Imperial College London

London Institute of Medical Sciences

Telomeres, Sex and Epigenetics:

The role of mouse HP1γ on telomere stability and sexual dimorphism

by

Emmanouil Stylianakis

Thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy of Imperial College London

2022

Copyright declaration

The following way of citation is suggested:

Emmanouil Stylianakis. *Telomeres, Sex and Epigenetics: The role of mouse HP1γ on telomere stability and sexual dimorphism*. PhD thesis. Imperial College London, 2022.

Statement of originality

I hereby declare that the work presented in this PhD thesis is my own, except where due acknowledgement is made.

Acknowledgements

First and foremost, I would like to express my absolute gratitude to my two supervisors, Dr. Jean-Baptiste Vannier and Prof. Richard Festenstein for their continuous guidance and mentorship throughout these years. JB, your unique ability to break down a complex biological problem and easily discern the best route of action to tackle it, has not only been inspiring, but also tremendously helpful for my scientific coming of age. Richard, you have been one, if not the most creative mind that I have encountered in my life. It has always been stimulating to hear your thoughts and seeing your science-driven curiosity. I hope this beautiful chaos called the brain of Richard Festenstein never stops to galvanize others.

To all the current and past lab members of Vannier and Festenstein groups I owe a big thank you! It has been a pleasure to spend the time of my PhD with you. You were always there, teaching me not only how to be a better scientist, but also a better human being. I will always cherish the never-ending scientific discussions and the daily encouragement, especially through the rough days that nothing seemed to work.

Special thanks to Dr. Rosa-Maria Porreca, Dr. Emilia Herrera Moyano and Dr. Jackson Ping-Kei Chan for their immense help and support throughout this study. Also, I am verry happy to have shared this experience with Soteroulla Ellina. While working on our PhDs we developed a true friendship! Camilla Valente was there at the right time to convince me that doing this PhD was my best decision to date and Hannah Danielle Shepherd showed me how one can balance fun and work. Carmen Ramirez Moncayo made me see again the things that truly matter. Science, coffee, music, books and Sunflowers.

I would also like to acknowledge our collaborators at the LMS Core facilities, including the LMS Genomics facility (Ivan Andrew, Jaspreet Haywood and Dr. Laurence Game), LMS Bioinformatics facility (Sanjay Khadayate and Dr. Mahdi Karimi), LMS Proteomics and Metabolomics Facility (Alex Montoya and Dr. Pavel Shliaha). Also, to Prof. Jesus Gil and Virinder Reen, to Prof. Boris Lenhard, Dr. Nevena Cvetesic and Radina

Georgieva for their continuous guidance, support and advice with the experimental work.

I am thankful to the UKRI-Medical Research Council for funding this studentship and to Imperial College London (and its magical community) for providing me with the necessary support to complete this PhD.

The journey of this PhD would have been very different if certain people did not make it as unique as it has been. To *my friends,* who made this, the best time of my life, thank you!

Last but definitely not least, I want to express my deepest gratitude to my family, Giorgos, Lia and Eleni Stylianakis. I would not have made it without their unconditional love and support. They have been the bright burning beacon that allows me to see life as an adventure, dream big and constantly push my limits, while knowing there is always a place that I can call *home*.

Abstract

Heterochromatin Protein 1 (HP1) is a protein family of epigenetic modifiers that are integral units of heterochromatin establishment with three mammalian isoforms HP1α, HP1β and HP1γ. All three isoforms are found in heterochromatic regions where they assist with chromatin compaction and gene silencing, yet, HP1γ is also enriched on euchromatic regions with a suggestive gene activating role. In mice, HP1γ has been shown to influence genes that differ between the sexes, nonetheless, the implications of this sexual dimorphism and the molecular mechanisms underlying it are poorly understood. Here, I show that HP1γ is important for male cell proliferation and its absence causes earlier onset of cellular senescence in both sexes. The effect of HP1γ loss is further reflected in male embryo growth rate. Cleavage under targets and release using nuclease (CUT & RUN) analysis revealed that HP1γ is binding on genes differentially expressed among the sexes affecting their expression. Preliminary data using Super-Low Input Carrier Cap analysis of gene expression (SLIC-CAGE) suggests that HP1γ most likely does not regulate cryptic transcription in mice. HP1γ was also enriched on repetitive DNA sequences at the end of chromosomes termed telomeres. Given the recruitment of HP1γ at telomeres for heterochromatin formation and the direct link between senescence and telomere length, we examined the role of this factor on telomere maintenance. Loss of mouse HP1γ leads to a downregulation of various telomere and telomere-accessory transcripts, including shelterin protein TRF1. This transcriptional and protein downregulation is associated with increased telomere replication stress and DNA damage, both effects more profound in females. My analysis suggests that the source of the impaired telomere replication is the increase in telomeric DNA:RNA hybrids due to the upregulation of TElomeric Repeat-containing RNA (TERRA) arising from mouse chromosome 18 and chromosome X. Overall, this PhD thesis showcases the important role of HP1γ on sexual dimorphism and telomere stability during early mouse development.

Table of Contents

List of Figures

List of Tables

Abbreviations

ALT – alternative lengthening of the telomeres
APB – ALT-associated PML bodies
BIR – break induced replication
bp – base pairs
CAGE – cap analysis gene expression
CDS – Coding sequence
CFS – common fragile site
ChIP – chromatin immunoprecipitation
CHIRT assay – Combination of ChIRP and CHART assays
CST complex – CTC1-STN1-TEN1 complex
CUT & RUN – cleavage under targets and release using nuclease
D-loop – displacement loop
DAPI – 4',6-diamidino-2-phenylindole
DDR – DNA damage response
DMEM – Dulbecco's modified eagle medium
DMSO – dimethyl sulphoxide
DNA – deoxyribonucleic acid
DNMT – DNA methyltransferase
dNTP – Deoxynucleotide
DRIP –DNA-RNA immunoprecipitation
DSB – double strand break
dsDNA – double stranded DNA
FACS – fluorescence-activated cell sorting
FBS – fetal bovine serum
FISH – fluorescence in-situ hybridisation
G4 – G-quadruplex
HDAC– histone deacetylase
HDM – Histone demethylase
HMT– histone methyltransferase
HP1– heterochromatin protein 1

HR – homologous recombination

IF – immunofluorescence

IP – immunoprecipitation

LFQ – label-free quantification

MEF – mouse embryonic fibroblast

MNase – micrococcal nuclease

MS – mass spectrometry

NHEJ – non-homologous end joining

NMD – non-sense mediated decay

ncRNA – non-coding RNA

PBS – phosphate buffered saline

PCR – Polymerase chain reaction

PEV – position effect variegation

qRT-PCR – quantitative real time polymerase chain reaction

RNA – ribonucleic acid

SD – standard deviation

SET domain – Su(var)3-9, Enhancer-of-zeste and Trithorax domain

SEM – standard error of the mean

shRNA – short hairpin RNA

SILAC – stable isotope labelling by amino acids in cell culture

SSB – single-stranded break

Ss DNA – single-stranded DNA

SV40 – simian virus 40

T-loop – telomeric loop

TES – transcript end site

TIF – Telomere Dysfunction Induced Foci

TRF – terminal restriction fragment

TSS – transcription start site

UTR – Untranslated region

WB – western blot

Chapter 1. General introduction

1.1 From DNA to higher-order genome organisation

1.1.1 Chromatin structure

In eukaryotic nuclei, genomic DNA is packaged with histone and non-histone proteins in a complex assemblage known as chromatin. The fundamental unit of chromatin is the nucleosome (Luger *et al.*, 1997), which is formed by the wrapping of 145–147 base pairs (bp) of DNA around a histone octamer made up by two copies of the core histones H2A, H2B, H3 and H4 (Luger *et al.*, 1997). Adjacent nucleosomes are connected by short DNA segments termed "linker DNA" to form uniform arrays of nucleosomes that are visible under the microscope as "beads on a string" (McKnight and Miller, 1976). Unlike the budding yeast *Saccharomyces cerevisiae*, where nucleosomes are mostly evenly spaced throughout the genome (Yuan *et al.*, 2005), in animals, nucleosomes are not that evenly distributed (Valouev *et al.*, 2011). The spacing of the nucleosomal array appears to correlate with certain genomic features like silent domains (regular spacing) or active gene promoters and enhancers (irregular spacing), at least in *Drosophila melanogaster* (Baldi *et al.*, 2018).

Linker histone H1 attaches to the nucleosomal structure, keeping the wrapped DNA in place with ~10 bp of DNA at both the entry and the exit sites of the nucleosome core, regulating the orientation of linker DNA and enabling the organisation of the nucleosome arrays into a higher chromatin order (Simpson, 1978; Hamiche *et al.*, 1996). At high molecular concentrations and at salt conditions found in the nucleus, *trans*-interactions across different parts of the chromatin fiber dominate and generate a largely unstructured globule (Maeshima *et al.*, 2016), while under high salt conditions the chromatin fibre gets compacted in a 30 nm structure (Allan *et al.*, 1984). Further long-range fibre-fibre interactions leads to the formation of the ~100 nm interphase

chromosomes (Belmont *et al.*, 1989)and the metaphasic chromosomes of ~600 nm (Belmont *et al.*, 1999). The chromatin state and compaction of genomic DNA affects many nuclear processes, including DNA replication and transcription (Bickmore and van Steensel, 2013) and is essential for normal cell division.

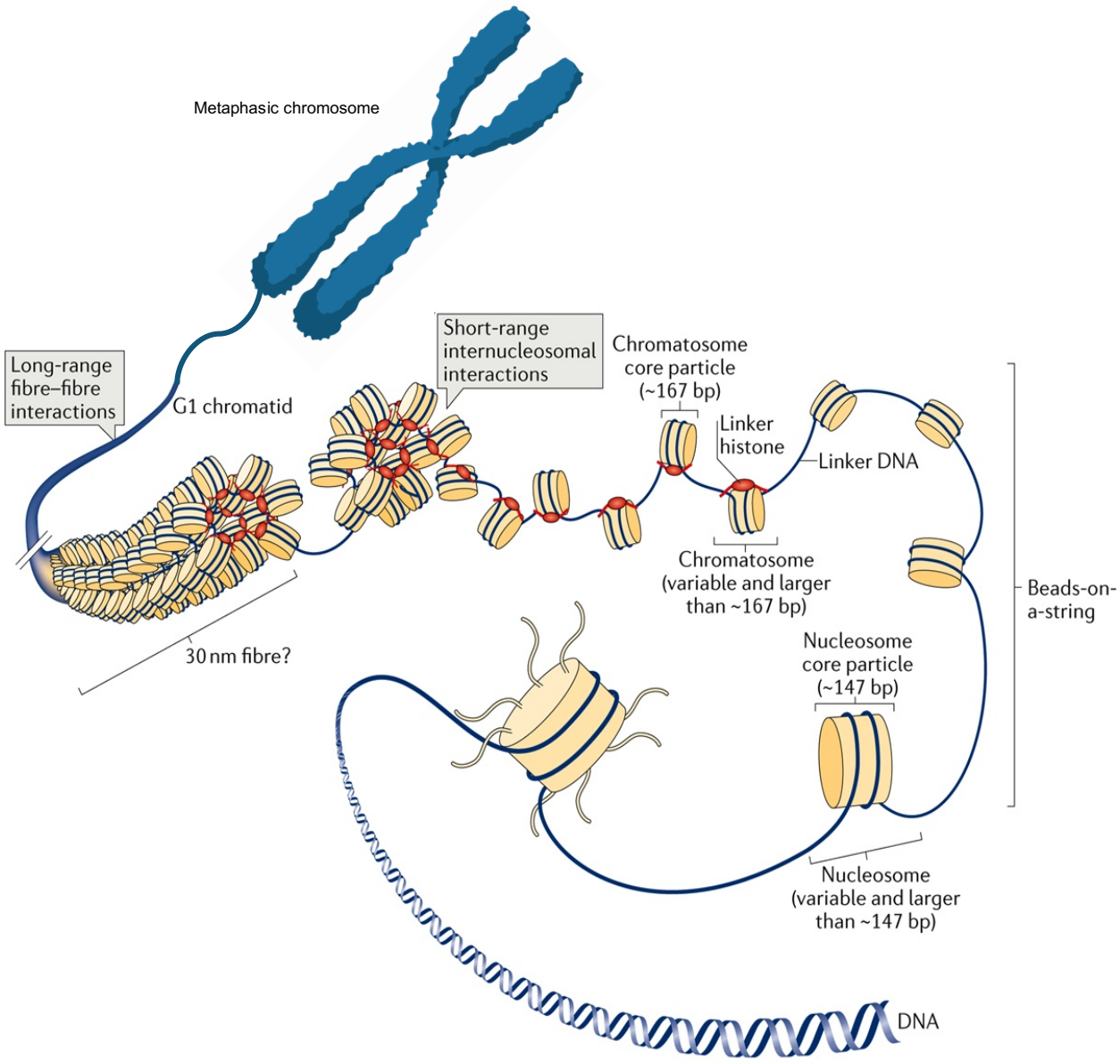


Figure 1.1 Schematic diagram of chromatin folding. DNA compaction and assembly into chromatin occurs mainly via histone-dependent interactions. These include the hierarchical formation of the nucleosome particle, nucleosomal arrays connected by linker DNA ("beads-on-a-string"), formation of the 30nm fiber (at least in vitro), and eventually the association of individual fibres produces tertiary structures and allows the formation of metaphasic chromosomes. This figure was adapted from Fyodorov et al., 2017.

1.1.2 Higher genome organisation

A lot of effort has been put recently into understanding the relationship between chromatin structure, three-dimensional (3D) organisation of the genome and how this relates to functional molecular processes. The development of techniques like Hi-C and Genome Architecture Mapping (GAM) that allow the elucidation of DNA's spatial architecture using next generation sequencing provided a clearer picture of how the genome folds and what are the functional consequences of this (Lieberman-Aiden *et al.*, 2009; Beagrie *et al.*, 2017).

At large scales, interphasic chromosomes occupy distinct subnuclear territories (Cremer and Cremer, 2010), with preferred positioning that depends on the cell type (Parada, McQueen and Misteli, 2004). At megabase scale, the genome segregates into distinct compartments named A and B which correlate with the chromatin state (Rao *et al.*, 2014; J. Wang *et al.*, 2016). Compartment A resides in the internal part of the nucleus having an open structure chromatin related with active transcription. Compartment B on the other hand is close to the periphery and is related with a compact chromatin structure and repressed transcription. At sub-megabase scales, Topologically Associated Domains (TADs) have been identified (Lieberman-Aiden *et al.*, 2009). TADs refer to regions of DNA residing on the same chromosomes that display increased interaction frequency, compared to regions of different TADs that associate less frequently with each other (Dixon *et al.*, 2012; Nora *et al.*, 2012; Sexton *et al.*, 2012). The presence of TADs has been described in many species including humans, mice and flies (Dixon *et al.*, 2012; Hou *et al.*, 2012), but also among different cell types such as ESCs and fibroblasts (Dixon *et al.*, 2015), suggesting that genome organisation at this level is a conserved feature.

Two of the factors enabling the formation of TADs is the zinc-finger protein CTCF (CCCTC-binding factor) which was originally described as a transcriptional repressor of the c-myc oncogene (Klenova *et al.*, 1993) and cohesin, the protein complex necessary for sister chromatid cohesion during metaphase (Guacci, Koshland and Strunnikov, 1997; Michaelis, Ciosk and Nasmyth, 1997). An intriguing "loop extrusion" model has been put forward addressing how CTCF and cohesin associate with other architectural proteins to establish the 3D organisation of the genome (Sanborn *et al.*, 2015; Fudenberg *et al.*, 2016). According to this model, two convergent CTCF binding

sites act as the anchor to stop the chromatin-loop fibre from extruding through the cohesion-ring complex, which has been proved to be crucial for loop formation, thereby defining the TAD boundaries. Supportive evidence is provided by the co-localisation of CTCF with cohesin, and enrichment of their binding sites at the chromatin loop forming regions (Parelho *et al.*, 2008; Wendt *et al.*, 2008; Rao *et al.*, 2014). At the same time, it is worth noting that while CTCF and cohesin are required for chromatin loops and TAD establishment, their depletion have only mild effects on the transcriptome (Nora *et al.*, 2017; Rao *et al.*, 2017; Schwarzer *et al.*, 2017; Wutz *et al.*, 2017).

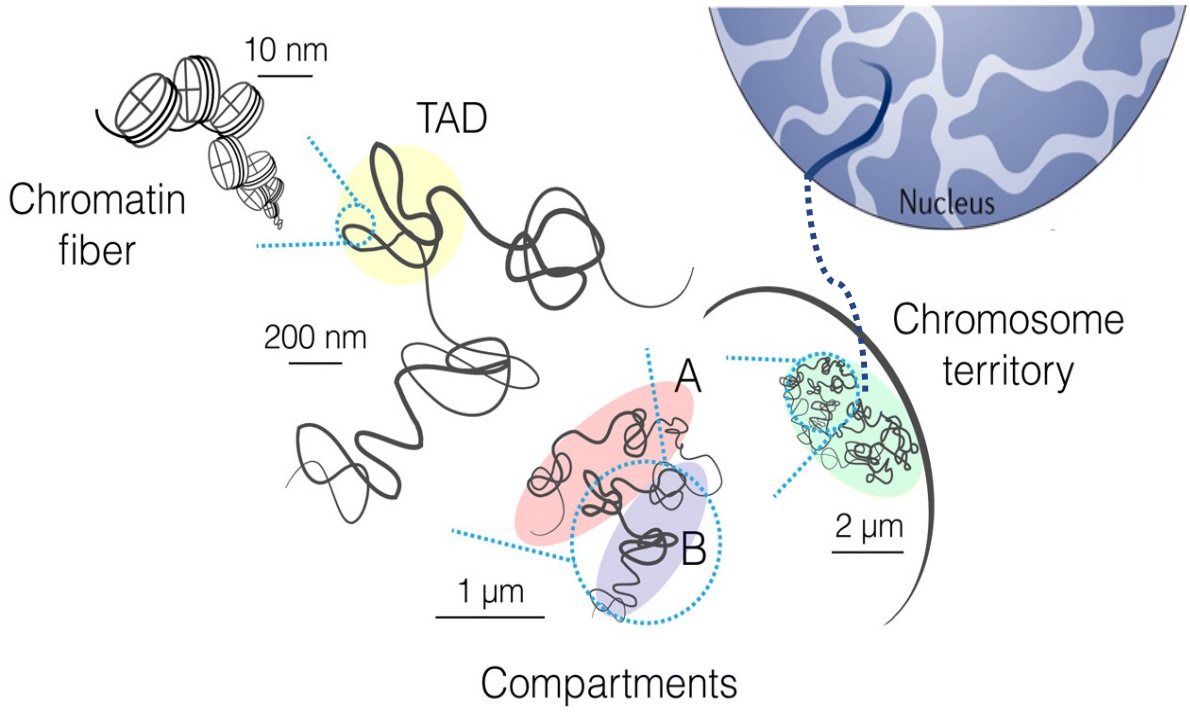


Figure 1.2 Principles of higher chromatin organization in eukaryotes. Chromatin is packed into fibers with different nucleosome densities sensitive to gene regulation and folds at the sub-megabase scale into higher-order domains referred to as TADs. At the chromosomal scale, chromatin is organized into active "A" and silent "B" compartments. Individual chromosomes show distinct subnuclear localization, forming chromosome territories during the interphase. Figure adapted from Szabo et al., 2019.

1.1.3 Epigenetics – historical overview

Conrad Waddington introduced the term epigenetics in the early 1940s by making use of the Greek prefix epi- (ἐπι-) meaning "over/around" the traditional term of genetics, which implies features that are "on top of" the traditional basis for inheritance. According to him, epigenetics refers to the investigation of the causal interactions between genes and their products which bring the phenotype into being (Waddington H., 1942). At the same time, the focus of Waddington's contemporaries involved with this newly founded branch of biology would vary, with Nanney and Lederberg for instance, being more interested in the stability of expression and cellular inheritance (Lederberg J., 1958; Nanney, 1958). These divergent opinions led to a long-lasting definition crisis of the epigenetics field.

Nowadays, the term epigenetics is used to describe the branch of biology which studies the alterations of gene functions that are mitotically or meiotically heritable, yet do not involve changes of the DNA sequence (Dupont, Armant and Brenner, 2009). Epigenetics therefore refer to the modifications made directly on the DNA or on the histone proteins and are facilitated by regulatory factors that write, read and erase these modifications.

The two major epigenetic mechanisms of gene expression regulation are DNA methylation and histone modifications, while recently RNA has been also recognized as a facilitator of epigenetic processes. DNA methylation was discovered even before the characterisation of DNA as the unit of genetic information (Hotchkiss, 1948; Watson & Crick, 1953), while the existence of histone modifications was identified almost 20 years later (Allfrey et al., 1964). The importance of these chemical modifications on chromatin was not fully appreciated until the late 1970s when it was first observed that cells showed certain expression patterns, based on their DNA methylation status (McGhee & Ginder, 1979), while later in the 1980s, it was shown that histones also affect gene transcription (Knezetic and Luse, 1986).

1.1.4 Euchromatin and heterochromatin

Apart from the spatiotemporal organisation of the genome, the state of chromatin accessibility reflects regulatory capacity, with a dynamic biophysical nature. Chromatin has two distinct types of organisation, first described by the botanist Emil Heitz in 1928 (Heitz, 1928). These two “flavours” of chromatin show distinct features which contribute to the regulation of gene expression.

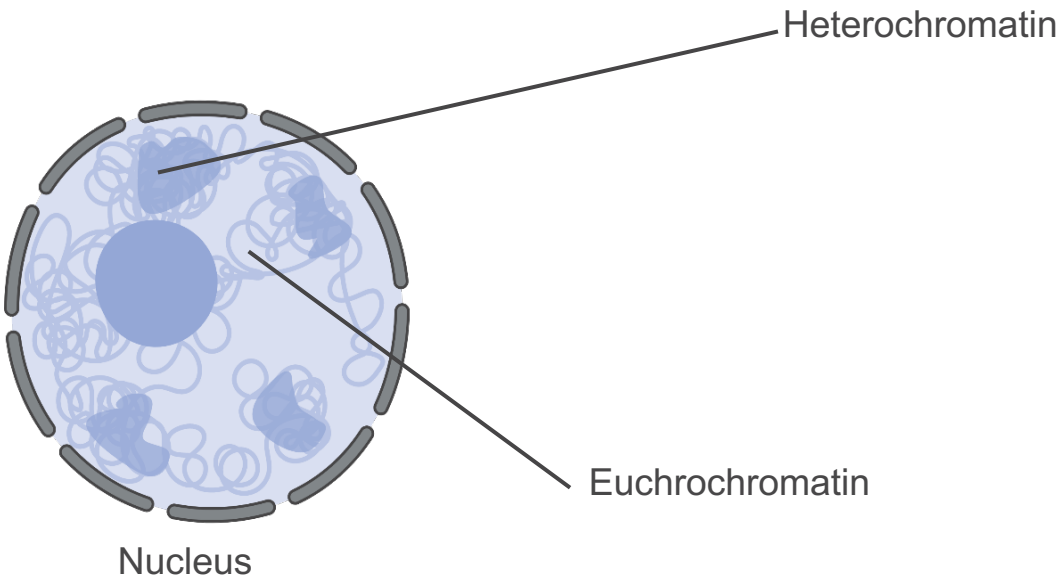


Figure 1.3 Chromatin states - Heterochromatin vs Euchromatin. Illustration of a mammalian cell nucleus during interphase. DNA cytological staining reveals different types of chromatin. The darker regions represent condensed, transcriptionally inactive chromatin (heterochromatin), whereas lighter regions consist of open, transcriptional active chromatin (euchromatin).

On the one hand, euchromatin appears to have an open structure with increased nuclease accessibility, that is replicating early during the S-phase and is rich in actively transcribed genes. On the other hand, heterochromatin appears to be a condensed chromatin structure with limited nuclease accessibility, rich in repetitive DNA sequences such as transposable elements and satellite repeats that replicates late

during the S-phase. Unlike euchromatin, heterochromatin is poor in actively transcribed genes. In lower eukaryotes only regions important for genome integrity such as telomeres, centromeres and their surrounding areas are stably heterochromatinised (Grunstein, 1998; Grewal and Jia, 2007). The same chromosomal regions are also considered heterochromatic in higher eukaryotes; however, due to the massive increase in genome size and organismal complexity, smaller blocks of heterochromatin can be also found interspersed throughout the genome (Dillon and Festenstein, 2002; Grewal and Moazed, 2003; S. Wang *et al.*, 2016).

Furthermore, the two states of chromatin are characterised by distinct epigenetic profiles. Euchromatin displays hypomethylated DNA at promoter of transcribed genes, enrichment of acetylated histones H3 and H4 and high levels of methylated lysine 4 of histone H3 (H3K4me), whereas heterochromatin is characterised by DNA methylation, hypoacetylated histones and histone modifications such as methylation of histone 3 lysine 9 (H3K9me), histone 3 lysine 27 (H3K27me) and histone 4 lysine 20 (H4K20me) (Okano *et al.*, 1999; Lehnertz *et al.*, 2003; Schotta *et al.*, 2004).

Heterochromatin can be further divided into two states, constitutive and facultative. Constitutive heterochromatin is usually composed of non-coding repetitive elements such as satellite DNA that comprises up to 10% of the human genome (Garrido-Ramos, 2017). Satellite DNA is primarily found in pericentromeric regions and in humans the common satellite elements include α-satellites and γ-satellites (Lin *et al.*, 1993), whereas in mice it is major satellites (Sart *et al.*, 1997). These regions stain darker with the C-banding technique owing to their highly condensed nature and are commonly localised near the periphery of the nucleus (Holla *et al.*, 2020). Constitutive heterochromatin is marked by histone 3 lysine 9 tri-methylation (H3K9me3) and histone 4 lysine 20 tri-methylation (H4K20me3). Meanwhile, facultative heterochromatin is marked by high levels of tri-methylated histone 4 lysine 27 (H3K27me3) and consists of non-repetitive DNA elements that are silenced through various mechanisms such as histone deacetylation. It is less compacted compared to constitutive heterochromatin and under specific environmental or developmental stimuli, facultative heterochromatin can become transcriptionally active by losing its condensed nature (Trojer & Reinberg, 2007). The inactive X chromosome is a typical example of facultative heterochromatin which forms the condensed Barr body in

mammalian female cells. The Barr body is visible in the interphase nucleus as a small dark mass in close proximity with the nuclear membrane (Barr & Bertram, 1949).

1.1.5 The epigenetic code

The “epigenetic code” is a proposed code in eukaryotic cells determined by a distinct set of epigenetic modifications including DNA methylation and histone post-translational modifications (PTMs), providing an extra layer of DNA regulation, on top of the genetic code.
While for an individual the genetic code in each cell is the same, the epigenetic code can be cell-type specific (Turner, 2007). In this thesis, I will be focusing on how histone methylation affects the properties of chromatin and hence regulates gene activity.

1.1.6 Histone modifications overview

The core histone proteins are composed of a globular domain and a flexible N-terminal “histone tail” which protrudes from the nucleosome. Histones are extensively modified with PTMs found on both DNA–histone and histone–histone interfaces (Mersfelder and Parthun, 2006). Advances of mass spectrometry (MS) technology has enabled the identification of at least sixteen different types of PTMs (acetylation, methylation, phosphorylation, ubiquitinylation, sumoylation, propionylation, crotonylation, butyrylation, hydroxylation, neddylation, O-GlcNAc, ADP ribosylation, N-formylation, proline isomerization, and citrullination) in over 60 different histone residues (Kouzarides, 2007; Zhao and Shilatifard, 2019). These modifications can occur sequentially and/or simultaneously on one or more histone tails within the same nucleosome with a synergistic or antagonistic effect, resulting to a unique cellular

response. This crosstalk between histone marks has been suggested to fine-tune gene regulation and has been termed “histone code” (Strahl and Allis, 2000), while recently it has been shown that TADs can be predicted via chromatin marks emphasizing the importance of the histone code on genome architecture (Ho *et al.*, 2014).

While PTMs of the globular domain have mostly a structural impact on nucleosome dynamics, modifications on histone tail residues can also regulate chromatin by acting as chemical signposts for PTM-specific binding proteins that act as effectors of the epigenetic code and facilitate chromatin-related processes such as transcription, DNA repair, replication and recombination or alternative splicing (Bannister and Kouzarides, 2011; Tropberger and Schneider, 2013).

These epigenetic regulators are divided in three main categories: writers, readers, and erasers. Writers include factors that add modifications, such as histone acetyltransferases (HATs), histone methyl transferases (HMTs), kinases, etc. Readers are proteins containing domains that recognize specific PTMs, with the ability to recruit other complexes assigned with downstream functions like DNA repair or DNA transcription. Erasers remove epigenetic modifications, including acetyl groups (HDACs) and methyl groups (HDMs).

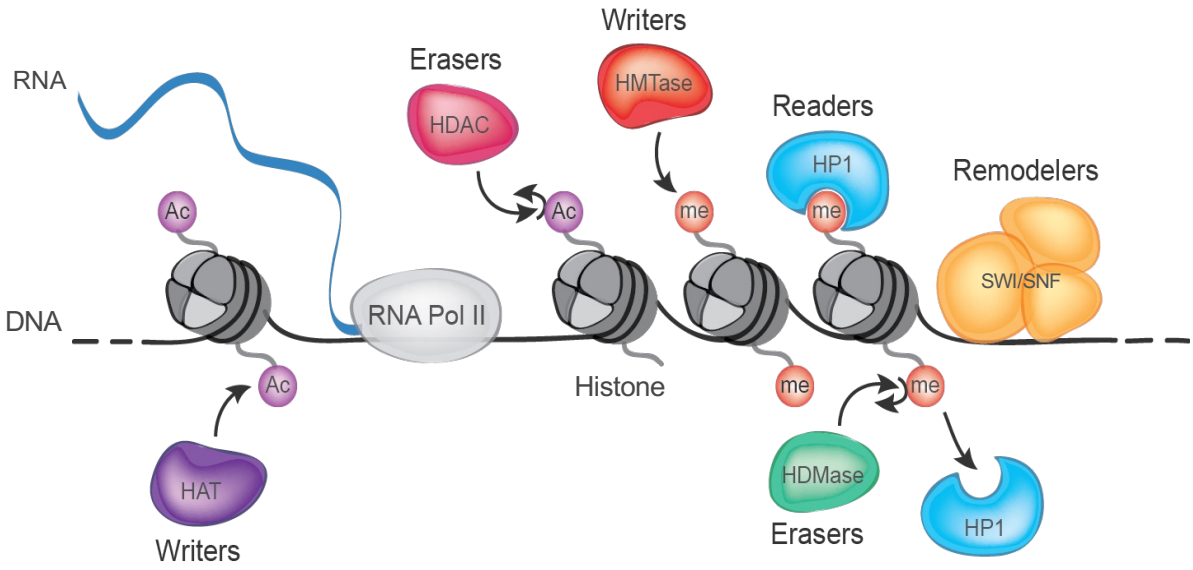


Figure 1.4 A representation of epigenetic regulators. Epigenetic regulators include writers, readers and erasers. These enzymes are responsible for most epigenetic modifications on histones and DNA. HAT – Histone acetyltransferase, RNA Pol II – RNA polymerase II, HDAC – Histone deacetylase, HMTase – Histone methyltransferase, HP1 – Heterochromatin Protein 1, HDMase – Histone demethylase, SWI/SNF – SWItch/Sucrose Non-Fermentable. Taken from the Kaushik Ragunathan lab website.

1.1.7 Histone methylation

In histones, lysine and arginine serve as the most common acceptor sites of methylation marks, with histone H3 being the primary site of histone methylation. Unlike lysine acetylation that neutralises the positive charges of the histones and thereby weakens the histone-DNA interactions, methylation does not alter the charge of histone proteins (Kouzarides, 2007). It rather assists with the recruitment of other effector proteins for downstream chromatin functions.

Histone lysine methylation comes in three states: mono-, di- or tri-methylation and depending on the target lysine residue can exert a gene-activating or a gene-repressive effect (Bannister and Kouzarides, 2011). For instance, mono-methylation of H3K4 (H3K4me1) is an activating mark defining enhancers (Heintzman *et al.*, 2007). Di- and tri-methylation of histone 3 lysine 36 (H3K36), histone 3 lysine 79 (H3K79) and H3K4 is usually considered transcription-activating. H3K4 tri-methylation (H3K4me3) marks promoters (Bernstein *et al.*, 2002; Barski *et al.*, 2007), while methylation of H3K79 and H3K36 occurs primarily over gene bodies (Bannister *et al.*, 2005; Nagalakshmi *et al.*, 2008). Of note, tri-methylation of (H3K36me) has been suggested to protect against erroneous, cryptic transcription (Carrozza *et al.*, 2005). Mono-methylation of histone 3 lysine 56 (H3K56me1) is implicated with DNA replication and tri-methylation of H3K56 with gene silencing (Yu *et al.*, 2012; Jack *et al.*, 2013). Methylation of H3K9 and H3K27 is generally considered gene-repressive (Bernstein *et al.*, 2002; Barski *et al.*, 2007). Figure 1.5 summarises the suggested functions of different methylation marks of histone lysine.

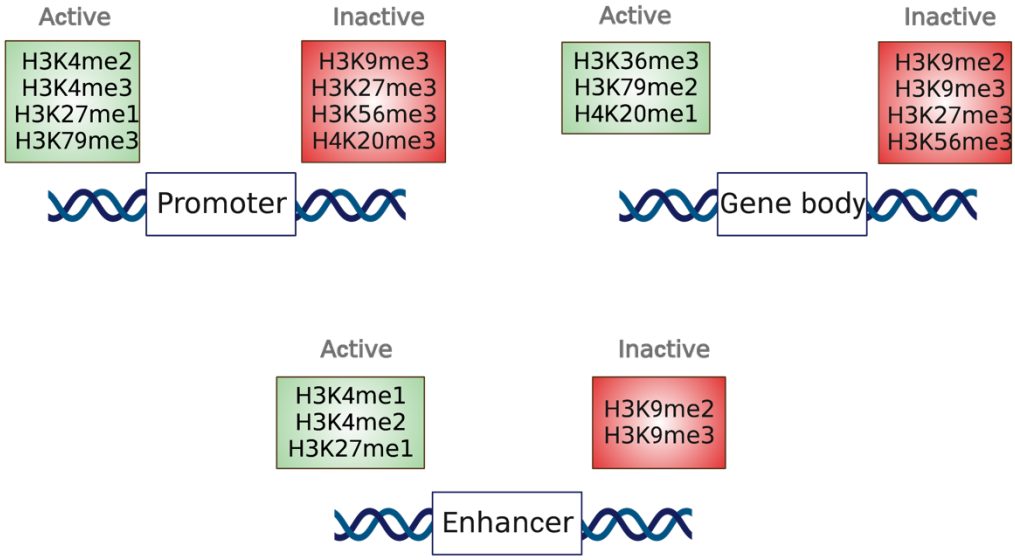


Figure 1.5 Histone methylation function in genomic features. At promoters, histone methylation contributes to fine-tuning expression levels, while at gene bodies, it discriminates between active and inactive states. At distal sites, histone methylation correlates with levels of enhancer activity.

1.1.8 Histone 3 lysine 9 methylation

As mentioned above, histone 3 lysine 9 methylation is typically associated with gene repression and extensively decorates constitutive heterochromatin, including pericentromeric and telomeric regions (Peters et al., 2003). However, the role of H3K9me is not as straightforward as initially thought. Based on the methylation state, the distribution of this PTM on the different chromatin regions can vary in heterochromatin while depleted from their transcription start sites (TSSs) (Ho et al., 2014). H3K9me1 has been also associated with gene activation, as it is enriched at promoters of transcribing regions (Barski et al., 2007; Peters et al., 2003). More light has been shone on the role of H3K9me on gene regulation by studies from the Blobel,

Zhao, Muchardt and Proudfoot groups that have implicated H3K9me with transcriptional active genes (Barski et al., 2007; Vakoc et al., 2005, 2006), alternative splicing (Saint-André et al., 2011) and transcriptional termination (Skourti-Stathaki et al., 2014).. For instance, in mammals, both H3K9me2 and H3K9me3 predominantly mark silent genes, yet the first is mostly found on euchromatic regions while the second localises almost exclusively in heterochromatic regions. Interestingly, both in humans and flies, H3K9me3 has been found enriched on the gene body of expressed genes that reside

In mammals, several HMTs responsible for the H3K9 methylation have been identified. These HMTs contain a Su(var)3-9, Enhancer-of-zeste and Trithorax (SET) protein domain that facilitates the transfer of a methyl group from Sadenosylmethionine (SAM) to the ε-amino group of the histone residues. The *Drosophila* Suppressor of variegation 3-9 (Su(var)3-9) and the mammalian Suppressor of variegation 3-9 homolog (SUV39H) with its two isoforms SUV39H1 & SUV39H2 are the first identified HMTs that catalyse H3K9me3 (Rea *et al.*, 2000; Peters *et al.*, 2001; Schotta *et al.*, 2004). SUV39H-mediated methylation has been implicated with constitutive heterochromatin (Rice *et al.*, 2003), retrotransposon elements (Bulut-Karslioglu *et al.*, 2014) and in some cases with the silencing of euchromatic genes (Nielsen *et al.*, 2001; Ait-Si-Ali *et al.*, 2004).

SET Domain Bifurcated Histone Lysine Methyltransferase 1 (SETDB1), also known as ESET is another H3K9-specific HMT. It catalyses H3K9me2 and H3K9me3 (Schultz *et al.*, 2002; Yang *et al.*, 2002) and can convert di-methylated to tri-methylated H3K9 when found in a complex with the ATFα-associated mAM (Wang *et al.*, 2003).

While SETDB1 is thought to associate mostly with euchromatin (Schultz *et al.*, 2002; Li *et al.*, 2006), a recent study showed that in murine embryonic stem cells (mESCs) SETDB1 catalyses H3K9me3 at telomeres which are considered heterochromatic (Tachibana *et al.*, 2001, 2002, 2005; Gauchier *et al.*, 2019).

Euchromatic histone-lysine N-methyltransferase 1 (EHMT1), also known as GLP and Euchromatic histone-lysine N-methyltransferase 2 (EHMT2), also known as G9a is another pair of closely related paralog HMTs that catalyse the mammalian euchromatic H3K9me1 and H3K9me2 marks (Tachibana *et al.*, 2001, 2002, 2005). Initially, this pair of EHMTs was believed to function solely as a hetero-dimeric

complex, but recent studies have shown that they can also act independently (Tachibana *et al.*, 2008; Zhang *et al.*, 2016).

H3K9-specific HDMs have been also identified that act as erasers of the methylation PTM. KDM3A (JHDM2A) has been shown to demethylate H3K9me1 and H3K9me2 while KDM4A (JHDM3A) and KDM4C (JMJD2C) are responsible for H3K9me3 demethylation (Fodor *et al.*, 2006; Klose, Kallin and Zhang, 2006; Whetstine *et al.*, 2006).

1.1.9 Position effect variegation (PEV)

One of the first observed phenomena that provided a palpable link between the conformational state of chromatin and the transcriptional status of genes is "position effect variegation''. PEV describes the phenomenon where, if a gene is abnormally juxtaposed to a location near heterochromatin due to rearrangement from its original euchromatic position, it exhibits variegation of expression. This rearrangement is believed to remove the pre-existing genetic "boundary" elements, promoting the spreading of heterochromatin, leading to the silencing of the affected genes. Consequently, a fraction of the cells that would normally express the variegated gene, now exhibits a stochastic silencing of it (Elgin and Reuter, 2013). PEV was first described in Drosophila studies with the observation that translocation of the *white* gene near centromeric heterochromatin resulted in variegation of the eye colour, with a proportion of the eye cells being red (Muller and Altenburg, 1930). It has since been observed in other organisms including mammals and yeast (Cattanach, 1974; Allshire *et al.*, 1994; Festenstein *et al.*, 1996; Milot *et al.*, 1996). Examples of PEV in mammals include the variegation of mouse coat's colour due to the silencing of the gene responsible for this trait, upon translocation onto the inactive X chromosome (Cattanach, 1961; Russel L. B. & and the silencing of β-globin and CD2 transgenes when integrated into pericentromeric regions (Festenstein *et al.*, 1996). Intriguingly, when the CD2 transgene was put close to DNA triplet repeats found in human diseases

including the GAA repeats associated with Friedreich's Ataxia, it also exhibited variegation of expression, which led the authors of the study to suggest that the transcriptional silencing caused by pathological repeat expansions may be exerted by a mechanism similar to heterochromatin-mediated PEV (Saveliev *et al.*, 2003).
Many PEV readers and writers were identified by mutagenesis screens in *Drosophila*, which either enhance or suppress variegation. Of note, two of the identified PEV suppressors are Su(var)3-9 and heterochromatin protein 1 (HP1) that are homologues of mammalian SUV39H and HP1, respectively (Wallrath, 1998).

1.1.10 Heterochromatin protein 1 (HP1)

HP1 is a family of conserved non-histone proteins, first discovered in *Drosophila* as a PEV modifier, that appears to be an elementary unit of chromatin packaging (Singh *et al.*, 1991; Bannister *et al.*, 2001; Lachner *et al.*, 2001). Indeed, HP1 can be found at heterochromatic sites of nearly all eukaryotes with the exception of *Saccharomyces cerevisiae* in which the homologous silent information regulator (SIR) proteins serve a similar function (Palladino *et al.*, 1993; Kueng, Oppikofer and Gasser, 2013). HP1 has 3 mammalian isoforms: HP1α, HP1β and HP1γ that display a high level of structural similarity and are encoded by the CBX5, CBX1 and CBX3 genes, respectively (Singh *et al.*, 1991). Human and mouse HP1 proteins share more than 97% similarity, with HP1γ isoform for instance, sharing a 99.5 % protein similarity among those species (Altschul *et al.*, 1997). In humans, a shorter HP1γ splice variant was recently reported (Mathison *et al.*, 2020). HP1 proteins have been implicated in a plethora of regulatory processes essential for cell physiology including gene activation/repression, chromosomal segregation, PEV, DNA repair, telomere maintenance, and RNA processing (Canzio, Larson and Narlikar, 2014).

HP1 proteins possess three main domains: First, the chromodomain domain (CD) which is near the N-terminus, and is responsible for the binding of HP1 proteins on di-

or tri-methylated H3K9 (Bannister *et al.*, 2001; Lachner *et al.*, 2001; Jacobs and Khorasanizadeh, 2002). Binding of HP1 on H3K9me3 has the higher affinity compared to H3K9me2 (Nielsen *et al.*, 2002). Abolishing or mutating the CD, disrupts HP1's localisation to heterochromatic regions (Cheutin *et al.*, 2003). The binding of HP1 on H3K9me3 is crucial for localisation of HP1 to heterochromatin, however it is not sufficient to ensure a steady bound presence. Indeed, fluorescence recovery after photobleaching (FRAP) experiments revealed that HP1 is highly mobile and the pool of chromatin-bound molecules is actively renewed (Festenstein et al., 2003; Cheutin et al., 2003). Several sites of H3K9me3 may be necessary and many low-affinity interactions may come into play, maintaining high levels of HP1 at a steady state. Interestingly, the different HP1 isoforms have variable binding affinities to H3K9 methylation. For instance, the mouse HP1 isoforms binding affinities range from ~5–40 μM, with HP1α having the highest affinity (Machado, Dans and Pantano, 2010; Hiragami-Hamada *et al.*, 2011).

Secondly, the HP1 hinge domain contains a nuclear localisation signal (NLS) and sites for PTMs (Lomberk *et al.*, 2006) and is the least conserved domain of HP1 proteins indicating a potential role in the selectivity of HP1 specific isoforms and their respective interacting partners (Meehan, Kao and Pennings, 2003; Canzio, Larson and Narlikar, 2014). Several studies have shown that there is direct interaction between the hinge domain and nucleic acids (Muchardt *et al.*, 2002; Perrini *et al.*, 2004; Machida *et al.*, 2018). However, it is not clear whether this is simply a consequence of HP1 binding on the methylated H3 tail, leading to stochastic proximity with nucleosomal DNA or if this interaction is occurring with linker DNA, hinting to a functional role. A recent study argues that the DNA binding ability of the hinge domain enables HP1α and HP1γ to recognise distinct nucleosomes in humans (Mishima *et al.*, 2015). Third, the chromoshadow domain (CSD) on the C-terminus is responsible for protein-protein interactions, including homo- and heterodimerization (Ye *et al.*, 1997; Nielsen *et al.*, 2001) through recognition and interaction with a PxVxL penta-peptide motif found on HP1's interacting partners (Cowieson *et al.*, 2000; Thiru *et al.*, 2004). Chromatin regulators including transcriptional, replication and DNA repair proteins, DNA and histone transferases or nuclear structure proteins comprise the long list of HP1 binding partners (Rosnoblet *et al.*, 2011).

Finally, the CD and CSD domains are flanked by stretches of intrinsically disordered regions (IDRs), regarded as the amino (NTE) and carboxyl (CTE) terminal extensions (Nishibuchi and Nakayama, 2014) and appear to play a role in the ability of HP1 to phase separate (see below).

HP1 isoforms show a distinct nuclear distribution (Horsley *et al.*, 1996; Nielsen *et al.*, 2001) that even varies throughout the different phases of the cell cycle (Minc *et al.*, 1999) and the developmental stage (Mattout *et al.*, 2015). Generally, during interphase HP1α and HP1β are located at constitutive heterochromatin such as centromeres and telomeres (Zeng, Ball and Yokomori, 2010), while HP1γ can be found in both heterochromatic and euchromatic regions (Minc, Courvalin and Buendia, 2000). In humans, HP1β and HP1γ dissociate almost completely from mitotic chromosomes, whereas HP1α is sparsely found at centromeric chromatin(Hayakawa *et al.*, 2003).

HP1 is extensively decorated with histone-code like PTMs including phosphorylation, acetylation, methylation, sumoylation and formylation providing a regulatory mechanism to HP1's localisation and function (LeRoy *et al.*, 2009). Phosphorylation of HP1α's hinge and NTE domains is critical for liquid-liquid phase separation (LLPS) (Larson *et al.*, 2017; Strom *et al.*, 2017), while phosphorylation of the NTE also results in a five-fold affinity increase for H3K9me3, rendering it necessary for the establishment of heterochromatin and in extension genome integrity (Hiragami-Hamada *et al.*, 2011). HP1β's phosphorylation at threonine 51 by casein kinase 2 (CK2) has been shown to be important during initiation of the DNA damage response (DDR) mechanism (Ayoub, Jeyasekharan and Venkitaraman, 2009). A modification of serine 93 (previously regarded as serine 83) of HP1γ defines a subpopulation of this isoform that is exclusive to euchromatin (Lomberk *et al.*, 2006) with suggestive roles on gene activation (Harouz *et al.*, 2014) and alternative splicing (Saint-André *et al.*, 2011), while sumoylation of mouse HP1α was shown to be important for the initial targeting of this isoform to pericentromeres (Maison *et al.*, 2011).

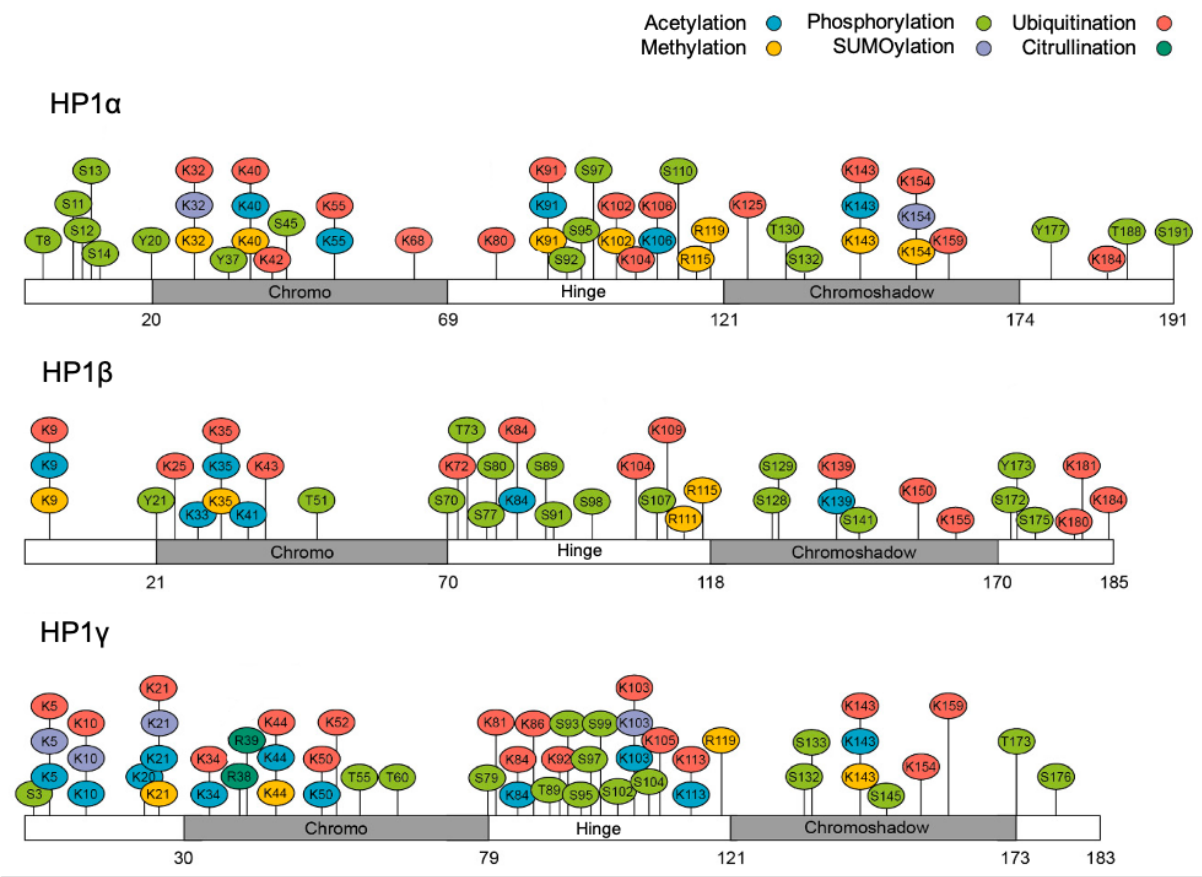


Figure 1.6 Heterochromatin Protein 1 (HP1) post-translational modifications (PTMs). 35 PTM sites have been detected for HP1α, 34 for HP1β, and 37 for HP1γ. Acetylation (blue), methylation (yellow), phosphorylation (light green), SUMOylation (purple), ubiquitination (red) and citrullination (dark green). Grey boxes highlight the chromo- and chromoshadow domains of HP1; the numbers below indicate the amino acids. Adapted from **Sales-Gil & Vagnarelli, 2020.**

Heterochromatin formation is initiated at nucleation centres by chromatin-binding proteins like HP1 and non-coding RNAs, both of which recruit and stabilise HDACs and HMTs, for deacetylation and hypermethylation of the appropriate histone residues at the nucleation sites (Volpe *et al.*, 2002; Bulut-Karslioglu *et al.*, 2012; Maeda and Tachibana, 2022). Since all HP1 isoforms interact with di- and tri-methylated H3K9 that predominantly mark silenced elements and because it can directly recruit H3K9-specific HMTs, HP1 has been generally regarded as a facilitator of heterochromatin establishment and propagation (Maison and Almouzni, 2004). In a range of organisms, including humans and flies, HP1 was recently shown to play a major role in LLPS formation, dictating the nuclear macromolecule interactions through their

inclusion or exclusion depending on the context and chromatin state (Figure 1.7), revealing an extra layer of molecular regulation (Larson *et al.*, 2017; Strom *et al.*, 2017). Specifically, it has been suggested that HP1α's driven LLPS assists with chromatin compaction and heterochromatin formation initially via nucleation, and subsequently via macromolecular oligomerisation (Larson *et al.*, 2017; Strom *et al.*, 2017; Keenen *et al.*, 2021) and nucleosomal core reshaping (Sanulli *et al.*, 2019).

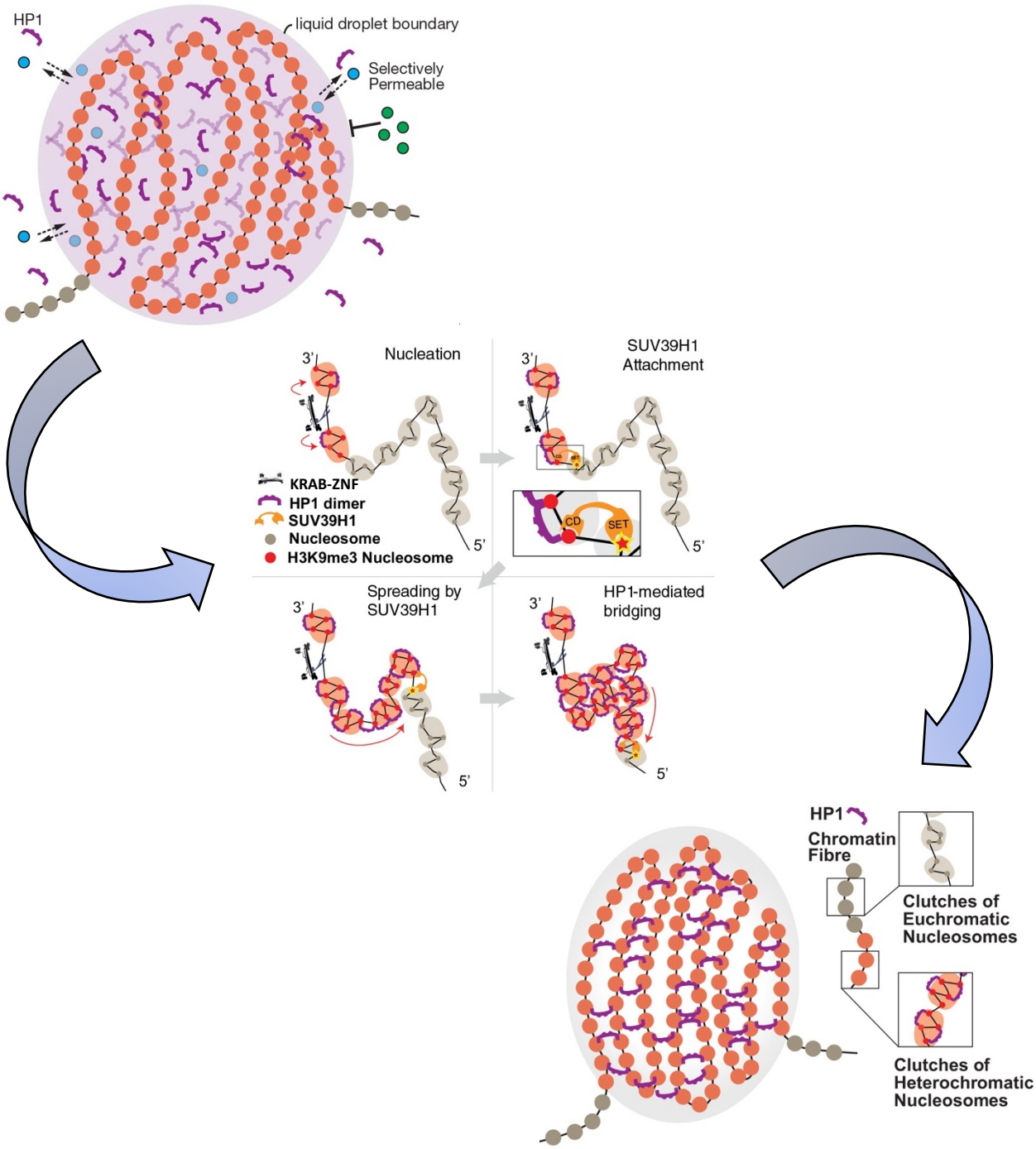


Figure 1.7 A model for HP1-driven LLPS, forming constitutive heterochromatin. HP1 binds on H3K9me3-marked nucleosomes (orange circles). High concentration of HP1 molecules drives soluble nucleosomes into droplets forming a boundary that is selectively permeable to molecules based on their chemical properties. Depicted are molecules (blue circles) that can cross the droplet boundary and others that cannot (green circles). HP1-mediated recruitment of SUV39H1 catalyzes the methylation of H3K9, providing a positive feedback loop that enables chromatin compaction and spreading of the heterochromatin away from the nucleation site. Adapted from Singh and Newman, 2022

1.1.11 HP1γ

Among the three HP1 isoforms, my PhD work had a particular interest in HP1γ. Due to its unique subnuclear distribution in both euchromatin and heterochromatin (Minc *et al.*, 1999; Nielsen *et al.*, 1999; Minc, Courvalin and Buendia, 2000) and its specific interacting partners (Table 1.1), HP1γ has been attributed with several non-redundant roles (Brown *et al.*, 2010; Ha *et al.*, 2014; Bosch-Presegué *et al.*, 2017; Zaidan *et al.*, 2018). Targeted deletion of HP1's isoforms, reveals that HP1β and HP1γ are essential for physiological organismal development, while HP1α's loss is not, which could be attributed to a compensation mechanism by the other two isoforms. Triple knockout of HP1 isoforms in hepatocytes, while proved dispensable for cell-survival, resulted in a significant increase of mouse liver tumourigenesis *in vivo* (Saksouk *et al.*, 2020). Mice that lack HP1β show perinatal lethality and the few newborns exhibited severe developmental problems and most died due to acute respiratory failure (Aucott *et al.*, 2008). Depletion of HP1γ also results in mouse neonatal lethality (Naruse *et al.*, 2007; Takada *et al.*, 2011) and the very few mice that survive, show a range of defects including growth retardation and severe infertility (Brown *et al.*, 2010; Abe *et al.*, 2011; Takada *et al.*, 2011; Aydin *et al.*, 2015).

HP1γ has been implicated with many processes regarding genome regulation, chromatin organisation and cell differentiation during development. HP1γ can regulate RNA splicing (Saint-André *et al.*, 2011; Salton, Voss and Misteli, 2014; Yearim *et al.*, 2015), and efficient transcriptional termination (Skourti-Stathaki, Kamieniarz-Gdula and Proudfoot, 2014). Moreover, HP1γ has been involved with the maintenance of constitutive heterochromatin by ensuring proper chromatin cohesion at telomeres (Canudas *et al.*, 2011) and centromeres (Yi *et al.*, 2018). Loss of HP1γ results in mitotic aberrations (Shimura *et al.*, 2011) and this isoform has been also associated with the DDR in several studies (Soria and Almouzni, 2012; Akaike *et al.*, 2015; Wu *et al.*, 2015).

Probably the most peculiar feature of HP1γ is its involvement in both gene repression and gene activation. The exact nature and molecular mechanisms dictating this duality remain unclear. Examples of HP1γ in gene silencing can be readily found:

Repression of MYC- and E2F- responsive genes in quiescent human fibroblasts (Ogawa *et al.*, 2002). Inhibition of HIV-1 expression in different human cell lines with the assistance of SUV39H1 (Chéné *et al.*, 2007). Repression of the mouse mammary tumour virus (MMTV) in breast cancer cells by binding on its promoter (Vicent *et al.*, 2006) and direct inhibition of NCOR2 and ZBTB7A expression, resulting in upregulation of tumour-promoting proteins in lung adenocarcinomas (Alam *et al.*, 2018). Repression of the pluripotency factors like Nanog, Gata4 and Bmp1 for cell-fate commitment (Sridharan *et al.*, 2013; Ostapcuk *et al.*, 2018) and repression of FOXP3 during immune T-cell differentiation (Liu *et al.*, 2010). Examples implicating HP1γ with gene activation include: Exchange of HP1β for HP1γ at the promoter of an inducible HIV1 gene upon its activation (Mateescu *et al.*, 2008). Enrichment of HP1γ on gene bodies of actively transcribed genes in both human and mouse cells and interaction with RNA polymerase II (RNAPII) (Vakoc *et al.*, 2005; Lomberk *et al.*, 2006; Smallwood *et al.*, 2008), the mediator complex (Sridharan *et al.*, 2013) and the H3.3 histone variant that is associated with expressed genes (Kim *et al.*, 2011).

Another intriguing role of the HP1 proteins seems to be their contribution in sexual dimorphism, with loss of mouse HP1γ affecting cell proliferation differently between the two sexes (Law et al., 2019).

Protein name	Species	HP1 domain	References
Chromatin associated factors			
H3 (K9me2 & 3)	Human, mouse	CD	1
H3.3	Human	nd	2, 3
HP1α	Human, mouse	CSD	4, 5
HP1β	Human, mouse	CSD	1, 6
HP1γ	Human, mouse	CSD	1, 6
INCENP	Human	nd	7
TIN2	Human	CSD	8
Chromatin modifiers			
DNMT1	Human	CD	9
DNMT2	Human	nd	9
DNTM3b	Human	nd	9
G9a	Human	nd	10
KDM2A	Human	nd	11
PIM1	Human	CSD	12
SUV39H1	Human	nd	13, 14
SUV4-20H2	Mouse	nd	15
SETDB1	Mouse	nd	16
CHD4	Human	nd	17
Transcription associated factors			
BRCA1	Human	nd	18
E2F-6	Human	nd	19
MED29	Mouse	nd	20
PAX3	Human	nd	21

RNAPII	Human, mosue	nd	22-25
TFIID	Human	nd	26
TIF1α	Mouse	CSD	27
TIF1β	Human, mouse	CSD	27, 28
ADNP	Human	CSD	17
Splicing associated factors			
AGO2	Human	nd	29
SRFSF1	Human	nd	30
DNA repair associated factors			
BARD1	Human	nd	31
HIPK2	Human	nd	32
Ku70	Human	CSD	24
POGZ	Human	nd	33

Table 1.1 Examples of mammalian HP1γ interacting partners. CD, chromodomain; CSD, chromoshadow domain; nd, not determined. References used are: (Nielsen *et al.*, 2001)[1], (Kim *et al.*, 2011)[2], (Loyola *et al.*, 2006)[3], (Lachner *et al.*, 2001)[4], (Bannister *et al.*, 2001)[5], (Rosnoblet *et al.*, 2011)[6], (Ainsztein *et al.*, 1998)[7], (Canudas *et al.*, 2011)[8], (Smallwood *et al.*, 2007)[9], (Ruan *et al.*, 2012)[10], (Frescas *et al.*, 2008)[11], (Koike *et al.*, 2000)[12], (Fritsch *et al.*, 2010)[13], (Stewart, Li and Wong, 2005)[14], (Schotta *et al.*, 2004)[15], (Schultz *et al.*, 2002)[16], (Ostapcuk *et al.*, 2018)[17], (Choi, Park and Lee, 2012)[18], (Ogawa *et al.*, 2002)[19], (Sridharan *et al.*, 2013)[20], (Hsieh *et al.*, 2006)[21], (Mateescu *et al.*, 2008)[22], (Smallwood *et al.*, 2012)[23], (Lomberk *et al.*, 2006)[24], (Vakoc *et al.*, 2005)[25], (Vassallo and Tanese, 2002)[26], (Lechner *et al.*, 2000)[27], (Brasher *et al.*, 2000)[28], (Ameyar-Zazoua *et al.*, 2012)[29], (Salton, Voss and Misteli, 2014)[30], (Wu *et al.*, 2015)[31], (Akaike *et al.*, 2015)[32], (Heath *et al.*, 2022)[33]

1.2 Sexual dimorphism

1.2.1 Sex differences in physiology and disease

Sexual dimorphism describes the differentiation at a genotypic and phenotypic level between males and females of the same species. It occurs in many animals, including humans and these differences are commonly attributed to hormonal and developmental differences between the two sexes. Sexual dimorphism in humans can include differentiation of the endocrine systems, alongside the behavioural and physiological effects, internal and external genitalia, height and muscle mass. These differences can play a role not only in physiology, but also in the prevalence, cause and outcome of many common pathologies, including cancer (Clocchiatti *et al.*, 2016), autoimmune diseases like Alzheimer's (Sala Frigerio *et al.*, 2019) or diabetes (Alejandro, 2019) and COVID-19 infection (Peckham *et al.*, 2020). Sexual dimorphism is even evident in cultured cells. For instance, female muscle-derived stem cells regenerate more efficiently compared to their male equivalent (Deasy *et al.*, 2007).

1.2.2 Molecular mechanisms underlying sexual dimorphism

According to the *unified model* by Arnold, "*all ontogenetic sex differences in phenotype derive from the differences in the effects of sex chromosome genes, which are the only factors that differ, on average, in the male and female zygote*" (Arnold, 2009). Indeed, extensive research has proved that sex differences at a molecular level are primarily caused by biochemical modifiers that are expressed in males from the Y chromosome and contribute to male sexual development, which are not present in females. The most important of these modifiers is the sex-determining region Y (SRY) as it controls the sexual differentiation of the gonads from ovaries to testis, hence, setting up life-long differences in production and secretion of gonadal hormones (Gilbert, 2000).

These hormones, especially estradiol and testosterone, act throughout the body in an organizational (permanent) and/or an activational (reversible) fashion and cause most known sexually dimorphic differences (Arnold, Chen and Itoh, 2012). Moreover, while, the X chromosome can be found in both sexes, in female mammals one of the two X chromosomes is randomly inactivated (Xi) during early development by epigenetic mechanisms for dosage compensation (Arnold, Chen and Itoh, 2012). In females, the presence of two X chromosomes can lead to differences in X chromosome gene expression, since some genes, known as "X chromosome escapees" can escape the random X chromosome inactivation and be expressed at double the dose compared to their male counterparts. Around 15% of the X-linked genes in humans and 3% in mice can become X escapees (Carrel and Willard, 1999; Berletch, Yang and Disteche, 2010) therefore leading to sex differences.

The chromatin state of X and Y chromosomes may also affect the availability of epigenetic factors which modulate the chromatin status of autosomal chromosomes and thereby control the expression of genes in a sex-specific manner. The existence of this sex-related mechanism is reinforced by studies in Drosophila indicating that the heterochromatic state of the Y chromosome can act as a "sink" for epigenetic factors and hence alter gene expression of autosomal genes located near heterochromatin by reducing the availability of these factors due to their finite number at a given time (Jiang, Hartl and Lemos, 2010; Lemos, Branco and Hartl, 2010). In the case of mammals, the heterochromatic Xi can cause sexually dimorphic PEV, as a recent study showed that heterochromatin formation is responsible for the expression of autosomal genes that differ among the sexes, with HP1 playing a role in that (Wijchers *et al.*, 2010). Apart from autosomal genes, suppression of repetitive DNA elements including γ-satellites was also regulated by HP1 in a sexually dimorphic manner (Law PhD Thesis, 2015).

1.3 Telomere biology

1.3.1 Telomeric DNA

Telomeres are sequences of repetitive DNA, present at the end of linear chromosomes in eukaryotes. The name derives from the Greek nouns telos (τέλος) meaning "end" and meros (μέρος) meaning "part". The first telomeres to be characterised were from the protozoa Tetrahymena thermophila (Blackburn and Gall, 1978). In vertebrates, the telomeric sequence is comprised of $(TTAGGG)_n$ tandem repeats yet, other organisms have similar but not identical sequences. For instance telomeric sequence in *Arabidopsis thaliana* (plant) is $(TTTAGGG)_n$, in *Schizosaccharomyces pombe* (fission yeast*)* is $(GTTACA)_n$, while *Saccharomyces cerevisiae (*budding yeast) has an even more irregular telomeric sequ(Richards and Ausubel, 1988; Meyne, Ratliff and Moyzis, 1989; Mceachern and Blackburn, 1994; Lejnine, Makarov and Langmore, 1995; Cooper *et al.*, 1997)more, 1995; Cooper et al., 1997). In mammals, the telomeric tract ends with a 3' single-stranded protrusion (3' overhang) which is guanine rich (G-rich). The complementary DNA strand (3'-5') is cytosine rich (C-rich).

Figure 1.8 Mammalian telomeric sequence. Schematic of a mammalian chromosome end with a telomeric array of TTAGGG repeats in the leading strand (brown) and AATCCC repeats in the lagging strand (yellow). The 3' overhang can be seen at the extremity of the telomeric sequence.

Although the telomeric sequence is a highly conserved feature of evolution, telomeric length is not only different among species, but there can be also significant differences among individuals. Mammals present sharp contrasts regarding telomere length. Mice telomeres are quite heterogeneous among different strains and particular strains have telomeres as long as 200 kilobases (kb), while telomeric length in humans is 5-10 times shorter, ranging from 10-12kb in newborns to 4-6 kb in those older than 60 years old (Hastie *et al.*, 1990; Kipling and Cooke, 1990; Lange *et al.*, 1990). *Arabidopsis thaliana* presents a 2-9kb telomeric tract (Richards and Ausubel, 1988), and fission and budding yeast have telomeres only 300-400 base pairs (bp) long (Richards and Ausubel, 1988; Kibe *et al.*, 2003).

Interestingly, telomere length can vary even among the sexes of the same organism as it is the case of humans (*Homo sapiens*), rodents (*Rattus norvegicus*), and reptiles (*Liasis fuscus*). When telomere length differences characterise the two sexes, males displaying shorter telomeres is the norm (Barrett and Richardson, 2011). What appears to be an exception to the rule, comes in the form of a critically endangered species of parrots, namely *Strigops habroptilus,* where females have shorter telomeres than males (Horn *et al.*, 2011).

Telomeres are not only linear structures but can adopt secondary structures. Due to the G-rich nature of telomeric DNA, it has been proposed that there is favouring of G-quadruplex secondary structure (G4) formation. G4s are helical structures containing G tetrads stack atop of each other that have the potential to form from one, two or four strands (Burge *et al.*, 2006). The quadruplex structure is further stabilized by the presence of a cation, usually potassium (K^+), which is located in the centre of each tetrad pair (Sen and Gilbert, 1990). Efforts have been made to unravel the G4 landscape and detection of these structures at telomeres *in vivo* was first achieved with the use of Sty49, a G4-specific antibody to stain algae telomeres (Schaffitzel *et al.*, 2001). G4s were subsequently visualised in telomeres of human cells (Biffi *et al.*, 2013; Henderson *et al.*, 2014). Another secondary structure that telomeres may adopt is the T-loop. The G-rich overhang can fold and invade the telomeric tract to form a T-loop that resembles a lasso-like structure, while this invasion into the double-stranded telomeric DNA generates a smaller, displacement loop (D-loop), thus masking the free end from being recognised as single stranded (Griffith *et al.*, 1999).

1.3.2 Telomeric chromatin

Similar to the rest of the genome, telomeric chromatin of higher eukaryotes is composed of DNA wrapped around nucleosomes. Telomeric nucleosomes are organized in a tightly packed, regularly spaced array, separated by linker DNA that is

about 40 bp shorter compared to bulk DNA (Makarov *et al.*, 1993; Tommerup, Dousmanis and Lange, 1994; Lejnine, Makarov and Langmore, 1995). Interestingly, both *in vitro* and *in vivo* studies suggest that telomeric DNA disfavours nucleosome formation, having the least stable nucleosomes among the tested DNA sequences (Cacchione, Cerone and Savino, 1997; Filesi *et al.*, 2000; Thåström, Bingham and Widom, 2004). This could be connected with an underrepresentation of the linker histone H1, which is normally found in bulk chromatin (Bedoyan *et al.*, 1996; Déjardin and Kingston, 2009).

Telomeric chromatin is generally considered heterochromatic with two main reasons for this attestation. First, when a reporter gene is placed in close proximity to long telomeric regions, it tends to be silenced. This is a phenomenon equivalent to PEV that is known as Telomere Position Effect (TPE) and has been described in different organisms including yeast and mammals (Gottschling *et al.*, 1990; Baur *et al.*, 2001). Proximity in this case refers to spatial vicinity as the gene reporter may be distally located in the linear genome but still silenced from TPE due to DNA looping of telomeric chromatin (Kim & Shay, 2018). Secondly, characterization of telomeres as heterochromatic is based on extensive studies on yeast, and flies. *Drosophila melanogaster* chromosomal ends that consist of telomere-specific retrotransposons instead of telomeric tandem repeats are enriched in histone marks such as H3K9me3 which is bound by HP1. HP1 is essential for (Fanti *et al.*, 1998) integrity (Fanti *et al.*, 1998) and for the recruitment of the SUVAR3-9 HMT that helps heterochromatin propagation (Schotta *et al.*, 2002).

