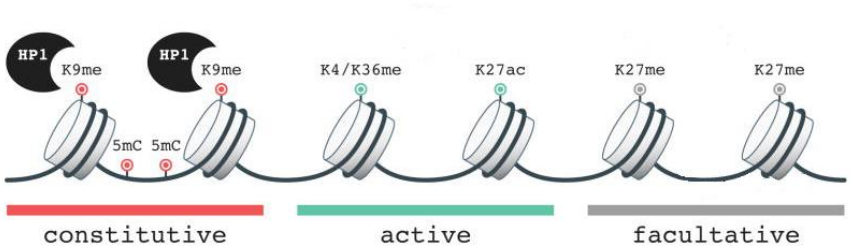


Figure I.7. Epigenetic modifications associated with different chromatin domains. Repressive features of constitutive heterochromatin including DNA methylation (5mC), H3K9me (K9me) and HP1 binding are depicted in red. Histone modifications associated with active transcription include H3K4me and H3K36me (K4/K36me) as well as H3K27ac (K27ac) and are depicted in blue. Finally, facultative heterochromatin is marked with H3K27me (K27me) as depicted in grey. Adapted[23].

2.1. Pericentromeric heterochromatin

The bulk of constitutive heterochromatin is localized in pericentromeric regions, which are regions that flank centromeres that are usually rich in A/T (major) satellites arranged in tandem arrays up to 2 Mb in length[75]. In contrast to telomeres,

pericentromeric heterochromatin varies greatly between species and even between chromosomes from the same species, suggesting that pericentromeric functions can be regulated epigenetically[62]. At the molecular level, pericentromeric heterochromatin is characterized by epigenetic repressive marks, such as DNA methylation and H3K9 tri-methylation[76].

In the 1960s, it was demonstrated that pericentromeric DNA was transcriptionally silent in differentiated tissues[77]. However, in the same decade, new insights of a possible DNA transcription of these regions were presented[78]. Currently, multiple studies have confirmed transcription of pericentromeric repeats in a multitude of organism and in different contexts, such as development[79].

Mouse cells have been widely used as a study model for constitutive heterochromatin in mammals, as they contain large blocks of pericentromeric heterochromatin. Heterochromatin flanking centromere regions from different chromosomes organize into large clusters in interphase cells, called chromocenters[13,80], which are not randomly organized inside the nucleus but instead preferably localize next to the nuclear periphery[81,82]. In mouse cells, these chromocenters can be easily visualized by DAPI, H3K9me3 and HP1α staining as well as by specific staining of repeats (Figure I.8). Nevertheless, in human cells, such clusters do not exist, as pericentromeric regions are more dispersed.

Chromocenter formation and maintenance of pericentromeric heterochromatin are likely to be important factors for separating silent heterochromatin from euchromatin, bringing in chromatin modifier enzymes to set repressive marks and coordinating proper chromosome segregation[54]. In fact, during development, the

organization of pericentromeric heterochromatin into chromocenters is dynamic[83]. This likely reflect global chromatin changes, for instance by allowing the formation of euchromatin loops in nuclear space to potentially contribute to gene expression regulation[84].

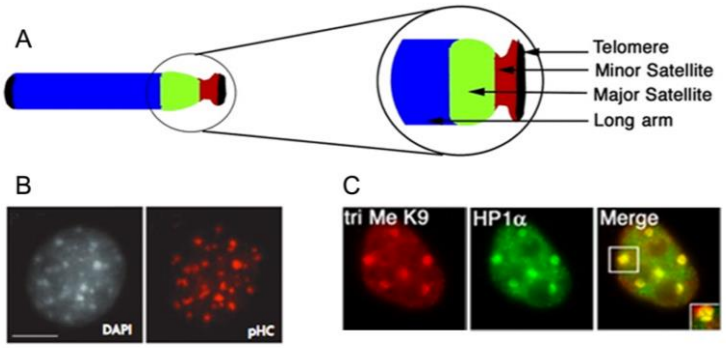


Figure I.8. Mouse cells clusters of pericentromeric heterochromatin in chromocenters. (A) Schematic representation of a mouse chromosome. Pericentromeric heterochromatin is positioned next to the centromere, composed by repetitive sequences termed major satellites (green). Minor satellites are repeats located in centromere regions (red), and telomeres are located at the end of the chromosomes (black). **(B)** Chromocenters, which are clusters of pericentromeric heterochromatin (pHC), are easily visualized by DAPI staining (left) or by DNA fluorescent in situ hybridization (FISH), using a probe for pericentromeric repeats (right). **(C)** Immunofluorescent staining of interphase nuclei with antibodies against H3K9me3 or HP1α, or the merge of both. Adapted from [13,14].

2.2. Heterochromatin protein 1 (HP1)

Heterochromatin protein 1 (HP1) is a small, non-histone chromosomal protein that is phylogenetically conserved and found highly enriched in almost all eukaryotes in pericentromeric heterochromatin[85]. In mammals, three different isoforms have been identified: HP1α, HP1β and HP1γ[86]. The main differences between these isoforms is in their localization: HP1α is mainly located in pericentromeric heterochromatin, while HP1β and HP1γ have been reported to also be associated with euchromatin regions[87].

HP1 proteins are mainly formed by two conserved domains; the amino-terminal chromodomain and the carboxy-terminal chromoshadow domain, which flank a variable hinge domain. The chromodomain of HP1α has a high affinity for H3K9 trimethylated histones[88], which are catalyzed by the histone methyltransferases Suv39H1 and H2[69]. On the other hand, the chromoshadow domain functions as a dimerization module, as it contains a vital consensus sequence found in many HP1 interacting proteins[89]. In this way, HP1α can interact with itself as well as with HP1β[90], Suv39H1/2 and the DNA methyltransferases Dnmt1 and Dnmt3a[91]. Association of HP1 with its partners allows specific histone marks to be set and thus maintains pericentromeric heterochromatin formation[20]. Finally, the hinge region binds to RNA, DNA and chromatin without sequence specificity. Indeed, the hinge domain is the least conserved domain, which might indicate that it could contribute to the proper localization of the different HP1 isoforms[20] (Figure I.9).

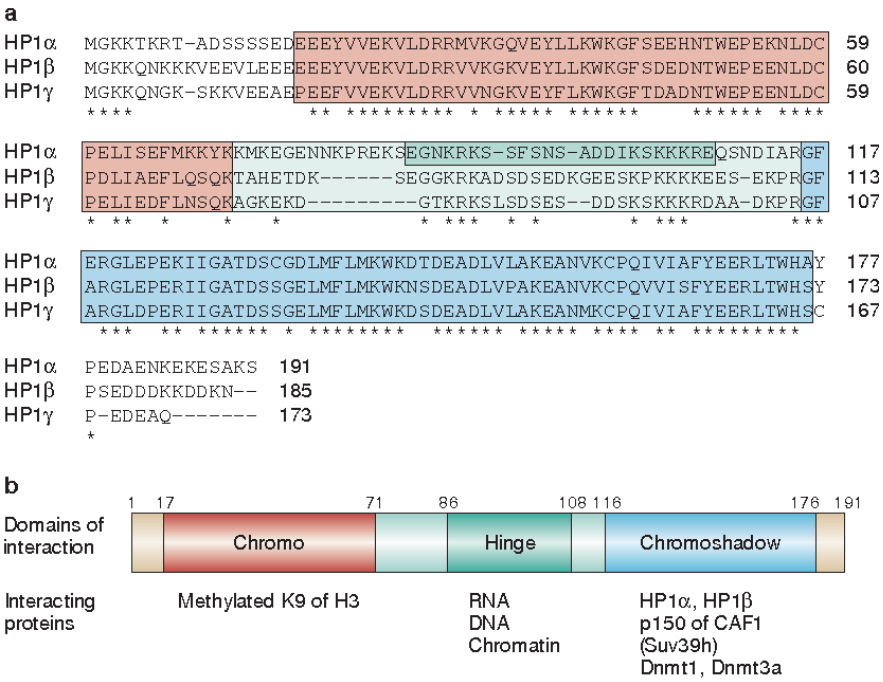


Figure I.9. Mouse HP1 interaction domains and interaction partners. (A) Amino-acid sequences of mouse HP1 isoforms α, β, γ. Asterisks indicate the amino acids that are conserved between three subtypes. Colors indicate different HP1 domains. **(B)** Schematic representation of HP1 interaciton domains and interacting proteins. Chromodomain is depicted in red, hinge region in green and chromoshadow domain in blue[20].

A few years ago, various studies were carried out that demonstrated that, in contrast to what was previously believed, the maintenance of stable heterochromatin domains in living cells requires the transient binding and dynamic exchange of HP1 from chromatin[92]. Therefore, heterochromatin might be much more dynamic than initially imagined, with transcription occurring and helping to set up heterochromatic state.

3. Nuclear architecture

The eukaryotic nucleus is a three-dimensional complex environment in which the genome must organize[8,93]. Genome functions and, consequently, transcriptional output depend not only on local modifications of the chromatin fiber but also on spatial organization inside the nucleus[7,94]. Although the localization of chromosomes is highly variable between cells, a general picture of chromosome organization has recently begun to emerge. Fluorescent in situ hybridization (FISH) experiments with chromosome paints have revealed that chromosomes are not randomly organized inside the nucleus but have preferential positions depending on their gene density, forming the chromosome territories in which the DNA of each chromosome occupies a spatial well-defined region[8]. Thus, gene-dense chromosomes tend to localize in the nuclear interior, whereas gene-poor chromosomes reorganize in the nuclear periphery[95,96]. However, this organization is dynamic, and particularly highly expressed genes tend to protrude outside their chromosome territory, allowing intermingling between different chromosomes located in different territories[97] (Figure I.10).

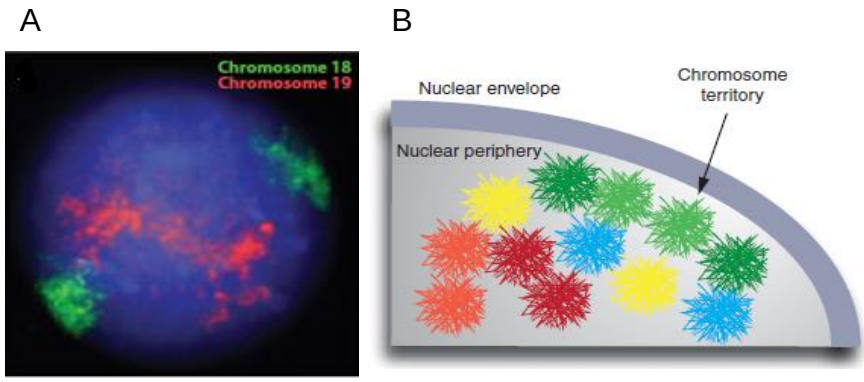


Figure I.10. Spatial organization of chromosome territories. (A) Two different chromosome territories obtained by FISH of a human nucleus hybridized with chromosome paints for gene-rich chromosome 19 (red) and gene-poor chromosome 18 (green). DNA is stained with DAPI (blue). **(B)** Schematic representation of chromosome territories. The distribution of chromosomes is not random with some chromosomes preferentially occupying internal regions (red) and others peripheral positions (green). Intermingling of different chromosome territories is also showed. Adapted[7,8].

3.1. Nuclear lamins

3.1.1. Structure and organization

Nuclear lamins are V-type intermediate filament (IF) proteins only present in metazoans[98]. Like all IFs, lamin polypeptides contain a small amino-terminal (N-terminal) “head” domain, a long central “rod” domain and a large carboxy-terminal (C-terminal) “tail” domain. The C-terminal domain contains a nuclear localization signal (NLS), an immunoglobulin (Ig)-fold motif and a CaaX motif (whereby C is cysteine, a is an aliphatic amino acid and X is any amino acid)[9] (Figure I.11.A).

The organization of lamins into higher order structures allows them to assemble into lamin filaments, which are important for the assembly, structure and mechanical stability of metazoan nuclei, which in turn regulate chromatin organization and gene expression[10]. For this organization, lamins first dimerize through their rod domain and then dimers polymerize to form a head-to-tail linear polymer. Lateral assembly of two polymers can then form a protofilament. Finally, groups of three to four protofilaments further assemble into 10 nm diameter filaments, which are called intermediate filaments[99] (Figure I.11.B).

Although numerous studies have been performed over the last decades to better characterize nuclear lamins, the exact *in vivo* organization of the lamin network in somatic cells is still elusive. Increasing evidence suggests that different lamins form separate filamentous networks that interact with each other, yet how they are interconnected is still not known[100].

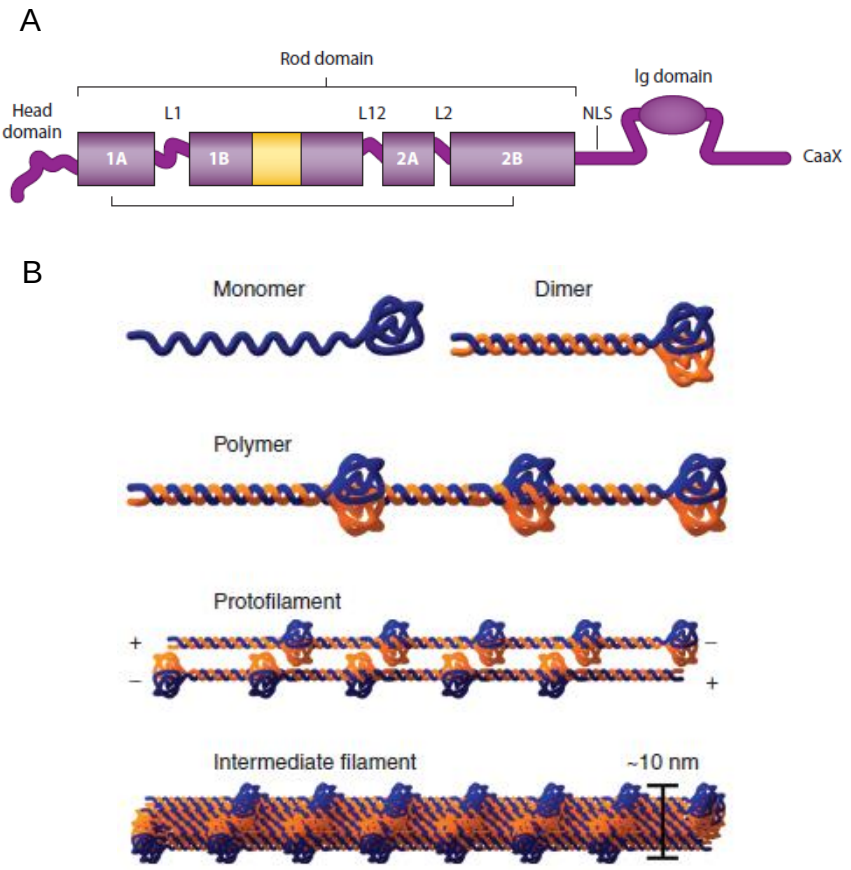


Figure I.11. Structure and assembly of lamins. (A) Domain organization of a lamin monomer showing the "head" domain; the central "rod" domain composed of four α-helices (1A, 1B, 2A and 2B) separated by three linker regions (L1, L12 and L2); and the C-terminal "tail" domain, which includes the NLS, an Ig domain and the CaaX motif. **(B)** Schematic model of lamin polymers showing dimerization through two rod domains, head-to-tail polymers formation and parallel assembly of two dimers to generate the protofilament; in turn, three or four of these forms the 10 nm intermediate filament. Adapted from [9,10].

3.1.2. Types of nuclear lamins

The nuclear lamins are mainly formed by so-called A- and B-type lamins[101,102]. In mammals, B-type lamins are encoded by two different human genes: *LMNB1* (encoding lamin B1) and *LMNB2* (encoding lamin B2 and lamin B3). Lamin B1 and lamin B2 are the two major B-type lamins in most vertebrates and are widely expressed in somatic cells[103]. In contrast, lamin B3 is generated through alternative splicing of *LMNB2*, and its expression is restricted to the testis[104]. The two major A-type lamin isoforms, lamin A and lamin C, are derived from a single human gene (*LMNA*) by alternative splicing[105] and are expressed also in somatic cells. Alternative splicing of *LMNA* in some vertebrates can also produce two less abundant isoforms: lamin C2, which is only expressed in testis, and lamin AΔ10, which is found in a variety of cell types[106,107]. Although different lamins can be expressed in different cell types, the main distinction between them is the timing of their protein expression. Specifically, B-type lamins are ubiquitously expressed, whereas A-type lamins are developmentally regulated and are generally expressed in differentiated cells[108,109].

Structurally, A- and B-type lamins differ in their C-terminal “tail” domain. While lamin B1 and lamin B2 have almost identic structures, lamin A has a longer “tail” domain, and lamin C lacks the entire CaaX motif, which is a region that usually undergoes posttranslational modifications. Thus, lamins A, B1 and B2 are expressed as prelamins that require modifications in their C-terminal domain to become mature lamins.

Specifically, during maturation, the cysteine residue of the CaaX box is first farnesylated by a farnesyltransferase. This modification allows the removal of –aaX by a protease and the final carboxymethylation of the carboxy-terminal cysteine[110-112]. B-type lamins remain farnesylated and carboxymethylated permanently, in contrast to lamin A, which needs an additional removal of the last 15 amino acids upstream of the farnesylated cysteine of the C-terminal domain. This additional step results in a mature lamin A that lacks the farnesylated and carboxymethylated modifications in the C-terminal tail domain[113]. It is thought that farnesylation and carboxymethylation leads to stable integration of lamins in the nuclear membrane; however, this is not a prerequisite for tethering lamins to the nuclear envelope, as experiments inhibiting –CaaX modifications do not likewise inhibit their incorporation into the nuclear membrane[24].

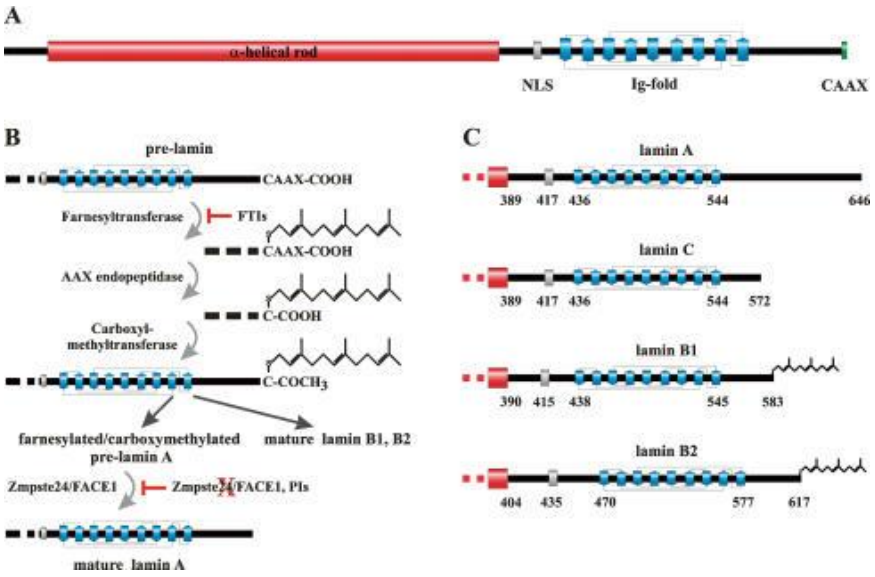


Figure I.12. Structure and posttranslational modifications of different types of nuclear lamins. (A) Schematic drawing of a pre-lamin polypeptide chain. **(B)** Posttranslational processing of lamin B1, B2 and A. A farnesyl group is attached to the cysteine residue of the –CaaX motif by a farnesyltransferase (FTIs), the last three residues –aaX are cleaved off by an AAX endopeptidase and the carboxylic acid group (–COOH) of the C-terminal cysteine residue is methylated by a carboxyl methyltransferase. These steps are common for all type lamins and lead to mature lamin B1 and B2. However, lamin A needs an additional cleavage of the 15 C-terminal residues by Zmpst24/FACE2, including the farnesylated/carboxymethylated cysteine to become mature[24].

3.1.3. Nuclear localization

The nuclear envelope (NE) represents one of the structural domains of the nucleus. The NE is mainly formed by two concentric membranes, called “inner” and “outer” nuclear membranes (INM and ONM, respectively), which are separated by a lumen. In addition, nuclear pore complexes (NPC) are able to cross the

nuclear envelope to mediate the movement of particles from the cytoplasm to the nucleoplasm[114,115] (Figure I.14).

Nuclear lamins are mostly found in an insoluble pool forming the nuclear lamina, which is located in close contact with the inner face of the nuclear envelope membrane in metazoans (INM)[24]. Importantly, lamins represents one of the major structural components of the nucleus related with maintenance, shape and mechanical stability, since alterations in any of the nuclear lamins lead to misshaped nuclei and dysfunction of many nuclear processes, such as premature senescence and altered proliferation rates[116,117]. Therefore, nuclear lamina forms a filament network surrounding nuclear envelope that contact with proteins located in the INM as well as with chromatin, providing structural support to the nuclear envelope and helping chromatin organization inside the nucleus[118]. As chromatin located near the periphery is generally heterochromatin, nuclear lamina mainly interacts with this repressed and gene-poor type of chromatin.

Although the major fraction of nuclear lamins has been reported to be associated with the nuclear lamina, these proteins are also present in the nucleoplasm, where they form a soluble network throughout the nucleus that is distinct from the filamentous meshwork of the nuclear lamina[119-121]. Indeed, antibodies against nuclear lamins stain mainly the nuclear periphery but also (to a lesser degree) the nuclear interior of many metazoan nuclei. For a long time, this nucleoplasm staining was considered an artifact of immunostaining procedures, as it was assumed that nuclear lamins were only present in the nuclear periphery. Later, novel antibodies

and techniques shed light on the authenticity of nuclear lamins located in the nucleoplasm[9,122] (Figure I.13).

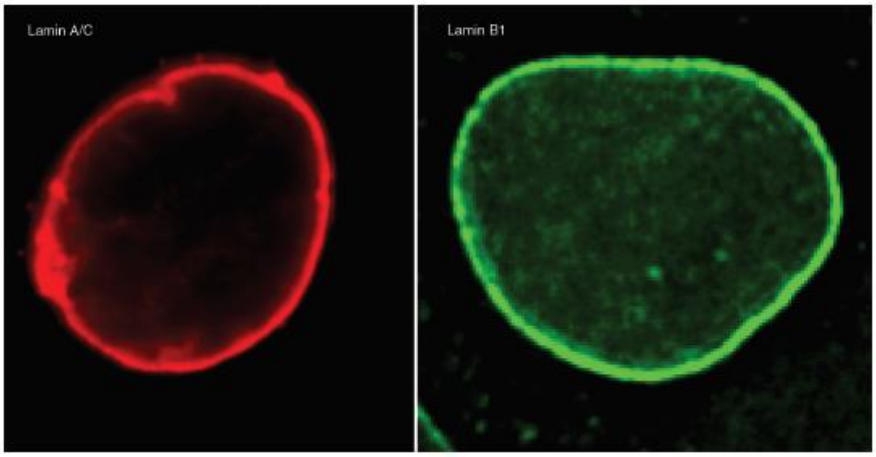


Figure I.13. Nuclear lamins localize in the nuclear periphery and in the nuclear interior. Immunofluorescence staining of lamin A/C (red) in U2OS human osteosarcoma cell nuclei and lamin B1 (green) in mouse embryonic fibroblast (MEF) cell nuclei[9].

The exact role of the small fraction of nucleoplasmic lamins is still unclear[100]. Recent studies suggest that nucleoplasmic lamins may have a role in DNA replication and transcription. Moreover, as nucleoplasmic lamins are more soluble than nuclear lamina, they may function as an intermediate state of immature lamins on their way to be incorporated into the nuclear lamina[24,123]. In any case, nuclear lamins represent an essential structural component of nuclear architecture involved in most nuclear functions[100].

3.2. Chromatin organization

Cell-type specific transcriptional regulation is crucial for maintaining cell identity throughout the lifetime of an organism. This regulation

is mediated not only by molecular factors (such as histone and DNA modifications) and specific transcription factors but also by the spatial compartmentalization and organization of the genome[37].

3.2.1. Lamina-associated domains (LADs)

During the past decade, genome-wide mapping methods have identified regions of chromatin in close contact with the nuclear lamina, termed lamina-associated domains (LADs)[124,125] (Figure I.14). LADs have been identified primarily using the DamID technique, which is based on introducing a bacterial adenine methyltransferase (Dam) tethered to an NL protein (typically lamin B1) to lead to adenine methylation of regions that are in contact with this NL protein. These modifications can be visualized by microscopy or mapped genome-wide[124]. Chromatin immunoprecipitation (ChIP) experiments can be also used to identify LADs by using a specific antibody against the NL protein of interest[126].

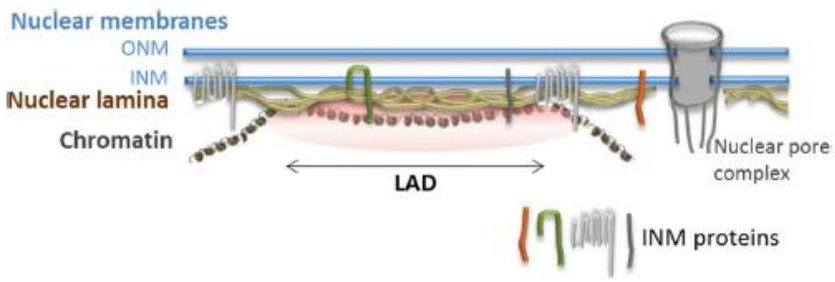


Figure I.14. Schematic representation of nuclear envelope and LADs. The outer (ONM) and inner (INM) nuclear membranes are perforated with the nuclear pore complexes (NPC). The INM harbors a specific set of proteins some of which interact with nuclear lamina or directly with chromatin. The nuclear lamina interacts with chromatin through relative large domains termed lamina-associated domains (LADs)[3].

Mouse and human cells contain between 1000 and 1500 LADs, which are typically 0.1 to 10 Mb in size; as LADs cover one-third of the entire genome, they represent one of the major key features of epigenome[127]. As expected by its position in the nuclear periphery, chromatin regions in contact with nuclear lamina possess several molecular characteristics related to heterochromatin. Indeed, most genes in LADs are transcriptionally silent, overlap with compact regions that replicate late during S-phase and are enriched in A/T sequences and repressive histone marks, mainly H3K9me2 and H3K9me3[128].

Several genome-wide studies have revealed that LADs are divided into constitutive LADs (cLADs), which are cell-type invariant, and facultative LADs (fLADs), which contact the NL in only certain cell types[129]. cLADs represent the majority subset of total LADs and are gene-poor and A/T-rich. Their genomic position and size, but not their actual DNA sequences, are highly conserved between

species. fLADs represent approximately 30% of LADs and are more gene-rich as well as less conserved than cLADs. Indeed, during cell differentiation, many genes change their positions relative to the NL, demonstrating the dynamism of genome spatial organization[130] (Figure I.15).

Even though new advances in genome-wide studies over the last decades have shed light on how the genome is organized with respect to nuclear lamina, many questions still remain enigmatic. Indeed, how LAD chromatin regions drive NL interactions is not fully understood. Several reports indicated that H3K9 methylation might have a prominent role, as inhibition of methyltransferases that specifically methylate these residues leads to at least partial detachment from the NL[131,132].

Just as LADs are located in the nuclear periphery, chromatin regions between LADs (inter-LADs) must be located away from the NL. In the nuclear interior, RNA polymerase II has been found to cluster in distinct foci, named “transcription factories”, which contact specific genes primarily located in inter-LADs and allow their active transcription[133]. In the same way, changes in the chromatin state—but not in transcription *per se*—are responsible for the detachment of specific regions from the NL[134]. Indeed, studies performed in embryonic stem cells (ESCs) during differentiation to different cellular lineages demonstrated that not all genes that detach from the NL becomes immediately activated but are more prone to activation in subsequent differentiation stages, suggesting that release from the NL may “unlock” these genes for activation[129].

LADs tend to have sharply defined borders that are normally located 5–10 Kb outside LADs. In mammals, LAD borders are enriched in active promoters and thus in the H3K4me3 histone mark. These frontiers are also enriched in CpG islands and are bound by the architectural CTCF DNA-binding protein, which is involved in the establishment of chromatin domain boundaries, suggesting possible mechanisms of borders in LAD confinement to the nuclear periphery. Intriguingly, the facultative heterochromatin histone mark H3K27me3 is also present in the regions immediately adjacent to LAD borders, where it probably prevents the spread of active chromatin from neighboring regions[128] (Figure I.15).

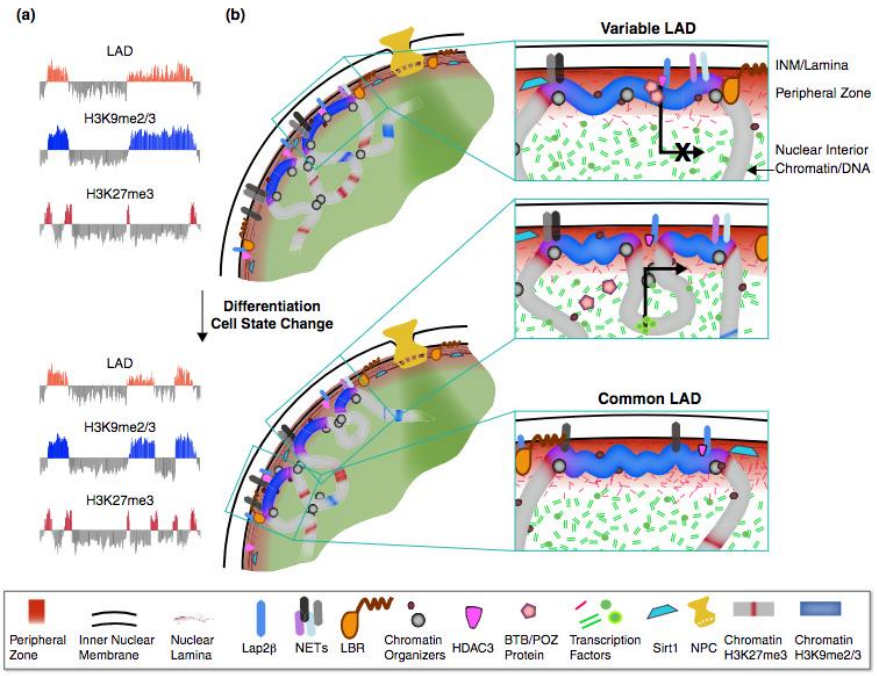


Figure I.15. Representation of cLADs and fLADs characteristics during differentiation. (A) LADs, H3K9me2/3 and H3K27me3 molecular profiles of chromatin are depicted and dynamically change during cell state changes or differentiation, with H3K9me2/3 typically located inside LADs and H3K27me3 located at borders. **(B)** Schematic representation of cLADs and fLADs (variable LADs). cLADs regions remain at the nuclear periphery before (top) and after (bottom) cell differentiation or cell state changes, whereas variable LADs move to or away from peripheral zone. Components involved in providing specific characteristics to LAD regions are depicted in the bottom square[25].

3.2.2. Topologically associating domains (TADs)

Genomes do not function in a sequential fashion but rather are folded into three-dimensional (3D) space; this intricate folding allows genomic regions located very remotely to contact and regulate each other[11].

Cutting-edge technology, including chromosome conformation capture (3C), as well as conventional and super-resolution imaging have helped to unravel the complexity of 3D chromatin architecture in the past decade. Specifically, 3C is a biochemical procedure based on the analysis of *in vivo* contact frequencies between different pair of selected regions[135]. The original 3C technology revealed the existence of contacts between disperse regulatory elements, such as enhancers, with their target genes in a tissue-specific manner dynamic during development, thus generating regulatory enhancer-promoter loops[136]. Thus, the original 3C technology allowed interactions between a pair of loci ("one-to-one" approach) to be studied. The more-recent development of high-throughput methods and their application to the 3C technology have greatly improved the understanding of genome folding and architecture in a more systematic and genome-wide fashion[137]. These methods include 4C or circularized 3C ("one-to-all")[138], 5C or 3C carbon copy ("many-to-many")[139] and Hi-C or chromosome capture followed by sequencing ("all-to-all")[140] technologies (Figure I.16).

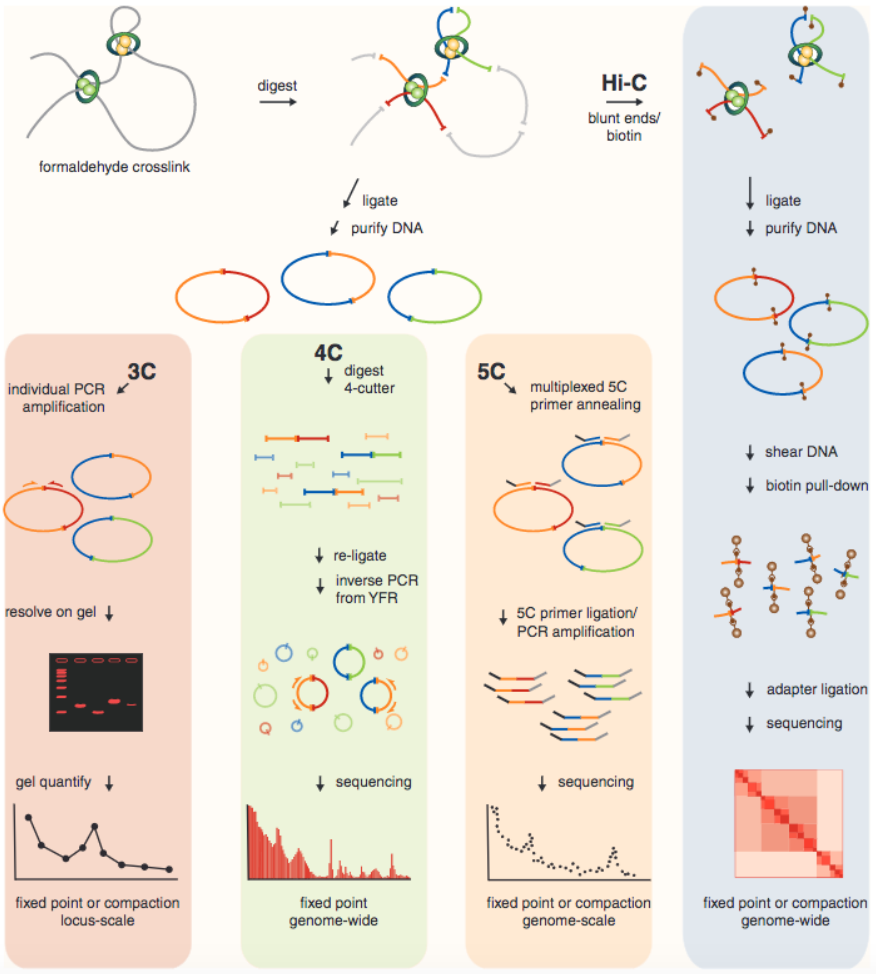


Figure I.16. Principles of 3C-derived methods. All methods derive from the Chromosome Conformation Capture (3C) protocol and are based in a first formaldehyde crosslinking step followed by an enzymatic digestion and a ligation step. DNA is then purified and the final ligated product detected. However, each method uses different approaches to generate genomic libraries. 3C employs secondary digestion. 4C employs inverse PCR to detect all fragments ligated to a locus of choice. 5C uses multiplexed ligation amplification to detect large numbers of interactions simultaneously using a pool of primers for thousands of genomic regions of interest. Finally, Hi-C employs biotinylated nucleotides prior to DNA ligation and posterior precipitation with streptavidin beads of DNA fragments containing ligation junctions. Ligation products are quantified by real-time PCR in the first method and by high-throughput sequencing in the others[15].

