Characterisation of HP1γ in mammalian cells

Meike Wiese

The degree of chromatin compaction plays a fundamental role in controlling the accessibility of DNA to the transcription machinery as well as other DNA-dependent biological pathways. The mammalian HP1 (Heterochromatin protein 1) protein family consists of three members: HP1α, β and γ. Each paralogue regulates formation and maintenance of heterochromatin by binding to the repressive chromatin marks H3K9me2/3 with their chromodomains (CDs). Despite high sequence conservation, each HP1 paralogue possess specific functions, which are likely to be cell type specific. The aim of my thesis was to find novel functions for HP1γ in mouse embryonic stem cells (mESCs) and breast cancer cells.

Mass spectrometry analysis identified citrullination of residues R38 and R39 within the CD of HP1γ. I show that these residues are citrullinated by peptidyl arginine deiminase 4 (PADI4) *in vitro* and *in vivo*. Mutations in HP1γ (R38/9A), designed to mimic the loss of charge accompanied with citrullination, affect HP1γ's binding to H3K9me3 peptides and reduce its residence time on chromatin in differentiated mESCs, indicating a role for citrullination in regulating HP1γ binding to chromatin during differentiation.

Furthermore, I studied the phenotype of HP1γ depletion in two human breast cancer models and found that HP1γ is essential for cell proliferation and viability of cancer, but not of normal epithelial cells. I performed whole transcriptome analysis in breast cancer cells depleted of HP1γ and cross-referenced it with its genomic localisation, which identified increased expression of interferon/antiviral defense genes and activation of pro-apoptotic pathways. Whilst genes involved in these pathways were not directly bound by HP1γ, this analysis also identified HP1γ as a novel regulator of zinc finger (ZNF) genes.

In summary, I identified novel post-translational modifications in HP1γ and characterised them in mESCs. I further demonstrated a role for HP1γ regulating breast cancer cell viability and identified HP1γ as a novel regulator of ZNF genes. My findings highlight HP1γ as a potential target for breast cancer therapy.

Declaration

This thesis is the result of my own work and includes nothing, which is the outcome of work done in collaboration except where specifically indicated in the text.

The number of words in this thesis does not exceed the limit set by the Degree Committee. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other university or similar institution.

Meike Wiese

February 2018

Acknowledgements

I owe a debt of gratitude to my supervisor, Tony Kouzarides, for giving me the opportunity to complete my PhD work in his lab. Tony always strikes the right balance between hard work and play, which makes his lab a truly unique place to work.

I am also very grateful to the other members of the TK lab. In particular I'd like to thank Andy, who was my day-to-day supervisor throughout the PhD. Thank you for your encouragement over the last four years and thank you for imparting your wisdom – even when I thought I didn't need it! I couldn't have done it without you. I'd also like to thank Tomasso, Luca and Sam for analysing the sequencing data and for helpful discussions. Without you, Luca, I still wouldn't know what a rubber policeman really is. Valentina, thank you for helping me settle into the lab, for your technical advice and beyond. Thank you to everyone else in the lab for your support, advice and friendship.

I'd like to thank my friends in Cambridge and afar for their support and friendship. Micha, Schuppi, Lissy and Jana - our holidays, weekends somewhere in Europe and our Skype chats mean a lot to me. I was very fortunate to have one of my dearest friends from Heidelberg living nearby. Hannah, what would I have done without you? I'd also like to thank Mike, who has supported me in every possible way over the years this thesis has taken. Thank you for your patience, tranquility and words of encouragement when it was most required. I am glad our paths crossed in Cambridge.

Finally, I want to thank my family, in particular Mama, Tina and Ilse, who have always believed in me. This wouldn't have been possible without you. Danke!

I gratefully acknowledge the funding of this work by Cancer Research UK.

Table of contents

Chapter 1

Introduction

1.1. Chromatin

1.1.1. A brief history of chromatin

Circa 1840, nuclear components of eukaryotic cells were visualised for the first time using basophilic aniline dyes. Decades later, Walther Flemming named this stainable material within eukaryotic nuclei 'chromatin'[1]. We now know that chromatin is a macromolecular complex of DNA and histone proteins, which provides the packaging scaffold for the entire genome of eukaryotic cells. Given the limited technology of the late 19th century, Flemming and others made incredible observations simply through staining biological material studied by light microscopy. In the early 20th century, Theodor Boveri noted that chromatin condenses into chromosomes during mitosis and decondenses into less defined structures after cell division[2]. About 20 years later, Emil Heitz named these differentially stained regions corresponding to condensed and decondensed chromatin in plant cell nuclei 'heterochromatin' and 'euchromatin' and proposed the revolutionary idea that euchromatin exclusively contains active and heterochromatin repressed genes[3]. Although this model is not entirely accurate anymore, these scientific contributions lined the path for major discoveries in chromatin biology.

