PhD

PROGRAM IN TRANSLATIONAL
AND MOLECULAR MEDICINE

DIMET

University of Milano-Bicocca
School of Medicine and Faculty of Science

Chromatin solubility as a novel determinant of epigenome dysfunction in prostate cancer

Coordinator: Prof. Andrea Biondi
Tutor: Prof. Chiara Lanzuolo

Dr. Valentina Rosti
Matr. No. 849310

XXXIV CYCLE
ACADEMIC YEAR
2020-2021

CHAPTER 1. GENERAL INTRODUCTION **1**

1.1 The Epigenome *2*

1.1.1 Three-dimensional organization of the genome 3

1.1.2 DNA methylation and Histone Code 7

1.1.3 Euchromatin and Heterochromatin 9

1.1.4 Nuclear lamina 11

1.1.5 Liquid-Liquid Phase Separation (LLPS) 12

1.2 Nuclear abnormalities in cancer *15*

1.3 Prostate cancer *17*

1.3.1 PCa diagnosis 19

1.3.2 Pathology and staging of PCa 21

1.4 The genomic landscape of PCa *23*

1.5 Epigenetic alterations in PCa *25*

1.5.1 DNA methylation 26

1.5.2 Histone Modification Machinery 26

1.5.3 3D genome structure 29

1.6 Epithelial-to-Mesenchymal Transition (EMT) *31*

1.7 Scope of the thesis *33*

References to chapter 1 *34*

CHAPTER 2. DYSFUNCTIONAL POLYCOMB TRANSCRIPTIONAL REPRESSION CONTRIBUTES TO LAMIN A/C DEPENDENT MUSCULAR DYSTROPHY **59**

Abstract *61*

Introduction *62*

Results *63*

Discussion *74*

Methods *76*

Figures 87

Supplementary Figures 98

Additional methods 112

References to chapter 2 120

CHAPTER 3. CHROMATIN SOLUBILITY AS A NOVEL DETERMINANT OF EPIGENOME DYSFUNCTION IN PROSTATE CANCER 134

Abstract 135

3.1 4 fractions Sequential Analysis of MacroMolecules accessibilitY (4fSAMMY-seq) allows mapping of heterochromatic and euchromatin regions in human fibroblasts 138

Figure 1 – 4fSAMMY-seq isolates specific genomic regions in human fibroblasts. 139

Supplementary Figure 1 – Separation of S2 small (S2S) and S2 large (S2L) subfractions. 141

Figure 2 - 4fSAMMY-seq differentially separates heterochromatin and euchromatin in human fibroblasts. 143

3.2 4fSAMMY-seq identifies functional chromatin solubility states in normal prostate biopsies 145

Figure 3 – Overview of experimental design. 146

Supplementary Figure 2 – Specimen cohort. 148

Supplementary Figure 3 – Dissociation of prostatic needle biopsies into single-cell suspensions. 149

Supplementary Figure 4 – Quality check. 151

Figure 4 – 4fSAMMY-seq discriminates heterochromatic and euchromatic regions in normal prostate biopsies. 152

3.3 Identification of malignant chromatin solubility rearrangements in PCa biopsies by 4fSAMMY-seq 155

Figure 5 – 4fSAMMY-seq reveals chromatin solubility changes between normal prostate and PCa biopsies. 156

3.4 4fSAMMY solubility profile is not influenced by relative amount of distinct cell types *158*
Supplementary Figure 5 – Flow Cytometry gating strategy. 158
Figure 6 – Flow Cytometry characterization of prostate bioptic tissues. 160
Supplementary Figure 6 – Specimen cohort. 161
Figure 7 – PCa-related 4fSAMMY-seq profile is not dependent on tissue composition. 163
Figure 8 – PCa-specific 4fSAMMY-seq fingerprint. 165
3.5 Prostate cancer tissues exhibit two distinct chromatin solubility profiles *166*
Supplementary Figure 7 – Clinicopathological features. 166
Figure 9 – 4fSAMMY-seq discriminates two distinct solubility clusters among PCa tissues. 168
3.6 Integrative transcriptome and chromatin organization analysis reveals aberrant solubility compartmentalization in PCa tissues *169*
Supplementary Figure 8 – PCa molecular subtyping. 170
Figure 10 - Transcriptomic differences in PCa samples. 172
Figure 11 - Correlation between solubility and gene expression. 175
3.7 Severe solubility remodeling is associated to ECM plasticity and EMT reprogramming *176*
Figure 12 – Recurrent solubility alterations in PCa tissues. 177
Figure 13 – Transcriptomic analysis of subtype 2 PCa tissues. 180
Materials and methods *182*
References to chapter 3 *196*
CHAPTER 4. FINAL CONSIDERATIONS **201**
4.1. Summary and conclusions *202*
4.2 Future perspectives *208*

References to chapter 4 *210*

PUBLICATIONS **215**

CHAPTER 1.
GENERAL INTRODUCTION

1.1 The Epigenome

The majority of hereditary information in eukaryotes is stored, transcribed, and replicated in the cell nucleus. A human being is composed by trillions of cells containing the same genome in all nuclei[1]. The differences among cells are determined by expressing only a small part of their genetic information without altering the primary sequence of DNA. The epigenetic mechanisms regulate how and when turning on or off specific genetic programs, to define and preserve the cell identity[2]. The epigenetic factors not only modulate the transcriptional state of genomic regions[3] but also prevent DNA damage[4] and reduce errors during replication and cell division[5]. The epigenome is plastic: it can adapt genomic function in response to internal or external cues. This biological adaptive dynamic mediates fundamental processes such as cell fate choice during tissue regeneration[6]. An example is provided by the activation of muscle stem cells (MuSCs), which is governed by the activity of epigenetic factors controlling the proper space and temporal gene expression for the production of new functional muscle fibers[7]. On the contrary, any dysfunction in this finely tuned epigenetic regulation can cause tissue alterations and pathology development[8]. We observed in a murine model of Emery-Dreifuss muscular dystrophy (EDMD), that the deregulation of Polycomb group (PcG) of proteins epigenetic repressors in MuSCs causes premature exhaustion of quiescent stem cells and accumulation of intramuscular fat, contributing to the muscular dystrophy progression (for in depth details see Chapter

2)[9]. Recently, a multitude of data has further underlined the importance of epigenetic mechanisms in the field of human cancer[10]. Novel research branches have developed clinical translational application of epigenetic studies[11]. The possibility of reprogramming the epigenetic landscape in cancer is one of the most promising therapeutic strategies. To date, several epigenetic drugs (epidrugs) were released both for treatment and reversibility of drug resistance[12]. In parallel, the exploration of epigenetic biomarkers represents a rapidly developing frontier in clinical application improving the diagnosis and prediction of clinical outcomes[13]. The epigenome is constituted by a plethora of components and layered on different levels as described in the following paragraph.

1.1.1 Three-dimensional organization of the genome

The genome is topologically and hierarchically organized at multiple levels and length scales in the nucleus (Figure 1)[14].

This 3D architecture has emerged as being crucial for regulation of transcription, replication, DNA repair and splicing[14,15]. Chromatin, the ensemble of DNA and proteins, represents the basic-folded structure of genetic information. As a first compaction level, genomic DNA is tightly condensed together with a complex set of proteins called histones, generating the structural unit of chromatin named nucleosome. The nucleosome core particle consists in the wrapping of 145-147 base pairs (bp) of linear DNA around an octamer of four histone-fold heterodimers, two each of H2A/H2B and H3/H4[16]. These core

histones are basic, highly conserved proteins among eukaryotes. Nucleosomes are connected one another by linker histones, called H1, which stabilize DNA positioning and allow to establish condensed 10 nm chromatin fibers visually represented by the “beads-on-a-string” model[17].

The electrostatic interactions between nucleosomes induce self-interactions among chromatin fibers which folds into different size of loops ranging from kbs to Mbs[18,19]. At the smallest scale, it is commonly accepted that chromatin looping brings distal regulatory regions, often enhancer, spatially close with promoter of target genes incrementing their level of transcription[20].

Larger loops, of Mbs in length, not only contribute to chromatin compaction but also drive the fine regulation of gene clusters[21]. This mechanism has been shown in the regulation of developmental genes which require precise temporal and spatial expression[22]. The next layer of genome organization are chromatin domains, as reviewed in[21,23]. They were first visually described by in situ hybridization experiments[24] and subsequent experimentally identified by genome-wide sequencing technologies, like chromosome conformation capture (Hi-C)[25]. The Hi-C allowed the discovery of the Topological Associated Domains (TADs), chromatin regulatory domains that organize the chromatin folding and function[23,26]. TADs are typically several hundred kb in size and are extremely conserved across cell-type and species. TADs are characterized by a central region of high chromatin interactions where reside tissue-specific gene[26,27]. On the other hand, the boundary regions of TADs are enriched in

housekeeping genes and CTCF (CCCTC-binding factor) binding sites, a main insulator protein that preserves functional integrity of domains by blocking the spread of regulatory signals coming from the adjacent regions[26,27]. Conventionally, TADs are organized into transcriptionally active (A) compartments and inactive (B) compartments[28]. Active compartments are enriched in RNA Polymerase II and epigenetic marks for transcriptionally competent chromatin, while inactive compartments have a higher degree of chromatin compaction and are characterized by gene poor regions of transcriptionally silent genes[29]. The active compartments host the “expression hubs”, where transcribed genes and distant regulatory elements (enhancers) are brought together by the specific folding of the chromosome[26]. The higher-order chromatin domain folding allows to establish precise mutual positions of enhancers and promoters.

The last level of genome organization is the non-random position of chromosomes in the 3D space of interphase nuclei. In fact, the chromosomes occupy distinct nuclear regions, termed chromosome territories (CTs)[30,31].

Two typical examples are the well-characterized territories of human chromosomes 18 and 19 (HSA18 and HSA19). These chromosomes of comparable dimension (85 and 67 Mb, respectively) occupy opposite spaces in the nucleus: the gene-dense HSA19 (with a mean of 20.5 genes per Mbp) is localized in the nuclear center whereas the gene-poor HSA18 (4.3 genes per Mbp, as a mean) is characterized by an evolutionarily conserved peripheral nuclear localization[32,33]. Finally, recent

evidences show that a major contribution to intranuclear organization comes from so-called liquid-liquid phase separation (LLPS), a phenomenon driven by non-covalent intramolecular interaction that allows the formation of chromatin subcompartments in a self-organizing manner and in the absence of physical barrier, such as a membranes (for in depth details see LLPS paragraph)[34]. The chromatin architecture is not a static structure, in fact it undergoes dynamic remodeling during cell differentiation, accompanying the specification of cell fate and during cell life to adapt to the environmental changes.

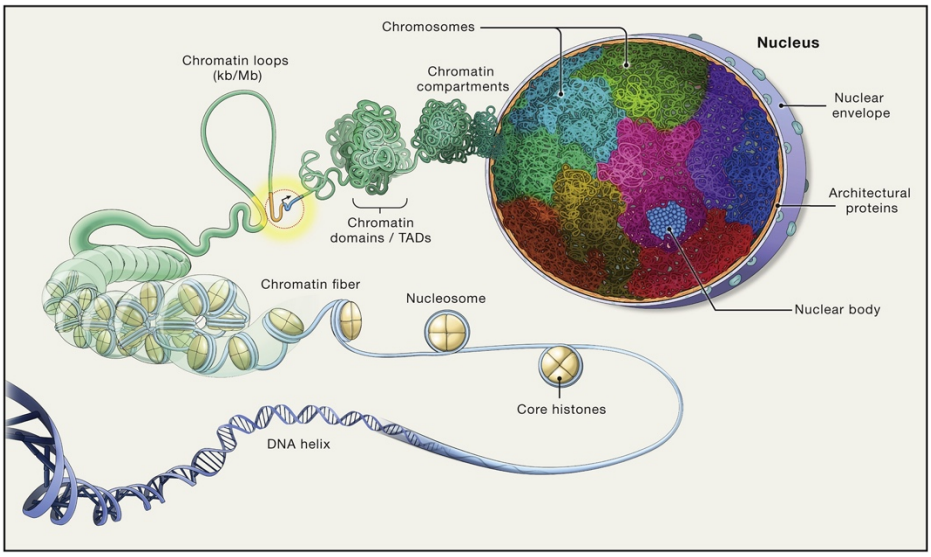


Figure 1 - The functional and hierarchical organization of the genome. Genomic DNA engages with multiple proteins and RNAs, forming the chromatin. The linear sequence of DNA regularly wraps around protein octamers called nucleosomes forming a 10-nm beads-on-a-string fiber which further folds to bring close together co-regulated regions, generating loops and, on a bigger scale, topologically associated domains (TADs) where many regions interacting among them cluster together. The chromosomes occupy specific nuclear regions known as chromosomes territories (CTs)[14].

1.1.2 DNA methylation and Histone Code

DNA methylation mostly occurs by covalent addition of a methyl group position C5 of cytosine (5 methylcytosine, 5mC), and constitutes the only covalent modification associated to the primary level of DNA sequence in mammals[35]. Operated by DNA methyl transferases (DNMTs), the presence of 5mC is almost completely confined to CpG dinucleotide contexts. This DNA epigenetic modifications is known to be involved in fundamental physiological processes, such as embryonic development, X chromosome inactivation, and genomic imprinting[36,37,38]. Cytosine methylation occurs on certain specific CpG sites, resulting in a cell-type, or developmental stage-specific pattern of DNA methylation[39]. Mutations in DNMTs and concomitant dysregulation of DNA methylation pattern often leads to murine embryonic lethality[40], and to a wide-spectrum of disorders in humans[41]. Additionally, alteration of the DNA methylation machinery is common in a variety of tumors, and its pivotal role in malignancy transformation and progression has become a topic of extensive interest in the last years[42].

Another epigenetic level of chromatin regulation is represented by the histones. In addition to their direct interaction with the DNA, histones are subject to post-translational modifications (PTMs) which provide a fine mechanism for the regulation of gene expression[43]. The nonglobular N-terminal protruding tails serve as hubs for wide array of modifications, including acetylation of lysines, methylation of lysines and arginines, phosphorylation of serines and threonines and ubiquitination of

lysines[43]. Histone tails are freely accessible to two major classes of enzymes that add (writers) or remove (erasers) PTMs. A third group of proteins, commonly recruited to the modified residues, are known as “readers”. Typically, writers and erasers also contain reader domains which determine substrate specificity and regulate the enzymatic activity[44]. PTMs can produce structural alterations inside the nucleosome, affecting inter-nucleosomal contacts, finally influencing genome conformation and gene expression[45]. Alternatively, they represent docking sites for various chromatin modifiers. It has been described that histone acetylation neutralize the positive charges of the N-terminal tails, thereby weakening electrostatic interactions with DNA[46] and increasing its accessibility[47]. On the contrary, de-acetylation catalyzed by the histone deacetylase (HDAC) is generally associated with transcriptionally inactive chromatin[48]. Histone lysine methylation can exist in three different states (mono-,di- or tri-methylation) and, according to the position on the histone tails, can mediate transcription activation or repression. H3K27 and H3K9 methylations, associated respectively with Polycomb regions[49] and nuclear lamina-associated domains[50], correlate with gene silencing. H3K9me3 is a high affinity docking site for heterochromatin proteins, as heterochromatin protein 1α (HP1α), that mediate the formation of constitutive repressed domains[51]. Polycomb repressive complex 2 (PRC-2)-mediated H3K27me3 marks dynamically regulated genes that need to be switched on and off in response to developmental and environmental signals[49]. On the other hand,

histone methylation has also been associated to activation: di- and tri-methylation at lysine 4 on histone H3 (H3K4me2/3) are specifically enriched at the proximity of transcription starting sites (TSSs) in active genes and correlates with promoter activation[52]; trimethylation of lysine 36 in histone H3 (H3K36me3) is specifically associated with actively transcribed gene bodies[53]; mono-methylation of H3K4 marks the enhancers[54,55]. Multiple PTMs coexisting on the same histone tail can produce different biological outputs, depending on the combination displayed, often defined as “chromatin state”[56]. The chromatin state allows a good prediction of the transcriptional activity of genomic regions. One example of this combinatorial histone marks has been firstly described in embryonic stem cells, where the two functionally opposite modifications, H3K4me3 and H3K27me3 can coexist on some gene promoters[57]. These sites, called “bivalent domains”, contain developmental and differentiation genes poised for activation[57].

1.1.3 Euchromatin and Heterochromatin

Based on the histone PTMs and the DNA packaging the chromatin exists in two main structural and functional conformations: heterochromatin (HT), which is the highly packaged and inaccessible organization enriched with repressed genes, and euchromatin (EC), which is the open-state chromatin associated with transcriptionally active regions[58]. Due to their characteristics EC and HC occupy different positions inside the nuclei; in particular, EC is found in the nuclear interior, while HC

is localized at the nuclear periphery and in small dots in the nuclear interior (Figure 2)[59,60]. Euchromatin encompasses different subtypes of condensates deputed to specific function. Two widely studied examples are the nucleolus[61] and the discrete sites of RNA polymerase clustering known as transcription factories[62,63]. Heterochromatin can be further divided into permanently constitutive (cHC)[64] and facultative heterochromatin (fHC)[65]. cHC is typically found at pericentromeric and telomeric regions of chromosomes where it contributes to chromosome segregation and telomeric protection processes[66]. It is also present in large domains positioned along the nuclear lamina (LADs, for in depth review see next paragraph)[50] or around the nucleoli (NADs)[67]. Facultative heterochromatin involves the action of Polycomb group (PcG) proteins[68], which organize in multimeric complexes and bind specific sites on the DNA, forming disseminated foci in the nucleus[69]. These proteins can mediate the formation of high order chromatin structures, which regulate gene expression and allow the transmission of the transcriptional program throughout the cell cycle[70]. Interestingly, direct visualization of DNA in the nucleus with nanometer-scale resolution has shown that Polycomb-repressed compartments exhibit more compact packing and higher degree of chromatin intermixing than constitutively inactive compartment[71].

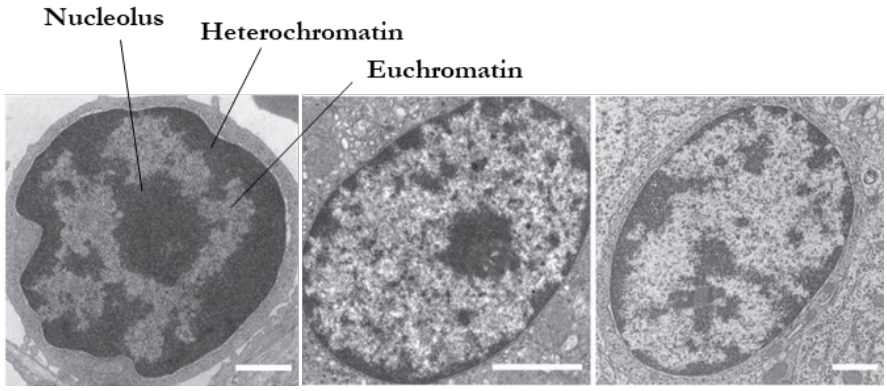


Figure 2 – Electron microscopy visualization of mammalian interphase nuclei. Electron micrographs of lymphocyte (left), epithelial cell (center), glial cell (right). Scale bars: 1 μm. Heterochromatin and nucleolus are darkly stained, while euchromatin is lightly stained[60].

1.1.4 Nuclear lamina

Nuclear Lamina (NL), a complex meshwork of protein localized under the inner nuclear membrane, not only contribute to nuclear structural and mechanical stability but also interacts with nearly 40% of the genome, governing the principal transcriptionally inactive nuclear compartment[72]. The lamin proteins are the main component of the NL and are classified as type V intermediate filament proteins[73,74]. Lamins can be divided further into A and B type, based on their sequence homology[75]. Whereas B-type lamins are ubiquitously expressed, lamin-A/C is absent in early embryos and developmentally regulated in lineage-committed cells and during differentiation[76]. Thus, lamins distribution in nuclei is cell-type specific, with lamin-B being predominantly found at the nuclear periphery, while lamin-A/C is also present in the nucleoplasm where it exists as a detergent-soluble pool

forming the nucleoplasmic lamin[77]. Lamins are involved in the spatial organization of the genome interacting with specific regions called lamina-associated domains (LADs)[50,72]. These large regions (from 0.1 Mb to 10 Mb, with a median of 0,5 Mb) suggest that LAD-NL interactions inflict constraints on the shape and location of chromosome at the nuclear envelope[78]. LADs are heterochromatic gene-poor regions, with low levels of transcription and enriched by the typical repressive epigenetic markers like H3K9me2, H3K9me3 and H3K27me3[79]. Additionally, LADs can be classified as constitutive LADs, highly conserved among species, and facultative LADs, variable and cell-type specific[80].

1.1.5 Liquid-Liquid Phase Separation (LLPS)

In recent years, liquid-liquid phase separation (LLPS), the phenomenon driving the formation of membraneless organelles by macromolecules organization, has been shown to play a key role in the spatio-temporal compartmentalization of the genome (Figure 3)[81,82,83,84]. The micro-scale non-membrane-bound compartments appears as droplets within the cell and are, collectively, known as biomolecular condensates[85]. Each condensate demonstrates a distinct phase ensembled and maintained by multivalent interactions among molecules. Proteins that drive phase separation are composed of modular interaction domains including self-interacting and/or protein interaction domain and nucleic acid binding domains[81]. Additionally, phase-separating proteins are typically interspersed

with large intrinsically disordered regions (IDRs) that provide multiple weakly adhesive sequence elements of interaction with multiple biomolecular species[86]. RNA and DNA, which harbor multiple binding regions with other nucleic acid and proteins, also contribute to phase separation[87,88]. Nuclear condensates are different in their dimensionality, physical properties, biological molecular composition and function and are emerging as a new epigenetic level of chromatin regulation[85]. In the nucleus, well-studied examples of phase separated structures are nucleolus[89], nuclear speckles[90] and histone locus bodies[91]. Recently, Strickfaden et al. indicate that the nucleosomal arrays, with no requirement for nonhistone proteins, within the chromatin condensates are packaged in a constrained, solid-like state[92]. In another study, it has been shown that the C-terminal tail of linker histone H1 forms a complex coacervate with DNA in vitro thereby promoting the formation of chromatin condensates[93]. Several nuclear factors have recently been shown to exhibit LLPS capacity, among which is BRD4 (Bromodomain Containing 4), the major bromodomain containing transcription factor. It has been shown that BRD4 bind to acetylated nucleosomes, forming a new phase-separated state and thereby regulating chromatin compartmentation via LLPS[94,95]. Additionally, acetylation of the core histone tails leads to droplet dissolution facilitating the activation of transcription[95]. BRD4-induced phase separation is impeded by the bromodomain inhibitor drug JQ1[96].

Many proteins associated with the transcription machinery, including RNA polymerase II (Pol II), contains multiple IDRs and

form transcriptional condensates by LLPS at the promoter, but also at the super enhancers [94,97]. An example of heterochromatin phase separation includes H3K9me3 reader protein HP1. Modulated by phosphorylation, HP1 interact with SUV39H1 (H3K9me2/3 writer) and TRIM28 (scaffolding protein) forming droplets in vitro that resemble properties of heterochromatin[98]. Moreover, it was found that the conserved sterile alpha motif (SAM) of the Ph subunit of Polycomb repressive complex 1 (PRC-1) can undergo LLPS, forming condensates that mediate the chromatin compaction, enhancing PRC1 activity[99].

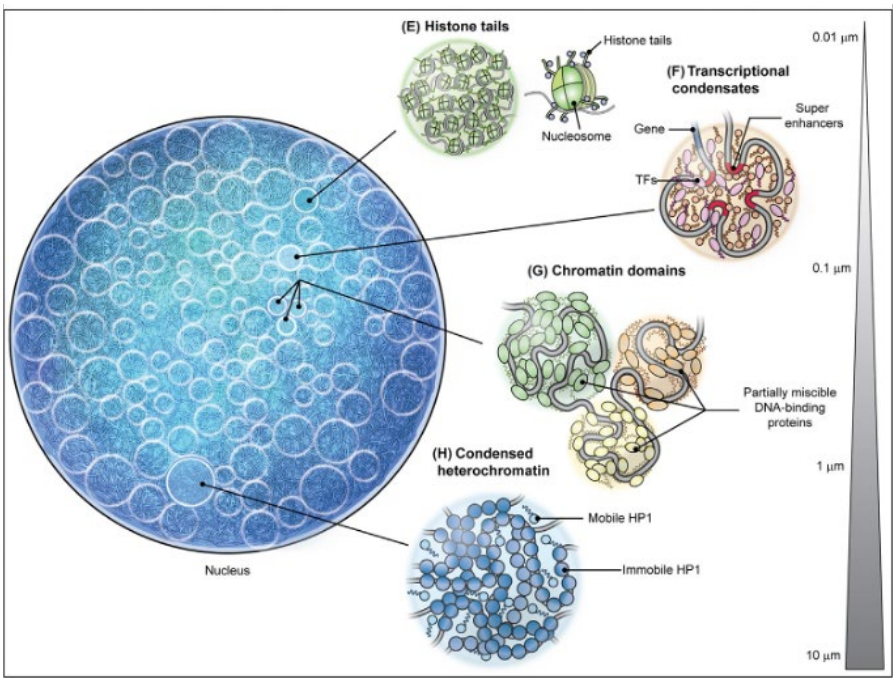


Figure 3 - Phase separation of the mammalian nuclear genome. Levels of LLPS in the nucleus include histone tails (**E**), transcriptional condensates (**F**), partial miscibility between neighboring chromatin domains (**G**), and heterochromatin blocks (**H**). Bar indicates length scale of interactions[81].

1.2 Nuclear abnormalities in cancer

Morphological changes in nuclear organization represent hallmarks of cancer aggressiveness[100,101]. The nuclear atypia remains a gold standard for diagnosing and staging of different cancers, as extensively described in breast, lung and prostate cancer[100]. The microscopic examination of cancer cells allows the identification of alterations in the nuclear architecture. The common structural changes in cancer consist of nuclear dysmorphia, prominent nucleoli, “hyperchromatic” chromatin, and high nuclear/cytoplasmic ratio (Figure 4)[101]. Of note, not all these abnormalities are observed simultaneously but specific changes could be characteristic of different cancer types[101]. The typical irregular nuclear outline of cancer nuclei is mainly the result of invaginations and convolutions of the nuclear lamina[102]. The perinucleolar compartment (PNC), which form a meshwork closely associated with the nucleolus, is rarely detected in normal nuclei, but it is present in different cancer type, as breast, thyroid and prostate tumors[103]. Nuclear abnormalities are often associated with altered heterochromatin organization. In colorectal cancer (CRC) has been demonstrated that neoplastic nuclei exhibit a loss of characteristic peripheral localization of heterochromatin and the acquisition of coarse and asymmetric aggregation of densely packed chromatin[104]. In line with this, the repositioning of LADs influences the activation of oncogenes or the inactivation of tumor-suppressor programs finally influencing cancer cell’s fate[79]. The promyelocytic leukemia (PML) nuclear bodies (NBs), involved in apoptosis and tumor suppression

pathways, are mislocalized and misregulated not only in leukaemia but also in various solid tumors[105]. In particular, the disruption of PML bodies correlates with the invasiveness of prostate and breast cancer[106]. Although the nuclear abnormalities in cancer cells are well documented, the prevalence and the functional consequences of such nuclear alterations remain poorly characterized. It is widely accepted that changes in nuclear morphology, visible under a microscope, affect chromatin organization and gene positioning, with resultant changes in transcriptomic programs. Therefore, one of the major challenges in cancer research is to link these structural changes with genome functions.

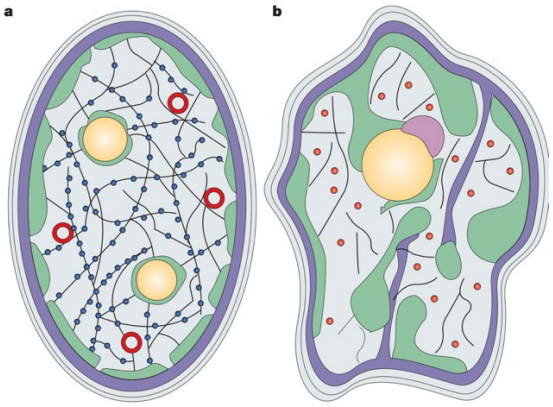


Figure 4 - Nuclear organization in normal and cancer cells. **a.** Normal nucleus. Nuclear lamina is marked in purple, the heterochromatin in green, while the nucleoli in yellow. The nuclear matrix (black internal network with nuclear matrix proteins as blue circles) is the non-chromatin scaffolding that provides spatial organization of nuclear substructures like PML bodies (red). **b**. Cancer nuclei are characterized by irregular morphology and coarse heterochromatin aggregates. Nucleoli are often enlarged and nuclear bodies mislocalize in microspeckles. Perinucleolar compartment (PNC, pink) is observed in cancer nuclei[101].

1.3 Prostate cancer

Prostate cancer (PCa) is the second most common cause of male cancer-related mortality. It has been ranked as one the most frequent new cancer diagnoses in males, with a typical late age-of-onset (median 66 years of age)[107]. Multiple predisposing factors, including age, family history, genetic susceptibility, race and lifestyle, contribute to the high incidence of prostate cancer[108,109]. Prostate cancer is usually slow-growing, and symptoms may not occur in its early stages. It is estimated that two-thirds of the newly diagnosed patients are asymptomatic[110]. The prostate is an androgen-dependent tissue involved in male reproductive function. Normal prostate tissue consists of tubular-alveolar epithelia glands surrounded by fibromuscular stroma[111]. The prostate ducts are normally made up of two well-organized layers of epithelial cells, with neuroendocrine cells dispersed throughout these layers. The secretory luminal cells directly line the lumen of the prostate ducts, while the basal cells lie on the basement membrane. The vast majority of PCa arises from the epithelial cells and are therefore defined as adenocarcinomas[112]. Histological studies have identified that the prostate cancer starts from specific pre-neoplastic lesions, known as prostatic intraepithelial neoplasia (PIN)[113]. This prostate cancer precursor is characterized by a luminal epithelial hyperplasia together with a reduction in basal cells while preserving the basement membrane. The transition from PIN lesions to adenocarcinomas implies the loss of the basal cell layer and basement membrane[114]. Most of the prostate cancer tend to be quite

indolent growing so slowly that it will not become life-threatening, rendering the early diagnosis quite difficult. However, a small subset of patients progresses to an aggressive and metastatic state[115] (Figure 5). Prostate cancer cells first spread to the regional lymph nodes and then preferentially metastasize to the bone[116]. The overall survival of patients with distant metastatic disease is variable, ranging from a few months to a few year (5-year survival rate of 30.5%)[107]. The prostate is divided in three morphologically zones, the peripheral (PZ), central (CZ) and transition zone (TZ), each having a different PCa incidence[112]. The majority of prostatic adenocarcinomas arise from the peripheral zone (almost 70%), which is also highly susceptible to inflammation, while BPH typically occur in the transition zone[117]. Additionally, tumors originate in PZ are associated with poorer clinical outcomes than TZ tumors[118].

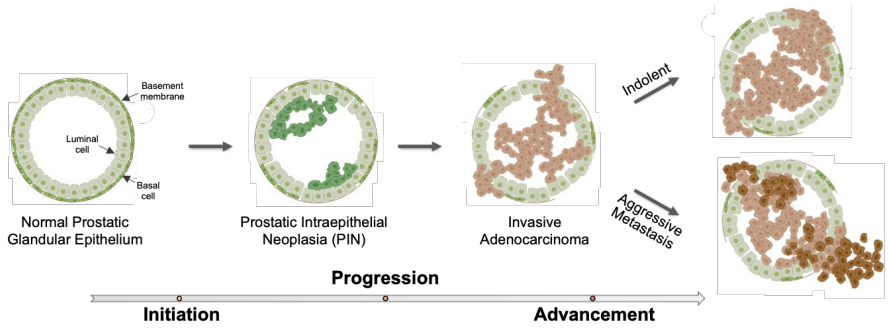


Figure 5 - Prostate cancer developmental stages. PCa initiates from a prostatic intraepithelial neoplasia (PIN) until it becomes an invasive adenocarcinoma. Then, it could be a slow growing and quite indolent cancer or evolve into an aggressive and metastatic state (picture created in BioRender.com).

1.3.1 PCa diagnosis

The current PCa screening guidelines are based on digital rectal examination (DRE) and serum levels of the prostate-specific antigen (PSA)[119]. A DRE checks directly anomalies in the consistency and size of the prostate gland, while elevated PSA levels, more than 4 ng/ml, indicate a greater likelihood of prostate cancer[120]. However, other common and non-malignant conditions like benign prostatic hypertrophy (BPH), prostatic inflammation, and even urinary tract infections raise the quantity of this antigen in the blood. Hence, the lack of specificity and sensitivity of PSA screening has generated considerable controversy within the scientific community[121]. To date, PSA test is not enough to diagnose prostate cancer and requires interpretation with the integration of the clinical situation. Several studies have identified some PCa-specific biomarkers from blood and urine sampling, but so far none of these biomarkers is widely used in clinical practice[122].

Currently, transrectal ultrasound-guided (TRUS) needle biopsy of the prostate and histological control still represents the gold standard for definitive diagnosis and cytological grading of PCa[123]. However, it is an invasive procedure and still multiple cores are required for histopathological examination. Moreover, the number of biopsy samples necessary for the optimal detection of prostate cancer is still debated[124]. One of the major difficulties in detecting significant PCa is its frequent multifocality; contrary to the other types of solid carcinomas, most primary prostate cancer consists of multiple tumours foci within the

prostate[125]. Several studies have demonstrated that increasing the number of samples taken improve the tumor detection rate[126,127], but it also increases the risk of side-effects[128]. It is generally accepted that 10-12-core systematic sampling protocol is sufficient for detecting significant PCa (Figure 6)[123].

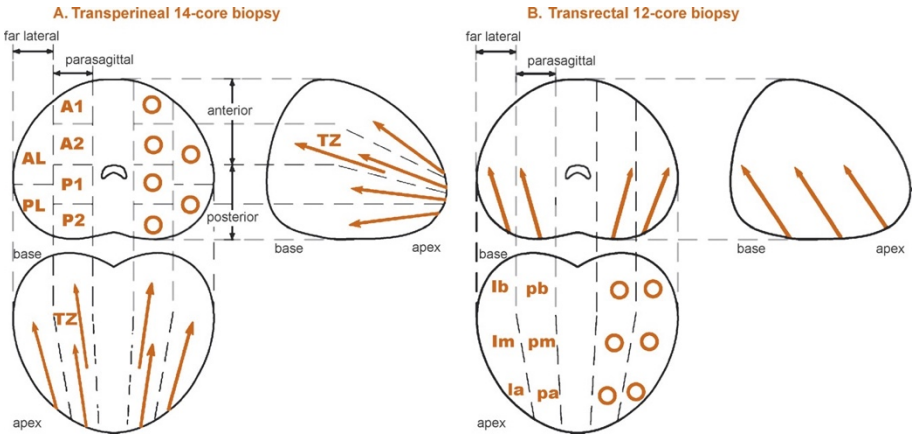


Figure 6 – TRUS biopsy scheme.
Transverse, sagittal, and coronal projections of the 12-core (**A**) and 14-core (**B**) prostate biopsy scheme. Transrectal 12-core biopsy including conventional 6 sextant cores and 6 cores collected from the peripheral zone. Transperineal 14-core biopsy including 12 cores from peripheral zone and 2 cores taken from the transition zone (TZ). The sampling sites are indicated as follows: anterior (A1, A2), posterior (P1, P2), anterolateral (AL)[129].

1.3.2 Pathology and staging of PCa

The most commonly used system to assess the histological grading of PCa is the Gleason score[130]. The tissue architecture is graded on a scale of 1 to 5, with higher numbers indicating higher abnormality and aggressiveness (Figure 7). To be considered, a grade pattern needs to occupy more than 5% of the biopsy tissue. Gleason score (GS) of biopsy-detected PCa derives from the sum of the two most extensive pattern and provides clinical implication for treatment and prognosis[131]. Additionally, the TNM system is also widely used to clinical cancer staging; it refers to local tumor growth (T), spread to regional lymph nodes (N) and distant metastases (M)[132].

In clinical practice, both Gleason grade and TNM stage are used as predictors for prognosis; however, there is highly variable clinical course among primary PCa. As mentioned above, the majority of localized prostate cancer is quite indolent and is cured with initial therapy or may be safely monitored (active surveillance), while a small subset of patient progresses to an aggressive and metastatic state[115,133].

Several factors have been described to be involved in the development of advance prostate cancer[115,134]; however, it is still unknown what exactly drives prostate cancer towards an aggressive and fatal form. For this reason, multiple risk stratification tools have been proceeded, combining the currently available clinical and pathological features (PSA levels, Gleason score and TNM classification)[135]; although these systems still do not uniquely predict PCa outcome.

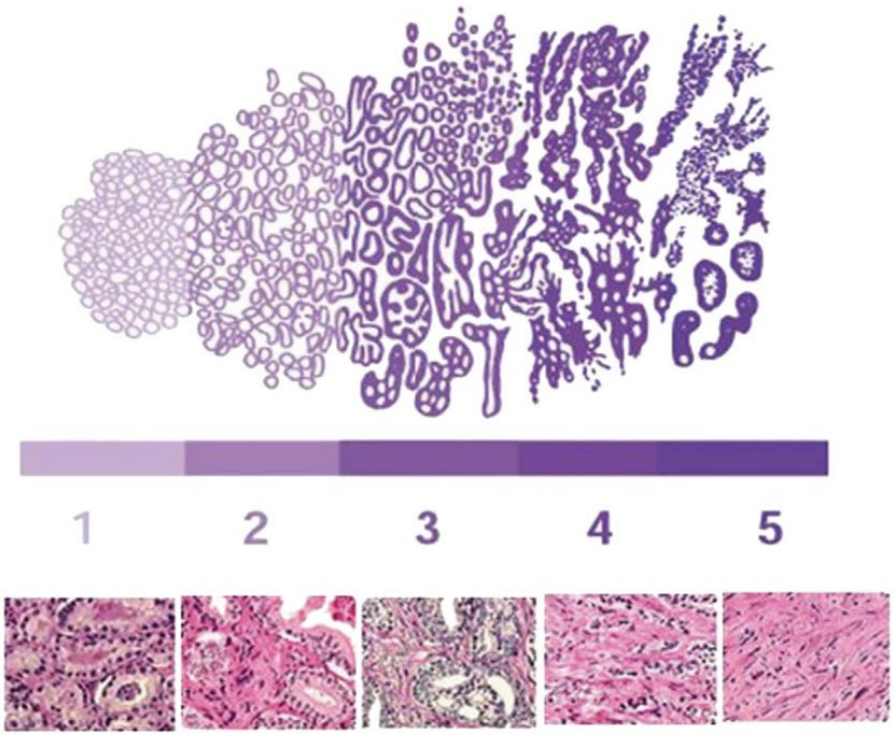


Figure 7 – Schematic diagram of Gleason grading system. Numbers indicate Gleason grades. The original Gleason diagram of each grade (above) and the corresponding stained micrographs (below)[136].

Knowledge on risk factors required for distinguish indolent from aggressive prostate cancer represents an important public health need. Comprehensive analyses of PCa genome and epigenome are increasingly growing to subtype clinically heterogeneous cancers and to guide more precisely targeted therapeutic strategies. In the following sections, an overview of well-established molecular genomic and epigenomic alterations of PCa will be presented.

1.4 The genomic landscape of PCa

Large-scale molecular characterization projects of primary and metastatic prostate cancer have provided critical new insight into the genomic drivers of prostate cancer progression. In contrast to other solid tumor types, primary prostate cancer harbors a paucity of driver single-nucleotide variants (SNVs) but exhibits multiple recurrent structural variants including DNA copy number aberrations (CNAs) and genomic rearrangements (GRs)[137, ,138,139]. One of the earliest rearrangements in prostate cancer genomes are fusions of androgen-responsive promoters with E26 transformation-specific (ETS) family of transcription factors, most notably the TMPRSS2:ERG fusion (T2E)[140]. This T2E fusion occurs in approximately 50% of prostate cancers, while another 10% harbor a fusion involving other ETS-family proto-oncogenes (e.g., ETV1, ETV4)[137]. Several studies examining fusion-bearing tumors have been reported a correlation between TMPRSS2-ERG rearrangement and higher Gleason grade and clinical stage[141,142].

Multiple genomic loci are recurrently amplified or deleted in PCa. These structural alterations accumulate early in prostate carcinogenesis and include losses of NKX3.1[143] and PTEN[144], gains of the androgen receptor gene (AR)[145]. A high level of copy number alterations (CNAs) in prostate primary tumors is associated with the progression toward aggressive state[137]. As mentioned above, several well-powered whole-exome sequencing (WES) studies have revealed few point mutations at rates exceeding 5% in localized PCa[146]. The most frequently

mutated genes in localized prostate cancers are *SPOP* and *FOXA1* which encode transcriptional regulators involved in DNA damage and cell proliferation signaling pathways[147,148,149]. Point mutations in the canonical tumor suppressor gene, TP53, are also relatively common in localized prostate cancer (3% recurrence)[146].

The more aggressive and metastatic prostate tumors are associated with higher mutational burden and genomic instability[150] which commonly drive androgen-independent growth[151,152]. Indeed, primary PCa is often hormone-dependent and critically relies on androgen receptor (AR) signaling[153].

In the nucleus, the androgen-activated receptor binds specific regions of the genome, called androgen-response elements, regulating the expression of pivotal genes involved in cellular growth and survival[154]. The majority of patients is responsive to front-line androgen deprivation therapy (ADT)[155], although approximately 20% of cases develops treatment resistance and progress to the incurable stage, also known as castration-resistant prostate cancer (CRPC)[156,157]. Several mechanisms underlie the development of CRPC, including AR point mutation (15% to 30% recurrence)[158], AR amplification (30% to 50% recurrence)[159] and constitutively active AR splicing variants (e.g., AR-V1 and AR-V7)[160].

1.5 Epigenetic alterations in PCa

Recently, the continued advances in technologies, such as epigenome-wide mapping and fluorescence microscopy studies, have enriched our understanding of the epigenomic landscapes in prostate cancer (Figure 8)[161,162,163].

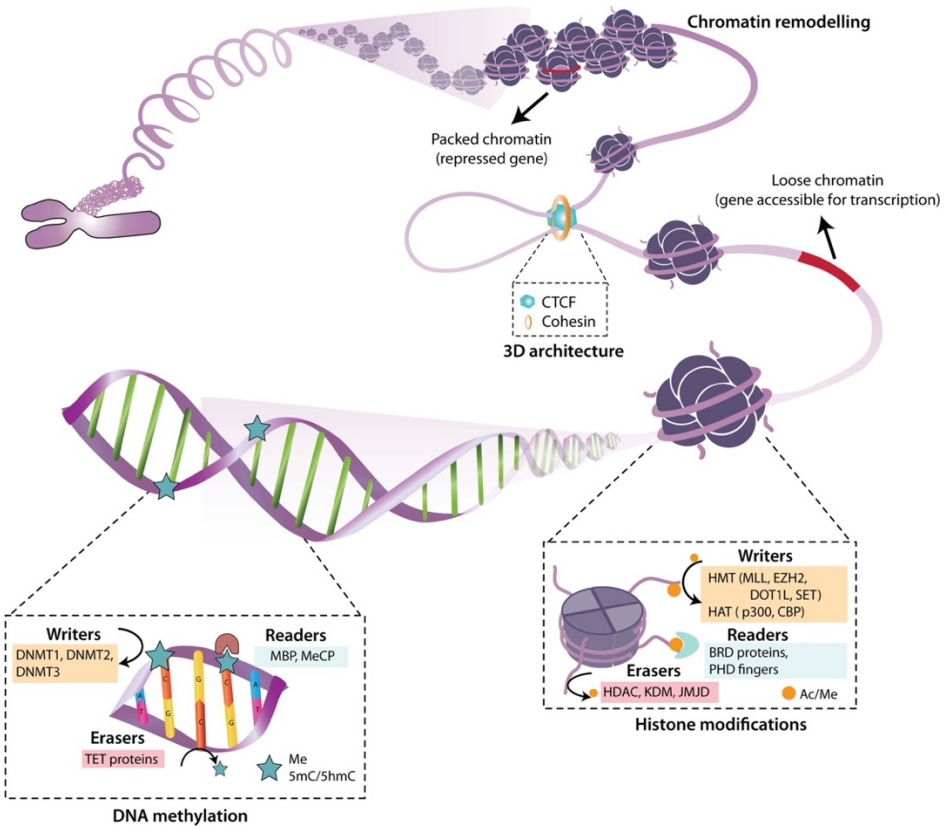


Figure 8 – Epigenetics alterations in prostate cancer cells. Schema showing pivotal epigenetic alterations observed in PCa at the three regulatory levels including DNA methylation, histone modification and chromatin remodelling[163]

1.5.1 DNA methylation

Global changes in DNA methylation are hallmarks of all cancer and are extensive studied also in PCa[164]. Prostate cancer nuclei simultaneously exhibit global hypomethylation and gene-promoter-specific hypermethylation. The second is associated with transcriptional silencing, providing a different mechanism for the inactivation of tumor-suppressor genes. Overall, more than 50 genes have been reported to be frequently subjected to promoter hypermethylation during prostate carcinogenesis[137,165]. Among them, keys regulatory genes controlling the cell cycle and DNA repair, such as APC, BRCA1/2, PTEN and GSTP1 were reported. On the basis of these studies, the differential methylation pattern has been proposed as a novel non-invasive diagnostic biomarker for prostate cancer[166]. In particular, the hypermethylation in the regulatory region of GSTP1, occurring in more than 90% of prostate tumor, is highly explored for its potential diagnostic and prognostic value[164,167].

The variable methylation profiles in PCa cells are explained by the increased activity of methyltransferases (DNMT1 and DNMT2) and decreased expression of demethylases such as methyl-CpGbinding protein (MBD2) and Ten-eleven translocation (TET) proteins[168].

1.5.2 Histone Modification Machinery

Prostate cancer is characterized by widespread changes of histone modifications which might underlie its heterogeneity and clinical variability[169]. Several epigenetic components involved in

chromatin conformation organization are dysregulated in PCa. Among them, the repressive activity of PcG complexes plays a key role in prostate cancer aggressiveness and resultant clinical outcome[170,171]. Ezh-2, the catalytic subunit of Polycomb repressive complex 2 (PRC-2), is upregulated in metastatic CRPC and its overabundance is detected in primary tumors with a higher risk of recurrence after radical prostatectomy[170,172]. PRC2 dysregulation leads to H3K27me3-mediated silencing of tumor suppressors, including ADRB2, CDH1, SLIT2, and TIMP2/3, finally stimulating cell growth and tumor progression[172,173]. In addition to the well-described histone methylation catalytic function, a novel PRC2-independent role of Ezh-2 as a AR transcriptional activator has been identified in advanced prostate cancer[174,175]. An interesting study, also confirmed a PRC2-independent interaction between AR and Ezh-2 that activates AR transcription in CRPC cell lines, enhancing AR signaling, and triggering the resistance to enzymatic EZH2 inhibitors[172]. The key step for the functional switch of Ezh-2 to a PRC2-independent function resides in the phosphorylation of serine 21 (pS21) mediated directly or indirectly by PI3K/AKT pathway. This modification plays a pivotal role in Ezh2-mediated androgen-independent growth; indeed, the level of both Ezh-2 and pS21-Ezh-2, are especially increased in CRPC[175]. Recently, it has been reported FOXA1, a pioneer factor of the androgen receptor, as a nonhistone substrate of PRC2-mediated methylation[176]. Ezh-2 methylates FOXA1 at lysine-295 which recruits deubiquitinase to block FOXA1

degradation, thereby increasing its transcriptional regulation of genes related to cell cycle process and apoptosis and promoting tumor growth[176]. Polycomb repressive complex 1 (PRC-1) is also involved in PCa progression, and its oncogenic activity depends on Bmi-1 which is a central player for cancer initiation and progression through regulation of stem-cells self-renewal pathways[177]. Like Ezh-2, Bmi-1 can be a transcriptional activator of AR by acting independently from the PRC-1 complex[178]. In fact, Bmi-1 can bind and stabilize the interaction with the AR, inhibiting the degradation pathways of AR and inducing the transcription of its target genes[178]. Other remarkable epigenetic changes in prostate cancer involve histone acetylation. It has been reported that histone acetyltransferase p300 and its highly homologous CREB-binding protein (CBP) are implicated in the prostate cancer aggressiveness and that deletion of p300 in mice reduces tumor progression and improves survival[179]. The oncogenic roles of p300/CBP in the progression of PCa are usually related to enhancing AR transcriptional activity and preventing its degradation, consequently promoting tumor growth[180]. Recent studies have shown that targeting p300/CBP using novel small-molecule inhibitors, reduced both androgen-responsive and CRPC cancer cells growth *in vitro* and *in vivo* [181,182]. In addition to chromatin modifying enzymes, the epigenetic readers such as the bromodomain proteins which read lysine acetylation marks can also contribute to prostate cancer. BRD4, as described above, is a well-reported member of the bromodomain and extra terminal (BET) family and it is an

epigenetic activator of special relevance to prostate cancer progression. The over-expression of BRD4 is associated with metastatic invasion and clinical stage of PCa[183,184]. A number of small-molecules BET inhibitors, such as JQ1, have revealed great promise as therapeutic agents in CRPC[185]. In cell line models of CRPC, JQ1 has been reported to repress AR-V7 expression[186] and AR-mediated gene transcription programs[187]. Currently, at least 10 BET inhibitors are used in clinical trials[188]. The repositioning of epigenetic marks at distinct loci suggests that the epigenetic context of the genome is necessary for specific tumorigenic functions.