The epigenetic state of telomeric chromatin in mammals is much more controversial. One of the reasons for this can be traced back to technical limitations of ChIP approaches where there is not an easy distinction between real telomeres and TTAGG repeats that are located at subtelomeric regions or at internal sites on the genome, known as internal telomeric sequences (ITSs) (Meyne *et al.*, 1990). ITSs vary in size and are widespread throughout the chromosomes of different vertebrate species, but the majority of them tends to gather around pericentromeric heterochromatin and subtelomeric regions (Azzalin, Nergadze and Giulotto, 2001). Early studies failed to identify ITSs in humans and mice, yet the development of more sensitive probes

allowed the characterization of ITSs in those species as well (Nergadze *et al.*, 2007). ChIP experiments of mouse telomeres demonstrated that they are enriched in heterochromatic marks, including H3K9me3, H4K20me3, and hypoacetylated histones H3 and H4. Meanwhile, genome-wide ChIP-seq supported that these modifications are enriched at telomeres of mESCs, and mouse embryonic fibroblasts (MEFs) (Mikkelsen *et al.*, 2007). HP1 has been also shown to interact with Suppressor of variegation 4-20 homolog 1 (SUV4-20H1) and Suppressor of variegation 4-20 homolog 2 (SUV4-20H2) HMTs that catalyse H4K20me3, thus helping for heterochromatin spreading at telomeres (Schotta *et al.*, 2004). Loss of HMTs (SUV4-20H1/H2 and SUV39H1/H2) affiliated with heterochromatin establishment, results in defective telomeres with increased length, yet the mechanism of telomere elongation is unclear (García-Cao *et al.*, 2(García-Cao *et al.*, 2003; Gonzalo *et al.*, 2005, 2006). On the other hand, a recent study using an unbiased proteomic approach termed Proteomics of isolated chromatin segments (PICh) (Déjardin and Kingston, 2009) showed low levels of H3K9me3 at telomeres in mouse ES cells (Gauchier *et al.*, 2019). In the same study, the authors show that telomeric H3K9me3 deposition is dependent on SETDB1 and not on SUV39H1/H2, which instead primarily act at pericentromeric regions.

Even more obscure is the epigenetic state of human telomeres. ChIP experiments on human fibroblasts indicate that the heterochromatic marks such as H3K9me3, H4K20me3, and H3K2me3 display unexpectedly low levels at telomeres (O'Sullivan *et al.*, 2010). This finding is further supported by a recent study testing nine different cell lines and interestingly showed that H3K9me3 is only enriched at telomeres of cells with the alternative lengthening of telomeres (ALT) pathway activated (Cubiles *et al.*, 2018). H3K9me3 is under-represented at telomeres, whereas it is enriched at subtelomeres, although at lower levels than at centromeric and pericentromeric regions (Cubiles *et al.*, 2018). EZH2 protein, a member of the Polycomb Repressive Complex 2 (PRC2), has been shown to target human telomeres where it catalyses H3K27me3, which appears to be necessary for deposition of other heterochromatic markers including H3K9me3, H4K20me3 and HP1 (Benetti, García-Cao and Blasco, 2007). Genome-wide ChIP-seq analyses have shown that H3K27me3 and H3K9me3 do not largely overlap, even at repetitive DNA sequences (Mikkelsen *et al.*, 2007; Kato,

Takemoto and Shinkai, 2018) making this site-specific interplay between PCR2, H3K9me3 & HP1 highly atypical.

1.3.3 The Shelterin complex

Shelterin is a protein complex essential for telomere maintenance (Figure 1.9), structure and protection that is composed of six different proteins (de Lange, 2009).Telomere Repeat binding Factor 1 (TRF1) (Chong *et al.*, 1995) and Telomere Repeat binding Factor 2 (TRF2) (Bilaud *et al.*, 1997), bind double-stranded telomeric DNA with their C-terminal Myb domains(de Lange, 2009). DNA binding of TRF1 and TRF2 is further enhanced by homodimerization (Fairall *et al.*, 2001) domains (Fairall *et al.*, 2001). Protection of Telomere 1 (POT1) (Baumann and Cech, 2001) binds the single-stranded 3' overhang via its N-terminal OB-fold domains (Loayza *et al.*, 2004). Unlike human telomeres, mouse telomeres have two POT1 paralogs with slightly different functions. POT1a is required for DNA damage signal repression, while POT1b regulates the amount of single-stranded DNA at telomeres (Hockemeyer *et al.*, 2006). The other three shelterin components, namely TRF1- and TRF2-Interacting Nuclear Protein 2 (TIN2), Adrenocortical Dysplasia Protein (ACD or TPP1) and the Repressor/Activator Protein 1 (RAP1) do not bind directly on DNA, but stabilize the complex via protein-protein interactions. TIN2 acts as a bridging component, tethering TRF1, TRF2 and TPP1 together. POT1 association with the shelterin depends on its binding partner TPP1 (Ye *et al.*, 2004). Finally, the recruitment of RAP1 to telomeres relies on its exclusive binding to TRF2 since it lacks a telomeric-DNA binding domain (Li et al., 2000).

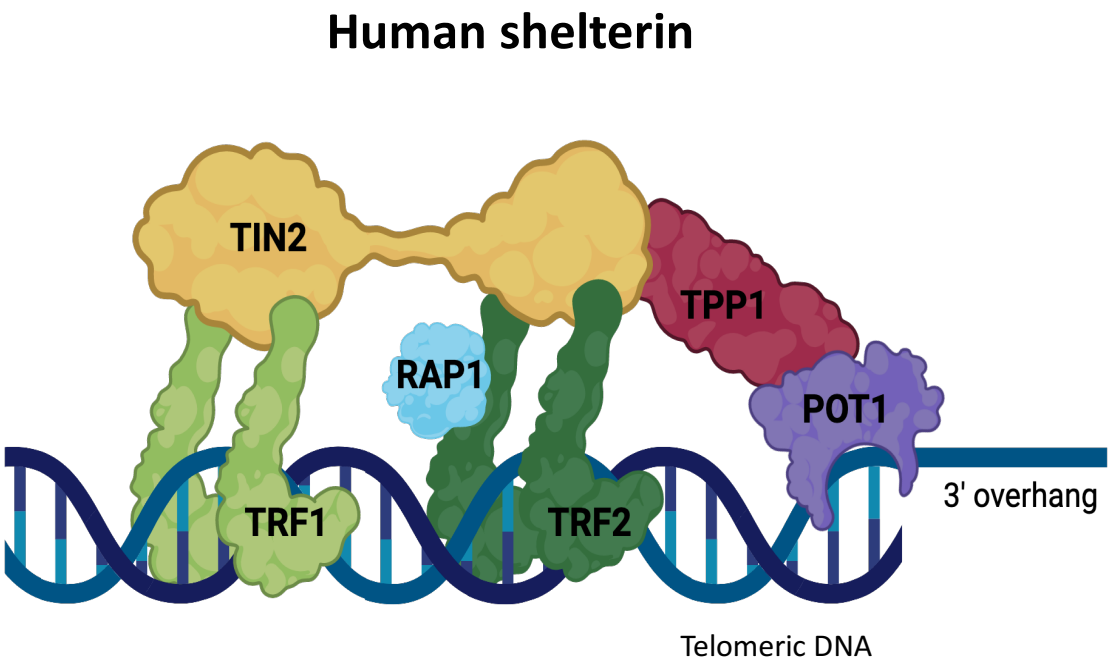


Figure 1.9 The human shelterin complex. Shelterin is composed of six protein subunits. TRF1 (light green) and TRF2 (dark green) form homodimers and bind double-stranded telomeric DNA, while POT1 (purple) binds on the single-stranded 3' overhang. TIN2 (yellow) acts as a bridge to tether the whole complex together with RAP1 (light blue) and TPP1 (burgundy) stabilising the complex via protein-protein interactions.

1.3.4 Telomere function

Genomic stability requires the solution of at least two important biological problems: the chromosome end-protection problem and the chromosome end-replication problem. The first problem refers to the quandary of the DNA damage repair machinery to distinguish the ends of linear chromosomes from DSB lesions. If this distinction is not accurately made, unscheduled DNA repair can lead to fusions of different chromosomes, with catastrophic genomic instability as a consequence.

The second problem refers to the gradual loss of genomic material through every cell cycle, owing to incomplete lagging strand DNA synthesis and exonucleolytic processing reactions (Watson J D, 1972; Olovnikov, 1973).

1.3.5 How do telomeres solve the end-protection problem?

The end-protection problem requires masking the chromosomal end from the DDR machinery (Figure 1.10). One of the mechanisms that appears to be in place for telomeres to avoid recognition as DSBs is the T-loop formation. It has been suggested that this conformation does not allow MRN complex (MRE11, RAD50 and NBS1) to identify the 3' overhang as a region of DNA damage (Palm and de Lange, 2008). The binding of MRN complex to DSBs promotes ATM-dependent DDR. T-loops were initially observed with electron microscopy (Griffith *et al.*, 1999) but only in a small number of telomeres (de Lange, 2004), making it unlikely to be the only mechanism protecting telomeres from the DNA repair pathways. It has been shown that for T-loop formation, TRF2 is essential (Doksani *et al.*, 2013), suggesting that shelterin holds together the structure, yet T-loops dynamics throughout the cell cycle are poorly understood (Poulet *et al.*, 2009; Benarroch-Popivker *et al.*, 2016; van Ly *et al.*, 2018).

The role of the shelterin complex as a DDR inhibitor is also evident and its function could be particularly important in linear telomeres lacking the T-loops. TRF2 can directly inhibit Ataxia Telangiectasia Mutated (ATM) kinase activation, which is a prime effector of DSB response activation and also suppresses the E3 ubiquitin ligase RNF168, which is found downstream of this DNA-damage signalling cascade (van Steensel, Smogorzewska and de Lange, 1998; Okamoto *et al.*, 2013). POT1, another member of the shelterin complex can counteract Ataxia Telangiectasia and Rad3-related (ATR) kinase activation, enabling telomeres to avoid the DNA damage surveillance. While ATM is crucial for recruitment of DSB repair machinery, ATR is also responsible for initiating the single strand break (SSB) response (Hurley, Wilsker and Bunz, 2006). POT1 also regulates the active trimming of the 5' strand. Its tethering to telomeres occurs via TPP1 and TIN2 interaction (Takai *et al.*, 2010). Finally, TRF1, by reducing replication stress, ensures correct replication of telomeres and therefore prevents DDR activation (both ATM and ATR pathways)(Martínez *et al.*, 2009; Sfeir *et al.*, 2009).
Interestingly, a recent report suggested that shelterin further safeguards against DDR activation by compacting telomeric chromatin into dense globular structures, rendering it inaccessible to DDR factors (Bandaria *et al.*, 2016). However, other reports

argue against this, dissociating sheleterin-mediated telomere compaction from DDR protection (Timashev *et al.*, 2017; Vancevska *et al.*, 2017).

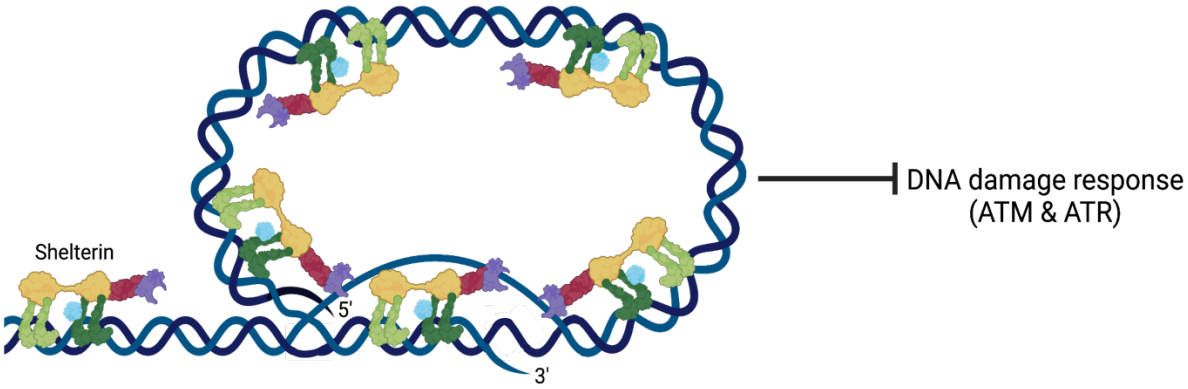


Figure 1.10 Inhibition of DNA damage response by the telomeres. Representation of telomeric DNA configured in a T-loop. The T-loop, in combination with the shelterin complex prevents the access of the DDR kinases ATM and ATR at the end of the chromosomes. Shelterin factor POT1 (purple) binds on the ssDNA region of the D-loop.

1.3.6 How do telomeres solve the end-replication problem?

DNA at the end of the linear eukaryotic chromosomes cannot be fully copied during each cell division resulting in gradual chromosomal trimming. This is primarily due to an inherent limitation of the DNA polymerases. Specifically, DNA polymerases exclusively synthesise DNA in the 5'-3' direction. During DNA replication, the leading strand is synthesised continuously, while the lagging strand is created in small pieces called Okazaki fragments. Each Okazaki fragment is generated from an RNA primer and then joined together in a continuous strand upon removal of the RNA primers. Nonetheless, at the end of the chromosomes, the very last stretch of DNA cannot generate an Okazaki fragment as the necessary RNA primer would have to be located beyond the chromosome end (Olovnikov, 1973; Watson, 1972). DNA polymerase fails also to fill in the gap left behind by the RNA primer of the outermost Okazaki fragment that does get generated. In human cells, this 5' most RNA primer in fact is positioned as much as 60 to 100 bps away from the chromosome end, resulting in the loss of the corresponding length from the lagging strand with every cell division (Harley, Futcher and Greider, 1990; Chow *et al.*, 2012). Furthermore, there is an active resection of telomeric nucleotides from the lagging strand, exerted by the nucleases EXO1 and APOLLO (Lingner, Cooper and Cech, 1995). The combinational effect of passive and active shortening of the lagging strand during DNA replication results in the 3' single-stranded DNA overhang formation.

Due to their repetitive nature, telomeres can act as a buffer, protecting the genome from losing essential genetic information while it has been also proposed that they act as "mitotic clocks". By shortening in a replication dependent manner, telomeres could provide a means for the cells to control and limit the number of their divisions. Indeed, somatic cells have a limited number of divisions which is dictated by the Hayflick limit. Hayflick and Moorhead first proposed that a minimum telomere length is necessary for cells to continue dividing, while below that critical length (the Hayflick limit), cells either enter senescence or die through programmed cell death, termed apoptosis (Hayflick and Moorhead, 1961).

1.3.7 Cellular senescence

It is well established that populations of human fibroblasts can proliferate only for a finite period of time in culture. After a certain number of divisions, cells change in size and morphology, and ultimately become incapable of proliferation (Hayflick and Moorhead, 1961). This phenomenon is known as cellular or replicative senescence and describes the state of irreversible cell growth arrest, without loss of cell viability. Unlike cell death, during senescence there is continuation of metabolic activity (Cristofalo and Pignolo, 1993).

Telomeres are directly linked with replicative senescence. Indeed, studies have shown that in human cells four to five critically short telomeres are sufficient to trigger senescence (Kaul *et al.*, 2012), while in yeast only one critically short telomere is required (Abdallah *et al.*, 2009; Bourgeron *et al.*, 2015). Upon reaching a critical short length, telomeres lose their normal structure, are recognised as sites of DNA damage and there is subsequent activation of the DDR pathway. The two key players of the DDR pathway implicated with telomere-induced senescence are the tumour protein 53 (p53) and the tumour protein 16 (p16, also known as $p16^{INK4a}$). Both of these factors can arrest cell proliferation. Specifically, p53 does not allow the G1/S progression by activating p21 which in turn deactivates cyclin-dependent kinase 2 (CDK2). Without CDK2, retinoblastoma protein (pRB) remains active, inhibiting E2F transcription factor 1 (E2F1), a crucial transcription factor that allows cell transition from G1 to S phase. p16 also activates pRB through inactivation of cyclin-dependent kinase 4 (CDK4) and cyclin-dependent kinase 6 (CDK6) (Rayess, Wang and Srivatsan, 2012).

Bypassing of senescence allows cells to proliferate indefinitely, gaining “immortality”, which is a hallmark for cancer (Hanahan and Weinberg, 2011), thus senescence is regarded as a natural defensive mechanism against tumourigenesis. Overcoming senescence may be achieved due to defects of senescence effectors or activation of telomere maintenance mechanisms.

1.3.8 Telomere maintenance mechanisms

There are two main pathways that allow telomere length maintenance. The first is based on an enzyme known as telomerase and the other pathway is based on homologous recombination (HR), and is known as the alternative lengthening of telomeres (ALT) pathway. Some cancer cells have the ability to exploit a combination of the two pathways to gain immortality (Recagni *et al.*, 2020).

1.3.9 Telomerase function and structure

Telomerase is a ribonucleoprotein complex that can synthesize telomere repeats *de novo* (Greider and Blackburn, 1985). Telomerase is composed by two essential subunits; the catalytic subunit which is the telomerase reverse transcriptase (TERT) and the RNA component termed TR or TERC which act as a primer for telomere lengthening
(Feng *et al.*, 1995; Weinrich *et al.*, 1997). Other accessory proteins such as dyskerin (DKC1), TCAB1, NHP2, NOP10 and GAR1 complete the telomerase holoenzyme (Nguyen *et al.*, 2018). Telomerase access to the 3' overhang requires the transient dismantle of t-loops, a process that takes place in S phase and is catalysed by the Regulator of Telomere Length 1 (RTEL1) (Sarek *et al.*, 2015). This enables the extension of the G-rich leading strand during DNA replication by 50-60 nucleotides in a physiological context. The newly extended strand functions as a template for the follow-up elongation of the lagging C-rich telomeric strand. This process is initiated by the CTC1-STN1-TEN1 (CST) complex that displaces telomerase from the G-rich sequence and recruits DNA polymerase alpha to the complementary telomeric strand (Chen, Redon and Lingner, 2012), resulting eventually in the elongation of both telomere tracts (Y. Zhao et al., 2009).
Several factors appear to influence telomerase stability and localisation, with a recent report also implicating histones H2A and H2B in the folding and function of telomerase (Ghanim *et al.*, 2021). The assembly of telomerase initiates in the membranelles organelles called Cajal bodies that also appear to assist with the delivery of telomerase to telomeres (Cristofari *et al.*, 2007). It is not clear what drives

Cajal bodies to telomeres but it has been shown that the colocalisation occurs in S phase (Tomlinson *et al.*, 2006). Telomerase has been also shown to diffuse in the nucleus, and localise on telomeres not necessarily proximal to the Cajal bodies that telomerase was assembled in (Schmidt, Zaug and Cech, 2016).

Telomerase recruitment at telomeres is facilitated by the shelterin complex. Specifically, there is a direct interaction of TERT with TPP1 (Xin *et al.*, 2007; Abreu *et al.*, 2010; Nandakumar *et al.*, 2012), while TRF1 needs to be also displaced in order for telomerase to gain access to the underlying telomeric DNA (Tong *et al.*, 2015). Tertiary telomeric structures, including G4s might also affect telomerase function. It has been proposed that G4s inhibit telomerase activity (Zahler *et al.*, 1991; Wang *et al.*, 2012), yet another report argues that G4s may also be contributing to its recruitment (Moye *et al.*, 2015). Accordingly, *in vitro* experiments have shown that POT1 can help with telomerase function by G4 dissolution (Kelleher, Kurth and Lingner, 2005), while when stably bound on telomeres it could interfere with telomerase activity (Kelleher, Kurth and Lingner, 2005).

1.3.10 Telomerase regulation

Because of telomerase importance in cell physiology, its activity is tightly regulated. Unlike in human cells, telomerase activity is not repressed in normal mouse cells (Gorbunova and Seluanov, 2009) and while TERC is expressed ubiquitously in normal human cells, TERT is highly repressed in this context. It is, nevertheless, present in germ and ES cells and highly expressed in most cancer types, especially those in an advanced stage (Kim et al., 1994; Shay & Bacchetti, 1997). Usually, cancer cells have short telomeres and, in many cases, even shorter than normal cells, owing to late activation of telomerase. Nonetheless, once telomerase is activated, it can keep telomere length constant, allowing cells to circumvent senescence or apoptosis. It has been suggested that under these conditions, every telomere is extended by telomerase during the S phase (Zhao *et al.*, 2009). A study showed that the average

telomere is bound by telomerase more than 2,000 times over the course of a single S-phase (Schmidt, Zaug and Cech, 2016).
CST has been also described as an inhibitor of telomerase, by terminating its activity. CST displays increased association with telomeres during late S and early G2 phase, the period of time where telomerase action has been completed, while mutations of CST result in elongated telomeres (Chen, Redon and Lingner, 2012).

Drugs targeting telomerase have been widely used in the clinic for cancer treatment since 1990. The main advantage of targeting telomerase is substantiated by the fact that normal cells do not express telomerase and at the same time, there is a decreased possibility of resistance development, compared to targeting signalling transduction enzymes. Recent advances in the structure of telomerase holoenzyme promise to improve drug efficacy (Nguyen *et al.*, 2018; Ghanim *et al.*, 2021). However, around 10-15% of human cancers can maintain their telomere length via a telomerase-independent manner, thus rendering inhibitors of telomerase ineffective in this context (Harley, 2008). These types of cancer usually include astrocytomas, glioblastomas and sarcomas (Heaphy *et al.*, 2011).

1.3.11 Alternative lengthening of the telomeres (ALT)

The other pathway that enables telomere length maintenance is known as Alternative lengthening of the telomeres (ALT). During ALT, it is presumed that telomere length is maintained through HR, instead of a telomerase-dependent mechanism. Telomeric HR could arise by strand invasion of the 3’ overhang into the adjacent telomere with the newly hybridized telomeric sequence acting as a template for elongation. Otherwise, the template can be the telomere of the sister chromatid or extrachromosomal (linear or circular) telomeric DNA (Cesare and Reddel, 2010). Human ALT telomeres, range from 3 kb to more than 50 kb (Bryan *et al.*, 1997) and ALT has been almost exclusively been observed in pathological context including human tumours or cancer cell lines (Neumann *et al.*, 2013). Replication stress has

been attributed as a main contributor for the induction of this pathway (Brault and Autexier, 2011). Interestingly, ALT cells do not display elevated genomic recombination events *per se*, rather a telomere-specific dysfunction (Bechter *et al.*, 2003).

At the cellular level, ALT is characterised by different markers including striking telomere length heterogeneity, the presence of a specific subclass of promyelocytic leukaemia (PML) bodies termed ALT-associated PML bodies (APBs), abundant extrachromosomal telomeric DNA (C-circles) and high rate of telomeric sister-chromatid exchanges (T-SCE) events (Conomos, Pickett and Reddel, 2013a). Recent studies also displayed that ALT cells are enriched with non-canonical telomere variant repeats (TVRs) at regions proximal to the subtelomeres. TVRs are sequences where nuclear factors like NR2C2 (TR4) and NR2F2 (COUP-TF2) can bind and promote break-induced telomere extension (Déjardin and Kingston, 2009; Conomos *et al.*, 2012).

PML bodies are spherical structures that exist in the nucleus of many cell types and have been associated with a variety of cellular functions, including response to DNA damage, cellular stress, gene transcription or apoptosis. PML bodies are mainly composed by PML proteins but they also contain factors that are recruited after specific stimuli (Bernardi and Pandolfi, 2014). While PML bodies are not usually linked to telomere biology, APBs contain additionally to the usual components, such as PML and Sp100, telomere-related and DNA recombination factors. While their role is not clear, they have been shown to localise on telomeres of ALT cells and their numbers peak in G2 phase when normally recombination events occur (Grobelny, Godwin and Broccoli, 2000). Therefore, they have been considered as “platforms” for telomeric HR while also serving as reservoirs of telomere-associated factors (Draskovic *et al.*, 2009).

Early FISH studies revealed that there is a significant presence of extrachromosomal telomeric DNA in ALT cells (Tokutake *et al.*, 1998), which can either take the form of partially single-stranded (C-circles or G-circles) or double-stranded (T-circle) circles (Nabetani and Ishikawa, 2009). T- circles seem to arise from improper resolution of T-loops, resulting in truncated telomeres. Unlike T-circles that have been also

foun(Pickett *et al.*, 2009)ve cells (Pickett *et al.*, 2009), C-circles and G-circles are the only ALT-specific feature. In humans, G-circles are barely detectable, with their levels being 100-fold lower compared to C-circles. C-circles are readily detected by a rolling circle amplification assay, employing the Φ29 DNA polymerase called C-circle assay (CCA) and their levels are significantly increased in the blood of ALT-positive osteosarcoma patients, suggesting that C-circles could act as clinical biomarkers (Henson *et al.*, 2009). Their formation is considered to be a consequence of replication stress during the telomeric HR events that take place in ALT cells.

The chromatin status at the telomeric and subtelomeric region seems to play a role in the emergence of ALT with an overall a loss of heterochromatin marks associating with increased ALT activity (Lovejoy *et al.*, 2012). There is elevated T-SCE and recombination frequency at telomeres in human cancer cell lines after treatment with the DNA demethylating agent 5-aza-deoxycitidine suggesting that hypomethylation of subtelomeric regions positively correlates with ALT induction (Vera *et al.*, 2008). Moreover, deficiency of the H3.3 histone in mESCs leads to a decrease of the H3K9me3 and H4K20me3 levels, which in turn results in elevated TElomeric Repeat containing RNA (TERRA) expression, and increased recombination events (Udugama *et al.*, 2015).

Another factor playing a prime role in telomere length homeostasis is the Kruppel-like zinc finger protein TZAP. It does so, by directly binding on longer telomeres that are not occupied by the shelterin factors TRF1 and TRF2. This binding promotes active telomere trimming, which involves the dissolution of T-loops (Griffith *et al.*, 1999) and formation of extra chromosomal telomeric DNA (Pickett *et al.*, 2009), thus preventing the accumulation of aberrantly long telomeres. The mechanism of TZAP action is exacerbated in ALT cells (Li *et al.*, 2017).

1.3.12 TERRA transcription, structure and regulation

Due to the presence of heterochromatin markers at telomeres, it was long believed that these sites were transcriptionally silent. This dogma was questioned for the first time in 1989, where telomeres of *Trypanosoma brucei* were observed to transcribe into species of non-coding RNA (ncRNA) (Rudenko *et al.*, 1989). More recently, with the discovery of TElomeric Repeat containing RNA (TERRA), this dogma was confuted (Azzalin *et al.*, 2007; Schoeftner and Blasco, 2007). TERRAs are ncRNAs containing the canonical telomeric sequence $(UUAGGG)_n$ as well as sequences unique to the subtelomeric regions of the different chromosomes that they derive from (Figure 1.11). A plethora of organisms was found to express these molecules, including mammals, zebrafish, worms and budding yeast (Azzalin *et al.*, 2007; Schoeftner and Blasco, 2007; Luke *et al.*, 2008).

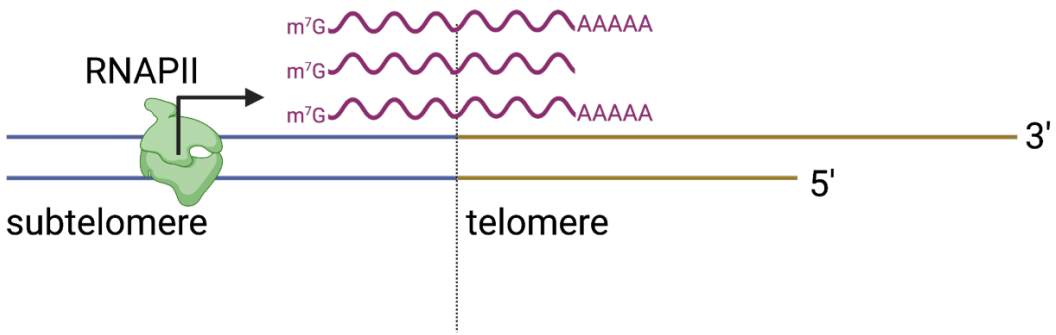


Figure 1.11 Overview of TElomeric Repeat-containing RNA (TERRA) transcription. RNA Polymerase II (RNAPII) is mainly responsible for the expression of TERRAs and the direction of their expression is 5' to 3' direction (from the centromere to telomere). TERRA transcription starts from the subtelomeric tract. Adapted from Barral and Déjardin, 2020.

TERRAs are transcribed from the C-rich strand, with their transcription start sites (TSSs) located in the subtelomeric tract and their transcription termination sites (TES) residing in the telomeric tract (Maicher, Lockhart and Luke, 2014). The chromatin of TERRA promoters is enriched for active histone marks including H3K4me3, even though the rest of the subtelomeric chromatin retains the heterochromatin-associated

markers like H4K20me3, H3K9me3 and HP1. In humans, TERRA TSSs can be found in CpG island promoters (Nergadze *et al.*, 2009). CpG islands are DNA regions, rich in cytosine-guanine dinucleotides repeats that characterise gene promoters and their methylation status provides an extra layer for transcription regulation (Illingworth *et al.*, 2010; Maunakea *et al.*, 2010). Indeed, mutations of various DNA methyltransferases (DNMTs) can affect TERRA expression (Nergadze *et al.*, 2009). For instance, patients suffering from the Immunodeficiency, Centromere instability, Facial (ICF) syndrome in which the DNMT3b is mutated, display high levels of TERRA and short telomeres (Sagie *et al.*, 2017), while loss of methylation at subtelomeric CpG islands in HeLa cells caused an upregulation of TERRA expression (Berre *et al.*, 2019).

It appears that RNAPII is the main RNA polymerase responsible for TERRA transcription. Specific inhibition of RNAPII by *α*-amanitin reduces TERRA levels in both mouse and human cells (Schoeftner and Blasco, 2007). Similar results have been reported for yeast RNAPII, where when mutated, there is a reduction of TERRA transcripts (Luke *et al.*, 2008). It was proposed that CCCTC-binding factor (CTCF) recruits RNAPII and cohesin at TERRA's promoters to orientate transcription towards the telomere (Stong *et al.*, 2014). Similar to other transcripts generated by RNAPII, 5' ends of TERRAs contain a 7-methylguanylate (m^7G) cap, at least in humans (Porro *et al.*, 2010).

Additionally, it has been shown that TERRAs can be polyadenylated at their 3' end. In budding yeast, the polyadenylation of TERRAs occurs by the poly(A) polymerase Pap1, yet the mechanism of polyadenylation remains unclear as a poly(A) signal is missing from the telomere tract (Luke *et al.*, 2008). Polyadenylation provides stabilization to the TERRA molecules, as polyadenylated TERRAs have a half-life of 8 h, while the non-polyadenylated fraction is significantly less stable with a half-life of 3 h (Porro *et al.*, 2010). All yeast TERRAs and around 7% of human TERRAs are polyadenylated (Azzalin and Lingner, 2008). In mouse, the polyadenylated fraction of TERRAs is around 2.5-fold enriched compared to the non-polyadenylated fraction (de Silanes *et al.*, 2014). The polyadenylation status also affects TERRA localisation. In mESCs the non-polyadenylated molecules localise *in cis* to telomeres while the polyadenylated TERRAs that also comprises the bigger fraction of TERRA molecules localises *in trans*

to other genomic regions, including intergenic regions, introns and gene promoters (H. P. Chu *et al.*, 2017).

A consensus has not been reached on whether TERRAs arise from all chromosome ends or from only a few of them. One team has shown that in U-2 OS (human ALT cell line) the majority of TERRAs arise from a single locus on chromosome 20q but there is also lower expression from other loci including the X chromosome (Montero *et al.*, 2018). The same team provided evidence that in mice TERRAs primarily arise from the subtelomeres of chromosome 18 and to a shorter extend from chromosome 9 (de Silanes *et al.*, 2014). On the contrary, other teams have shown that TERRA can stem from multiple chromosome ends both in mice (Mazzolini *et al.*, 2018; Liu *et al.*, 2019) and in humans (Arnoult, van Beneden and Decottignies, 2012; Deng *et al.*, 2012; Feretzaki, Nunes and Lingner, 2019). TERRA expression has been observed at least from chromosomes 1q, 2p, 9p, 10q, 12q, 13q, 15q,16p, 17p,18q, 20q, Xq and Yq in human lung fibroblasts (HLF), HeLa and U-2 OS cell lines (Feretzaki, Nunes and Lingner, 2019). Moreover, another recent study argues that a significant fraction of TERRAs in mESCs arise from the subtelomeric, pseudoautosomal regions (PARs) of the sex chromones, instead of the autosomal chromosomes, and thus they term this fraction PAR-TERRA (Chu et al., 2017).
TERRA length is less of a debate as it has been accepted that these transcripts are very heterogeneous in length, with reports showing that TERRAs range from 100 bp to at least 9 kb in mammals (Azzalin *et al.*, 2007). A recent report estimates their length to be over a thousand kilobases in mice (Chu et al., 2017). TERRA size also correlates with their polyadenylation status, as almost exclusively, transcripts larger than 2 kb are polyadenylated (Azzalin and Lingner, 2008).

As an important telomeric component, TERRA transcription is finely regulated during the cell cycle to guarantee that telomeres can overcome the end-protection and end-replication problem. Due to the complementarity of TERRA molecules with telomeric DNA, TERRAs can hybridize to the exposed C-rich lagging strand during replication and DNA:RNA hybrids may form at telomeres, which in turn, by displacing the G-rich strand can give rise to R-loops (Balk *et al.*, 2013). High TERRA levels found in ICF patients have been associated with telomere shortening, as they can give rise to the

aforementioned structures that may interfere with the semiconservative replication of telomeres (Sagie *et al.*, 2017).

While Ribonucleases H (RNase H1 and RNase H2) have the ability to resolve DNA:RNA hybrids, TERRA transcription is programmed to not overlap with replication through the cell cycle as an extra safety net. TERRA accumulates in early G1 and drops in S phase reaching the lowest expression levels as cells progress from late S to G2 phase (Porro *et al.*, 2010). This staggered process is realised owing to the nonsense-mediated mRNA decay (NMD) surveillance pathway. Problems of this pathway can lead to loss of entire telomeric tracts (Azzalin *et al.*, 2007). The NMD pathway ensures the displacement of TERRAs from telomeres but not regulation of their degradation speed or their total numbers, as loss of key components of the NMD pathway led to permanent binding of TERRAs to telomeres and telomere instability. Telomere instability is presumed to arise due to replication stress derived from the high levels of DNA:RNA hybrids (Balk *et al.*, 2013; Sagie *et al.*, 2017).

Moreover, the shelterin complex seems to be playing a role in TERRA transcription regulation, as depletion of TRF1(Schoeftner and Blasco, 2007), or TRF2 leads to lower levels of TERRA (Porro *et al.*, 2010). Interestingly, studies in *S. pombe* show limited transcription of the G-rich telomeric strand, resulting in the production of $(CCCUAA)_n$ RNA molecules termed ARIA but these species are barely detectable, possibly due to fast degradation (Bah *et al.*, 2012).

1.3.13 TERRA function

TERRA functionality is as broad as its interactome of chromatin, telomere, DNA transcription/replication, RNA-binding, nuclear matrix and cell cycle-related factors (H. P. Chu *et al.*, 2017). At the same time, the diverse localisation of TERRA fractions, which is affected by their polyadenylation status, hints for different roles.

While TERRA effects in *trans* have not been extensively studied, an interesting role for the PAR-TERRA fraction is homologous sex chromosome pairing. X-chromosome pairing is necessary early in development for the internal chromosome counting mechanism that determines the number of X chromosomes and initiates X chromosome inactivation (XCI) in case there are two or more X chromosomes. PAR-TERRAs play a role in this mechanism by anchoring the X inactivation center (Xic) to the PAR, creating a "tetrad" of homologous interactions which in turn initiates XCI. Depleting PAR-TERRA leads to pairing abrogation, thus stopping the initiation of XCI. Interestingly, male cells also undergo PAR-PAR pairing, as they still need to count their X chromosomes. However, due to the lack of a second X chromosome and thus a second Xic, initiation of XCI does not occur (Chu et al., 2017).

Moreover, TERRAs appear to act as epigenome modulators at least in mESCs. Through functional antagonism with ATRX, an important chromatin regulator, TERRAs regulate distal gene expression. When TERRAs are bound on shared gene targets, they promote gene transcription, while when ATRX is bound, it suppresses it. For genes where binding of these two factors is not co-occurring, TERRAs have either an activating or a silencing effect, likely based on context-specific transcription factors. TERRAs' actions are not limited to distal gene regions but are also affecting genes at subtelomeres, where loss of TERRAs cause(H. P. Chu *et al.*, 2017) genes (H. P. Chu *et al.*, 2017). At the same time, TERRAs play a major *in cis* role, by regulating multiple levels of telomere physiology, including chromatin structure, replication and telomerase recruitment. The antagonistic relationship of TERRA and ATRX is even apparent at telomeres. There is a direct interaction of TERRAs with ATRX which displaces the latter from telomeric sequences, while depletion of TERRAs leads to increased localisation of ATRX at telomeres.

TERRA has been also shown to regulate telomeric chromatin structure. TERRA recruits HP1*α* and the Polycomb complex (PRC2) at telomeres which facilitates the deposition of the H3K27me3 on this region. This marker sequentially recruits SETDB1 methylase that facilitates H3K9me3 (Gauchier *et al.*, 2019). The presence of HP1α in combination with the elevated H3K9me3 density causes further heterochromatinization by the recruitment of other isoforms of the HP1 family,

including HP1γ and further deposition of heterochromatin markers including H4K20me3 (Montero *et al.*, 2018). Additionally, TERRAs assist with the stabilisation and maintenance of the heterochromatic landscape at telomeres by recruiting TRF2 and origin of replication complex (ORC), a known interactor of HP1 (Z. Deng *et al.*, 2009).

TERRAs are also essential for proper telomere replication. TERRAs can interact with the heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1) and regulate its telomeric localisation. During early S phase, the telomeric 3'-overhang is coated with the replication protein A (RPA). In late S phase, when TERRA levels are low, hnRNPA1 displaces RPA from the chromatin ends. In G2/M TERRA levels are elevated again and TERRAs sequester hnRNPA1 away from the 3'-overhang, allowing the binding of POT1 and the stabilisation of the telomeres (Flynn *et al.*, 2011). Furthermore, CTCF-driven TERRA transcription facilitates chromosome stability and telomere replication but the mechanism is poorly understood (Beishline *et al.*, 2017).

Mouse and human *in vitro* experiments demonstrated that TERRAs inhibit telomerase recruitment to telomeres by binding to both TR and the TERT subunits (Schoeftner and Blasco, 2007), while a recent mouse study showed that knockdown of TERRAs stimulated telomerase activity (Chu et al., 2017). Nevertheless, recent *in vivo* studies in yeast showed a positive TERRA role in telomerase recruitment to short telomeres to promote their preferential elongation, while, the upregulation of TERRA transcription resulted in increased telomerase activity (Cusanelli, Romero and Chartrand, 2013; Moravec *et al.*, 2016). The combination of these findings potentially suggests a species-specific mechanism of action.

Overall aim and specific hypotheses

The overall aim of this PhD project is to investigate the role of HP1γ on sexual dimorphism and its function in telomere maintenance. This is addressed experimentally by a classical gene knockout approach using mouse as a model organism. Specifically, primary mouse embryonic fibroblasts (MEFs) from day 13.5 day of mouse development were employed.

While HP1γ was shown to be important for the regulation of genes that differ in expression among the sexes, the underlying molecular mechanisms are less understood. The involvement of HP1γ in both repression and activation of genes in a context-specific manner could account for these differences. In Chapter 3 of the results, I will show and explain my work on how HP1γ exerts transcriptional regulation through its binding to gene promoters and what are the physiological consequences *in vivo* and *ex vivo* for both sexes. To this end, the following objectives have been examined:

- What is the effect of HP1γ on cell proliferation, cellular senescence and mouse embryo growth in a sexually dimorphic manner?

- What is the mechanism of action of HP1γ on gene expression with a focus on genes that differ among the sexes?

In Chapter 4, the importance of HP1γ on telomere stability has been investigated focusing on its role as transcription regulator, not only regarding the expression of the long-non coding RNAs (TERRAs) arising from telomeres, but also from interstitial genes necessary for telomere maintenance. The potential sex-specific consequences at telomeres following the loss of such an important chromatin regulator were examined by setting these objectives:

- Analysis of publicly available RNA-sequencing data of MEFs lacking HP1γ for factors that are a necessary for telomere maintenance and could differ among the sexes.
- What is the effect of HP1γ on TERRA regulation?
- Characterisation of the effects of HP1γ depletion on mouse telomeres while considering sex as a biological variable.

Chapter 2. Materials and methods

2.1 Animal handling and transgenic mice genotyping

Animals used in this study were maintained and handled according to the Imperial College London guidelines for Animal Research and the regulations of the British Home Office.

DNA was extracted from mouse ear punches or mouse embryo heads following the HotSHOT protocol (Truett *et al.*, 2018). Briefly, 75 µl of Alkaline Lysis Solution (25 mM NaOH, 0.2 mM Na_2EDTA) was added to the samples, that were then incubated at 95 °C for 30 minutes (min). Samples were cooled down on ice and 75 µl of Neutralizing Solution (40 mM Tris-HCl) was added. DNA samples were analysed with Polymerase Chain Reaction (PCR) and samples were resolved with a 1.5 % weight per volume (w/v) agarose gel run at 100 V for 40 min. The list of primers and the PCR conditions can be found in Tables 2.1-2.3.

Table 2.1 PCR primers used in this study

Primer Name	**Primer Sequence 5' to 3'**
HP1γ common Fw	GAGTGATTACCGACACCACCA
HP1γ wild-type Rev	TTTAATCGGAGACTTGAAGAGC
HP1γ mutant Rev	GTTCGCTTCTCGCTTCTGTT
KDM5D Fw	ACAAAGTGGGGGCAAAAAGT
KDM5D Rev	AGTTATGACCCTCACCACAAGA

Table 2.2 HP1γ PCR parameters

Cycles	Temperature (°C)	Time
1	96	3 min
5	96 70 (-1 °C per cycle) 72	15 sec 15sec 40 sec
30	96 58 72	15 sec 15 sec 45 sec
1	72	10 min
1	12	hold

Table 2.3 KDM5D PCR parameters

Cycles	Temperature (°C)	Time
1	96	5 min
35	96 60 72	30 sec 30 sec 15 sec
1	72	10 min
1	12	hold

PCR of HP1γ wild-type ($HP1\gamma^{+/+}$) mice produces a single band of 501 bp with primers testing the wild-type allele, and no band with primers for the mutant allele. PCR of homozygous HP1γ knockout ($HP1\gamma^{-/-}$) mice gives a single band of 525 bp with primers for the mutant allele, and no band with primers for the wild-type allele. Heterozygous HP1γ ($HP1\gamma^{+/-}$) mice, when genotyped, will produce both 501 bp and 525 bp bands. Supplementary Figure S1 shows examples of HP1γ PCR.

In order to determine the sex of the embryos, PCR for the lysine demethylase 5D (Kdm5d) gene residing on the Y chromosome was performed. Male embryos produce a single band of 597 bp, while female embryos show no band due to the lack of the Y chromosome. Examples of PCR results can be found in Supplementary Figure S1.

2.2 Weighing of mouse embryos, generation of mouse embryonic fibroblasts (MEFs) and cell culturing

HP1γ$^{+/+}$ and HP1γ$^{-/-}$ embryos were generated by crossing male and female HP1γ$^{+/-}$ mice. Embryonic day 13.5 (E13.5) embryos were dissected out of the uterus and the removal of the yolk sac and placenta followed. The embryos were weighed with a precision balance.

For the generation of mouse embryonic fibroblasts (MEFs), the embryo head and internal body organs were removed and the remaining body was finely minced in 1.5 ml of ice-cold 0.25 % (w/v) trypsin- ethylenediaminetetraacetic acid (EDTA) (Sigma). The samples were incubated overnight at 4 °C, where trypsin is allowed to diffuse in the tissue while being practically inactive. The following day, samples were incubated at 37 °C for 30 min and culture medium Dulbecco's modified Eagle's medium (DMEM) (Thermo Fisher Scientific) supplemented with 15 % volume per volume (v/v) fetal bovine serum (FBS) (Thermo Fisher Scientific), 1 % (v/v) penicillin/streptomycin (P/S) (Gibco) was added to the samples to deactivate trypsin. Vigorous resuspension allowed for isolation of the cells, which were subsequently cultured at 37 °C and 5 % (v/v) CO_2.

Human cell lines including the ALT bone osteosarcoma U-2 OS cells (ATCC) were maintained in DMEM, supplemented with 10 % (v/v) FBS. Cells were incubated at 37 °C and 5 % (v/v) CO_2 and when at a confluency of around 80 %, they were split using 0.25 % (v/v) Trypsin-EDTA solution (Sigma-Aldrich) with a 5 min incubation at 37 °C. Cells were usually re-plated at a 60 % confluency.

2.3 Generation of immortalised 13.5 MEFs

For the generation of immortalized cells, 2 x 10^6 primary E13.5 MEFs from each genotype (HP1γ$^{+/+}$ male, HP1γ$^{+/+}$ female, HP1γ$^{-/-}$ male and HP1γ$^{-/-}$ female) were transfected with 20 µg of plasmid DNA, carrying the simian virus 40 (SV40) large T antigen gene, using Lipofectamine 2000 (Thermo Fisher Scientific), according to manufacturer's instructions. The SV40-transformed cells grow extremely quickly, rapidly overtaking the entire cell population; thus, the non-transformed cells were competed out by repeated splitting for 3 weeks.

2.4 Freezing and thawing cells

1-4 x 10^6 cells were harvested by dissociation with trypsin and pelleting with a 300 relative centrifugal force (rcf) spin down for 5 min at room temperature. The cell pellet was resuspended in 1 ml of freezing solution (10 % v/v Dimethyl sulphoxide (DMSO) (Sigma), 90 % v/v FBS) which was then transferred to a 2 ml Cryovial tube. The cells were frozen down in a Nalgene "Mr. Frosty" container at -80 °C. Cryovials were transferred to liquid nitrogen (LN_2) to following day.

To restart the cells, the samples were taken out from LN_2 and placed immediately at 37 °C for 2 min and then transferred to pre-warmed culture medium. A centrifugation at 300 rcf for 5 min followed and the supernatant was discarded. The cells were resuspended in fresh culture medium and plated in a 10 cm dish. Their culture continued as previously described.

2.5 Senescence-associated β-galactosidase assay (SA-β-gal)

Cells were washed twice with 1x phosphate-buffered saline (PBS) before being fixed with 0.5 % (v/v) Glutaraldehyde (Sigma-Aldrich, #G7776) in 1x PBS for 15 min, followed by 2 washes with 1x PBS/1mM $MgCl_2$. X-gal solution (1 mg/ml X-gal, 5 mM potassium ferrocyanide, 5 mM potassium ferricyanide, 1 mM $MgCl_2$ in 1x PBS) was used to stain the cells by incubating at 37 °C for 16 hours (h). After staining, the cells were washed with 1x PBS and the proportion of cells with SA-β-gal activity was quantified using the Olympus CKX41 microscope.

2.6 Bromodeoxyuridine (BrdU) staining

12 x 10^4 MEFs were cultured for 24 h with 100 µl Dulbecco's modified Eagle's medium (DMEM) supplemented with 15 % (v/v) FBS, 1 % (v/v) P/S (Gibco) and the following day, 1x BrdU (#11647229001, Merck) was added to the medium. An 18 h incubation of the plates at 37 °C followed, cells were washed once with 1x PBS and fixed with 4 % (v/v) formaldehyde for 45 min. Cells were washed three times with 1x PBs and permeabilization followed with 0.2 % (v/v) Triton X-100 in 1x PBS. Samples were then incubated for 1 h with blocking solution (1 % w/v BSA and 0.2 % w/v fish gelatin in 1X PBS) and primary antibody (anti-BrdU, BD Biosciences, #555627, 1:2,000 diluted in blocking solution with the addition of 0.5u/ul DNAse (Roche) and 1 mM $MgCl_2$) incubation followed for 30 min at 37 °C. Three washes with 1x PBS took place and the samples were incubated with secondary antibody (anti-mouse Alexa594, A-11032, 1:2,000 in blocking solution) for 1 h at 37 °C. Samples were washed three times with 1x PBS and mounted with ProLong Gold antifade reagent containing DAPI (Invitrogen) and images were captured with InCell Analyzer 2000 (GE healthcare) and quantified with the InCell Investigator software (v1.5).

2.7 Quantitative fluorescence *in situ* hybridization (Q-FISH)

For metaphase spread preparation, MEFs were incubated for 4 h with 10 ng/ml colcemid (Roche, 10295892001). Cells were then collected and incubated for 15 min at 37 °C, in hypotonic buffer (KCl 75 mM). Fixation was performed with ethanol: glacial acetic acid (3:1, v/v), followed by three washes with the same fixative. The metaphase suspensions were subsequently dropped on glass slides. Slides were left to dry overnight at room temperature.

Q-FISH was performed as previously described (Ourliac-Garnier and Londoño-Vallejo, 2011). Briefly, metaphase spreads were fixed with 4 % (v/v) formaldehyde for 2 min, washed three times in 1x PBS for 5 min each, and treated with pepsin (1 mg/ml in 0.05 M citric acid pH=2) for 10 min at 37 °C. They were post-fixed for 2 min with 4 % (v/v) formaldehyde, washed three times with 1x PBS and incubated in increasing ethanol concentration baths (70-100 % v/v). Each slide was then covered with hybridising solution containing 10 nM of Cy3-O-O-$(CCCTAA)_3$ probe (PNA bio) in 70 % (v/v) formamide, 10 mM Tris pH=7.4 and 1 % blocking reagent (Roche, 11096176001). This step was followed by denaturation for 3 min at 80 °C. Hybridisation was performed for 2 h at room temperature and the slides were washed twice for 15 min in 70 % (v/v)formamide, 20 mM Tris pH=7.4. Three 5 min washes in 50 mM Tris pH 7.4, 150 mM NaCl, 0.05 % Tween-20 followed and the slides were dehydrated in successive ethanol baths (70-100 % v/v) and air-dried. Slides were mounted with ProLong Gold antifade reagent containing DAPI (Invitrogen) and images were captured with Zeiss microscope using Carl Zeiss software. Telomeric signal was quantified using the ImageJ FIJI software.

2.8 Immunofluorescence (IF)

MEFs seeded on culture slides, were treated with permeabilisation solution (50 mM NaCl, 3 mM MgCl2, 20 mM Tris pH=8, 0.5 % (v/v) Triton X-100, 300 mM sucrose) for cytoplasm removal and fixed for 15 min in fixative solution (4 % (v/v) formaldehyde, 2 % (w/v) sucrose) and washed three times with 1x PBS. Another 10 min permeabilization followed, slides were washed once with 1x PBS and were incubated

for 30 min with blocking solution (10 % (v/v) goat or donkey serum (Stratech Scientific Ltd) in 1x PBS) at 37 °C. Then, the primary antibody (anti-γH2AX, Millipore, 05-636, 1:500; or anti-53BP1, Thermo Fisher Scientific, 1:400) was added in blocking solution and incubated for 1 hour at 37 °C. Three 5 min washes with 1x PBS, followed and the slides were incubated with secondary antibody (1:400 in blocking buffer, donkey a-rabbit Alexa 488 antibody, Invitrogen A-21206; or goat a-mouse Alexa488, Invitrogen A-11001) for 30 min at 37 °C. ProLong Gold antifade reagent containing DAPI (Invitrogen) was used to stain DNA and the images were captured with Zeiss microscope using Carl Zeiss software. Signal was quantified with the ImageJ FIJI software.

For the p16 IF experiments, since they were performed in collaboration with Gil's lab [Veerinder Reen] the same conditions as for BrdU staining were employed, with the only difference that primary antibodies (mouse anti-p16, SC-56330, 1:200) were diluted in blocking solution without the addition of 0.5 U/μl DNase (Roche) and 1 mM $MgCl_2$ and primary antibody incubation time was 1 h instead of 30 min at 37 °C (see BrdU staining section).

2.9 Immunofluorescence - fluorescence *in situ* hybridization (IF-FISH)

IF took place as previously described. Instead of staining the DNA with DAPI immediately after the secondary antibody incubation step, slides were post-fixed for 10 min using 4 % (v/v) formaldehyde, 2 % (w/v) sucrose and washed three times with 1x PBS for 3 min. The follow-up FISH protocol was performed similar to metaphases, using the Cy3-O-O-$(CCCTAA)_3$ probe. Finally, samples were mounted with ProLong Gold antifade reagent containing DAPI and samples were visualized with Zeiss microscope using Carl Zeiss software and quantification was performed using the CellProfiler 3.1.9 software

2.10 Telomere restriction fragment (TRF) analysis

Cells were lysed using the Hirt buffer (10 mM Tris pH=7.6, 100 mM NaCl, 10 mM EDTA) with the addition of 0.5 % (w/v) sodium dodecyl sulfate (SDS) and 200 µg/ml of RNase A (Roche, 10109169001). The samples were heated at 37 °C for 60 min, followed by the addition of 150 µg/ml Proteinase K (Sigma, #P4850) and further incubating the samples at 55 °C overnight. DNA extraction followed with the addition of 1 volume of UltraPure Phenol:Chloroform:Isoamyl Alcohol (25:24:1, v/v) (ThermoFisher), and incubating samples for 20 min on a rotating wheel and spinning down for 15 min at 16,000 rcf. DNA was precipitated with 1/10 volumes of 3 M NaAc pH=5.2 and 2.5 volumes of 100 % EtOH and centrifuging the samples at 21,000 rcf at 4 °C. Samples were washed with 70 % EtOH and samples were resuspended in water overnight. 5 µg of DNA was digested with HinfI and RsaI restriction enzymes (New England BioLabs, 50 U each) at 37 °C overnight, followed by an extra 2h digestion with 10U of each enzyme the next day. Digested DNA was purified with phenol-chloroform and 5µg were loaded on a 0.8 % SeaKem gold agarose gel (Lonza, 50150) and run at 40 V for 20 h. The gel was stained with 1µg/ml ethidium bromide (Sigma, E1510) to establish the size of the samples based on the 1 kb Plus DNA Ladder (Thermo Fisher Scientific). DNA was then depurinated for 15 min with 250 mM HCl, denatured for 30 min in 500 mM NaOH,1.5 M NaCl, neutralized for 30 min in 1.5 M NaCl, 500 mM Tris pH=7.5 and transferred overnight on an Amersham Hybond N+ membrane with 20x saline-sodium citrate (SSC). DNA was UV-crosslinked (Stratalinker, 7000 kJ) and the membrane was neutralized with 2x SSC. The membrane was hybridized with a $TAA(CCCTAA)_4$ probe which was conjugated with digoxigenin (DIG oligonucleotide 3'-end labelling kit, Roche). The probe was diluted 1:2,500 in the Hybridization buffer (6x SSC, 0.1 % w/v SDS, 1 % w/v milk) and the hybridization took place at 42 °C overnight. The next day, the membrane was washed twice for 5 min with 2x SSC, 0.1 % (w/v) SDS at 37 °C, 2 min with 0.2x SSC, 0.1 % (w/v) SDS at 37 °C and for 5 min with 2x SSC at room temperature. The membrane was blocked with the Blocking solution (100 mM maleic acid, 150 mM NaCl, 1 % w/v milk, pH=7.5) for 30 min and anti-DIG antibody incubation took place (1:20000, Roche, 11093274910) for another 30 min. The membrane was then washed twice for 15 min with the Washing buffer (100 mM maleic acid, 150 mM NaCl, 3 % Tween-20, pH=7.5) and the detection followed by using the Detection

Solution (100 mM Tris, 100 mM NaCl, pH=9.5) for 5 min and applying 0.5 ml of CPD-Star solution (Roche) on the membrane. Images were captured using the Amersham Imager 680.

2.11 Northern blot

RNA extraction was performed using the RNeasy Mini Kit (Qiagen, 74104) and DNA was eliminated by on-column treatment with DNase I (Qiagen, 79254). 10 µg of total RNA was denatured for 10 min at 65 °C in 1x MOPS (0.2M MOPS, 50 mM NaOAc, 10 mM EDTA, RNase-free water) with 50 % v/v formamide, and 2.2 M formaldehyde, followed by a 5 min incubation on ice. 10X dye buffer (50 % v/v Glycerol, 0.3 % w/v Bromophenol Blue, 4 mg/ml Ethidium Bromide) was added to the samples, which were then run on a formaldehyde agarose gel (0.8% w/v agarose, 1x MOPS, 6.5 % v/v formaldehyde) at 5 V/cm in 1x MOPS buffer for 3.5 h. The gel was subsequently washed twice with RNase-free water and three times with 20X SSC, before transferring the RNA to an Amersham Hybond N+ membrane (Cytiva) using a neutral transfer in 20X SSC overnight. The membrane was UV-crosslinked (Stratalinker, 2000 kJ) and baked for 45 min at 80 °C, followed by pre-hybridisation with UltraHyb-Oligo solution (ThermoFisher #AM8663) at 42 °C for 1 h. Hybridization followed with 10 µM of ^{32}-P labelled TAA$(CCCTAA)_4$ probe in UltraHyb-Oligo solution (ThermoFisher #AM8663) at 42 °C overnight. The next day, the membrane was washed with 2x SSC, 0.1 % (w/v) SDS for 10 min at room temperature, 5 min with 0.2x SSC, 0.1 % (w/v) SDS at 42 °C and for 5 min with 2x SSC, 0.1 % (w/v) SDS at room temperature. Signal was detected with phosphor-imager (Amersham Biosciences).

2.12 RNA dot blot

RNA extraction was carried out similar to Northern blotting. 2 µg of the extracted RNA were treated with RNase A (500 µg/ml) for 3 h at 37 °C to serve as negative control. RNase A treated and RNase A non-treated samples were denatured in 10 mM NaOH and 1mM EDTA by heating at 65°C for 10 min, incubated 5 min on ice and spotted on an Amersham Hybond N+ membrane using the Bio-Dot apparatus (Bio-Rad), following manufacturer's instructions. Membrane was rinsed once in 2x SSC, RNA was UV-crosslinked (Stratalinker, 2000 kJ) and the membrane was baked for 45 min at 80°C. Hybridization of the membrane followed similarly to the TRF assay using the DIG-labeled telomeric C-rich probe. The membrane was imaged using the Amersham Imager 680 and analysed using the Image Studio Lite software.
The membrane was then stripped by using two 15 min washes of 0.5 % (w/v) SDS at 60 °C, followed by two 15 min washes with 0.2 M NaOH, 0.1 % (w/v) SDS at 37 °C and neutralized with two 5min washes with 2X SSC. The re-probing took place with an 18s rRNA targeting probe (5'-CCATCCAATCGGTAGTAGCG-3', Sigma) which was used for signal normalization.

2.13 C-circles Assay (CCA)

CCA was adapted from Henson et al., 2017. 5 µg of DNA was isolated from cells, digested, extracted and precipitated using the same conditions as the TRF assay. 7.5ng of the digested genomic DNA were used to amplify the C-circle sequence with a Φ29 DNA polymerase that is exploited for rolling circle amplification (12h incubation at 30 °C. 4mM dATP, dGTP and dTTP were used for the amplification, while dCTP was omitted to ensure the specificity for C-circles which are comprised entirely of telomeric DNA). After sample amplification, the DNA polymerase was heat inactivated and the samples were transferred with 2x SSC on an Amersham Hybond N+ membrane using a dot blot apparatus (Bio-Rad). DNA was then UV-crosslinked (Stratalinker, 2000 kJ) and the membrane was probed similar with the TRF assay using the DIG-labeled telomeric C-rich oligonucleotide. The membrane was imaged using the Amersham Imager 680 and analysed using the Image Studio Lite software.

2.14 TERRA reverse transcription quantitative real-time PCR (RT-qPCR)

RT-qPFeretzaki & Lingner, 2017. rom Feretzaki & Lingner, 2017. RNA was extracted similar to Northern blotting with two additional DNase I digestions. 1U DNase I (Roche) per µg of RNA was used, followed by an on-column treatment with DNase I (Qiagen) according to the manufacturer's protocols. 3 µg of total RNA was reverse transcribed using 200 U of SuperScript III Reverse Transcriptase (Invitrogen) with either random hexamers (Invitrogen, N8080127) or TERRA specific oligonucleotides $(CCCTAA)_4$ (de Silanes *et al.*, 2014). qPCR reactions were performed using Power SYBR Green PCR Master Mix (Applied Biosystems, 4368708) and the Bio-Rad CFX96 system with the following parameters: 40 cycles of 15s of denaturation at 95 °C followed by 1 min of annealing and extension at 60 °C. RT-qPCR primers are listed in Table 2.4 (de Silanes *et al.*, 2014). The $2_t^{-\Delta\Delta Ct}$ method was employed for relative TERRA quantification, using the β-actin housekeeping gene for normalisation.

Table 2.4 RT-qPCR primers used in this study

Primer Name	Primer Sequence 5' to 3'
beta-actin Fw	CTGTCCCTGTATGCCTCTG
beta-actin Rev	ATGTCACGCACGATTTCC
L-chr18_1	GGGGGTTAGGGGTAAGGTTT
R-chr18_1	AGGAATCACTGCTGGCATTT
L-ChrX-1	GAGGTTCCTGTAAGTCTCCA
R-ChrX-1	CCTATGATGATGTGCATGTG
L-Chr9-2	TGCCTCTCAAGTGCTGTT
R-Chr9-2	GTAGGCATTGTGTCAGTCTCA
L-qPCR-Chr10	TCAGCAAATCATGGTTCAGAT
R-qPCR-Chr10	TGCATTGCATTTGACAACAG

2.15 DNA:RNA Immunoprecipitation (DRIP)

DRIP was performed as previously described (Herrera-Moyano *et al.*, 2014; Sanz and Chédin, 2019), with the following modifications. 4 x 10^6 cells were resuspended in 800 µl of 1x TE Buffer (10 mM Tris, pH=8, 1 mM EDTA) and lysis took place with the addition of 25 µl 20 % (w/v) SDS and 2.5 µl of 24 mg/ml Proteinase K (P4850) and incubation of the samples at 37 °C overnight. Genomic DNA was extracted from MEFs with 2ml phase-lock tubes (5Prime, #2302830), using 1 volume Phenol:Chloroform:Isoamyl Alcohol (25:24:1, v/v) and spinning down at 16,000 rcf for 10 min. For DNA precipitation, the addition of 1/10 volumes 3M NaAC pH=5.2 and 2.5 volumes 100 % EtOH followed, and then DNA was washed twice with 80 % EtOH. DNA was eluted in 1x TE buffer by incubation for 2 h on ice and was then digested at 37 °C overnight with

the following restriction enzymes: SspI, EcoRI, HindIII, XhoI and BsrGI (New England BioLabs, 44 U each). The next day, samples were further digested for 2 h at 37 °C and digested nucleic acids were precipitated with 1 volume Phenol:Chloroform:Isoamyl Alcohol (25:24:1, v/v) and by being centrifuged at 16,000 rcf for 35 min at 4 °C. Samples were then EtOH-precipitated with 1.5 µl of glycogen (Thermo Fisher Scientific, R0561) and resuspended in 50 µl of 1x TE buffer. 6 µg of digested nucleic acids were incubated overnight at 37 °C with either the addition of 10 µl of RNase H (New England BioLabs, M029L) or without the addition of RNase H, in 1x RNase H buffer and 1/10^{th} of the samples was saved as input. RNase H treated and non-treated samples were incubated with 10 µl of the S9.6 antibody (Kerafast ENH001) for 16 h at 4 °C in 1x DRIP binding buffer (10 mM sodium phosphate, pH=7, 140 mM NaCl and 0.05 % v/v Triton X-100). Samples were then incubated with 50 µl of a 2:1 (A:G) mixture of Dynabeads Protein A and G (Invitrogen, 10001D and 10004D) for 2 h at 4 °C that were washed twice with 1x DRIP binding buffer and the precipitated samples were eluted in 300 µl of elution buffer (10 mM EDTA pH= 8, 50 mM Tris pH=8, 0.5 % SDS) and treated with 7 µl of proteinase K (24 mg/mL) for 45 min at 55 °C. A treatment with 50 µg/ml RNase A (Roche, 10109169001) for 1 h at 37 °C plus 1 h at 65 °C followed and cleaned samples were resuspended in 100 µl of 1x TE buffer. DNA:RNA hybrids were detected using at the indicated subtelomeric regions with the corresponding

qPCR primers listed in Table 2.4. qPCR was performed with the same parameters employed for the TERRA RT-qPCR and percentage of signal to input was calculated.

2.16 Protein extraction and western blotting (WB)

Cell pellets were incubated for 10 min on ice in lysis buffer (25 mM Tris pH=8, 40 mM NaCl, 2 mM $MgCl_2$, 0.05 % w/v SDS) supplemented with 100 units/ml Benzonase (Sigma-Aldrich, E1014-25KU) and cOmplete, EDTA-free Protease Inhibitor Cocktail (Roche). The samples were lysed by being forced ten times through a 25 G needle and incubated on ice for a further 10 min. Proteins were quantitated using the Pierce BCA Protein Assay kit (Thermo Fisher Scientific, #23225) and 5 to 30 µg from each sample were denatured for 5 min at 100 °C with the addition of 4x Laemmli buffer (50 mM Tris pH=6.8, 2 % w/v SDS, 100 mM DTT, 10 % v/v glycerol, 0.1 % w/v bromophenol blue). Proteins were resolved in 4–12 % Bis-Tris gels (Invitrogen) at 100 V for 2.5 h and then transferred to a nitrocellulose membrane (Amersham Protran 0.2 µm NC), employing a wet transfer apparatus (Bio-Rad) at 90 V for 2h at 4 °C. After blocking with 5 % (w/v) non-fat milk in 1x PBS, 0.1 % (v/v) Tween (PBST) for 1 h at room temperature the membranes were probed overnight with primary antibodies at 4 °C. The membranes were washed three times with 1x PBST for 5 min and incubated with secondary antibodies for 1 h at room temperature, followed by 1x PBST washes. Finally, the signal was visualized using ECL Western blotting reagent (Sigma-Aldrich, RPN2106) with the Amersham Imager 680 (GE Healthcare). Quantification of the signal was performed with the ImageStudio Lite software.

For reblotting, the membranes were washed five times for 3 min with TBST. Afterwards, membranes were incubated with Restore™ Western Blot Stripping Buffer (Life technologies, 21059) for 10 minutes in a shaker at room temperature. Three 5 min washes with TBST followed and a second WB commenced by blocking with 5 % (w/v) milk.

A list of Antibodies used in WB can be found in Tables 2.5 and 2.6

Table 2.5 Primary antibodies used for WB

Protein	Source	Reference	Dillution
HP1γ	Mouse	Thermo Fisher Scientific #MA3-054	1:2,500
TRF1	Rabbit	Gift from Titia de Lange [Rockefeller University]	1:1,000
α-Tubulin	Mouse	Sigma-Aldrich #T6199	1:5,000
Lamin B1	Rabbit	Abcam #16048	1:1,000
H3	Rabbit	Abcam #1791	1:2,500

Table 2.6 Secondary antibodies used for WB

Specificity	Conjugation	Reference	Dilluition
Mouse	Horseradish peroxidase	Agilent Dako #P0447	1:10,000
Rabbit	Horseradish peroxidase	Agilent Dako #P0217	1:10,000

2.17 Fractionation assay

Cell fractionation was performed similar to Gillotin, 2018. Briefly, MEFs were collected by pelleting at 130 rcf for 3 min at 4 °C and were lysed by pipetting gently 5 times with 5 volumes of E1 buffer 50 mM HEPES-KOH pH=7.5, 140 mM NaCl, 1 mM EDTA pH=8.0, 10 % v/v glycerol, 0.5 % v/v NP-40, 0.25 % v/v Triton X-100, 1 mM Dithiothreitol (DTT), supplemented with cOmplete, EDTA-free Protease Inhibitor Cocktail) and spinning down at 1,100 rcf at 4 °C for 2 min. The supernatant was collected (cytoplasmic fraction) and the pellet was resuspended in E1 buffer, similar to the previous step and

a centrifugation at 1,100 rcf at 4 °C for 2 min, followed. The pellet was resuspended in E1 buffer and incubated on ice for 10 min, followed by another centrifugation at 1,100 rcf at 4 °C for 2 min. The pellet was resuspended in 2 volumes E2 buffer (10 mM Tris-HCl pH=8.0, 200 mM NaCl, 1 mM EDTA pH=8.0, 0.5 mM EGTA pH=8, supplemented with cOmplete, EDTA-free Protease Inhibitor Cocktail) and centrifuged at 1,100 rcf at 4 °C for 2 min. The supernatant was collected (nuclear soluble fraction) and the pellet was resuspended with 2 volumes of E2 buffer, followed by a centrifugation at 1,100 rcf at 4 °C for 2 min. The pellet was washed once by using the same volume of E2 as the previous step and incubating on ice for 10 min. After a centrifugation at 1,100 rcf at 4 °C for 2 min the pellet was resuspended in 5 volumes of E3 buffer (500 mM Tris-HCl pH=6.8, 500 mM NaCl, supplemented with cOmplete, EDTA-free Protease Inhibitor Cocktail) and a sonication for 5 min (30 sec ON/30 sec OFF) on maximum power at 4 °C using the Bioruptor Plus (Diagenode) took place. The isolation of the chromatin bound fraction occurred by spinning down at 16,000 rcf at 4 °C for 10 min and collecting the supernatant.

Protein quantification of the different fractions was performed as previously described and the samples were analysed by WB and liquid chromatography–mass spectrometry (LC/MS) (see below). MaxQuant software package (Cox and Mann, 2008) was used to analyse the mass-spectrometric data and quantification was performed with the Perseus platform (Tyanova *et al.*, 2016).

2.18 Liquid chromatography–mass spectrometry (LC/MS)

Sample for LC/MS were processed similar to Hughes *et al.*, 2018. Briefly, 50 µl of each fraction was denatured, reduced and alkylated for 30 min at 60 °C, with 5µl 10 % SDS, 2.5 µl EPPS 1M pH=8.5, 1 µl TCEP 0.5M, 4 µl chloracetamide 0.25 M. The samples were precipitated on 100 µg of MagReSyn Hydroxyl beads (Resynbio) with 75 µl EtOH at room temperature for 5 min. The beads were washed 3 times with 100 µl of 80 % EtOH and the proteins were digested with 1 µg trypsin and 0.5 µg lysC overnight shanking at 1200 rpm. Supernatant was collected for injection and the beads were

washed with 10 µl of 1 % TFA and the wash was pulled with the supernatant for a final volume of 60 µl of digested, acidified sample.

LC-MS/MS analysis was performed using an Ultimate 3000 RSLC nano-flow liquid chromatography system (Thermo Scientific) coupled to a Q-Exactive HFX mass spectrometer (Thermo Scientific) via an EASY spray source (Thermo Scientific). Protein digests were injected and loaded onto a trap column (Acclaim PepMap 100 C18, 100 µm × 2 cm) for desalting and concentration at a flow rate of 8 µL/min (loading pump buffer: 2 % acetonitrile, 0.1 % TFA). Final on-column digest amount was 400 ng per injection. Peptides were eluted on-line to an analytical column (Acclaim Pepmap RSLC C18, 75 µm × 50 cm) at a flow rate of 250 nl/min. Peptides were separated using a 90-minute gradient, 1-30 % of buffer A (5% dimethylsulfoxide (DMSO), 0.1% FA) for 80 minutes followed by 30-40 % buffer B (75% acetonitrile, 20% water, 5 % DMSO, 0.1 % FA) for another 10 minutes and subsequent column conditioning and equilibration. Total method length was 125 minutes. Eluted peptides were analysed by the mass spectrometer operating in positive polarity using a data-dependent acquisition mode. Ions for fragmentation were determined from an initial MS1 survey scan at 120,000 resolution, followed by HCD (Higher Energy Collision Induced Dissociation) fragmentation of the top 20 most abundant ions at 15,000 MS2 resolution. MS1 and MS2 scan AGC targets were set to 3e6 and 5e4, allowing maximum ion injection times of 25 ms and 110 ms, respectively. A survey scan m/z range of 350-1750 was used, normalised collision energy set to 27 %, charge state exclusion enabled with unassigned and +1 charge states rejected and a minimal AGC target of 8e3. Dynamic exclusion was set to 50 s.