These new 3C methods confirmed the presence of chromosome territories as well as the limited intermingling between them and the spatial segregation of active and inactive chromatin regions in the nucleus. Indeed, analyzing Hi-C data made it evident that the genome is divided into two major compartments, termed the A and B compartments. The A compartment is generally gene-rich, transcriptionally active and accessible, while the B compartment represents a more repressed environment with fewer genes[140] (Figure I.17A). However, this bimodal distribution does not totally reflect *in vivo* situations, as some genomic regions might switch from one compartment to another during cell differentiation[141].

Beyond chromosomal compartments, genomes subdivide into domains that display high internal contact frequencies, which are named topologically associating domains (TADs). Human and mouse genomes contain more than 2000 TADs, which cover more than 90% of the entire genome and have a median size of less than 1 Mb[142]. It is now widely believed that TADs not only form structural entities but represent the basic functional unit of chromosome organization. Indeed, in mouse cells, long-range interactions occur mainly within the same TAD (intra-TAD interactions), with fewer than 4% taking places between different TADs (inter-TAD interactions)[143].

Regions between TADs are called TAD borders, which are highly stable structures with specific topological characteristics[144]. In mammals, TAD borders are enriched in transcription start sites (TSS)[142]. Accordingly, transcription rates in these regions are also high, and active histone marks (H3K4me3) prevail. Just as for LADs, TAD borders are enriched for the CTCF insulator protein, an

architectural protein involved in the strengthening of intra-TAD interaction and in preventing long-range interactions between adjacent TADs, and thereby in maintaining TAD border integrity[145]. Therefore, the strength of TAD borders is defined as the ratio between intra-TAD and inter-TAD interactions around border sequences, and it correlates with the amount of architectural proteins present[146].

Given that TAD borders are highly stable between cell types and even between species, the dynamism observed in long-range contacts during cell differentiation must take place within TADs. Indeed, *in situ* Hi-C experiments have identified about 10,000 loops (mostly between enhancer-promoters), with an average size of 185 Kb, that were termed sub-TADs (Figure I.17B).

Globally, genome may be hierarchically organized into TADs, which are stable units maintained throughout cell differentiation and development that, at the same time, are organized into sub-TADs domains or chromatin loops that undergo more dynamic changes in cell-type specific rearrengements[147].

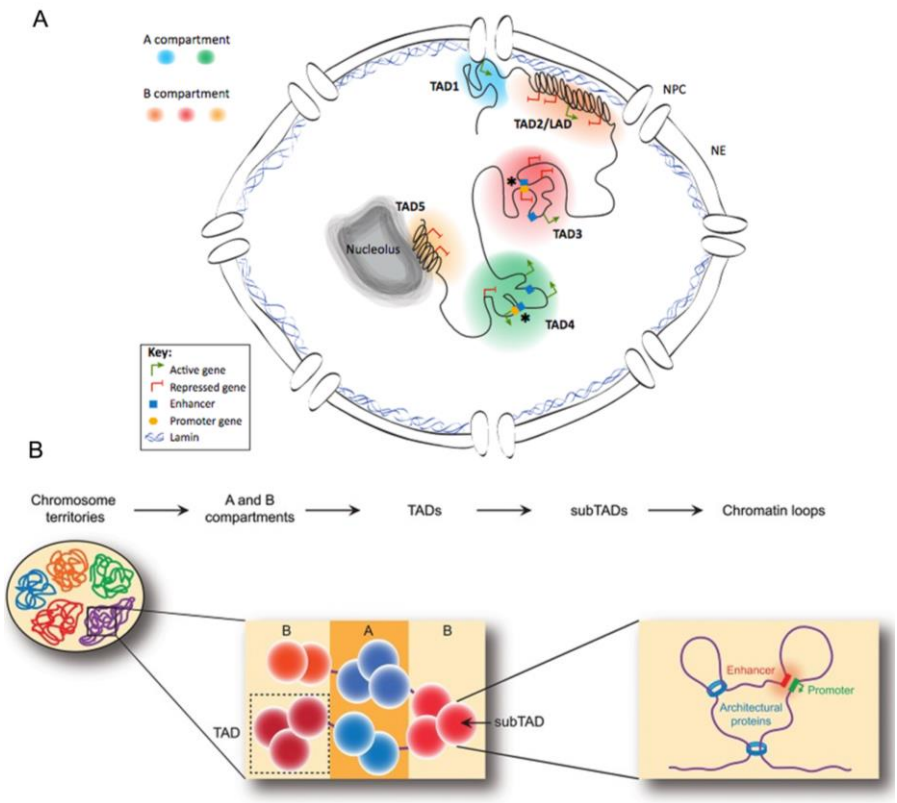


Figure I.17. Representation of genome organization into different domains. (A) Genomes divide into A (red and orange) and B (green and blue) compartments mainly depending on the activity of the genomic region. Thus, B compartment contains mainly repressed regions, such as those in contact with the NL, while A compartment primarily represents an active genomic region. At the same time, inside compartments, chromatin is organized into unit blocks termed Topologically-Associating Domains (TADs). **(B)** Hierarchical genome organization starting from chromosome territories and finishing with chromatin loops or sub-TADs harboring contacts mainly between enhancers and promoters. Adapted[11,12].

3.2.3. Architectural proteins

Architectural proteins, also known as insulator proteins, have been detected in organisms ranging from yeast to human and play a critical role in the 3D organization of the genome by regulating long-range contacts and therefore gene expression. As previously mentioned, the strength of TAD borders is partially influenced by the presence of architectural proteins. Hence, sites with few architectural proteins are weak borders, and *vice versa*. Indeed, many functional studies have reached the same conclusion: depleting architectural proteins from TAD borders results in increased inter-TAD interactions and, consequently, in decreased border strength and aberrant chromatin interactions[145,149]. These data imply that architectural proteins may regulate TAD structure.

The best characterized architectural protein in vertebrates is CTCF, which requires association with its partner cohesin for its function[5]. However, a wide range of proteins that interact or co-localize with CTCF have been also described, such as Ying Yang 1 (YY1), MYC-associated zing finger protein (MAZ) and TFIIIC, among others[148]. Although a clear correlation between TAD borders and the presence of architectural proteins exists, only 15% of total CTCF sites are located in TAD borders in mouse and human cells, while about 85% are located within TADs. Indeed, the original function ascribed to architectural proteins was their ability to insulate promoters from the effect of regulatory regions. Recent evidence suggests that architectural proteins located inside TADs may facilitate interactions between enhancers and gene promoters, while those present in TAD borders may function as insulators,

preventing interactions between genes and regulatory regions located in different TADs[150] (Figure I.18).

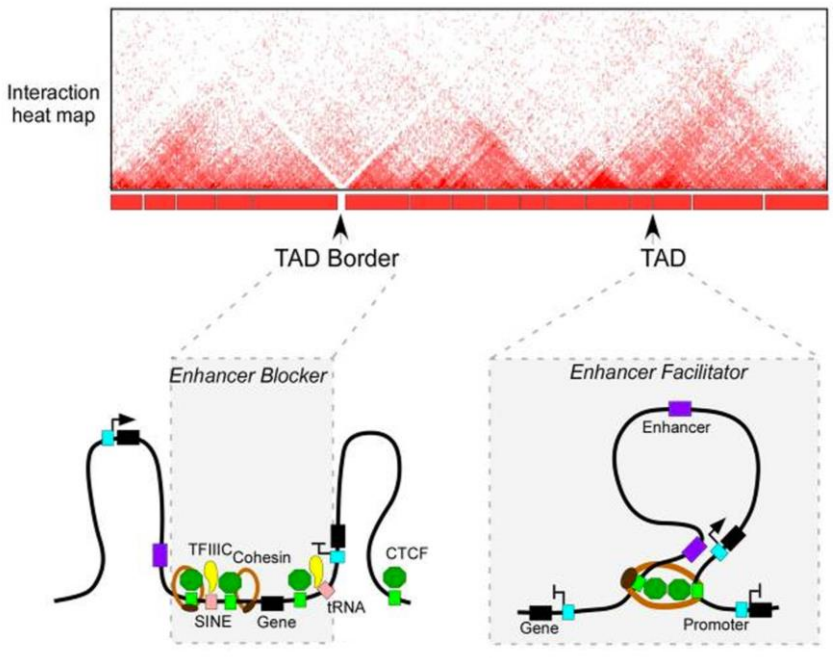


Figure I.18. CTCF regulates three-dimensional genome architecture. Interaction heat map of a chromosome segment around 2.5 Mb in length, depicting data from Hi-C experiment in a mammalian cell line. TAD and TAD border are depicted. The presence of architectural proteins such as CTCF, contribute to the establishment of the border acting as an enhancer blocker (left). On the other hand, CTCF binding sites within TAD, facilitates the interaction between long distance enhancer-promoter regions (right)[5].

Interestingly, architectural proteins are also associated with genomic elements that are important regulators of gene expression other than enhancers and promoters, such as Polycomb response elements (PREs) and sites involved in sister chromatid cohesion[151].

Even though most architectural proteins are ubiquitously expressed, the mechanisms that influence their genomic

positioning are still poorly understood. Recent results suggest that the occupancy of architectural proteins may be modulated by DNA methylation, which may negatively influence the binding of architectural proteins to DNA, or other posttranslational modifications that can modulate both their interactions with DNA and interactions between different architectural proteins. Poly-ADP-ribosylation of CTCF, for example, regulates DNA binding of CTCF in mammals and interactions between CTCF and another partner (CP190) in *Drosophila.* However, architectural proteins are regulated not only by covalent modifications but also by other factors; for instance, the presence of long non-coding RNAs (lncRNAs) is necessary for the formation of some of architectural protein complexes[152].

A wide range of proteins have been characterized to be important for loop formation within a specific genomic locus in mammals. However, whether these proteins also function as genome-wide regulators of chromatin interactions is still not known. Thus, beyond CTCF and cohesin, other potential architectural proteins important for long-range interactions might be present in mammals, but this requires further investigation[93].

4. Epithelial-to-mesenchymal transition (EMT)

The EMT program describes a series of events in which epithelial cells lose their main characteristics and acquire a mesenchymal phenotype. Accordingly, cells that undergo this process undergo deep changes in cell behavior and architecture[153]. Moreover, as this process is a transition, an intermediate phenotype exists in which cells have only passed partially through EMT and maintain both epithelial and mesenchymal markers[154]. Changes between epithelial to mesenchymal cells were first discovered in the primitive streak of chicken embryos[155], although the reverse process—mesenchymal-to-epithelial transition (MET)—has also been reported to be necessary in many cellular contexts[156,157].

Epithelial cells form apical-basal polarized structures that are held together to form cell-cell junctions through cell adhesion proteins, such as claudins and E-cadherin. Vertebrate epithelial cells contact to each other through tight junctions, desmosomes, gap junctions and adherent junction, which are essential for epithelial integrity[158]. Upon the initiation of the EMT process, a key event in all tissue contexts is the dissolution of epithelial cell-cell junctions accompanied by the loss of either function or expression of cell adhesion proteins, such as E-cadherin, claudins and occludins, among others. For instance, during the destabilization of adherent junctions, E-cadherin is cleaved at the plasma membrane and subsequently degraded. As EMT progress, the expression of junction proteins is repressed, which stabilizes the loss of epithelial junctions to reduce adhesion and increase cell motility[159,160]. Moreover, loss of adherent junctions between cells triggers a

reorganization of cortical actin cytoskeleton, which enables dynamic cell elongation and motility[161].

The repression of genes encoding proteins responsible to form adherent junctions is accompanied by activation of mesenchymal genes, such as N-cadherin (mesenchymal neural cadherin), fibronectin, vimentin, Snail and αSMA (α-smooth muscle actin), among others. Finally, there is also an increase of expression of metalloproteases (MMP2 and MMP9), which are responsible for degradation of the extracellular matrix (ECM) proteins, enabling invasion into underlying tissues through a matrix of fibronectin[154] (Figure I.19).

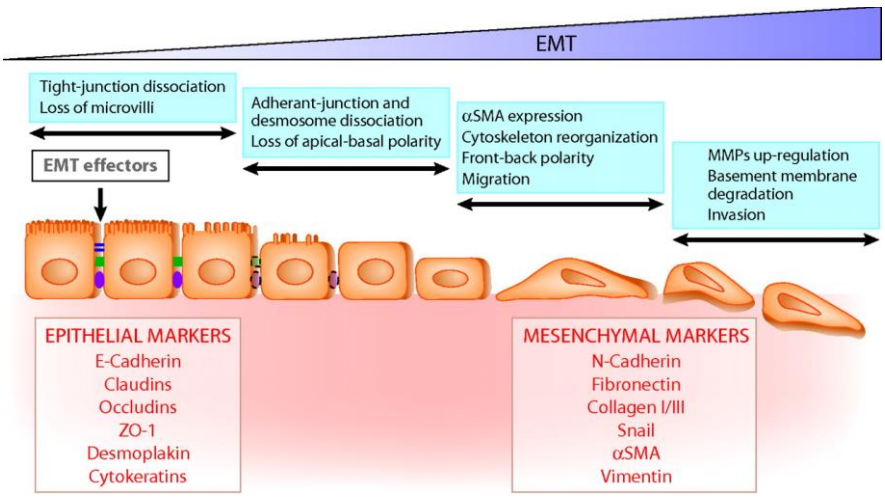


Figure I.19. Cellular events during EMT. Representation of the key events occurring during the EMT process. First, epithelial junctions are dissociated concomitantly with the repression of epithelial markers, which leads to less of cell-cell contacts and apical basal polarity. Consequently, cytoskeleton reorganization takes place allowing cells to migrate and subsequently expression of mesenchymal genes leads to base membrane degradation and acquisition of invasive capacities[4].

Conversion of epithelial cells to mesenchymal cells is a crucial process that takes place during embryonic development as well as in pathological conditions, such as fibrosis and cancer.

4.1. Classification of EMT

A consensus was reached several years ago to classify EMT into three different biological subtypes, which have distinct functional consequences based on their biological context: types 1, 2 and 3.

Type 1 EMT occurs during implantation, embryo formation and organ development, and is associated with endoderm and mesoderm formation from the primitive streak as well as with mobile neural crest cells[155]. Primitive epithelium undergoes EMT to give rise to mesenchymal cells (primitive mesenchyme), which have the potential to undergo MET to generate the secondary epithelium[162]. Thus, type 1 EMT consists in successive waves of EMT-MET to generate embryonic tissues and organs.

The second type of EMT is associated with wound healing, tissue regeneration and organ fibrosis. In contrast to type 1, type 2 EMT begins with an inflammation process caused by a trauma or an injury. In this case, epithelial cells undergo EMT to generate fibroblasts and other related cells in order to regenerate tissues, and primarily kidney, liver, lung and intestine tissue[163]. Normally, this type of EMT concludes once the inflammation is attenuated. However, in organ fibrosis, type 2 EMT can continually respond to persistent inflammatory signals (chronic inflammation) with detrimental effects that can even include organ destruction[154].

Finally, type 3 EMT is related to cancer progression and metastasis. Carcinoma cells normally undergo epigenetic changes in both tumor suppressor genes and oncogenes, leading to outcomes far different from those observed in the other two types of EMT. In this case, epithelial tumor cells became more invasive after undergoing EMT, accessing the circulatory system through intravasation. This results in dissemination of cancer cells to sites distant from the primary tumor. Once these circulating cells achieve secondary sites, they can undergo the inverse process of MET, which is termed as metastatic colonization of secondary organs[164]. Therefore, activation of the EMT process has been proposed as a mechanism responsible for the acquisition of malignant phenotypes in different types of carcinomas[165] (Figure I.20).

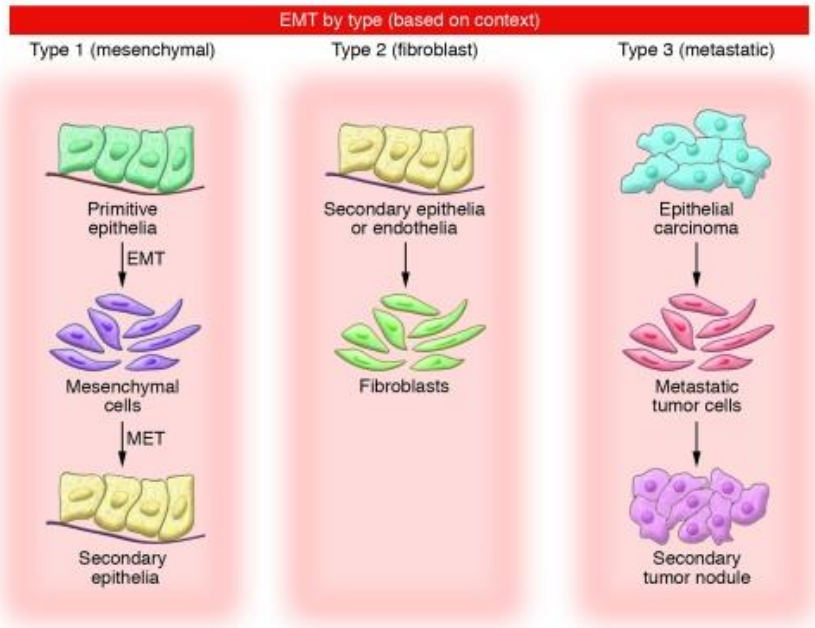


Figure I.20. Three different types of EMT depending on its biological context. Type 1 is observed when primitive epithelial cells transition into primitive mesenchymal cells that have the potential to develop the reverse process, MET, to generate the secondary epithelia. Type 2 EMT takes place in inflammatory tissues where secondary epithelia cells generate fibroblasts through EMT process to repair the injured tissue. Finally, type 3 EMT is a part of the metastatic process where epithelial cells from a primary tumor undergo EMT to migrate to a new tissue site and form a secondary tumor through MET. Adapted[22].

4.2. EMT signal inducers

A huge variety of signaling pathways are responsible for the induction of the EMT program in both physiological and pathological situations. Many of these signaling pathways are triggered by soluble ligands capable of binding to specific receptors, which in turn activate transcription factors able to induce or repress specific genes involved in the EMT process. Among EMT signal inducers, the most common are the transforming growth factor β (TGFβ) family, ligands that bind tyrosine kinase receptors (including FGF, EGF and IGF), Wnt, Hedgehog and Notch. In addition, EMT can also be induced by components of the ECM, such as collagen via its binding to integrin receptors, which are partially responsible for cell adhesion[153,166,167] (Figure I.21).

The TGFβ pathway is the most well-characterized pathway that triggers EMT. TGFβ signals either through SMAD proteins (SMAD-dependent) or other proteins (SMAD-independent), such as GTPases of the Ras family[168] (Figure I.21). In both cases, signaling is activated by a superfamily of TGFβ ligands, which includes three isoforms of TGFβ (TGFβ 1, 2 and 3), among others. Signaling occurs through the formation of heteromeric receptor complex, formed by two type 1 serine-threonine kinase receptors (TGFβRI) and two type 2 receptors (TGFβRII). In response to ligand binding, TGFβRII phosphorylates and activates TGFβRI, which leads to activation of cellular responses to TGFβ signaling pathway. In the case of SMAD-dependent signaling, phosphorylated TGFβRI is responsible for phosphorylating the C-terminal domain of the Smad 2 and Smad 3 proteins[169]. Phosphorylated Smad 2 and -3 then

form trimers with Smad 4, which translocate into the nucleus to specifically regulate and cooperate with transcription factors involved in EMT progression[170].

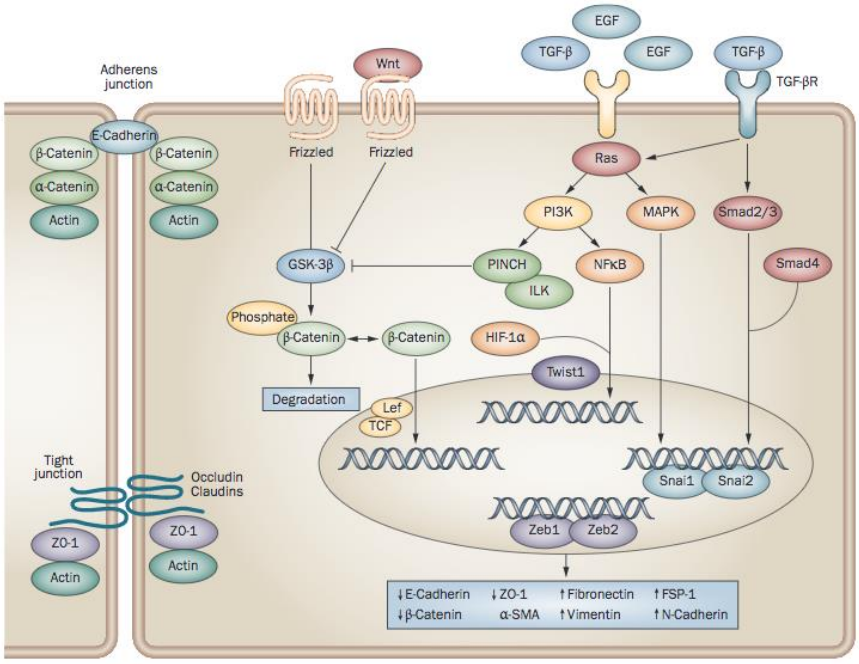


Figure I.21. Signaling pathways involved in EMT. Simplified representation of some EMT induced signaling pathways. The cooperation between pathways is also represented, in particular, between TGFβ and EGF signals as well as between EGF and Wnt pathways. TGFβ, EGF and Wnt bind to their receptors in the cell surface to induce specific signaling inside the cell leading to activation of Smad proteins, small GTPases of the Ras family, among others. Finally, translocation inside the nucleus allows both activation and repression of specific transcription factors involved in EMT, such as Snail, Fibronectin and E-cadherin[19].

4.3. EMT transcriptional regulation

Repression of epithelial markers and activation of mesenchymal genes is mediated by master regulators, including Snail, zinc-finger E-box-binding (Zeb) and the Twist transcription factors. These master regulators are activated early in the EMT process, normally by some of the signaling pathways mentioned above. Their expression is often controlled by each other, and they usually cooperate to regulate the expression of target genes. Accordingly, additional transcription factors are needed to activate transcription programs necessary to drive EMT progression[171].

Initiation of EMT is often mediated by downregulation of epithelial markers responsible for cell-cell adhesion, and in particular by E-cadherin, which encoded by *CDH1* gene. The Snail, Twist and Zeb family members are able to directly inhibit the expression of *CDH1* through the binding to E-box regions located in its promoter. Nonetheless, other regulators of EMT, such as the Twist family members, downregulate E-cadherin expression through indirect mechanisms[172-174].

The Snail family of transcription factors has been extensively studied and found to play a critical role in regulating EMT[175]. Three Snail family proteins have been described; Snail1, Snail2 (also known as Slug) and the recently characterized Snail3. They share structural similarities, including a C-terminal domain with zinc fingers that mediates binding to DNA E-box elements (of 5′-CACCTG-3′ or 5′-CAGGTG-3′) located in the promoters of its target genes[176]. The Snail proteins also contain an N-terminal SNAG (Snail/Gfi) domain. Both of these domains are responsible

for the Snail protein family activities[177]. Snail1, the most well-characterized family member, is induced in response to TGFβ as well as by other growth factors, such as epidermal and fibroblast growth factors (EGF and FGF, respectively)[175]. Upon induction of its expression, Snail1 binds to the E-box elements of target gene promoters (for instance, at the E-cadherin promoter) through its C-terminal domain, and interacts with histone modifying enzymes through its N-terminal SNAG domain. In order to repress gene expression of its target genes, Snail1 recruits the co-repressor Sin3a together with histone deacetylases (HDAC1/2) and induces deacetylation of histone H3 and H4[178]. Moreover, Snail1 can recruit Polycomb repressive complex 2 (PRC2) and lysine-specific histone demethylase 1 (LSD1), which trimethylate H3K27 and remove di- and tri-methyl marks of H3K4, respectively[179,180].

In addition to the well-established EMT transcription factors, other proteins have recently been shown to induce or at least regulate EMT. For instance, previous studies from our group revealed that a new histone-modifying enzyme, LOXL2, is recruited by Snail1 to the *CDH1* promoter allowing deamination of H3K4me3 and thus leading to gene repression[181]. Deamination of H3K4me3 by LOXL2 causes the deposition of an aldehyde group in the histone H3 tail, giving rise to oxidized H3 (oxH3). Further studies improved the mechanistic understanding of the LOXL2 function in EMT. We now know that LOXL2 is crucial for regulating heterochromatin transcription during EMT. Indeed, at the onset of EMT, and coincident with Snail1 upregulation, LOXL2 deaminates H3K4me3 on pericentromeric regions, leading to downregulation of major satellites transcripts. This downregulation induces HP1α release from pericentromeric regions and allows chromatin reorganization,

a crucial step to acquire the final mesenchymal phenotype. Once the mesenchymal phenotype is well established, HP1α localizes again in pericentromeric regions, and major satellite transcription is restored[6] (Figure I.22).

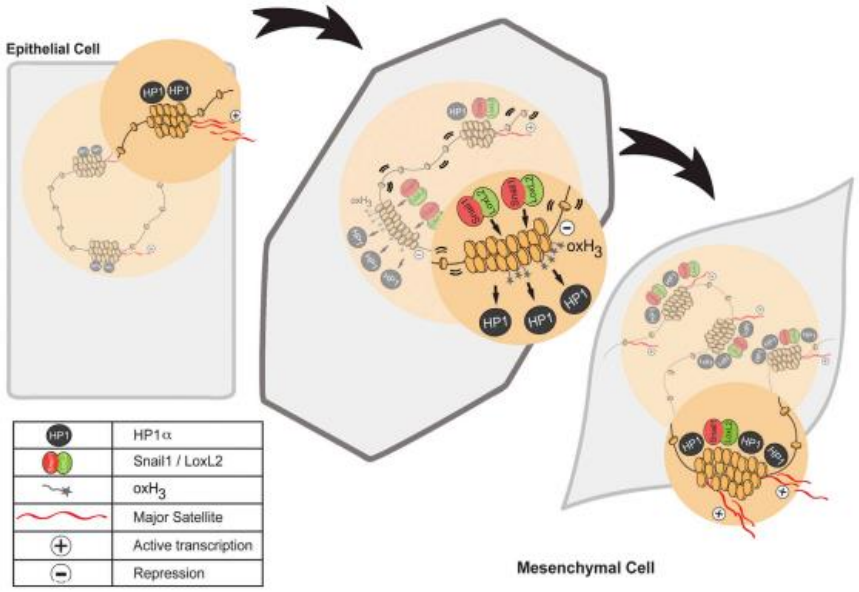


Figure I.22. LOXL2 involvement in heterochromatin reorganization during EMT. Upon TGFβ induction of EMT in normal murine mammary epithelial cells (NMuMG), Snail1 is up-regulated, binds to pericentromeric regions and recruits LOXL2. This histone-modifying enzyme then, oxidizes histone H3 leading to major satellite transcription repression and HP1α release from pericentromeric regions. After 24h of TGFβ induction, once the cells are completely mesenchymal, major satellite levels and HP1α binding to heterochromatin regions are recovered[6].

OBJECTIVES

Prior to this work, chromatin reorganization was thought likely to be crucial for the proper development of EMT[6]. However, specifics about which genomic regions reorganize during this process, and how this reorganization occurs, were still largely unknown.

The main objective of this thesis was therefore to determine whether chromatin reorganization plays a key role during EMT, and to determine which genomic regions are rearranged during this cellular transformation. As genomic organization relies mainly in its interactions with the nuclear lamina, we focused on:

1. Finding genomic regions that contact lamin B1 during EMT;

2. Characterizing TADs as well as changes in active/inactive (A/B) compartments during EMT by Hi-C techniques; and

3. Assessing the biological relevance of chromatin reorganization during EMT.

RESULTS

1. Heterochromatin changes during EMT

Previous data obtained in our laboratory showed that, during EMT, changes occur in heterochromatin that are essential for the transformation process[6]. We thus first studied pericentromeric heterochromatin, which is conserved during evolution and clusters into chromocenters in mouse cells[54], which can easily be visualized by DAPI staining.

In order to check whether heterochromatin structure changes take place during EMT, we first studied chromocenter organization at the onset of the EMT process. For this, we used the well-established model of mouse mammary epithelial NMuMG cells, which undergo EMT upon TGFβ treatment[182]. NMuMG cells were either untreated, or treated with TGFβ for 8 hours (giving rise to an intermediate state) or 24 hours (leading to the acquisition of a mesenchymal phenotype). DAPI staining revealed a clear increase in the number of heterochromatin foci per nucleus at 24 hours upon TGFβ treatment (Figure R.1A and R.1B), suggesting that the chromocenter reorganized during EMT.

Constitutive heterochromatin has a particular epigenetic signature characterized by specific histone modifications, such as H3K9me3[62]. For that reason, we decided to further study the heterochromatin state during EMT by immunostaining using H3K9me3 antibody. As expected, we observed a significant increase of H3K9me3 deposition 24 hours after EMT induction (Figure R.1C and R.1D).

We next used super-resolution 3D structured illumination microscopy (SIM) to corroborate the previous results (Figure R.1E). Images were depicted as a merge of DAPI and H3K9me3 signals showing that after 24 hours of TGFβ induction, not only the number of foci per nucleus increased but also the association of H3K9me3 at chromocenters.

In sum, these observations suggested that heterochromatin is reorganized during the EMT process.

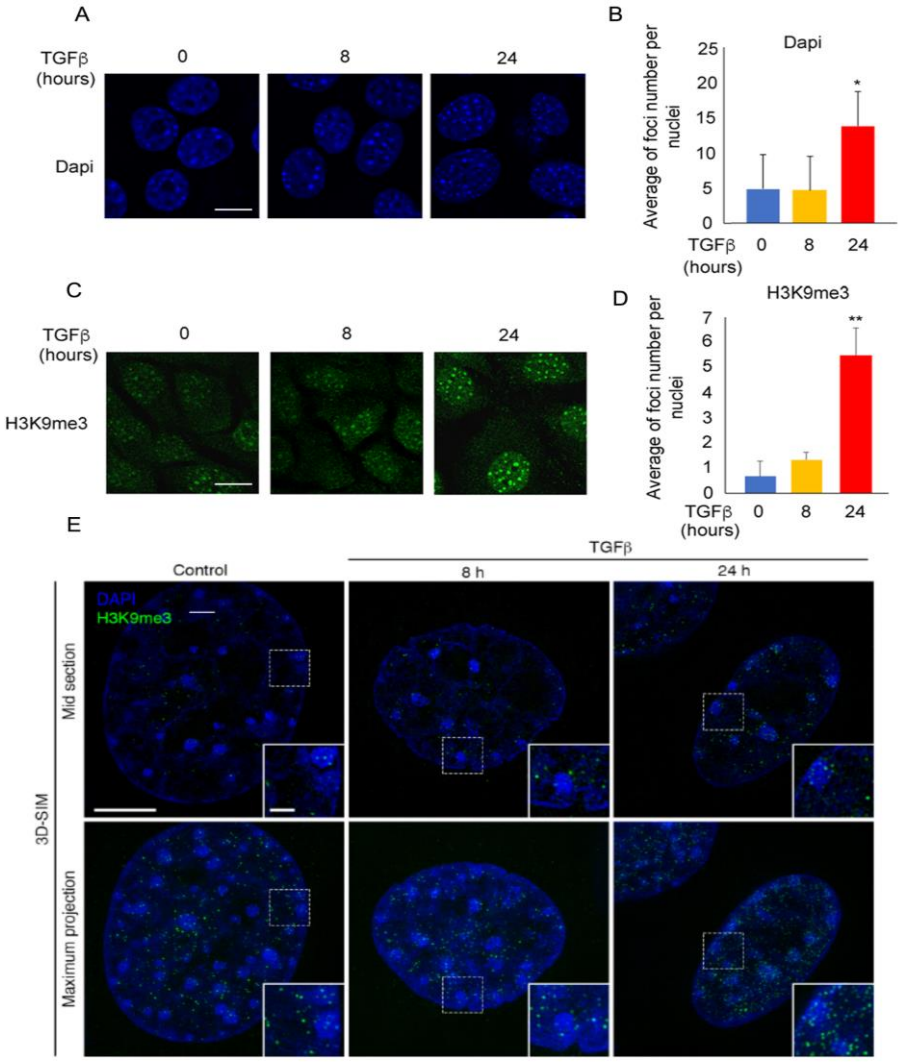


Figure R.1. Heterochromatin regions are reorganized during EMT. **(A)** Distribution of DAPI staining before and after TGFβ-induced EMT. Scale bar, 15 μm. **(B)** The graph depicts the average of DAPI foci number per nucleus in the same conditions. **(C)** Immunostaining of H3K9me3 histone mark during EMT. Scale bar, 10 μm. **(D)** Average of H3K9me3 foci number per nucleus. Error bars indicate SD of at least three different experiments; one asterisk indicates $p < 0.05$ and two asterisk indicates $p < 0.01$. **(E)** Super-resolution 3D-SIM images showing DAPI and H3K9me3 staining merged. Row 1 displays single optical mid-section, row 2 displays a maximum projection over a central 3 μm z-range. Scale bar, 5 μm (inset, 1 μm).

Constitutive heterochromatin formation ensures a condensed and transcriptionally inactive chromatin conformation[62]. As we had already described changes in heterochromatin during the EMT (such as increased numbers of DAPI and H3K9me3 foci per nucleus upon TGFβ treatment), we used the uridine analog of 5-ethynyl uridine (EU) to detect *de novo* RNA synthesis in NMuMG cells during EMT transition. This method is based on treating the cells for 24 hours to allow the biosynthetic incorporation of EU into newly transcribed RNAs. Labelled RNA is then detected by reaction with fluorescent azide (Alexa488-azide) through the presence of cooper (I)-catalyzed alkyne-azide cycloaddition (also known as the "Click" reaction)[183]. FACs analysis revealed a slight decrease in RNA transcription upon TGFβ treatment (Figure R.2), as shown by a decrease in FITC staining. These results are in agreement with previously observed increase in heterochromatin foci during EMT transition. Thus, the higher the number of heterochromatin foci per nucleus, the lower the transcription rates.

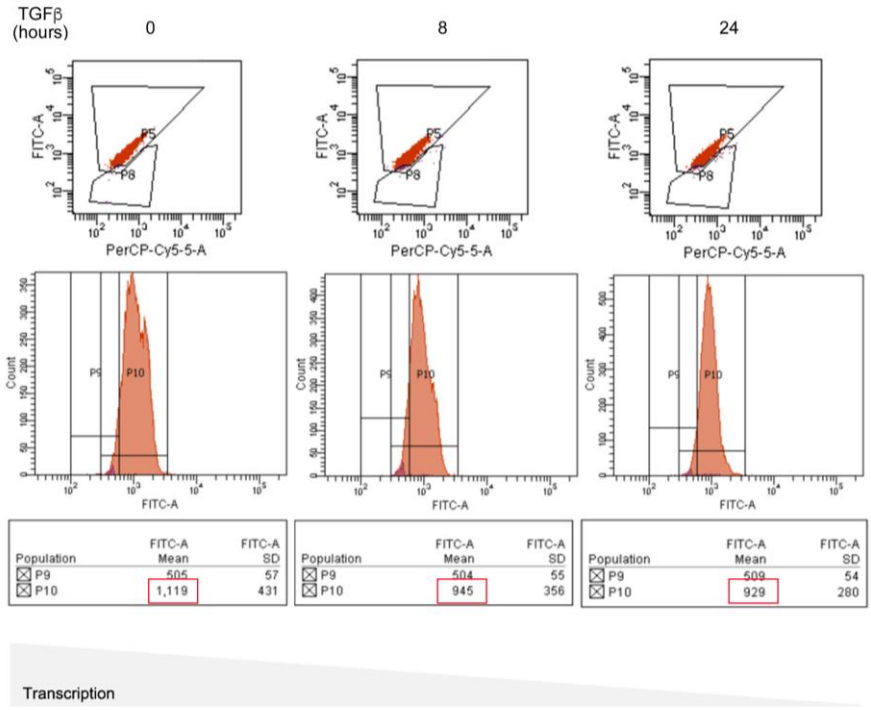


Figure R.2. Global transcription rates. FACs analysis of EU incorporation to newly formed RNAs. P5 (upper panel) and P10 (middle panel) represent the population of cells that has been labelled with Alexa488-azide (FITC staining). The lower panel shows the statistical analysis for each region in the figures.