1.5.3 3D genome structure

Only recently the dysregulation of chromatin architecture has been identified as a major player in the initiation and progression of human cancer[189]. An example is Colorectal Cancer that exhibits profoundly alteration in the spatial partitioning of closed and open genome compartments[190]. Additionally, interesting studies have shown that focal topological alterations rewire specific enhancer–promoter interactions, impacting gene expression toward oncogenic programs in gastrointestinal stromal tumors and acute lymphoblastic leukemia[191,192]. Of note, prostate tumor commonly harbor mutations and copy number loss of CTCF and other cohesion-interacting proteins involved in the establishment and maintenance of higher-order chromosomal structure[193,194]. The resultant chromatin disruption may lead to inappropriate 3D looping that drive oncogenic

program. The recent advent of new methods, the C-Technologies, able to detect the conformation of the genome[195], revolutionized the field of chromatin structure and allowed to evaluate the impact of chromatin structural organization in cancer progression. The boundaries of megabase-sized topologically associated domains (TADs) are conserved across cell types[196] and cancer cells retain the ability to segment their genome into TADs[197]. However, several studies have identified PCa-specific enhancer-promoter chromatin loops and transcription factors (such as pioneering factors like FOXA1) specifically enriched in these chromatin regulatory structures[198,199]. Moreover, PCa risk region located at 7p15.2 is involved in a repressive regulatory loop that control the expression of HOXA13 and HOTTIP In the HOXA locus[200]. Certain TADs change in size and epigenetic states in prostate cancer cells compared to normal cells[199]. The transition between epigenetic state affects the general expression level of genes within a given TAD finally promoting the oncogenic transcriptional programs. In particular, the AR locus exhibits multiple cancer-specific chromatin loops[199]. Further studies have identified several gene clusters that switch from the inactive B compartment to the active A compartment in prostate cancer cell line models[201]. Interestingly some structural changes including CDK14, AR and TMPRSS2-ERG locus, are persistent throughout the cell line metastatic models, thus highlighting the pivotal role of genome structure changes in prostate cancer aggressiveness[201].

1.6 Epithelial-to-Mesenchymal Transition (EMT)

As mentioned above, the majority of PCa-related deaths occur due to dissemination of cancer cells to distant metastatic sites. Epithelial-to-Mesenchymal Transition (EMT) represents a central mechanism for cancer cells migration[202].

EMT is composed of serial phenotypic changes through which epithelial cells alter their apical-basal polarity while acquire mesenchymal properties, enhancing cell motility (Figure 9)[203].

One of the critical molecular feature of EMT is the downregulation of cell adhesion molecules such as E- cadherin (CDH1) essential for the maintenance of epithelial integrity[204].

Moreover, EMT is promoted by various signaling pathways which often involved the activity of transforming growth factor β (TGFβ)[205]. The activity of a core set of transcription factors is also involved in EMT, like ZEB1 and ZEB2, the zinc finger Snail (SNAI1 and SNAI2), and the basic helix-loop- helix (Twist1 and Twist2)[202,206,207]. The transcription activator BRD4 directly regulates the expression of EMT-related transcription factors like SNAI1 and SNAI2 in metastatic castration resistant prostate cancer[184].

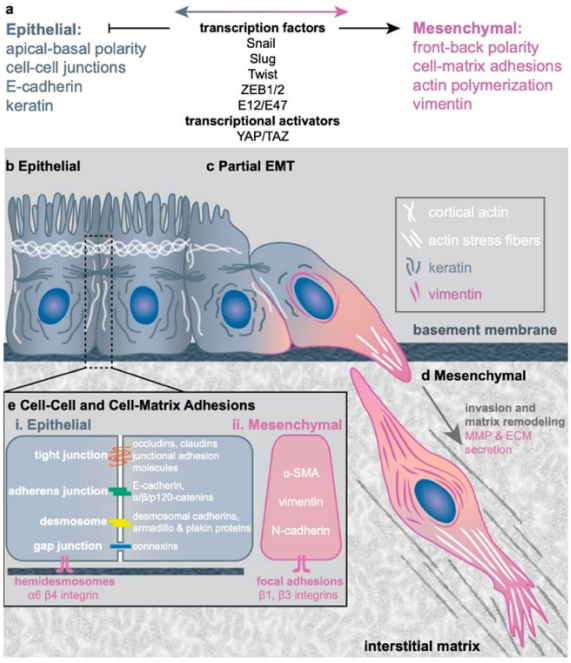


Figure 9 – Schematic of EMT progression. Phenotype of epithelial and mesenchymal cells and TFs that drive EMT (**a**). Morphology and cytoskeletal organization for epithelial (**b**), "partial" EMT (**c**), and mesenchymal cells (**d**). Detailed descriptions of cell–cell and cell–matrix adhesions (**e**)[203].

Recently, it has been shown that also the epigenetic regulation of genes involved in extracellular matrix (ECM) remodeling, such as MMP9 and ADAM19, contributes to EMT[208]. In line with this observation, prostate cancer cells undergoing EMT exhibit extensive nuclear morphology and texture alterations[209]. Interestingly, it has been reported that the epigenetic reprogramming during EMT is mediated by wide-spread chromatin changes mainly specific to Large Organized Chromatin K9-modifications (LOCKs)[210].

1.7 Scope of the thesis

In this work, we are proposing to investigate the complex nuclear architecture dynamics in different cell-type populations. Moreover, we aim to delineate some chromatin pathological remodelling that lead to genome dysfunction.

Our research efforts are identified in the following chapters:

Chapter 2 – in our first project, we focus on muscle stem cells (MuSCs) of a murine model of Emery-Dreifuss muscular dystrophy (EDMD) lacking the structural protein Lamin A/C. We show that in dystrophic MuSCs the uncontrolled spreading of Polycomb group (PcG) of proteins repressors from their targets triggers an aberrant differentiation program. The dysfunctionality in chromatin organization produces premature exhaustion of quiescent MuSCs and lack of cell identity, finally contributing to the disease progression.

Chapter 3 - we present advances in our chromatin fractionation technique, named 4 fractions Sequential Analysis of MacroMolecules accessibilitY (4fSAMMY-seq). With a unique protocol we are able to distinguish distinct soluble-associated fractions of euchromatin and heterochromatin (Manuscript in preparation). We used the novel 4fSAMMY-seq technique to dissect the chromatin structure alterations in primary prostate cancer using needle biopsy tissues.

References to chapter 1

1. D'Urso, A. & Brickner, J. H. Mechanisms of epigenetic memory. *Trends Genet* **30**, 230–236 (2014).

2. Schoenfelder, S. *et al.* Divergent wiring of repressive and active chromatin interactions between mouse embryonic and trophoblast lineages. *Nat Commun* **9**, 4189 (2018).

3. Gorkin, D. U., Leung, D. & Ren, B. The 3D genome in transcriptional regulation and pluripotency. *Cell Stem Cell* **14**, 762–775 (2014).

4. Takata, H. *et al.* Chromatin compaction protects genomic DNA from radiation damage. *PLoS One* **8**, e75622 (2013).

5. Piskadlo, E., Tavares, A. & Oliveira, R. A. Metaphase chromosome structure is dynamically maintained by condensin I-directed DNA (de)catenation. *Elife* **6**, e26120 (2017).

6. Ganesan, A. Epigenetics: the first 25 centuries. *Philos Trans R Soc Lond B Biol Sci* **373**, 20170067 (2018).

7. Dilworth, Fj. & Blais, A. Epigenetic regulation of satellite cell activation during muscle regeneration. *Stem Cell Research & Therapy* **2**, 18 (2011).

8. Pagiatakis, C., Musolino, E., Gornati, R., Bernardini, G. & Papait, R. Epigenetics of aging and disease: a brief overview. *Aging Clin Exp Res* **33**, 737–745 (2021).

9. Bianchi, A. *et al.* Dysfunctional polycomb transcriptional repression contributes to lamin A/C–dependent muscular dystrophy. *J Clin Invest* **130**, 2408–2421 (2020).

10. Jones, P. A. & Baylin, S. B. The Epigenomics of Cancer. *Cell* **128**, 683–692 (2007).

11. Selcuklu, S. D. & Spillane, C. Translational epigenetics: clinical approaches to epigenome therapeutics for cancer. *Epigenetics* **3**, 107–112 (2008).

12. Lu, Y. *et al.* Epigenetic regulation in human cancer: the potential role of epi-drug in cancer therapy. *Molecular Cancer* **19**, 79 (2020).

13. Berdasco, M. & Esteller, M. Clinical epigenetics: seizing opportunities for translation. *Nat Rev Genet* **20**, 109–127 (2019).

14. Misteli, T. The Self-Organizing Genome: Principles of Genome Architecture and Function. *Cell* **183**, 28–45 (2020).

15. Girelli, G. *et al.* GPSeq reveals the radial organization of chromatin in the cell nucleus. *Nat Biotechnol* **38**, 1184–1193 (2020).

16. Felsenfeld, G. & Groudine, M. Controlling the double helix. *Nature* **421**, 448–453 (2003).

17. Fyodorov, D. V., Zhou, B.-R., Skoultchi, A. I. & Bai, Y. Emerging roles of linker histones in regulating chromatin structure and function. *Nat Rev Mol Cell Biol* **19**, 192–206 (2018).

18. Dekker, J. & Misteli, T. Long-Range Chromatin Interactions. *Cold Spring Harb Perspect Biol* **7**, a019356 (2015).

19. Vermunt, M. W., Zhang, D. & Blobel, G. A. The interdependence of gene-regulatory elements and the 3D

genome. *Journal of Cell Biology* **218**, 12–26 (2018).

20. Heintzman, N. D. & Ren, B. Finding distal regulatory elements in the human genome. *Curr Opin Genet Dev* **19**, 541–549 (2009).

21. Pombo, A. & Dillon, N. Three-dimensional genome architecture: players and mechanisms. *Nat Rev Mol Cell Biol* **16**, 245–257 (2015).

22. Darbellay, F. *et al.* The constrained architecture of mammalian Hox gene clusters. *PNAS* **116**, 13424–13433 (2019).

23. Dixon, J. R., Gorkin, D. U. & Ren, B. Chromatin Domains: The Unit of Chromosome Organization. *Molecular Cell* **62**, 668–680 (2016).

24. Shopland, L. S. *et al.* Folding and organization of a contiguous chromosome region according to the gene distribution pattern in primary genomic sequence. *Journal of Cell Biology* **174**, 27–38 (2006).

25. Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009).

26. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).

27. Bianchi, A. & Lanzuolo, C. Into the chromatin world: Role of nuclear architecture in epigenome regulation. *AIMSBPOA* **2**, 585–612 (2015).

28. Zheng, H. & Xie, W. The role of 3D genome organization in

development and cell differentiation. *Nat Rev Mol Cell Biol* **20**, 535–550 (2019).

29. Cremer, T. *et al.* The 4D nucleome: Evidence for a dynamic nuclear landscape based on co-aligned active and inactive nuclear compartments. *FEBS Lett* **589**, 2931–2943 (2015).

30. Cremer, T. *et al.* Rabl's model of the interphase chromosome arrangement tested in Chinese hamster cells by premature chromosome condensation and laser-UV-microbeam experiments. *Hum Genet* **60**, 46–56 (1982).

31. Cremer, T. *et al.* Role of chromosome territories in the functional compartmentalization of the cell nucleus. *Cold Spring Harb Symp Quant Biol* **58**, 777–792 (1993).

32. Croft, J. A. *et al.* Differences in the Localization and Morphology of Chromosomes in the Human Nucleus. *Journal of Cell Biology* **145**, 1119–1131 (1999).

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

34. Erdel, F. & Rippe, K. Formation of Chromatin Subcompartments by Phase Separation. *Biophysical Journal* **114**, 2262–2270 (2018).

35. Feng, S., Jacobsen, S. E. & Reik, W. Epigenetic reprogramming in plant and animal development. *Science* **330**, 622–627 (2010).

36. Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**,

247–257 (1999).

37. Pontier, D. B. & Gribnau, J. Xist regulation and function explored. *Hum Genet* **130**, 223–236 (2011).

38. Moore, L. D., Le, T. & Fan, G. DNA Methylation and Its Basic Function. *Neuropsychopharmacol* **38**, 23–38 (2013).

39. Greenberg, M. V. C. & Bourc'his, D. The diverse roles of DNA methylation in mammalian development and disease. *Nat Rev Mol Cell Biol* **20**, 590–607 (2019).

40. Li, E., Bestor, T. H. & Jaenisch, R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915–926 (1992).

41. Velasco, G. & Francastel, C. Genetics meets DNA methylation in rare diseases. *Clin Genet* **95**, 210–220 (2019).

42. Locke, W. J. *et al.* DNA Methylation Cancer Biomarkers: Translation to the Clinic. *Front Genet* **10**, 1150 (2019).

43. Kouzarides, T. Chromatin modifications and their function. *Cell* **128**, 693–705 (2007).

44. Musselman, C. A., Lalonde, M.-E., Côté, J. & Kutateladze, T. G. Perceiving the epigenetic landscape through histone readers. *Nature structural & molecular biology* **19**, 1218–27 (2012).

45. Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res* **21**, 381–395 (2011).

46. Hong, L., Schroth, G. P., Matthews, H. R., Yau, P. & Bradbury, E. M. Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies

reveal that acetylation markedly reduces the binding constant of the H4 'tail' to DNA. *J Biol Chem* **268**, 305–314 (1993).

47. Shahbazian, M. D. & Grunstein, M. Functions of site-specific histone acetylation and deacetylation. *Annu Rev Biochem* **76**, 75–100 (2007).

48. Seto, E. & Yoshida, M. Erasers of Histone Acetylation: The Histone Deacetylase Enzymes. *Cold Spring Harb Perspect Biol* **6**, a018713 (2014).

49. Zhao, X. D. *et al.* Whole-Genome Mapping of Histone H3 Lys4 and 27 Trimethylations Reveals Distinct Genomic Compartments in Human Embryonic Stem Cells. *Cell Stem Cell* **1**, 286–298 (2007).

50. van Steensel, B. & Belmont, A. S. Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. *Cell* **169**, 780–791 (2017).

51. Bannister, A. J. *et al.* Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120–124 (2001).

52. Bernstein, B. E. *et al.* Methylation of histone H3 Lys 4 in coding regions of active genes. *Proc Natl Acad Sci U S A* **99**, 8695–8700 (2002).

53. Mikkelsen, T. S. *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560 (2007).

54. Heintzman, N. D. *et al.* Distinct and predictive chromatin

signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet* **39**, 311–318 (2007).

55. Heintzman, N. D. *et al.* Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108–112 (2009).

56. Prakash, K. & Fournier, D. Evidence for the implication of the histone code in building the genome structure. *Biosystems* **164**, 49–59 (2018).

57. Bernstein, B. E. *et al.* A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell* **125**, 315–326 (2006).

58. Gilbert, N. *et al.* Chromatin architecture of the human genome: gene-rich domains are enriched in open chromatin fibers. *Cell* **118**, 555–566 (2004).

59. Ou, H. D. *et al.* ChromEMT: Visualizing 3D chromatin structure and compaction n interphase and mitotic cells. *Science (New York, N.Y.)* **357**, (2017).

60. Pueschel, R., Coraggio, F. & Meister, P. From single genes to entire genomes: the search for a function of nuclear organization. *Development* **143**, 910–923 (2016).

61. Pederson, T. The Nucleolus. *Cold Spring Harb Perspect Biol* **3**, a000638 (2011).

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

63. Rieder, D., Trajanoski, Z. & McNally, J. G. Transcription factories. *Front Genet* **3**, 221 (2012).

64. Saksouk, N., Simboeck, E. & Déjardin, J. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics Chromatin* **8**, 3 (2015).

65. Trojer, P. & Reinberg, D. Facultative heterochromatin: is there a distinctive molecular signature? *Mol Cell* **28**, 1–13 (2007).

66. Janssen, A., Colmenares, S. U. & Karpen, G. H. Heterochromatin: Guardian of the Genome. *Annu Rev Cell Dev Biol* **34**, 265–288 (2018).

67. van Koningsbruggen, S. *et al.* High-resolution whole-genome sequencing reveals that specific chromatin domains from most human chromosomes associate with nucleoli. *Mol Biol Cell* **21**, 3735–3748 (2010).

68. Tiwari, V. K. *et al.* PcG proteins, DNA methylation, and gene repression by chromatin looping. *PLoS Biol* **6**, 2911–2927 (2008).

69. Lanzuolo, C., Roure, V., Dekker, J., Bantignies, F. & Orlando, V. Polycomb response elements mediate the formation of chromosome higher-order structures in the bithorax complex. *Nat Cell Biol* **9**, 1167–1174 (2007).

70. Lanzuolo, C., Lo Sardo, F. & Orlando, V. Concerted epigenetic signatures inheritance at PcG targets through replication. *Cell Cycle* **11**, 1296–1300 (2012).

71. Boettiger, A. N. *et al.* Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* **529**, 418–422 (2016).

72. Guelen, L. *et al.* Domain organization of human

chromosomes revealed by mapping of nuclear lamina interactions. *Nature* **453**, 948–951 (2008).

73. McKeon, F. D., Kirschner, M. W. & Caput, D. Homologies in both primary and secondary structure between nuclear envelope and intermediate filament proteins. *Nature* **319**, 463–468 (1986).

74. Turgay, Y. *et al.* The molecular architecture of lamins in somatic cells. *Nature* **543**, 261–264 (2017).

75. Dittmer, T. A. & Misteli, T. The lamin protein family. *Genome Biol* **12**, 222 (2011).

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

77. Hozák, P., Sasseville, A. M., Raymond, Y. & Cook, P. R. Lamin proteins form an internal nucleoskeleton as well as a peripheral lamina in human cells. *J Cell Sci* **108 (Pt 2)**, 635–644 (1995).

78. Kind, J. *et al.* Single-Cell Dynamics of Genome-Nuclear Lamina Interactions. *Cell* **153**, 178–192 (2013).

79. Lochs, S. J. A., Kefalopoulou, S. & Kind, J. Lamina Associated Domains and Gene Regulation in Development and Cancer. *Cells* **8**, E271 (2019).

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

81. Feric, M. & Misteli, T. Phase separation in genome organization across evolution. *Trends Cell Biol* **31**, 671–685

(2021).

82. Hyman, A. A., Weber, C. A. & Jülicher, F. Liquid-Liquid Phase Separation in Biology. *Annual Review of Cell and Developmental Biology* **30**, 39–58 (2014).

83. Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* (2017).

84. Alberti, S., Gladfelter, A. & Mittag, T. Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* **176**, 419–434 (2019).

85. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat Rev Mol Cell Biol* **18**, 285–298 (2017).

86. Borcherds, W., Bremer, A., Borgia, M. B. & Mittag, T. How do intrinsically disordered protein regions encode a driving force for liquid-liquid phase separation? *Curr Opin Struct Biol* **67**, 41–50 (2021).

87. Tang, S.-J. Potential Role of Phase Separation of Repetitive DNA in Chromosomal Organization. *Genes* **8**, 279 (2017).

88. Guo, Q., Shi, X. & Wang, X. RNA and liquid-liquid phase separation. *Non-coding RNA Research* **6**, 92–99 (2021).

89. Feric, M. *et al.* Coexisting Liquid Phases Underlie Nucleolar Subcompartments. *Cell* **165**, 1686–1697 (2016).

90. Lu, H. *et al.* Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. *Nature* **558**, 318–323 (2018).

91. Hur, W. *et al.* CDK-Regulated Phase Separation Seeded by

Histone Genes Ensures Precise Growth and Function of Histone Locus Bodies. *Dev Cell* **54**, 379-394.e6 (2020).

92. Strickfaden, H. *et al.* Condensed Chromatin Behaves like a Solid on the Mesoscale In Vitro and in Living Cells. *Cell* **183**, 1772-1784.e13 (2020).

93. Turner, A. L. *et al.* Highly disordered histone H1−DNA model complexes and their condensates. *PNAS* **115**, 11964–11969 (2018).

94. Sabari, B. R. *et al.* Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958 (2018).

95. Gibson, B. A. *et al.* Organization of Chromatin by Intrinsic and Regulated Phase Separation. *Cell* **179**, 470-484.e21 (2019).

96. Filippakopoulos, P. *et al.* Selective inhibition of BET bromodomains. *Nature* **468**, 1067–1073 (2010).

97. Boija, A. *et al.* Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell* **175**, 1842-1855.e16 (2018).

98. Larson, A. G. *et al.* Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. *Nature* **547**, 236–240 (2017).

99. Seif, E. *et al.* Phase separation by the polyhomeotic sterile alpha motif compartmentalizes Polycomb Group proteins and enhances their activity. *Nat Commun* **11**, 5609 (2020).

100. Fischer, A. H. *et al.* The cytologic criteria of malignancy *Cell Biochem* **110**, 795–811 (2010).

101. Zink, D., Fischer, A. H. & Nickerson, J. A. Nuclear structure in cancer cells. *Nat Rev Cancer* **4**, 677–687 (2004).

102. Dey, P. Nuclear margin irregularity and cancer: a review. *Anal Quant Cytol Histol* **31**, 345–352 (2009).

103. Huang, S., Deerinck, T. J., Ellisman, M. H. & Spector, D. L. The Dynamic Organization of the Perinucleolar Compartment in the Cell Nucleus. *J Cell Biol* **137**, 965–974 (1997).

104. Cherkezyan, L. *et al.* Nanoscale changes in chromatin organization represent the initial steps of tumorigenesis: a transmission electron microscopy study. *BMC Cancer* **14**, 189 (2014).

105. Bernardi, R. & Pandolfi, P. P. Structure, dynamics and functions of promyelocytic leukaemia nuclear bodies. *Nat Rev Mol Cell Biol* **8**, 1006–1016 (2007).

106. Gurrieri, C. *et al.* Loss of the tumor suppressor PML in human cancers of multiple histologic origins. *J Natl Cancer Inst* **96**, 269–279 (2004).

107. Siegel, R. L., Miller, K. D., Fuchs, H. E. & Jemal, A. Cancer Statistics, 2021. *CA: A Cancer Journal for Clinicians* **71**, 7–33 (2021).

108. Rawla, P. Epidemiology of Prostate Cancer. *World J Oncol* **10**, 63–89 (2019).

109. Al Olama, A. A. *et al.* A meta-analysis of 87,040 individuals identifies 23 new susceptibility loci for prostate cancer. *Nat Genet* **46**, 1103–1109 (2014).

110. Miller, D. C., Hafez, K. S., Stewart, A., Montie, J. E. & Wei,

J. T. Prostate carcinoma presentation, diagnosis, and staging: an update form the National Cancer Data Base. *Cancer* **98**, 1169–1178 (2003).

111. Aaron, L., Franco, O. & Hayward, S. W. Review of Prostate Anatomy and Embryology and the Etiology of BPH. *Urol Clin North Am* **43**, 279–288 (2016).

112. McNeal, J. E., Redwine, E. A., Freiha, F. S. & Stamey, T. A. Zonal distribution of prostatic adenocarcinoma. Correlation with histologic pattern and direction of spread. *Am J Surg Pathol* **12**, 897–906 (1988).

113. Bostwick, D. G., Liu, L., Brawer, M. K. & Qian, J. High-Grade Prostatic Intraepithelial Neoplasia. *Rev Urol* **6**, 171–179 (2004).

114. Tzelepi, V. *et al.* Contemporary Grading of Prostate Cancer: The Impact of Grading Criteria and the Significance of the Amount of Intraductal Carcinoma. *Cancers* **13**, 5454 (2021).

115. Tomlins, S. A. *et al.* Integrative molecular concept modeling of prostate cancer progression. *Nat Genet* **39**, 41–51 (2007).

116. Bubendorf, L. *et al.* Metastatic patterns of prostate cancer: an autopsy study of 1,589 patients. *Hum Pathol* **31**, 578–583 (2000).

117. De Marzo, A. M. *et al.* Inflammation in prostate carcinogenesis. *Nat Rev Cancer* **7**, 256–269 (2007).

118. Ali, A. *et al.* Prostate zones and cancer: lost in transition? *Nat Rev Urol* 1–15 (2021).

119. Mottet, N. *et al.* EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent. *Eur Urol* **71**, 618–629 (2017).

120. Hernández, J. & Thompson, I. M. Prostate-specific antigen: a review of the validation of the most commonly used cancer biomarker. *Cancer* **101**, 894–904 (2004).

121. Crawford, E. D. & Abrahamsson, P.-A. PSA-based screening for prostate cancer: how does it compare with other cancer screening tests? *Eur Urol* **54**, 262–273 (2008).

122. Patel, D. N. & Freedland, S. J. New Prostate Cancer Biomarkers: The Search Continues. *European Urology* **72**, 950–951 (2017).

123. Bjurlin, M. A. & Taneja, S. S. Standards for prostate biopsy. *Curr Opin Urol* **24**, 155–161 (2014).

124. Ashley, R. A. *et al.* Reassessing the diagnostic yield of saturation biopsy of the prostate. *Eur Urol* **53**, 976–981 (2008).

125. Djavan, B. *et al.* Predictability and significance of multifocal prostate cancer in the radical prostatectomy specimen. *Tech Urol* **5**, 139–142 (1999).

126. Abd, T. T. *et al.* Comparison of 12-core versus 8-core prostate biopsy: multivariate analysis of large series of US veterans. *Urology* **77**, 541–547 (2011).

127. Guo, G., Xu, Y. & Zhang, X. TRUS-guided transperineal prostate 12+X core biopsy with template for the diagnosis of prostate cancer. *Oncol Lett* **13**, 4863–4867 (2017).

128. Papagiannopoulos, D. *et al.* Predictors of Infectious

Complications after Targeted Prophylaxis for Prostate Needle Biopsy. *J Urol* **199**, 155–160 (2018).

129. Remzi, M. Editorial comment on: Improved accuracy in predicting the presence of gleason pattern 4/5 prostate cancer by three-dimensional 26-core systematic biopsy. *Eur Urol* **52**, 1668–1669 (2007).

130. Gleason, D. F. & Mellinger, G. T. Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging. *J Urol* **111**, 58–64 (1974).

131. Epstein, J. I. *et al.* The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma: Definition of Grading Patterns and Proposal for a New Grading System. *The American Journal of Surgical Pathology* **40**, 244–252 (2016).

132. Bertero, L. *et al.* Eighth Edition of the UICC Classification of Malignant Tumours: an overview of the changes in the pathological TNM classification criteria-What has changed and why? *Virchows Arch* **472**, 519–531 (2018).

133. Borley, N. & Feneley, M. R. Prostate cancer: diagnosis and staging. *Asian J Androl* **11**, 74–80 (2009).

134. Rybak, A. P., Bristow, R. G. & Kapoor, A. Prostate cancer stem cells: deciphering the origins and pathways involved in prostate tumorigenesis and aggression. *Oncotarget* **6**, 1900–1919 (2015).

135. Cooperberg, M. R., Broering, J. M. & Carroll, P. R. Risk assessment for prostate cancer metastasis and mortality at

the time of diagnosis. *J Natl Cancer Inst* **101**, 878–887 (2009).

136. Harnden, P., Shelley, M. D., Coles, B., Staffurth, J. & Mason, M. D. Should the Gleason grading system for prostate cancer be modified to account for high-grade tertiary components? A systematic review and meta-analysis. *The Lancet Oncology* **8**, 411–419 (2007).

137. Abeshouse, A. *et al.* The Molecular Taxonomy of Primary Prostate Cancer. *Cell* **163**, 1011–1025 (2015).

138. Espiritu, S. M. G. *et al.* The Evolutionary Landscape of Localized Prostate Cancers Drives Clinical Aggression. *Cell* **173**, 1003-1013.e15 (2018).

139. Berger, M. F. *et al.* The genomic complexity of primary human prostate cancer. *Nature* **470**, 214–220 (2011).

140. Tomlins, S. A. *et al.* Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* **310**, 644–648 (2005).

141. Attard, G. *et al.* Duplication of the fusion of TMPRSS2 to ERG sequences identifies fatal human prostate cancer. *Oncogene* **27**, 253–263 (2008).

142. Demichelis, F. *et al.* TMPRSS2:ERG gene fusion associated with lethal prostate cancer in a watchful waiting cohort. *Oncogene* **26**, 4596–4599 (2007).

143. He, W. W. *et al.* A novel human prostate-specific, androgen-regulated homeobox gene (NKX3.1) that maps to 8p21, a region frequently deleted in prostate cancer. *Genomics* **43**, 69–77 (1997).

144. Cairns, P. *et al.* Frequent inactivation of PTEN/MMAC1 in primary prostate cancer. *Cancer Res* **57**, 4997–5000 (1997).

145. Visakorpi, T. *et al.* In vivo amplification of the androgen receptor gene and progression of human prostate cancer. *Nat Genet* **9**, 401–406 (1995).

146. Fraser, M. *et al.* Genomic hallmarks of localized, non-indolent prostate cancer. *Nature* **541**, 359–364 (2017).

147. Barbieri, C. E. *et al.* Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. *Nat Genet* **44**, 685–689 (2012).

148. Blattner, M. *et al.* SPOP Mutation Drives Prostate Tumorigenesis In Vivo through Coordinate Regulation of PI3K/mTOR and AR Signaling. *Cancer Cell* **31**, 436–451 (2017).

149. Adams, E. J. *et al.* FOXA1 mutations alter pioneering activity, differentiation and prostate cancer phenotypes. *Nature* **571**, 408–412 (2019).

150. Armenia, J. *et al.* The long tail of oncogenic drivers in prostate cancer. *Nat Genet* **50**, 645–651 (2018).

151. Hieronymus, H. *et al.* Copy number alteration burden predicts prostate cancer relapse. *Proc Natl Acad Sci U S A* **111**, 11139–11144 (2014).

152. Quigley, D. A. *et al.* Genomic Hallmarks and Structural Variation in Metastatic Prostate Cancer. *Cell* **174**, 758-769.e9 (2018).

153. Wong, Y. N. S., Ferraldeschi, R., Attard, G. & de Bono, J.

Evolution of androgen receptor targeted therapy for advanced prostate cancer. *Nat Rev Clin Oncol* **11**, 365–376 (2014).

154. Davey, R. A. & Grossmann, M. Androgen Receptor Structure, Function and Biology: From Bench to Bedside. *Clin Biochem Rev* **37**, 3–15 (2016).

155. Lu-Yao, G. L. *et al.* Fifteen-year survival outcomes following primary androgen-deprivation therapy for localized prostate cancer. *JAMA Intern Med* **174**, 1460–1467 (2014).

156. Davies, A., Conteduca, V., Zoubeidi, A. & Beltran, H. Biological Evolution of Castration-resistant Prostate Cancer. *Eur Urol Focus* **5**, 147–154 (2019).

157. Feldman, B. J. & Feldman, D. The development of androgen-independent prostate cancer. *Nat Rev Cancer* **1**, 34–45 (2001).

158. Grasso, C. S. *et al.* The mutational landscape of lethal castration-resistant prostate cancer. *Nature* **487**, 239–243 (2012).

159. Labrecque, M. P. *et al.* Molecular profiling stratifies diverse phenotypes of treatment-refractory metastatic castration-resistant prostate cancer. *J Clin Invest* **129**, 4492–4505 (2019).

160. Antonarakis, E., Armstrong, A., Dehm, S. & Luo, J. Androgen receptor variant-driven prostate cancer: clinical implications and therapeutic targeting. *Prostate Cancer Prostatic Dis* **19**, 231–241 (2016).

161. Kirby, M. K. *et al.* Genome-wide DNA methylation

measurements in prostate tissues uncovers novel prostate cancer diagnostic biomarkers and transcription factor binding patterns. *BMC Cancer* **17**, 273 (2017).

162. Xu, J. *et al.* Super-resolution imaging reveals the evolution of higher-order chromatin folding in early carcinogenesis. *Nat Commun* **11**, 1899 (2020).

163. Goel, S., Bhatia, V., Biswas, T. & Ateeq, B. Epigenetic reprogramming during prostate cancer progression: A perspective from development. *Semin Cancer Biol* S1044-579X(21)00023–7 (2021).

164. Ce, M., Ig, M. & Ag, L. The importance of DNA methylation in prostate cancer development. *The Journal of steroid biochemistry and molecular biology* **166**, (2017).

165. Kim, J. H. *et al.* Deep sequencing reveals distinct patterns of DNA methylation in prostate cancer. *Genome Res* **21**, 1028–1041 (2011).

166. Pavan, N., Grassi, G. & Scaggiante, B. An update of aberrant methylation detection on circulating cell-free DNA as a tool to improve prostate cancer diagnosis and prognosis. *Journal of Translational Genetics and Genomics* **5**, 173–181 (2021).

167. Bakavicius, A. *et al.* Urinary DNA methylation biomarkers for prediction of prostate cancer upgrading and upstaging. *Clinical Epigenetics* **11**, 115 (2019).

168. Smeets, E. *et al.* The role of TET-mediated DNA hydroxymethylation in prostate cancer. *Molecular and Cellular Endocrinology* **462**, 41–55 (2018).

169. Seligson, D. B. *et al.* Global histone modification patterns predict risk of prostate cancer recurrence. *Nature* **435**, 1262–1266 (2005).

170. Varambally, S. *et al.* The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature* **419**, 624–629 (2002).

171. Richly, H., Aloia, L. & Di Croce, L. Roles of the Polycomb group proteins in stem cells and cancer. *Cell Death Dis* **2**, e204 (2011).

172. Liu, Q. *et al.* Polycomb group proteins EZH2 and EED directly regulate androgen receptor in advanced prostate cancer. *Int J Cancer* **145**, 415–426 (2019).

173. Yu, J. *et al.* Integrative genomics analysis reveals silencing of beta-adrenergic signaling by polycomb in prostate cancer. *Cancer Cell* **12**, 419–431 (2007).

174. Xu, K. *et al.* EZH2 Oncogenic Activity in Castration Resistant Prostate Cancer Cells is Polycomb-Independent. *Science* **338**, 1465–1469 (2012).

175. Park, S. H. *et al.* Going beyond Polycomb: EZH2 functions in prostate cancer. *Oncogene* **40**, 5788–5798 (2021).

176. Park, S. H. *et al.* Posttranslational regulation of FOXA1 by Polycomb and BUB3/USP7 deubiquitin complex in prostate cancer. *Science Advances* (2021)

177. Bansal, N. *et al.* BMI-1 Targeting Interferes with Patient-Derived Tumor-Initiating Cell Survival and Tumor Growth in Prostate Cancer. *Clin Cancer Res* **22**, 6176–6191 (2016).

178. Zhu, S. *et al.* BMI1 regulates androgen receptor in prostate

cancer independently of the polycomb repressive complex 1. *Nat Commun* **9**, 500 (2018).

179. Zhong, J. *et al.* p300 acetyltransferase regulates androgen receptor degradation and PTEN-deficient prostate tumorigenesis. *Cancer Res* **74**, 1870–1880 (2014).

180. Lasko, L. M. *et al.* Discovery of a selective catalytic p300/CBP inhibitor that targets lineage-specific tumours. *Nature* **550**, 128–132 (2017).

181. Jin, L. *et al.* Therapeutic Targeting of the CBP/p300 Bromodomain Blocks the Growth of Castration-Resistant Prostate Cancer. *Cancer Res* **77**, 5564–5575 (2017).

182. Liu, J. *et al.* p300/CBP inhibition enhances the efficacy of programmed death-ligand 1 blockade treatment in prostate cancer. *Oncogene* **39**, 3939–3951 (2020).

183. Tan, Y. *et al.* Inhibition of BRD4 suppresses tumor growth in prostate cancer via the enhancement of FOXO1 expression. *Int J Oncol* **53**, 2503–2517 (2018).

184. Shafran, J. S., Jafari, N., Casey, A. N., Győrffy, B. & Denis, G. V. BRD4 regulates key transcription factors that drive epithelial–mesenchymal transition in castration-resistant prostate cancer. *Prostate Cancer Prostatic Dis* **24**, 268–277 (2021).

185. Raina, K. *et al.* PROTAC-induced BET protein degradation as a therapy for castration-resistant prostate cancer. *Proc Natl Acad Sci U S A* **113**, 7124–7129 (2016).

186. Asangani, I. A. *et al.* BET Bromodomain Inhibitors Enhance Efficacy and Disrupt Resistance to AR Antagonists in the

Treatment of Prostate Cancer. *Mol Cancer Res* **14**, 324–331 (2016).

187. Asangani, I. A. *et al.* Therapeutic targeting of BET bromodomain proteins in castration-resistant prostate cancer. *Nature* **510**, 278–282 (2014).

188. Shu, S. & Polyak, K. BET Bromodomain Proteins as Cancer Therapeutic Targets. *Cold Spring Harb Symp Quant Biol* **81**, 123–129 (2016).

189. Corces, M. R. & Corces, V. G. The three-dimensional cancer genome. *Curr Opin Genet Dev* **36**, 1–7 (2016).

190. Johnstone, S. E. *et al.* Large-Scale Topological Changes Restrain Malignant Progression in Colorectal Cancer. *Cell* **182**, 1474-1489.e23 (2020).

191. Flavahan, W. A. *et al.* Altered chromosomal topology drives oncogenic programs in SDH-deficient GISTs. *Nature* **575**, 229–233 (2019).

192. Kloetgen, A. *et al.* Three-dimensional chromatin landscapes in T cell acute lymphoblastic leukemia. *Nat Genet* **52**, 388–400 (2020).

193. Filippova, G. N. *et al.* A widely expressed transcription factor with multiple DNA sequence specificity, CTCF, is localized at chromosome segment 16q22.1 within one of the smallest regions of overlap for common deletions in breast and prostate cancers. *Genes Chromosomes Cancer* **22**, 26–36 (1998).

194. Filippova, G. N. *et al.* Tumor-associated zinc finger mutations in the CTCF transcription factor selectively alter

tts DNA-binding specificity. *Cancer Res* **62**, 48–52 (2002).

195. Denker, A. & de Laat, W. The second decade of 3C technologies: detailed insights into nuclear organization. *Genes Dev* **30**, 1357–1382 (2016).

196. Pope, B. D. *et al.* Topologically associating domains are stable units of replication-timing regulation. *Nature* **515**, 402–405 (2014).

197. Taberlay, P. C. *et al.* Three-dimensional disorganization of the cancer genome occurs coincident with long-range genetic and epigenetic alterations. *Genome Res* **26**, 719–731 (2016).

198. Cai, M. *et al.* 4C-seq revealed long-range interactions of a functional enhancer at the 8q24 prostate cancer risk locus. *Sci Rep* **6**, 22462 (2016).

199. Rhie, S. K. *et al.* A high-resolution 3D epigenomic map reveals insights into the creation of the prostate cancer transcriptome. *Nat Commun* **10**, 4154 (2019).

200. Luo, Z., Rhie, S. K., Lay, F. D. & Farnham, P. J. A Prostate Cancer Risk Element Functions as a Repressive Loop that Regulates HOXA13. *Cell Reports* **21**, 1411–1417 (2017).

201. San Martin, R. *et al.* Chromosome compartmentalization alterations in prostate cancer cell lines model disease progression. *J Cell Biol* **221**, e202104108 (2022).

202. Montanari, M. *et al.* Epithelial-mesenchymal transition in prostate cancer: an overview. *Oncotarget* **8**, 35376–35389 (2017).

203. Leggett, S. E., Hruska, A. M., Guo, M. & Wong, I. Y. The

epithelial-mesenchymal transition and the cytoskeleton in bioengineered systems. *Cell Communication and Signaling* **19**, 32 (2021).

204. Ma, H., N, K. & H, B. Molecular requirements for epithelial-mesenchymal transition during tumor progression. *Current opinion in cell biology* **17**, (2005).

205. Xu, J., Lamouille, S. & Derynck, R. TGF-β-induced epithelial to mesenchymal transition. *Cell Res* **19**, 156–172 (2009).

206. Yang, J. *et al.* Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell* **117**, 927–939 (2004).

207. Thiery, J. P. Epithelial-mesenchymal transitions in tumour progression. *Nat Rev Cancer* **2**, 442–454 (2002).

208. Peixoto, P. *et al.* EMT is associated with an epigenetic signature of ECM remodeling genes. *Cell Death Dis* **10**, 1–17 (2019).

209. Verdone, J. E., Parsana, P., Veltri, R. W. & Pienta, K. J. Epithelial-mesenchymal transition in prostate cancer is associated with quantifiable changes in nuclear structure. *Prostate* **75**, 218–224 (2015).

210. 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 (2011).

CHAPTER 2.

DYSFUNCTIONAL POLYCOMB TRANSCRIPTIONAL REPRESSION CONTRIBUTES TO LAMIN A/C DEPENDENT MUSCULAR DYSTROPHY

Andrea Bianchi, Chiara Mozzetta, Gloria Pegoli, Federica Lucini, Sara Valsoni, Valentina Rosti, Cristiano Petrini, Alice Cortesi, Francesco Gregoretti, Laura Antonelli, Gennaro Oliva, Marco De Bardi, Roberto Rizzi, Beatrice Bodega, Diego Pasini, Francesco Ferrari, Claudia Bearzi, Chiara Lanzuolo*

JCI The Journal of Clinical Investigation

Authorship note: BA and MC contributed equally to this work.

Conflict of interest: The authors have declared that no conflict of interest exists.

Submitted: February 14, 2019; Accepted: January 22, 2020; **Published**: April 6, 2020.

Reference information: J Clin Invest. 2020;130(5):2408–2421. https://doi.org/10.1172/JCI128161.

Dysfunctional Polycomb transcriptional repression contributes to Lamin A/C dependent muscular dystrophy

Andrea Bianchi[1, 2*], Chiara Mozzetta[2*§], Gloria Pegoli[3], Federica Lucini[1,3], Sara Valsoni[1,3, #], Valentina Rosti[4], Cristiano Petrini[5], Alice Cortesi[1], Francesco Gregoretti[6], Laura Antonelli[6], Gennaro Oliva[6], Marco De Bardi[3], Roberto Rizzi[1, 4], Beatrice Bodega[1], Diego Pasini[7], Francesco Ferrari[5, 8], Claudia Bearzi[1, 9] and Chiara Lanzuolo[3, 4]

1. Istituto Nazionale di Genetica Molecolare “Romeo ed Enrica Invernizzi”, Milan, Italy
2. CNR Institute of Cellular Biology and Neurobiology, Rome, Italy
3. IRCCS Santa Lucia Foundation, Rome, Italy
4. CNR Institute of Biomedical Technologies, Milan, Italy
5. IFOM, the FIRC Institute of Molecular Oncology, Milan, Italy
6. CNR Institute for High Performance Computing and Networking, Naples, Italy
7. European Institute of Oncology, Milan, Italy
8. CNR Institute of Molecular Genetics "Luigi Luca Cavalli-Sforza", Pavia, Italy
9. CNR Istituto di Biochimica e Biologia Cellulare, Rome, Italy

§. Present address: CNR Institute of Molecular Biology and Pathology (IBPM), Rome, Italy

#. Present address: San Raffaele Telethon Institute for Gene Therapy (SR-Tiget), IRCCS San Raffaele Scientific Institute, Milan, Italy

* Equal contribution

Conflict of interest statement

The authors have declared that no conflict of interest exists.

Keywords

Muscle Stem Cells (MuSCs), Lamin A, Polycomb, cell fate, differentiation, premature senescence, muscular dystrophy.

Running head

PcG dysfunctions in lamin dystrophy

Abstract

Lamin A is a component of the inner nuclear membrane that, together with epigenetic factors, organizes the genome in higher order structures required for transcriptional control. Mutations in the Lamin A/C gene cause several diseases, belonging to the class of laminopathies, including muscular dystrophies. Nevertheless, molecular mechanisms involved in the pathogenesis of Lamin A-dependent dystrophies are still largely unknown. Polycomb group of proteins (PcG) are epigenetic repressors and Lamin A interactors, primarily involved in the maintenance of cell identity. Using a murine model of Emery-Dreifuss Muscular Dystrophy (EDMD), we showed here that Lamin A loss deregulated PcG positioning in muscle satellite stem cells leading to de-repression of non-muscle specific genes and $p16^{INK4a}$, a senescence driver encoded in the *Cdkn2a* locus. This aberrant transcriptional programme caused impairment in self-renewal, loss of cell identity and premature exhaustion of quiescent satellite cell pool. Genetic ablation of *Cdkn2a* locus restored muscle stem cell properties in Lamin A/C null dystrophic mice. Our findings established a direct link between Lamin A and PcG epigenetic silencing and indicated that Lamin A-dependent muscular dystrophy can be ascribed to intrinsic epigenetic dysfunctions of muscle stem cells.

Introduction

The nuclear lamina (NL) is located in the inner part of the nuclear membrane and is made up of a complex network of type V filament proteins, the lamins (1, 2). In vertebrates lamin proteins are divided into A and B types, based on sequence homologies. A growing body of evidence suggests that lamins are directly involved in the functional control of the genome, by organizing its three dimensional positioning in the nuclear space, through the association with transcriptionally repressed large genomic regions, called Lamina-associated domains (LADs) (3). The crucial function of lamins is attested by an entire class of genetic diseases, called *laminopathies*, where specific components of the NL are altered (4). In particular, the study of Lamin A/C is gaining an increasing interest for three reasons: i) Lamin A/C plays an undisputed role in several cellular processes from mechanotransduction to cell differentiation; ii) Lamin A/C has a peculiar intranuclear distribution being present in the nucleoplasm as well as in the nuclear periphery (5); iii) Lamin A/C interacts with several epigenetic factors, exerting a functional control on transcriptional regulation (3, 6). One of the most studied Lamin A/C dependent cellular process is myogenesis because mutations in the *LMNA* gene lead to muscular dystrophies, as in the case of Emery Dreifuss Muscular Dystrophy (EDMD) (7). However, epigenetic mechanisms involved in lamin-dependent dystrophy are still largely unknown. PcG proteins are epigenetic repressors originally discovered for their central roles in development and cell differentiation (8) and recently described as functional partners of Lamin A/C (9-14). In the last years several evidence demonstrated that PcG proteins are involved in the regulation of adult stem cells (15, 16), safeguarding cell identity and preventing cell fate transition. In multipotent stem cells, PcG proteins ensure the correct balance between self-renewal and lineage-specific

differentiation, promptly responding to the environmental changes. At the molecular level this is achieved through PcG binding at bivalent domains, genomic regions containing active and repressive epigenetic signatures and a poised RNA polymerase II (17). This epigenetic condition allows a rapid transition from one transcriptional state to another, ensuring the correct expression of unique and specific cell lineage genes. Defects in these fine-tuned mechanisms lead to lack of cell identity (18) or pathological reprogramming (19).