Data were processed using the MaxQuant software platform (v1.6.10.43). A reverse decoy search approach was used at a 1% false discovery rate (FDR) for both peptide spectrum matches and protein groups. Search parameters included: maximum missed cleavages set to 3, fixed modification of cysteine carbamidomethylation and variable modifications of methionine oxidation, protein N-terminal acetylation, asparagine deamidation and glutamine conversion to pyro-glutamate. Label-Free Quantification (LFQ) was enabled with a LFQ minimum ratio count of 1. ‘Match

between runs' function was enabled, with alignment and matching time windows of 20 and 0.7 minutes respectively.

2.19 Cleavage Under Targets and Release Using Nuclease (CUT & RUN)

The CUT & RUN assay was performed as described in Skene et al., 2018) with minor adaptations. 1 x 10^6 MEFs were employed per condition, 5 µg of Abs were used to target the proteins of interest and optimal digitonin (Enzo Life Sciences, UK) concentration for membrane permeabilization was determined at 0.02 %. Cells were resuspended in 1 ml of Wash Buffer (20 mM HEPES pH=7.5, 150 mM NaCl, 0.5 mM spermidine, supplemented with 1 tablet of cOmplete, EDTA-free Protease Inhibitor Cocktail per 50 ml of Wash Buffer) and pelleted at 600 rcf for 3 min. This step was repeated two more times and cells were incubated on a rotating wheel for 10 min with 10 µl of Concanavalin A (ConA) beads that have been previously washed twice with Binding Buffer (20 mM HEPES pH=7.9, 10 mM KCl, 1 mM $CaCl_2$, 1 mM $MnCl_2$). Bound cells were resuspended in 150 µl of Antibody Buffer (Wash Buffer, 2mM EDTA, 0.02 % digitonin), containing the antibody of interest (mouse anti-HP1γ and mouse anti-IgG included in the same set, Millipore #17-646; or mouse anti-H3K27me3, Abcam, #6002) to a 1:100 dilution and the samples were incubated at 4 °C overnight on a rotating wheel. The following day, cells were washed twice with 1 ml of Digitonin Wash Buffer (Wash Buffer containing 0.02 % digitonin) and resuspended in 150 µl of the same buffer. 700 ng/ml of Protein A conjugated with MNase (pA-MNase) (gift from Steven Henikoff [Fred Hutch Cancer Center] was added to the reaction and the samples were incubated at 4 °C for 1 h. Samples were washed twice with Digitonin Wash Buffer before being resuspended in 100 µl of Digitonin Wash Buffer. Samples were placed at 0 °C for 2 min and 2 µl of 100 mM $CaCl_2$ were added to the reaction for the activation of the pA-MNase. After a 30 min incubation at 0 °C, 100 µl of 2x STOP Buffer (340 mM NaCl, 20 mM EDTA, 4 mM EGTA, 0.025 % digitonin, 50 µg/ml RNase A, 10 µg/ml GlycoBlue (Thermo Fisher Scientific), 100 pg yeast spike-in DNA) were added to the samples. Samples were incubated at 37 °C for 30 min to release the digested

chromatin fragments from the insoluble nuclear chromatin and standard phenol-chloroform DNA extraction followed. Quality of the samples was tested with the Agilent 2100 Bioanalyzer using High Sensitivity DNA Analysis Kits (#5067-4626, Agilent Technologies) and DNA libraries were prepared with the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina (New England BioLabs) following the manufacturer's instructions. Paired-end next generation sequencing (NGS) using the Illumina HiSeq 2500 platform followed and the analysis of the sequencing data was performed in conjunction with the Genomics Laboratory of the MRC LMS as described in (Meers, Tenenbaum and Henikoff, 2019).

Sequenced CUT & RUN reads were aligned to the mouse genome version mm9 using Bowtie2. For mapping the spike-in fragments, the --no-overlap --no-dovetail options were selected to avoid cross-mapping of the experimental genome to that of the yeast DNA. Quality controls and statistics were generated using the ChIPQC package (ChIPQC:1.18.2). Graphs were generated using the ggplot2 visualisation R package and sequencing tracks were visualised with the IGV genome browser.

2.20 Super low input carrier- Cap analysis of gene expression (SLIC-CAGE)

The SLIC-CAGE assay was carried out according to Cvetesic et al., 2018. RNA extraction from MEFs was performed similar to Northern blotting. RNA quality was tested with the Agilent 2100 Bioanalyzer using the Agilent RNA 6000 Nano Kit (#5067-1511, Agilent Technologies). 1 µg of high-quality extracted RNA was mixed with 4 µg of carrier RNA (generated as described in Cvetesic et al., 2018) for a total of 5 µg of RNA material and reversed transcribed using Random Primer 6 (New England BioLabs) and SuperScript III Reverse Transcriptase (Invitrogen) following the manufacturer's suggestions. Samples were cleaned with the RNAClean XP (#A63987, Beckman Coulter) following the manufacturer's instructions, oxidised using 1.1 mM $NaIO_4$ by incubating on ice for 45 min and the reaction was stopped by addition of 260 mM Tris-HCl pH=8.5. Samples were cleaned with the RNAClean XP and biotinylation

followed with the addition of 0.8 mM biotin solution (Biotium). Biotinylated samples were treated with 5 U of RNase I (Promega) to select only full-length RNA:cDNA hybrids and the samples were incubated at 37 °C for 30 min. Purification of the samples was performed with the Dynabeads M-270 Streptavidin (Thermo Fisher Scientific) according to manufacturer's suggestions. Release and isolation of cDNA from the cap trapped mRNA followed, by incubating the samples that were treated with 6 U of RNase H (Takara, #2150A) and 10 U of RNase I at 37 °C for 15 min. Single-stranded cDNA was cleaned with RNAClean XP, ligated with 10 μM of 5' adapters for 16 h at 16 °C followed by ligation with 3' adapters (10 μM) by incubating samples for 16 h at 16 °C. Samples were dephosphorylated with 1 U of Shrimp Alkaline Phosphatase (#783901000UN, Thermo Fisher Scientific) and second strand synthesis took place with the Deep Vent (exo-) DNA Polymerase (New England Biolabs) following the manufacturer's protocol. Incubating samples with 20 U Exonuclease I (#M0293L, New England Biolabs) at 37 °C for 30 min led to the degradation of the 2nd strand synthesis primer and samples were subsequently purified with the RNAClean XP kit. Removal of the carrier RNA was performed with two rounds of degradation using the endonucleases I-SceI and I-CeuI (5 U each, New England Biolabs) by incubating samples at 37°C for 3 h and removal of degraded carrier RNA occurred with the use of AMPure XP beads (Beckman Coulter, #A6388) according to manufacturer's protocol. An extra round of AMPure XP beads purification enabled size selection of the SLIC-CAGE fragments that were subsequently quantified with the quant-iT, PicoGreen dsDNA Assay (Thermo Fisher Scientific, #P7589) following manufacturer's instructions, PCR-amplified to increase the amount of material and quality tested with the Agilent 2100 Bioanalyzer using High Sensitivity DNA Analysis Kits (Agilent Technologies, #5067-4626). Samples were sequenced with Illumina HiSeq 2500 and the follow-up analysis was performed in conjunction with Boris Lenhard group [Nevena Cvetesic and Radina Georgieva]. Sequenced reads were aligned to the mm10 version of the mouse genome using Bowtie2 and samples were analysed with the CAGEr (v1.20.0) (Haberle *et al.*, 2015), the DESeq2 Bioconductor package (Love, Huber and Anders, 2014) and the GREAT tool (v4.0.4) (McLean *et al.*, 2010). Sequencing tracks were visualised with the IGV genome browser.

2.21 Differential gene expression analysis and statistical analysis

Previously generated RNA-sequencing data (Law *et al.*, 2019), was examined with the DESeq2 Bioconductor package (Love, Huber and Anders, 2014). Raw p-values were adjusted (padj) for multiple testing using the Benjamini-Hochberg procedure.

Statistical analysis for non-NGS experimental procedures was performed using GraphPad Prism. Error bars, number of replicates (n) and statistical methods are indicated in figure legends and/or the corresponding chapters. Biological replicates for all experiments were based on embryos from independent litters.

Chapter 3. The role of HP1γ in sexual dimorphism in mouse

3.1 Introduction

3.1.1 Epigenetic regulation of sexual dimorphism early in development by HP1γ

Sex dimorphism during early development has been observed in many animal species including mice (Thornhill and Burgoyne, 1993; Lowe *et al.*, 2015) and humans (Pergament *et al.*, 1994; Ray *et al.*, 1995; Ménézo *et al.*, 1999). Growth rate of mouse embryos for instance appears to have a sexually dimorphic character with males developing faster in the early days of embryonic growth compared to their sister littermates (Burgoyne, 1993; Thornhill and Burgoyne, 1993; Burgoyne *et al.*, 1995). Using the "Four Core Genotypes" (FCG) mouse model, a number of studies have shown the importance of sex chromosome complement in defining sex discrepancies (Wijchers *et al.*, 2010; Kuljis *et al.*, 2013). By deleting the testis-determining *Sry* gene from the Y chromosome (Y^- chromosome) and inserting an *Sry* transgene onto an autosome, the FCG model produces XY mice with and without the *Sry* transgene (XY^-*Sry*, XY^-) and XX mice with and without the *Sry* transgene (XX*Sry*, XX). This allows researchers to study the effect of an XY and XX genotype in mice with the same gonadal type (Lovell-Badge and Robertson, 1990).

A study showed that the expression of 369 mouse autosomal genes is sensitive to sex chromosome complement, rather than sex. Differences in heterochromatin formation between the two sexes, appears to be responsible for this differential expression (Wijchers *et al.*, 2010). Interestingly, the sex chromosome complement sensitive genes were enriched for genes that are repressed by HP1β, linking the family of epigenetic modifiers HP1 with the regulation of sexually dimorphic genes. HP1 proteins are generally regarded as silencing factors due to their interaction with the H3K9me modification that predominantly marks constitutive heterochromatin

(Lachner *et al.*, 2001; Peters *et al.*, 2003). However, unlike the other two HP1 isoforms α and β, HP1γ is found in both euchromatic and heterochromatic regions in interphase nuclei (Minc *et al.*, 1999; Nielsen *et al.*, 1999; Minc, Courvalin and Buendia, 2000). Because of its unique distribution, this protein is considered to regulate gene expression in a context-specific manner, with both an activating and suppressive role. Indeed, HP1γ appears to regulate genes that are expressed differently among the sexes, early in mouse development (Figure 3.1). Specifically, HP1γ is required for keeping the expression of 62 autosomal genes higher in males than females (termed “male higher” genes) and the expression of 114 genes lower in males than females (termed “male lower” genes). Loss of HP1γ led to a downregulation of the “male higher” genes suggesting that these genes require HP1γ to be expressed at a higher level than in females. On the other hand, abolishing HP1γ resulted in upregulation of the “male lower” genes, showing that HP1γ represses this group of genes in wild type conditions. Sex-biased differences of several biological processes arise due to the differential effect of HP1γ with males being more affected by the loss of HP1γ than females (Law *et al.*, 2019).

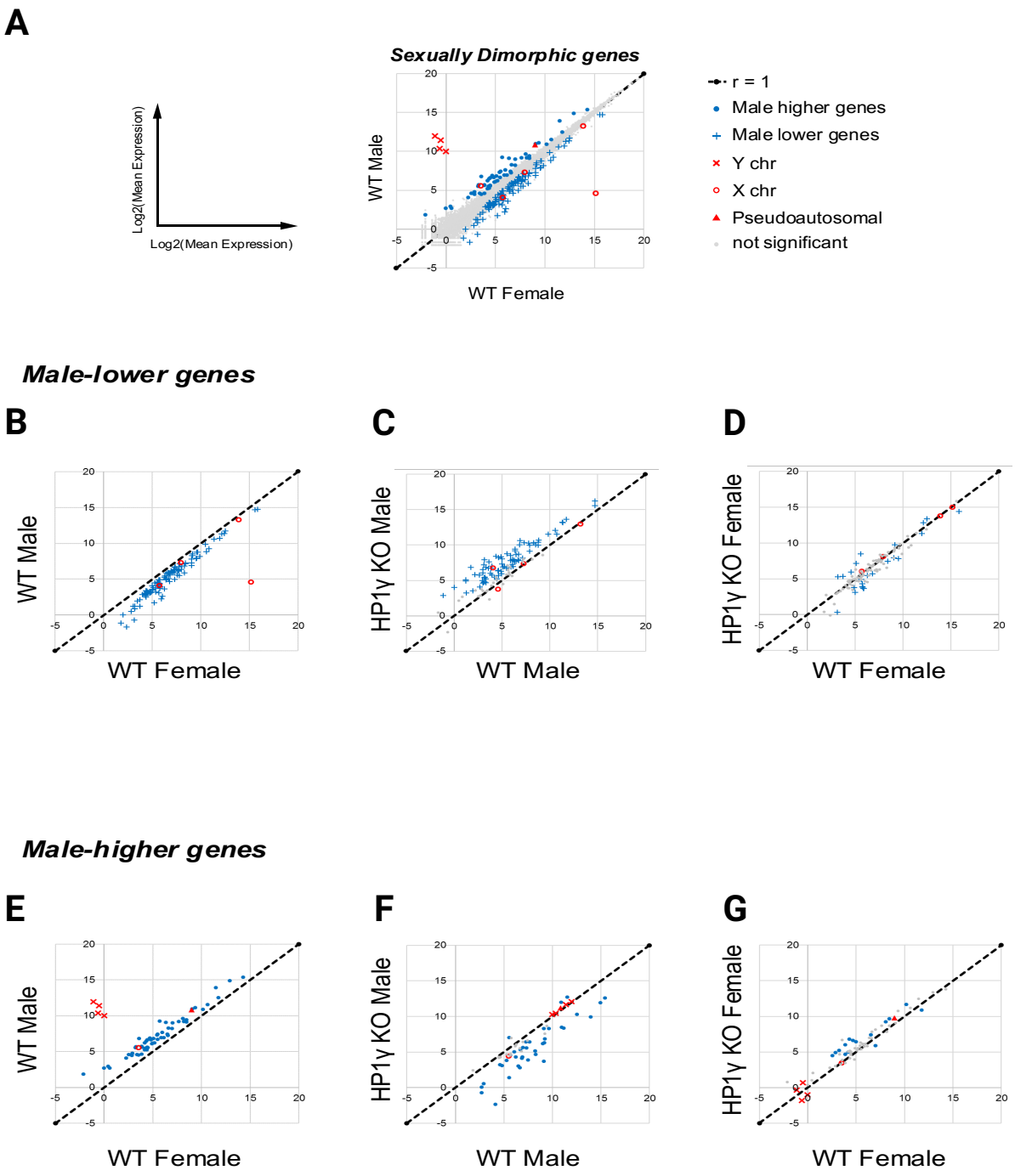


Figure 3.1 Sexual dimorphic gene expression in HP1γ wild-type (WT) MEFs and their sexual dimorphic response to HP1γ knockout (KO). (A) Density plot of differentially expressed genes in WT HP1γ male and female MEFs. 62 autosomal genes present higher expression in males (termed "male higher" genes) and 114 autosomal genes present lower expression in males (termed "male lower" genes). (B) Expression level of "male lower" genes in HP1γ WT males compared to females. (C) Expression level of "male lower" genes in HP1γ WT compared to HP1γ KO males. (D) Expression level of "male lower" genes in HP1γ WT compared to HP1γ KO females. (E) Expression level of "male-higher" genes in HP1γ WT male compared to females. (F) Expression level of "male-higher" genes in HP1γ WT compared to HP1γ KO males. (G) Expression level of "male-higher" genes in HP1γ WT compared to HP1γ KO females MEFs. This figure is adapted with permission from Law *et al.*, 2019

3.1.2. Rationale & aim

Because of the importance of sex differences in physiology and disease, one main aim of this work has been to further elucidate the molecular mechanisms underlying HP1γ-regulated sexually dimorphic genes and what is the effect of this sexual dimorphism at the cellular and organismal level. We are specifically interested in the role of HP1γ early in development, therefore we are employing mouse embryonic fibroblasts (MEFs) derived from embryonic day 13.5 (E13.5) embryos as our model system.

3.2 Results

3.2.1 MEFs as an experimental model for studying HP1γ's role in regulation of sexually dimorphic genes

As mentioned in Section 1.1.11, the importance of HP1γ in physiology is highlighted by the fact that its homozygous deletion leads to neonatal lethality in mouse bred on a pure genetic background (Naruse *et al.*, 2007). Even mice bred on a mixed genetic background (C57BL/6 and 129/Ola) show a minimal survival rate of 1% upon loss of HP1γ (Brown *et al.*, 2010; Takada *et al.*, 2011).

Generation of an HP1γ knockout mouse by Naruse and colleagues allows the investigation of HP1γ functions *in vivo*. The insertion of a retroviral construct (ROSAN β-geo) into intron 1, located 998 bp downstream of exon 1 of the *Cbx3* gene acts as a gene-trap for HP1γ. A fusion transcript containing β-geo and *HP1γ's* first exon is generated, with transcription being prematurely terminated at the inserted polyadenylation site and HP1γ protein not being synthesised (Naruse *et al.*, 2007). Mice with one mutated *Cbx3* allele ($HP1\gamma^{+/-}$) present no apparent phenotype and develop normally into adulthood. By breeding $HP1\gamma^{+/-}$ mice, I was able to isolate mouse embryonic fibroblasts (MEFs) deriving from embryonic day 13.5 (E13.5) embryos, where at this developmental stage, sex hormone levels are minimal (Gondos, 1980). From this cross, there is 25% chance of getting wild type ($HP1\gamma^{+/+}$) embryos and 25% chance of getting homozygous knockout ($HP1\gamma^{-/-}$) embryos (Figure 3.2).

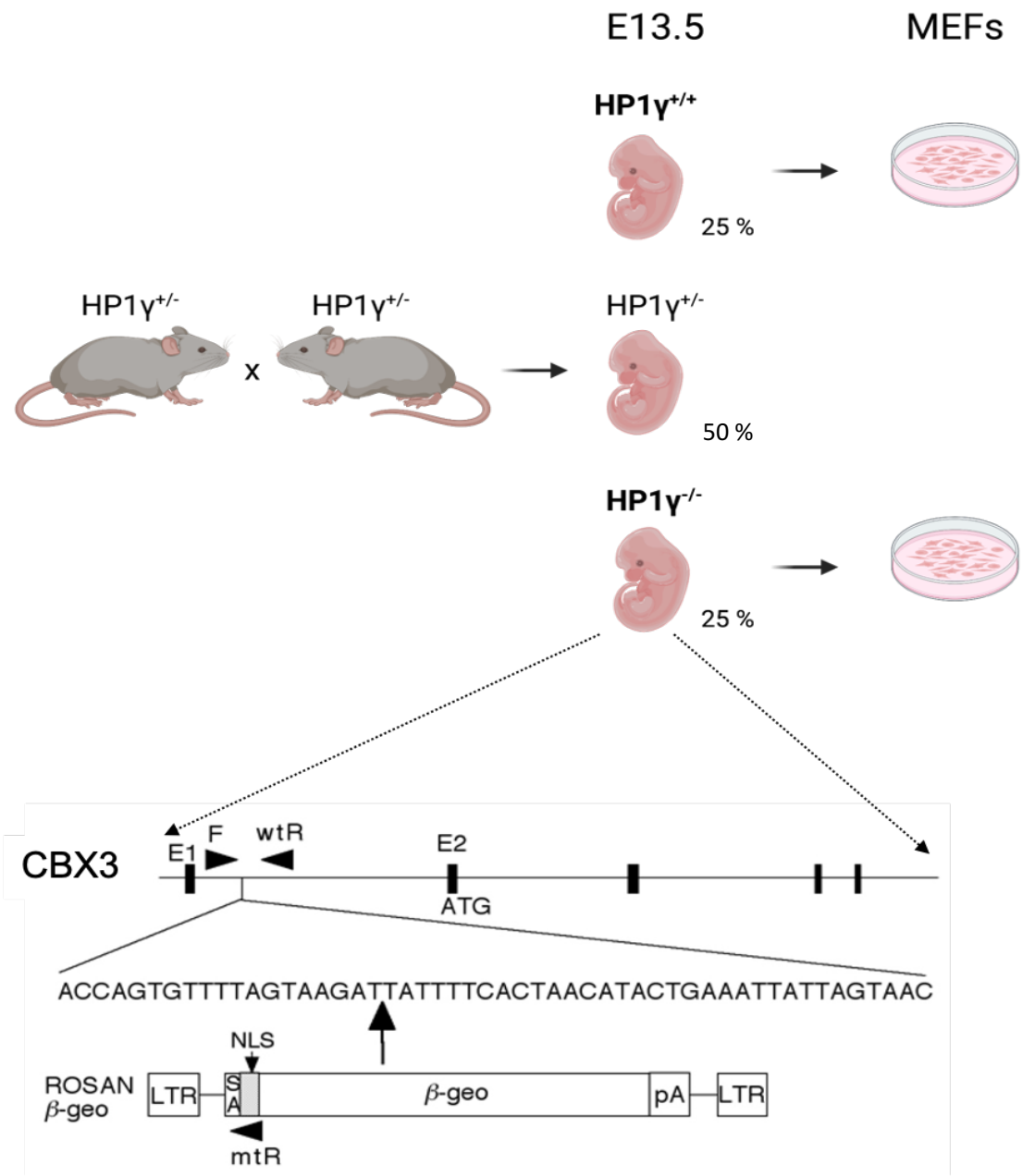


Figure 3.2 Schematic diagram demonstrating the breeding of HP1γ mice and the insertion site of the gene trap in the *Cbx3* gene. Breeding of heterozygous HP1γ knockout (HP1γ$^{+/-}$) mice results in a 25% chance of HP1γ$^{+/+)}$ embryos, 25% chance of homozygous HP1γ knockout (HP1γ$^{-/-}$) embryos and 75% chance of HP1γ$^{+/-}$ embryos. Mouse embryonic fibroblasts (MEFs) are generated from embryonic day 13.5 (E13.5) embryos. The insertion site of the gene trap vector in the CBX3 gene is also shown for the HP1γ$^{-/-}$ mice. A retroviral vector (ROSAN β-geo) is inserted into intron 1 which leads to the generation of non-functional transcript. Genotyping PCR primers are indicated by black arrows. HP1γ common forward (F), HP1γ wild-type reverse (wtR), HP1γ mutant reverse (mtR). Schematic was adapted from Takada *et al.*, 2011.

The MEFs generated from these embryos are employed in every experiment performed for this thesis. HP1γ RNA and protein levels are similar for both sexes in HP1γ$^{+/+}$ MEFs, while in HP1γ$^{-/-}$ MEFs, HP1γ is undetectable (Figure 3.3), making this a good model for studying loss of function effects of HP1γ.

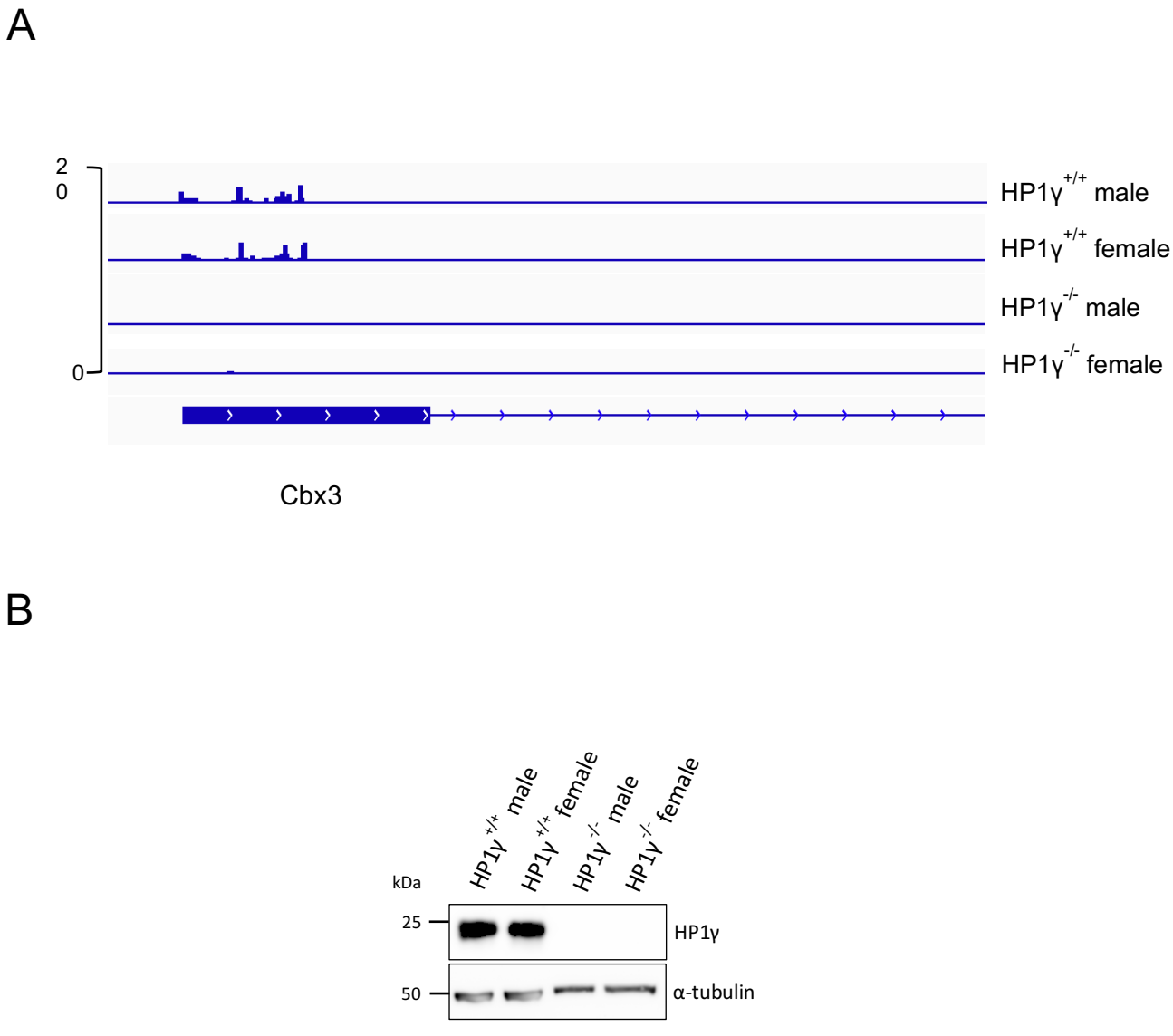


Figure 3.3 Undetectable RNA and protein levels upon HP1γ knockout. (A) SLIC-CAGE tracks of the corresponding genotypes over the CBX3 gene (start codon in blue and arrows indicate direction of transcription). The height of each track is proportional to the expression level (log scale) and is the overlay of two biological replicates. No expression of the CBX3 gene is observed in HP1γ$^{-/-}$ samples. Visualisation with the IGV genome browser. (B) Western blot of HP1γ (30 μg of protein) where HP1γ$^{-/-}$ MEFs show undetectable protein levels. α-tubulin is used as a loading control.

3.2.2 Loss of HP1γ results in slower replication and upregulation of cell cycle regulator *Cdkn2a* in males

The sex-biased differences of physiological functions that are dependent upon HP1γ give rise to discrepancies that are evident at the cellular level. It has been shown that male MEFs have a proliferation advantage compared to female MEFs and removal of HP1γ led to a reduction of the proliferation rate of male MEFs to comparable levels of females, while the proliferation rate of female MEFs remained largely unaffected by perturbation of HP1γ (Law *et al.*, 2019).

To further characterise the sex-biased dependency of cellular proliferation on HP1γ at the molecular level we performed experiments with the synthetic thymidine analogue bromodeoxyuridine (BrdU). BrdU staining allows the detection of DNA replication, by being incorporated in the newly synthesized DNA molecules of dividing cells (Kee *et al.*, 2002). For this experiment, cells from Passage 3 (P3) and Passage 4 (P4) of sex-matched HP1γ$^{+/+}$ and HP1γ$^{-/-}$ primary MEFs were used. To ensure that the observable signal arises only due to the binding of the secondary antibody to the primary antibody, a secondary-antibody-only staining was included in the experimental setup. This control revealed that the antibody labelling is specific, as no signal was detected with the secondary antibody alone (Figure 3.4 A). HP1γ$^{+/+}$ males and females showed similar levels of BrdU incorporation with BrdU positive cells ranging from 32 % to 38 % for both passages. Interestingly, P3 male MEFs showed a clear proliferation halting upon HP1γ depletion (34 % male HP1γ$^{+/+}$ BrdU positive cells compared to 18 % male HP1γ$^{-/-}$ BrdU positive cells). The difference became significant in P4, with 38 % male HP1γ$^{+/+}$ BrdU positive cells, compared to 16 % male HP1γ$^{-/-}$ BrdU positive cells. Female MEFs do not appear to be affected by HP1γ loss, as in P3 32 % of female HP1γ$^{+/+}$ MEFs were BrdU positive compared to 30 % female HP1γ$^{-/-}$ MEFs, while in P4 both HP1γ wild-type and knockout conditions had 37 % of BrdU positive cells (Figure 3.4 B). This outcome, is in agreement with the observation that HP1γ displays male-biased transcriptional defects (Law *et al.*, 2019).

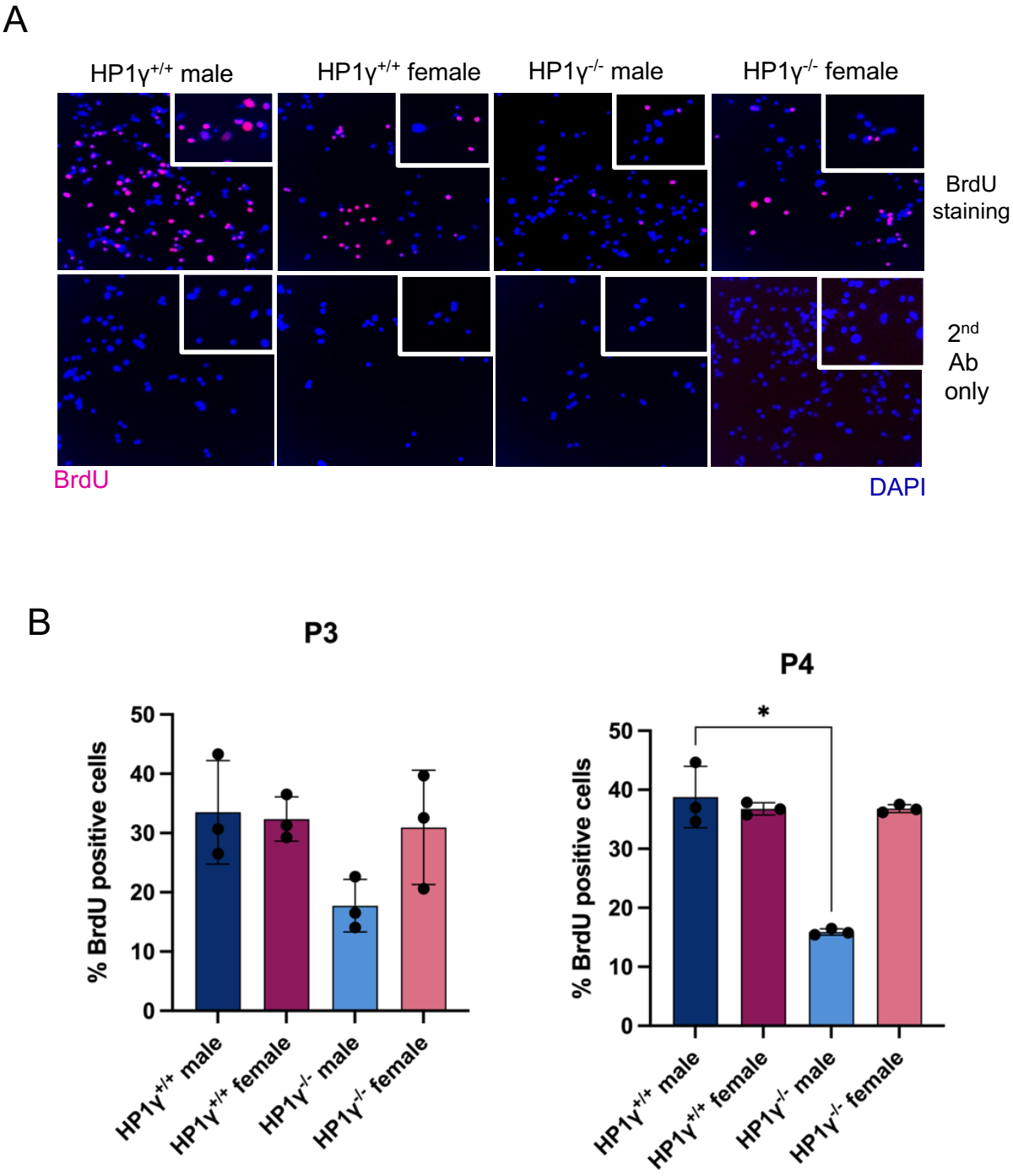


Figure 3.4 Loss of HP1γ results in slower replication in E13.5 male MEFs. (A) Representative images of BrdU immunofluorescence using passage 4 (P4) MEFs of the indicated genotypes. (B) Graphs represent the mean ± SD from three independent biological replicates (shown as dots) of passage 3 (P3) and P4 MEFs. Statistical significance was tested with one-way ANOVA and Tukey's *post-hoc* test, * $p<0.05$.

From the analysis of genes that are sensitive to HP1γ and differ between the sexes, *Cdkn2a* was shown to be expressed at lower levels in males compared to females (Law et al., 2019). *Cdkn2a* encodes the tumour suppressors p16 and p14ARF (p19ARF in mouse). p16 is an inhibitor of the cyclin dependent kinases 4 and 6 (CDK4 and CDK6), thereby blocking the G1-S phase transition (Tam et al., 1994), while p14ARF blocks cell cycle progression by activating p53 and halting cells in G1 phase (Eymin *et al.*, 2003). The differential expression of this gene between the sexes could be a possible explanation for the sexual dimorphism in male cell proliferation. *Post-hoc* analysis of the transcriptomic data (Law *et al.*, 2019), reveals that upon loss of HP1γ, the expression of *Cdkn2a* is upregulated in males by 2.8-fold (Figure 3.5) while it remains at the same level in females.

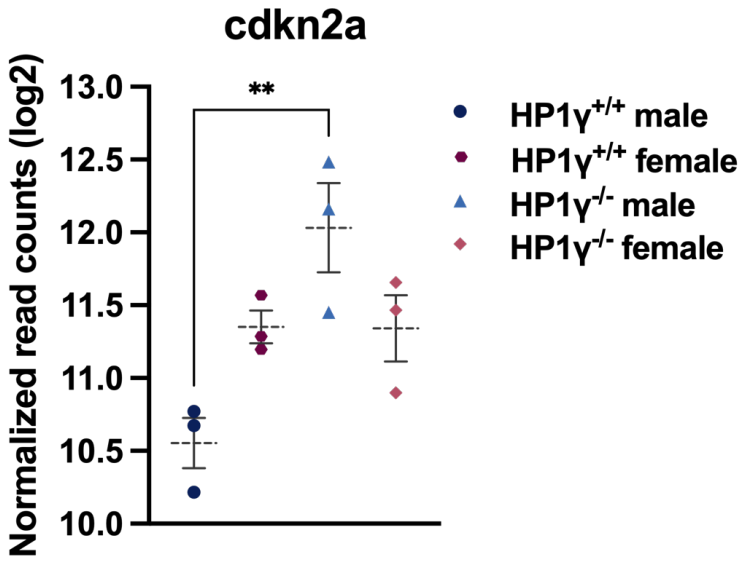
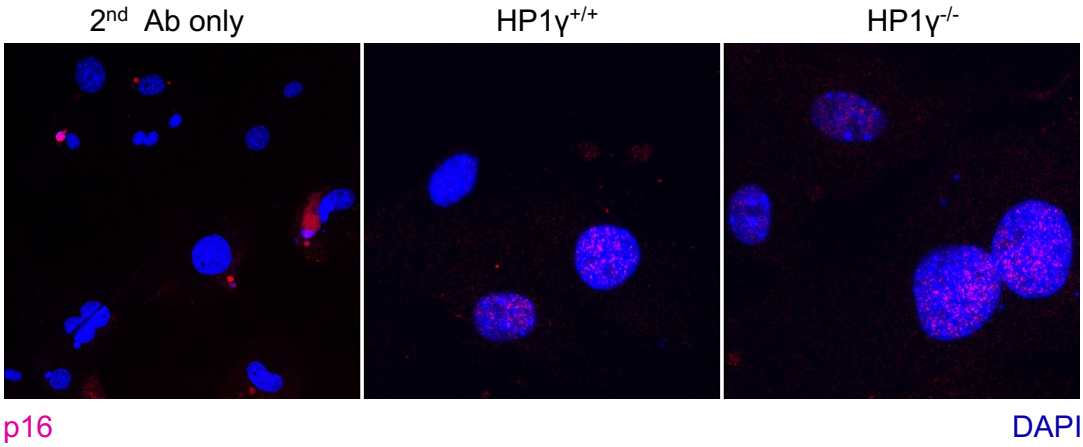


Figure 3.5 Depletion of HP1γ leads to upregulation of *Cdk2na* in E13.5 male MEFs. Interleaved scatter plot of *Cdkn2a* normalized read counts in respect to library size (Law et al 2019), where each data point represents one biological replicate. Dotted lines indicate mean, with error bars indicating ± SEM. Statistical significance was tested with one-way ANOVA and Tukey's *post-hoc* test, ** p<0.01.

This was also verified by RT-qPCR experiments performed by H. Zhi, another member of Festenstein's group (Supplementary Figure S2). To examine the protein levels of *Cdkn2a*, we performed p16 immunofluorescence experiments for both sexes using, P3 and P4 HP1γ$^{+/+}$ and HP1γ$^{-/-}$ primary MEFs. Similar to BrdU staining, a secondary-antibody-only staining was included that showed signal specificity (Figure 3.6 A). The analysis showed similar percentage of p16 positive wild-type males and females for both passages with a small increase of p16 positive cells in P4 (Figure 3.6 B). Consistent with the transcriptomic data previously acquired at P3 (Law *et al.*, 2019), there was a p16 signal increase upon deletion of HP1γ at P3 male MEFs (18 % p16 positive HP1γ$^{+/+}$ males and 31 % p16 positive HP1γ$^{-/-}$ males), but this trend was not recapitulated at P4 where the p16 protein levels are very similar between wild-type and HP1γ knockout conditions (22 % p16 positive HP1γ$^{+/+}$ males and 24 % p16 positive HP1γ$^{-/-}$ males). Comparison of HP1γ$^{+/+}$ and HP1γ$^{-/-}$ females showed that p16 is not affected in the absence of HP1γ for either passage.

A



B

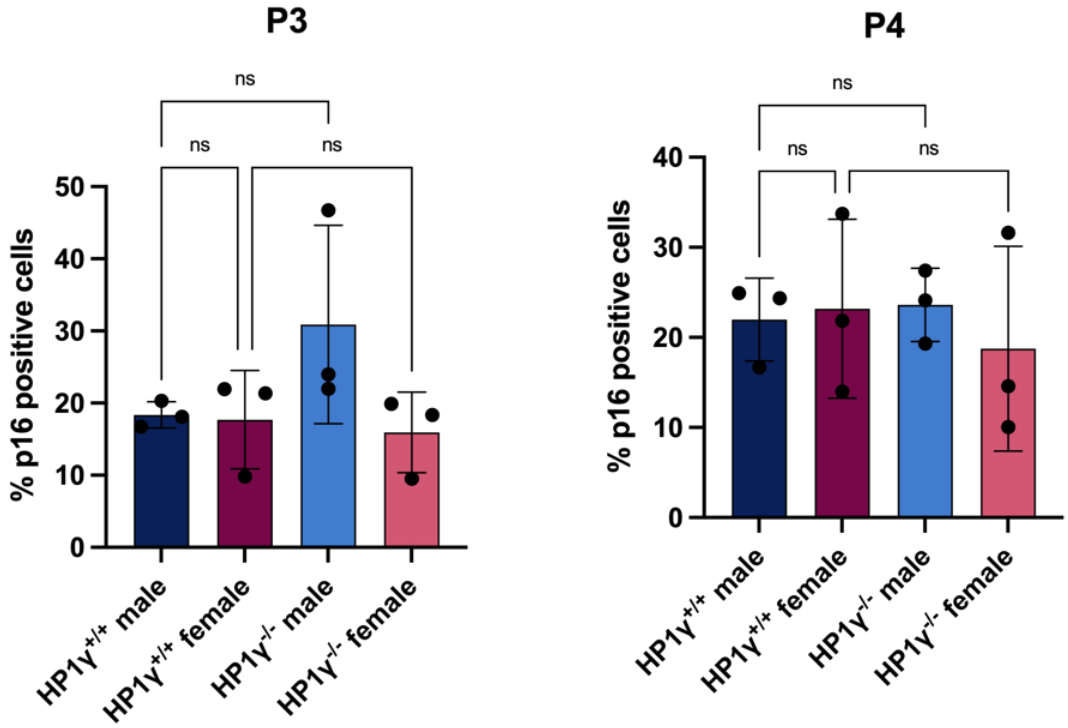


Figure 3.6 HP1γ loss does not affect p16 protein levels in E13.5 MEFs. (A) Representative images of p16 immunofluorescence using passage 4 (P4) MEFs of the indicated genotypes. (B) Graphs represent the mean ± SD from three independent biological replicates (shown as dots) of passage 3 (P3) and P4 MEFs. Statistical significance was tested with one-way ANOVA and Tukey's *post-hoc* test, ns= non-significant.

3.2.3 Earlier onset of senescence for both sexes upon HP1γ depletion

Cellular senescence is an important defence mechanism against the development of cancer (Lowe, Cepero and Evan, 2004). Cells under physiological conditions can proliferate only for a finite period of time, after which they reach senescence, a state where there is loss of proliferative capacity, despite continued metabolic activity and viability (Hayflick and Moorhead, 1961). MEFs are a classic system for studying cell senescence, with clear parallels to key genetic human elements (Hahn and Weinberg, 2002). When cultured *ex vivo,* MEFs initially replicate rapidly, but after five to six passages they undergo senescence attributed mostly to oxidative stress elicited by non-physiological oxygen levels of standard culturing conditions (Parrinello *et al.*, 2003). Apart from slower replication, the proliferation differences observed could also be arising due to earlier onset of senescence.

Proliferation halting and a change in size and morphology of E13.5 MEFs lacking HP1γ was noticeable as soon as P3, with cells becoming larger and more spindle-shaped, suggestive of a senescent state (Kuilman *et al.*, 2010). By testing for β-galactosidase (β-gal) activity, one of the most commonly used senescence biomarkers (Debacq-Chainiaux *et al.*, 2009), I addressed whether earlier onset of senescence arises due to loss of HP1γ. β-galactosidase is an enzyme that catalyses the hydrolysis of β-galactosides into monosaccharides and is accumulated in senescent cells due to overexpression of *GLB1*, the gene encoding lysosomal β-D-galactosidase (Lee *et al.*, 2006). By treating senescent cells with a chromogenic substrate like X-gal, the β-gal activity can be measured *in situ* (Dimri *et al.*, 1995), as seen in Figure 3.7 A. Cells from early passages including P3 and P4 and a later passage (P8), were included in my analysis. As shown in Figure 3.7 B, less than 10 % of the HP1γ$^{+/+}$ male and female MEFs are senescent at P3 and P4. Interestingly, at P3 more than 15 % of the HP1γ$^{-/-}$ female cells are senescent. At P4, the lack of HP1γ leads to senescent cells comprising 1/5th of the cell population for both sexes and by P8, a massive increase in senescence is observed, where almost 90 % of HP1γ$^{-/-}$ cells are senescent. As a negative control for the β-gal assay, P4 HP1γ$^{+/+}$ MEFs were treated with low concentration (0.4 %) DMSO, known to have negligible effects on senescence

(Kakolyri, Margaritou and Tiligada, 2016). The experiment showed comparable levels of wild-type and DMSO- treated senescent cells. 20 μM of etoposide (EP), a known inducer of senescence at low doses (Tamamori-Adachi *et al.*, 2018), was also used as a positive experimental control. There is a clear increase in the percentage of "blue" (indicative of β-gal activity) EP-treated MEFs (35 %), compared to wild-type MEFs (8 %), confirming that the staining is indeed specific to senescence. We can conclude that for both sexes, loss of HP1γ leads to earlier onset of senescence, with females being slightly more sensitive to this perturbation, accentuated at P3 MEFs.

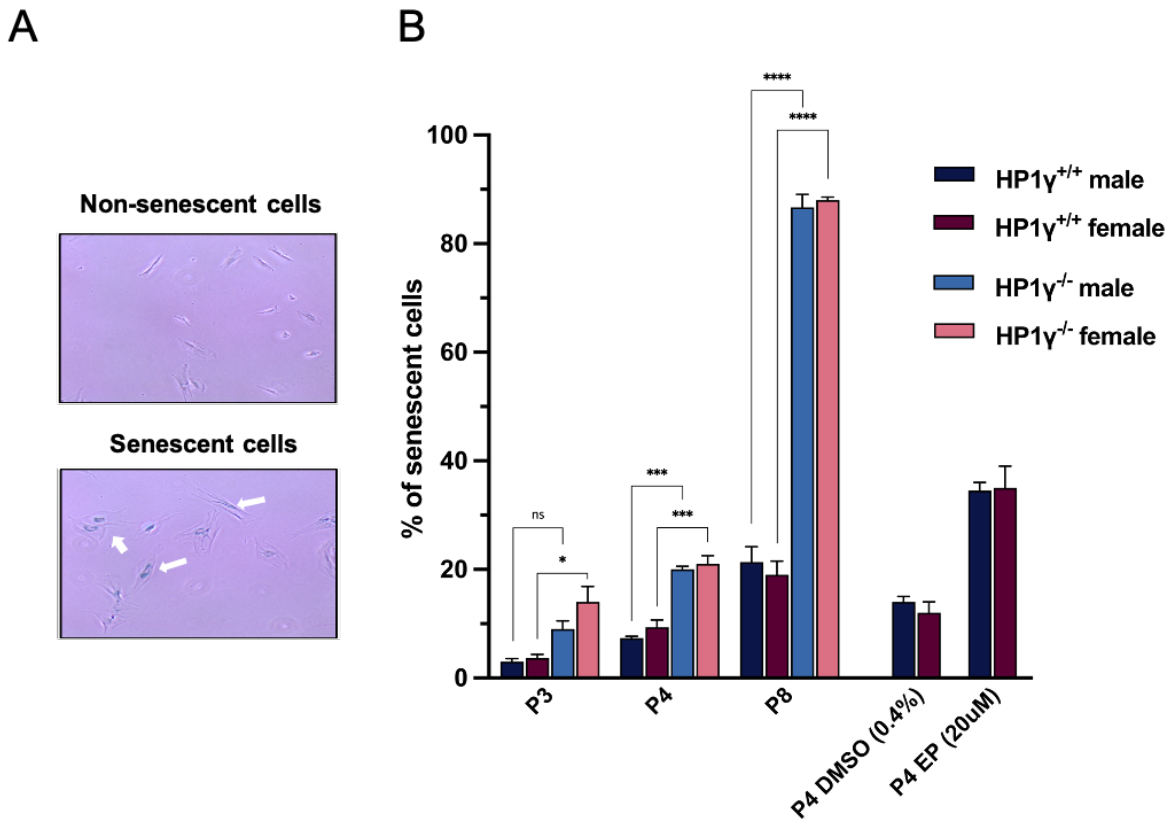


Figure 3.7 Earlier onset of senescence in E13.5 MEFs upon loss of HP1γ. (A) Representative images of SA-β-gal staining for senescent and non-senescent cells. (B) Bars represent the mean ± SEM for the percentage of senescent cells in Passage 3, Passage 4 and Passage 8 from three biological replicates (at least 100 cells analysed per replicate). Two biological replicates of Passage 4 cells were also treated with either DMSO (0.4%) or Etoposide (20uM) as negative and positive controls for senescence, respectively. Statistical significance was tested with one-way ANOVA and Tukey's *post-hoc* test, * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001, ns=non-significant.

3.2.4 Male embryo growth rate at E13.5 is dependent upon HP1γ

Gene ontology (GO) analysis of transcriptomic data from E13.5 HP1γ$^{-/-}$ MEFs reveals that cell cycle-related genes are downregulated in males but not in females (Law *et al.*, 2019), reflected at cellular level by the male-biased slowing of proliferation measured by BrdU staining (see Section 3.2.2). Furthermore, biological processes that are disrupted upon HP1γ depletion include "embryonic morphogenesis" "skeletal system development" and "tissue development", which suggests that the effects of HP1γ seen *ex vivo*, may be reflected at the organismal level, thus playing a role in developmental sex differences and corroborated by the finding that loss of HP1γ induces earlier onset of senescence (see Section 3.2.3)

Embryonic weight of E13.5 mouse embryos was measured. 27 HP1γ$^{+/+}$ male embryos, 17 HP1γ$^{+/+}$ female embryos, 14 HP1γ$^{-/-}$ male embryos and 8 HP1γ$^{-/-}$ female embryos were weighed. The average weight for both HP1γ$^{+/+}$ male and female embryos is 92 mg, while upon loss of HP1γ the average weight is reduced to 72 and 78 mg, respectively (Figure 3.8). The measurements confirm that the male growth is dependent on HP1γ, whereas HP1γ deficiency has a modest effect in female growth at this developmental stage.

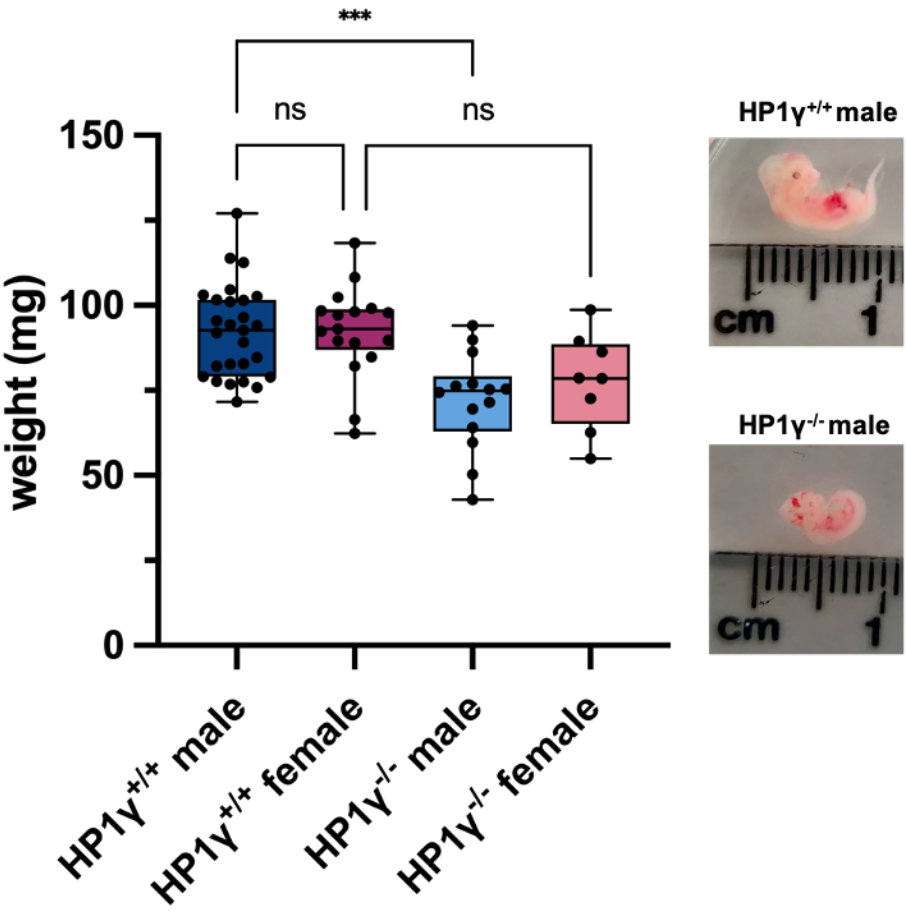


Figure 3.8 Embryo growth rate is dependent on HP1γ in males. Each dot represents an E13.5 embryo with box whiskers indicating the minimum and maximum values. Number of independent biological replicates: HP1γ$^{+/+}$ males=27, HP1γ$^{+/+}$ females =17, HP1γ$^{-/-}$ males =14, HP1γ$^{-/-}$ females =8. Statistical significance was tested with was tested with one-way ANOVA and Tukey's *post-hoc* test, *** p<0.001, ns=non-significant.

3.2.5 Examination of HP1γ's genome-wide binding pattern

As HP1γ was previously shown to be important for the expression of sexually dimorphic genes (See Section 3.1.1), we hypothesized that by determining the chromatin binding sites of HP1γ we would gain a better understanding as to whether the expression of these genes is directly or indirectly regulated by this protein. Binding of HP1γ on the promoters and/or the coding regions of sex dimorphic genes would suggest a direct HP1γ effect, while absence of signal on these regions should hint towards regulation by another sex-specific targeting factor responsible for their differential expression.

Cleavage Under Targets and Release Using Nuclease (CUT & RUN) followed by next generation sequencing was conducted. CUT & RUN is a method for studying genome-wide protein-DNA interactions similar to ChIP, which is the current standard. CUT & RUN offers several advantages over ChIP, including the absence of a fixation step that can potentially generate false-positive binding sites or epitope masking. At the same time, with CUT & RUN, only $1/10^{th}$ of the sequencing depth is required, significantly reducing the necessary starting material compared to ChIP (Skene and Henikoff, 2017). Figure 3.9 provides an overview of the CUT & RUN assay.

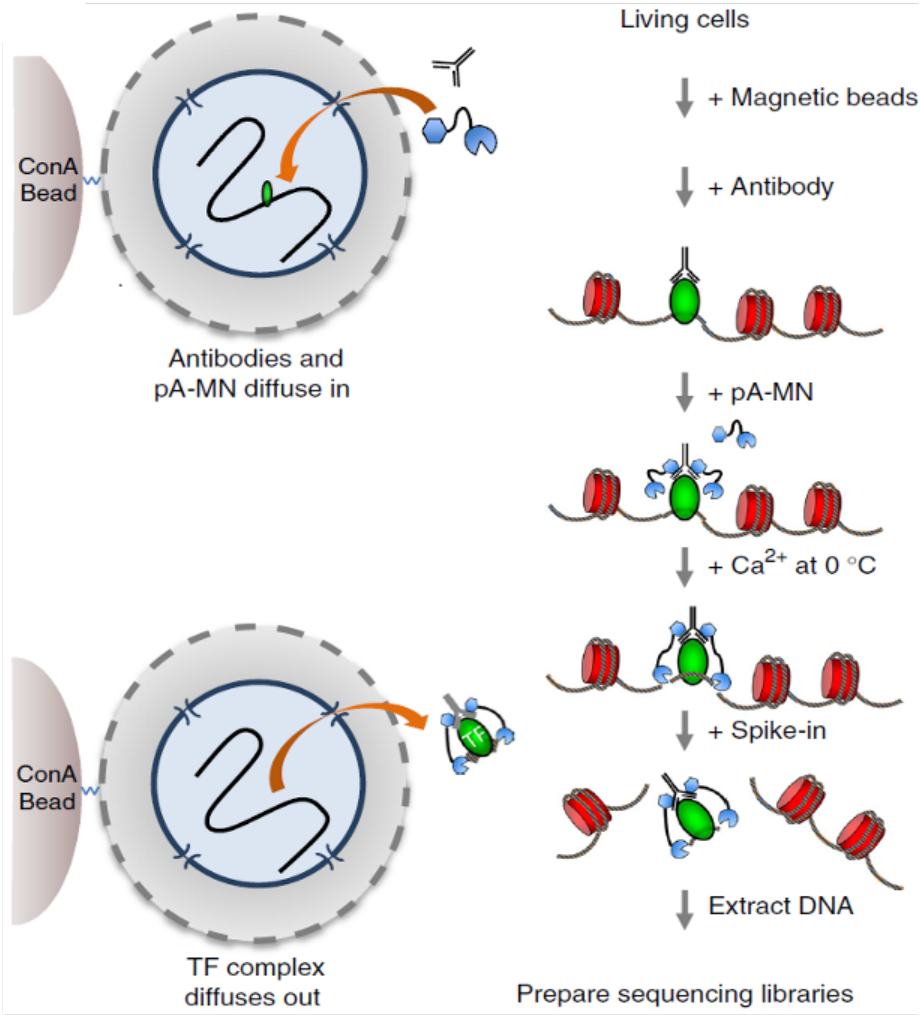


Figure 3.9 Schematic overview of cleavage under targets and release using nuclease (CUT&RUN). Briefly, cells are harvested and bound to concanavalin A–coated (conA) magnetic beads. Membranes permeabilisation follows with digitonin to allow the antibody to find its target. After incubation with the antibody, beads are briefly washed and then incubated with proteinase A conjugated with a micrococcal nuclease (pA-MN). Cells are cooled down to 0 °C, and chromatin digestion occurs with the addition of Ca^{2+}. Reactions are stopped by chelation, yeast spike-in DNA is added and the DNA fragments released into solution by cleavage are extracted from the supernatant. Adapated from Skene *et al.*, 2018.

To target HP1γ, the CUT & RUN conditions had to be optimized. To this end, immortalised MEFs (iMEFs) were employed as these are more easily to obtain and cultured and digitonin's permeability efficiency was initially tested. Cells were treated with 0.01%, 0.02%, 0.05% and 0.1% digitonin for 10 min, stained with Trypan Blue as a measure of permeability and then counted under a brightfield microscope. The experiment showed that 0.02% of digitonin was sufficient to permeabilise MEFs (data

not shown), so this concentration was used for the CUT & RUN experiment. According to Skene et al., 2018, CUT & RUN can provide high-quality binding profiles with as few as 1,000 cells when interrogating transcription factors and 100 cells for abundant histone modifications. We decided to perform the experiment with 1 x 10^6 cells as starting material, since other studies have shown that this number works well with factors of similar kinetic properties as HP1 (e.g CTCF) (Liu *et al.*, 2018; Hainer *et al.*, 2019; Oomen *et al.*, 2019). The choice of temperature for the permeabilisation and chromatin release steps of the experiment also needed to be considered, since native chromatin is analysed in CUT & RUN and transient interactions of mobile factors with a quick on-off binding rate on chromatin like HP1γ (Festenstein *et al.*, 2003) are majorly affected by temperature. I tested two different temperatures (4 °C and room temperature) and used Bioanalyzer signal as a proxy for a successful experiment (Figure 3.10 A). The expected HP1γ-bound DNA signal of approximately nucleosomal size (~150-200 bp), was stronger at room temperature conditions so the CUT & RUN reactions were performed at room temperature, similar to Skene et al., 2018.

As soon as the conditions were optimised, I proceeded to conduct the CUT & RUN experiment using primary MEFs from two male and two female HP1γ$^{+/+}$ embryos to target HP1γ. The abundant histone mark H3K27me3 was included as a positive control. We used one male (primary) and two female samples (one primary and one iMEF sample) targeting this histone mark. Two primary female samples were also used, where we targeted IgG. IgG randomly binds to chromatin at very low efficiency with no specific sequence, thus acting as negative control. Prior to PCR amplification, HP1γ CUT & RUN produced a small peak of approximately mono-nucleosomal size and signal coming from long undigested DNA species, while H3K27me3 CUT & RUN produced strong peaks of mono- di- tri-nucleosomal sizes, as well as signal from long undigested DNA fragments (Figure 3.10 B), similar to Skene et al., 2018. As expected, IgG CUT & RUN showed no observable peaks on the Bioanalyzer. DNA signal was then amplified, sequencing libraries were generated by adding 50 bp adaptors to the 3’ and 5’ ends (Figure 3.10 C) and paired-end sequencing followed.

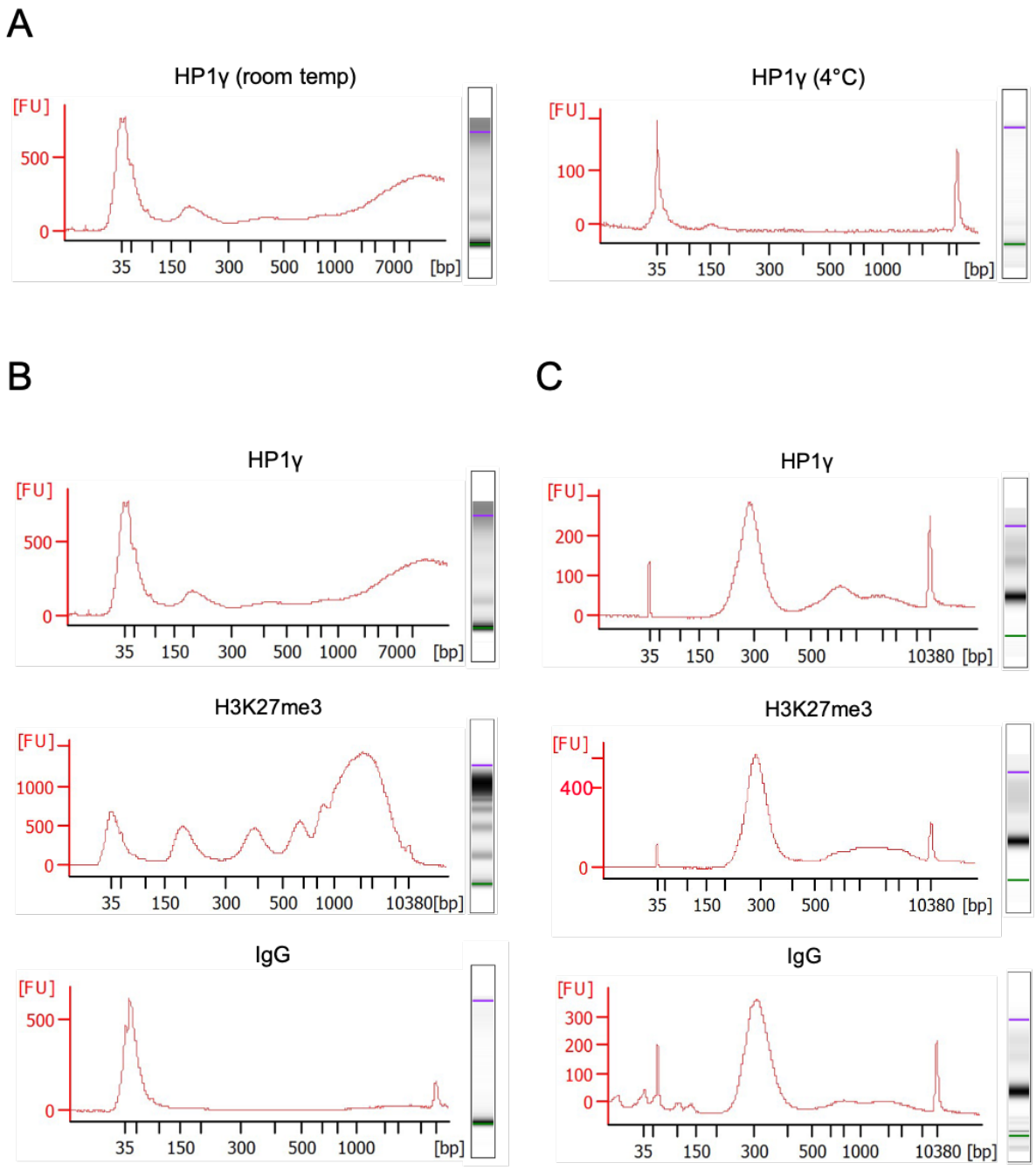


Figure 3.10 Bioanalyzer analysis as a positive indication for a successful CUT & RUN experiment. (A) HP1γ-bound DNA shows a peak of around 150-200 bp (mono-nucleosomal size). Stronger signal is observed at room temperature; hence follow-up experiments were performed accordingly. (B) Representative HP1γ, H3K27me3 and IgG CUT & RUN traces pre-amplification. HP1γ-bound DNA of mono-nucleosomal size, H3K27me3-bound DNA of mono- di- tri-nucleosomal sizes and IgG-bound DNA with no apparent peaks are observable. (C) Ligation of the adapters (50 bp) on 3' and 5' ends, PCR amplification of the signal and selection against small and large DNA fragments, results in DNA enrichment of 300 bp for all conditions. Peaks of 35 bp and 10,380 bp on the electropherograms, represent the lower and upper marker respectively.

As any next generation sequencing experiment, normalisation and quality control (QC) assessment of the data is necessary before proceeding to downstream analysis (Zhou *et al.*, 2018). In CUT & RUN, normalisation of the signal occurs with the use of spike-in yeast DNA instead of the use of chromatin input like in ChIP, allowing for direct comparison of samples. Following peak calling (genome-wide signal enrichment), profiles of HP1γ and H3K27me3 show a similar shape at the centre of their peak summits, indicative of signal enrichment (Figure 3.11 A). On the other hand, IgG peak signal remains flat. The distribution of reads across known genomic features such as genes and their subcomponents can allow further evaluation of CUT & RUN success. Figure 3.11 B shows that for both sexes there is a genome-wide enrichment of HP1γ, H3K27me3 and IgG signal over the 5' untranslated regions (5'UTRs) and the gene bodies, also known as protein coding sequences (CDSs). Unlike IgG, HP1γ is binding near the TSS (up to 500 bp upstream), similar to H3K27me3, where the signal is stronger in female. The latter appears to be also binding at regions as far as 20,000 bp upstream the TSS.

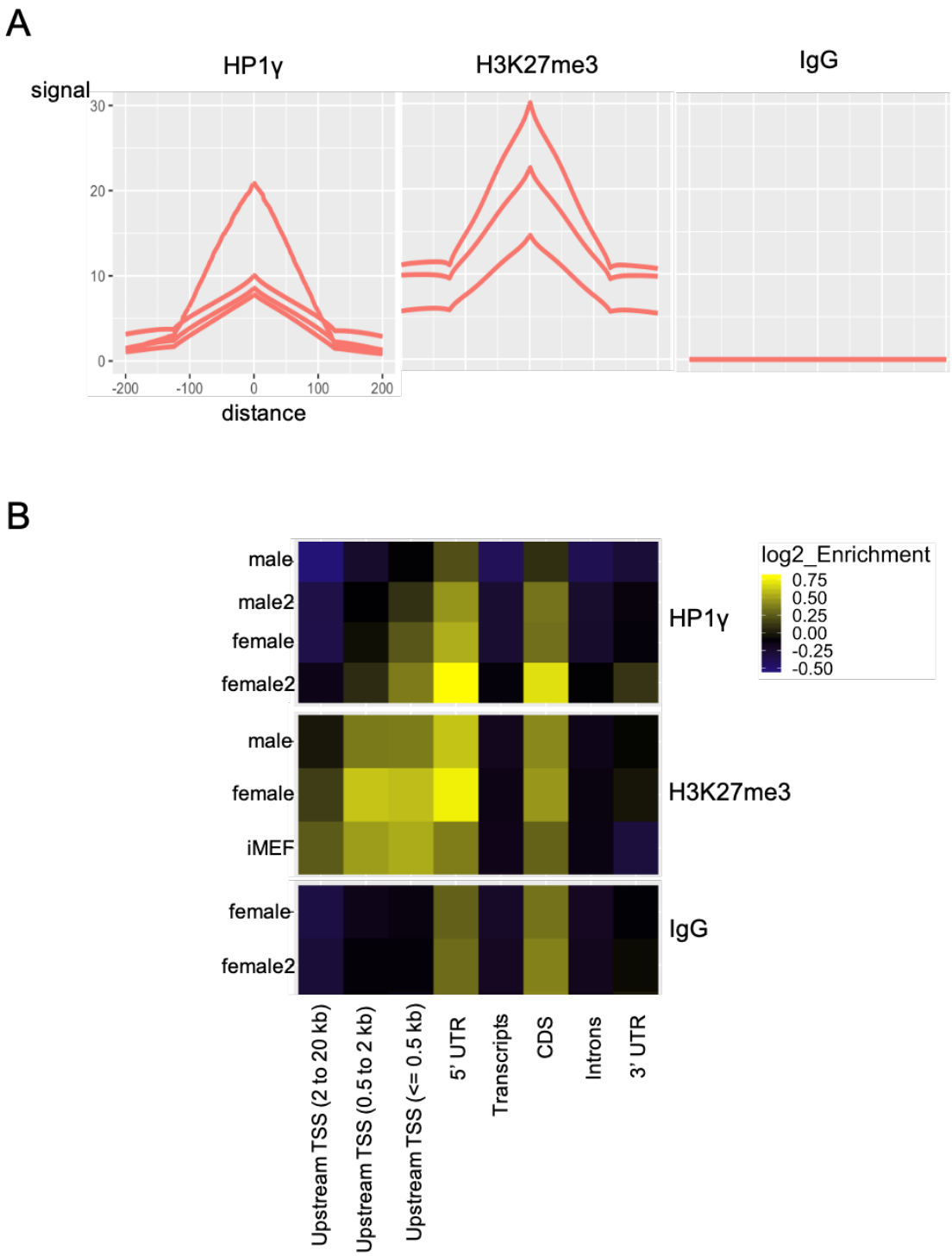


Figure 3. 11 Peak profile plot and Heatmap of signal enrichment over genomic features act as quality controls for the CUT & RUN. (A) The percentage of signal for a 400 bp window upstream and downstream of the summit is computed and averaged for all peaks in the sample. Notice the distinct patterns of enrichment for HP1γ and H3K27me3 replicates in these peaks, with IgG signal remaining flat. (B) Heatmap shows the log2 signal enrichment for specified genomic features with regions of greater enrichment depicted in yellow and lower enrichment in black. iMEF - immortalized MEF; UTR – untranslated region; CDS - coding sequence; TSS – Transcription start site.

Principal component analysis (PCA) (Figure 3.12) and heatmaps (Figure 3.13) of normalised HP1γ reads, reveal that female replicates cluster close together when examined for a 2,000 bp window around the TSSs (Figure 3.12 A and 3.13 A) and over the body of genes (Figure 3.12 B and 3.13 B). Similar is the case for males, suggesting for low variability among the replicates.