2. Lamin B1 associates with euchromatin regions

Our results, together with previous studies performed in our laboratory, showed that heterochromatin is reorganized during EMT. Changes in heterochromatin might lead to dynamic changes in euchromatin, where most of the genes are located. Therefore, we wondered whether euchromatin could also be reorganized during this cellular transformation.

Some of the main players in genome organization are the nuclear lamins: lamin A/C and B. As both type of lamins (lamin A/C and B) form an interconnected network, and as recent studies have shown that lamin A/C contacts euchromatin regions to form a more dynamic pool highly involved in gene expression[184], we wondered whether lamin B1 could also associate with active regions of the chromatin located in the nuclear interior; and if so, whether it forms more dynamic structures than LADs that are also important for the genome architecture.

Confocal immunofluorescence of lamin B1 confirmed the presence of this intermediate filament not only in the nuclear envelope but also in the nuclear interior during EMT transformation (Figure R.3A). Moreover, colocalization of lamin B1 with emerin, an integral protein of the NL, was only observed in the nuclear envelope as shown by confocal microscopy as well as by quantification of antibodies signals in both fractions (nuclear envelope and nuclear interior) (Figure R.3B).

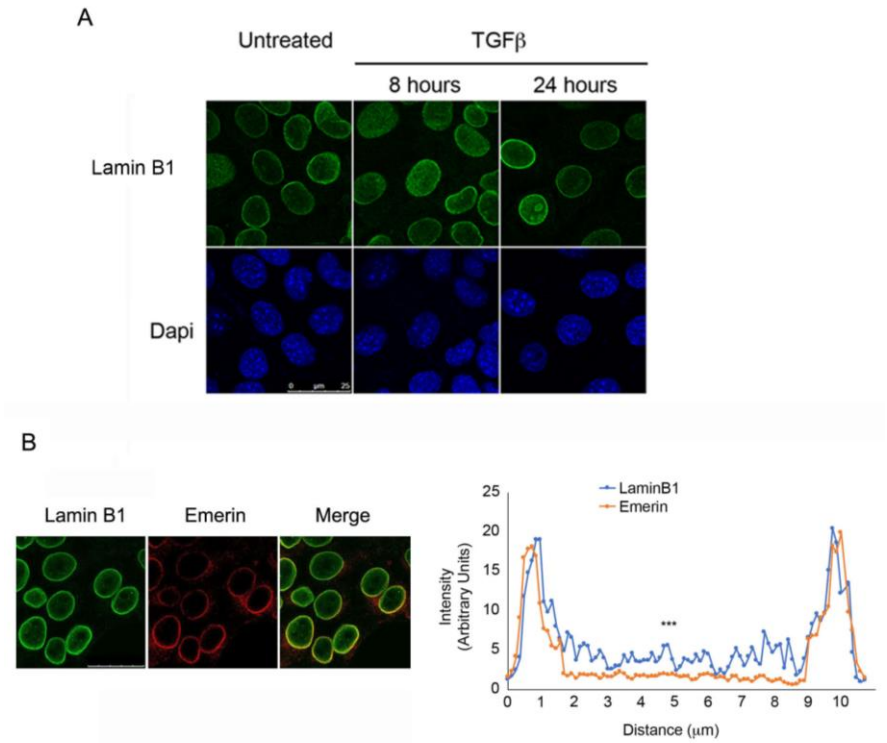


Figure R.3. Lamin B1 localizes in the nuclear interior. (A) Representative image of NMuMG cells at different time points of TGFβ induction stained with DAPI and immunolabelled with anti-lamin B1. Scale bar, 25 μm. **(B)** NMuMG cells in untreated conditions were immunolabelled with lamin B1 and emerin antibodies. The colocalization of both signals is shown in the third column as a merge. Scale bar, 25 μm (upper panel). A graphic representation of the antibodies signal intensities among the nuclei is shown. Regions of higher intensity belong to the nuclear envelope, and regions of lesser intensity, to the nucleoplasm (lower panel). All images were obtained by confocal microscopy. ***$p < 0.001$.

In order to detect the putative genomic regions in close contact with lamin B1 present in the nuclear interior, we decided to perform chromatin immunoprecipitation assay followed by deep sequencing (ChIP-seq). As the specificity of the antibody used is crucial, a ChIP followed by Western blot of NMuMG cells infected with

lentivirus carrying either an irrelevant short hairpin RNA (shRNA) as a control (C) or one specific for lamin B1 (knockdown, KD) was performed. The specificity of the antibody was confirmed both by the presence of lamin B1 only in the immunoprecipitated samples, and by the strongly decreased signal in KD conditions (Figure R.4A). Importantly, we also detected a region in contact with previously published canonical LADs (cLADs) by using ChIP-qPCR approach with the lamin B1 antibody as compared with an irrelevant IgG (Figure R.4B).

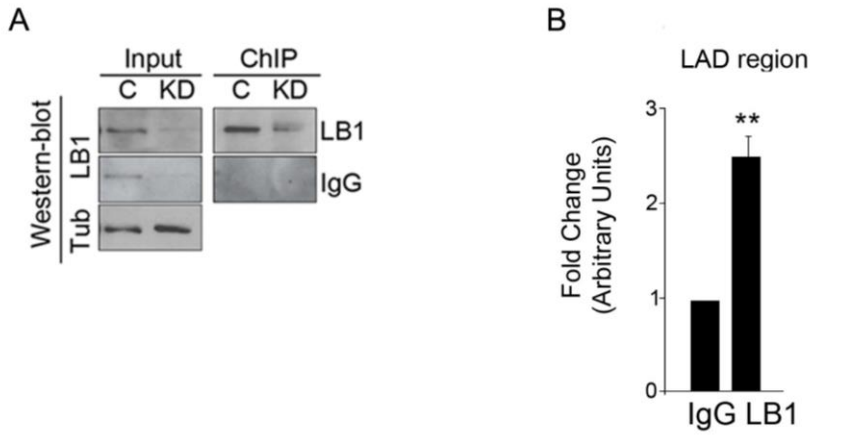


Figure R.4. Validation of the specificity of lamin B1 antibody. (A) Western blot of chromatin immunoprecipitated with anti-lamin B1 (LB1) or irrelevant IgG (as control) from control or knockdown NMuMG cells. Tubulin is shown as a loading control. **(B)** ChIP-qPCR of a selected canonical LAD region. Data are expressed as the fold-change relative to the data obtained from IgG. Error bars indicate the SD of at least three independent experiments. $^{**}p < 0.01$.

Heterochromatin is mainly formed in gene-poor regions located in the nuclear periphery, whereas euchromatin is mainly present in the nuclear interior forming decondensed structures allowing gene transcription. Therefore, ChIP-seq of lamin B1 was performed in an

euchromatin-enriched fraction; due to its nuclear localization, the genomic fraction highly likely to be in contact with the nuclear interior lamin B1.

As chromatin preparation prior to immunoprecipitation is crucial for identifying genomic-associated regions[126], we sheared the chromatin by low sonication conditions to obtain DNA fragments of 300 to 600 bp. Euchromatin is less condensed and thus easily sheared by sonication, while heterochromatin is highly condensed and requires more rounds of shearing to become completely fragmented[184]. Importantly, fragments higher than 1 Kb were excluded from Illumina sequencing, which mainly corresponded to heterochromatin (Figure R.5A), while shorter fragments, which primarily contain euchromatin, were successfully sequenced.

Interestingly, after deep sequencing of low sonicated chromatin, we detected significant chromatin lamin B1 positive sites (LB1+) in NMuMG cells that increased during the EMT process (Figure R.5B). Specifically, a total of 4645 LB1+ peaks did not overlap with cLAD peaks in untreated cells, 10,484 peaks in cells with intermediate transition (i.e., at 8 hours of TGFβ treatment) and 7,083 peaks in fully-transitioned cells (at 24 hours of TGFβ treatment) (Table A.1 and Figure R.5B). These results indicated we detected a chromatin fraction that contacts lamin B1 that is distinct from the “traditional” heterochromatin regions located in the nuclear periphery; these “novel” peaks may correspond to nuclear interior euchromatin.

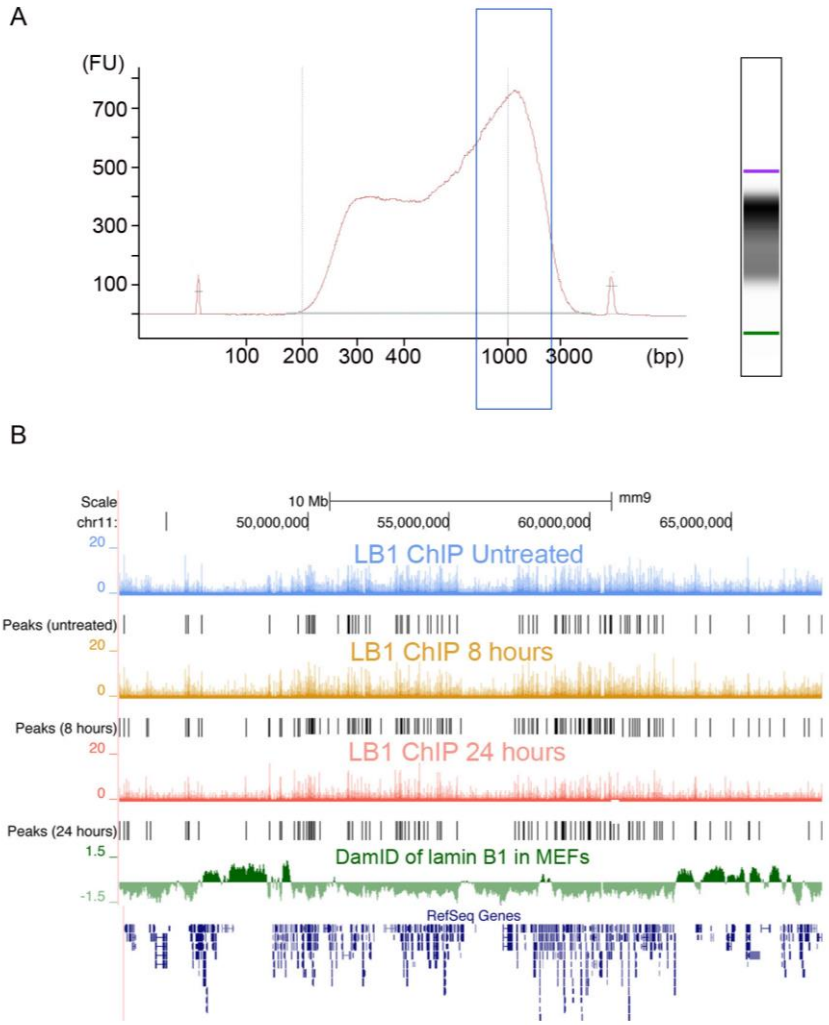


Figure R.5. Lamin B1 positive sites represents the specular image of canonical LADs due to heterochromatin exclusion from the sequencing. (A) Representative bioanalyzer intensity profile of LB1 ChIP sample. Fragments excluded from the sequencing are within the blue rectangle. **(B)** UCSC Genome Browser overview of one region across chromosome 11 (mm9) of the lamin B1 ChIP-seq profiles in NMuMG cells that were untreated (blue) or treated with TGF-β for 8 hours (orange) or 24 hours (red). Lamin B1+ sites identified at each condition (black), previously published lamina-associated domains (LADs) from mouse embryonic fibroblast cells (green), and the RefSeq gene track from UCSC Genome Browser (dark blue) are shown.

Genome distribution of ChIP-seq peaks showed that LB1+ sites were mainly located around the transcription start site (TSS) of specific genes (Figure R.6A). Specifically, 4454 LB1+ genes did not overlap with cLAD peaks in untreated conditions, 8,755 genes in cells with intermediate transition (i.e., after 8 hours of TGFβ treatment) and 6,595 genes in fully-transitioned cells (at 24 hours of TGFβ treatment) (Table A.1). Moreover, in contrast to A/T rich cLADs, these newly identified sites were highly enriched in G/C sequences at each time point (Figure R.6B).

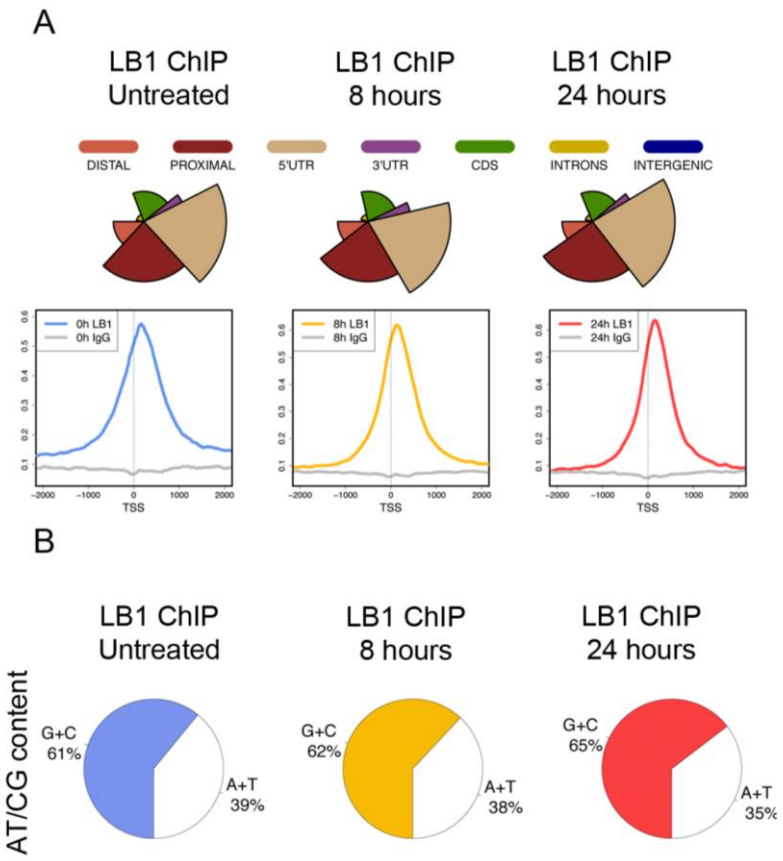


Figure R.6. Characterization of LB1+ sites. (A) Genome distribution of LB1 ChIP-seq peaks in untreated cells or those treated for 8 or 24 hours. Distal region is that within 2.5 Kb and 0.5 Kb upstream of a gene´s TSS, and the proximal region, within 0.5 Kb of a gene´s TSS. The 5′ UTR and 3′ UTR sequences are indicated. CDS, protein coding sequence; intergenic, all genome regions besides introns (top). The average distribution is shown of LB1 ChIP-seq reads, and the corresponding IgG control experiments, at 2 Kb around the TSS of LB1+ genes from untreated cells or those treated for 8 or 24 hours (bottom). **(B)** A/T and C/G content of the sequences of LB1+ sites.

Further characterization of LB1+ genes was performed to determine whether these genes were associated with canonical euchromatin or were instead decorated with heterochromatin histone marks. Bioinformatic Enrich analysis revealed how histone marks associated to these genes were characteristic for euchromatin (Figure R.7A). Interestingly, at each time point, genes that were found not to be enriched in lamin B1, which mainly corresponding to cLADs, were clearly associated with H3K9me2/3, the classical heterochromatin histone mark (Figure R.7B).

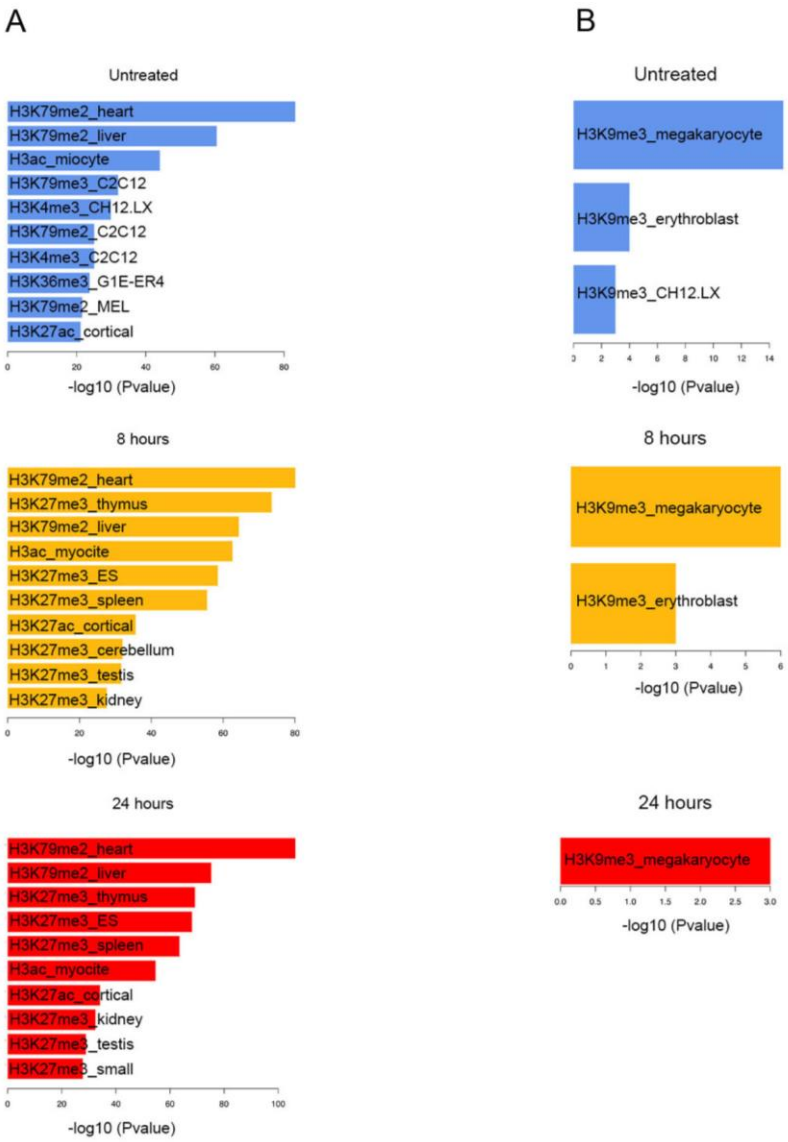


Figure R.7. LB1+ genes, in contrast to cLADs, are enriched in euchromatin histone marks. (A) Functional analysis of LB1+ genes showing the statistically significant association with histone marks in NMuMG cells that were untreated (blue) or treated with TGF-β for 8 hours (orange) or 24 hours (red). **(B)** The same analysis as in **(A)** was applied to genes not associated with LB1+ sites under the same conditions: untreated (blue), or treated with TGF-β for 8 hours (orange) or 24 hours (red).

To assess the genome transcription profile of NMuMG cells during the EMT, we analyzed these samples by RNA-seq. Lamin B1-enriched genes were actively transcribed at each time point compared with the rest of the genes in the genome (Figure R.8A).

In addition to studying the transcription profile of these genes, we also carried out ATAC sequencing (ATAC-seq) experiments, which allowed us to determine chromatin accessibility genome-wide. This technique is based on using a hyperactive Tn5 transposase, which is able to cut and insert adaptors into accessible regions of the genome at the same time. These adaptors are subsequently used to sequence the regions of the genome with more chromatin accessibility, thereby leading to sequencing of open (rather than closed) chromatin regions[185]. As for the RNA-seq data, we crossed ATAC-seq data with LB1+ genes; we observed that promoter regions of these genes were located in accessible chromatin regions as compared to the rest of the genes in the genome (Figure R.8B).

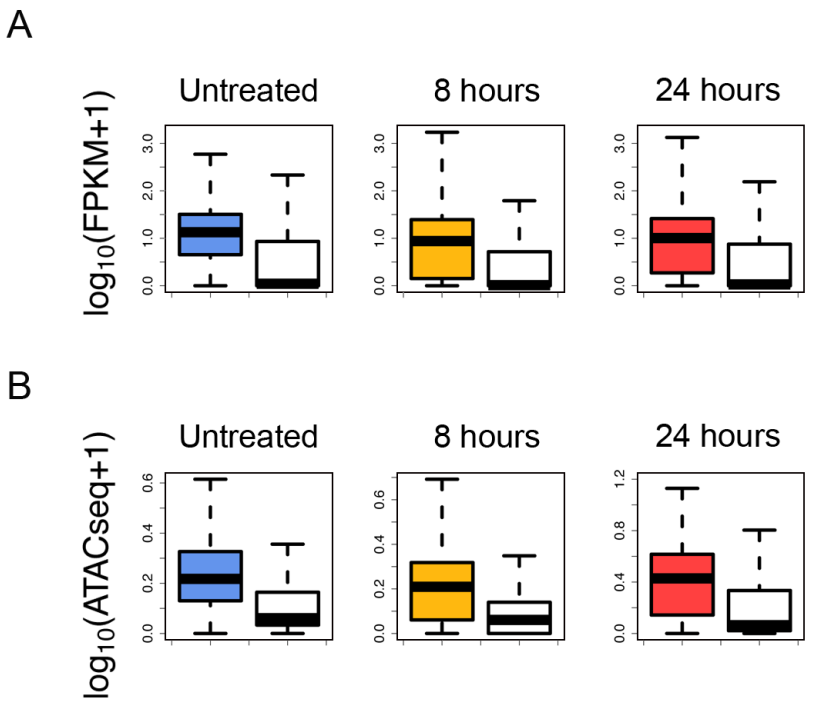


Figure R.8. Genes enriched in lamin B1 are actively transcribed and located in accessible chromatin regions. (A) Expression of LB1+ genes computed in FPKM (fragment per kilobase of transcript per million mapped reads) in untreated cells (blue) or those treated with TGF-β for 8 h (orange) or 24 h (red), as compared to the rest of genes in the genome (white) **(B)** Promoter ATAC-seq enrichment of LB1+ genes measured in number of normalized reads in untreated cells (blue) or those treated with TGF-β for 8 h (orange) or 24 h (red), as compared to the rest of genes in the genome (white). In the boxes, the bottom and top fractions represent the first and third quartiles, and the line, the median. Whiskers denote the interval between 1.5 times the interquartile range (IQR) and the median.

Importantly, we observed a correlation between lamin B1 binding and gene expression. The full set of mouse genes were stratified into four groups at each time point based on its transcription rates by RNA-seq, of no expression (silent) or low, medium or high expression. Lamin B1 levels obtained by ChIP-seq were then

plotted for each gene set; this showed a strong correlation between expression and lamin B1 binding around the TSS. Thus, the higher the levels of expression, the higher the lamin B1 enrichment (Figure R.9).

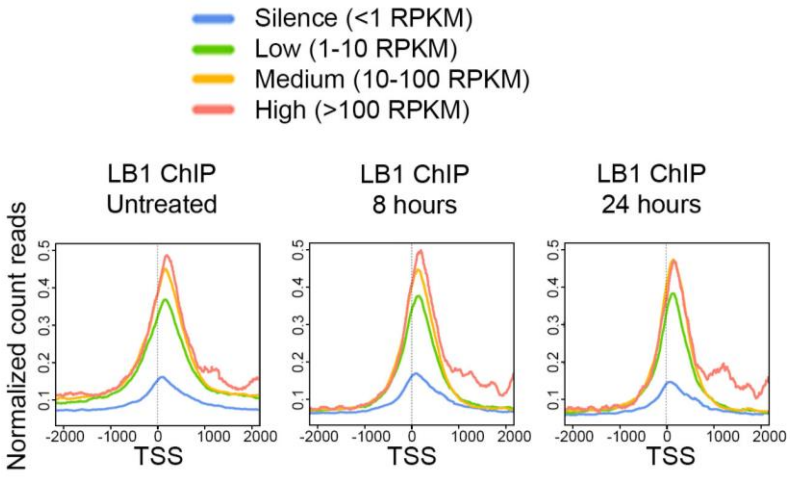


Figure R.9. Lamin B1 binding correlates with gene expression during EMT. Average distribution of lamin B1 ChIP-seq reads at 2 Kb around the TSS of mouse genes that were classified into four categories (silent, low, medium or high) according to their expression levels in untreated cells or cells treated for 8 or 24 hours with TGFβ.

Taken together, these results demonstrate that lamin B1 not only contacts heterochromatin regions, as widely described, but also is present in the euchromatin fraction. Indeed, lamin B1 can be found in expressed euchromatin regions associated with G/C region that are gene-rich, accessible and decorated with euchromatin histone marks. Moreover, lamin B1 binding highly correlates with gene expression during EMT.

3. Euchromatin lamin B1 regions are dynamic during EMT

As previously mentioned, cLADs are conserved structures that do not change during differentiation or between species[130]. However, as the euchromatin fraction *per se* is more dynamic than heterochromatin, we reasoned that euchromatin lamin B1 regions could change their structures during certain processes, such as EMT.

Interestingly, comparison of different LB1+ sites during EMT revealed that, of the total 10,350 target genes identified in the three time points, only 2,933 genes (28%) maintained lamin B1 throughout the EMT process, while a vast majority (7417; 72%) changed at the onset of EMT (Figure R.10A). Gene ontology (GO) enrichment analysis of LB1+ genes at each time point showed that, in untreated conditions, LB1+ genes belonged to chromatin modification and gene transcription categories. However, once the EMT transition started (i.e., at 8 and 24 hours after treatment with TGFβ), genes enriched in lamin B1 were related to the EMT process, with functions belonging to embryonic morphogenesis and pattern specification (Figure R.10B).

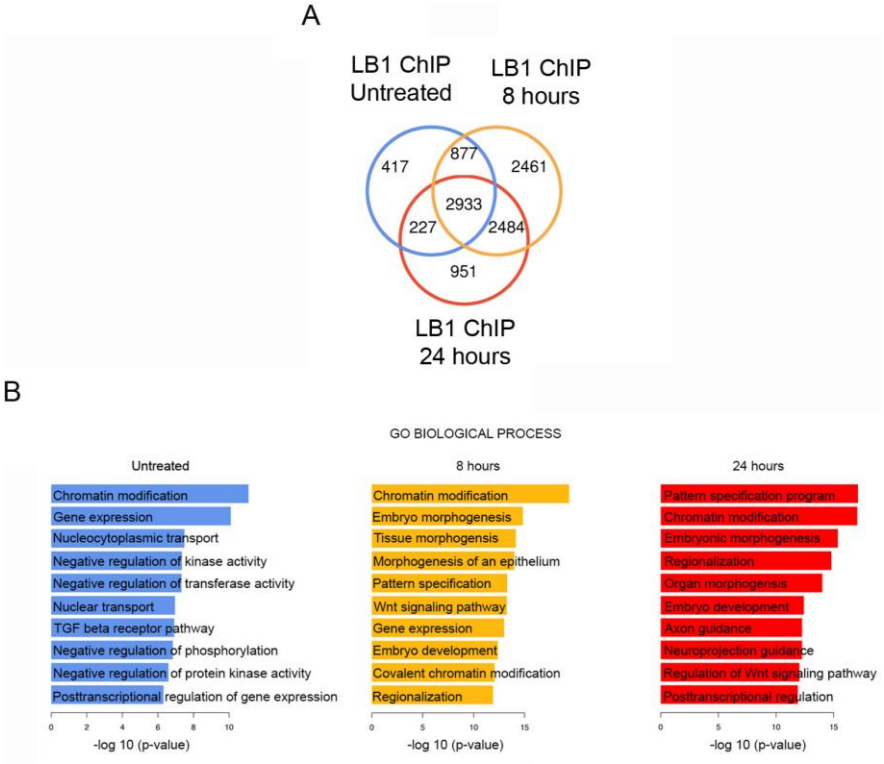


Figure R.10. LB1+ genes dynamically change during EMT. (A) Venn diagram showing overlap between LB1+ genes in untreated cells (blue) or those treated with TGF-β for 8 hours (orange) or 24 hours (red). **(B)** Functional analysis of statistically significant biological processes of lamin B1+ genes in NMuMG cells untreated (blue) or treated with TGF-β for 8 hours (orange) or 24 hours (red).

Further analysis revealed that not only LB1+ genes in untreated conditions but also those genes that maintain their level of lamin B1 during the entire EMT process (2933; 28%), also belonged to chromatin modification and general gene transcription, characteristics that are common in a wide variety of cellular processes and therefore not specific of the EMT pathway. However, the total set of genes that emerged as new lamin B1

targets upon EMT induction by TGFβ treatment (i.e., genes only present at 8 hours and/or 24 hours after TGFβ treatment but not in untreated conditions), which corresponded to 57% of the total (5,896 *de novo* LB1+ genes), encoded proteins with functions related to the EMT process (Figure R11.A).

Considering the amount of newly formed LB1+ genes upon induction (5896; 57%), we further investigated the enrichment of transcription factor motifs at these new LB1+ sites. As seen by the E value, the most highly enriched factors were involved in developmental and differentiation process, including some members of the TGFβ signaling pathway[186,187] (Figure R.11B and Figure R.11C). These data suggest that TGFβ-activated transcription factors might play a role in recruiting lamin B1 to these specific sites.

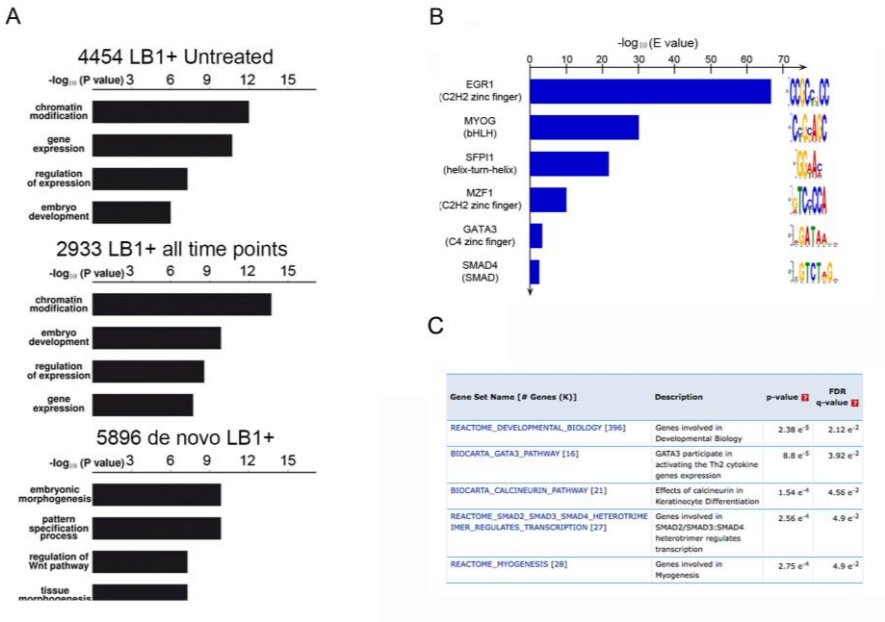


Figure R.11. TGFβ-activated transcription factors might recruit lamin B1 to EMT-related newly formed LB1+ sites. (A) Gene ontology (GO) terms of lamin B1+ genes only present in untreated conditions, present in all three time points, or newly formed at the onset of EMT are depicted. **(B)** Statistically significant transcription factor motif analysis of 5,896 *de novo* lamin B1+ genes. **(C)** Signature of pathways associated to the list of previously identified transcription factors. Only statistically significant signature pathways are shown.

Finally, we assessed whether LB1+ genes changed their expression during cellular transformation. We found that 42% of LB1+ genes changed their expression at the onset of EMT in untreated conditions (1,872 genes of total 4454 LB1+ genes), 37% at 8 hours of TGFβ treatment (3269 genes of total 8755 LB1+ genes) and 38% (2454 genes of total 6595 LB1+ genes at 24 hours of treatment) at 24 hours of TGFβ treatment (Figure R.12).

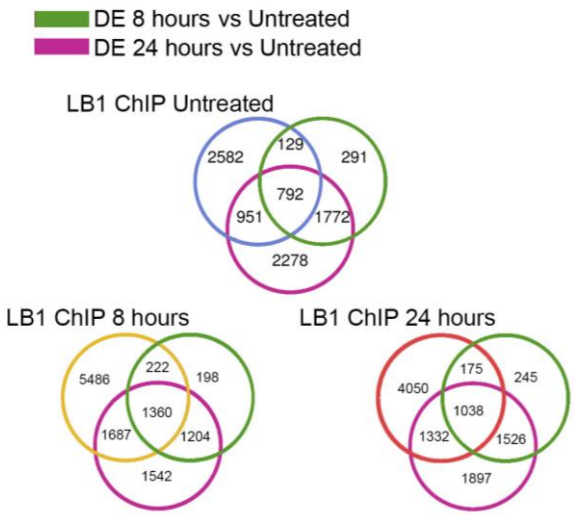


Figure R.12. LB1+ genes change their expression at the onset of the EMT. Venn diagram showing the overlap between LB1+ genes in untreated conditions (blue) or after treatment with TGF-β for 8 hours (orange) or 24 hours (red) and genes differentially expressed (DE) at 8 hours (green) and 24 hours (red) after TGF-β treatment.

In sum, these data reveal that lamin B1 found in expressed euchromatin changes dynamically during the EMT process, and indicate that expression of LB1+ genes is also dynamic at the onset of this cellular transformation. Moreover, the data suggest that TGFβ-activated transcription factors might be responsible for lamin B1 deposition at EMT-related genes.

4. eLAD characterization

Similar to cLADs, which were originally characterized as regions enriched in nuclear lamin proteins due to comparisons in the corresponding control experiments[128,184], we defined new euchromatin lamina B1-associated domains (eLADs), as clusters of neighboring LB1+ sites that occur within a delimited region in the genome (of less than 250 Kb).

Importantly, we were able to characterize eLADs in all three conditions and determined that these not overlap with cLADs (Figure R.13A). We identified 2051 eLADs in untreated cells, 2429 eLADs in 8 hour-treated cells and 2949 eLADs in 24 hour-treated cells (Table A.2). These newly-associated domains differ in structure from the ones previously known, as the average size of eLADs was 0.34 Mb, which is smaller than that of cLADs (about 1 Mb[128]). Moreover, the coverage of the mouse genome for each set of eLADs was 25.9% in untreated cells, 31.7% in cells treated with TGFβ for 8 hours and 40.1% in cells treated with TGFβ for 24 hours (Table A.2). This coverage at the onset of the EMT (at the 8-hour timepoint) is also lower than that of cLADs, which covers about 40% of the genome.

Similar to LB1+ sites, these newly characterized domains are active and located in accessible chromatin regions during all the EMT process, as shown by the analysis of RNA-seq and ATAC-seq data in comparison with the other regions of the genome (Figure R.13B).

Previous studies reported that lamin A/C can also contact euchromatin regions, as detected by different levels of sonication[184]. To further confirm the presence of eLADs within euchromatin regions, we analyzed the degree of overlap between lamin B1 and lamin A/C ChIP-seq data. As expected, eLADs showed a high overlap with lamin A/C euchromatin regions, which were detected under low sonication conditions and with a much lower overlap with heterochromatin regions enriched in lamin A/C, which were identified after high chromatin sonication (Figure R.13C).

In sum, we have identified new lamin B1-associated domains (eLADs) that contact euchromatin, which are accessible, active and change dynamically during the EMT process.