Given their key role in regulating stem cells fate decisions and tissue homeostasis, it is conceivable that PcG dysfunctions contribute to lamin-dependent, tissue-specific human diseases. Here, we examined how the absence of Lamin A/C impacts muscle stem (satellite) cells (MuSCs) in vivo, and the role of PcG proteins in lamin muscular dystrophy. We found that MuSCs lacking Lamin A/C redistribute PcG-dependent histone marks leading to transcriptional upregulation of crucial PcG-target genes, such as non-muscle related genes. This leads to lack of MuSC cell identity and cellular senescence, determining a premature exhaustion of the muscular stem cell niche. Genetic ablation of the PcG-regulated *Cdkn2a* locus in lamin dystrophic mice restores MuSCs defects.

Results

Lamin A is required to preserve the quiescent MuSC pool

We analyzed the MuSC niche composition in the severely dystrophic *Lmna*Δ8–11$^{-/-}$ mice (homozygous, hom), together with their unaffected littermates, wild-type (WT, *Lmna*Δ8–11$^{+/+}$) or heterozygous (het, *Lmna*Δ8–11$^{+/-}$), during dystrophy progression at 10, 14, 16, and 19 days after birth. In early stages of postnatal growth (day 10 and d14) no differences were found in the relative amounts of quiescent stem

cells (QSCs; PAX7^{+}MYOD^{-}) and activated stem cells (ASCs; PAX7^{+}MYOD^{+}) (Figure 1, A and B, and Supplemental Figure 1A; supplemental material available online with this article; https://doi.org/10.1172/JCI128161DS1) among MuSCs with distinct genotypes. Conversely, starting from d16, an imbalance of MuSCs becomes evident in *Lmna*Δ8–11$^{-/-}$ muscles, with a decreased proportion of QSCs compared with ASCs, mirroring a decline in myofiber cross-sectional area (CSA) (Figure 1C). Of note, the overall amount of PAX7^{+} MuSCs was not significantly altered across the different genotypes (Supplemental Figure 1B) and Ki67 staining at d19 confirmed that in *Lmna*Δ8–11$^{-/-}$ muscles a lower amount of QSCs (PAX7^{+}Ki67^{-}) is present (Supplemental Figure 1C). These findings suggested that *Lmna*Δ8–11$^{-/-}$ MuSCs may be deficient in self-renewal capacity. To test this hypothesis, we isolated single myofibers at d19 and cultured them for 96 hours, monitoring their ability to give rise to self-renewing PAX7^{+}MYOD^{-}, activated PAX7^{+}MYOD^{+}, and differentiating PAX7^{-}MYOD^{+} cells (Figure 1, D and E). In fibers isolated from *Lmna*Δ8–11$^{-/-}$ muscles, we observed a decrease in the number of self-renewing PAX7^{+}MYOD^{-} cells compared with WT, paralleled by a diminished number of differentiating cells (PAX7^{-}MYOD^{+}) and an increased number of activated satellite cells. Immunostaining with the myogenic marker MYOG, which is involved in later stages of differentiation, highlighted a lower number of MYOG^{+}PAX7^{-} cells in *Lmna*Δ8–11$^{-/-}$ (Figure 1, F and G) accompanied by proliferation defects ex vivo (Supplemental Figure 1D). These findings suggest a defect in muscle differentiation, as described in Cohen et al. (20), and a previously unreported self-renewal impairment. Interestingly, the healthy heterozygous *Lmna*Δ8–11$^{+/-}$ mice, although not developing muscular dystrophy (21), presented an intermediate self-renewal phenotype between WT and homozygous *Lmna*Δ8–11$^{-/-}$ (Figure 1, D

and E), suggesting that proper lamin A levels are important for MuSC homeostasis to preserve their self-renewal capacity. To further address this, we repeatedly injured muscle in adult heterozygous *Lmna*Δ8–11$^{+/-}$ mice, which show less lamin A at the mRNA and protein level (Supplemental Figure 2, A and B; see complete unedited blots in the supplemental material). Analysis of MuSC populations revealed a lower amount of QSCs in *Lmna*Δ8–11$^{+/-}$ muscles upon repeated injuries (Figure 2, A and B) and a decline in Pax7^{+} cells (Figure 2C), suggesting that lamin A affects MuSC self-renewal in a dose-dependent manner.

*Lmna*Δ8-11$^{-/-}$ dystrophic MuSCs display chromatin redistribution of PcG dependent signature

Our recent results showed a lamin A/C–PcG crosstalk during in vitro myogenesis (10). We thus wondered if the altered MuSC balance observed in *Lmna*Δ8–11$^{-/-}$ muscles might be ascribed to aberrant PcG functions. We first performed immunostaining of Ezh2, the catalytic subunit of polycomb repressive complex 2 (PRC2) (Supplemental Figure 3A) in d19 MuSCs. We fixed MuSCs before FACS isolation to preserve the nuclear architecture of lamin A–deficient cells (see Methods). We found a general intranuclear diffusion of Ezh2 in *Lmna*Δ8–11$^{-/-}$ MuSCs, ascertained by measuring PcG body parameters (22) (Supplemental Figure 3, A–C). We also measured Ezh2 expression both in MuSCs and whole muscles (Supplemental Figure 3D) and we analyzed Ezh2 protein levels in whole muscles (Supplemental Figure 3, E and F; see complete unedited blots in the supplemental material). We found no major differences between *Lmna*Δ8–11$^{+/+}$ and *Lmna*Δ8–11$^{-/-}$ mice. To further analyze the Ezh2 intranuclear distribution in QSCs, we performed triple

immunostaining on muscle cryosections (Supplemental Figure 3, G and H). Ezh2 levels, assessed measuring fluorescence intensity, were similar in *Lmna*Δ8–11$^{+/+}$ and *Lmna*Δ8–11$^{-/-}$ MuSCs in PAX7^{+}Ki67^{-} and PAX7^{+}Ki67^{+} cells. Because Ezh2 is hardly detectable in adult quiescent satellite cells (23–25), this result suggests that during postnatal growth developmental signals might instead contribute to maintaining Ezh2 expression in nonproliferating MuSCs.

On the other hand, evaluation of the number of PcG bodies on the same sections highlighted a decrease in the number of Ezh2 bodies in the mutant (Supplemental Figure 3, I and J), leading us to conclude that the absence of lamin A/C does not affect Ezh2 protein levels but influences its nuclear distribution. To gain further insight into possible PcG-dependent transcriptional defects, we performed RNA sequencing (RNA-seq) on freshly isolated MuSCs at d19, finding 1424 upregulated genes and 1842 downregulated genes in the *Lmna*Δ8–11$^{-/-}$ MuSCs compared with WT (Supplemental Figure 4A). Interestingly, performing a gene set enrichment analysis (GSEA) based on differential expression generated after conditional ablation of Ezh2 in MuSCs (24) and *Lmna*Δ8–11$^{-/-}$ upregulated genes, we found a significant association between the 2 data sets, suggesting that lamin A absence impairs Ezh2 function (Supplemental Figure 4B). We also followed the deposition of the Ezh2-dependent H3K27me3 histone mark in *Lmna*Δ8–11 mice by quantitative spike-in ChIP-seq (26) (see Supplemental Methods and Supplemental Figure 4, C and D). Integrative analysis of RNA-seq and ChIP-seq revealed that upregulated genes in the *Lmna*Δ8–11$^{-/-}$ condition are significantly enriched for H3K27me3 targets (identified in the WT condition) (Figure 3A). Indeed, analysis of H3K27me3 distribution around the transcription start sites (TSSs) and along the body of genes indicated a decrease in this repressive mark in *Lmna*Δ8–11$^{-/-}$ MuSCs compared

with WT (Figure 3, B and C), which was not accompanied by a statistically significant decrease in global H3K27me3 levels in MuSCs (Supplemental Figure 5, A and B; see complete unedited blots in the supplemental material) and in whole muscle (Supplemental Figure 5, C and D; see complete unedited blots in the supplemental material). In contrast, a deep analysis of H3K27me3 ChIP-seq read coverage in the intergenic genomic regions between the known H3K27me3 enrichment peaks interestingly showed a higher average coverage in the *Lmna*Δ8–11$^{-/-}$ MuSCs compared with WT counterparts (Figure 3D). These results are compatible with a diffusion of PcG proteins along the chromatin fibers rather than a complete PcG displacement. To identify the PcG targets mostly affected by lamin A deficiency, genes were grouped according to their transcription level in WT MuSCs. We thus defined 4 equally sized groups of genes based on expression level quartiles (Figure 3E). For each expression category, we reanalyzed the H3K27me3 distribution along the body of genes and at the TSS and the percentage of upregulated genes in the *Lmna*Δ8–11$^{-/-}$ MuSCs (Figure 3, E and F, and Supplemental Figure 5E). In quartile I we found only a small percentage (0.65%) of upregulated genes in *Lmna*Δ8–11$^{-/-}$, suggesting that the H3K27me3 decrease/redistribution is not sufficient to activate transcription in highly repressed genes (Figure 3F, Supplemental Figure 5E, and Supplemental Figure 6A). In contrast, quartiles II, III, and IV are more affected by the diminished H3K27me3 levels in *Lmna*Δ8–11$^{-/-}$ (Figure 3F, Supplemental Figure 5E, and Supplemental Figure 6B), showing a percentage of upregulated genes between 5% and 9%. Specifically, we noticed that in WT MuSCs, H3K27me3 ChIP-seq signal enrichment around the TSS and within the body of genes is progressively lower in quartiles of higher expression, as expected (Figure 3F, quartiles III and IV, and Supplemental Figure 5E).

However, for *Lmna*Δ8–11$^{-/-}$ mice the decrease in H3K27me3 signal inside the gene body is relatively less marked than in WT mice; in fact, the average enrichment is slightly higher. We quantified and confirmed this observation by considering the ratio of H3K27me3 ChIP-seq enrichment signal at the TSS and 2.5 kb downstream of the TSS, for each gene, in WT and *Lmna*Δ8–11$^{-/-}$ mice (Supplemental Figure 7A), showing that this ratio is significantly different for higher expression quartiles (Supplemental Figure 7B).

Lamin A-dependent PcG redistribution determines de-repression of non-muscle related bivalent genes

The altered PcG binding observed in *Lmna*Δ8–11$^{-/-}$ MuSCs prompted us to examine in greater detail the bivalent genes, a subgroup of PcG targets whose expression is more susceptible to variations in PcG occupancy (27). Bivalent genes are characterized by the concurrent presence of both H3K27me3 and H3K4me3 marks around TSSs and have an intermediate gene expression state (28). We first performed H3K27me3 and H3K4me3 ChIP-seq in WT MuSCs (Supplemental Figure 8, A and B) and we defined bivalent genes using the parameters described in Bernstein et al. (17) for the H3K4me3 window at the TSS (Supplemental Figure 8C). Then, we tested the association between bivalent and upregulated genes in the *Lmna*Δ8–11$^{-/-}$ MuSCs by means of Fisher's exact test. We observed a significant overrepresentation of bivalent genes among upregulated ones in the *Lmna*Δ8–11$^{-/-}$ MuSCs (Figure 4A). To gain more insight into the biological relevance of deregulated genes in the mutant mice, we performed semantic similarity analysis of all Gene Ontology (GO) terms associated with upregulated genes (Figure 4B) together with GSEA (Supplemental Figure 8, D and E). These analyses showed a

negative correlation with muscle specification (Supplemental Figure 8D) together with an acquisition of markers related to lipid metabolic processes (Figure 4B and Supplemental Figure 8E). Notably, Fisher's exact test analysis highlighted a significant overlap between genes with bivalent promoters and the *Lmna*Δ8–11$^{-/-}$ MuSC upregulated genes involved in adipogenesis (Figure 4C), suggesting that lamin A is involved in preserving MuSC identity by ensuring the correct PcG-mediated transcriptional repression of non-muscle genes.

PPARγ is aberrantly expressed in *Lmna*Δ8-11$^{-/-}$ dystrophic MuSCs

Given this strong association between bivalent gene reactivation and adipogenesis markers (Figure 4C), we analyzed different lipid-related GO categories, finding among the top GO terms peroxisome proliferator–activated receptor γ (*Pparg*) (Supplemental Table 1). This master transcription factor for adipose cell differentiation (29, 30) was found to be significantly upregulated in *Lmna*Δ8–11$^{-/-}$ MuSCs (FDR < 0.05). Moreover, the *Pparg* gene is a polycomb target and has a bivalent signature in WT MuSCs (Supplemental Table 1 and Supplemental Figure 9A). These observations prompted us to analyze *Pparg* transcriptional deregulation. We stained muscles for PAX7 and PPARγ to directly test if MuSCs displayed aberrant expression of PPARγ in the absence of lamin A. Strikingly, we found approximately 10% of *Lmna*Δ8–11$^{-/-}$ MuSCs that simultaneously express both muscular and adipogenesis markers, being PAX7^{+}PPARγ^{+} (Figure 5, A and B). Accordingly, the genomic region of the *Pparg* gene showed a decrease in H3K27me3 enrichment around the TSS in the *Lmna*Δ8–11$^{-/-}$ MuSCs, accompanied by transcriptional upregulation (Supplemental Figure 9B). To evaluate if the aberrant

expression of adipogenic genes in MuSCs of mutant mice culminates with fatty infiltration we performed immunofluorescence staining for perilipin 1, a protein present on the surface of lipid droplets (ref. 31 and Figure 5, C and D), on cryosections of muscles derived from d19 *Lmna*Δ8–11 mice. We found large areas of adipose accumulation between myofibers of lamin A/C$^{+/-}$ and lamin A/C$^{-/-}$ muscles, which were instead undetectable in the WT mice. Considering the key role of PcG proteins in mediating the formation of chromatin loop structures (32, 33), we reasoned that the loss of H3K27me3 and transcriptional upregulation of the *Pparg* locus could be related to the alteration of chromatin 3D structure. The genomic 3D architecture is organized in structurally separated topologically associated domains (TADs), chromosomal structures that favor intradomain looping interactions (34). TADs can be identified by genome-wide chromosome conformation capture (Hi-C) and are largely conserved across different cell types. We verified that the *Pparg* locus is included in a TAD encompassing a region extending also upstream of the *Pparg* locus itself, using high-resolution Hi-C data on mouse embryonic stem cells (35) and the 3D Genome Browser (ref. 36 and Supplemental Figure 10). Then, we performed 3D multicolor DNA FISH analysis on prefixed MuSCs using 1 BAC probe overlapping the *Pparg* promoter and a second probe annealing at the other TAD border. We observed an overlap of the signals from the 2 regions in the WT *Lmna*Δ8–11$^{+/+}$ MuSCs that indicates the presence of a DNA looping in *cis* (Figure 5, E and F). By contrast, in *Lmna*Δ8–11$^{-/-}$ MuSCs the distance between the signals was higher, definitely suggesting the lack of DNA/DNA interactions. Indeed, from the analysis of H3K27me3 ChIP-seq tracks we noticed in the *Lmna*Δ8–11$^{-/-}$ MuSCs a reduction in H3K27me3 peaks upstream of the *Pparg* locus (Supplemental Figure 10). FISH analysis also highlighted that in WT the entire

genomic region is close to the nuclear periphery (Figure 5, E and G), whereas in *Lmna*Δ8–11$^{-/-}$ it is relocated in the nuclear interior, suggesting that lamin A absence interferes with chromatin anchoring to the nuclear lamina and PcG-dependent DNA conformation.

*Lmna*Δ8-11$^{-/-}$ MuSCs undergo premature senescence

Taken together, these results clearly point toward a role of lamin A in mediating PcG-transcriptional repression in MuSCs to safeguard their identity and regenerative capacity. This lack of cell identity and the impairment of self-renewal displayed by *Lmna*Δ8–11$^{-/-}$ MuSCs are all features reminiscent of the phenotype described for Ezh2-null MuSCs (24). Moreover, the impairment in self-renewal and the progressive decline of the MuSC pool are also typical traits of aged MuSCs (37) in which both lamin A/C and PcG proteins play a key role (38, 39). A major cellular mechanism that ensures self-renewal and hence the maintenance of the MuSC pool is asymmetric division (40). At the molecular level, in aged mice, the accumulation of the activated form of p38 (phospho-p38 [p-p38]) and its symmetric distribution in MuSC doublets heavily compromise the self-renewal capacity, leading to MuSC functional decline (41, 42). To test whether premature exhaustion of quiescent *Lmna*Δ8–11$^{-/-}$ MuSCs could be ascribed to defective asymmetric division, we stained myofiber-associated MuSCs for p-p38 after 48 hours of culture, a timing at which d19 myofibers formed MuSC-derived doublets (Figure 6A). In contrast to WT, *Lmna*Δ8–11$^{-/-}$ MuSC doublets showed a preferential symmetric distribution of p-p38, quantified by relative fluorescence intensity (Figure 6, A and B), often accompanied by a planar orientation with respect to myofibers (see Supplemental Methods and Figure 6C). This highlights problems in asymmetric division, which should be instead characterized by apico-basal orientation (43). In line with this result,

in *Lmna*Δ8–11$^{-/-}$ muscle sections we found higher amounts of p-p38^{+} (Supplemental Figure 11, A and B) MuSCs and signs of genomic instability, as measured by increased γH2AX DNA repair signal foci (Supplemental Figure 11, C and D), not accompanied by apoptosis or necrosis as evidenced by annexin staining (Supplemental Figure 12A). To test if defective asymmetric division is associated with premature senescence we then analyzed RNA-seq to determine if *Lmna*Δ8–11$^{-/-}$ MuSCs share the same transcriptional signature of MuSCs isolated from aged mice. We performed a GSEA using 2 different RNA data sets from MuSCs of 24-month-old mice (25, 44) and *Lmna*Δ8–11$^{-/-}$ upregulated genes. In line with our hypothesis, we found that d19 *Lmna*Δ8–11$^{-/-}$ MuSCs present an upregulated transcriptome similar to 20- to 24-month-old MuSCs (Figure 6D and Supplemental Figure 12B), but different from geriatric 28- to 32-month-old MuSCs (Supplemental Figure 12C). At the molecular level, the senescence program is supported by upregulation of some PcG-regulated cyclin-dependent kinase inhibitors (CDKIs) (45) such as p21, which is involved in cellular senescence and in cell cycle arrest. p21 maintains the viability of DNA damage–induced senescent cells (46) and aberrant expression of p21 has been observed in EDMD-derived human myoblasts (47). ChIP-seq and RNA-seq analyses of the *Cdkn1a*/p21 locus showed a displacement of Ezh2 from the promoter accompanied by an upregulation of p21 transcripts in *Lmna*Δ8–11$^{-/-}$ MuSCs (Figure 6E).

Genetic ablation of *Cdkn2a* locus partially rescues self-renewal defects in *Lmna* Δ8-11 -/- dystrophic mice

To further corroborate our findings, we also analyzed the *Cdkn2a* locus, a PcG target primarily involved in muscular

senescence (ref. 44 and Supplemental Figure 13A). Two transcripts, $p16^{INK4a}$ and $p19^{ARF}$, originate from the *Cdkn2a* locus (48). Interestingly, it was recently reported that $p16^{INK4a}$ expression is a second event, subsequent to p21 upregulation, in cellular senescence progression (49). In line with these observations, $p16^{INK4a}$ expression is specifically induced in geriatric 28- to 32-month-old MuSCs (but not in 24-month-old MuSCs) (44). Moreover, depletion of $p16^{INK4a}$ is sufficient to reduce senescence-associated gene expression in geriatric MuSCs. RNA-seq and quantitative reverse transcription PCR (qRT-PCR) on d19 *Lmna*Δ8–$11^{-/-}$ MuSCs did not reveal any transcription of $p16^{INK4a}$ in WT or *Lmna*Δ8–$11^{-/-}$ (Figure 7A). However, qRT-PCR analysis performed in older mice (d26) revealed higher levels of $p16^{INK4a}$ transcripts in *Lmna*Δ8–$11^{-/-}$ MuSCs and whole muscles compared with the WT counterpart (Figure 7A), suggesting a transition during dystrophy progression toward a geriatric condition. We thus decided to test whether genetic ablation of the *Cdkn2a* locus could reverse *Lmna*Δ8–$11^{-/-}$ MuSC premature aging, by crossing *Lmna*Δ8–$11^{+/-}$ with *Cdkn2a*$^{-/-}$ mice (50). Analysis of the *Lmna*Δ8–$11^{+/+}$ background showed no differences in the percentages of QSCs and ASCs, nor in CSA, suggesting that *Cdkn2a* is dispensable for postnatal muscle development (Figure 7, B–D; *LMNA*$^{+/+}$). On the other hand, *Cdkn2a*$^{-/-}$ *Lmna*Δ8–$11^{-/-}$ mice partially rescued the quiescent MuSC pool and CSA defects observed in the absence of lamin A (Figure 7, B–D; *LMNA*$^{-/-}$; Supplemental Figure 13, B and C), emphasizing that lamin A–dependent muscular dystrophy might be due to progressive MuSC functional decline caused by acquisition of premature aging features.

Discussion

Lamin A–dependent muscular dystrophy pathogenesis has been classically ascribed to nuclear fragility that renders myonuclei more prone to mechanical stress and damage imposed by myofiber contraction (4). However, the evidence that lamin A/C is expressed also by MuSCs has led to the suggestion that satellite cell dysfunction might contribute to EDMD progression (51), yet experimental evidence in support of this hypothesis is still lacking.

Cell fate choice during muscle differentiation is governed by epigenetic factors controlling the sequential restriction of transcriptional programs (52). Any dysfunction in this finely tuned epigenetic regulation could lead to impaired or aberrant cell fate determination (53). Here, we show that lamin A/C is indeed crucial to preserve MuSC identity and regenerative capacity. We demonstrate that cell-autonomous lamin A–dependent polycomb dysfunction leads to MuSC functional decline, which culminates with impaired regenerative capacity and a dystrophic phenotype (Figure 1). Traditionally, the role of lamin A/C in muscle differentiation has been considered to cause defects in muscle differentiation (54, 55). However, in other conditions, MuSCs from lamin A/C–null mice showed a normal ability to differentiate and to form myotubes (20, 56). By moving the viewpoint from differentiation to cell identity we now propose that in the absence of lamin A/C, a portion of MuSCs are derailed from their fate, affecting the quiescent MuSC pool. At the molecular level, we describe a mechanism of lamin A–dependent deregulation of PcG targets showing the spreading of repressive marks along the chromatin fiber (Figure 3), with lack of the expected higher order structures and consequent derepression of bivalent genes (Figures 4, 5, and 8). Recently, it has been proposed that PcG domains can have distinct size and boundary characteristics (57); upon differentiation, loci directly involved in fate specification lose

PcG-mediated looping interactions, allowing new active promoter/enhancer interactions. In parallel, other PcG domains, such as the Hox clusters, do not change their 3D architecture. Our findings further corroborate this hypothesis, showing that the stability of PcG-interacting domains correlates with PcG occupancy and depends on lamin A (Figures 5 and 8). PcG dysfunctions drive *Lmna*Δ8–11$^{-/-}$ MuSCs toward 2 not mutually exclusive fates (Figure 8): lack of cell identity, highlighted by the presence of MuSCs coexpressing muscle and adipogenic markers (Figures 4 and 5); and premature senescence, as shown by defects in asymmetric division and accumulation of p-p38 and γH2AX (Figure 6 and Supplemental Figure 11). These epigenetic alterations determine a progressive decline in MuSC self-renewal that accompanies the muscular dystrophy progression (Figure 1), ultimately leading to a geriatric condition characterized by the expression of p16^{INK4a} from the *Cdkn2a* locus (Figure 7 and ref. 58). Genetic ablation of the *Cdkn2a* locus can recover some muscular dystrophy defects of the *Lmna*Δ8–11$^{-/-}$ mouse (Figure 7), thus supporting the hypothesis that dystrophic and aging muscles share dysfunction in epigenetic mechanisms controlling cell cycle and fate decisions of MuSCs. Our findings corroborate recent evidence of PcG dysfunction in human disease (19), showing that PcG alterations contribute to pathology progression and severity in EDMD. This will further stimulate future studies on the role of PcG proteins in the dynamics of the stem cell niche, when embedded in a pathological environment.

Methods

Animals

Heterozygous B6.129S1(Cg)-*Lmnatm1Stw*/BkknJ mice (*Lmna* Δ8-11 +/-) (21) and *Cdkn2a* +/- mice (50) were used.

Satellite cells extraction, apoptosis evaluation and multiple injuries

Hind-limb muscle were isolated from sacrificed mice and digested 120 minutes in 2,4 U/ml of Dispase II (Roche, 04942078001), 2 ug/ml of Collagenase A (Roche, 1013586001), 0,2 mM CaCl2 (Sigma, C5670), 4 mM MgCl2 (Sigma, M8266), 10 ng/ml DNase I (Roche, 1014159001) in PBS1X (Euroclone, ECB4004L) at 37°C in a water bath. The sample were resuspended in HBSS (Gibco, 14025-050) supplemented with 0,1% BSA (Sigma, A7030). Cell suspension was serially filtered with 70 (Falcon, 352350) and 40 μm (Falcon, 352340), stained with antibodies indicated in supplementary table S2: PB-CD45 1:50, PB-CD31 1:50, PB-Ter119 1:50, FITC-Sca1 1:50, APC-α7integrin 1:200 and sorted with BD FACS ARIA III for: PB-CD45^{-}/ PB-CD31^{-}/ PB-Ter119^{-}/ FITC-Sca1^{-}/ APC-α7integrin^{+}. For multiple injury experiment, 20 μl of Cardiotoxin (CTX) 10 μM (Latoxan, L8102) were injected in TA muscle each week for 3 weeks. TA muscle was harvested one week after the last CTX injection. For apoptosis assay we stained MuSCs with Annexin. 15-20K of satellite cells were washed in 1 mL cold PBS 1X up, centrifuged 7 minutes at 400g and incubated 20 minutes at room temperature in the dark with 100 μL Annexin V FITC buffer (FITC Annexin V Apoptosis Detection Kit I-ref. 556547). Then, samples were washed with 500 μL of cold PBS 1X, centrifuged 7 minutes at 400g and incubated 15 minutes at room temperature in the dark with 300 μL of

cold PBS 1X containing 5 μL of Propidium iodide. Samples were analyzed at BD FACS-CANTO (voltage FSC=357, SSC=462, medium flow to acquire, 300-400 events/second).

Immunofluorescence

On muscle sections: *Tibialis Anterior* (TA) muscles were embedded in Killik (Bio-Optica, 05-9801), immediately frozen in pre-cooled isopentane (Sigma, 277258) and sectioned (Leica CM1850 cryostat) at 8 μm thin. Sections were fixed 20 minutes in paraformaldehyde (PFA) 4% (Sigma, P6148) and washed 3x5 minutes in PBS1X (Euroclone, ECB4004L). To permeabilize tissues pre-cooled methanol (Sigma, 322415) at -20°C was added for 6 minutes. Antigen retrieval was performed 2x5 minutes in hot citric acid 80°C (Sigma, C0759) pH6.0 and washed 2x5 minutes in PBS1X. Sections were blocked 1 h in BSA 4% (Sigma, A7030) followed by incubation 45 minutes with FAB mouse fragment 1:100 (Jackson Immuno Research, 115-007-003)/ PBS1X. Primary antibodies were diluted 1:100 in blocking solution, except for pparγ diluted 1:75, and incubated O/N at 4°C. The day after, sections were washed 3x10 minutes in PBS1X/0,1% BSA and incubated with secondary antibodies in blocking solution 1:200 1h at RT in the dark. Then, sections were washed 3x10 minutes in 0,1% BSA/PBS1X and incubated 2h at RT with Laminin and Pax7 (1:20). After washing 3x10 minutes in PBS1X/0,1% BSA sections were incubated 45 minutes with Cy5 (1:300) and Biotin (1:500) for Pax7 signal amplification. After washing 3x10 minutes in PBS1X, sections were incubated 45 minutes with secondary antibodies Cy3-streptavidin (1:1250). The sections were finally washed 3x10 minutes in PBS1X, stained 5 minutes with dapi (Sigma, D9542), briefly washed twice in PBS1X and mounted on slide with a drop of Prolong Antifade (Life P36930).

On single myofibers: *Tibialis Anterior* (TA), *Soleus* (S), *Gastrocnemius* (G) and *Exstensor Digitorium Longus* (EDL) were isolated from mice and digested 45-50 minutes in 0,2% collagenase type I (Sigma, C0130)/DMEM (Gibco, 10569010) at 37°C. 2 rounds of myofibers washes were performed in pre-coated dishes with 20% FBS (Gibco, 10500064)/DMEM. Myofibers were let grown in DMEM supplemented with 20% FBS, 1% Chicken Embryo extract (Seralab, CE650-DL) and 1% Penicillin/Streptomycin (Euroclone, ECB 3001) for 48h or 96h changing the medium only after 72h. Myofibers were collected in 2ml tubes pre-coated with 10% FBS/PBS1X and fixed 15 minutes with 4%PFA followed by 3 washes in PBS1X. Permeabilization was performed 5 minutes with 0,5% Triton X-100 (Sigma, 93443)/PBS1X followed by 2 washes in PBS1X. Myofibers were incubated 1h in blocking solution (10% of FBS/PBS1X). Primary antibodies were incubated in blocking solution ON at 4°C. The day after myofibers were washed in 0,25% tween/PBS1X twice and incubated 60 minutes with secondary antibodies in blocking solution. Fibers were washed in 0,1% tween/PBS1X, incubate 5 min with dapi, briefly washed twice in PBS1X and mounted on slide with a drop of Prolong Antifade (Thermofisher, P36970). On satellite cells: In order to preserve the integrity of chromatin architecture, Muscular stem (Satellite) Cells (MuSCs) suspension was fixed in 1% PFA for 9 minutes and quenched with 125 mM Glycine (Sigma, 8898) before FACS staining and sorting. SC cells were placed on pre- Poly-L-lisined coverslips (Sigma, P8920) at density of 100.000/mL for 30 minutes at RT. Coverslips were fixed with PFA at 4%/PBS for 10 minute at RT. Then, cells were washed in 2X5 minutes in 0,05% Triton X-100/PBS 1X, permeabilized with 0,5% Triton X-100/PBS1X for 10 minutes and rinsed in PBS 1X. The slides were let in 20% Glycerol (Sigma, G5516)/PBS 1X at least 60 minutes followed by 4 round of freeze and thaw: freezed on dry ice (30 seconds)

and thawed in 20% Glycerol/PBS 1X at RT. Slides were washed 2X5 minutes in 0,05% Triton X-100 in PBS 1X, 1x5 minutes 0,5% Triton X-100/PBS1X, incubated in HCl 0,1M (Sigma, H1758) for 15 minutes and rinsed in PBS 1X. Aspecific signals were blocked with 1% BSA/PBS1X for 30 minutes at room temperature. Reaction with primary antibodies Ezh2 diluted 1:100 in blocking solution was performed 12–16 h at 4°C; Lamin A/C diluted 1:200 in blocking solution was performed the day after at room temperature for 2 hour. Secondary antibodies were diluted 1:200 in blocking solution for 1 h at room temperature. Washes were done in PBS1X. DNA was counterstained with dapi, and glasses were mounted in Prolong Antifade. Primary antibodies are listed in supplementary table S2. Secondary antibodies were from Jackson Immuno Research: Alexa 488 (711-545-152); Cy5 1:300 (111-115-144); Cy3-streptavidin (016-160-084); Alexa 594 (711-545-150).

Three-dimensional multicolor DNA FISH

To produce probes for 3D multicolor DNA FISH we used the following BAC DNA clones (BACPAC Resources Program, CHORI): CH29-101F16 (for Pparγ) and CH29-555O5 (for the upstream region of Pparγ). 1-3 µg of BAC DNA were labelled with dig-dUTP (Roche, 11093088910) (for the upstream region of Pparγ) or cy3-dUTP (Thermo Fisher Scientific, C11401) (for Pparγ) through nick translation in 50 µl of Labelling mix buffer (0.02 mM C-G-A dNTPs: Euroclone, EMR273025, EMR274025, EMR272025 respectively, 0.01 mM dTTP Euroclone EMR275025, 0.01 mM labelled dUTP, 50 mM Tris-HCl pH 7.8, 5 mM MgCl2, 10 mM b-mercaptoethanol, 10 ng/µl Bovine serum albumin (BSA), 0.05-0.1 U/µl DNA Polymerase I (Thermo Fisher Scientific, 18010-017), 0.004-0.001 U/µl Amplification Grade DNase I (Sigma, D5307) for 30 min-2 h at 16 °C, to obtain an average probe

size of 50 bp. Probes were collected by ethanol precipitation, resuspended in 10 mM Tris-HCl pH 7.5 and then quantified using a Nanodrop 1000 Spectrophotometer (Thermo Fisher Scientific). For a single experiment 100-300 ng of each probe was precipitated with 3.5 µg of Mouse Cot-1 DNA (Thermo Fisher Scientific, 18440-016) and 20 µg of Deoxyribonucleic acid, single stranded from salmon testes (Sigma, D7656), and then resuspended in 6 µl of Hybridization solution (50% formamide pH 7.0 (FA)/2X SSC/10% Dextran sulfate). Pre-fixed Satellite cells Cells were plated directly on poly-L-Lisyned coverslips and fixed with 4% Paraformaldehyde (PFA) in 1X PBS and TWEEN 20 0.1% (PBS-T) for 10 min at room temperature. During the last minute few drops of 0.5% Triton X-100 in 1X PBS (PBS) were added and then cells were washed with 0.01% Triton X-100 in PBS three times for 3 min at room temperature. Cells were first permeabilized with 0.5% Triton X-100 in PBS for 10 min at room temperature. In order to remove RNA, samples were treated with RNase Cocktail Enzyme Mix (Thermo Fisher Scientific, AM2288) for 1 h at 37 °C. Cells were subjected to other steps of permeabilization with 20% Glycerol in PBS overnight at room temperature, followed by four cycles of freeze and thaw interleaved by soak with 20% Glycerol in PBS. Permeabilized cells were washed with PBS three times for 10 min at room temperature. Cells were then incubated in 0.1 M HCl for 5 min at room temperature, followed by a rinse with 2X SSC and then incubated in 50% FA in 2X SSC for at least 30 min at room temperature. Slides were equilibrated in 2X SSC for 2 min, washed in PBS for 3 min and then treated with 0.0025-0.0075% pepsin in 0.01-0.03 N HCl for 2-4 min at room temperature to eliminate cytoskeleton. Pepsin was inactivated with 50 mM MgCl2 in PBS twice for 5 min. Nuclei were post-fixed with 1% PFA in PBS for 1 min, washed with PBS for 5 min and with 2X SSC twice, and then back to 50% FA in 2X SSC for at least 30 min at room

temperature. Hybridization solution was loaded on a clean microscopic slide, coverslip with nuclei was turned upside down on the drop of hybridization mixture and sealed with rubber cement. Samples were denatured for 4 min at 75 °C and leaved to hybridize in a metallic box floating in a 37 °C water bath overnight. Samples were washed with 2X SSC three times for 5 min at 37 °C and with 0.1X SSC three times for 5 min at 60 °C, followed a rinse with 0.2% TWEEN 20 in 4X SSC. Aspecific binding sites were blocked with Blocking solution (4% BSA in 4X SSC, 0.2% TWEEN 20) for 20 min at 37 °C. Samples were then incubated in the appropriate concentration of Streptavidin, DyLight 488 Labeled Anti-Digoxigenin/Digoxin (Vector Laboratories, DI-7488) (1:100) diluted in Blocking solution for 35 min in a dark and wet chamber at 37 °C. Samples were washed with 0.2% TWEEN 20 in 4X SSC three times for 3 min at 37 °C, equilibrated in PBS and post-fixed with 2% formaldehyde in PBS for 10 min at room temperature. Finally, the 3D-fixed nuclei were washed with PBS three times for 5 min at room temperature, counterstained with 1 ng/µl DAPI in PBS for 10 min at room temperature and washed with PBS two times for 5 min at room temperature. Coverslips were mounted. An Eclipse Ti-E (Nikon Instruments) microscope was used to scan the nuclei, with an axial distance of between 0.2-0.25 micron consecutive sections.

Histone extraction and Western blot

Total proteins were prepared starting from quadriceps muscle. Muscles were homogenized on ice with tissue ruptor (Qiagen, 902756) in 1ml of extraction buffer (50 mM Tris HCl, pH 7.5, 150mM NaCl, 1 mM EDTA pH 8.0, 0,1% SDS, 1% Np40 (Sigma, 74385), 0,5% Sodium deoxycholate (Sigma, D6750) 1× protease Inhibitors (Roche, 04693132001) 2 mM PMSF (Sigma, 93482), 1 mM NaF (Sigma, s7920); 1 mM Na3VO4). After 30 minutes in ice 3 pulses of 10 seconds

sonication at 30% amplitude were performed with Branson digital sonifier 250 to allow dissociation of protein from chromatin and solubilization. Extracts were analyzed by SDS-PAGE using a 8–10% gel (29:1 acrylamide solution Sigma, 01708). For histone extractions approximately 2.000.000 of freshly isolated MuSCs was collected and resuspended in 15uL of 0.2 N HCl with 1× protease Inhibitors (Roche, 04693132001) 2 mM PMSF (Sigma, 93482), 1 mM NaF (Sigma, s7920); 1 mM Na3VO4) and incubated overnight at 4°C with constant rocking. The supernatant was run on 4-12% bis-tris Acrylamide gel (Thermo Fisher Scientific, NW04125).
Primary antibodies are listed in supplementary table S2.

RNA-seq

Total RNA from freshly isolated satellite cells from 3 mice for each genotype was extracted in TRI-Reagent (Sigma, T9424) following the relative guidelines. Libraries were prepared using the Illumina TruSeq Stranded Total RNA with Ribo-Zero GOLD kit and were quality controlled with an Agilent Bioanalyzer at the Sequencing Facility of the Institute of Applied Genomics (IGA, Udine, Italy). 125 bp reads were produced using an Illumina HiSeq2500 machine in paired-end mode to reach a sequencing depth of about 30 million reads for each sample. RTPCR were performed on total RNA extracted from single mouse. Primers sequences: gapdh: gtatgtcgtggagtctactgg, tcgtggttcacacccatcac; p16: gtgtgcatgacgtgcggg, cagttcgaatctgcaccgtag; p19: gctctggctttcgtgaacatg, tcgaatctgcaccgtagttgag; p21: acggaggcagaccagcct, acacagagtgagggctaagg.

ChIP-seq

For ChIP analysis, satellite cells from pools of 6-9 mice were used and prefixed as previously described (Immunofluorescence section). Isolated satellite cells were resuspended in SDS buffer (100 mM NaCl, 50mM TrisHCl pH8.1, 5 mM EDTA, 0.2% NaN_3, 0.5% SDS) and stored at -80°C (100 µL/10^6 cells). 2-2.5*10^6 of fixed cells/experiment were thawed on ice, resuspended in fresh SDS buffer and incubated at 4°C in mild agitation for 3 hours, passing them through a 0.50x16mm syringe needle every hour. Solution was then adjusted to IP Buffer composition (100 mM Tris pH 8.6, 0.3% SDS, 1.7% Triton X-100, 5 mM EDTA) and cells were sonicated with Branson Digital Sonifier to 200 bp fragments. For spike-in experiments, a 5% of sheared drosophila chromatin was added to each sample. The 2% of the total volume from each sample was taken as input chromatin. The remaining fragmented chromatin was incubated with 1 mM PMSF (Sigma, 93482) and 4 µg of the antibody of interest on a rotating wheel at 4°C overnight. Primary antibodies used were: H3K27me3 (Millipore 07-449) and H3K4me3 (Millipore 07-473). The next day, protein G beads (Life Technology, 1004D) were added (80 µL) and samples were incubated for additional 2 hours on the rotating wheel at 4°C. The beads were washed with Low Salt solution (150 mM NaCl, 20 mM TrisHCl pH 8.0, 2 mM EDTA, 0.1% SDS, 1% Triton X-100), High Salt solution (500 mM NaCl, 20 mM TrisHCl pH 8.0, 2 mM EDTA, 0.1% SDS, 1% Triton X-100), Low Salt solution and then TE NaCl (50 mM NaCl, 10 mM TrisHCl pH8.0, 1 mM EDTA). Crosslinking was reversed by incubating the beads at 65°C overnight in Elution buffer (50 mM TrisHCl pH 8.0, 10 mM EDTA, 1% SDS). Input chromatin was also decrosslinked in Elution buffer overnight at 65°C. The next day, all samples were diluted with one volume of TE 10:1, treated with 0.2 ug/mL RNase A (Sigma R6513) for 2 hours at 37˚C and then with

0.2ug/mL Proteinase K (Sigma P2308) for 2 hours at 55°C. DNA was isolated through standard phenol/chloroform extraction, followed by precipitation and resuspension in 31 µL of 10 mM Tris-HCl pH 8.0. Before library construction, ChIP were validated using the primers: hoxd9 ggataatcgcctaggtgtgactt, catctcttcttgcctctctggg and pax7prom gcgacccccctgaggaaaa, cgaaaagaagtctccaacgagtatt. Libraries for IP and reference input DNA were created for each sample using the automation instrument Biomek FX (Beckman Coulter), then qualitatively and quantitatively checked using Agilent High Sensitivity DNA Kit (Agilent Technologies, 5067-4627) on a Bioanalyzer 2100 (Agilent Technologies). Libraries with distinct adapter indexes were multiplexed and, after cluster generation on FlowCell, were sequenced for 50 bases in the single read mode on a HiSeq 2000 sequencer at the IEO Genomic Unit in Milan.

Real-time PCR analysis

Total RNA was extracted from satellite cells using Maxwell RSC miRNA tissue kit (Promega, AS1460), while total RNA from muscle tissues was extracted using TriReagent (Sigma-Aldrich, T9424). 1ug of RNA from each sample was subjected to cDNA synthesis using QuantiTect reverse transcription kit (Qiagen, 205313) and amplified in the presence of 8 ul of SYBR select master mix (Thermo fisher, 4472908). Expression was calculated by normalizing on GAPDH and relative to the average of the wild type controls samples. Primer sequences used for transcriptional analyses, mouse: mRTP21 5'-acggaggcagaccagcct-3', 5'-acacagagtgagggctaagg-3'; mrtezh2 5'-caaatacatgtgcagctttctg-3', 5'-atgcctatcctgtggtcacc-3'; mrtpparg 5'-ttgctgaacgtgaagcccatcgag-3', 5'-gtccttgtagatctcctggagcag-3'; mrtp16 5'-gtgtgcatgacgtgcggg-3', 5'-cagttcgaatctgcaccgtag-3';

mrtp19 5’-gctctggctttcgtgaacatg-3’, 5’-tcgaatctgcaccgtagttgag-3’; mrtgapdh 5’-gtatgtcgtggagtctactgg-3’, 5’-tcgtggttcacacccatcac-3’.

Data availability

The datasets generated during the current study are available in the GEO repository with accession number GSE123725.

Statistics

All the data are represented using Graph Pad prism 6. The sample size (n) is described for each experiment in the relative figure legend. Multiple comparison between three o more groups were made using one way Anova or two way Anova with significance as P value <0,05 (*), P value < 0,01 (**), P value <0,001 (***).

Study approval

Heterozygous B6.129S1(Cg)-*Lmnatm1Stw*/BkknJ mice (*Lmna* Δ8-11 +/-) (21) and *Cdkn2a* +/- mice (50) were used. All the experimental procedures were performed under the ethical approval of the Italian Ministry of Health and the Institutional Animal Care and Use Committee (authorization n. 83/2019-PR). The animals were maintained in an authorized facility at San Raffaele Hospital, Milan (authorization n. 127/2012-A) and Santa Lucia Foundation, Rome (authorization n. 9/2006-A).

Author Contributions

C.L. conceived the project and designed experiments. A.B. and C.M. performed experiments and analysed data. F. L. and D.P. performed and interpreted ChIP-seq experiments. S.V. performed bioinformatic analysis. G.P., V.R., R.R. and C.B. performed experiments in mice. A.B., A.C. and B.B. performed and interpreted FISH experiments.

L.A., F.G. and G.O. quantified immunofluorescence and FISH images. M.D.B. performed FACS. F.F. and C.P. analysed ChIPseq data. C.L. supervised the project and analysed data. C.L., A.B. and C.M. wrote the manuscript. All authors edited and approved the manuscript.

Acknowledgements

We thank Maria Vivo, Gisèle Bonne, Federica Marasca, Pierluigi Manti, Davide Gabellini, Lorenzo Puri, Daniela Palacios, Giovanna Lattanzi, the Italian network of Laminopathies and members of the laboratory for stimulating discussions and constructive criticisms. We are grateful to Chiara Cordiglieri and INGM Imaging Facility for assistance during images acquisition and to Mariacristina Crosti, Monica Moro and INGM FACS Facility for assistance in cell sorting. This work was supported by grants from the Italian Minister of Health n. GR-2013-02355413 to C.L., My First AIRC Grant (MFAG) n. 18535 to C.L., AFM-Telethon n. 21030 to C.L. and F.F. and Cariplo 2017-0649 to C.L. and F.F.

Figures

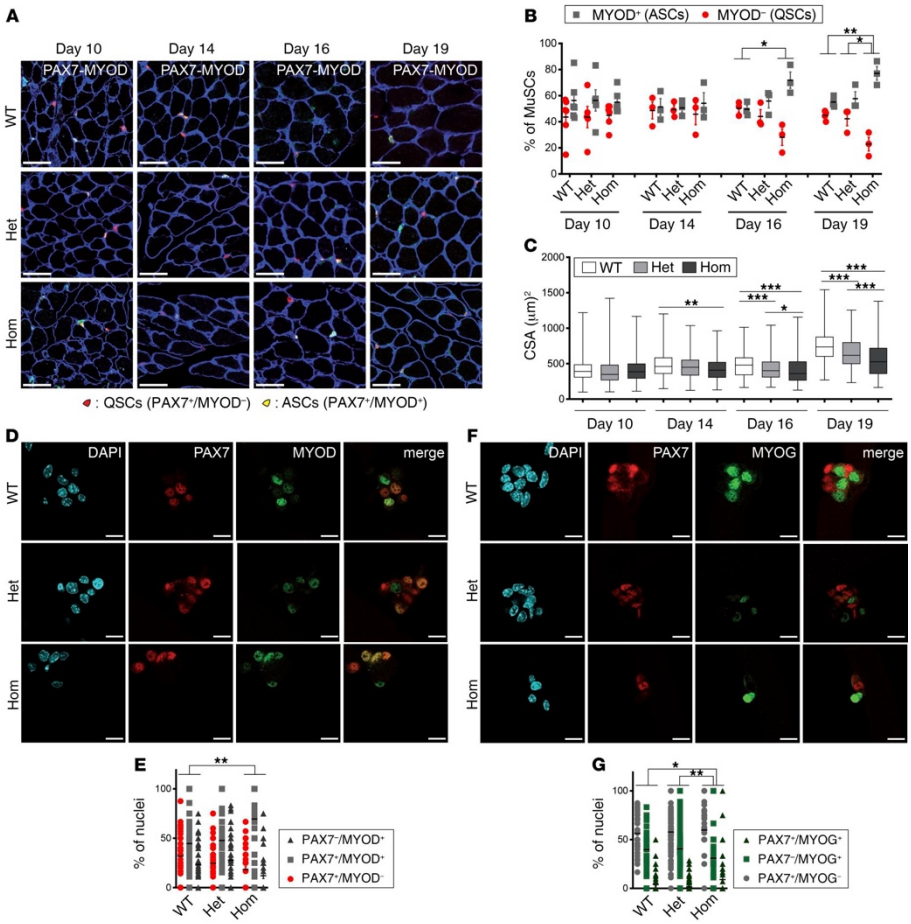


Figure 1. Lamin A regulates MuSC self-renewal.