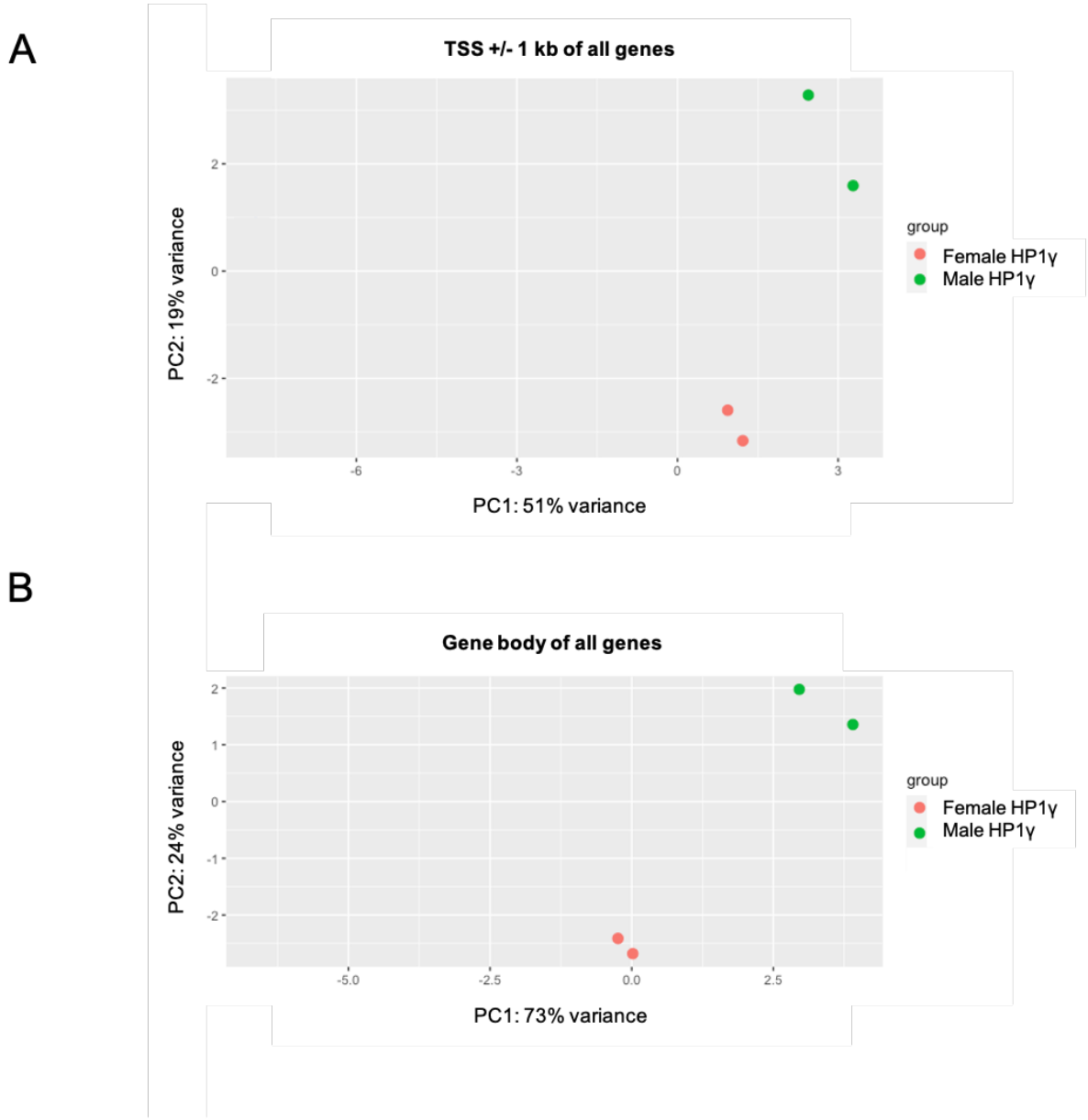


Figure 3.12 Principal component analysis (PCA) shows high similarity of HP1γ male and female replicates for certain genomic features. PCA plots of male and female replicates for HP1γ binding on (A) the TSS ± 1 kb and (B) the body of all genes. Raw counts have been Rlog transformed to minimize differences between samples in respect to library size. The first two principal components (PC1 and PC2) are plotted and coloured according to sex. Percentage of variation for each principal component is shown.

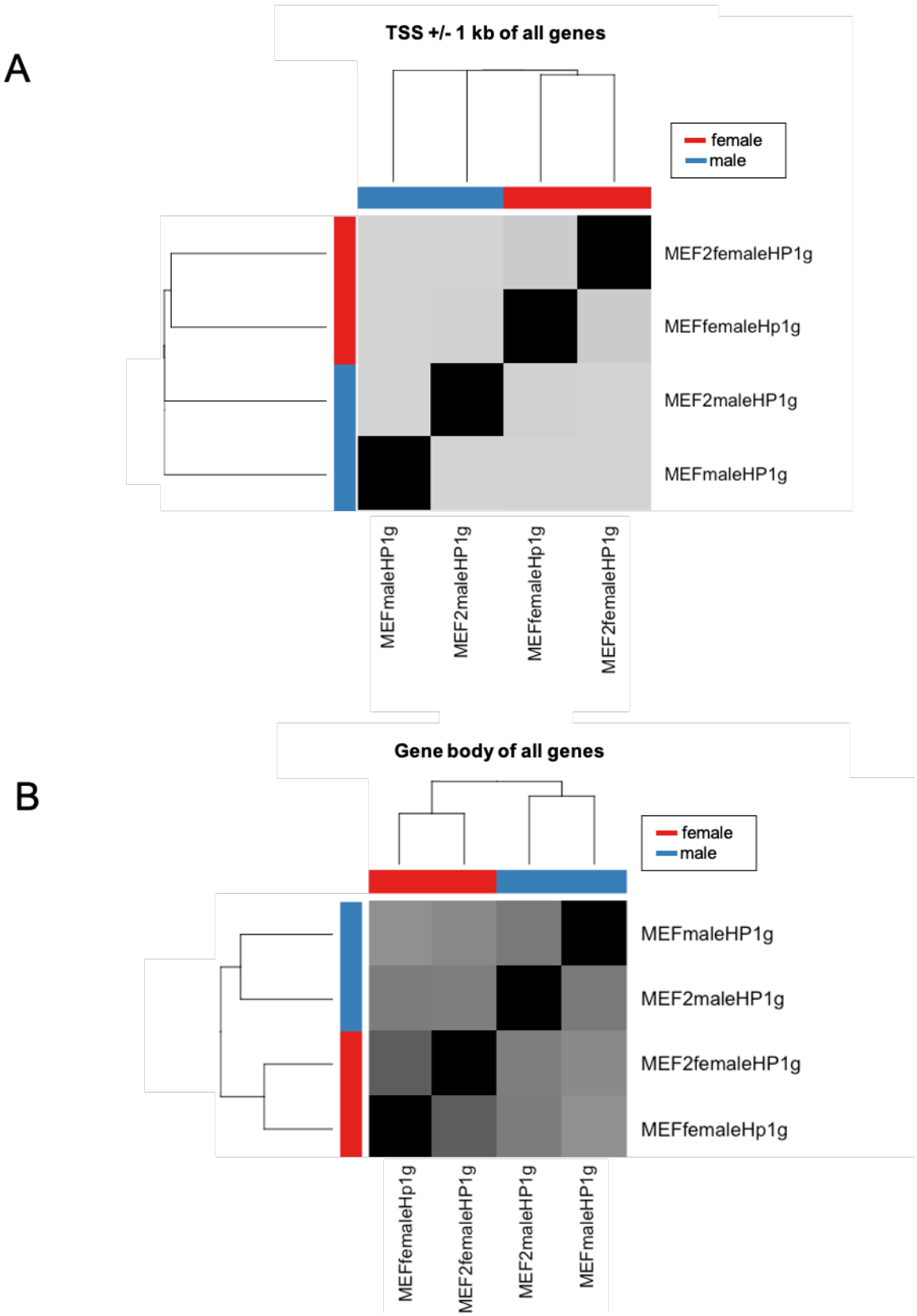


Figure 3.13 Heatmaps of male and female HP1γ show good clustering among the replicates. Heatmaps of male and female replicates for HP1γ binding on (A) the TSS ± 1 kb and (B) the body of all genes. Raw counts have been Rlog transformed to minimize differences between samples in respect to library size.

Downstream CUT & RUN analysis followed, by exploiting the available transcriptomic data (Law *et al.*, 2019) and dividing the total number of genes in quartiles, based on their expression levels. This analysis revealed that in both sexes, HP1γ presents a preferential binding on highly expressed genes (Figure 3.14). Specifically, there is an enrichment of HP1γ at the promoter region, with an absence of signal observable a few bp upstream of the TSS. HP1γ binding over the gene body follows a “high gene expression-high enrichment” pattern and an HP1γ peak is observed downstream of the transcript end site (TES). Of note, perhaps surprisingly, the total HP1γ binding signal appears to be higher in females compared to males (note Y axis of Figure 3.14).

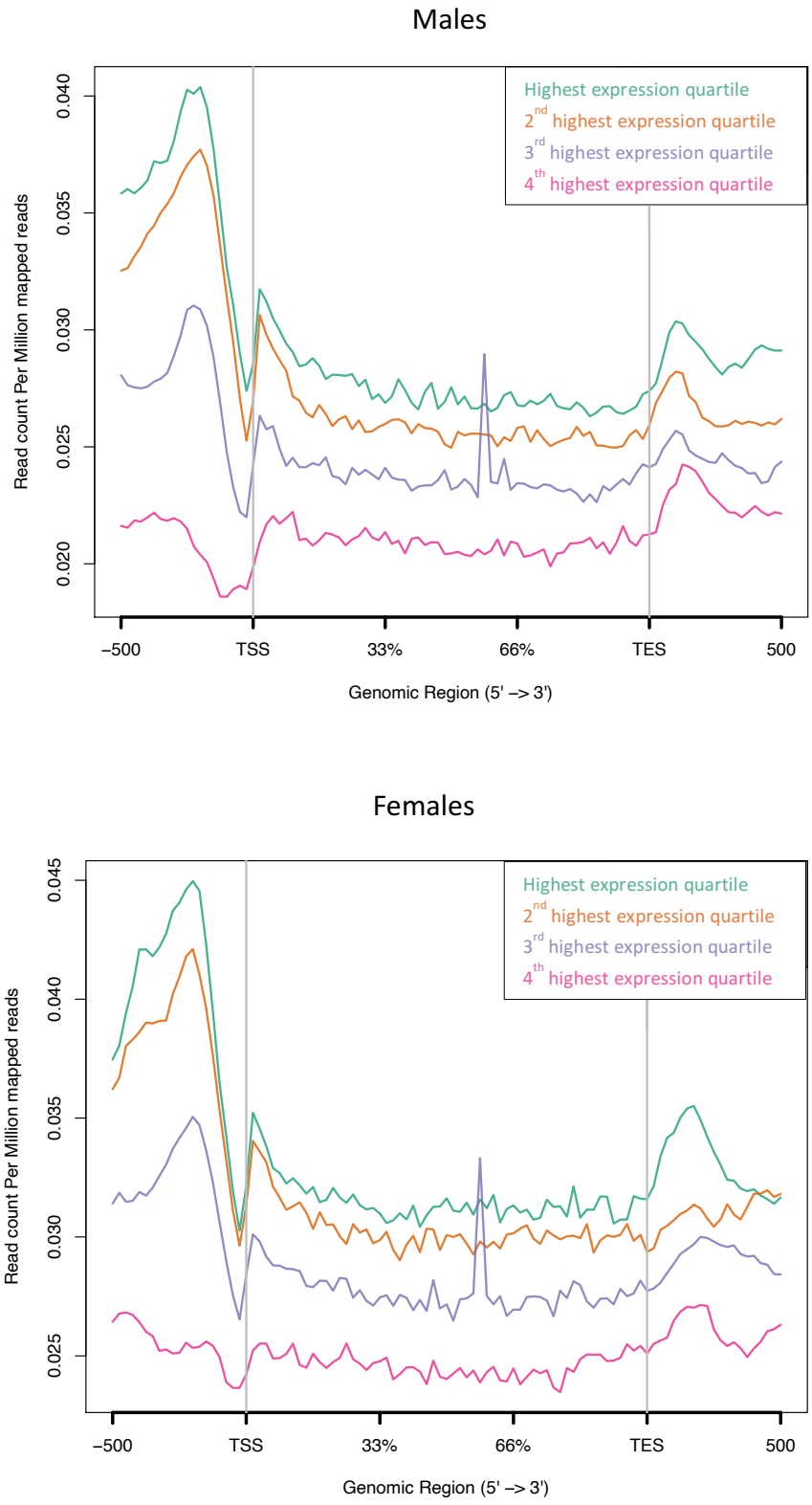


Figure 3.14 HP1γ is enriched over highly expressed genes in both sexes. Quantitation trend plots show HP1γ binding over the gene body ± 500 bp, for the highest, 2nd highest, 3rd highest and 4th highest gene expression quartiles in HP1γ$^{+/+}$ males and females (merged replicates, n=2 for both sexes). Notice the signal enrichment at the promoter region and downstream of the TES.

As we are particularly interested in elucidating the molecular mechanism of how HP1γ regulates sexually dimorphic genes, we focused on the binding of HP1γ specifically at sexually dimorphic genes (see Section 3.1.1). By plotting heatmaps of normalised HP1γ signal we visualised HP1γ's sex-related enrichment on "male lower" and "male higher" genes. Some "male lower" genes show preferential binding of HP1γ in males (e.g *Havcr2, Spp1, Lnsl3*), others show preferential binding in females (e.g *Bzrap1, Tnfrsf18*) and a third group of genes show very similar HP1γ enrichment among the sexes (e.g *Rex2, Snord68, Cdkn2a*) (Figure 3.15).

The same is true for the "male higher" group, where genes like *Agtr1b*, *Lpl* and *Wnt6* are enriched in males, genes like *Stac, Adamtsl2* and *Lrcc10b* are enriched in females and genes like *Col8a2* and *Pappa2* show similar levels of HP1γ binding in both sexes (Figure 3.16).

Figure 3.15 Heatmap of HP1γ binding on "male lower" genes. Unsupervised hierarchical clustering of the "male lower" genes in males and females (merged replicates, n=2 for both sexes), where each row represents the normalised HP1γ signal for the corresponding gene. Greater enrichment is depicted in red and lower enrichment in white.

Figure 3.16 Heatmap of HP1γ binding on "male higher" genes. Unsupervised hierarchical clustering of the "male higher" genes in males and females (merged replicates, n=2 for both sexes), where each row represents the normalised HP1γ signal for the corresponding gene. Greater enrichment is depicted in red and lower enrichment in white.

During this analysis, it came to our attention that *Sfi1* which is one of the "male lower" genes, was a binding hotspot for all targets, including the negative control IgG, thus skewing the results' interpretation (Supplementary Figure S3). Trend plots where *Sfi1* gene was removed, revealed that in males, HP1γ binds more on the gene body of "male lower" genes compared to the "male higher" genes, a phenomenon observed in females as well (Figure 3.17).

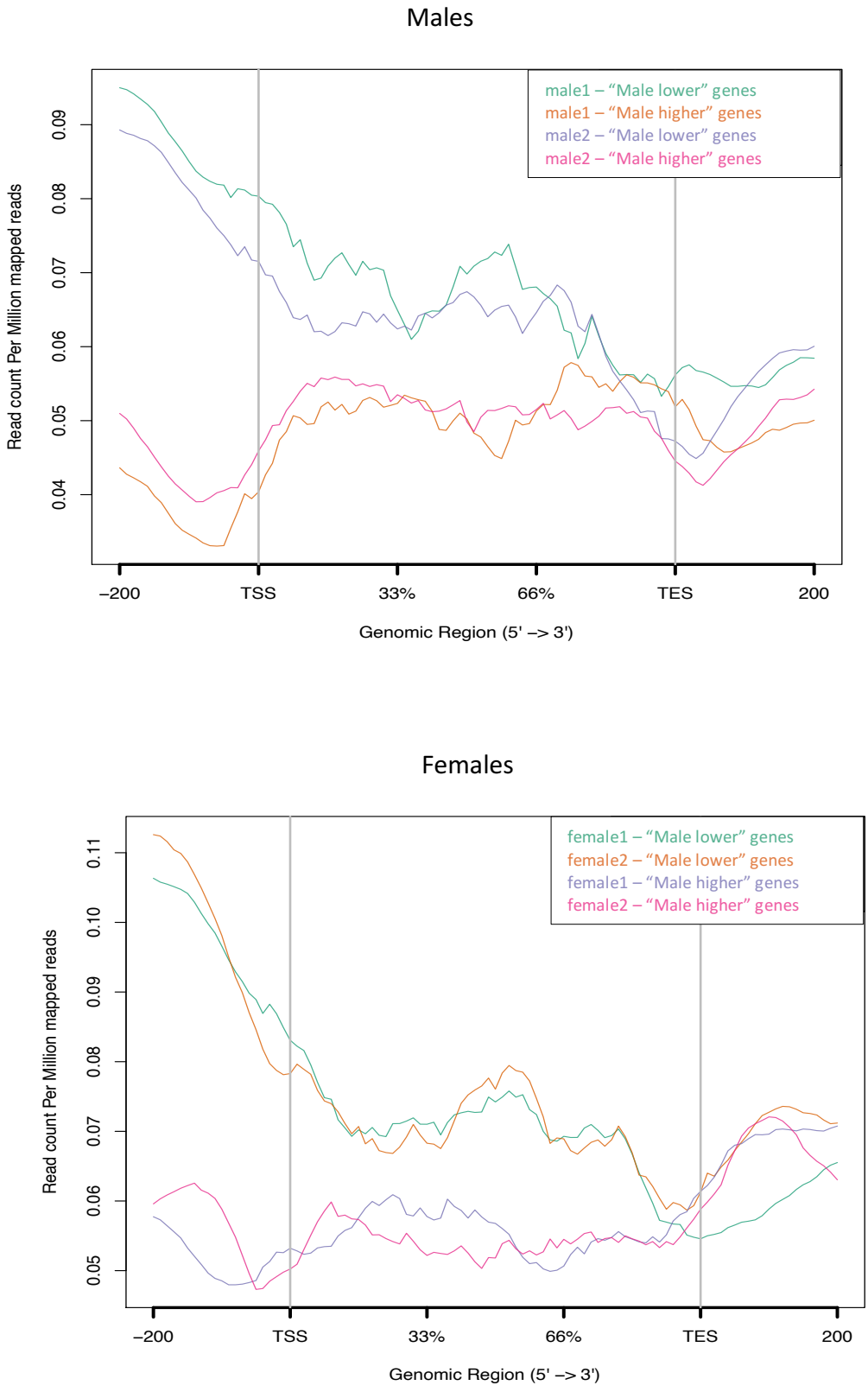


Figure 3.17 HP1γ binding over sexually dimorphic genes in both sexes. Quantitation trend plots show HP1γ enrichment over the gene body ± 200 bp, for "male higher" and "male lower" genes in HP1γ$^{+/+}$ males and females. HP1γ binds more on "male lower" genes compared to "male higher" genes in both sexes.

Taking advantage of the CUT & RUN data, we also questioned HP1γ's localisation on heterochromatic regions and whether there are sex differences regarding its binding. We focused on telomeres as they are prime examples of mouse heterochromatin. These sites are normally not included in sequencing analysis owing to their repetitive nature. Therefore, we generated an *in silico* probe with length of 150 bp, that is comprised of 25 tandem repeats of the telomeric sequence "TTAGGG". Since mouse telomeres can be as long as 150 kb (Zijlmans *et al.*, 1997), by aligning this probe to the sequencing reads, we can identify the HP1γ binding on telomeric sequences. In Figure 3.18, we can observe a small enrichment of HP1γ on female telomeres compared to male telomeres suggesting a potential sexually dimorphic nature, however it is important to mention that the variation between the replicates is substantial.

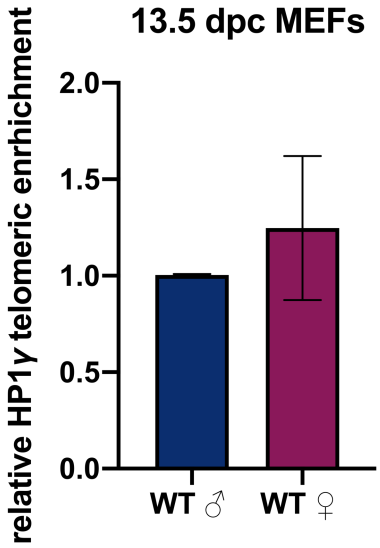


Figure 3.18 Small enrichment of HP1γ on female telomeres. Bars show the relative enrichment of HP1γ on male and female telomeres. Data have been normalized to yeast spike-in and error bars represent ± SEM (n=2 for both sexes).

Investigation of HP1γ's binding on specific telomeric sites with the IGV genome browser, reveals that HP1γ primarily localises at the q arm ends of chromosome X and chromosome 18 (Figure 3.19). Strongest HP1γ enrichment is observed on the telomeres of chromosome 18 with signal distribution showing a similar pattern for both sexes. Other chromosome ends showed little to no HP1γ binding including chromosome 9 and chromosome 10 (Figure 3.20).

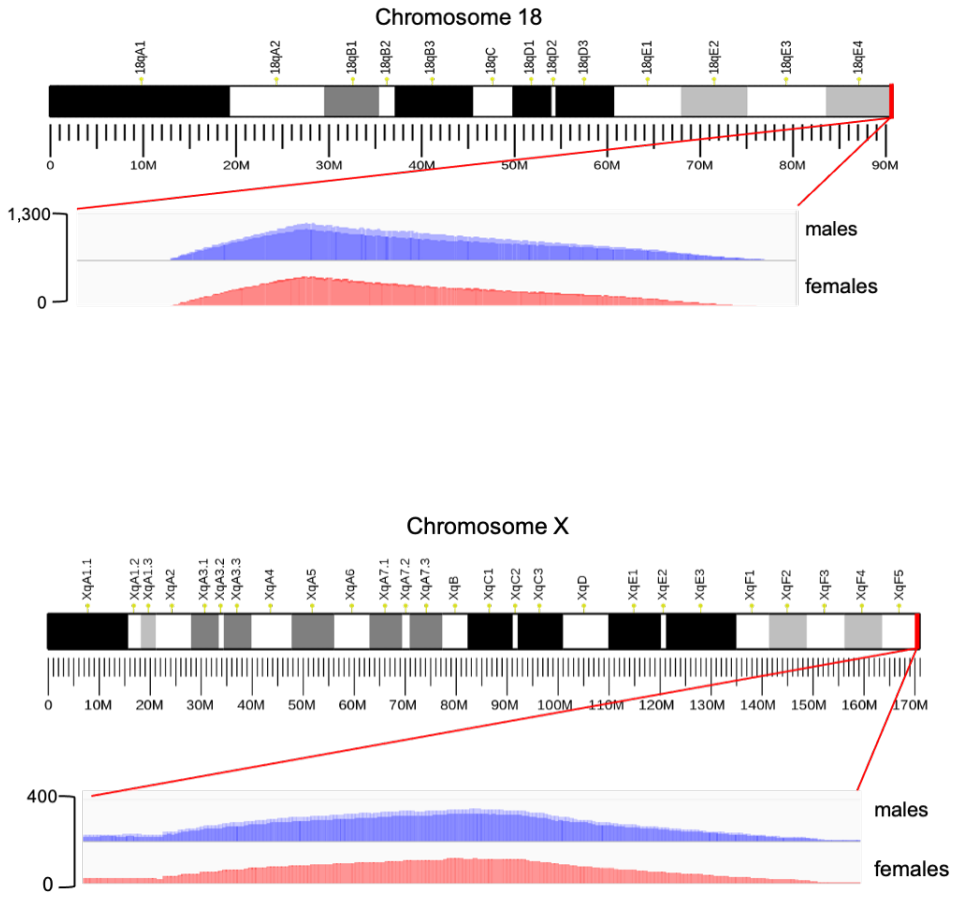


Figure 3.19 HP1γ is enriched on chromosome 18 and chromosome X telomeres. IGV genome browser snapshots of the indicated positions on chromosome 18 and X. High HP1γ CUT & RUN signal is observed for both males and females (merged replicates, n=2 for both sexes).

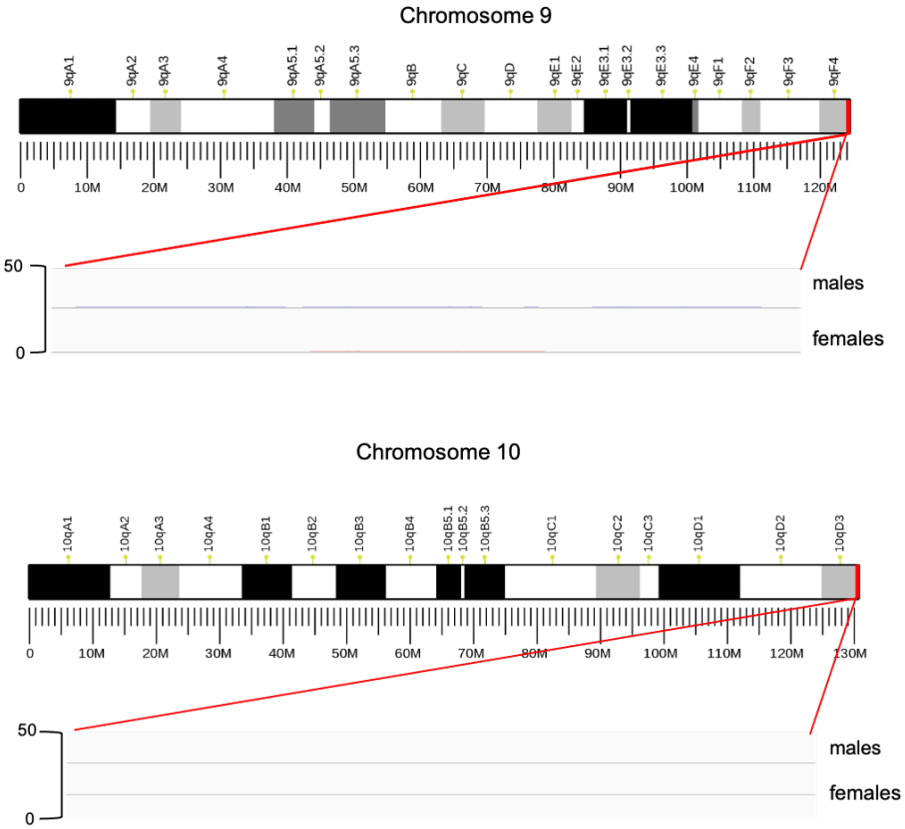


Figure 3.20 Lack of HP1γ binding on chromosome 9 and chromosome 10 telomeres. IGV genome browser snapshots of the indicated positions on chromosome 9 and 10. There is lack of HP1γ CUT & RUN signal for both males and females (merged replicates, n=2 for both sexes).

3.2.6 Investigating the potential involvement of HP1γ in cryptic transcription regulation

The intriguing observation that HP1γ is enriched at highly expressed genes (see Section 3.2.5), alongside the known association of RNA polymerase II with methylated H3K9 and HP1γ during transcription elongation (Vakoc *et al.*, 2005), suggests an epigenetic, context-specific role of this factor in gene regulation. Why would this isoform of HP1 that normally marks heterochromatin be also found at the transcribed portions of active genes? Aiming to address this paradox, we hypothesized that HP1γ inhibits cryptic transcriptional initiation. Cryptic transcription can occur both in sense and antisense orientation, when RNA Polymerase II transcribes intragenic regions or nucleosome-free loci located in the direct vicinity of the TSS that are normally inaccessible to the transcriptional machinery (Katayama *et al.*, 2005; Seila *et al.*, 2008). Antisense transcripts have also been shown to arise in yeast from the body of coding genes where the presence of H3K36me3 over their start sites results in their suppression (Carrozza *et al.*, 2005; Venkatesh *et al.*, 2016). The cryptic transcripts usually are extremely short lived, hence common RNA assays have previously underestimated their levels, yet recent transcriptomic studies showed that cryptic transcripts can comprise as much as 13% of the mapped transcripts (Neil *et al.*, 2009; Xu *et al.*, 2009).

To examine whether HP1γ regulates cryptic transcription in mice, we performed a preliminary experiment utilising the latest version of cap analysis of gene expression (CAGE). CAGE is used for 5'-centered expression profiling of RNA polymerase II transcripts (Shiraki *et al.*, 2003) and allows the genome-wide mapping of TSSs at a single nucleotide resolution. This information can be used for various analyses, including the studying of promoter architecture (Carninci *et al.*, 2006). For instance, CAGE experiments revealed genomic regions that extend around 40 bp upstream and downstream the TSSs, where the transcription machinery binds to direct initiation of gene expression (Smale and Kadonaga, 2003). The initial CAGE protocol required 50 μg of total RNA as starting material (Takahashi *et al.*, 2012), while the recently-developed super low input carrier - cap analysis of gene expression (SLIC-CAGE), has enabled the generation of high-complexity libraries from only 5 ng of total RNA, with

longer reads and increased mappability (Cvetesic *et al.*, 2018). Figure 3.21 shows a schematic representation of the SLIC-CAGE technique.

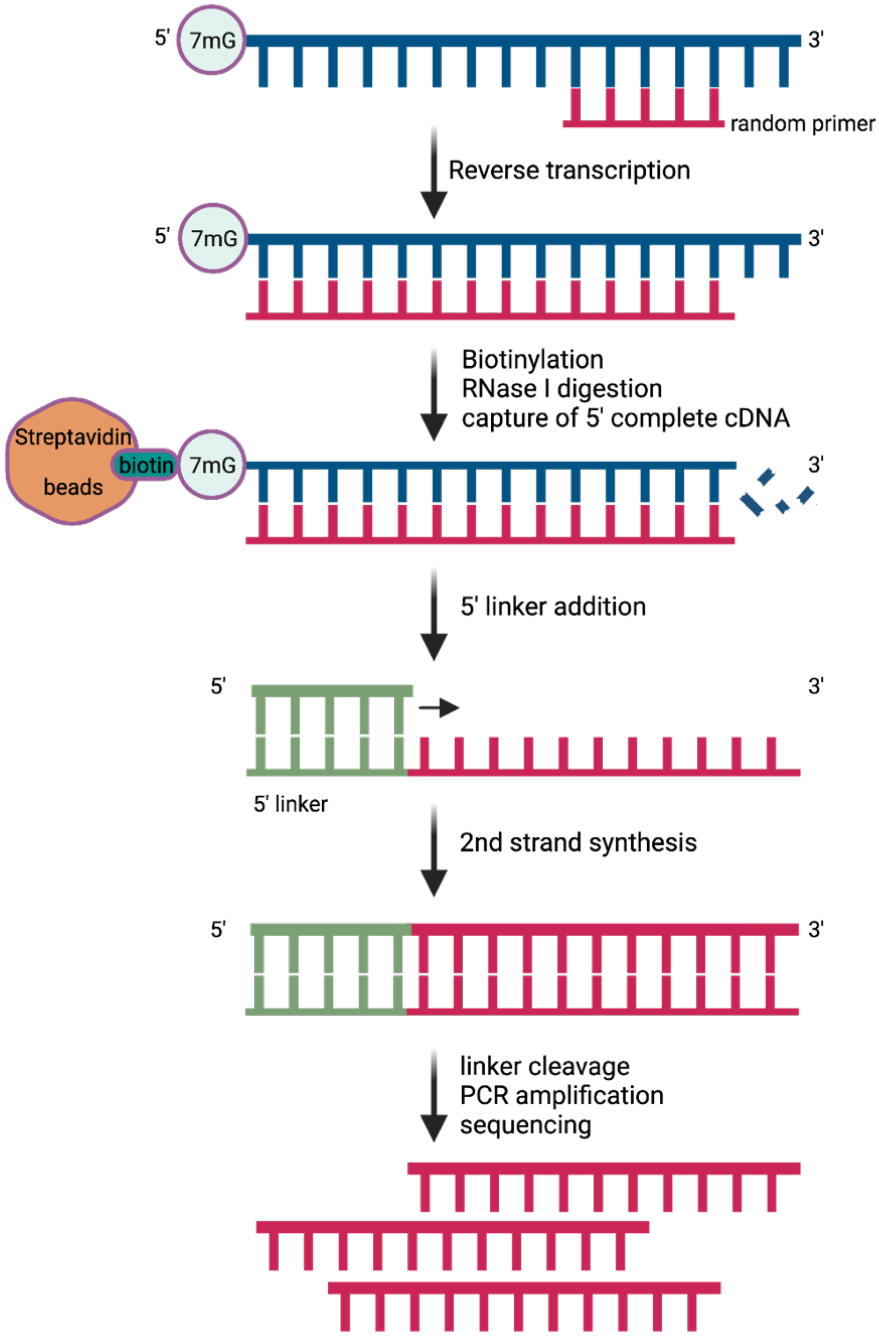


Figure 3. 21 Schematic of the Cap analysis of gene expression (CAGE) experiment. cDNA (red strand) is synthesized through reverse transcription of total RNA and 7-Methylguanosine (7mG) cap is oxidized using sodium periodate. Oxidation allows the attachment of biotin which enables the isolation of the fragments with affinity purification on streptavidin magnetic beads. The samples are treated with RNase I for RNA degradation and cDNA is released and

5’ linkers are ligated. Reverse transcription follows and the library molecules are PCR-amplified to increase the amount of material for sequencing.

Total RNA from two biological replicates of HP1γ$^{+/+}$ male and female and HP1γ$^{-/-}$ male and female MEFs was isolated and RNA integrity was tested by running the samples on the Bioanalyzer (Figure 3.22). RNA of high quality was employed for the SLIC-CAGE experiment with all the samples having RNA integrity number (RIN) values of ≥ 9.5 (Schroeder *et al.*, 2006).

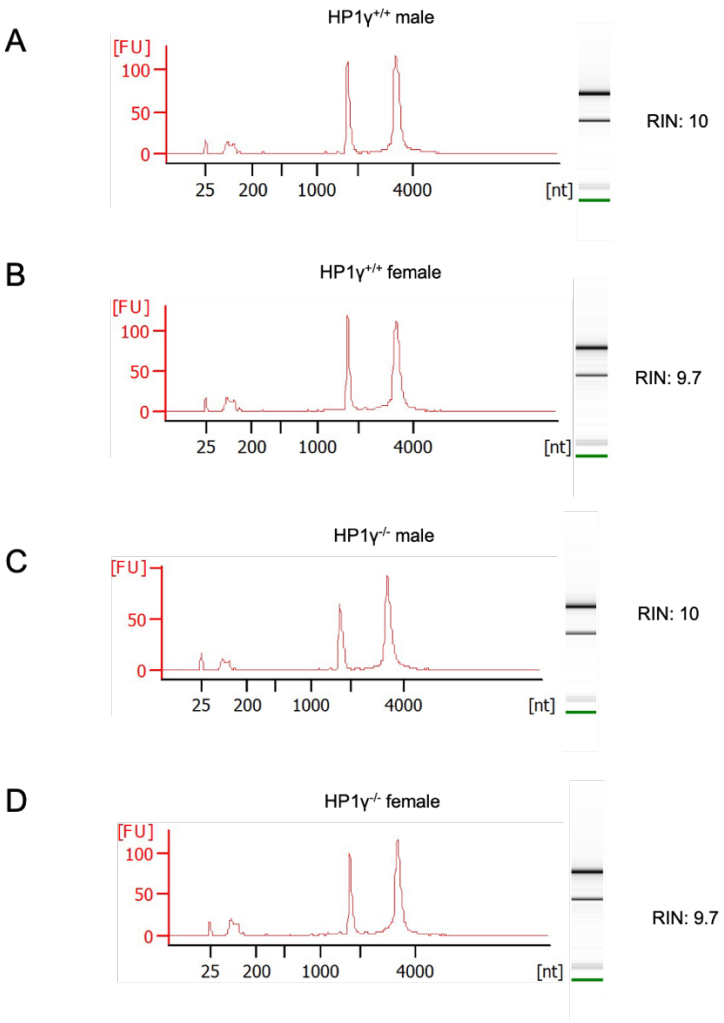


Figure 3. 22 Bioanalyzer analysis for total RNA quality. Representative electropherograms of the indicated genotypes (A-D), alongside the RNA integrity number (RIN), a classification system of RNA quality developed by Agilent Technologies. RIN of 1 refers to fully degraded

RNA, while a RIN of 10 refers to, intact, high-quality RNA. The peak of 25 bp, represents the lower marker, while the peaks of 2,000 bp and 4,500 bp represent the 18S and 28S RNA species, respectively.

SLIC-CAGE was performed in collaboration with Boris Lenhard's group and quality of PCR amplified samples was tested with the Bioanalyzer (Figure 3.23), before proceeding to paired-end sequencing. Due to technical reasons during the SLIC-CAGE experiment (limitation of available indexes at the time), one of the HP1γ$^{-/-}$ male replicates was processed as two separate samples, as it would not constitute a problem for post-sequencing analysis (see below).

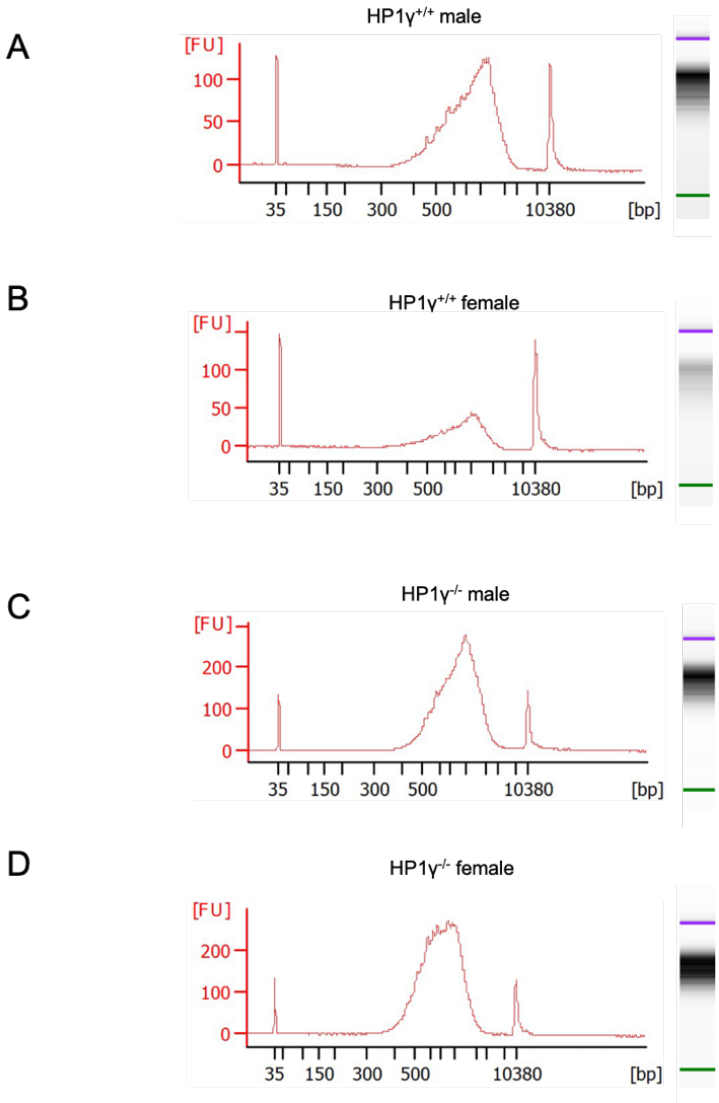


Figure 3.23 Bioanalyzer analysis for PCR-amplified, SLIC-CAGE samples. Representative electropherograms of the indicated genotypes (A-D), from SLIC-CAGE, PCR-amplified

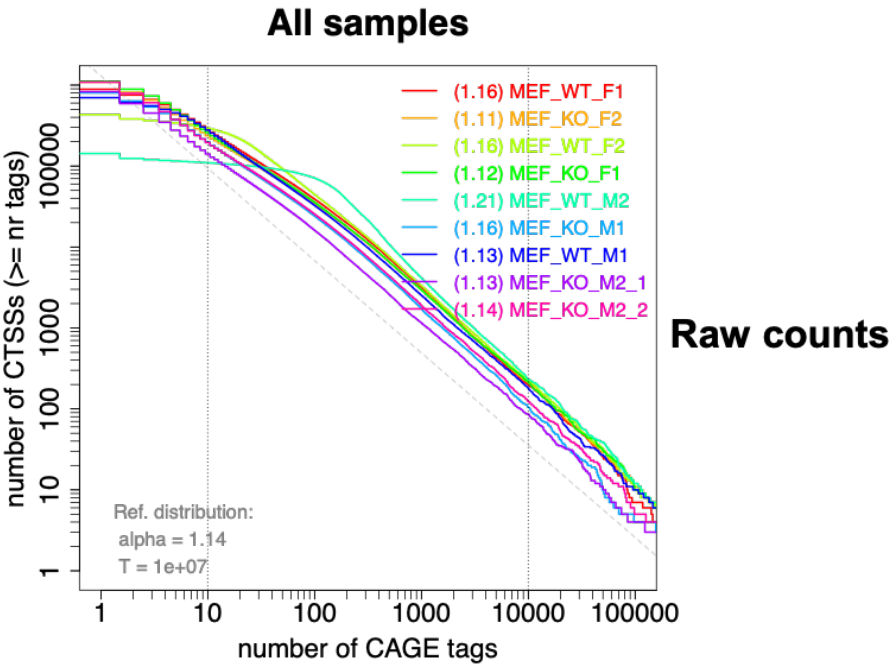
samples. PCR amplification results in enrichment of transcripts with sizes of 300 bp to 3,000 bp for all conditions. The peaks of 35 bp and 10380 bp, represent the lower and upper marker, respectively.

Table showing the abbreviations used for SLIC-CAGE analysis:

Sample name	**Abbreviation**
HP1γ$^{+/+}$ male replicate 1	MEF_WT_M1
HP1γ$^{+/+}$ male replicate 2	MEF_WT_M2
HP1γ$^{+/+}$ female replicate 1	MEF_WT_F1
HP1γ$^{+/+}$ female replicate 2	MEF_WT_F2
HP1γ$^{-/-}$ male replicate 1	MEF_KO_M1
HP1γ$^{-/-}$ male replicate 2	MEF_KO_M2_1 and MEF_KO_M2_2
HP1γ$^{-/-}$ female replicate 1	MEF_KO_F1
HP1γ$^{-/-}$ female replicate 2	MEF_KO_F2

Raw sequencing counts were normalized with a power-law distribution to a typical total of 10^6 tags (Balwierz *et al.*, 2009), resulting in normalized tags per million (TPMs) (Figure 3.24) and SLIC-CAGE signal was annotated to specific genomic features like promoters, 5' UTRs, introns, exons, 3' UTRs but also distal intergenic regions (Figure 3.25).

A



B

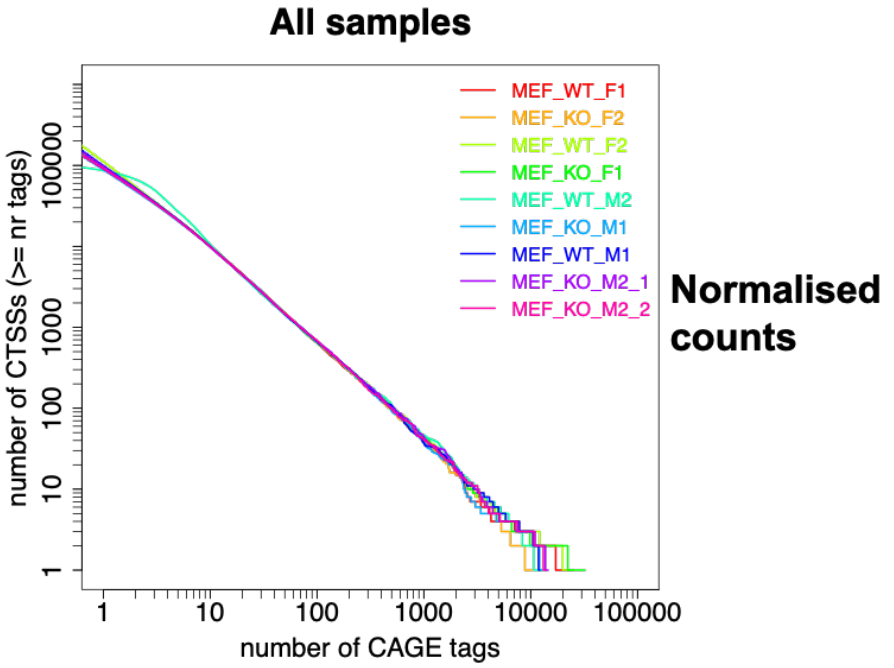


Figure 3.24 Power-law normalization of SLIC-CAGE reads. (A) Plotting the number of SLIC-CAGE tags (X-axis) against the number of TSSs that are supported by ≤ of the number of tags (Y-axis) results in a distribution that can be approximated by a power-law. (B) Power-law is fitted to the reverse cumulative distribution of a total number of tags equal to 10^6, resulting in normalized tags per million (TPM) values.

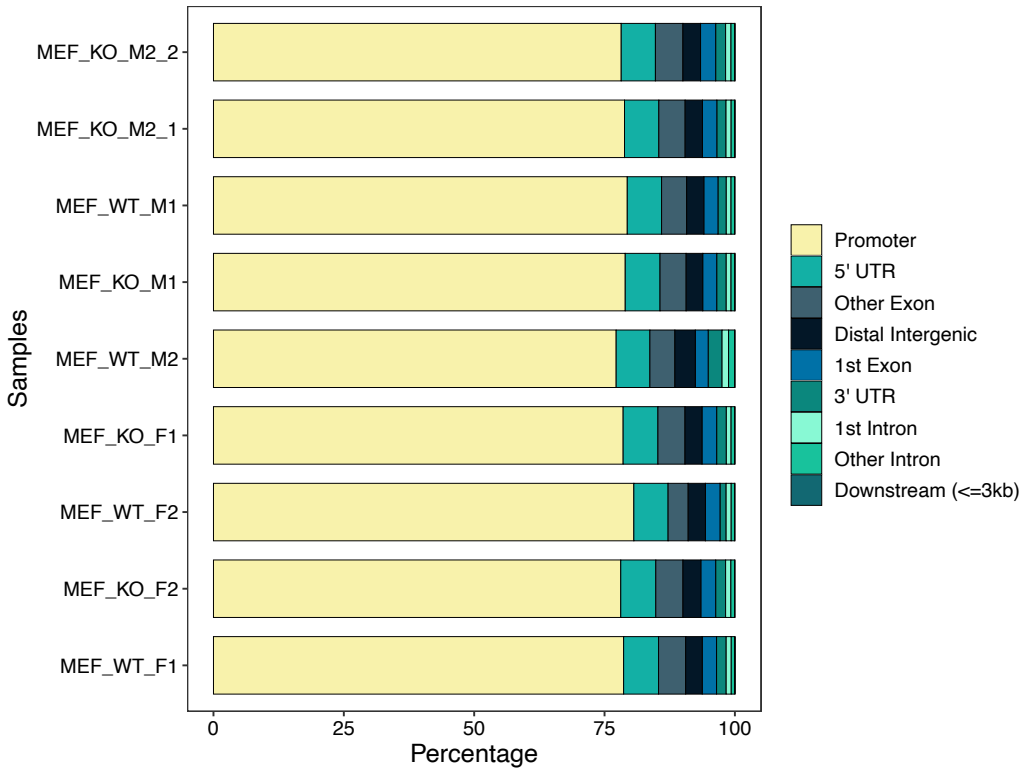
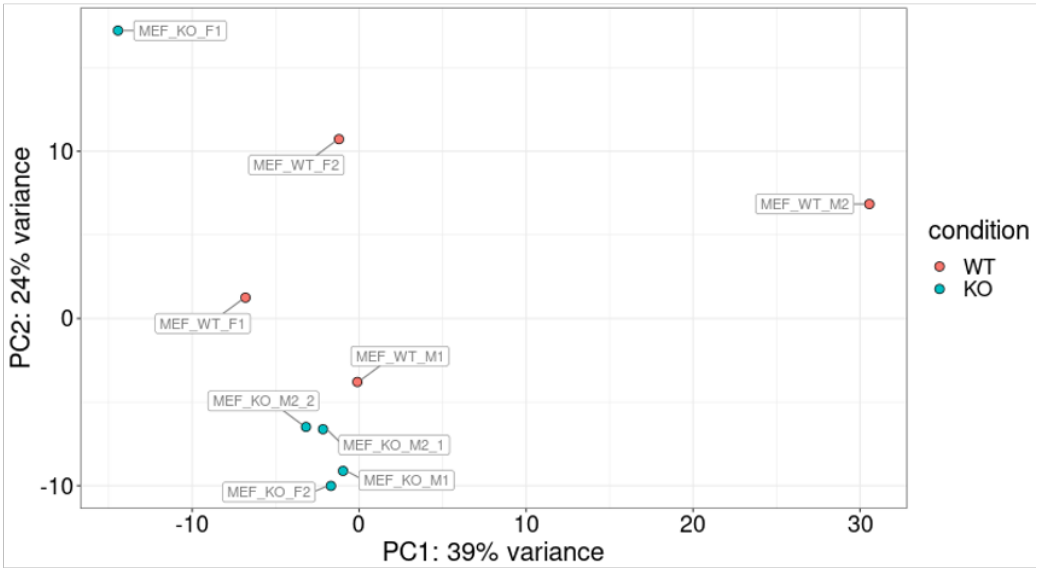


Figure 3.25 Annotation of SLIC-CAGE samples to genomic features. Stacked bar plots show the percentage of normalised SLIC-CAGE residing in the specified genomic features for the corresponding samples.

The majority of the reads (more than 75 %) map at promoters, with 5' UTR being the second most-mapped feature. Overall, there are little to none differences regarding the genome annotations, irrespective of samples' sex or HP1γ genotyping. PCA (Figure 3.26 A) and Pearson's correlation scatterplots were then generated (Figure 3.26 B) and revealed two main things. First, that MEF_KO_M2_1 and MEF_KO_M2_2 samples show low variability and second, that MEF_WT_M2 does not correlate well with the rest of the samples.

A



B

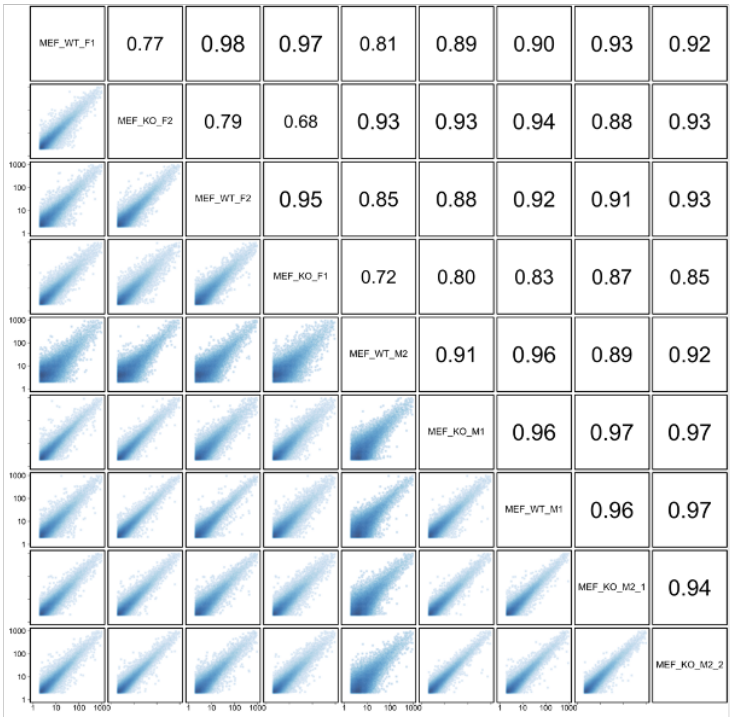
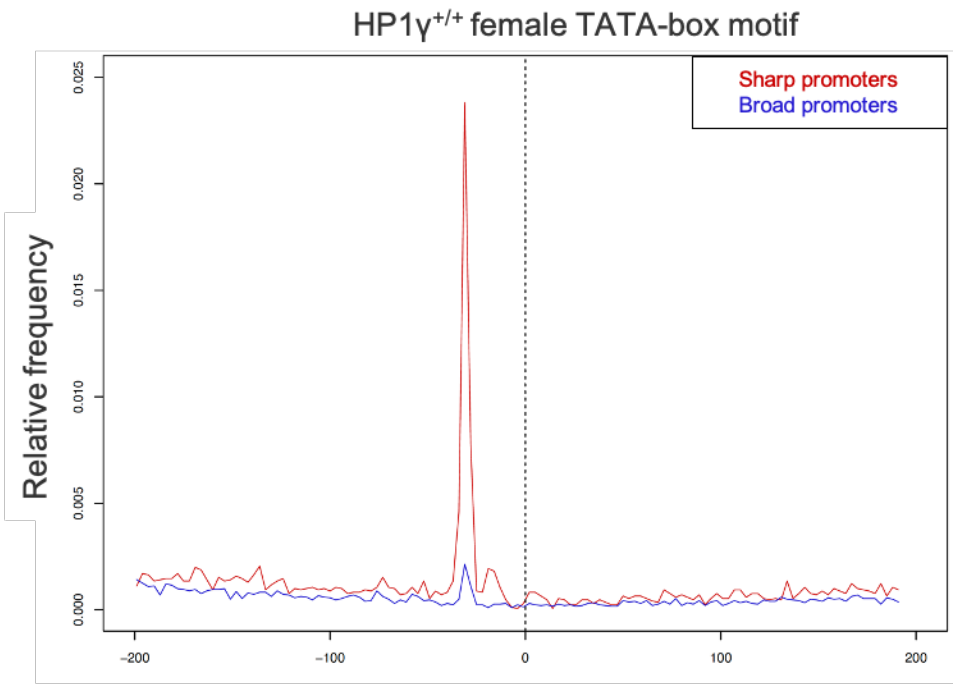


Figure 3.26 MEF_WT_M2 does not cluster with the rest of the samples. (A) PCA plots of the normalised SLIC-CAGE samples. The first two principal components (PC1 and PC2) are plotted and coloured according to HP1γ genotyping (WT= HP1γ wild-type, KO= HP1γ knockout). Percentage of variation for each principal component is shown. (B) Pairwise Pearson's correlation plots of raw SLIC-CAGE tag counts per TSS.

With that in mind, we proceeded to examine promoter architecture features. TSSs in the close proximity of each other give rise to a functionally equivalent set of transcripts and thus can be clustered into transcriptional units, called tag clusters (TCs) (Haberle *et al.*, 2015). TCs correspond essentially to promoters. In mammals, there are two main set of promoters. “Broad” promoters with multiple TSSs containing a CpG island, which are associated with widely expressed or developmentally regulated genes and “sharp” promoters with one dominant TSS often associated with a TATA-box at a fixed upstream distance, which often regulate tissue-specific transcription (Ponjavic *et al.*, 2006). Plotting the relative frequency of TATA-box motif for sharp and broad promoters for HP1γ$^{+/+}$ and HP1γ$^{-/-}$ samples reveals that, as expected, the majority of TATA boxes are found in sharp promoters with a much smaller fraction detectable in broad promoters. The preliminary analysis reveals that HP1γ does not have an effect on the mapping of TATA-boxes among the two promoter classes (Figure 3.27).

A



B

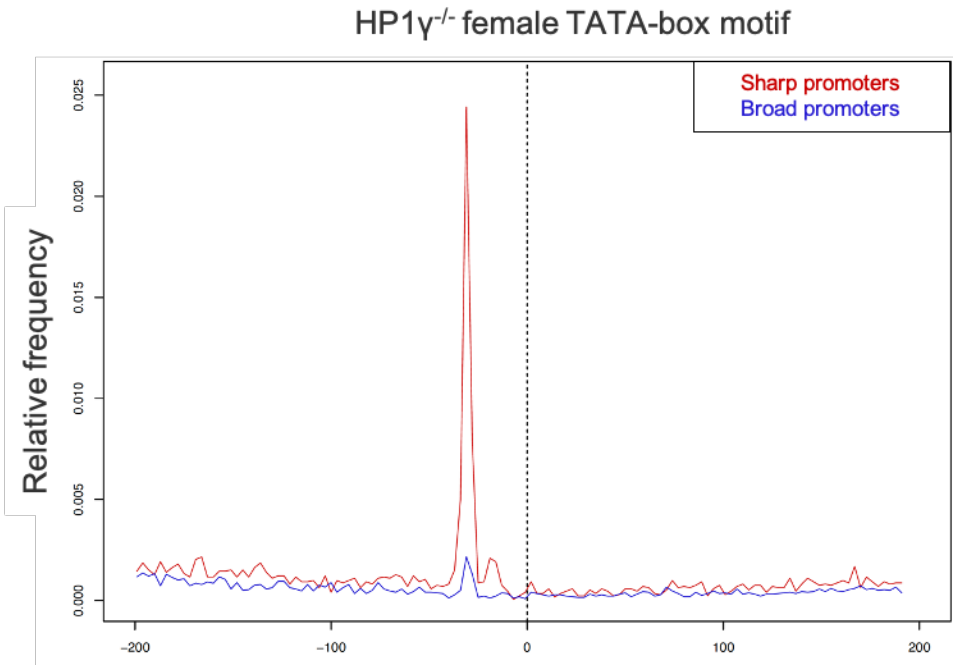


Figure 3.27 Sharp promoters have a higher occurrence of TATA-box motifs, irrespective of HP1γ. Cumulative Plots of A) HP1γ$^{+/+}$ females (n=2) and (B) HP1γ$^{-/-}$ females (n=2) show the average TATA-box motif occurrence for sharp and broad promoters in a window of 200 bp upstream and downstream the TSS.

Promoter width can be measured by taking into account the mapped CAGE signal at TSSs along the tag clusters. This information can further reveal the complexity of CAGE-derived libraries. Taking the cumulative distribution of CAGE signal along the TC and defining an interquantile promoter width as the bp distance between the lower (10^{th}) and the upper (90^{th}) quantile, I generated interquantile width histograms of the TCs. The histograms revealed very similar promoter width patterns among the samples, with the exception of MEF_WT_M2 (Figure 3.28). This analysis further highlighted the problematic nature of MEF_WT_M2 replicate, hinting to a low-complexity sample, while it emphasized the similarity of MEF_KO_M2_1 with MEF_KO_M2_2. Therefore, MEF_WT_M2 was excluded from subsequent analysis and the latter two samples were merged into one (now MEF_KO_M2_merged). This analysis also suggests that HP1γ does not affect gene promoter width.

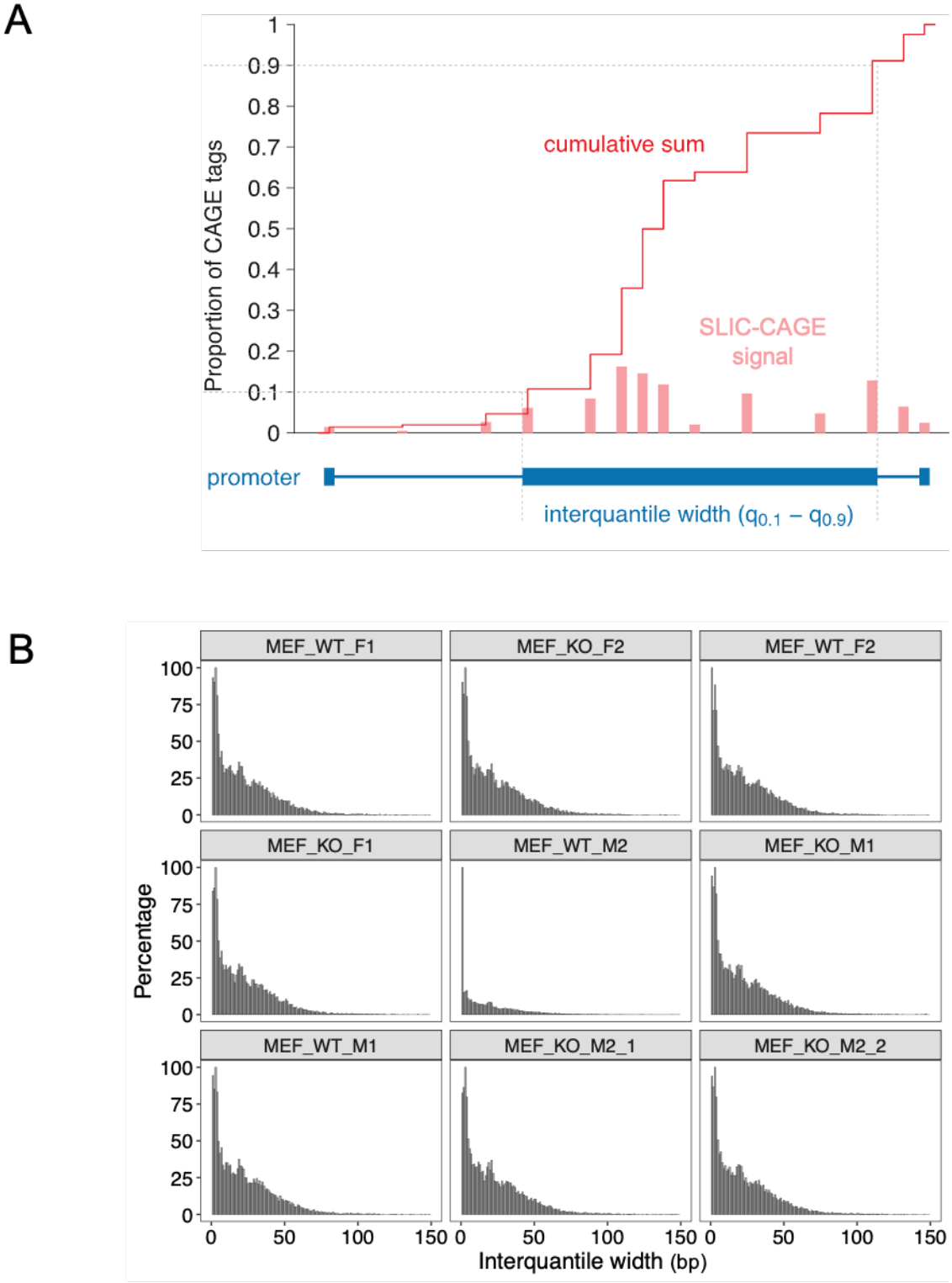


Figure 3.28 HP1γ does not affect gene promoter width. (A) Schematic representation of the interquantile promoter width computation procedure. (B) Histograms reveal very similar promoter width patterns across the different samples, with the exception of MEF_WT _M2, hinting to a low-complexity sample.

TSSs within the same promoter region can be used differently depending on the cell context. Thus, although the overall transcription signal from a promoter does not change among samples, the differential usage of TSSs, also known as promoter shifting may indicate changes in the regulation of transcription, which cannot be detected by conventional RNA-sequencing (Haberle et al., 2014). For all promoters, a shifting score is calculated based on the difference in the cumulative distribution of the SLIC-CAGE signal. The analysis of HP1γ$^{+/+}$ and HP1γ$^{-/-}$ samples suggests that HP1γ has a minimal effect on the differential usage of TSSs. HP1γ deficiency affected the promoters of 29 genes in males and 21 genes in females (Supplementary Table S1 & S2). Representative images of HP1γ-dependent promoter shifting are shown in Figure 3.29 for males and females. None of the identified promoters belonged to the sexually dimorphic genes, suggesting that differential usage of TSS is not playing a role in the regulation of these genes.

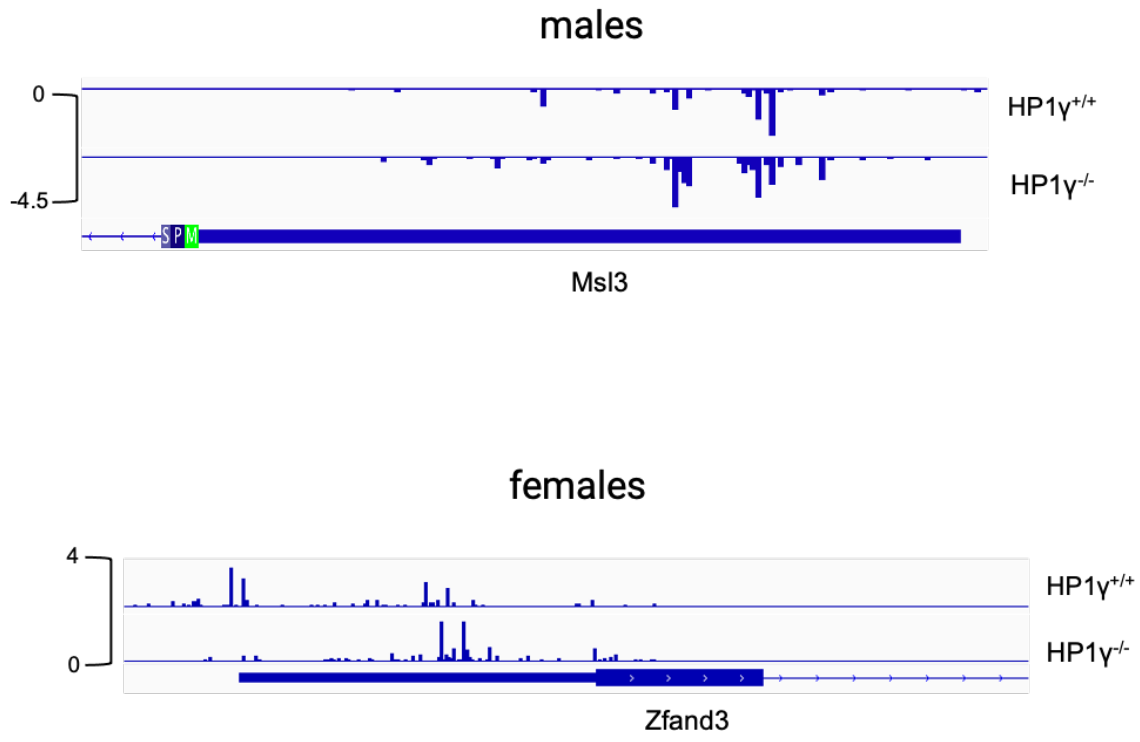


Figure 3.29 Examples of promoter shifting sensitive to HP1γ. IGV genome browser snapshots of genes that HP1γ is responsible for differential usage of their TSSs in males (merged files, n=1 for HP1γ$^{+/+}$ and n=2 for HP1γ$^{-/-}$) and females (merged files, n=2).

3.2.7 Questioning's HP1γ's higher enrichment on female chromatin

The slight enrichment of genome-wide HP1γ CUT & RUN signal observed in females compared to males, could hint to stronger or more abundant binding of this factor on female chromatin. Higher HP1γ signal in females has been also observed previously with ChIP-seq experiments performed by P. Law in Festenstein's group. However, as chromatin input was used for signal normalisation in the ChIP-seq experiment, it did not allow the comparison of HP1γ's overall binding between the sexes because previous studies have shown that the detected amount of binding is directly related to the sequencing depth which may vary between different samples (Kharchenko, Tolstorukov and Park, 2008). The normalisation of the signal provided by the spike-in control in the CUT & RUN experiment overcomes this problem as it is constant between samples, irrespective of sequencing depth (Chen *et al.*, 2016). Therefore, the higher female HP1γ signal according to CUT & RUN, could suggest that there is a sex-biased difference in the distribution of HP1γ in the nuclear fractions (e.g chromatin-bound and nuclear soluble HP1γ fractions).

In that respect, I performed a cell fractionation assay (similar to Gillotin, 2018), that allowed the isolation of cytosolic and sub-nuclear protein fractions and examined HP1γ's distribution in the different subcellular compartments and how this distribution may vary among the sexes. Figure 3.30 shows a schematic diagram of the fractionation assay

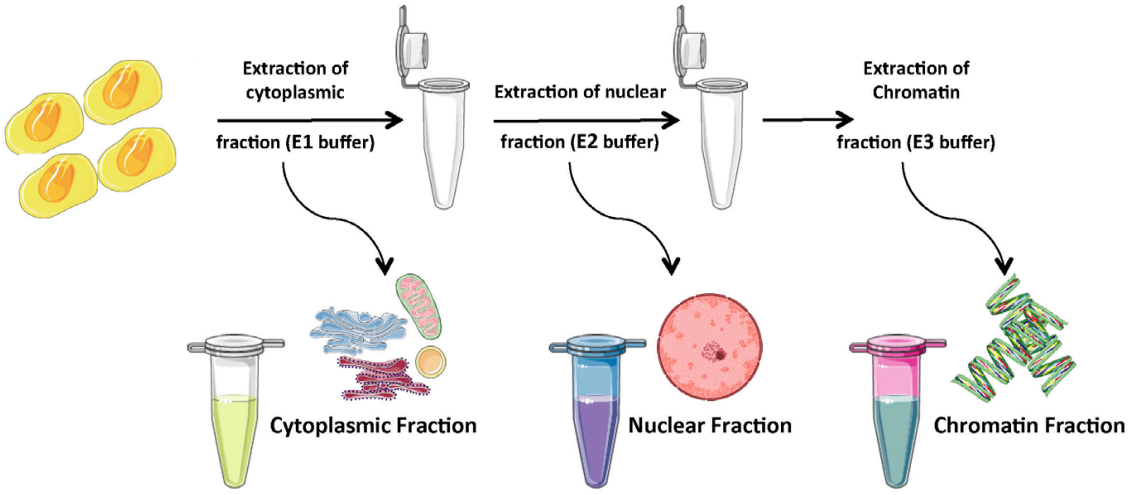


Figure 3.30 Cell fractionation assay workflow. The collection of each subcellular fraction is performed by successive protein extractions using the corresponding buffers. The purity of each fraction is controlled by Western blot for α-Tubulin, Lamin B and Histone H3 depicting the cytoplasmic, the nuclear and the chromatin fraction respectively.

Three biological replicates of HP1γ$^{+/+}$ males and females were included in the experiment and the purity of each subcellular fraction was controlled by Western blot for α-tubulin, Lamin B and Histone H3, depicting the cytoplasmic, the nuclear-soluble and the chromatin-bound fraction, respectively. An additional sample where all the subcellular fractions were combined at the end of the assay was included as a positive control, with all the loading factors being present (termed "combined fractions"). Following the fractionation assay, HP1γ protein levels were measured by Western blot (Figure 3.31 A). HP1γ signal was quantitated by normalising to α-tubulin for the cytoplasmic fraction, to Lamin B for the nuclear-soluble fraction and H3 for the chromatin-bound fraction. The analysis shows that both sexes have similar HP1γ levels regarding the cytoplasmic and nuclear soluble fractions. Interestingly, the chromatin-bound fraction of HP1γ shows a 1.9-fold enrichment in females compared to males (Figure 3.31 B), in agreement with the signal enrichment observed by CUT & RUN (see Section 3.2.5). The levels of HP1γ bound on female chromatin identified with the fractionation assay are much higher, compared to the signal enrichment observed with CUT & RUN. This is likely due to the exclusion of repetitive sequences during the analysis of CUT & RUN data which are natural sites of HP1γ binding, therefore "dampening" the sexually dimorphic differences.

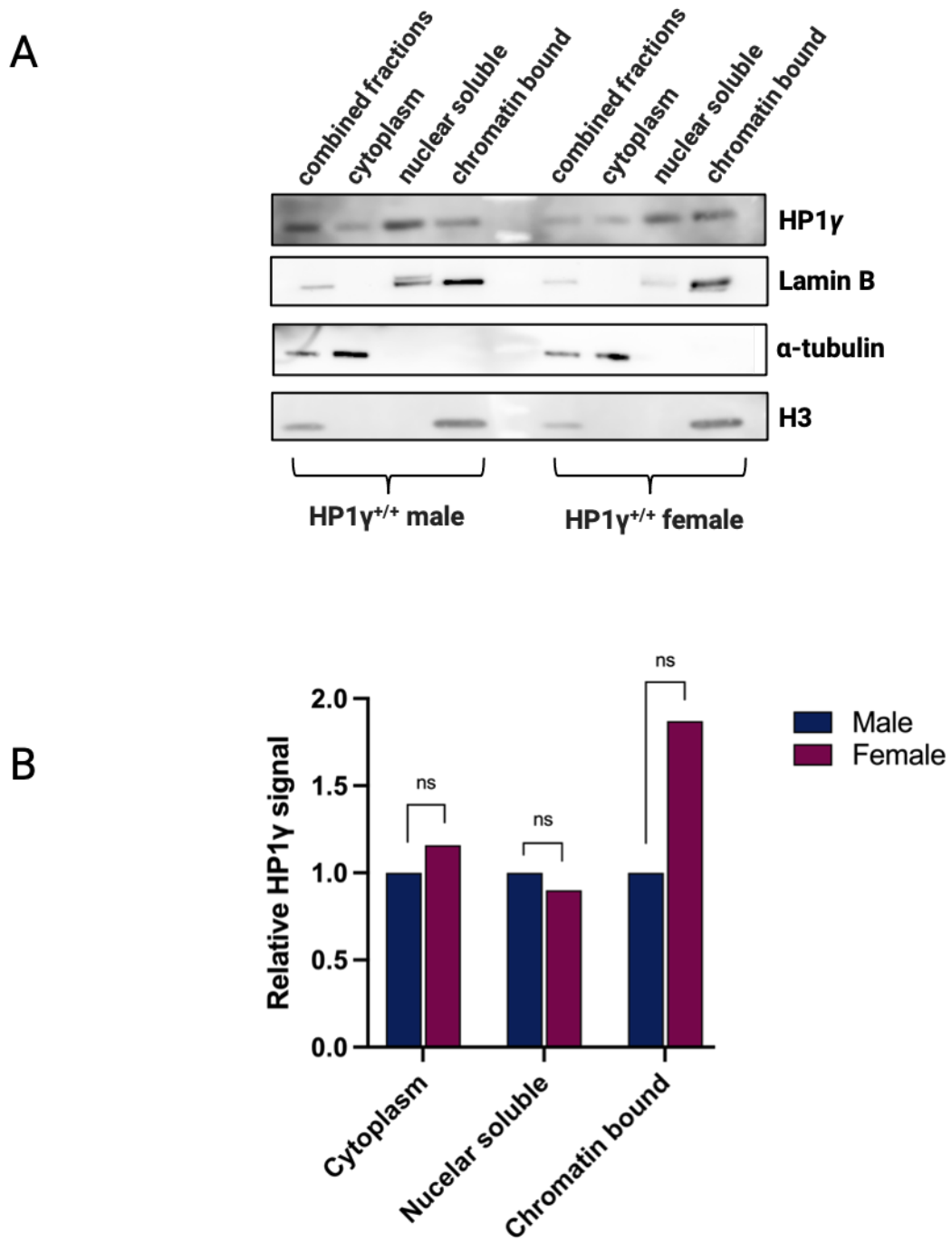


Figure 3.31 HP1γ enrichment in the chromatin-bound fraction in females by WB. (A) Representative image of WB, following the cell fractionation assay. 5 µg of proteins were loaded. (B) Each bar represents the average normalised signal of HP1γ from three HP1γ$^{+/+}$ male and female biological replicates. α-Tubulin, Lamin B and Histone H3 were used for normalisation of the cytoplasmic, the nuclear and the chromatin fraction respectively. Statistical significance was tested with Student's t-test. ns= non-significant.