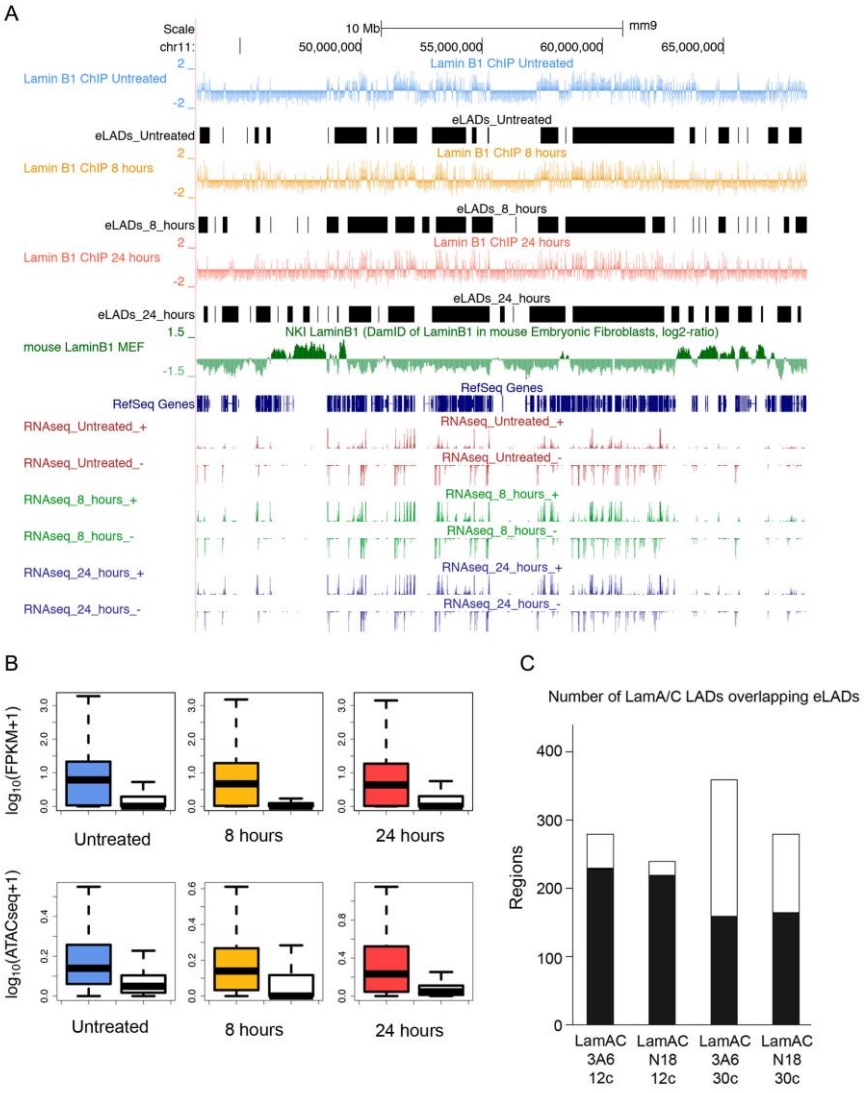

A
Scale
chr11:
10 Mb
mm9
50,000,000
55,000,000
60,000,000
65,000,000
Lamin B1 ChIP Untreated
eLADs_Untreated
Lamin B1 ChIP 8 hours
eLADs_8_hours
Lamin B1 ChIP 24 hours
eLADs_24_hours
NKI LaminB1 (DamID of LaminB1 in mouse Embryonic Fibroblasts, log2-ratio)
mouse LaminB1 MEF
RefSeq Genes
RNAseq_Untreated_+
RNAseq_Untreated_-
RNAseq_8_hours_+
RNAseq_8_hours_-
RNAseq_24_hours_+
RNAseq_24_hours_-
B
log10(FPKM+1)
log10(ATACseq+1)
Untreated
8 hours
24 hours
C
Number of LamA/C LADs overlapping eLADs
Regions
LamAC 3A6 12c
LamAC N18 12c
LamAC 3A6 30c
LamAC N18 30c

Figure R.13. Definition of eLADs. (A) UCSC Genome Browser overview of one region across the chromosome 11 (mm9) containing the following information (from top to bottom): lamin B1 ChIP-seq profiles subtracting the IgG control in untreated cells (blue) or in cells treated with TGF-β for 8 hours (orange) or 24 hours (red); eLADs identified in each condition (black); previously published LADs from mouse embryonic fibroblast cells (green); and RNA-seq strand-specific expression profiles at each time point (red for untreated cells, green for cells treated with TGF-β for 8 hours, and blue for cells treated for 24 hours). **(B)** Gene expression within eLADs (given in FPKM) in cells that were untreated (blue) or treated with TGF-β for 8 hours (orange) or 24 hours (red), as compared to the rest of genes (white). Promoter ATAC-seq enrichment of LB1+ genes (measured in number of normalized reads) in cells that were untreated (blue) or treated with TGF-β for 8 h (orange) or 24 h (red), as compared to the rest of genes (white). In the boxes, the bottom and top fractions represent the first and third quartiles, and the line, the median. Whiskers denote the interval between 1.5 times the interquartile range (IQR) and the median. **(C)** Bar plot showing the overlap between eLADs and previously published lamin A/C LADs obtained after low or high sonication conditions (12 cycles or 30 cycles, respectively) and with two different antibodies (3A6 and N18).

5. Role of Lamin B1 in 3D chromatin architecture

Intermediate filaments play a crucial role in cellular architecture[188]. In addition, cLADs were found to be involved in the spatial organization of the genome[127]. As genome function relies on their spatial architecture[189], we wondered whether lamin B1 that we found in contact with euchromatin regions could be important for acquiring the mesenchymal genome structure. To address this, we used genome-wide chromatin conformation capture (Hi-C), a method to study the 3D architecture of whole genomes[140], to analyze the EMT transformation. Genomes separate in two

different spatial compartments, termed A and B compartments, which have been described to correspond to active and inactive compartments, respectively. Regions within each compartment tend to be more closely in contact with each other rather than with regions of other compartments, suggesting a clear compartmentalization of the genome[140].

Deep sequencing of Hi-C libraries of NMuMG cells that were untreated or treated with TGFβ for 8 or 24 hours (to induce EMT) resulted in 89–106 million valid interactions per time point (Table A.3). These data allowed us to generate chromosome interaction maps with a resolution of 40 Kb and 100 Kb for TADs and genome segregation into A and B compartments, respectively. Comparison between different TGFβ treatment timepoints showed overall compartment changes. In the intermediate state (8 hours of TGFβ treatment), approximately 1200 bins (with domain sizes of 100 Kb) moved from A to B (1281 bins) or from B to A (1206 bins). Interestingly, in the mesenchymal state, these structural changes were reinforced, with 942 bins of EMT-related genes moving from A to B, and 2228 bins from B to A (Figure R.14A and Figure R.14B).

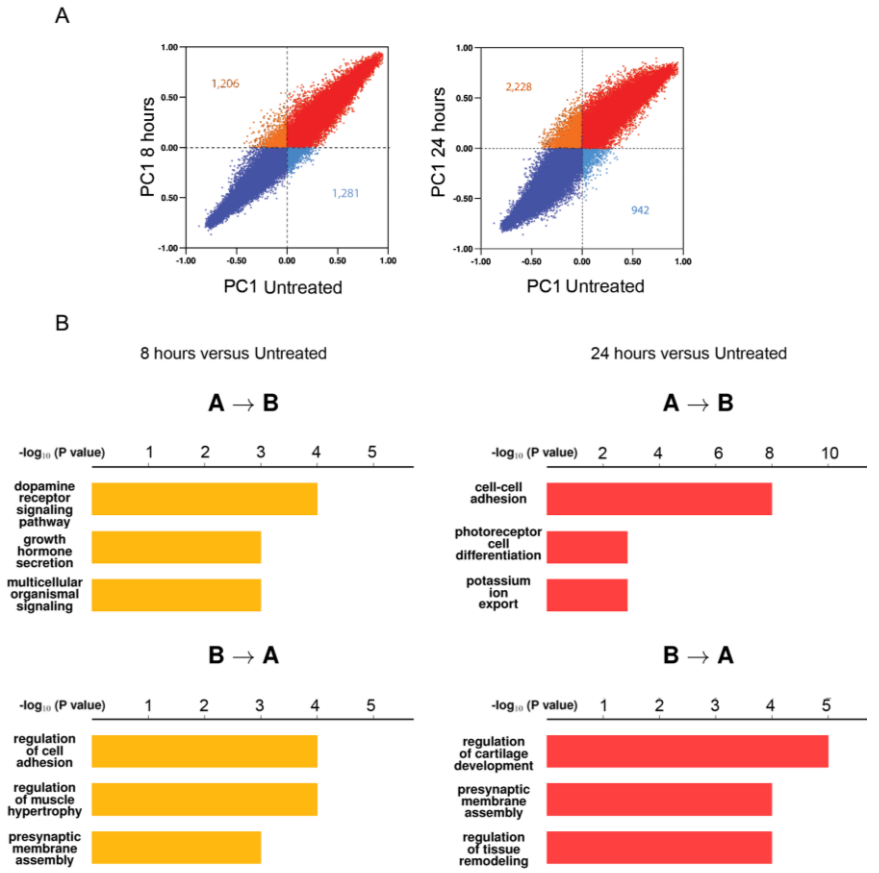


Figure R.14. Genome compartmentalization during EMT. (A) A/B compartment dynamics. Scatterplots of PC1 values indicate compartment changes between untreated cells and those treated with TGF-β for 8 or 24 hours. Dots indicate 100 Kb bins that changed compartment with respect to untreated cells (from B to A, shown in orange; from A to B, in cyan). Numbers indicate the total bins that changed compartment. **(B)** GO analysis of genes that changed from A to B compartment, or from B to A compartment, at 8 or 24 hours after TGF-β treatment, as compared to untreated NMuMG cells.

These results might indicate changes in genome architecture during EMT. Regions located between TADs, named borders, are enriched in active genes and contain binding sites for architectural

proteins, such as CTCF, which normally acts as insulators[146,149]. The strength of these borders is determined by the ratio between intra-TAD and inter-TAD interactions and is highly influenced by the presence of architectural proteins. Therefore, an increase in intra-TAD interactions leads to increased border strength[146].

Although 50% of TAD borders were conserved during EMT, the overall strength of these TADs increased at the onset of the EMT process (Figure R.15A), suggesting an increase in intra-TAD interactions that might be related with previously described changes in genome architecture. Moreover, enhanced border strength correlates with an increase in the percentage of lamin B1-containing borders (with *p* values of 10^{-28}, 10^{-68} and 10^{-32} in untreated, 8-hour treated and 24-hour treated cells, respectively) (Figure R.15B). This correlation suggests that lamin B1 could act as an architectural protein.

In some cases, changes in border strength can be attributed to changes in gene transcription[149]. As we found increased border strength following TGFβ treatment, we wanted to elucidate whether could be attributed to changes in the transcription profile. We thus checked the transcription rates during TGFβ treatment in TAD borders that had increased levels of lamin B1. Importantly, transcription was maintained in these regions during the EMT process (Figure R.15C).

In sum, the increased number of LB1+ sites that we previously identified (Table A.1) correlated with an increase in TAD border strength and with an enhanced percentage of lamin B1 occupancy at TAD borders. This indicated that lamin B1 might contribute to the

3D genome organization during EMT as an architectural protein and, interestingly, regardless of direct transcription regulation.

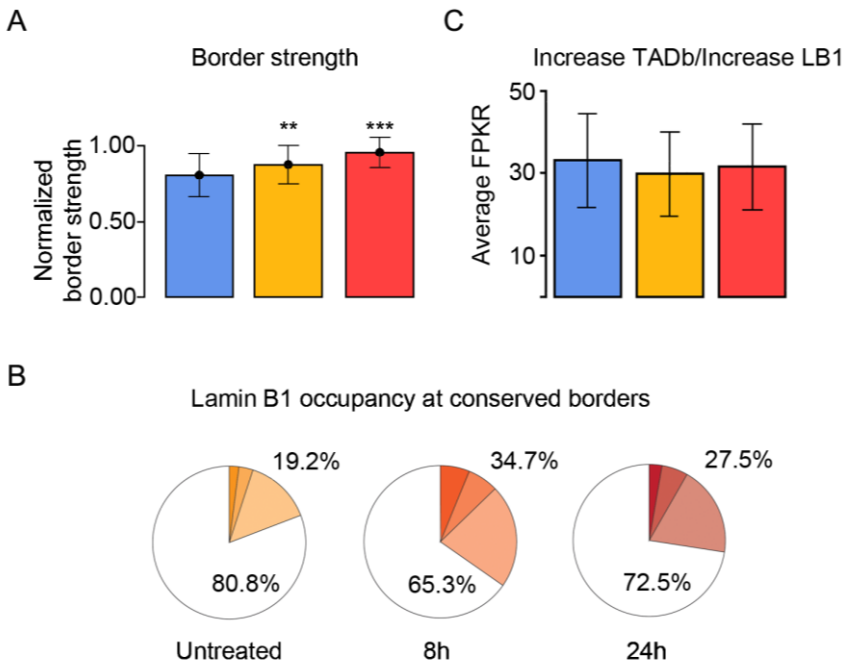


Figure R.15. Lamin B1 as an architectural protein. (A) TAD border dynamics. Bar plot of normalized border strength for conserved borders in untreated cells (blue) or those treated with TGF-β for 8 hours (orange) or 24 hours (red). **(B)** Border occupancy by lamin B1 in conserved TADs. Pie charts of 1, 2 or ≥3 called peak (dark to light color) of lamin B1 ChIP-seq in a conserved border. **(C)** Expression of genes within TAD borders in which lamin B1 increased (computed in FPKM) in untreated cells (blue) or cells treated with TGF-β for 8 hours (orange) or 24 hours (red).

Finally, we sought to confirm that lamin B1 is involved in changes in genome architecture. As expected, transcription, ATAC signals, LB1+ sites and eLADs were enriched in the A compartment, while cLADs were located in the B compartment (Figure R.16).

Interestingly, we observed that LB1+ sites and eLADs decreased in the A compartment during the EMT, with an increase in the B compartment (Figure R.16) that corresponded to newly formed eLADs rather than to changes between A and B compartments (Table A.4 and Table A.5), further suggesting that lamin B1 could contribute to the 3D reorganization of the genome during EMT

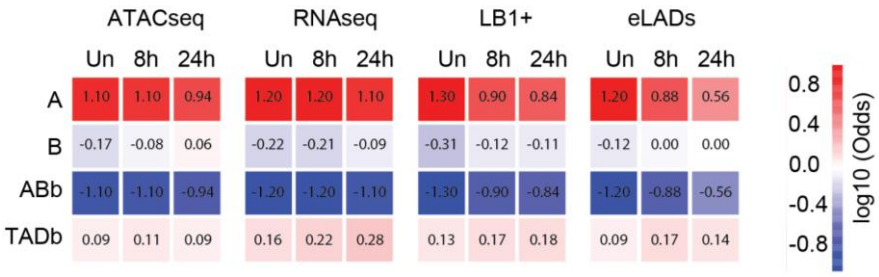


Figure R.16. Dynamics of A/B compartments as well as TAD borders during EMT in terms of transcription, accessibility and regions enriched in lamin B1. Contingency table analysis for enriched and depleted features according to distinct genome architectural landmarks. Red to blue colours indicate positive to negative log odds of the feature in genomic bins assigned to A, B or TAD borders. All log odds are statistically different than zero ($p < 0.05$, Fisher's exact *t*-test).

6. Downregulation of lamin B1 does not affect cellular viability

We have demonstrated so far that lamin B1 could be an architectural protein important for folding the 3D structure of the genome during the EMT process. However, we also wanted to determine its functional relevance. For that, we performed lamin B1

downregulation functional assays. Most of the previous studies induced loss of lamin B1 with either high efficient knockdown systems or by directly knocking-out the protein[190,191]. In order to avoid indirect effects due to the massive loss of lamin B1, we developed a 50% efficient knockdown (KD) approach that we maintained for all experiments. The 50% efficiency of the KD was verified by mRNA expression and Western blot analysis (Figure R.17).

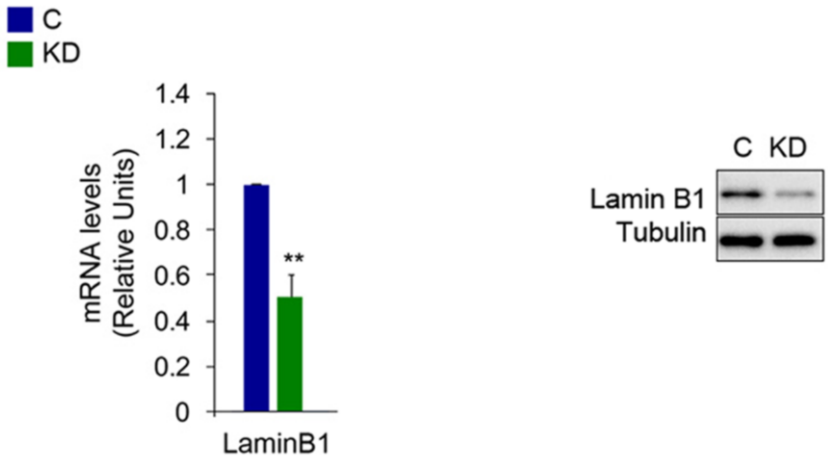


Figure R.17. 50% efficiency of lamin B1 knockdown. qRT-PCR showing mRNA levels of lamin B1 in KD NMuMG cells. Expression levels were normalized to an endogenous control and expressed relative to the control-infected cells, which was set as 1 (left panel). Representative Western blot for lamin B1 and tubulin (as a loading control) from control and KD NMuMG cells (right).

Previous studies demonstrated that highly efficient loss of lamin B1 leads to decreased cell proliferation rates[192] as well as misshapen nuclei[191]. NMuMG cells with a 50% depletion of lamin B1 displayed

neither disrupted nuclear lamina (Figure R.18A) nor changes in growth rates as compared to control cells (Figure R.18B).

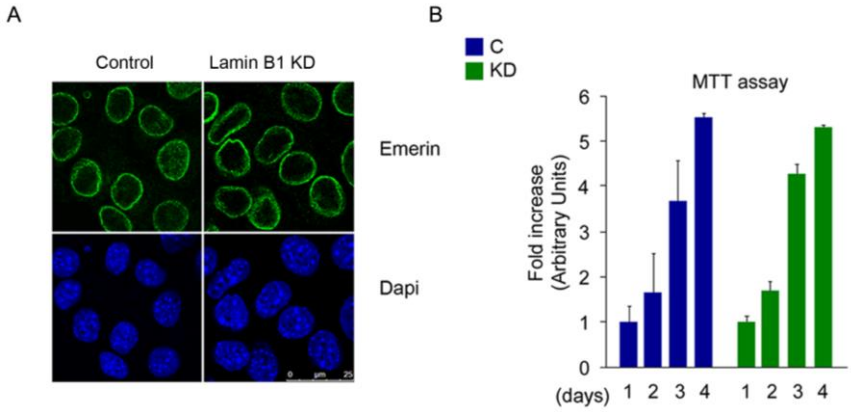


Figure R.18. Lamin B1 knockdown does not alter normal cellular characteristics. (A) Control (C) and KD NMuMG cells were stained with DAPI (blue) and immunolabelled with an antibody against the nuclear envelope protein emerin (green). Scale bar, 25 μm. **(B)** Quantification of cell viability by the MTT assay in NMuMG cells infected with either control or lamin B1 shRNA (KD). Measurements were obtained over four consecutive days after selection. Error bars indicate the SD from at least three independent experiments.

The best characterized effect of lamin B1 depletion is the premature induction of senescence[190,192]. However, 50% deficient lamin B1 NMuMG cells did not show senescence at any point during the EMT process as compared to cells treated with doxorubicin at different concentrations (as a positive control) (Figure R.19).

In sum, these results indicate that KD cells are not morphologically altered and behave as normal NMuMG cells.

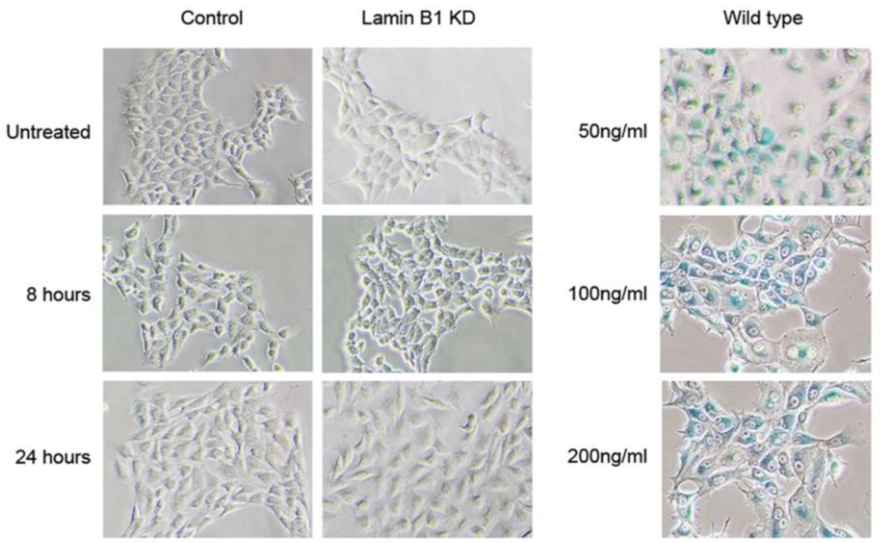


Figure R.19. NMuMG lamin B1 deficient cells do not show premature senescence. SA-β-gal (senescence-associated galactosidase) staining of control and KD NMuMG cells that were untreated or treated with TGF-β for 8 or 24 hours after selection (left). SA-β-gal staining of wild-type NMuMG cells treated with different concentrations of doxorubicin for 48 hours was used as a positive control of senescence (right).

7. Downregulation of lamin B1 mainly affects the nuclear interior fraction

Given the low rate of protein turnover of nuclear lamins once they have been stably integrated into the nuclear envelope[193], and the use of the less efficient NMuMG knockdown model that we intentionally developed, we reasoned that the first fraction to be affected after lamin B1 KD would be the non-NL or the less-integrated NL lamin B1, located in the nuclear interior fraction. To test this, we next extracted nuclei from control and KD NMuMG

cells to further separate soluble from insoluble proteins. Interestingly, under KD conditions, primarily the soluble fraction of lamin B1 was affected (Figure R.20A).

We then used fluorescence recovery after photobleaching (FRAP) experiments to further investigate the idea that KD affected mainly the non- or less-stably integrated NL lamin B1 fractions rather than the high-integrated-NL fraction. FRAP consists of specifically bleaching a region of the cell with the maximum potency of the microscope laser and monitoring the recovery time of the fluorescence over a certain period. For this, NMuMG cells transfected with human mCerulean-lamin B1 infected with shRNA lentivirus (using either control or human specific-lamin B1 shRNA). While the recovery after bleaching of the high-NL integrated fraction was similar in mCerulean-lamin B1 NMuMG cells in both control and KD conditions, the non-NL or less-integrated NL fraction recovered faster in control cells than in KD cells. Quantification data of the mean recoveries for 10 cells are shown in Figure R.20B and Figure R.20C.

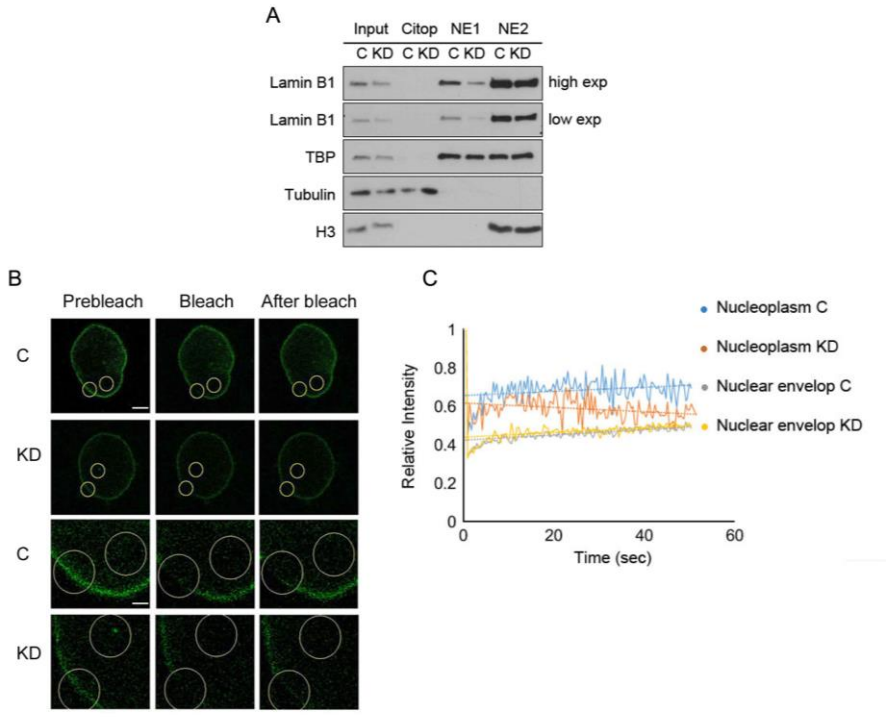


Figure R.20. Downregulation of lamin B1 occurs primarily in the nucleoplasm. (A) Fractionation of nuclei from control and KD NMuMG cells to obtain cytoplasmic proteins, soluble proteins (NE1) and insoluble proteins (NE2), as monitored by western blot using TBP as a NE1 control, tubulin as a cytoplasmic control and H3 as a NE2 control. **(B)** Representative images of FRAP experiment. Circles denote the areas of bleaching. Each column displays prebleach, bleached and post-bleached images. Row 3 and 4 are magnifications of row 1 and 2, respectively. Scale bar, 5 μm (inset, 1 μm). **(C)** Graphic representation of relative intensity (after 60 seconds) of nuclear envelope and nucleoplasm recovery from bleaching in control (C) and KD cells.

Lamin B1 occupancy in KD conditions was reduced by 50% in three ChIP-seq selected LB1+ sites (eLADs) as compared with control cells (Figure R.21A). This is in agreement with previous results that indicated that keeping the KD at about 50% drastically minimized indirect effects. Concomitantly, lamin B1 occupancy did

not change upon KD induction in cLADs (Figure R.21B), further supporting the idea that downregulation of lamin B1 mainly affects the non-NL or less-integrated NL lamin B1 fractions (which would include eLADs) rather than the high-integrated-NL fraction (which would include cLADs).

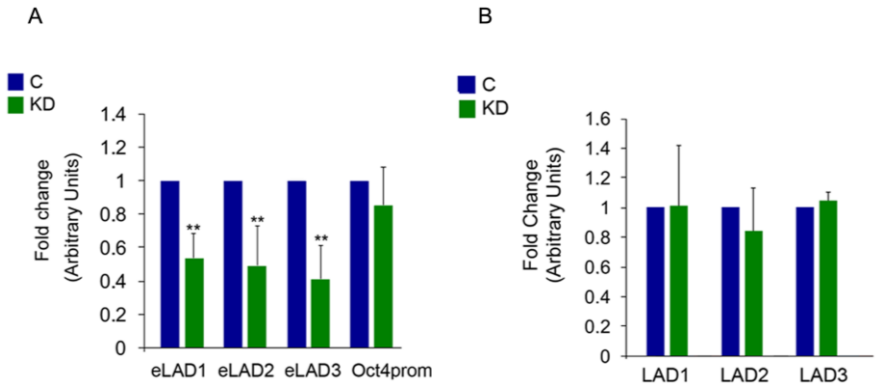


Figure R.21. Lamin B1 occupancy is reduced in eLADs but not in cLADs. (A) ChIP-qPCR of three selected LB1+ regions and a negative control (Oct4prom), selected from a negative LB1 region, in control (C) and KD NMuMG cells. Data from real-time PCR (qPCR) amplifications were normalized to the input and expressed as the fold-change relative to data obtained in control condition, which was set as 1. **(B)** ChIP-qPCR of three cLAD-selected regions in control and KD NMuMG cells. Data from qPCR amplifications were normalized to the input and expressed as the fold-change relative to data obtained in control condition, which was set as 1. Error bars indicate the SD from at least three independent experiments. *$p < 0.05$, **$p < 0.01$.

8. Downregulation of lamin B1 impairs EMT

Considering the putative role of dynamic eLADs in reorganizing the chromatin to acquire the final mesenchymal phenotype, and as a clear link exists between genome architecture and transcription[194],

we next sought to analyze the lamin B1–dependent transcriptome profile.

Interestingly, differentially expressed genes showed a maximum dependence of lamin B1 at 8 hour of TGFβ treatment, with more than 2000 genes differentially expressed in KD cells as compared with control cells. Strikingly, comparing the differentially expressed genes with the genes enriched in lamin B1 obtained by ChIP-seq revealed that around 50% of these differentially expressed genes were direct targets of lamin B1 (Figure R.22A).

Strikingly, principle component analysis (PCA) revealed that KD cells at 8 hours of TGFβ induction were the most highly divergent from control cells (Figure R.22B). Importantly, this is also the time point at which we observed the highest number of newly formed LB1+ regions (Table A.1).

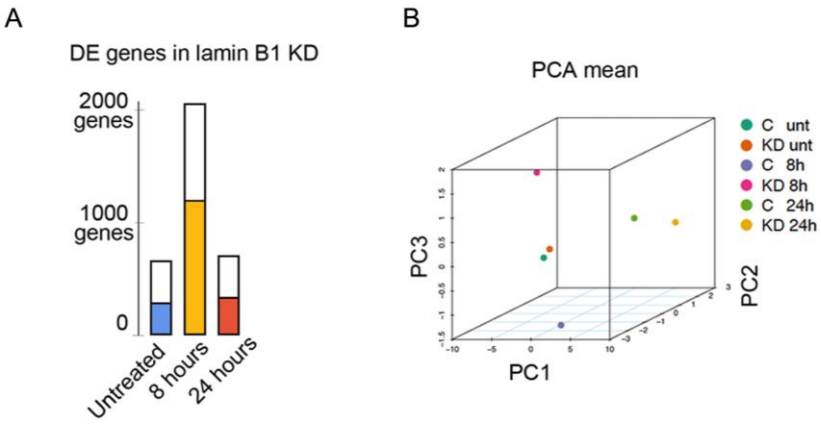


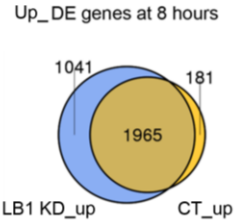
Figure R.22. Lamin B1 dependent genes increase at the intermediate state of the EMT. (A) Number of differentially expressed (DE) genes in KD cells during the EMT process. The proportion of genes on each list that are direct targets of lamin B1 in the ChIP-seq experiment of NMuMG cells untreated (blue), or treated with TGF-β for 8 hours (orange) or 24 hours (red), are shown inside each bar. **(B)** Principal component analysis (PCA) plot of RNA-seq based expression profiling in KD conditions. Each time point was normalized to its respective control.

At the 8-hour time point, 2000 genes were differentially expressed, of which about 1000 genes were upregulated in KD cells as compared to control conditions; therefore, 1000 genes were downregulated in KD cells (Figure R.23A).

These results suggest that the absence of lamin B1 indeed altered transcription rates. However, as half of the lamin B1–dependent genes were upregulated, and half were downregulated, this protein did not seem to act as a repressor or an activator. Rather, it is more likely to be an architectural protein, allowing changes in the general transcriptome profile that are necessary in EMT for cells to acquire the final mesenchymal status. Moreover, differentially

expressed genes that were upregulated in KD conditions did not belong to the EMT pathway (Figure R.23B), suggesting that downregulation of lamin B1 completely changes the EMT transcriptional program.

A



B

1041 up-regulated genes in KD cells upon 8 hours TGFβ treatment

Gene Set Name [# Genes (K)]	Description	# Genes in Overlap (k)	k/K	p-value	FDR q-value
HALLMARK_MTORC1_SIGNALING [200]	Genes up-regulated through activation of mTORC1 complex.	20		2.49 e^{-9}	1.25 e^{-7}
	secretion pathway.				
HALLMARK_HEME_METABOLISM [200]	Genes involved in metabolism of heme (a cofactor consisting of iron and porphyrin) and erythroblast differentiation.	16		1.97 e^{-6}	1.68 e^{-5}
HALLMARK_MITOTIC_SPINDLE [200]	Genes important for mitotic spindle assembly.	16		1.97 e^{-6}	1.68 e^{-5}
HALLMARK_TNFA_SIGNALING_VIA_NFKB [200]	Genes regulated by NF-kB in response to TNF [GeneID=7124].	16		1.97 e^{-6}	1.68 e^{-5}
HALLMARK_UNFOLDED_PROTEIN_RESPONSE [113]	Genes up-regulated during unfolded protein response, a cellular stress response related to the endoplasmic reticulum.	12		2.02 e^{-6}	1.68 e^{-5}
HALLMARK_ADIPOGENESIS [200]	Genes up-regulated during adipocyte differentiation (adipogenesis).	15		8.96 e^{-6}	5.6 e^{-5}

Figure R.23. Lamin B1 is necessary for the regulation of genes involved in EMT. (A) Venn diagram depicting the overlap between upregulated genes after 8 hour treatment with TGF-β in KD and control conditions. **(B)** GO analysis of 1041 genes upregulated in KD conditions at 8 hours after TGF-β treatment.

To further investigate the importance of lamin B1 in EMT, we checked the expression of classical EMT markers in NMuMG cells with reduced levels of lamin B1 by Western blot. As expected, we detected altered expression patterns for all proteins analyzed; these altered patterns included a lack of fibronectin, upregulation of N-cadherin and a loss of E-cadherin downregulation (Figure R.24A). Moreover, KD cells were not able to migrate or to invade, which are well-known abilities of cells that have acquired the mesenchymal phenotype (Figure R.24B). These results indicate a clear role of lamin B1 in the EMT process, as its downregulation functionally impaired this process.

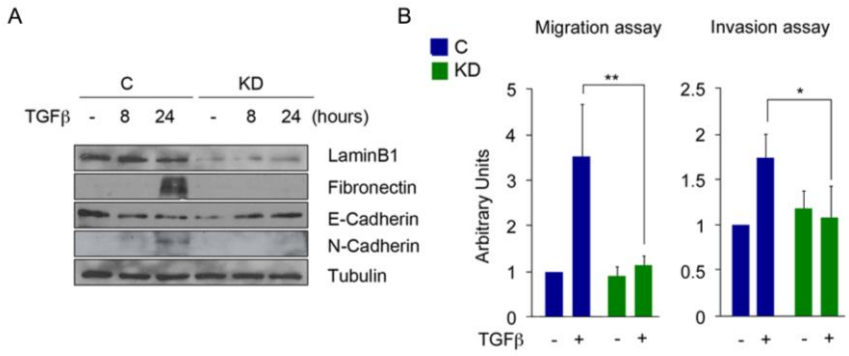


Figure R.24. Lamin B1 downregulation impairs classical EMT markers. (A) Western blot showing protein levels of different EMT markers and lamin B1 in control and KD NMuMG cells that were untreated or treated with TGF-β for 8 or 24 hours. Tubulin was used as a loading control. **(B)** Migration (left panel) and invasion (right panel) assays performed with control and KD NMuMG cells after 24 hours of TGF-β treatment. Error bars indicate the SD from at least three independent experiments. *$p < 0.05$; **$p < 0.01$.

The causative role of lamin B1 gene association with EMT impairment was revealed by ChIP-qPCR to be LB1+ EMT genes at 8 hours of TGFβ induction, such as fibronectin, vimentin and Twist 2. Specifically, enrichment of lamin B1 in these specific genes was lost in KD conditions as compared to control cells at the 8-hour time point (Figure R.25).