(**A**) Immunohistochemical staining in *Lmna*Δ8–11 mice of PAX7 and MYOD markers at the indicated days of postnatal growth (d10–d19). Basement membrane of muscle fibers was stained with anti-laminin. Activated, ASC ($PAX7^{+}MYOD^{+}$) and self-renewing, QSC ($PAX7^{+}MYOD^{-}$) MuSCs are shown. Scale bars: 50 µm. (**B**) Quantification of MuSC pool composition in **A**; *n* = 3–6 animals per genotype (**C**) Quantification of myofiber size during postnatal growth, evaluated by the cross-sectional area (CSA). *n* > 350 fibers, *n* = 3–4

animals per genotype. Horizontal lines within the boxes represent the medians, upper and lower bounds of the boxes represent quartiles Q3 (75th percentile) and Q1 (25th percentile), respectively, and the whiskers min to max. (**D**) Immunohistochemical staining of single myofibers extracted from *Lmna*Δ8–11 mice at d19 and cultured 96 hours. Activated ($PAX7^{+}MYOD^{+}$), self-renewing ($PAX7^{+}MYOD^{-}$), and differentiating ($PAX7^{-}MYOD^{+}$) cells are shown. Scale bars: 20 μm. (**E**) Quantification of MuSC pool composition in **D**; $n >$ 50 muscle fibers/genotype, $n =$ 5–8 animals per genotype. (**F**) Immunohistochemical staining of single myofibers as in **D**. $PAX7^{+}MYOG^{+}$, $PAX7^{+}MYOG^{-}$, and $PAX7^{-}MYOG^{+}$ cells are shown. Scale bars: 20 μm. (**G**) Quantification of MuSC pool composition in **F**; $n >$ 50 muscle fibers/group, $n =$ 3–5 animals per genotype. Data in **B**, **E**, and **G** are the mean ± SEM. $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$ by 1-way (**C**) or 2-way (**B**, **E**, and **G**) ANOVA with multiple comparisons. WT = *Lmna*$\Delta 8–11^{+/+}$; het = *Lmna*$\Delta 8–11^{+/-}$; hom = *Lmna*$\Delta 8–11^{-/-}$.

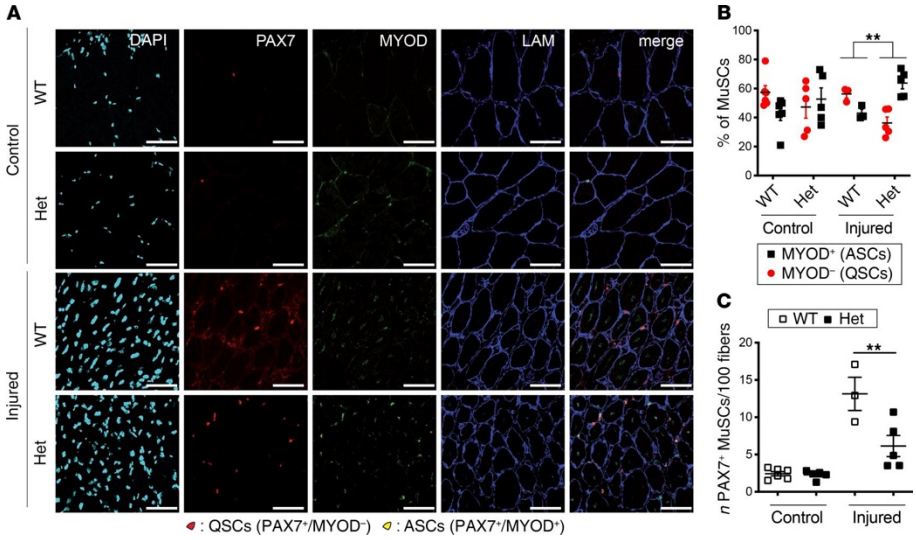


Figure 2. Lamin A levels influence muscle regeneration.

(**A**) Immunohistochemical staining of PAX7 and MYOD markers in adult, injured *Lmna*Δ8–11 mice. Basement membrane of muscle fibers was stained with anti-laminin. Activated, ASC ($PAX7^+MYOD^+$) and self-renewing, QSC ($PAX7^+MYOD^-$) MuSCs are shown. Scale bars: 100 μm. (**B**) Quantification of MuSC pool composition in **A**; *n* = 3–6 animals per group. (**C**) Number of $Pax7^+$ MuSCs on 100 fibers for the experiment shown in **A**. Data in **B** and **C** are the mean ± SEM. Statistics by 1-way (**C**) or 2-way (**B**) ANOVA with multiple comparisons. Statistical comparisons between controls and injured (**B** and **C**) are not shown. $^{**}P < 0.01$; WT = $Lmna\Delta8–11^{+/+}$; het = $Lmna\Delta8–11^{+/-}$.

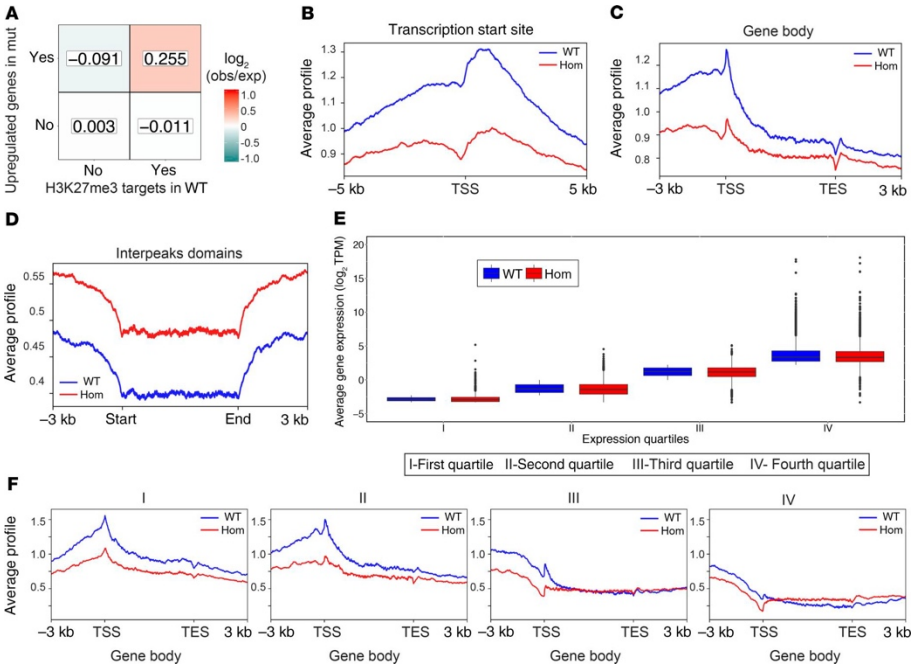


Figure 3. *Lmna* Δ8-11$^{-/-}$ dystrophic MuSCs display PcG displacement.

(**A**) Heatmap reporting $\log_2$ ratios of observed over expected (colored bar) number of genes in the intersections between H3K27me3 targets identified in *Lmna*Δ8–11$^{+/+}$ mice and the upregulated genes in *Lmna*Δ8–11$^{-/-}$ mice. Fisher's exact test $P = 2.38 \times 10^{-5}$. (**B**–**D**) Average profile of H3K27me3 ChIP-seq signal calculated as the IP/input ratio over annotated mouse genes. (**B**) Average profile of H3K27me3 signal around the TSS. (**C**) Average profile of H3K27me3 signal along the gene body. TES, annotated transcript end. (**D**) Average profile of H3K27me3 signal in regions outside H3K27me3 peaks and outside annotated genes. (**E**) Normalized expression distribution of genes stratified using WT expression level in the 3 biological replicates (see Supplemental Methods). Distribution of average $\log_2$ transcripts per million (TPM + 0.1) values is plotted for WT and hom. Data in the boxes extend from the 25th to the 75th

percentiles with the median indicated. The upper whisker extends from the hinge to the highest value that is within 1.5 × IQR of the hinge, where IQR is the interquartile range, or distance between the first and third quartiles. The lower whisker extends from the hinge to the lowest value within 1.5 × IQR of the hinge. Data beyond the end of the whiskers are outliers and plotted as points. (**F**) Average profile of H3K27me3 signal (IP/input) along the gene body using gene categories as in **E**. WT = *Lmna*Δ8–11$^{+/+}$; hom = *Lmna*Δ8–11$^{-/-}$.

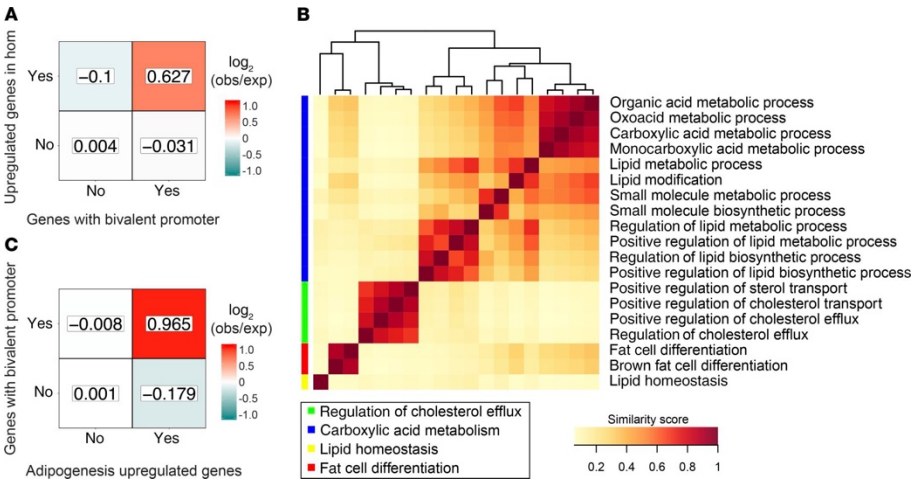


Figure 4. Lamin A-PcG-mediated transcriptional repression preserves MuSCs identity.

(**A**) Heatmap reporting $\log_2$ ratios of observed over expected number of genes (colored bar) in the intersections between bivalent promoters identified in WT satellite cells and the upregulated genes in *Lmna*Δ8–11$^{-/-}$ mice. Fisher's exact test $P = 4.57 \times 10^{-7}$. (**B**) Semantic similarity analysis of GO terms enriched in upregulated genes in hom versus WT comparison (FDR < 0.05) with macrocategories identified using the REVIGO web tool (http://revigo.irb.hr/). (**C**) Heatmap reporting

log_2 ratios of observed over expected number of genes (colored bar) in the intersections between upregulated genes in *Lmna*Δ8–11$^{-/-}$ mice in *Pparg*-related GO terms and the bivalent genes identified as above. Fisher's exact test $P = 6.73 \times 10^{-6}$. WT = *Lmna*Δ8–11$^{+/+}$; hom = *Lmna*Δ8–11$^{-/-}$.

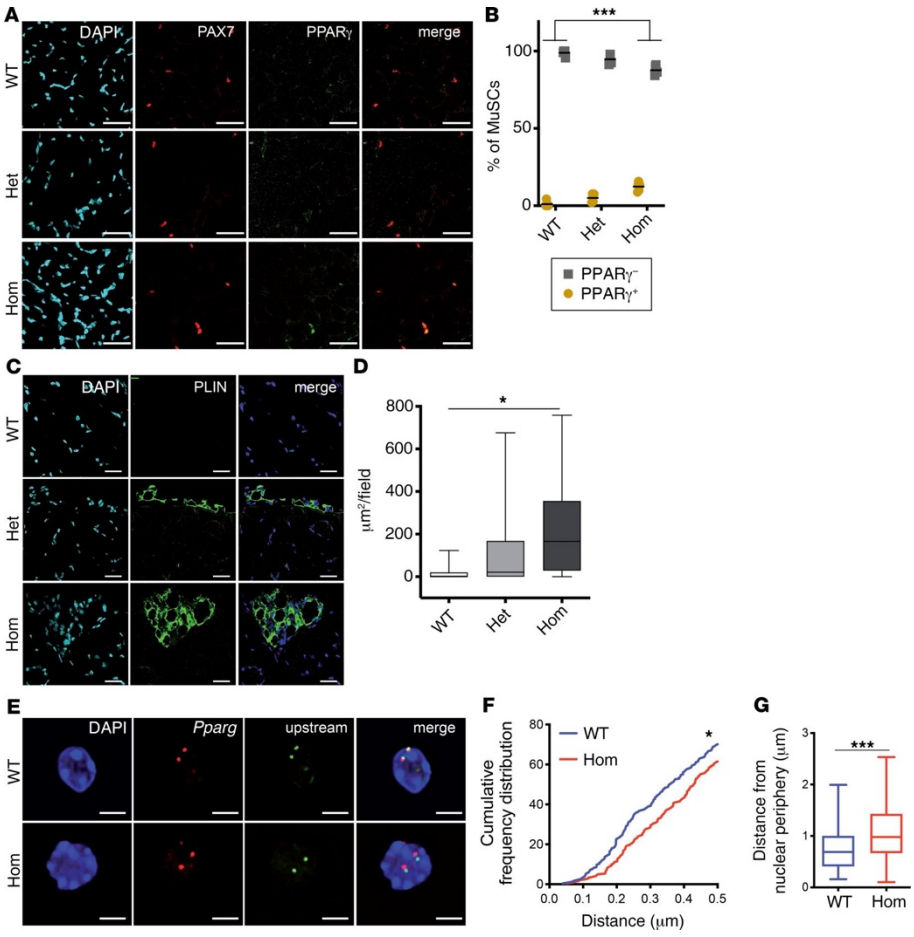


Figure 5. Deregulation of Pparγ locus in *Lmna* Δ8-11 -/- MuSCs.

(**A**) Immunohistochemical staining of PAX7 and PPARγ markers in *Lmna*Δ8–11 mice at 19 days of postnatal growth (d19). Scale bars: 50 µm. (**B**) Quantification of PPARγ$^+$ MuSCs in **A**; *n* = 4–5 animals per

genotype. Data are the mean ± SEM. (**C**) Immunohistochemical staining of perilipin (PLIN) in *Lmna*Δ8–11 muscles at d19. Nuclei of muscle fibers were stained with DAPI. Scale bars: 25 μm. (**D**) Quantification of perilipin staining in **C**; *n* = 5 animals per genotype. (**E**) Representative image of FISH analysis of fixed and sorted MuSCs from *Lmna*Δ8–11 mice at d19 with probes indicated in Supplemental Figure 10. Scale bars: 2 μm. (**F**) Quantification of FISH probe (represented in Supplemental Figure 10) distances (*x* axis) versus cumulative frequency distributions (*y* axis). Only probes with distances of 0.5 μm or less are reported. *n* = 1–2 animals per genotype. (**G**) Quantification of FISH probe position with respect to the nuclear envelope. In the box-and-whisker plots in **D** and **G**, horizontal lines within the boxes represent the medians, upper and lower bounds of the boxes represent quartiles Q3 (75th percentile) and Q1 (25th percentile), respectively, and whiskers min to max. **P* < 0.05; ****P* < 0.001 by 2-way ANOVA with multiple comparisons (**B** and **D**) or Kolmogorov-Smirnov test (**F** and **G**). WT = *Lmna*Δ8–11$^{+/+}$; het = *Lmna*Δ8–11$^{+/–}$; hom = *Lmna*Δ8–11$^{–/–}$.

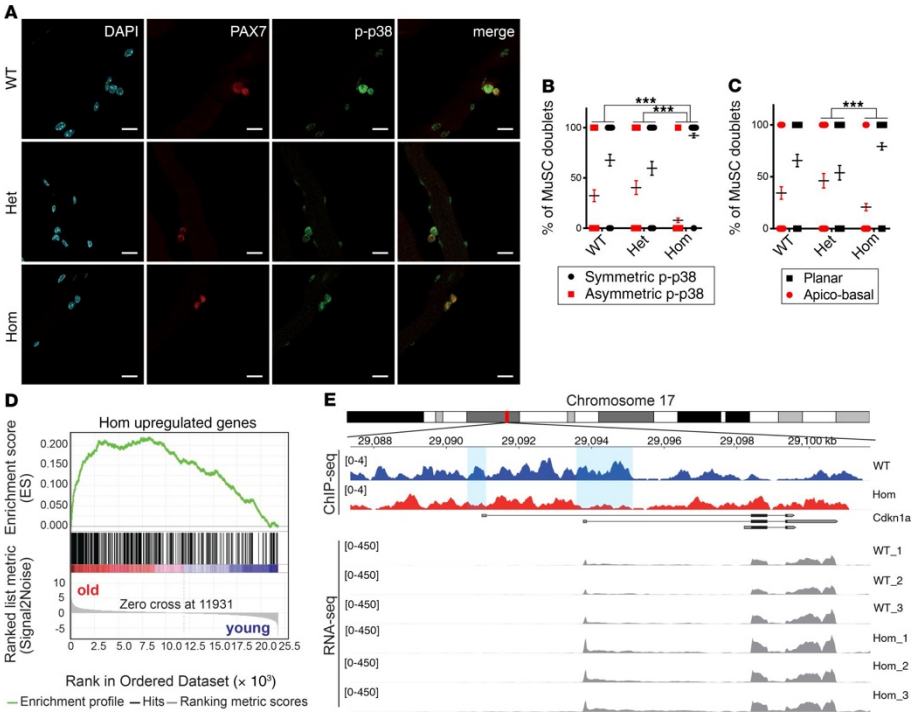


Figure 6. *Lmna* Δ8-11 -/- MuSCs acquire senescence transcriptional traits.

(**A**) Representative image of myofiber-derived MuSCs from *Lmna*Δ8–11 mice at d19 immunostained for p-p38 and PAX7 after 48 hours of culture. Scale bars: 25 μm. (**B**) Quantification of asymmetric and symmetric divisions assessed by p-p38 distribution as shown in **A**. (**C**) Quantification of asymmetric apico-basal division versus symmetric planar divisions. *n* = 46 ± 6 doublets of MuSCs per genotype, *n* = 7–9 mice per group. Data in **B** and **C** are the mean ± SEM. (**D**) GSEA of expression data from old and young mouse quiescent satellite cells (25). Upregulated (log[fold change] > 1) genes in hom versus WT comparison added to Biocarta mouse pathways from the gskb R package were used as gene sets (NES = 4.70, FDR < 1×10^{-4}). (**E**) ChIP-seq of H3K27me3 mark and RNA-seq signal tracks on the *Cdkn1a*/p21 locus. Promoter regions are highlighted by light blue

rectangles. Statistics by 2-way ANOVA with multiple comparisons. $^{**}P < 0.01$; $^{***}P < 0.001$. WT = *Lmna*Δ8–11 $^{/+}$; het = *Lmna*Δ8–11$^{+/-}$; hom = *Lmna*Δ8–11$^{-/-}$.

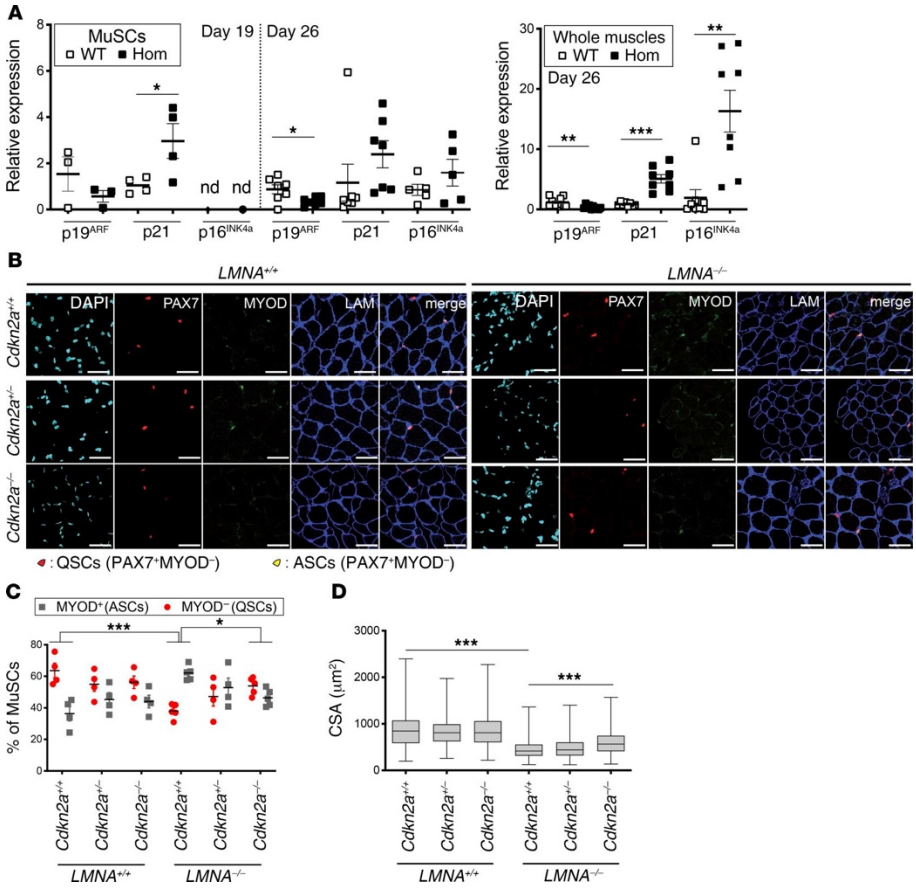


Figure 7. *Cdkn2a* genetic ablation restores regenerative capacity of *Lmna* Δ8-11 -/- dystrophic mice.

(**A**) Transcriptional analysis of p16^{INK4} and p19^{ARF} at the *Cdkn2a* locus in *Lmna*Δ8–11 mouse MuSCs (left graph) at d19 and d26 and whole muscles (right graph) at d26. Values were normalized to *Gapdh* and compared with the average of WT amplification. nd, not detected. *n* = 3–10 animals per genotype. (**B**) Immunohistochemical staining of

PAX7 and MYOD markers in *Cdkn2a*/*Lmna*Δ8–11 mice at d19. Basement membrane of muscle fibers was stained with anti-laminin. Activated, ASC (PAX7^{+}/MYOD^{+}) and self-renewing, QSC (PAX7^{+}MYOD^{-}) MuSCs are shown. Scale bars: 50 μm. (**C**) Quantification of MuSC pool composition in **B**. *n* = 4–5 animals per genotype. Data in **A** and **C** are the mean ± SEM. (**D**) Quantification of myofiber size, evaluated by the cross-sectional area (CSA). *n* = 600 muscle fibers. *n* = 4–5 animals per genotype. The horizontal lines within the boxes represent the medians, upper and lower bounds of the boxes represent quartiles Q3 (75th percentile) and Q1 (25th percentile), respectively, and whiskers min to max. **P* < 0.05; ***P* < 0.01; ****P* < 0.001 by unpaired *t* test (**A**) or by 1-way (**D**) or 2-way (**C**) ANOVA with multiple comparisons. WT = *Lmna*Δ8–11$^{+/+}$; hom = *Lmna*Δ8–11$^{-/-}$.

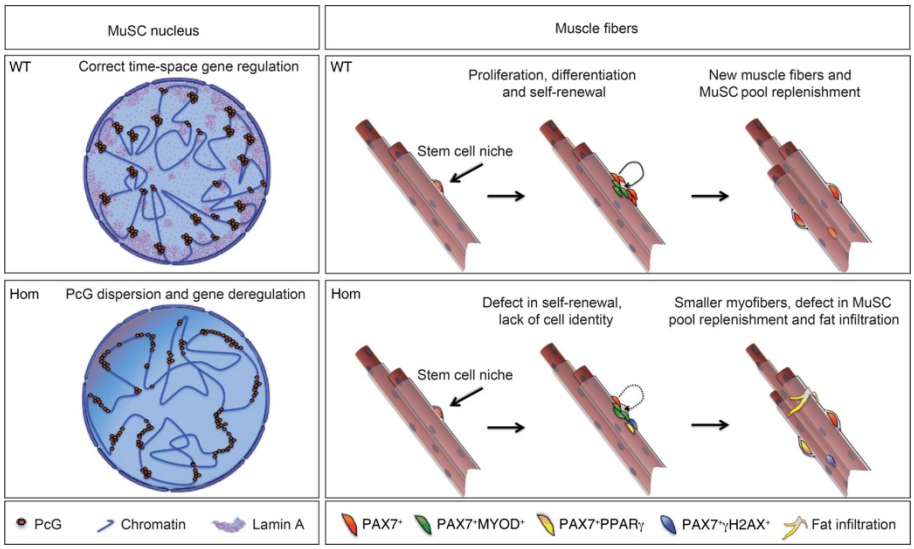
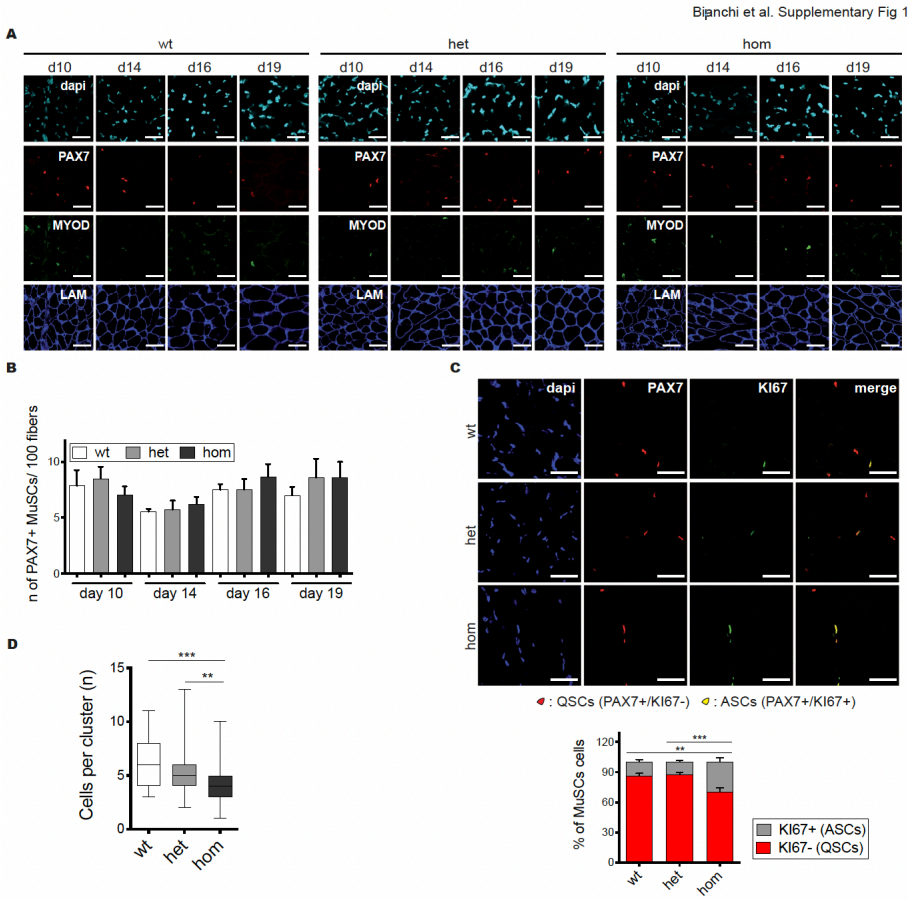


Figure 8. Lamin A/C-Polycomb crosstalk in Lamin dependent muscular dystrophy.

In WT MuSCs the lamin A–PcG interplay sustains the chromatin higher order structure at differentiation loci, ensuring proper spatio-temporal

gene regulation during muscle differentiation. The absence of lamin A/C determines PcG displacement and relaxation of PcG-mediated higher-order chromatin structure. In *Lmna*Δ8–11$^{-/-}$ MuSCs lamin A/C–PcG dysfunctional crosstalk causes a lack of cellular identity and premature senescence, finally leading to an exhaustion of the stem cell niche and infiltration of adipogenic cells. WT = *Lmna*Δ8–11$^{+/+}$; hom = *Lmna*Δ8–11$^{-/-}$.

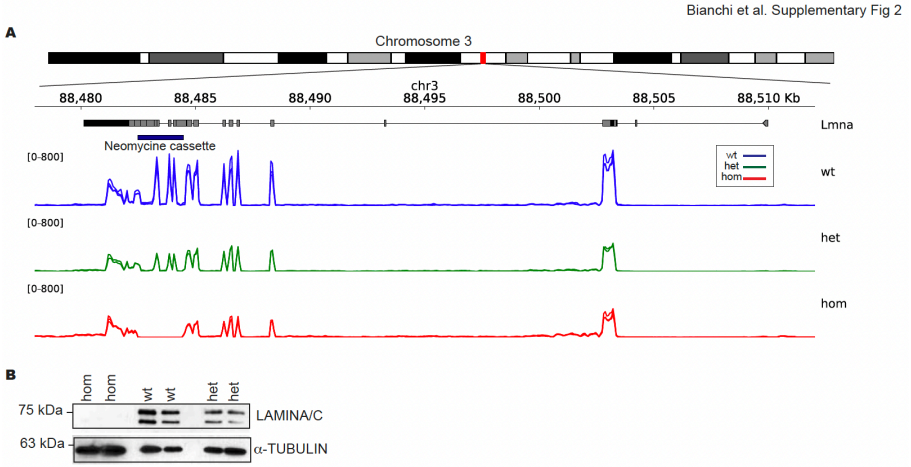
Supplementary Figures



Supplementary Figure 1. Dystrophic mice show unbalanced stem cell niche.

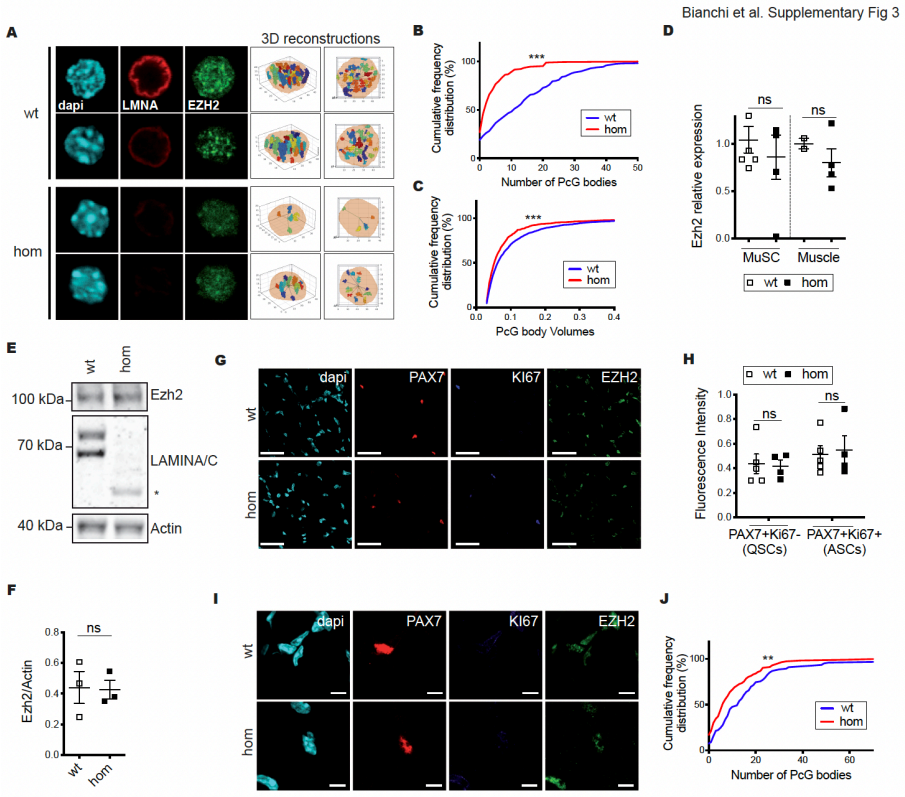
(**A**) Single channel images for the merged panels shown in Figure 1A, for of PAX7, MYOD and Laminin immunohistochemical staining in Lmna Δ8-11 mice of PAX7 and MYOD markers at the indicated days of post-natal growth. Basement membrane of muscle fibers was stained with Laminin. Scale bars, 50 μm. (**B**) Quantification of the % of PAX7+ MuSCs per 100 fibers at the indicated days of post-natal growth in (A). n =3-6 animals per genotype. (**C**) Immunohistochemical staining in Lmna Δ8-11 mice of activated, ASCs (PAX7+/KI67+) and quiescent

QSCs (PAX7+/Ki67-) MuSCs at d19 and relative quantification (below). n= 4-6 animals per genotype. Scale bars, 50 μm. (**D**) Quantification of the number of cells per cluster in single myofibers extracted from d19 Lmna Δ8-11 mice and cultured 96h. n= 4-5 animals per group. Data are box with median and whiskers min to max. B, C, Data are mean ± s.e.m. Statistics by one-way (B) or two-way (C, D) analysis of variance (ANOVA) with multiple comparisons. * * P < 0.01, * * * P < 0.001. wt= Lmna Δ8-11 +/+; het= Lmna Δ8-11 +/; hom= Lmna Δ8-11 -/-.



Supplementary Figure 2. Heterozygous mice show intermediate Lamin A levels.

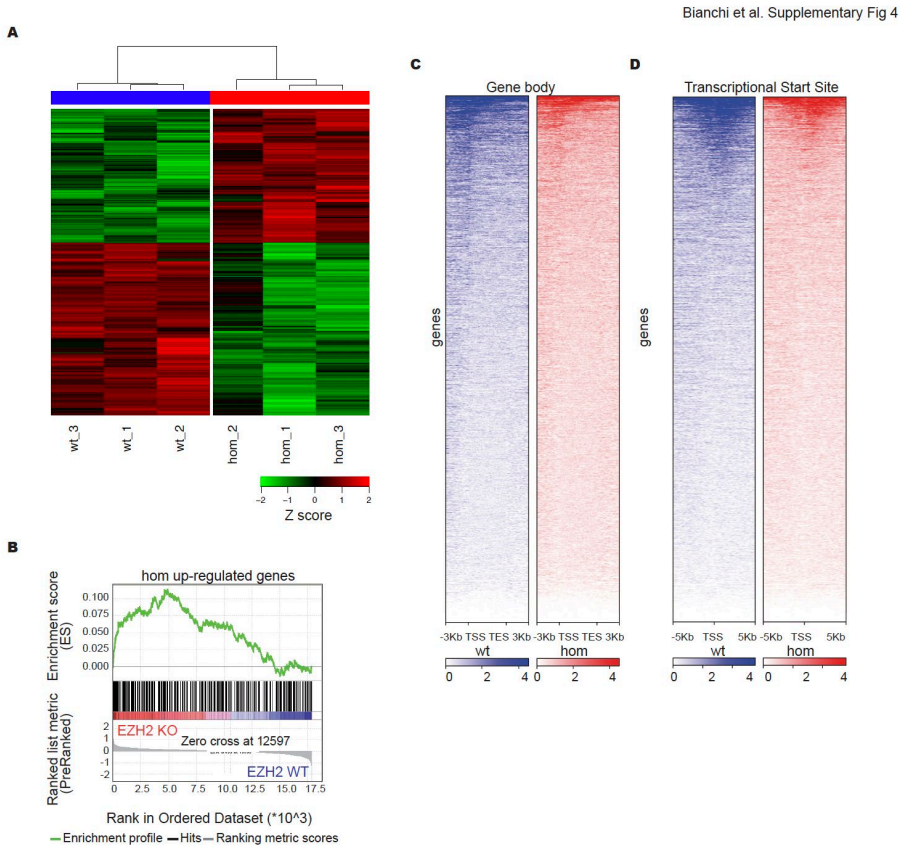
(**A**) RNA-seq signal tracks as the effective genome size normalized coverage of each biological replicate of Lmna Δ8-11 mice on Lmna locus. Neomycine cassette is indicated as a dark blue rectangle. (**B**) Western blot of total protein extracted from the whole Lmna Δ8-11 muscles at d19 hybridized with indicated antibodies. α-tubulin was used as loading control. wt= Lmna Δ8-11 +/+; het= Lmna Δ8-11 +/; hom= Lmna Δ8-11 -/-.



Supplementary Figure 3. Lmna Δ8-11 dystrophic mice show alteration of PcG nuclear distribution.

(**A**) Representative images of Ezh2/Lamin A immunofluorescence analysis on fixed and sorted Lmna Δ8-11 MuSCs at d19. On the right, 3D reconstruction of PcG bodies (see additional methods). Of note, in some Lmna Δ8-11 -/- nuclei we could see low levels of the truncated form of Lamin A/C (1) recognizable by the Lamin A/C antibody. (**B**) Quantification of number of PcG bodies/nucleus (x-axis) in (A) by cumulative frequency distributions (y-axis). n=4 animals per genotype. (**C**) Quantification of number of PcG body volumes (x-axis) in (A) by cumulative frequency distributions (y-axis). n=4 animals per genotype. (**D**) RTPCR analysis of Ezh2 in satellite cells and whole muscles from

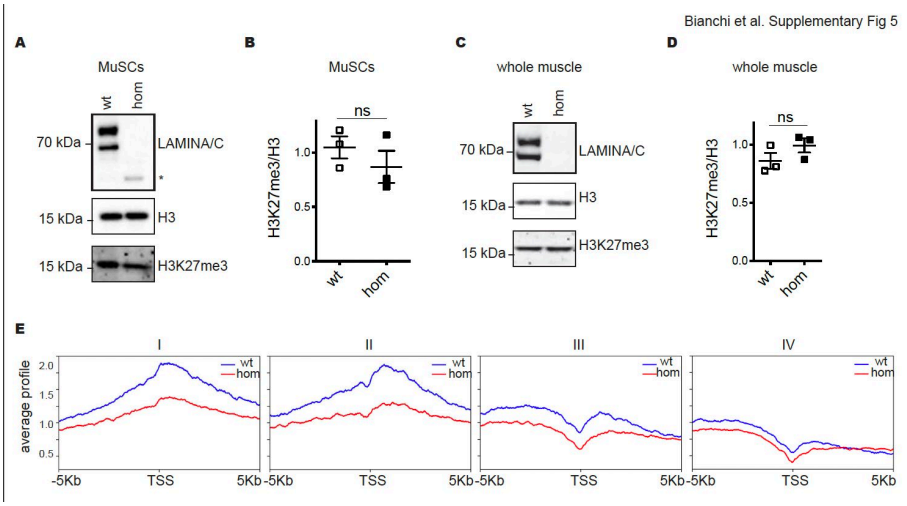
Lmna Δ8-11 +/+ and Lmna Δ8-11 -/- mice at d19. n=2-6 animals per genotype. Value were normalized on gapdh and compared with the average of wt amplification. (**E**) Western blot of total extracts from Lmna Δ8-11 MuSCs at d19 hybridized with indicated antibodies. The * indicates the truncated form of Lamin A detected in mutant mice (1) (**F**) Band density quantification of Ezh2 normalized on actin levels. n=3 animals per genotype. (**G**) Immunohistochemical staining in Lmna Δ8-11 mice of Ezh2 in Activated, ASCs (PAX7+/KI67+) and quiescent, QSCs (PAX7+/Ki67-) MuSCs at d19. Scale bars, 50 μm. (**H**) Quantification of Ezh2 fluorescence intensity in ASCs (PAX7+/KI67+) and QSCs (PAX7+/Ki67-) MuSCs in (G); n =4-5 animals per genotype (**I**) Magnification of the immunohistochemical staining described in (G). Scale bars, 10 μm. (**J**) Quantification of number of PcG bodies/nucleus (x-axis) in (I) by cumulative frequency distributions (y-axis). n =5 animals per genotype. **D**, **F**, **H**, Data are mean ± s.e.m. Comparisons by unpaired t-test. **B**, **C**, **J**, Comparisons by KolmogorovSmirnov test. ns: not statistical. * * P < 0.01, * * * P < 0.001. wt= Lmna Δ8-11 +/+; hom= Lmna Δ8-11 -/-.



Supplementary Figure 4. Transcriptional deregulation in Lmna Δ8-11 dystrophic mice.

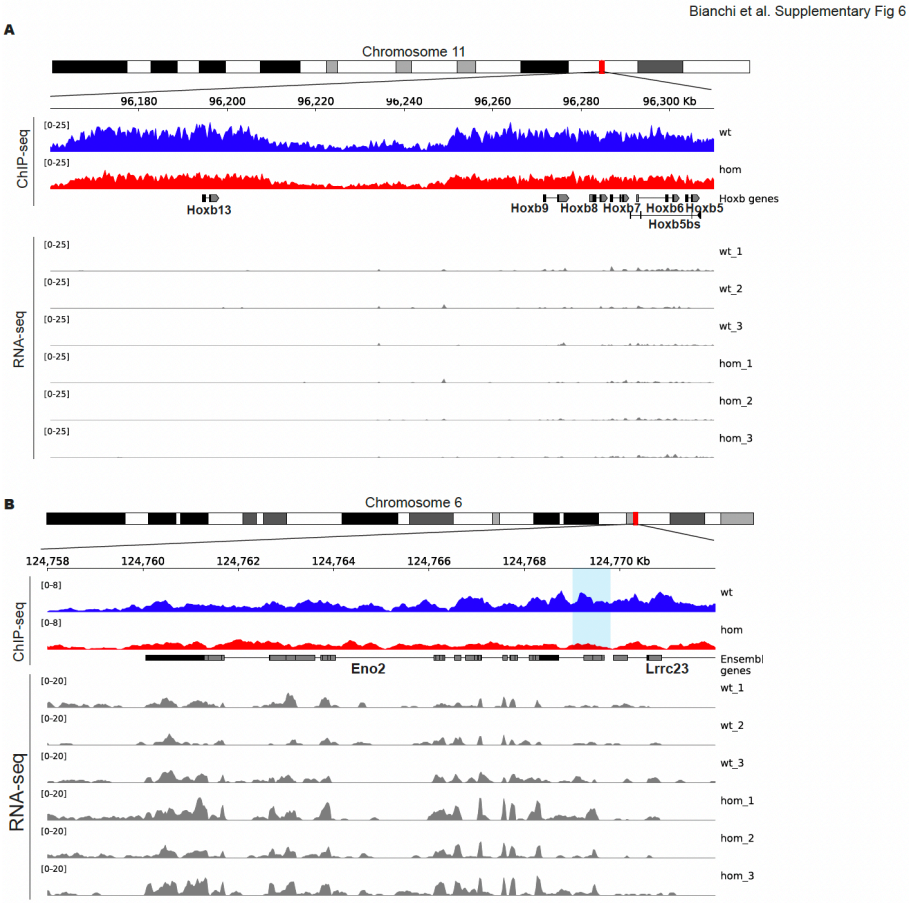
(**A**) Heatmap of differentially expressed genes Z-scores in Lmna Δ8-11 -/- vs Lmna Δ8-11 +/+ comparison (FDR < 0.05). (**B**) pre-ranked GSEA on log Fold Change from Ezh2 KO vs WT satellite cells (2). Up-regulated (logFC > 1) in hom vs wt comparison added to Biocarta mouse pathways from gskb R package were used as gene sets (NES = 2.31, FDR = 0.020). (**C**, **D**) heatmaps showing H3K27me3 spike-in normalized linear ratio IP/input signal along the gene body (C) or around the TSS (D) (coordinates along x-axis) of individual annotated mouse genes (y-axis). Genes are sorted by computing the average signal for each gene and then sorting them in descending order based

on the mean value. TES= annotated transcript end. wt= Lmna Δ8-11 +/+; hom= Lmna Δ8-11 -/-.



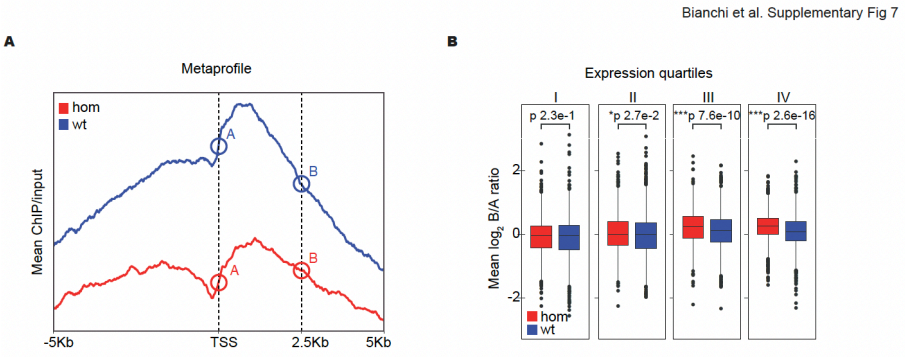
Supplementary Figure 5. H3K27me3 levels in Lmna Δ8-11 dystrophic mice.

(**A**) Western blot of total histones extracted from Lmna Δ8-11 MuSCs at d19 hybridized with indicated antibodies. The * indicates the truncated form of Lamin A detected in mutant mice (1). (**B**) Band density quantification of H3K27me3 normalized on H3 levels. n=3 animals per genotype. (**C**) Western blot of total histones extracted from Lmna Δ8-11 whole muscles at d19 hybridized with indicated antibodies. (**D**) Band density quantification of H3K27me3 normalized on H3 levels. n=3 animals per genotype. (**E**) Average profile of H3K27me3 signal as the spike-in normalized linear ratio IP/input at the TSS using genes grouped as in panel (Figure 3E). **B**, **D**, Data are mean ± s.e.m. Comparisons by unpaired t-test. ns: not statistical. wt= Lmna Δ8-11 +/+; hom= Lmna Δ8-11 -/-.



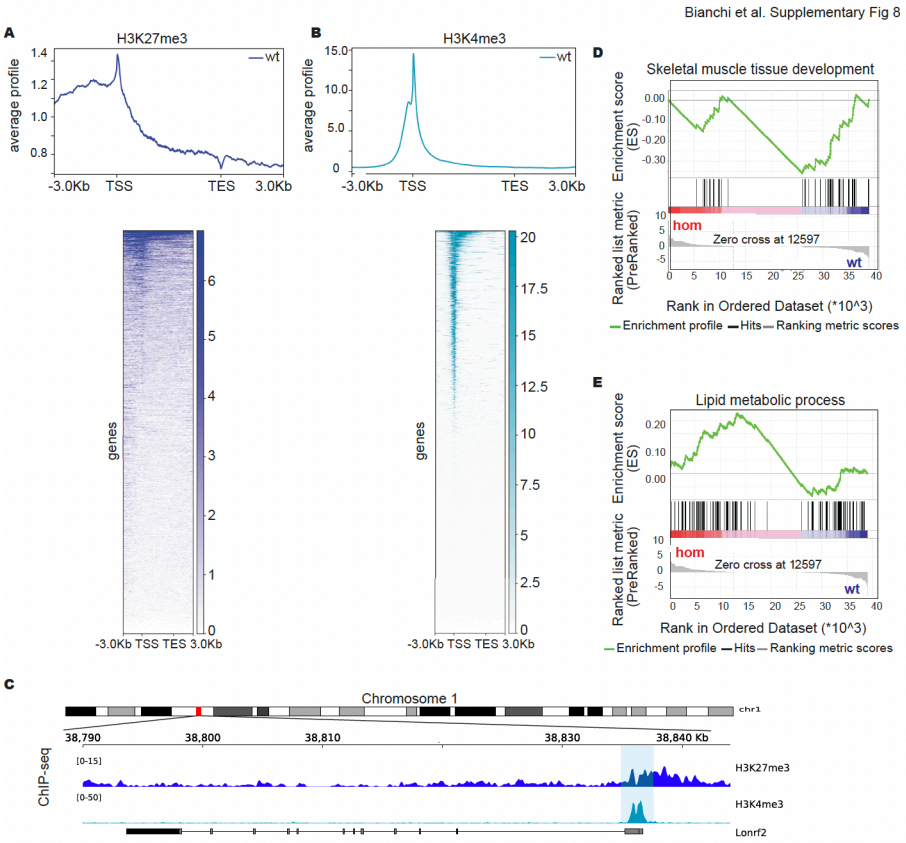
Supplementary Figure 6. Aberrant H3K27me3 deposition in Lmna Δ8-11 -/- MuSCs

(**A**) ChIP-seq of H3K27me3 mark as the spike-in normalized linear ratio IP/input and RNA-seq signal tracks as the effective genome size normalized coverage on Hoxb genes, belonging to the I/II quartiles of expression. (**B**) ChIP-seq of H3K27me3 mark as the spike-in normalized linear ratio IP/input and RNA-seq signal tracks as the effective genome size normalized coverage on Eno2 locus, belonging to the III quartile of expression. Eno2 promoter is highlighted by a light blue rectangle.