While WB is a widely used technique with many applications in the protein field, the quantitation power of this assay is limited. Moreover, due to the nature of the cell fractionation assay, small pipetting errors can lead to differences during the extraction of the fractions, rendering the comparison by WB semi-quantitative. For this reason, we decided to repeat the cell fractionation assay with five biological replicates of male and female HP1γ$^{+/+}$ MEFs and exploit the sensitivity of the liquid chromatography–mass spectrometry (LC/MS) approach to observe subtle changes in composition of the different subcellular fractions that are unlikely to be detected by western blot.

The LC/MS samples were subjected to normalisation using the MaxLFQ algorithm that allows for label-free quantification (LFQ) of proteins and importantly normalisation of proteins that belong to the same subcellular fraction and not to total protein levels (Cox *et al.*, 2014). LFQ intensity corresponds to the average intensity of a protein, which is computed by summing the intensity signal of its identified peptides. During the analysis, fraction identifiers for sub-cellular samples were specified. For example, chromatin-bound and nuclear-soluble samples were assigned as adjacent fractions, while whole cell extract, cytoplasmic and pool samples were specified as adjacent fractions. This is to avoid the transfer of peptide identifications between samples, where there may not be a substantial overlap of the proteome. The normalised protein intensities of the different fractions, alongside other metrics like numbers of peptides is reported in Supplementary Table S3. The “pool” group refers to a positive control that was included to ensure non-erroneous LC/MS running. As we aimed to identify subtle changes among the different fractions of the two sexes, we specifically looked at the Y-linked Eif2s3y protein to ensure that the sex allocation post LC/MS was correct. This factor was enriched only in male fractions verifying the sex allocation (Supplementary Figure S4). As an additional control, we also looked at Lamin B. As expected, Lamin B was enriched only in the nuclear fractions (nuclear soluble and nuclear-bound) while it was absent from the cytoplasmic fraction (Supplementary Figure S5), similar to the WB analysis. Pearson’s correlation matrixes of raw and normalised protein intensities also reveal that the different subcellular fractions correlate closely with each other for both sexes (Figure 3.32), further supported by PCA plots for protein intensities (Figure 3.33).

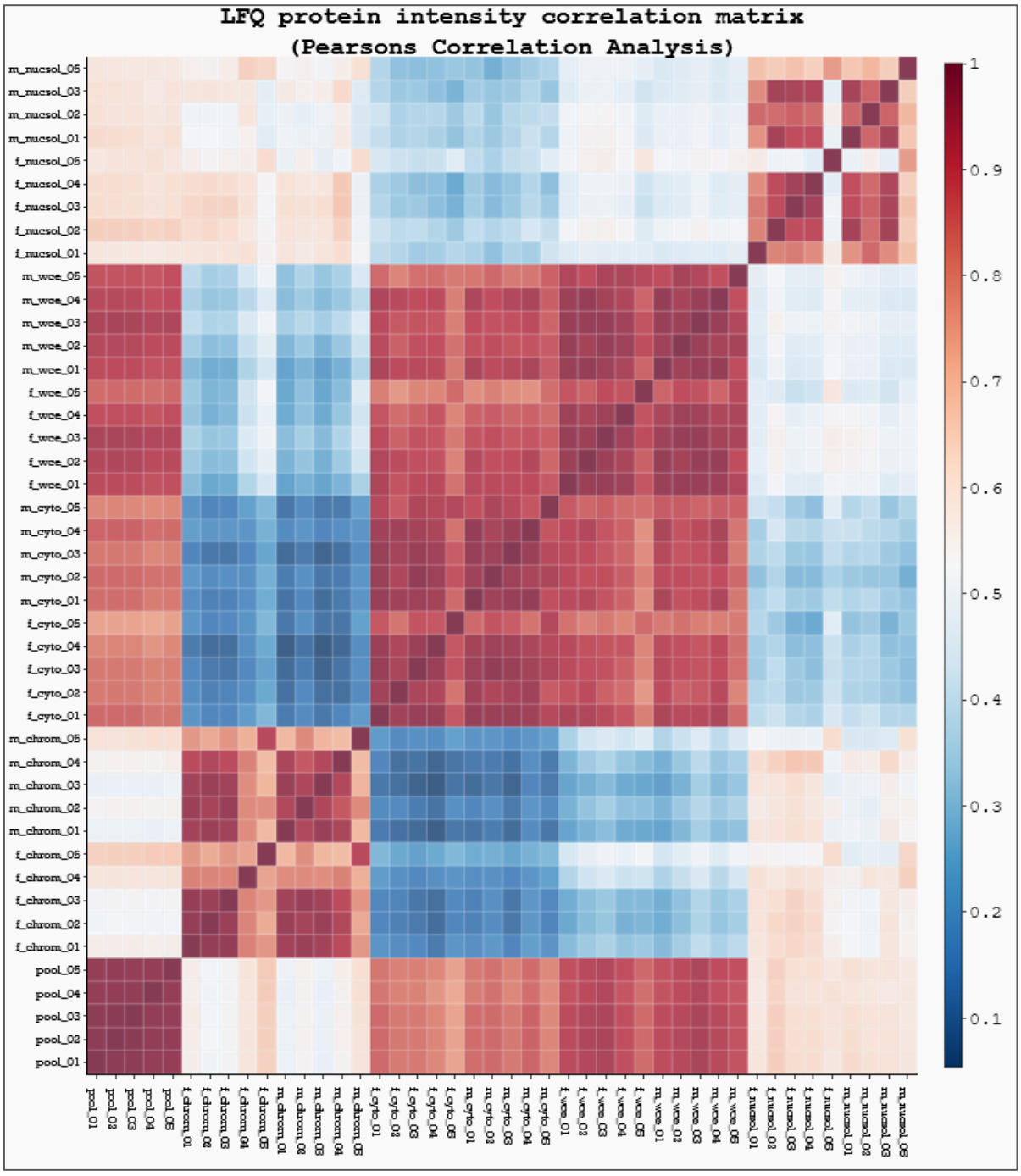
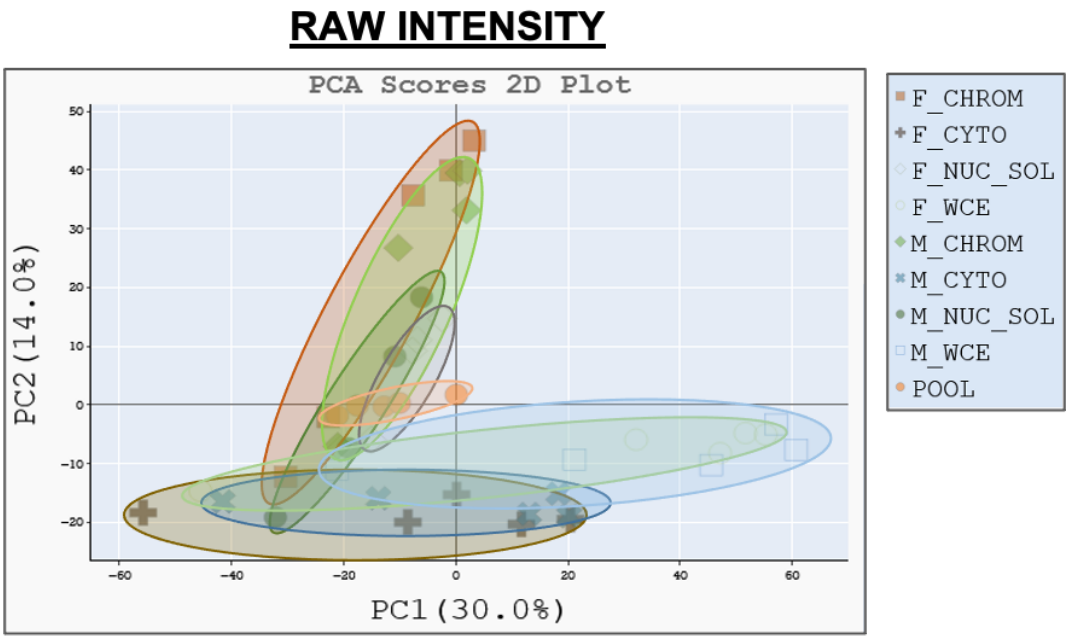


Figure 3.32 High correlation among the subcellular fractions. Pearson's correlation matrixes of normalised protein intensities of the different subcellular fractions of male and female MEFs. Greater enrichment is depicted in red and lower enrichment in blue.

A



B

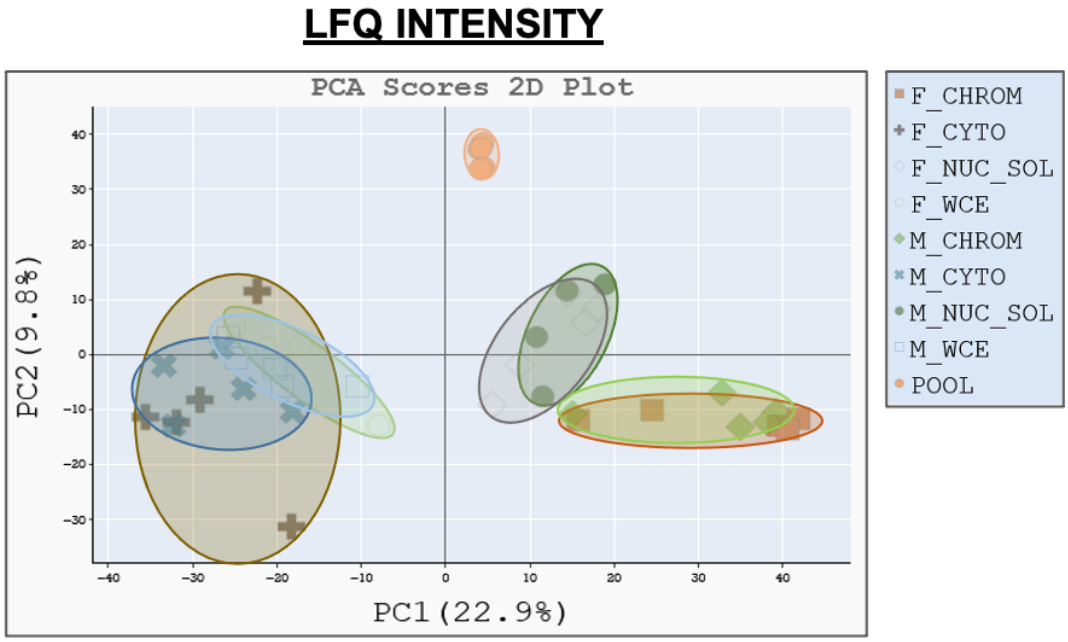


Figure 3.33 Male and female subcellular fractions cluster well after normalisation. PCA plots of male and female (A) raw and (B) normalised protein intensities of the different subcellular fractions. The first two principal components (PC1 and PC2) are plotted and coloured according to the corresponding fraction. Percentage of variation for each principal component is shown.

Therefore, we looked specifically at the distribution of HP1γ among the different subcellular fractions. The nuclear-soluble fraction of HP1γ shows a minor female-biased enrichment, but high variability among the male samples is also observed for this fraction. Contradictory to the previous analysis by WB, HP1γ is slightly more enriched in the male chromatin-bound fraction, while similarly to the WB, HP1γ levels are very similar in the cytoplasm for both sexes (Figure 3.34).

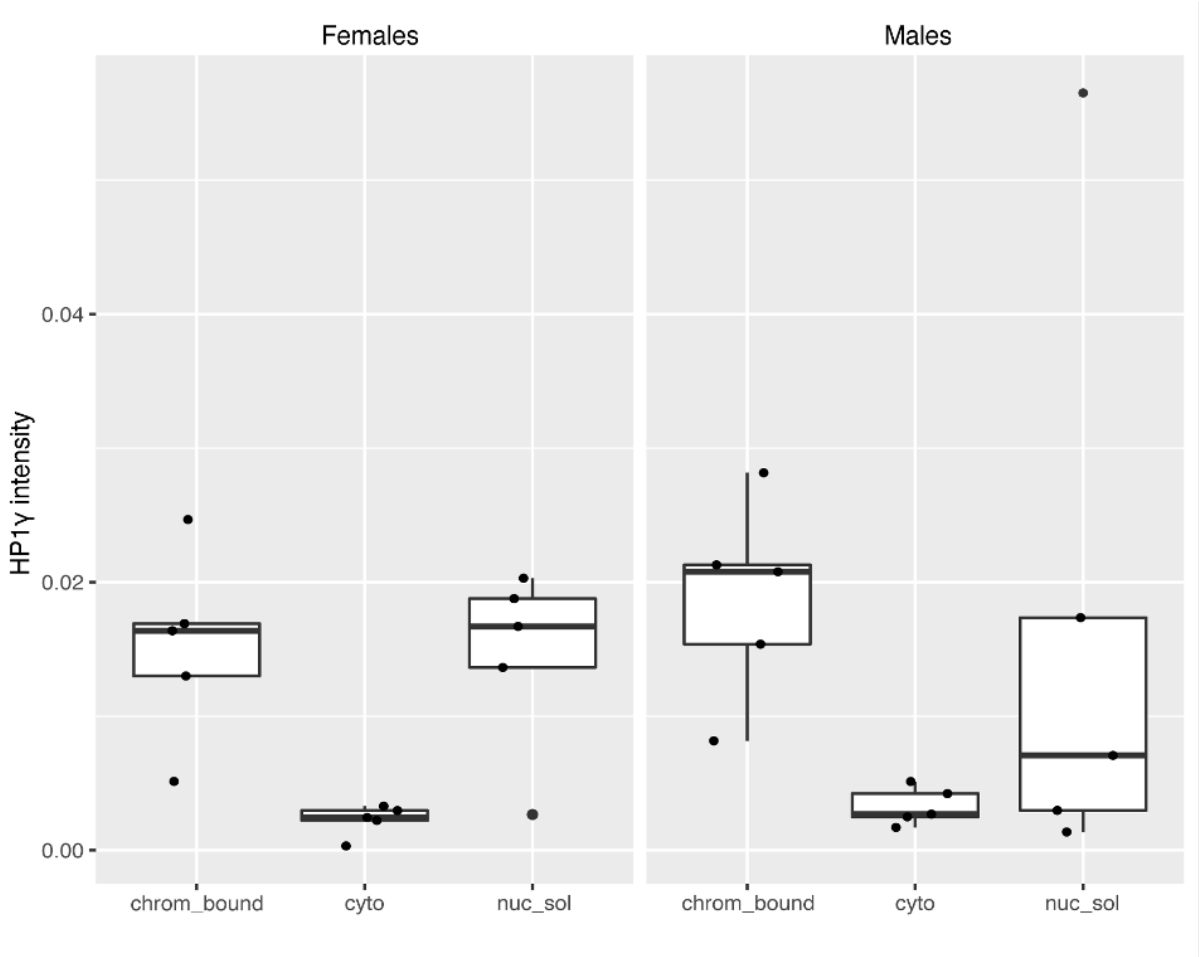


Figure 3.34 HP1γ subcellular distribution is similar between the sexes by LC/MS. Each dot represents a biological replicate with the floating bars showing interquantile range and the black line indicating the median.

As a sex-biased enrichment of chromatin-bound HP1γ was observed by WB, we directed our analysis on male and female chromatin-bound-only factors, which revealed that none is significantly enriched in either sex (FDR = 0.05) (Figure 3.35).

Our analysis suggests for high similarities at the levels of chromatin-bound proteins among the sexes at E13.5 MEFs.

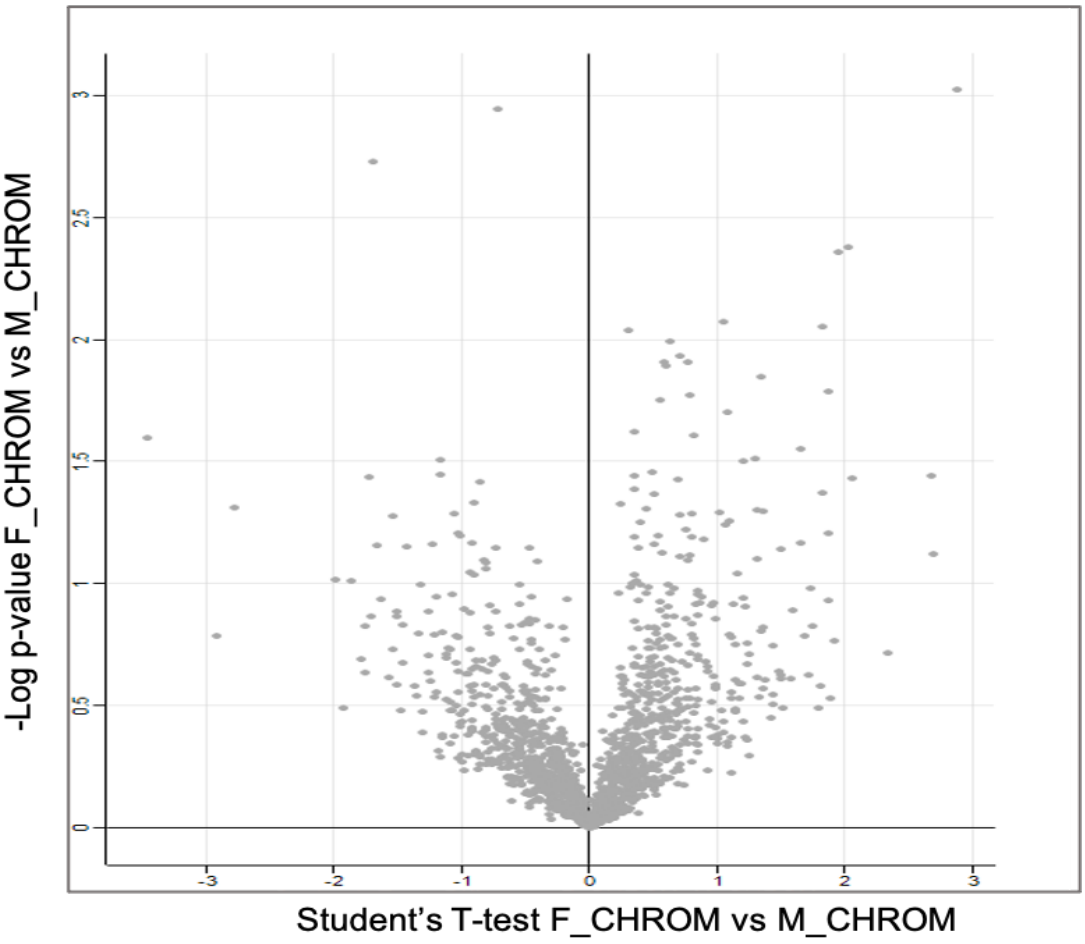


Figure 3.35 No sex-biased protein enrichment is observed in the chromatin-bound fraction by LC/MS. Comparison of 1,543 female and male chromatin-bound proteins (merged, n=5), visualised as volcano plot. FDR=0.05.

3.3 Discussion

The 2015 guidelines from the National Institutes of Health urge that sex should be considered in research designs and reporting in vertebrate animal studies (NOT-OD-15-102). In mammals, adult mortality tends to be higher in males and it has been suggested that this can occur because of the unguarded expression of deleterious recessive alleles residing on the sex chromosomes (Liker and Székely, 2005). A timely example of sexual dimorphism in disease is COVID-19, with the increased severity of this infection in males (Nguyen *et al.*, 2021). Elucidating the sex-biased patterns of gene regulation is crucial not only for understanding how sex-specific biological processes play a role in health and disease but also for the development of precision therapeutics.

HP1γ is an isoform of the HP1 family of proteins that is distinguished by its localisation pattern in both euchromatin and heterochromatin and has been suggested to have both an activating and a repressive role on gene expression (see Section 1.1.11). Conditional knockout of HP1 preferentially reduces male viability in *Drosophila*, while microarray analysis showed that HP1-regulated genes in males are almost twice as many as in females (Liu *et al.*, 2005). HP1γ has also been implicated in the regulation of sex-biased gene expression in mouse, where the effect of HP1γ depletion on gene expression was noticeably greater in males than females, suggesting that the male genome is more dependent on HP1γ to fine-tune genome-wide transcription (Law et al., 2019). How HP1γ exerts its role on sexual dimorphism and what the implications are of the sex differences in physiology and/or disease is poorly understood.

The work described in this Chapter, shines light on the role of the epigenetic modifier HP1γ on sexual dimorphism. The molecular mechanisms of how HP1γ exerts its role have been investigated early in mouse development, with the importance of HP1γ on defining sex differences being reflected both *ex vivo* and *in vivo*.

This study utilised an HP1γ knockout mouse model where no functional mRNA transcript is produced to ensure complete removal of HP1γ protein. Since we are

interested in differences of gene expression that can be very similar among the sexes, incomplete deletion of HP1γ could confound the results. If an HP1γ knockdown for example was utilised, some of the sex dimorphic genes may not be affected, as the low levels of HP1γ present in the knockdown may still be able to establish and/or maintain their expression levels. At the same time, by utilising MEFs coming from E13.5 embryos to study the sex-specific effects of HP1γ, we can presume that the observed sex differences are most likely due to sex chromosome complement effects rather than sex hormones as the latter are produced in significant quantities only later in mouse development (Wijchers and Festenstein, 2011).

The sex-biased differences of physiological functions that are dependent upon HP1γ give rise to discrepancies that are evident at the cellular level. It has been shown that male MEFs have a proliferation advantage compared to female MEFs and removal of HP1γ led to a reduction of the proliferation rate of male MEFs comparable to that found in female MEFs, while the proliferation rate of female MEFs remained largely unaffected by perturbation of HP1γ (Law *et al.*, 2019).

One of the first observations while culturing these cells was a reduction in the proliferation rate of MEFs lacking HP1γ. To investigate this phenotype further I employed BrdU staining which revealed that male MEFs are sensitive to HP1γ loss with a significant replication retardation (Figure 3.4), while similar to Law et al., 2019, the replication rate of female MEFs remained unaffected by the loss of HP1γ. This suggests that in the absence of HP1γ, male cells are not progressing efficiently through the S phase and may be accumulating at G1. A hypothesis supported by Propidium iodide (PI) staining followed by Fluorescence-activated cell sorting (FACS) analysis performed previously in Festenstein's group that showed a small increase of $HP1\gamma^{-/-}$ MEFs in G1 phase compared to $HP1\gamma^{+/+}$ MEFs (Liang PhD Thesis, 2019).

One of the HP1γ-sensitive sexually dimorphic genes is *Cdkn2a* (Rayess, Wang and Srivatsan, 2012; Law *et al.*, 2019), which in mice encodes two master regulators of the cell cycle, p16 and p19ARF. Both proteins can inhibit the G1-S phase transition by controlling CDK and pRB levels, ultimately leading to cell senescence (Rayess, Wang and Srivatsan, 2012). A male-specific downregulation of *Cdk4* and an upregulation of

Rb1 was also observed in the absence of HP1γ (Law *et al.*, 2019), further supporting the hypothesis that this pathway may be involved in the male-biased proliferation differences. In Section 3.2.2 I showed that loss of HP1γ leads to a male-specific upregulation of *Cdkn2a*. Examining p16 with immunofluorescence experiments revealed that the protein levels are not affected by the loss of HP1γ for either sex. On one hand, this could be arising due to technical limitations, where the sensitivity of the assay does not allow to distinguish between small differences at the protein levels. p16 could be further examined by a more sensitive method like Enzyme-linked immunosorbent assay (ELISA). On the other hand, rather than p16, p19ARF might be affected by the depletion of HP1γ, which could be responsible for the slower replication observed in male HP1γ$^{-/-}$ MEFs. p19ARF protein levels could be tested by ELISA or WB.

The sexually dimorphic proliferation discrepancies observed in the MEFs could be arising due to accelerated senescence in the absence of HP1γ. Cellular senescence refers to the state of irreversible cell growth arrest without loss of cell viability. Cellular senescence refers to the state of irreversible cell growth arrest without loss of cell viability. Since *Glb1*, the gene encoding β-galactosidase (β-gal) is not a sexually dimorphic gene (Law *et al.*, 2019), I examined the activity of β-gal, a well characterised senescent marker. This analysis revealed that upon depletion of HP1γ cells senesce faster with females being more affected in P3 than males. At later passages, HP1γ$^{-/-}$ male MEFs show similar levels of senescence to their female equivalents (Figure 3.7). To further support this finding, other commonly used senescence markers could be tested, including p53 and/or p21 (Hernandez-Segura, Nehme and Demaria, 2018). Intriguingly, human HP1 proteins are enriched in senescence associated heterochromatin foci (SAHF), which are domains of compacted facultative heterochromatin that form in senescent cells (Aird and Zhang, 2013). SAHF is believed to contribute to senescence by repressing the expression of cell cycle promoting genes such as cyclin A and silencing of rRNA synthesis. HP1 proteins, are transiently localized to Promyelocytic leukaemia protein (PML) bodies prior to incorporation into SAHF and HP1γ becomes phosphorylated on serine 93 for efficient SAHF deposition (Zhang et al., 2007). Our observation from cells lacking HP1γ seems to be at odds with this previously suggested role of SAHF in senescence, nonetheless, a recent

study demonstrated that SAHF do not form in mouse fibroblasts (Kennedy *et al.*, 2010), hinting towards a different role of chromatin in human and mouse senescence.

Burgoyne et al showed that sex differences in growth rate are more evident in early embryonic development, with the male and female embryo size converging closer to birth (Burgoyne *et al.*, 1995). Therefore, sex differences might be identified if one looked earlier in development. Apart from the body size, these differences could include, variation of the somite numbers, which also acts as an indicator of the embryonic stage (Theiler, 1989; Downs and Davies, 1993). Mouse embryo measurements reveal that at E13.5, male embryo growth is dependent on HP1γ as its deletion led to a significant body weight reduction, not observed in females (Figure 3.8). The slower proliferation of male MEFs in combination with the earlier onset of senescence could be one of the reasons why the male embryos present growth retardation in the absence of HP1γ. Naturally, the embryo body is composed of different cell populations that could also be sensitive to HP1γ deletion, including but not limited to epithelial and endothelial cells. Heart, brain and liver are well developed tissues by E13.5. How the size or the proliferation rate of cells comprising these tissues is contributing to the growth differences has not been investigated, since these tissues are discarded during the generation of MEFs (Tan and Lei, 2019). Magnetic resonance histology (Petiet *et al.*, 2008) and immunohistochemistry of whole-mount HP1γ$^{+/+}$ and HP1γ$^{-/-}$ embryos (Wong, 2021) could shine light upon how different cells are affected by HP1γ in an *in vivo* setting.

To interrogate the molecular mechanism driving HP1γ-dependent sexual dimorphism I performed CUT & RUN experiments using E13.5 MEFs. CUT & RUN allows the elucidation of the chromatin binding sites of a protein of interest and in combination with transcriptomic data, one can gain better understanding on how a factor might regulate gene expression. Our analysis reveals that HP1γ is enriched at the promoter of genes with high levels of expression for both sexes, in agreement with other studies (Smallwood *et al.*, 2012). A dip of HP1γ signal is particularly evident right before the TSS, which could be explained by the unique nature of the first nucleosome located upstream of the TSS (-1 nucleosome). This nucleosome covers a region from −300 to

−150 bp relative to the TSS, and can affect the accessibility of regulatory factors in that region. During transcription, the −1 nucleosome undergoes changes that affect its functionality, including histone replacement, swap of methylation for acetylation marks, repositioning, and ultimately eviction shortly after the pre-initiation complex (PIC) formation (Jiang & Pugh, 2009) resulting in the absence of a binding site for HP1γ.

Focused analysis on the sexually dimorphic genes (Law *et al.*, 2019) reveals that HP1γ binds more on the "male lower genes" compared to the "male higher genes", a phenomenon observed in both sexes (Figure 3.17). It is important to point out that the "male lower genes" are the genes that present a higher expression level in females, and respectively, the "male higher genes" are the genes with a lower expression in females. Therefore, it seems that HP1γ binding on sexually dimorphic genes in females is in agreement with the enrichment of HP1γ on genes with higher expression levels, unlike in males. Our data suggest for opposing roles of HP1γ regarding the regulation of this set of genes, where in males it appears to have a suppressive role, while in females an activating one which further illustrates the contrasting role of HP1γ, acting both as a suppressor and an activator of gene expression in a gene-specific context. HP1γ is not the only factor known to have contrasting roles regarding gene regulation. CTCF for instance, was first described acting as a suppressor of the chicken *c-myc* gene (Klenova *et al.*, 1993), while it was shown that it can act as an activator in the case of the amyloid β-protein precursor (APP) gene (Vostrov and Quitschke, 1997). Similar to CTCF, PTMs and/or binding partners likely enable HP1γ to act either as an activator or suppressor (Kim et al., 2015).

Improving the HP1γ CUT & RUN signal-to-noise ratio could allow us to pinpoint the HP1γ localisation on the gene region which leads to sexually dimorphic transcription regulation. For example, HP1γ's binding near the TES of the sexual dimorphic genes could be leading to an R-loop-induced premature termination resulting in differential gene expression (Skourti-Stathaki, Kamieniarz-Gdula and Proudfoot, 2014). Investigations into the DNA sequences of the sexually dimorphic genes, can further reveal how HP1γ is recruited to these sites. Indeed, work from H. Shepherd, showed an enrichment at the promoters of sexually dimorphic genes for KRAB zinc finger proteins DNA motifs (Shepherd MSc Thesis, 2018). These motifs and their associated

factors could be targeting HP1γ to these promoters and specifically regulate the gene expression in a sex-specific manner. The low signal-to-noise ratio could also explain the lack of HP1γ enrichment over intronic regions (see Figure 3.11 B), which is similar to the negative control (IgG).

HP1γ has been shown to interact with both the initiating and the elongating forms of RNAPII (Vakoc *et al.*, 2005; Lomberk *et al.*, 2006; Mateescu *et al.*, 2008), while it has also been suggested to stabilise ongoing transcription by regulating the rate of RNAP II elongation (Saint-André *et al.*, 2011). Another possibility of why HP1γ is binding on the promoter and gene bodies of highly expressed genes (Figure 3.14) is to supress initiation of cryptic transcription. In yeast, H3K36me3 has been associated with the inhibition of cryptic transcription initiation (Carrozza *et al.*, 2005; Venkatesh *et al.*, 2016), hence high-resolution binding profiles of HP1γ and H3K36me3 in combination with 5'-centered transcriptomic analysis can elucidate the potential HP1γ involvement in this mechanism. To examine this further, we performed SLIC-CAGE experiments that allows mapping of TSSs at bp resolution (Cvetesic *et al.*, 2018). Our preliminary analysis reveals that, irrespective of the sex, HP1γ deficiency did not lead to an alteration of basic promoter architectural features. HP1γ loss did not affect gene promoter width and, as expected, TATA-boxes mapped in "sharp" promoters that normally have a single dominant TSS (Ponjavic *et al.*, 2006). Notably, in the absence of HP1γ, usage of multiple TSSs, that could mark spurious transcription, was not observed (Figure 3.27 and Figure 3.28). Moreover, differential usage of TSSs (also known as promoter shifting) only occurred in a handful of genes in HP1γ$^{-/-}$ samples (Supplementary Tables S1 & S2). Since none of the affected promoters belonged to sexually dimorphic genes, it strongly suggests against the usage of this mechanism by HP1γ for the regulation of their sex-biased expression. Collectively, the preliminary SLIC-CAGE results argue against the inhibition of cryptic transcription initiation by HP1γ, as no major differences at the TSS landscape is observed between HP1γ$^{+/+}$ and HP1γ$^{-/-}$ MEFs. However, it is important to note, that due to the nature of the SLIC-CAGE assay, only transcripts produced by RNAPII with a 5'- m^{7}G cap can be identified. Loss of HP1γ could lead to de-repression of intergenic chromatin and spurious transcription of repetitive elements, not necessarily having a m^{7}G cap and hence not captured by SLIC-CAGE. The vast majority of Alu elements for example, are

transcribed by RNA polymerase III and if a dysregulation of their expression was to be occurring by HP1γ deletion, this would not be captured by SLIC-CAGE (Zhang, Gingeras and Weng, 2019). To study non-RNAPII transcripts that could be regulated by HP1γ and contributing to sexual dimorphism, RNA Annotation and Mapping of Promoters for the Analysis of Gene Expression (RAMPAGE) assay could be employed (Batut et al., 2013).

The higher total HP1γ signal observed by CUT & RUN in wild-type females (Figure 3.14), prompted me to investigate the subnuclear distribution of HP1γ by performing a cell fractionation assay. WB analysis of the different cell fractions revealed an enrichment of the chromatin-bound fraction of HP1γ in females compared to males (Figure 3.21) but this finding was not recapitulated by LC/MS analysis. Despite the high sensitivity of the LC/MS, the "peptide-to-protein" approach renders the distinction of highly similar isoforms problematic and can yield inconsistent results in the determination of protein levels, especially in a label-free quantification (Ezkurdia et al., 2012). As HP1α and HP1γ percentage protein identity is 62%, and HP1β and HP1γ percentage protein identity is 74% (BLASTP 2.13.0, Altschul et al., 1997), even partial miss-assignment of the peptides to the corresponding isoforms could have confounded the results. Stable isotope labelling approaches like SILAC, might overcome this limitation (Ong *et al.*, 2002).

There are several ways in which the sex chromosome complement effect might be contributing to the sexually dimorphic effect of HP1γ, described in this Chapter. Firstly, the Y chromosome carries male specific genes some of which are known chromatin modifiers like *Kdm5d*. KDM5D demethylates the active modification H3K4me3 and would hence be a good candidate for acting as a transcriptional repressor, strengthened by the observation that knockdown of *Kdm5d* upregulates major satellite repeats in mouse T cells (Silva MRes Thesis, 2015). It is therefore possible that a synergistic effect of HP1γ and KDM5D exists to repress genes that are expressed in lower levels in males. Secondly, one or more of the "X chromosome escapee" genes (Carrel and Willard, 2005) might be limiting the repressive effect of HP1γ on "male lower genes" in females. In mouse, around 3–7% of X-linked genes escape from XCI (Balaton and Brown, 2016). The genetic interaction of HP1γ and Y

or X-linked factors could be investigated in the future by utilising CRISPR/Cas9 targeted deletion or transgenic overexpression for functional studies in HP1γ$^{+/+}$ and HP1γ$^{-/-}$ MEFs. Thirdly, it is possible that one of the sex chromosomes acts as a heterochromatic "sink". For instance, *Drosophila* males contain a large heterochromatic Y chromosome that alters the stoichiometric balance of euchromatin/heterochromatin between the sexes. Hence, the Y chromosome may be acting as a "sink" for sex-specific epigenetic modifications resulting in a dimorphic expression of genes (Deng et al., 2009; Lemos et al., 2010; Zhou & Bachtrog, 2012). In mammals, the opposite could happen, where the inactive X chromosome sequesters silencing proteins and females could have lower levels of heterochromatin-like features in the rest of their genome as compared with males. Therefore, HP1γ is sequestered to the inactive X (Xi) in females, thus limiting its availability in other genomic regions, especially considering that HP1γ localises on the Xi in humans (Chadwick and Willard, 2003) and H3K9me3 extensively decorates it (Heard *et al.*, 2001). To distinguish between effects that arise from the X chromosome "sink" compared to the effects of "X chromosome escapees", XY cells that carry an autosomal *Xist* transgene (Loda *et al.*, 2017) could be employed. The Xist would provide a potential autosomal "sink", while the lack of XCI will not give rise to escapee genes.

Finally, it is important to note that while the sexually dimorphic genes regulated by mouse HP1γ are likely to be species specific, the epigenetic mechanism underlying sexual dimorphism is likely to be conserved between mice and humans. shRNA-mediated HP1γ knockdown can be utilised to test whether the effect of HP1γ in regulating sexually dimorphic genes is conserved in humans, which can further improve our understanding of physiological sex differences and sex bias in disease.

Chapter 4. The role of HP1γ in telomere maintenance

4.1. Introduction

4.1.1 The telomeric landscape and its function

Telomeres are repetitive sequences at the end of chromosomes necessary for genomic stability. The inability of DNA polymerases to fully replicate the linear ends of chromosomes could lead to a gradual loss of necessary genetic information. Telomeres, with their repetitive nature prevent this from happening (Olovnikov, 1973). Moreover, telomeres protect the genome by inhibiting the DNA damage repair (DDR) machinery that would normally recognize chromosome ends as DNA lesions and in the attempt of repairing these lesions, could give rise to detrimental chromosomal fusion events. Various mechanisms appear to be in place for telomeres to avoid being recognized as DNA damage sites, including the actions of the shelterin complex (de Lange, 2005). Shelterin acts as a DDR inhibitor by masking telomeres from being sensed as double-strand breaks (DSBs). POT1 and TRF2 shelterin subunits are mainly attributed with this function (Hurley, Wilsker and Bunz, 2006; Okamoto *et al.*, 2013), while TRF1, another shelterin subunit, by ensuring proper telomere replication provides an extra layer of DDR protection (Martínez *et al.*, 2009; Sfeir *et al.*, 2009). TRF1 appears to also affect the levels of RNA Polymerase II (RNAPII)-transcribed noncoding RNAs termed TElomeric Repeat containing RNAs (TERRAs) (Porro *et al.*, 2014; H. P. Chu *et al.*, 2017; Marión *et al.*, 2019; Porreca *et al.*, 2020). TERRAs are crucial for normal maintenance of telomeres by comprising an integral component of telomeric chromatin. In mice, TERRAs recruit the Polycomb complex (PRC2) to telomeres which facilitates the H3K27me3, leading further to the establishment of H3K9me3 and the recruitment of HP1 (Montero et al., 2018), a phenomenon highly atypical, since H3K27me3 and H3K9me3 rarely overlap in the rest of the genome (Mikkelsen *et al.*, 2007; Kato, Takemoto and Shinkai, 2018). Neverhteless, the combination of high H3K9me3 density and HP1 occupancy leads to telomeric heterochromatinization

(Gauchier *et al.*, 2019). The importance of HP1 on telomeres is also highlighted by the fact that in humans, the HP1γ isoform interacts with the Tin2 subunit of shelterin, and assists with sister-telomere cohesion (Canudas *et al.*, 2011). While all three HP1 isoforms are associated with heterochromatin (Minc *et al.*, 1999) , HP1γ is also found enriched on transcribing regions of the genome (Smallwood *et al.*, 2012), suggesting for a role in transcription regulation. This is enhanced by the fact that HP1γ interacts with RNA polymerase II (Vakoc *et al.*, 2005; Smallwood *et al.*, 2008; Kwon *et al.*, 2010) and widespread dysregulation of genes that differ among the sexes occurs, upon loss of this isoform (Law *et al.*, 2019). It is unclear if HP1γ's transcriptional role is important for telomere maintenance.

4.1.2. Rationale & aims

Telomeric chromatin is an important component of telomere biology. The synergistic effect of epigenetic modifications, epigenetic factors and telomere-associated proteins contributes to the maintenance of these regions. The role of HP1γ at telomeres has not been investigated through the prism of its transcriptional effect on TERRAs and telomeric proteins. Moreover, little is known about sex differences that could be characterising telomeres at the epigenetic level. Hence, it is important to shed light on the interplay between TERRAs, HP1γ and HP1γ's mediated transcriptional regulation of telomere-associated factors, while taking into account how sex could be playing a role in this cross-talk. To this end, I used primary mouse embryonic fibroblasts (MEFs), which is an established model for telomere studies. The reason for this is two-fold. First, there is high conservation of molecular mechanisms underlying telomeric maintenance that is shared among mammals (Blasco, 2007). Second, the short length of human telomeres in primary somatic cells makes their study limited resulting in the need of establishing sublines of immortalised cells with extended telomeres (Takai *et al.*, 2010). The mouse telomeres are of sufficient length while providing a physiological system.

4.2 Results

4.2.1 HP1γ regulates the expression of TRF1 and other telomere-associated factors

Since HP1γ is involved in transcription regulation (Smallwood *et al.*, 2012) and telomere maintenance (Canudas *et al.*, 2011), we decided to look how HP1γ affects the expression of factors that are necessary for telomere stability, by taking advantage of the available transcriptomic data of MEFs lacking HP1γ (Law *et al.*, 2019).

GO term analysis reveals that biological processes associated with dysregulated genes upon HP1γ knockout, includes “chromosome organisation” (GO term: 0051276). One subcategory of this umbrella term is “telomere organization”. Closer investigation with differential gene expression analysis (padj < 0.05) reveals that especially in males, the absence of HP1γ leads to a widespread dysregulation of factors involved with telomere maintenance (Glousker and Lingner, 2021), including TRF1 of the shelterin complex and telomerase-associated Dyskerin (DKC1) (Figure 4.1), but also telomere-associated helicases RTEL1, BLM, FANCM and DNA repair proteins including RAD51 and its paralogs (Figure 4.2). Regardless of the sex, the majority of these factors are downregulated upon HP1γ knockout, suggesting that HP1γ has an activating role for these genes’ expression in wild-type conditions. A minority of factors like the exonuclease APOLLO, which are upregulated in the absence of HP1γ, suggest for a silencing role of HP1γ in both sexes.

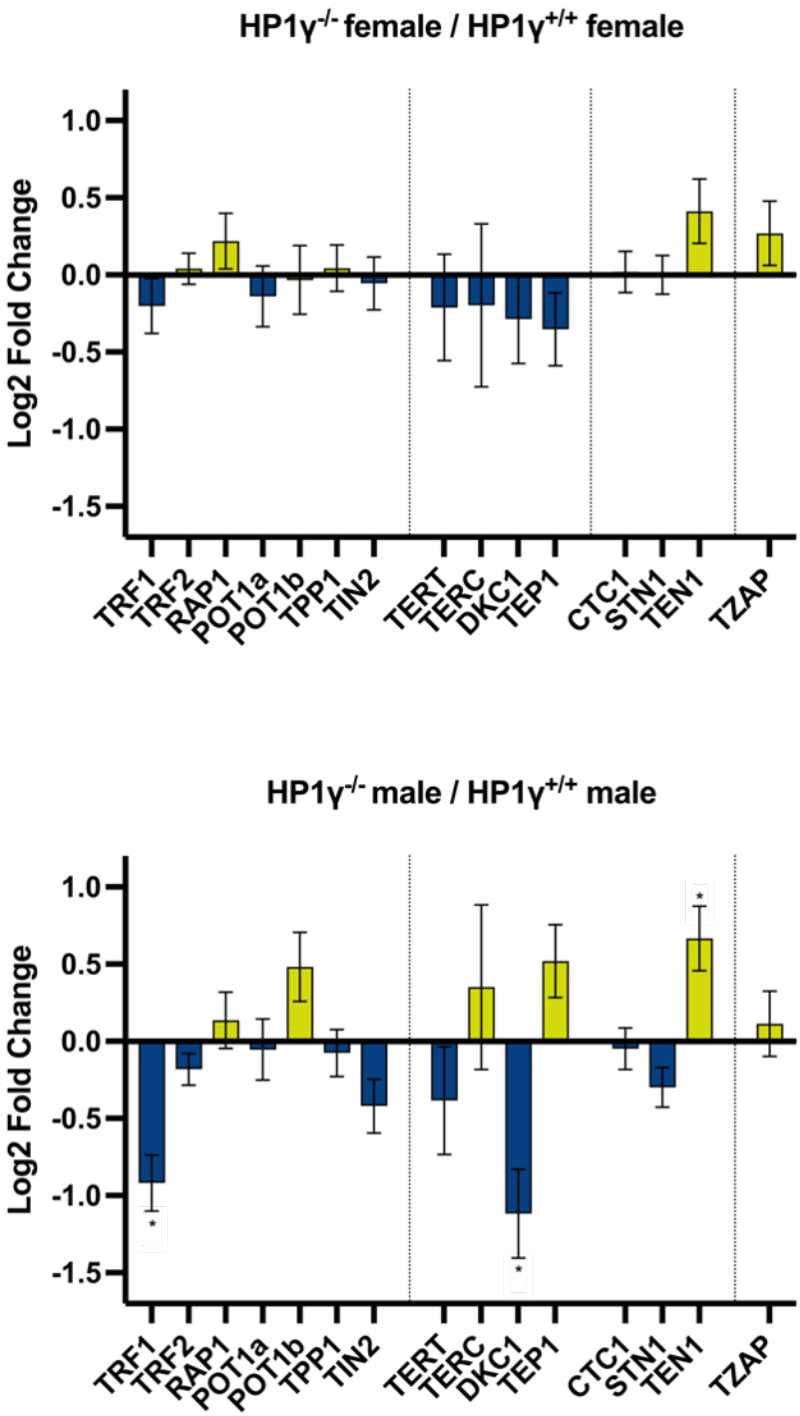


Figure 4.1 Changes in expression of factors that are necessary for telomere stability upon HP1γ depletion. Analysis of mouse shelterin genes (TRF1, TRF2, RAP1, POT1a, POT1b, TTP1, TIN2), telomerase-associated genes (TERT, TERC, DKC1, TEP1) CST complex (CTC1, STN1, TEN1) and TZAP. Bars represent the Mean log_2 fold changes ± SEM of three biological replicates for each genotype (Law *et al.,* 2019). * padj < 0.05.

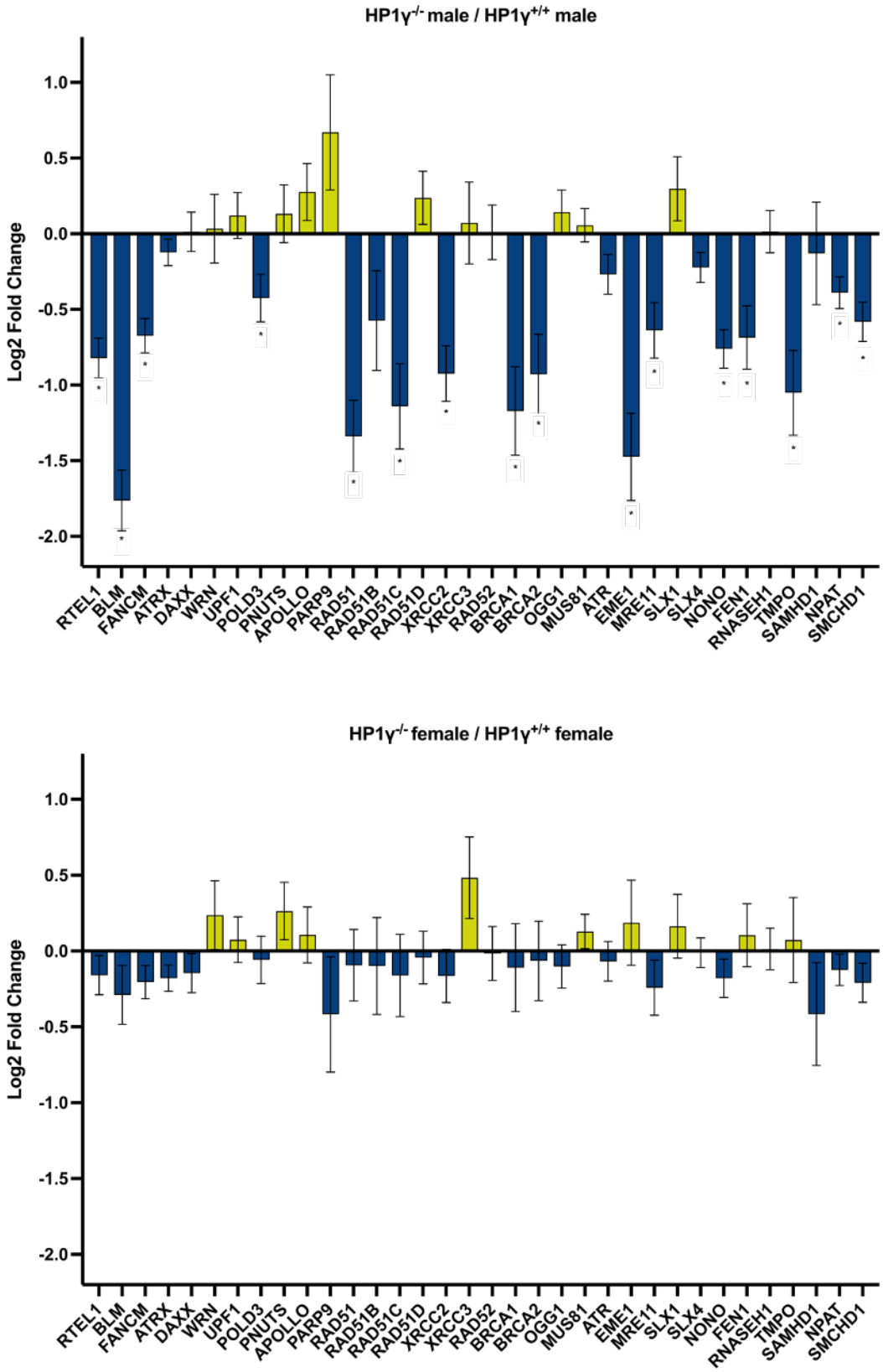


Figure 4.2 HP1γ loss results in widespread dysregulation of factors involved with telomere maintenance in both sexes. Bars represent the mean log_2 fold changes ± SEM of three biological replicates for each genotype (Law *et al.*, 2019). * padj < 0.05.

TRF1 is an essential telomere factor facilitating telomere replication (Martínez *et al.*, 2009; Sfeir *et al.*, 2009) which averts mouse telomere chromatin remodelling (Porreca *et al.*, 2020). HP1γ binding on the TRF1 locus and expression of TRF1 is similar between the sexes in wild-type conditions (Figure 4.3 A). However, downregulation of TRF1 transcripts was particularly evident in male HP1γ$^{-/-}$ cells, a finding which was further confirmed by WB for both males and females MEFs (Figure 4.3 B). Male-biased dysregulation of telomere-essential factors is in agreement with the widespread dysregulation of genes observed previously (Law *et al.*, 2019).

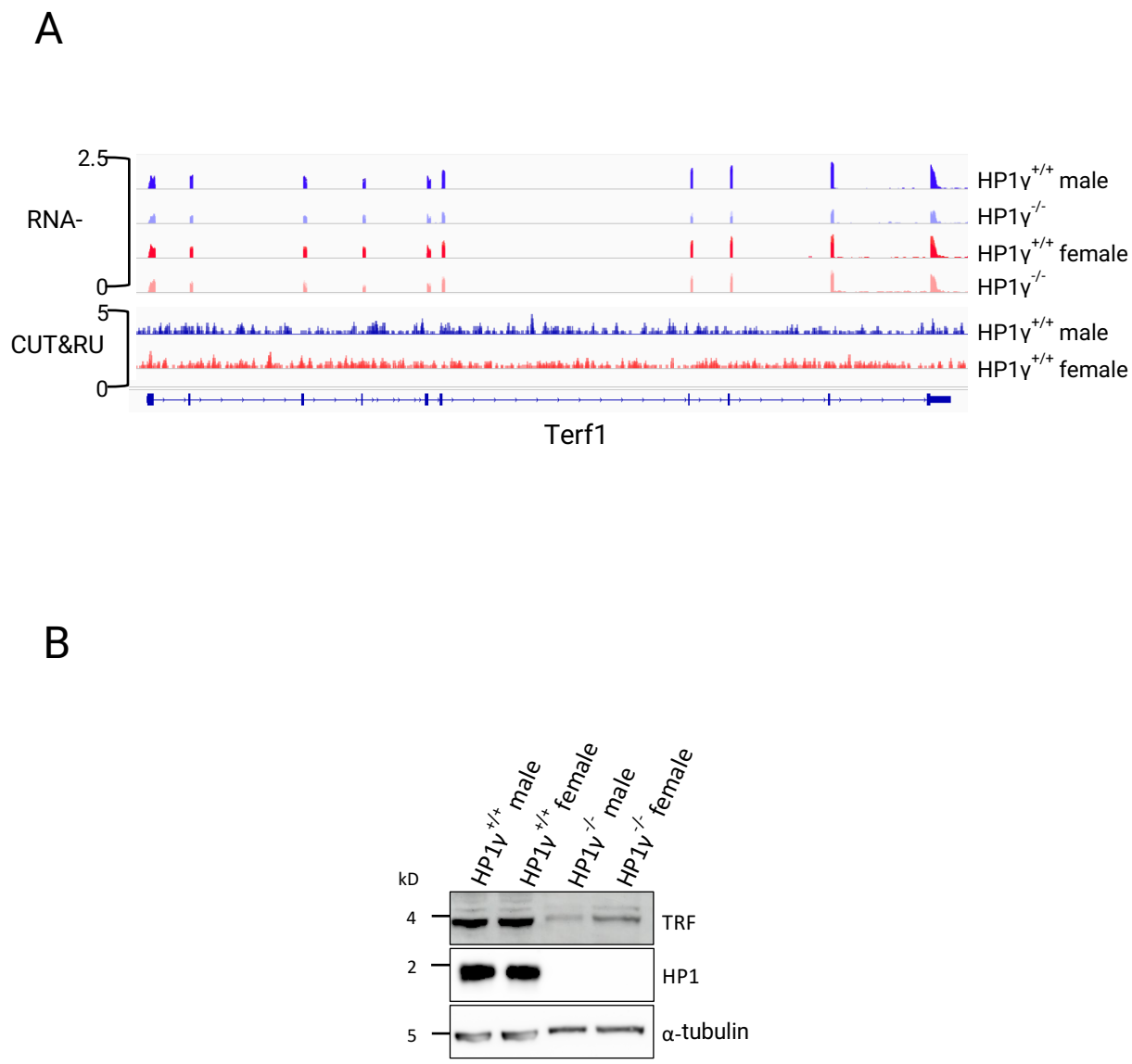


Figure 4.3 TRF1 transcript and protein levels decrease upon HP1γ knockout. (A) IGV genome browser snapshot of Terf1 gene (encoding TRF1, blue boxes indicate exons). RNA-seq (Law *et al.*, 2019) shows Terf1 expression: merged files, n=3 for all genotypes. CUT & RUN shows HP1γ binding on Terf1 locus: merged files, n=2 for both sexes. (B) Western blot of TRF1 and HP1γ (30 µg of protein). α-tubulin is used as a loading control.

4.2.2 HP1γ depletion results in higher TERRA levels

HP1γ's role on transcription regulation is still an open question, with some reports suggesting an activating role (Vakoc *et al.*, 2005; Lomberk *et al.*, 2006; Kwon *et al.*, 2010; Smallwood *et al.*, 2012), while others attributing a silencing character to it (Ostapcuk *et al.*, 2018). Remarkably, we observed a significant downregulation of TRF1 transcripts in male HP1γ$^{-/-}$ cells, and less of an effect in females but both male and female MEFs showed significant reduction in TRF1 proteins, to levels of a knockdown (Figure 4.3 B). Some of the known consequences of TRF1 dysregulation are increased TERRA levels (Porro *et al.*, 2014; Sadhukhan *et al.*, 2018)and telomere replication stress (Porreca *et al.*, 2020).

To gain better insight on the effect of HP1γ's depletion and the subsequent downregulation of TRF1 on TERRA expression, we utilised previously generated RNA-sequencing data (Law *et al.*, 2019). In mouse, the polyadenylated TERRAs are around 2.5-fold enriched compared to the non-polyadenylated fraction (de Silanes *et al.*, 2014). Since the sequencing library was generated by capturing total polyadenylated RNA, aligning an *in silico* TERRA probe to the sequencing reads, allowed us to identify the polyadenylated TERRA transcripts. Our analysis reveals that there are not major differences regarding this fraction of TERRAs among the wild-type sexes. Interestingly, upon loss of HP1γ, the polyadenylated TERRA transcripts almost doubles in males [1.8-fold difference] and more than triples in females [3.5-fold difference] (Figure 4.4).

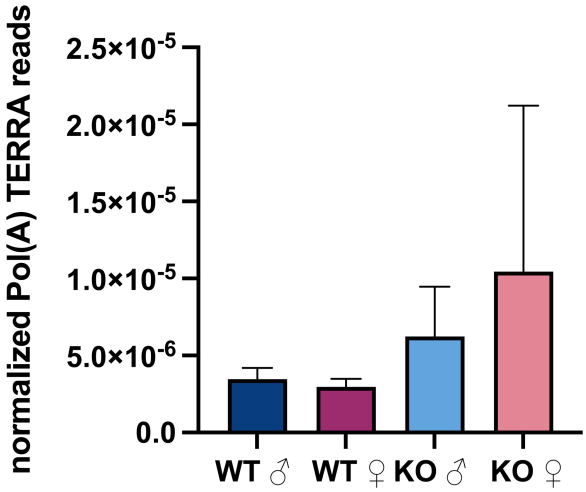


Figure 4.4 Depletion of HP1γ leads to more pol(A) transcripts, especially in females. Bars represent the mean of pol(A) TERRA normalized read counts of three biological replicates, from RNA-sequencing (Law *et al.,* 2019). Error bars indicate SD. No statistical significance as tested with one-way ANOVA and Tukey's *post-hoc* test.

To further characterise the effects of HP1γ and TRF1 dysregulation on TERRA levels, RNA dot blot and Northern blot analyses were performed using total RNA from E13.5 MEFs. During the RNA dot blot, RNA was transferred on a positively charged membrane which then was probed with a telomere-specific C-rich probe. RNase A treated samples served as negative control to indicate potential binding of the telomeric C-rich probe on undigested telomeric DNA instead of RNA species. An 18s rRNA probe was also used for normalization of the signal. In agreement with the RNA-sequencing analysis, RNA dot blot revealed an increase of RNA molecules containing telomeric repeats in both male and female HP1γ$^{-/-}$ conditions (Figure 4.5).

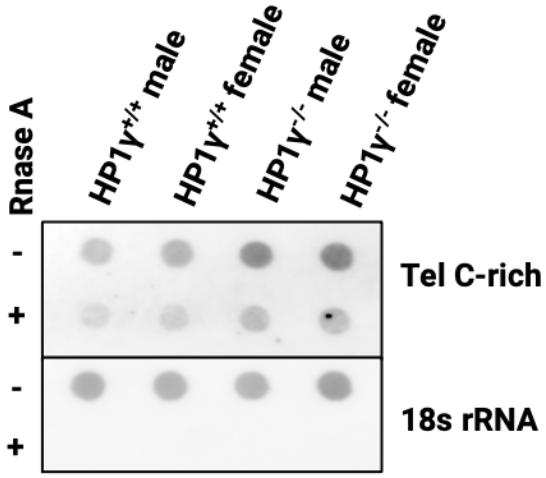


Figure 4. 5 HP1γ loss leads to elevated TERRA levels. RNA dot blot analysis of total MEFs RNA (2 µg). The blot was revealed with a DIG-Tel-C-rich probe and an 18 s rRNA probe, which served as a loading control.

For Northern blot analysis, total RNA was separated by electrophoresis and then transferred on a Hybond N^+ membrane which was then probed with the C-rich telomeric probe, revealing RNA molecules containing telomeric repeats. Once again, RNase A treatment served as negative control. A main advantage of Northern blot over RNA dot blot analysis is that it also provides information regarding the size of the RNA molecules. Northern blot showed a sharp increase of high molecular weight TERRAs (over 9kb - black arrow), particularly evident in HP1$\gamma^{-/-}$ females, with HP1$\gamma^{-/-}$ males also showing an increase compared to their wild-type counterparts (Figure 4.6).

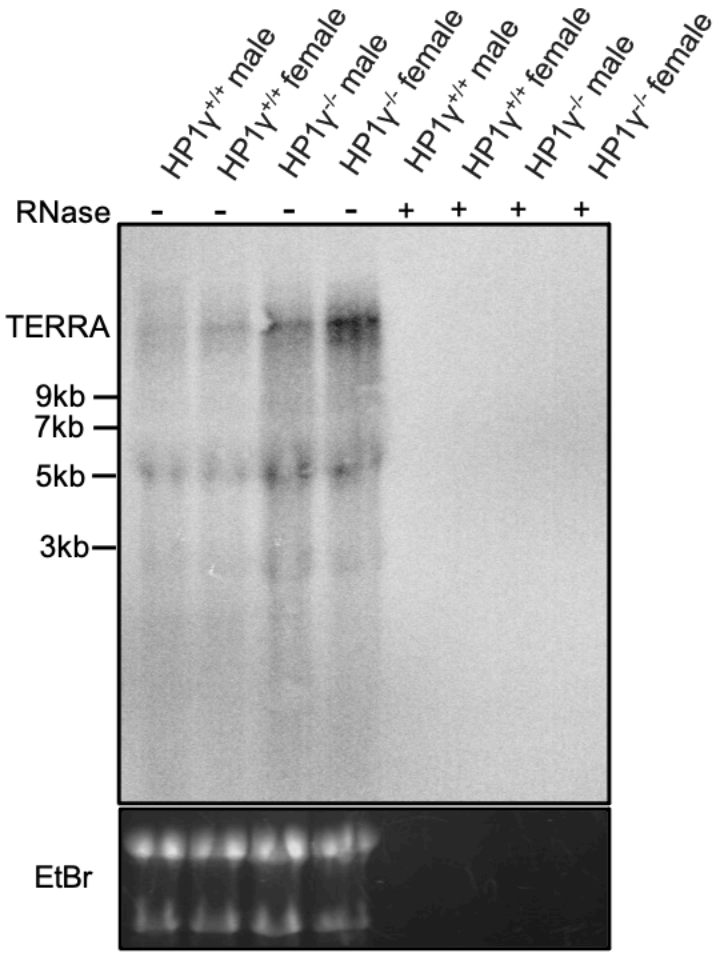


Figure 4.6 HP1γ loss leads to elevated TERRA levels, especially in females. Northern blotting of total MEFs RNA (10 μg) using a ^{32}P-labelled tel-C-rich probe for hybridization. RNase A treatment is used as a negative control, with ethidium bromide (EtBr) staining acting as a loading control.

In mice, the vast majority of TERRAs arise from the subtelomere of chromosome 18, while a smaller fraction may arise from other subtelomeric regions including those on chromosome 9, 10 and X (de Silanes *et al.*, 2014; Mazzolini *et al.*, 2018; Liu *et al.*, 2019). Utilizing specific primers for the subtelomeres of chromosomes 9, 10,18 and X, residing in immediate proximity to the telomeric repeats, we observed by RT-qPCR that TERRAs arising from chromosome 18 and chromosome X are significantly increased in female HP1γ$^{-/-}$ MEFs. Specifically, there is a 3.8-fold increase of chromosome 18 TERRA transcripts and a 3.2-fold increase of chromosome X TERRAs in females depleted of HP1γ (Figure 4.7). We could not detect a significant variation

of TERRAs arising from chromosome 18 and X subtelomeres for the males lacking HP1γ (Figure 4.7), while chromosome 9 and chromosome 10 TERRAs remained at undetectable levels, irrespective of the genotype (data not shown).

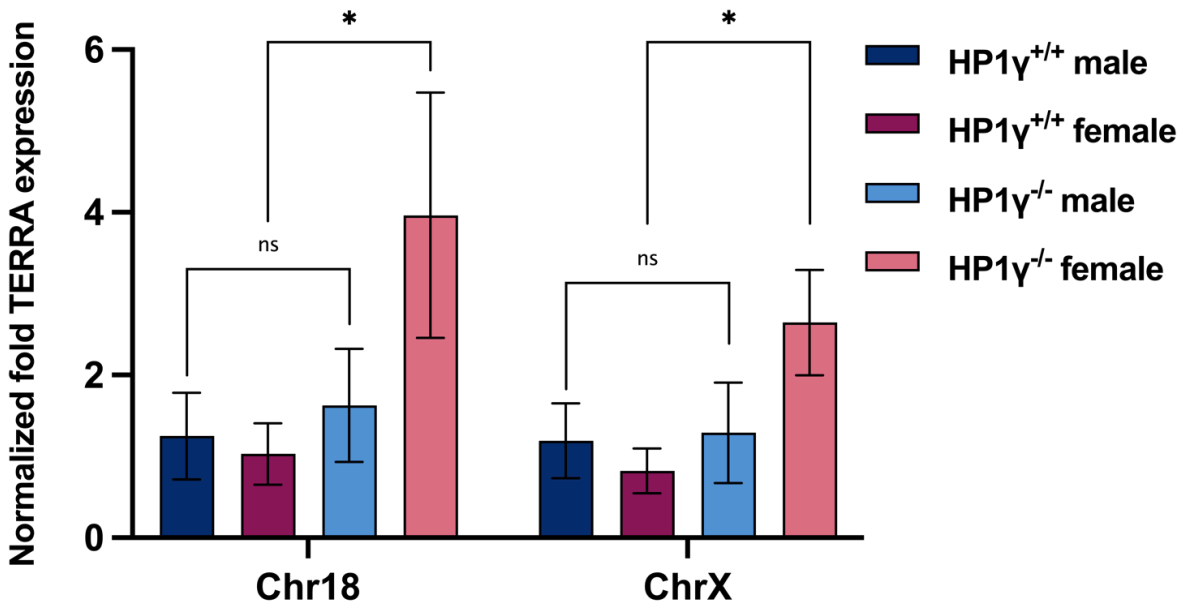


Figure 4.7 Depletion of HP1γ leads to upregulation of chromosome 18 and chromosome X TERRAs in females. RT-qPCR of total MEFs RNA (3 µg). The graphs represent the average value from three biological and three technical replicates for each sample after normalization to β-actin. Error bars represent the SD. Statistical significance was tested with one-way ANOVA and Tukey's post-hoc test, * $p<0.05$ ns=non-significant.

Finally, taking advantage of the SLIC-CAGE data (see Section 3.2.6), I interrogated whether there is differential usage of TSS by TERRAs, in the presence or absence of HP1γ. No shifting promoters came up at subtelomeric and telomeric sites from this analysis (data not shown), suggesting that this is not the reason for the enrichment of long TERRA transcripts seen by HP1γ knockout.

Taken together, these experiments suggest that HP1γ plays an important role in regulating the expression of TERRA transcripts, which appears to have a female-biased expression pattern at this developmental stage. The TRF1 downregulation in the absence of HP1γ likely contributes to the elevated levels of TERRAs.

4.2.3 Elevated levels of telomeric R-loops in the absence of HP1γ

Due to the complementarity of TERRA molecules with telomeric DNA, TERRAs can hybridize to the exposed C-rich lagging strand during DNA replication and generate DNA:RNA hybrids at telomeres, which in turn can give rise to R-loops by displacing the G-rich strand (Balk *et al.*, 2013). While these structures appear to be important for physiological telomere homeostasis (Balk *et al.*, 2013; Pfeiffer *et al.*, 2013; Graf *et al.*, 2017), persistent DNA:RNA hybrids can cause replication fork stalling and hence replication stress (Crossley, Bocek and Cimprich, 2019; García-Muse and Aguilera, 2019; Niehrs and Luke, 2020). For instance, in Immunodeficiency, Centromere region instability, Facial anomalies syndrome (ICF) patients, abnormally high level of TERRAs have been suggested to give rise to R-loops that cause telomeric dysfunction and genome instability (Sagie *et al.*, 2017).

We examined whether the increased TERRA levels observed in HP1γ depleted cells lead to more telomeric R-loops by performing DNA:RNA immunoprecipitation (DRIP) experiments (See Figure 4.8). During DRIP, genomic DNA is extracted and a restriction enzyme cocktail is used to fragment the genome. Then, the S9.6 monoclonal antibody is employed to specifically immunoprecipitate DNA:RNA hybrids (Boguslawski *et al.*, 1986; Bou-Nader *et al.*, 2022). Characterisation of the localisation of these structures can be achieved with subsequent qPCRs.

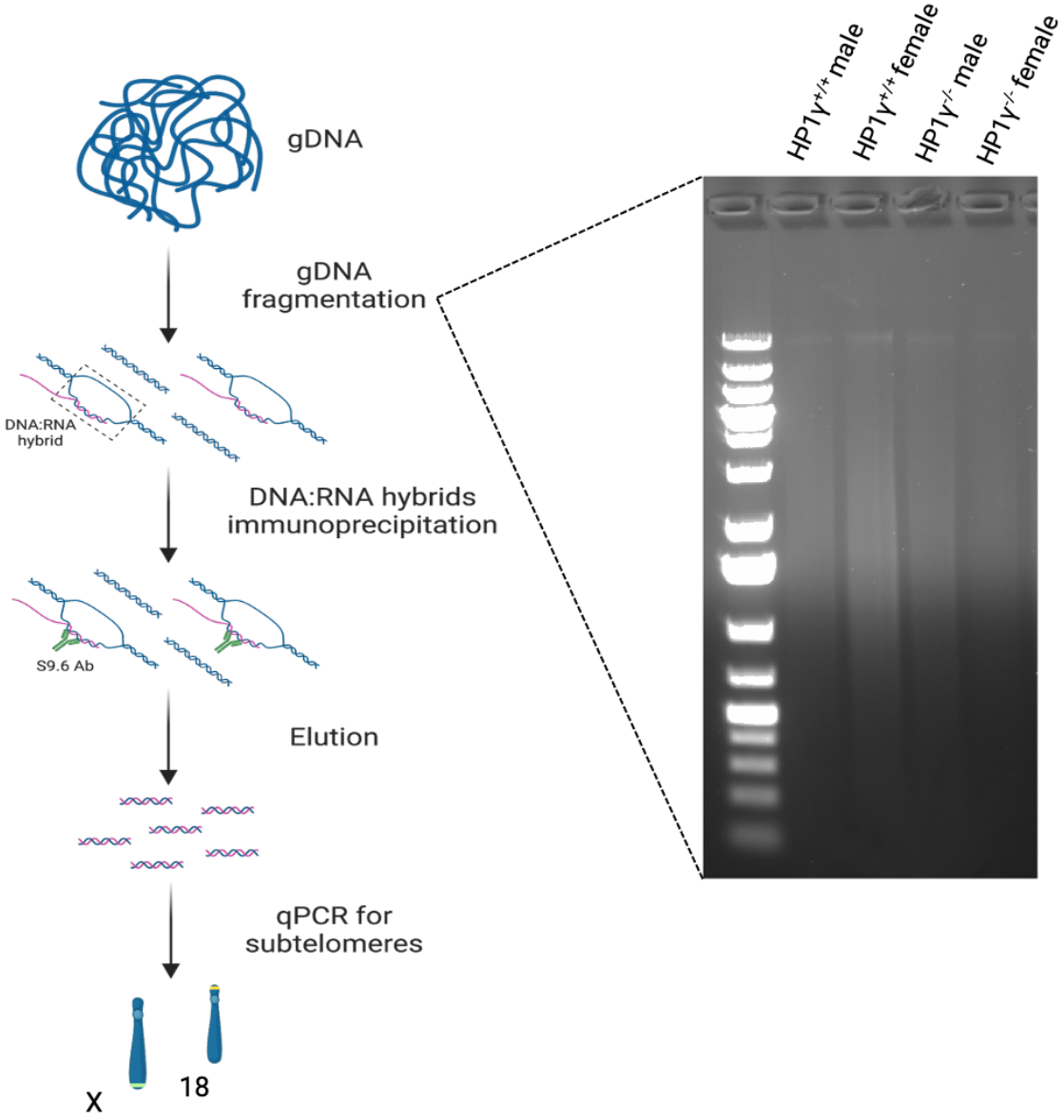


Figure 4.8 DNA:RNA immunoprecipitation (DRIP) is used to test R-loops at telomeres. Left: Outline of the DRIP experiment. Briefly, genomic DNA (gDNA) is isolated and fragmented by a combination of restriction enzymes. DNA:RNA hybrids are immunoprecipitated with the S9.6 antibody, the fragments are isolated and quantification of site-specific DNA:RNA hybrids is performed by qPCR. Right: Agarose gel showing digested genomic DNA.