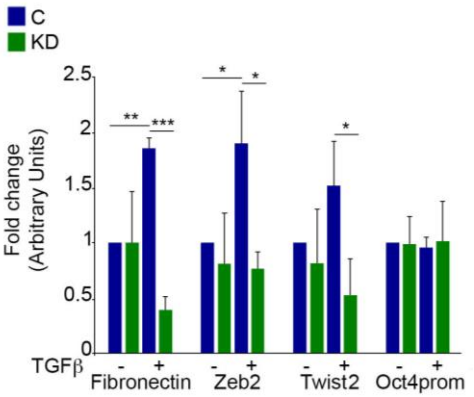


Figure R.25. Lamin B1 binding in EMT related genes is lost in KD cells. Chip-qPCR of three selected EMT-related LB1+ regions at 8 hours after TGF-β treatment in control and KD conditions. Oct4prom was used as a negative control. Data from qPCR were normalized to the input and expressed as the fold-change relative to the untreated C condition, which was set as 1. Error bars indicate the SD from at least three independent experiments. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

Previous studies performed in our group described an important role of HP1α in chromatin reorganization during EMT[6]. Specifically, HP1α is transiently released from chromocenters at 8 hours after TGFβ induction, coincident with transcription factor Snail1 upregulation, and is subsequently reloaded into chromocenters. It was proposed that a few hours after EMT induction (e.g., 8 hours),

HP1α is first released from chromatin and subsequently accumulates in the nucleoplasm.

As both localization of HP1α as well as the presence of lamin B1 seem to be relevant for the induction of the EMT process, we decided to check whether HP1α is still released from chromocenters under our KD conditions, Immunofluorescence of HP1α in KD conditions at the onset of EMT showed an early release of HP1α from chromocenters as compared to control conditions, which was maintained during the process (Figure R.26). Therefore, these results suggest that when lamin B1 is downregulated, EMT is impaired, probably due to alterations in chromatin reorganization, as HP1α localization might be important for genome assembly.

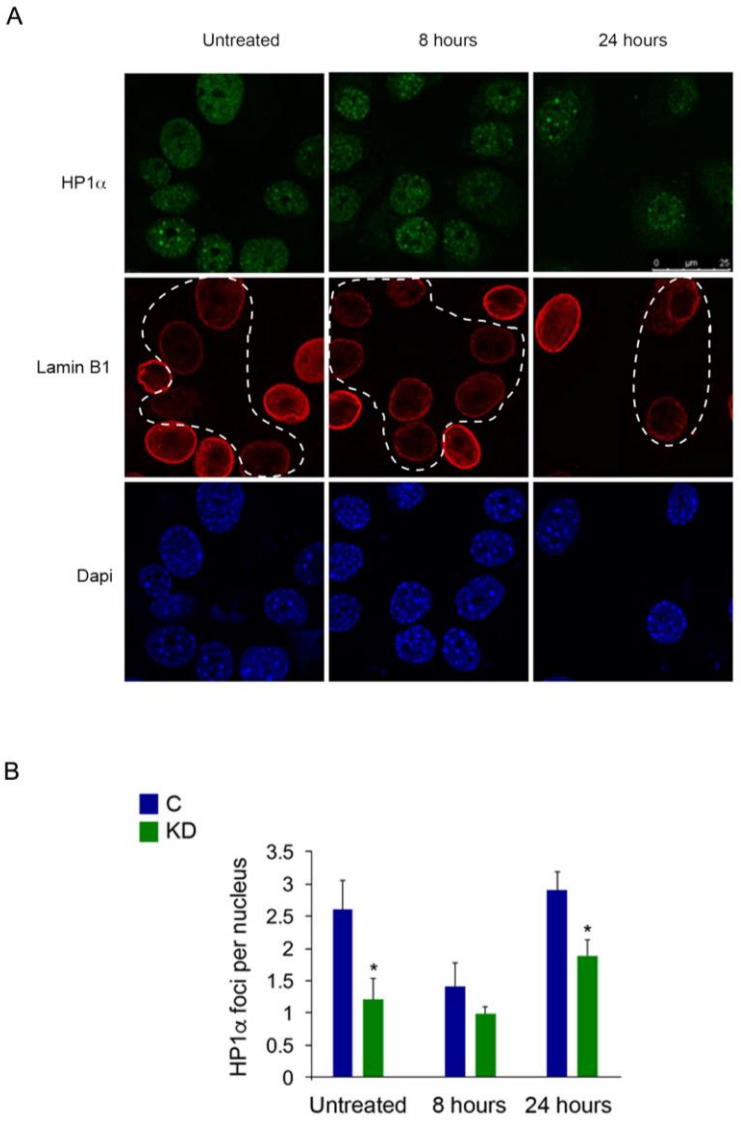


Figure R.26. Downregulation of lamin B1 allows premature release of HP1α. (A) NMuMG cells infected with shLB1 (KD) were treated with TGF-β and then immunolabelled with HP1α and lamin B1 antibodies and stained with DAPI. Scale bar, 25 μm. Dashed lines highlight the KD cells. **(B)** The graph depicts the average of HP1α-foci number per nucleus. Lamin B1 staining was used to select control or KD cells at different time points of TGF-β treatment. Error bars indicate the SD from at least three independent experiments. *$p < 0.05$.

Importantly, the release of HP1α from chromocenters did not lead to a decrease in fluorescence intensity. Indeed, total HP1α protein levels did not change in KD conditions (Figure R.27A). This might indicate that HP1α was not downregulated but rather was delocalized from pericentromeric heterochromatin regions.

To address a possible delocalization of HP1α, we extracted isolated nuclei with high salt. As previously published[6], endogenous HP1α in control cells was mostly released at 8 hours after TGFβ treatment, as compared with untreated or 24-hour treated cells. However, total HP1α levels did not change, as shown in the input by Western blot (Figure R.27A). This suggests that HP1α was loosely bound to the chromatin at this 8-hour time point. In contrast, salt extraction of nuclei from KD cells showed that HP1α were released mainly in untreated cells as compared with the other time points. These results are concurrent with our previous results showing that HP1α is prematurely released from chromocenters (Figure R.26). Importantly, similar to nuclei isolated from control cells, no changes in total HP1α protein levels were observed from KD nuclei when an input fraction was loaded in the Western blot (Figure R.27A).

MNase digestion assay allowed us To study the total chromatin compaction levels, we next used the MNase digestion assay; the micrococcal enzyme makes double-strand cuts between nucleosomes, such that the compaction level of the chromatin is determined by the enzyme accessibility and digestion efficiency. Similar to the HP1α release at 8 hours in control conditions, we observed that chromatin was more compacted at 8 hours after EMT induced as compared with untreated or fully mesenchymal

cells. These results fit perfectly with the idea that HP1α is a crucial protein involved in chromatin reorganization. In contrast, in KD cells, chromatin became highly compacted in untreated conditions, consistent with the high level of release of HP1α from chromocenters (Figure R.27B). Thus, these results reinforce our hypothesis that an early release of HP1α in KD conditions allows aberrant chromatin reorganization.

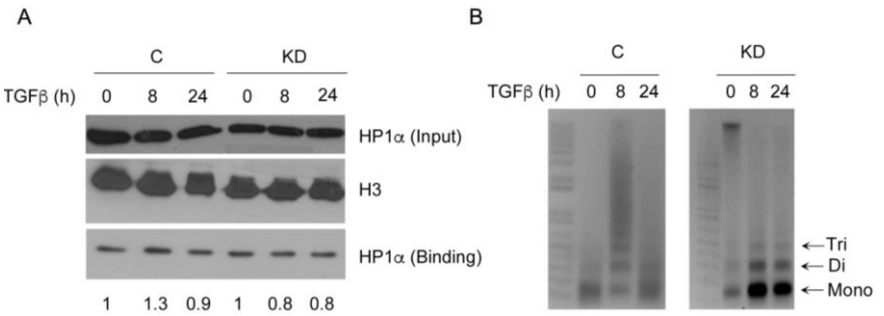


Figure R.27. Lamin B1 downregulation leads to early release of HP1α from chromatin and early chromatin compaction. (A) Salt extraction assay in control and KD cells treated with TGFβ. Purified nuclei of NMuMG cells were extracted with NaCl (0.5 M). The extracted fractions and an input fraction were analyzed by Western blot. H3 was used as a loading control. Quantification of the extracted chromatin fractions is shown in the bottom. **(B)** Nuclei isolated from control and KD NMuMG cells treated with TGFβ were digested with Micrococcal enzyme. Total DNA was analyzed by agarose gel electrophoresis. Positions of the nucleosomes are indicated (mono-, di- and tri-).

Even in light of these results, we could still not rule out that the effects observed with the downregulation of lamin B1 protein are due to indirect proteins alterations. To address this doubt, we next performed lamin B1 rescue experiments. Human lamin B1-GFP (hLB1-GFP) or empty-GFP vectors were efficiently overexpressed in control and KD NMuMG cells (as monitored by Western blot), with the endogenous knockdown of lamin B1 maintained (Figure R.28A). Immunofluorescence analysis verified the nuclear localization of hLB1-GFP, which contrasted with the widespread empty-GFP signal (Figure R.28B). Importantly, cells rescued with hLB1-GFP had restored migratory capacities. Thus, we suggest that lamin B1 is responsible for the acquisition of the final mesenchymal phenotype, functioning by helping the 3D chromatin rearrangements that are known to occur during EMT.

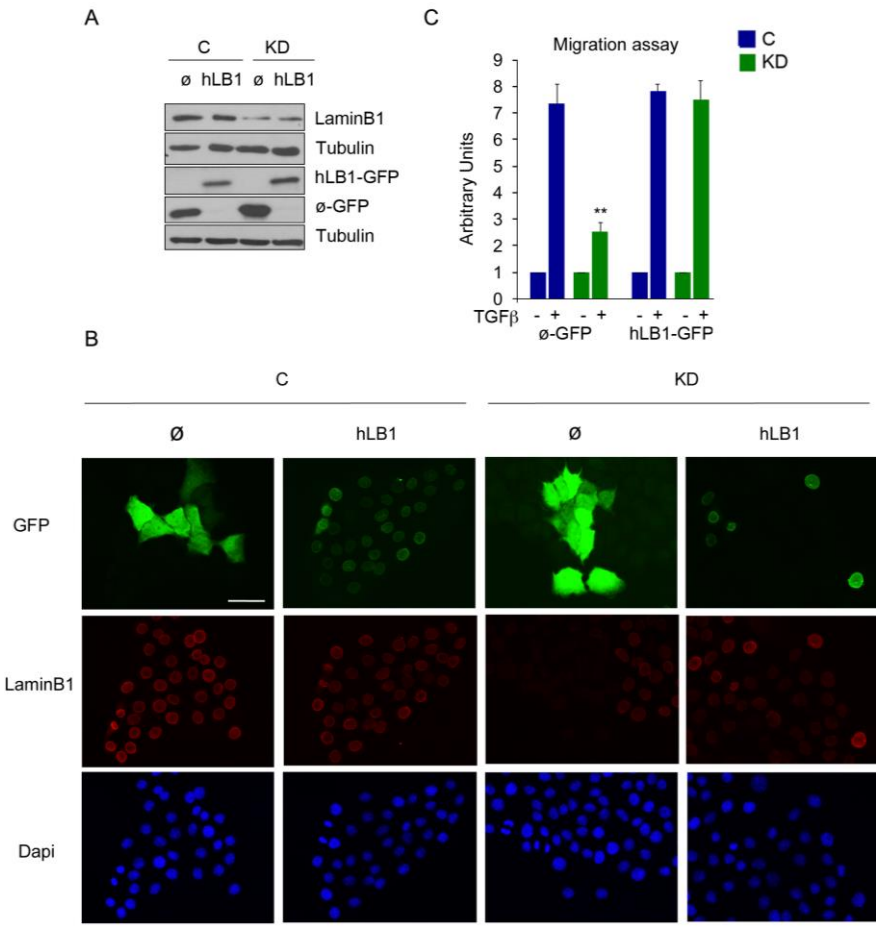


Figure R.28. Overexpession of human lamin B1-GFP restored the migratory capacities of KD cells. (A) Western blot of overexpressed empty-GFP or hLB1-GFP from control or KD NMuMG cells. Tubulin was used as a loading control. **(B)** Epifluorescence analysis of DAPI staining and immunofluorescence against GFP and lamin B1 in C and KD cells transfected with either empty-GFP or hLB1-GFP. **(C)** Migration assay of C and KD NMuMG cells transfected with either empty-GFP or hLB1-GFP. Scale bar, 50 μm.

DISCUSSION

1. Heterochromatin organization and EMT

Changes in chromatin organization occur during genome reprogramming and cellular differentiation[195]. Pericentromeric repeats from different chromosomes tend to cluster together into structures called chromocenters, contributing to the repressed status of specific regions, and in particular, to heterochromatin regions. Interestingly, this clustering is cell-type specific, with chromocenter number and size differing between cell types[196]. In general, chromocenters in differentiated cells are smaller, more numerous and more condensed than in undifferentiated cells[62].

We have shown that chromocenter re-organization take place during the EMT cellular differentiation process. In particular, we showed a clear increase in the number of chromocenters in NMuMG cells that had been treated with TGFβ to induce EMT, as compared with untreated cells. Different cell types exhibit different chromocenter organization; therefore, it is not surprising that during EMT process, epithelial and mesenchymal cells differ in their chromocenter organization, as these are completely distinct cell types. Specifically, mesenchymal-like cells have smaller but more numerous chromocenters than epithelial cells, in agreement with other differentiation pathways, such as embryonic stem cell differentiation[62].

Interestingly, we also showed an increased association between constitutive a heterochromatin histone mark (H3K9me3) and chromocenters during EMT cellular transformation. In addition to re-organization of chromocenters during EMT, specific chromatin domains across the genome are also reprogrammed[197]. Thus, it is

tempting to speculate that during EMT, heterochromatin regions cluster together to form domains of increased repression that might be essential for the acquisition of the mesenchymal phenotype.

We also observed that increased numbers of chromocenters as well as increased association with the H3K9me3 histone mark coincided with a slight decrease of transcription rates. These observations are in agreement with the well-established role of heterochromatin in gene repression. Nevertheless, it is worth noting that the decrease in transcription rates was very subtle with respect to those expected, given the clear association between H3K9me3 and chromocenters at 24 hours after EMT induction by TGFβ. Pericentromeric heterochromatin, although generally repressed, is actively transcribed to generate a new subtype of lncRNAs[6] that are known to be tightly regulated during EMT. Thus, the slightly decrease we observed at 24 hours, when the number of chromocenters had increased, and the increased levels of H3K9me3, might be due to transcription of these pericentromeric regions.

More accurate experiments are necessary in the future to determine whether transcription rates decrease during EMT. Technology using RNA approaches to equilibrium sequencing (e.g., RATE-seq) would be a useful tool for clearly measuring the abundance of transcripts by calculating absolute RNA degradation and synthesis rates[198].

Finally, Hi-C data revealed a higher number of bins that moved from the A to the B compartment at 24 hours after TGFβ treatment as compared to those moving from B to A or to those at the 8-hour time point. Interestingly, those genes that moved from an active

compartment (A compartment) to an inactive one (B compartment) in the mesenchymal-like state (at the 24-hour time point) are highly related with epithelial characteristics (cell-cell adhesion). A recent publication demonstrated that pericentromere-associated domains (PADs) are associated with the B compartment[199]. Thus, it is worth considering the possibility that, during EMT, the heterochromatin reorganization that takes place might occur mainly in chromocenters, where epithelial genes that have to be inactive in the mesenchymal state cluster together and associate with the inactive B compartment. Moreover, as H3K9me3 domains do not allow differentiation genes to be activated, therefore being restrictive for cell fate control[200] and reinforcing the idea the chromocenter formation would be necessary to establish the final mesenchymal phenotype. However, further experiments, such as FISH to determine the movement of these specific genes and H3K9me3 ChIP-sequencing, would be necessary to demonstrate that.

2. Lamin B1 in contact with euchromatin regions

Changes in heterochromatin might lead to dynamic changes in euchromatin. Therefore, it is highly likely that euchromatin will also be reorganized during this cellular transformation.

Nuclear lamins form separate but interacting filamentous protein networks in the INM from the nuclear envelope to give rise to the nuclear lamina[201]. Regions of chromatin in contact with nuclear lamina – LADs – are mainly heterochromatin regions that are relatively static and are highly conserved between species. Besides

the important role lamins play in maintaining structural properties of the nuclei and tethering heterochromatin to the nuclear periphery, the presence of lamins in the nuclear interior has also been reported[202,203], but their specific functions have been poorly understood until now. Indeed, nuclear interior lamins have long been considered as solely a transient pool on the way to assembly into the nuclear lamina. Recent studies have now revealed important regulatory roles of nucleoplasmic A–type lamins in gene expression and chromatin regulation, through their association with euchromatin regions[184]. In agreement with this, we have observed with confocal immunofluorescence that lamin B1 is located in the interior of the nucleus. Interestingly, colocalization with emerin, an integral protein of the inner nuclear membrane, is only observed in the nuclear membrane, which strongly suggests specific lamin B1 localization in the nuclear interior.

We now propose that functional units termed euchromatin lamina-associated domains (eLADs) exist and are distinct from the better characterized cLADs. This proposal is based on numerous data: the specular image of chromatin regions in contact with lamin B1 in lamin B1 ChIP-seq; the enrichment of these regions with G/C rich sequences; the association of lamin B1 with euchromatin histone marks; the location of lamin B1 in active and accessible chromatin regions, the clustering of lamin B1–positive sites; and the structural distinctions between eLADs and cLADs (with eLADs being smaller and having less genome coverage). Thus, we have shown, for the first time, that lamin B1 can not only interact with heterochromatin, as previously published, but also with euchromatin. As lamins form distinct but interconnected filamentous meshworks, and as nucleoplasmic A-type lamins have been recently found to contact

euchromatin regions, it is not surprising that B-type lamins interact also with chromatin primarily located in the nuclear interior of cells. The overlap we observed between euchromatin lamin A/C regions and LB1+ regions further support this suggestion. Notably, the specificity of the antibody was validated, allowing us to discard that large amounts of artefacts interfered in our analysis of the lamin B1 ChIP-seq data.

Chromatin digestion is a crucial step for sequencing ChIP samples. The association between lamin A/C and euchromatin presented in previous studies[126,184] was reached through low levels of chromatin shearing before sequencing. Thus, low-to-moderate shearing conditions leads to shearing of "open" regions (euchromatin) but not of "closed" regions (heterochromatin), which yield larger fragments that are usually excluded from sequencing. On the other hand, high sonication conditions (for more rigorous shearing) lead to heterochromatin enrichment, while the majority of euchromatin regions are lost, probably due to hyperfragmentation. Our results from the bioanalyzer profile show a high portion of fragments larger than 1 Kb after low sonication conditions; these large fragments mainly correspond to heterochromatin and are excluded from sequencing, leading to detection of lamin B1 that is primarily in contact with euchromatin regions. Importantly, in an assay that does not select for size but analyzes all chromatin fragments (ChIP-qPCR), we detected lamin B1 in both euchromatin as well as heterochromatin regions.

In this work, we have demonstrated that lamin B1: *i*) binds to gene-rich regions; *ii*) associates with euchromatin histone marks; *iii*) is transcriptionally active; and *iv*) is located in accessible chromatin

regions. All of these characteristics can be used to describe euchromatin. Nonetheless, we cannot conclude where these newly identified lamin B1 sites are located in the nucleus. In this context, three scenarios may be plausible:

1. Lamin B1 that contacts euchromatin may be located in the nucleoplasm, forming a pool independent from the nuclear lamina. Although B-type lamins remain permanently farnesylated and carboxymethylated (in contrast to A-type lamins), and therefore are tightly associated to nuclear membranes[204], they have also been reported to localize in the nuclear interior[117].

2. Euchromatin-bound lamin B1 could also be located in nuclear lamina invaginations that reach into the nucleoplasm, yet are still membrane bound. Recently developed super-resolution microscopy techniques allow nuclear lamina to be studied in more detail for features that are below the optical diffraction limit[205]. Notably, these invaginations contain inner and outer nuclear membrane, as well as nuclear pore complexes and both A and B type lamins[206].

3. Finally, a peripheral pool of lamin B1, which would be less well integrated to the nuclear lamina filament network, could also be in contact with genomic regions associated with active genes. Given that we detect euchromatin regions in contact with lamin B1, and that the permanent modified state of lamin B1 implies its association with nuclear lamina, this last scenario is the one that might fit better with our observations. Nevertheless, further experiments are needed to elucidate where these new identified lamin B1–euchromatin contacts are located within the nucleus.

Strikingly, the results we present in this thesis regarding lamin B1 binding euchromatin go against the well-established dogma of LADs. Indeed, Gesson *et al.* did not detect any lamin B1 in euchromatin regions but only in heterochromatin regions, although they were also working with similar low sonication conditions[184]. This apparent discrepancy could be explain by the distinct configurations of epithelial and mesenchymal cell types and the use of different lamin B1 antibodies for the two studies. Nevertheless, Gesson *et al.* claim that moderate sonication conditions lead to completely exclusion of heterochromatin fragments from sequencing, allowing them to detect lamin A/C in contact with euchromatin. Thus, it is difficult to imagine a scenario in which heterochromatin is excluded but nonetheless detected, and that it is only in contact with lamin B1 but not with lamin A/C.

On the other hand, it is still unclear why the presence of lamins (lamin A/C and B1) in contact with euchromatin regions cannot be detected by the DamID (DNA adenine methyltransferase identification) method, which was the first one used to characterize LADs. This technique is based on the expression of a fusion protein carrying DNA adenine methyltransferase (Dam) and protein of interest (in this case, lamins). Dam adds methyl groups to adenines in GATC sequences, a modification that is not present in most eukaryotes, near genomic regions where fusion protein interacts with DNA[207]. While the DamID technique is very powerful and useful, it is known that methylation by Dam proteins is not perfect, and that this technique introduces a high level of bias: first, the intrinsic affinity of Dam for GATC sequences causes considerable background levels; and second, methylation levels are also affected by chromatin accessibility and protein turnover.

To distinguish specific binding from this background signal, it is important to compare the methylation profile obtained with Dam alone (i.e., Dam not fused to a DNA binding protein). The binding is expressed as the ratio of methylation for Dam fusion:Dam, thereby normalizing the background methylation[208]. Moreover, DamID was recently proposed to favor the identification of stable interactions between lamins and chromatin, which are likely to occur within heterochromatin regions[209]. If we take this into consideration, together with the fact that are two different pools of lamin B1 (one highly associated to the nuclear lamina, which is more abundant and stable, and the other less integrated into the nuclear lamina, which is less abundant and has a faster turnover), it is thus plausible that specific euchromatin-lamin B1 interactions are discarded during the normalization for background methylation. Although DamID has been also used to map DNA binding sites of transcription factors in euchromatin, it is important to note that these proteins do not have the same level of localization complexity as that observed for lamin B1 (which has an extra pool with a distinct nuclear localization).

3. Role of euchromatin–lamin B1 during EMT

At the onset of EMT, most LB1+ sites completely changed, showing a dynamic association between nuclear interior lamin B1 and euchromatin. During differentiation, such as embryonic stem cell (ESC) differentiation, most genes maintain their patterns of interaction with the nuclear lamina; only genes involved in specific lineages reorganize over time, according to the specific differentiation stage[129]. Thus, tethering genes to the nuclear lamina is required for stable repression of genes during differentiation[42,210]. The dynamism between lamin B1 and euchromatin that we observe during EMT might indicate more complex gene regulation function than the one occurring at the nuclear periphery. Indeed, lamins are known to bind DNA, histones and histone-binding proteins. Recently, the dynamic pool of lamin A/C in contact with euchromatin regions has been demonstrated to have important roles in gene regulation, mainly affecting epigenetic pathways and chromatin remodeling[184]. It is thus conceivable that all type of lamins in the nuclear interior contribute to chromatin organization, although still little has been known about their potential roles in gene regulation.

Notably, our results now demonstrate a role for nuclear interior lamin B1 in gene regulation during the EMT process. We have shown a clear association between lamin B1 binding and gene expression, which is maintained during all the process. Interestingly, in untreated conditions, LB1+ genes are involved in general gene regulation and chromatin organization, but LB1+ genes formed upon TGFβ induction are clearly associated with the

EMT process. This might indicate that lamin B1 has an important role in regulating the expression of genes important for EMT. Moreover, loss of 50% of the lamin B1 protein level mainly affected nuclear interior lamin B1 yet led to an impaired EMT due to changes in the EMT transcriptional program. Furthermore, we showed that, after 8 hours of TGFβ treatment (to induce EMT), the transcription profile was altered, with 50% of altered genes being lamin B1 direct targets. This indicates a relationship between lamin B1, gene regulation and the EMT process. But how can lamin B1 mechanistically regulate gene expression during EMT? The control of transcription programs, and thus gene regulation, is generally mediated by three mechanisms that might act concomitantly:

First, transcription can be regulated through the association of transcription factors (TF) to specific chromatin regions, to regulate different sets of genes. Hundreds of proteins are able to interact with lamins, but lamina-interacting proteins can vary between tissues or cells[211]. Accordingly, our results show that upon TGFβ induction, new LB1+ sites are formed and enriched in TF binding motifs belonging to developmental processes, including the TGFβ signaling pathway. These observations suggest a putative role for these TFs in recruiting lamin B1 to regions that must be transcriptionally regulated during, in this case, EMT program. Nevertheless, further experiments would be needed to demonstrate this causative role.

Second, DNA methylation and posttranslational modifications of histones also influence gene regulation. Consequently, we observed a clear association between LB1+ sites and a mixture of euchromatin histone marks during all the EMT transition. In

particular, H3K79me2, a histone mark mainly located in gene bodies and related with gene transcription[212], was observed to associate with LB1+ sites. Importantly, sites that are not enriched in nuclear interior lamin B1 (which might correspond to cLADs) only associated with H3K9me3, the histone mark characteristic of constitutive heterochromatin, mainly present in cLADs. These results are thus in agreement with the idea that nuclear interior lamin B1, in contrast to lamin B1 in the nuclear periphery, is in close contact with euchromatin regions rather than with heterochromatin.

However, it is important to note that at the onset of the EMT, genes positive for lamin B1 appear to be associated with a repressive mark associated with Polycomb, namely, H3K27me3. This is not the first time nuclear interior lamins have been demonstrated to associate with this histone mark. Recent studies show that nucleoplasmic lamin A/C is required for the correct assembly of Polycomb protein (PcG) foci[213,214], which are involved in the repression of developmentally regulated genes. PcG proteins are transcription repressors that play central roles in development, located in the nucleus and forming the multimeric Polycomb repressive complexes (PRC), which post-translationally modify histones to repress specific genes involved in development. In mammals, there are two characterized complexes: PRC1 and PRC2. One of the proteins within PRC2, EZH2, is responsible for di- and trimethylation of lysine 27 in histone H3 (H3K27me3), an epigenetic mark recognized by PRC1[215]. During ESC differentiation, regions that reorganize from the periphery mainly involve genes related to cell identity, although the disassembly from the nuclear lamina does not induce direct activation of these

genes; rather, they become “unlocked” and can be activated at any stage[129]. Hence, as many developmental genes are only active in particular cell lines or tissues during development and are silent in the rest of the organism, and nucleoplasmic lamin A/C is also associated with PcG foci, it is not surprising to find nuclear interior LB1+ sites in association with facultative heterochromatin at the onset of the EMT differentiation process. Thus, the third option of a mechanisms by which nuclear interior lamin B1 could regulate gene expression would be through posttranslational modifications of histones (although this has to be further characterized). It is tempting to speculate that nuclear interior lamin B1 might play a role not only in general gene transcription but also in regulating expression of genes involved in differentiation, or at least those involved in the EMT program. Thus, in untreated conditions, genes in contact with nuclear interior lamin B1 would be related with general transcription and chromatin organization, while genes in contact with the nuclear lamina would be in a static repressive state. However, once the EMT process starts, nuclear interior lamin B1 might contact specific differentiation-related genes and thereby regulate its expression during different steps of the transition process.

Finally, gene transcription can also be regulated by dynamic higher-order chromatin organization, leading to active genes being “relocated” into a repressive compartment or *vice versa*. Given that intermediate filaments are the major contributors of cell architecture[188], and specifically, that nuclear lamins provide structural stability to the nucleus[216], we propose that lamin B1 and, at a higher level, eLADs, are at least in part responsible for 3D chromatin rearrangements during EMT. Nevertheless, not only

lamins in the nuclear periphery has been proposed to be involved in chromatin organization, but also nucleoplasmic lamin A/C, although specific mechanisms remain still elusive[217].

Several results we obtained support this role of lamin B1 in helping chromatin reorganization during EMT. Hi-C data in combination with lamin B1 ChIP-seq show that eLADs are mainly present in the A compartment during the entire process, and thus in a transcriptionally active domain. Importantly, during EMT, there is a significant increase of eLADs in the B compartment, which correspond to newly formed eLADs, rather than simply reflecting a movement of eLADs from the A to B compartment. Interestingly, genes present in these new eLADs are required to be in a repressive state when cells become mesenchymal, further supporting lamin B1 function as a chromatin organizer. Recently published studies show that the lamin meshwork is important for regulating 3D chromatin organization through the maintenance of proper inter- and intra-TAD interactions during ESC differentiation[218]. Accordingly, we have shown a correlation between lamin B1 enrichment and border strength, which suggests lamin B1 might be involved in the establishment of new intra-TAD interactions during EMT and, in turn, in the establishment of new genomic architecture during this cellular transformation. Importantly, the increase in the border strength we observe during the EMT is not due to transcription since we do not observe transcriptional changes at TAD borders enriched in lamin B1, further supporting the idea that lamin B1 might be directly involved in chromatin rearrangements. Nevertheless, we cannot discard that other proteins are also involved in establishing long-range interactions. Notably, RNA polymerase II and/or its associated

proteins (but not transcription *per se*) have been identified as mediators of interactions throughout the A compartment[219].

These different mechanisms that we propose here are not mutually exclusive. Indeed, it is likely that they coexist and together contribute to lamin B1's role in gene regulation during the EMT.

4. Lamin B1 as an architectural protein

We observed that lamin B1 enrichment correlated with increased TAD border strength during EMT. This suggests that lamin B1, acting as an architectural protein, has an important role in helping to establishing a new genomic architecture during this cellular transformation. The strength of TAD borders is measured as the ratio between inter- and intra-TAD interactions around border sequences and strongly correlates with the presence of architectural proteins. Thus, high occupancy of architectural proteins around border sequences leads to stronger borders, and *vice versa*[93]. Moreover, as previously mentioned, lamin B1 enrichment and border strength are not related with transcription changes in these regions, strongly supporting the idea that lamin B1 acts as an architectural protein to regulate chromatin rearrangements that in turn regulate gene expression.

Importantly, comparison between ChIP-seq results obtained here for lamin B1 during EMT and previously published ones for CTCF (one of the best characterized architectural proteins) in mouse embryonic fibroblasts (MEFs) reveal a clear overlap between both

proteins. Approximately 70% of LB1+ genes in all three time points studied overlap with CTCF genes present in MEFs. Concomitantly, CTCF+ sites highly correlate with LB1+ sites, as well as eLADs (Figure D.1). Interestingly, only 15% of CTCF binding sites are located within TAD borders, where they restrain interactions between promoters and regulatory regions located in different TADs; the majority (85%), however, are located inside TADs, where they facilitate interactions between promoters and their regulatory regions within the same TAD. Importantly, the functional differences between both regions rely on the number of architectural proteins present[143]. Thus, the correlation we observe (Figure D.1) might be in agreement with previous studies that have suggested that clustering of architectural proteins at TAD borders is necessary to generate new intra-TAD as well as constraining inter-TAD interactions, which establishes new patterns of gene expression during cell differentiation[146]. Further bioinformatics analyses are required to verify that the correlation occurs mainly at TAD borders, which would reinforce our theory.

In this light, it is important to also note that, while architectural proteins can recognize a unique DNA motif, they can also interact with accessory proteins that do not recognize specific DNA motifs; this can help architectural proteins in the 3D re-organization of the genome[220]. As we have not identified specific DNA motifs for lamin B1, we cannot discard that nuclear interior lamin B1 acts as an accessory protein rather than as a *bona fide* architectural protein. Thus, lamin B1 could interact with CTCF, allowing increased intra-TAD interactions during EMT and, in turn, 3D chromatin reorganization during this cellular transformation.

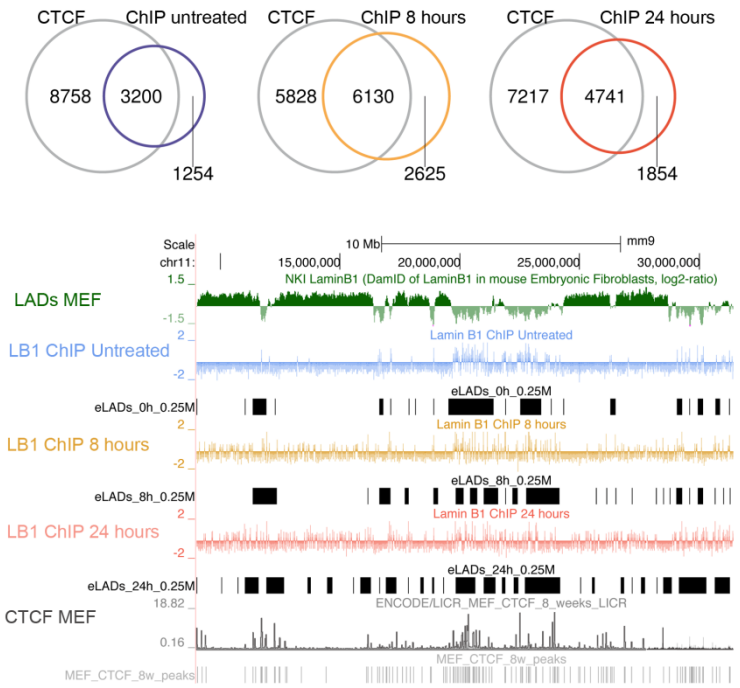


Figure D.1. CTCF and lamin B1 are present in the same regions during EMT. Venn diagram showing the overlap between CTCF genes and LB1+ genes in untreated cells (blue) and cells treated with TGFβ for 8 hours (orange) and 24 hours (red) (upper panel). UCSC Genome Browser overview of one region across chromosome 11 (mm9) of the lamin B1 ChIP-seq profiles in NMuMG cells, eLADs identified in each condition, previously published LADs and previously published CTCF ChIP-seq in mouse embryonic fibroblasts (MEF) (lower panel).

Generation of Hi-C maps in lamin B1 knockdown conditions, among other experiments, would be needed to determine whether lamin B1 plays a role as architectural protein. If it does, the correlation between lamin B1 and TAD border strength should decrease under knockdown conditions, leading to weaker border strength.

5. Relevance of nuclear interior lamin B1 for EMT

We have implemented a partial downregulation of lamin B1 with a specific short hairpin RNA, with a 50% decrease of both mRNA and protein levels that was maintained during all the experiments performed in this work. Intriguingly, we found that this lamin B1 downregulation mainly affected the nuclear interior protein, which could be located in the nucleoplasm or be less-stably integrated in the nuclear lamina (NL). This result is not surprising, given the low protein turnover/high stability of nuclear lamins once these are stably integrated into the nuclear lamina[193]. As only 50% of lamin B1 was downregulated in KD conditions, it is conceivable that the first subsets of lamin B1 to be altered in KD conditions are the nucleoplasmic or less-stably NL-integrated lamin B1 subsets.

Nuclear interior lamin B1 depletion impaired the EMT process, as shown by abrogation of migration and invasion capacities as well as altered protein levels of EMT markers. Restoration of migration capability upon lamin B1 rescue further demonstrates the crucial role of lamin B1 in this cellular transformation. Although is the first time that alterations in lamin B1 levels have been related to EMT, lamin B1 polymorphisms have been implicated in neural tube defects[221], a development stage in which EMT takes place[222]. Moreover, elevated lamin B1 levels have also been reported in human liver cancer patients[223].

