UNIVERSITÀ DEGLI STUDI DI TORINO

Doctoral School in Life and Health Sciences

PhD Program in Complex Systems for Life Sciences

Cycle: XXXIV

Role of long non-coding RNAs in cellular senescence

Thesis's author: Isabelle Laurence Polignano

Tutor: Prof. Salvatore Oliviero

Code of scientific discipline: BIO/11
Academic years of enrolment: AA 2018-2021

Contents

Abstract

Senescence is known to have a key physiological role to ensure the removal of damaged cells and tissue renewal. The signalling pathways involved and responsible for its activation are well-studied and include a number of different cellular stress and signalling. More recently the senescence phenomenon was investigated revealing a profound chromatin remodelling while comparing proliferating and senescent cells.

The role of non-coding transcripts involved in cellular senescence has been so far poorly investigated although few lncRNAs are known to interact and/or modulate known senescence-related transcription factors.

My project aims to unveil novel lncRNAs playing a key role in the activation of cellular senescence. To this attempt, we first set up in the lab and characterised the cellular senescence model. Next, we identified differentially expressed lncRNAs. We then selected nine lncRNAs among the upregulated genes in senescent conditions to be subjected to a CRISPR/Cas9 knockout to evaluate their effect on the cellular life-span. Our screening revealed five lncRNAs that displayed a delay in the senescence occurrence, suggesting their involvement in senescence activation. A preliminary characterization consisting in the inspection of their subcellular localization, revealed that they belong to different cellular compartments, and therefore we expected them to act and intervene on gene expression regulation at different levels.

Further experiments are required for the functional characterization of these five lncRNAs, in order to define in detail how they affect senescence activation and, in particular, which are their molecular partners and targets.

Introduction

Cellular senescence

Cellular senescence is a physiological process uncovered sixty years ago by Leonard Hayflick and Paul Moorhead (1961) from the observation that normal human diploid fibroblasts cease to proliferate after a fixed number of population doublings, thereafter called Hayflick limit (Hayflick and Moorhead, 1961). Based on the Nomenclature Committee on Cell Death (NCCD), cellular senescence is the irreversible loss of proliferative potential associated with specific morphological and biochemical features, including the senescence-associated secretory phenotype (SASP) (Galluzzi et al., 2018). From the *in vivo* point of view, it represents a fundamental process that enables the removal of damaged cells and the tissue renewal through remodelling and re-population. Senescent cells undergo a durable state of cell cycle arrest and thus are hyporeplicative cells, but they still retain and explicate some functions of their replicative version.

Various kinds of stress and signals can trigger this phenomenon, such as telomere shortening and DNA damage response (DDR), telomeric structure modifications, epigenetic changes, inflammation, ionization, chemotherapeutic drugs, oncogene activation; these diverse causes lead to different types of senescence, generally classified as DNA damage-induced, oncogene-induced senescence (OIS), chemotherapy-induced, oxidative stress-induced senescence, or replicative senescence (Hernandez-Segura et al., 2018; Muñoz-Espín and Serrano, 2014). As a result of the senescence-related molecular pathways, senescent cells show altered morphological characteristics, like enlarged size and irregular shape, loss of lamin B1, which affects nuclear lamina integrity, dysfunctional mytochondria, leading to higher reactive oxygen species (ROS) level, increased lysosomal content, altered composition of the plasma membrane, enriched in caveolin-1.

Hallmarks of senescence

DNA damage

DNA damage, in particular double-strand breaks (DSB), is a strong activator of DDR, which represents a robust and evolutionarily conserved response to damaged DNA (Figure 1.1). The stimuli that can trigger this mechanism are both extrinsic, such as chemotherapeutic agents, ultraviolet radiation, and intrinsic, such as telomere attrition, oxidation and hyperproliferation. During replicative senescence a progressive telomere shortening takes place, activating DDR machinery in reaction to uncapped chromosome exposed. In OIS the hyperproliferation of cells leads to accumulation of DNA damage triggering DDR. Both replicative and oncogene-induced senescence processes involve DDR, respectively in response to a telomeric and non-telomeric DNA damage, and in both cases the persistence of DDR triggers the phosphorylation and subsequent activation of p53 which in turn can activate several genes. In fact, single-strand breaks and double-strand breaks are detected by specific complexes which recruit ataxia telangiectasia mutated (ATM) and ataxia telangiectasia and RAD3-related (ATR) to the site of DNA damage. The signal is subsequently amplified by the phosphorylation of histone H2AX (γ-H2AX), that aids in the formation of other DNA repair complexes, leading to nuclear foci assembly where several DDR proteins accumulate. In parallel, also histone methylation, such as trimethylation of lysine 9 histone 3 (H3K9me3), functions as marker for DSB and locus of specific repair complexes assembly. Since the majority of stimuli inducing senescence have an impact on DNA damage, DDR characteristics such as γ-H2AX and phospho-p53 are considered and employed as markers of senescence (Kumari and Jat, 2021).

Cell Cycle Arrest

Cell cycle arrest is probably the main hallmark of senescence, although this feature is in common with another form of growth arrest, called quiescence. One important difference between senescence and quiescence is that the cell cycle interrupts in two different stages, respectively in G1/G2 and in G0. Moreover, while senescence is considered an irreversible process, quiescence is reversible, meaning that cells can restart proliferating under specific conditions. In fact, this latter phenomenon is induced by nutrient deprivation and involves mechanistic target of Rapamycin (mTOR) inhibition while when

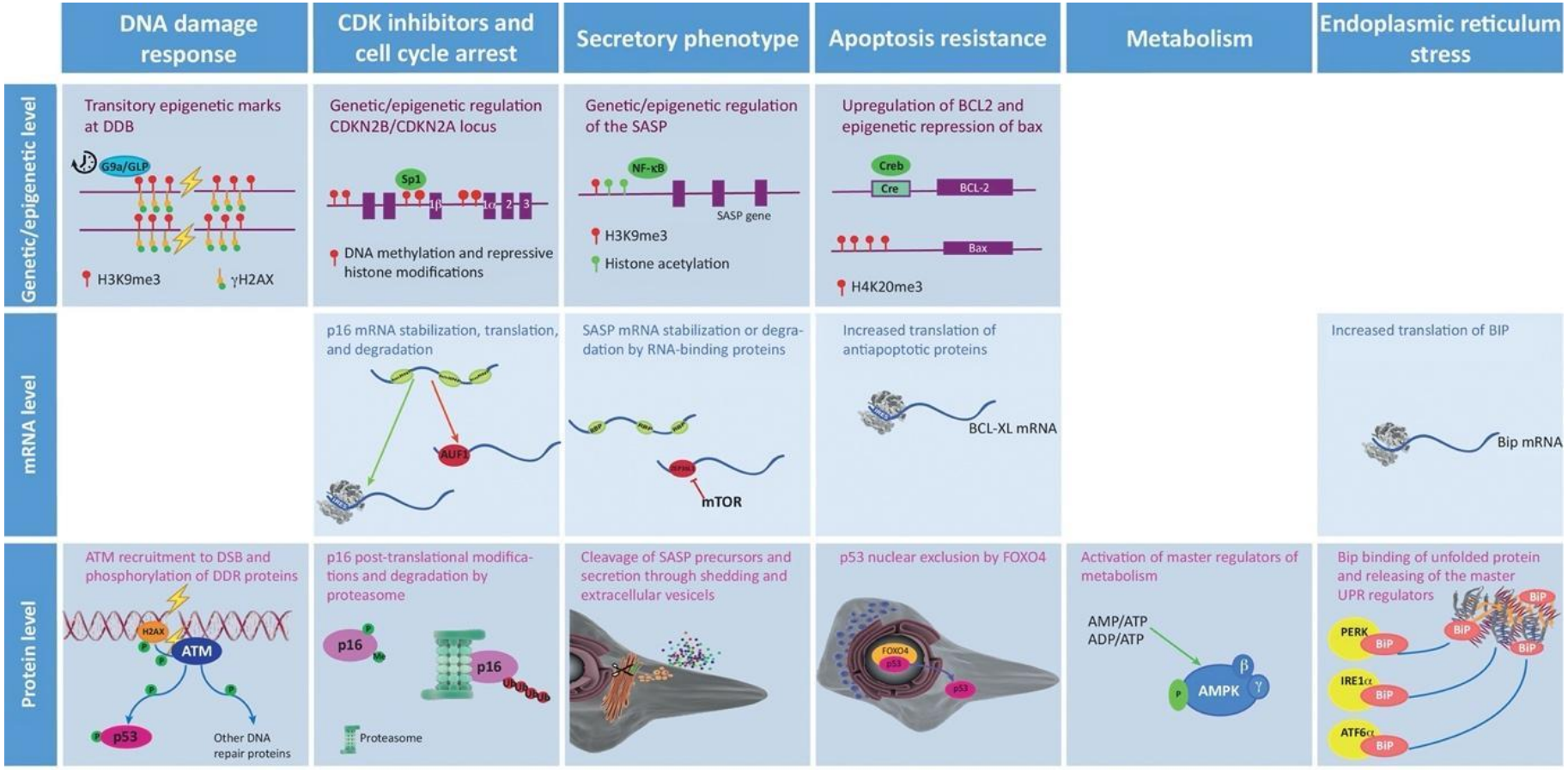


Figure 1.1: The main regulation steps for key molecular players and signaling pathways of senescence. The regulation occurs at three different levels (i) genetic and/ or epigenetic level, which includes mechanisms that modify transcription; (ii) mRNA level, which includes mechanisms such as mRNA stabilization or degradation and even recruitment of ribosomes, all of which affect translation; and (iii) protein level, which includes regulatory mechanisms that occur after translation. Figure from Hernandez-Segura et al., 2018.

mTOR is maintained activated, senescence is induced. Another type of proliferation arrest is represented by apoptosis, the process of programmed cell death. In this case, cells are committed to death and engulfment, which enables the rapid removal of dead cells by macrophages, while senescent cells remain partially metabolically active after proliferation arrest (Kumari and Jat, 2021).

Cell cycle arrest is mainly mediated by cyclin-dependent kinase inhibitors (CDKis) such as CDKN2A ($p16^{INK4a}$, hereafter p16) and CDKN2B ($p15^{INK4b}$, hereafter p15), belonging to INK4 family, and CDKN1A ($p21^{CIP}$, hereafter p21), member of KIP/CIP family (Figure 1.1).

In particular, p16, a unique and specific marker of senescence, is activated as a consequence of epigenetic changes. In fact, its genomic locus, named CDKN2A, is kept methylated in young proliferating cells either by DNA methyl-transferase 1 (DNMT1) or Polycomb group repressive complexes 1 and 2 (PCR1, PCR2). Also, the macroH2A1 repressive histone variant is depleted in the active p16 locus in senescent cells. The derepression of CDKN2A locus is related to the interplay between transcription factors, that trigger its transcriptional activation (i.e., AP1, Ets, Sp1), and repressive mechanisms. The molecular details of this overall mechanism are still not completely understood (Hernandez-Segura et al., 2018).

The function of p15 in senescence is less studied compared to p16. Encoded by CDKN2B gene locus, it is known to be involved in TGF-β-mediated cell cycle arrest, since it was demonstrated that p15 is an effector of TGF-β, which stabilizes the protein thus increasing p15-CDK4 complex (Sandhu et al., 1997).

The CDKi p21 was the first identified target of p53 (Wafik S. El-Deiry, 1993) and it actually plays a dual and opposite role in cell cycle progression: when it is expressed at high levels, it is able to inhibit cyclin D-CDK4/6 complex, resulting in inhibition of cell cycle progression; on the other side, when p21 is expressed at low levels it can participate to cell cycle progression acting as scaffold for cyclin D-CDK4/6. The activation of p21 can trigger the arrest of the cell cycle at any stage, since it is able to inhibit all CDKs; this represents a main difference with INK4 family CDKis that can bind only CDK4 and CDK6, interrupting cell cycle progression in G0/G1. p21 is also responsive to other stimuli and can be activated in a p53-independent manner, for example via retinoid, androgen, and vitamin D receptors (Kumari and Jat, 2021).

Downstream effector of CDKis is the retinoblastoma (RB) family. RB family members when dephosphorylated are able to bind to E2F family forming a repressive complex that inhibits cell-cycle progression. CDKs are responsible of phosphorylating RB family thus enabling cell-cycle progression. With the upregulation of CDKis in senescence, RB family proteins are activated and inhibit E2F, leading to proliferation arrest (Gorgoulis et al., 2019).

Apoptosis resistance

Senescent cells undergo durable cell cycle arrest but maintain activated some metabolic functions and are not immediately removed from the tissue, as happens with dead cells. This is due to the inhibition of apoptosis (Figure 1.1), thanks to the activation of the anti-apoptotic protein BCL-2, whose inactivation is prevented by the expression of cAMP response element-binding protein (CREB). Moreover, the repression of the proapoptotic protein Bax is guaranteed by the repressive histone mark trimethylated lysine 20 of histone H4 (H4K20me3). Other factors participate to this pro-survival program, such as other members of the BCL-2 family, like BCL-XL and BCL-W; FOXO4 responsible of p53 hijacking thus preventing cell death; and p21 that reduces caspase signalling pathway when induced by a persistent DNA damage (Hernandez-Segura et al., 2018).

SASP

Senescence-associated secretory phenotype (SASP) is considered a hallmark of senescence and includes all the inflammatory cytokines, chemokines, growth factors, extracellular matrix components that are secreted by senescent cells (Figure 1.1). It is a highly heterogeneous program that varies depending on cell type, pro-senescence stimuli origin and duration of senescence; this aspect makes SASP unspecific to employ it as unequivocal marker for senescence. SASP is implicated in physiological processes such as tissue remodelling, embryonic structures reorganization and immunosurveillance, but it also plays a detrimental role as it is related to chronic inflammation, cancer recurrence and, more generally, to aging phenotype (Figure 1.2) (Hao et al., 2022).

The main regulators of SASP program are nuclear factor kappa-light-chain-enhancer of activated B cells (NF-kB), CCAAT-enhancer-binding protein β (C/EBPβ), GATA binding protein 4 (GATA4), mTOR and p38/mitogen-activated protein kinase (MAPK) signalling pathway.

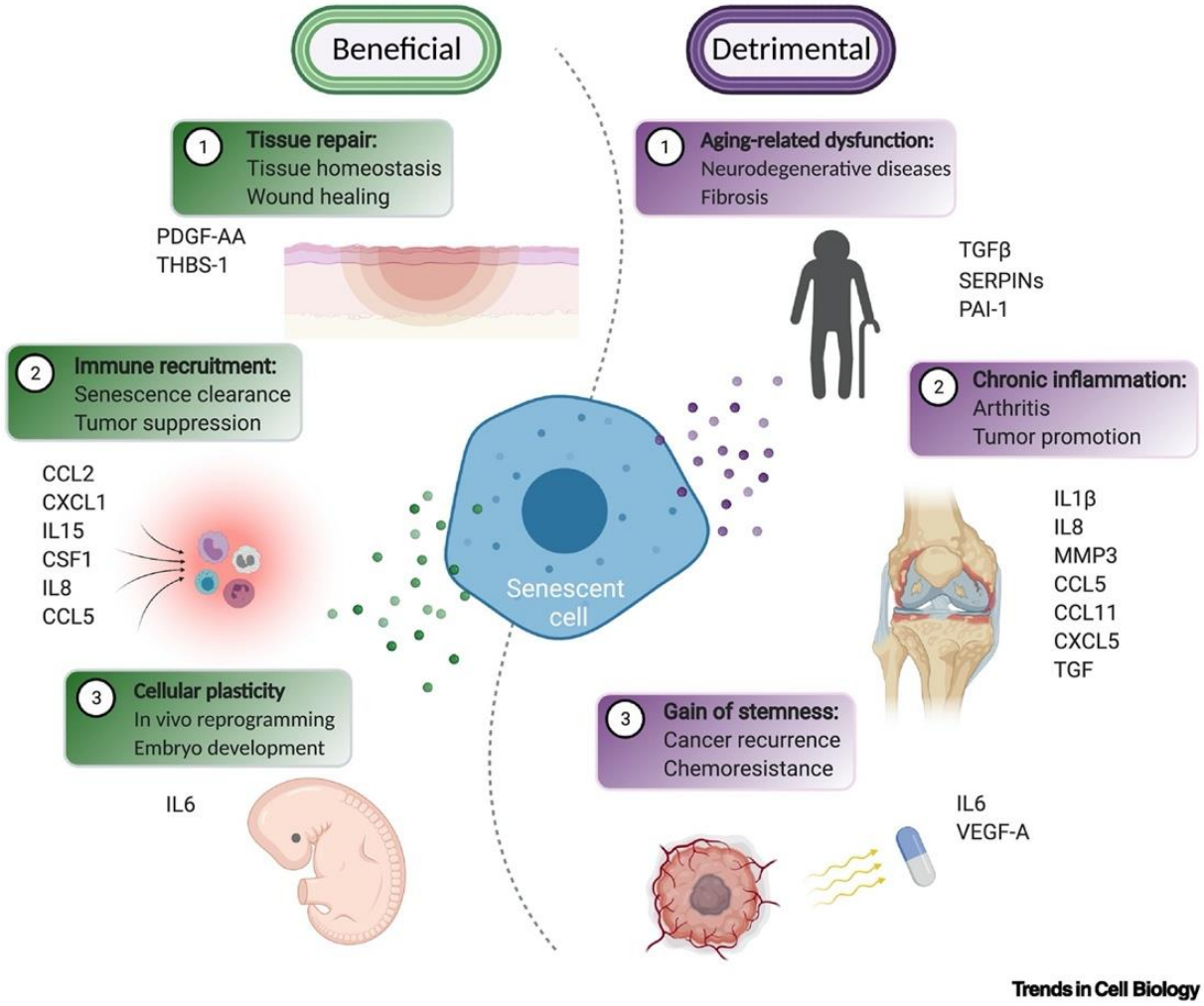


Figure 1.2: The multifaceted functions of the senescence-associated secretory phenotype (SASP). The beneficial (in green) and the detrimental (in purple) roles of the SASP are summarized. The specific SASP factors that function in these processes are indicated. Note that the beneficial and detrimental roles played by the specific SASP factors are context dependent. For instance, IL6 is critical for senescence-associated cellular reprogramming during embryo development, while contributing to the relapse and cancer stemness during therapy-induced senescence. Figure from Hao et al., 2022.

All different events leading to DDR, such as telomers dysfunction, oxidative stress, proliferative signals, DNA damage, trigger the activation of SASP. The main transcription factor involved in DDR-induced SASP is NF-kB, but also the cyclic GMP-AMP synthase (cGAS)-stimulator of interferon genes (STING) pathway is a known alternative. Interestingly, it was demonstrated that Notch receptor 1 (NOTCH1) signalling plays a role in OIS, driving a TGF-β-rich secretome and inhibiting the pro-inflammatory secretome typical of replicative senescence and induced by C/EBPb (Kumari and Jat, 2021).

In case of mitochondrial dysfunction-associated senescence, SASP induction is milder or absent; this reinforces the hypothesis that DNA damage is a main driver of SASP.

Epigenetic changes can also influence the transcription of SASP genes. In particular, promoters of cytokines genes, such as interleukin (IL)-6 and IL-8, are depleted of their usual repressive mark histone 3 lysine 9 (H3K9) bimethylation. This is due to two opposite events: on one side the proteosomal degradation of methyltransferases; on the contrary, histone deacetylases, such as sirtuin 1 (SIRT1), that are downregulated during senescence leading to an increased histone acetylation at the promoters of the same genes. The repressive histone variant macroH2A1 during senescence is depleted from genomic regions encoding for SASP genes, but it also plays an additional role enabling the expression of the main cytokines IL-1A, IL-1B, IL-6 and IL-8 and the matrix metallopeptidase 1 (MMP1). In fact, a feedback loop takes place starting from stressed endoplasmic reticulum (ER) caused by the activation of SASP program, that lead to increased ROS production and DDR (Figure 1.1). This latter, in turn, triggers the mobilization of macroH2A1 in order to silence SASP genes and decrease ER stress. SASP transcriptional program is positively regulated also by another histone variant named H2A.J. Finally, from the epigenetic point of view, three-dimensional chromatin rearrangement contributes to the regulation of secretome genes, in particular the expression of high-mobility group box 2 (HMGB2) limits the spreading of heterochromatin (Hao et al., 2022; Hernandez-Segura et al., 2018).

Post-transcriptionally in OIS, SASP program is known to be regulated by mTOR. A role was proposed for p38/MAPK signalling that, promoted by ROS, acts phosphorylating RNA-binding proteins that stabilize mRNAs of SASP genes.

The regulation of SASP occurs also at post-translational level. One example is the cleavage of IL-1b by caspase-1, that activates IL-1 pathway.

SASP factors have both cell autonomous and non-cell autonomous functions, since they exert their role reinforcing the senescent state of the cell it-self but some of them are soluble and secreted in the extracellular environment influencing the behaviour of neighbouring cells (Stow and Murray, 2013).

Altered metabolism

Mitochondrial dysfunction is another key aspect of senescence. In fact, mitochondria number is increased as well as their size, but their functions are compromised, in particular maintenance of membrane potential and ATP production. In fact, senescent cells show proton leak, higher level of ROS, which lead to lipid and proteins damage and contributes to telomere shortening and DDR activation, and increased ratio of AMP:ATP and ADP:ATP, contributing to cell cycle arrest because of the AMPK (AMP kinase) activation (Figure 1.1). Mitochondrial changes are also related to SASP. Senescent cells present lower levels of NAD+/NADH ratio that affect the activity of sirtuins, known to be involved in SASP regulation. Despite these observations about mitochondria dysfunction *in vitro*, very few is known on the *in vivo* side, and moreover this phenomenon regards also other cellular processes. For these reasons they are not used as senescence biomarker (Gorgoulis et al., 2019).

Another aspect of altered metabolism in senescence regards lysosomes, which are more abundant and bigger in size, they are actually responsible of the typical granular aspect of senescent cells cytoplasm. Known to be responsible for degradation of biological macromolecules through endocytic, autophagic and phagocytic pathways, their biogenesis is regulated by energetic and degradative needs of the cell. The hypothesis behind their increased number is balancing the accumulation of dysfunctional lysosomes by producing more new organelles, to try to maintain the equilibrium between anabolism and catabolism. Nonetheless, lysosomes abundance is not proportional to their activity and efficiency, actually in senescent cells autophagy process is decreased. Lysosome dysfunction is also related to the mitochondrial one. In fact, a negative feedback loop occurs since the alteration of lysosome functions affects mitochondrial turnover, leading to increased levels of ROS, which in turn are detrimental not only for macromolecules but also for cellular structures like lysosomes (Park et al., 2018). Their increased content

has been related to the lysosomal enzyme senescence-associated β-galactosidase (SA-β-gal). SA-β-gal staining is probably the most common marker for detecting senescent cells, even though the mechanism behind its overexpression in actually unknown (Hernandez-Segura et al., 2018).