Since dysregulation of TERRA expression was mainly observed on the end of chromosomes 18 and X, we analysed the abundance of R-loops at these two specific chromosomes. DNA:RNA hybrids signal increased upon HP1γ loss at the ends of chromosomes 18 and X, especially in HP1γ$^{-/-}$ female MEFs. Chromosome 18 telomeric DNA:RNA hybrids went up by 2.6-fold and chromosome X telomeric DNA:RNA hybrids

went up by 4.9-fold in HP1γ$^{-/-}$ females compared to HP1γ$^{+/+}$ females (Figure 4.9). In accordance with the TERRA expression levels, the effect of HP1γ ablation was less dramatic in males, where in HP1γ knockout conditions, chromosome 18 telomeric DNA:RNA hybrids showed a 2.8-fold increase, while chromosome X telomeric DNA:RNA hybrids remained at similar levels to wild-type conditions. An essential specificity control for DRIP is the use of RNase H as this endonuclease catalyses the cleavage of RNA in an RNA:DNA context (Ohtani *et al.*, 1999). Indeed, the RNase H-treated samples show very low levels of DNA:RNA hybrids at the interrogated loci, compared to the non-treated samples (Figure 4.9).

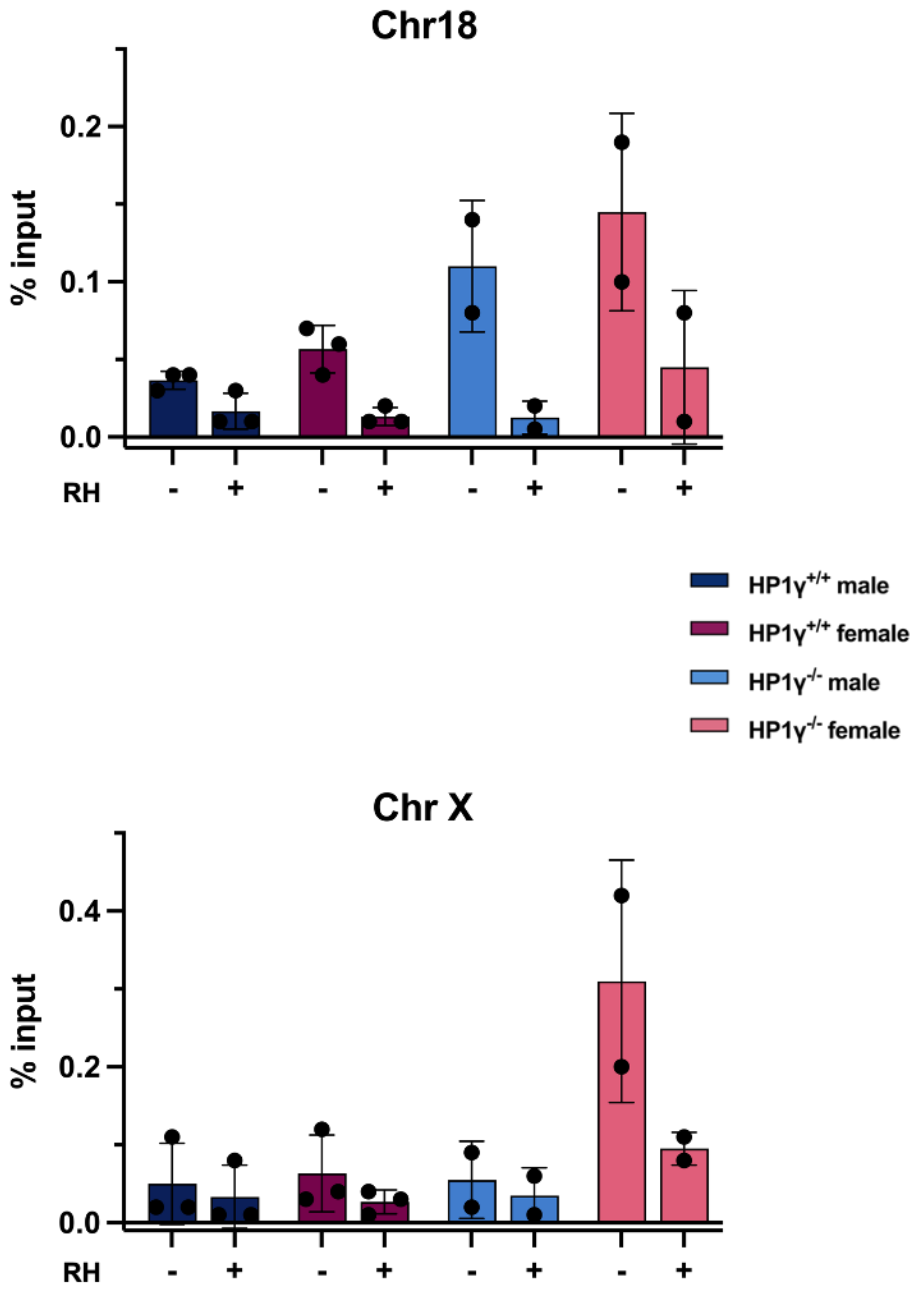


Figure 4.9 Upregulation of subtelomeric chromosome 18 and chromosome X DNA:RNA hybrids, in the absence of HP1γ. DRIP RT-qPCR assay at chromosome 18 and chromosome X subtelomeric regions. Bars represent the mean ± SD for the percentage of DNA:RNA hybrids in respect to input for three biological replicates of HP1γ$^{+/+}$ male and HP1γ$^{+/+}$ female MEFs and two biological replicates of HP1γ$^{-/-}$ male and HP1γ$^{-/-}$ females MEFs. RH: RNase H treatment.

This indicates that HP1γ regulation of TRF1 is essential to prevent increased TERRA levels that can form excessive telomeric R-loops at the ends of chromosome 18 and X and again suggest that female cells might be more sensitive to such dysregulation.

4.2.4 HP1γ depletion does not induce C-circle DNA levels

In a small percentage of cancers, telomere length is maintained not via the re-activation of telomerase, but rather through break-induced telomeric HR. The telomerase-independent maintenance of telomere length is known as alternative lengthening of telomeres (ALT) pathway (Chang *et al.*, 2003; Neumann *et al.*, 2013). High levels of TERRA transcripts characterise ALT cells where it is believed that they render telomeres more prone to recombination by forming RNA:DNA hybrids (Arora *et al.*, 2014).

Given the upregulation of TERRA and enrichment of telomeric R-loops upon HP1γ deletion, we decided to test the levels of C-circles in primary MEFs. C-circles are molecules of partially single-stranded telomeric DNA and they constitute the key ALT-specific biomarker (Conomos et al., 2013). While, the exact mechanism of their formation remains unclear, some reports suggest that they mainly arise due to replication stress at telomeres (Henson *et al.*, 2009). C-circle levels can be measured by the C-circle assay (CCA) that utilises the rolling circle amplification Φ29 DNA polymerase (Henson *et al.*, 2017). Φ29 DNA polymerase is auto-primed by the partial telomeric G-strand (TTAAGGG), producing long telomeric ssDNA concatemers. dATPs, dGTPs and dTTPs are used for the amplification of the C-circles, while dCTP is omitted to ensure the specificity of the assay for C-circles which are comprised entirely of telomeric DNA. The amplified products are transferred to a N^+ membrane and hybridised with a telomere-specific C-rich oligonucleotide. (Figure 4.10)

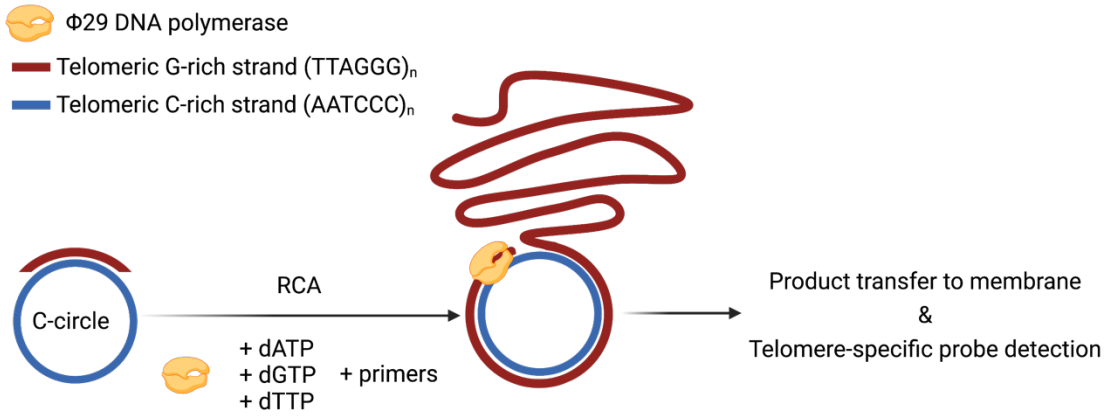


Figure 4.10 C-circle assay (CCA) workflow. Φ29 DNA polymerase is primed by the partial telomeric G-rich strand, and through a rolling circle amplification there is the production of long telomeric ssDNA concatemers. The amplified products are transferred to a N^+ membrane and hybridised with a telomere-specific C-rich oligonucleotide.

DNA from one male HP1γ$^{+/+}$ sample, one female HP1γ$^{+/+}$ sample and one female HP1γ$^{-/-}$ sample were used for a preliminary CCA. A human cell line which employs the ALT pathway to maintain its telomere length (U-2 OS) was included as a positive control, while a no-DNA sample served as negative control. Moreover, samples where Φ29 DNA polymerase was omitted, served as a control for sample-specific background. The analysis reveals that C-circle levels of primary MEFs are very low, as Φ29 DNA polymerase amplified samples have very similar signal with samples lacking Φ29 DNA polymerase, while the U-2 OS sample (positive control) shows a clear signal enrichment in the Φ29 DNA polymerase (Figure 4.11 A). Quantification of the signal shows a small female-biased signal enrichment in wild-type conditions in a single experiment, while C-circles in HP1γ knockout conditions do not present elevated levels, as one might have expected (Figure 4.11 B).

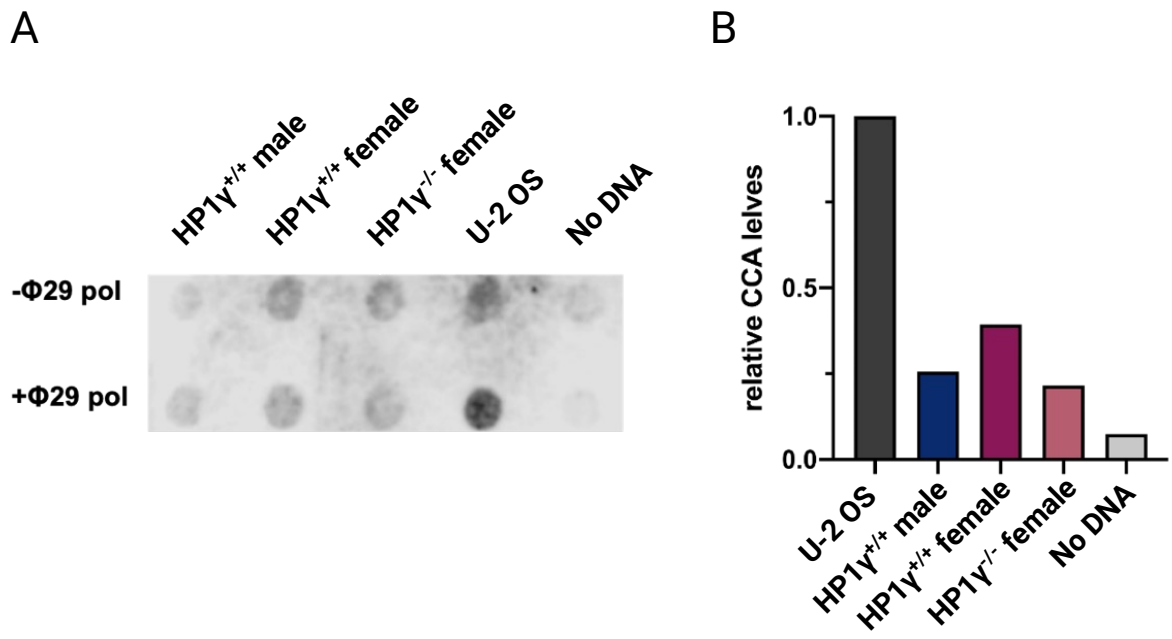


Figure 4.11 C-circle assay (CCA) using E13.5 MEFs DNA. (A) 7.5 ng of DNA from one independent biological replicate was used for CCA. The blot was revealed with a DIG-Tel-C-rich probe U-2 OS, an ALT cell line, serves as a positive control, while the absence of Φ29 polymerase serves as negative control. 3 dNTPs are used to ensure sole amplification of the c-circles (B) Relative quantification of CCA signal intensity.

4.2.5 Increased telomeric fragility upon loss of HP1γ

Chromosomal regions sensitive to DNA replication stress are known as common fragile sites (CFS) and are considered hotspots for genomic instability and rearrangements in cancers (Li & Wu, 2020). CFSs contain difficult-to-replicate sequences such as repetitive DNA elements, which tend to form secondary structures that stall DNA replication. Due to their repetitive nature, their unique structure and mechanism of replication (via telomerase or ALT), telomeres can exhibit a CFS phenotype that is characterised by aberrant signal that is spatially separated from the chromatid end of metaphasic chromosomes, when examined by Quantitative Fluorescence in situ hybridisation (Q-FISH) (Sfeir et al., 2009).

To perform the Q-FISH assay, the microtubule destabiliser colcemid is used to arrest cells in metaphase (Dewey and Miller, 1969; Zijlmans *et al.*, 1997). Cells are then

collected, hypotonic shock is performed, the suspension is fixed and dropped on slides from a vertical position, resulting in the breakage of the cell membranes and release of the DNA content. In this stage of the cell cycle, where chromatin is highly condensed in chromosomes, telomeres can be visualised at their extremities with a telomere-specific DNA probe conjugated to a fluorophore (Figure 4.12). Proper labelling by Q-FISH, enabling the discrimination of the telomeric shape is possible, since the inbred mouse telomere length ranges from around 30 to 150 kb (Zijlmans *et al.*, 1997). Furthermore, to ensure unbiased analysis, characterisation of the telomeric phenotypes after image acquisition, took place by being blinded to the different genotypes.

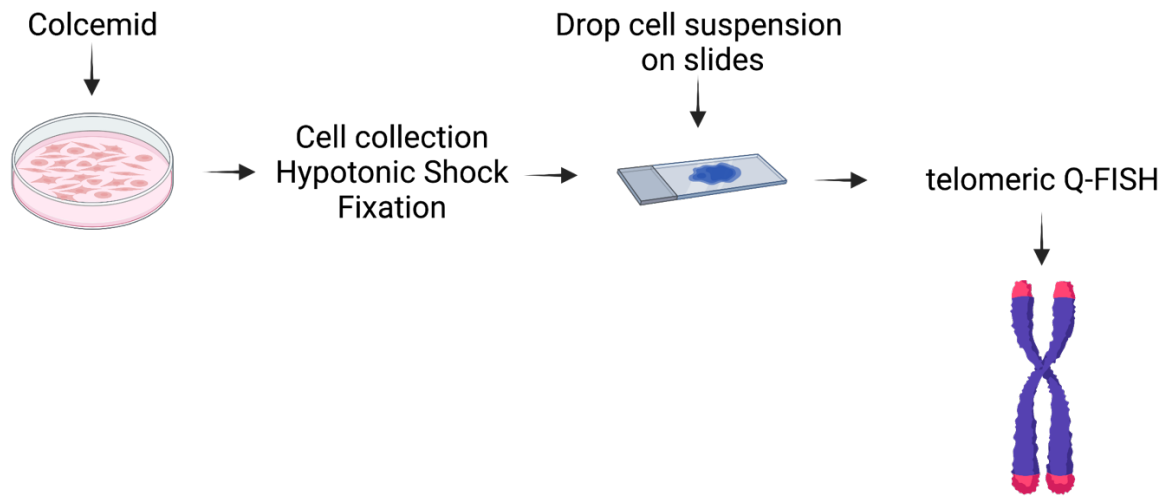


Figure 4.12 Telomeric Q-FISH experimental workflow. Cells are arrested in metaphase via incubation with colcemid. The cell suspension is collected and undergoes hypotonic shock and is later preserved by several fixation steps. Cell suspensions are dropped on slides, where the DNA is released. Slides are then hybridised with a telomere-specific DNA probe conjugated with a fluorophore and the rest of the genome is stained with DAPI.

Telomeres of male and female HP1γ$^{+/+}$ and HP1γ$^{-/-}$ MEFs were analysed by Q-FISH. Aberrant, or spatially separated signals, due to telomeric DNA being broken or failed to condense was scored as fragility. Telomeric loss was characterized by the complete absence of signal from the chromatid ends, while foci that appeared as single dots with twice the signal intensity deriving from the co-localization of the two chromatid ends were scored as sister chromatid fusion. My analysis shows that there

are no significant differences regarding the stability of telomeres between wild-type male and female MEFs (Figure 4.13). However, comparison of HP1γ$^{+/+}$ and HP1γ$^{-/-}$ telomeres reveals a significant higher incidence of fragility upon loss of HP1γ. The effect is stronger in females where 17 % of the HP1γ$^{-/-}$ chromosomes have fragile telomeres compared to 9 % in HP1γ$^{+/+}$ females and milder in males where 11 % of the HP1γ$^{+/+}$ chromosomes have fragile telomeres compared to 14 % in HP1γ$^{-/-}$ males. Telomere fragility is the main telomere defect in HP1γ$^{-/-}$ MEFs as neither telomere fusions nor telomere loss were observed (both phenotypes below 1 % of the total telomeres, irrespective of the genotype) (Figure 4.14), which is also consistent with a TRF1-deletion effect in these cells (Porreca *et al.*, 2020). Together the results suggest that the increased telomeric fragility observed more evidently in females arises not only due to the TRF1 downregulation that occurs in both sexes upon loss of HP1γ, but also due to the increased TERRAs and R-loops particularly evident in HP1γ$^{-/-}$ females.

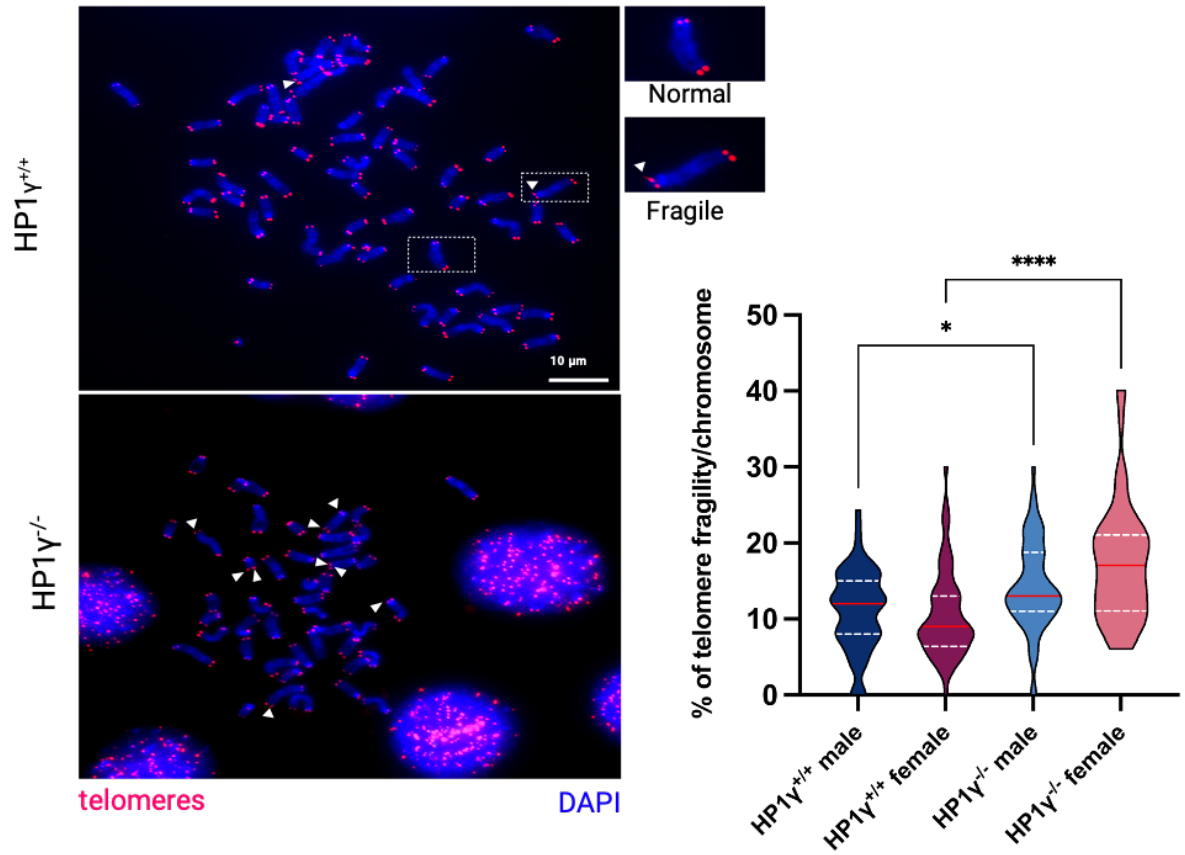


Figure 4.13 Loss of HP1γ causes increased telomeric fragility in E13.5 MEFs. Left: Representative images of mouse metaphasic chromosomes (blue), hybridized with a telomere-specific probe, conjugated with a Cy3 fluorophore (red). The arrows indicate distorted signal (fragile telomeric site). Right: Percentage of telomere fragility per metaphase

spread. Violin plots represent the combined analysis of 60 metaphases from three independent biological replicates (20 metaphases per biological replicate). Median is underlined in red and quartiles in white. Statistical significance was tested with one-way ANOVA and Tukey's *post-hoc* test, * p<0.05, **** p<0.0001.

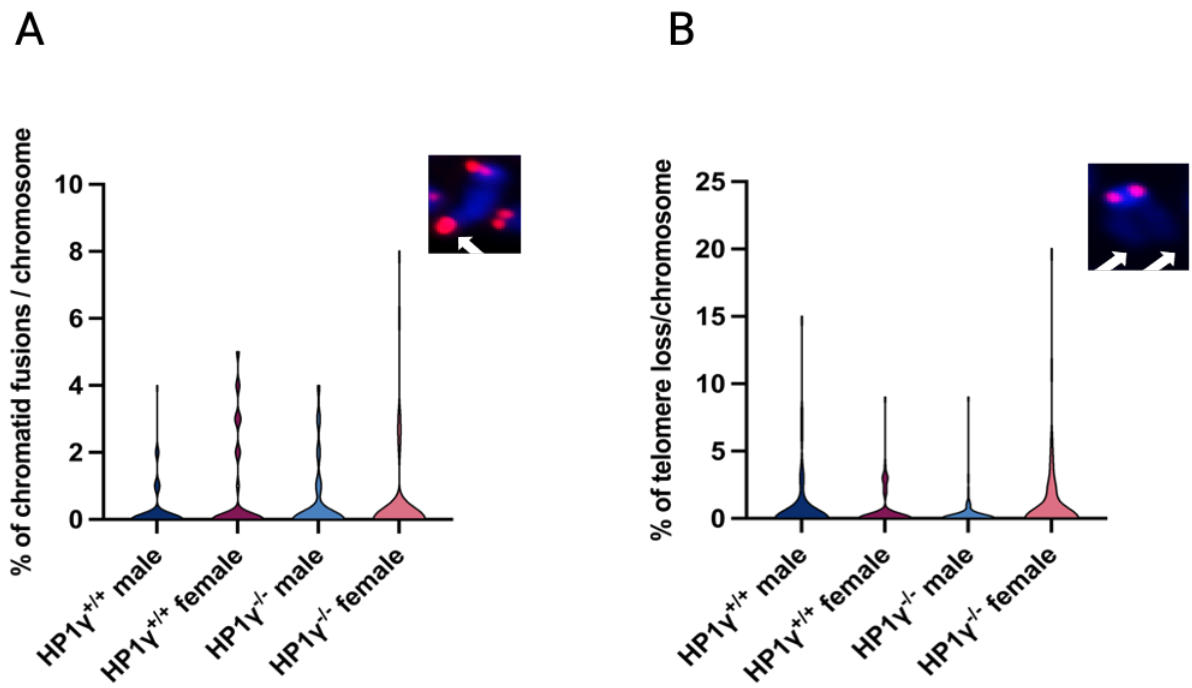


Figure 4.14 Telomere fusions and telomere loss is not affected by HP1γ. Percentage of (A) telomere fusions and (B) telomere loss per chromosome spread. Violin plots represent the combined analysis of 60 metaphases from three independent biological replicates (20 metaphases per biological replicate). Median is underlined in red and quartiles in white. No statistical significance (one-way ANOVA and Tukey's *post-hoc* test).

4.2.6 HP1γ depletion induces telomeric DNA damage

Telomeres guard against the recognition of linear chromosome ends as DSBs, while telomere fragility is a marker of replication stress that is often linked to the induction of the DNA damage response (DDR) (Martínez *et al.*, 2009; Sfeir *et al.*, 2009; Vannier *et al.*, 2012). Therefore, we hypothesised that the telomeric fragility phenotype observed in the absence of HP1γ would be accompanied by a potential recruitment of

DSB-associated factors to telomere. Hence, an IF followed by FISH was performed to examine the localisation of DNA damage factors to telomeres in the absence of HP1γ. A telomere-targeting DNA probe conjugated with a fluorophore was used to mark telomeric DNA of MEFs. Co-localisation of the factors of interest to telomeres was only considered positive when signal superposition existed between telomeres and the interrogated candidate. Signal quantification was automated with the use of an image analysis software (CellProfiler).

Accumulation of the phosphorylated histone variant H2AX (γH2AX) has been suggested to act as an early sensor of DSBs. To allow DNA repair, the chromatin must be remodelled and γH2AX, is involved in the steps leading to chromatin decondensation (Anderson, Henderson and Adachi, 2001). IF-FISH analysis reveals that HP1γ deletion leads not only to a nuclear-wide increase of γH2AX signal, but also to more telomere dysfunction-induced foci (TIFs) (Figure 4.15 A). Specifically, deletion of HP1γ results in a 1.8-fold increase of nuclear γH2AX foci in males and a 1.7-fold increase in females (Figure 4.15 B). Moreover, γH2AX telomere-specific signal doubles in males and triples in females in the absence of HP1γ (Figure 4.15 C).

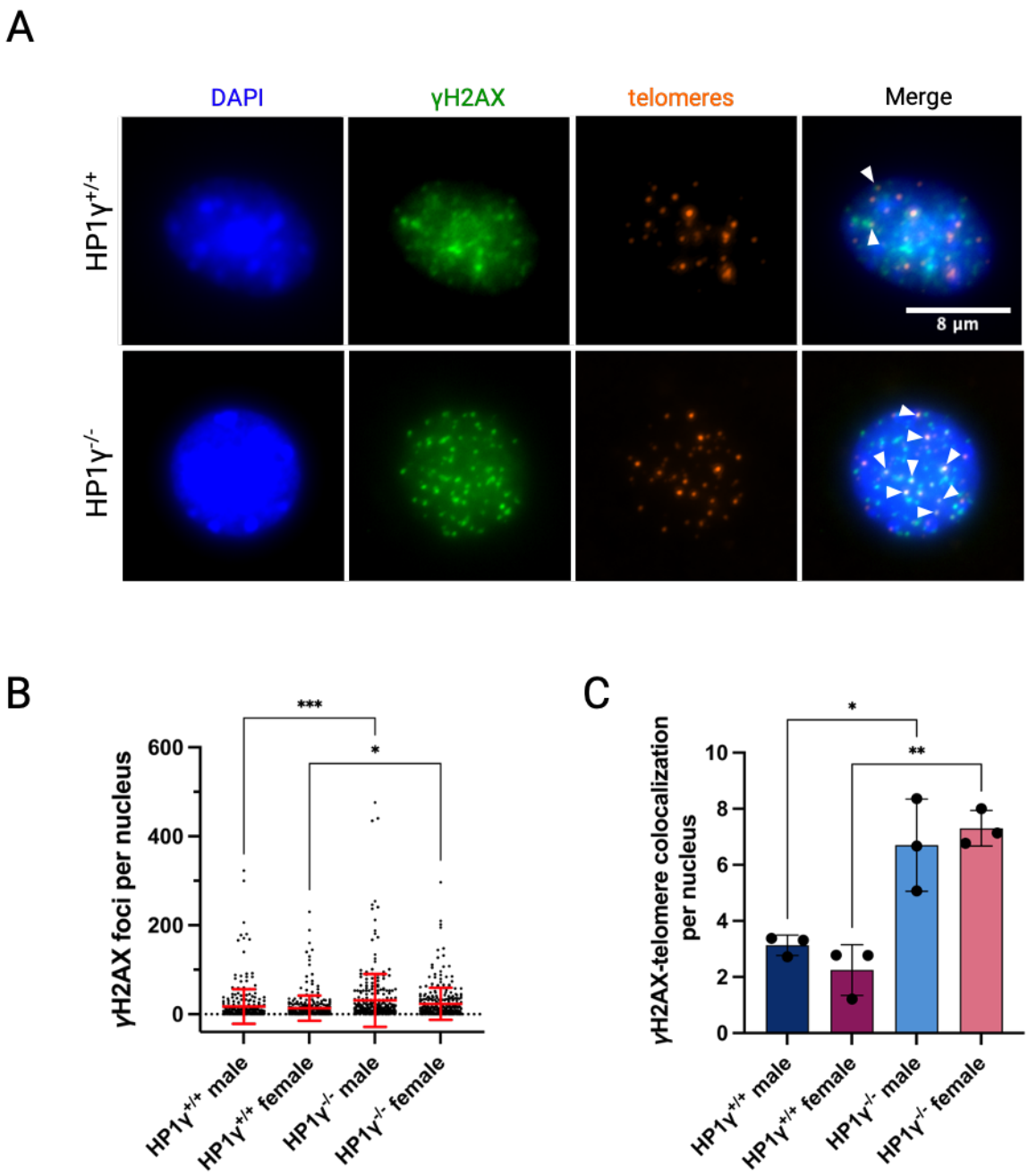


Figure 4.15 Elevated levels of γH2AX nuclear-wide signal and more γH2AX TIFs, upon HP1γ deletion. (A) Representative image of IF-FISH showing co-localisation of telomeres (red) with γH2AX (green) in HP1γ$^{+/+}$ and HP1γ$^{-/-}$ MEFs nuclei (DAPI). (B) Quantification of γH2AX foci per nucleus and of (C) γH2AX co-localization to telomeres per nucleus (n = 300 nuclei). Data are represented as mean ± SD from three independent biological replicates. Statistical significance was tested with one-way ANOVA and Tukey's *post-hoc* test, * p<0.05, ** p<0.01, *** p<0.001.

p53-binding protein 1 (53BP1) is another important factor involved with the cellular response to DSBs. When a DSB is detected, 53BP1 rapidly accumulates on the chromatin surrounding the break site, following the signalling cascade that is initiated

by γH2AX (Anderson, Henderson and Adachi, 2001) and acts as a key determinant of DSB repair pathway choice, promoting the non-homologous end-joining (NHEJ) repair pathway, over the homologous recombination (HR) (Bouwman *et al.*, 2010; Bunting *et al.*, 2010). Accordingly, I investigated the possible genome-wide accumulation of 53BP1 foci, as well as, its localisation on telomeric DNA, which would reinforce the idea of the presence of DSBs and would indicate NHEJ as the repair pathway of choice. Unlike γH2AX, 53BP1 levels do not show a difference among HP1γ$^{+/+}$ and HP1γ$^{-/-}$ conditions (Figure 4.16 A). Total 53BP1 foci per nucleus are similar among all conditions with a mean number of foci of 55 in HP1γ$^{+/+}$ males, 60 in HP1γ$^{+/+}$ females, 57 in HP1γ$^{-/-}$ males and 71 in HP1γ$^{-/-}$ females (Figure 4.16 B). 53BP1 localisation to telomeres is also not affected by the loss of HP1γ with 53BP1 foci ranging from 8 to 12 at chromosome ends of the different genotypes (Figure 4.16 C).

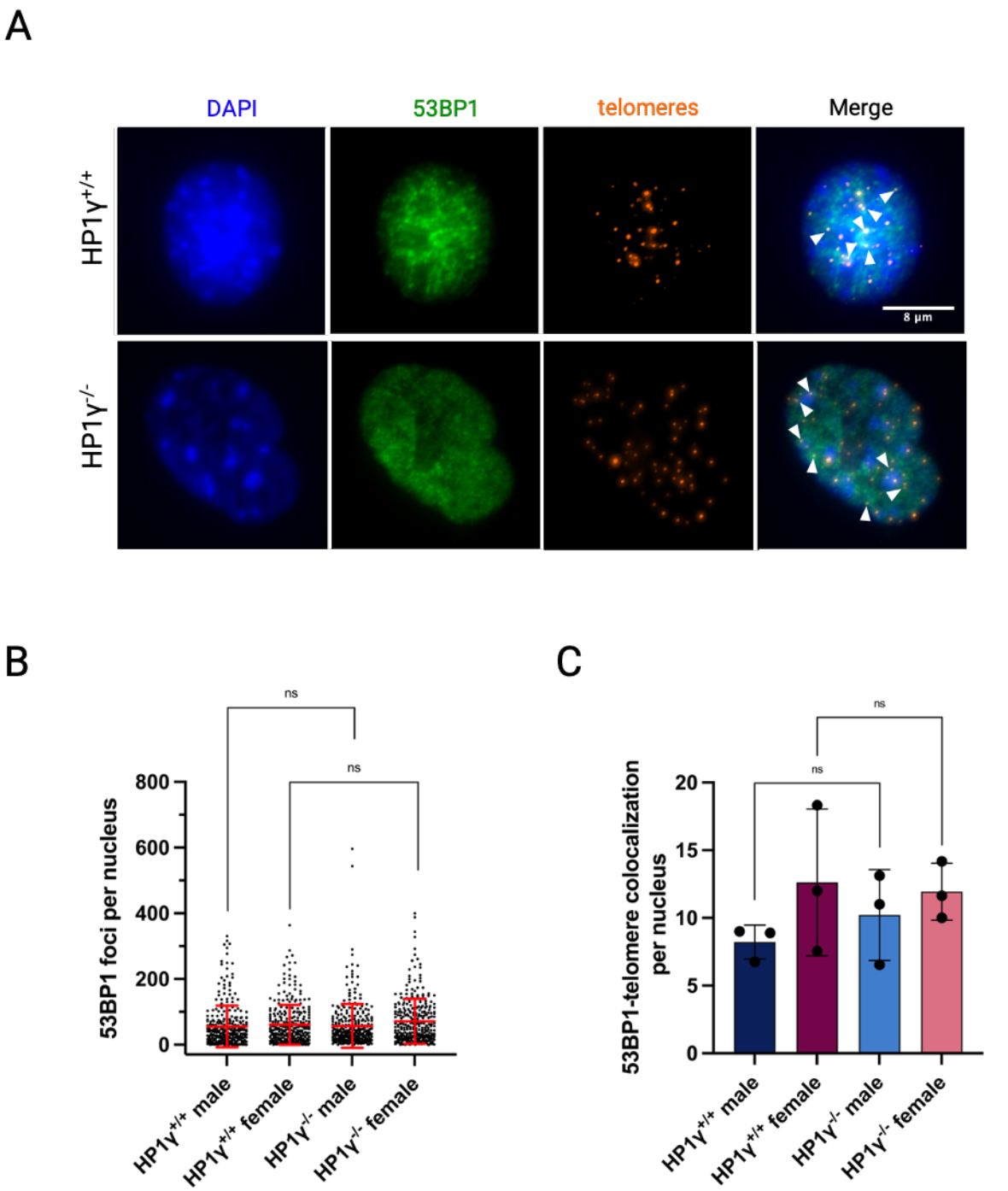


Figure 4.16 HP1γ loss does not affect 53BP1 levels. (A) Representative image of IF-FISH showing co-localisation of telomeres (red) with 53BP1 (green) in HP1γ$^{+/+}$ and HP1γ$^{-/-}$ MEFs nuclei (DAPI). (B) Quantification of 53BP1 foci per nucleus and of (C) 53BP1 co-localization to telomeres per nucleus (n = 300 nuclei). Data are represented as mean ± SD from three independent biological replicates. Statistical significance was tested with one-way ANOVA and Tukey's *post-hoc* test, ns=non-significant.

The increased presence of γH2AX TIFs indicates that the DNA damage signalling is activated at telomeres upon HP1γ depletion in both sexes further showing the compromise of the telomeric structure, alongside the fragile phenotype.

4.2.7 Telomere length remains unaffected by the loss of HP1γ

The molecular basis of senescence is directly mandated by the Hayflick limit, where this cellular process is triggered when telomeres reach a critical minimal length (Hayflick and Moorhead, 1961; Harley, Futcher and Greider, 1990). The telomere shortening leads to a degradation of the normal telomeric structure, recognized as DNA damage with the subsequent activation of cell cycle regulators including p16 and p53 leading to cell cycle arrest (Kuilman *et al.*, 2010). Given the earlier onset of senescence, upregulation of Cdkn2a, the compromised structure of telomeres observed by Q-FISH and increased γ-H2AX TIFS in MEFs lacking HP1γ, we decided to test mouse telomeric length aberrations upon loss of HP1γ and whether there is a sexually dimorphic nature in these cells. Telomere Restriction Fragment (TRF) analysis (Lansdorp *et al.*, 1996) is the "gold-standard" for assessing the average telomere length. The assay includes genomic DNA digestion by a combination of restriction enzymes while telomeres remain intact, followed by a Southern blot. The visualization of telomere length is achieved by employing a telomere-specific probe. The TRF experiment demonstrates that despite the transcriptional effect of HP1γ on telomere-regulatory factors like TRF1, the average telomere length, remains unchanged among the different genotypes and sexually unbiased. The majority of telomeres migrated into a strong band above the 20 kb mark and the smearing of shorter telomeres (around 3.5 kb) remained faint in all conditions (Figure 4.17). HP1γ loss has no major noticeable effect on telomere length. This is also consistent with

previous reports showing an absence of effect on telomere length when mouse TRF1 is downregulated (Martínez *et al.*, 2009; Sfeir *et al.*, 2009).

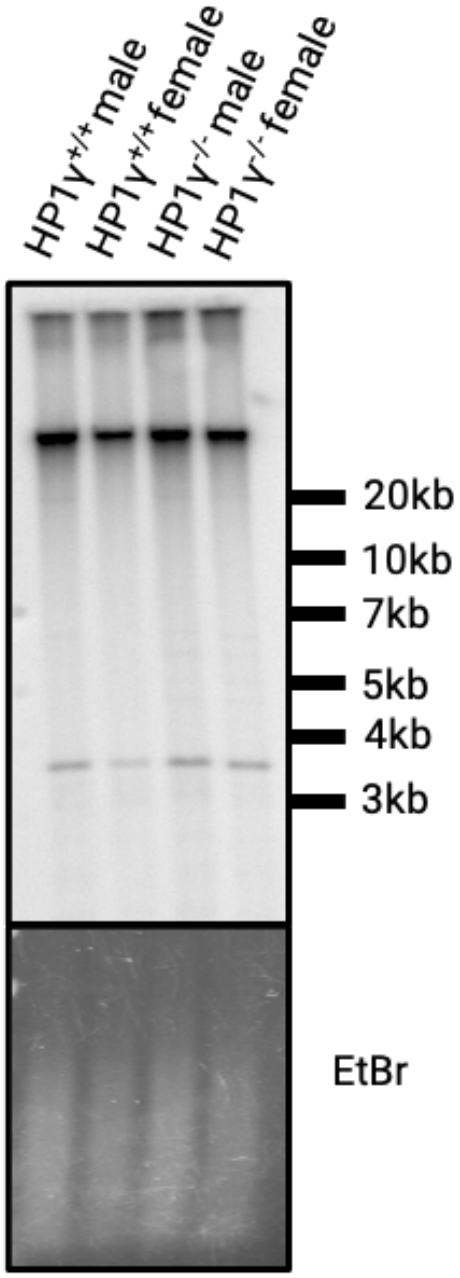


Figure 4.17 Loss of HP1γ does not result in telomere length heterogeneity. Telomere Restriction Fragments (TRF) analysis followed by Southern blot. The blot was revealed with a DIG-Tel-C-rich probe (top). Ethidium bromide (EtBr) staining (bottom) is used as loading control.

4.3 Discussion

Telomere stability is largely influenced by the chromatin state. While recent studies show that human telomeres display lower than expected levels of H3K9me3 (O'Sullivan *et al.*, 2010; Cubiles *et al.*, 2018), mouse telomeres are naturally enriched for this heterochromatic mark (García-Cao *et al.*, 2003), and have a highly compacted chromatin state (Hübner *et al.*, 2022). HP1 can affect chromatin compaction by regulating the histone methyltransferases (HMTs), responsible for the establishment of H3K9me3 with important implications on telomere integrity. A recent report showed that the histone methyltransferase (HMT) SETDB1, a known interactor of HP1 (Schultz *et al.*, 2002), is responsible for the establishment of H3K9me3 on mouse telomeres (Gauchier *et al.*, 2019), while loss of other H3K9me3-affiliated HMTs like SUV39H1/H2, also results in defective telomeres with increased length (García-Cao *et al.*, 2003). The role of HP1γ at telomeres has not been previously investigated through the prism of its transcriptional effect on TERRAs and telomeric proteins. Moreover, little is known about sex differences that could be characterising telomeres at the epigenetic level.

In this Chapter, I have investigated the role of HP1γ on telomere maintenance at early mouse development. This work showcases the important transcriptional control by HP1γ on various telomeric factors including TRF1 and TERRAs that has profound consequences for telomere stability, with an underlying sexually dimorphic nature.

Genes involved in telomere homeostasis have been mapped to both autosomal (Andrew *et al.*, 2006) and sex chromosomes (Connor *et al.*, 1986). In order for telomeres to overcome the several hindrances that arise during their replication, the coordination of a large set of factors is necessary. For instance, the *Dkc1* gene that encodes dyskerin, which is involved in the maturation of TERC, resides on the X chromosome of both mouse and humans (Sarek *et al.*, 2015). Meanwhile, RTEL1 helicase that assists with T-loop resolution in order for the replication fork to progress (Sarek *et al.*, 2015), or the BLM helicase that resolves G4 structures that may arise during telomere replication are both encoded by autosomal genes. Both of these helicases are recruited at telomeres by subunits of the shelterin complex (Glousker

and Lingner, 2021), respectively by TRF2 and TRF1. While, other subunits of shelterin, like TRF2 and POT1 may still be able to associate with telomeres in TRF1 null cells (Celli, Denchi and de Lange, 2006; Palm *et al.*, 2009), TRF1 loss results in replication stress at telomeres (Martínez *et al.*, 2009; Sfeir *et al.*, 2009). My analysis shows that HP1γ regulates the expression of many of the telomere-related factors (Glousker and Lingner, 2021) and importantly TRF1, whose levels are reduced in the absence of HP1γ, especially in males (Figure 4.1 and Figure 4.2). HP1γ's interaction with TIN2 has been shown to be necessary for the establishment of cohesion at human telomeres in S phase (Canudas *et al.*, 2011). The lower levels of TRF1 in the absence of HP1γ could further disrupt shelterin formation, exacerbating this phenotype, implicating HP1γ with telomere cohesion through its transcriptional role. Since most of the telomere-associated genes are downregulated upon HP1γ knockout, it is likely that for these genes HP1γ acts as a transcriptional activator.

Telomeres were long considered silent, heterochromatic regions of the genomes, but the discovery of TERRAs, transcripts from these regions, broke the dogma (Azzalin *et al.*, 2007; Schoeftner and Blasco, 2007). Both the total TERRA numbers and the polyadenylated TERRA fraction, which comprises the majority of TERRA molecules in mice (de Silanes *et al.*, 2014) are elevated in the absence of HP1γ (Figure 4.4 and Figure 4.5). As the polyadenylation status of TERRAs can affect their stability (Porro *et al.*, 2010) and localisation (Chu et al., 2017), HP1γ affects TERRA functionality by influencing their numbers.

The constant telomeric length among the different genotypes suggests that altered TERRA levels observed in the HP1γ$^{-/-}$ MEFs and especially females, do not arise due to longer telomeres which could act as extended substrates for TERRA transcription (Arnoult, van Beneden and Decottignies, 2012), or shorter telomeres that as shown in yeast, can lead to a lower number of transcription repressors and in turn increased TERRA expression (Cusanelli, Romero and Chartrand, 2013; Moravec *et al.*, 2016).

A difference in the length of TERRA transcripts would indicate for a role of HP1γ in the transcriptional termination which could be happening at multiple sites within the

telomeric tract and/or differential processing of their polyadenylation, which does not appear to be the case. It could also suggest an involvement of HP1γ on TERRA processing by splicing regulation (Saint-André *et al.*, 2011; Ameyar-Zazoua *et al.*, 2012). Loss of HP1γ would have led to miss-spliced transcripts with different lengths that would have been visualised by Northern blot if they were stable enough. If the splicing variants were unstable, they would have been degraded at a faster rate (Azzalin *et al.*, 2007) and the total TERRA levels would have been lower in HP1γ$^{-/-}$ MEFs. Both scenarios are opposite to what we observe.

On the other hand, the CUT & RUN experiments (Figure 3.19) suggest a direct role of HP1γ on TERRA transcription, since we observed direct HP1γ binding at the ends of chromosome 18 and X in both sexes, where the majority of mouse TERRA arise (de Silanes *et al.*, 2014) and where the transcription is most dysregulated (Figure 4.7). The TERRA-focused analysis of the preliminary SLIC-CAGE experiment suggested that HP1γ does not influence TERRA levels by regulating differential usage of TSSs residing in the subtelomeric track. However, HP1γ has been shown to counteract CTCFs effect, when found in a complex with ADNP and CHD4 (termed ChAHP) (Kaaij *et al.*, 2019). This could mean, that in wild-type conditions HP1γ is bound at TERRA promoters and supresses their expression. Loss of HP1γ leads to misformation of the ChAHP complex, enabling CTCF to localise at TERRA promoters and induce their expression (Beishline *et al.*, 2017).

Another possibility of how HP1γ may contribute to the suppression of TERRAs is by promoting the formation of a locally condensed chromatin structure. Since HP1γ physi(Schultz *et al.*, 2002; Stewart, Li and Wong, 2005; Fritsch *et al.*, 2010)ong, 2005; Fritsch et al., 2010), their recruitment can induce further methylation of H3K9 at the target locus which in turn can provide more binding sites for the HP1 isoforms. Dense HP1 binding and dimerisation can result in the re-configuration of the histone core with the normally buried residues now exposed and able to participate in weak multivalent interactions, promoting the formation of phase-separated liquid condensates and tight crosslinking of nucleosomes (Canzio *et al.*, 2011; Sanulli *et al.*, 2019). While HP1γ has not been shown to form liquid droplets on its own (Larson *et al.*, 2017; Strom *et al.*, 2017), heterodimerisation with other HP1 isoforms (Machida *et al.*, 2018) may result in LLPS-driven exclusion of transcription factors and ultimately silencing of transcription.

A big portion of the total mouse TERRAs arise from chromosome X (de Silanes *et al.*, 2014; Viceconte N *et al.*, 2021) and since females have two X chromosomes, one would expect that expression of X-linked TERRAs would be double in females. However, according to RT-qPCR analysis this is not the case, with both sexes showing comparable levels in wild-type conditions (Figure 4.7). The inactivation of one of the two X chromosomes could be acting as a dosage compensation mechanism for these long-non coding RNAs. Apart from the *in cis* effects of TERRA dysregulation by the loss of HP1γ, it would be worth investigating by CHIRT experiments (Chu et al., 2017) their potential re-distribution in distal genomic regions and the downstream effects which could result in combination with HP1γ to sex differences.

To further assess the effects of HP1γ depletion on telomeres, a Q-FISH approach was employed and the telomeres of metaphase mouse chromosomes were analysed. As expected, the initial comparison of telomeres from wild-type males and females did not reveal significant differences regarding their stability. Fragile telomeres comprised a very small portion of the total telomeres tested (Figure 4.13), similar to the levels reported previously (Sfeir *et al.*, 2009), while telomeric loss and telomere fusions were barely detectable (Figure 4.14). If a different outcome had been observed, that would suggest an evolutionary discrepancy between the sexes, as telomeres are essential for the overall genome integrity and normal cell function (Fajkus, Sýkorová and Leitch, 2005). Interestingly, loss of HP1γ led to an evident increase of telome fragility in MEFs and this phenotype was stronger in females where almost double of the total telomeres were fragile compared to their wild-type counterparts (Figure 4.13). Fragile sites represent genomic regions where replication forks stall and collapse with their repair requiring breakage and re-establishment of functional DNA synthesis(Margalef *et al.*, 2018). Problems during this process result in fragile sites being hotspots for chromosomal rearrangements (Min, Wright and Shay, 2017).

Microscopy experiments followed to further investigate the potential DNA damage caused by the deletion of HP1γ by targeting proteins like γH2AX and 53BP1, early markers of DNA damage (Schultz *et al.*, 2000). γH2AX has been shown to instigate a

local chromatin environment that is permissive to the assembly of the DNA repair machinery through association with the RNF8 protein and the subsequent interaction with the nucleosome remodeller CHD4 (Luijsterburg *et al.*, 2012). Via 53BP1 recruitment at DNA damage sites, the cell can direct DDR towards either NHEJ (in G1) or HR (in S–G2), by controlling the extent to which a DSB is resected (Panier and Boulton, 2013).

My analysis reveals that upon depletion of HP1γ, γH2AX and 53BP1 levels rise in the nucleus, while there is a significant γH2AX enrichment at telomeric sites with a potential female-bias (Figure 4.15 and Figure 4.16). The increased presence of γH2AX TIFs, the absence of 53BP1 at telomeres and the lack of telomeric fusions in HP1γ$^{-/-}$ MEFs (Figure 4.14), suggests for accumulation of SSBs at telomeres (Podhorecka, Skladanowski and Bozko, 2010). The SSBs likely arise from telomere replication fork stalling/collapsing reflected as fragility (Figure 4.13), caused by the lack of TRF1 and/or elevated DNA:RNA intermediates. Since ATR is a major DDR component involved in SSB response and ATR is recruited at telomeres by the shelterin subunit POT1 (Denchi and de Lange, 2007), whose levels are unaffected by loss of HP1γ (Figure 4.1), ATR microscopy experiments could reveal if the telomeric DNA damage in the absence of HP1γ is due to single-stranded DNA lesions.
Unresolved SSB may lead to detrimental DSBs, rendering the selection of the damage repair pathway critical for the maintenance of genomic stability (Chapman *et al.*, 2013). The lack of 53BP1 TIFs and decreased TRF1 levels in HP1γ$^{-/-}$ MEFs, would indicate for HR to be the repair pathway of choice, where the repair machinery is recruited following telomeric replication stress (Scully and Xie, 2013; Porreca *et al.*, 2020). As BLM has been shown to drive BIR-mediated HR at telomeres (Yang *et al.*, 2020), its downregulation in HP1γ depleted cells (Figure 4.2), argues against the activity of this pathway, that can potentially exacerbate the fragile phenotype. However, the obvious candidates to test whether HR is functional in HP1γ$^{-/-}$ MEFs would be RAD51 and RAD52 (Claussin and Chang, 2015). As both of these factors are downregulated in the absence of HP1γ (Figure 4.2) that could mean that the accumulated telomeric DNA damage is not repaired efficiently, ultimately leading to senescence, as seen in Section 3.2.3.

The elevated telomeric R-loop levels observed in HP1γ$^{-/-}$ conditions could be also contributing to the telomere fragility and the TIF phenotype (Balk et al., 2013; Sagie et al., 2017). R-loops are dynamic structures that can occupy up to 5% of mammalian genomes (Sanz *et al.*, 2016) and are typically linked to ongoing transcription, known to form in response to transcriptional or replicative induced torsional stress in double-stranded DNA (Malig *et al.*, 2020). It is unclear at which stage of the cell cycle the TERRA-mediated R-loops on chromosome 18 and chromosome X ends are formed, considering that a recent study showed that TERRAs can form telomeric R-loops post-transcriptionally and *in trans* (Feretzaki *et al.*, 2020). While Ribonucleases H (RNase H1 and RNase H2) have the ability to resolve DNA:RNA hybrids, TERRA transcription is programmed to not overlap with replication through the cell cycle as an extra safety net, in human cells. Highest TERRA levels are observed in early G1 and drop to their lowest point in late S phase (Porro *et al.*, 2010). This staggered process is realised owing to the nonsense-mediated mRNA decay (NMD) pathway that ensures the displacement of TERRAs from telomeres. Problems of this pathway can lead to telomere instability and even loss of entire telomeric tracts. Deletion of HP1γ did not affect the expression of the NMD helicase UPF1 (Figure 4.2) that is responsible for regulating TERRA levels (Azzalin *et al.*, 2007). However, since the levels of UPF1 do not match the elevated TERRA levels this means that the additional TERRA molecules cannot be efficiently processed resulting in the accumulation of DNA:RNA hybrids at telomeres (Figure 4.9). This can be further exacerbated by the downregulation of FANCM helicase in HP1γ $^{-/-}$ cells (Figure 4.2) that normally resolves these structures (Pan *et al.*, 2019). To identify the stage of the cell cycle that upregulation of TERRA transcription occurs upon HP1γ depletion, cells would have to be synchronised and the TERRA levels could be tested by RT-qPCR. Finally, several studies have attributed a repressive role to TRF1 regarding TERRA regulation with several reports showing that loss of TRF1 lead to upregulation of TERRAs with decreased co-localization at telomeres (Porro *et al.*, 2014; Sadhukhan *et al.*, 2018; Feretzaki *et al.*, 2020). Therefore, the combination effect of HP1γ depletion and TRF1 downregulation most likely causes the accumulation of TERRAs at telomeres which results in the increased number of telomeric R-loops. Rescue experiments with exogenous TRF1 or HP1γ in HP1γ$^{-/-}$ MEFs could be performed to test whether the elevated levels of R-loops are

reversed and replication stress is relieved, which would ascertain direct or indirect function of HP1γ in this regulation.

Telomere length appears to be the same between wild-type and HP1γ$^{-/-}$ conditions, and so are C-circle levels, a well-established biomarker for the ALT pathway (Plantinga *et al.*, 2013). This suggests that traceable HR-mediated telomere elongation has not occurred which could have enabled cells to bypass senescence (Figure 4.17). In mouse somatic cells, ALT activity has been observed as a normal component of telomere biology, complementing telomerase activity for telomeric length maintenance (Neumann *et al.*, 2013). The lack of a clear phenotype in the absence of HP1γ in the preliminary CCA prompted me to not follow this phenotype further, also supported by the unaffected levels of ATRX, which has been extensively associated with ALT prevalence (Heaphy *et al.*, 2011; Li *et al.*, 2019). On the other hand, HP1γ knockout resulted in a downregulation of *Dkc1* which was stronger in males. It would be interesting to assess telomerase activity and its recruitment at telomeres and determine whether there is a sexual dimorphism underlying it.

Collectively, the results presented in this Chapter, suggest for a model (Figure 4.18) where HP1γ is necessary for a protected, telomeric chromatin state, through the regulation of telomere-accessory factors including TRF1 and TERRAs. Loss of HP1γ results in a deprotection of telomeres, uncontrolled expression of TERRAs and accumulation of telomeric R-loops. The lower levels of TRF1, in combination with the elevated R-loops levels cause replication stress which in turn cause DNA damage. The accumulated DNA damage at telomeres can contribute to the onset of senescence, observed in HP1γ$^{-/-}$ cells.

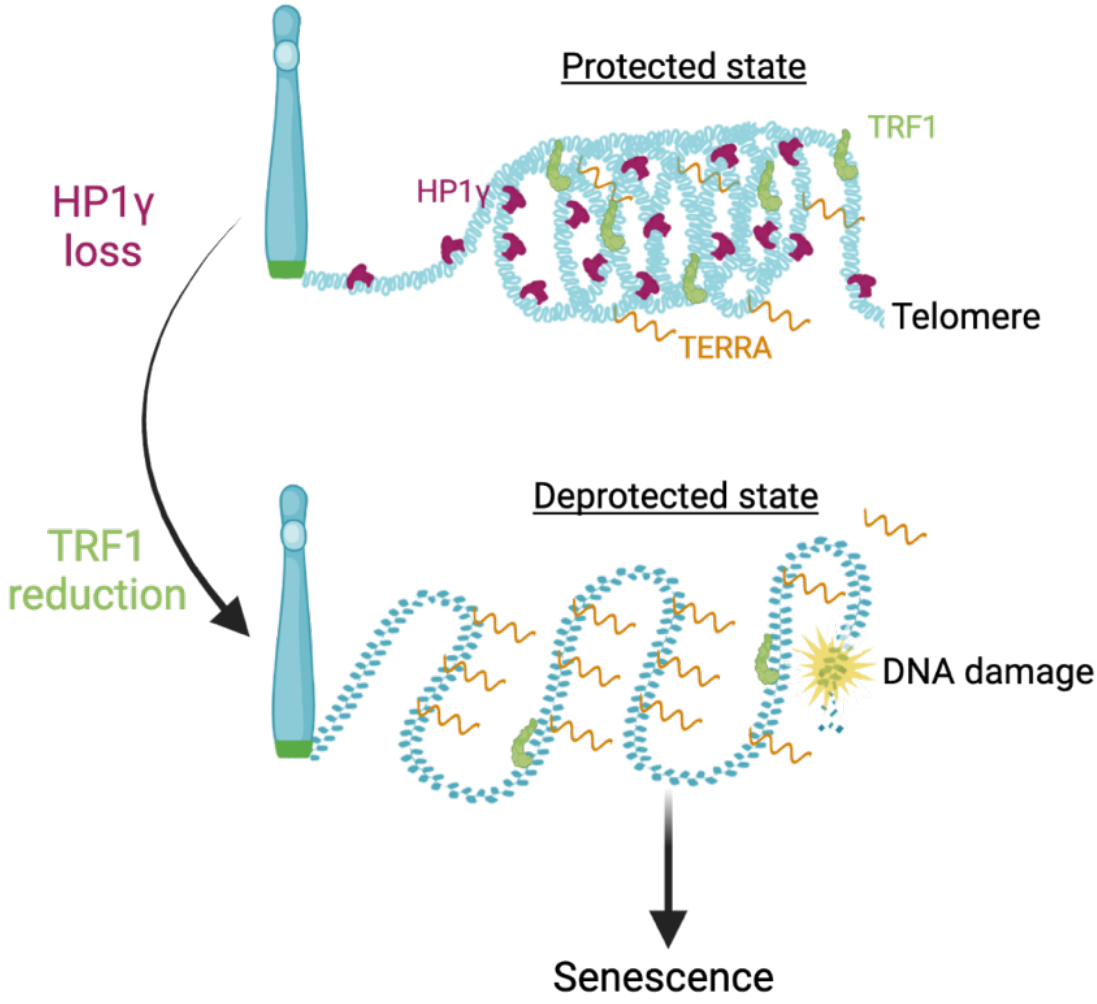


Figure 4.18 Model of the role of HP1γ on telomere maintenance. Illustration of the protective role of HP1γ on telomere stability. Loss of HP1γ results in a decrease of TRF1 levels. In combination with upregulation of TERRA and formation of telomeric DNA:RNA hybrids, telomeres undergo replication stress resulting in DNA damage and compromise of their integrity. Compromise of the protected telomeric chromatin state can lead to earlier onset of senescence.

Chapter 5. Main Findings

The focus of this PhD thesis has been the examination HP1γ's role on the regulation of sexually dimorphic differences at an early developmental stage and the involvement of this epigenetic factor in telomere maintenance. To this end, mouse embryonic fibroblasts (MEFs) deriving from E13.5 embryos were employed and the effect of HP1γ depletion was questioned among the sexes.

In Chapter 3, I focused on the effects of HP1γ on sexual dimorphism and how the differences among the sexes can have important implications early in development.

The main findings are:

- Male cell proliferation is dependent on HP1γ as its depletion leads to slower DNA replication, revealed by BrdU staining.

- Depletion of HP1γ results in upregulation of the cell cycle regulator *cdkn2a* in males but this is not recapitulated at protein level, at least, with the resolution of microscopic analysis.

- For both sexes, loss of HP1γ leads to earlier onset of senescence, measured by β-gal activity.

- Male embryo growth is dependent on HP1γ as its depletion results in smaller embryo size

- CUT & RUN analysis, revealed that HP1γ is enriched on the promoters of highly expressed genes in both sexes. HP1γ also binds on the bodies of sexually dimorphic genes, suggesting for either an activating, or a suppressive role, depending on the sex.

- Preliminary data of SLIC-CAGE suggest a dispensable role of HP1γ in inhibiting cryptic transcription

In Chapter 4, I interrogated the effects of HP1γ on telomere stability, especially through its transcriptional role. Potential sexually dimorphic differences were also taken into account.

The main findings include:

- The expression of a plethora of telomere-associated factors is regulated by HP1γ, with males being more sensitive to HP1γ depletion. Importantly, TRF1 of the shelterin complex, is downregulated in the absence of HP1γ in both sexes.

- Loss of HP1γ and TRF1 downregulation leads to a female-biased upregulation of TERRA expression. This upregulation in turn results in elevated levels of telomeric R-loops, as revealed by DRIP RT-qPCR.

- Increased telomeric R-loops and abnormal levels of TRF1 result in replication stress, detected as fragile telomeres. This is associated with telomeric DNA damage, revealed by co-localisation of γH2AX foci with telomeres.

- Accumulated replication stress and DNA damage at telomeres could be an underlying cause of the onset of senescence observed at early cell passages, detected by β-gal activity in the absence of HP1γ.

References

Abdallah, P. *et al.* (2009) 'A two-step model for senescence triggered by a single critically short telomere', *Nature Cell Biology 2009 11:8*, 11(8), pp. 988–993. Available at: https://doi.org/10.1038/ncb1911.

Abe, K. *et al.* (2011) 'Loss of Heterochromatin Protein 1 Gamma Reduces the Number of Primordial Germ Cells via Impaired Cell Cycle Progression in Mice1', *Biology of Reproduction*, 85(5), pp. 1013–1024. Available at: https://doi.org/10.1095/biolreprod.111.091512.

Abreu, E. *et al.* (2010) ' TIN2-Tethered TPP1 Recruits Human Telomerase to Telomeres In Vivo ', *Molecular and Cellular Biology*, 30(12), pp. 2971–2982. Available at: https://doi.org/10.1128/MCB.00240-10/ASSET/BE03BCA1-8A28-446B-904D-26CABC94C265/ASSETS/GRAPHIC/ZMB9991086440009.JPEG.

Ainsztein, A.M. *et al.* (1998) 'INCENP Centromere and Spindle Targeting: Identification of Essential Conserved Motifs and Involvement of Heterochromatin Protein HP1', *Journal of Cell Biology*, 143(7), pp. 1763–1774. Available at: https://doi.org/10.1083/JCB.143.7.1763.

Aird, K.M. and Zhang, R. (2013) 'Detection of senescence-associated heterochromatin foci (SAHF)', *Methods in molecular biology (Clifton, N.J.)*, 965, p. 185. Available at: https://doi.org/10.1007/978-1-62703-239-1_12.

Ait-Si-Ali, S. *et al.* (2004) 'A Suv39h-dependent mechanism for silencing S-phase genes in differentiating but not in cycling cells', *The EMBO Journal*, 23(3), pp. 605–615. Available at: https://doi.org/10.1038/SJ.EMBOJ.7600074.

Akaike, Y. *et al.* (2015) 'Homeodomain-interacting protein kinase 2 regulates DNA damage response through interacting with heterochromatin protein 1γ', *Oncogene*, 34(26), pp. 3463–3473. Available at: https://doi.org/10.1038/onc.2014.278.

Alam, H. *et al.* (2018) 'HP1γ Promotes Lung Adenocarcinoma by Downregulating the Transcription-Repressive Regulators NCOR2 and ZBTB7A', *Cancer Research*, 78(14), pp. 3834–3848. Available at: https://doi.org/10.1158/0008-5472.CAN-17-3571.

Alejandro, E.U. (2019) 'Males require estrogen signaling too: Sexual dimorphism in the regulation of glucose homeostasis by nuclear era', *Diabetes*, 68(3), pp. 471–473. Available at: https://doi.org/10.2337/DBI18-0046/-/DC1.

Allan, J. *et al.* (1984) 'Higher order structure in a short repeat length chromatin', *The Journal of Cell Biology*, 98(4), p. 1320. Available at: https://doi.org/10.1083/JCB.98.4.1320.

Allfrey V. G., Faulner, R., and Mirsky A. E. (1964) 'Acetylation and Methylation of Histones and their possible role in the regulation of RNA synthesis', *Proceedings of the National Academy of Sciences of the United States of America*, 51(5), p. 786. Available at: https://doi.org/10.1073/PNAS.51.5.786.

Allshire, R.C. *et al.* (1994) 'Position effect variegation at fission yeast centromeres', *Cell*, 76(1), pp. 157–169. Available at: https://doi.org/10.1016/0092-8674(94)90180-5.

Altschul, S.F. *et al.* (1997) 'Gapped BLAST and PSI-BLAST: a new generation of protein database search programs', *Nucleic acids research*, 25(17), pp. 3389–3402. Available at: https://doi.org/10.1093/NAR/25.17.3389.

Ameyar-Zazoua, M. *et al.* (2012) 'Argonaute proteins couple chromatin silencing to alternative splicing', *Nature Structural & Molecular Biology 2012 19:10*, 19(10), pp. 998–1004. Available at: https://doi.org/10.1038/nsmb.2373.
Anderson, L., Henderson, C. and Adachi, Y. (2001) 'Phosphorylation and Rapid Relocalization of 53BP1 to Nuclear Foci upon DNA Damage', *Molecular and Cellular Biology*, 21(5), p. 1719. Available at: https://doi.org/10.1128/MCB.21.5.1719-1729.2001.
Andrew, T. *et al.* (2006) 'Mapping Genetic Loci That Determine Leukocyte Telomere Length in a Large Sample of Unselected Female Sibling Pairs', *American Journal of Human Genetics*, 78(3), p. 480. Available at: https://doi.org/10.1086/500052.
Arnold, A.P. (2009) 'The organizational–activational hypothesis as the foundation for a unified theory of sexual differentiation of all mammalian tissues', *Hormones and Behavior*, 55(5), pp. 570–578. Available at: https://doi.org/10.1016/J.YHBEH.2009.03.011.
Arnold, A.P., Chen, X. and Itoh, Y. (2012) 'What a difference an X or Y makes: Sex chromosomes, gene dose, and epigenetics in sexual differentiation', *Handbook of Experimental Pharmacology*, 214, pp. 67–88. Available at: https://doi.org/10.1007/978-3-642-30726-3_4/FIGURES/5.
Arnoult, N., van Beneden, A. and Decottignies, A. (2012) 'Telomere length regulates TERRA levels through increased trimethylation of telomeric H3K9 and HP1α', *Nature Structural & Molecular Biology 2012 19:9*, 19(9), pp. 948–956. Available at: https://doi.org/10.1038/nsmb.2364.
Arora, R. *et al.* (2014) 'RNaseH1 regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells', *Nature Communications 2014 5:1*, 5(1), pp. 1–11. Available at: https://doi.org/10.1038/ncomms6220.
Aucott, R. *et al.* (2008) 'HP1-β is required for development of the cerebral neocortex and neuromuscular junctions', *Journal of Cell Biology*, 183(4), pp. 597–606. Available at: https://doi.org/10.1083/jcb.200804041.
Aydin, E. *et al.* (2015) 'A hypomorphic Cbx3 allele causes prenatal growth restriction and perinatal energy homeostasis defects', *Journal of Biosciences*, 40(2), pp. 325–338. Available at: https://doi.org/10.1007/s12038-015-9520-x.
Ayoub, N., Jeyasekharan, A.D. and Venkitaraman, A.R. (2009) 'Mobilization and recruitment of HP1β: A bimodal response to DNA breakage', *http://dx.doi.org/10.4161/cc.8.18.9486*, 8(18), pp. 2946–2951. Available at: https://doi.org/10.4161/CC.8.18.9486.
Azzalin, C.M. *et al.* (2007) 'Telomeric repeat-containing RNA and RNA surveillance factors at mammalian chromosome ends', *Science*, 318(5851), pp. 798–801. Available at: https://doi.org/10.1126/SCIENCE.1147182/SUPPL_FILE/AZZALIN.SOM.PDF.
Azzalin, C.M. and Lingner, J. (2008) 'Telomeres: The silence is broken', *http://dx.doi.org/10.4161/cc.7.9.5836*, 7(9), pp. 1161–1165. Available at: https://doi.org/10.4161/CC.7.9.5836.
Azzalin, C.M., Nergadze, S.G. and Giulotto, E. (2001) 'Human intrachromosomal telomeric-like repeats: sequence organization and mechanisms of origin', *Chromosoma 2001 110:2*, 110(2), pp. 75–82. Available at: https://doi.org/10.1007/S004120100135.
Bah, A. *et al.* (2012) 'The telomeric transcriptome of Schizosaccharomyces pombe', *Nucleic Acids Research*, 40(7), p. 2995. Available at: https://doi.org/10.1093/NAR/GKR1153.

Balaton, B.P. and Brown, C.J. (2016) 'Escape Artists of the X Chromosome', *Trends in Genetics*, 32(6), pp. 348–359. Available at: https://doi.org/10.1016/J.TIG.2016.03.007.
Baldi, S. *et al.* (2018) 'Genome-wide Rules of Nucleosome Phasing in Drosophila', *Molecular Cell*, 72(4), pp. 661-672.e4. Available at: https://doi.org/10.1016/J.MOLCEL.2018.09.032.
Balk, B. *et al.* (2013) 'Telomeric RNA-DNA hybrids affect telomere-length dynamics and senescence', *Nature Structural & Molecular Biology 2013 20:10*, 20(10), pp. 1199–1205. Available at: https://doi.org/10.1038/nsmb.2662.
Balwierz, P.J. *et al.* (2009) 'Methods for analyzing deep sequencing expression data: Constructing the human and mouse promoterome with deepCAGE data', *Genome Biology*, 10(7), pp. 1–21. Available at: https://doi.org/10.1186/GB-2009-10-7-R79/FIGURES/15.
Bandaria, J.N. *et al.* (2016) 'Shelterin Protects Chromosome Ends by Compacting Telomeric Chromatin', *Cell*, 164(4), pp. 735–746. Available at: https://doi.org/10.1016/J.CELL.2016.01.036.
Bannister, A.J. *et al.* (2001) 'Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain', *Nature*, 410(6824), pp. 120–124. Available at: https://doi.org/10.1038/35065138.
Bannister, A.J. *et al.* (2005) 'Spatial Distribution of Di- and Tri-methyl Lysine 36 of Histone H3 at Active Genes', *Journal of Biological Chemistry*, 280(18), pp. 17732–17736. Available at: https://doi.org/10.1074/JBC.M500796200.
Bannister, A.J. and Kouzarides, T. (2011) 'Regulation of chromatin by histone modifications', *Cell Research 2011 21:3*, 21(3), pp. 381–395. Available at: https://doi.org/10.1038/cr.2011.22.
Barr M. L and Bertram E. G. (1949) 'A Morphological Distinction between Neurones of the Male and Female, and the Behaviour of the Nucleolar Satellite during Accelerated Nucleoprotein Synthesis', *Nature*, 163(4148), pp. 676–677. Available at: https://doi.org/10.1038/163676a0.
Barral, A. and Déjardin, J. (2020) 'Telomeric Chromatin and TERRA', *Journal of Molecular Biology*, 432(15), pp. 4244–4256. Available at: https://doi.org/10.1016/J.JMB.2020.03.003.
Barrett, E.L.B. and Richardson, D.S. (2011) 'Sex differences in telomeres and lifespan', *Aging Cell*, 10(6), pp. 913–921. Available at: https://doi.org/https://doi.org/10.1111/j.1474-9726.2011.00741.x.
Barski, A. *et al.* (2007) 'High-Resolution Profiling of Histone Methylations in the Human Genome', *Cell*, 129(4), pp. 823–837. Available at: https://doi.org/10.1016/J.CELL.2007.05.009.
Batut, P. *et al.* (2013) 'High-fidelity promoter profiling reveals widespread alternative promoter usage and transposon-driven developmental gene expression', *Genome Research*, 23(1), p. 169. Available at: https://doi.org/10.1101/GR.139618.112.
Baumann, P. and Cech, T.R. (2001) 'Pot1, the putative telomere end-binding protein in fission yeast and humans', *Science*, 292(5519), pp. 1171–1175. Available at: https://doi.org/10.1126/SCIENCE.1060036/SUPPL_FILE/1060036S2_THUMB.GIF.
Baur, J.A. *et al.* (2001) 'Telomere position effect in human cells', *Science*, 292(5524), pp. 2075–2077. Available at: https://doi.org/10.1126/SCIENCE.1062329/SUPPL_FILE/105942018S2_THUMB.GIF.