Previous results from our lab have suggested that chromatin reorganization takes place during EMT cellular transformation. Indeed, upon TGFβ induction of EMT, there is a window of time

during which HP1α (a protein essential for maintaining heterochromatin structure) is released from heterochromatin, and major satellite transcription is downregulated, allowing chromatin organization and acquisition of mesenchymal traits[6]. Therefore, it has been proposed that initial steps of EMT induction are responsible for chromatin reorganization rather than for final mesenchymal cells properties.

Importantly, we have shown that at 8 hours after TGFβ treatment, NMuMG cells gained the highest number of newly formed lamin B1+ regions. Furthermore, differentially expressed genes showed a maximum dependence of lamin B1 at this time point, with more than 2000 genes differentially expressed in KD conditions as compared with control cells. Analyses of the differentially expressed genes upregulated in KD conditions show functions unrelated to EMT, indicating that the lamin B1 plays a crucial role in allowing induction of the EMT program at the initial steps of this transformation process.

Although HP1α delocalization from chromocenters depends on different factors, such as Snail upregulation and LOXL2 binding to specific regions leading to heterochromatin transcripts repression, it is a crucial step that must occur in order to allow cells to undergo EMT. Downregulation of lamin B1 not only functionally impairs EMT but also induces the earliest release of HP1α from chromocenters. Thus, in lamin B1 KD conditions, HP1α is released when cells are still untreated rather than after TGFβ treatment. Moreover, earliest release of HP1α from chromocenters leads to the earliest compaction of chromatin, which is in agreement with previous

findings indicating that the release of HP1α from chromocenters is key for chromatin reorganization during EMT.

In sum, as lamin B1 downregulation induces earliest chromatin compaction, it is tempting to speculate that lamin B1 acts as an architectural protein. In this case, lamin B1 would play an important role during EMT by facilitating the chromatin rearrangements that take place during this transformation process to acquire the final mesenchymal phenotype.

6. Role of lamin B1 in other cell contexts

We have demonstrated a clear role of nuclear interior lamin B1 in the EMT differentiation process. However, we cannot discard that lamin B1 might be involved in other differentiation and/or other cell contexts. Indeed, the GO categories found in genes enriched in lamin B1 in the epithelial state (untreated conditions) belong to general gene expression and chromatin modification, thus suggesting a general role of this intermediate filament in gene regulation in response to a given stimulus.

Nuclear lamins have been characterized as essential for proper development of organs such as lungs and brain, thus playing critical roles in organogensis[224-226]. Importantly, B-type lamins primarily play roles in neurodevelopment[227,228]. For that reason, we decided to check whether lamin B1 could play an important role in mouse embryonic stem cell (mESC) differentiation to neuroectoderm. mESCs were induced to differentiate to

neuroectoderm through serum-free medium containing N2B27 during 5 days (Figure D.2). Importantly, mESC lamin B1 KD led to an almost complete protein depletion of both RNA and protein levels using two different clones (Figure D.2A and Figure D.2B), in contrast with NMuMG cells, in which still 50% of RNA and protein levels were detected upon lamin B1 KD. This is not surprising, as we were using completely different cell lines that likely have distinct levels of lamin B, which might respond differently to protein downregulation.

Interestingly, upon 5 days of differentiation to neuroectoderm, lamin B1 in control (shCTRL 5d) conditions was slightly upregulated, whereas in KD conditions (shLB1 5d), it was almost completely abrogated, suggesting lamin B1 might play a role in this differentiation process. Pluripotency genes were not affected in lamin B1 KD conditions (Figure D.2B). Moreover, neither morphology nor proliferation of mESCs were altered upon KD of lamin B1 (data not shown). These results are in agreement with previous studies demonstrating that lamins are not essential for ESCs[224].

B-type lamins are required for proper organogenesis but are not necessary for the differentiation of ESCs to trophectoderm[224]. Surprisingly, after 5 days of differentiation to neuroectoderm, the upregulation of neural markers observed in control conditions was abolished in KD cells (Figure D.2C). This indicates that lamin B1 might be necessary for a correct neuroectoderm specification, and thus for induction of target lineage genes, rather than for repression of ESC pluripotent genes. It has been previously suggested that nuclear lamins evolved to facilitate the integration of different cell

types to complex tissue architecture, and therefore are only essential during organogenesis but not in previous stages of ESCs differentiation[229]. However, the role of B-type lamins in ESC differentiation has only been studied in trophectoderm differentiation. Given the fact that trophectoderm differentiation comes before neuroectoderm specification[230], it could be possible that B-type lamins play important roles in more advanced stages of differentiation.

Additional experiments to better characterize the lamin B1 roles in the differentiation of mESCs to neuroectoderm and even other lineages would be necessary to conclude that this intermediate filament is not only involved in EMT but also in other cell contexts. Moreover, it would be very interesting to determine, using euchromatin-enriched ChIP-seq, whether nuclear interior lamin B1 domains are also found in mESCs and are involved in neuroectoderm -specification process.

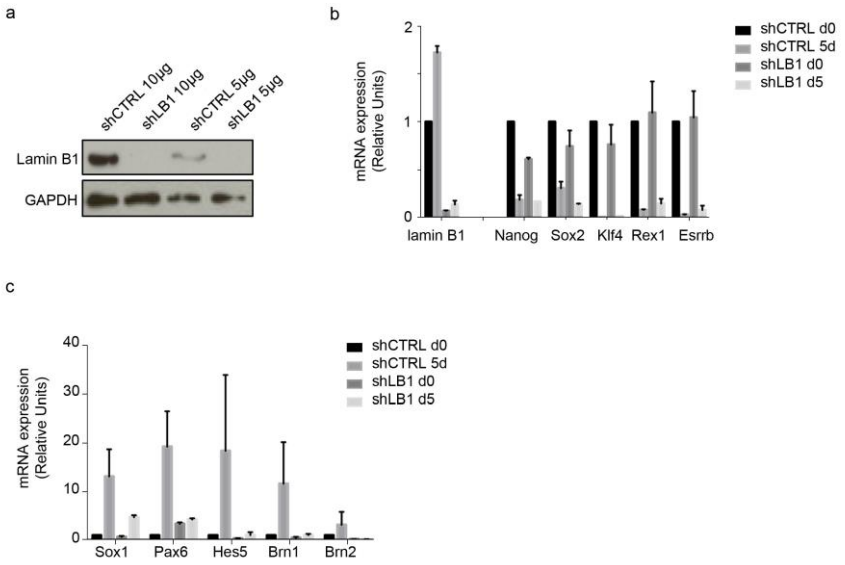


Figure D.2. Lamin B1 is also involved in mouse embryonic stem cells (mESC) differentiation to neuroectoderm. A. Western blot analysis of lamin B1 in control (shCTRL) and laminB1-depleted (shLB1) mouse Embryonic Stem cells (mESCs). GAPDH was used as a loading control. **B.** RT-qPCR analysis of pluripotency markers in proliferating condition (d0) or after 5 days of differentiation (serum-free medium containing N2B27) (d5) of control (shCTRL) and laminB1-depleted (shLB1) mESCs. **C.** RT-qPCR analysis of neural markers in proliferating condition (d0) or after 5 days ofdifferentiation (serum-free medium containing N2B27) (d5) of control (shCTRL) and laminB1-depleted (shLB1) mESCs. esults were normalized to the housekeeping *RplpO*. Error bars represent the standard error of the mean (SEM) of two independent experiments.

CONCLUDING REMARKS

Nuclear lamins have been classically believed to be only located in the nuclear periphery, in contact with the inner nuclear membrane and forming a stable structure named nuclear lamina. However, it has become evident that a substantial fraction of lamins also reside throughout the nuclear interior. Recent reports have demonstrated that A-type lamins not only bind to heterochromatin regions, as previously well established, but also to euchromatin regions in the nuclear interior, with important roles in gene regulation and chromatin organization[184]. Nevertheless, the exact functions and mechanisms of these new-identified nuclear interior lamins are just beginning to emerge. Importantly, A and B-type lamins form separated but interconnected meshwork of filaments with different roles. Furthermore, there is also evidence for the existence of a nuclear interior pool of lamins that are assembled into stable structures, showing different characteristics from those located in the nuclear lamina[117].

In this thesis, we show for the first time that lamin B1 can also interact with euchromatin regions, to form euchromatin lamina-associated domains (eLADs). Interestingly, in contrast to cLADs, which are located in the inactive B compartment, eLADs are mostly located in the A compartment. Importantly, these two types of lamin B1 domains (cLADs and eLADs) behave differently, which is consistent with the fact that nuclear interior lamins and those located in the nuclear periphery differs in their characteristics. While cLADs are static structures, eLADs are dynamic and change during EMT. In untreated conditions, eLADs are mainly located in the A compartment, but upon TGFβ treatment (e.g., EMT induction), new eLADs are formed that tend to localize in the B compartment. These observations correlate with changes in

chromatin organization which might be involved in gene regulation. Moreover, the correlation between lamin B1 enrichment and border strength during EMT suggests a possible participation of these domains in establishing new genomic architecture necessary for the acquisition of the final mesenchymal phenotype (Figure CR.1).

In conclusion, eLADs in untreated cells are located in the active compartment, where they help to regulate general gene expression. Once cells are treated with TGFβ to induce EMT, these domains might be responsible for the correct EMT process, probably through association with TAD borders, and thus through reorganization of genomic architecture to maintain expression and repression of EMT-related genes.

Many questions still remain to be addressed: is this lamin B1 pool located in the nucleoplasm as a soluble fraction or is it less-stably but still integrated with the nuclear membrane? Are the molecular structures of lamin B1 filament networks distinct in cLADs and eLADs? Do eLADs have a role in all types of cellular transformation processes? Which are the signals and/or transcription factors responsible for recruiting lamin B1 to different loci? Are the changes in interactions the cause or consequence of lamin B1 enrichment at TAD borders? Answers to these questions will contribute to a better understanding of genome reorganization during cellular transformation and to understanding the roles of euchromatin–lamin B1 contacts in gene regulation and, therefore, chromatin reorganization.

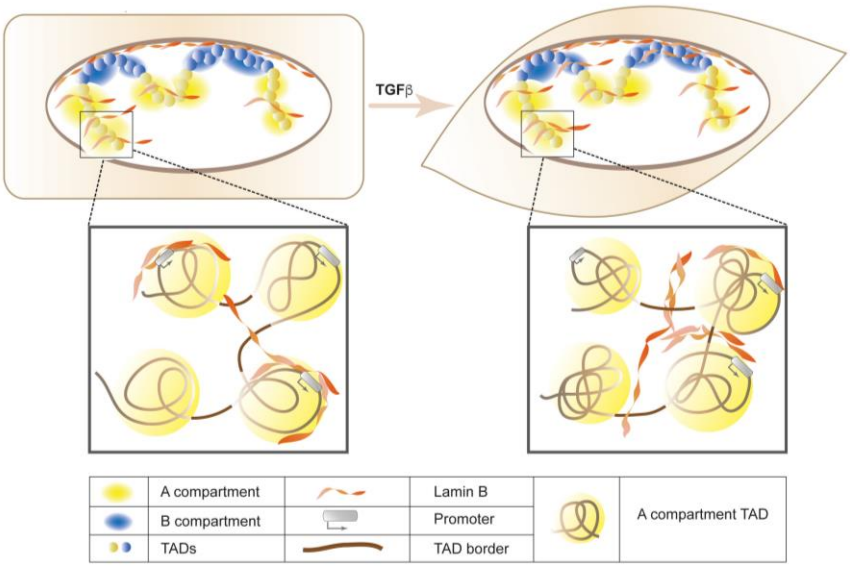


Figure CR.1. Representation of the working model. Lamin B1 contacts with active chromatin regions in the A compartment to form eLADs, which are dynamic and change during EMT. Upon TGFβ induction, some eLADs that are still located in the A compartment relocate, and new eLADs are formed in the B compartment. Moreover, during EMT, lamin B1 enrichment at TAD borders contribute to increasing border strength, thus participating in establishing novel genomic architecture during EMT.

MATERIALS & METHODS

1. Cell lines

The following cell lines were used in this study:

- **HEK293T:** human embryonic kidney cells, used for virus production;

- **NMuMG:** mouse epithelial mammary gland cells, used for all experiments performed. These cells undergo EMT when treated with TGFβ.

Both cell lines were maintained in Dulbecco's modified Eagle's medium (Invitrogen) supplemented with 100 units/ml penicillin, 100 µg/ml streptomycin, 2 mM glutamine and 10% FBS (Invitrogen) at 37ºC in 5% CO_2. NMuMG cells were also supplemented with 10 µg/ml insulin (Sigma). TGF-β was added to a final concentration of 1–5 ng/ml to stimulate NMuMG cells and induce EMT. Cells were collected after 0, 8 and 24 hours of treatment.

All cell lines were regularly tested for the absence of mycoplasma using standard PCR with the following primers: F: 5′-GGCGAATGGGTGAGTAACACG-3′ and R: 5'-CGGATAACGCTTGCGACCTATG-3′.

2. Lentivirus production and infection

For lentiviral infections, HEK293T cells were used to produce viral particles. Cells were seeded in p150 plates with 14 ml medium, grown to 70% confluency and then transfected (day 0) by adding, dropwise, a mixture of 3.3 ml of NaCl 150 mM, DNA (7.5 µg of mouse shLB1 [TRCN0000091906] or human shLB1 [TRCN0000029270]/ shCT, 1.5 µg pCMV-VSVG, 4.5 µg pMDLg/pRRE and 1.5 µg pRSV rev) and 166 µl of polyethylenimine polymer (Polysciences Inc.) that had been pre-incubated for 15 min at room temperature. At 24 hours after transfection (day 1), transfection medium was replaced with 12 ml of fresh medium. The cell-conditioned medium at days 2 and 3 was filtered with a 0.45 µm filter unit (Merck Millipore) and collected together. The mixture was stored at 4ºC during at least 30 minutes and concentrated using Lenti-X Concentrator product (#631232, Clonetech). Following manufacturer's instructions, the concentrated mixture was centrifuged at 1500 g for 45 min at 4ºC, and the pellet was resuspended in 1 ml fresh medium, aliquoted in 100 µl and stored at –80ºC.

NMuMG cells were infected with a single aliquot of the concentrated virus. Infected cells were selected with puromycin at a concentration of 1µg/ml for 48 hours.

3. Protein analysis

3.1. Total extracts

Total cell extracts from NMuMG cells were obtained in **SDS lysis buffer** (2% SDS, 50 mM Tris-HCl, 10% glycerol). Samples were kept at room temperature to avoid SDS precipitation and passed through a syringe five times to homogenize them. Protein was quantified using both the *DC* Protein Array kit (Lowry method; Bio-Rad) and Nanodrop analysis.

3.2. Western blots

Western blots were performed according standard procedures. Briefly, samples were mixed with 5× loading buffer and boiled at 95ºC during 5 minutes. Proteins were then loaded in sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) at different percentages of polyacrylamide concentration, ranging from 7.5% to 15%. Gels were run in TGS buffer, and proteins were transferred to a nitrocellulose membrane for 60 to 90 minutes (depending on the molecular weight of the protein of interest) in transfer buffer.

To ensure that the proteins were well transferred, membranes were incubated prior to blocking with a protein staining solution, Ponceau S. The solution was removed, and several washes with distilled water were performed to remove residual staining solution.

Membranes were blocked with 5% non-fat milk in TBS-T buffer for 1 hour. Primary antibodies were added to fresh blocking solution and incubated overnight at 4ºC. After three 10-minute washes with TBS-T buffer, membranes were incubated for 1 hour at room temperature (RT) with horseradish peroxidase (HRP) secondary antibodies diluted with fresh blocking solution. After further washes with TBS-T buffer, membranes were incubated with a substrate for HRP-enhanced chemiluminiscence (ECL) and exposed to autoradiography films. For proteins that were more difficult to detect, more sensitive ECL and films were used.

Western blot buffers

5× loading buffer: 250 mM Tris-HCl pH 6.8, 10% SDS, 0.02% bromophenol blue, 50% glycerol, 20% β-mercaptoethanol

TGS buffer: 25 mM Tris-OH pH 8.3, 192 mM glycine, 5% SDS

Transfer buffer: 50 mM Tris-OH, 396 mM glycine, 0.1% SDS, 20% methanol

TBS-T: 25 mM Tris-HCl pH 7.5, 137 mM NaCl, 0.1% Tween

Ponceau S: 0.5% Ponceau, 1% acetic acid

4. Immunofluorescence

4.1. Cell preparation

Cells were seeded in coverslips placed inside p60 dishes with 5 ml medium. The following day, TGFβ was added at different moments to each plate, depending on the time point treatment (e.g., at 0, 8 or 24 hours).

Cells were washed with PBS three times and fixed with 4% PFA for 15 minutes at RT. After fixation, cells were blocked for 1 hour with 1% PBS-BSA (bovine serum albumin). Cells were then incubated either at RT for 2 hours or at 4ºC overnight with the primary antibody, followed by 3 washes with PBS, and then incubated 1 hour at RT with the secondary antibody conjugated with a fluorescent dye (Alexa Fluor®). Cells were then washed three times with PBS and incubated 5 minutes with PBS-DAPI (0.25 µg/mL) to stain cell nuclei before mounting them with fluoromount.

4.2. Image acquisition

Fluorescent images were acquired with a Leica TCS SPE confocal microscope at 63x magnification using a Leica DFC300 FX camera and the Leica IM50 software.

Alternatively, GFP images corresponding to the rescue experiment were acquired with the InCell 2000 automated epifluorescence microscope at 40x magnification.

4.3. Image analysis

Image analysis was performed using ImageJ software. For quantification of emerin and lamin B1 signal, intensity was measured for both nuclear envelope signal and nuclear interior signal on each cell.

Foci number (based on staining of H3K9me3, DAPI, and HP1α) was also quantified using ImageJ software. To define the expression of HP1α foci per cell, the average intensity of pixels in the reference channel (Alexa 488) was measured. A threshold filter was used to define positive and negative HP1α foci per cell, and the same threshold was used for all experiments and all conditions. At least ten cells/condition were quantified in three independent replicates.

5. EU detection by Click reaction

NMuMG cells were seeded in 6-well plates with 2 ml of fresh medium. Cells were treated with TGFβ for 0, 8 and 24 hours. At 4 hours prior to cell collection, the uridine analogue 5-ethynyl uridine (EU) was added to plates to a 500 μM final concentration. Cells were trypsinized and collected in 15 ml falcon tubes. Paraformaldehyde (PFA) was added to a final concentration of 4% to fix cells, and the suspension was incubated at room temperature for 10 minutes in a tube rotator. Glycine was added to a final concentration of 0.125 M to stop fixation, and cells were incubated

for 5 minutes at room temperature in a tube rotator. The suspension was centrifuged 5 minutes at 700 g, and the pellet was washed twice with PBS. Cells were counted using Neubauer's chamber, and 500,000 cells were collected in Eppendorf tubes. After centrifugation at 700 g for 5 minutes, 1 ml of PBT buffer (50 ml PBS, 0.5 g BSA, 0.5 ml Triton X-100 10%) was added and incubated at room temperature for 30 minutes in a rube rotator. The suspension was then centrifuged 5 minutes at 700 g, and 1 ml of Click reaction (680 μl H20, 100 μl 1 M Tris-HCl pH 8, 20 μl 100 mM $CuSO_4$, 0.2 μl 10 mM Alexa-azide, 200 μl 0.5 M ascorbic acid) was added to each tube, and mixtures were incubated for 30 minutes at room temperature in a tube rotator protected from light. Pellets were then washed twice with PBS 1%-BSA and once with PBS-DAPI. Finally, cells were resuspended in 500 μl PBS 1%-BSA and analyzed by FACS in a LSRII instrument.

6. MTT assay

Cells that had been previously infected and selected for 48 hours with puromycin were counted with Neubauer's chamber. About 10,000 cells/condition were seeded in 96-well plates with triplicates.

MTT assays were performed by adding 0.5 mg of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT; Sigma) per ml of Dulbecco's modified Eagle's medium without FBS for 3 hours at 37ºC to determine the percentage of viable cells. Cells were solubilized with dimethyl sulfoxide-isopropanol (1:4). The

absorbance of insoluble formazan (purple) at 590 nm, which is proportional to the number of viable cells, was then determined. Cell viability was quantified during four consecutive days.

7. Cytochemical staining for senescence-associated galactosidase (SA-β-gal)

Cytochemical staining for SA-β-galactosidase was performed using a senescence β-galactosidase staining kit (Cell Signaling Technology) at pH 6.0.

NMuMG cells that had been infected and selected with puromycin were seeded in 6-well plates; after 48 hours, the assay was performed following manufacturer's instructions. In brief, culture medium was removed from the plates, and the plates were washed twice with PBS. Cells were then fixed with fixative solution (provided with the kit) for 15 min at room temperature and then washed twice with PBS prior to the addition of β-Galactosidase staining solution. Plates were sealed with parafilm and incubated overnight at 37ºC in a dry incubator. To avoid evaporation, the corners of the plates were filled with PBS.

8. Nuclear fractionation

For nuclear fractionation, NMuMG cells that had been infected and selected with puromycin were seeded in p150 plates and grown to 70% confluency. The nuclear fractionation assay was performed at 4ºC using a Nuclear Extract Kit (ab219177, Abcam) following the manufacturer's instructions. Briefly, cells were scraped in cold PBS, centrifuged for 5 minutes at 1000 rpm at 4ºC, and washed with cold PBS. To purify cytoplasm, cell pellets were resuspended in 500 μl of ice-cold cytoplasm extraction buffer, incubated on ice for 10 minutes and then centrifuged at 1000 g for 3 minutes at 4ºC; the supernatant was the cytoplasmic protein extract. Soluble nuclear proteins were extracted by resuspending cell pellets obtained before in 500 μl of ice-cold nuclear extraction buffer 1. Mixtures were incubated on ice for 15 minutes with vortexing every 5 minutes and centrifuged for 3 minutes at 5000 g at 4ºC; the supernatant was enriched for soluble nuclear proteins. To obtain insoluble nuclear proteins, cell pellets were resuspended in 500 μl of ice-cold nuclear extraction buffer 2. Samples were briefly sonicated on ice until the sample was homogenous, to shear all genomic DNA. This fraction contained all remaining nuclear proteins that were not soluble.

Each sample was quantified using Pierce™ BCA Protein Assay Kit and analyzed by western blot.

9. Micrococcal digestion

NMuMG cells that had been infected and selected with puromycin were seeded in p150 plates and treated with TGFβ. Cells were counted using a Neubaure's chamber, and 4×10^6 cells per condition were used for the assay. Cell pellets were lysed in 500 µl of buffer A for 10 minutes at 4ºC. After adding NP-40 to a final concentration of 0.2% (v/v) and incubation for 10 minutes at 4ºC, the lysate was centrifuged at 1200 rpm for 10 minutes at 4ºC. The resulting pellet was resuspended in 100 µl of buffer A (containing $CaCl_2$ to a final concentration of 10 mM and MNase (0.08 units) for 10 minutes. The enzyme was inactivated with 50 mM EDTA. Finally, pellets were treated with RNase A for 2 minutes at room temperature prior to proteinase K treatment for 10 minutes at 56ºC. DNA was purified using GFX™ PCR DNA Purification Kit (GE Healthcare), and digestion products were analyzed by agarose gel electrophoresis. Total DNA was visualized by SYBR® safe staining.

Buffer A: 10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$ and 0.3 M sacarose, supplemented with protease inhibitors

10. Salt extraction

Salt extraction experiments were performed as previously described[231]. Briefly, NMuMG cells were washed in PBS, harvested, dounced in buffer A, layered over a cushion of high-sucrose buffer A (30% sucrose) and centrifuged for 15 minutes at 3000 g. Pelleted nuclei were resuspended in buffer B and incubated with 500 mM NaCl at 4ºC for 30 minutes. Supernatants were analyzed by Western blot.

Salt extraction buffers

Buffer A: 0.32 M sucrose, 15 mM HEPES pH 7.9, 60 mM KCl, 2 mM EDTA, 0.5 mM EGTA, 0.5% BSA, 0.5 mM spermidine, 0.15 mM spermidine, 0.5 mM DTT

Buffer B: 15 mM HEPES pH 7.9, 60 mM KCl, 15 mM NaCl, 0.34 mM sucrose, 10% glycerol

11. Rescue experiment

NMuMG-infected cells were seeded 48 hours after selection for transfection with Lipofectamine 2000 reagent (Invitrogen) with either GFP empty vector or human mPA-GFP-LaminB1-10 (a gift from Michael Davidson; originally purchased from Addgene plasmid #57141). The transfection medium was replaced with fresh medium after 24 hours (day 1). At 48 hours after transfection (day 2), cells were selected with G418 sulfate (Merk Millipore) for 10

days at a final concentration of 0.5 µg/µl, together with puromycin (1 µg/ml) to maintain the levels of LB1 shRNA.

12. Migration and invasion assays

For migration experiments, 50,000 cells were resuspended in DMEM 0.1% FBS-0.1% BSA, reseeded on a transwell filter chamber (Costar 3422) and incubated for 6 to 8 hours. For invasion assays, cells were placed in Matrigel-coated transwell filter (BD356234) and incubated for 12 to 16 hours. In both cases, DMEM with 10% FBS was added to the lower chamber and used as a chemoattractant. Non-migrating and non-invading cells were removed from the upper surface of the membrane, while cells that adhered to the lower surface were fixed with PFA 4% for 15 min. Nuclei were stained with PBS-DAPI (0.25 µg/mL). DAPI-stained nuclei were counted in four fields per filter by ImageJ software.

13. Fluorescence recovery after photobleaching (FRAP) assay

NMuMG cells were first transfected with human mCerulean-LaminB1-10, which was a gift from Michael Davidson (Addgene plasmid #55380)[232] using *Trans*IT-X2® Dynamic Delivery System (Mirus Bio LLC). After 8 hours, transfected cells were infected using irrelevant shRNA and human shRNA LB1 and selected using

puromycin (1 µg/ml) for 24 hours. Transfected and infected cells were seeded in 35-mm MatTek dishes. After 24 hours, Leica TCS SP5 confocal system with the LAS-AF application Wizard was used to perform FRAP experiment. Cells were kept in a fully incubated (CO_2 and 37ºC) chamber while imaging with a 63×1.4 objective and 458 nm laser line of the Argon laser for excitation. Selected nuclear areas were bleached five times using maximum laser intensity, and 100 frames after the photobleach were collected in 370 ms intervals, using the minimal laser power required. Prior to photobleaching, 10 frames were collected as an internal control of the experiment. Signal recovery was measured by ImageJ software using pre-bleached (pb), background (bg), and non-bleached (nb) areas (ROIs) to normalize data. For every time point, data were normalized according to the following formula: $(ROI_b - ROI_{bg})/(ROI_{nb} - ROI_{bg})$ / $(pbRO_{lb} - pbROI_{bg})/(pbROI_{nb} - (pbROI_{bg})$[233].

14. Mouse embryonic stem cell experiments

E14TG2a mouse embryonic stem cells (ESCs) were kept under feeder-free, 20% serum conditions as previously described[234]. Cells were infected with lentivirus produced in 293T to generate stable cell lines expressing shRNA against lamin B1 or a control shRNA. Cells were selected with puromycin (2 µg/ml) for three days. For neuroectodermal specification, lamin B1 knockdown and control cells were plated in serum/LIF condition (100,000 cells per 100×20 NUNC dish). On the following day, the medium was

changed with serum-free medium supplemented with N2 and B27. Samples were collected after five days of differentiation.

15. RNA analysis

15.1. Phenol-chloroform RNA extraction

To extract RNA, cells were seeded in p60 plates and washed three times with PBS. Cells were lysed with 800 µl of TRIzol® reagent (Invitrogen). The lysate was vortexed, and 200 µl of RNase-free chloroform was added, mixed and incubated at room temperature for 2 minutes. The solution was centrifuged at 13,000 rpm for 15 minutes at 4ºC, and the supernatant was then mixed with 500 µl RNase-free isopropanol and incubated at RT for 10 minutes. RNA was precipitated by centrifugation at 13,000 rpm for 10 minutes at 4ºC. Pellets were washed once with 1 ml of RNase-free ethanol and centrifuged at 7,500 rpm for 5 minutes at 4ºC. After evaporation of ethanol traces in a dry bath at 60ºC, the pellet was resuspended in DEPC-water and dissolved for 10 minutes at 60ºC prior to quantification by Nanodrop.

15.2. Quantitative RT-PCR

RNA was retrotranscribed with the Transcriptor First-Strand cDNA Synthesis kit (Roche) following manufacturer's instructions. Analyses were carried out in triplicates with 25 ng of cDNA using

Light Cycler PCR machine (Roche). Primers used for quantitative RT-PCR are listed in *Table MM.2.*

15.3. RNA extraction for sequencing

RNA-seq experiments were performed with two biological replicates of NMuMG cells (with or without shRNA of either control or KD) that were untreated or treated with TGF-β for 8 or 24 hours. After RNA extraction with GenElute™ Mammalian Total RNA Miniprep Kit (Sigma-Aldrich), samples were sequenced using the Illumina HiSeq 2500 system.

15.4. Analysis of RNA-seq data

RNA-seq samples were mapped against the mm9 mouse genome assembly using TopHat[235] with the option –g 1 to discard those reads that could not be uniquely mapped in just one region. Cufflinks and Cuffdiff were run to quantify the expression in FPKMs of each annotated transcript in RefSeq and to identify the list of differentially expressed genes for each case[236].

16. Chromatin immunoprecipitation (ChIP)

ChIP experiments were performed as described[179]. In brief, NMuMG cells were crosslinked in 1% formaldehyde for 5–10 minutes at 37ºC. Crosslinking was stopped by adding glycine to a final concentration of 0.125 M for 2 minutes at room temperature. For nuclear fractions, cells were scraped with cold soft-lysis buffer supplemented with protease inhibitors. Samples were then centrifuged at 3,000 rpm for 15 minutes, and nuclei pellets were lysed with SDS-lysis buffer supplemented with protease inhibitors. Extracts were sonicated to generate 200–600 bp DNA fragments, incubated on ice for 20 minutes, centrifuged at 13,000 rpm for 10 minutes, and then diluted 1:10 with dilution buffer. The primary antibody or an irrelevant antibody (IgG) was added to the sample, and the mixture was incubated overnight with rotation at 4ºC. Chromatin bound to the antibody was then immunoprecipitated using unblocked protein A beads (40 µl/sample, Diagenode) for 3 hours with rotation at 4ºC. Precipitated samples were then washed three times with low salt buffer, three times with high salt buffer, and then twice with LiCl buffer, using columns.

ChIP buffers

Soft-lysis buffer: 50 mM Tris-HCl, 10 mM EDTA, 0.1% NP-40, and 10% glycerol, supplemented with protease inhibitors

SDS lysis buffer: 1% SDS, 10 mM EDTA, and 50 mM Tris pH 8, supplemented with protease inhibitors

Dilution buffer: 0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris pH 8, 167 mM NaCl

Low salt buffer: 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.0, 150 mM NaCl

High salt buffer: 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.0, 500 mM NaCl

LiCl buffer: 250 mM LiCl, 1% Nonidet P-40, 1% sodium deoxycholate, 1 mM EDTA, 10 mM Tris-HCl pH 8.0

16.1. ChIP-Western blot

To verify antibody specificity, washed samples were eluted with 2× Western blot loading buffer for 5 minutes at 95ºC. Proteins were analyzed by Western blot using anti-lamin B1 antibody.

16.2. ChIP-qPCR

For qPCR detection of genomic regions, washed samples were treated with elution buffer (100 mM Na_2CO_3 and 1% SDS) for 1 hour at 37ºC and then incubated at 65ºC overnight with the addition of a final concentration of 200 mM NaCl to reverse the formaldehyde crosslinking. After proteinase K solution (0.4 mg/ml proteinase K [Roche], 50 mM EDTA, 200 mM Tris-HCl pH 6.5) treatment for 1 hour at 55ºC, DNA was purified with MinElute PCR purification kit (Qiagen) and eluted in nuclease-free water. Genomic regions were detected by qPCR. Primers used are listed in *Table MM.*2. Results were quantified relative to the input and the amount of irrelevant IgG immunoprecipitated in each condition.

16.3. ChIP-seq

For ChIP sequencing (ChIP-seq) analysis, two parallel ChIPs were performed and mixed after elution with nuclease-free water. The NEBNext Ultra DNA library Prep Kit for Illumina was used to prepare the libraries, and samples were sequenced using Illumina HiSeq 2500 system.

16.4. Analysis of ChIP-seq data

ChIP-seq samples were mapped against the mm9 mouse genome assembly using BowTie with the option –m 1 to discard those reads that could not be uniquely mapped to just one region[237]. MACS was run with the default parameters but with the shift-size adjusted to 100 bp to perform the peak calling against the corresponding control sample[238]. The genome distribution of each set of peaks was calculated by counting the number of peaks fitted on each class of region according to RefSeq annotations. Distal region was the region within 2.5 Kbp and 0.5 Kbp upstream of the transcription start site (TSS). Proximal region was the region within 0.5 Kbp of the TSS. UTR, untranslated region; CDS, protein coding sequence; intronic regions, introns; and the rest of the genome, intergenic. Peaks that overlapped with more than one genomic feature were proportionally counted the same number of times. Pie charts were generated by calculating the genome distribution of all features in the full genome, and the R caroline package was used to combine the piechart of each set of peaks with the full genome distribution[239]. Each set of target genes was retrieved by matching the ChIP-seq peaks in the region 2.5 Kbp upstream of the TSS until

the end of the transcripts as annotated in RefSeq. Reports of functional enrichments of gene ontology and ENCODE ChIP-seq histone marks categories were generated using the EnrichR tool[240] and the Molecular Signatures Database (MSigDB)[241]. Plots showing the average distribution of ChIP-seq reads 2 Kbp around the TSS of each target gene were generated by counting the number of reads for each region according to RefSeq and then averaging the values for the total number of mapped reads of each sample and the total number of genes in the particular gene set. The heatmaps displaying the density of ChIP-seq reads 5 Kb around the TSS of each target gene set were generated by counting the number of reads in this region for each individual gene and normalizing this value with the total number of mapped reads of the sample. Genes on each ChIP heatmap were ranked by the logarithm of the average number of reads in the same genomic region. To generate the lists of eLADs, the signal of the corresponding IgG was subtracted from each LB1 ChIP-seq sample to score each bin of 100 bps along the genome in terms of normalized reads. Next, a threshold of 1 unit was set to filter out those bins without peaks. Finally, two or more individual peaks of LB1 were clustered at a distance less than 0.25 Mb. The UCSC genome browser was used to generate the screenshots of each group of experiments along the thesis[242].