Senescence-associated transcription networks

In the last decade, the majority of efforts in studying senescence focused on the definition of morphological and metabolic changes, or the analysis of the main pathways activated to induce this program. Senescence process is highly heterogeneous and variable, therefore these characteristics cannot be considered as universal and unequivocal. To the attempt of profiling senescent cells in a more robust manner, transcriptome analysis is fundamental. This has been recently performed comparing several cellular models, mainly fibroblasts and endothelial cells, undergoing different types of senescence, replicative, oncogene-induced, exposed to radiation or drugs. The intersection of the different conditions and strains allowed to identify the common up-regulated and down-regulated genes which can be considered as a sort of transcriptional signature of senescence (Casella et al., 2019; Marthandan et al., 2016). Interestingly, in some cases, mRNA and protein levels were not correlated in all the strains, so that known senescence-associated genes such as p16 and p21 were not among the significantly up-regulated genes. Therefore, based on cell type, the regulation of key factors in senescence can occur at the transcriptional level or by other down-stream mechanisms (Marthandan et al., 2016).

In particular, one study identified 55 core senescence genes that were in common between the different samples and surprisingly no one of the common marker of senescence was is this list. Within the core signature, genes correlated to cell cycle machinery were found down-regulated, while those linked to inflammation, p53 signalling or apoptosis modulation were found up-regulated. One of the parameters influencing the expression profile is the temporal dynamic, so the time of detection of the gene expression has to be considered as a key variable in senescence studies (Hernandez-Segura et al., 2017).

Among the up-regulated genes, the majority were protein-coding genes, but also some non-coding RNAs (ncRNAs) were found to be overexpressed in senescence conditions. Among them, micro-RNAs (miRNAs) have an influence on senescence program modulating key factors such as p53, p21 and SIRT1. They exert their role alone as well

as acting together, sometimes also miRNAs participate to intricate feedback loops to regulate senescence-associated factors. In parallel, also long non-coding RNAs (lncRNAs) play a role in senescence process (Hernandez-Segura et al., 2017).

Long non-coding RNAs

With the ENCODE (Encyclopedia of DNA Elements) project and FANTOM (Functional Annotation of Mammalian genome) consortium the concept of "junk" DNA was challenged and novel transcripts were identified from those regions and called ncRNAs. LncRNAs are recognized as novel regulators participating in multiple cellular processes, such as cell cycle control, differentiation and development, imprinting and chromosome remodelling.

LncRNAs are defined as non-coding transcripts longer than 200 nucleotides, they are known to be mainly transcribed by RNA polymerase II, and they are usually spliced and polyadenylated. However, they are a heterogeneous group presenting different characteristics, since some lncRNAs, on the contrary, lack the polyadenylated (poly(A)) tail or splicing capacity and others can even code for small peptides. Their heterogeneity also regards the multiplicity of functions and mechanisms they exert thanks to their ability to interact with different biomolecules. In fact, lncRNAs can bind DNA as well as RNA molecules through base pairing, and thus act as sponges for micro-RNAs (miRNAs) or splicing controllers. On the other hand, their capacity of interaction with proteins renders lncRNAs versatile molecules, being able to modulate the role of epigenetic factors and chromatin modifiers, as well as transcription factors enhancing their recruitment to different loci (Mirzadeh Azad et al., 2021).

Classification of lncRNAs

Till now, an official classification comprising all lncRNAs does not exist. One of the ways to classify them is based on their genomic position relative to other genes (Figure 1.3): (A) Divergent and antisense lncRNAs overlap or are located in close proximity to a sense gene and are transcribed in the opposite direction; (B) intronic lncRNAs are transcribed from the intron of a gene; and (C) intergenic lncRNAs do not overlap annotated coding genes and are autonomously transcribed. LncRNAs can also be categorized based on their mechanisms of action and regulation, including for example the way they regulate their target gene, by *cis*-acting or *trans*-acting lncRNAs (Figure 1.3); their molecular role, such as competitive endogenous RNAs and enhancer RNAs; their own transcriptional regulation, for example stress-induced promoter-associated

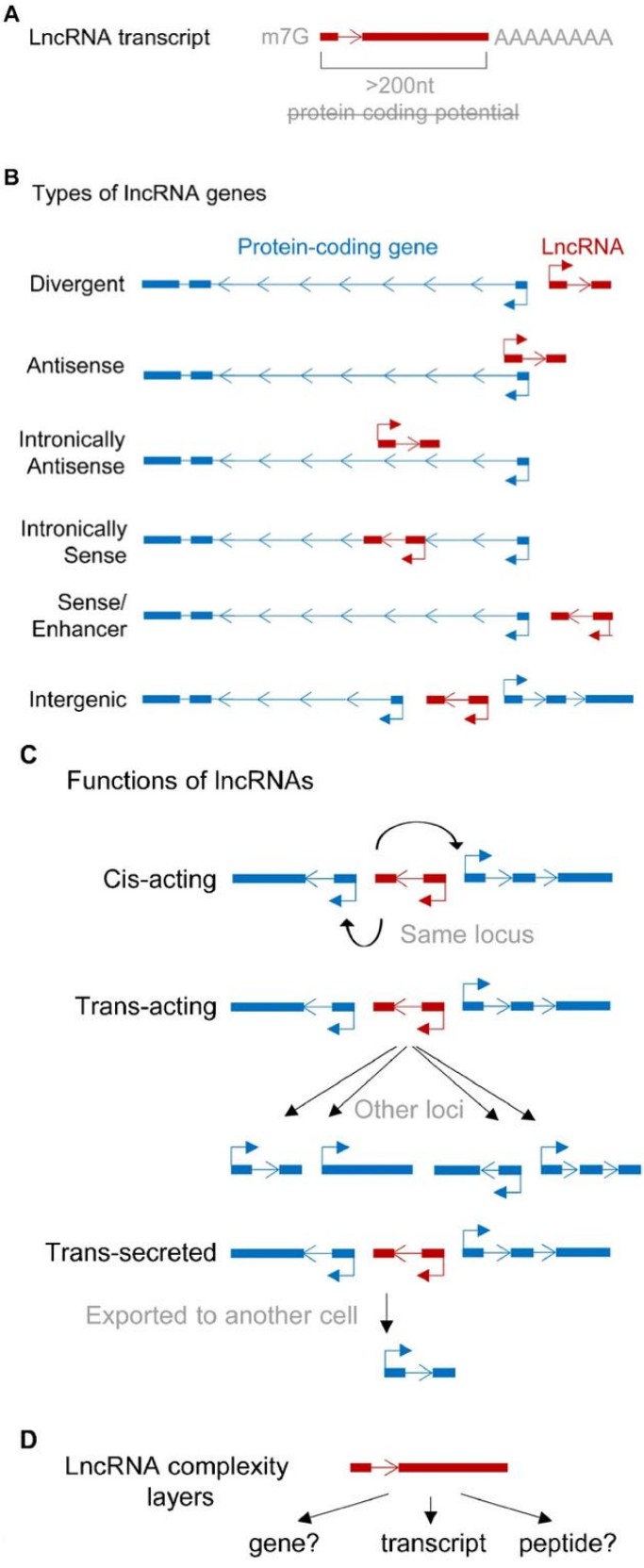


Figure 1.3: Research outlines and lncRNA characteristics. A LncRNA transcripts are defined as non-coding RNAs longer than 200nt apparently lacking protein-coding potential. Typically, the majority of lncRNAs are mRNA-like RNAs harbouring a 5'Cap and a polyA tail. B Genomic location of lncRNA genes. C LncRNAs can act *in cis* to regulate the immediate locus from which the lncRNA was transcribed, *in trans* to function elsewhere in the cell or trans-secreted. D For some lncRNA genes, functions on their gene itself, their transcript or peptide are known increasing the layer of complexity for their mode of operations. Figure modified from Oo et al., 2022.

antisense lncRNAs, damage-induced lncRNAs (Oo et al., 2022). More recently, it was proposed that, beyond the role and mechanism of the lncRNA transcript, also the process of transcription itself can independently have a function. In fact, the genomic locus encoding for the lncRNA could be part of a 3D nuclear architecture that could be modified by the transcriptional activity of the genomic locus; in particular it was seen that both transcription initiation and elongation can affect 3D genome architecture. The "functional microdomains" model was proposed, in which transcription factors, chromatin remodelers, RNA polymerase II (Pol II), and non-coding transcripts interact and collaborate together in these 3D microdomains; the transcriptional activity participate too in gene regulation influencing chromatin structure at different levels (Ali and Grote, 2020).

Furthermore, it was pointed out that lncRNAs are not completely non-coding but, even though their coding potential is low, they can originate small functional peptides (called micropeptides); in this case, of course, to be considered and classified as lncRNA they have to retain an independent RNA function (Figure 1.3) (Statello et al., 2021).

All these characteristics, functions and possible mechanisms increase the complexity of trying to classify the lncRNA transcripts group in an ordered and clear way, making it still a proper challenge.

LncRNAs in chromatin regulation

LncRNAs can influence chromatin structure either acting in *cis*, when they modulate chromatin topology locally, or in *trans*, when they participate to inter-chromosomal nuclear architecture regulation. In both situations they can exploit chromatin looping and interaction with chromatin-modifying complexes, in order to enforce activation or repression of the gene transcription.

HOXA transcript at the distal tip (*HOTTIP)* is an example of lncRNA actin *in cis* that promotes HOXA gene cluster expression (Figure 1.4). In particular, *HOTTIP* interacts with the 5´ region of the HOXA locus via chromatin looping, and it carries the adaptor protein WD repeat-containing protein 5 (WDR5)-myeloid/ lymphoid or mixed-lineage leukaemia protein 1 (MLL1; also known as KMT2A) histone lysine methyltransferase complex to the promoter, inducing histone 3 lysine 4 (H3K4) trimethylation and HOXA transcription (Ransohoff et al., 2018).

A well-studied *trans*-acting lncRNA is HOX transcript antisense RNA (*HOTAIR*), whose gene is part of the HOXC locus, and is responsible for the repression of the distal HOXD genes (Figure 1.4). The silencing is due to *HOTAIR* acting as scaffold for PRC2 with the KDM1A–coREST–REST complex, thereby facilitating the histone 3 lysine 27 (H3K27) trimethylation and H3K4 demethylation that combined together lead to transcriptional repression. More recently, *HOTAIR* mechanism of action and function in PRC2 recruitment was challenged by demonstration that the HOTAIR–PRC2 interaction is dispensable for *HOTAIR* function; the suggested model is that the affinity of chromatin modifiers for RNA is important to compete with chromatin, thus preventing low affinity binding (Portoso et al., 2017).

Very recently, a novel mechanism of action was demonstrated for *lncSmad7* (Figure 1.5). This lncRNA, previously known to be involved in pluripotency maintenance, is shown to regulate gene expression in-*trans* by recruiting p300 to enhancer and promoter regions to perform H3K27 acetylation. In particular *lncSmad7* is shown to directly interact with p300 at the C-terminal domain of this enzyme, and on the other side it is able to bind DNA via Hoogsteen base pairing (Maldotti et al., 2022).

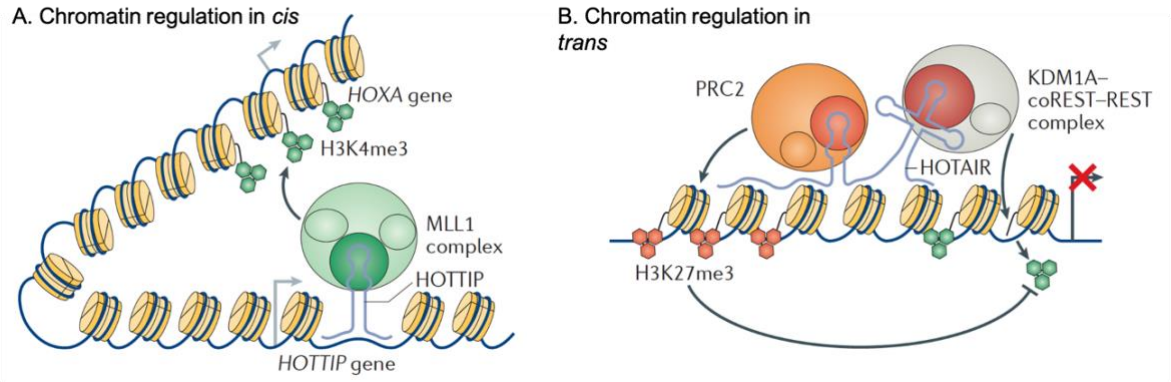


Figure 1.4: Regulation of chromatin structure and function in cis and in trans. A HOTTIP associates with MLL1 complex, which catalyses H3K4 trimethylation to activate HOTTIP transcription in cis (left). B HOTAIR interacts in trans with the polycomb recessive complex 2 (PRC2) to mediate its deposition of the repressive H3K27me3 modification and with the KDM1A–CoREST–REST complex to mediate H3K4 demethylation, to coordinate transcription repression at target loci (right). Figure modified from Ransohoff et al., 2018.

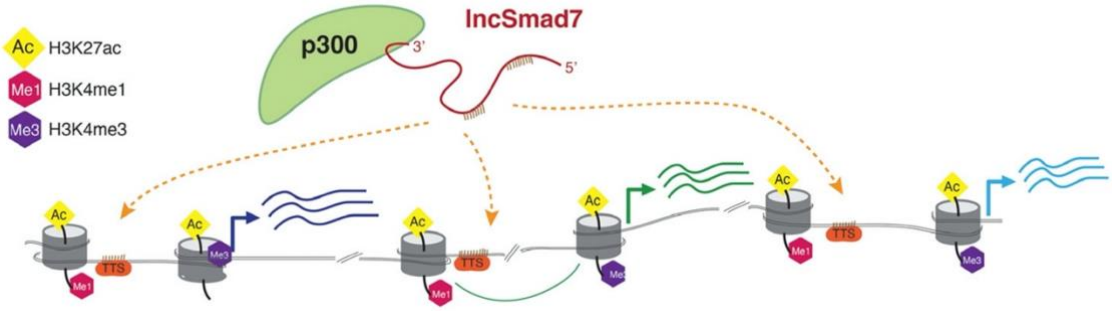


Figure 1.5: Model of p300 recruitment by lncSmad7 on genomi loci. The enhancers and promoters are marked by histone modifications. The RNA–DNA triplexes derived from the indicate regions in the lncSmad7 transcripts (brown bars) and the TTS regions near the enhancers (orange) of target genes. Figure from Maldotti et al., 2022.

LncRNAs in transcription regulation

The genomic position of lncRNAs relatively to the neighbouring genes can represent a key factor for the transcript role and function. In fact, lncRNAs seem to be non-randomly distributed along the genome and their diffused antisense and bidirectional transcription is actually evolutionarily conserved; it was proposed that this could be considered as an evolutionary adaptation in order to regulate gene expression in a context-related way.

A main example of this subgroup of lncRNAs is the well-known X-inactive specific transcript (*Xist*), which plays a key role in X chromosome inactivation in female mammalian cells. *Xist* lncRNA is able to coat one of the two X chromosomes silencing it; several diverse mechanisms contribute to this attempt, in fact *Xist* is able to exploit the 3D chromatin organization, spreading from spatially proximal loci to distant sites, meanwhile it also interacts with chromatin modifiers to alter the target chromatin structure. It is considered as an initiator of epigenetic memory, since this process was demonstrated to persist later even in absence of *Xist*, maintained by protein complexes assembled on the X chromosome (Engreitz et al., 2013).

Besides the recruitment of repressive protein complexes, lncRNAs can also act interfering with the transcription machinery. One example is antisense of IGF2R non-protein coding RNA (*Airn*) lncRNA in mouse, responsible for the allele-specific expression during mouse embryonic stem cells (mESCs) differentiation. *Airn* overlaps the protein-coding gene Igf2r, therefore, when *Airn* is transcribed from the paternal allele, it displaces Pol II from Igf2r promoter causing the silencing of the gene (Figure 1.6A) (Santoro and Pauler, 2013).

In some cases, an even more close relationship can exist between lncRNAs and their neighbouring genes. The conserved lncRNA CHD2 adjacent, suppressive regulatory RNA (*Chaserr*) is positioned upstream of the Chd2 chromatin remodeller gene and act in *cis* with a particular mechanism. CHD2 binds nascent RNAs, including *Chaserr* RNA, promoting their expression; on the contrary, *Chaserr* affects the accessibility of Chd2 and other genes promoters. Thus, a feedback loop takes place, in which *Chaserr* is exploited by CHD2 as a sensor for its own expression levels and regulation (Rom et al., 2019).

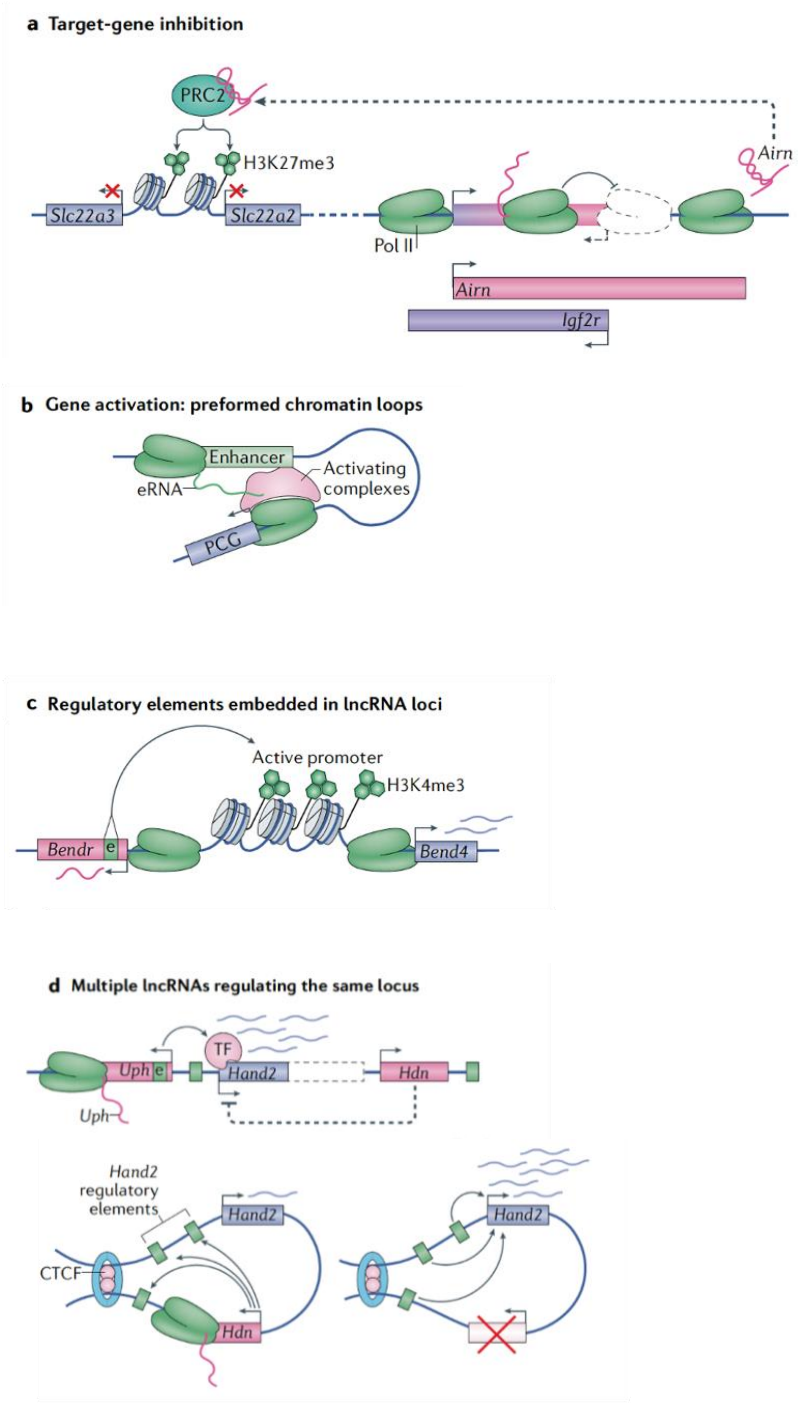


Figure 1.6: Transcription regulation by long non-coding RNAs. A In mouse, *Airn* functions in trans as it is guided to the promoters of two distal target genes, where it recruits PRC2, leading to gene silencing. *Airn* also functions in cis through its own transcription causing steric hindrance for RNA Pol II at the transcription start site of Igfr2r. **B** lncRNAs and eRNAs)can promote the expression of protein coding genes that are in close proximity to their enhancers through chromatin loops, allowing the recruitment of chromatin-activating complexes. **C** Transcription of *Bendr* activates enhancer elements (e) embedded in its locus, which promotes the formation of an active chromatin state at the promoter of the proximal gene Bend4. **D** Example of a complex regulatory unit formed by the lncRNAs *Upperhand* (*Uph*) and *Handsdown* (*Hdn*) in regulating the PCG Hand2. Figure modified from Statello et al., 2021.

LncRNAs at enhancers

Two types of transcripts can originate from active enhancer regions: eRNAs and enhancer-associated lncRNAs (elncRNAs). They display different characteristics: eRNAs are short transcripts, often bidirectionally transcribed, usually non-polyadenylated, unspliced and unstable; elncRNAs are mostly unidirectional transcripts, spliced and polyadenylated. Notably, eRNAs expression correlates with enhancer activity, but they were also proposed to be able to function independently. Some eRNAs participate to chromatin looping binding scaffold proteins, like Mediator, creating contacts between promoters and enhancers located even at great distances (Figure 1.6B) (Statello et al., 2021). The role of elncRNAs is often related to their splicing, which positively influence enhancer activity and the expression level of the neighbouring protein-coding genes. Moreover, elncRNAs were found to be related to human diseases, as the lncRNA SWI/SNF interacting GAS6 enhancer non-coding RNA (*SWINGN*), which facilitates the assembly of the chromatin remodelling complex SWI/SNF at the promoter of its neighbouring gene GAS6, but also at other more distant loci having pro-oncogenic role (Grossi et al., 2020).

Furthermore, there are some specific situations in which the lncRNA locus hosts a regulatory element of a near gene, regulating its expression by its own transcription. An example is the lncRNA Bend4-regulating effects not dependent on the RNA (*Bendr*) whose locus comprises enhancer elements of its neighbouring gene BEND4 (Figure 1.6C) (Statello et al., 2021).