First indications for the repressive properties of heterochromatin originated from work by Hermann Mueller in 1930, who observed that X-ray-irradiated *Drosophila* mutants showed variegated (mosaic) expression of the *white* gene in the eyes, evidenced by changes in eye color in a subset of cells. This effect was later directly related to the translocation of the *white* gene from its normal euchromatic location into the vicinity of pericentric heterochromatin and was termed position effect variegation (PEV). PEV represented the first link between chromatin state and expression of a gene.

The following decades were largely devoted to identifying nuclear components encoding the genetic information. Research was greatly determined by the realisation of DNA as the genetic material by Avery, MacLeod and MacCarty[4], the decoding of the molecular structure of DNA by Watson and Crick[5], and the discovery of mRNA and the principle of transcription by Jacob and Monod[6].

With the discovery of the nucleosome - the basic building block of chromatin - in the 1970s, a new era in the field of transcription arose[7-10]. It became increasingly evident that gene expression is strongly impacted by the chromatin structure.

1.1.2. A definition of epigenetics

The term 'epigenetics' was originally proposed by Conrad Waddington in the 1950s to describe heritable changes in a cellular phenotype that were independent of alterations in the DNA sequence[11]. However, despite decades of research, a consensus definition of epigenetics remains controversial. Most commonly, epigenetics is used to describe chromatin-based events that regulate DNA-templated processes. An epigenetic phenotype is most likely triggered by a change in the environment of a cell. Environmental cues can then initiate signaling cascades that lead to the activation of factors that are capable of determining the precise chromatin location and/or DNA location such as DNA-binding proteins, noncoding RNAs or other factors, which then establish an epigenetic pathway[12]. To maintain an epigenetic state over succeeding generations these pathways involve the recruitment of chromatin modifying enzymes or histone variants, resulting in nucleosome repositioning, gene expression changes or changes in higher order chromatin structure.

1.1.3. The nucleosome

The basic functional unit of chromatin is the nucleosome, which consists of 147 bp of DNA wrapped around a histone octamer containing two copies of histones H2A, H2B, H3 and H4[13-16]. Nucleosome units are further compacted into larger 30nm chromatin fiber, at least *in vitro*, by incorporating the linker histone H1 and other architectural chromatin proteins. The exact mechanisms of chromatin architecture beyond the 30nm fiber are less well known. However, it is clear that condensation and decondensation of higher order chromatin structures is a

highly regulated process that controls key DNA based processes including transcription, DNA repair and replication[16].

The first X-ray crystal structure of the nucleosome confirmed that the interaction of DNA with histones occurs primarily with the phosphodiester backbone on the inner surface of the DNA superhelix. The interaction is multifaceted and consists of electrostatic and hydrogen bonds with the DNA phosphates and non-polar interactions with the deoxyribose groups[13]. Contact between DNA and histone proteins is made every 10 base pairs where the minor groove faces inwards[13]. As DNA bases are not in direct contact with histones, DNA packaging by histones is not directly dependent upon the DNA sequence, although it can affect where nucleosomes are positioned (see below). The histone octamer structure contains a highly structured globular domain, an unstructured flexible amino-terminal and carboxy-terminal tail for each of the core histones. Histones tails extend from the nucleosome and can contact DNA and histones of adjacent nucleosomes. All components of the nucleosome are subject to a plethora of covalent modifications, which profoundly impact the organisation and function of these basic modules of chromatin[17].

1.1.4. Nucleosome positioning and remodeling

Nucleosomes need to retain a certain degree of flexibility to allow DNA based processes to occur. Homeostatic histone concentrations ensure that only ~75-90% of the DNA is wrapped in nucleosomes with consecutive nucleosomes classically separated by ~20-50 bp of unwrapped 'linker' DNA. This limiting number of nucleosomes creates competition among different regions of the genomic DNA[18]. Nucleosomal DNA is much less accessible to proteins than linker DNA, hence nucleosome positioning strongly influences the ability of proteins to bind target sequences within DNA. Several different factors can influence nucleosome positioning *in vivo* such as the DNA sequence preference of nucleosomes themselves, DNA methylation, histone variants and covalent modifications of either DNA or histones[18].

DNA sequence preference is not mediated through direct interaction between DNA base pairs and the histones, but rather results from sequence-dependent mechanics of the wrapped DNA. The almost circular wrapping of nucleosomal DNA requires a sharp bending every 10bp and

specific dinucleotide or longer DNA motifs facilitate this conformation[19]. Also modifications on DNA such as cytosine methylation can influence nucleosome positioning by decreasing DNA flexibility[20]. Genomic regions that are critical for transcriptional regulation including promoters and transcription factor (TF) binding sites are usually nucleosome free.