Supplementary Figure 7. Statistical analysis of H3K27me3 distribution at TSS

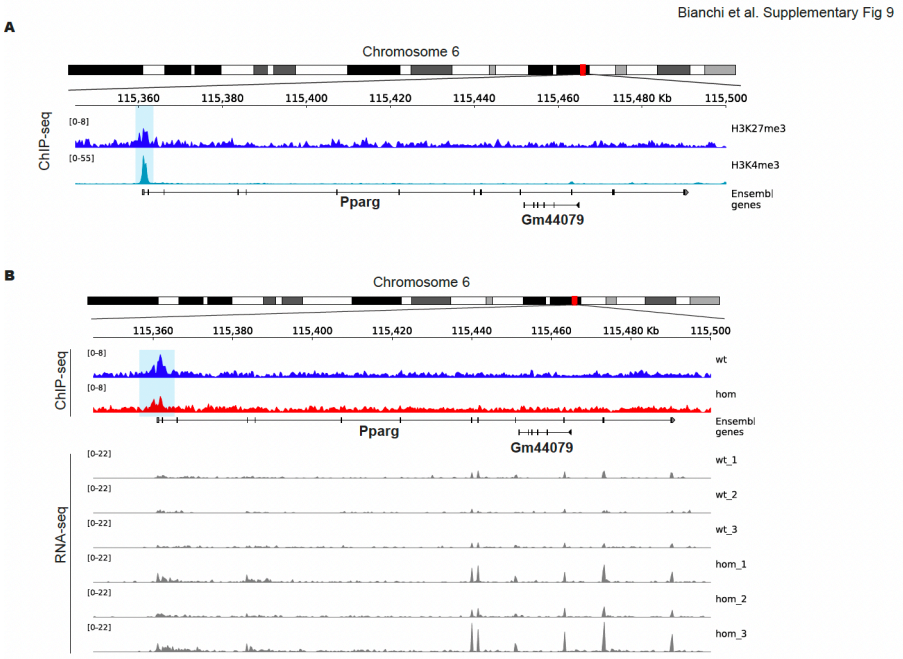
(**A**) Cartoon to illustrate the approach used to measure relative differences in H3K27me3 ChIP over input enrichment at TSS (A point) vs 2.5Kb downstream (B point). For each sample (wt or hom), for each individual gene, the ratio H3K27me3 ChIP over input enrichment at point A over point B is computed. (**B**) The boxplots report the distribution of A/B ChIP-seq enrichment ratios (defined as in panel C) in each sample, for each gene, grouped as defined in panel (Figure 3E). The p-values reported are based on non-parametric Wilcoxon test.



Supplementary Figure 8. Bivalent genes deregulation in Lmna Δ8-11 -/- MuSCs

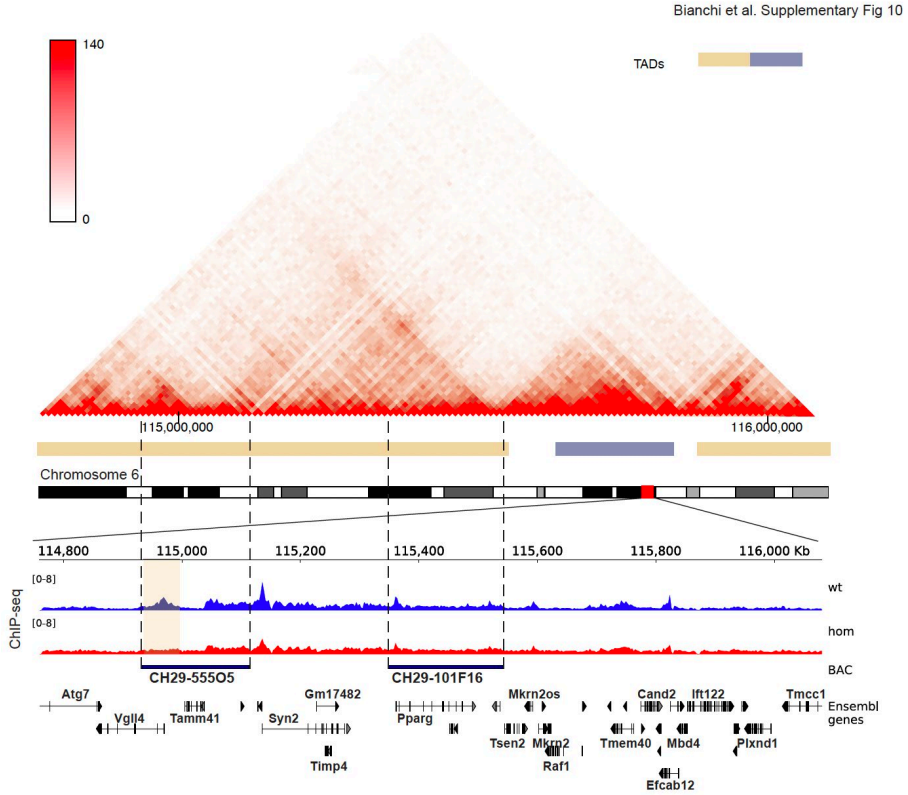
(**A, B**) Average profile and heatmap of H3K27me3 (A) and H3K4me3 (B) ChIP-seq signal along the gene body calculated as the ratio IP/input over annotated mouse genes. TES= annotated transcript end. (C) ChIP-seq signal tracks of H3K27me3 and H3K4me3 marks as the library size normalized linear ratio IP/input on Lonrf2 locus, showing a bivalent promoter. Lonrf2 promoter is highlighted by a light blue rectangle. (**D, E**) Results of pre-ranked GSEA using log fold change from RNA-seq differential analysis in hom vs wt comparison. Representative muscle development-related NES = -3.11, FDR = 1.86e-5 (**D**) and lipid metabolism-related NES = 2.91, FDR = 1.54e-4

(**E**) gene sets among Gene Ontology terms from gskb R package are shown.

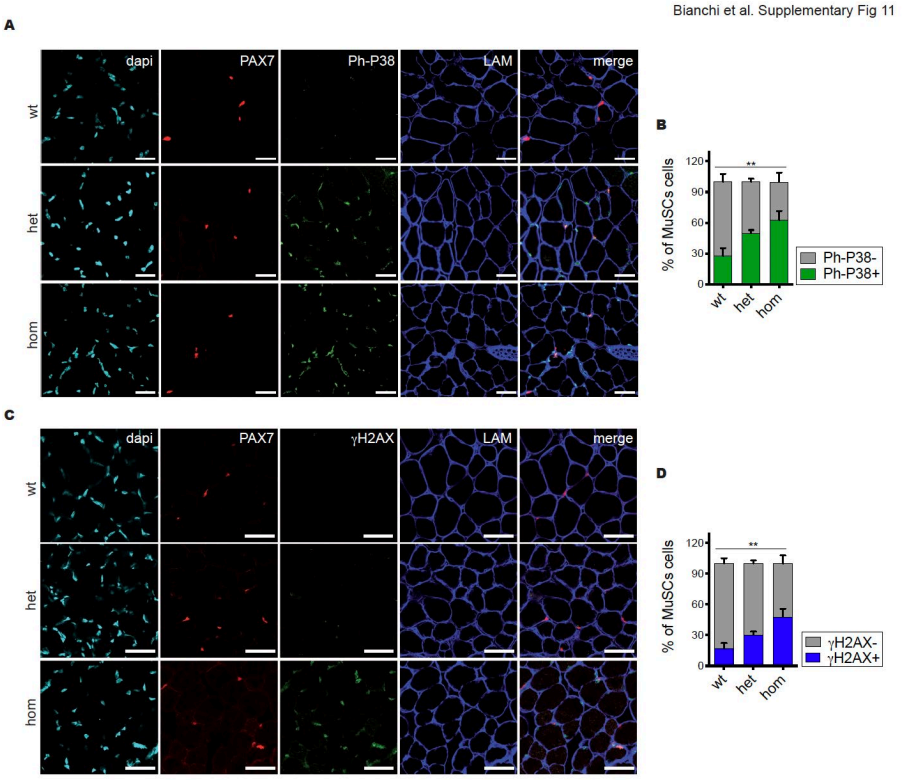


Supplementary Figure 9. PPARG is a bivalent gene

(**A**) ChIP-seq signal tracks of H3K27me3 and H3K4me3 marks as the library size normalize linear ratio ChIP/input on PPARG locus, showing the bivalent PPARG promoter, highlighted by a light blue rectangle. (**B**) ChIP-seq of H3K27me3 mark as the spike-in normalized linear ratio IP/input and RNA-seq signal tracks as the effective genome size normalized coverage on PPARG locus. PPARG promoter is highlighted by a light blue rectangle.

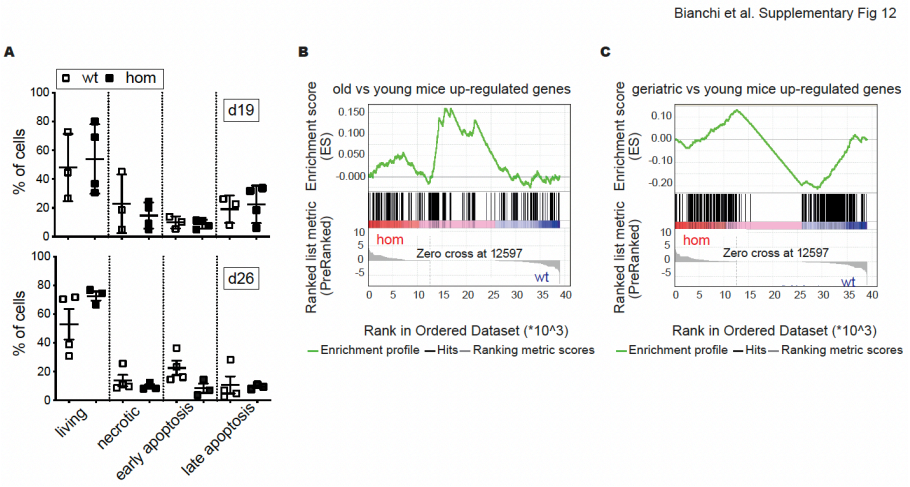


Supplementary Figure 10. Hi-C contact map at the PPARG locus. At the top, the Hi-C contact map of mouse embryonic stem cells ((3), resolution 10kb) showing the TADs compartmentalization (horizontal light yellow and blue bars) in the region chr6:114,760,000-116,080,000. At the bottom, H3K27me3 ChIP-seq spike-in normalized linear IP/input ratio of the same genomic region. In the middle of two dashed lines are highlighted as dark blue rectangles the mapping positions of BAC clones used for FISH experiments described in Figure 5E. H3K27me3 peak lost in the Lmna Δ8-11 -/- MuSCs is highlighted by light red rectangle.



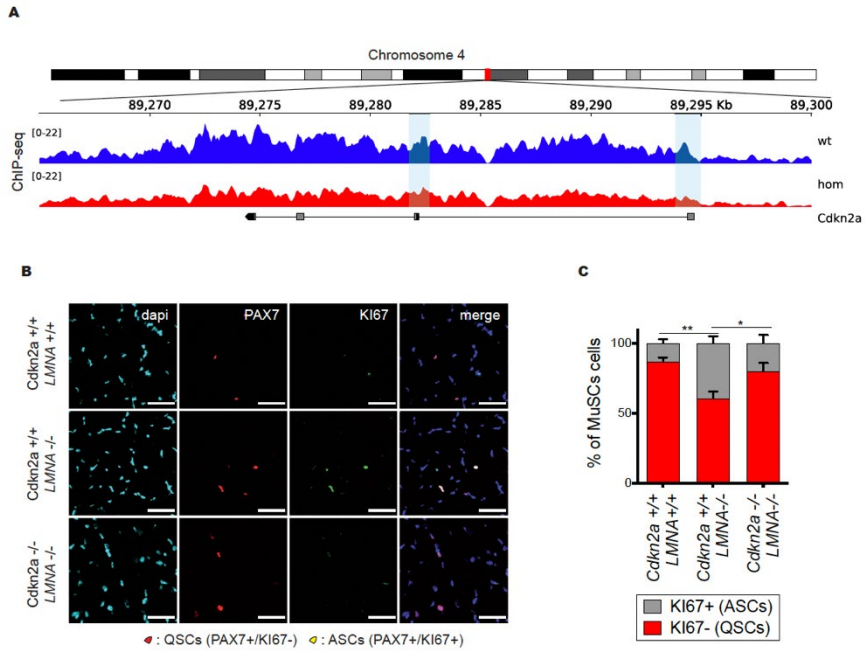
Supplementary Figure 11. Lmna Δ8-11 -/- MuSCs undergo premature senescence.

(**A**) Immunohistochemical staining in Lmna Δ8-11 mice of PAX7 and Ph-P38 markers at d19. Basement membrane of muscle fibers was stained with Laminin. Scale bars, 50 μm. (**B**) Quantification of MuSCs pool in (A); n= 5 animals per genotype. (C) Immunohistochemical staining in Lmna Δ8-11 mice of PAX7 and γH2AX markers at d19. Basement membrane of muscle fibers was stained with Laminin. Scale bars, 50 μm. (**D**) Quantification of MuSCs pool in (C); n= 3 animals per genotype. **B, D**, Data are mean ± s.e.m. Statistics by two-way analysis of variance (ANOVA) with multiple comparisons. * * P < 0.01. wt= Lmna Δ8-11 +/+; het= Lmna Δ8-11 +/-; hom= Lmna Δ8-11 -/



Supplementary Figure 12. Lmna Δ8-11 -/- MuSCs do not show apoptosis.

(**A**) Apoptosis assay performed on MuSCs of Lmna Δ8-11 +/+ and -/- at 19 and 26 days of post-natal growth evidenced living, apoptotic or necrotic cells. n=3-4 animals per genotype. (**B, C**) pre-ranked GSEA on log Fold Change from RNA-seq differential analysis in hom vs wt comparison. Up-regulated genes (q-value < 0.05) in old vs young (B) and geriatric vs young (C) mice comparisons from (4) were added to Biocarta mouse pathways from gskb R package and used as gene sets. (**B**) old vs. young mice up-regulated genes are enriched in hom vs wt up-regulated ones (NES = 2.86, FDR = 3.81e-04). (**C**) geriatric vs young mice up-regulated genes are enriched in hom vs wt down-regulated ones (NES = -5,09, FDR < 10e-4). **A**, Data are mean ± s.e.m. Statistics by Unpaired t test with Welch's correction. wt= Lmna Δ8-11 +/+; hom= Lmna Δ8-11 -/-.



Supplementary Figure 13. Cdkn2a absence restores MuSCs quiescence in Lmna Δ8-11 -/- mice.

(**A**) ChIP-seq signal tracks of H3K27me3 mark as the spike-in normalized linear ratio IP/input on Cdkn2a locus. Promoters are highlighted by a light blue rectangle. (**B**) Immunohistochemical staining in Cdkn2a/Lmna Δ8-11 mice of Activated, ASCs (PAX7+/KI67+) and quiescent, QSCs (PAX7+/Ki67-) of MuSCs at d19. Scale bars, 50 μm. (**C**) Quantification of MuSCs pool in (B). n=4-5 animals per genotype. Data are mean ± s.e.m. Statistics by two-way analysis of variance (ANOVA) with multiple comparisons. * P < 0.05, ** P < 0.01. wt= Lmna Δ8-11 +/+; hom= Lmna Δ8-11 -/-.

Additional methods

Bioinformatics analysis

RNA-seq: the overall quality of the sequenced reads was assessed using FastQC tool (version 0.11.2; http:/www.bioinformatics.babraham.ac.uk/projects/fastqc), then reads were trimmed and the adapters removed using Trimmomatic software (version 0.33;(5)). The remaining reads were aligned to mouse genome (version GRCm38; Ensembl release 87) providing the murine transcriptome (Ensembl release 87) with Spliced Transcripts Alignment to a Reference (STAR; version 2.5.2b;(6)). STAR alignments were used as input for HTSeq (version 0.6.1;(7)) to quantify the amount of reads per gene in the annotation with -s reverse and -t exon parameters. The differentially expressed genes (DEGs) between Lmna homozigous and wild type mice were identified using edgeR (version 3.16.0;(8) in R environment version 3.3.2); the significance threshold was set to False Discovery Rate (FDR) ≤ 0.05. Raw expression values normalization was computed calculating log2 count per million (cpm) with the same tool. Z-scores of hom vs. wt DEGs were calculated and visualized using gplots R package (version 3.0.1 in R environment version 3.3.2). A functional enrichment on DEGs between Lmna homozygous and wild type mice was conducted using the Gene Set Enrichment Analysis (GSEA) software (9) and the mouse Gene Ontology (GO) terms of the “Biological Processes” section retrieved from the R package “gskb” (version 1.6.1 in R environment 3.3.2; (10)). GSEA tool was used in pre-ranked mode as suggested for RNA-seq experiments using as input the log fold change (log FC) for all the genes in the annotation and the “classic” metric for gene ranking. The significance threshold was set to FDR ≤ 0.05. To compare our results with public data a GSEA analysis was performed on Liu et al., 2013

dataset (11), Sousa-Victor et al., 2014 dataset (4) and Juan et al., 2011 dataset (2). The first was retrieved from Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/; GSE47177), the second one from supplementary data, whereas the latter was obtained from the authors in form of raw data. On Liu et al., 2013: gene expression level of quiescent satellite cells from old and young mice was calculated starting from raw data using affy and annotate R packages (genelevel summarization using Brainarray Entrez gene custom cdf for mouse gene 1.0 st microarrays; http://brainarray.mbni.med.umich.edu/Brainarray/Database/CustomCDF/CDF_downl oad.asp) and was used as expression dataset, up-regulated genes in homozygous mice filtered for log FC ≥ 1 were added to Biocarta mouse pathways retrieved from gskb R package and used as gene sets database; moreover the old vs young class comparison was used as phenotype labels and Signal2Noise as metric for gene ranking as defined in (9). On Sousa-Victor et al., 2014: differential expressed genes from old vs. young mice and geriatric vs. young mice were filtered for q-value ≤ 0.05 and added to Biocarta mouse pathways retrieved from gskb R package and used as gene sets database and hom vs. wt log FC of all genes for a pre-ranked analysis with classic as metric for gene ranking. For Juan et al., 2011: raw data were analyzed using affy and annotate R packages (gene-level summarization using Brainarray Entrez gene custom cdf for mouse 430.2 microarrays; http://brainarray.mbni.med.umich.edu/Brainarray/Database/CustomDF/CDF_download.asp); log FC were calculated for all genes to perform a pre-ranked analysis with classic as metric for gene ranking; up-regulated genes in homozygous mice filtered for log FC ≥ 1 were added to Biocarta mouse pathways retrieved from gskb R package and used as gene sets database. To correlate H3K27me3 changes with gene expression, genes were divided into groups based on the

expression level in wt condition. In this case we quantified gene expression using kallisto (version 0.44.0; (12)). Genes with 0 coverage, i.e. 0 transcript per million (tpm) in all of the 3 replicates were classified as “zero-expressed” were excluded from the analysis. This group includes genes not expressed and not detected for technical reasons, i.e.either low abundance resulting in technical dropout, or low coverage due to low mappability of the sequence. Remaining genes were divided into 4 equally sized groups (i.e. “I”, “II”, “III” and “IV” categories) based on quartiles of the mean expression values in the 3 replicates. RNA-seq genome coverage was determined using bamCoverage function of deepTools suite (version 2.5.1;(13)) with 1 base bin size and effective genome size normalization (excluding X and Y chromosomes) and visualized using pyGenomeTracks (version 2.0;(13)). To refine the functional analysis on DEGs, an additional enrichment of GO terms was performed using the “FGNet” R package (version 3.8;(14)) followed by a semantic similarity analysis on significantly enriched terms given and FDR threshold of 0.05 taking advantage of the “GOSemSim” R package (version 2.0.4;(15)). Then the macro-categories grouping GO terms taking part of the same biological process were identified using Revigo web tool (16). Similarity scores heatmap was visualized using gplots R package (version 3.0.1 in R environment version 3.3.2). To test the enrichment of upregulated genes among those with a bivalent promoter and among H3K27me3 target genes a Fisher exact test was performed using fisher.test function of stats R package (R environment version 3.3.2; the p-value for significant threshold was set to 0.05) considering only mappable genes as reference background (i.e. genes with at least one reported alignment in one of the RNA-seq/ChIP-seq samples). Moreover, the same procedure was used to test the enrichment of Pparγ-related up-regulated genes among those with a bivalent promoter selecting up-

regulated genes in significantly enriched adipogenesis-related GO terms. The expected number of genes was determined by means of the product of marginal distributions and the log2 ratio of observed over expected number of genes was computed and visualized using gplots R package (version 3.0.1 in R environment version 3.3.2). ChIP-seq: FastQC (version 0.11.3; http:/www.bioinformatics.babraham.ac.uk/projects/fastqc) and Trimmomatic (version 0.32;(5)) were respectively used for the quality control of the sequenced reads and trimming/adapter removal steps. Reads were aligned to mouse genome (GRCm38; Ensembl release 87) using Bowtie (version 1.0.1;(17)) and retaining only those reads that map to unique loci. PCR duplicates were removed using samtools (version0.1.19;(18)). Peak calling step was performed with Model-based Analysis of ChIPSeq2 (MACS2; version 2.1.1;(19)) using input samples as control and –broad parameter for H3K27me3 ChIP-seq samples. Signal tracks were constructed calculating the linear fold enrichment (IP/input) with the same tool. Annotation of called peaks was performed with ChIPpeakAnno R package (version 3.8.9 in R environment version 3.3.2;(20)) in both mode (nearest start plus overlapping annotation) and calculating the distance between the middle point of peaks and the transcription start site (TSS) of genes. Genes with bivalent promoter were identified as those genes having an H3K27me3 peak that overlap for at least the 5% an H3K4me3 peak in a window between -0.5kb and +2.0kb around the TSS (similarly to what originally proposed by (21)) in wt condition (bedtools version 2.24.0;(22)). For ChIP-seq with spike-in approach reads were aligned also to Drosophila melanogaster genome (BDGP6; Ensembl release 87) and filtered as those mapping to mouse genome; then ambiguous reads aligning to both genomes were removed using BBMap tool (BBMap - Bushnell B. - sourceforge.net/projects/bbmap/). Reads

mapping to Dm6 genome were used to calculate spike-in normalization factors for IP samples as in (23) which were employed as scaling factors to normalize signal tracks(MACS2 version 2.1.1). H3K27me3 target genes were identified as those genes with a H3K27me3 peak in a window of +/- 5 kb around the TSS in the wt condition. We defined intergenic and interpeaks regions as those outside H3K27me3 peaks and outside annotated genes (bedtools version 2.24.0). Average IP/input ChIP-seq profiles on the indicated lists of genes were calculated and visualized using computeMatrix and plotProfile functions of deepTools suite (version 2.5.1; (13)). Heatmap of average IP/input ChIP-seq signal on indicated lists of genes were calculated and visualized using computeMatrix and plotHeatmap functions of deepTools suite (version 2.5.1; (13)). Both average profiles and average heatmaps were performed after excluding genes with length less than 2kb and genes overlapping each other without respect of the strand in the window of visualization. Signal tracks on specific loci were visualized using pyGenomeTracks tool (version 2.0;(13)). The HiC matrix, showing the TAD information in the region chr6:114,760,000-116,080,000 (UCSC coordinates style) has been downloaded form the “3D Genome Browser website” (http://promoter.bx.psu.edu/hi-c/). The query has been inserted in the “HiC” section with the following parameters: “Hi-C” as “Dataset Type” (step 1), “mouse” as Species and “mm10” as “assembly” (step 2), the checkbox “Browse Available Hi-C Data” has been flagged and “mESC” as “Tissue” (step 3), “Bonev_2017-raw” as “Type” and “10kb” as “Resolution” have been chosen, finally “chr6” as “Chromosome”, “114760000” as “Start” and “116080000” as “End” have been chosen in the box “Option 2: Search by Location” (step 4).

Image analysis

Image analysis derived from (24) with some adaptation. The algorithm performs the 2D segmentation of cell nuclei and the detection of PcG bodies for each slice of the stack, followed by the 3D reconstruction and identification of nuclei and PcG bodies. It then measures the volume of nuclei and the number and volume of the PcG bodies. The algorithm has been implemented in MATLAB following this scheme:

$[nuclei_n, avg_{Ezh2,n}]$ = nuclei_seg($I_{dapi,lamin,n}$,$I_{Ezh2,n}$) %Performs 2D nuclei segmentation of the middle section n of the stack

$I_{avg,n}$ = imfilter($I_{Ezh2,n}$,fspecial('average', [3,3]))

$I_{filt,n} = I_{Ezh2,n} - I_{avg,n}$

t_n = isodata_thresh($I_{filt,n}$,$avg_{Ezh2,n}$) %Applies ISODATA method to separate PcG bodies from nuclei regions in the middle section n of the stack

for each slice n of the stack

$nuclei_n$ = nuclei_seg($I_{dapi,lamin,n}$) %Performs 2D nuclei segmentation

$I_{avg,n}$ = imfilter($I_{Ezh2,n}$,fspecial('average', [3,3]))

$I_{filt,n} = I_{Ezh2,n} - I_{avg,n}$

$pcg_n = I_{filt,n} > t_n$ %Performs thresholding operation for 2D PcG bodies detection

$pcg\_vol(:,:,n) = pcg_n(:,:)$

$nuclei\_vol(:,:,n) = nuclei_n(:,:)$

endfor

nuclei_CC = bwconncomp(nuclei_vol)

nuclei_L = labelmatrix(nuclei_CC)

compute volume for each nucleus object in nuclei_CC

exclude nuclei whose volume is less than 10% of mean volumes

$\{NCL\}_M$ <- identified 3D nuclei

for each nucleus m in $\{NCL\}_M$
NCL_m.PcG = pcg_vol .* $\{NCL\}_M$ %3D positions of
detected PcG bodies within the nucleus NCL_m
NCL_m.PcG = bwareaopen(NCL_m.PcG,17,6)
PcG_CC = bwconncomp(NCL_m.PcG)
compute number of PcG bodies in PcG_CC
compute volume for each PcG body in PcG_CC
endfor

Idapi,lamin is the image obtained by the sum of both images showing the fluorescence of nucleus and lamin. The function nuclei_seg performs a partition of cell image Idapi,lamin in nuclei regions and background implementing a region based segmentation algorithm (25). avgEzh2 is the mean intensity value of the nuclei regions in the image IEzh2 that shows the fluorescence of PcG bodies. In order to better enhance PcG areas we subtract from the original IEzh2 image its smoothed version obtained by applying an averaging filter of size 3, obtaining the image Ifilt. The function isodata_thresh implements the ISODATA classification algorithm (26) and uses relevant values computed by nuclei_seg function in order to extract PcG bodies from the nuclei regions. It sets the initial threshold value of ISODATA method as avgEzh2. For each slice of the stack, the algorithm separates PcG bodies from nuclei regions by means of a thresholding operation using the threshold value estimated by the function isodata_thresh applied to the middle section. pcg_vol and nuclei_vol are 3D arrays that contain the positions of the detected PcG bodies and nuclei from all slices. 3D reconstructions of nuclei are obtained through the connected components algorithm (bwconncomp MATLAB function, using a connectivity of 26). 3D nuclei are then labeled by applying the labelmatrix MATLAB function so they can easily

separated each from the others. The algorithm computes the volume of each 3D reconstruction, discarding objects whose volume is less than 10% of mean volumes which are just noise. The algorithm uses the bwareaopen function In order to discard too small detected PcG objects which are probably just noise. 3D reconstructions of PcG bodies are obtained through the connected components algorithm (bwconncomp MATLAB function, using a connectivity of 6). The algorithm computes the number of the PcG bodies and the volume of any PcG body. For the 3D multicolor DNA FISH image analysis we used the tool NuCLεD described in (27), which is capable of automatically detecting and localizing fluorescent spots in 3D reconstructed nuclei. We adopted the following parameters: size 9 and standard deviation 7 for the Laplacian of Gaussian (LoG) operator and h=0.5 with a neighborhood size of 15x15 for the h-dome transformation. Fluorescence intensity measurements (Fig 6A and S3G) were executed with imageJ software. Asymmetric or symmetric distribution of pp38 protein in single myofibers was assigned by measuring the relative fluorescence intensity of the two MuSCs with asymmetry achieved when the ratio is higher than 2. For the apico-basal evaluation were considered asymmetric only cell doublets forming an angle between 50 and 90 with the fiber.

References to chapter 2

1. Shimi T, Kittisopikul M, Tran J, Goldman AE, Adam SA, Zheng Y, Jaqaman K, and Goldman RD. Structural organization of nuclear lamins A, C, B1, and B2 revealed by superresolution microscopy. *Mol Biol Cell.* 2015;26(22):4075-86.
2. Turgay Y, Eibauer M, Goldman AE, Shimi T, Khayat M, Ben-Harush K, Dubrovsky-Gaupp A, Sapra KT, Goldman RD, and Medalia O. The molecular architecture of lamins in somatic cells. *Nature.* 2017;543(7644):261-4.
3. van Steensel B, and Belmont AS. Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. *Cell.* 2017;169(5):780-91.
4. Zaremba-Czogalla M, Dubinska-Magiera M, and Rzepecki R. Laminopathies: the molecular background of the disease and the prospects for its treatment. *Cell Mol Biol Lett.* 2011;16(1):114-48.
5. Kolb T, Maass K, Hergt M, Aebi U, and Herrmann H. Lamin A and lamin C form homodimers and coexist in higher complex forms both in the nucleoplasmic fraction and in the lamina of cultured human cells. *Nucleus.* 2011;2(5):425-33.
6. Serebryannyy L, and Misteli T. Protein sequestration at the nuclear periphery as a potential regulatory mechanism in premature aging. *J Cell Biol.* 2018;217(1):21-37.

7. Dubinska-Magiera M, Zaremba-Czogalla M, and Rzepecki R. Muscle development, regeneration and laminopathies: how lamins or lamina-associated proteins can contribute to muscle development, regeneration and disease. *Cell Mol Life Sci.* 2013;70(15):2713-41.
8. Schuettengruber B, Bourbon HM, Di Croce L, and Cavalli G. Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cell.* 2017;171(1):34-57.
9. Briand N, and Collas P. Laminopathy-causing lamin A mutations reconfigure lamina-associated domains and local spatial chromatin conformation. *Nucleus.* 2018;9(1):216-26.
10. Cesarini E, Mozzetta C, Marullo F, Gregoretti F, Gargiulo A, Columbaro M, Cortesi A, Antonelli L, Di Pelino S, Squarzoni S, et al. Lamin A/C sustains PcG protein architecture, maintaining transcriptional repression at target genes. *J Cell Biol.* 2015;211(3):533-51.
11. Marullo F, Cesarini E, Antonelli L, Gregoretti F, Oliva G, and Lanzuolo C. Nucleoplasmic Lamin A/C and Polycomb group of proteins: an evolutionarily conserved interplay. *Nucleus.* 2016:0.
12. Oldenburg A, Briand N, Sorensen AL, Cahyani I, Shah A, Moskaug JO, and Collas P. A lipodystrophy-causing lamin A mutant alters conformation and epigenetic regulation of the anti-adipogenic MIR335 locus. *J Cell Biol.* 2017;216(9):2731-43.

13. Salvarani N, Crasto S, Miragoli M, Bertero A, Paulis M, Kunderfranco P, Serio S, Forni A, Lucarelli C, Dal Ferro M, et al. The K219T-Lamin mutation induces conduction defects through epigenetic inhibition of SCN5A in human cardiac laminopathy. *Nat Commun.* 2019;10(1):2267.
14. Zheng X, Hu J, Yue S, Kristiani L, Kim M, Sauria M, Taylor J, Kim Y, and Zheng Y. Lamins Organize the Global Three-Dimensional Genome from the Nuclear Periphery. *Mol Cell.* 2018;71(5):802-15 e7.
15. Chiacchiera F, Rossi A, Jammula S, Piunti A, Scelfo A, Ordonez-Moran P, Huelsken J, Koseki H, and Pasini D. Polycomb Complex PRC1 Preserves Intestinal Stem Cell Identity by Sustaining Wnt/beta-Catenin Transcriptional Activity. *Cell Stem Cell.* 2016;18(1):91-103.
16. Lapthanasupkul P, Feng J, Mantesso A, Takada-Horisawa Y, Vidal M, Koseki H, Wang L, An Z, Miletich I, and Sharpe PT. Ring1a/b polycomb proteins regulate the mesenchymal stem cell niche in continuously growing incisors. *Dev Biol.* 2012;367(2):140-53.
17. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell.* 2006;125(2):315-26.
18. Morata G, and Herrera SC. Cell reprogramming during regeneration in Drosophila: transgression of compartment boundaries. *Curr Opin Genet Dev.* 2016;40(11-6.

19. Lu TT, Heyne S, Dror E, Casas E, Leonhardt L, Boenke T, Yang CH, Sagar, Arrigoni L, Dalgaard K, et al. The Polycomb-Dependent Epigenome Controls beta Cell Dysfunction, Dedifferentiation, and Diabetes. *Cell Metab.* 2018;27(6):1294-308 e7.
20. Cohen TV, Gnocchi VF, Cohen JE, Phadke A, Liu H, Ellis JA, Foisner R, Stewart CL, Zammit PS, and Partridge TA. Defective skeletal muscle growth in lamin A/C-deficient mice is rescued by loss of Lap2alpha. *Hum Mol Genet.* 2013;22(14):2852-69.
21. Sullivan T, Escalante-Alcalde D, Bhatt H, Anver M, Bhat N, Nagashima K, Stewart CL, and Burke B. Loss of A-type lamin expression compromises nuclear envelope integrity leading to muscular dystrophy. *J Cell Biol.* 1999;147(5):913-20.
22. Gregoretti F, Cesarini E, Lanzuolo C, Oliva G, and Antonelli L. An Automatic Segmentation Method Combining an Active Contour Model and a Classification Technique for Detecting Polycomb-group Proteinsin High-Throughput Microscopy Images. *Methods Mol Biol.* 2016;1480(181-97.
23. Boonsanay V, Zhang T, Georgieva A, Kostin S, Qi H, Yuan X, Zhou Y, and Braun T. Regulation of Skeletal Muscle Stem Cell Quiescence by Suv4-20h1-Dependent Facultative Heterochromatin Formation. *Cell Stem Cell.* 2016;18(2):229-42.

24. Juan AH, Derfoul A, Feng X, Ryall JG, Dell'Orso S, Pasut A, Zare H, Simone JM, Rudnicki MA, and Sartorelli V. Polycomb EZH2 controls self-renewal and safeguards the transcriptional identity of skeletal muscle stem cells. *Genes Dev.* 2011;25(8):789-94.
25. Liu L, Cheung TH, Charville GW, Hurgo BM, Leavitt T, Shih J, Brunet A, and Rando TA. Chromatin modifications as determinants of muscle stem cell quiescence and chronological aging. *Cell Rep.* 2013;4(1):189-204.
26. Orlando DA, Chen MW, Brown VE, Solanki S, Choi YJ, Olson ER, Fritz CC, Bradner JE, and Guenther MG. Quantitative ChIP-Seq normalization reveals global modulation of the epigenome. *Cell Rep.* 2014;9(3):1163-70.
27. Minoux M, Holwerda S, Vitobello A, Kitazawa T, Kohler H, Stadler MB, and Rijli FM. Gene bivalency at Polycomb domains regulates cranial neural crest positional identity. *Science.* 2017;355(6332).
28. Bernstein E, Duncan EM, Masui O, Gil J, Heard E, and Allis CD. Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. *Mol Cell Biol.* 2006;26(7):2560-9.
29. Barak Y, Nelson MC, Ong ES, Jones YZ, Ruiz-Lozano P, Chien KR, Koder A, and Evans RM. PPAR gamma is required for placental, cardiac, and adipose tissue development. *Mol Cell.* 1999;4(4):585-95.

30. Rosen ED, Sarraf P, Troy AE, Bradwin G, Moore K, Milstone DS, Spiegelman BM, and Mortensen RM. PPAR gamma is required for the differentiation of adipose tissue in vivo and in vitro. *Mol Cell.* 1999;4(4):611-7.
31. Tong J, Li W, Vidal C, Yeo LS, Fatkin D, and Duque G. Lamin A/C deficiency is associated with fat infiltration of muscle and bone. *Mech Ageing Dev.* 2011;132(11-12):552-9.
32. Bantignies F, Roure V, Comet I, Leblanc B, Schuettengruber B, Bonnet J, Tixier V, Mas A, and Cavalli G. Polycomb-Dependent Regulatory Contacts between Distant Hox Loci in Drosophila. *Cell.* 2011;144(2):214-26.
33. Lanzuolo C, Roure V, Dekker J, Bantignies F, and Orlando V. Polycomb response elements mediate the formation of chromosome higher-order structures in the bithorax complex. *Nat Cell Biol.* 2007;9(10):1167-74.
34. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, and Ren B. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature.* 2012;485(7398):376-80.
35. Bonev B, Mendelson Cohen N, Szabo Q, Fritsch L, Papadopoulos GL, Lubling Y, Xu X, Lv X, Hugnot JP, Tanay A, et al. Multiscale 3D Genome Rewiring during Mouse Neural Development. *Cell.* 2017;171(3):557-72 e24.
36. Wang Y, Song F, Zhang B, Zhang L, Xu J, Kuang D, Li D, Choudhary MNK, Li Y, Hu M, et al. The 3D Genome

Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions. *Genome Biol.* 2018;19(1):151.

37. Chakkalakal JV, Jones KM, Basson MA, and Brack AS. The aged niche disrupts muscle stem cell quiescence. *Nature.* 2012;490(7420):355-60.
38. Gonzalo S, Kreienkamp R, and Askjaer P. Hutchinson-Gilford Progeria Syndrome: A premature aging disease caused by LMNA gene mutations. *Ageing Res Rev.* 2017;33(18-29.
39. Jacobs JJ, Kieboom K, Marino S, DePinho RA, and van Lohuizen M. The oncogene and Polycomb-group gene bmi-1 regulates cell proliferation and senescence through the ink4a locus. *Nature.* 1999;397(6715):164-8.
40. Chang NC, Chevalier FP, and Rudnicki MA. Satellite Cells in Muscular Dystrophy - Lost in Polarity. *Trends Mol Med.* 2016;22(6):479-96.
41. Bernet JD, Doles JD, Hall JK, Kelly Tanaka K, Carter TA, and Olwin BB. p38 MAPK signaling underlies a cell-autonomous loss of stem cell self-renewal in skeletal muscle of aged mice. *Nat Med.* 2014;20(3):265-71.
42. Cosgrove BD, Gilbert PM, Porpiglia E, Mourkioti F, Lee SP, Corbel SY, Llewellyn ME, Delp SL, and Blau HM. Rejuvenation of the muscle stem cell population restores strength to injured aged muscles. *Nat Med.* 2014;20(3):255-64.

43. Kuang S, Kuroda K, Le Grand F, and Rudnicki MA. Asymmetric self-renewal and commitment of satellite stem cells in muscle. *Cell.* 2007;129(5):999-1010.
44. Sousa-Victor P, Gutarra S, Garcia-Prat L, Rodriguez-Ubreva J, Ortet L, Ruiz-Bonilla V, Jardi M, Ballestar E, Gonzalez S, Serrano AL, et al. Geriatric muscle stem cells switch reversible quiescence into senescence. *Nature.* 2014;506(7488):316-21.
45. Beguelin W, Rivas MA, Calvo Fernandez MT, Teater M, Purwada A, Redmond D, Shen H, Challman MF, Elemento O, Singh A, et al. EZH2 enables germinal centre formation through epigenetic silencing of CDKN1A and an Rb-E2F1 feedback loop. *Nat Commun.* 2017;8(1):877.
46. Yosef R, Pilpel N, Papismadov N, Gal H, Ovadya Y, Vadai E, Miller S, Porat Z, Ben-Dor S, and Krizhanovsky V. p21 maintains senescent cell viability under persistent DNA damage response by restraining JNK and caspase signaling. *EMBO J.* 2017;36(15):2280-95.
47. Kandert S, Wehnert M, Muller CR, Buendia B, and Dabauvalle MC. Impaired nuclear functions lead to increased senescence and inefficient differentiation in human myoblasts with a dominant p.R545C mutation in the LMNA gene. *Eur J Cell Biol.* 2009;88(10):593-608.
48. Quelle DE, Ashmun RA, Hannon GJ, Rehberger PA, Trono D, Richter KH, Walker C, Beach D, Sherr CJ, and Serrano M. Cloning and characterization of murine

p16INK4a and p15INK4b genes. *Oncogene.* 1995;11(4):635-45.

49. Ito T, Teo YV, Evans SA, Neretti N, and Sedivy JM. Regulation of Cellular Senescence by Polycomb Chromatin Modifiers through Distinct DNA Damage- and Histone Methylation-Dependent Pathways. *Cell Rep.* 2018;22(13):3480-92.
50. Serrano M, Lee H, Chin L, Cordon-Cardo C, Beach D, and DePinho RA. Role of the INK4a locus in tumor suppression and cell mortality. *Cell.* 1996;85(1):27-37.
51. Gnocchi VF, Ellis JA, and Zammit PS. Does satellite cell dysfunction contribute to disease progression in Emery-Dreifuss muscular dystrophy? *Biochem Soc Trans.* 2008;36(Pt 6):1344-9.
52. Robinson DCL, and Dilworth FJ. Epigenetic Regulation of Adult Myogenesis. *Curr Top Dev Biol.* 2018;126(235-84.
53. Tosic M, Allen A, Willmann D, Lepper C, Kim J, Duteil D, and Schule R. Lsd1 regulates skeletal muscle regeneration and directs the fate of satellite cells. *Nat Commun.* 2018;9(1):366.
54. Frock RL, Kudlow BA, Evans AM, Jameson SA, Hauschka SD, and Kennedy BK. Lamin A/C and emerin are critical for skeletal muscle satellite cell differentiation. *Genes Dev.* 2006;20(4):486-500.
55. Solovei I, Wang AS, Thanisch K, Schmidt CS, Krebs S, Zwerger M, Cohen TV, Devys D, Foisner R, Peichl L, et al. LBR and lamin A/C sequentially tether peripheral

heterochromatin and inversely regulate differentiation. *Cell.* 2013;152(3):584-98.

56. Melcon G, Kozlov S, Cutler DA, Sullivan T, Hernandez L, Zhao P, Mitchell S, Nader G, Bakay M, Rottman JN, et al. Loss of emerin at the nuclear envelope disrupts the Rb1/E2F and MyoD pathways during muscle regeneration. *Hum Mol Genet.* 2006;15(4):637-51.
57. Kundu S, Ji F, Sunwoo H, Jain G, Lee JT, Sadreyev RI, Dekker J, and Kingston RE. Polycomb Repressive Complex 1 Generates Discrete Compacted Domains that Change during Differentiation. *Mol Cell.* 2017;65(3):432-46 e5.
58. Sousa-Victor P, Perdiguero E, and Munoz-Canoves P. Geroconversion of aged muscle stem cells under regenerative pressure. *Cell Cycle.* 2014;13(20):3183-90.

Additional references

1. Jahn D, Schramm S, Schnolzer M, Heilmann CJ, de Koster CG, Schutz W,Benavente R, and Alsheimer M. A truncated lamin A in the Lmna -/-mouse line: implications for the understanding of laminopathies. Nucleus.2012;3(5):463-74.
2. Juan AH, Derfoul A, Feng X, Ryall JG, Dell'Orso S, Pasut A, Zare H, SimoneJM, Rudnicki MA, and Sartorelli V. Polycomb EZH2 controls self-renewal and safeguards the transcriptional identity of skeletal muscle stem cells.Genes Dev. 2011;25(8):789-94.

3. Bonev B, Mendelson Cohen N, Szabo Q, Fritsch L, Papadopoulos GL,Lubling Y, Xu X, Lv X, Hugnot JP, Tanay A, et al. Multiscale 3D Genome Rewiring during Mouse Neural Development. Cell. 2017;171(3):557-72e24.
4. Sousa-Victor P, Gutarra S, Garcia-Prat L, Rodriguez-Ubreva J, Ortet L, Ruiz-Bonilla V, Jardi M, Ballestar E, Gonzalez S, Serrano AL, et al. Geriatric muscle stem cells switch reversible quiescence into senescence. Nature. 2014;506(7488):316-21.
5. Bolger AM, Lohse M, and Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114-20.
6. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15-21.
7. Anders S, Pyl PT, and Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31(2):166-9.
8. Robinson MD, McCarthy DJ, and Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139-40.
9. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al. Gene set enrichment analysis: a

knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545-50.

10. Lai Liming HJ, Bares Valerie, Son Eun Woo, Ban Yuguang, Wang Wei, Qi Jianli, Jiang Gaixin, Liberzon Arthur, Ge Steven. GSKB: A gene set database for pathway analysis in mouse bioRxiv. 2016.
11. Liu L, Cheung TH, Charville GW, Hurgo BM, Leavitt T, Shih J, Brunet A, and Rando TA. Chromatin modifications as determinants of muscle stem cell quiescence and chronological aging. Cell Rep. 2013;4(1):189-204.
12. Bray NL, Pimentel H, Melsted P, and Pachter L. Near-optimal probabilisticRNA-seq quantification. Nat Biotechnol. 2016;34(5):525-7.
13. Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dundar F, and Manke T. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 2016;44(W1):W160-5.
14. Aibar S, Fontanillo C, Droste C, and De Las Rivas J. Functional Gene Networks: R/Bioc package to generate and analyse gene networks derived from functional enrichment and clustering. Bioinformatics. 2015;31(10):1686-8.
15. Yu G, Li F, Qin Y, Bo X, Wu Y, and Wang S. GOSemSim: an R package for measuring semantic similarity among

GO terms and gene products. Bioinformatics. 2010;26(7):976-8.

16. Supek F, Bosnjak M, Skunca N, and Smuc T. REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS One. 2011;6(7):e21800.
17. Langmead B, Trapnell C, Pop M, and Salzberg SL. Ultrafast and memoryefficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10(3):R25.
18. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, and Genome Project Data Processing S. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-9.
19. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008;9(9):R137.
20. Zhu LJ, Gazin C, Lawson ND, Pages H, Lin SM, Lapointe DS, and Green MR. ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIPchip data. BMC Bioinformatics. 2010;11(2)37.
21. Ku M, Koche RP, Rheinbay E, Mendenhall EM, Endoh M, Mikkelsen TS, Presser A, Nusbaum C, Xie X, Chi AS, et al. Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. PLoS Genet. 2008;4(10):e1000242.

22. Quinlan AR, and Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26(6):841-2.
23. Orlando DA, Chen MW, Brown VE, Solanki S, Choi YJ, Olson ER, Fritz CC, Bradner JE, and Guenther MG. Quantitative ChIP-Seq normalization reveals global modulation of the epigenome. Cell Rep. 2014;9(3):1163-70.
24. Gregoretti F, Cesarini E, Lanzuolo C, Oliva G, and Antonelli L. An Automatic Segmentation Method Combining an Active Contour Model and a Classification Technique for Detecting Polycomb-group Proteinsin High-Throughput Microscopy Images. Methods Mol Biol. 2016;1480(1)81-97.
25. Goldstein T, Bresson X, and Osher S. Geometric Applications of the Split Bregman Method: Segmentation and Surface Reconstruction. Journal of Scientific Computing. 2010;45(5):272-93.
26. Ball G, and Hall D. Menlo Park: Stanford Research Institute; 1965.
27. Cortesi A, Pesant M, Sinha S, Marasca F, Sala E, Gregoretti F, Antonelli L, Oliva G, Chiereghin C, Solda G, et al. 4q-D4Z4 chromatin architecture regulates the transcription of muscle atrophic genes in facioscapulohumeral muscular dystrophy. Genome Res. 2019;29(6):883-95.