Beagrie, R.A. *et al.* (2017) 'Complex multi-enhancer contacts captured by genome architecture mapping', *Nature*, 543(7646), pp. 519–524. Available at: https://doi.org/10.1038/nature21411.
Bechter, O.E. *et al.* (2003) 'Homologous recombination in human telomerase-positive and ALT cells occurs with the same frequency', *EMBO reports*, 4(12), pp. 1138–1143. Available at: https://doi.org/10.1038/SJ.EMBOR.7400027.
Bedoyan, J.K. *et al.* (1996) 'Condensation of Rat Telomere-specific Nucleosomal Arrays Containing Unusually Short DNA Repeats and Histone H1', *Journal of Biological Chemistry*, 271(31), pp. 18485–18493. Available at: https://doi.org/10.1074/JBC.271.31.18485.
Beishline, K. *et al.* (2017) 'CTCF driven TERRA transcription facilitates completion of telomere DNA replication', *Nature Communications 2017 8:1*, 8(1), pp. 1–10. Available at: https://doi.org/10.1038/s41467-017-02212-w.
Belmont, A.S. *et al.* (1989) 'Large-scale chromatin structural domains within mitotic and interphase chromosomes in vivo and in vitro', *Chromosoma*, 98(2), pp. 129–143. Available at: https://doi.org/10.1007/BF00291049.
Belmont, A.S. *et al.* (1999) 'Large-scale chromatin structure and function', *Current Opinion in Cell Biology*, 11(3), pp. 307–311. Available at: https://doi.org/10.1016/S0955-0674(99)80041-6.
Benarroch-Popivker, D. *et al.* (2016) 'TRF2-Mediated Control of Telomere DNA Topology as a Mechanism for Chromosome-End Protection', *Molecular Cell*, 61(2), pp. 274–286. Available at: https://doi.org/10.1016/J.MOLCEL.2015.12.009.
Benetti, R., García-Cao, M. and Blasco, M.A. (2007) 'Telomere length regulates the epigenetic status of mammalian telomeres and subtelomeres', *Nature Genetics 2007 39:2*, 39(2), pp. 243–250. Available at: https://doi.org/10.1038/ng1952.
Berletch, J.B., Yang, F. and Disteche, C.M. (2010) 'Escape from X inactivation in mice and humans', *Genome Biology*, 11(6), pp. 1–7. Available at: https://doi.org/10.1186/GB-2010-11-6-213/FIGURES/2.
Bernardi, R. and Pandolfi, P.P. (2014) 'A dialog on the first 20 years of PML research and the next 20 ahead', *Frontiers in Oncology*, 4 FEB, p. 23. Available at: https://doi.org/10.3389/FONC.2014.00023/BIBTEX.
Bernstein, J.A. *et al.* (2002) 'Global analysis of mRNA decay and abundance in Escherichia coli at single-gene resolution using two-color fluorescent DNA microarrays', *Proceedings of the National Academy of Sciences of the United States of America*, 99(15), pp. 9697–9702. Available at: https://doi.org/10.1073/PNAS.112318199/SUPPL_FILE/3181TABLE6.XLS.
Berre, G. le *et al.* (2019) 'Repression of TERRA Expression by Subtelomeric DNA Methylation Is Dependent on NRF1 Binding', *International Journal of Molecular Sciences 2019, Vol. 20, Page 2791*, 20(11), p. 2791. Available at: https://doi.org/10.3390/IJMS20112791.
Bickmore, W.A. and van Steensel, B. (2013) 'Genome architecture: domain organization of interphase chromosomes', *Cell*, 152(6), pp. 1270–1284. Available at: https://doi.org/10.1016/J.CELL.2013.02.001.
Biffi, G. *et al.* (2013) 'Quantitative visualization of DNA G-quadruplex structures in human cells', *Nature Chemistry 2013 5:3*, 5(3), pp. 182–186. Available at: https://doi.org/10.1038/nchem.1548.
Bilaud, T. *et al.* (1997) 'Telomeric localization of TRF2, a novel human telobox protein', *Nature Genetics*, 17(2), pp. 236–239. Available at: https://doi.org/10.1038/ng1097-236.

Blackburn, E.H. and Gall, J.G. (1978) 'A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in Tetrahymena', *Journal of Molecular Biology*, 120(1), pp. 33–53. Available at: https://doi.org/10.1016/0022-2836(78)90294-2.
Blasco, M.A. (2007) 'The epigenetic regulation of mammalian telomeres', *Nature Reviews Genetics 2007 8:4*, 8(4), pp. 299–309. Available at: https://doi.org/10.1038/nrg2047.
Boguslawski, S.J. *et al.* (1986) 'Characterization of monoclonal antibody to DNA · RNA and its application to immunodetection of hybrids', *Journal of Immunological Methods*, 89(1), pp. 123–130. Available at: https://doi.org/10.1016/0022-1759(86)90040-2.
Bosch-Presegué, L. *et al.* (2017) 'Mammalian HP1 Isoforms Have Specific Roles in Heterochromatin Structure and Organization', *Cell Reports*, 21(8), pp. 2048–2057. Available at: https://doi.org/10.1016/J.CELREP.2017.10.092.
Bou-Nader, C. *et al.* (2022) 'Structural basis of R-loop recognition by the S9.6 monoclonal antibody', *Nature Communications*, 13(1), pp. 1–14. Available at: https://doi.org/10.1038/s41467-022-29187-7.
Bourgeron, T. *et al.* (2015) 'The asymmetry of telomere replication contributes to replicative senescence heterogeneity', *Scientific Reports 2015 5:1*, 5(1), pp. 1–11. Available at: https://doi.org/10.1038/srep15326.
Bouwman, P. *et al.* (2010) '53BP1 loss rescues BRCA1 deficiency and is associated with triple-negative and BRCA-mutated breast cancers', *Nature structural & molecular biology*, 17(6), p. 688. Available at: https://doi.org/10.1038/NSMB.1831.
Brasher, S. v. *et al.* (2000) 'The structure of mouse HP1 suggests a unique mode of single peptide recognition by the shadow chromo domain dimer', *The EMBO Journal*, 19(7), pp. 1587–1597. Available at: https://doi.org/10.1093/EMBOJ/19.7.1587.
Brault, M.E. and Autexier, C. (2011) 'Telomeric recombination induced by dysfunctional telomeres', *Molecular Biology of the Cell*, 22(2), pp. 179–188. Available at: https://doi.org/10.1091/MBC.E10-02-0173/ASSET/IMAGES/LARGE/179FIG5.JPEG.
Brown, J.P. *et al.* (2010) 'HP1γ function is required for male germ cell survival and spermatogenesis', *Epigenetics & Chromatin*, 3(1), p. 9. Available at: https://doi.org/10.1186/1756-8935-3-9.
Bryan, T.M. *et al.* (1997) 'Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines', *Nature Medicine 1997 3:11*, 3(11), pp. 1271–1274. Available at: https://doi.org/10.1038/nm1197-1271.
Bulut-Karslioglu, A. *et al.* (2012) 'A transcription factor–based mechanism for mouse heterochromatin formation', *Nature Structural & Molecular Biology*, 19(10), pp. 1023–1030. Available at: https://doi.org/10.1038/nsmb.2382.
Bulut-Karslioglu, A. *et al.* (2014) 'Suv39h-Dependent H3K9me3 Marks Intact Retrotransposons and Silences LINE Elements in Mouse Embryonic Stem Cells', *Molecular Cell*, 55(2), pp. 277–290. Available at: https://doi.org/10.1016/J.MOLCEL.2014.05.029.
Bunting, S.F. *et al.* (2010) '53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks', *Cell*, 141(2), pp. 243–254. Available at: https://doi.org/10.1016/J.CELL.2010.03.012.
Burge, S. *et al.* (2006) 'Quadruplex DNA: sequence, topology and structure', *Nucleic Acids Research*, 34(19), p. 5402. Available at: https://doi.org/10.1093/NAR/GKL655.

Burgoyne, P.S. (1993) ‘A Y-chromosomal effect on blastocyst cell number in mice’, *Development*, 117(1), pp. 341–345. Available at: https://doi.org/10.1242/DEV.117.1.341.
Burgoyne, P.S. *et al.* (1995) ‘The genetic basis of XX-XY differences present before gonadal sex differentiation in the mouse’, *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 350(1333). Available at: https://doi.org/10.1098/RSTB.1995.0159.
Cacchione, S., Cerone, M.A. and Savino, M. (1997) ‘In vitro low propensity to form nucleosomes of four telomeric sequences’, *FEBS Letters*, 400(1), pp. 37–41. Available at: https://doi.org/https://doi.org/10.1016/S0014-5793(96)01318-X.
Canudas, S. *et al.* (2011) ‘A role for heterochromatin protein 1γ at human telomeres’, *Genes & Development*, 25(17), p. 1807. Available at: https://doi.org/10.1101/GAD.17325211.
Canzio, D. *et al.* (2011) ‘Chromodomain-mediated oligomerization of HP1 suggests a nucleosome bridging mechanism for heterochromatin assembly’, *Molecular cell*, 41(1), p. 67. Available at: https://doi.org/10.1016/J.MOLCEL.2010.12.016.
Canzio, D., Larson, A. and Narlikar, G.J. (2014) ‘Mechanisms of functional promiscuity by HP1 proteins’, *Trends in Cell Biology*, 24(6), pp. 377–386. Available at: https://doi.org/10.1016/J.TCB.2014.01.002.
Carninci, P. *et al.* (2006) ‘Genome-wide analysis of mammalian promoter architecture and evolution’, *Nature Genetics*, 38(6), pp. 626–635. Available at: https://doi.org/10.1038/ng1789.
Carrel, L. and Willard, H.F. (1999) ‘Heterogeneous gene expression from the inactive X chromosome: An X-linked gene that escapes X inactivation in some human cell lines but is inactivated in others’, *Proceedings of the National Academy of Sciences of the United States of America*, 96(13), pp. 7364–7369. Available at: https://doi.org/10.1073/PNAS.96.13.7364/ASSET/8792E1DB-6F0F-44A6-88FD-08DA28EE070D/ASSETS/GRAPHIC/PQ1390844003.JPEG.
Carrel, L. and Willard, H.F. (2005) ‘X-inactivation profile reveals extensive variability in X-linked gene expression in females’, *Nature 2005 434:7031*, 434(7031), pp. 400–404. Available at: https://doi.org/10.1038/nature03479.
Carrozza, M.J. *et al.* (2005) ‘Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription’, *Cell*, 123(4), pp. 581–592. Available at: https://doi.org/10.1016/j.cell.2005.10.023.
Cattanach, B.M. (1974) ‘Position effect variegation in the mouse’, *Genetics Research*, 23(3), pp. 291–306. Available at: https://doi.org/10.1017/S0016672300014932.
Celli, G.B., Denchi, E.L. and de Lange, T. (2006) ‘Ku70 stimulates fusion of dysfunctional telomeres yet protects chromosome ends from homologous recombination’, *Nature Cell Biology*, 8(8), pp. 885–890. Available at: https://doi.org/10.1038/ncb1444.
Cesare, A.J. and Reddel, R.R. (2010) ‘Alternative lengthening of telomeres: models, mechanisms and implications’, *Nature Reviews Genetics*, 11(5), pp. 319–330. Available at: https://doi.org/10.1038/nrg2763.
Chadwick, B.P. and Willard, H.F. (2003) ‘Chromatin of the Barr body: histone and non-histone proteins associated with or excluded from the inactive X chromosome’, *Human Molecular Genetics*, 12(17), pp. 2167–2178. Available at: https://doi.org/10.1093/HMG/DDG229.

Chang, S. *et al.* (2003) 'Telomere-based crisis: functional differences between telomerase activation and ALT in tumor progression', *Genes & Development*, 17(1), p. 88. Available at: https://doi.org/10.1101/GAD.1029903.
Chapman, J.R. *et al.* (2013) 'RIF1 Is Essential for 53BP1-Dependent Nonhomologous End Joining and Suppression of DNA Double-Strand Break Resection', *Molecular Cell*, 49(5), p. 858. Available at: https://doi.org/10.1016/J.MOLCEL.2013.01.002.
Chen, K. *et al.* (2016) 'The Overlooked Fact: Fundamental Need for Spike-In Control for Virtually All Genome-Wide Analyses', *Molecular and Cellular Biology*, 36(5), pp. 662–667. Available at: https://doi.org/10.1128/MCB.00970-14/ASSET/1024259B-B9E5-4D3E-8689-6E6DD5D31F20/ASSETS/GRAPHIC/ZMB9991011380003.JPEG.
Chen, L.Y., Redon, S. and Lingner, J. (2012) 'The human CST complex is a terminator of telomerase activity', *Nature 2012 488:7412*, 488(7412), pp. 540–544. Available at: https://doi.org/10.1038/nature11269.
Chéné, I. du *et al.* (2007) 'Suv39H1 and HP1γ are responsible for chromatin-mediated HIV-1 transcriptional silencing and post-integration latency', *The EMBO Journal*, 26(2), pp. 424–435. Available at: https://doi.org/https://doi.org/10.1038/sj.emboj.7601517.
Cheutin, T. *et al.* (2003) 'Maintenance of Stable Heterochromatin Domains by Dynamic HP1 Binding', *Science*, 299(5607), pp. 721–725. Available at: https://doi.org/10.1126/science.1078572.
Choi, J.D., Park, M.A. and Lee, J.S. (2012) 'Suppression and recovery of BRCA1-mediated transcription by HP1γ via modulation of promoter occupancy', *Nucleic Acids Research*, 40(22), pp. 11321–11338. Available at: https://doi.org/10.1093/NAR/GKS947.
Chong, L. *et al.* (1995) 'A Human Telomeric Protein', *Science*, 270(5242), pp. 1663–1667. Available at: https://doi.org/10.1126/SCIENCE.270.5242.1663.
Chow, T.T. *et al.* (2012) 'Early and late steps in telomere overhang processing in normal human cells: the position of the final RNA primer drives telomere shortening', *Genes & Development*, 26(11), pp. 1167–1178. Available at: https://doi.org/10.1101/GAD.187211.112.
Chu, H.-P. *et al.* (2017) 'PAR-TERRA directs homologous sex chromosome pairing', *Nature Structural & Molecular Biology*, 24(8), pp. 620–631. Available at: https://doi.org/10.1038/nsmb.3432.
Chu, H.P. *et al.* (2017) 'TERRA RNA antagonizes ATRX and protects telomeres', *Cell*, 170(1), p. 86. Available at: https://doi.org/10.1016/J.CELL.2017.06.017.
Claussin, C. and Chang, M. (2015) 'The many facets of homologous recombination at telomeres', *Microbial Cell*, 2(9), p. 308. Available at: https://doi.org/10.15698/MIC2015.09.224.
Clocchiatti, A. *et al.* (2016) 'Sexual dimorphism in cancer', *Nature Reviews Cancer*, 16(5), pp. 330–339. Available at: https://doi.org/10.1038/nrc.2016.30.
Connor, J.M. *et al.* (1986) 'Assignment of the gene for dyskeratosis congenita to Xq28', *Human Genetics*, 72(4), pp. 348–351. Available at: https://doi.org/10.1007/BF00290963.
Conomos, D. *et al.* (2012) 'Variant repeats are interspersed throughout the telomeres and recruit nuclear receptors in ALT cells', *The Journal of Cell Biology*, 199(6), p. 893. Available at: https://doi.org/10.1083/JCB.201207189.
Conomos, D., Pickett, H.A. and Reddel, R.R. (2013a) 'Alternative lengthening of telomeres: Remodeling the telomere architecture', *Frontiers in Oncology*, 3 FEB, p. 27. Available at: https://doi.org/10.3389/FONC.2013.00027/BIBTEX.

Conomos, D., Pickett, H.A. and Reddel, R.R. (2013b) 'Alternative lengthening of telomeres: remodeling the telomere architecture', *Frontiers in Oncology*, 3. Available at: https://doi.org/10.3389/FONC.2013.00027.
Cooper, J.P. *et al.* (1997) 'Regulation of telomere length and function by a Myb-domain protein in fission yeast', *Nature*, 385(6618), pp. 744–747. Available at: https://doi.org/10.1038/385744a0.
Cowieson, N.P. *et al.* (2000) 'Dimerisation of a chromo shadow domain and distinctions from the chromodomain as revealed by structural analysis', *Current Biology*, 10(9), pp. 517–525. Available at: https://doi.org/10.1016/S0960-9822(00)00467-X.
Cox, J. *et al.* (2014) 'Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ', *Molecular & cellular proteomics : MCP*, 13(9), pp. 2513–2526. Available at: https://doi.org/10.1074/MCP.M113.031591.
Cox, J. and Mann, M. (2008) 'MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification', *Nature Biotechnology 2008 26:12*, 26(12), pp. 1367–1372. Available at: https://doi.org/10.1038/nbt.1511.
Cremer, T. and Cremer, M. (2010) 'Chromosome Territories', *Cold Spring Harbor Perspectives in Biology*, 2(3). Available at: https://doi.org/10.1101/CSHPERSPECT.A003889.
Cristofalo, V.J. and Pignolo, R.J. (1993) 'Replicative senescence of human fibroblast-like cells in culture', *https://doi.org/10.1152/physrev.1993.73.3.617*, 73(3), pp. 617–638. Available at: https://doi.org/10.1152/PHYSREV.1993.73.3.617.
Cristofari, G. *et al.* (2007) 'Human Telomerase RNA Accumulation in Cajal Bodies Facilitates Telomerase Recruitment to Telomeres and Telomere Elongation', *Molecular Cell*, 27(6), pp. 882–889. Available at: https://doi.org/10.1016/J.MOLCEL.2007.07.020.
Crossley, M.P., Bocek, M. and Cimprich, K.A. (2019) 'R-loops as cellular regulators and genomic threats', *Molecular cell*, 73(3), p. 398. Available at: https://doi.org/10.1016/J.MOLCEL.2019.01.024.
Cubiles, M.D. *et al.* (2018) 'Epigenetic features of human telomeres', *Nucleic Acids Research*, 46(5), pp. 2347–2355. Available at: https://doi.org/10.1093/NAR/GKY006.
Cusanelli, E., Romero, C.A.P. and Chartrand, P. (2013) 'Telomeric noncoding RNA TERRA is induced by telomere shortening to nucleate telomerase molecules at short telomeres', *Molecular cell*, 51(6), pp. 780–791. Available at: https://doi.org/10.1016/J.MOLCEL.2013.08.029.
Cvetesic, N. *et al.* (2018) 'SLIC-CAGE: High-resolution transcription start site mapping using nanogram-levels of total RNA', *Genome Research*, 28(12), pp. 1931–1942. Available at: https://doi.org/10.1101/GR.235937.118/-/DC1.
Deasy, B.M. *et al.* (2007) 'A role for cell sex in stem cell–mediated skeletal muscle regeneration: female cells have higher muscle regeneration efficiency', *Journal of Cell Biology*, 177(1), pp. 73–86. Available at: https://doi.org/10.1083/jcb.200612094.
Debacq-Chainiaux, F. *et al.* (2009) 'Protocols to detect senescence-associated beta-galactosidase (SA-βgal) activity, a biomarker of senescent cells in culture and in vivo', *Nature Protocols*, 4(12), pp. 1798–1806. Available at: https://doi.org/10.1038/nprot.2009.191.

Déjardin, J. and Kingston, R.E. (2009) 'Purification of Proteins Associated with Specific Genomic Loci', *Cell*, 136(1), pp. 175–186. Available at: https://doi.org/10.1016/J.CELL.2008.11.045.
Denchi, E.L. and de Lange, T. (2007) 'Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1', *Nature 2007 448:7157*, 448(7157), pp. 1068–1071. Available at: https://doi.org/10.1038/nature06065.
Deng, X. *et al.* (2009) 'Dosage Regulation of the Active X Chromosome in Human Triploid Cells', *PLoS Genetics*, 5(12), p. 1000751. Available at: https://doi.org/10.1371/JOURNAL.PGEN.1000751.
Deng, Z. *et al.* (2009) 'TERRA RNA Binding to TRF2 Facilitates Heterochromatin Formation and ORC Recruitment at Telomeres', *Molecular Cell*, 35(4), pp. 403–413. Available at: https://doi.org/https://doi.org/10.1016/j.molcel.2009.06.025.
Deng, Z. *et al.* (2012) 'A role for CTCF and cohesin in subtelomere chromatin organization, TERRA transcription, and telomere end protection', *The EMBO Journal*, 31(21), pp. 4165–4178. Available at: https://doi.org/10.1038/EMBOJ.2012.266.
Dewey, W.C. and Miller, H.H. (1969) 'X-ray induction of chromatid exchanges in mitotic and G1 Chinese hamster cells pretreated with Colcemid', *Experimental Cell Research*, 57(1), pp. 63–70. Available at: https://doi.org/10.1016/0014-4827(69)90366-8.
Dillon, N. and Festenstein, R. (2002) 'Unravelling heterochromatin: competition between positive and negative factors regulates accessibility', *Trends in Genetics*, 18(5), pp. 252–258. Available at: https://doi.org/10.1016/S0168-9525(02)02648-3.
Dimri, G.P. *et al.* (1995) 'A biomarker that identifies senescent human cells in culture and in aging skin in vivo.', *Proceedings of the National Academy of Sciences of the United States of America*, 92(20), p. 9363. Available at: https://doi.org/10.1073/PNAS.92.20.9363.
Dixon, J.R. *et al.* (2012) 'Topological domains in mammalian genomes identified by analysis of chromatin interactions', *Nature*, 485(7398), pp. 376–380. Available at: https://doi.org/10.1038/nature11082.
Dixon, J.R. *et al.* (2015) 'Chromatin architecture reorganization during stem cell differentiation', *Nature 2015 518:7539*, 518(7539), pp. 331–336. Available at: https://doi.org/10.1038/nature14222.
Doksani, Y. *et al.* (2013) 'Super-Resolution Fluorescence Imaging of Telomeres Reveals TRF2-Dependent T-loop Formation', *Cell*, 155(2), pp. 345–356. Available at: https://doi.org/10.1016/J.CELL.2013.09.048.
Downs, K.M. and Davies, T. (1993) 'Staging of gastrulating mouse embryos by morphological landmarks in the dissecting microscope', *Development*, 118(4), pp. 1255–1266. Available at: https://doi.org/10.1242/DEV.118.4.1255.
Draskovic, I. *et al.* (2009) 'Probing PML body function in ALT cells reveals spatiotemporal requirements for telomere recombination', *Proceedings of the National Academy of Sciences of the United States of America*, 106(37), pp. 15726–15731. Available at: https://doi.org/10.1073/PNAS.0907689106/SUPPL_FILE/0907689106SI.PDF.
Dupont, C., Armant, D.R. and Brenner, C.A. (2009) 'Epigenetics: Definition, mechanisms and clinical perspective', *Seminars in Reproductive Medicine*, 27(5), pp. 351–357. Available at: https://doi.org/10.1055/S-0029-1237423/ID/44.
Elgin, S.C.R. and Reuter, G. (2013) 'Position-Effect Variegation, Heterochromatin Formation, and Gene Silencing in Drosophila', *Cold Spring Harbor Perspectives in*

Biology, 5(8), p. a017780. Available at: https://doi.org/10.1101/CSHPERSPECT.A017780.
Eymin, B. *et al.* (2003) 'p14ARF induces G2 arrest and apoptosis independently of p53 leading to regression of tumours established in nude mice', *Oncogene*, 22(12), pp. 1822–1835. Available at: https://doi.org/10.1038/sj.onc.1206303.
Ezkurdia, I. *et al.* (2012) 'Comparative Proteomics Reveals a Significant Bias Toward Alternative Protein Isoforms with Conserved Structure and Function', *Molecular Biology and Evolution*, 29(9), p. 2265. Available at: https://doi.org/10.1093/MOLBEV/MSS100.
Fairall, L. *et al.* (2001) 'Structure of the TRFH Dimerization Domain of the Human Telomeric Proteins TRF1 and TRF2', *Molecular Cell*, 8(2), pp. 351–361. Available at: https://doi.org/10.1016/S1097-2765(01)00321-5.
Fajkus, J., Sýkorová, E. and Leitch, A.R. (2005) 'Telomeres in evolution and evolution of telomeres', *Chromosome Research*, 13(5), pp. 469–479. Available at: https://doi.org/10.1007/s10577-005-0997-2.
Fanti, L. *et al.* (1998) 'The Heterochromatin Protein 1 Prevents Telomere Fusions in Drosophila', *Molecular Cell*, 2(5), pp. 527–538. Available at: https://doi.org/10.1016/S1097-2765(00)80152-5.
Feng, J. *et al.* (1995) 'The RNA Component of Human Telomerase', *Science*, 269(5228), pp. 1236–1241. Available at: https://doi.org/10.1126/SCIENCE.7544491.
Feretzaki, M. *et al.* (2020) 'RAD51-dependent recruitment of TERRA lncRNA to telomeres through R-loops', *Nature*, 587(7833), pp. 303–308. Available at: https://doi.org/10.1038/s41586-020-2815-6.
Feretzaki, M. and Lingner, J. (2017) 'A practical qPCR approach to detect TERRA, the elusive telomeric repeat-containing RNA', *Methods*, 114, pp. 39–45. Available at: https://doi.org/10.1016/J.YMETH.2016.08.004.
Feretzaki, M., Nunes, P.R. and Lingner, J. (2019) 'Expression and differential regulation of human TERRA at several chromosome ends', *RNA*, 25(11), pp. 1470–1480. Available at: https://doi.org/10.1261/RNA.072322.119/-/DC1.
Festenstein, R. *et al.* (1996) 'Locus Control Region Function and Heterochromatin-Induced Position Effect Variegation', *Science*, 271(5252), pp. 1123–1125. Available at: https://doi.org/10.1126/science.271.5252.1123.
Festenstein, R. *et al.* (2003) 'Modulation of heterochromatin protein 1 dynamics in primary mammalian cells', *Science*, 299(5607), pp. 719–721. Available at: https://doi.org/10.1126/SCIENCE.1078694/SUPPL_FILE/FESTENSTEIN2.PDF.
Filesi, I. *et al.* (2000) 'The main role of the sequence-dependent DNA elasticity in determining the free energy of nucleosome formation on telomeric DNAs', *Biophysical Chemistry*, 83(3), pp. 223–237. Available at: https://doi.org/10.1016/S0301-4622(99)00143-X.
Flynn, R.L. *et al.* (2011) 'TERRA and hnRNPA1 orchestrate an RPA-to-POT1 switch on telomeric single-stranded DNA', *Nature 2011 471:7339*, 471(7339), pp. 532–536. Available at: https://doi.org/10.1038/nature09772.
Fodor, B.D. *et al.* (2006) 'Jmjd2b antagonizes H3K9 trimethylation at pericentric heterochromatin in mammalian cells', *Genes & Development*, 20(12), p. 1557. Available at: https://doi.org/10.1101/GAD.388206.
Frescas, D. *et al.* (2008) 'KDM2A represses transcription of centromeric satellite repeats and maintains the heterochromatic state', *http://dx.doi.org/10.4161/cc.7.22.7062*, 7(22), pp. 3539–3547. Available at: https://doi.org/10.4161/CC.7.22.7062.

Fritsch, L. *et al.* (2010) 'A subset of the histone H3 lysine 9 methyltransferases Suv39h1, G9a, GLP, and SETDB1 participate in a multimeric complex', *Molecular cell*, 37(1), pp. 46–56. Available at: https://doi.org/10.1016/J.MOLCEL.2009.12.017.
Fudenberg, G. *et al.* (2016) 'Formation of Chromosomal Domains by Loop Extrusion', *Cell Reports*, 15(9), pp. 2038–2049. Available at: https://doi.org/10.1016/J.CELREP.2016.04.085.
Fyodorov, D. v. *et al.* (2017) 'Emerging roles of linker histones in regulating chromatin structure and function', *Nature Reviews Molecular Cell Biology 2017 19:3*, 19(3), pp. 192–206. Available at: https://doi.org/10.1038/nrm.2017.94.
García-Cao, M. *et al.* (2003) 'Epigenetic regulation of telomere length in mammalian cells by the Suv39h1 and Suv39h2 histone methyltransferases', *Nature Genetics 2004 36:1*, 36(1), pp. 94–99. Available at: https://doi.org/10.1038/ng1278.
García-Muse, T. and Aguilera, A. (2019) 'R Loops: From Physiological to Pathological Roles', *Cell*, 179(3), pp. 604–618. Available at: https://doi.org/10.1016/J.CELL.2019.08.055.
Garrido-Ramos, M.A. (2017) 'Satellite DNA: An Evolving Topic', *Genes 2017, Vol. 8, Page 230*, 8(9), p. 230. Available at: https://doi.org/10.3390/GENES8090230.
Gauchier, M. *et al.* (2019) 'SETDB1-dependent heterochromatin stimulates alternative lengthening of telomeres', *Science Advances*, 5(5). Available at: https://doi.org/10.1126/SCIADV.AAV3673/SUPPL_FILE/AAV3673_TABLE_S2.XLSX.
Ghanim, G.E. *et al.* (2021) 'Structure of human telomerase holoenzyme with bound telomeric DNA', *Nature*, 593(7859), pp. 449–453. Available at: https://doi.org/10.1038/s41586-021-03415-4.
Gilbert SF. Developmental Biology. (2000) 'Developmental Biology', *Development Biology: 6th edition*, (Hall 1988), p. http://www.ncbi.nlm.nih.gov/books/NBK10056/. Available at: https://www.ncbi.nlm.nih.gov/books/NBK9983/ (Accessed: 26 October 2022).
Gillotin, S. (2018) 'Isolation of Chromatin-bound Proteins from Subcellular Fractions for Biochemical Analysis', *Bio-protocol*, 8(19). Available at: https://doi.org/10.21769/BIOPROTOC.3035.
Glousker, G. and Lingner, J. (2021) 'Challenging endings: How telomeres prevent fragility', *BioEssays*, 43(10), p. 2100157. Available at: https://doi.org/10.1002/BIES.202100157.
Gondos B. (1980) *Development and differentiation of the testis and male reproductive tract*, *Raven Press, New York*. New York: Raven Press. Available at: https://cir.nii.ac.jp/crid/1570009751282394240 (Accessed: 25 October 2022).
Gonzalo, S. *et al.* (2005) 'Role of the RB1 family in stabilizing histone methylation at constitutive heterochromatin', *Nature Cell Biology 2005 7:4*, 7(4), pp. 420–428. Available at: https://doi.org/10.1038/ncb1235.
Gonzalo, S. *et al.* (2006) 'DNA methyltransferases control telomere length and telomere recombination in mammalian cells', *Nature Cell Biology 2006 8:4*, 8(4), pp. 416–424. Available at: https://doi.org/10.1038/ncb1386.
Gorbunova, V. and Seluanov, A. (2009) 'Coevolution of telomerase activity and body mass in mammals: From mice to beavers', *Mechanisms of Ageing and Development*, 130(1–2), pp. 3–9. Available at: https://doi.org/10.1016/J.MAD.2008.02.008.
Gottschling, D.E. *et al.* (1990) 'Position effect at S. cerevisiae telomeres: Reversible repression of Pol II transcription', *Cell*, 63(4), pp. 751–762. Available at: https://doi.org/10.1016/0092-8674(90)90141-Z.

Graf, M. *et al.* (2017) 'Telomere Length Determines TERRA and R-Loop Regulation through the Cell Cycle', *Cell*, 170(1), pp. 72-85.e14. Available at: https://doi.org/10.1016/J.CELL.2017.06.006.
Greider, C.W. and Blackburn, E.H. (1985) 'Identification of a specific telomere terminal transferase activity in tetrahymena extracts', *Cell*, 43(2), pp. 405–413. Available at: https://doi.org/10.1016/0092-8674(85)90170-9.
Grewal, S.I.S. and Jia, S. (2007) 'Heterochromatin revisited', *Nature Reviews Genetics*, 8(1), pp. 35–46. Available at: https://doi.org/10.1038/nrg2008.
Grewal, S.I.S. and Moazed, D. (2003) 'Heterochromatin and epigenetic control of gene expression', *Science*, 301(5634), pp. 798–802. Available at: https://doi.org/10.1126/SCIENCE.1086887/ASSET/AAD926D4-CF54-4847-9961-425023981479/ASSETS/GRAPHIC/SE3131756004.JPEG.
Griffith, J.D. *et al.* (1999) 'Mammalian Telomeres End in a Large Duplex Loop', *Cell*, 97(4), pp. 503–514. Available at: https://doi.org/10.1016/S0092-8674(00)80760-6.
Grobelny, J. v., Godwin, A.K. and Broccoli, D. (2000) 'ALT-associated PML bodies are present in viable cells and are enriched in cells in the G(2)/M phase of the cell cycle', *Journal of Cell Science*, 113(24), pp. 4577–4585. Available at: https://doi.org/10.1242/JCS.113.24.4577.
Grunstein, M. (1998) 'Yeast Heterochromatin: Regulation of Its Assembly and Inheritance by Histones', *Cell*, 93(3), pp. 325–328. Available at: https://doi.org/10.1016/S0092-8674(00)81160-5.
Guacci, V., Koshland, D. and Strunnikov, A. (1997) 'A Direct Link between Sister Chromatid Cohesion and Chromosome Condensation Revealed through the Analysis of MCD1 in S. cerevisiae', *Cell*, 91(1), pp. 47–57. Available at: https://doi.org/10.1016/S0092-8674(01)80008-8.
Ha, N. *et al.* (2014) 'HP-1γ controls high-affinity antibody response to T-dependent antigens', *Frontiers in Immunology*, 5(JUN), p. 271. Available at: https://doi.org/10.3389/FIMMU.2014.00271/BIBTEX.
Haberle, V. *et al.* (2014) 'Two independent transcription initiation codes overlap on vertebrate core promoters', *Nature 2014 507:7492*, 507(7492), pp. 381–385. Available at: https://doi.org/10.1038/nature12974.
Haberle, V. *et al.* (2015) 'CAGEr: precise TSS data retrieval and high-resolution promoterome mining for integrative analyses', *Nucleic Acids Research*, 43(8), pp. e51–e51. Available at: https://doi.org/10.1093/NAR/GKV054.
Hahn, W.C. and Weinberg, R.A. (2002) 'Modelling the molecular circuitry of cancer', *Nature Reviews Cancer 2002 2:5*, 2(5), pp. 331–341. Available at: https://doi.org/10.1038/nrc795.
Hainer, S.J. *et al.* (2019) 'Profiling of Pluripotency Factors in Single Cells and Early Embryos', *Cell*, 177(5), pp. 1319-1329.e11. Available at: https://doi.org/10.1016/j.cell.2019.03.014.
Hamiche, A. *et al.* (1996) 'Interaction of the histone (H3-H4)2 tetramer of the nucleosome with positively supercoiled DNA minicircles: Potential flipping of the protein from a left- to a right-handed superhelical form.', *Proceedings of the National Academy of Sciences of the United States of America*, 93(15), p. 7588. Available at: https://doi.org/10.1073/PNAS.93.15.7588.
Hanahan, D. and Weinberg, R.A. (2011) 'Hallmarks of Cancer: The Next Generation', *Cell*, 144(5), pp. 646–674. Available at: https://doi.org/10.1016/J.CELL.2011.02.013.
Harley, C.B. (2008) 'Telomerase and cancer therapeutics', *Nature Reviews Cancer 2008 8:3*, 8(3), pp. 167–179. Available at: https://doi.org/10.1038/nrc2275.

Harley, C.B., Futcher, A.B. and Greider, C.W. (1990) 'Telomeres shorten during ageing of human fibroblasts', *Nature 1990 345:6274*, 345(6274), pp. 458–460. Available at: https://doi.org/10.1038/345458a0.
Harouz, H. *et al.* (2014) 'Shigella flexneri targets the HP1γ subcode through the phosphothreonine lyase OspF', *The EMBO Journal*, 33(22), p. 2606. Available at: https://doi.org/10.15252/EMBJ.201489244.
Hastie, N.D. *et al.* (1990) 'Telomere reduction in human colorectal carcinoma and with ageing', *Nature 1990 346:6287*, 346(6287), pp. 866–868. Available at: https://doi.org/10.1038/346866a0.
Hayakawa, T. *et al.* (2003) 'Cell cycle behavior of human HP1 subtypes: distinct molecular domains of HP1 are required for their centromeric localization during interphase and metaphase', *Journal of Cell Science*, 116(16), pp. 3327–3338. Available at: https://doi.org/10.1242/JCS.00635.
Hayflick, L. and Moorhead, P.S. (1961) 'The serial cultivation of human diploid cell strains', *Experimental Cell Research*, 25(3), pp. 585–621. Available at: https://doi.org/https://doi.org/10.1016/0014-4827(61)90192-6.
Heaphy, C.M. *et al.* (2011) 'Altered Telomeres in Tumors with ATRX and DAXX Mutations', *Science*, 333(6041), p. 425. Available at: https://doi.org/10.1126/SCIENCE.1207313.
Heard, E. *et al.* (2001) 'Methylation of Histone H3 at Lys-9 Is an Early Mark on the X Chromosome during X Inactivation', *Cell*, 107(6), pp. 727–738. Available at: https://doi.org/10.1016/S0092-8674(01)00598-0.
Heath, J. *et al.* (2022) 'POGZ promotes homology-directed DNA repair in an HP1-dependent manner', *EMBO reports*, 23(1), p. e51041. Available at: https://doi.org/10.15252/EMBR.202051041.
Heintzman, N.D. *et al.* (2007) 'Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome', *Nature Genetics*, 39(3), pp. 311–318. Available at: https://doi.org/10.1038/ng1966.
Heitz, E. (1928) *Das heterochromatin der moose*. Bornträger.
Henderson, A. *et al.* (2014) 'Detection of G-quadruplex DNA in mammalian cells', *Nucleic Acids Research*, 42(2), pp. 860–869. Available at: https://doi.org/10.1093/NAR/GKT957.
Henson, J.D. *et al.* (2009) 'DNA C-circles are specific and quantifiable markers of alternative-lengthening-of-telomeres activity', *Nature Biotechnology 2009 27:12*, 27(12), pp. 1181–1185. Available at: https://doi.org/10.1038/nbt.1587.
Henson, J.D. *et al.* (2017) 'The C-Circle Assay for alternative-lengthening-of-telomeres activity', *Methods*, 114, pp. 74–84. Available at: https://doi.org/10.1016/J.YMETH.2016.08.016.
Hernandez-Segura, A., Nehme, J. and Demaria, M. (2018) 'Hallmarks of Cellular Senescence', *Trends in Cell Biology*, 28(6), pp. 436–453. Available at: https://doi.org/10.1016/J.TCB.2018.02.001.
Herrera-Moyano, E. *et al.* (2014) 'The yeast and human FACT chromatin-reorganizing complexes solve R-loop-mediated transcription–replication conflicts', *Genes & Development*, 28(7), pp. 735–748. Available at: https://doi.org/10.1101/GAD.234070.113.
Hiragami-Hamada, K. *et al.* (2011) 'N-Terminal Phosphorylation of HP1α Promotes Its Chromatin Binding', *Molecular and Cellular Biology*, 31(6), p. 1186. Available at: https://doi.org/10.1128/MCB.01012-10.

Ho, J.W.K. *et al.* (2014) 'Comparative analysis of metazoan chromatin organization', *Nature*, 512(7515), pp. 449–452. Available at: https://doi.org/10.1038/nature13415.
Hockemeyer, D. *et al.* (2006) 'Recent Expansion of the Telomeric Complex in Rodents: Two Distinct POT1 Proteins Protect Mouse Telomeres', *Cell*, 126(1), pp. 63–77. Available at: https://doi.org/10.1016/J.CELL.2006.04.044.
Holla, S. *et al.* (2020) 'Positioning Heterochromatin at the Nuclear Periphery Suppresses Histone Turnover to Promote Epigenetic Inheritance', *Cell*, 180(1), pp. 150-164.e15. Available at: https://doi.org/10.1016/J.CELL.2019.12.004.
Horn, T. *et al.* (2011) 'Inheritance of Telomere Length in a Bird', *PLOS ONE*, 6(2), p. e17199. Available at: https://doi.org/10.1371/JOURNAL.PONE.0017199.
Horsley, D. *et al.* (1996) 'M32, a murine homologue of Drosophila heterochromatin protein 1 (HP1), localises to euchromatin within interphase nuclei and is largely excluded from constitutive heterochromatin', *Cytogenetic and Genome Research*, 73(4), pp. 308–311. Available at: https://doi.org/10.1159/000134363.
Hotchkiss, R.D. (1948) 'The quantitative separation of Purines, Pyrimidines, and Nucleosides by Paper Chromatography', *Journal of Biological Chemistry*, 175(1), pp. 315–332. Available at: https://doi.org/10.1016/S0021-9258(18)57261-6.
Hou, C. *et al.* (2012) 'Gene Density, Transcription and Insulators Contribute to the Partition of the Drosophila Genome into Physical Domains', *Molecular cell*, 48(3), p. 471. Available at: https://doi.org/10.1016/J.MOLCEL.2012.08.031.
Hsieh, M.J. *et al.* (2006) 'Transcriptional repression activity of PAX3 is modulated by competition between corepressor KAP1 and heterochromatin protein 1', *Biochemical and Biophysical Research Communications*, 349(2), pp. 573–581. Available at: https://doi.org/10.1016/J.BBRC.2006.08.064.
Hübner, B. *et al.* (2022) 'Ultrastructure and nuclear architecture of telomeric chromatin revealed by correlative light and electron microscopy', *Nucleic Acids Research*, 50(9), p. 5047. Available at: https://doi.org/10.1093/NAR/GKAC309.
Hughes, C.S. *et al.* (2018) 'Single-pot, solid-phase-enhanced sample preparation for proteomics experiments', *Nature Protocols 2018 14:1*, 14(1), pp. 68–85. Available at: https://doi.org/10.1038/s41596-018-0082-x.
Hurley, P.J., Wilsker, D. and Bunz, F. (2006) 'Human cancer cells require ATR for cell cycle progression following exposure to ionizing radiation', *Oncogene 2007 26:18*, 26(18), pp. 2535–2542. Available at: https://doi.org/10.1038/sj.onc.1210049.
Illingworth, R.S. *et al.* (2010) 'Orphan CpG Islands Identify Numerous Conserved Promoters in the Mammalian Genome', *PLOS Genetics*, 6(9), p. e1001134. Available at: https://doi.org/10.1371/JOURNAL.PGEN.1001134.
Jack, A.P.M. *et al.* (2013) 'H3K56me3 Is a Novel, Conserved Heterochromatic Mark That Largely but Not Completely Overlaps with H3K9me3 in Both Regulation and Localization', *PLOS ONE*, 8(2), p. e51765. Available at: https://doi.org/10.1371/JOURNAL.PONE.0051765.
Jacobs, S.A. and Khorasanizadeh, S. (2002) 'Structure of HP1 Chromodomain Bound to a Lysine 9-Methylated Histone H3 Tail', *Science*, 295(5562), pp. 2080–2083. Available at: https://doi.org/10.1126/science.1069473.
Jiang, C. and Pugh, B.F. (2009) 'Nucleosome positioning and gene regulation: advances through genomics', *Nature Reviews Genetics 2009 10:3*, 10(3), pp. 161–172. Available at: https://doi.org/10.1038/nrg2522.
Jiang, P.P., Hartl, D.L. and Lemos, B. (2010) 'Y Not a Dead End: Epistatic Interactions Between Y-Linked Regulatory Polymorphisms and Genetic Background Affect Global

Gene Expression in Drosophila melanogaster', *Genetics*, 186(1), pp. 109–118. Available at: https://doi.org/10.1534/GENETICS.110.118109.
Kaaij, L.J.T. *et al.* (2019) 'The ChAHP Complex Counteracts Chromatin Looping at CTCF Sites that Emerged from SINE Expansions in Mouse', *Cell*, 178(6), pp. 1437-1451.e14. Available at: https://doi.org/10.1016/J.CELL.2019.08.007.
Kakolyri, M., Margaritou, A. and Tiligada, E. (2016) 'Dimethyl sulphoxide modifies growth and senescence and induces the non-revertible petite phenotype in yeast', *FEMS Yeast Research*, 16(2). Available at: https://doi.org/10.1093/FEMSYR/FOW008.
Katayama, S. *et al.* (2005) 'Antisense transcription in the mammalian transcriptome', *Science*, 309(5740), pp. 1564–1566. Available at: https://doi.org/10.1126/SCIENCE.1112009/SUPPL_FILE/P1-2CAGE_MOUSE_ACCESSION.LIST.ZIP.
Kato, M., Takemoto, K. and Shinkai, Y. (2018) 'A somatic role for the histone methyltransferase Setdb1 in endogenous retrovirus silencing', *Nature Communications 2018 9:1*, 9(1), pp. 1–13. Available at: https://doi.org/10.1038/s41467-018-04132-9.
Kaul, Z. *et al.* (2012) 'Five dysfunctional telomeres predict onset of senescence in human cells', *EMBO reports*, 13(1), pp. 52–59. Available at: https://doi.org/10.1038/EMBOR.2011.227.
Kee, N. *et al.* (2002) 'The utility of Ki-67 and BrdU as proliferative markers of adult neurogenesis', *Journal of Neuroscience Methods*, 115(1), pp. 97–105. Available at: https://doi.org/10.1016/S0165-0270(02)00007-9.
Keenen, M.M. *et al.* (2021) 'HP1 proteins compact dna into mechanically and positionally stable phase separated domains', *eLife*, 10. Available at: https://doi.org/10.7554/ELIFE.64563.
Kelleher, C., Kurth, I. and Lingner, J. (2005) 'Human Protection of Telomeres 1 (POT1) Is a Negative Regulator of Telomerase Activity In Vitro', *Molecular and Cellular Biology*, 25(2), pp. 808–818. Available at: https://doi.org/10.1128/MCB.25.2.808-818.2005/ASSET/08BEBE16-0613-47B5-A430-2A5DABF3FB85/ASSETS/GRAPHIC/ZMB0020547040006.JPEG.
Kennedy, A.L. *et al.* (2010) 'Senescent mouse cells fail to overtly regulate the HIRA histone chaperone and do not form robust Senescence Associated Heterochromatin Foci', *Cell Division*, 5, p. 16. Available at: https://doi.org/10.1186/1747-1028-5-16.
Kharchenko, P. v., Tolstorukov, M.Y. and Park, P.J. (2008) 'Design and analysis of ChIP-seq experiments for DNA-binding proteins', *Nature Biotechnology 2008 26:12*, 26(12), pp. 1351–1359. Available at: https://doi.org/10.1038/nbt.1508.
Kibe, T. *et al.* (2003) 'Fission yeast Rhp51 is required for the maintenance of telomere structure in the absence of the Ku heterodimer', *Nucleic Acids Research*, 31(17), pp. 5054–5063. Available at: https://doi.org/10.1093/NAR/GKG718.
Kim, H. *et al.* (2011) 'Histone variant H3.3 stimulates HSP70 transcription through cooperation with HP1γ', *Nucleic Acids Research*, 39(19), pp. 8329–8341. Available at: https://doi.org/10.1093/nar/gkr529.
Kim, N.W. *et al.* (1994) 'Specific Association of Human Telomerase Activity with Immortal Cells and Cancer', *Science*, 266(5193), pp. 2011–2015. Available at: https://doi.org/10.1126/SCIENCE.7605428.
Kim, S., Yu, N.K. and Kaang, B.K. (2015) 'CTCF as a multifunctional protein in genome regulation and gene expression', *Experimental & Molecular Medicine*, 47(6), p. e166. Available at: https://doi.org/10.1038/EMM.2015.33.

Kim, W. and Shay, J.W. (2018) 'Long-range telomere regulation of gene expression: Telomere looping and telomere position effect over long distances (TPE-OLD)', *Differentiation*, 99, pp. 1–9. Available at: https://doi.org/10.1016/J.DIFF.2017.11.005.
Kipling, D. and Cooke, H.J. (1990) 'Hypervariable ultra-long telomeres in mice', *Nature 1990 347:6291*, 347(6291), pp. 400–402. Available at: https://doi.org/10.1038/347400a0.
Klenova, E.M. *et al.* (1993) 'CTCF, a conserved nuclear factor required for optimal transcriptional activity of the chicken c-myc gene, is an 11-Zn-finger protein differentially expressed in multiple forms.', *Molecular and Cellular Biology*, 13(12), p. 7612. Available at: https://doi.org/10.1128/MCB.13.12.7612.
Klose, R.J., Kallin, E.M. and Zhang, Y. (2006) 'JmjC-domain-containing proteins and histone demethylation', *Nature Reviews Genetics 2006 7:9*, 7(9), pp. 715–727. Available at: https://doi.org/10.1038/nrg1945.
Knezetic, J.A. and Luse, D.S. (1986) 'The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II in vitro', *Cell*, 45(1), pp. 95–104. Available at: https://doi.org/10.1016/0092-8674(86)90541-6.
Koike, N. *et al.* (2000) 'Identification of heterochromatin protein 1 (HP1) as a phosphorylation target by Pim-1 kinase and the effect of phosphorylation on the transcriptional repression function of HP1 1', *FEBS Letters*, 467(1), pp. 17–21. Available at: https://doi.org/10.1016/S0014-5793(00)01105-4.
Kouzarides, T. (2007) 'Chromatin Modifications and Their Function', *Cell*, 128(4), pp. 693–705. Available at: https://doi.org/10.1016/J.CELL.2007.02.005.
Kueng, S., Oppikofer, M. and Gasser, S.M. (2013) 'SIR Proteins and the Assembly of Silent Chromatin in Budding Yeast', *Annual Review of Genetics*, 47(1), pp. 275–306. Available at: https://doi.org/10.1146/annurev-genet-021313-173730.
Kuilman, T. *et al.* (2010) 'The essence of senescence', *Genes & Development*, 24(22), pp. 2463–2479. Available at: https://doi.org/10.1101/GAD.1971610.
Kuljis, D.A. *et al.* (2013) 'Gonadal- and Sex-Chromosome-Dependent Sex Differences in the Circadian System', *Endocrinology*, 154(4), pp. 1501–1512. Available at: https://doi.org/10.1210/EN.2012-1921.
Kwon, S.H. *et al.* (2010) 'Heterochromatin protein 1 (HP1) connects the FACT histone chaperone complex to the phosphorylated CTD of RNA polymerase II', *Genes & Development*, 24(19), p. 2133. Available at: https://doi.org/10.1101/GAD.1959110.
Lachner, M. *et al.* (2001) 'Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins', *Nature 2001 410:6824*, 410(6824), pp. 116–120. Available at: https://doi.org/10.1038/35065132.
Lange, T. de *et al.* (1990) 'Structure and variability of human chromosome ends', *Molecular and Cellular Biology*, 10(2), pp. 518–527. Available at: https://doi.org/10.1128/MCB.10.2.518-527.1990.
de Lange, T. (2004) 'T-loops and the origin of telomeres', *Nature Reviews Molecular Cell Biology 2004 5:4*, 5(4), pp. 323–329. Available at: https://doi.org/10.1038/nrm1359.
de Lange, T. (2005) 'Shelterin: the protein complex that shapes and safeguards human telomeres', *Genes & development*, 19(18), pp. 2100–2110. Available at: https://doi.org/10.1101/GAD.1346005.
de Lange, T. (2009) 'How telomeres solve the end-protection problem', *Science*, 326(5955), pp. 948–952. Available at: https://doi.org/10.1126/SCIENCE.1170633/SUPPL_FILE/948.MP3.

Lansdorp, P.M. *et al.* (1996) 'Heterogeneity in telomere length of human chromosomes', *Human molecular genetics*, 5(5), pp. 685–691. Available at: https://doi.org/10.1093/HMG/5.5.685.
Larson, A.G. *et al.* (2017) 'Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin', *Nature*, 547(7662), pp. 236–240. Available at: https://doi.org/10.1038/nature22822.
Law P. (2015) 'Investigation into the role of Heterochromatin Protein 1 gamma (HP1γ) in Gene Regulation in Mammals', *PhD thesis* [Preprint].
Law, P.-P. *et al.* (2019) 'Sex differences in gene expression and proliferation are dependent on the epigenetic modifier HP1γ', *bioRxiv*, p. 563940. Available at: https://doi.org/10.1101/563940.
Lechner, M.S. *et al.* (2000) 'Molecular Determinants for Targeting Heterochromatin Protein 1-Mediated Gene Silencing: Direct Chromoshadow Domain–KAP-1 Corepressor Interaction Is Essential', *Molecular and Cellular Biology*, 20(17), pp. 6449–6465. Available at: https://doi.org/10.1128/MCB.20.17.6449-6465.2000/ASSET/0F9F5937-695A-4CE4-B893-F77BB3971658/ASSETS/GRAPHIC/MB1700395010.JPEG.
Lederberg J. (1958) 'Genetic approaches to somatic cell variation: Summary comment', *Journal of Cellular and Comparative Physiology*, 52(S1), pp. 383–401. Available at: https://doi.org/10.1002/JCP.1030520418.
Lee, B.Y. *et al.* (2006) 'Senescence-associated beta-galactosidase is lysosomal beta-galactosidase', *Aging cell*, 5(2), pp. 187–195. Available at: https://doi.org/10.1111/J.1474-9726.2006.00199.X.
Lehnertz, B. *et al.* (2003) 'Suv39h-Mediated Histone H3 Lysine 9 Methylation Directs DNA Methylation to Major Satellite Repeats at Pericentric Heterochromatin', *Current Biology*, 13(14), pp. 1192–1200. Available at: https://doi.org/10.1016/S0960-9822(03)00432-9.
Lejnine, S., Makarov, V.L. and Langmore, J.P. (1995) 'Conserved nucleoprotein structure at the ends of vertebrate and invertebrate chromosomes.', *Proceedings of the National Academy of Sciences*, 92(6), pp. 2393–2397. Available at: https://doi.org/10.1073/PNAS.92.6.2393.
Lemos, B., Branco, A.T. and Hartl, D.L. (2010) 'Epigenetic effects of polymorphic Y chromosomes modulate chromatin components, immune response, and sexual conflict', *Proceedings of the National Academy of Sciences of the United States of America*, 107(36), pp. 15826–15831. Available at: https://doi.org/10.1073/PNAS.1010383107/SUPPL_FILE/PNAS.201010383SI.PDF.
LeRoy, G. *et al.* (2009) 'Heterochromatin Protein 1 Is Extensively Decorated with Histone Code-like Post-translational Modifications', *Molecular & Cellular Proteomics : MCP*, 8(11), p. 2432. Available at: https://doi.org/10.1074/MCP.M900160-MCP200.
Li, B., Oestreich, S. and de Lange, T. (2000) 'Identification of Human Rap1: Implications for Telomere Evolution', *Cell*, 101(5), pp. 471–483. Available at: https://doi.org/10.1016/S0092-8674(00)80858-2.
Li, F. *et al.* (2019) 'ATRX loss induces telomere dysfunction and necessitates induction of alternative lengthening of telomeres during human cell immortalization', *The EMBO Journal*, 38(19). Available at: https://doi.org/10.15252/EMBJ.201796659.
Li, H. *et al.* (2006) 'The Histone Methyltransferase SETDB1 and the DNA Methyltransferase DNMT3A Interact Directly and Localize to Promoters Silenced in Cancer Cells', *Journal of Biological Chemistry*, 281(28), pp. 19489–19500. Available at: https://doi.org/10.1074/JBC.M513249200.

Li, J.S.Z. *et al.* (2017) 'TZAP: A telomere-associated protein involved in telomere length control', *Science*, 355(6325), pp. 638–641. Available at: https://doi.org/10.1126/SCIENCE.AAH6752/SUPPL_FILE/LI.SM.PDF.
Li, S. and Wu, X. (2020) 'Common fragile sites: protection and repair', *Cell & Bioscience*, 10(1), pp. 1–9. Available at: https://doi.org/10.1186/S13578-020-00392-5.
Lieberman-Aiden, E. *et al.* (2009) 'Comprehensive mapping of long range interactions reveals folding principles of the human genome', *Science (New York, N.Y.)*, 326(5950), p. 289. Available at: https://doi.org/10.1126/SCIENCE.1181369.
Liker, A. and Székely, T. (2005) 'Mortality costs of sexual selection and parental care in natural populations of birds', *Evolution*, 59(4), pp. 890–897. Available at: https://doi.org/https://doi.org/10.1111/j.0014-3820.2005.tb01762.x.
Lin, C.C. *et al.* (1993) 'Isolation and identification of a novel tandemly repeated DNA sequence in the centromeric region of human chromosome 8', *Chromosoma*, 102(5), pp. 333–339. Available at: https://doi.org/10.1007/BF00661276.
Lingner, J., Cooper, J.P. and Cech, T.R. (1995) 'Telomerase and DNA End Replication: No Longer a Lagging Strand Problem?', *Science*, 269(5230), pp. 1533–1534. Available at: https://doi.org/10.1126/science.7545310.
Liu, B. *et al.* (2010) 'The Ligase PIAS1 Restricts Natural Regulatory T Cell Differentiation by Epigenetic Repression', *Science (New York, N.Y.)*, 330(6003), p. 521. Available at: https://doi.org/10.1126/SCIENCE.1193787.
Liu, B. *et al.* (2019) 'Telomere shortening by transgenerational transmission of TNF-α-induced TERRA via ATF7', *Nucleic Acids Research*, 47(1), pp. 283–298. Available at: https://doi.org/10.1093/NAR/GKY1149.
Liu, L.-P. *et al.* (2005) 'Sex-specific role of Drosophila melanogaster HP1 in regulating chromatin structure and gene transcription', *Nature Genetics*, 37(12), pp. 1361–1366. Available at: https://doi.org/10.1038/ng1662.
Liu, N. *et al.* (2018) 'Direct Promoter Repression by BCL11A Controls the Fetal to Adult Hemoglobin Switch', *Cell*, 173(2), pp. 430-442.e17. Available at: https://doi.org/10.1016/j.cell.2018.03.016.
Loayza, D. *et al.* (2004) 'DNA Binding Features of Human POT1: A NONAMER 5'-TAGGGTTAG-3' MINIMAL BINDING SITE, SEQUENCE SPECIFICITY, AND INTERNAL BINDING TO MULTIMERIC SITES', *Journal of Biological Chemistry*, 279(13), pp. 13241–13248. Available at: https://doi.org/10.1074/JBC.M312309200.
Loda, A. *et al.* (2017) 'Genetic and epigenetic features direct differential efficiency of Xist-mediated silencing at X-chromosomal and autosomal locations', *Nature Communications 2017 8:1*, 8(1), pp. 1–16. Available at: https://doi.org/10.1038/s41467-017-00528-1.
Lomberk, G. *et al.* (2006) 'Evidence for the existence of an HP1-mediated subcode within the histone code', *Nature Cell Biology 2006 8:4*, 8(4), pp. 407–415. Available at: https://doi.org/10.1038/ncb1383.
Love, M.I., Huber, W. and Anders, S. (2014) 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2', *Genome Biology*, 15(12), pp. 1–21. Available at: https://doi.org/10.1186/S13059-014-0550-8/FIGURES/9.
Lovejoy, C.A. *et al.* (2012) 'Loss of ATRX, Genome Instability, and an Altered DNA Damage Response Are Hallmarks of the Alternative Lengthening of Telomeres Pathway', *PLOS Genetics*, 8(7), p. e1002772. Available at: https://doi.org/10.1371/JOURNAL.PGEN.1002772.

Lovell-Badge, R. and Robertson, E. (1990) 'XY female mice resulting from a heritable mutation in the primary testis-determining gene, Tdy', *Development*, 109(3), pp. 635–646. Available at: https://doi.org/10.1242/DEV.109.3.635.
Lowe, R. *et al.* (2015) 'Sexually dimorphic gene expression emerges with embryonic genome activation and is dynamic throughout development', *BMC Genomics*, 16(1), pp. 1–13. Available at: https://doi.org/10.1186/S12864-015-1506-4/FIGURES/5.
Lowe, S.W., Cepero, E. and Evan, G. (2004) 'Intrinsic tumour suppression', *Nature 2004 432:7015*, 432(7015), pp. 307–315. Available at: https://doi.org/10.1038/nature03098.
Loyola, A. *et al.* (2006) 'PTMs on H3 Variants before Chromatin Assembly Potentiate Their Final Epigenetic State', *Molecular Cell*, 24(2), pp. 309–316. Available at: https://doi.org/10.1016/J.MOLCEL.2006.08.019.
Luger, K. *et al.* (1997) 'Crystal structure of the nucleosome core particle at 2.8 Å resolution', *Nature 1997 389:6648*, 389(6648), pp. 251–260. Available at: https://doi.org/10.1038/38444.
Luijsterburg, M.S. *et al.* (2012) 'DDB2 promotes chromatin decondensation at UV-induced DNA damage', *The Journal of cell biology*, 197(2), pp. 267–281. Available at: https://doi.org/10.1083/JCB.201106074.
Luke, B. *et al.* (2008) 'The Rat1p 5' to 3' Exonuclease Degrades Telomeric Repeat-Containing RNA and Promotes Telomere Elongation in Saccharomyces cerevisiae', *Molecular Cell*, 32(4), pp. 465–477. Available at: https://doi.org/10.1016/J.MOLCEL.2008.10.019.
van Ly, D. *et al.* (2018) 'Telomere Loop Dynamics in Chromosome End Protection', *Molecular Cell*, 71(4), pp. 510-525.e6. Available at: https://doi.org/10.1016/J.MOLCEL.2018.06.025.
Machado, M.R., Dans, P.D. and Pantano, S. (2010) 'Isoform-specific determinants in the HP1 binding to histone 3: insights from molecular simulations', *Amino Acids*, 38(5), pp. 1571–1581. Available at: https://doi.org/10.1007/s00726-009-0371-3.
Machida, S. *et al.* (2018) 'Structural Basis of Heterochromatin Formation by Human HP1', *Molecular cell*, 69(3), pp. 385-397.e8. Available at: https://doi.org/10.1016/J.MOLCEL.2017.12.011.
Maeda, R. and Tachibana, M. (2022) 'HP1 maintains protein stability of H3K9 methyltransferases and demethylases', *EMBO reports*, 23(4), p. e53581. Available at: https://doi.org/10.15252/EMBR.202153581.
Maeshima, K. *et al.* (2016) 'Liquid-like behavior of chromatin', *Current Opinion in Genetics & Development*, 37, pp. 36–45. Available at: https://doi.org/10.1016/J.GDE.2015.11.006.
Maicher, A., Lockhart, A. and Luke, B. (2014) 'Breaking new ground: Digging into TERRA function', *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 1839(5), pp. 387–394. Available at: https://doi.org/10.1016/J.BBAGRM.2014.03.012.
Maison, C. *et al.* (2011) 'SUMOylation promotes de novo targeting of HP1α to pericentric heterochromatin', *Nature Genetics*, 43(3), pp. 220–227. Available at: https://doi.org/10.1038/ng.765.
Maison, C. and Almouzni, G. (2004) 'HP1 and the dynamics of heterochromatin maintenance', *Nature Reviews Molecular Cell Biology 2004 5:4*, 5(4), pp. 296–305. Available at: https://doi.org/10.1038/nrm1355.

Makarov, V.L. *et al.* (1993) 'Nucleosomal organization of telomere-specific chromatin in rat', *Cell*, 73(4), pp. 775–787. Available at: https://doi.org/10.1016/0092-8674(93)90256-P.
Malig, M. *et al.* (2020) 'Ultra-deep Coverage Single-molecule R-loop Footprinting Reveals Principles of R-loop Formation', *Journal of molecular biology*, 432(7), p. 2271. Available at: https://doi.org/10.1016/J.JMB.2020.02.014.
Margalef, P. *et al.* (2018) 'Stabilization of Reversed Replication Forks by Telomerase Drives Telomere Catastrophe', *Cell*, 172(3), pp. 439-453.e14. Available at: https://doi.org/10.1016/J.CELL.2017.11.047.
Marión, R.M. *et al.* (2019) 'TERRA regulate the transcriptional landscape of pluripotent cells through TRF1-dependent recruitment of PRC2', *eLife*, 8. Available at: https://doi.org/10.7554/ELIFE.44656.
Martínez, P. *et al.* (2009) 'Increased telomere fragility and fusions resulting from TRF1 deficiency lead to degenerative pathologies and increased cancer in mice', *Genes & development*, 23(17), pp. 2060–2075. Available at: https://doi.org/10.1101/GAD.543509.
Mateescu, B. *et al.* (2008) 'Regulation of an inducible promoter by an HP1β–HP1γ switch', *EMBO Reports*, 9(3), p. 267. Available at: https://doi.org/10.1038/EMBOR.2008.1.
Mathison, A. *et al.* (2020) 'Discovery, expression, cellular localization, and molecular properties of a novel, alternative spliced HP1γ isoform, lacking the chromoshadow domain', *PLOS ONE*, 15(2), p. e0217452. Available at: https://doi.org/10.1371/JOURNAL.PONE.0217452.
Mattout, A. *et al.* (2015) 'Heterochromatin Protein 1β (HP1β) has distinct functions and distinct nuclear distribution in pluripotent versus differentiated cells', *Genome Biology*, 16(1), p. 213. Available at: https://doi.org/10.1186/s13059-015-0760-8.
Maunakea, A.K. *et al.* (2010) 'Conserved role of intragenic DNA methylation in regulating alternative promoters', *Nature*, 466(7303), pp. 253–257. Available at: https://doi.org/10.1038/nature09165.
Mazzolini, R. *et al.* (2018) 'Snail1 transcription factor controls telomere transcription and integrity', *Nucleic Acids Research*, 46(1), pp. 146–158. Available at: https://doi.org/10.1093/NAR/GKX958.
Mceachern, M.J. and Blackburn, E.H. (1994) 'A conserved sequence motif within the exceptionally diverse telomeric sequences of budding yeasts.', *Proceedings of the National Academy of Sciences*, 91(8), pp. 3453–3457. Available at: https://doi.org/10.1073/PNAS.91.8.3453.
McGhee J. D. and Ginder G. D. (1979) 'Specific DNA methylation sites in the vicinity of the chicken β-globin genes', *Nature*, 280(5721), pp. 419–420. Available at: https://doi.org/10.1038/280419a0.
McKnight, S.L. and Miller, O.L. (1976) 'Ultrastructural patterns of RNA synthesis during early embryogenesis of Drosophila melanogaster', *Cell*, 8(2), pp. 305–319. Available at: https://doi.org/10.1016/0092-8674(76)90014-3.
McLean, C.Y. *et al.* (2010) 'GREAT improves functional interpretation of cis-regulatory regions', *Nature Biotechnology*, 28(5), pp. 495–501. Available at: https://doi.org/10.1038/nbt.1630.
Meehan, R.R., Kao, C.F. and Pennings, S. (2003) 'HP1 binding to native chromatin in vitro is determined by the hinge region and not by the chromodomain', *The EMBO Journal*, 22(12), p. 3164. Available at: https://doi.org/10.1093/EMBOJ/CDG306.

Meers, M.P., Tenenbaum, D. and Henikoff, S. (2019) 'Peak calling by Sparse Enrichment Analysis for CUT&RUN chromatin profiling', *Epigenetics and Chromatin*, 12(1), pp. 1–11. Available at: https://doi.org/10.1186/S13072-019-0287-4/FIGURES/6.
Ménézo, Y.J.R. *et al.* (1999) 'Birth weight and sex ratio after transfer at the blastocyst stage in humans', *Fertility and Sterility*, 72(2), pp. 221–224. Available at: https://doi.org/10.1016/S0015-0282(99)00256-3.
Mersfelder, E.L. and Parthun, M.R. (2006) 'The tale beyond the tail: histone core domain modifications and the regulation of chromatin structure', *Nucleic Acids Research*, 34(9), pp. 2653–2662. Available at: https://doi.org/10.1093/nar/gkl338.
Meyne, J. *et al.* (1990) 'Distribution of non-telomeric sites of the (TTAGGG)n telomeric sequence in vertebrate chromosomes', *Chromosoma*, 99(1), pp. 3–10. Available at: https://doi.org/10.1007/BF01737283.
Meyne, J., Ratliff, R.L. and Moyzis, R.K. (1989) 'Conservation of the human telomere sequence (TTAGGG)n among vertebrates.', *Proceedings of the National Academy of Sciences*, 86(18), pp. 7049–7053. Available at: https://doi.org/10.1073/PNAS.86.18.7049.
Michaelis, C., Ciosk, R. and Nasmyth, K. (1997) 'Cohesins: Chromosomal Proteins that Prevent Premature Separation of Sister Chromatids', *Cell*, 91(1), pp. 35–45. Available at: https://doi.org/10.1016/S0092-8674(01)80007-6.
Mikkelsen, T.S. *et al.* (2007) 'Genome-wide maps of chromatin state in pluripotent and lineage-committed cells', *Nature*, 448(7153), pp. 553–560. Available at: https://doi.org/10.1038/nature06008.
Milot, E. *et al.* (1996) 'Heterochromatin Effects on the Frequency and Duration of LCR-Mediated Gene Transcription', *Cell*, 87(1), pp. 105–114. Available at: https://doi.org/10.1016/S0092-8674(00)81327-6.
Min, J., Wright, W.E. and Shay, J.W. (2017) 'Alternative Lengthening of Telomeres Mediated by Mitotic DNA Synthesis Engages Break-Induced Replication Processes', *Molecular and cellular biology*, 37(20). Available at: https://doi.org/10.1128/MCB.00226-17.
Minc, E. *et al.* (1999) 'Localization and phosphorylation of HP1 proteins during the cell cycle in mammalian cells', *Chromosoma 1999 108:4*, 108(4), pp. 220–234. Available at: https://doi.org/10.1007/S004120050372.
Minc, E., Courvalin, J.C. and Buendia, B. (2000) 'HP1γ associates with euchromatin and heterochromatin in mammalian nuclei and chromosomes', *Cytogenetic and Genome Research*, 90(3–4), pp. 279–284. Available at: https://doi.org/10.1159/000056789.
Mishima, Y. *et al.* (2015) 'Nucleosome compaction facilitates HP1γ binding to methylated H3K9', *Nucleic Acids Research*, 43(21), p. 10200. Available at: https://doi.org/10.1093/NAR/GKV841.
Montero, J.J. *et al.* (2018) 'TERRA recruitment of polycomb to telomeres is essential for histone trymethylation marks at telomeric heterochromatin', *Nature Communications 2018 9:1*, 9(1), pp. 1–14. Available at: https://doi.org/10.1038/s41467-018-03916-3.
Moravec, M. *et al.* (2016) 'TERRA promotes telomerase-mediated telomere elongation in Schizosaccharomyces pombe', *EMBO reports*, 17(7), pp. 999–1012. Available at: https://doi.org/10.15252/EMBR.201541708.