Nucleosome dynamics can be significantly influenced by chromatin remodeling complexes, histone chaperones and post-translational modifications of histones. Histone chaperones were initially characterised as passive mediators that import newly synthesised histones to be incorporated into newly replicated DNA during S phase of the cell cycle[21]. It is now clear that these factors are not passive chaperones, but rather play an active role in the assembly and disassembly of nucleosomes during transcription[21].

Chromatin remodelers play an essential role in DNA based processes. The majority of them catalyse the sliding of nucleosomes along DNA in an ATP dependent manner resulting in nucleosome repositioning, sliding, ejection or editing[22]. Chromatin remodelers and indeed many of the proteins that function on the chromatin interface are recruited, at least in part, by specific interactions with covalent modifications on nucleosomes.

1.1.5. Covalent modifications of the nucleosome

1.1.5.1. DNA methylation

The most abundant and one of the best-characterised epigenetic marks is DNA methylation of carbon 5 of cytosine (5mC), which occurs almost exclusively in CpG dinucleotides. DNA methylation is a key regulator of embryonic development and genomic imprinting, genome stability and chromatin structure in mammals. 5mC is primarily detected within centromeres, pericentromeres, telomeres, the inactive X chromosome and repetitive genomic regions and plays a crucial role in heterochromatin formation and transcriptional repression at these genomic loci[23-25]. Areas of high CpG density are called CpG islands, which colocalise with approximately 60-70% of all promoters annotated in the human genome. Methylation of CpG islands plays a fundamental role in transcriptional regulation. CpG islands at most active genes in normal cells tend not to be methylated and this process is commonly deregulated during malignant transformation[26]. Furthermore, loss of normal imprinting due to aberrant DNA methylation contributes to several inherited diseases in humans[25].

In higher eukaryotes three DNA methyltransferases (DNMT) have been identified[27]. DNMT1 is a maintenance methyltransferase that mainly functions after DNA replication to copy the methylation pattern of the parental strand[28]. Conversely, DNMT3a and DNMT3b, function predominantly as *de novo* methyltransferases to establish DNA methylation during embryogenesis[29].

DNA methylation was generally considered to be a relatively stable chromatin modification. Conversely, studies assessing the global distribution of this modification during embryogenesis had identified a global loss of DNA methylation in the male pronucleus of the early zygote, suggesting the existence of an enzymatic activity within mammalian cells that either removes or alters methylated DNA[30]. The identification of 5-hydroxymethylcytosine (5hmC) presented first insights into the metabolism of 5mC[31]. It is well established now that the TET 1–3 (ten-eleven translocation 1-3) family of proteins are mammalian DNA hydroxylases responsible for converting 5mC to 5hmC[32]. 5hmC plays an important role in transcriptional regulation and TET proteins are crucial for diverse biological processes, including zygotic epigenetic reprogramming, pluripotent stem cell differentiation and hematopoiesis[32]. Iterative oxidation of 5hmC mediated by TET proteins can result in further oxidation derivatives, including 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC), which can be followed by replication-dependent dilution or thymine DNA glycosylase (TDG)-dependent base excision repair to integrate an unmodified cytosine.

DNA methylation provides a platform for several methyl-binding proteins, which function to recruit other chromatin modifying enzymes regulating changes in chromatin structure, including histone deacetylation, methylation, and local chromatin compaction[23,33,34]. Most of these effects are mediated by one of two families of methylated CpG DNA binding proteins, containing the MBD family (MeCP2 and MBD1–4) and the BTB/POZ family (Kaiso and ZBTB4/38). For example, MeCP2 interacts with histone methyltransferases SUV39H1/2, which methylate histone H3K9[35]. Like DNA methylation, trimethylation of H3K9 (H3K9me3) is generally associated with transcriptionally repressed gene regulatory elements and heterochromatic regions (see Section 1.1.8.5). Indeed, MeCP2 and H3K9me3 appear concentrated at pericentromeric regions of heterochromatin in immunofluorescence studies of mammalian nuclei[34]. Conversely, H3K9 methylation can also promote DNA methylation at pericentric heterochromatic regions. In these regions DNMT3a/b can be recruited by the

H3K9me3 binding protein HP1 (heterochromatin protein 1) resulting in increased DNA methylation within heterochromatic regions[36,37]. The interplay of H3K9 methylation and DNA methylation provides one of the best examples for cross-talk between histone and DNA modifications.