LncRNAs regulatory networks acting in *cis*

With the increasing of knowledge about lncRNAs and their functions, it becomes clear that they are part of complex regulatory units, in which different lncRNAs cooperate to regulate the expression of one protein-coding gene.

The lncRNAs *Upperhand* and *Handsdown* regulate the expression of the transcription factor Hand2, a heart development gene whose unbalancing leads to severe malformations. *Upperhand* hosts an enhancer of the proximal Hand2 gene, therefore when the lncRNA gene is transcribed, Hand2 transcription is activated, without requiring chromatin reorganization. On the other side, *Handsdown* functions by chromatin looping formation, moving its promoter close to Hand2-activating enhancers. When *Handsdown*

transcription is activated, the Hand2 enhancers are hijacked and, thus, Hand2 expression results inhibited (Figure 1.6D) (Anderson et al., 2016).

Scaffolding and condensates

Biomolecular condensates are membrane-less compartments present in eukaryotic cells that function to concentrate proteins and nucleic acids. These condensates are involved in diverse processes, including RNA metabolism, ribosome biogenesis, DNA damage response and signal transduction (Banani et al., 2017). Several abundant lncRNAs participate to the nuclear condensates acting as scaffolds essential for their assembly and functions. A main example is the lncRNA nuclear paraspeckle assembly transcript 1 (*NEAT1*), fundamental for the organization of the so-called paraspeckles, defined as nuclear condensates enriched for proteins and mRNAs, having a role in many biological processes (Figure 1.7A). In particular, the long isoform of *NEAT1* contains the subdomains that recruit the paraspeckle core proteins initiating the assembly of the condensates through liquid-liquid phase separation. The mechanisms by which this process takes place are currently under study (Statello et al., 2021).

Another lncRNA playing a key role in this context is metastasis-associated lung adenocarcinoma transcript 1 (*MALAT1*), which localizes in the so-called nuclear speckles, nuclear domains situated in inter-chromosomal regions and enriched in splicing factors (Figure 1.7B). The depletion of *MALAT1* affects the composition of the nuclear speckles, spatially organized with the splicing factors at the center and the lncRNA at the periphery. The exact manner by which this structural organization of *MALAT1* influence the function of these nuclear domains is still unclear. This lncRNA was also proposed to function as an RNA hub, based also on its ability to form long-range structures; *NEAT1* and the small nuclear RNA U1 are among the transcripts interacting with MALAT1 (Statello et al., 2021).

The scaffolding function of lncRNAs is also operated by other subtypes of RNAs which are small nucleolar RNA-related lncRNAs (sno-lncRNAs) and 5′ small nucleolar RNA-capped and 3′ polyadenylated lncRNAs (SPAs) (Figure 1.7C). These lncRNAs are abundant in human embryonic stem cells (hESCs), and are known to accumulate at their own transcription site and sequester several spicing factors, forming a nuclear body involved in the regulation of alternative splicing. They are also associated with the Prader-Willi syndrome (PWS) since it was shown that induced pluripotent stem cells (iPSCs)

from patients with PWS are depleted of these lncRNAs and display an alteration in the splicing patterns and protein binding to pre-mRNAs associated with neuronal functions.

Analogously, the lncRNA pyrimidine-rich non-coding transcript (*PNCTR*) is highly expressed in cancer and localises in the perinuclear compartment (Figure 1.7D). Along its sequence, it presents several polypyrimidine tract-binding protein 1 (PTBP1) binding motifs, therefore sequestering PTBP1 and suppressing its splicing activity (Statello et al., 2021).

Finally, lncRNAs are also able to put in spatial proximity different chromosomes forming nuclear domains. One example is functional intergenic RNA repeat element (*Firre*), a lncRNA that binds the nuclear matrix factor heterogeneous nuclear ribonucleoprotein (hnRNP) U to maintain a nuclear domain and functioning as a chromosome scaffold to tether together four different chromosomes (Figure 1.7E) (Statello et al., 2021).

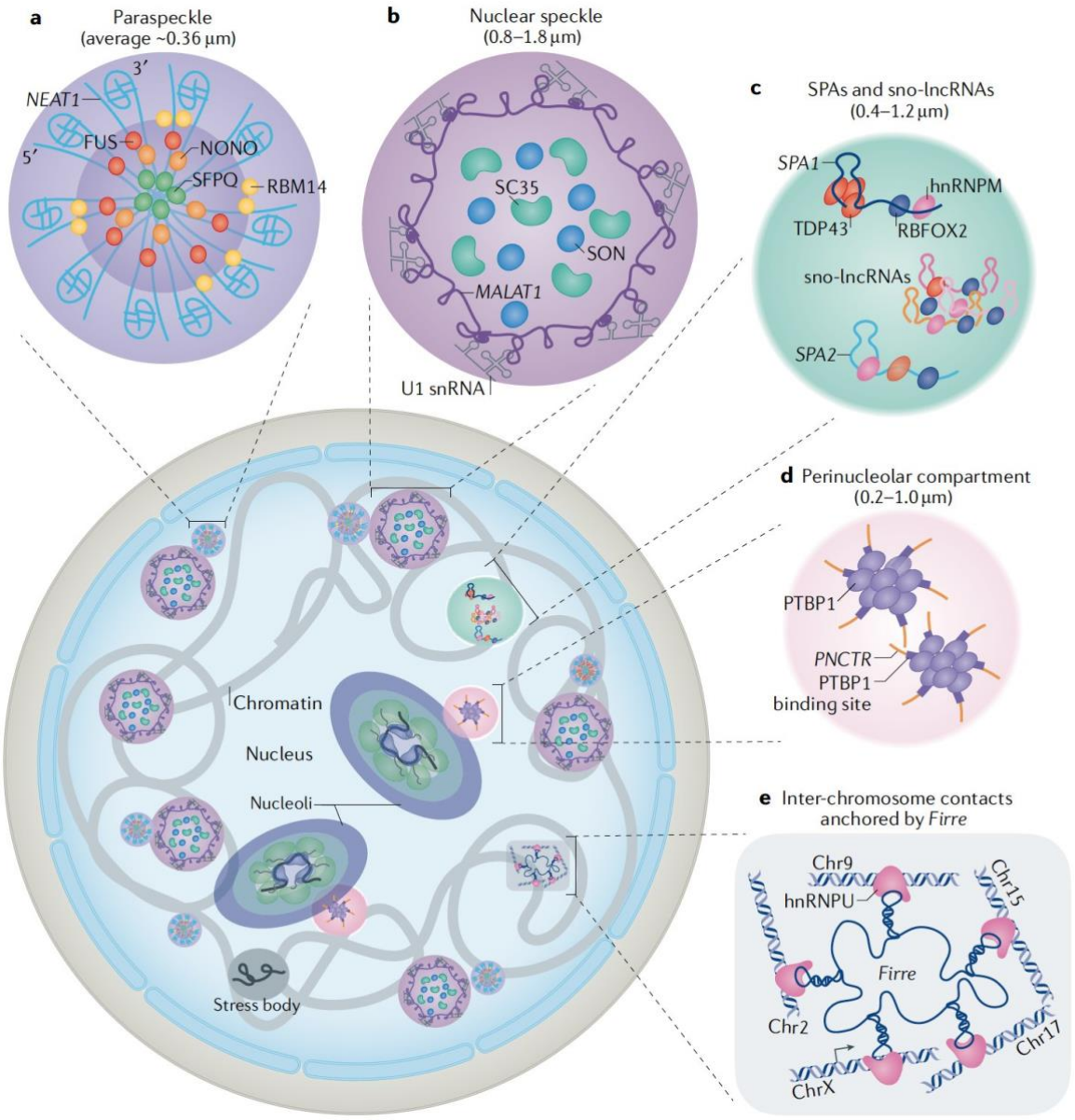


Fig. 1.7 Roles of long non-coding RNAs in nuclear organization. A The lncRNA *NEAT1* sequesters numerous paraspeckle proteins to form a highly organized core–shell (dark and light purple, respectively) spheroidal nuclear body. **B** The lncRNA *MALAT1* is localized at the periphery of nuclear speckles and is involved in the regulation of pre-mRNA splicing. At the periphery, *MALAT1* interacts with the U1 small nuclear RNA, whereas splicing components are localized at the centre of nuclear speckles. **C** SPAs and sno-lncRNAs accumulate at their transcription sites and interact with several splicing factors to form a microscopically visible nuclear body that is involved in the regulation of alternative splicing. **D** The perinucleolar compartment contains the lncRNA *PNCTR*, which sequesters PTBP1 and, thus, suppresses PTPBP1-mediated pre-mRNA splicing elsewhere in the nucleoplasm. **E** The lncRNA *Firre* is transcribed from the mouse X chromosome and interacts with the nuclear matrix factor hnRNPU to tether chromosome X, 2, 9, 15 and 17 into a nuclear domain. The size of each type of nuclear body is indicated in parts A-D. Figure from Statello et al., 2021

LncRNAs in post-transcriptional regulation

LncRNAs display mechanisms to regulate gene expression also at the post-trascriptional level, that we could subdivide into post-transcriptional, translational and post-translation level.

Direct lncRNA-protein interaction

LncRNAs can act sequestering proteins by RNA sequence motifs or structures and assemble into lncRNA-protein complexes (lncRNPs), to influence splicing and turnover processes. Another manner in which they regulate splicing is the modulation of post-translational modifications of the splicing factors, or the binding to pre-mRNAs interfering with their splicing.

On the cytoplasmic side, lncRNA non-coding RNA activated by DNA damage (*NORAD*) sequesters Pumilio, which is responsible for mRNA decay through deadenylation and decapping; *NORAD* is activated by DNA damage and hijacks Pumilio from mRNAs involved in genomic stability (Figure 1.8Ab) (Statello et al., 2021).

LncRNAs are known to be able to fold forming structures that can interact with proteins and modulate signalling pathways. One example is FOXD3 antisense transcript 1 (*FAST*), a lncRNA demonstrated to be necessary for pluripotency maintenance by binding through secondary structures the E3 ubiquitin ligase β-TrCP and therefore avoid the degradation of phospho-β-catenin (Figure 1.8Ac). This latter is free to translocate to the nucleus and activate WNT-dependent pluripotency genes transcription (Statello et al., 2021).

Protein complexes recruitment through lncRNA-RNAs pairing

LncRNAs containing Alu retroelements or short interspersed elements (SINEs) can trigger the so-called Staufen-mediated mRNA decay: binding through base pairing mRNA transcripts, they recruit the double-stranded RNA-binding protein STAU1. By contrast, the lncRNA terminal differentiation-induced non-coding RNA (*TINCR*), harbours several 25-nucleotide motifs that base pair with complementary sequences in mRNAs of genes involved in epidermal differentiation; *TINCR* recruits STAU1, and the lncRBP complex stabilizes the differentiation-associated mRNAs (Figure 1.8Ba) (Gong and Maquat, 2011).

In another context, the lncRNA antisense to ubiquitin carboxyterminal hydrolase L1 (*AS-Uchl1*) contains a SINEB2 repeat; stress signalling pathways trigger the lncRNA translocation from nucleus to cytoplasm, where its SINEB2 element binds the 5′ end of Uchl1 by base pairing enforcing its translation (Figure 1.8Bb) (Statello et al., 2021).

LncRNAs acting as sponge for microRNAs

Some lncRNAs can regulate microRNAs (miRNAs) abundance acting as competitive endogenous RNAs (ceRNAs) or "sponges". One example is represented by the phosphatase 1 nuclear targeting subunit lncRNA (*PNUTS*), involved in cancer, which harbours several binding sites for miR-205, which in turn is a repressor of ZEB1 and ZEB2 factors maintaining epithelial cells (Figure 1.8Bc). When the miRNA is sequestered by the lncRNA, as a result ZEB1 and ZEB2 are released and epithelial-mesenchymal transition is promoted (Statello et al., 2021).

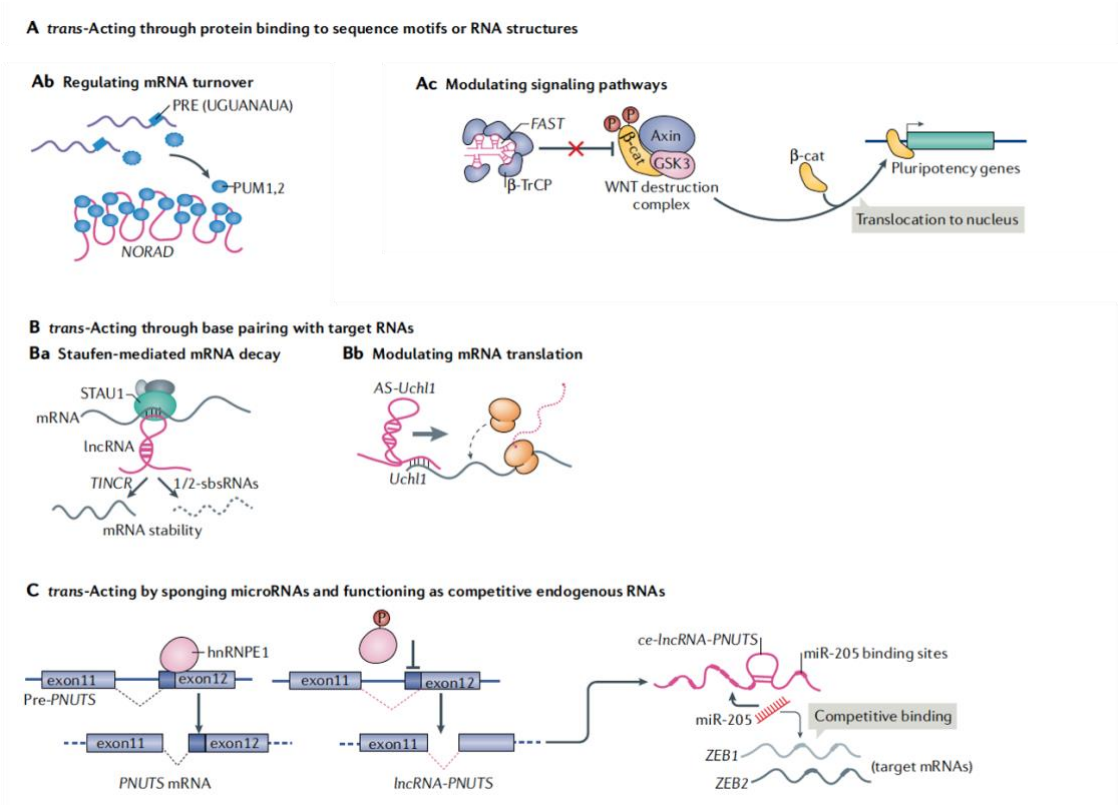


Fig. 1.8: Post-transcriptional functions of trans-acting long non-coding RNAs. A trans-Acting lncRNAs interact with RNA-binding proteins through sequence motifs or by forming unique structural motifs. **Ab** In the cytosol, *NORAD* sequesters Pumilio, which repress the stability and translation of mRNAs to which they bind. **Ac** Human *FAST* binds the E3 ligase β-TrCP, blocking the degradation of its substrate β-catenin, leading to activation of WNT signalling in human embryonic stem cells. **B** trans-Acting lncRNAs directly interact with RNAs through base pairing. **Ba** *TINCR* promotes mRNA stability by forming intermolecular duplexes that bind STAU1, the key protein of Staufen-mediated mRNA decay. **Bb** | The SINEB2 repeat of mouse *AS-Uchl1* complementarily binds the Uchl1 mRNA and promotes polysome association with Uchl1 and translation. **C** Some abundant lncRNAs functions as ce-RNAs. LncRNA-*PNUTS* contains seven miR-205 binding sites, which reduce the availability of miR-205 to bind and suppress ZEB1 and ZEB2 mRNAs. Figure from Satello et al., 2021.

lncRNAs in senescence

LncRNAs in p16 pathway

ANRIL is located in the INK4b-ARF-INK4a locus that encodes for p15, p16 and p14, sharing a bidirectional promoter with the latter transcription factor (Figure 1.9). *ANRIL* acts *in cis* negatively regulating the expression of this locus by PCR1 and PRC2 recruitment. When senescence occurs, the expression of CDKis inversely correlates with *ANRIL*, resulting in repressive marks depletion on the locus (Puvvula, 2019).

With a similar mechanism involving PRC complexes, nuclear lncRNA-MIR31host gene (*MIR31HG*) represses the expression of p16 in proliferating fibroblasts (Figure 1.9). Furthermore, *MIR31HG* downregulation dislodges the repressor complex and restore p16 expression. In senescent fibroblast B-Raf Proto-Oncogene, Serine/Threonine Kinase (BRAF) ectopic expression downregulate *MIR31HG* leading to p16 activation (Puvvula, 2019).

The lncRNA promoter of CDKN1A antisense DNA damage activated RNA (*PANDAR*) is expressed at high levels in breast cancer cells and regulates the G1/S transition and p16 expression by recruitment of the polycomb protein complex BMI-1 (Figure 1.9). *PANDAR* silencing derepresses the transcription of p16, suggesting that the BMI-1 complex associates with the p16 locus in a lncRNA dependent manner. In the same model, Myocardial infarction associated transcript (*MIAT*) downregulation leads to p16 upregulation with cellular senescence and apoptosis increment as consequence and in parallel migration of breast cancer cells decrement (Puvvula, 2019).

In OIS the lncRNA called Very long RNA Antisense to Dimethylarginine dimethylaminohydrolase1 (*VAD*) regulates INK4 locus promoting the dislodgment of H2A.Z repressor mark from the promoter, leading the gene expression activation (Figure 1.9) (Degirmenci and Lei, 2016).

Aside the expression regulation at the transcriptional level, other lncRNAs affect p16 expression acting on splicing and mRNA stability modulation. Urothelial Cancer-Associated 1 (*UCA1*) is highly expressed in senescent cells where it stabilizes p16 mRNA sequestering the hnRNPA1, that is responsible for the rapid degradation mRNAs (Figure 1.9) (Degirmenci and Lei, 2016).

LncRNAs in p53/p21 signalling

Long intergenic non-coding RNA-p21 (*LincRNA-p21*) plays key roles in senescence by targeting p21 (Figure 1.9). This lncRNA acts as scaffold for hnRNP-K recruitment on p21 promoter helping p53 in p21 transcriptional activation. In doxorubicin induced cardiac senescence *lincRNA-p21* is highly expressed and interacts with β-catenin modulating the Wnt/ β-catenin signalling pathway and senescence (Degirmenci and Lei, 2016).

P21-associated ncRNA DNA damage activated (*PANDA*) is directly transcribed by p53 activation. *PANDA* interacts with PRC complexes and scaffold-attachment factor A (SAFA) and nuclear transcription factor Y subunit alpha (NFYA) transcription factors (Figure 1.9). In proliferating cells, *PANDA* forms a repressor complex with SAFA and PRC1/PRC2 regulating the expression of pro-senescence markers like CDKN1A and IL-8. During senescence, SAFA is downregulated therefore the expression of senescence-associated genes is activated, including *PANDA* itself. In parallel, when upregulated *PANDA* acts as decoy lncRNA binding NFYA and preventing the activation of proliferating genes (Puvvula et al., 2014).

Maternally expressed gene 3 (*MEG3*) is a conserved lncRNA known to have a role in several cellular processes (Figure 1.9). Its mechanism of action involves p53, whose activation can be triggered by *MEG3* in different ways: as transcriptional activator, reducing p53 degradation through inhibition of MDM2 expression, or stimulating p53-dependent transcription from a p53-responsive promoter (Puvvula, 2019).

P53 Induced Noncoding Transcript (*Pint*) is a p53 responsive lncRNA that negatively regulates TGF-β, MAPK and p53 autoregulatory pathways through PRC2 targeting (Figure 1.9). Similarly, P53 Regulation Associated LncRNA (*PRAL*) inhibits cell proliferation modulating p53 transcription (Puvvula, 2019).

Other lncRNAs regulate p53 affecting mRNA translation. In particular p53 translation is fine-tuned by two counteracting factors, human antigen R (HuR) and *7SL* lncRNA (Figure 1.9). This lncRNA is highly expressed in proliferating cells, in which it dislodges its competitor HuR reducing p53 translation. In contrast, when 7SL is downregulated during senescence, p53 translation efficiency increases with higher levels of p53 protein production (Puvvula, 2019).

LncRNAs in telomere regulation

Two lncRNAs named telomerase RNA component (*TERC*) and telomeric repeat-containing RNA (*TERRA*) are encoded by telomeric regions (Figure 1.9). TERRA protects chromosome ends by recruitment of heterochromatin protein 1 (HP1α), hnRNPA1, and shelterin components telomeric repeat factor 1 (TRF1) and telomeric repeat-binding factor 2 (TRF2). *TERC* is involved in the maintenance of telomere length and heterochromatin assembly. *TERRA* mechanism of action also involves PRC complexes which are recruited and lead to trimethylation of H3K9, H4K20 and H3K27 marks deposition and heterochromatin formation at telomers (Abdelmohsen and Gorospe, 2015).

GUARDIN is a p53 responsive lncRNA which plays a role in both DNA repair and telomere protection (Figure 1.9). In particular, *GUARDIN* acts as an endogenous competitive lncRNA activating TRF2 through miR-23a inhibition, in order to preserve telomere ends from DDR (Puvvula, 2019).

LncRNAs in stress responses

Recent studies focused on the correlation between lncRNAs and UV-radiation induced photodamage senescence. *MALAT1* lncRNA was found upregulated in fibroblasts upon UVB irradiation and correlates with MMP-1 secretion and senescence phenotype induction (Figure 1.9).

HOTAIR acts as scaffold for PRC1/2 complexes leading to the repression of target genes (Figure 1.9). Its overexpression in cancer is correlated with DDR activation and subsequent induction of NF-kB which in turn positively regulates the transcription of *HOTAIR*.

Similarly, *ANRIL* is involved in a positive feed-back loop for constitutive activation and maintenance of SASP mediated senescence (Figure 1.9). Also, *lincRNA-p21* was reported to participates in a feed-forward loop involving p53 to activate p21 expression and supporting the senescence phenotype (Figure 1.9).