CHAPTER 3.

CHROMATIN SOLUBILITY AS A NOVEL DETERMINANT OF EPIGENOME DYSFUNCTION IN PROSTATE CANCER

Abstract

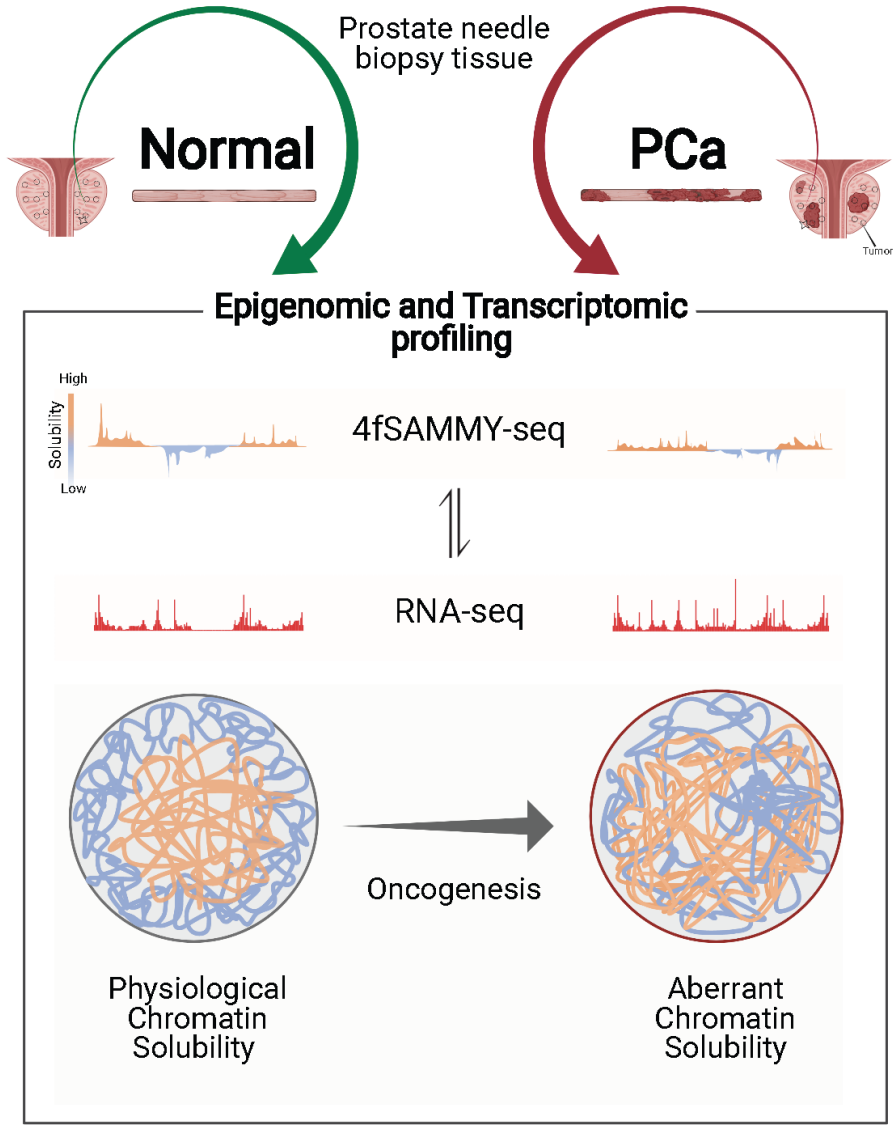
Prostate cancer (PCa) is the second most recurrent male tumor, often characterized by an unfavorable outcome due to its high clinical and molecular heterogeneity and its frequent multifocality. Increasingly growing epigenetic research efforts are implicated in the detection of novel biomarkers improving PCa diagnosis and prognosis. Among the epigenetic layers, the chromatin nuclear architecture abides tight rules that ensure proper genome functions and cell identity maintenance, and its remodeling has emerged as a novel crucial player during malignant transformation. Yet, the occurrence and the functional implications of chromatin structural changes in cancer remains poorly characterized. We have recently developed a high-throughput sequencing based technique, named 4fSAMMY-seq (4 fractions Sequential Analysis of MacroMolecules accessibilitY), consisting of the sequential fractionation of the chromatin according to its accessibility and solubility status to isolate and characterize both heterochromatin and euchromatin. A key advantage of our method is that it only requires few thousands of living cells, making it suitable for analyses on limited materials as biopsy samples derived from clinical practice. In this study, we applied 4fSAMMY-seq on fresh biopsy specimens from patients undergoing explorative prostate mapping biopsy for prostate cancer diagnosis.

We examined the genome conformation and its association with transcriptional profile in control tumor-free biopsies and prostate

cancer biopsies. In bulk healthy tissues, containing a mix of epithelial cells, leukocytes and stroma, 4fSAMMY-seq allowed the detection of a conserved pattern of euchromatin and heterochromatin regions associated with the respective epigenetic signatures. We also found a quantitative correlation between chromatin solubility and gene expression.

On the other hand, cancer biopsies are characterized by different degrees of genome architectural and transcriptome alterations, reflecting the heterogeneous nature of prostate tumors. On the basis of chromatin solubility, we identified a group of PCa tissues characterized by extensive chromatin remodeling and transcriptomic reprogramming involved in tumor migratory capacity. Our data highlight the impact of chromatin architecture on aberrant genome usage and propose 4fSAMMY-seq as an eligible tool to shed light on the epigenetic remodulations driving primary prostate tumor aggressiveness.

Graphical Abstract



3.1 4 fractions Sequential Analysis of MacroMolecules accessibilitY (4fSAMMY-seq) allows mapping of heterochromatic and euchromatin regions in human fibroblasts

(Manuscript in preparation)

Our recently published SAMMY-seq technology (patent n. 18200482.0) allows genome-wide profiling of multiple chromatin fractions characterized by different solubility and accessibility properties[1]. SAMMY-seq is based on the sequential isolation of distinct chromatin fractions containing DNase-sensitive chromatin (S2 fraction), salt-sensible chromatin (S3 fraction) and salt-resistant chromatin (S4 fraction). Differentially enriched genomic regions are identified by pairwise comparisons between fractions after reads alignment, filtering and quality check. The correlation with public epigenomic data-sets deriving from other chromatin-related technologies on the same cell type highlights a different distribution of genomic regions based on the chromatin states between the different SAMMY-seq fractions. This first publicly available version of the protocol, 3 fractions SAMMY-seq (3fSAMMY-seq), is capable of mapping the heterochromatic regions tightly associated to insoluble Nuclear Lamina (NL) components by using pairwise comparisons between less accessible and more accessible fractions (S3vsS2, S4vsS3 and S4vsS2)[1].

Building on this, we recently developed an improved protocol, called 4 fractions SAMMY-seq (4fSAMMY-seq), schematically

represented in Figure 1A (Manuscript in preparation), capable to capture the whole spectrum of chromatin solubility and accessibility.

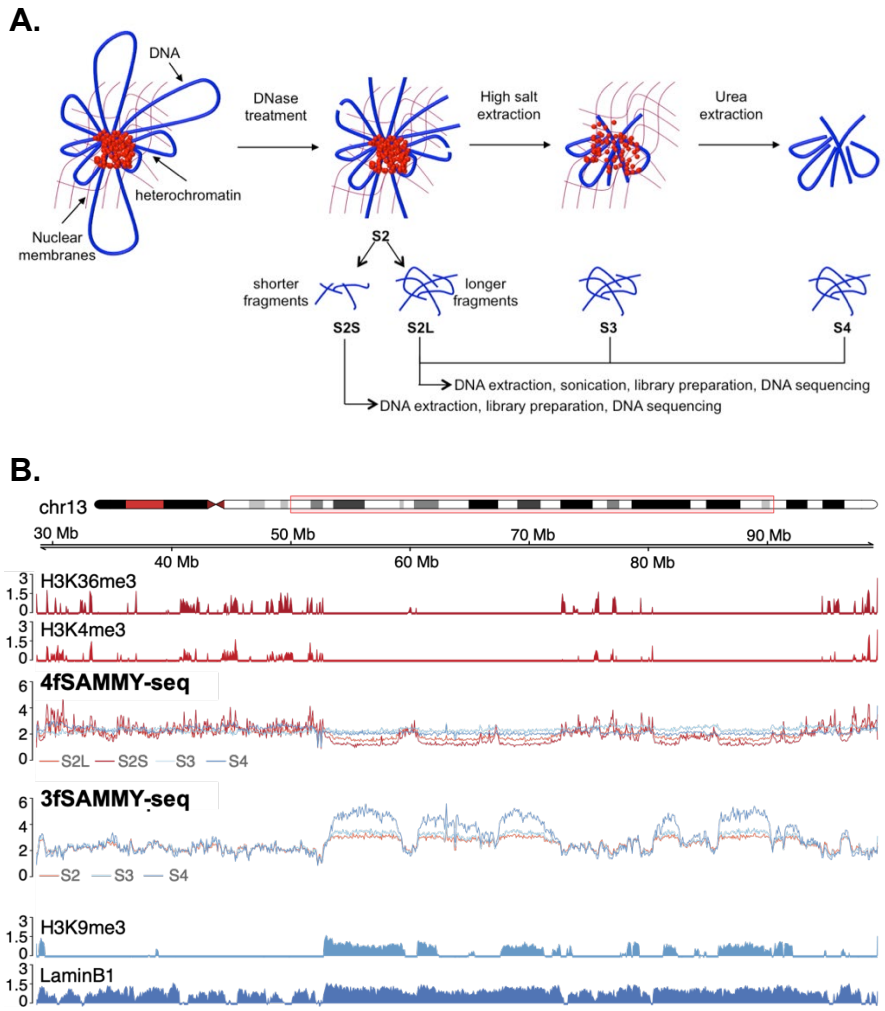
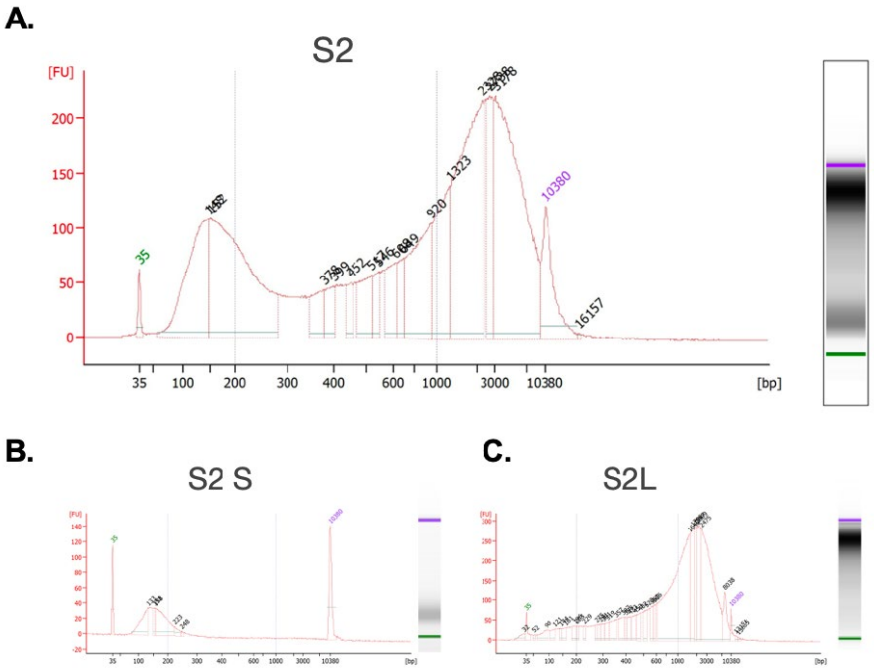


Figure 1 – 4fSAMMY-seq isolates specific genomic regions in human fibroblasts.
A. Flowchart of 4fSAMMY-seq procedure. Chromatin fractions sequentially isolated after DNase treatment (S2), high salt (S3) and urea buffers (S4) extractions, correspond to the most soluble parts of the chromatin, the DNase-resistant chromatin and DNA associated with nuclear lamins, respectively. The S2 fraction is

further divided in two sub-fractions based on the size distribution of DNA fragments, generating S2 small fraction (S2S) and S2 large fraction (S2L). The associated genomic DNA from all fractions is extracted and processed for high-throughput sequencing. Fractions S2L, S3 and S4 are sonicated. **B.** Distribution of 4fSAMMY-seq and 3fSAMMY-seq reads in each sequenced fraction of healthy skin primary fibroblast sample (CTRL004) along a representative genomic region (chr13:28,643,272-99,222,399). Library size normalized read counts over 10 kb genomic bins are shown for each sequenced fraction. From top to bottom: ChIP-seq tracks for histone marks (HMs) associated to active chromatin (H3K36me3 and H3K4me3 - red); 4fSAMMY-seq linear tracks (S2L - orange, S2S - dark red, S3 - light blue, S4 – blue); 3fSAMMY-seq linear tracks (S2 - red, S3 - light blue, S4 – blue); ChIP-seq tracks of heterochromatin mark (H3K9me3 –blue and LaminB1- dark blue).

In the novel version of the SAMMY-seq technology, set up on primary healthy skin fibroblasts, we decided to modulate the DNase-I digestion step by opting for a less-active enzyme: after digestion we observed the existence of two distinct S2 populations composed of genomic fragments with different length (Suppl. Fig. 1A). Hence, we used SPRI affinity beads to physically separate S2-associated DNA fragments according to their size (see methods), recovering two subfractions: the S2 small fraction (S2S, Suppl. Fig. 1B) containing DNA fragments ranging from 50 to 300 bps and S2 large fraction (S2L, Suppl. Fig. 1C) gathering longer fragments. These two fractions are then separately processed. S2L-associated DNA, together with S3- and S4-associated DNA is further sonicated before library

preparation while S2S-associated DNA is directly used for library constructions.



Supplementary Figure 1 – Separation of S2 small (S2S) and S2 large (S2L) subfractions.
Representative bioanalyzer assay of S2 purified DNA before size-selection (**A**) and after separation of S2S (**B**) and S2L (**C**) subfractions. The lower marker (35bp) and the upper marker (10.380bp) are labelled in green and purple, respectively.

Visual inspection of read coverage profiles of each sequenced fraction highlights the improvements of the novel 4fSAMMY-seq technology (Fig. 1B). As mentioned above, 3fSAMMY-seq shows megabase-scale “bumps” in insoluble S3 fraction and, even more prominently, S4 fraction highlighting Lamin Associated Domains (LADs) that overlap with H3K9me3 and LaminB1 ChIP-seq

signals. On the other hand, the distribution of 3fSAMMY-seq reads in S2 is similar to the S3 fraction and shows no enrichment for accessible chromatin. The 4fSAMMY-seq shows a different picture with S2 subfractions (S2S and S2L) clearly enriched in chromatin region marked with euchromatin-specific histone modifications (H3K36me3, H3K4me3) and “bumps” previously observed in S3 and S4 substituted by “valleys” in S2S and S2L subfracions (Fig. 1B).

4fSAMMY-seq comparison of soluble S2S fraction versus more insoluble fraction S3 isolates a reproducible set of regions in the positive y-axis correlating with H3K36me3 and H3K4me3 ChIP-seq signals and in the negative y-axis correlating with heterochromatic regions (H3K9me3 and Lamin B marks) (Fig. 2A).

A.

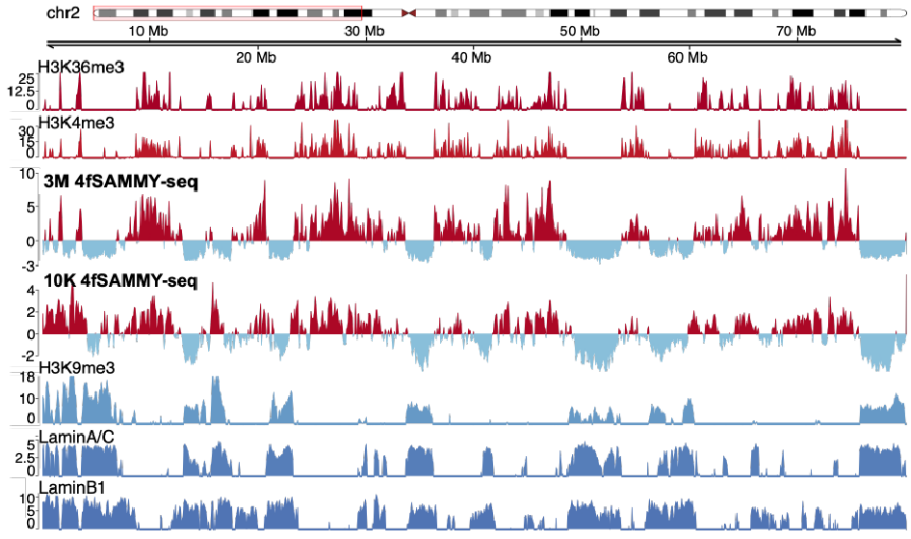


Figure 2 continued on the next page.

Figure 2 continued.

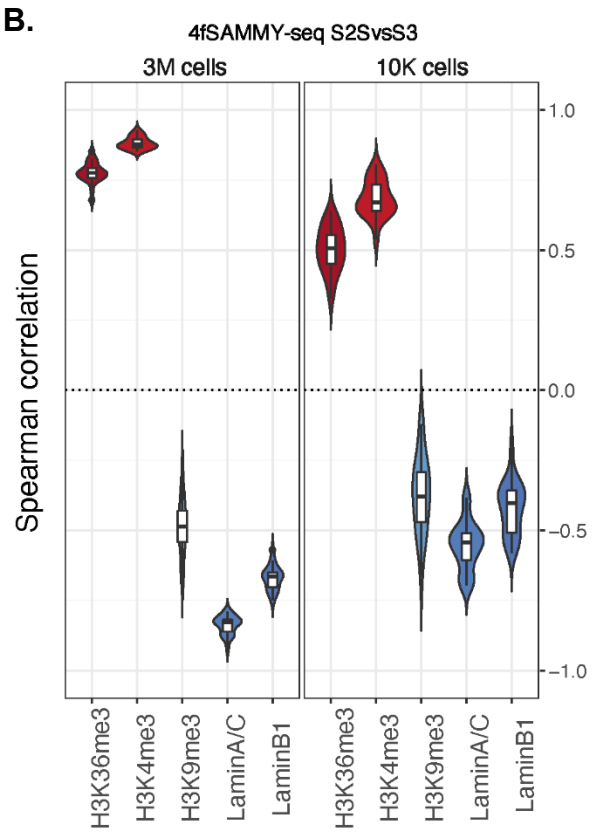


Figure 2 - 4fSAMMY-seq differentially separates heterochromatin and euchromatin in human fibroblasts.
A. Differential read distributions across pairwise comparison of 4fSAMMY-seq fractions (S2SvsS3) in primary human fibroblasts aligned on a representative region of chromosome 2 (80Mb on chr2:1-80.055.066). The smoothed differential signal is calculated with the SPP package. Positive signals are marked in red and represent the S2S enriched regions, whereas negative signals are marked in light-blue and represent the S3 enrichments. ChIP-seq tracks of euchromatin marks H3K36me3 and H3K4me3 (red) and heterochromatin mark H3K9me3 (blue), Lamin A/C and Lamin B (dark blue) of the same region are shown. **B.** Box-Violin plot showing the chromosome by chromosome Spearman correlation coefficient of 4fSAMMY-seq

S2SvsS3 comparison with euchromatin-associated histone marks (HMs) (H3K36me3 and H3K4me3), heterochromatin-associated HM (H3K9me3) and ChIP-seq of LaminA/C and LaminB1.

We confirmed these observations at the genome-wide level using Spearman correlation analysis (Fig. 2B), which shows reproducible positive correlation of S2S vs S3-enriched regions with open chromatin and anticorrelation with closed chromatin. To further prove the robustness of the technique, this analysis was performed on 10K cells showing that 4fSAMMY-seq maintains the capability of discriminating between euchromatin and heterochromatin in a scale-down test (Fig. 2A, Fig. 2B).
In conclusion, the novel 4fSAMMY-seq allows a genome-wide mapping of both lamina-associated heterochromatic regions and active euchromatin regions in a single assay on living cells, without fixation, reconstructing the genome functional organization of the cells.

3.2 4fSAMMY-seq identifies functional chromatin solubility states in normal prostate biopsies

Prostate cancer (PCa herein) is the second most common male malignancy[2], often characterized by an unfavorable outcome due to an asymptomatic nature in its earliest stages. Of note, this type of tumor is characterized by multifocal random growth in the prostate, in fact the collection of multiple biopsy tissues remains the standard for diagnostic and prognostic assessment of PCa. In addition to the well-established role of genomic alterations during PCa progression toward more malignant forms[3,4,5,6,7], several works have recently highlighted epigenetic rearrangements as key events in driving prostate carcinogenesis and subsequent poor outcome [8,9,10,11,12].

Taking advantage of our technology capable to simultaneously isolate heterochromatin and euchromatin regions starting from limited amount of cells, we applied 4fSAMMY-seq on fresh prostate biopsy needle specimens.

To this end, we gathered prostate tissues from patients undergoing standard prostate mapping biopsy for putative PCa diagnosis and processed them as depicted in Figure 3. Briefly, tissues were divided in two subsections destined either to tissue inclusion or enzymatic dissociation. After tissue digestion (Suppl. Fig. 3) specimens were subjected to genome-wide analyses as 4fSAMMY-seq and RNA-seq (Fig. 3). According to the current standard for the detection of PCa, 14 ultrasound-guided cores, seven from each side of the prostate gland, were collected from

each patient for the clinical diagnosis. During the same surgery, one more specimen, destined to our research purpose, was taken directly adjacent to one clinical core of the diagnostics biopsy set (Fig. 3).

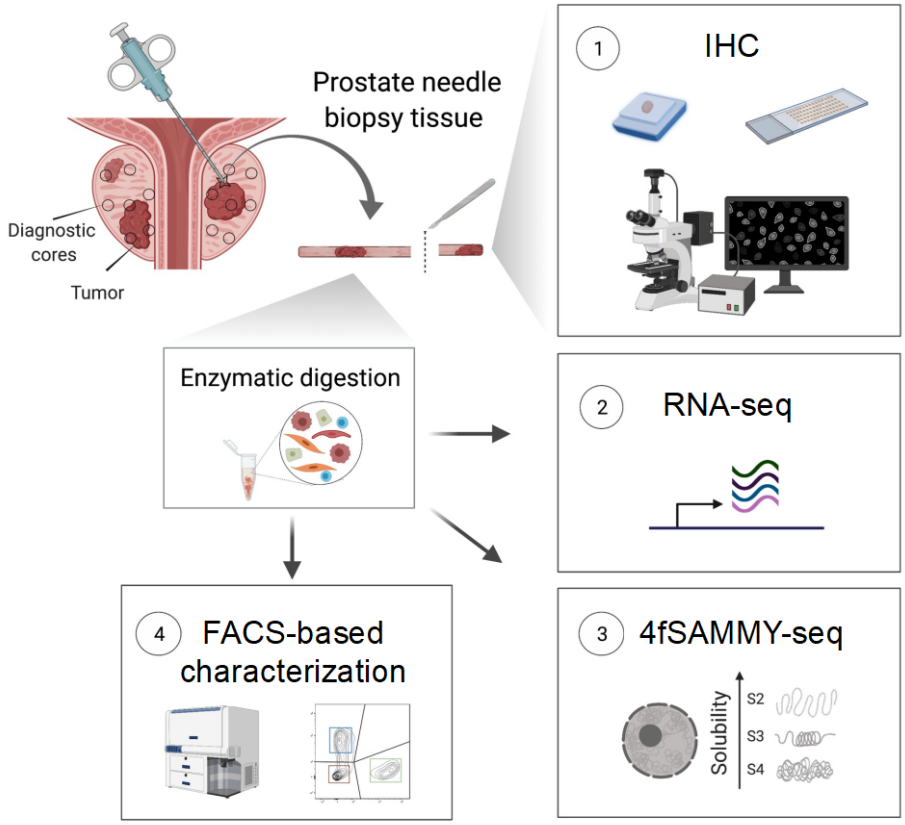


Figure 3 – Overview of experimental design.
Fresh tissues were dissected into 2 portions: one third was flash frozen and embedded in OCT for histopathological and immunohistological analyses (1) while the rest was enzymatically digested into single cells suspensions that were further divided for RNA sequencing (2) 4fSAMMY sequencing (3) and Flow cytometry-based phenotyping assays (4). The 14 clinical transrectal ultrasonography (TRUS)-guided core biopsies, taken from the transition zone, apex, mid and basal regions of the prostate, are indicated as small circles. The research-dedicated

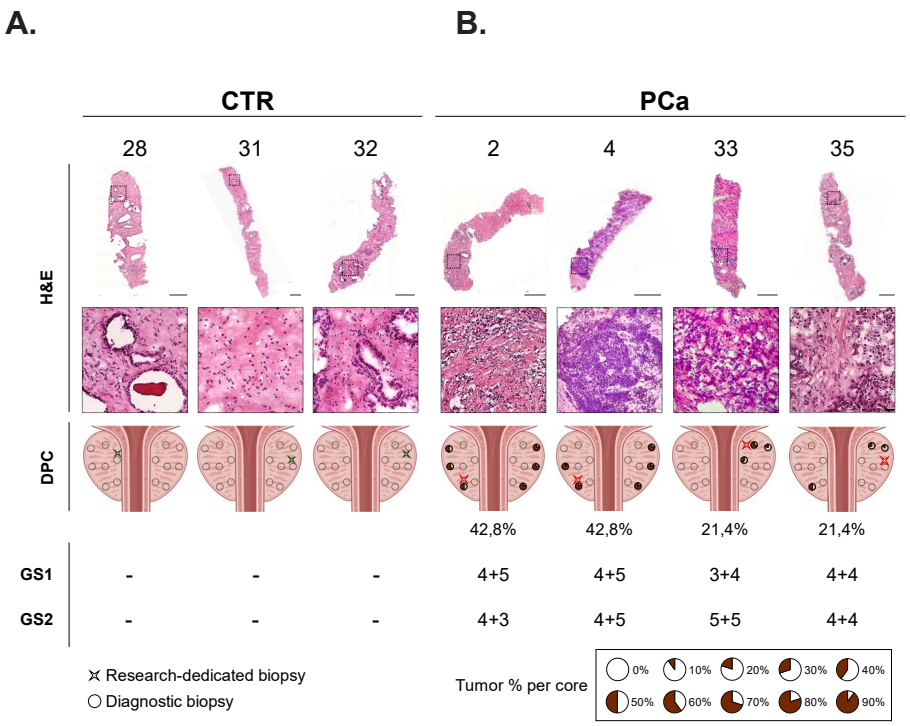
tissue, taken in the proximity of one diagnostic clinical core, is labelled with a star. The schematic representation is created with BioRender.com.

The medical diagnosis of prostate cancer derived from the histologically evaluation of all 14 cores of the clinical biopsy set. Gleason Score (GS) were assigned considering cores containing cancer as described in[13]. Since the multifocality of PCa, also the research-dedicated biopsy was evaluated for the tumor differentiation grade. To standardize the histopathological analysis of the tissue, all biopsies, including the one we used, were examined by the same pathologist.

Based on histology and spatial distribution of the positive cores we distinguished two different groups: histologically normal control biopsies from cancer-free patients (all negative cores) (Suppl. Fig. 2A) and PCa biopsies with histologically confirmed percentage of tumor in the research-dedicated biopsy and/or in the nearby clinical biopsy (Suppl. Fig. 2B).

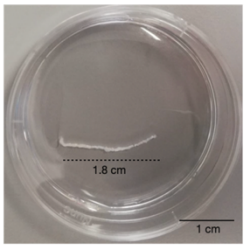
Given the limiting size (around 2 cm in length) of the prostate bioptic tissue used for genome conformations assays (Suppl. Fig. 3A), we developed a specific dissociation protocol by combining mechanical manipulation with a gentle collagenase digestion of the fresh tissue (see methods).

This enzymatic digestion optimized cell yield and viability without compromising their surface antigenicity (Suppl. Fig. 3B). This ensured an adequate number of viable cells (about 3000 cells per mg of prostate biopsy) for subsequent sequencing analysis (Suppl. Fig. 3C-D).

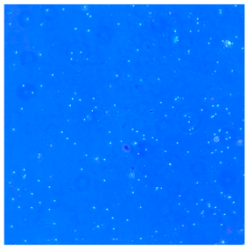


Supplementary Figure 2 – Specimen cohort.
Photomicrographs reconstruction at Leica DMI6000 (hematoxylin-eosin staining: magnification, x10) of histologically normal prostatic biopsies (**A**) and histologically confirmed PCa biopsies (**B**). Inset in the dashed grey boxed area shows a 20x magnification. The circles represent biopsies used for clinical diagnosis. The research-dedicated tissue, taken in the proximity of one diagnostic clinical core, is labelled with a star.
The quantification of tumor is marked with a pie chart showing the total percentage of each needle biopsy tissue positive for prostate cancer. Spatial distribution of positive cores (DPC) and Percentage of Positive Cores (PPC) are also shown.
GS1 indicate Gleason score assigned to our research-dedicated biopsy. GS2 indicate diagnostic Gleason score assigned after evaluation of 14 clinical cores. Scale bars: 500 µm.
The representative pictures of the transrectal ultrasound-guided (TRUS) prostate biopsy scheme are created with BioRender.com.

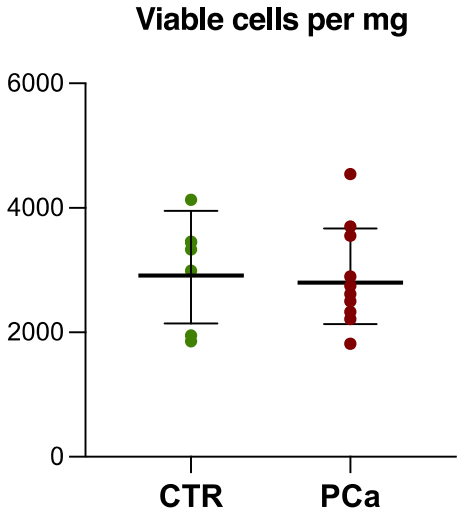
A.



B.



D.



C.

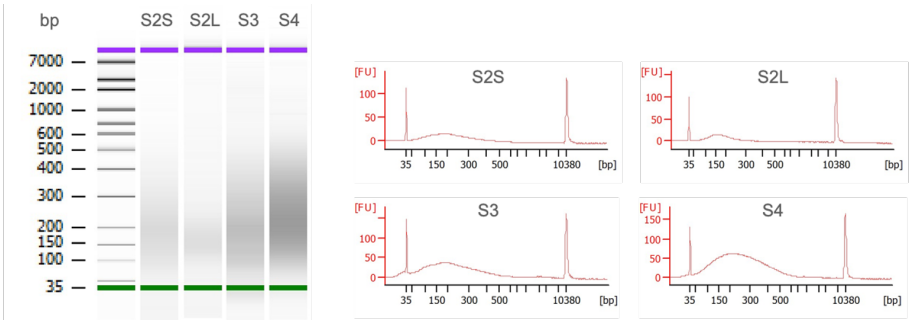
	Biopsy weight [mg] **Median** (Range)	Viable Cells **Median** (Range)	Viable Cells per mg **Median** (Range)
CTR 7	**16** (22-10,3)	**47.000** (20.000-71.000)	**3000** (1860-4100)
PCa 10	**22** (11-39)	**60.000** (30.000-100.000)	**2900** (1820-4500)

Supplementary Figure 3 – Dissociation of prostatic needle biopsies into single-cell suspensions.
A. Representative image of needle prostate biopsy specimen in a petri dish (35x15mm). Bar length=1 cm. **B**, Representative microscope image of prostatic single cells suspension after enzymatic digestion (10x magnification). **C**. Table showing viable cells (second column) quantified using Trypan Blue staining, normalized to the initial tissue weight (first column), and reported as thousands of live cells per mg of tissue (third column). **D.** Number of viable cells recovered per mg of each tissue biopsy. 7 control and 10 PCa samples are indicated in green and red, respectively.

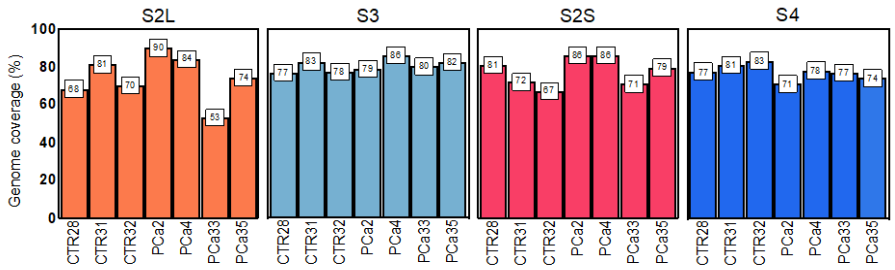
We initially applied 4fSAMMY-seq on 3 tumor-free specimens (CTR28, CTR31, CTR32).

Despite the limited number of tissue-derived cells (between 10-80K cells), we observed a correct pattern of fractionation, size selection of S2-digested fragments and sonication, as shown in Suppl. Fig. 4A. Moreover, we generated a mean yield of 35 million sequenced reads per fraction and on average 73% (S2L), 73% (S2S), 79% (S3) and 80% (S4) of the genome were covered by at least one read (Suppl. Fig. 4B). Taken together these data suggest that 4fSAMMY-seq is suitable for analysis on limited materials as human samples derived from clinical practice.

A.



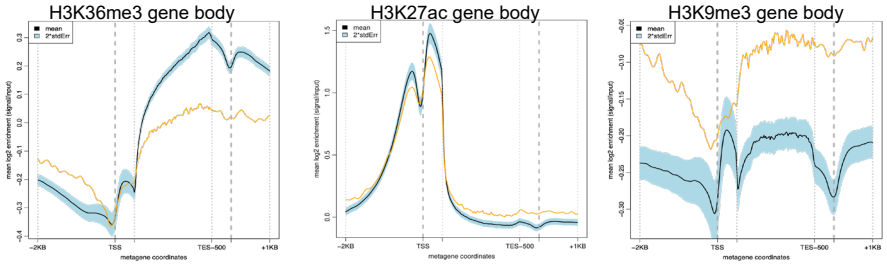
B.



Supplementary Figure 4 continued on the next page.

Supplementary Figure 4 continued.

C.



Supplementary Figure 4 – Quality check.
A, Representative bioanalyzer assay after S2 size selection and sonication for each fraction (S2S, S2L, S3, S4). **B.** Percentage of the human genome covered by at least one read in each 4fSAMMY-seq CTR and PCA sample (S2L, S3, S2S, S4). **C.** ChIP-seq metaprofile enrichment from published datasets of H3K36me3 (ENCSR768PFZ), H3K27ac (ENCSR763IDK) and H3K9me3 (ENCSR133QBG) in prostate gland tissue constructed by ChIP-seq quality Control framework (ChIC)[14]. Average log2 transform normalized (ChIP over input) enrichment signals is shown for tested ENCODE samples (yellow line) compared to the mean signal of reference compendium datasets for the same chromatin mark (grey line) along with its ±2 standard error interval (blue shadow).

Visual inspection of a representative genomic locus (Fig. 4A) of 4fSAMMY-seq comparison of soluble S2L versus insoluble S3 shows a conserved and robust presence of differentially enriched genomic regions in histologically control biopsies (CTR28, CTR31, CTR32).

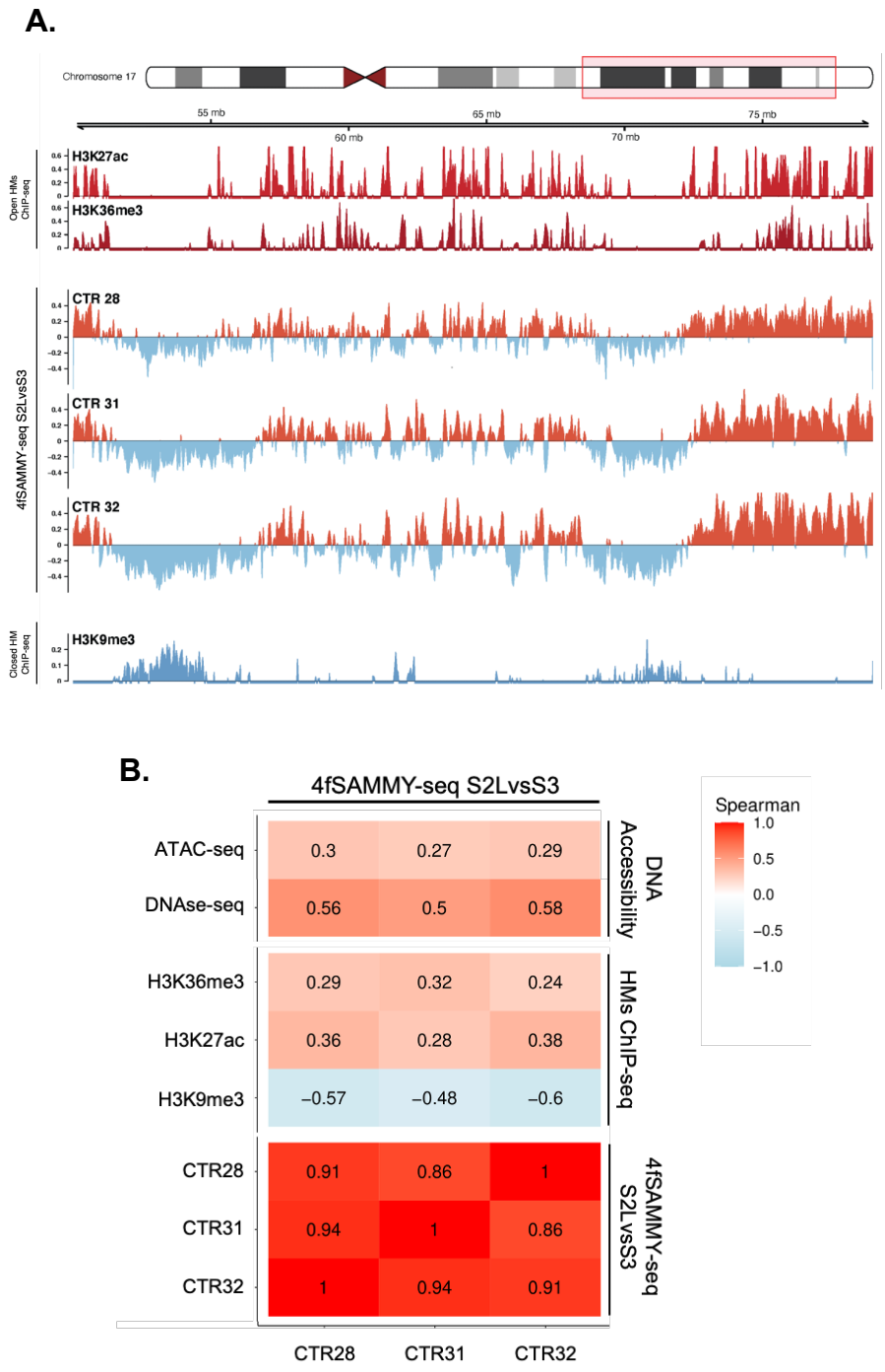


Figure 4 – 4fSAMMY-seq discriminates heterochromatic and euchromatic regions in normal prostate biopsies.
A. Differential reads distributions across pairwise comparison of 4fSAMMY-seq (S2LvsS3) in each histologically normal tissue

(CTR28, CTR31, CTR32) aligned on a representative region of chromosome 17 (29 Mb on chr17:50.000.000-79.000.000). The smoothened differential signal was calculated with SPP package. Positive signal enrichments are marked in orange and correspond to the more soluble S2L regions, whereas negative signal enrichments are marked in light-blue and represent the insoluble fraction S3. Publicly available ChIP-seq data from normal prostate gland of euchromatin-associated marks H3K27ac (red) and H3K36me3 (dark red) and of heterochromatin mark (H3K9me3, blue) are shown. For ChIP-seq data, the (ChIP–input) reads distribution was computed with SPP package (y axis range to zero as a minimum value). **B.** Genome-wide Spearman correlation for 4fSAMMY-seq (S2LvsS3 enrichment) between three control biopsies (CTR28, CTR31, CTR32) and publicly available ATAC-seq (ENCSR999NKW), Dnase-seq (ENCSR564FZH), and histone modifications ChIP-seq (H3K27ac, H3K36me3, and H3K9me3) from normal prostate glands.

To functionally characterize these conserved regions among control prostate biopsies, we compared the SAMMY-seq enrichments with publicly available ChIP-seq tracks from normal prostate gland tissues. To this end, we first checked the quality of the published datasets analysing the metagene profile for the corresponding chromatin marks by using a recent tool named ChIP-seq quality Control framework (ChIC)[14]. As expected, we found sharp enrichments of H3K27ac around the TSS, broad enrichments of H3K36me3 along the entire gene body and depletion of H3K9me3 signal around the TSS and TES (Suppl. Fig. 4C), confirming the validity of the selected reference data-set. As anticipated in[14], some differences in shape and distribution of the enrichment profile in the broad mark H3K9me3

are accepted due to its preferential localization in gene-desert regions.

In control prostate tissues, we observed a reproducible enrichment of open chromatin marks (H3K27ac, H3K36me3) coinciding with soluble, S2LvsS3 regions. At the same time, negative enrichments correspond to very broad regions of the H3K9me3 heterochromatin marker (Fig. 4A).

These results were further validated at the genome-wide level by pairwise correlation between control 4fSAMMY-seq specimens (S2LvsS3 comparison) and considering 4fSAMMY-seq enrichments and ChIP-seq signal (Fig. 4B). We found a strong correlation of genome distribution among healthy samples (on average r = 0,9). Regarding histones, as expected we found high anticorrelations of soluble domains with heterochromatin (between -0,48 and -0,6), while the correlations with euchromatin associated histone marks are lower reflecting the interindividual variability (between 0,24 and 0,38).

Histone marks are relevant epigenetic determinants, but represent a fraction of all the factors contributing to chromatin solubility features captured by 4fSAMMY-seq. We further characterized our data by comparison with ATAC-seq and DNase-seq, two techniques developed to capture accessible chromatin. As shown in Figure 4B, the soluble compartment captured by 4fSAMMY-seq (S2LvsS3 comparison) shows a positive correlation with both these techniques, ranging between 0,27 and 0,58 values.

Overall, these observations prove the reliability of 4fSAMMY-seq to comprehensively map the 3D genome organization of cells deriving from limited biological material as prostate human biopsies.

3.3 Identification of malignant chromatin solubility rearrangements in PCa biopsies by 4fSAMMY-seq

To understand how nuclear architecture is altered in PCa, we applied 4fSAMMY-seq on 4 primary prostate cancer tissues (PCa2, PCa4, PCa33, PCa35) with different grade of aggressiveness and local invasiveness (Suppl. Fig. 2B). In line with control samples, the sequencing quality check in PCa samples showed on average 75% (S2L), 80% (S2S), 82% (S3) and 75% (S4) of the genome covered by at least one read (Suppl. Fig. 4B).

From the observation of 4fSAMMY-seq tracks on a representative genomic locus, we noticed that malignant tissues exhibit patient-specific alterations of chromatin solubility pattern compared with healthy ones (Fig. 5A). Indeed, 4fSAMMY-seq enriched regions appeared more variable and scattered across PCa tissues respect to controls. Accordingly, PCa tissues do not share a high degree of similarity via genome-wide Spearman correlation of 4fSAMMY-seq signal tracks (Fig. 5B).

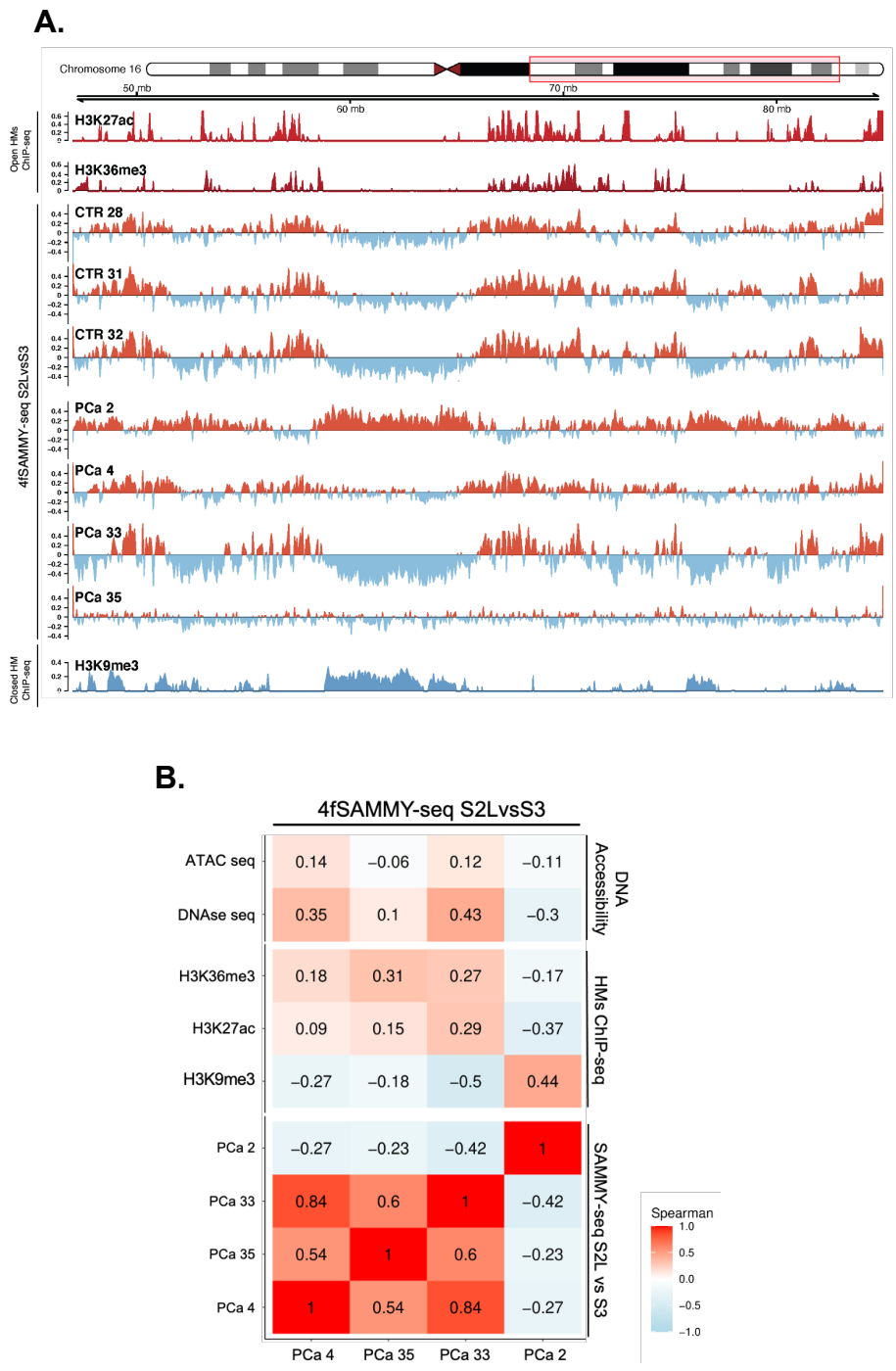


Figure 5 – 4fSAMMY-seq reveals chromatin solubility changes between normal prostate and PCa biopsies.
A. Differential reads distributions across pairwise comparison of 4fSAMMY-seq (S2LvsS3) in each histologically normal tissue

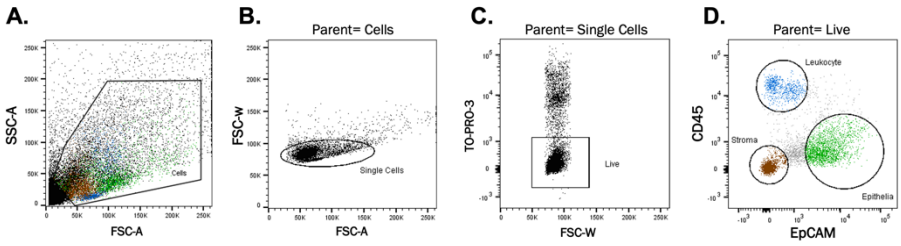
(CTR28, CTR31, CTR32) and prostate cancer tissue (PCa2, PCa4, PCa33 and PCa35) aligned on a representative region of chromosome 16 (38Mb on chr16:47.000.000-85.000.000). The smoothened differential signal was calculated with SPP package. Positive signal enrichments are marked in orange and correspond to the more soluble S2L regions, whereas negative signal enrichments are marked in light-blue and represent the insoluble fraction S3. Publicly available ChIP-seq data from normal prostate gland of euchromatin-associated marks H3K27ac (red) and H3K36me3 (dark red) and of heterochromatin mark (H3K9me3, blue) are shown. For ChIP-seq data, the (ChIP–input) reads distribution was computed with SPP package (y axis range to zero as a minimum value). **B.** Genome-wide Spearman correlation for 4fSAMMY-seq (S2LvsS3 enrichment) of prostate cancer biopsies (PCa2, PCa4, PCa33, PCa35) compared to publicly available ATAC-seq, DNase-seq and ChIP-seq (H3K27ac, H3K36me3, H3K9me3) data from normal prostate gland.