Moye, A.L. *et al.* (2015) 'Telomeric G-quadruplexes are a substrate and site of localization for human telomerase', *Nature Communications 2015 6:1*, 6(1), pp. 1–12. Available at: https://doi.org/10.1038/ncomms8643.
Muchardt, C. *et al.* (2002) 'Coordinated methyl and RNA binding is required for heterochromatin localization of mammalian HP1α', *EMBO reports*, 3(10), pp. 975–981. Available at: https://doi.org/https://doi.org/10.1093/embo-reports/kvf194.
Muller, H.J. and Altenburg, E. (1930) 'The Frequency of Translocations Produced by X-rays in Drosophila', *Genetics*, 15(4), pp. 283–311. Available at: https://doi.org/10.1093/GENETICS/15.4.283.
Nabetani, A. and Ishikawa, F. (2009) 'Unusual Telomeric DNAs in Human Telomerase-Negative Immortalized Cells', *Molecular and Cellular Biology*, 29(3), pp. 703–713. Available at: https://doi.org/10.1128/MCB.00603-08/SUPPL_FILE/NABETANISUPPREV.PDF.
Nagalakshmi, U. *et al.* (2008) 'The Transcriptional Landscape of the Yeast Genome Defined by RNA Sequencing', *Science*, 320(5881), pp. 1344–1349. Available at: https://doi.org/10.1126/science.1158441.
Nandakumar, J. *et al.* (2012) 'The TEL patch of telomere protein TPP1 mediates telomerase recruitment and processivity', *Nature 2012 492:7428*, 492(7428), pp. 285–289. Available at: https://doi.org/10.1038/nature11648.
Nanney, D.L. (1958) 'Epigenetic Control Systems', *Proceedings of the National Academy of Sciences*, 44(7), pp. 712–717. Available at: https://doi.org/10.1073/PNAS.44.7.712/ASSET/45720238-503A-4AFA-86DE-96AB7E0D8D61/ASSETS/PNAS.44.7.712.FP.PNG.
Naruse, C. *et al.* (2007) 'A novel gene trapping for identifying genes expressed under the control of specific transcription factors', *Biochemical and Biophysical Research Communications*, 361(1), pp. 109–115. Available at: https://doi.org/10.1016/J.BBRC.2007.06.161.
Neil, H. *et al.* (2009) 'Widespread bidirectional promoters are the major source of cryptic transcripts in yeast', *Nature 2009 457:7232*, 457(7232), pp. 1038–1042. Available at: https://doi.org/10.1038/nature07747.
Nergadze, S.G. *et al.* (2007) 'Contribution of telomerase RNA retrotranscription to DNA double-strand break repair during mammalian genome evolution', *Genome Biology*, 8(12), pp. 1–13. Available at: https://doi.org/10.1186/GB-2007-8-12-R260/FIGURES/4.
Nergadze, S.G. *et al.* (2009) 'CpG-island promoters drive transcription of human telomeres', *RNA*, 15(12), pp. 2186–2194. Available at: https://doi.org/10.1261/RNA.1748309.
Neumann, A.A. *et al.* (2013) 'Alternative lengthening of telomeres in normal mammalian somatic cells', *Genes & Development*, 27(1), p. 18. Available at: https://doi.org/10.1101/GAD.205062.112.
Nguyen, N.T. *et al.* (2021) 'Male gender is a predictor of higher mortality in hospitalized adults with COVID-19', *PLoS ONE*, 16(7). Available at: https://doi.org/10.1371/JOURNAL.PONE.0254066.
Nguyen, T.H.D. *et al.* (2018) 'Cryo-EM structure of substrate-bound human telomerase holoenzyme', *Nature 2018 557:7704*, 557(7704), pp. 190–195. Available at: https://doi.org/10.1038/s41586-018-0062-x.
Niehrs, C. and Luke, B. (2020) 'Regulatory R-loops as effectors of gene expression and genome stability', *Nature reviews. Molecular cell biology*, 21(3), p. 167. Available at: https://doi.org/10.1038/S41580-019-0206-3.

Nielsen, A.L. *et al.* (1999) 'Interaction with members of the heterochromatin protein 1 (HP1) family and histone deacetylation are differentially involved in transcriptional silencing by members of the TIF1 family', *The EMBO Journal*, 18(22), pp. 6385–6395. Available at: https://doi.org/10.1093/EMBOJ/18.22.6385.
Nielsen, P.R. *et al.* (2002) 'Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9', *Nature 2002 416:6876*, 416(6876), pp. 103–107. Available at: https://doi.org/10.1038/nature722.
Nielsen, S.J. *et al.* (2001) 'Rb targets histone H3 methylation and HP1 to promoters', *Nature*, 412(6846), pp. 561–565. Available at: https://doi.org/10.1038/35087620.
Nishibuchi, G. and Nakayama, J. (2014) 'Biochemical and structural properties of heterochromatin protein 1: understanding its role in chromatin assembly', *The Journal of Biochemistry*, 156(1), pp. 11–20. Available at: https://doi.org/10.1093/jb/mvu032.
Nora, E.P. *et al.* (2012) 'Spatial partitioning of the regulatory landscape of the X-inactivation centre', *Nature*, 485(7398), pp. 381–385. Available at: https://doi.org/10.1038/nature11049.
Nora, E.P. *et al.* (2017) 'Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization', *Cell*, 169(5), pp. 930-944.e22. Available at: https://doi.org/10.1016/J.CELL.2017.05.004.
Ogawa, H. *et al.* (2002) 'A complex with chromatin modifiers that occupies E2f- and Myc-responsive genes in G0 cells', *Science*, 296(5570), pp. 1132–1136. Available at: https://doi.org/10.1126/SCIENCE.1069861/SUPPL_FILE/OGAWASUPPL.DOC.
Ohtani, N. *et al.* (1999) 'Identification of the genes encoding Mn2+-dependent RNase HII and Mg2+-dependent RNase HIII from Bacillus subtilis: Classification of RNases H into three families', *Biochemistry*, 38(2), pp. 605–618. Available at: https://doi.org/10.1021/BI982207Z/ASSET/IMAGES/LARGE/BI982207ZF00010.JPEG.
Okamoto, K. *et al.* (2013) 'A two-step mechanism for TRF2-mediated chromosome end protection', *Nature*, 494(7438), p. 502. Available at: https://doi.org/10.1038/NATURE11873.
Okano, M. *et al.* (1999) 'DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development', *Cell*, 99(3), pp. 247–257. Available at: https://doi.org/10.1016/S0092-8674(00)81656-6.
Olovnikov, A.M. (1973) 'A theory of marginotomy: The incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon', *Journal of Theoretical Biology*, 41(1), pp. 181–190. Available at: https://doi.org/https://doi.org/10.1016/0022-5193(73)90198-7.
Ong, S.E. *et al.* (2002) 'Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics', *Molecular & cellular proteomics : MCP*, 1(5), pp. 376–386. Available at: https://doi.org/10.1074/MCP.M200025-MCP200.
Oomen, M.E. *et al.* (2019) 'CTCF sites display cell cycle–dependent dynamics in factor binding and nucleosome positioning', *Genome Research*, 29(2), pp. 236–249. Available at: https://doi.org/10.1101/GR.241547.118.
Ostapcuk, V. *et al.* (2018) 'Activity-dependent neuroprotective protein recruits HP1 and CHD4 to control lineage-specifying genes', *Nature*, 557(7707), pp. 739–743. Available at: https://doi.org/10.1038/s41586-018-0153-8.
O'Sullivan, R.J. *et al.* (2010) 'Reduced histone biosynthesis and chromatin changes arising from a damage signal at telomeres', *Nature Structural & Molecular Biology*

2010 17:10, 17(10), pp. 1218–1225. Available at: https://doi.org/10.1038/nsmb.1897.
Ourliac-Garnier, I. and Londoño-Vallejo, A. (2011) 'Telomere length analysis by quantitative fluorescent in situ hybridization (Q-FISH).', *Methods in molecular biology (Clifton, N.J.)*, 735, pp. 21–31. Available at: https://doi.org/10.1007/978-1-61779-092-8_3/FIGURES/1_3.
Palladino, F. *et al.* (1993) 'SIR3 and SIR4 proteins are required for the positioning and integrity of yeast telomeres', *Cell*, 75(3), pp. 543–555. Available at: https://doi.org/10.1016/0092-8674(93)90388-7.
Palm, W. *et al.* (2009) 'Functional Dissection of Human and Mouse POT1 Proteins', *Molecular and Cellular Biology*, 29(2), pp. 471–482. Available at: https://doi.org/10.1128/MCB.01352-08/SUPPL_FILE/MCB1352_08SUPPL.PDF.
Palm, W. and de Lange, T. (2008) 'How Shelterin Protects Mammalian Telomeres', *Annual Review of Genetics*, 42(1), pp. 301–334. Available at: https://doi.org/10.1146/annurev.genet.41.110306.130350.
Pan, X. *et al.* (2019) 'FANCM suppresses DNA replication stress at ALT telomeres by disrupting TERRA R-loops', *Scientific Reports 2019 9:1*, 9(1), pp. 1–14. Available at: https://doi.org/10.1038/s41598-019-55537-5.
Panier, S. and Boulton, S.J. (2013) 'Double-strand break repair: 53BP1 comes into focus', *Nature Reviews Molecular Cell Biology 2013 15:1*, 15(1), pp. 7–18. Available at: https://doi.org/10.1038/nrm3719.
Parada, L.A., McQueen, P.G. and Misteli, T. (2004) 'Tissue-specific spatial organization of genomes.', *Genome biology*, 5(7), pp. 1–9. Available at: https://doi.org/10.1186/GB-2004-5-7-R44/FIGURES/4.
Parelho, V. *et al.* (2008) 'Cohesins Functionally Associate with CTCF on Mammalian Chromosome Arms', *Cell*, 132(3), pp. 422–433. Available at: https://doi.org/10.1016/J.CELL.2008.01.011.
Parrinello, S. *et al.* (2003) 'Oxygen sensitivity severely limits the replicative lifespan of murine fibroblasts', *Nature Cell Biology 2003 5:8*, 5(8), pp. 741–747. Available at: https://doi.org/10.1038/ncb1024.
Peckham, H. *et al.* (2020) 'Male sex identified by global COVID-19 meta-analysis as a risk factor for death and ITU admission', *Nature Communications*, 11(1), p. 6317. Available at: https://doi.org/10.1038/s41467-020-19741-6.
Pergament, E. *et al.* (1994) 'Fertilization and early embryology: Sexual differentiation and preimplantation cell growth', *Human Reproduction*, 9(9), pp. 1730–1732. Available at: https://doi.org/10.1093/OXFORDJOURNALS.HUMREP.A138783.
Perrini, B. *et al.* (2004) 'HP1 Controls Telomere Capping, Telomere Elongation, and Telomere Silencing by Two Different Mechanisms in Drosophila', *Molecular Cell*, 15(3), pp. 467–476. Available at: https://doi.org/10.1016/J.MOLCEL.2004.06.036.
Peters, A.H.F.M. *et al.* (2001) 'Loss of the Suv39h Histone Methyltransferases Impairs Mammalian Heterochromatin and Genome Stability', *Cell*, 107(3), pp. 323–337. Available at: https://doi.org/10.1016/S0092-8674(01)00542-6.
Peters, A.H.F.M. *et al.* (2003) 'Partitioning and Plasticity of Repressive Histone Methylation States in Mammalian Chromatin', *Molecular Cell*, 12(6), pp. 1577–1589. Available at: https://doi.org/10.1016/S1097-2765(03)00477-5.
Petiet, A.E. *et al.* (2008) 'High-resolution magnetic resonance histology of the embryonic and neonatal mouse: A 4D atlas and morphologic database', *Proceedings of the National Academy of Sciences of the United States of America*, 105(34), p. 12331. Available at: https://doi.org/10.1073/PNAS.0805747105.

Pfeiffer, V. *et al.* (2013) 'The THO complex component Thp2 counteracts telomeric R-loops and telomere shortening', *The EMBO Journal*, 32(21), p. 2861. Available at: https://doi.org/10.1038/EMBOJ.2013.217.
Pickett, H.A. *et al.* (2009) 'Control of telomere length by a trimming mechanism that involves generation of t-circles', *The EMBO Journal*, 28(7), pp. 799–809. Available at: https://doi.org/10.1038/EMBOJ.2009.42.
Plantinga, M.J. *et al.* (2013) 'Telomerase suppresses formation of ALT-associated single-stranded telomeric C-circles', *Molecular Cancer Research*, 11(6), pp. 557–567. Available at: https://doi.org/10.1158/1541-7786.MCR-13-0013/80258/AM/TELOMERASE-SUPPRESSES-FORMATION-OF-ALT-ASSOCIATED.
Podhorecka, M., Skladanowski, A. and Bozko, P. (2010) 'H2AX Phosphorylation: Its Role in DNA Damage Response and Cancer Therapy', *Journal of Nucleic Acids*, 2010. Available at: https://doi.org/10.4061/2010/920161.
Ponjavic, J. *et al.* (2006) 'Transcriptional and structural impact of TATA-initiation site spacing in mammalian core promoters', *Genome Biology*, 7(8), pp. 1–18. Available at: https://doi.org/10.1186/GB-2006-7-8-R78/FIGURES/11.
Porreca, R.M. *et al.* (2020) 'Trf1 averts chromatin remodelling, recombination and replication dependent-break induced replication at mouse telomeres', *eLife*, 9. Available at: https://doi.org/10.7554/ELIFE.49817.
Porro, A. *et al.* (2010) 'Molecular Dissection of Telomeric Repeat-Containing RNA Biogenesis Unveils the Presence of Distinct and Multiple Regulatory Pathways', *Molecular and Cellular Biology*, 30(20), p. 4808. Available at: https://doi.org/10.1128/MCB.00460-10.
Porro, A. *et al.* (2014) 'Functional characterization of the TERRA transcriptome at damaged telomeres', *Nature Communications 2014 5:1*, 5(1), pp. 1–13. Available at: https://doi.org/10.1038/ncomms6379.
Poulet, A. *et al.* (2009) 'TRF2 promotes, remodels and protects telomeric Holliday junctions', *The EMBO Journal*, 28(6), pp. 641–651. Available at: https://doi.org/https://doi.org/10.1038/emboj.2009.11.
Rao, S.S.P. *et al.* (2014) 'A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping', *Cell*, 159(7), pp. 1665–1680. Available at: https://doi.org/10.1016/j.cell.2014.11.021.
Rao, S.S.P. *et al.* (2017) 'Cohesin Loss Eliminates All Loop Domains', *Cell*, 171(2), pp. 305-320.e24. Available at: https://doi.org/10.1016/J.CELL.2017.09.026.
Ray, P.F. *et al.* (1995) 'Increased number of cells and metabolic activity in male human preimplantation embryos following in vitro fertilization', *Journal of Reproduction and Fertility*, 104, pp. 165–171. Available at: https://doi.org/https://doi.org/10.1530/jrf.0.1040165.
Rayess, H., Wang, M.B. and Srivatsan, E.S. (2012) 'Cellular senescence and tumor suppressor gene p16', *International Journal of Cancer. Journal International du Cancer*, 130(8), p. 1715. Available at: https://doi.org/10.1002/IJC.27316.
Rea, S. *et al.* (2000) 'Regulation of chromatin structure by site-specific histone H3 methyltransferases', *Nature*, 406(6796), pp. 593–599. Available at: https://doi.org/10.1038/35020506.
Recagni, M. *et al.* (2020) 'The Role of Alternative Lengthening of Telomeres Mechanism in Cancer: Translational and Therapeutic Implications', *Cancers 2020, Vol. 12, Page 949*, 12(4), p. 949. Available at: https://doi.org/10.3390/CANCERS12040949.

Rice, J.C. *et al.* (2003) 'Histone Methyltransferases Direct Different Degrees of Methylation to Define Distinct Chromatin Domains', *Molecular Cell*, 12(6), pp. 1591–1598. Available at: https://doi.org/10.1016/S1097-2765(03)00479-9.
Richards, E.J. and Ausubel, F.M. (1988) 'Isolation of a higher eukaryotic telomere from Arabidopsis thaliana', *Cell*, 53(1), pp. 127–136. Available at: https://doi.org/10.1016/0092-8674(88)90494-1.
Rosnoblet, C. *et al.* (2011) 'Analysis of the human HP1 interactome reveals novel binding partners', *Biochemical and Biophysical Research Communications*, 413(2), pp. 206–211. Available at: https://doi.org/10.1016/J.BBRC.2011.08.059.
Ruan, J. *et al.* (2012) 'Structural Basis of the Chromodomain of Cbx3 Bound to Methylated Peptides from Histone H1 and G9a', *PLOS ONE*, 7(4), p. e35376. Available at: https://doi.org/10.1371/JOURNAL.PONE.0035376.
Rudenko, G. *et al.* (1989) 'Alpha-amanitin resistant transcription of protein coding genes in insect and bloodstream form Trypanosoma brucei.', *The EMBO Journal*, 8(13), pp. 4259–4263. Available at: https://doi.org/10.1002/J.1460-2075.1989.TB08611.X.
Sadhukhan, R. *et al.* (2018) 'Expression of Telomere-Associated Proteins is Interdependent to Stabilize Native Telomere Structure and Telomere Dysfunction by G-Quadruplex Ligand Causes TERRA Upregulation', *Cell Biochemistry and Biophysics*, 76(1–2), pp. 311–319. Available at: https://doi.org/10.1007/S12013-017-0835-0/FIGURES/6.
Sagie, S. *et al.* (2017) 'Telomeres in ICF syndrome cells are vulnerable to DNA damage due to elevated DNA:RNA hybrids', *Nature Communications*, 8(1), pp. 1–12. Available at: https://doi.org/10.1038/ncomms14015.
Saint-André, V. *et al.* (2011) 'Histone H3 lysine 9 trimethylation and HP1γ favor inclusion of alternative exons', *Nature Structural & Molecular Biology 2011 18:3*, 18(3), pp. 337–344. Available at: https://doi.org/10.1038/nsmb.1995.
Saksouk, N. *et al.* (2020) 'The mouse HP1 proteins are essential for preventing liver tumorigenesis', *Oncogene*, 39(13), pp. 2676–2691. Available at: https://doi.org/10.1038/s41388-020-1177-8.
Sala Frigerio, C. *et al.* (2019) 'The Major Risk Factors for Alzheimer's Disease: Age, Sex, and Genes Modulate the Microglia Response to Aβ Plaques', *Cell Reports*, 27(4), pp. 1293-1306.e6. Available at: https://doi.org/10.1016/J.CELREP.2019.03.099.
Sales-Gil, R. and Vagnarelli, P. (2020) 'How HP1 Post-Translational Modifications Regulate Heterochromatin Formation and Maintenance', *Cells 2020, Vol. 9, Page 1460*, 9(6), p. 1460. Available at: https://doi.org/10.3390/CELLS9061460.
Salton, M., Voss, T.C. and Misteli, T. (2014) 'Identification by high-throughput imaging of the histone methyltransferase EHMT2 as an epigenetic regulator of VEGFA alternative splicing', *Nucleic Acids Research*, 42(22), pp. 13662–13673. Available at: https://doi.org/10.1093/nar/gku1226.
Sanborn, A.L. *et al.* (2015) 'Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes', *Proceedings of the National Academy of Sciences of the United States of America*, 112(47), pp. E6456–E6465. Available at: https://doi.org/10.1073/PNAS.1518552112/SUPPL_FILE/PNAS.1518552112.ST01.XLSX.
Sanulli, S. *et al.* (2019) 'HP1 reshapes nucleosome core to promote phase separation of heterochromatin', *Nature*, 575(7782), pp. 390–394. Available at: https://doi.org/10.1038/s41586-019-1669-2.

Sanz, L.A. *et al.* (2016) 'Prevalent, Dynamic, and Conserved R-Loop Structures Associate with Specific Epigenomic Signatures in Mammals', *Molecular cell*, 63(1), pp. 167–178. Available at: https://doi.org/10.1016/J.MOLCEL.2016.05.032.
Sanz, L.A. and Chédin, F. (2019) 'High-resolution, strand-specific R-loop mapping via S9.6-based DNA:RNA ImmunoPrecipitation and high-throughput sequencing.', *Nature protocols*, 14(6), p. 1734. Available at: https://doi.org/10.1038/S41596-019-0159-1.
Sarek, G. *et al.* (2015) 'TRF2 Recruits RTEL1 to Telomeres in S Phase to Promote T-Loop Unwinding', *Molecular Cell*, 57(4), p. 622. Available at: https://doi.org/10.1016/J.MOLCEL.2014.12.024.
Sart, D. du *et al.* (1997) 'A functional neo-centromere formed through activation of a latent human centromere and consisting of non-alpha-satellite DNA', *Nature Genetics*, 16(2), pp. 144–153. Available at: https://doi.org/10.1038/ng0697-144.
Saveliev, A. *et al.* (2003) 'DNA triplet repeats mediate heterochromatin-protein-1-sensitive variegated gene silencing', *Nature*, 422(6934), pp. 909–913. Available at: https://doi.org/10.1038/nature01596.
Schaffitzel, C. *et al.* (2001) 'In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with Stylonychia lemnae macronuclei', *Proceedings of the National Academy of Sciences of the United States of America*, 98(15), pp. 8572–8577. Available at: https://doi.org/10.1073/PNAS.141229498/ASSET/48F41601-EAB4-48A9-B4F2-79635649D9EB/ASSETS/GRAPHIC/PQ1412294004.JPEG.
Schmidt, J.C., Zaug, A.J. and Cech, T.R. (2016) 'Live Cell Imaging Reveals the Dynamics of Telomerase Recruitment to Telomeres', *Cell*, 166(5), pp. 1188-1197.e9. Available at: https://doi.org/10.1016/J.CELL.2016.07.033.
Schoeftner, S. and Blasco, M.A. (2007) 'Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II', *Nature Cell Biology 2008 10:2*, 10(2), pp. 228–236. Available at: https://doi.org/10.1038/ncb1685.
Schotta, G. *et al.* (2002) 'Central role of Drosophila SU(VAR)3–9 in histone H3-K9 methylation and heterochromatic gene silencing', *The EMBO Journal*, 21(5), p. 1121. Available at: https://doi.org/10.1093/EMBOJ/21.5.1121.
Schotta, G. *et al.* (2004) 'A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin', *Genes & Development*, 18(11), pp. 1251–1262. Available at: https://doi.org/10.1101/GAD.300704.
Schroeder, A. *et al.* (2006) 'The RIN: An RNA integrity number for assigning integrity values to RNA measurements', *BMC Molecular Biology*, 7(1), pp. 1–14. Available at: https://doi.org/10.1186/1471-2199-7-3/FIGURES/8.
Schultz, D.C. *et al.* (2002) 'SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins', *Genes & Development*, 16(8), pp. 919–932. Available at: https://doi.org/10.1101/GAD.973302.
Schultz, L.B. *et al.* (2000) 'p53 binding protein 1 (53BP1) is an early participant in the cellular response to DNA double-strand breaks', *The Journal of cell biology*, 151(7), pp. 1381–1390. Available at: https://doi.org/10.1083/JCB.151.7.1381.
Schwarzer, W. *et al.* (2017) 'Two independent modes of chromatin organization revealed by cohesin removal', *Nature*, 551(7678), pp. 51–56. Available at: https://doi.org/10.1038/nature24281.
Scully, R. and Xie, A. (2013) 'Double strand break repair functions of histone H2AX', *Mutation research*, 750(0), pp. 5–14. Available at: https://doi.org/10.1016/J.MRFMMM.2013.07.007.

Seila, A.C. *et al.* (2008) 'Divergent transcription from active promoters', *Science*, 322(5909), pp. 1849–1851. Available at: https://doi.org/10.1126/SCIENCE.1162253/SUPPL_FILE/SEILA_SOM.PDF.
Sen, D. and Gilbert, W. (1990) 'A sodium-potassium switch in the formation of four-stranded G4-DNA', *Nature*, 344(6265), pp. 410–414. Available at: https://doi.org/10.1038/344410a0.
Sexton, T. *et al.* (2012) 'Three-dimensional folding and functional organization principles of the Drosophila genome', *Cell*, 148(3), pp. 458–472. Available at: https://doi.org/10.1016/j.cell.2012.01.010.
Sfeir, A. *et al.* (2009) 'Mammalian telomeres resemble fragile sites and require TRF1 for efficient replication', *Cell*, 138(1), p. 90. Available at: https://doi.org/10.1016/J.CELL.2009.06.021.
Shay, J.W. and Bacchetti, S. (1997) 'A survey of telomerase activity in human cancer', *European Journal of Cancer*, 33(5), pp. 787–791. Available at: https://doi.org/10.1016/S0959-8049(97)00062-2.
Shepherd H. D. (2018) 'Epigenetics and Sex: Investigating the role of HPγ in human and mouse sexual dimorphisms', *MSc Thesis* [Preprint].
Shimura, M. *et al.* (2011) 'Epigenetic displacement of HP1 from heterochromatin by HIV-1 Vpr causes premature sister chromatid separation', *Journal of Cell Biology*, 194(5), pp. 721–735. Available at: https://doi.org/10.1083/jcb.201010118.
Shiraki, T. *et al.* (2003) 'Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage', *Proceedings of the National Academy of Sciences of the United States of America*, 100(26), pp. 15776–15781. Available at: https://doi.org/10.1073/PNAS.2136655100/ASSET/DBBD343B-DCFC-4955-AAAE-06D359959797/ASSETS/GRAPHIC/PQ2536655002.JPEG.
de Silanes, I.L. *et al.* (2014) 'Identification of TERRA locus unveils a telomere protection role through association to nearly all chromosomes', *Nature Communications*, 5(1), pp. 1–13. Available at: https://doi.org/10.1038/ncomms5723.
Silva R S. A. S (2015) 'Influence of sex chromosome complement and hormones on heterochromatin silencing in mice', *MPhil Thesis* [Preprint].
Simpson, R.T. (1978) 'Structure of the Chromatosome, a Chromatin Particle Containing 160 Base Pairs of DNA and All the Histones', *Biochemistry*, 17(25), pp. 5524–5531. Available at: https://doi.org/10.1021/BI00618A030/ASSET/BI00618A030.FP.PNG_V03.
Singh, P.B. *et al.* (1991) 'A sequence motif found in a Drosophila heterochromatin protein is conserved in animals and plants.', *Nucleic Acids Research*, 19(4), p. 789. Available at: https://doi.org/10.1093/NAR/19.4.789.
Singh, P.B. and Newman, A.G. (2022) 'HP1-Driven Micro-Phase Separation of Heterochromatin-Like Domains/Complexes', *Epigenetics Insights*, 15, p. 25168657221109770. Available at: https://doi.org/10.1177/25168657221109766.
Skene, P.J., Henikoff, J.G. and Henikoff, S. (2018) 'Targeted in situ genome-wide profiling with high efficiency for low cell numbers', *Nature Protocols 2018 13:5*, 13(5), pp. 1006–1019. Available at: https://doi.org/10.1038/nprot.2018.015.
Skene, P.J. and Henikoff, S. (2017) 'An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites', *eLife*, 6. Available at: https://doi.org/10.7554/ELIFE.21856.

Skourti-Stathaki, K., Kamieniarz-Gdula, K. and Proudfoot, N.J. (2014) 'R-loops induce repressive chromatin marks over mammalian gene terminators', *Nature*, 516(7531), p. 436. Available at: https://doi.org/10.1038/NATURE13787.
Smale, S.T. and Kadonaga, J.T. (2003) 'The RNA Polymerase II Core Promoter', *Annual Review of Biochemistry* , 72, pp. 449–479. Available at: https://doi.org/10.1146/ANNUREV.BIOCHEM.72.121801.161520.
Smallwood, A. *et al.* (2007) 'Functional cooperation between HP1 and DNMT1 mediates gene silencing', *Genes & Development*, 21(10), pp. 1169–1178. Available at: https://doi.org/10.1101/GAD.1536807.
Smallwood, A. *et al.* (2008) 'HP1-mediated silencing targets Pol II coactivator complexes', *Nature Structural & Molecular Biology 2008 15:3*, 15(3), pp. 318–320. Available at: https://doi.org/10.1038/nsmb.1385.
Smallwood, A. *et al.* (2012) 'CBX3 regulates efficient RNA processing genome-wide', *Genome Research*, 22(8), pp. 1426–1436. Available at: https://doi.org/10.1101/GR.124818.111.
Soria, G. and Almouzni, G. (2012) 'Differential contribution of HP1 proteins to DNA end resection and homology-directed repair', *https://doi.org/10.4161/cc.23215*, 12(3), pp. 422–429. Available at: https://doi.org/10.4161/CC.23215.
Sridharan, R. *et al.* (2013) 'Proteomic and genomic approaches reveal critical functions of H3K9 methylation and heterochromatin protein-1γ in reprogramming to pluripotency', *Nature Cell Biology*, 15(7), pp. 872–882. Available at: https://doi.org/10.1038/ncb2768.
van Steensel, B., Smogorzewska, A. and de Lange, T. (1998) 'TRF2 Protects Human Telomeres from End-to-End Fusions', *Cell*, 92(3), pp. 401–413. Available at: https://doi.org/10.1016/S0092-8674(00)80932-0.
Stewart, M.D., Li, J. and Wong, J. (2005) 'Relationship between Histone H3 Lysine 9 Methylation, Transcription Repression, and Heterochromatin Protein 1 Recruitment', *Molecular and Cellular Biology*, 25(7), p. 2525. Available at: https://doi.org/10.1128/MCB.25.7.2525-2538.2005.
Stong, N. *et al.* (2014) 'Subtelomeric CTCF and cohesin binding site organization using improved subtelomere assemblies and a novel annotation pipeline', *Genome Research*, 24(6), pp. 1039–1050. Available at: https://doi.org/10.1101/GR.166983.113.
Strahl, B.D. and Allis, C.D. (2000) 'The language of covalent histone modifications', *Nature 2000 403:6765*, 403(6765), pp. 41–45. Available at: https://doi.org/10.1038/47412.
Strom, A.R. *et al.* (2017) 'Phase separation drives heterochromatin domain formation', *Nature*, 547(7662), p. 241. Available at: https://doi.org/10.1038/NATURE22989.
Szabo, Q., Bantignies, F. and Cavalli, G. (2019) 'Principles of genome folding into topologically associating domains', *Science Advances*, 5(4). Available at: https://doi.org/10.1126/SCIADV.AAW1668/ASSET/1920B784-9140-4F28-B7E2-EFF7540ECA00/ASSETS/GRAPHIC/AAW1668-F3.JPEG.
Tachibana, M. *et al.* (2001) 'SET Domain-containing Protein, G9a, Is a Novel Lysine-preferring Mammalian Histone Methyltransferase with Hyperactivity and Specific Selectivity to Lysines 9 and 27 of Histone H3', *Journal of Biological Chemistry*, 276(27), pp. 25309–25317. Available at: https://doi.org/10.1074/JBC.M101914200.
Tachibana, M. *et al.* (2002) 'G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early

embryogenesis', *Genes & Development*, 16(14), pp. 1779–1791. Available at: https://doi.org/10.1101/GAD.989402.
Tachibana, M. *et al.* (2005) 'Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9', *Genes & Development*, 19(7), pp. 815–826. Available at: https://doi.org/10.1101/GAD.1284005.
Tachibana, M. *et al.* (2008) 'G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription', *The EMBO Journal*, 27(20), pp. 2681–2690. Available at: https://doi.org/10.1038/EMBOJ.2008.192.
Takada, Y. *et al.* (2011) 'HP1γ links histone methylation marks to meiotic synapsis in mice', *Development*, 138(19), pp. 4207–4217. Available at: https://doi.org/10.1242/DEV.064444.
Takahashi, H. *et al.* (2012) '5' end–centered expression profiling using cap-analysis gene expression and next-generation sequencing', *Nature Protocols 2012 7:3*, 7(3), pp. 542–561. Available at: https://doi.org/10.1038/nprot.2012.005.
Takai, K.K. *et al.* (2010) 'In vivo stoichiometry of shelterin components', *The Journal of biological chemistry*, 285(2), pp. 1457–1467. Available at: https://doi.org/10.1074/JBC.M109.038026.
Tam S, Shay J and Pagano M (1994) 'Differential expression and cell cycle regulation of the cyclin-dependent kinase 4 inhibitor p16Ink4', *Cancer Res.*, 22, pp. 5816–5820.
Tamamori-Adachi, M. *et al.* (2018) 'DNA damage response induced by Etoposide promotes steroidogenesis via GADD45A in cultured adrenal cells', *Scientific Reports 2018 8:1*, 8(1), pp. 1–13. Available at: https://doi.org/10.1038/s41598-018-27938-5.
Tan, Y.S. and Lei, Y.L. (2019) 'Generation and Culture of Mouse Embryonic Fibroblasts', *Methods in molecular biology (Clifton, N.J.)*, 1960, p. 85. Available at: https://doi.org/10.1007/978-1-4939-9167-9_7.
Thåström, A., Bingham, L.M. and Widom, J. (2004) 'Nucleosomal Locations of Dominant DNA Sequence Motifs for Histone–DNA Interactions and Nucleosome Positioning', *Journal of Molecular Biology*, 338(4), pp. 695–709. Available at: https://doi.org/10.1016/J.JMB.2004.03.032.
Theiler, K. (1989) 'The House Mouse', *Springer-Verlag* [Preprint]. Available at: https://doi.org/10.1007/978-3-642-88418-4.
Thiru, A. *et al.* (2004) 'Structural basis of HP1/PXVXL motif peptide interactions and HP1 localisation to heterochromatin', *The EMBO Journal*, 23(3), pp. 489–499. Available at: https://doi.org/10.1038/SJ.EMBOJ.7600088.
Thornhill, A.R. and Burgoyne, P.S. (1993) 'A paternally imprinted X chromosome retards the development of the early mouse embryo', *Development*, 118(1), pp. 171–174. Available at: https://doi.org/10.1242/DEV.118.1.171.
Timashev, L.A. *et al.* (2017) 'The DDR at telomeres lacking intact shelterin does not require substantial chromatin decompaction', *Genes & development*, 31(6), pp. 578–589. Available at: https://doi.org/10.1101/GAD.294108.116/-/DC1.
Tokutake, Y. *et al.* (1998) 'Extra-Chromosomal Telomere Repeat DNA in Telomerase-Negative Immortalized Cell Lines', *Biochemical and Biophysical Research Communications*, 247(3), pp. 765–772. Available at: https://doi.org/10.1006/BBRC.1998.8876.
Tomlinson, R.L. *et al.* (2006) 'Cell cycle-regulated trafficking of human telomerase to telomeres', *Molecular Biology of the Cell*, 17(2), pp. 955–965. Available at: https://doi.org/10.1091/MBC.E05-09-0903/ASSET/IMAGES/LARGE/ZMK0020675170006.JPEG.

Tommerup, H., Dousmanis, A. and Lange, T. de (1994) 'Unusual chromatin in human telomeres', *Molecular and Cellular Biology*, 14(9), pp. 5777–5785. Available at: https://doi.org/10.1128/MCB.14.9.5777-5785.1994.
Tong, A.S. *et al.* (2015) 'ATM and ATR Signaling Regulate the Recruitment of Human Telomerase to Telomeres', *Cell Reports*, 13(8), pp. 1633–1646. Available at: https://doi.org/10.1016/J.CELREP.2015.10.041.
Trojer, P. and Reinberg, D. (2007) 'Facultative Heterochromatin: Is There a Distinctive Molecular Signature?', *Molecular Cell*, 28(1), pp. 1–13. Available at: https://doi.org/10.1016/J.MOLCEL.2007.09.011.
Tropberger, P. and Schneider, R. (2013) 'Scratching the (lateral) surface of chromatin regulation by histone modifications', *Nature Structural & Molecular Biology 2013 20:6*, 20(6), pp. 657–661. Available at: https://doi.org/10.1038/nsmb.2581.
Truett, G.E. *et al.* (2018) 'Preparation of PCR-Quality Mouse Genomic DNA with Hot Sodium Hydroxide and Tris (HotSHOT)', *https://doi.org/10.2144/00291bm09*, 29(1), pp. 52–54. Available at: https://doi.org/10.2144/00291BM09.
Turner, B.M. (2007) 'Defining an epigenetic code', *Nature Cell Biology 2007 9:1*, 9(1), pp. 2–6. Available at: https://doi.org/10.1038/ncb0107-2.
Tyanova, S. *et al.* (2016) 'The Perseus computational platform for comprehensive analysis of (prote)omics data', *Nature Methods*, 13(9), pp. 731–740. Available at: https://doi.org/10.1038/nmeth.3901.
Udugama, M. *et al.* (2015) 'Histone variant H3.3 provides the heterochromatic H3 lysine 9 tri-methylation mark at telomeres', *Nucleic Acids Research*, 43(21), pp. 10227–10237. Available at: https://doi.org/10.1093/nar/gkv847.
Vakoc, C.R. *et al.* (2005) 'Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin', *Molecular cell*, 19(3), pp. 381–391. Available at: https://doi.org/10.1016/J.MOLCEL.2005.06.011.
Valouev, A. *et al.* (2011) 'Determinants of nucleosome organization in primary human cells', *Nature*, 474(7352), p. 516. Available at: https://doi.org/10.1038/NATURE10002.
Vancevska, A. *et al.* (2017) 'The telomeric DNA damage response occurs in the absence of chromatin decompaction', *Genes and Development*, 31(6), pp. 567–577. Available at: https://doi.org/10.1101/GAD.294082.116/-/DC1.
Vannier, J.B. *et al.* (2012) 'RTEL1 Dismantles T Loops and Counteracts Telomeric G4-DNA to Maintain Telomere Integrity', *Cell*, 149(4), pp. 795–806. Available at: https://doi.org/10.1016/J.CELL.2012.03.030.
Vassallo, M.F. and Tanese, N. (2002) 'Isoform-specific interaction of HP1 with human TAFII130', *Proceedings of the National Academy of Sciences of the United States of America*, 99(9), pp. 5919–5924. Available at: https://doi.org/10.1073/PNAS.092025499/ASSET/ED8EBEFB-EAB8-4F88-9C5C-5B27DCDA06D1/ASSETS/GRAPHIC/PQ0920254005.JPEG.
Venkatesh, S. *et al.* (2016) 'Selective suppression of antisense transcription by Set2-mediated H3K36 methylation', *Nature Communications*, 7. Available at: https://doi.org/10.1038/NCOMMS13610.
Vera, E. *et al.* (2008) 'Epigenetic regulation of telomeres in human cancer', *Oncogene 2008 27:54*, 27(54), pp. 6817–6833. Available at: https://doi.org/10.1038/onc.2008.289.

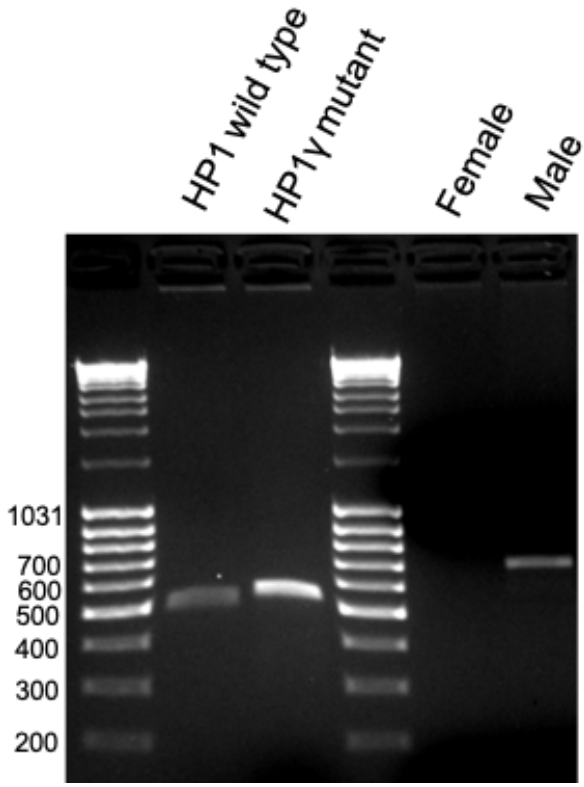
Viceconte N *et al.* (2021) 'PAR-TERRA is the main contributor to telomeric repeat-containing RNA transcripts in normal and cancer mouse cells', *RNA*, 27(1), pp. 106–121. Available at: https://doi.org/10.1261/RNA.076281.120/-/DC1.
Vicent, G.P. *et al.* (2006) 'Induction of Progesterone Target Genes Requires Activation of Erk and Msk Kinases and Phosphorylation of Histone H3', *Molecular Cell*, 24(3), pp. 367–381. Available at: https://doi.org/10.1016/J.MOLCEL.2006.10.011.
Volpe, T.A. *et al.* (2002) 'Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi', *Science*, 297(5588), pp. 1833–1837. Available at: https://doi.org/10.1126/SCIENCE.1074973/SUPPL_FILE/VOLPESOM.PDF.
Vostrov, A.A. and Quitschke, W.W. (1997) 'The Zinc Finger Protein CTCF Binds to the APBβ Domain of the Amyloid β-Protein Precursor Promoter', *Journal of Biological Chemistry*, 272(52), pp. 33353–33359. Available at: https://doi.org/10.1074/jbc.272.52.33353.
Waddington H. (1942) 'Canalization of development and the inheritance of acquired characters', *Nature*, 150(3811), pp. 563–565. Available at: https://doi.org/10.1038/150563a0.
Wallrath, L.L. (1998) 'Unfolding the mysteries of heterochromatin', *Current Opinion in Genetics & Development*, 8(2), pp. 147–153. Available at: https://doi.org/10.1016/S0959-437X(98)80135-4.
Wang, F. *et al.* (2012) 'Telomere- and telomerase-interacting protein that unfolds telomere G-quadruplex and promotes telomere extension in mammalian cells', *Proceedings of the National Academy of Sciences of the United States of America*, 109(50), pp. 20413–20418. Available at: https://doi.org/10.1073/PNAS.1200232109/SUPPL_FILE/PNAS.201200232SI.PDF.
Wang, H. *et al.* (2003) 'mAM Facilitates Conversion by ESET of Dimethyl to Trimethyl Lysine 9 of Histone H3 to Cause Transcriptional Repression', *Molecular Cell*, 12(2), pp. 475–487. Available at: https://doi.org/10.1016/J.MOLCEL.2003.08.007.
Wang, J. *et al.* (2016) 'Genome-wide association analysis identifies variation in vitamin D receptor and other host factors influencing the gut microbiota', *Nature Genetics 2016 48:11*, 48(11), pp. 1396–1406. Available at: https://doi.org/10.1038/ng.3695.
Wang, S. *et al.* (2016) 'Spatial organization of chromatin domains and compartments in single chromosomes', *Science*, 353(6299), pp. 598–602. Available at: https://doi.org/10.1126/SCIENCE.AAF8084/SUPPL_FILE/WANG-SM.PDF.
Watson J D (1972) 'Origin of Concatemeric T7DNA', *Nature New Biology*, 239(94), pp. 197–201. Available at: https://doi.org/10.1038/newbio239197a0.
Watson J D and Crick F H C (1953) 'Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid', *Nature*, 171(4356), pp. 737–738. Available at: https://doi.org/10.1038/171737a0.
Weinrich, S.L. *et al.* (1997) 'Reconstitution of human telomerase with the template RNA component hTR and the catalytic protein subunit hTRT', *Nature Genetics 1997 17:4*, 17(4), pp. 498–502. Available at: https://doi.org/10.1038/ng1297-498.
Wendt, K.S. *et al.* (2008) 'Cohesin mediates transcriptional insulation by CCCTC-binding factor', *Nature 2008 451:7180*, 451(7180), pp. 796–801. Available at: https://doi.org/10.1038/nature06634.
Whetstine, J.R. *et al.* (2006) 'Reversal of Histone Lysine Trimethylation by the JMJD2 Family of Histone Demethylases', *Cell*, 125(3), pp. 467–481. Available at: https://doi.org/10.1016/J.CELL.2006.03.028.

Wijchers, P.J. *et al.* (2010) 'Sexual dimorphism in mammalian autosomal gene regulation is determined not only by sry but by sex chromosome complement as well', *Developmental Cell*, 19(3), pp. 477–484. Available at: https://doi.org/10.1016/j.devcel.2010.08.005.
Wijchers, P.J. and Festenstein, R.J. (2011) 'Epigenetic regulation of autosomal gene expression by sex chromosomes', *Trends in Genetics*, 27(4), pp. 132–140. Available at: https://doi.org/10.1016/J.TIG.2011.01.004.
Wong, F.C.K. (2021) 'Whole-mount immunofluorescence staining of early mouse embryos', *Methods in Molecular Biology*, 2214, pp. 143–155. Available at: https://doi.org/10.1007/978-1-0716-0958-3_10/FIGURES/3.
Wu, W. *et al.* (2015) 'Interaction of BARD1 and HP1 Is Required for BRCA1 Retention at Sites of DNA Damage', *Cancer Research*, 75(7), pp. 1311–1321. Available at: https://doi.org/10.1158/0008-5472.CAN-14-2796.
Wutz, G. *et al.* (2017) 'Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins', *The EMBO Journal*, 36(24), pp. 3573–3599. Available at: https://doi.org/10.15252/EMBJ.201798004.
Xin, H. *et al.* (2007) 'TPP1 is a homologue of ciliate TEBP-β and interacts with POT1 to recruit telomerase', *Nature 2006 445:7127*, 445(7127), pp. 559–562. Available at: https://doi.org/10.1038/nature05469.
Xu, Z. *et al.* (2009) 'Bidirectional promoters generate pervasive transcription in yeast', *Nature 2009 457:7232*, 457(7232), pp. 1033–1037. Available at: https://doi.org/10.1038/nature07728.
Yang, L. *et al.* (2002) 'Molecular cloning of ESET, a novel histone H3-specific methyltransferase that interacts with ERG transcription factor', *Oncogene 2002 21:1*, 21(1), pp. 148–152. Available at: https://doi.org/10.1038/sj.onc.1204998.
Yang, Z. *et al.* (2020) 'Break-induced replication promotes fragile telomere formation', *Genes and Development*, 34(19–20), pp. 1392–1405. Available at: https://doi.org/10.1101/GAD.328575.119/-/DC1.
Ye, J.Z.S. *et al.* (2004) 'POT1-interacting protein PIP1: a telomere length regulator that recruits POT1 to the TIN2/TRF1 complex', *Genes & Development*, 18(14), pp. 1649–1654. Available at: https://doi.org/10.1101/GAD.1215404.
Ye, Q. *et al.* (1997) 'Domain-specific Interactions of Human HP1-type Chromodomain Proteins and Inner Nuclear Membrane Protein LBR', *Journal of Biological Chemistry*, 272(23), pp. 14983–14989. Available at: https://doi.org/10.1074/JBC.272.23.14983.
Yearim, A. *et al.* (2015) 'HP1 Is Involved in Regulating the Global Impact of DNA Methylation on Alternative Splicing', *Cell Reports*, 10(7), pp. 1122–1134. Available at: https://doi.org/10.1016/j.celrep.2015.01.038.
Yi, Q. *et al.* (2018) 'HP1 links centromeric heterochromatin to centromere cohesion in mammals', *EMBO reports*, 19(4), p. e45484. Available at: https://doi.org/https://doi.org/10.15252/embr.201745484.
Yu, Y. *et al.* (2012) 'Histone H3 Lysine 56 Methylation Regulates DNA Replication through Its Interaction with PCNA', *Molecular Cell*, 46(1), pp. 7–17. Available at: https://doi.org/10.1016/J.MOLCEL.2012.01.019.
Yuan, G.C. *et al.* (2005) 'Molecular biology: Genome-scale identification of nucleosome positions in S. cerevisiae', *Science*, 309(5734), pp. 626–630. Available at: https://doi.org/10.1126/SCIENCE.1112178/SUPPL_FILE/YUAN.SOM.PDF.
Zahler, A.M. *et al.* (1991) 'Inhibition of telomerase by G-quartet DMA structures', *Nature*, 350(6320), pp. 718–720. Available at: https://doi.org/10.1038/350718a0.

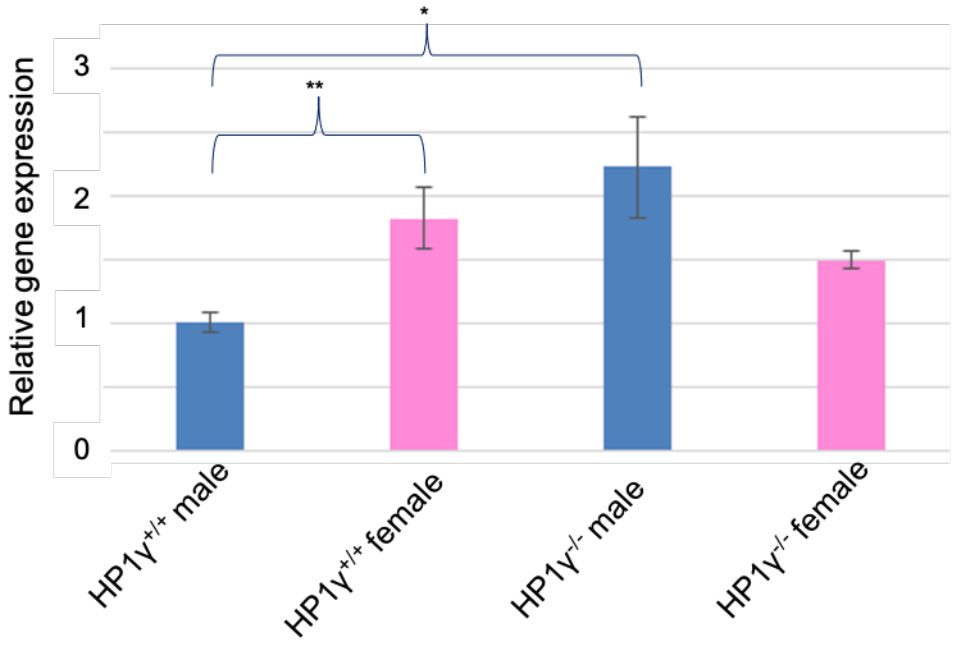
Zaidan, N.Z. *et al.* (2018) 'Compartmentalization of HP1 Proteins in Pluripotency Acquisition and Maintenance', *Stem Cell Reports*, 10(2), p. 627. Available at: https://doi.org/10.1016/J.STEMCR.2017.12.016.
Zeng, W., Ball, A.R. and Yokomori, K. (2010) 'HP1: Heterochromatin binding proteins working the genome', *Epigenetics : official journal of the DNA Methylation Society*, 5(4), p. 287. Available at: https://doi.org/10.4161/EPI.5.4.11683.
Zhang, R., Chen, W. and Adams, P.D. (2007) 'Molecular Dissection of Formation of Senescence-Associated Heterochromatin Foci', *Molecular and Cellular Biology*, 27(6), p. 2343. Available at: https://doi.org/10.1128/MCB.02019-06.
Zhang, R.H. *et al.* (2016) 'The lysine methyltransferase Ehmt2/G9a is dispensable for skeletal muscle development and regeneration', *Skeletal Muscle*, 6(1). Available at: https://doi.org/10.1186/S13395-016-0093-7.
Zhang, X.O., Gingeras, T.R. and Weng, Z. (2019) 'Genome-wide analysis of polymerase III–transcribed Alu elements suggests cell-type–specific enhancer function', *Genome Research*, 29(9), pp. 1402–1414. Available at: https://doi.org/10.1101/GR.249789.119.
Zhao, Y. *et al.* (2009) 'Telomere Extension Occurs at Most Chromosome Ends and Is Uncoupled from Fill-In in Human Cancer Cells', *Cell*, 138(3), pp. 463–475. Available at: https://doi.org/10.1016/J.CELL.2009.05.026.
Zhao, Z. and Shilatifard, A. (2019) 'Epigenetic modifications of histones in cancer', *Genome Biology 2019 20:1*, 20(1), pp. 1–16. Available at: https://doi.org/10.1186/S13059-019-1870-5.
Zhou, Q. *et al.* (2018) 'RNA-QC-chain: comprehensive and fast quality control for RNA-Seq data', *BMC Genomics*, 19(1). Available at: https://doi.org/10.1186/S12864-018-4503-6.
Zhou, Q. and Bachtrog, D. (2012) 'Sex-specific adaptation drives early sex chromosome evolution in Drosophila', *Science*, 337(6092), pp. 341–345. Available at: https://doi.org/10.1126/SCIENCE.1225385/SUPPL_FILE/ZHOU-SOM.PDF.
Zijlmans, J.M.J.M. *et al.* (1997) 'Telomeres in the mouse have large inter-chromosomal variations in the number of T2AG3 repeats', *Proceedings of the National Academy of Sciences of the United States of America*, 94(14), pp. 7423–7428. Available at: https://doi.org/10.1073/PNAS.94.14.7423/ASSET/5919DD39-BD2B-4C87-856A-05BF26312D06/ASSETS/GRAPHIC/PQ1471010004.JPEG.

Appendix

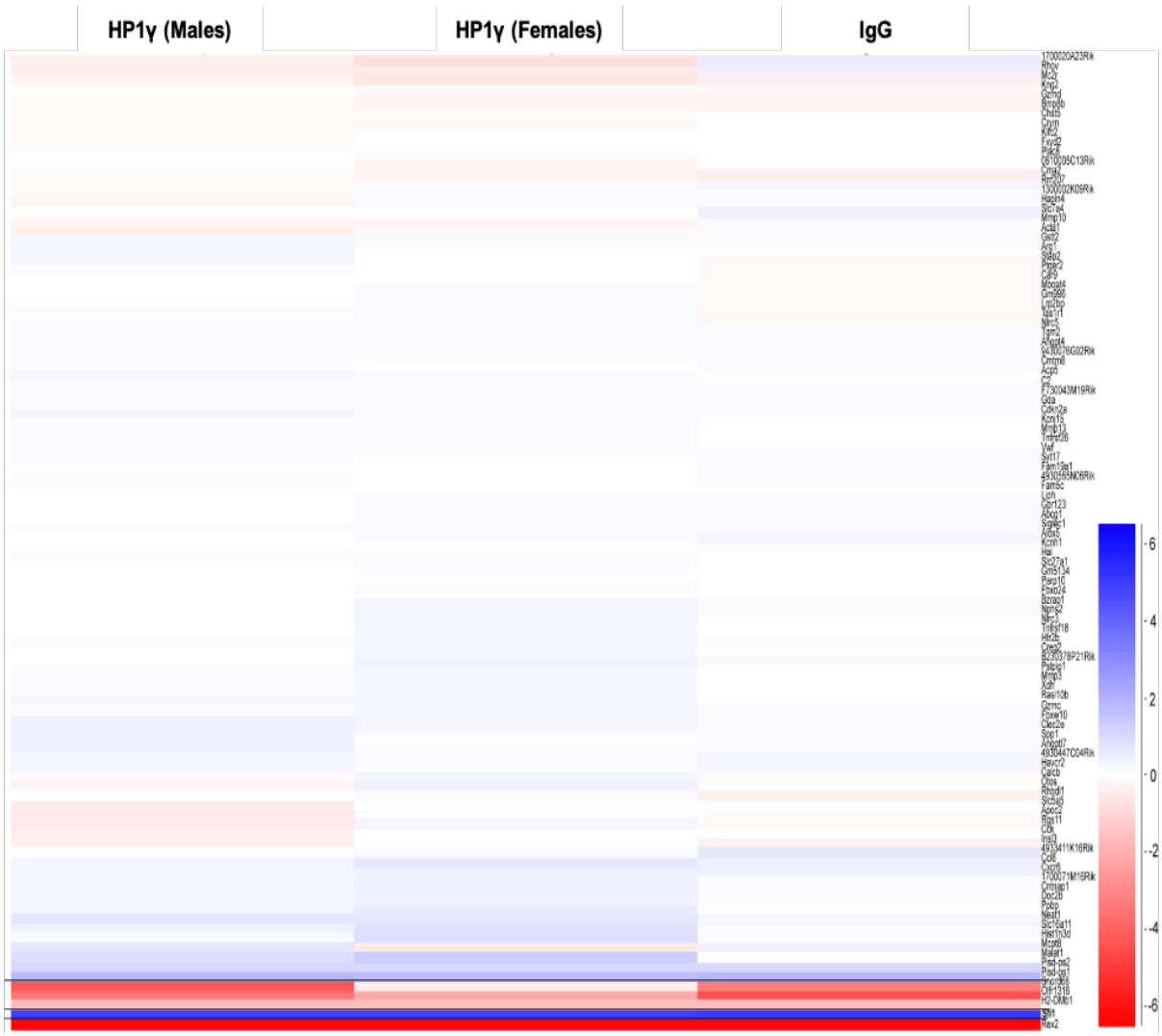
Supplementary Figures



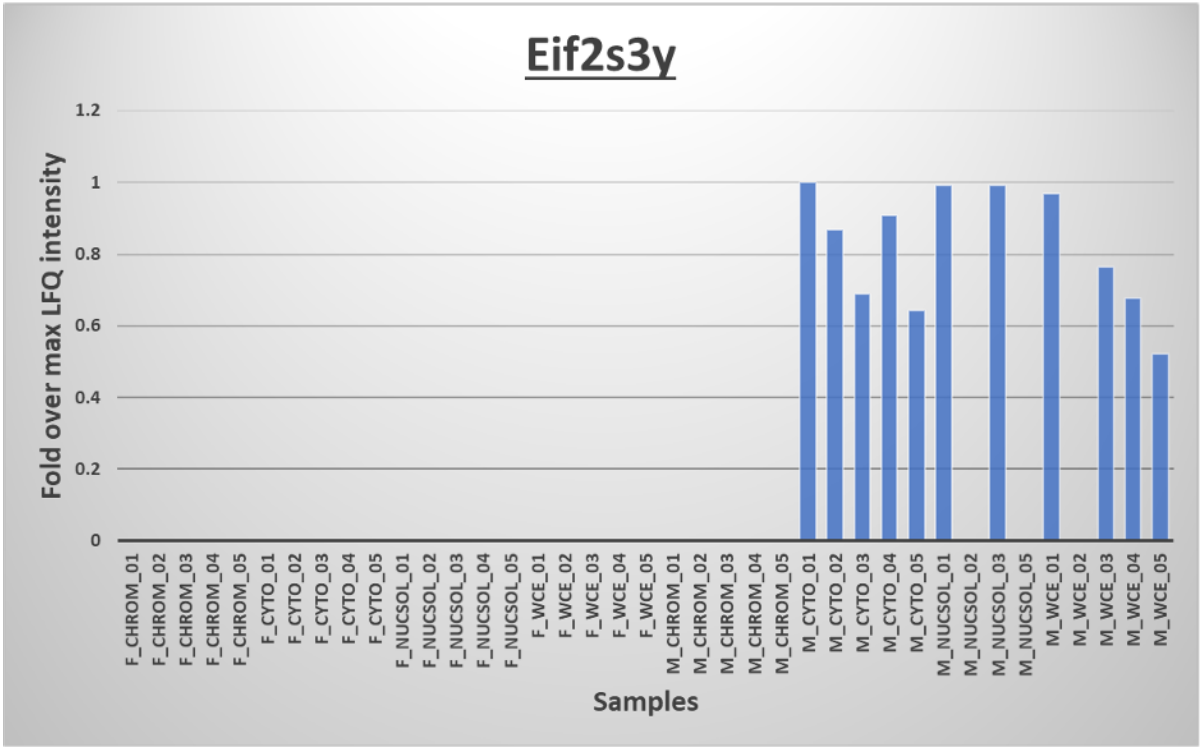
Supplementary Figure S1 Genotyping of mice and E13.5 embryos. PCR for HP1γ genotyping of mice and embryos shows a 501bp band for the wild-type allele and a 525bp band for mutant allele, while PCR for Kdm5d genotyping is used for sex determination of embryos. Male animals show a 597bp band while female animals show no bands.



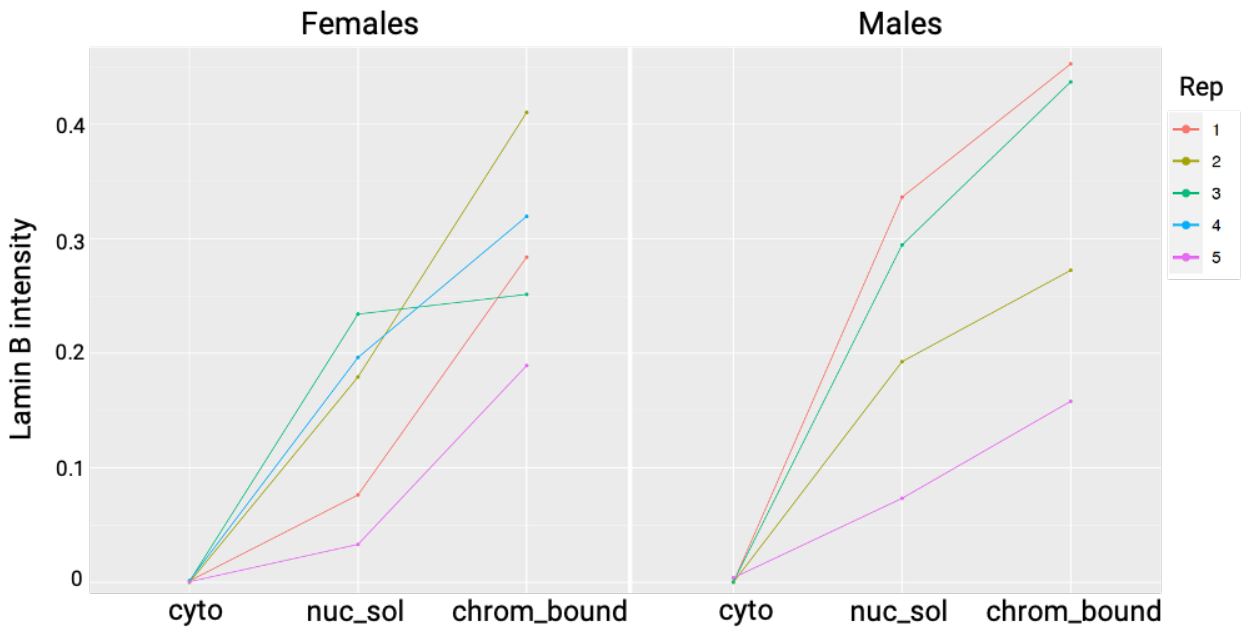
Supplementary Figure S2 *Cdkn2a* is upregulated upon HP1γ depletion in males. RT-qPCR of total RNA extracted from E13.5 MEFs. The graphs represent the average value from three biological replicates for each sample after normalization to 18S rRNA. Error bars represent ± SEM. Statistical significance was tested with Student's t-test, * p<0.05, ** p<0.01, n.s=non-significant. Figure is adapted from Zhi MSc thesis, 2015.



Supplementary Figure S3 *Sfi1* is a binding hotspot. Signal intensity heatmap over the body of "male lower" genes ± 200 bp in the indicated samples. Each row represents a gene. Greater enrichment is depicted in blue and lower enrichment in red. Notice signal enrichment on *Sfi1* (2nd to last row) across the samples, including the negative control (IgG).



Supplementary Figure S4 Eif2s3y is enriched in male samples. Eif2s3y is a Y-linked protein mainly located in the cytoplasm, hence only male samples are expected to show enrichment for this factor.



Supplementary Figure S5 Lamin B is enriched only in nuclear fractions. Lamin B localizes in the nucleus and no enrichment is expected in the cytoplasmic fraction for either sex. Each line represents a different biological replicate with each dot showing the normalised protein intensity of Lamin B in the corresponding subcellular fraction.

Supplementary Tables

Chromosome	Start	End	Strand	Shifting score	p-value	FDR
chr7	142658728	142659493	-	0.517265311	0	0
chr16	35981743	35981778	-	0.286143327	1.69E-14	1.43E-11
chr14	54894121	54894300	+	0.258429652	9.4E-09	4.48E-06
chr9	108339165	108339330	+	0.192113667	0	0
chr8	102865864	102865875	-	0.170824432	5.25E-07	0.0002
chr14	26638018	26638248	+	0.146229414	2.48E-05	0.006758
chr5	33018801	33019055	+	0.124252254	1.55E-05	0.004706
chr13	44870012	44870019	-	0.124078816	0	0
chr3	5860609	5860726	+	0.111410078	1.68E-13	1.28E-10
chr1	63177093	63177228	+	0.097545194	8.35E-06	0.002895
chr16	84834850	84835013	-	0.08995238	6.51E-09	3.31E-06
chr17	39847163	39847338	+	0.083532007	3.34E-13	2.32E-10
chr17	39847008	39847060	+	0.083303424	0	0
chr17	39847480	39848047	+	0.068677276	0	0
chr11	53350744	53351083	+	0.056147999	1.75E-05	0.004936
chr11	109011689	109012121	-	0.04862924	0	0
chr9	45935874	45935996	-	0.019951736	1.13E-05	0.00375
chr13	97190585	97190623	-	0.006990361	1.58E-06	0.000574
chr5	125389689	125389758	-	0.006827839	1.22E-15	1.16E-12
chr17	39846855	39846902	+	0.005279855	8.93E-08	3.59E-05
chr7	142387787	142387972	-	0.004476973	1.6E-05	0.004706
chr17	39845542	39846761	+	0.004252319	3.15E-10	1.85E-07
chr5	142906697	142906758	-	0.00353004	3.15E-10	1.85E-07
chr3	90613090	90613184	+	0.001586454	1.11E-08	4.99E-06
chr14	75845249	75845402	+	0.001024608	3.72E-05	0.00978
chr18	52529124	52529741	-	0.000445338	1.34E-05	0.004263
chr16	57391489	57391680	-	3.013E-05	0	0
chr7	45324939	45324960	+	0	1.31E-09	7.14E-07
chr14	105681826	105681830	+	-0.00397334	1.21E-08	5.15E-06

Supplementary Table S1 Genomic regions which display "promoter shifting" in HP1γ$^{+/+}$ *vs*. HP1γ$^{-/-}$ male MEFs (SLIC-CAGE experiment).

Chromosome	Start	End	Strand	Shifting score	p-value	FDR
chr4	116685543	1.17E+08	+	0.183713	1.62E-09	7.91E-07
chr7	142658728	1.43E+08	-	0.16052	2.79E-06	0.001076
chr13	97190585	97190623	-	0.143323	0	0
chr3	90613090	90613184	+	0.128786	0	0
chr4	43523395	43523571	-	0.120341	5.83E-08	2.67E-05
chr7	141178507	1.41E+08	-	0.10288	3.78E-07	0.000154
chr17	39846855	39846902	+	0.09806	0	0
chr8	128685646	1.29E+08	+	0.085597	1.45E-13	1.06E-10
chr11	109011689	1.09E+08	-	0.084396	0	0
chr19	47854971	47855051	+	0.081608	2.28E-07	9.85E-05
chrX	167209153	1.67E+08	-	0.067657	0	0
chr13	64370276	64370368	-	0.065543	7.92E-06	0.002906
chr3	5860609	5860726	+	0.026937	6.56E-12	3.7E-09
chr7	45324939	45324960	+	0.025985	5.85E-14	4.77E-11
chr16	57391489	57391680	-	0.024075	2.44E-15	2.24E-12
chr17	39847163	39847338	+	0.004655	1.49E-05	0.005212
chr11	52231990	52232064	+	0.003369	2.14E-12	1.31E-09
chr17	39845542	39846761	+	0.002701	1.88E-13	1.26E-10
chr17	39847480	39848047	+	0.000757	0	0
chr5	142906697	1.43E+08	-	0.000239	4.33E-10	2.27E-07
chr19	34255331	34255352	-	-2.9E-06	0	0

Supplementary Table S2 Genomic regions which display "promoter shifting" in HP1γ$^{+/+}$ *vs*. HP1γ$^{-/-}$ female MEFs (SLIC-CAGE experiment).

Experimental class	Sample names	Number of proteins - LFQ intensity	Number of peptides	Sum of protein raw intensities	Sum of protein LFQ intensities
POOL	pool_01	3514	20305	2.84e+11	4.89e+11
POOL	pool_02	3602	21681	3.22e+11	4.93e+11
POOL	pool_03	3594	21814	3.33e+11	4.89e+11
POOL	pool_04	3227	16564	3.01e+11	4.72e+11
POOL	pool_05	3503	21127	3.76e+11	4.65e+11
F_CHROM	f_chrom_01	1889	15570	8.47e+11	8.03e+11
F_CHROM	f_chrom_02	1903	15653	1.01e+12	7.70e+11
F_CHROM	f_chrom_03	1867	15233	6.33e+11	8.91e+11
F_CHROM	f_chrom_04	1319	8183	1.10e+11	7.13e+11
F_CHROM	f_chrom_05	1334	7608	4.57e+10	5.23e+11
M_CHROM	m_chrom_01	1702	12499	7.96e+11	6.40e+11
M_CHROM	m_chrom_02	1663	11697	4.51e+11	7.27e+11
M_CHROM	m_chrom_03	1638	12511	9.48e+11	6.22e+11
M_CHROM	m_chrom_04	1538	10117	6.86e+11	4.95e+11
M_CHROM	m_chrom_05	1107	5576	8.13e+10	4.65e+11
F_CYTO	f_cyto_01	3709	23434	2.40e+11	3.08e+11
F_CYTO	f_cyto_02	2951	14125	2.90e+11	3.78e+11
F_CYTO	f_cyto_03	3015	15061	1.53e+11	2.66e+11
F_CYTO	f_cyto_04	3248	18274	2.62e+11	2.74e+11
F_CYTO	f_cyto_05	2440	10816	4.81e+10	2.73e+11
M_CYTO	m_cyto_01	3389	18656	2.85e+11	3.02e+11
M_CYTO	m_cyto_02	2775	13788	1.39e+11	2.63e+11
M_CYTO	m_cyto_03	3253	16875	2.32e+11	2.78e+11
M_CYTO	m_cyto_04	3146	15573	3.03e+11	3.09e+11
M_CYTO	m_cyto_05	2365	11423	6.27e+10	2.54e+11
F_WCE	f_wce_01	3332	18291	6.02e+11	3.33e+11
F_WCE	f_wce_02	3710	20742	6.43e+11	3.61e+11
F_WCE	f_wce_03	3328	17438	4.95e+11	3.56e+11
F_WCE	f_wce_04	3300	18638	6.52e+11	3.23e+11
F_WCE	f_wce_05	2439	11480	8.34e+10	3.32e+11
M_WCE	m_wce_01	3517	18006	5.37e+11	3.28e+11
M_WCE	m_wce_02	3210	16212	4.01e+11	3.35e+11
M_WCE	m_wce_03	3618	19935	6.81e+11	3.45e+11
M_WCE	m_wce_04	3635	21562	6.38e+11	3.35e+11
M_WCE	m_wce_05	2908	15879	1.57e+11	3.24e+11
F_NUC_SOL	f_nucsol_01	724	3580	4.33e+10	3.07e+11
F_NUC_SOL	f_nucsol_02	1615	10218	2.10e+11	4.04e+11
F_NUC_SOL	f_nucsol_03	1396	9070	1.64e+11	3.48e+11
F_NUC_SOL	f_nucsol_04	1570	10589	2.34e+11	3.43e+11
F_NUC_SOL	f_nucsol_05	633	2128	1.63e+10	2.33e+11
M_NUC_SOL	m_nucsol_01	1722	10145	1.60e+11	3.84e+11
M_NUC_SOL	m_nucsol_02	1228	5577	5.08e+10	2.90e+11
M_NUC_SOL	m_nucsol_03	1989	14658	2.05e+11	4.07e+11
M_NUC_SOL	m_nucsol_05	1028	4645	1.69e+10	3.63e+11

Supplementary Table S3 LC/MS metrics of the MEF subcellular protein fractions. Metrics for the different fractions subjected to LC/MS are presented in the table. LFQ: Label-free quantification; F_CHROM: female chromatin-bound fraction; M_CHROM: male chromatin-bound fraction; F_CYTO: female cytoplasmic fraction; M_CYTO: male cytoplasmic fraction; F_WCE: female combined fractions; M_WCE: male combined fractions; F_NUC_SOL: female nuclear soluble fraction; M_NUC_SOL: male nuclear soluble fraction.