17. ATAC sequencing

The ATAC experiment was performed as described[185]. Briefly, NMuMG cells were either untreated or treated with TGF-β for 8 or 24 hours and then harvested and treated with transposase Tn5 (Nextera DNA Library Preparation Kit, Illumina). DNA was purified using MinElute PCR Purification Kit (Qiagen). All samples were then amplified by PCR using NEBNextHigh-Fidelity 2× PCR Master Mix (New Englands Labs) with primers containing a barcode to generate the libraries. DNA was again purified using MinElute PCR Purification kit, and samples were sequenced using Illumina HiSeq 2500 system. Primers used to generate the libraries were:

Sample	Forward	Reverse
0h	5′AATGATACGGCGACCACCGAGATCTACACTCGTCGGCAGCGTCAGATGTG3′	5′CAAGCAGAAGACGGCATACGAGATTCGCCTTAGTCTCGTGGGCTCGGAGATGT3′
8h	5′AATGATACGGCGACCACCGAGATCTACACTCGTCGGCAGCGTCAGATGTG3′	5′CAAGCAGAAGACGGCATACGAGATCTAGTACGGTCTCGTGGGCTCGGAGATGT3′
24h	5′AATGATACGGCGACCACCGAGATCTACACTCGTCGGCAGCGTCAGATGTG3′	5′CAAGCAGAAGACGGCATACGAGATTTCTGCCTGTCTCGTGGGCTCGGAGATGT3′

17.1. Analysis of ATAC-seq data

The ATAC-seq samples were mapped against the mm9 mouse genome assembly using BowTie with the option –m 1 to discard those reads that could not be uniquely mapped to just one region, and with the option –X 2000 to define the maximum insert size for paired-end alignment[237]. MACS was run with the default parameters but with the shift size adjusted to 100 bp to perform peak calling[238]. Each set of target genes was retrieved by matching the ChIP-seq peaks in the region 2.5 Kb upstream of the TSS until the end of the transcripts as annotated in RefSeq. Reports of functional enrichments of gene ontology and ENCODE ChIP-seq histone marks categories were generated with the EnrichR tool[240].

18. Hi-C

Hi-C libraries were generated from NMuMG cells (that were untreated or treated with TGF-β for 8 or 24 hours) according the previously published Hi-C protocol[147], with minor adaptations. Five million cells were crosslinked with 1% formaldehyde for 10 minutes at room temperature. Prior to permeabilization, cells were treated for 5 minutes with trypsin to obtain single cells. DNA was digested with 400 units of Dpn II, and the ends of restriction fragments were labelled using biotinylated nucleotides and ligated in a small volume (*in situ* Hi-C). Libraries were generated independently in the three conditions (e.g. treated with TGF-β for 8 or 24 hours or

untreated), controlled for quality and sequenced on an Illumina HiSeq 2000 sequencer.

18.1. Analysis of Hi-C data

Hi-C data was processed using TADbit[243] for read quality control, read mapping, interaction detection, interaction filtering and matrix normalization. First, reads were checked using an implemented FastQC protocol (http://www.bioinformatics.babraham.ac.uk/projects/fastqc) in TADbit. This allowed problematic samples to be discarded and systematic artefacts to be detected. Next, a fragment-based strategy in TADbit was used to map the remaining reads to the reference mouse genome (reference mm9). The mapping strategy resulted in about 85% of reads mapped uniquely to the genome (Table A.3). Non-informative contacts between two reads, including self-circles, dangling-ends, errors, random breaks, or duplicates, were filtered as previously described[243]. The final interaction matrices resulted in 89–106 M of valid interactions per time point (Table A.3). These valid interactions were then used to generate genome-wide interaction maps at 100, 40 and 5 Kb to segment the genome into the so-called A/B compartments or topologically associating domains (TADs) or to perform a meta-analysis, respectively[140,244]. A/B compartments were calculated using normalized and decay-corrected matrices[141] by calculating the first component of a PCA of chromosome-wide matrices of the Pearson product-moment internal correlation as implemented in HOMMER[245]. TADs were identified using 100 Kb resolution normalized and decay-corrected matrices as input to a series of

scripts from the Dekker lab with default parameters[246]. TAD border localization and strength was calculated and used to identify conserved borders and their strengths. A border was considered conserved if it was localized in the same bin in the three experiments. Finally, to assess whether particular parts of the Hi-C interaction matrices had common structural features, a meta-analysis of the region was performed by merging individual local submatrices at 5 Kb resolution in similar fashion as previously published[247]. 4C-like profiles centered at the TSS of several selected genes were extracted from the meta-matrices and compared between untreated and treated samples. The resulting differential 4C-like profiles were calculated by subtracting the treated from the untreated interaction counts, which were represented as line plots or heatmaps.

Accession number

Sequencing samples (raw data and processed files) are available at NCBI GEO under the accession code GSE96033.

19. Antibodies used

Protein	Species	Provider	Reference	Dilution
Lamin B1	Rabbit	Abcam	Ab16048	WB 1:1000 IF 1:1000 ChIP 10µg/500µg chromatin
GFP	Rabbit	Abcam	Ab 6556	WB 1:2000 IF 1:2000
Tubulin	Mouse	Sigma	T9026	IF 1:50000
Emerin	Mouse	Leica	CE-S	IF 1:50
Fibronectin	Rabbit	Dako	A0245	WB 1:5000
E-cadherin	Mouse	BD Transduction	610182	WB 1:2000
N-cadherin	Mouse	BD Transduction	610920	WB 1:1000
TBP	Rabbit	Abcam	Ab63766	WB 1:2000
H3	Rabbit	Abcam	Ab1791	WB 1:10000
HP1	Rabbit	Sigma	H2164	WB 1:2000

Table MM.1. Antibodies used and their applications. Antibodies used in this thesis, their commercial information and dilutions used for each application are listed.

20. Primers used

Primer	Direction	Sequence	Use
LAD region	Forward	CAAGCTGCACTGGGACAAAG	ChIP-qPCR
LAD region	Reverse	CAAATGTATGGTGTCCTGAAGGTT	
LAD 1	Forward	TCCATGGGTGACAGGGAC	ChIP-qPCR
LAD 1	Reverse	TCTTTGGGCATCATTTGCTT	
LAD 2	Forward	CCAAAGCTGTTCAGTGAGAGG	ChIP-qPCR
LAD 2	Reverse	CGATAGGGAAGACAGGAGACAC	
LAD 3	Forward	CAAGCTGCACTGGGACAAAG	ChIP-qPCR
LAD 3	Reverse	CAAATGTATGGTGTCCTGAAGGTT	
eLAD 1	Forward	TGCTCTCTCCCTTTGGACC	ChIP-qPCR
eLAD 1	Reverse	CGGGGGTAGGGCATCATAT	
eLAD 2	Forward	GACGTCTTGTGACCGGGTT	ChIP-qPCR
eLAD 2	Reverse	CGGCAGCAGTAGGAGCAGT	
eLAD 3	Forward	GCGCCCGTCGTCCTTCTCGTC	ChIP-qPCR
eLAD 3	Reverse	CTTCCGCGACTGGGGGTCCT	
Fibronectin	Forward	TGAGCATCTTGAGTGGATGG	ChIP-qPCR
Fibronectin	Reverse	GTGTGAGCCGGACAACTTCT	
Zeb2	Forward	GGGCCTCTTCTTACCGTTTT	ChIP-qPCR
Zeb2	Reverse	CGCTGTGTTTGGTTGCTAGA	
Twist2	Forward	GCCTCGAAATCAGAGCCTTT	ChIP-qPCR
Twist2	Reverse	TCCAGCTCTTCCTCACTGGT	
prOct4	Forward	ACCAACCTGGACAACACAAGATG	ChIP-qPCR
prOct4	Reverse	GCTTACCCACCCGTCTAGAGTCC	

Lamin B1	Forward	CTGCTGCTCAATTATGCCAAGAAG	qRT-PCR
Lamin B1	Reverse	GGCAGATAAGGATGCTTCTAGCT	
Pumilio	Forward	CGGTCGTCCTGAGGATAAAA	qRT-PCR
Pumilio	Reverse	CGTACGTGAGGCGTGAGTAA	

Table MM.2. Mouse primers used for mRNA and ChIP analysis. The sequence of all primers used in this thesis are listed (given as 5′-to-3′).

ANNEX

1. Lamin B1 ChIP sequencing data

Sample Name	Number of Peaks	Number of Genes
0 h	4645	4454
8 h	10484	8755
24 h	7083	6595

Table A.1. Number of lamin B1 positive regions and genes. List of peaks and genes identified after peak calling analysis of the lamin B1 ChIP-seq in untreated or treated (for 8 or 24 hours) NMuMG cells.

Sample Name	Number of eLADs	Average size (Mb)	Coverage (%)
0 h	2051	0.33	25.9
8 h	2429	0.34	31.7
24 h	2949	0.36	40.1

Table A.2. Characterization of eLADs. List of eLADs detected after clustering lamin B1 peaks, the average size (in Mb) of representative eLADs extracted from the size of mm9 mouse genome and the coverage of the mouse genome for each set of eLADs in untreated NMuMG cells or in cells treated with TGFβ for 8 or 24 hours.

1. Hi-C sequencing data

Sample and read Name	Number of reads	Number of mapped reads	Number of mapped interactions	Number of valid interactions
0 h read 1	211679877	185780142		
0 h read 2	211679877	179013420	163087664	88644054
8 h read 1	196125007	171021097		
8 h read 2	196125007	166820416	152109607	105589112
24 h read 1	182595645	159768798		
24 h read 2	182595645	152521578	138758327	90046184

Table A.3. Reads obtained with the Hi-C sequencing. Total number of reads and mapped reads in untreated and treated NMuMG cells per each per each pair of reads. Mapped interactions and valid ones in every time point are also listed.

Sample Name	A compartment	eLADs	A +eLADs	Increment (%)	Expressed Genes
0 h	21534	15822	13049	0	7972
8 h	21660	19273	13774	6	8493
24 h	23463	24220	15144	16	8170

Table A.4. New eLADs appearing in A compartment. Number of bins in the A compartment, in eLADs and increment in number and in percentage of eLADs bins appearing in the A compartment as well as number of expressed genes in these bins.

Sample Name	B compartment	eLADs	B +eLADs	Increment (%)	Expressed Genes
0 h	23720	15822	1418	0	189
8 h	23839	19273	3714	162	271
24 h	22872	24220	7123	402	391

Table A.5. New eLADs appearing in B compartment. Number of bins in the B compartment, in eLADs and increment in number and in percentage of eLADs bins appearing in the B compartment as well as number of expressed genes in these bins.

REFERENCES

1 Ou, H. D. *et al.* ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells. *Science (New York, N.Y.)* **357**, doi:10.1126/science.aag0025 (2017).

2 Lawrence, M., Daujat, S. & Schneider, R. Lateral Thinking: How Histone Modifications Regulate Gene Expression. *Trends in genetics : TIG* **32**, 42-56, doi:10.1016/j.tig.2015.10.007 (2016).

3 Collas, P., Lund, E. G. & Oldenburg, A. R. Closing the (nuclear) envelope on the genome: how nuclear lamins interact with promoters and modulate gene expression. *BioEssays : news and reviews in molecular, cellular and developmental biology* **36**, 75-83, doi:10.1002/bies.201300138 (2014).

4 Aroeira, L. S. *et al.* Epithelial to mesenchymal transition and peritoneal membrane failure in peritoneal dialysis patients: pathologic significance and potential therapeutic interventions. *Journal of the American Society of Nephrology : JASN* **18**, 2004-2013, doi:10.1681/asn.2006111292 (2007).

5 Ong, C. T. & Corces, V. G. CTCF: an architectural protein bridging genome topology and function. *Nature reviews. Genetics* **15**, 234-246, doi:10.1038/nrg3663 (2014).

6 Millanes-Romero, A. *et al.* Regulation of heterochromatin transcription by Snail1/LOXL2 during epithelial-to-

mesenchymal transition. *Mol Cell* **52**, 746-757, doi:10.1016/j.molcel.2013.10.015 (2013).

7 Bickmore, W. A. The spatial organization of the human genome. *Annual review of genomics and human genetics* **14**, 67-84, doi:10.1146/annurev-genom-091212-153515 (2013).

8 Dekker, J. & Misteli, T. Long-Range Chromatin Interactions. *Cold Spring Harb Perspect Biol* **7**, a019356, doi:10.1101/cshperspect.a019356 (2015).

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

10 Gruenbaum, Y. & Foisner, R. Lamins: nuclear intermediate filament proteins with fundamental functions in nuclear mechanics and genome regulation. *Annu Rev Biochem* **84**, 131-164, doi:10.1146/annurev-biochem-060614-034115 (2015).

11 Denker, A. & de Laat, W. The second decade of 3C technologies: detailed insights into nuclear organization. *Genes Dev* **30**, 1357-1382, doi:10.1101/gad.281964.116 (2016).

12 Gonzalez-Sandoval, A. & Gasser, S. M. On TADs and LADs: Spatial Control Over Gene Expression. *Trends Genet* **32**, 485-495, doi:10.1016/j.tig.2016.05.004 (2016).

13 Guenatri, M., Bailly, D., Maison, C. & Almouzni, G. Mouse centric and pericentric satellite repeats form distinct

functional heterochromatin. *The Journal of cell biology* **166**, 493-505, doi:10.1083/jcb.200403109 (2004).

14 Probst, A. V., Dunleavy, E. & Almouzni, G. Epigenetic inheritance during the cell cycle. *Nat Rev Mol Cell Biol* **10**, 192-206, doi:10.1038/nrm2640 (2009).

15 Wang, X. Q. D. & Dostie, J. Chromosome folding and its regulation in health and disease. *Current opinion in genetics & development* **43**, 23-30, doi:10.1016/j.gde.2016.10.006 (2017).

16 McGinty, R. K. & Tan, S. Nucleosome structure and function. *Chemical reviews* **115**, 2255-2273, doi:10.1021/cr500373h (2015).

17 Grewal, S. I. & Elgin, S. C. Transcription and RNA interference in the formation of heterochromatin. *Nature* **447**, 399-406, doi:10.1038/nature05914 (2007).

18 Musselman, C. A., Lalonde, M. E., Cote, J. & Kutateladze, T. G. Perceiving the epigenetic landscape through histone readers. *Nature structural & molecular biology* **19**, 1218-1227, doi:10.1038/nsmb.2436 (2012).

19 Grande, M. T. & Lopez-Novoa, J. M. Fibroblast activation and myofibroblast generation in obstructive nephropathy. *Nature reviews. Nephrology* **5**, 319-328, doi:10.1038/nrneph.2009.74 (2009).

20 Maison, C. & Almouzni, G. HP1 and the dynamics of heterochromatin maintenance. *Nature reviews. Molecular cell biology* **5**, 296-304, doi:10.1038/nrm1355 (2004).

21 Maeshima, K., Hihara, S. & Eltsov, M. Chromatin structure: does the 30-nm fibre exist in vivo? *Current opinion in cell biology* **22**, 291-297, doi:10.1016/j.ceb.2010.03.001 (2010).

22 Zeisberg, M. & Neilson, E. G. Biomarkers for epithelial-mesenchymal transitions. *The Journal of clinical investigation* **119**, 1429-1437, doi:10.1172/jci36183 (2009).

23 Wiles, E. T. & Selker, E. U. H3K27 methylation: a promiscuous repressive chromatin mark. *Current opinion in genetics & development* **43**, 31-37, doi:10.1016/j.gde.2016.11.001 (2017).

24 Dechat, T. *et al.* Nuclear lamins: major factors in the structural organization and function of the nucleus and chromatin. *Genes Dev* **22**, 832-853, doi:10.1101/gad.1652708 (2008).

25 Luperchio, T. R., Wong, X. & Reddy, K. L. Genome regulation at the peripheral zone: lamina associated domains in development and disease. *Current Opin Genetics & Dev* **25**, 50-61, doi:10.1016/j.gde.2013.11.021 (2014).

26 Arents, G., Burlingame, R. W., Wang, B. C., Love, W. E. & Moudrianakis, E. N. The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-

handed superhelix. *Proc Natl Acad Sci U S A* **88**, 10148-10152 (1991).

27 Olins, A. L. & Olins, D. E. Spheroid chromatin units (v bodies). *Science (New York, N.Y.)* **183**, 330-332 (1974).

28 Richmond, T. J., Finch, J. T., Rushton, B., Rhodes, D. & Klug, A. Structure of the nucleosome core particle at 7 A resolution. *Nature* **311**, 532-537 (1984).

29 Arents, G. & Moudrianakis, E. N. Topography of the histone octamer surface: repeating structural motifs utilized in the docking of nucleosomal DNA. *Proc Natl Acad Sci USA* **90**, 10489-10493 (1993).

30 Cutter, A. R. & Hayes, J. J. A brief review of nucleosome structure. *FEBS letters* **589**, 2914-2922, doi:10.1016/j.febslet.2015.05.016 (2015).

31 Pepenella, S., Murphy, K. J. & Hayes, J. J. Intra- and inter-nucleosome interactions of the core histone tail domains in higher-order chromatin structure. *Chromosoma* **123**, 3-13, doi:10.1007/s00412-013-0435-8 (2014).

32 Syed, S. H. *et al.* Single-base resolution mapping of H1-nucleosome interactions and 3D organization of the nucleosome. *Proc Natl Acad Sci USA* **107**, 9620-9625, doi:10.1073/pnas.1000309107 (2010).

33 Allan, J., Hartman, P. G., Crane-Robinson, C. & Aviles, F. X. The structure of histone H1 and its location in chromatin. *Nature* **288**, 675-679 (1980).

34 Caterino, T. L., Fang, H. & Hayes, J. J. Nucleosome linker DNA contacts and induces specific folding of the intrinsically disordered H1 carboxyl-terminal domain. *Mol Cell Biol* **31**, 2341-2348, doi:10.1128/mcb.05145-11 (2011).

35 Whitlock, J. P., Jr. & Simpson, R. T. Removal of histone H1 exposes a fifty base pair DNA segment between nucleosomes. *Biochemistry* **15**, 3307-3314 (1976).

36 Ozer, G., Luque, A. & Schlick, T. The chromatin fiber: multiscale problems and approaches. *Current Opin Struct Biol* **31**, 124-139, doi:10.1016/j.sbi.2015.04.002 (2015).

37 Hubner, M. R., Eckersley-Maslin, M. A. & Spector, D. L. Chromatin organization and transcriptional regulation. *Curr Opin Genet Dev* **23**, 89-95, doi:10.1016/j.gde.2012.11.006 (2013).

38 Nishino, Y. *et al.* Human mitotic chromosomes consist predominantly of irregularly folded nucleosome fibres without a 30-nm chromatin structure. *EMBO J* **31**, 1644-1653, doi:10.1038/emboj.2012.35 (2012).

39 Moore, L. D., Le, T. & Fan, G. DNA methylation and its basic function. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology* **38**, 23-38, doi:10.1038/npp.2012.112 (2013).

40 Chen, Z., Li, S., Subramaniam, S., Shyy, J. Y. & Chien, S. Epigenetic Regulation: A New Frontier for Biomedical

Engineers. *Annu Rev Biomed Eng* **19**, 195-219, doi:10.1146/annurev-bioeng-071516-044720 (2017).

41 Peschansky, V. J. & Wahlestedt, C. Non-coding RNAs as direct and indirect modulators of epigenetic regulation. *Epigenetics* **9**, 3-12, doi:10.4161/epi.27473 (2014).

42 Harr, J. C., Gonzalez-Sandoval, A. & Gasser, S. M. Histones and histone modifications in perinuclear chromatin anchoring: from yeast to man. *EMBO Rep* **17**, 139-155, doi:10.15252/embr.201541809 (2016).

43 Narlikar, G. J., Sundaramoorthy, R. & Owen-Hughes, T. Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes. *Cell* **154**, 490-503, doi:10.1016/j.cell.2013.07.011 (2013).

44 Zentner, G. E. & Henikoff, S. Regulation of nucleosome dynamics by histone modifications. *Nat Struct & Mol Bio* **20**, 259-266, doi:10.1038/nsmb.2470 (2013).

45 Cosgrove, M. S., Boeke, J. D. & Wolberger, C. Regulated nucleosome mobility and the histone code. *Nat Struct & Mol Biol* **11**, 1037-1043, doi:10.1038/nsmb851 (2004).

46 Mersfelder, E. L. & Parthun, M. R. The tale beyond the tail: histone core domain modifications and the regulation of chromatin structure. *NAR* **34**, 2653-2662, doi:10.1093/nar/gkl338 (2006).

47 Garcia, B. A., Shabanowitz, J. & Hunt, D. F. Characterization of histones and their post-translational

modifications by mass spectrometry. *Current opinion in chemical biology* **11**, 66-73, doi:10.1016/j.cbpa.2006.11.022 (2007).

48 Kouzarides, T. Chromatin modifications and their function. *Cell* **128**, 693-705, doi:S0092-8674(07)00184-5 [pii] 10.1016/j.cell.2007.02.005 (2007).

49 Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell research* **21**, 381-395, doi:10.1038/cr.2011.22 (2011).

50 Tan, M. *et al.* Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* **146**, 1016-1028, doi:10.1016/j.cell.2011.08.008 (2011).

51 Heitz, E. *Das Heterochromatin der Moose I*. Vol. 69 (1927).

52 Sun, F. L., Cuaycong, M. H. & Elgin, S. C. Long-range nucleosome ordering is associated with gene silencing in Drosophila melanogaster pericentric heterochromatin. *Molecular and cellular biology* **21**, 2867-2879, doi:10.1128/mcb.21.8.2867-2879.2001 (2001).

53 Turner, B. M. Histone acetylation and an epigenetic code. *BioEssays : news and reviews in molecular, cellular and developmental biology* **22**, 836-845, doi:10.1002/1521-1878(200009)22:9<836::aid-bies9>3.0.co;2-x (2000).

54 Almouzni, G. & Probst, A. V. Heterochromatin maintenance and establishment: lessons from the mouse pericentromere. *Nucleus* **2**, 332-338, doi:10.4161/nucl.2.5.17707 (2011).

55 Huisinga, K. L., Brower-Toland, B. & Elgin, S. C. The contradictory definitions of heterochromatin: transcription and silencing. *Chromosoma* **115**, 110-122, doi:10.1007/s00412-006-0052-x (2006).

56 Grunstein, M. Yeast heterochromatin: regulation of its assembly and inheritance by histones. *Cell* **93**, 325-328 (1998).

57 Noma, K., Allis, C. D. & Grewal, S. I. Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. *Science* **293**, 1150-1155, doi:10.1126/science.1064150293/5532/1150 [pii] (2001).

58 Schubeler, D. *et al.* The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev* **18**, 1263-1271, doi:10.1101/gad.1198204 (2004).

59 Jones, B. *et al.* The histone H3K79 methyltransferase Dot1L is essential for mammalian development and heterochromatin structure. *PLoS genetics* **4**, e1000190, doi:10.1371/journal.pgen.1000190 (2008).

60 Tschiersch, B. *et al.* The protein encoded by the Drosophila position-effect variegation suppressor gene Su(var)3-9 combines domains of antagonistic regulators of homeotic gene complexes. *The EMBO journal* **13**, 3822-3831 (1994).

61 Jenuwein, T. & Allis, C. D. Translating the histone code. *Science* **293**, 1074-1080, doi:10.1126/science.1063127293/5532/1074 [pii] (2001).

62 Saksouk, N., Simboeck, E. & Dejardin, J. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics Chromatin* **8**, 3, doi:10.1186/1756-8935-8-3 (2015).

63 Le, H. D., Donaldson, K. M., Cook, K. R. & Karpen, G. H. A high proportion of genes involved in position effect variegation also affect chromosome inheritance. *Chromosoma* **112**, 269-276, doi:10.1007/s00412-003-0272-2 (2004).

64 Eymery, A., Callanan, M. & Vourc'h, C. The secret message of heterochromatin: new insights into the mechanisms and function of centromeric and pericentric repeat sequence transcription. *The International journal of developmental biology* **53**, 259-268, doi:10.1387/ijdb.082673ae (2009).

65 Schueler, M. G. & Sullivan, B. A. Structural and functional dynamics of human centromeric chromatin. *Annual review of genomics and human genetics* **7**, 301-313, doi:10.1146/annurev.genom.7.080505.115613 (2006).

66 Trojer, P. & Reinberg, D. Facultative heterochromatin: is there a distinctive molecular signature? *Mol Cell* **28**, 1-13, doi:10.1016/j.molcel.2007.09.011 (2007).

67 Becker, J. S., Nicetto, D. & Zaret, K. S. H3K9me3-Dependent Heterochromatin: Barrier to Cell Fate Changes. *Trends in genetics : TIG* **32**, 29-41, doi:10.1016/j.tig.2015.11.001 (2016).

68 Nakayama, J., Rice, J. C., Strahl, B. D., Allis, C. D. & Grewal, S. I. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science* **292**, 110-113, doi:10.1126/science.10601181060118 [pii] (2001).

69 Schultz, D. C., Ayyanathan, K., Negorev, D., Maul, G. G. & Rauscher, F. J., 3rd. SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev* **16**, 919-932, doi:10.1101/gad.973302 (2002).

70 Tachibana, M. *et al.* Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. *Genes Dev* **19**, 815-826, doi:10.1101/gad.1284005 (2005).

71 Bannister, A. J. *et al.* Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120-124, doi:10.1038/35065138 35065138 [pii] (2001).

72 Lee, T. I. *et al.* Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* **125**, 301-313, doi:10.1016/j.cell.2006.02.043 (2006).

73 Breiling, A., Turner, B. M., Bianchi, M. E. & Orlando, V. General transcription factors bind promoters repressed by Polycomb group proteins. *Nature* **412**, 651-655, doi:10.1038/35088090 (2001).

74 Soufi, A., Donahue, G. & Zaret, K. S. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* **151**, 994-1004, doi:10.1016/j.cell.2012.09.045 (2012).

75 Vissel, B. & Choo, K. H. Mouse major (gamma) satellite DNA is highly conserved and organized into extremely long tandem arrays: implications for recombination between nonhomologous chromosomes. *Genomics* **5**, 407-414 (1989).

76 Peters, A. H. *et al.* Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107**, 323-337 (2001).

77 Flamm, W. G., Walker, P. M. & McCallum, M. Some properties of the single strands isolated from the DNA of the nuclear satellite of the mouse (Mus musculus). *Journal of molecular biology* **40**, 423-443 (1969).

78 Harel, J., Hanania, N., Tapiero, H. & Harel, L. RNA replication by nuclear satellite DNA in different mouse cells. *Biochemical and biophysical research communications* **33**, 696-701 (1968).

79 Rudert, F., Bronner, S., Garnier, J. M. & Dolle, P. Transcripts from opposite strands of gamma satellite DNA are differentially expressed during mouse development. *Mammalian genome : official journal of the International Mammalian Genome Society* **6**, 76-83 (1995).

80 Maison, C. *et al.* Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. *Nat Genet* **30**, 329-334, doi:10.1038/ng843 (2002).

81 Fang, Y. & Spector, D. L. Centromere positioning and dynamics in living Arabidopsis plants. *Molecular biology of the cell* **16**, 5710-5718, doi:10.1091/mbc.E05-08-0706 (2005).

82 Andrey, P. *et al.* Statistical analysis of 3D images detects regular spatial distributions of centromeres and chromocenters in animal and plant nuclei. *PLoS computational biology* **6**, e1000853, doi:10.1371/journal.pcbi.1000853 (2010).

83 Benoit, M., Layat, E., Tourmente, S. & Probst, A. V. Heterochromatin dynamics during developmental transitions in Arabidopsis - a focus on ribosomal DNA loci. *Gene* **526**, 39-45, doi:10.1016/j.gene.2013.01.060 (2013).

84 Simon, L., Voisin, M., Tatout, C. & Probst, A. V. Structure and Function of Centromeric and Pericentromeric Heterochromatin in Arabidopsis thaliana. *Frontiers in plant science* **6**, 1049, doi:10.3389/fpls.2015.01049 (2015).

85 James, T. C. & Elgin, S. C. Identification of a nonhistone chromosomal protein associated with heterochromatin in Drosophila melanogaster and its gene. *Molecular and cellular biology* **6**, 3862-3872 (1986).

86 Singh, P. B. *et al.* A sequence motif found in a Drosophila heterochromatin protein is conserved in animals and plants. *Nucleic acids research* **19**, 789-794 (1991).

87 Minc, E., Courvalin, J. C. & Buendia, B. HP1gamma associates with euchromatin and heterochromatin in mammalian nuclei and chromosomes. *Cytogenetics and cell genetics* **90**, 279-284, doi:10.1159/000056789 (2000).

88 Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. & Jenuwein, T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116-120, doi:10.1038/35065132 (2001).

89 Smothers, J. F. & Henikoff, S. The HP1 chromo shadow domain binds a consensus peptide pentamer. *Current biology : CB* **10**, 27-30 (2000).

90 Jones, D. O., Cowell, I. G. & Singh, P. B. Mammalian chromodomain proteins: their role in genome organisation and expression. *BioEssays : news and reviews in molecular, cellular and developmental biology* **22**, 124-137, doi:10.1002/(sici)1521-1878(200002)22:2<124::aid bies4>3.0.co;2-e (2000).

91 Fuks, F., Hurd, P. J., Deplus, R. & Kouzarides, T. The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Res* **31**, 2305-2312 (2003).

92 Cheutin, T. *et al.* Maintenance of stable heterochromatin domains by dynamic HP1 binding. *Science (New York, N.Y.)* **299**, 721-725, doi:10.1126/science.1078572 (2003).

93 Cubenas-Potts, C. & Corces, V. G. Architectural proteins, transcription, and the three-dimensional organization of the genome. *FEBS Lett* **589**, 2923-2930, doi:10.1016/j.febslet.2015.05.025 (2015).

94 Cremer, T. & Cremer, C. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nature reviews. Genetics* **2**, 292-301, doi:10.1038/35066075 (2001).

95 Boyle, S. *et al.* The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. *Human molecular genetics* **10**, 211-219 (2001).

96 Bolzer, A. *et al.* Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. *PLoS biology* **3**, e157, doi:10.1371/journal.pbio.0030157 (2005).

97 Branco, M. R. & Pombo, A. Chromosome organization: new facts, new models. *Trends in cell biology* **17**, 127-134, doi:10.1016/j.tcb.2006.12.006 (2007).

98 Parry, D. A., Conway, J. F. & Steinert, P. M. Structural studies on lamin. Similarities and differences between lamin and intermediate-filament proteins. *The Biochemical journal* **238**, 305-308 (1986).

99 Ben-Harush, K. *et al.* The supramolecular organization of the C. elegans nuclear lamin filament. *Journal of molecular biology* **386**, 1392-1402, doi:10.1016/j.jmb.2008.12.024 (2009).

100 Gruenbaum, Y. & Medalia, O. Lamins: the structure and protein complexes. *Curr Opin Cell Biol* **32**, 7-12, doi:10.1016/j.ceb.2014.09.009 (2015).

101 Vorburger, K., Lehner, C. F., Kitten, G. T., Eppenberger, H. M. & Nigg, E. A. A second higher vertebrate B-type lamin. cDNA sequence determination and in vitro processing of chicken lamin B2. *Journal of molecular biology* **208**, 405-415 (1989).

102 Peter, M. *et al.* Cloning and sequencing of cDNA clones encoding chicken lamins A and B1 and comparison of the primary structures of vertebrate A- and B-type lamins. *Journal of molecular biology* **208**, 393-404 (1989).

103 Lin, F. & Worman, H. J. Structural organization of the human gene (LMNB1) encoding nuclear lamin B1. *Genomics* **27**, 230-236, doi:10.1006/geno.1995.1036 (1995).

104 Furukawa, K. & Hotta, Y. cDNA cloning of a germ cell specific lamin B3 from mouse spermatocytes and analysis of its function by ectopic expression in somatic cells. *The EMBO journal* **12**, 97-106 (1993).

105 Lin, F. & Worman, H. J. Structural organization of the human gene encoding nuclear lamin A and nuclear lamin C.

The Journal of biological chemistry **268**, 16321-16326 (1993).

106 Nakajima, N. & Abe, K. Genomic structure of the mouse A-type lamin gene locus encoding somatic and germ cell-specific lamins. *FEBS letters* **365**, 108-114 (1995).

107 Machiels, B. M. *et al.* Abnormal A-type lamin organization in a human lung carcinoma cell line. *European journal of cell biology* **67**, 328-335 (1995).

108 Benavente, R., Krohne, G. & Franke, W. W. Cell type-specific expression of nuclear lamina proteins during development of Xenopus laevis. *Cell* **41**, 177-190 (1985).

109 Lehner, C. F., Stick, R., Eppenberger, H. M. & Nigg, E. A. Differential expression of nuclear lamin proteins during chicken development. *The Journal of cell biology* **105**, 577-587 (1987).

110 Beck, L. A., Hosick, T. J. & Sinensky, M. Isoprenylation is required for the processing of the lamin A precursor. *The Journal of cell biology* **110**, 1489-1499 (1990).

111 Kitten, G. T. & Nigg, E. A. The CaaX motif is required for isoprenylation, carboxyl methylation, and nuclear membrane association of lamin B2. *The Journal of cell biology* **113**, 13-23 (1991).

112 Farnsworth, C. C., Gelb, M. H. & Glomset, J. A. Identification of geranylgeranyl-modified proteins in HeLa cells. *Science (New York, N.Y.)* **247**, 320-322 (1990).

113 Corrigan, D. P. *et al.* Prelamin A endoproteolytic processing in vitro by recombinant Zmpste24. *The Biochemical journal* **387**, 129-138, doi:10.1042/bj20041359 (2005).

114 Gruenbaum, Y., Margalit, A., Goldman, R. D., Shumaker, D. K. & Wilson, K. L. The nuclear lamina comes of age. *Nature reviews. Molecular cell biology* **6**, 21-31, doi:10.1038/nrm1550 (2005).

115 Stewart, C. L., Roux, K. J. & Burke, B. Blurring the boundary: the nuclear envelope extends its reach. *Science (New York, N.Y.)* **318**, 1408-1412, doi:10.1126/science.1142034 (2007).

116 Burke, B. & Stewart, C. L. The nuclear lamins: flexibility in function. *Nature reviews. Molecular cell biology* **14**, 13-24, doi:10.1038/nrm3488 (2013).

117 Moir, R. D. *et al.* Review: the dynamics of the nuclear lamins during the cell cycle-- relationship between structure and function. *J Struct Biol* **129**, 324-334, doi:10.1006/jsbi.2000.4251 (2000).

118 Moir, R. D., Spann, T. P. & Goldman, R. D. The dynamic properties and possible functions of nuclear lamins. *International review of cytology* **162b**, 141-182 (1995).

119 Goldman, A. E., Moir, R. D., Montag-Lowy, M., Stewart, M. & Goldman, R. D. Pathway of incorporation of microinjected lamin A into the nuclear envelope. *J Cell Biol* **119**, 725-735 (1992).

120 Moir, R. D., Montag-Lowy, M. & Goldman, R. D. Dynamic properties of nuclear lamins: lamin B is associated with sites of DNA replication. *J Cell Biol* **125**, 1201-1212 (1994).

121 Dechat, T. *et al.* Lamina-associated polypeptide 2alpha binds intranuclear A-type lamins. *J Cell Sci* **113 Pt 19**, 3473-3484 (2000).

122 Dyer, J. A. *et al.* Cell cycle changes in A-type lamin associations detected in human dermal fibroblasts using monoclonal antibodies. *Chromosome Res* **5**, 383-394 (1997).

123 Dechat, T., Adam, S. A., Taimen, P., Shimi, T. & Goldman, R. D. Nuclear lamins. *Cold Spring Harb Perspect Biol* **2**, a000547, doi:10.1101/cshperspect.a000547 (2010).

124 Pickersgill, H. *et al.* Characterization of the Drosophila melanogaster genome at the nuclear lamina. *Nature genetics* **38**, 1005-1014, doi:10.1038/ng1852 (2006).

125 Kind, J. & van Steensel, B. Genome-nuclear lamina interactions and gene regulation. *Current opinion in cell biology* **22**, 320-325, doi:10.1016/j.ceb.2010.04.002 (2010).

126 Lund, E. G., Duband-Goulet, I., Oldenburg, A., Buendia, B. & Collas, P. Distinct features of lamin A-interacting chromatin domains mapped by ChIP-sequencing from sonicated or micrococcal nuclease-digested chromatin. *Nucleus* **6**, 30-39, doi:10.4161/19491034.2014.990855 (2015).

127 van Steensel, B. & Belmont, A. S. Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. *Cell* **169**, 780-791, doi:10.1016/j.cell.2017.04.022 (2017).