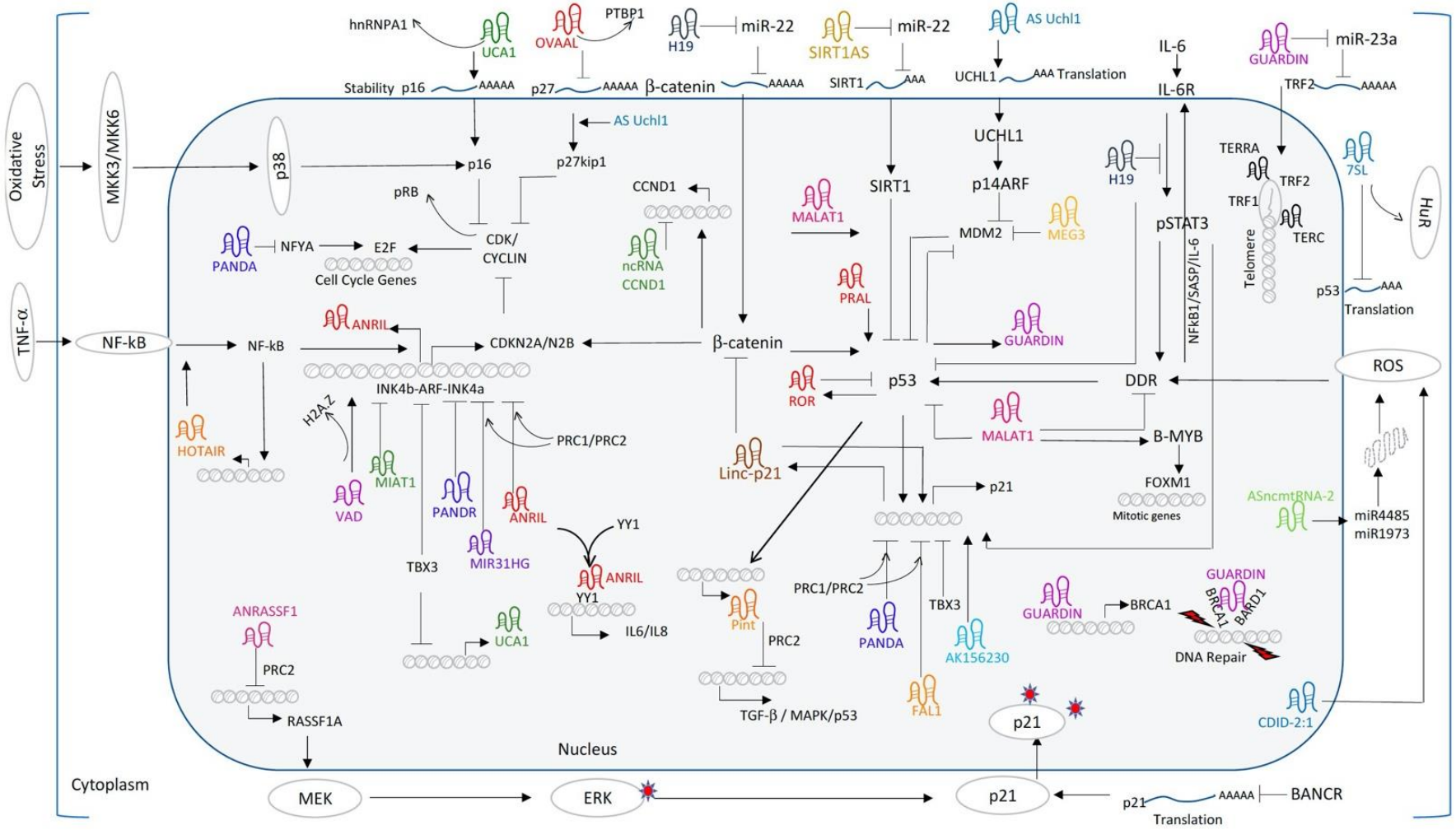


Figure 1.9: lncRNAs affect senescence pathways. External and internal factors trigger intracellular signalling pathways, including MAPK, NF-kB, p53/p21, Rb/p16, IL-6/STAT3, b-catenin, and DDR, whose activation initiates cell cycle inhibition and promotes pro-senescence markers. lncRNAs function at different stages either by activator or inhibitor of gene regulation. Figure from Puvvula et al., 2019.

Aim of the study

Senescence process is known to have a key physiological role to ensure the removal of damaged cells and the renewal of the tissues.

This thesis project aims to unveil novel lncRNAs playing a key role in the activation of cellular senescence. After an initial characterization of our senescence model, also from the epigenetic point of view, nine lncRNAs were selected among the upregulated genes in senescent conditions to perform a CRISPR/Cas9 KO and evaluate the effect on cellular life-span. Five of them displayed to delay the senescence occurrence, suggesting their involvement in senescence activation. A preliminary characterization consisting in the inspection of their subcellular localization, revealed that they belong to different cellular compartments, and therefore we expected them to act and intervene on gene expression regulation at different levels.

Further experiments are required for the functional characterization of these five lncRNAs, in order to define in detail how they affect senescence activation and, in particular, which are their molecular partners and targets.

Results

Senescence model characterization

Senescence state establishment

Cellular senescence is defined as a durable state of cell cycle arrest in which cells cease to proliferate, but maintain some functions of their replicative counterpart. For the purpose of unveiling the role of lncRNAs in senescence activation, we began setting up a replicative senescence model. We chose as cellular model human diploid IMR-90 fibroblasts, since they are currently employed for studying senescence phenomenon. Replicative senescence can be recapitulated *in vitro* passaging cells until replicative exhaustion.

Therefore, proliferating IMR-90 cells were passaged once a week and the cell number at each passage was recorded in order to monitor the population doubling (hereafter PDs). This parameter refers to the total number of times the cells in the population have doubled since their primary isolation in vitro. As shown in Figure 2.1A, IMR-90 fibroblasts duplicated their number for a fixed number of times, roughly between 26.5 and 27 PD, from the first thawing. We performed the experiment in two biological replicates and both times the cells reached the same final PD.

Observing the cells during the transition towards senescence, the morphological changes between the young proliferating fibroblasts and the old senescent ones were evident, in particular the enlarged size and the flatten shape of the cells with the increasing of the passage number. In order to visualize these differences, we performed immunofluorescence with phalloidin to stain actin filaments, and tubulin to visualize microtubules (Figure 2.1B).

The senescent state establishment was evaluated by the senescence-associated β-galactosidase (SA-β-gal) assay. This is probably the most common assay to verify the senescence level of a cell culture, that takes advantage of the higher number of lysosomes in senescent cells, which corresponds to an increase in the SA-β-gal activity. As visible from Figure 2.1C and 2.1D, only 5% of early PD IMR-90 were blue, while final PD

fibroblasts, considered as senescent, were almost all blue, as expected (Figure 2.1C and 2.1D).

Finally, to confirm the senescent state achievement of our cellular model, we checked the kinetics of three well-known senescence markers at three PDs, from early passages until senescence:1) lamin B1, that decreases because of alterations of the nuclear lamina (Freund et al., 2012); 2) p21, known key player in senescence and thus expected to be upregulated in late PDs and 3) histone H3, as it was reported that with senescence the expression of histones is depleted (O'Sullivan et al., 2010) (Figure 2.1E).

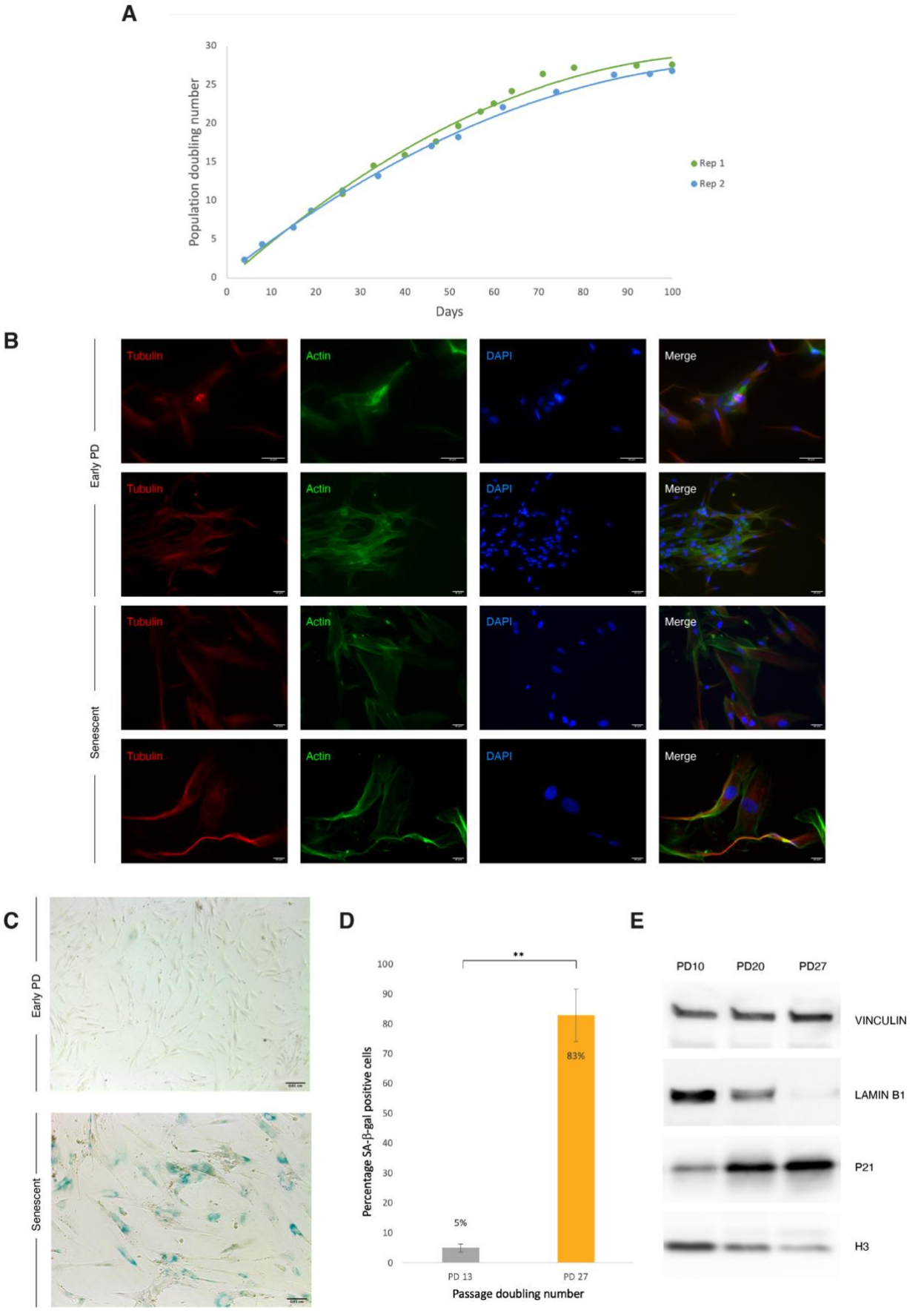


Figure 2.1 (legend in the next page)

Figure 2.1: senescence model characterization

A. Replicative life span curves of two biological replicates of IMR-90 cells maintained in culture until they naturally reached senescence. Y-axis represents the cumulative population doubling, that is the number of divisions that the cells underwent in culture. X-axis represents time (in days).
B. Representative images of early PD and senescent IMR-90 cells. Microtubules are visualized by tubulin staining and actin by phalloidin staining.
C. Representative images of SA-β-gal stained IMR-90 cells, at early PD and senescent stage.
D. The number of b-gal-positive cells represented as a percentage of total counted cells. Data are presented as mean values ± SD of 3 independent experiments; **P<0.01, *P<0.05 by Student's t-test.
E. Western blots showing lamin B1, p21 and histone H3 levels in early PD and senescent IMR-90 cells.

Transcriptome analysis of senescent IMR-90

In order to better characterize our senescence model also from the transcriptomic point of view, we performed RNA-sequencing (RNA-seq) on proliferating (PD 7) and senescent (PD 27) IMR90 cells, with three biological replicates per each condition. From the principal component analysis (PCA, Figure 2.2A), it was possible to affirm that samples clustered separately based on the cell state. Performing differential expression (DE) analysis, we could identify 1403 upregulated and 888 downregulated genes in old versus young (Figure 2.2B). Gene Ontology (GO) analysis on the DE genes, highlighted that the downregulated genes were enriched in known downregulated senescence-associated processes, such as cell cycle, DNA repair and DNA replication; among the others also metabolic processes and ribosomes biogenesis as already reported in literature (Marthandan et al., 2016; Park et al., 2021). In contrast, the upregulated genes were enriched in secretion, glycosylation, extracellular structure organization and cell matrix adhesion, terms related to known characteristics or processes of cellular senescence, in particular the hyperadhesive cell phenotype of senescent cells and the activation of SASP program (Park et al., 2021; Shin et al., 2020).

Recently new gene sets were proposed as senescence transcriptome signature, obtained merging data from different cell types and senescence induction conditions, and therefore proposed as more reliable than traditional markers. We verified the behavior of these genes in our RNA-seq data and from the results we were able to confirm the robustness of our senescence model (Figure 2.2D).

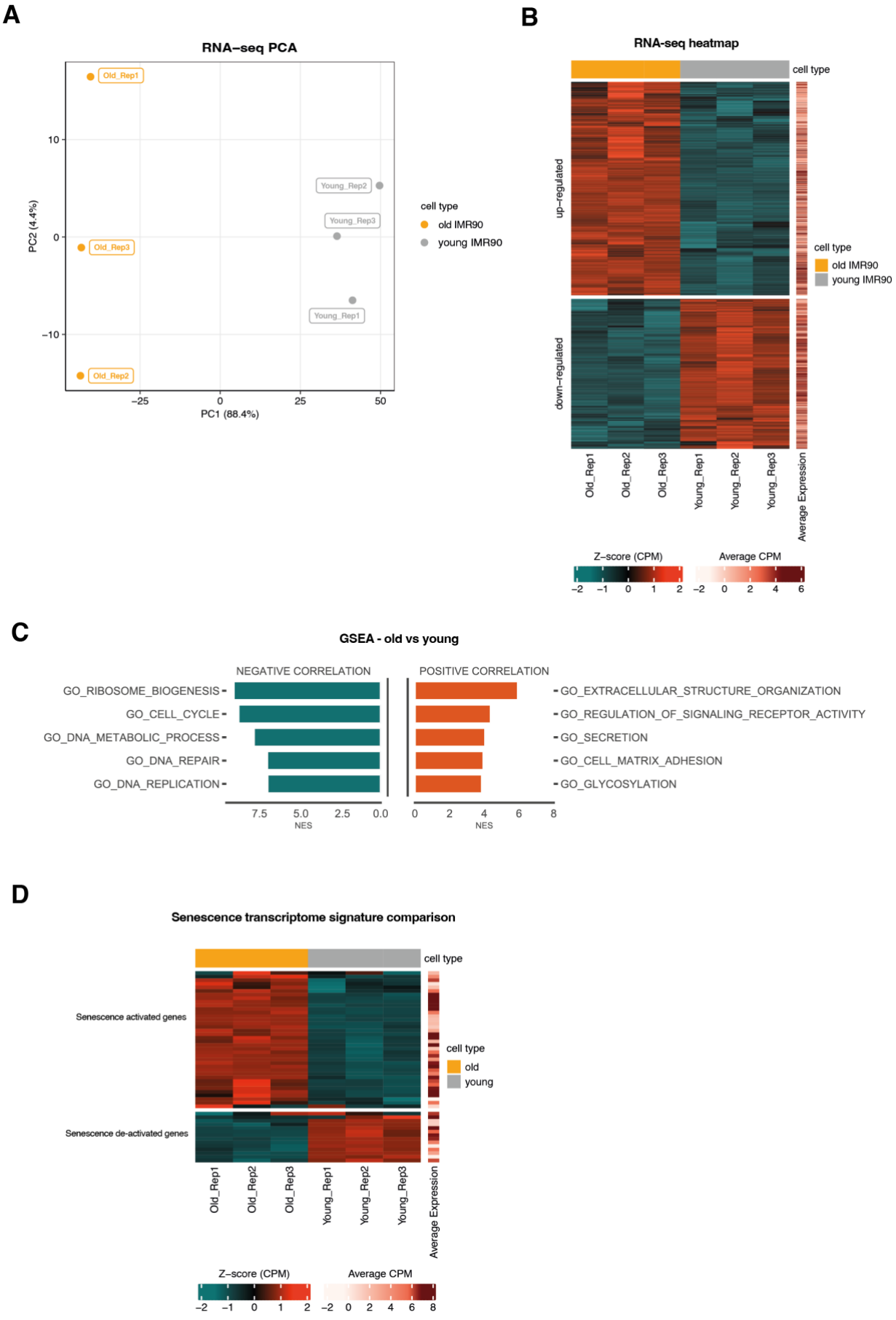


Figure 2.2 (legend in the next page)

Figure 2.2: RNA-seq analysis of proliferating and senescent IMR-90 cells

A. Principal component analysis (PCA) plot showing the distribution of the biological triplicates of the proliferating (grey) and senescent (orange) IMR-90 cells.
B. Heatmap reporting all the differentially expressed genes between proliferating and senescent IMR-90 cells. Upregulated genes are shown in red, downregulated genes in green.
C. Top correlation gene sets from gene set enrichment analysis (GSEA) in GO biological pathways.
D. Heatmap showing the behaviour in our senescence model of the senescence transcriptome signature genes published by Casella et al.. As shown, senescence-activated genes are upregulated in our system, and senescence-deactivated genes are downregulated (Casella et al., 2019).

Chromatin state assessment

Data from literature suggests that senescence involves a profound reorganization of the chromatin with major changes in the enhancer landscape. These changes consist in the inactivation of regulatory regions close to genes related to cell cycle progression and proliferation, and in parallel in the activation of new super-enhancer regions that allow the progression of the senescence program (Martínez-Zamudio et al., 2020; Sen et al., 2019; Tasdemir et al., 2016).

In order to better characterize our model from the epigenetic point of view, we mapped the accessible chromatin sites by assay for transposase-accessible chromatin sequencing (ATAC-seq), performing the experiment in three biological replicates on both proliferating and senescent IMR-90 cells. From differential accessibility analysis, we obtained 19040 gained peaks in senescent IMR-90 cells compared to proliferating ones, and 13806 lost peaks (Figure 2.3A).

In parallel we profiled the acetylation changes at H3K27 (H3K27ac) in young proliferating and senescent cells. For this purpose, we performed chromatin immunoprecipitation-sequencing (ChIP-seq) in two biological replicates on both cell conditions. The differential analysis highlighted 5934 gained peaks and 2179 lost peaks in senescent versus proliferating fibroblasts (Figure 2.3C).

Heatmap and boxplot of ATAC-seq and H3K27ac ChIP-seq (Figures 2.3B and 2.3D) of proliferating and senescent fibroblasts showed strong enrichment of the signal in the second condition compared to early passage cells.

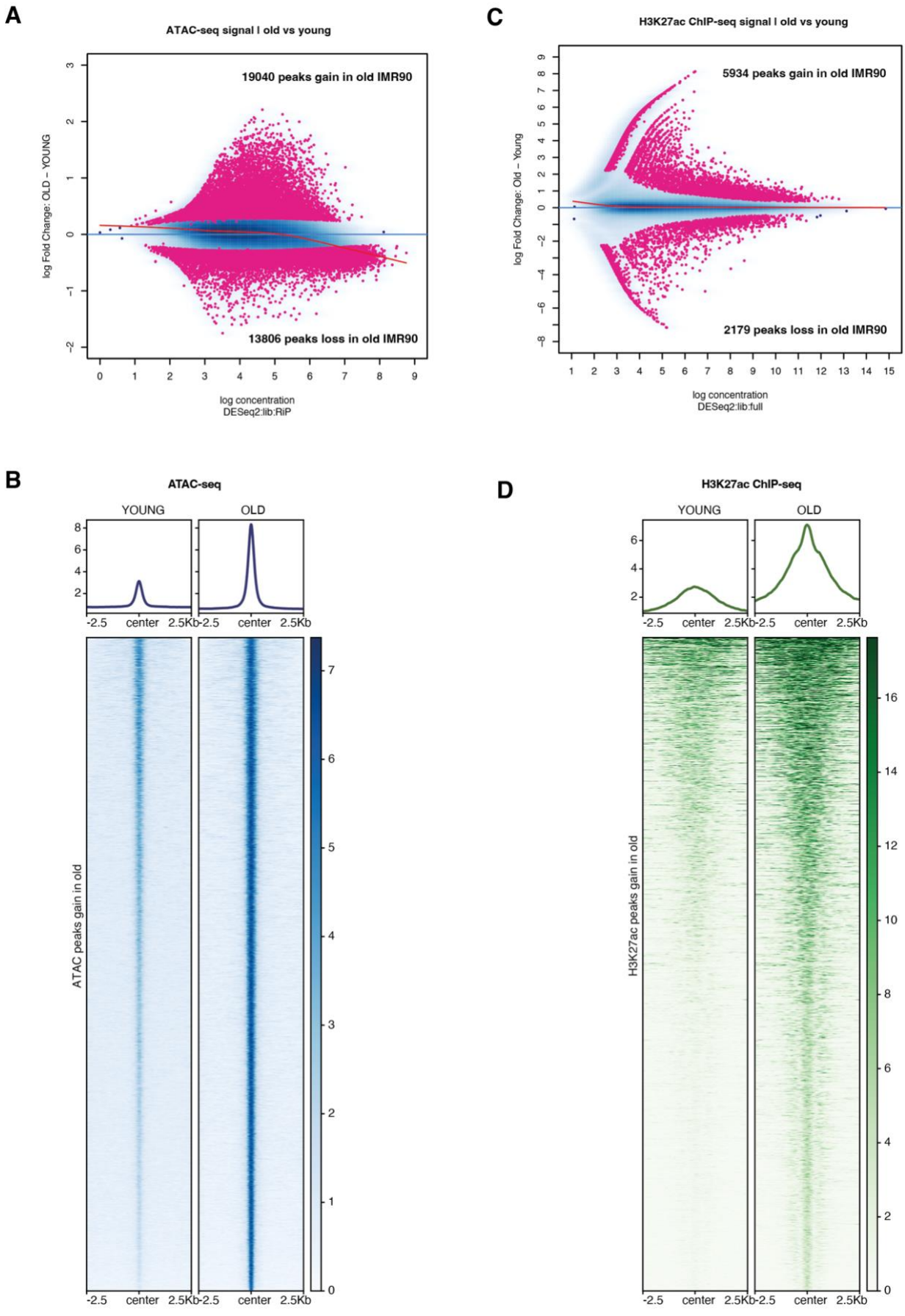


Figure 2.3 (legend in the next page)

Figure 2.3: chromatin state inspection of proliferating and senescent IMR-90 cells

A. MA plot of ATAC-seq signal which is a scatter plot of log2 fold changes of old compared to young IMR-90 signal (M, on the y-axis) versus the average expression signal (A, on the x-axis). Differentially accessible peaks are shown in magenta, with 13806 lost peaks and 19040 gained peaks in senescence.
B. Metaplot and heatmap of ATAC-seq gained peaks in senescence in young and old IMR-90 cells.
C. MA plot of H3K27 ac ChIP-seq signal which is a scatter plot of log2 fold changes of old compared to young IMR-90 signal (M, on the y-axis) versus the average expression signal (A, on the x-axis). Differentially accessible peaks are shown in magenta, with 2179 lost peaks and 5934 gained peaks in senescence.
D. Metaplot and heatmap of H3K27ac ChIP-seq gained peaks in senescence in young and old IMR-90 cells.