We next explored the Spearman correlation for 4fSAMMY-seq S2LvsS3 comparison to reference controls for accessible chromatin (ATAC-seq and Dnase-seq) and chromatin HMs (H3K27ac, H3K36me3 and H3K9me3) finding three distinct conditions: a complete loss of correlation in 1 out of 4 patients (PCa2), lower coefficient values, more prominently with heterochromatin, respect to control samples in 2 out of 4 patients (PCa4, PCa35) and a pattern of solubility similar to the controls in the last patient (PCa33) (Fig. 5B).

Taken together, these observations confirm the highly heterogeneous and variable nature of PCa and suggest that 4fSAMMY-seq can detect nuclear architecture aberrancies on bulk tissues.

3.4 4fSAMMY solubility profile is not influenced by relative amount of distinct cell types

Biopsy tissue consist of a heterogeneous cellular composition, mainly represented by two generic cell types: fibromuscular stroma and epithelial cells. Pathological prostate specimens also contain a high number of infiltrating leukocytes[15]. The observed bulk tissue solubility picture thus arises from the combination of individual cell-type profiles. To properly interpret epigenetic data obtained on bulk tissues and to assess if observed chromatin solubility variations are dependent on cell-types composition, we processed additional 13 biopsies adding a new step in our flowchart: a FACS-based profiling on 10K cells from the tissue-digested suspension (Fig. 3). We adopted a well-established gating strategy for human prostate subpopulations[16] that, using the surface expression of the pan-epithelial marker CD326 along with the pan-leukocyte marker CD45, discriminates the following populations: epithelial, stromal and leukocyte cells (see methods, Suppl. Fig. 5).



Supplementary Figure 5 – Flow Cytometry gating strategy. Primary cells from prostate tissues were stained with CD326 (EpCAM) and CD45 (PTPRC) antibodies together with the viability marker TO-PRO-3. **A.** Forward and side scatter gating

(FSC-A vs SSC-A) to identify tissue-digested prostatic cells and to remove debris. **B.** FSC-A vs FSC-W plot allows doublets removal. **C**. FSC-W vs TO-PRO-3 plot excludes low viability cells. **D.** Tissue resident leukocytes (blue) are identified by CD45, epithelia are identified by EpCAM (green), and stroma (brown) are the double negative cells.

No significant differences were found in the relative proportion of the major cell populations between control and PCa groups (Fig. 6A). In general, we identified a prevalence of stroma in all prostate tissues while we found an average of 16% and 15% of epithelial cells in control and PCa biopsies, respectively. As expected, malignant biopsy tissues contain on average a higher amount of infiltrating leukocytes respect to controls, however control specimens exhibit a large range of the percentage of immune cells within the samples (Fig. 6A).

The flow cytometry profiling suggests that the cellular composition of the biopsy *per se* is not sufficient to discriminate the prostate cancer state.

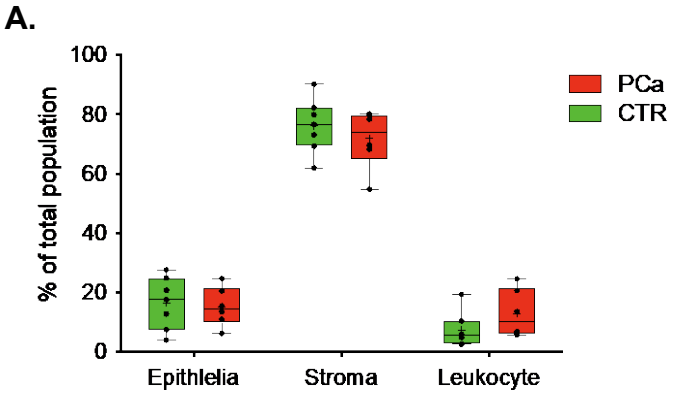


Figure 6 continued on the next page.

Figure 6 continued.

B.

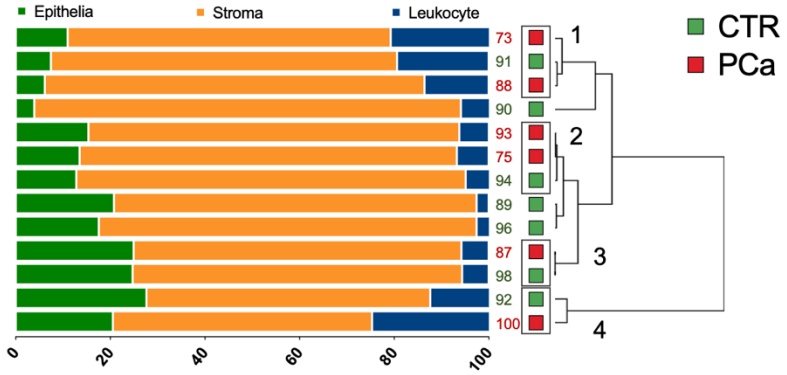
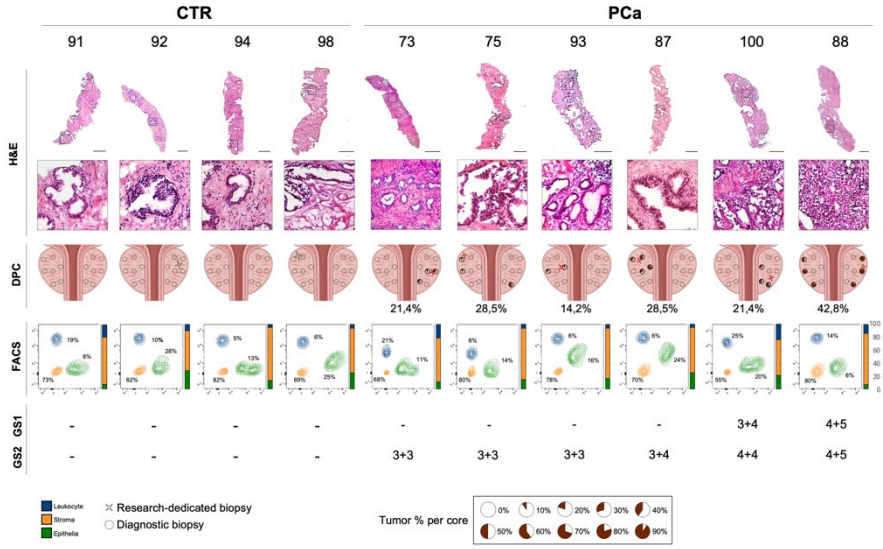


Figure 6 – Flow Cytometry characterization of prostate bioptic tissues.
A. Box plot showing the percentage of 3 distinct prostate cell types (epithelial cells, stromal cells and leukocytes) in the histologically normal samples (green box) and prostate cancer samples (red box). In the box plots, the whiskers of the box extend from the minimum value to the maximum value, the box extends from the 25th to the 75th percentiles, the middle line of the box is plotted at the median and the mean is indicated by the black plus. **B.** The relative percent of different cell subsets (epithelial cells, stromal cells and leukocytes) in each histologically normal control (green) and prostate cancer sample (red). Dendrogram (on the right) showing the unsupervised clustering of the biopsy tissues based on cell-type composition. The boxes indicating the 4 cluster of tissues with comparable cell composition designated for sequencing. CTR91-PCa73-PCa88 cluster is characterized by high representative level of infiltrating leukocytes. Two clusters (CTR94-PCa75-PCa93 and CTR98-PCa87) include tissues that exhibit poorer immune response. CTR92-PCa100 cluster consist of tissues with a higher presence of both epithelia and infiltrating leukocyte.

Using clustering analysis, we discriminated groups of control and PCa specimens with comparable cell composition (Fig. 6B) and we selected 10 additional samples to sequence (4 CTR and 6 PCa) (Suppl. Fig. 6).



Supplementary Figure 6 – Specimen cohort.
Photomicrographs reconstruction at Leica DMI6000 (hematoxylin-eosin staining: magnification, x10) of histologically normal prostatic biopsies and histologically confirmed PCa biopsies. Inset in the dashed grey boxed area shows a x20 magnification. The 14 clinical cores of a biopsy set are indicated as small circles and study cohort samples as stars. Scale bars: 500 µm. Tumor areas are marked with a pie chart showing the total percentage of each needle biopsy tissue involved by PCa. Spatial distribution of positive cores (DPC) and consequent percentage of positive cores are also shown. FACS plots show the relative percent of different cell-type subsets (epithelia, stroma and leukocytes). GS1 indicate Gleason score assigned to our research-dedicated biopsy. GS2 indicate diagnostic Gleason score assigned after evaluation of 14 clinical cores.

Despite differences in cell-type composition among control samples, 4fSAMMY-seq allowed the detection of a conserved pattern of soluble/insoluble regions associated with the respective epigenetic histone modifications (Fig. 7A). On the other hand, 4fSAMMY-seq tracks of PCa specimens confirmed previous results, showing cancer-specific alteration of chromatin compartmentalization.

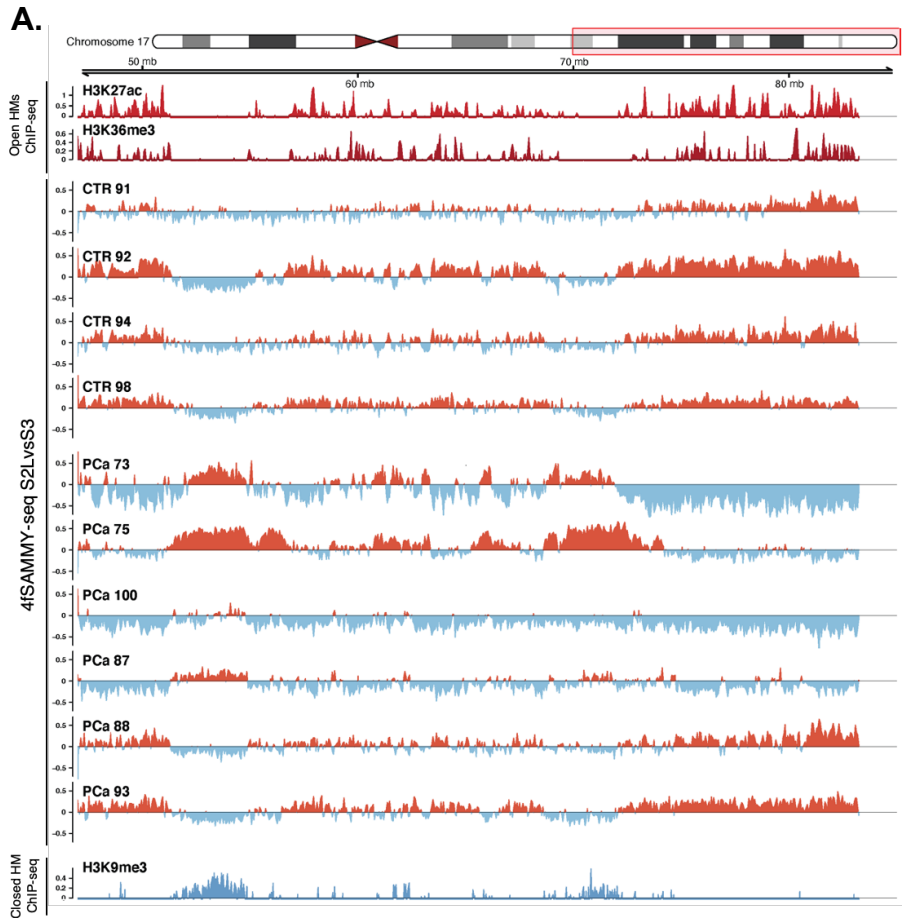


Figure 7 continued on the next page.

Figure 7 continued.

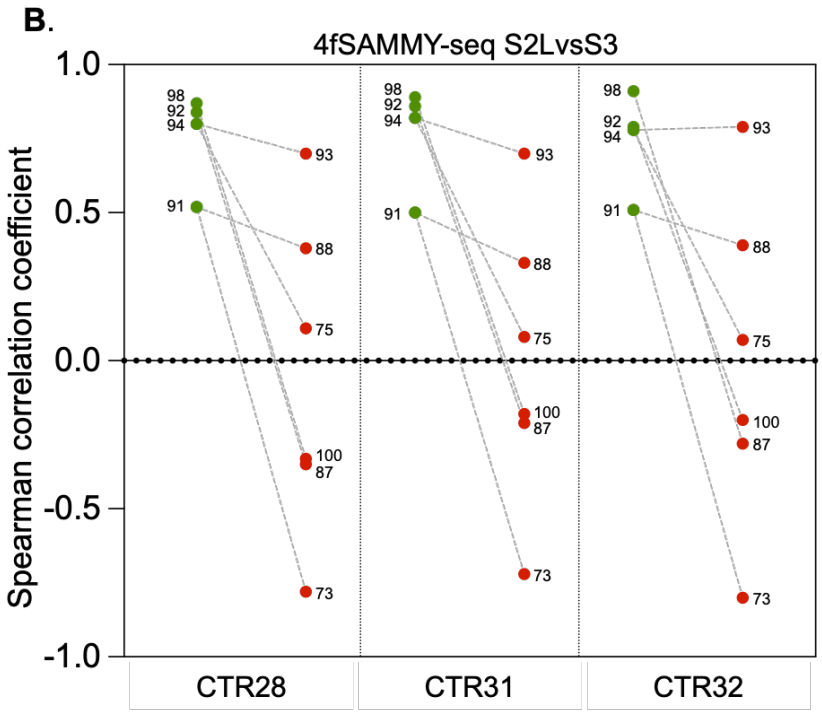


Figure 7 – PCa-related 4fSAMMY-seq profile is not dependent on tissue composition.
A. Visualization of 4fSAMMY-seq comparison (S2LvsS3) of histologically normal prostate (CTR91, CTR92, CTR94, CTR98) and prostate cancer (PCa73, PCa75, PCa93, PCa87, PCa100, PCa88) biopsies along a representative region of chromosome 17 (30Mb on chr17:47.000.000-85.000.000). The smoothed differential signal is calculated with SPP. The S2LvsS3 signal enrichment is labelled in orange and the S3vsS2L signal enrichment is marked in light-blue. Publicly available ChIP-seq tracks on normal prostate gland of euchromatin marks H3K27ac (red) and H3K36me3 (dark red) and heterochromatin mark H3K9me3 (blue) in the same region are also shown. **B.** Genome-wide Spearman correlation values for 4fSAMMY-seq (S2LvsS3 enrichment) of 4 histologically normal prostate biopsies labelled in green and 6 prostate cancer biopsies marked in red compared

to 4fSAMMY-seq (S2LvsS3 comparison) data on control prostate biopsies of the first cohort (Fig. 4, CTR28, CTR31, CTR32). The dashed gray line associates each control normal tissue to distinct prostate cancer biopsies based on cell-type composition subset (CTR91-PCa73-PCa88; CTR92-PCa100; CTR94-PCa75-PCa93; CTR98-PCa87).

Genome-wide Spearman correlation for 4fSAMMY-seq (S2LvsS3) between controls of the first cohort (CTR 28, CTR 31, CTR 32) and all samples of the second cohort showed a conserved solubility profile among control biopsies with a clear compartmentalization of euchromatin and heterochromatin and a strong variability of PCa tissues, not ascribable to different cell-types composition as indicated by dashed grey lines (Fig. 7B).

To further support this data, we have divided the human reference genome into discrete units of 250 kb, called bins, and calculated the score of 4fSAMMY-seq (S2LvsS3 comparison) in each bin. Then, we have ranked the bins according to the mean values of 4fSAMMY-seq between control samples, referred as consensus. We noticed that PCa tissues exhibit cancer-specific 4fSAMMY-seq fingerprint in the distribution and shape of solubility enrichment compared with healthy ones. Indeed, except for one cancer tissue (PCa93), we discriminated in each PCa sample distinct differently enriched bins clearly separate from the grey bins of control samples. Of note, PCa93 showed the lowest percentage of positive cores (Suppl. Fig. 6).

We highlighted an almost completely inverted pattern for 5 PCa tissues (PCa2, PCa100, PCa73, PCa75, PCa87) and a general

altered enrichments score in PCa33 becoming less evident in the other tumor tissues (PCa4, PCa35, PCa88) (Fig. 8).

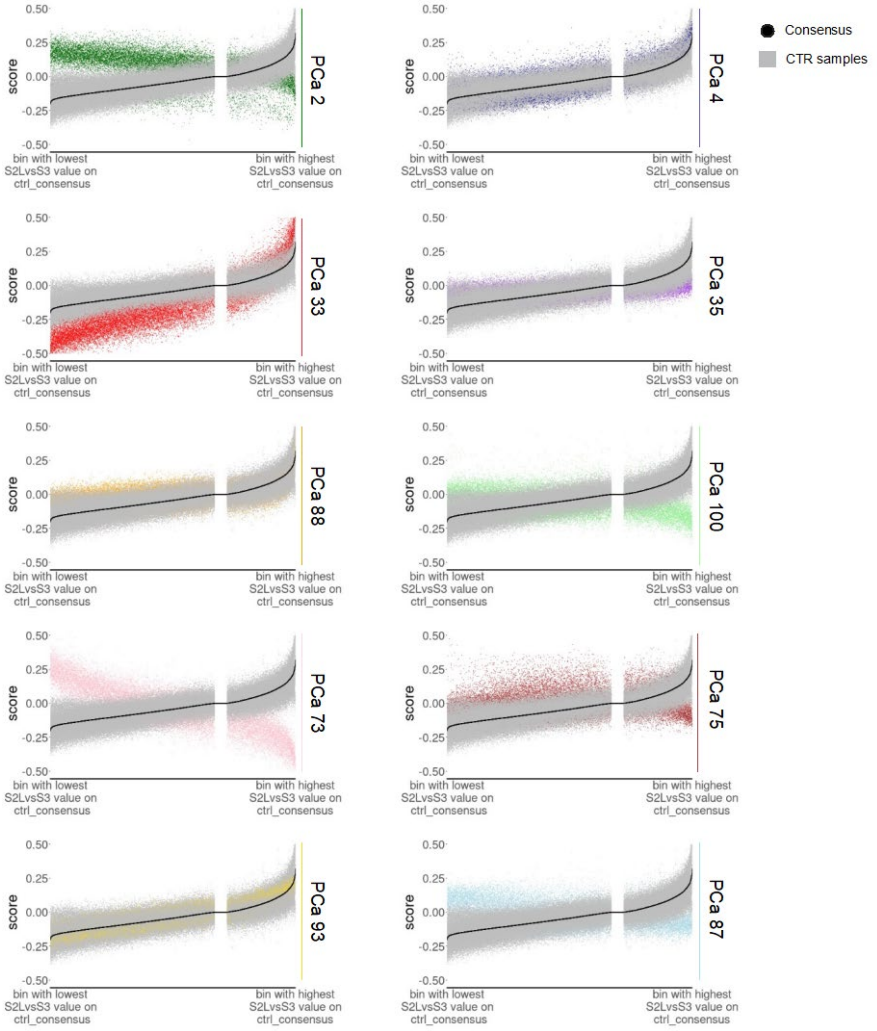
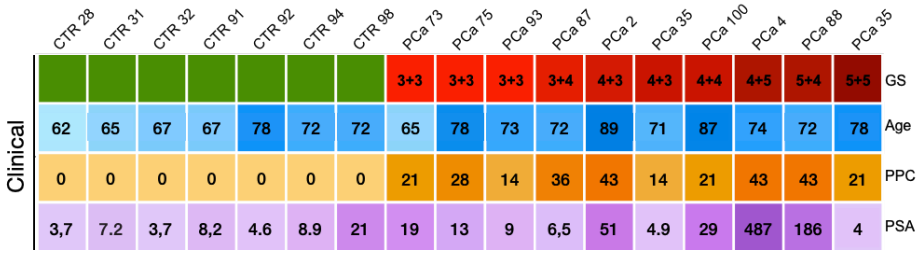


Figure 8 – PCa-specific 4fSAMMY-seq fingerprint.
4fSAMMY-seq values (y-axis) of 7 control sample (CTR28, CTR31, CTR32, CTR91, CTR92, CTR94, CTR98 grey) and 10 PCa tissues (PCa2 dark-green, PCa 4 blue, PCa 33 red, PCa 35 purple, PCa 88 orange, PCa100 green, PCa 73 pink, PCa 75 dark-red, PCa 93 yellow, PCa 87 light-blue) using a 250 kb bins size. The bins have been sorted according to the value of the S2LvsS3 consensus of 7 control patients (black line).

Taken together, these results show that the 4fSAMMY-seq technology holds the key to profile the vast heterogeneous chromatin remodeling of prostate tumor tissues.

3.5 Prostate cancer tissues exhibit two distinct chromatin solubility profiles

The patients enrolled in our study are characterized by a high degree of clinical heterogeneity in terms of age, prostate specific antigen (PSA) test, Gleason score (GC) and percentage of positive biopsy cores (Suppl. Fig. 7). The patients are men aged 62-89, with a median of 72 years. Our specimens have a primary Gleason pattern from 3 (low-grade) to 5 (high-grade). In all PCa patients, biopsy spatial sampling identified multiple distinct loci of cancer with varying histologic grades (Suppl. Fig. 2-6).



Supplementary Figure 7 – Clinicopathological features. The study participants undergoing prostate biopsy are ordered according to Gleason score (GS). Columns are individual sample and rows are specific features. Each sample is annotated for Gleason Score (GS), Age, Percentage of Positive Cores (PPC) and Prostate Specific Antigen (PSA).

We decided to test whether differential cancer alterations in chromatin solubility underlie specific oncogenic programs. We determined the genome-wide Spearman correlation for 4fSAMMY-seq S2LvsS3 comparison between PCa biopsies to cluster them in different subgroups (Fig. 9A).

We observed two distinct PCa-related solubility remodeling patterns: subtype 1 (PCa4, PCa33, PCa35, PCa88, PCa93) contains cancer tissues that exhibit slight solubility changes respect to controls and subtype 2 (PCa2, PCa73, PCa75, PCa87, PCa100) characterized by severe solubility remodelling (Fig. 9B).

A.

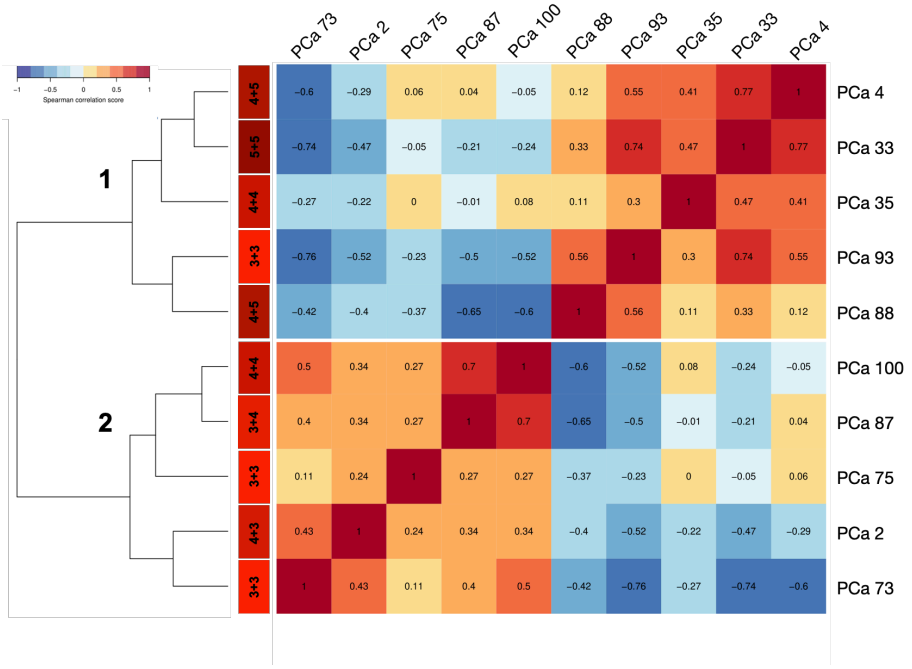


Figure 9 continued on the next page.

Figure 9 continued.

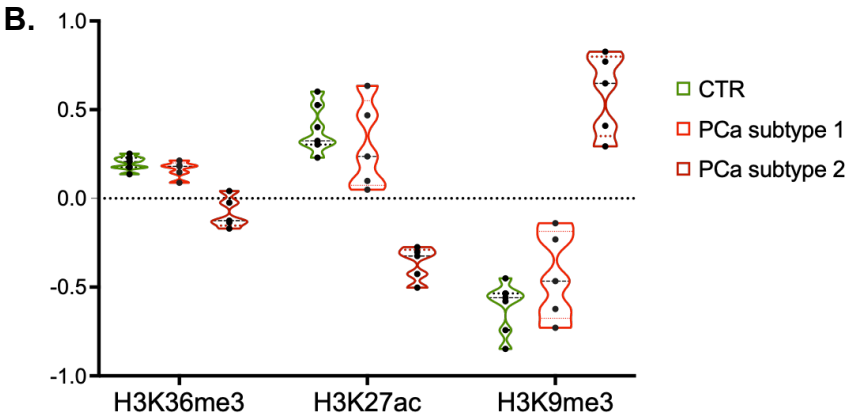


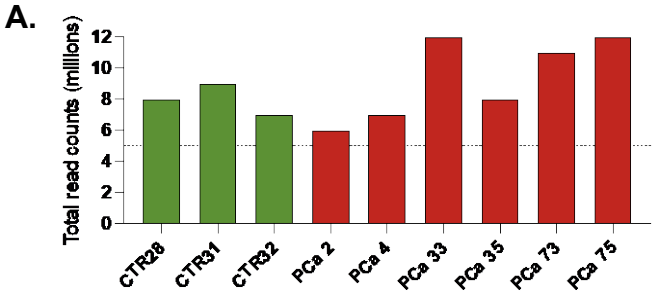
Figure 9 – 4fSAMMY-seq discriminates two distinct solubility clusters among PCa tissues.
A. Genome-wide clustering of 4fSAMMY-seq (S2LvsS3 comparison) in prostate cancer biopsies (PCa2, PCa4, PCa33, PCa35, PCa73, PCa87, PCa100, PCa75, PCa93, PCa88) using the pairwise genome-wide Spearman correlation. Cluster 1 (PCa4, PCa33, PCa35, PCa88, PCa93) contains tissues characterized by low solubility rearrangements. Cluster 2 (PCa2, PCa73, PCa75, PCa87, PCa100) include tissues that exhibit severe solubility rearrangements. Gleason score (GS) associated to each PCa sample is also shown (top). **B.** Violin plot distribution of genome-wide Spearman correlation values for 4fSAMMY-seq (S2LvsS3 enrichment) between control biopsies (CTR28, CTR31, CTR32, CTR91, CTR92, CTR94, CTR98, green), PCa biopsies of subtype 1 (PCa4, PCa33, PCa35, PCa88, PCa93 red) and PCa biopsies of subtype 2 (PCa2, PCa73, PCa75, PCa87, PCa100, dark-red) compared to ChIP-seq (H3K36me3, H3K27ac, H3K9me3) data on normal prostate gland.

Of note, the separation between euchromatin and heterochromatin even in subtype 1 PCa tissues is less clear than in controls groups (Fig. 9B). Notably, no associations between solubility alterations and tumor grade measured with Gleason score were detected (Fig.9A), suggesting that we are capturing

chromatin remodelling not detectable with histopathological analysis of biopsies.

3.6 Integrative transcriptome and chromatin organization analysis reveals aberrant solubility compartmentalization in PCa tissues

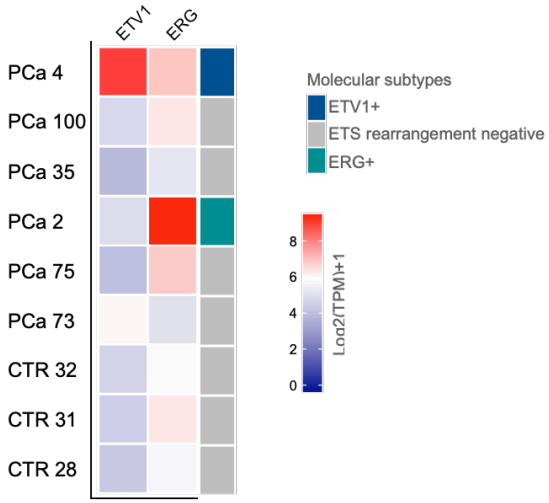
To identify the interplay between chromatin solubility and genome function, we analyzed cancer-induced alterations in the transcriptome. In our RNA-seq dataset, we initially successfully sequenced 10 tissues (59%) with an optimal bare minimum of mapped reads (5 million) for a differential gene expression (DGE) analysis (Suppl. Fig. 8A). First, we performed molecular subtyping assay with STAR-fusion package[17]. We have found that two cancer tissues (PCa2, PCa4) exhibit oncogenic activation of ETS-related genes (ERG, ETV1) due to gene fusion events, in agreement with the well-established notion that such events represent one of the most common genomic alterations in PCa[18] (Suppl. Fig 8B).



Supplementary Figure 8 continued on the next page.

Supplementary Figure 8 continued.

B.



Supplementary Figure 8 – PCa molecular subtyping.
A. Millions of RNA-seq reads per sequenced sample. **B**. Heatmap of Log_2 (TPM) expression values of ERG and ETV1 transcription factors and molecular tumor subtypes defined by ERG/ETV1 gene rearrangements (detected by STAR-fusion package[17]).

Differential expression analyses between tumor (PCa2, PCa4, PCa33, PCa35, PCa73, PCa75) and normal tissues (CTR28, CTR31, CTR32) uncovered only 147 DEGs either upregulated or downregulated (72 and 75, respectively), once again highlighting the heterogeneity of PCa samples. We performed principal component analysis (PCA) to identify groups of transcriptome similarity among samples via the iDEP tool[19]. The resulting PCA reveals that the genes expression profiling clearly segregate

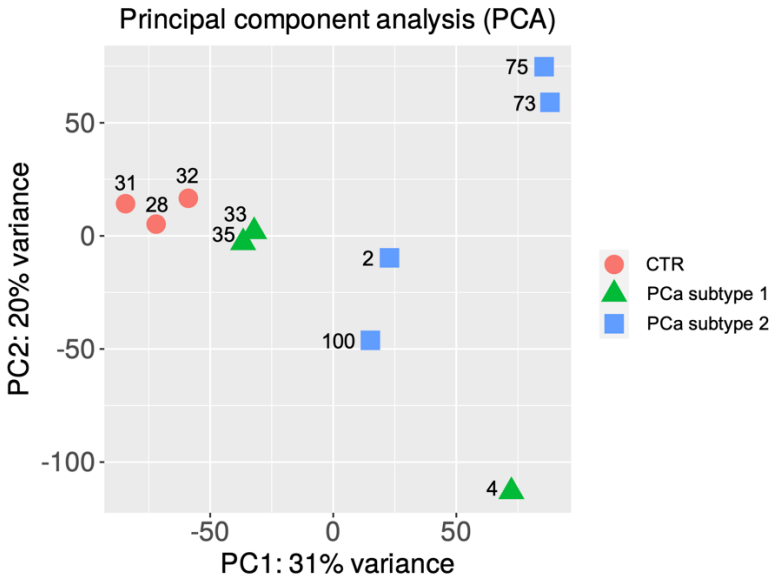
primary neoplastic samples from normal prostatic tissues (Fig. 10A).

Pathway analysis of PCA rotation with the PGSEA package revealed that the substantial difference of genes expression variance in the sample set is captured by canonical hallmarks of PCa, such as androgen response and MYC signalling (Fig. 10B). Moreover, this analysis identified several differential genes involved in terms of TNF signalling via the NFKB pathway, apical junction organization, Epithelial-to-Mesenchymal Transition (EMT) and hypoxia.

Interestingly, PCA on RNA-seq data (Fig. 10A) separates the PCa specimens in two subgroups, matching the clustering we already found with 4fSAMMY-seq (Fig. 9), suggesting a connection between chromatin architecture rearrangements and transcriptome remodulation in PCa. Tissues characterized by slight solubility changes (subtype 1) showed a transcriptome more comparable to control samples, while tissue of subtype 2, exhibiting a severe chromatin remodeling also have a compromised transcription.

An exception of this scheme is represented by the PCa4 specimen, belonging to the subtype 1, that occupies a position distant to controls in the PCA plot (Fig. 10A). However, this sample harbors gene fusion involving ETS family transcription factors (Suppl Fig. 8B) that can *per se* mainly capture the transcriptomic oncogenic program.

A.



B

.

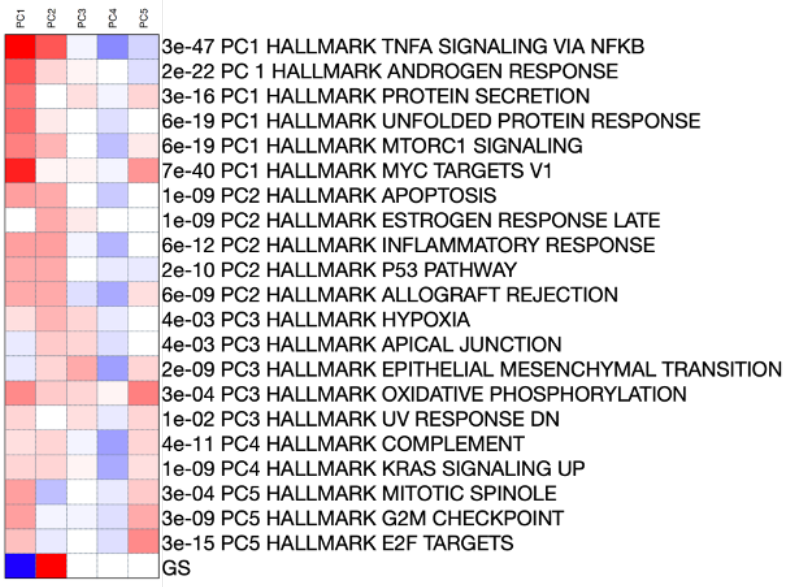


Figure 10 - Transcriptomic differences in PCa samples.
A. Principal Component Analysis (PCA) on RNA-seq data of 3 control samples (CTR28, CTR31, CTR32 orange), 3 PCa samples of subtype 1 (PCa4, PCa33, PCa35 green) and 4 PCa samples of subtype 2 (PCa2, PCa73, PCa75, PCa100 blue).

PCA plot project samples into two-dimensional space. The first components (PC1) points to the direction of the most variation among samples. **B.** Detection of differential gene pathways in the principal components in the comparison between tumor (PCa2, PCa4, PCa33, PCa35, PCa73, PCa75) and normal tissues (CTR28, CTR31, CTR32). For each pathway, this runs the PAGE algorithm which performs one-sample t-test on each gene set in the MSigDB database. The adjusted P-values are used to rank the pathways for each of the first 5 principal components. The pathways are labeled with FDR first, followed by the principal components (PC1, PC2 and so on). PC1 captures the highest fraction of the total variance in the sample set (PC1=31%), whereas the rest of components are sorted in descending order of their contribution to the variance 20% (PC2), 14% (PC3), 11% (PC4) and 8%(PC5).

Next, to deeper assess the correspondence between chromatin texture compartments and transcriptional activity, we further quantitively subdivided the solubility spectrum *in silico*. To this end, the degree of solubility has been ranked according to the 4fSAMMY-seq (S2LvsS3 comparison) negative and positive enrichments at 150 kb resolution. Then, the S2L coverage (positive enrichment of S2LvsS3 comparison) and S3 coverage (negative enrichment of S2LvsS3 comparison) were both divided in quartile ranges.

In prostate control biopsies (CTR28, CTR31, CTR32), we noticed low genes content in the upper quartile of S3 enrichments, and the highest number of genes when the S2L signals is above the third quartile (Fig. 11A). Interestingly, the gene expression level, measured as TPM, increase progressively from the highest S3 quartile to the quartile of highest S2L signals correlating with the

solubility degree (Fig. 11B). Thus, we observed a quantitative association between solubility, as measured by 4fSAMMY-seq, and the level of gene expression.

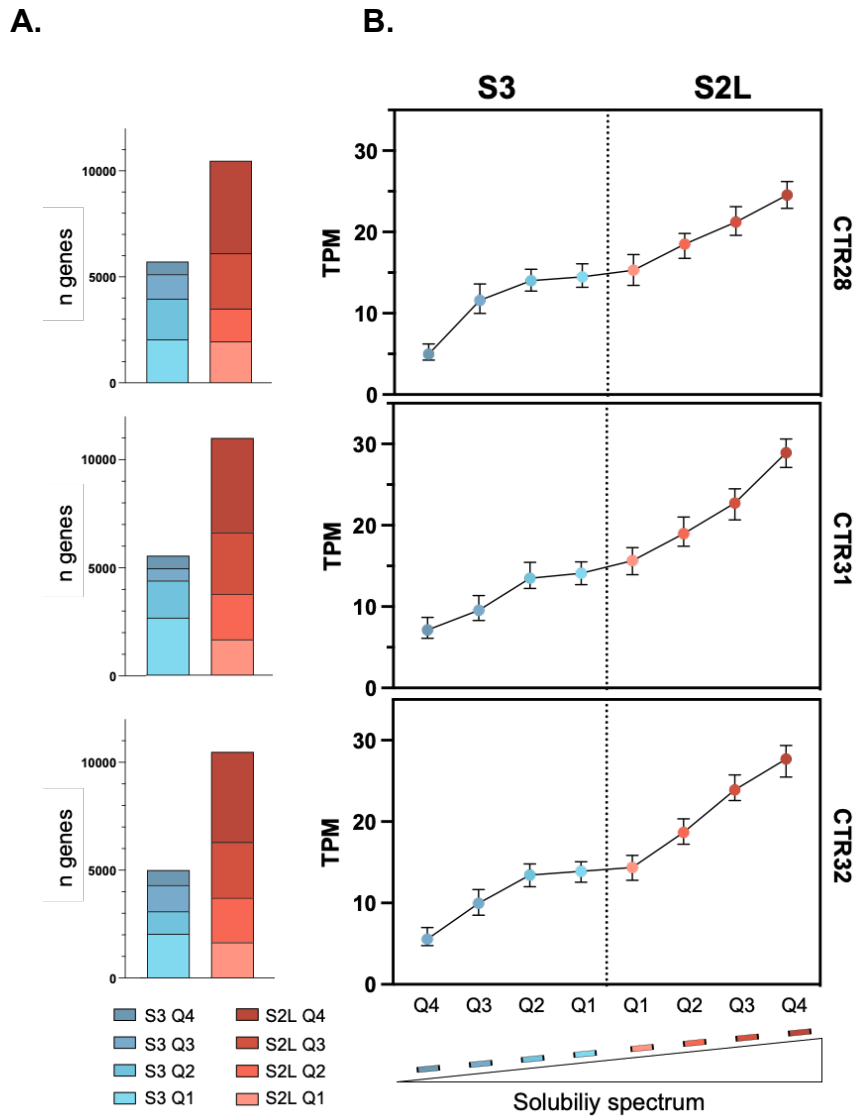


Figure 11 continued on the next page.

Figure 11 continued.

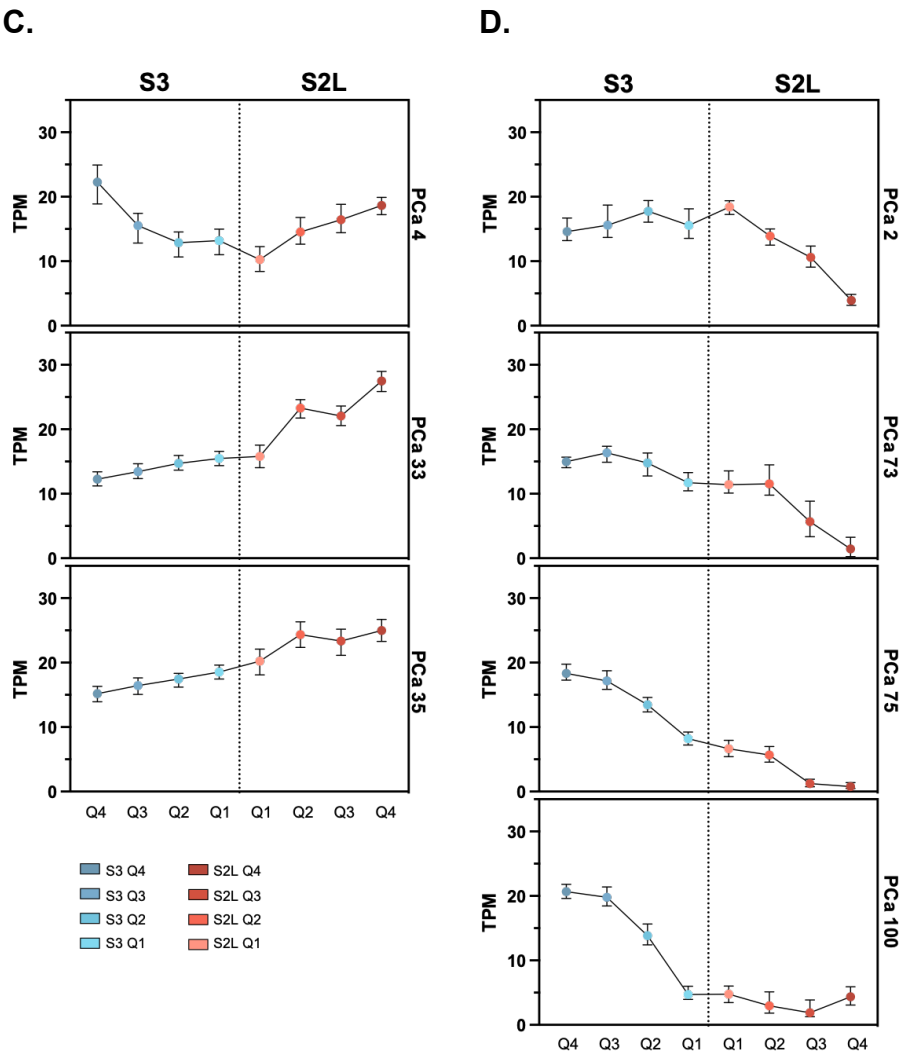


Figure 11 - Correlation between solubility and gene expression.
A. Number of genes for each quartile of S3 signals (shades of blue) and S2L signals (shades of red) in histologically normal tissue (from top to bottom CTR28, CTR 31, CTR32). **B-D.** Plots showing the median TPM value of gene expression (Y-axis) across the distinct quartile of solubility in each sample (X-axis):

histologically normal tissue (B), PCa tissues of subtype 1 (C) and PCa tissues of subtype 2 (D). The line extends the 95% confidence interval for the median labelled as dot.

PCa33 and PCa35 samples, belonging to the control-like subtype 1, maintained functional solubility compartmentalization showing a gradual increase of transcripts expression from the more insoluble quartile (S3-Q4) to the more soluble quartile (S2L-Q4). As expected from previous analyses, PCa4 shows an outlying profile, as in this sample the correlation between genes activity and solubility is lost, with a more pronounced difference in the heterochromatin regions (Fig. 11C). Subtype 2 of PCa tissues exhibit abnormal solubility segregation showing opposite profiles compared with healthy ones, finally losing the correspondence between genome accessibility and transcriptional activity (Fig. 11D).
All together, these data highlight a drastic solubility rewiring of subtype 2 PCa tissues which may reflect specific malignant reprogramming.

3.7 Severe solubility remodeling is associated to ECM plasticity and EMT reprogramming

We next aimed to identify oncogenic programs associated with the severe solubility alterations. To this end, we first assessed primary PCa looking for genomic regions with a deviation of solubility higher than two standard deviations (corresponding to

a confidence level of 95%) from the control mean S2LvsS3 comparison 4fSAMMY-seq value at 150kb resolution. This allowed the preliminary characterization of varying regions at a single PCa sample level: we classified these switching variants as higher solubility variants (HSV), or lower solubility variants (LSV) (Fig. 12). In line with previous results, while subtype 1 PCa samples exhibit limited solubility alterations, subtype 2 PCa biopsies are characterized by elevate levels of locus-specific solubility alterations, which led us to first focus on this subtype. We started our analyses on subtype 2 PCa SVs by looking at regions that are commonly altered in all samples of this subtype. Visual inspection has shown that most HSVs correspond to broad H3K9me3-decorated heterochromatin regions in normal prostate samples (Fig. 12).

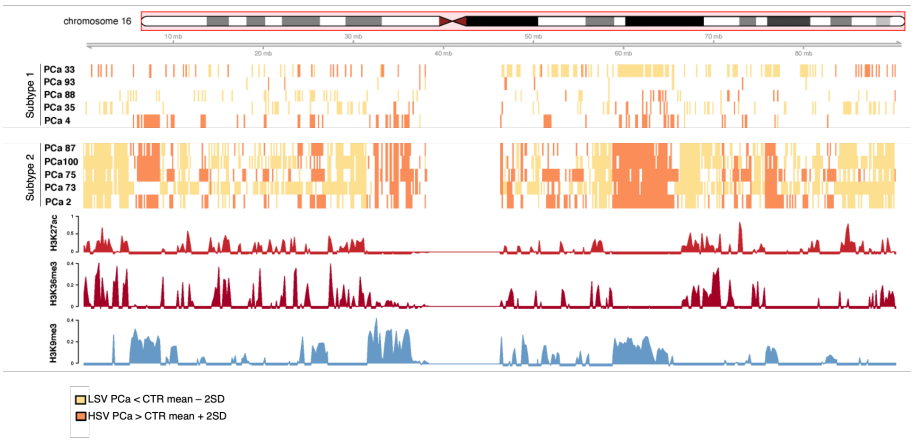


Figure 12 – Recurrent solubility alterations in PCa tissues. Visualization of solubility variants in each cancer tissue of subtype 1 (PCa33, PCa35, PCa4, PCa88, PCa93) and subtype 2 (PCa2, PCa73, PCa75, PCa87, PCa100). Using a cutoff corresponding to 2 standard deviations (SD) in the cancer sample relative to mean of controls, higher solubility variants,

HSV (orange), and lower solubility variants, LSV (yellow), were defined. Publicly available ChIP-seq data from normal prostate gland of euchromatin-associated marks H3K27ac (red) and H3K36me3 (dark red) and of heterochromatin mark (H3K9me3, blue) are also shown.

Next, we analysed the transcriptomic differences between control samples and 4fSAMMY subtype 2 of PCa tissues. Using a FDR q value < 0.01 with DESeq2[20], we identified 706 genes that were up-regulated and 467 genes that were down-regulated in all subtype2 PCa samples (Fig. 13A). We performed an enrichment analysis of DEGs using Metascape[21], to identify pathways which are differentially organized in subtype2 PCa samples. We revealed that downregulated genes were enriched in extracellular matrix (ECM) organization and cell-cell adhesion processes (Fig. 13B), while upregulated genes were primally involved in immune-related signals, such as TNF-α signaling via NFKB, inflammatory responses and Epithelial-Mesenchymal Transition (EMT) (Fig. 13C).