128 Guelen, L. *et al.* Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* **453**, 948-951, doi:10.1038/nature06947 (2008).

129 Peric-Hupkes, D. *et al.* Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol Cell* **38**, 603-613, doi:10.1016/j.molcel.2010.03.016 (2010).

130 Meuleman, W. *et al.* Constitutive nuclear lamina-genome interactions are highly conserved and associated with A/T-rich sequence. *Genome Res* **23**, 270-280, doi:10.1101/gr.141028.112 (2013).

131 Bian, Q., Khanna, N., Alvikas, J. & Belmont, A. S. beta-Globin cis-elements determine differential nuclear targeting through epigenetic modifications. *The Journal of cell biology* **203**, 767-783, doi:10.1083/jcb.201305027 (2013).

132 Harr, J. C. *et al.* Directed targeting of chromatin to the nuclear lamina is mediated by chromatin state and A-type lamins. *The Journal of cell biology* **208**, 33-52, doi:10.1083/jcb.201405110 (2015).

133 Edelman, L. B. & Fraser, P. Transcription factories: genetic programming in three dimensions. *Current opinion in*

genetics & development **22**, 110-114, doi:10.1016/j.gde.2012.01.010 (2012).

134 Chuang, C. H. *et al.* Long-range directional movement of an interphase chromosome site. *Curr Biol* **16**, 825-831, doi:10.1016/j.cub.2006.03.059 (2006).

135 Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. Capturing chromosome conformation. *Science (New York, N.Y.)* **295**, 1306-1311, doi:10.1126/science.1067799 (2002).

136 Palstra, R. J. *et al.* The beta-globin nuclear compartment in development and erythroid differentiation. *Nature genetics* **35**, 190-194, doi:10.1038/ng1244 (2003).

137 de Wit, E. & de Laat, W. A decade of 3C technologies: insights into nuclear organization. *Genes Dev* **26**, 11-24, doi:10.1101/gad.179804.111 (2012).

138 Simonis, M. *et al.* Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nature Gen* **38**, 1348-1354, doi:10.1038/ng1896 (2006).

139 Dostie, J. *et al.* Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome research* **16**, 1299-1309, doi:10.1101/gr.5571506 (2006).

140 Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human

genome. *Science* **326**, 289-293, doi:10.1126/science.1181369 (2009).

141 Imakaev, M. *et al.* Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat Methods* **9**, 999-1003, doi:10.1038/nmeth.2148 (2012).

142 Ea, V., Baudement, M. O., Lesne, A. & Forne, T. Contribution of Topological Domains and Loop Formation to 3D Chromatin Organization. *Genes* **6**, 734-750, doi:10.3390/genes6030734 (2015).

143 Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376-380, doi:10.1038/nature11082 (2012).

144 Sexton, T. *et al.* Three-dimensional folding and functional organization principles of the Drosophila genome. *Cell* **148**, 458-472, doi:10.1016/j.cell.2012.01.010 (2012).

145 Zuin, J. *et al.* Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc Natl Acad Sci USA* **111**, 996-1001, doi:10.1073/pnas.1317788111 (2014).

146 Van Bortle, K. *et al.* Insulator function and topological domain border strength scale with architectural protein occupancy. *Genome Biol* **15**, R82, doi:10.1186/gb-2014-15-5-r82 (2014).

147 Rao, S. S. *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665-1680, doi:10.1016/j.cell.2014.11.021 (2014).

148 Phillips-Cremins, J. E. *et al.* Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* **153**, 1281-1295, doi:10.1016/j.cell.2013.04.053 (2013).

149 Li, L. *et al.* Widespread rearrangement of 3D chromatin organization underlies polycomb-mediated stress-induced silencing. *Mol Cell* **58**, 216-231, doi:10.1016/j.molcel.2015.02.023 (2015).

150 Seitan, V. C. *et al.* Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments. *Genome Res* **23**, 2066-2077, doi:10.1101/gr.161620.113 (2013).

151 Dowen, J. M. *et al.* Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* **159**, 374-387, doi:10.1016/j.cell.2014.09.030 (2014).

152 Ong, C. T., Van Bortle, K., Ramos, E. & Corces, V. G. Poly(ADP-ribosyl)ation regulates insulator function and intrachromosomal interactions in Drosophila. *Cell* **155**, 148-159, doi:10.1016/j.cell.2013.08.052 (2013).

153 Lamouille, S., Xu, J. & Derynck, R. Molecular mechanisms of epithelial-mesenchymal transition. *Nature reviews. Molecular cell biology* **15**, 178-196, doi:10.1038/nrm3758 (2014).

154 Kalluri, R. & Weinberg, R. A. The basics of epithelial-mesenchymal transition. *The Journal of clinical investigation* **119**, 1420-1428, doi:10.1172/jci39104 (2009).

155 Hay, E. D. An overview of epithelio-mesenchymal transformation. *Acta Anat (Basel)* **154**, 8-20 (1995).

156 Li, R. *et al.* A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. *Cell stem cell* **7**, 51-63, doi:10.1016/j.stem.2010.04.014 (2010).

157 Samavarchi-Tehrani, P. *et al.* Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell* **7**, 64-77, doi:10.1016/j.stem.2010.04.015 (2010).

158 Huang, R. Y., Guilford, P. & Thiery, J. P. Early events in cell adhesion and polarity during epithelial-mesenchymal transition. *Journal Cell Sci* **125**, 4417-4422, doi:10.1242/jcs.099697 (2012).

159 De Craene, B. & Berx, G. Regulatory networks defining EMT during cancer initiation and progression. *Nat Rev Cancer* **13**, 97-110, doi:10.1038/nrc3447 (2013).

160 Qin, Y., Capaldo, C., Gumbiner, B. M. & Macara, I. G. The mammalian Scribble polarity protein regulates epithelial cell adhesion and migration through E-cadherin. *Journal Cell Biol* **171**, 1061-1071, doi:10.1083/jcb.200506094 (2005).

161 Yilmaz, M. & Christofori, G. EMT, the cytoskeleton, and cancer cell invasion. *Cancer Metastasis Rev* **28**, 15-33, doi:10.1007/s10555-008-9169-0 (2009).

162 Hay, E. D. The mesenchymal cell, its role in the embryo, and the remarkable signaling mechanisms that create it. *Developmental dynamics : an official publication of the American Association of Anatomists* **233**, 706-720, doi:10.1002/dvdy.20345 (2005).

163 Zeisberg, M. *et al.* Fibroblasts derive from hepatocytes in liver fibrosis via epithelial to mesenchymal transition. *The Journal of biological chemistry* **282**, 23337-23347, doi:10.1074/jbc.M700194200 (2007).

164 Hugo, H. *et al.* Epithelial--mesenchymal and mesenchymal--epithelial transitions in carcinoma progression. *Journal of cellular physiology* **213**, 374-383, doi:10.1002/jcp.21223 (2007).

165 Thiery, J. P. Epithelial-mesenchymal transitions in tumour progression. *Nature reviews. Cancer* **2**, 442-454, doi:10.1038/nrc822 (2002).

166 Xu, J., Lamouille, S. & Derynck, R. TGF-beta-induced epithelial to mesenchymal transition. *Cell Res* **19**, 156-172, doi:10.1038/cr.2009.5 (2009).

167 Gonzalez, D. M. & Medici, D. Signaling mechanisms of the epithelial-mesenchymal transition. *Science Signaling* **7**, re8, doi:10.1126/scisignal.2005189 (2014).

168 Derynck, R. & Zhang, Y. E. Smad-dependent and Smad-independent pathways in TGF-beta family signalling. *Nature* **425**, 577-584, doi:10.1038/nature02006 (2003).

169 Itoh, S., Itoh, F., Goumans, M. J. & Ten Dijke, P. Signaling of transforming growth factor-beta family members through Smad proteins. *European journal of biochemistry* **267**, 6954-6967 (2000).

170 Feng, X. H. & Derynck, R. Specificity and versatility in tgf-beta signaling through Smads. *Annual review of cell and developmental biology* **21**, 659-693, doi:10.1146/annurev.cellbio.21.022404.142018 (2005).

171 Peinado, H., Olmeda, D. & Cano, A. Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? *Nat Rev Cancer* **7**, 415-428, doi:nrc2131 [pii]10.1038/nrc2131 (2007).

172 Batlle, E. *et al.* The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nat Cell Biol* **2**, 84-89, doi:10.1038/35000034 (2000).

173 Remacle, J. E. *et al.* New mode of DNA binding of multi-zinc finger transcription factors: deltaEF1 family members bind with two hands to two target sites. *EMBO J* **18**, 5073-5084, doi:10.1093/emboj/18.18.5073 (1999).

174 Yang, M. H. *et al.* Bmi1 is essential in Twist1-induced epithelial-mesenchymal transition. *Nature Cell Biol* **12**, 982-992, doi:10.1038/ncb2099 (2010).

175 Barrallo-Gimeno, A. & Nieto, M. A. The Snail genes as inducers of cell movement and survival: implications in development and cancer. *Development* **132**, 3151-3161, doi:132/14/3151 [pii]10.1242/dev.01907 (2005).

176 Chiang, C. & Ayyanathan, K. Characterization of the E-box binding affinity to snag-zinc finger proteins. *Molekuliarnaia biologiia* **46**, 907-914 (2012).

177 Nieto, M. A. The snail superfamily of zinc-finger transcription factors. *Nat Rev Mol Cell Biol* **3**, 155-166, doi:10.1038/nrm757nrm757 [pii] (2002).

178 Peinado, H., Ballestar, E., Esteller, M. & Cano, A. Snail mediates E-cadherin repression by the recruitment of the Sin3A/histone deacetylase 1 (HDAC1)/HDAC2 complex. *Mol Cell Biol* **24**, 306-319 (2004).

179 Herranz, N. *et al.* Polycomb complex 2 is required for E-cadherin repression by the Snail1 transcription factor. *Mol Cell Biol* **28**, 4772-4781, doi:MCB.00323-08 [pii]10.1128/MCB.00323-08 (2008).

180 Lin, T., Ponn, A., Hu, X., Law, B. K. & Lu, J. Requirement of the histone demethylase LSD1 in Snai1-mediated transcriptional repression during epithelial-mesenchymal transition. *Oncogene* **29**, 4896-4904, doi:10.1038/onc.2010.234 (2010).

181 Herranz, N. *et al.* Lysyl Oxidase-Like 2 (LOXL2) Oxidizes Trimethylated Lysine 4 in Histone H3. *FEBS J*, doi:10.1111/febs.13922 (2016).

182 Miettinen, P. J., Ebner, R., Lopez, A. R. & Derynck, R. TGF-beta induced transdifferentiation of mammary epithelial cells to mesenchymal cells: involvement of type I receptors. *J Cell Biol* **127**, 2021-2036 (1994).

183 Jao, C. Y. & Salic, A. Exploring RNA transcription and turnover in vivo by using click chemistry. *Proc Natl Acad Sci U S A* **105**, 15779-15784, doi:10.1073/pnas.0808480105 (2008).

184 Gesson, K. *et al.* A-type lamins bind both hetero- and euchromatin, the latter being regulated by lamina-associated polypeptide 2 alpha. *Genome Res* **26**, 462-473, doi:10.1101/gr.196220.115 (2016).

185 Buenrostro, J. D., Wu, B., Chang, H. Y. & Greenleaf, W. J. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol* **109**, 21 29 21-29, doi:10.1002/0471142727.mb2129s109 (2015).

186 Hill, C. S. Transcriptional Control by the SMADs. *Cold Spring Harb Perspect Biol* **8**, doi:10.1101/cshperspect.a022079 (2016).

187 Seoane, J. & Gomis, R. R. TGF-beta Family Signaling in Tumor Suppression and Cancer Progression. *Cold Spring Harb Perspect Biol* **9**, doi:10.1101/cshperspect.a022277 (2017).

188 Lowery, J., Kuczmarski, E. R., Herrmann, H. & Goldman, R. D. Intermediate Filaments Play a Pivotal Role in Regulating

Cell Architecture and Function. *J Biol Chem* **290**, 17145-17153, doi:10.1074/jbc.R115.640359 (2015).

189 Rowley, M. J. & Corces, V. G. The three-dimensional genome: principles and roles of long-distance interactions. *Curr Opin Cell Biol* **40**, 8-14, doi:10.1016/j.ceb.2016.01.009 (2016).

190 Shah, P. P. *et al.* Lamin B1 depletion in senescent cells triggers large-scale changes in gene expression and the chromatin landscape. *Genes Dev* **27**, 1787-1799, doi:10.1101/gad.223834.113 (2013).

191 Yang, S. H., Jung, H. J., Coffinier, C., Fong, L. G. & Young, S. G. Are B-type lamins essential in all mammalian cells? *Nucleus* **2**, 562-569, doi:10.4161/nucl.2.6.18085 (2011).

192 Shimi, T. *et al.* The role of nuclear lamin B1 in cell proliferation and senescence. *Genes Dev* **25**, 2579-2593, doi:10.1101/gad.179515.111 (2011).

193 Broers, J. L. *et al.* Dynamics of the nuclear lamina as monitored by GFP-tagged A-type lamins. *Journal of cell science* **112 (Pt 20)**, 3463-3475 (1999).

194 Papantonis, A. & Cook, P. R. Genome architecture and the role of transcription. *Curr Opin Cell Biol* **22**, 271-276, doi:10.1016/j.ceb.2010.03.004 (2010).

195 Casanova, M. *et al.* Heterochromatin reorganization during early mouse development requires a single-stranded

noncoding transcript. *Cell reports* **4**, 1156-1167, doi:10.1016/j.celrep.2013.08.015 (2013).

196 Mayer, R. *et al.* Common themes and cell type specific variations of higher order chromatin arrangements in the mouse. *BMC cell biology* **6**, 44, doi:10.1186/1471-2121-6-44 (2005).

197 McDonald, O. G., Wu, H., Timp, W., Doi, A. & Feinberg, A. P. Genome-scale epigenetic reprogramming during epithelial-to-mesenchymal transition. *Nat Struct Mol Biol* **18**, 867-874, doi:10.1038/nsmb.2084 (2011).

198 Neymotin, B., Athanasiadou, R. & Gresham, D. Determination of in vivo RNA kinetics using RATE-seq. *RNA (New York, N.Y.)* **20**, 1645-1652, doi:10.1261/rna.045104.114 (2014).

199 Wijchers, P. J. *et al.* Characterization and dynamics of pericentromere-associated domains in mice. *Genome research* **25**, 958-969, doi:10.1101/gr.186643.114 (2015).

200 Becker, J. S. *et al.* Genomic and Proteomic Resolution of Heterochromatin and Its Restriction of Alternate Fate Genes. *Molecular cell* **68**, 1023-1037.e1015, doi:10.1016/j.molcel.2017.11.030 (2017).

201 Adam, S. A. & Goldman, R. D. Insights into the differences between the A- and B-type nuclear lamins. *Advances in biological regulation* **52**, 108-113, doi:10.1016/j.advenzreg.2011.11.001 (2012).

202 Shimi, T. *et al.* The A- and B-type nuclear lamin networks: microdomains involved in chromatin organization and transcription. *Genes Dev* **22**, 3409-3421, doi:10.1101/gad.1735208 (2008).

203 Dechat, T., Gesson, K. & Foisner, R. Lamina-independent lamins in the nuclear interior serve important functions. *Cold Spring Harb Symp Quant Biol* **75**, 533-543, doi:10.1101/sqb.2010.75.018 (2010).

204 Gerace, L. & Blobel, G. The nuclear envelope lamina is reversibly depolymerized during mitosis. *Cell* **19**, 277-287 (1980).

205 Schoen, I., Aires, L., Ries, J. & Vogel, V. Nanoscale invaginations of the nuclear envelope: Shedding new light on wormholes with elusive function. *Nucleus* **8**, 506-514, doi:10.1080/19491034.2017.1337621 (2017).

206 Fricker, M., Hollinshead, M., White, N. & Vaux, D. Interphase nuclei of many mammalian cell types contain deep, dynamic, tubular membrane-bound invaginations of the nuclear envelope. *The Journal of cell biology* **136**, 531-544 (1997).

207 Vogel, M. J., Peric-Hupkes, D. & van Steensel, B. Detection of in vivo protein-DNA interactions using DamID in mammalian cells. *Nature protocols* **2**, 1467-1478, doi:10.1038/nprot.2007.148 (2007).

208 Aughey, G. N. & Southall, T. D. Dam it's good! DamID profiling of protein-DNA interactions. *Wiley interdisciplinary*

reviews. Developmental biology **5**, 25-37, doi:10.1002/wdev.205 (2016).

209 Naetar, N., Ferraioli, S. & Foisner, R. Lamins in the nuclear interior - life outside the lamina. *Journal of cell science* **130**, 2087-2096, doi:10.1242/jcs.203430 (2017).

210 Amendola, M. & van Steensel, B. Mechanisms and dynamics of nuclear lamina-genome interactions. *Curr Opin Cell Biol* **28**, 61-68, doi:10.1016/j.ceb.2014.03.003 (2014).

211 de Leeuw, R., Gruenbaum, Y. & Medalia, O. Nuclear Lamins: Thin Filaments with Major Functions. *Trends in cell biology* **28**, 34-45, doi:10.1016/j.tcb.2017.08.004 (2018).

212 Papait, R. *et al.* Genome-wide analysis of histone marks identifying an epigenetic signature of promoters and enhancers underlying cardiac hypertrophy. *Proc Natl Acad Sci U S A* **110**, 20164-20169, doi:10.1073/pnas.1315155110 (2013).

213 Cesarini, E. *et al.* Lamin A/C sustains PcG protein architecture, maintaining transcriptional repression at target genes. *The Journal of cell biology* **211**, 533-551, doi:10.1083/jcb.201504035 (2015).

214 Marullo, F. *et al.* Nucleoplasmic Lamin A/C and Polycomb group of proteins: An evolutionarily conserved interplay. *Nucleus* **7**, 103-111, doi:10.1080/19491034.2016.1157675 (2016).

215 Lanzuolo, C. & Orlando, V. Memories from the polycomb group proteins. *Annual review of genetics* **46**, 561-589, doi:10.1146/annurev-genet-110711-155603 (2012).

216 Houben, F., Ramaekers, F. C., Snoeckx, L. H. & Broers, J. L. Role of nuclear lamina-cytoskeleton interactions in the maintenance of cellular strength. *Biochimica et biophysica acta* **1773**, 675-686, doi:10.1016/j.bbamcr.2006.09.018 (2007).

217 Bronshtein, I. *et al.* Loss of lamin A function increases chromatin dynamics in the nuclear interior. *Nature communications* **6**, 8044, doi:10.1038/ncomms9044 (2015).

218 Zheng, X. *et al.* Lamins Organize the Global Three-Dimensional Genome from the Nuclear Periphery. *Molecular Cell* **71**, 1-14, doi:10.1016/j.molcel.2018.05.017 (2018).

219 Rowley, M. J. *et al.* Evolutionarily Conserved Principles Predict 3D Chromatin Organization. *Mol Cell* **67**, 837-852.e837, doi:10.1016/j.molcel.2017.07.022 (2017).

220 Gurudatta, B. V., Yang, J., Van Bortle, K., Donlin-Asp, P. G. & Corces, V. G. Dynamic changes in the genomic localization of DNA replication-related element binding factor during the cell cycle. *Cell cycle (Georgetown, Tex.)* **12**, 1605-1615, doi:10.4161/cc.24742 (2013).

221 Hutchison, C. J. B-type lamins in health and disease. *Seminars in cell & developmental biology* **29**, 158-163, doi:10.1016/j.semcdb.2013.12.012 (2014).

222 Kalcheim, C. Epithelial-Mesenchymal Transitions during Neural Crest and Somite Development. *Journal of clinical medicine* **5**, doi:10.3390/jcm5010001 (2015).

223 Irianto, J., Pfeifer, C. R., Ivanovska, I. L., Swift, J. & Discher, D. E. Nuclear lamins in cancer. *Cellular and molecular bioengineering* **9**, 258-267, doi:10.1007/s12195-016-0437-8 (2016).

224 Kim, Y. *et al.* Mouse B-type lamins are required for proper organogenesis but not by embryonic stem cells. *Science (New York, N.Y.)* **334**, 1706-1710, doi:10.1126/science.1211222 (2011).

225 Bogdanova, M. A. *et al.* [Nuclear lamins regulate osteogenic differentiation of mesenchymal stem cells]. *Tsitologiia* **56**, 260-267 (2014).

226 Chen, H. & Zheng, Y. Nuclear lamina builds tissues from the stem cell niche. *Fly* **8**, 63-67, doi:10.4161/fly.28063 (2014).

227 Coffinier, C. *et al.* Deficiencies in lamin B1 and lamin B2 cause neurodevelopmental defects and distinct nuclear shape abnormalities in neurons. *Molecular biology of the cell* **22**, 4683-4693, doi:10.1091/mbc.E11-06-0504 (2011).

228 Mahajani, S. *et al.* Lamin B1 levels modulate differentiation into neurons during embryonic corticogenesis. *Scientific reports* **7**, 4897, doi:10.1038/s41598-017-05078-6 (2017).

229 Zuela, N., Bar, D. Z. & Gruenbaum, Y. Lamins in development, tissue maintenance and stress. *EMBO reports* **13**, 1070-1078, doi:10.1038/embor.2012.167 (2012).

230 Zhang, X. *et al.* Pax6 is a human neuroectoderm cell fate determinant. *Cell stem cell* **7**, 90-100, doi:10.1016/j.stem.2010.04.017 (2010).

231 Meshorer, E. *et al.* Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Developmental cell* **10**, 105-116, doi:10.1016/j.devcel.2005.10.017 (2006).

232 Rizzo, M. A., Davidson, M. W. & Piston, D. W. Fluorescent protein tracking and detection: fluorescent protein structure and color variants. *Cold Spring Harb Protoc* **2009**, pdb top63, doi:10.1101/pdb.top63 (2009).

233 Nissim-Rafinia, M. & Meshorer, E. Photobleaching assays (FRAP & FLIP) to measure chromatin protein dynamics in living embryonic stem cells. *J Vis Exp*, doi:10.3791/2696 (2011).

234 Morey, L. *et al.* Nonoverlapping functions of the Polycomb group Cbx family of proteins in embryonic stem cells. *Cell Stem Cell* **10**, 47-62, doi:10.1016/j.stem.2011.12.006 (2012).

235 Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105-1111, doi:10.1093/bioinformatics/btp120 (2009).

236 Roberts, A., Trapnell, C., Donaghey, J., Rinn, J. L. & Pachter, L. Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome Biol* **12**, R22, doi:10.1186/gb-2011-12-3-r22 (2011).

237 Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25, doi:10.1186/gb-2009-10-3-r25 (2009).

238 Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137, doi:10.1186/gb-2008-9-9-r137 (2008).

239 Feitelson, D. G. Comparing Partitions with Spie Charts. *Technical Report* 2003-2087 (2003).

240 Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* **44**, W90-97, doi:10.1093/nar/gkw377 (2016).

241 Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics (Oxford, England)* **27**, 1739-1740, doi:10.1093/bioinformatics/btr260 (2011).

242 Kent, W. J. *et al.* The human genome browser at UCSC. *Genome Res* **12**, 996-1006, doi:10.1101/gr.229102. Article published online before print in May 2002 (2002).

243 François Serra, D. B., Guillaume Filion, Marc A. Marti-Renom. Structural features of the fly chromatin colors

revealed by automatic three-dimensional modeling. *bioRxiv*, doi:doi.org/10.1101/036764 (2016).

244 Nora, E. P. *et al.* Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381-385, doi:10.1038/nature11049 (2012).

245 Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular cell* **38**, 576-589, doi:10.1016/j.molcel.2010.05.004 (2010).

246 Lajoie, B. R., Dekker, J. & Kaplan, N. The Hitchhiker's guide to Hi-C analysis: practical guidelines. *Methods* **72**, 65-75, doi:10.1016/j.ymeth.2014.10.031 (2015).

247 de Wit, E. *et al.* The pluripotent genome in three dimensions is shaped around pluripotency factors. *Nature* **501**, 227-231, doi:10.1038/nature12420 (2013).

RESEARCH ARTICLES

Research article resulting from this thesis:

Pascual-Reguant L, Blanco E, Galan S, Le Dily F, Cuartero Y, Serra-Bardenys G, Di Carlo V, Iturbide A, Cebrià-Costa JP, Nonell L, de Herreros AG, Di Croce L, Marti-Renom MA, Peiró S. Lamin B1 mapping reveals the existence of dynamic and functional euchromatin lamin B1 domains. **Nat Commun**, 2018, 9(1):3420.

Collaboration with other projects:

Izquierdo-Bouldstridge A, Bustillos A, Bonet-Costa C, Aribau-Miralbés P, García-Gomis D, Dabad M, Esteve-Codina A, **Pascual-Reguant** L, Peiró S, Esteller M, Murtha M, Millán-Ariño L, Jordan A. Histone H1 depletion triggers an interferon response in cancer cells via activation of heterochromatic repeats. **NAR**, 2017, 45(20):11622-11642.

Herranz N, Dave N, Millanes-Romero A, **Pascual-Reguant** L, Morey L, Díaz VM, Lórenz-Fonfría V, Gutierrez-Gallego R, Jerónimo C, Iturbide A, Di Croce L, García de Herreros A, Peiró S. Lysyl oxidase-like 2 (LOXL2) oxidizes trimethylated lysine 4 in histone H3. **FEBS J**, 2016, 283(23):4263-4273.

Iturbide A, **Pascual-Reguant L**, Fargas L, Cebriá JP, Alsina B, García de Herreros A, Peiró S. LOXL2 Oxidizes Methylated TAF10 and Controls TFIID-Dependent Genes during Neural Progenitor Differentiation. **Mol Cell**, 2015, 58(5): 755-766.

ACKNOWLEDGEMENTS

Primer de tot m'agradaria fer un balanç del que ha suposat per a mi aquesta tesi doctoral. Fer un doctorat, per suposat, no és una feina fàcil i per a mi no ha sigut una excepció. Ha estat com viure permanentment en una muntanya russa; baixades i pujades constants, moments de voler plegar i moments de sentir-te capaç de menjar-te el món. Tot i així, al mirar enrere i recordar moments viscuts me n'adono que el balanç que faig de tots aquests anys és molt positiu. Tinc molt presents moments difícils, tant científicament parlant com personalment, però si intento resumir els últims cinc anys, el primer que em ve al cap és tot el que he après. He après ciència (ara me n'adono que abans de començar al laboratori no sabia res), tot i que com més sé, més me n'adono de tot el que em falta per aprendre. He après què és treballar en un equip, on si falla un fallem tots, però sobretot he après moltíssim a nivell personal. He tingut la sort de poder conèixer persones increïbles que m'han ensenyat moltíssim i de les quals aprenc coses noves cada dia. He après a passar-m'ho extremadament bé a la feina, a disfrutar del que fem, a veure el cantó positiu de les coses i a intentar remuntar quan tot és de color negra.

Aquesta tesi la vaig començar gràcies a tu, Sandra. Vas ser la única persona que em va donar esperances tot i no tenir una nota excel·lent, però sobretot la vaig començar perquè la teva energia i les teves ganes de menjar-te el món són extremadament contagioses i admirables. Durant aquests cinc anys, les hem passat de tots colors; hem rigut, hem discutit de ciència, ens hem flipat i després hem tocat a terra...però, sobretot, m'has ensenyat tantíssimes coses; a pensar per mi mateixa, a treure sempre alguna cosa positiva del resultat més desastrós, a tirar endavant

un projecte que partia del “no res”, a ser positiva però crítica a la vegada, a no donar mai res per perdut, a apassionar-me pels experiments com més complicats millor... Però de totes les coses que has fet per mi, el que més t’agraeixo és la confiança que sempre m’has demostrat. Per a mi aquesta confiança ha sigut un dels motors imprescindible per tirar endavant el projecte i, en definitiva, acabar la tesi. Per tot això i probablement per molt més, t’agraeixo que m’hagis donat la oportunitat de formar part del teu (ja no tant petit) Chromatin Team. Com diu l’Antonio, ets la “big little boss”.

En aquesta feina, formar part d’un grup és indispensable per aprendre, per progressar i per millorar. Per això, vull donar les gràcies a tots els membres de Chromatin Team amb els que he treballat durant aquests cinc anys. Primer de tot, gràcies Alba per ensenyar-me i dedicar part del teu temps en un dels moments més complicats de la tesi, sé que no va ser fàcil. Ane, muchas gracias por adoptarme cuando me quedé un poco “colgada” y enseñarme de todo y más. Aprendí mucho de ti y aunque les revisiones son el peor momento posible, disfruté muchísimo trabajando contigo. Les cogí cariño a las ESC y todo. Gràcies també a tu, Gemma, per tenir tanta paciència amb mi quan et vaig intentar ensenyar sense saber gairebé res. Sóc conscient que no era la situació idònia però vaig aprendre molt durant aquell temps. Molts ànims per l’últim any de tesi, ja veuràs que és molt millor del que sembla, ja quasi ho tens. Marc, vas entrar fa només un any però amb un bon “rollo” que encomana, sempre de bon humor i amb un somriure, gràcies per fer-nos el dia a dia més fàcil. Gràcies també a tu, Queralt, fa molt poc que ens coneixem però t’asseguro que m’has ajudat molt durant aquestes setmanes, amb la teva energia i les teves ganes

de treballar has fet que m'oblidi de tot, fins i tot has aconseguit que enmig del dipòsit de la tesi, m'emocioni al revelar un Western. Jess, ets la persona més incondicional que he conegut mai. Portem cinc anys treballant juntes i és un autèntic plaer tant a nivell professional com personal, moltíssimes gràcies per tant. Tenir-te a tu al laboratori em fa la vida extremadament més fàcil; sempre disposada a ajudar i sempre buscant solucions. Estic contentíssima de veure't tant feliç amb el títol de pilot, t'ho mereixes, ets una autèntica crack.

Aquests anys no haguessin sigut ni la meitat de divertits sense Snail Team. Moltíssimes gràcies a tots pel bon ambient, sense vosaltres no hagués sigut el mateix; Raúl, Jordi, Lorena, David, Willy, Aida, Jelena, Jou i, per suposat a tu Rubén, no t'imagines la il·lusió que em va fer quan vas entrar a Snail, tot i que no haguem compartit gaire poiata, saps que has sigut imprescindible. Gràcies també als veïns de Biga's lab i als PN's lab, sobretot a tu Judit, que sempre has format part de la família Snail. També m'agradaria donar les gràcies a tots els jefes; Antonio, Jepi i Víctor. Els vostres consells i aportacions, sobretot en els lab meetings, m'han ajudat molt a enfocar les coses des de un altra punt de vista.

La incorporació a una nova institució no va ser fàcil però al VHIO he conegut a persones que han fet que valgués la pena; Albert, Oriol, Estefania (tot i que ja ens coneixíem, m'ha agradat retrobar-te), Judit, Isa, Gemma, Sandra, Ale (con tu energia "andaluza" es un placer ir a trabajar)...

Dos pilars fonamentals durant aquests anys heu estat, sens dubte, vosaltres, Joan Pau i Rumi. Només per tot el que heu aguantat us mereixeu un capítol a part. Conèixer a persones com vosaltres és

un dels motius pels quals ha valgut la pena fer la tesi. Vam començar a portar-nos bé, a fer unes birres, a fer uns riures... però de seguida us vau convertir en amics indispensables en el meu dia a dia. Joan Pau, només entrar al lab em vas acollir i em vas fer sentir part del grup a la primera setmana, amb les teves bromes i tonteries... per a mi va significat molt. Al cap de poc temps, les bromes es van convertir en converses d'hores davant del PRBB i en "pitis" durant el dia per desconnectar i tenir un moment per parlar. També hem discutit un munt sobre ciència, les teves crítiques constructives sobre el meu projecte i les divagacions sobre ciència m'han aportat molt més del que segurament et penses. Vull donar-te les gràcies per escoltar-me, per entendre'm (o al menys fer-ho veure...jeje), per creure que sóc molt millor del que em sento, per ajudar-me amb tot i més i, sobretot, per ser-hi absolutament sempre que t'he necessitat. Rumi, quan ens vam començar a conèixer era com si et conegués de tota la vida. Poques vegades he sentit la connexió que sento amb tu. Durant aquest temps sempre has estat present en els moments en què ho he necessitat, sempre a punt i sempre disposada, això només ho fa una amiga de veritat. Juntes hem passat de tot; hem parlat molt, hem rigut molt, també hem plorat alguna vegada... però sobretot hem viscut moltes coses juntes durant aquest temps. Et desitjo el millor en el teu postdoc, et mereixes tantíssim aquesta oportunitat... Me n'alegro moltíssim per tu però també saps que et trobaré molt a faltar. Per últim, només queda donar-te les gràcies; moltes gràcies per aguantar les meves "rallades", moltes gràcies per tenir sempre un somriure a punt, moltes gràcies per entendre'm, moltes gràcies per ser-hi sempre, moltes gràcies per tot.

Per descomptat, vull donar les gràcies a tots els meus amics de la Uni. Vosaltres heu fet que tot sigui més fàcil però sobretot més divertit. A tu Rubén, pel teu suport incondicional i per estar sempre disposat a escolar-me tot i que siguin les vuit del vespre i tinguis encara mil coses per fer. A tu David, per tot i estar lluny, estar a prop absolutament sempre que ho he necessitat, sense pensar-t'ho dues vegades. A tu Oscar, per pensar en mi sempre i pel teu humor negra (ja saps que m'encanta). A ti Cari, por estar también lejos pero hacer que no lo parezca cada vez que nos vemos. A tu Ali, per compartir amb mi el que penses i per les teves ganes de quedar tots junts per veure'ns. A ti Garchi, por ser siempre "el alma de la fiesta" y hacernos reír aunque tengamos un mal día. I a tu Sara, per ser una persona sempre disposada a tot. Sou els millors amics que podria tenir.

La meva família ha sigut imprescindible durant tot aquest temps, com, òbviament, també ho ha sigut al llarg de la meva vida. A vosaltres, papes, us vull agrair el suport incondicional que sempre m'heu transmès. Sóc conscient que no és fàcil entendre el món de la ciència però sempre heu fet l'esforç de comprendre'm i donar-me els consells oportuns. També us agraeixo infinitament els valors que m'heu inculcat des de petita, sobretot el valor de l'esforç, que ha estat indispensable durant aquests anys i que l'he après de vosaltres; no donar mai res per perdut i sempre lluitar pel què vulgui, encara que no sigui fàcil. Germana, sempre has sigut un mirall per a mi; que tu fas Biologia, jo faig Biologia, que tu fas un Doctorat, jo faig un Doctorat... Suposo que l'admiració fa que et vulguis assemblar a l'altra persona. Moltes gràcies per sentir-te sempre orgullosa de mi i per donar-me suport amb absolutament

totes les decisions que he près. Espero ser capaç algun dia de tornar-vos tot el que heu fet per a mi.

Per últim només quedes tu, Adri. Donar-te les gràcies és poc, no sabria com agrair-te tantíssim. Ets, indiscutiblement, el meu pilar fonamental. Des del principi, has celebrat les meves victòries, has plorat les meves derrotes i has afrontat les meves inseguretats. Et demano disculpes per haver-me hagut d'aixecar les vegades que m'he enfonsat però sobretot t'agraeixo que hagis compartit amb mi les alegries, vivint-les encara amb més intensitat que jo i fent-me veure que realment val la pena. Sempre que m'he sentit incapaç, m'has fet sentir capaç, sempre que m'he sentit impotent, m'has fet sentir forta i sempre que m'he sentit satisfeta, tu m'has fet sentir orgullosa. Infinites gràcies.

Aquesta tesi ha estat impresa amb el suport del VHIO (Vall d'Hebron Institut d'Oncologia)