Senescence-activated regions definition

Focusing on the regions that are activated in senescent state, we intersected the data from ATAC-seq and ChIP-seq analyses and identified 1800 regions that display a more open chromatin and a higher acetylation level at H3K27 in senescent cells, that we defined as senescence-associated activated regions (Figure 2.4A).

We checked how they were annotated on GeneHancer database and they resulted to be mainly indicated as enhancer regions (Figure 2.4B). This is in line with recent publications reporting the emergence of new super-enhancer regions in senescent cells and indicate the enhancer landscape as a crucial driver of senescence process (Martínez-Zamudio et al., 2020; Sen et al., 2019).

Inspecting the peaks resulting as gained from ATAC-seq or ChIP-seq, we defined which could be their putative target genes, in two possible ways. If the peak is already annotated on GeneHancer database as regulatory regions, and known to interact with a specific target, that gene was assigned as putative target of the peak. On the contrary, for peaks that were not found as already annotated, the nearest transcription start site (TSS) was picked up as putative target gene. Taking in account this association between peaks and targets, we inspected whether there could be a correlation between the accessibility signal or the acetylation signal at H3K27, and the differential expression of the target gene. We found out that there is a positive correlation between the differential expression of the target and the differential accessibility and acetylation state of the regulatory region. Therefore, the target genes of the 1800 senescence-activated regions are mostly upregulated genes in our senescence model (Figure 2.4C and 2.4D).

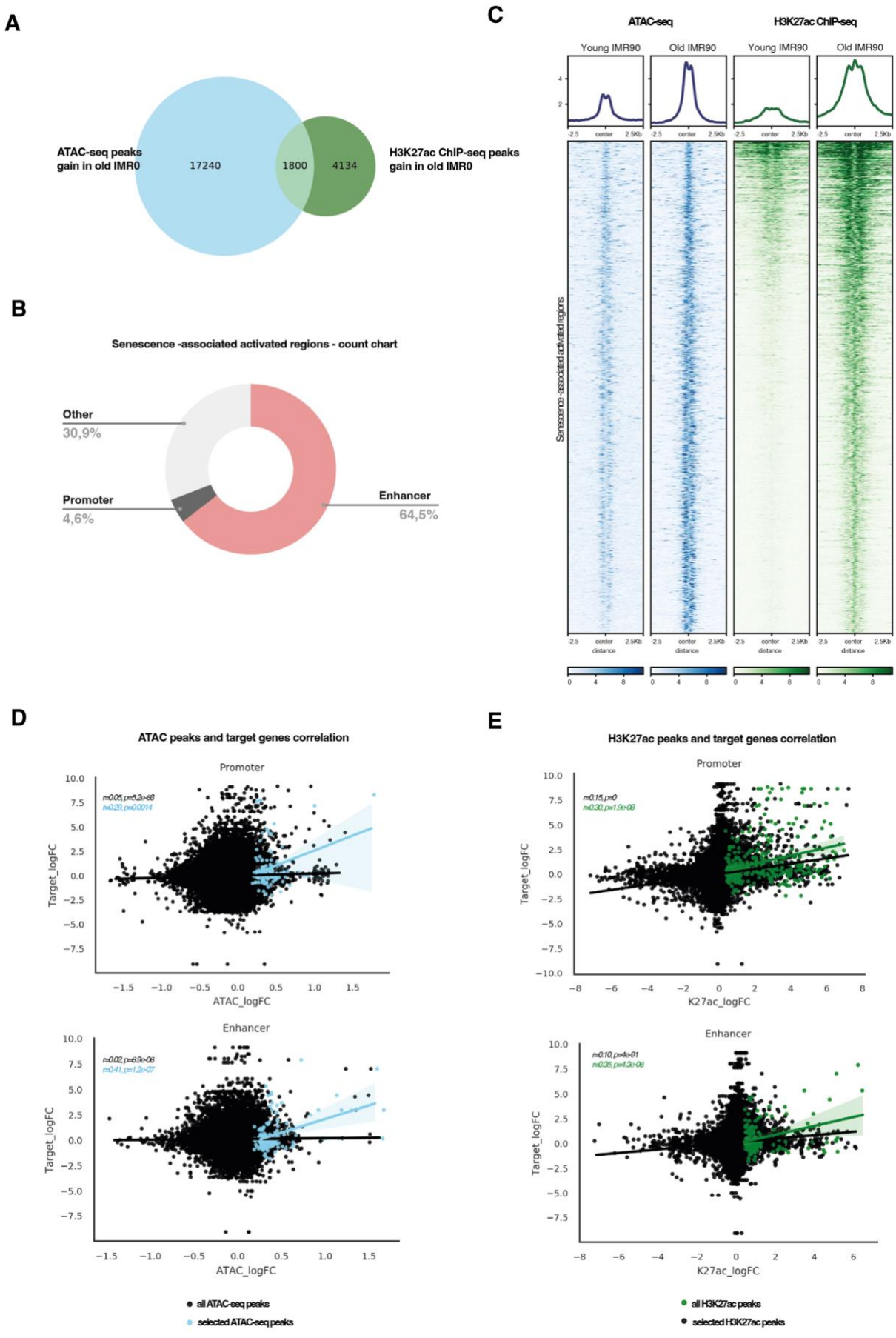


Figure 2.4 (legend in the next page)

Figure 2.4: Senescence-associated activated regions definition

A. Venn diagram representing the intersection between the gained peaks in senescence in ATAC-seq and H3K27ac ChIP-seq experiments. The 1800 regions in common are named senescence-associated activated regions.
B. Representation of the annotation of the 1800 regions on GeneHancer database as enhancer, promoter and other (exonic regions or intronic regions non annotated as enhancers).
C. Metaplot and heatmap of ATAC-seq (in blue) and H3K27ac ChIP-seq (in green) comprised in the 1800 senescence-associated activated regions.
D. Scatter plot showing the correlation between log2 fold change of ATAC-seq signal and log2 fold change of their putative targets. Correlations regarding the 1800 regions are highlighted in blue.
E. Scatter plot showing the correlation between log2 fold change of H3K27ac ChIP-seq signal and log2 fold change of their putative targets. Correlations regarding the 1800 regions are highlighted in green.

LncRNAs in cellular senescence

Selection of lncRNAs as candidates for knock-out screening

To proceed towards our goal of defining the role of lncRNAs in senescence process, we looked at our differentially expressed genes in RNA-seq focusing on lncRNA transcripts. We identified 39 lncRNAs significantly downregulated out of 888 total downregulated genes in senescence; on the other hand, we found out 171 lncRNAs significantly upregulated in senescence out of 1403 upregulated genes in total (Figure 2.5A). To focus on those that could be involved in the activation of senescence, we filtered the 171 upregulated transcripts based on three different parameters: 1) their expression level, choosing as arbitrary threshold 3 reads per kilobase of exon per million mapped reads (RPKM), in order to increase the accuracy of the fold-induction calculations; 2) the log2 fold change of their expression between early and late passages IMR-90 cells that was required to be at least 1.5; 3) the characteristics of the genomic loci were also took into account to select only intergenic, in order not to affect other genes with the editing for obtaining the knockouts, and multiexonic lncRNAs, as it was demonstrated that they are more probably conserved among mammals. After applying these filters, we found out 9 lncRNAs to be screened in order to evaluate their putative role in senescence, indicated as lncRNA1 to lncRNA9 (Figure 2.5B and 2.5C).

For the purpose of the screening, one additional lncRNA was selected as internal control of the screening, indicated as lncRNA0. We chose a transcript that based on the RNA-seq was not expressed in proliferating and in senescent IMR-90 cells (Figure 2.5B).

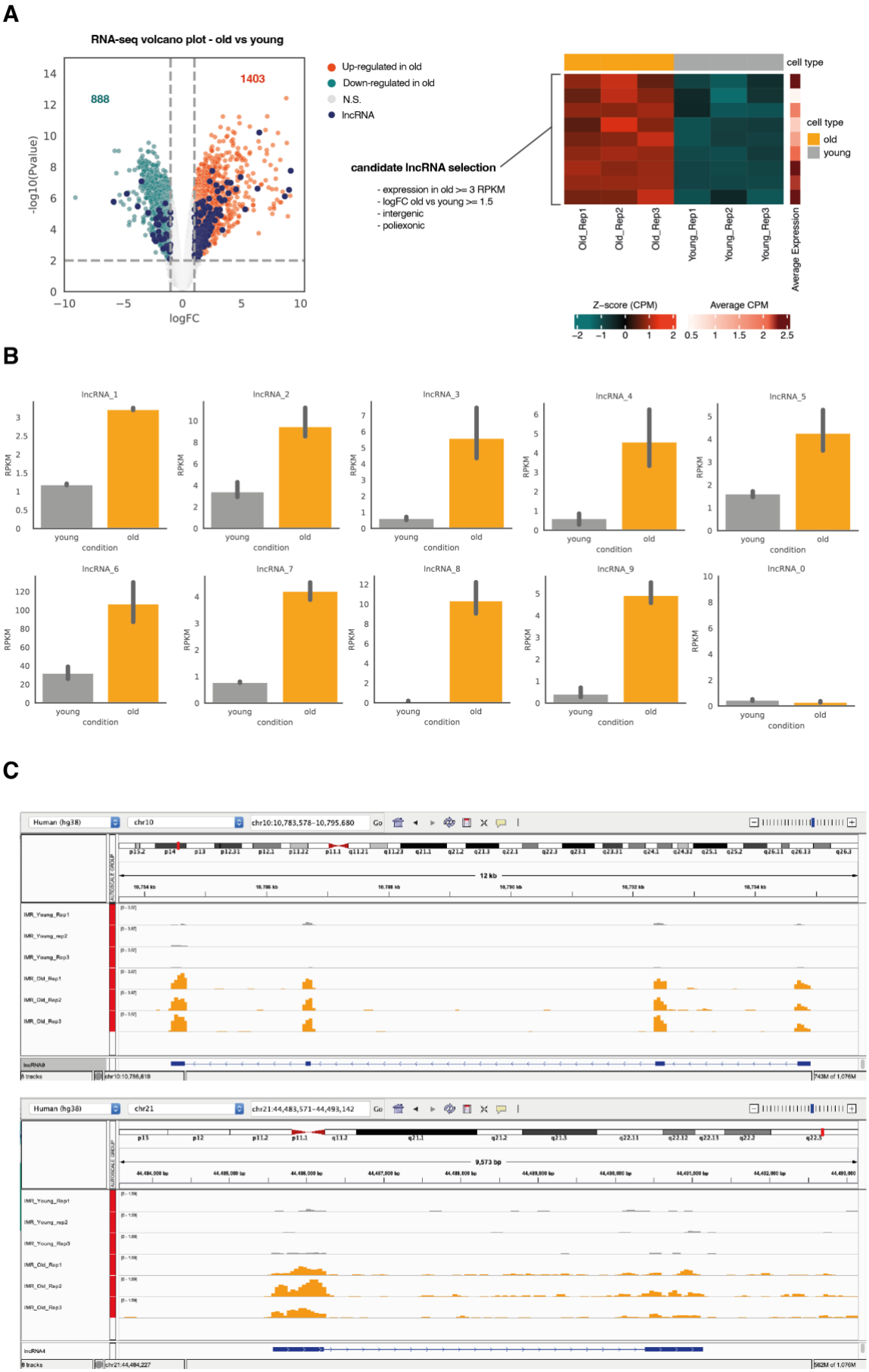


Figure 2.5 (legend in the next page)

Figure 2.5: selection of lncRNAs for knock-out screening

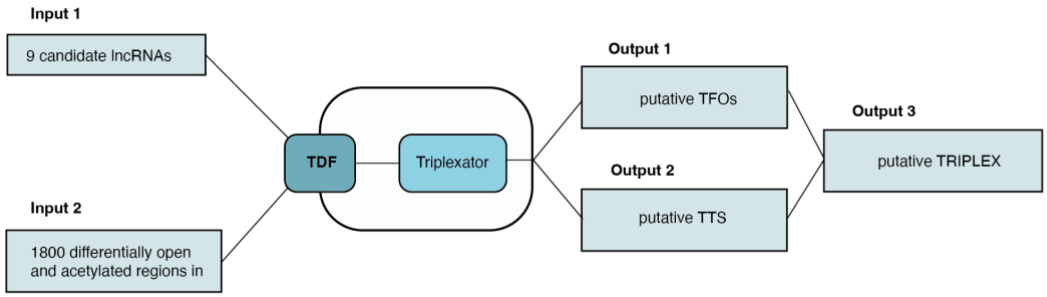
A. Volcano plot of the DEG with a log2 fold change of at least 1. Downregulated genes are shown in green, while upregulated ones are in red. DE lncRNAs are highlighted in dark blue. The filters applied to select the candidates are listed and heatmap of the differential expression signal of the nine chosen lncRNAs in shown on the right.
B. Box plot representing the expression level of the selected lncRNAs in young (grey) and old (orange) IMR-90 cells. lncRNA0 box plot shows that the internal control is not expressed in proliferating and in senescent IMR-90 cells.
C. Expression profiles of two representative lncRNAs (lncRNA9 on the top, lncRNA4 at the bottom) in the biological triplicates for young (grey) and old (orange) IMR-90 cells.

Triple helical structures prediction

One mechanism exploited by lncRNAs to regulate gene expression is targeting specific DNA sequences by forming RNA-DNA triple helical structures (triplexes). Therefore, we wondered whether the selected lncRNAs could be able to form triplexes with the 1800 senescence-activated regions. For this attempt, we exploited TDF (Triplex Domain Finder) that is a statistical framework for evaluating triplexes predicted by TRIPLEXATOR (triple-helix locator). This method firstly identifies potential triplex forming oligos (TFOs) within RNA transcripts, determines triplex target sites (TTS) within DNA regions, and evaluates the compatibility of TFOs and TTS potential partners based on the canonical triplex formation rules (Buske et al., 2012). Subsequently TDF defines the DNA binding domains (DBDs) by seeking for contiguous RNA regions with overlapping TFOs and it compares the number of TTS formed by a given DBD in target DNA regions, with the number of TTS of the same DBD in background regions.

We performed the analysis using as inputs the nine lncRNAs selected from our RNA-seq data and the senescence-associated activated regions, and obtaining from TDF-TRIPLEXATOR analysis the number of triplexes that each lncRNA could form, with the indication of the sequence region of the RNA that could interact. From the results, we could assess that all nine candidate lncRNAs could significantly form triplex interactions within the senescence-activated regions, with different propensities between the lncRNAs as indicated by the ranking based on z-score.

A



B

		target regions		non target regions			
lncRNA	**DBD**	**with TTS**	**without TTS**	**with TTS (average)**	**s.d.**	**p.value**	**z-score**
lncRNA_2	253-277	461	1339	350.8	17.0	0	6.49
lncRNA_7	142-155	331	1469	257.7	15.0	0	5.27
lncRNA_4	325-341	363	1437	292.2	15.6	0	4.53
lncRNA_9	0-17	417	1383	344.0	17.3	0	4.23
lncRNA_8	358-369	372	1428	312.4	16.8	0	3.56
lncRNA_6	532-543	409	1391	347.9	17.2	0	3.55
lncRNA_1	2072-2099	579	1221	515.6	20.2	0	3.14
lncRNA_5	396-415	346	1454	292.0	15.6	0,002	3.46
lncRNA_3	872-888	338	1462	289.1	15.6	0,001	3.14

Figure 2.6: triple helical structures prediction

A. Scheme of the triple helical structure prediction.

B. Chart containing the results (outputs) of the prediction for triplex formation between our nine lncRNAs and the 1800 senescence-associated activated regions. For every predicted triplex target site (TTS) the probability to form triplexes on random non-target regions was evaluated to calculate the significance of the triplex on the 1800 target regions.

Time course: from early passages to senescence

Before proceeding with the screening of the chosen lncRNAs, we wanted to evaluate their expression level along senescence to understand the moment in which their transcription is triggered. To do so, we thawed the earliest PD IMR-90 available, that were at PD 13, and splitted them every week until they reached senescence, at PD 27,5. In the meantime, every week, in addition to counting the cells to keep trace of the PD, cells were collected for protein and RNA extraction, from the so-called p2, corresponding to PD15, until p9, corresponding to the last PD.

As shown in figure 2.7A, p21, lamin B1 and histone H3 levels were checked, as done previously. As expected, the protein levels of lamin B1 and histone H3 were dramatically decreased during senescence process, while p21 showed a good upregulation (Figure 2.7B).

Subsequently, RNA-seq was performed in three biological replicates for all the collected time points from p2 to p9. Principal component analysis was performed including the previously analyzed young and old IMR-90 (Figure 2.7C). The distribution of the samples along the two components showed a good clustering by condition in particular at the earliest PDs; moreover, from p2 (PD 15) to p5 (PD 22,5) the samples showed a quite high similarity. From p6 (PD 24) to p8 (PD 27) the differences among the biological replicates were more evident; one explanation could be that in these steps the activation of senescence could start to take place, therefore the observed higher heterogeneity among the replicates could mirror a higher heterogeneity within the population of cells proceeding towards senescence. The last passage p9 (PD 27,5) was distributed in the PCA very close to the senescent IMR-90 previously sequenced and analyzed, which was a good indicator of reproducibility of our senescence model.

In Figure 2.7D the expression trend of the nine lncRNAs is shown. In particular some of the candidates, from lncRNA4 to lncRNA9, displayed an upregulation starting from p7 (PD 26,5) or p8, with a low expression level in the previous PDs. On the contrary, lncRNA1, lncRNA2 and lncRNA3 besides the upregulation in the last stages, also showed an increment in their expression level at p2 or p3 (PD 17,5). This could suggest the involvement of these three lncRNAs also in other processes other than senescence.

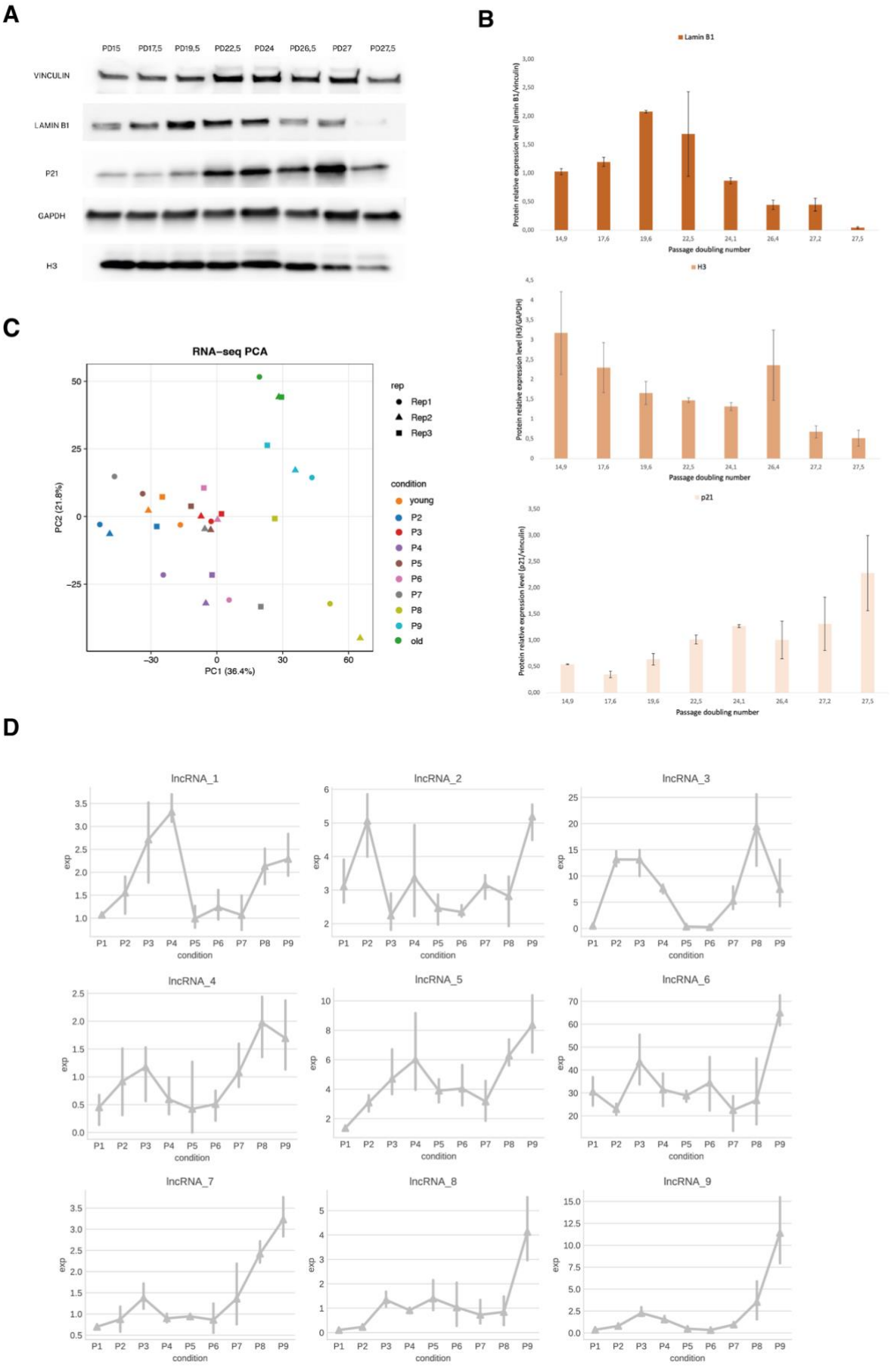


Figure 2.7 (legend in the next page)

Figure 2.7: time course from early passage IMR-90 to senescent cells

A. Western blots showing lamin B1, p21 and histone H3 levels in early PD and senescent IMR-90 cells. Vinculin and GAPDH were used as housekeeping genes. The PDs indicated on the top corresponds to passages from p2 to p9 of the time course of senescence performed in triplicate.
B. Quantification of the western blot band performed with Fiji-ImageJ, normalizing the signal of lamin B1 and p21 over vinculin, and histone H3 over GAPDH. Data are presented as mean values ± SD of 2 independent experiments; **P<0.01, *P<0.05 by Student's t-test.
C. Principal component analysis (PCA) plot showing the distribution of the biological triplicates of the time course of senescence, and in addition also the previously analyzed IMR90 proliferating (young) and senescent (old). Biological triplicates are shown as three different symbols.
D. Scatterplot showing for each lncRNA selected for KO screening the expression level across the time course from p2 to p9.