Interestingly, we identified a family of proteins involved in ECM remodeling and tumor migration, such as MMP9 and MMP1, also among the up-regulated DEGs, suggesting a key role of ECM stiffness in controlling drastic malignant chromatin alterations.

Among known EMT-regulators, TWIST2, TRPM7, SERPINE1, and EGR1 were upregulated, while suppressor EMT genes like CDH1, FBLN1, DPYSL3 were downregulated in PCa subtype 2 tissues. Further, we identified both an overexpression of the master EMT inducer, TGF-β1, and an over-representation of its signaling. Pathway analysis also identified a hypoxic tumor

microenvironment which is closely related to tumor EMT process (Fig.13C).

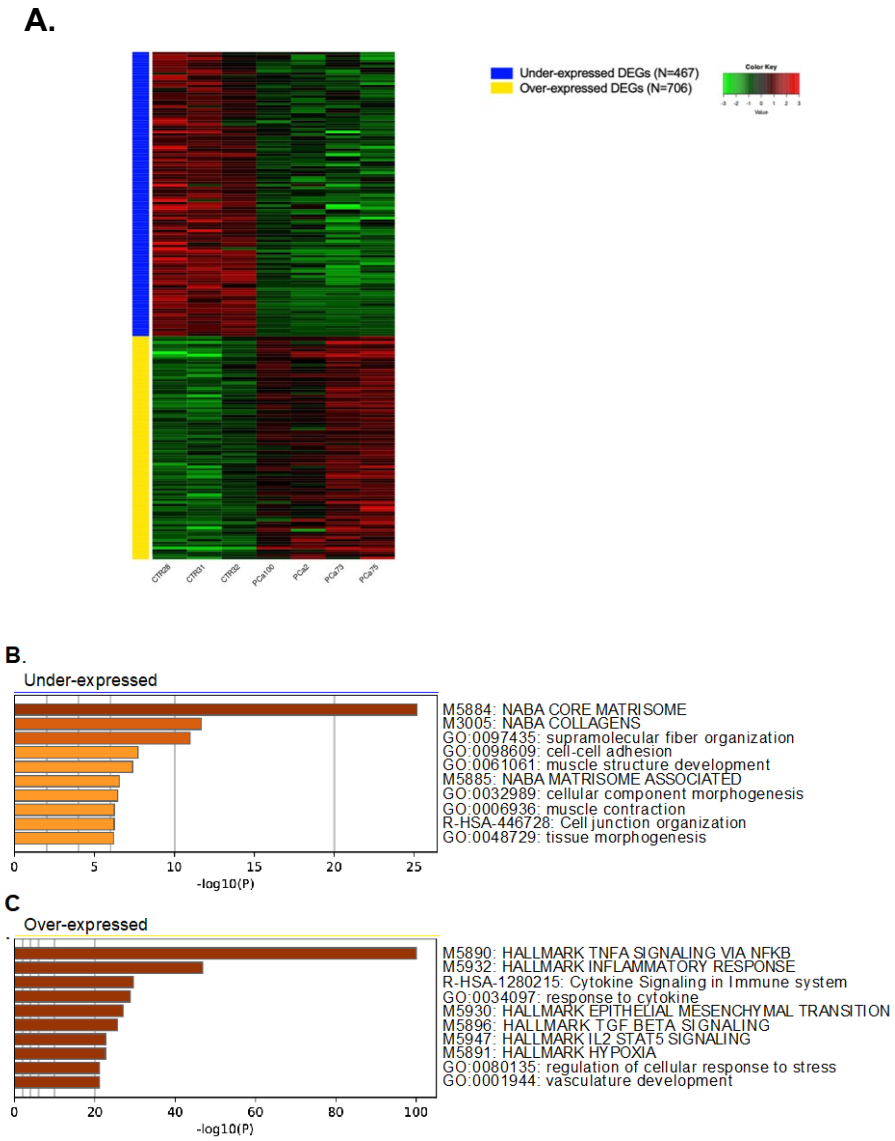


Figure 13 continued on the next page.

Figure 13 continued.

D.

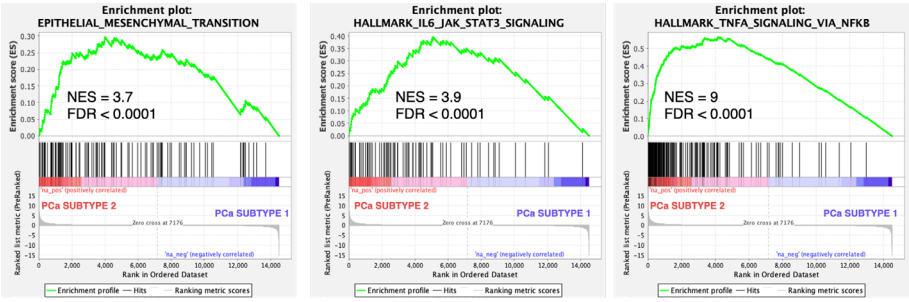


Figure 13 – Transcriptomic analysis of subtype 2 PCa tissues.
A. Heatmap of Differentially-Expressed Gene (DEG) comparing PCa tissues of subtype 2 (PCa2, PCa73, PCa75, PCa100) with control tissues (CTR28, CTR31, CTR32). Blue bar represent significantly down-regulated DEGs (n = 467) and yellow bar represent significantly up-regulated DEGs (n = 706) **B-C**. Enrichment analysis of PCa subtype 2 DEGs using Signatures Database (MSigDB), GO biological process and Canonical Pathways gene sets. Bar plots representing significantly enriched terms in down-regulated genes (B) and up-regulated genes (C). **D.** GSEA enrichment plots of EMTome prostate cancer-specific EMT signatures[22] and established IL6-JAK-STAT3 and TNFα signaling via NF-κB pathways signatures[23] within the expression data of PCa subtype 2 (PCa2, PCa73, PCa75, PCa100) compared to PCa subtype 1 tissues (PCa4, PCa33, PCa35).

On the contrary, PCa tissues of subtype 1 exhibit only 108 common DEGs (46 upregulated and 62 downregulated), compared to normal tissues, which do not correspond to any significantly enriched pathway (data not show). These data collectively showed a transcriptome rewiring of subtype 2 PCa tissues associated to relevant migratory processes. Corroborating this, a gene set enrichment analysis (GSEA) using

prostate cancer-specific EMT gene set validated the enrichment of transcriptional signatures of EMT in PCa tissues of subtype 2 compared to PCa of subtype 1. The hallmark GSEA results indicate that also IL6-JAK-STAT3 signaling and TNFα signaling via NF-κB pathways commonly associated with EMT were significantly enriched in PCa subtype 2 tissues (Fig. 13D).

All together, these data suggest that PCa tissues with a severe solubility remodeling might be associated with transcriptional reprogramming accompanying the migratory propensity.

Materials and methods

Cell cultures

Human primary fibroblast cell line CTRL004 (preputialfibroblast strain #2294, from 4-years old donor) was routinely grown in High Glucose Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 1% glutamax (Gibco,10566-016), 15% fetal bovine serum (Gibco,10270106) and 1% Penicillin-Streptomycin. Cell cultures have been kept under optimal growth conditions: 5% CO_2 at 37°C. CTRL004 was a generous gift from the Laboratory of Molecular and Cell Biology, Istituto Dermopatico dell'Immacolata (IDI)-IRCCS (Rome, Italy).

SAMMY-seq procedure

Chromatin fractionation on CTRL004 (10 thousand to 3 million cells) and tissue-digested single cell suspension was performed with minor adaptations to the protocol described in[1]. Cells were counted, washed in cold PBS and resuspended in cold cytoskeleton buffer CSK: 10 mM PIPES pH 6,8; 100 mM NaCl; 1 mM EGTA; 300 mM Sucrose; 3 mM $MgCl_2$; 1X Protease Inhibitor Cocktail (Roche, 04693116001); 1 mM PMSF (Sigma-Aldrich, 93482) supplemented with 1 mM DTT and 0,5% Triton X-100. After 10 minutes on a wheel at 4°C, samples were centrifugated for 3 minutes at 900g at 4°C and cytoplasmic and nucleoplasmic components were collected as S1 fraction. Pellets were washed for 10 minutes on the wheel at 4°C with an additional volume of the same buffer. Chromatin was then digested by using DNase I (Invitrogen, AM2222) (25U for more than 100 thousand cells and

12.5U for less than 100 thousand cells) in CSK buffer (10 mM PIPES pH 6.8; 100 mM NaCl; 1 mM EGTA; 300 mM Sucrose; 3 mM MgCl2; 1mM PMSF, with protease inhibitors) for 60 min at 37°C. To stop digestion, ammonium sulphate was added to samples to a final concentration of 250 mM and, after 5 min on ice, samples were pelleted at 900g for 3 min at 4°C and the supernatant was collected as S2 fraction. After a wash in CSK buffer, the pellet was further extracted with 2M NaCl in CSK buffer for 10 min at 4°C, centrifuged at 2300 g 3 min at 4°C and the supernatant was conserved as S3 fraction. This treatment removed the majority of histones from chromatin. Pellets were washed twice for 10 minutes on the wheel at 4°C with double volume of high salt CSK buffer. Finally, after 3 minutes of 3000g centrifugation at 4°C, pellets were solubilized in in 8M urea for 10 minutes at RT to denature any remaining protein and to dissolve membranes, and labelled as S4. Factions were stored at – 80°C until DNA extraction.

DNA extraction, library preparation and sequencing

Fractions were diluted 1:2 in TE buffer (10mM TrisHCl pH 8.0, 1 mM EDTA) and incubated with 61,5 U of RNAse cocktail (Ambion, AM2286) at 37° for 90 minutes, followed by 40µg of Proteinase K (Invitrogen, AM2548), at 55° for 150 minutes. Genomic DNA was then isolated using phenol/chloroform (Sigma-Aldrich, 77617) extraction followed by a back extraction of phenol/chloroform with additional volume of TE1X. DNA was precipitated in 3 volumes of cold ethanol, 0.3M sodium acetate and 20ug glycogen (Ambion AM9510) for 1 hour on dry ice or

overnight at -20°. Dry pellets were resuspended in 50 µl (S2) or 15 ul (S3 and S4) of nuclease-free water and incubated at 4°C overnight. On the next day, S2 was further purified using PCR DNA Purification Kit (Qiagen, 28106) and separated using AMPure XP paramagnetic beads (Beckman Coulter, A63880) with the ratio 0,90/0,95 to obtain smaller fragments conserved as S2S (< 300 bp) and larger fragments labelled as S2L (> 300bp) fractions. Both were suspended in 20 ul of nuclease-free water and then reduced to 15ul using a centrifugal vacuum concentrator. S2L, S3 and S4 fractions were sonicated in a Covaris M220 focused-ultrasonicator using screw cap microTUBEs (Covaris, 004078) to obtain a smear of DNA fragments peaking at 150-200 bp (waterbath 20°C, peak power 30.0, duty factor 20.0, cycles/burst 50). Prostate tissues: 150 seconds for S2L and 175 seconds for S3 and S4; Fibroblasts: 125 seconds for S2L and S3, 150 seconds for S4. Fractions were quantified using Qubit 4 fluorometer with Qubit dsDNA HS Assay Kit (Invitrogen, Q32854) and run on an Agilent 2100 Bioanalyzer using High Sensitivity DNA Kit (Agilent, 5067-4626). Libraries were created from each sample using NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB, E7645L) and Unique Dual Index NEBNextMultiplex Oligos for Illumina (NEB, E6440S); libraries were then qualitatively and quantitatively checked on Bioanalyzer 2100. Libraries with distinct adapter indexes were then multiplexed and, after cluster generation on FlowCell, sequenced for 50 bases in paired-ends mode on an IlluminaNovaSeq 6000 instrument at the IEO Genomic Unit in

Milan. A sequencing depth of at least 35 million reads was obtained for each sample.

Tissues Cohort

Our analyses involved cohorts of chemo-naïve patients followed-up in the Urology's Division of Fondazione IRCCS Ca' Granda-Ospedale Maggiore Policlinico (Milan) who underwent the transrectal ultrasound-guided systematic sampling of the prostate (TRUS biopsy) in relation to abnormal digital rectal examination or an elevated PSA. Fourteen multiregional cores TRUS prostate biopsy were collected from all enrolled patients for the diagnosis and one more core was taken for this study from the same surgery. The institutional ethics committee board authorized this study (authorization n 1063). All specimens were obtained after the patients had provided written informed consent following the ethical principles of biomedical researches on biospecimens. To ensure optimal tissue recovery in terms of quality, fresh surgical prostate biopsy specimens were placed in ice-cold saline buffer and then directly transported to the laboratory within 1 hour from sample collection. Tissues were then immediately processed to preserve the intranuclear genomic and protein architectures of prostatic nuclei. Samples selected for NGS consist of 18 fresh biopsies divided on the bases of the histology and the spatial distribution of the positive cores in two different groups: 10 PCa biopsies with histologically confirmed prostate cancer and 8 histologically negative biopsies from patients who had no cancer in any biopsy core. The enrolled

patients were men aged 62-87, with a median age of 72 years. All the clinical data were provided by the Urology division maintaining the patient confidentiality.

Tissues processing

The biopsy specimen was stored at 4-8°C for maximum 6 hours before dissociation, avoiding its freezing. For the digestion step, given the little size of the needle biopsy tissues (typically 20-30 mg) we determined the dissociation condition to ensure optimal cell populations recovery. Briefly, tissue was cut into small pieces (~1 mm) in ice-cold PBS with autoclaved surgical scissors and resuspended in 1 ml HBSS (Gibco, 14025) containing 200 units of collagenase type I (Life Technologies, 17018-029) plus 67 µg DNase I (Sigma-Aldrich, 10104159001) each ~10 mg of minced tissue. Digestion was carried out in a water bath at 37 °C shaking vigorously for 10 sec every 5 minutes; at its completion (usually after 1 hour of digestion) single cells were washed once by topping up to 2 ml with RPMI + 10% FBS and centrifugated at 300g. The cell pellet was resuspended in RPMI + 10% FBS and dispersed by passing through a 75µm cells strainer, followed by an additional wash of the filter with RPMI + 10% FBS. Finally, the cells were centrifugated and resuspended in 1 ml of ice-cold PBS for counting with a hemacytometer. The entire procedure of digestion takes approximately 3 hours.

Histological Evaluation

One third portion of each tissue specimen was embedded in Killik (Bio-Optica, 05-9801), immediately frozen in precooled isopentane (MilliporeSigma, 277258) and stored at -80°. OCT-embedded biopsy cores (one per patient) were serially sectioned, at 10 µm thick in a cryostat at -20°C. Ten glasses per patient, containing multiple sections representing distinct regions of the same tissue were prepared in parallel and stored at -80°, along with a corresponding hematoxylin and eosine stained slide. Hematoxylin and Eosin staining was performed using H&E Staining Kit (ab245880). The hematoxylin-eosin-stained slides were reviewed by an expert genitourinary pathologist to assign a Gleason score according to the International Society of Urinary Pathology grading system. The histopathological architectural pattern was also confirmed by comparing the Gleason score between our prostate biopsy tissue sample and the nearby biopsy used for clinical diagnosis of the patients. All the clinical data were provided by the Urology division maintaining the patient confidentiality.

Flow cytometry

To quantify the relative frequency of cell types for each biopsy, ten thousand cells form the digestion step were stained, acquired on a BD FACSCanto™ Flow Cytometer and analyzed with FlowJo software in the INGM FACS Facility. To avoid unspecific binding, antibodies were incubated with PBS-BSA 1% for 30 minutes at 4°C. TO-PRO®-3 stain is used to assess cell viability,

tissue resident leukocytes are identified by $CD45^+/CD326^-$, epithelia are identified by $CD326^+/CD45^-$, and stroma are double negative ($CD326^-/CD45^-$). The followed table displays information on antibodies used for flow cytometry.

Specificity	Clone	Fluorochrome	Identifier	Specificity
TO-PRO®-3	-	APC	Thermofisher, T3605	Viability
CD326	EBA-1	FITC	BD, 347197	Epithelia
CD45	H130	PB	Biolegend, 982306	Leukocyte

RNA extraction, library preparation and sequencing

Ten thousand cells form the digestion step were stabilized in 200µl of 1Thioglycerol/Homogenization Solution of the Maxwell® RSC miRNA Tissue Kit (Promega, AS1460) and stored frozen at -80°C for later total RNA automated purification using Maxwell® RSC 48 Instrument (Promega, AS8500) according to manufacturer's instructions. The automated purification on the Maxwell® RSC Instrument ensures the higher RNA yield from this low number of cells. Total RNA was quantified by Qubit 4 fluorometer with Qubit RNA HS Assay Kit (Invitrogen, Q32852) and assessed by Agilent 2100 Bioanalyzer using Agilent RNA 6000 Pico Kit (Agilent, 5067-1513) to inspect RNA integration. For each sample, 1 ng of total RNA was used to construct strand specific RNAseq library with SMARTer Stranded Total RNA-Seq Kit - Pico Input (Takara, 634487). The yield and quality of the libraries were evaluated on Agilent 2100 Bioanalyzer using High

Sensitivity DNA Kit (Agilent, 5067-4626). RNAseq libraries were sequenced on the Illumina NextSeq™ 550 system at the sequencing facilities of Humanitas or Division of Pathology of Fondazione IRCCS Ca' Granda-Ospedale Maggiore Policlinic in Milan to a minimal target of 40 million for 75 bases in paired-ends mode.

Literature data (fibroblasts)

We collected publicly available datasets from the following sources: Lamin A/C ChIP-seq[24] (SRR605493, SRR605494, SRR605495 and SRR605496), Lamin B1 ChIP-seq[25] (SRR2119331, SRR2119332, SRR2119335 and SRR2119336) and H3K9me3, H3K4me1, H3K36me3, H3K27ac, H3K4me3 from Roadmap Epigenomics (fibroblast sample: E055). These data have been downloaded as raw data and analyzed as described in the next sections.

Literature data (prostate)

The ChIP-seq data of prostate gland have been downloaded from the following datasets available on ENCODE[26]: ENCSR763IDK (H3K27ac), ENCSR768PFZ (H3K36me3) and ENCSR133QBG (H3K9me3) (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE143079). These data have been downloaded as raw data and analyzed as described in the next sections. The DNase-seq and the ATAC-seq data have been downloaded from the following datasets available on ENCODE[26]: ENCSR999NKW (ATAC-seq) and ENCSR564FZH (DNase-seq). These data have been used

for this work in already processed format (read depth normalized signal).

SAMMY-seq sequencing read analysis

Sequencing reads were quality controlled with fastqc (version 0.11.3;http://www.bioinformatics.babraham.ac.uk/projects/fastqc) and trimmed using Trimmomatic[27] (v0.39) with the following parameters for SAMMY-seq and literature data: 2 for seed_mismatch, 30 for palindrome_threshold, 10 for simple_threshold, 3 for leading, 3 for trailing and 4:15 for slidingwindow. The sequence minimum length threshold of 35 has been applied for all data. As clip file has been used the trimmomatic provided dataset “TruSeq3-SE.fa” (for single end) and “TruSeq3-PE-2.fa” (for paired end). After trimmed reads were aligned using BWA[28] (v0.7.17-r1188) setting –k parameter as 2 and using as reference genome the UCSC hg38[29]. The alignment duplicates were marked with Picard (v2.22) (https://github.com/broadinstitute/picard) MarkDuplicates option and then filtered using Samtools[30] (v1.9). We further filtered all the reads with mapping quality lower than 1. Each lane was analyzed separately and then merged at the end of the process. From the alignment of output a coverage analysis was performed using Deeptools[31] (version 3.4.3) bamCoverage function. For this analysis the genome was binned at 50bp, the reads extended up to 250 bp and RPKM normalization method has been used. Human genome size was considered of 2701495761 bp (value suggested in the Deeptools manualhttps://deeptools.readthedocs.io/en/latest/content/featur

e/effectiveGenomeSize.html), from the analysis were excluded regions known to be problematics in term of sequencing (the list have been downloaded from the ENCODE portal, https://www.encodeproject.org/files/ENCFF356LFX/@@download/ENCFF356LFX.bed.gz). The comparison between datasets was performed using the SPP[32] (v1.16.0) R (v3.5.2) library. The reads were imported from the bam files using the "read.bam.tags" function, then they were filtered using "remove.local.tag.anomalies" and finally the comparison were performed using the function "get.smoothed.enrichment.mle" setting "tag.shift = 0" and "background.density.scaling = TRUE The analysis were performed with the same paramenters for all the datasets.

Correlation analysis

Genome-wide Spearman correlations between SAMMY-seq samples and all the public dataset was conducted using "cor" R (v3.5.2) base function with "method = Spearman". The comparisons and the genome coverage tracks (DeepTools output, see section "SAMMY-seq sequencing read analysis") were imported in the R session using the function "import" of the rtracklayer (v1.42.2) library[33]. Then the files were rebinned using the function "tileGenome" and the correlation was performed per chromosome across the two comparisons in analysis. The analysis was performed identically for all the datasets. The data have been plotted using the R library "ggplot2" (v3.3.5) for

fibroblasts while for prostate the R library “gplot” have been used (in particular the function “heatmap.2”).

Track visual representation

The visual representation of tracks was performed thanks to the Gviz R library[34] (v1.26.5). The track profile was calculated using the function “DataTrack” (the input file was imported using the function “import” of the rtracklayer library) and plotted using the function “plotTracks” setting the value “window = 1000”. Line plots were drawn setting the parameter type as ‘a’ and overlayed using the function “OverlayTrack”; instead, mountain plots were obtained setting the parameter type as "polygon". Extra elements of these plot as chromosome ideogram (on top) and genome axis were plotted respectively using the functions “IdeogramTrack” and “GenomeAxisTrack”. The analysis were performed identically for all the datasets.

Cancer fingerprint

Data were imported in the R session from comparisons in the same way used for correlation analysis (see “Correlation analysis” section) and the value of each bin were reported on the plot. Consensus values per bin were obtained performing the arithmetic mean among the values of the control samples per each bin. The data were plotted using the R library “ggplot2” (v3.3.5).

Switching cancer variants

To evaluate which genomic bin is more different between a cancer patient and the control consensus value, the data were imported and rebinned as for the correlation analysis (see “Correlation analysis” section”). Then the consensus for each bin across controls were calculated as the same for the cancer fingerprint plot (see “Cancer fingerprint” section). Finally, each bin in each tumor were compared with the correspondent bin in the control consensus: the bin has been marked as: positively enriched if its value higher than the one in control consensus more 2 standard deviations and labeled as “higher solubility variant”, negatively enriched if its value lower than the one in control consensus less 2 standard deviations and labeled as “lower solubility variant”, not enriched in all the other cases. The switching bins have been plot using the R library Gviz using the function ”DataTrack” setting “type = heatmap”.

RNA-seq sequencing read analysis

The overall quality of the sequenced reads was assessed using FastQC tool (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). (Version 0.11.8), then reads were trimmed with Trimmomatic software[27] (version 0.39) removing the adapters (ILLUMINACLIP:Picov2smart-PE.fa), primer dimers, and low quality bases at the beginning and at the end of the reads (trimmomatic PE phred33 LEADING:3 TRAILING:3 SLIDINGWINDOWD:4:15 MINLEN:36). STAR[35](V. 2.7.0f_0328)

has been used to index (STAR --runMode genomeGenerate) the Human Genome (GENCODE Release 39, GRCh38 primary assembly genome[29]) and to align sequenced reads in Paired-End mode (--readFiles R1.FASTQ R2.FASTQ) on the previous indexed reference. Multimapping reads and PCR duplicate unique mappers were marked in the final output (--bamRemoveDuplicatesType Unique) and unaligned reads stored in a different file (--outReadsUnmapped Fastx). The reads counts that fall on genes were calculated using as a reference a GTF file with RefSeq annotation downloaded from UCSC (https://hgdownload.soe.ucsc.edu/goldenPath/hg38/bigZips/genes/hg38.ncbiRefSeq.gtf.gz). This file was further processed to remove non canonical and mitochondrial chromosome, selected only curated genes (NM, NR) and finally splitted in protein coding (NM) and noncoding (NR) files. Reads count was carried out with HTSeq-count (V. 0.13.5) on bam files (previously generated by STAR) using as a feature the union of all exons in a gene. The kind of library were specified with "-s reverse" parameter. The reads that align to more than position in reference were discarded (htseq-count --non-unique none). The full matrix with raw count reads for each sample were loaded in R 3.6.1 and normalized using DESeq2's[20] median of ratios. Differential expression analyses were performed with DESeq2[36] (V. 1.26) using Wald test and the Benjamini and Hochberg method, (correction for multiple testing) to compute p-values and p-values adjusted respectively.

TPM (Transcript Per Million) normalization was calculated for each sample in R 3.6.1 environment by applying the following steps: reads count for each gene have been divided by the length of the gene in kilobases (RPK, Reads per kilobase); Sum all RPK values in a sample and divided by 106 obtaining the scaling factor; Divided RPK of each individual gene by the scaling factor. The solubility fractionation in silico was performed starting from SAMMY-seq comparisons files (e.g S3/S2L), rebinning them at 150Kb resolution and taking the mean signal values for each window (V 3.5.0; multiBigwigSummary bins -bs 150000 --outRawCounts). The genomic windows with positive values and negative values (as absolute value) were then stored in two different files. For each file, the quantile function in R were used to classify the coverage of genomic regions in four quartiles according with coverage distribution. The genes within different genomic regions quartiles have been selected by intersecting the specific genomic coordinates with the RefSeq GTF reference (V.2.26.0; bedtools intersect; 1 bp is the minimum overlap required; different genomic regions quartiles may intersect different part of the same gene, in this case both are considered valid)

Gene Set Enrichment Analysis (GSEA)

The specifically expressed gene list was ranked according to log2 fold change and p-value adjusted (-log10(p-adjusted) * log2fc) and compared with Molecular Signatures Database (MSigDB) hallmark gene set collection using the following

parameters: Number permutations: 1000; Collapse: No; Enrichment Statistic: classic; Max size: 500; Min size:15.

References to chapter 3

1. Sebestyén, E. *et al.* SAMMY-seq reveals early alteration of heterochromatin and deregulation of bivalent genes in Hutchinson-Gilford Progeria Syndrome. *Nat. Commun.* **11**, 6274 (2020).

2. Siegel, R. L., Miller, K. D., Fuchs, H. E. & Jemal, A. Cancer Statistics, 2021. *CA. Cancer J. Clin.* **71**, 7–33 (2021).

3. Berger, M. F. *et al.* The genomic complexity of primary human prostate cancer. *Nature* **470**, 214–220 (2011).

4. Barbieri, C. E. *et al.* Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. *Nat. Genet.* **44**, 685–689 (2012).

5. Cancer Genome Atlas Research Network. The Molecular Taxonomy of Primary Prostate Cancer. *Cell* **163**, 1011–1025 (2015).

6. Fraser, M. *et al.* Genomic hallmarks of localized, non-indolent prostate cancer. *Nature* **541**, 359–364 (2017).

7. Viswanathan, S. R. *et al.* Structural Alterations Driving Castration-Resistant Prostate Cancer Revealed by Linked-Read Genome Sequencing. *Cell* **174**, 433-447.e19 (2018).

8. Pomerantz, M. M. *et al.* Prostate cancer reactivates

developmental epigenomic programs during metastatic progression. *Nat. Genet.* **52**, 790–799 (2020).

9. Yuan, H. *et al.* SETD2 Restricts Prostate Cancer Metastasis by Integrating EZH2 and AMPK Signaling Pathways. *Cancer Cell* **38**, 350-365.e7 (2020).

10. Labbé, D. P. *et al.* TOP2A and EZH2 Provide Early Detection of an Aggressive Prostate Cancer Subgroup. *Clin. Cancer Res.* **23**, 7072–7083 (2017).

11. Taberlay, P. C. *et al.* Three-dimensional disorganization of the cancer genome occurs coincident with long-range genetic and epigenetic alterations. *Genome Res.* **26**, 719–731 (2016).

12. Rhie, S. K. *et al.* A high-resolution 3D epigenomic map reveals insights into the creation of the prostate cancer transcriptome. *Nat. Commun.* **10**, 4154 (2019).

13. Egevad, L., Delahunt, B., Srigley, J. R. & Samaratunga, H. International Society of Urological Pathology (ISUP) grading of prostate cancer - An ISUP consensus on contemporary grading. *APMIS Acta Pathol. Microbiol. Immunol. Scand.* **124**, 433–435 (2016).

14. Livi, C. *et al. A ChIC solution for ChIP-seq quality assessment*. (2020). doi:10.1101/2020.05.19.103887.

15. Henry, G. H. *et al.* A Cellular Anatomy of the Normal Adult Human Prostate and Prostatic Urethra. *Cell Rep.* **25**, 3530-3542.e5 (2018).

16. Henry, G. H., Loof, N. & Strand, D. W. OMIP-040: Optimized gating of human prostate cellular

subpopulations. *Cytometry A* **91**, 1147–1149 (2017).

17. Haas, B. J. *et al. STAR-Fusion: Fast and Accurate Fusion Transcript Detection from RNA-Seq*. 120295 https://www.biorxiv.org/content/10.1101/120295v1 (2017)

18. Kumar-Sinha, C., Tomlins, S. A. & Chinnaiyan, A. M. Recurrent gene fusions in prostate cancer. *Nat. Rev. Cancer* **8**, 497–511 (2008).

19. Ge, S. X., Son, E. W. & Yao, R. iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data. *BMC Bioinformatics* **19**, 534 (2018).

20. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).

21. Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).

22. Vasaikar, S. V. *et al.* EMTome: a resource for pan-cancer analysis of epithelial-mesenchymal transition genes and signatures. *Br. J. Cancer* **124**, 259–269 (2021).

23. Liberzon, A. *et al.* The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).

24. McCord, R. P. *et al.* Correlated alterations in genome organization, histone methylation, and DNA-lamin A/C interactions in Hutchinson-Gilford progeria syndrome. *Genome Res.* **23**, 260–269 (2013).

25. Dou, Z. *et al.* Autophagy mediates degradation of nuclear

lamina. *Nature* **527**, 105–109 (2015).

26. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).

27. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinforma. Oxf. Engl.* **30**, 2114–2120 (2014).

28. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics **25**, 1754–1760 (2009).

29. Schneider, V. A. *et al.* Evaluation of GRCh38 and de novo haploid genome assemblies demonstrates the enduring quality of the reference assembly. *Genome Res* **27**, 849–864 (2017).

30. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinforma. Oxf. Engl.* **25**, 2078–2079 (2009).

31. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160-165 (2016).

32. Kharchenko, P. V., Tolstorukov, M. Y. & Park, P. J. Design and analysis of ChIP-seq experiments for DNA-binding proteins. *Nat. Biotechnol.* **26**, 1351–1359 (2008).

33. Lawrence, M., Gentleman, R. & Carey, V. rtracklayer: an R package for interfacing with genome browsers. *Bioinformatics* **25**, 1841–1842 (2009).

34. Hahne, F. & Ivanek, R. Visualizing Genomic Data Using Gviz and Bioconductor. *Methods Mol. Biol. Clifton NJ*

1418, 335–351 (2016).

35. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinforma. Oxf. Engl.* **29**, 15–21 (2013).

36. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

CHAPTER 4.

FINAL CONSIDERATIONS

4.1. Summary and conclusions

Chromatin architecture inside the nuclei of eukaryotic cells follows tight rules to fulfil its pivotal roles in both transcriptional regulation and maintenance of genomic stability[1]. Thus, a correct three-dimensional organization of interphase nuclei is a precondition for a correct genome usage and function, while its reorganization is often associated with pathological states, as highlighted by laminopathies[2], or severe alteration of cell fate, as in cancer[3]. The aim of our research group is the study of chromatin architecture reshaping during pathology insurgency within the context of lamin-dependent muscular dystrophy, aging and cancer. During my PhD, I worked in two pathological contexts, namely Emery Dreifuss Muscular Dystrophy and Prostate Cancer. I also contributed to the development of an improved high-throughput technology to monitor the chromatin organization.

In our first project (Chapter 2), we have analyzed the dynamics of muscle stem cells (MuSCs) population of a murine model of Emery-Dreifuss muscular dystrophy (EDMD)[4]. MuSCs cell fate determination requires a sequential restriction of transcriptional programs to acquire cell-specific competences[5]. The classical view of EDMD relies on nuclear fragility as the leading mechanism promoting insurgence of the pathology[6]. We previously showed a tight interplay between Lamin A/C and Polycomb group (PcG) of proteins, epigenetic repressors involved in the regulation of developmental genes[7]. Thus we

investigated the Lamin A/C-PcG axis in EDMD. We found that in the absence of Lamin A/C, PcG diffuse from their targets over the flanking genomic regions, leading to a de-repression of bivalent genes, key players in stemness preservation and cell identity determination[4]. The chromatin aberrant PcG displacement leads to a co-expression of distinct lineage-specific genes, as muscular and adipogenic markers, which culminates in an impaired self-renewal capacity and a dystrophic phenotype. In a second work (Chapter 3.1), we presented the advances of our previously published sequencing-based technology[8]. The new version of the technique, named 4fraction SAMMY-seq (4fSAMMY-seq), expanded our perspective on the whole spectrum of chromatin solubility and accessibility (Manuscript in preparations). Over the past decade, the advent of experimental techniques based on high-throughput sequencing (NGS) permitted important breakthroughs of mechanisms underlying 3D genome organization[9]. Most of these methods are specifically designed to characterize active regions, such as DNase-seq[10], ATAC-seq[11] and FAIRE-seq[12]. Few technologies were developed with the aim to capture the heterochromatin regions, probably for the lack of transcriptional activity. DamID-seq[13] and Protect-seq[14] are example of heterochromatin capturing techniques. However, the possible artefacts and biases caused by the exogenous gene expression, as in DamID-seq, or by the chemical modification of chromatin, as in Protect-seq, remains poorly considered.

The first publicly available version of our protocol, 3fSAMMY-seq, probably represents the simplest method for the identification of heterochromatic regions tightly associated to Nuclear Lamina (NL) in primary cells[8]. This technology is indeed robust, fast, easy-to-perform and avoids formaldehyde crosslinking or antibodies used in most other methods to explore chromatin architecture. However, 3fSAMMY-seq comes with some limitations, as it is exclusively exploitable to map the heterochromatin regions, not recovering accessible regions.

We modified the experimental setup of the protocol to develop the 4 fractions SAMMY-seq (4fSAMMY-seq), which further enhances the analytical potential of our technology.

The specificity of the signal in the soluble S2 fraction allows the identification of both active euchromatin regions in the track peaks and lamina-associated heterochromatic regions in the track valleys (Figure 1). Thus, our novel 4fSAMMY-seq technology represents a significant advancement in the field of nuclear architecture characterization as it provides the isolation of both euchromatin and heterochromatin in a single experimental protocol on primary cells. 4fSAMMY-seq is also suitable for application on low amounts of cells (10K) (Figure 2), a key advantage for analysis on limited materials as human biopsy tissues derived from clinical practice.

In my PhD project, I applied 4fSAMMY-seq on prostate biopsies to study chromatin architecture changes in primary prostate carcinoma (PCa) (Chapter 3.2). PCa is the second most common cause of male cancer-related mortality[15], due to its highly

variable clinical course[16] and frequent multifocality[17]. To date histopathology analysis of prostate biopsy tissues remains the primary basis for PCa diagnosing and nuclear atypia is taken into consideration to assess the tumor grade[18]. Of note, among the parameters of nuclear aberrations, the chromatin texture and the heterochromatin nuclear localization are carefully evaluated. However, the clinical and molecular heterogeneity among primary PCa and the subjective assessments by pathologists, render the histopathology-score grading not able enough to be a good predictor of PCa outcome[19]. To improve prognostication and to guide more precisely therapeutic strategies, studies on primary prostate cancer tumors are rapidly growing. Increasing evidence indicate that alterations in chromatin architecture and concomitant epigenetic reprogramming play a pivotal role in PCa[20]. Several data support the essential role of heterochromatin for the maintenance of genome stability[21]. Indeed, the global loss of H3K9 di- and tri-methylation culminates in higher rate of tumorigenesis in mouse models[22]. Heterochromatin deregulation is also associated to cancer cells phenotypic plasticity[23]. Recent evidence demonstrated that the majority of H3K9me3-enriched TADs changes epigenetic status in metastatic prostate cancer cells compared to normal cells, depicting more defective organization in heterochromatic regions[24]. However, this study adopted prostate cancer cell lines derived from metastatic and no data in primary tissue were reported. Moreover, the functional consequences of misregulation of heterochromatic blocks remain poorly defined.

Hence, reliable methods for comprehensive analyses of prostate cancer chromatin remodeling, including heterochromatin, using limited material as biopsy tissues represent an important advantage for biomedical scientific community.

We are interested in understanding the role of chromatin structure changes in primary prostate cancer that may provide useful insights in the prognostic evaluation.

We successfully set up our 4fSAMMY-seq protocol on prostate biopsies (between 10-80K cells after tissue digestion) to capture early chromatin remodelling events in primary prostate cancer. We found that 4fSAMMY-seq technology is capable to identify a reliable and conserved pattern of euchromatin/heterochromatin regions on bulk prostate control tissues (Figure 4), containing a heterogeneous mix of epithelial cells, stromal cells and leukocytes. This result confirms and further corroborates previous studies showing that chromatin compartments are highly conserved across cell types[25,26].

On the contrary, all tissues containing cancer cells exhibit a different picture, with tumor-specific chromatin remodeling (Figure 8). Noteworthy, integration between 4fSAMMY-seq and RNAseq indicates that cancer tissues show aberrant correlation between genomic solubility properties and transcriptional rates (Figure 11).

Our 4fSAMMY-seq technology further allows the clusterization of tumors into two distinct subtypes exhibiting a different extent of chromatin architecture rearrangements (Figure 9). Interestingly, we did not observe an association between different degree of

solubility alterations and tumor grade measured with Gleason score, suggesting that chromatin remodelling does not match the histopathological features. Preliminary assessment of regions recurrently altered among tumor tissues of subtype 2, identified normal H3K9me3-decorated regions (Figure 12), suggesting abnormal confining of heterochromatic territories (Figure 9). We started our analyses on subtype 2 PCa solubility variations by looking at regions that are commonly altered in all samples of this subtype, which allowed us to identify 388 genes in higher solubility variants and in 3304 genes in lower solubility variants (data not shown). Of note, the genes contained in the altered solubility regions showed little overlapping with differential expression genes (data not shown), suggesting that an aberrant solubility does not directly affect gene expressions. This is in line with our previous study, showing that changes in chromatin solubility are not accompanied by transcriptional changes of genes residing in the same genomic regions, rather represents a pivotal trigger event that functionally affect PcG epigenetic repression[8]. Transcriptome analyses of PCa subtype 2 tissues shows a deregulation of Extracellular Matrix (ECM) organization and Epithelial-Mesenchymal-Transition (EMT) process suggesting a migratory propensity. Accordingly, the enrichment analysis identified further key pathways regulating EMT such as TNF-α signaling via NFKB[27], hypoxia[28], IL6-JAK-STAT3[29] and, above all, TGFβ-pathway[30], (Figure 13). This is in agreement with recent findings that shown large-scale histone modifications

alterations[31] and severely altered nuclear structure[32] in cells undergoing TGFβ-induced EMT.

Overall, our results suggest that drastic chromatin remodeling may provide relevant information in the triggering of the invasive/migratory cascade and evolutionary advantage. Further analysis will clarify the epigenetic mechanisms involved.

4.2 Future perspectives

Emerging evidence indicated that EMT is not “all-or-none” processes but it is a gradual transformation of plastic states known as hybrid/partial/intermediate EMT[33]. Hybrid EMT is a multi-faceted process that exhibits higher metastatic risk than complete EMT[34]. Taken together, our results induce us to hypothesize that heterochromatin instability in primary prostate cancer may underlies the phenotypic transient state of hybrid EMT. Indeed. we highlighted a possible connection between solubility aberrations and cancer invasiveness capacity: preliminary integration analysis suggests that PCa biopsies exhibiting a severe solubility remodeling may be epigenetically reprogrammed to undergo the Epithelial-Mesenchymal Transition and, eventually, to develop invasive phenotype.

On the bases of these findings, we are first planning to sequence the remaining RNA-seq dataset, to extend our results to all cohort. Then, we will deeply analyze the hierarchy of deregulated genes to eventually find primary EMT players in the domains

more affected by solubility changes. We will also dissect the role of chromatin remodeling in the epigenetic control of transcription, examining the chromatin states of transcriptionally deregulated regions. This information may finally reveal the functional effect of chromatin remodeling on a subset of specific epigenetic regulators. Finally, we will investigate the effect of TGF-β-induced EMT on chromatin remodeling using both prostate cancer cell lines and primary tissues from healthy patients.

Overall, our studies shed light on the importance of chromatin remodeling in primary prostate cancer and with a future extension and validation on a bigger cancer sample size may provide useful insights in the diagnosis and prognosis assessments.

References to chapter 4

1. Sewitz, S. A., Fahmi, Z. & Lipkow, K. Higher order assembly: folding the chromosome. *Curr Opin Struct Biol* **42**, 162–168 (2017).

2. Ho, R. & Hegele, R. A. Complex effects of laminopathy mutations on nuclear structure and function. *Clin Genet* **95**, 199–209 (2019).

3. Feng, Y. & Pauklin, S. Revisiting 3D chromatin architecture in cancer development and progression. *Nucleic Acids Research* **48**, 10632–10647 (2020).

4. Bianchi, A. *et al.* Dysfunctional polycomb transcriptional repression contributes to lamin A/C–dependent muscular dystrophy. *J Clin Invest* **130**, 2408–2421 (2020)

5. Dilworth, Fj. & Blais, A. Epigenetic regulation of satellite cell activation during muscle regeneration. *Stem Cell Research & Therapy* **2**, 18 (2011).

6. Zaremba-Czogalla, M., Dubińska-Magiera, M. & Rzepecki, R. Laminopathies: the molecular background of the disease and the prospects for its treatment. *Cell Mol Biol Lett* **16**, 114–148 (2011).

7. Cesarini, E. *et al.* Lamin A/C sustains PcG protein architecture, maintaining transcriptional repression at target genes. *J Cell Biol* **211**, 533–551 (2015).

8. Sebestyén, E. *et al.* SAMMY-seq reveals early alteration of heterochromatin and deregulation of bivalent genes in Hutchinson-Gilford Progeria Syndrome. *Nat Commun* **11**,

6274 (2020).

9. Dirks, R. A. M., Stunnenberg, H. G. & Marks, H. Genome-wide epigenomic profiling for biomarker discovery. *Clin Epigenetics* **8**, 122 (2016).

10. Boyle, A. P. *et al.* High-Resolution Mapping and Characterization of Open Chromatin across the Genome. *Cell* **132**, 311–322 (2008).

11. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* **10**, 1213–1218 (2013).

12. Giresi, P. G., Kim, J., McDaniell, R. M., Iyer, V. R. & Lieb, J. D. FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. *Genome Res* **17**, 877–885 (2007).

13. van Steensel, B. & Henikoff, S. Identification of in vivo DNA targets of chromatin proteins using tethered dam methyltransferase. *Nat Biotechnol* **18**, 424–428 (2000).

14. Spracklin, G. & Pradhan, S. Protect-seq: genome-wide profiling of nuclease inaccessible domains reveals physical properties of chromatin. *Nucleic Acids Res* **48**, e16 (2020).

15. Siegel, R. L., Miller, K. D., Fuchs, H. E. & Jemal, A. Cancer Statistics, 2021. *CA: A Cancer Journal for Clinicians* **71**, 7–33 (2021).

16. Segura-Moreno, Y. Y., Sanabria-Salas, M. C., Varela, R., Mesa, J. A. & Serrano, M. L. Decoding the heterogeneous

landscape in the development prostate cancer (Review). *Oncology Letters* **21**, 1–13 (2021).

17. Djavan, B. *et al.* Predictability and significance of multifocal prostate cancer in the radical prostatectomy specimen. *Tech Urol* **5**, 139–142 (1999).

18. Fischer, A. H. *et al.* The cytologic criteria of malignancy. *J Cell Biochem* **110**, 795–811 (2010).

19. Andrén, O. *et al.* How well does the Gleason score predict prostate cancer death? A 20-year followup of a population based cohort in Sweden. *J Urol* **175**, 1337–1340 (2006).

20. Goel, S., Bhatia, V., Biswas, T. & Ateeq, B. Epigenetic reprogramming during prostate cancer progression: A perspective from development. *Semin Cancer Biol* S1044-579X(21)00023–7 (2021)

21. Janssen, A., Colmenares, S. U. & Karpen, G. H. Heterochromatin: Guardian of the Genome. *Annu Rev Cell Dev Biol* **34**, 265–288 (2018).

22. Braig, M. *et al.* Oncogene-induced senescence as an initial barrier in lymphoma development. *Nature* **436**, 660–665 (2005).

23. Wen, B., Wu, H., Shinkai, Y., Irizarry, R. A. & Feinberg, A. P. Large organized chromatin K9-modifications (LOCKs) distinguish differentiated from embryonic stem cells. *Nat Genet* **41**, 246–250 (2009).

24. Rhie, S. K. *et al.* A high-resolution 3D epigenomic map reveals insights into the creation of the prostate cancer transcriptome. *Nat Commun* **10**, 4154 (2019).

25. Pope, B. D. *et al.* Topologically associating domains are stable units of replication-timing regulation. *Nature* **515**, 402–405 (2014).

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

27. Li, C.-W. *et al.* Epithelial–Mesenchymal Transition Induced by TNF-α Requires NF-κB–Mediated Transcriptional Upregulation of Twist1. *Cancer Res* **72**, 1290–1300 (2012).

28. Saxena, K., Jolly, M. K. & Balamurugan, K. Hypoxia, partial EMT and collective migration: Emerging culprits in metastasis. *Translational Oncology* **13**, 100845 (2020).

29. Rojas, A. *et al.* IL-6 promotes prostate tumorigenesis and progression through autocrine cross-activation of IGF-IR. *Oncogene* **30**, 2345–2355 (2011).

30. Papageorgis, P. TGFβ Signaling in Tumor Initiation, Epithelial-to-Mesenchymal Transition, and Metastasis. *Journal of Oncology* **2015**, e587193 (2015).

31. 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 (2011).

32. Verdone, J. E., Parsana, P., Veltri, R. W. & Pienta, K. J. Epithelial-mesenchymal transition in prostate cancer is associated with quantifiable changes in nuclear structure. *Prostate* **75**, 218–224 (2015).

33. Bracken, C. P. & Goodall, G. J. The many regulators of

epithelial−mesenchymal transition. *Nat Rev Mol Cell Biol* **23**, 89–90 (2022).

34. M, S. Involvement of partial EMT in cancer progression. *Journal of biochemistry* **164**, (2018).

PUBLICATIONS

- Luminal lncRNAs Regulation by ERα-Controlled Enhancers in a Ligand-Independent Manner in Breast Cancer Cell [2018]
  V. Miano , G. Ferrero, V. Rosti, E. Manitta , J. Elhasnaoui, G. Basile, M. De Bortoli
  https://pubmed.ncbi.nlm.nih.gov/29462945/

- Dysfunctional polycomb transcriptional repression contributes to lamin A/C–dependent muscular dystrophy [2020]
  A. Bianchi, C. Mozzetta, G. Pegoli, F. Lucini , S. Valsoni, V. Rosti, C Petrini, A. Cortesi, F. Gregoretti, L. Antonelli, G.Oliva, M. De Bardi, R. Rizzi, B. Bodega, D. Pasini, F. Ferrari, C. Bearzi, C. Lanzuolo
  https://pubmed.ncbi.nlm.nih.gov/31999646/