CRISPR/Cas9 knock-out of selected lncRNAs

In order to evaluate the possible role of these nine lncRNAs in senescence process, we proceeded with the editing of proliferating IMR-90 by CRISPR/Cas9 to obtain knock-out (KO) cells for each transcript. The design of this experiment required to take into account the characteristics of our cellular model. In fact, it was not possible to obtain CRISPR/Cas9 single-cell clones, since IMR-90 are primary cells with a limited proliferation capacity. For this reason, the genome editing was carried out in bulk population, as previously reported in literature (Martufi et al., 2019). Moreover, the KO was performed interrupting the lncRNA transcription by polyadenylation (poly(A)) signal insertion instead of gene deletion, in order to make sure not to affect the expression and/or the regulation of other genes in proximity of the genomic locus.

To select the edited population, we cloned the homology arms, required for the homologous recombination and editing, into a donor vector carrying a neomycin resistance cassette (Figure 2.8A). In particular, the homology arms were obtained by "nested" PCR: two couples of primers were designed based on the wild type genome, one external pair of primers, one internal. The latter is also characterized by the presence of restriction enzymes sites at their ends, necessary for the subsequent cloning into the donor vector (Figure 2.8B).

IMR-90 were electroporated at PD18 (corresponding to P3 of the time course) with three different DNA plasmids: one encoding for the Cas9 protein; a second plasmid for the guide RNA (gRNA) to direct the Cas9 enzyme to cleave the target genomic location; and a third donor vector containing the homology arms, the poly(A) signal sequence, and the neomycin resistance cassette for cell selection. The KO of lncRNA0 was performed in parallel with the others in order to have an internal control.

After electroporation, IMR-90 cells were treated with Neomycin (200ug/ml) for 7 days, to select the edited cells, and subsequently they were passaged until they reached senescence. In the meantime, genome editing was checked by PRC on genomic DNA to verify the insertion of the poly(A) signal at the desired locus (Figure 2.8C and 2.87D). For lncRNA2, lncRNA7 and lncRNA8 KOs, it was not possible to obtain a positive amplification band (Figure 2.8E). This could be explained by the size of the amplicon and repetitive DNA sequence of these loci.

Additionally, when the lncRNAs started to be expressed in lncRNA0 KO IMR-90, real-time quantitative PCR (RT-qPCR) was performed for each KO and compared with the internal control in order to assess the effective silencing of the lncRNAs expression. As shown in Figure 2.8F, all the KOs for the selected lncRNAs displayed a significant reduction of the transcript expression level; it is important to consider that, differently from single-clone KOs, in a bulk population there is the possibility to have both homozygous and heterozygous edited cells, therefore the repression can be incomplete.

The only KO that didn't show a significant reduction of the signal was the one for lncRNA6; a possible explanation is the presence of alternative isoform for this transcript that were not repressed by the poly(A) signal insertion, as verified by genomic PCR and subsequent Sanger sequencing for the correct insertion into the desired locus.

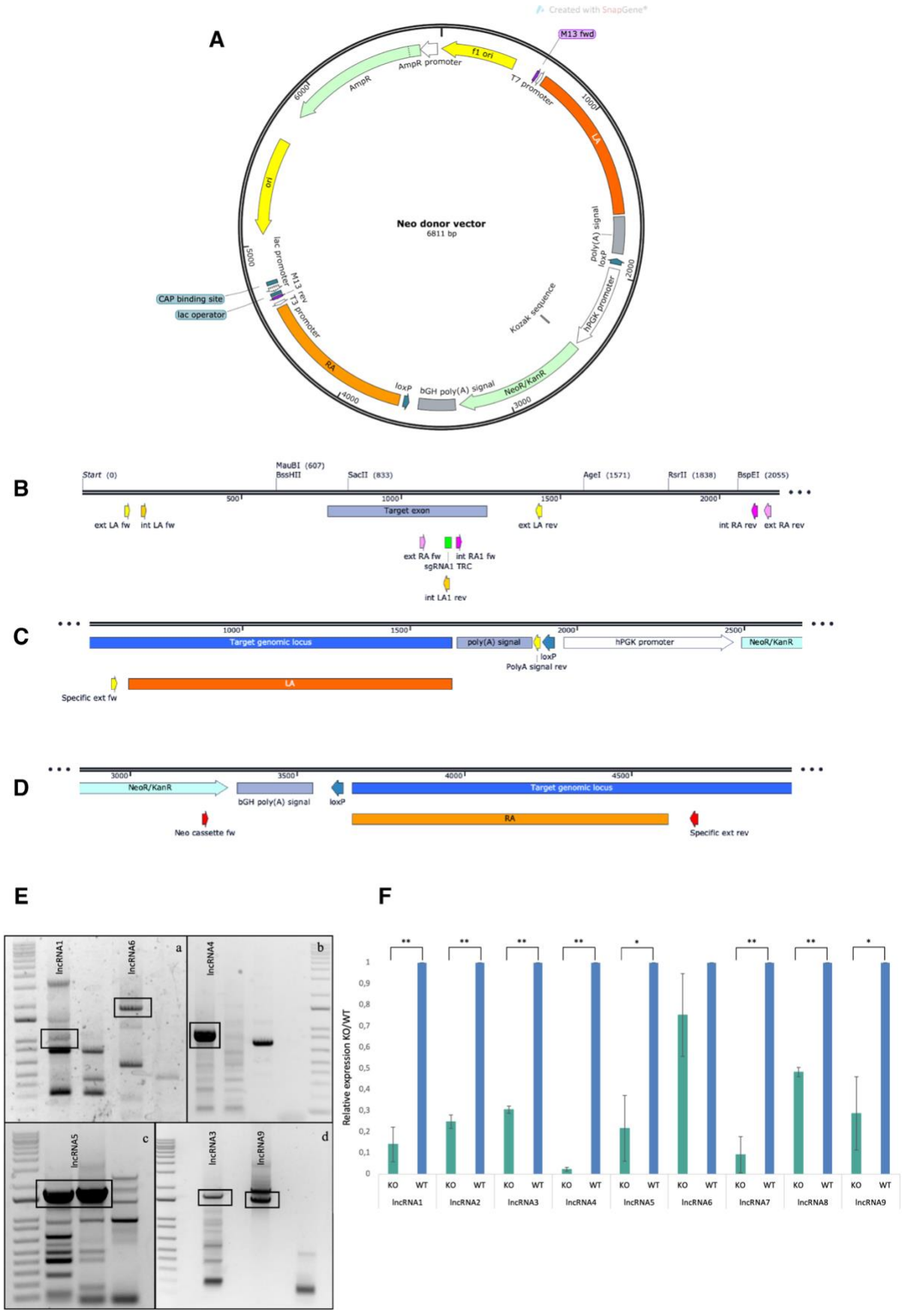


Figure 2.8 (legend in the next page)

Figure 2.8: knock out experiment design and validation

A. Map of the donor vector employed for the lncRNAs KO in IMR-90 cells. It contains the neomycin resistance cassette (light green) flanked by two loxP sites (blue arrows), the left and right arms for the homologous recombination (orange), and the poly(A) signal (grey).
B. Design of the primers for the nested PCR to amplify and clone the left and the right arm. External left arm primers, indicated in yellow as ext_LA_fw and ext_LA-rev, and the corresponding internal primers harboring the enzyme restriction site for the cloning into the donor vector, indicated in orange as int_LA_fw and int_LA1_rev. External right arm primers, indicated in pink as ext_RA_fw and ext_RA-rev, and the corresponding internal primers harboring the enzyme restriction site for the cloning into the donor vector, indicated in magenta as int_RA1_fw and int_RA_rev.
C. For the left arm side, design of a specific forward external primer and a reverse internal primer, to amplify specifically the region of the targeted locus and the inserted poly(A) signal.
D. For the right arm, design of a forward internal primer and a specific reverse external primer, to amplify specifically the region between the Neomycin resistance cassette and the specific targeted locus. Representative images of SA-b-gal stained IMR-90 cells, at early PD and senescent stage.
E. PCR performed on genomic DNA of IMR-90 cells from edited cells and wild type DNA used as negative control. The correct bands are squared in black. a. Correct bands for lncRNA1 and lncRNA6 KO; b. Correct bands for lncRNA5 KO; c. Correct bands for lncRNA3 and lncRNA9 KO; d. Correct bands for lncRNA4 KO.
F. RT-qPCR of lncRNAs KO IMR-90 cells showing a significant repression of the signal of the lncRNAs compared to the wild type, with the exception of lncRNA6. Data are presented as mean values ± SD of 3 independent experiments; **P<0.01, *P<0.05 by Student's t-test.

Phenotypic characterization of lncRNAs KO cells

IMR-90 KO cells were monitored after the neomycin selection step to understand whether there was a delay in replicative senescence establishment in comparison to the internal control lncRNA0 KO cells. To this attempt, the population doubling number was evaluated counting the cells at each passage, to finally compare the growth curves of KO populations in contrast with the control KO. The experiment was performed in two biological replicates. The growth curves are depicted in Figure 2.9A, in which the passages before the electroporation are reported, in order to show a more correct curve fitting. Compared to the control lncRNA0, five lncRNAs exhibited a relevant delay in senescence assessment: lncRNA2, lncRNA3, lncRNA4, lncRNA8 and lncRNA9. We considered as arbitrary threshold the passages over PD 25, since in the KO experiment the control ceased to proliferate at PD 24-24,5 (Figure 2.9B). The other lncRNAs, lncRNA1, lncRNA5, lncRNA6 and lncRNA7, either had a similar behavior to lncRNA0, or continued to proliferate beyond the internal control only in one of the two replicates (Figure 2.9B).

Furthermore, as validation of what observed from the growth curves, we also performed the SA-β-gal staining for each lncRNA KO. Comparing the percentage of stained cells of the lncRNA0 with the other KOs, the lncRNAs KOs that exhibited a delay in senescence also had a significantly lower percentage of positive cells, while the others had similar percentages to the control (Figure 2.9C).

Finally, we inspected the localization of these five lncRNAs. We performed a subnuclear fractionation followed by RT-qPCR. lncRNA2 and lncRNA3 showed equal distribution among the three cellular compartments, cytoplasm, nucleoplasm and chromatin; lncRNA4 and lncRNA8 were found in the chromatin fraction only; on the contrary, lncRNA9 was mostly cytoplasmic. Therefore, the five lncRNAs we found out to be involved in senescence, based on their diverse subcellular compartment localization, seemed to regulate the process at different levels.

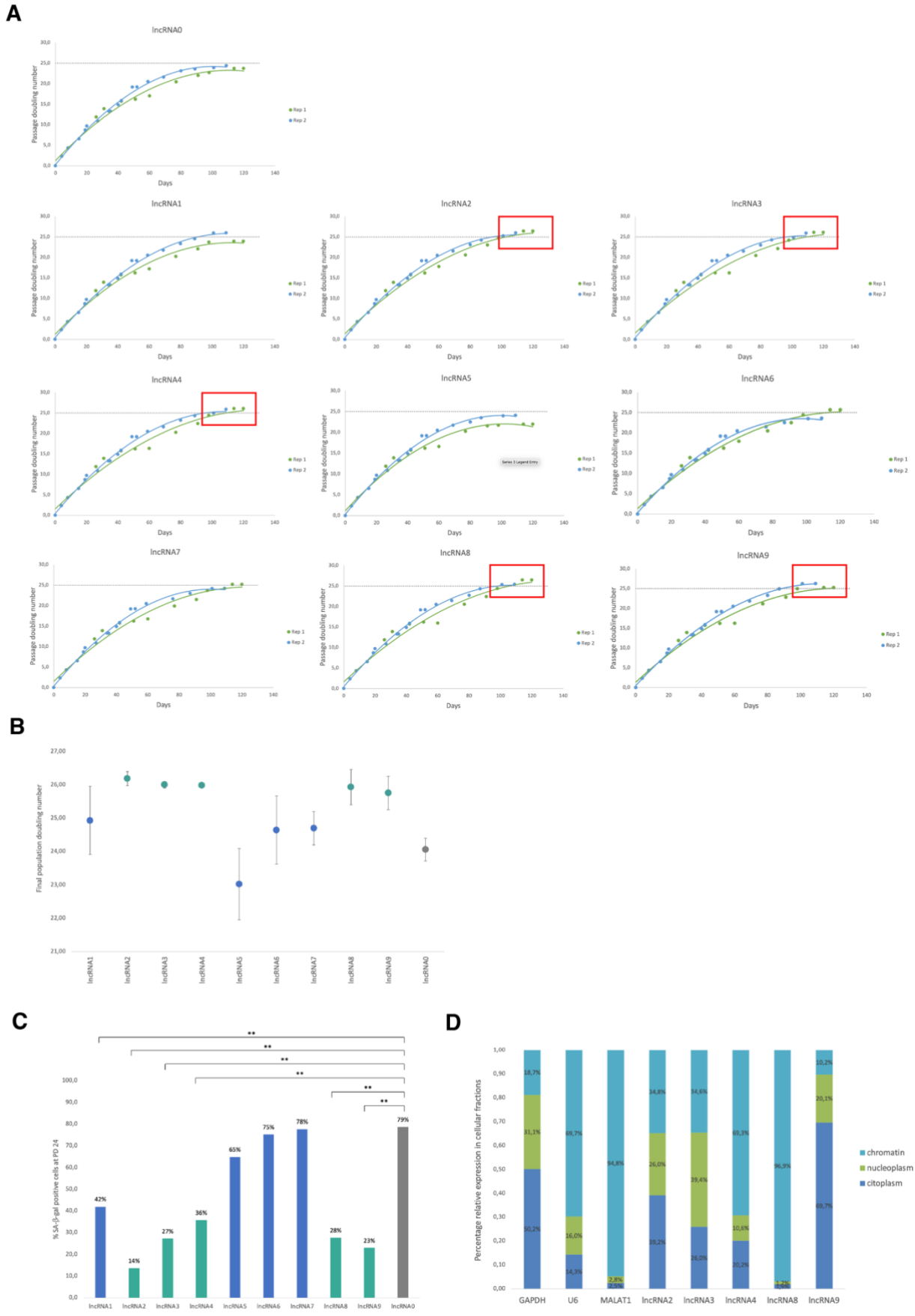


Figure 2.9 (legend in the next page)

Figure 2.9: phenotypical characterization of lncRNAs KOs

A. Replicative life span curves of two biological replicates of each lncRNA KO IMR-90 cells, maintained in culture until they naturally reached senescence. Y-axis represents the cumulative population doubling, that is the number of divisions that the cells underwent in culture. X-axis represents time (in days). The electroporation was performed at PD 16 for replicate 1 and at PD 19 for replicate 2 of the KO experiment.
B. Scatter plot of the final PDs reached by each lncRNA KO population. The internal control lncRNA0 is shown in grey, the selected five lncRNAs for next experiments are in green. Data are presented as mean values ± SD of the 2 biological replicates of each lncRNA KOs.
C. The number of b-gal-positive cells represented as a percentage of total counted cells for each lncRNA KO IMR-90. Data are presented as mean values ± SD of 2 independent experiments; **P<0.01, *P<0.05 by Student's t-test.
D. RT-qPCR of subcellular fractionation of senescent IMR-90 cells. Percentage of the five chosen lncRNAs in each fraction over the whole total expression levels represents the distribution of lncRNAs in fibroblasts. Controls of subcellular fractionation are GAPDH for cytoplasm, and U6 for nucleoplasm and MALAT1 for chromatin fraction.

Discussion and future perspectives

Senescence is triggered by the activation of a chronic DNA damage response (DDR), accompanied by the activation of cyclin-dependent kinase inhibitors (CDKi), increased production and secretion of proinflammatory and tissue-remodelling factors that constitute the so-called senescence associated secretory phenotype (SASP). The activation of these signalling pathways leads to structural alterations of the cells, such as an enlarged and more flattened morphology, accumulation of lysosomes and mitochondria, and nuclear changes (Hernandez-Segura et al., 2018; Kumari and Jat, 2021). To this end, we first started by the definition and characterization of our senescence model. This was performed by traditional β-galactosidase assay for staining of senescent cells and evaluating the kinetics of known senescence-related markers. We then proceeded with the analysis of the transcriptome of proliferating and senescent cells, to obtain a list of the differentially expressed genes (DEG) and we compared it to signature dataset published in literature to verify the senescent state of our cells (Casella et al., 2019; Marthandan et al., 2016).

From the literature it is known that a profound chromatin reorganization takes place during senescence, involving changes in the enhancer landscape that allow the repression of genes related to proliferation and cell cycle progression, and, on the contrary, the activation of regions related to senescence-associated genes (Hao et al., 2022).

We examined the epigenetic changes that occur with senescence activation, mapping the accessible chromatin sites by assay for transposase-accessible chromatin sequencing (ATAC-seq), and profile the acetylation changes at lysine 27 histone H3 (H3K27) in young proliferating and senescent cells. We intersected the senescence-gained peaks obtained from the two analyses and we obtained 1800 regions that showed a more open chromatin and contemporarily a gain in H3K27 acetylation, that we named senescence-activated regions. These regions are mainly annotated as enhancer regions. This finding is in line with recent publications about the activation of enhancer and super-enhancer regions during senescence (Sen et al., 2019).

Aiming to unveil lncRNAs involved in the activation of the senescence process, we selected from RNA-seq analyses the long non-coding transcripts upregulated in senescent

cells, based on the following criteria: 1) expression level above 3 RPKM and 2) Old versus Young log2 fold change greater than 1.5.

The involvement of the lncRNAs in senescence process was evaluated performing a knock-out for each candidate through CRISPR/Cas9 strategy on proliferating cells. As internal control we chose one lncRNA from the RNA-seq data among the non-expressed transcripts both in young and old cells. After selection of the knockout population, the cells were passaged until senescence to observe any variation in the life-span of KO compared to control cells. Five lncRNAs showed a delay in senescence occurrence measured in terms of two passages. This result was reinforced by quantification of cells stained for senescence-associated β-galactosidase activity, showing that these five lncRNAs had a lower percentage of senescent cells compared to the internal control at the same passage.

LncRNAs have been demonstrated to participate to gene expression regulation in diverse contexts with several mechanisms of action (Geisler and Coller, 2013). A comprehensive analysis about long non-coding transcript showed that they tend to be enriched in the nucleus if compared to protein-coding mRNAs (Derrien et al., 2012). A large number of them is known to interact with chromatin modifiers assembling in RNA-protein complexes to regulate gene expression at the transcriptional level.

Interestingly, four lncRNAs out of the five that we found to display a role in senescence delay were enriched in chromatin fraction and therefore could be involved in transcriptional regulation.

Triple helical structures (triplex) prediction performed to evaluate the probability for the selected lncRNAs to form triplexes with the 1800 senescence-activated genomic regions. We found out that all these lncRNAs had a good score as probability to form triplexes on our regions of interest, in comparison to random regions used as internal control. In particular, on the basis of our triplex formation prediction, we know they can form triple helices on the senescence-associated activated regions, that in our model are more accessible and showed enrichment of H3K27 acetylation in senescent IMR-90 compared to proliferating cells. Looking at the putative targets of these regions and focusing on the ones that are upregulated in senescence state, a subgroup is represented by inflammation-related genes, such as CXCR4, RORα, TLR4 (Bhattacharyya et al., 2018; Loewer et al., 2010; Tian et al., 2019). Consequently, we may hypothesize their involvement in

expression modulation of the expression of these genes occurring through interaction with transcription or epigenetic factors activating the senescence-associated regions. To validate this hypothesis, we will analyze through RNA-seq which are the signalling pathways affected by the lncRNAs repression. Therefore, changes in H3K27 acetylation profile could be inspected to check if chromatin changes occurred with the lncRNAs KO. Moreover, RNA antisense purification assay (RAP) or chromatin isolation by RNA purification (ChIRP) (Chu et al., 2011; Engreitz et al., 2015) will be performed in order to unveil the molecular interactors cooperating with our lncRNAs for the regulation of senescence-associated gene expression.

In the cytoplasm, lncRNAs were proposed to function as ceRNAs to fine tune miRNAs availability, hijacking them from their protein-coding target RNAs (Kallen et al., 2013; Thomson and Dinger, 2016). Therefore, we hypothesized a possible role for our lncRNA9, that localized in the cytoplasm of senescent IMR-90 fibroblasts. Looking at possible miRNAs interactors searching for recognition elements on its sequence, we found some miRNAs known to be involved in proliferation. Among these, miR-221 is actually downregulated in our model in the late PDs while it is expressed in young proliferating cells. From literature it is known to promote migration and proliferation, and to suppress apoptosis in a PTEN/Akt mediated manner (Sun et al., 2020). The opposite expression trend with respect to the lncRNA9 could suggest that our lncRNA may act as sponge for miR-221, counteracting its function and therefore promoting senescence assessment. To investigate this hypothesis, we will evaluate the number of predicted miRNA responsive elements on lncRNA9. Performing small RNA-seq on both lncRNA9 KO sample and lncRNA0 KO as control, we could be able to assess whether the expression levels of the predicted miRNAs vary while repressing the lncRNA. We would expect an increase of the miRNA transcript if it is sequestered by our lncRNA. Therefore, the effectiveness of the sponge activity could be confirmed ectopically expressing the lncRNA9 and the miRNA monitoring its expression by a sensor system.

Experimental procedures

Senescence model characterization

Cell culture

IMR90 cells (ATCC CCL-186) which are diploid female primary lung fibroblasts were grown in standard tissue culture medium (DMEM supplemented with 10% FBS and 1% penicillin/streptomycin; Invitrogen) at 3% oxygen. Replicative senescence was induced passaging the cells until growth ceased and cells were maintained in dishes for 2 weeks to ensure growth termination. Markers such as p21 expression and b-gal were tested to confirm senescence. For each passage, cells were washed with PBS, trypsinized at 37°C for 2 min, and plated on fresh 10 cm plates previously coated with 0,1% gelatin at 0.5-1e6/plate or fresh 15cm plates at 1-2e6/plate. Cells were counted with a Neubauer counting chamber, and the numbers were recorded for growth curve generation.

Immunofluorescence and image acquisition

IMR-90 cells were grown in monolayer on sterile coverslips in 96 well-plates. After 24 hours of seeding, cells were gently washed with PBS and fixed using 4% paraformaldehyde in PBS for 20 minutes. Cells were then gently washed with PBS three times and subsequently incubated 5 minutes with 0.3% Triton X-100 for permeabilization. After one wash in PBS, cells were blocked with PBS/1% BSA for 1 hour. Subsequently cells were incubated with tubulin diluted 1:400 in PBS/1% BSA for 1 hour. After three washes in PBS, secondary antibody diluted 1:1000 was added, together with TRITC–phalloidin diluted 1:200 for actin staining. After three washes with PBS, DAPI diluted 1:10000 was incubated for one minute for nuclei staining. Coverslips were rinsed with PBS and mounted using Prolong Gold. Images are representative of at least 15 images taken
from two independent experiments. Images were acquired on a Leica (Wetzlar, Germany) DMI3000B microscope equipped with a Leica-DFC310X camera using the LAS AF software.

SA-b-gal staining

SA-b-gal assays were performed using a cellular senescence assay kit (Chemicon KAA002), according to the manufacturer's protocol. Cells were incubated over night with b-gal detection solution at 37°C overnight and imaged by light microscopy. Quantification of the percentage of positive cells was performed counting the blue stained cell with Fiji ImageJ software, and staining the cells with Hoechst for the total number of cells.

Protein extraction and western blot

Cells were lysed in F-buffer (10 mM Tris–HCl pH 7.0, 50 mM NaCl, 30 mM Napyrophosphate, 50 mM NaF, 1% Triton X-100, protease inhibitor cocktail). The lysates were briefly sonicated, and cleared by centrifugation at max speed for 10 min at 4°C. Extracts were quantified using BCA assay (Pierce) and were run on SDS-PAGE gels in Biorad Mini-PROTEAN chambers, according to the manufacturer's protocol. Gels were transferred to nitrocellulose membranes, blocked in 5% milk in TBST for 1 hour at RT rocking platform and incubated with specific primary antibodies overnight 4°C, followed by 5 times washes with TBST and probed with secondary antibody for 1 hour at RT and later developed by using ECL reagent (GE Healthcare Amersham).

RNA-seq

Total RNA from proliferating and senescent IMR-90 cells (three biological replicates) was isolated using TRIzol reagent (Invitrogen), according to the manufacturer's protocol. Quantity and quality of the starting RNA were checked by Qubit and Agilent Bioanalyzer 2100. 1mg of total RNA was subjected to poly(A) selection, and libraries were prepared using the TruSeq RNA Sample Prep Kit (Illumina) following the manufacturer's instructions. Sequencing was performed on the Illumina NextSeq 500 platform.

RNA-seq analysis

Sequencing was performed on the Illumina NextSeq 500 platform. The sequenced reads were aligned to human reference genome (UCSC hg38) using STAR v2.7.1a (DOI: 10.1093/bioinformatics/bts635) with parameters -outFilterMismatchNmax 999 -outFilterMismatchNoverLmax 0.04 and providing a list of known splice sites extracted

from GENCODE 32 comprehensive annotation. Gene expression levels were quantified with featureCounts v1.6.3 (DOI:https://doi.org/10.1093/bioinformatics/btt656) with options -t exon -g gene_name using GENCODE 32 basic gene annotation. Multi-mapped reads were excluded from quantification. Gene expression counts were next analyzed using the edgeR package (DOI:https://doi.org/10.1093/bioinformatics/btp616). Normalization factors were calculated using the trimmed-mean of M-values (TMM) method implemented in the *calcNormFactors* function and RPKM were obtained using normalized library sizes and gene length. After filtering lowly expressed genes (below 1 CPM in three or more samples) differentially expression analysis was performed by fitting a GLM to all groups followed by a LF test applied to the interesting pairwise contrasts. Genes were defined as significantly differentially expressed if they showed a | logFC | > 1 and a Pvalue < 0.01 in each relevant comparison. Gene expression heatmaps were generated using the *seaborn.clustermap* function (DOI:https://doi.org/10.21105/joss.03021).

ATAC-seq

The transposition reaction and library construction were performed as previously described17. Briefly, 100,000 cells from proliferating and senescent IMR-90 fibroblasts (three biological replicates) were collected, washed in 100ml PBS and centrifuged at 500g at 4°C for 5 min. Nuclei were extracted by incubating cells in nuclear extraction buffer (containing 10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl2, 0.1% v/v NP-40) and immediately centrifuging at 500g at 4°C for 10 min. The supernatant was carefully removed by pipetting, and the transposition was performed by resuspending nuclei in 100 µl of Transposition Reaction Mix containing 1x TD Buffer (Illumina) and 5 µl Tn5 (Illumina) for 30 min at 37°C. DNA was extracted using a Qiagen MinElute Reaction Cleanup kit. Libraries were produced by PCR amplification (12 cycles) of tagmented DNA using a NEB Next High-Fidelity 2x PCR Master Mix (New England Biolabs). Library quality was assessed using an Agilent Bioanalyzer 2100. Paired-end sequencing was performed in an Illumina Hiseq 2500. Typically, 30–50 million reads per library were required for downstream analyses.

ATAC-seq analysis

The sequencing reads were processed with the ENCODE ATAC-seq pipeline (v1.9.0 https://github.com/ENCODE-DCC/atac-seq-pipeline) using the default parameters. Bowtie2 (DOI:https://doi.org/10.1038/nmeth.1923) was used in order to align reads to the human reference genome UCSC hg38. After the discard of duplicated, multi-mapping and poor-quality alignments, the peak calling was performed with MAC2 (DOI:https://doi.org/10.1186/gb-2008-9-9-r137) generating also the signal tracks as fold enrichment control. Both Young and Old samples were processed in the same way.

Differentially opened regions between Old and Young samples were identified using DiffBind (v3.0.3 DOI:https://doi.org/10.3389/fgene.2015.00169) with parameters: normalize = DBA_NORM_LIB, library = DBA_LIBSIZE_PEAKREADS, background = BKGR_FALSE, AnalysisMehod = DESEQ2. Peaks with pvalue < 0.05 and logFC > 0 were considered as regions showing significantly more open chromatin levels as compared to the Young samples.

ChIP-seq

A total of 15x10^6 cells from proliferating and senescent IMR-90 fibroblasts (two biological replicates) were fixed in 1% formaldehyde for 10 min, quenched in 0.125 M glycine for an additional 5 min, washed in cold PBS and pelleted by centrifugation at 1,500 r.p.m. at 4°C for 5 min. Cells were resuspended in Lysis Buffer (20 mM Tris-HCl pH 8.0, 2 mM EDTA, 150mM NaCl, 1% triton X-100, 0.15% SDS and proteinase inhibitor cocktail). Sonication was performed using a Diagenode Picoruptor for 20 cycles to obtain the desired average fragment size (100–500 bp). Soluble chromatin was obtained by centrifugation at 11,000 r.p.m. for 10 min at 4°C. Immunoprecipitation was performed overnight at 4°C with rotation using 5 µg antibodies. Streptavidin beads (Dynabeads Protein G, Life Technologies) were saturated with 1% BSA/PBS by overnight incubation at 4°C. Immunoprecipitated samples were incubated with saturated beads for 2 hours at 4°C. Successively immunoprecipitated complexes were washed one time in low-salt buffer (150 mM NaCl, 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.0), one time in high-salt buffer (500 mM NaCl, 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.0), one time in lithium chloride buffer (250 mM LiCl, 1% NP-40, 1% sodium deoxycholate, 1 mM EDTA, 10 mM Tris-HCl pH 8.0)

and two times in TE buffer (10 mM Tris-HCl, 1 mM EDTA). Elution was performed with Elution buffer (1% SDS, 150 mM NaCl, in TE buffer) at room temperature for 40 min with rotation. The decrosslinking was performed at 65°C overnight. Decrosslinked DNA was purified using QIAQuick PCR Purification Kit (QIAGEN) according to the manufacture's instruction.

ChIP-seq analysis

After quality controls performed with FastQC v0.11.2, sequencing reads were processed with Trim Galore! v0.5.0 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore) in order to perform quality and adapter trimming (parameters: --stringency=3). Trimmed reads were then analyzed with the ENCODE Transcription Factors and Histone Modifications ChIP-seq pipeline 2 (v1.6.1 from https://github.com/ENCODE-DCC/chip-seq-pipeline2) using the 'Histone Modification' processing mode with default parameters. In this step reads were aligned to the human reference genome UCSC hg38 using Bowtie2 (DOI:https://doi.org/10.1038/nmeth.1923). The peak calling was performed using MACS2 (DOI:https://doi.org/10.1186/gb-2008-9-9-r137) after the discard of duplicated, multi-mapping and poor-quality aligned reads, and using the input DNA as control library. Signal tracks were generated as fold enrichment control for both individual and pooled replicates using MACS2. Both Young and Old samples were processed in the same way.

Differentially H3K27-acetylated regions between Old and Young samples were identified using DiffBind (v3.0.3) (DOI:https://doi.org/10.3389/fgene.2015.00169) with parameters: normalize = DBA_NORM_LIB, library = DBA_LIBSIZE_FULL, background = BKGR_FALSE, AnalysisMethod = DBA_DESEQ2. Peaks with FDR< 0.05 and logFC > 0 were considered as significantly over-acetylated regions in Old samples with respect to the Young samples.

Senescence-activated regions definition

We defined as senescence-activated regions the DNA regions that resulted having a more open chromatin and showing an higher acetylation level on H3K27 in Old samples with respect to the Young ones (from H3K27ac ChIP-seq and ATAC-seq analysis of both Young and Old IMR90 cells).

Target gene of senescence-activated regions definition

The senescence-activated regions were associated to their target genes using as reference the GeneHancer (GH) database (https://doi.org/10.1093/database/bax028) (UCSC hg38) filtered for 'regulatory region-target gene' pairs in IMR90 cell type. The senescence-activated regions that were not mapped in GH have been defined as targeting the gene of the nearest TSS (within 5 kb).

LncRNAs in cellular senescence

Triplex forming predictions

Triplex Domain Finder (TDF) - Regulatory Analysis Toolbox (RGT) version: 0.13.1 (https://doi.org/10.1093/nar/gkz037) was used with the processing mode 'Genomic Region Test' and parameters -ccf 0.1 -n 1000 -mp 10 -l 10 -e 20 -par L_-1_-E_-1 in order to predict the capability of our selected lncRNA of interest to form triplex with DNA in senescence-activated regions.

Genomic DNA isolation

Cells were washed with 2 mL of PBS, then they were incubated at 37°C for 2 minutes with 120 µl trypsin to detach them. The action of trypsin was blocked with 1.5 mL of fresh medium, and the cells were collected in a 2 mL Eppendorf and finally centrifuged at 1000 rpm. The pellet of cells is then resuspended in 200 µL PBS and DNA was extracted with the DNeasy Blood & Tissue Kit (Qiagen, 69504 and 69506) according to manufacturer's instruction. The DNA quantification was performed with NanoDrop (ThermoFisher, ND-2000).

Generation of KO IMR-90

LncRNAs KO IMR-90 cells were generated by using a CRISPR/Cas9-based approach for inserting polyA signal. A donor plasmid containing a 5' HA with bGH polyA signal, the Neomycin resistance cassette and a 3' HA was built by cloning PCR fragments of each lncRNA genomic region from IMR-90 cells, into a modified version of PGKneolox2DTA plasmid (Addgene #13443). Briefly, around 1000 bp 3'HA and 5'HA were amplified by nested PCR with specific primers for each lncRNA and cloned into the PGKneolox2DTA

vector. The plasmid containing sgRNA targeting to the selected of each lncRNA were designed using the CRISPR Design Tool (http://crispr.mit.edu/). Oligonucleotides corresponding to the two strands the sgRNA were annealed and cloned into the BbsI-digested gRNA backbone (BB) previously cloned into TOPO™TA vector (Invitrogen) (Addgene #42335). Primers for amplification and cloning are listed in Table 1, which includes the corresponding references.

These plasmids, together with the Cas9- containing plasmid (Addgene #41815), were co-transfected into IMR-90 cells by electroporation using Neon Transfection System (Invitrogen), following manufacturer's protocol specific for IMR-90 cells. Subsequently, IMR-90 were selected with 200mg/ml G418 (Sigma) after 24 h and maintained under drug selection for one week.

Electroporation

IMR-90 cells were harvested and counted in order to have 2 x 10^6 cells per condition. After centrifugation, cells were resuspended in Resuspension Buffer R 1x10^7/ml.

DNA plasmids (donor vector, Cas9 vector, sgRNA vector) were added to the cells in a 1.5 microcentrifuge tube, not exceeding 10% of the total volume. The Neon™ tube was filled with 3 mL of Electrolytic Buffer.

The Neon™ Pipette equipped with the Neon™ Tip were used to aspirate the cells-plasmids suspension, which were finally inserted in the Neon™ Tube where the electroporation reaction occurred according to the manufacturer's instructions for IMR-90 cells (1 pulse, 30 ms, 1,500v). Once the electroporation occurred the cells from the Neon™ Tip were transferred into a 6-well plate, filled with warm complete medium.

The electroporated cells were incubated at 37°C for one day to recover, and then they were selected with 200ug/ml Neomycin for 7 days, at the end of selection they were passaged and counted in order to seed 200.000 cells/well.

PCR-based genotyping screening

PCR was performed on 200 ng of genomic DNA extracted from wild type and KO cells, with QuickLoad Taq 2X PCR Master mix (NEB, M0271L). Primers are listed in Table 2, which includes the corresponding references. The PCR product was separated on a 1.5% agarose gel, the gel was imaged through the ChemiDoc Imaging System.

RT-qPCR

RT-qPCR analysis Total RNA was extracted using TRIzol reagent (Invitrogen) and quantified by Nanodrop (Thermo Scientific). RealTime PCR (RT-qPCR) was performed as previously described (51) using the SensiFAST SYBR NO-ROX OneStep (BIOLINE) following the manufacturer's instructions. Briefly, RT-qPCR reactions were performed on a RotorGene Q 2plex HRM Platform (Qiagen, 9001560) and relative gene expression levels were determined using calculated concentration values, normalized to Pumilio and Gapdh as reference genes. Real-time PCR Ct values were analysed using the $2^{-\Delta\Delta Ct}$ method to calculate the fold expression (Livak and Schmittgen, 2001). For experiments with palbociclib the fold expression was calculated individually for each replicate. qPCR primers used are listed in Table 3, which includes the corresponding references.

Sub-nuclear fractionation

Cell fractionation was performed as described by Narita et al. (Narita et al., 2006). Briefly, 6x10^6 cells were resuspended in buffer A (10 mM HEPES pH 7.9, 10 mM KCl, 1.5 mM MgCl2, 0.34 M sucrose, 10% glycerol, 1 mM DTT, protease inhibitor cocktail, RNase inhibitor). 0.1% Triton X-100 was added, and the cells were incubated for 10 min on ice. Nuclei were collected by low-speed centrifugation at 1300xg at 4°C for 4 min. Supernatants were isolated as cytoplasmic fractions. Nuclei pellets were washed once in buffer A, then lysed in buffer B (3mM EDTA, 0.2mM EGTA, 1 mM DTT, protease inhibitor cocktail, RNase inhibitor) and incubated for 30 min on ice. The insoluble chromatin factions were isolated by low-speed centrifugation and supernatants were isolated as nucleoplasmic fractions. The chromatin fractions were washed once with buffer B. All fractions (whole cell, cytoplasm, nucleoplasm, and chromatin) were resuspended in Trizol reagent to proceed with RNA extraction.

Table 1: Oligonucleotide sequences used as primers for cloning sgRNA and homology arms with the indication of the corresponding genes.

Sequence	Gene	Strand	Note
CACCGCTGTCACACAAAGCCGCCGA	lncRNA1	Fw	sgRNA1
AAACTCGGCGGCTTTGTGTGACAGC	lncRNA1	Rev	sgRNA1
GCTTGCTCCCTCTGTCTC	lncRNA1	Fw	LA_ext
AGAGCGTCAGCTAGGTTG	lncRNA1	Rev	LA_ext
GCCTCACTCCTAGAGTGA	lncRNA1	Fw	RA_ext
CGGAGTTTCTCTCTTGTTG	lncRNA1	Rev	RA_ext
ACGCGTGCTCAAAGTTTCTTCGTGG	lncRNA1	Fw	LA_int_MluI
GCGGCCGCCTTTGTGTGACAGGCGTG	lncRNA1	Rev	LA_sgRNA1_int_NotI
GCTAGCTGGCAAGGCTGTAGGGCT	lncRNA1	Fw	RA_sgRNA1_int_NheI
GGTACCCGATCTCGGCTCACTGCA	lncRNA1	Rev	RA_int_KpnI
CCGCGGGCTCAAAGTTTCTTCGTGG	lncRNA1	Fw	LA_int
CCGCGGCTTTGTGTGACAGGCGTG	lncRNA1	Rev	LA_sgRNA1_int
CACCGAGAAGTAGGTGGTTTCCTAG	lncRNA2	Fw	sgRNA1
AAACCTAGGAAACCACCTACTTCTC	lncRNA2	Rev	sgRNA1
CTTGATGAGTCTGGTCCAAGT	lncRNA2	Fw	LA_ext
GAAGTCGTGCCTAGCAAATG	lncRNA2	Rev	LA_ext
CGCAAGTCTACCTGACTCTAA	lncRNA2	Fw	RA_ext
CCACAACATAACCCAGTTCTA	lncRNA2	Rev	RA_ext
CCGCGGGCATTTGACTTTGGCTTTAGC	lncRNA2	Fw	LA_int_SacII
GCGGCCGCGAAACCACCTACTTCTCCCAC	lncRNA2	Rev	LA1_int_NotI
GCTAGCTTGGGTAATGAAGCCACAGAG	lncRNA2	Fw	RA1_int_NheI
GGTACCCCCAGTTCTACAAACCTATAGC	lncRNA2	Rev	RA_int_KpnI
CACCGAGATGGCTGATCCTGCAAAG	lncRNA3	Fw	sgRNA1
AAACCTTTGCAGGATCAGCCATCTC	lncRNA3	Rev	sgRNA1
CACCGCTTTAACTTGAACCATGCC	lncRNA3	Fw	sgRNA2

AAACGGCATGGTTCAAGTTAAAGC	lncRNA3	Rev	sgRNA2
AACTTGCACCTAAGGAAAACC	lncRNA3	Fw	LA_ext
CACGACACAATCAAAACCTC	lncRNA3	Rev	LA_ext
GCATAGAACACTGGTCTCAA	lncRNA3	Fw	RA_ext
CAGGTTTTAATGGAGGGCAA	lncRNA3	Rev	RA_ext
CCGCGGAGTGTGTCCCAGTGTGTC	lncRNA3	Fw	LA1_int_SacII
GCGGCCGCCTTTGCAGGATCAGCCATC	lncRNA3	Rev	LA1_int_NotI
GCTAGCGGTGACATCAACAGTGACTC	lncRNA3	Fw	RA1_int_NheI
GATATCGATGAAAATACTTCTGCCGGT	lncRNA3	Rev	RA_int_EcoRV
CACCGAGCCGACGGAGGAGCCACAA	lncRNA4	Fw	sgRNA1
AAACTTGTGGCTCCTCCGTCGGCTC	lncRNA4	Rev	sgRNA1
ATTCACCTTCCGAGATTCC	lncRNA4	Rev	LA_ext
GAGATGCGGACTTTGGTG	lncRNA4	Fw	LA_ext
CCGCGGTTGGCACTAACAACGGGG	lncRNA4	Fw	LA_int_SacII_bis
GCGGCCGCCTCCTGGAGTTGAGCTGC	lncRNA4	Rev	LA1_int_NotI
GCTAGCTCCTCCGTCGGCTGAGAG	lncRNA4	Fw	RA1_int_NheI
GGTACCGGCACACTTGACCAACGC	lncRNA4	Rev	RA_int_KpnI
CCGCGGTTGGCACTAACAACGGGG	lncRNA4	Fw	LA_int_SacII_bis
CACCGGAGTTGGATGGACCGAATG	lncRNA5	Fw	sgRNA1
AAACCATTCGGTCCATCCAACTCC	lncRNA5	Rev	sgRNA1
CTATCGCTCCCAGATCAG	lncRNA5	Fw	LA_ext
TAGAAGCACGAGCAAGGCAG	lncRNA5	Rev	LA_ext
CACCGAGTGTTGTCCAAGAGTAAGG	lncRNA5	Fw	sgRNA2
AAACCCTTACTCTTGGACAACACTC	lncRNA5	Rev	sgRNA2
CCGCGGCTGCAATGGAGAGTCAAGCC	lncRNA5	Fw	LA_int_SacII
GCGGCCGCTTCCTGGAGACGTGGCAC	lncRNA5	Rev	LA_int_NotI
GCTAGCCATTCGGTCCATCCAACTCC	lncRNA5	Fw	RA_int_NheI
GATATCAAGTCCACCATAAGGGACC	lncRNA5	Rev	RA_int_EcoRV
CACCGCAAAGACTGACACCCCCTTG	lncRNA6	Fw	sgRNA1

AAACCAAGGGGGTGTCAGTCTTTGC	lncRNA6	Rev	sgRNA1
TTTCCAGCGATGTCACAGG	lncRNA6	Fw	LA_ext
AGGAGGTTGCTCTCATCG	lncRNA6	Rev	LA_ext
AGAAATTGTTGTTGTGGCGA	lncRNA6	Fw	LA_ext1
TAGTGCCCTCTGCAATTGAT	lncRNA6	Rev	LA_ext1
CCGCGGTGCACTCCTGGGTTACG	lncRNA6	Fw	LA1_int_SacII
GCGGCCGCGGGAAACACTCCAGTTTGC	lncRNA6	Rev	LA1_int_NotI
AAGCTTGTCTTTGTCGACTGACTCTG	lncRNA6	Fw	RA1_int_HindIII
GATATCGCGACTTGTTTAAGGTCACAC	lncRNA6	Rev	RA1_int_EcorV
CACCGTCTTCCAGATCCCACCACAG	lncRNA7	Fw	sgRNA2
AAACCTGTGGTGGGATCTGGAAGAC	lncRNA7	Rev	sgRNA2
TACAATGGGACCTTGCTCAC	lncRNA7	Fw	LA_ext
TGGCACAGACCTGTTAAGGAA	lncRNA7	Rev	LA_ext
CCGCGGGCTGGTAAAGCTTGAGGATCA	lncRNA7	Fw	LA_int_SacII
GCGGCCGCGGAAAAGAACTCCGGTGTCC	lncRNA7	Rev	LA2_int_NotI
TGTCAAGCTGAATCGCTACAT	lncRNA7	Fw	RA_ext
GGTACCTGGTGACTTTTCATTGCCAC	lncRNA7	Rev	RA_int_KpnI
GCTAGCGCTGTGAGTCTTTGGAAAGC	lncRNA7	Fw	RA2_int_NheI
TGGTTTTCGCCCTCTCTTCT	lncRNA8	Fw	LA_ext
TCCCACCTCTTTTGAATGCC	lncRNA8	Rev	LA_ext
CGTGGCACTCAGGAAATGC	lncRNA8	Fw	RA_ext
GAGGGCCATGCATCAACATT	lncRNA8	Rev	RA_ext
CGAGCTCTGTTGACTCAGAGGGTGGC	lncRNA8	Fw	Int_LA_SacI
GCGACGCGTTTCCAAGACAGGGTGGAAGC	lncRNA8	Rev	Int_LA_MluI
ACATCGGATATCATTGCATTCAAAAGAGGTGGGAGG	lncRNA8	Fw	Int_RA_EcoRV
CGCGTCGACCTAGATAATTTCAGAGTGCCG	lncRNA8	Rev	Int_RA_SalI
CACCGTGGAACACGGACACAGTCCA	lncRNA8	Fw	RA_ext
AAACTGGACTGTGTCCGTGTTCCAC	lncRNA8	Rev	sgRNA1
CTTGCAAAGGAAATCCC	lncRNA9	Fw	LA_ext

GTAGGCCCTCTAGATTGG	lncRNA9	Rev	LA_ext
GCATTATTTGTCTCTCTCTG	lncRNA9	Fw	RA_ext
ACCAACTGTGACAGGGTG	lncRNA9	Rev	RA_ext_
ACACCATTACCTACCTACC	lncRNA9	Rev	RA_ext
ACGCGTCTTGCAAAGGAAATCCC	lncRNA9	Fw	LA_int_MluI
GCGGCCGCCAGTTTTCCCAAGGGGTTCC	lncRNA9	Rev	LA2_int_NotI
GTCGACGGGTTCTTAACCTGCGGTG	lncRNA9	Fw	RA2_int_SalI
GGTACCACCAACTGTGACAGGGTG	lncRNA9	Rev	RA_int_KpnI
CACCGACTGTGTAGACAAACATTTG	lncRNA9	Fw	sgRNA2
AAACCAAATGTTTGTCTACACAGTC	lncRNA9	Rev	sgRNA2
CACCGTTCCCTCTTAAAAACAGATG	lncRNA10	Fw	sgRNA1
AAACCATCTGTTTTTAAGAGGGAAC	lncRNA10	Rev	sgRNA1
TGCACAGTGACTTGGCAG	lncRNA10	Fw	LA_ext
CTCACCACAGTGGGAAGT	lncRNA10	Rev	LA_ext
CCGCGGCTGCACCTGAATAGTTAGCC	lncRNA10	Fw	LA_int_SacII
GCGGCCGCGAGTCCTGTTTGTAAGCACAC	lncRNA10	Rev	LA1_int_NotI
TTTCCAGTCGGCCTTTGTTA	lncRNA10	Fw	RA_ext
GCCCATGTCCCAACTTTTA	lncRNA10	Rev	RA_ext
GCTAGCGAGGGAAATCTGAGTTTTCAAGG	lncRNA10	Fw	RA1_int_NheI
GCTAGCAGATGGAATTCCTAACCCGG	lncRNA10	Fw	RA2_int_NheI
GATATCGGCATCAGGTGACTACTGAG	lncRNA10	Rev	RA_int_EcoRV

Table 2: Oligonucleotide sequences used as primers for screening knockouts with the indication of the corresponding genes.

Sequence	Gene	Strand
CGGAGTTTCTCTCTTGTTG	lncRNA1	Rev
CAGGTTTTAATGGAGGGCAA	lncRNA3	Rev
CGTGCTTAGTTCTGATTGG	lncRNA4	Fw
CTCAGTTGGGCTGAGTCATC	lncRNA5	Rev
AGAAATTGTTGTTGTGGCGA	lncRNA6	Fw
ACTCCCTGGTGACCATCTC	lncRNA8	Rev
GCCCATGTCCCAACTTTTA	lncRNA10	Rev
AGGCTTACATTCCAGCTGTGT	lncRNA2	Rev
TACAATGGGACCTTGCTCAC	lncRNA7	Fw
GGGAGAGAAAGTGAGTTGTTC	lncRNA9	Rev
GGGGCTCGAATCAAGCTGAT	Neomycin cassette	Fw
TTCCGGATCTATGCATGAC	PolyA signal	Rev

Table 3: Oligonucleotide sequences used as primers for RT-qPCR with the indication of the corresponding genes.

Sequence	Gene	Strand
CAGCACAGCACTCATTCAGCT	lncRNA1	Fw
GGCACCACTCCTCAAAATGCC	lncRNA1	Rev
TTGAAGGCTTCCTGGTCTGAG	lncRNA2	Fw
AGGCTTACATTCCAGCTGTGT	lncRNA2	Rev
GCTAGCGCATGGTTCAAGTTAAAGCTG	lncRNA3	Fw
GCCGGAGTTGCAGTGATGGAC	lncRNA3	Rev
TCCAAGTGGACAGGCAACTGG	lncRNA4	Fw
GGGACCACACCCTCTTCTACC	lncRNA4	Rev
CAGGAACCCCCTCCTTACTC	lncRNA5	Fw
TAGAAGCACGAGCAAGGCAG	lncRNA5	Rev
AAGCTTGTCTTTGTCGACTGACTCTG	lncRNA6	Fw
TGCCAGATCCTGCCTGAGG	lncRNA6	Rev
ACAGGGAGCCAGGACACC	lncRNA7	Fw
GACCTTGGCACAGACCTG	lncRNA7	Rev
ATGTCCTGAATGGAAGGATCCC	lncRNA9	Fw
TCTCATCGACCCCATTCCCC	lncRNA9	Rev
TTGTGGCACGAGTAAGCCAA	lncRNA10	Fw
TCAAGGGCAATATTCCGGGT	lncRNA10	Rev
CCAGCATGGTTCTAGATTCATACAGCAAAA	Pum2	Fw
CCACGAATACGAGTAGCCAGGG	Pum2	Rev

References

Abdelmohsen, K., and Gorospe, M. (2015). Noncoding RNA control of cellular senescence. Wiley Interdisciplinary Reviews: RNA *6*, 615–629. https://doi.org/10.1002/wrna.1297.

Ali, T., and Grote, P. (2020). Beyond the RNA-dependent function of LncRNA genes. Elife *9*, 1–14. https://doi.org/10.7554/eLife.60583.

Anderson, K.M., Anderson, D.M., McAnally, J.R., Shelton, J.M., Bassel-Duby, R., and Olson, E.N. (2016). Transcription of the non-coding RNA upperhand controls Hand2 expression and heart development. Nature *539*, 433–436. https://doi.org/10.1038/nature20128.

Banani, S.F., Lee, H.O., Hyman, A.A., and Rosen, M.K. (2017). Biomolecular condensates: Organizers of cellular biochemistry. Nature Reviews Molecular Cell Biology *18*, 285–298. https://doi.org/10.1038/nrm.2017.7.

Bhattacharyya, S., Wang, W., Qin, W., Cheng, K., Coulup, S., Chavez, S., Jiang, S., Raparia, K., de Almeida, L.M. v., Stehlik, C., et al. (2018). TLR4-dependent fibroblast activation drives persistent organ fibrosis in skin and lung. JCI Insight *3*. https://doi.org/10.1172/jci.insight.98850.

Buske, F.A., Bauer, D.C., Mattick, J.S., and Bailey, T.L. (2012). Triplexator: Detecting nucleic acid triple helices in genomic and transcriptomic data. Genome Research *22*, 1372–1381. https://doi.org/10.1101/gr.130237.111.

Casella, G., Munk, R., Kim, K.M., Piao, Y., De, S., Abdelmohsen, K., and Gorospe, M. (2019). Transcriptome signature of cellular senescence. Nucleic Acids Res *47*, 7294–7305. https://doi.org/10.1093/nar/gkz555.

Chu, C., Qu, K., Zhong, F.L., Artandi, S.E., and Chang, H.Y. (2011). Genomic Maps of Long Noncoding RNA Occupancy Reveal Principles of RNA-Chromatin Interactions. Molecular Cell *44*, 667–678. https://doi.org/10.1016/j.molcel.2011.08.027.

Degirmenci, U., and Lei, S. (2016). Role of lncRNAs in cellular aging. Frontiers in Endocrinology *7*. https://doi.org/10.3389/fendo.2016.00151.

Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D.G., et al. (2012). The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. Genome Research *22*, 1775–1789. https://doi.org/10.1101/gr.132159.111.

Engreitz, J., Lander, E.S., and Guttman, M. (2015). RNA Antisense Purification (RAP) for Mapping RNA Interactions with Chromatin. (Humana Press, New York, NY), pp. 183–197.

Engreitz, J.M., Pandya-Jones, A., McDonel, P., Shishkin, A., Sirokman, K., Surka, C., Kadri, S., Xing, J., Goren, A., Lander, E.S., et al. (2013). The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. Science (1979) *341*. https://doi.org/10.1126/science.1237973.

Freund, A., Laberge, R.M., Demaria, M., and Campisi, J. (2012). Lamin B1 loss is a senescence-associated biomarker. Molecular Biology of the Cell *23*, 2066–2075. https://doi.org/10.1091/mbc.E11-10-0884.

Galluzzi, L., Vitale, I., Aaronson, S.A., Abrams, J.M., Adam, D., Agostinis, P., Alnemri, E.S., Altucci, L., Amelio, I., Andrews, D.W., et al. (2018). Molecular mechanisms of cell death: Recommendations of the Nomenclature Committee on Cell Death 2018. Cell Death and Differentiation *25*, 486–541. https://doi.org/10.1038/s41418-017-0012-4.

Geisler, S., and Coller, J. (2013). RNA in unexpected places: Long non-coding RNA functions in diverse cellular contexts. Nature Reviews Molecular Cell Biology *14*, 699–712. https://doi.org/10.1038/nrm3679.
Gong, C., and Maquat, L.E. (2011). LncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 39 UTRs via Alu eleme. Nature *470*, 284–290. https://doi.org/10.1038/nature09701.
Gorgoulis, V., Adams, P.D., Alimonti, A., Bennett, D.C., Bischof, O., Bishop, C., Campisi, J., Collado, M., Evangelou, K., Ferbeyre, G., et al. (2019). Cellular Senescence: Defining a Path Forward. Cell *179*, 813–827. https://doi.org/10.1016/j.cell.2019.10.005.
Grossi, E., Raimondi, I., Goñi, E., González, J., Marchese, F.P., Chapaprieta, V., Martín-Subero, J.I., Guo, S., and Huarte, M. (2020). A lncRNA-SWI/SNF complex crosstalk controls transcriptional activation at specific promoter regions. Nature Communications *11*. https://doi.org/10.1038/s41467-020-14623-3.
Hao, X., Wang, C., and Zhang, R. (2022). Chromatin basis of the senescence-associated secretory phenotype. Trends in Cell Biology https://doi.org/10.1016/j.tcb.2021.12.003.
Hayflick, L., and Moorhead, P.S. The serial cultivation of human diploid cell strains.
Hernandez-Segura, A., de Jong, T. v., Melov, S., Guryev, V., Campisi, J., and Demaria, M. (2017). Unmasking Transcriptional Heterogeneity in Senescent Cells. Current Biology *27*, 2652-2660.e4. https://doi.org/10.1016/j.cub.2017.07.033.
Hernandez-Segura, A., Nehme, J., and Demaria, M. (2018). Hallmarks of Cellular Senescence. Trends in Cell Biology *28*, 436–453. https://doi.org/10.1016/j.tcb.2018.02.001.
Kallen, A.N., Zhou, X.B., Xu, J., Qiao, C., Ma, J., Yan, L., Lu, L., Liu, C., Yi, J.S., Zhang, H., et al. (2013). The Imprinted H19 LncRNA Antagonizes Let-7 MicroRNAs. Molecular Cell *52*, 101–112. https://doi.org/10.1016/j.molcel.2013.08.027.
Kumari, R., and Jat, P. (2021). Mechanisms of Cellular Senescence: Cell Cycle Arrest and Senescence Associated Secretory Phenotype. Frontiers in Cell and Developmental Biology *9*. https://doi.org/10.3389/fcell.2021.645593.
Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. Methods *25*, 402–408. https://doi.org/10.1006/meth.2001.1262.
Loewer, S., Cabili, M.N., Guttman, M., Loh, Y.H., Thomas, K., Park, I.H., Garber, M., Curran, M., Onder, T., Agarwal, S., et al. (2010). Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. Nat Genet *42*, 1113–1117. https://doi.org/10.1038/ng.710.
Maldotti, M., Lauria, A., Anselmi, F., Molineris, I., Tamburrini, A., Meng, G., Polignano, I.L., Scrivano, M.G., Campestre, F., Simon, L.M., et al. (2022). The acetyltransferase p300 is recruited in trans to multiple enhancer sites by lncSmad7 . Nucleic Acids Research https://doi.org/10.1093/nar/gkac083.
Marthandan, S., Baumgart, M., Priebe, S., Groth, M., Schaer, J., Kaether, C., Guthke, R., Cellerino, A., Platzer, M., Diekmann, S., et al. (2016). Conserved senescence associated genes and pathways in primary human fibroblasts detected by RNA-seq. PLoS ONE *11*. https://doi.org/10.1371/journal.pone.0154531.
Martínez-Zamudio, R.I., Roux, P.F., de Freitas, J.A.N.L.F., Robinson, L., Doré, G., Sun, B., Belenki, D., Milanovic, M., Herbig, U., Schmitt, C.A., et al. (2020). AP-1 imprints a reversible transcriptional programme of senescent cells. Nature Cell Biology *22*, 842–855. https://doi.org/10.1038/s41556-020-0529-5.
Martufi, M., Good, R.B., Rapiteanu, R., Schmidt, T., Patili, E., Tvermosegaard, K., New, M., Nanthakumar, C.B., Betts, J., Blanchard, A.D., et al. (2019). Single-Step,

High-Efficiency CRISPR-Cas9 Genome Editing in Primary Human Disease-Derived Fibroblasts. The CRISPR Journal *2*, 31–40. https://doi.org/10.1089/crispr.2018.0047.
Mirzadeh Azad, F., Polignano, I.L., Proserpio, V., and Oliviero, S. (2021). Long Noncoding RNAs in Human Stemness and Differentiation. Trends in Cell Biology *31*, 542–555. https://doi.org/10.1016/j.tcb.2021.02.002.
Muñoz-Espín, D., and Serrano, M. (2014). Cellular senescence: From physiology to pathology. Nature Reviews Molecular Cell Biology *15*, 482–496. https://doi.org/10.1038/nrm3823.
Narita, M., Narita, M., Krizhanovsky, V., Nuñez, S., Chicas, A., Hearn, S.A., Myers, M.P., and Lowe, S.W. (2006). A Novel Role for High-Mobility Group A Proteins in Cellular Senescence and Heterochromatin Formation. Cell *126*, 503–514. https://doi.org/10.1016/j.cell.2006.05.052.
Oo, J.A., Brandes, R.P., and Leisegang, M.S. (2022). Long non-coding RNAs: novel regulators of cellular physiology and function. Pflugers Archiv European Journal of Physiology *474*, 191–204. https://doi.org/10.1007/s00424-021-02641-z.
O'Sullivan, R.J., Kubicek, S., Schreiber, S.L., and Karlseder, J. (2010). Reduced histone biosynthesis and chromatin changes arising from a damage signal at telomeres. Nature Structural and Molecular Biology *17*, 1218–1225. https://doi.org/10.1038/nsmb.1897.
Park, J.H., Ryu, S.J., Kim, B.J., Cho, H.J., Park, C.H., Choi, H.J.C., Jang, E.J., Yang, E.J., Hwang, J.A., Woo, S.H., et al. (2021). Disruption of nucleocytoplasmic trafficking as a cellular senescence driver. Experimental and Molecular Medicine *53*, 1092–1108. https://doi.org/10.1038/s12276-021-00643-6.
Park, J.T., Lee, Y.S., Cho, K.A., and Park, S.C. (2018). Adjustment of the lysosomal-mitochondrial axis for control of cellular senescence. Ageing Research Reviews *47*, 176–182. https://doi.org/10.1016/j.arr.2018.08.003.
Portoso, M., Ragazzini, R., Brenčič, Ž., Moiani, A., Michaud, A., Vassilev, I., Wassef, M., Servant, N., Sargueil, B., and Margueron, R. (2017). PRC 2 is dispensable for HOTAIR -mediated transcriptional repression . The EMBO Journal *36*, 981–994. https://doi.org/10.15252/embj.201695335.
Puvvula, P.K. (2019). Lncrnas regulatory networks in cellular senescence. International Journal of Molecular Sciences *20*. https://doi.org/10.3390/ijms20112615.
Puvvula, P.K., Desetty, R.D., Pineau, P., Marchio, A., Moon, A., Dejean, A., and Bischof, O. (2014). Long noncoding RNA PANDA and scaffold-attachment-factor SAFA control senescence entry and exit. Nature Communications *5*. https://doi.org/10.1038/ncomms6323.
Ransohoff, J.D., Wei, Y., and Khavari, P.A. (2018). The functions and unique features of long intergenic non-coding RNA. Nat Rev Mol Cell Biol *19*, 143–157. https://doi.org/10.1038/nrm.2017.104.
Rom, A., Melamed, L., Gil, N., Goldrich, M.J., Kadir, R., Golan, M., Biton, I., Perry, R.B.T., and Ulitsky, I. (2019). Regulation of CHD2 expression by the Chaserr long noncoding RNA gene is essential for viability. Nature Communications *10*. https://doi.org/10.1038/s41467-019-13075-8.
Sandhu, C., Garbe, J., Bhattacharya, N., Daksis, J., Pan, C.-H., Yaswen, P., Koh, J., Slingerland, J.M., and Stampfer, M.R. (1997). Transforming Growth Factor Stabilizes p15 INK4B Protein, Increases p15 INK4B-cdk4 Complexes, and Inhibits Cyclin D1-cdk4 Association in Human Mammary Epithelial Cells.
Santoro, F., and Pauler, F.M. (2013). Silencing by the imprinted Airn macro lncRNA: Transcription is the answer. Cell Cycle *12*, 711–712. https://doi.org/10.4161/cc.23860.
Sen, P., Lan, Y., Li, C.Y., Sidoli, S., Donahue, G., Dou, Z., Frederick, B., Chen, Q., Luense, L.J., Garcia, B.A., et al. (2019). Histone Acetyltransferase p300 Induces De

Novo Super-Enhancers to Drive Cellular Senescence. Molecular Cell *73*, 684-698.e8. https://doi.org/10.1016/j.molcel.2019.01.021.
Shin, E.-Y., Park, J.-H., You, S.-T., Lee, C.-S., Won, S.-Y., Park, J.-J., Kim, H.-B., Shim, J., Soung, N.-K., Lee, O.-J., et al. (2020). Integrin-mediated adhesions in regulation of cellular senescence.
Statello, L., Guo, C.J., Chen, L.L., and Huarte, M. (2021). Gene regulation by long non-coding RNAs and its biological functions. Nature Reviews Molecular Cell Biology *22*, 96–118. https://doi.org/10.1038/s41580-020-00315-9.
Stow, J.L., and Murray, R.Z. (2013). Intracellular trafficking and secretion of inflammatory cytokines. Cytokine and Growth Factor Reviews *24*, 227–239. https://doi.org/10.1016/j.cytogfr.2013.04.001.
Sun, L., Zhu, W., Zhao, P., Zhang, J., Lu, Y., Zhu, Y., Zhao, W., Liu, Y., Chen, Q., and Zhang, F. (2020). Down-Regulated Exosomal MicroRNA-221 – 3p Derived From Senescent Mesenchymal Stem Cells Impairs Heart Repair. Frontiers in Cell and Developmental Biology *8*. https://doi.org/10.3389/fcell.2020.00263.
Tasdemir, N., Banito, A., Roe, J.S., Alonso-Curbelo, D., Camiolo, M., Tschaharganeh, D.F., Huang, C.H., Aksoy, O., Bolden, J.E., Chen, C.C., et al. (2016). BRD4 connects enhancer remodeling to senescence immune surveillance. Cancer Discovery *6*, 613–629. https://doi.org/10.1158/2159-8290.CD-16-0217.
Thomson, D.W., and Dinger, M.E. (2016). Endogenous microRNA sponges: Evidence and controversy. Nature Reviews Genetics *17*, 272–283. https://doi.org/10.1038/nrg.2016.20.
Tian, X., Xie, G., Xiao, H., Ding, F., Bao, W., and Zhang, M. (2019). CXCR4 knockdown prevents inflammatory cytokine expression in macrophages by suppressing activation of MAPK and NF-κB signaling pathways. Cell and Bioscience *9*. https://doi.org/10.1186/s13578-019-0315-x.
Wafik S. El-Deiry, T.T.V.E.V.D.B.L.R.P.J.M.T.D.L.W.E.M.K.W.K.B.V. (1993). WAF1, a potential mediator of p53 tumor suppression. Cell *75*, 817–825. .

Acknowledgements

I would first like to thank my tutor, Professor Salvatore Oliviero, who gave me the opportunity to work in his lab and pursue my PhD. It has been a challenging experience and I really learnt a lot from our scientific discussions.

I want to acknowledge all my group colleagues for their support and the collaboration they provided to me. In particular, I would like to thank Stefania and Mara, who taught me everything that I know about molecular biology, and their patience. I want to thank Annalaura, my personal bioinformatician that I challenged too much with all my requests (especially about plot colours). Besides, I want to thank Hua who grew up with me in this lab, Valentina for her interesting suggestions and help along these years, and Fatemeh who was the best colleague possible in the worst period of my PhD. Last but not least, I want to thank Claudia and Corinne who worked with me at this project, providing me all the support that I needed.

In addition, I would like to acknowledge the reviewers of my PhD Thesis: prof. Maurizio Orlandini and prof. Claudio Santoro.

I want to acknowledge my family for supporting and encouraging me for my studies, and for trying to understand the subject of my project. Finally, I want to thank Fabio who was patiently at my side during these three years providing pleasant distractions to rest my mind outside of my research.