1.1.5.2. Histone modifications

In 1964, Vincent Allfrey was the first to hypothesise that histone modifications might have a functional influence on transcription[38]. Ever since, numerous post-translational modifications of histones have been identified and it has been the focus of intense study over the past two decades to unravel their roles, not only in transcription, but in all DNA based processes[39]. Generally histone modifications serve two main functions: they alter chromatin structure by enhancing or weakening the non-covalent histone-histone or DNA-histone interactions and/or provide docking sites for chromatin readers that specifically recognise these modifications and in turn recruit other chromatin modifiers and remodeling enzymes. Besides modifications of amino-terminal tails, a plethora of modifications of the histone core have been identified, which more directly alter protein-DNA interactions and lead to altered chromatin architecture and stability[40].

To date there are at least 8 different classes of modifications and over 60 distinct modifications described in the major core histones[39]. These modifications include lysine acetylation[41], lysine, arginine and glutamine methylation[42,43], serine, threonine and tyrosine phosphorylation[44], lysine ubiquitylation[45], glutamate poly-ADP ribosylation[46], lysine sumoylation[47], arginine deimination (citrullination)[48,49] and proline isomerization[50]. An overview of modifications on all 4 core histones is illustrated in Figure 1-1 and major cellular processes associated to each of these modifications are summarised Table 1-1. Histone methylation and deimination will be discussed in more detail in Section 1.1.8 and 1.19.

Chromatin mediated gene regulation of eukaryotic cells is highly complex and is controlled by extensive cross-talks between histone modifications. Cross-talk can either occur between residues on the same histone, between histones within the same nucleosome or across nucleosomes. An increasing number of histone modifying complexes are found to contain more than one enzymatic activity, which coordinate multiple histone modifications to

establish a specific chromatin environment. Histone modifications can act combinatorially in a context-dependent manner to facilitate or repress chromatin-mediated transcription. Here, the modification of one residue can act as binding site for enzymes mediating the modification of a second residue. A well-characterised example is the requirement of histone H2B mono-ubiquitination for proper H3K4 methylation by the H3K4 methylase complex COMPASS, which is an important mechanism mediating transcription elongation[45]. Conversely, modifications of neighboring residues can also prevent the recognition of another histone residue by an enzyme. For example, methylation of histone H3 arginine 2 (H3R2) interferes with H3K4 methylation by COMPASS-like complexes, thereby impeding transcriptional activation of target genes[51].

Many histones are modified by different classes of enzymes at the same residue, which can influence chromatin function in sometimes opposing ways. For instance, H3K9 can be both methylated, a modification associated with transcriptional silencing, as well as acetylated, which occurs predominantly at actively transcribed genes. *De novo* heterochromatin formation strongly depends on a precise cross-talk between these two modifications. SIRT1 (NAD-Dependent Protein Deacetylase Sirtuin-1) specifically deacetylates H3K9ac (and H4K16ac) at promoters. Chromatin compaction is subsequently achieved by further reduction of active marks and an increase of repressive histone marks such as H3K9me3 and H4K20me3 mediated by methyltransferases such as SUV39H1 (suppressor of variegation 3-9 homolog 1)[52].

1.1.5.3. Histone variants

Canonical histones can be replaced by variant histones to carry out distinct functions in transcription, replication or heterochromatin formation, all of which are mediated by histone chaperones, chromatin remodeling and histone-modifying enzymes[53]. Whilst canonical histones are expressed from non-intronic genes, variant histone genes contain introns and are therefore subject to alternative splicing[54]. Structural differences introduced by a core histone variant influence the interactions between histone proteins within a nucleosome, which can affect nucleosome stability as well as chromatin conformation. Histone variants H2A.Z and H3.3 for instance are predominantly linked to a more open chromatin conformation and found at promoters of active genes, whilst incorporation of macroH2A stabilizes the nucleosome

and is often associated with repressive chromatin[54]. Interestingly, although H3.3 differs from the canonical H3.1 by only 5 amino acids, it is recognised by different histone chaperones. Whilst H3.1 is incorporated during replication by CAF1 (chromatin assembly factor 1), H3.3 can be deposited replication-independently by chaperones such as HIRA or DAXX (death domain-associated protein)[55,56]. Whereas HIRA deposits H3.3 at the promoter region of active genes, DAXX was shown to incorporate H3.3 at telomeric and pericentric heterochromatin[55].

Besides structural differences, variant histones can be subject to specific post-translational modifications reflecting an additional layer of complexity in the modulation of chromatin. The best-characterised modification is the phosphorylation at Ser139 of histone variant H2A.X, which occurs specifically at sites of DNA double strand breaks (DSB) and is required for efficient DSB repair[57].

1.1.6. Histone modifying enzymes

Histones are modified by a large number of enzymes. As many of them have also non-histone targets, it has been challenging to deconvolute the cellular consequences of individual histone modifications from other targets of many of these enzymes[14]. Histone modifications are not static entities but a dynamically changing and complex landscape that evolves in a cell-context dependent fashion. For example, acetylation, phosphorylation or methylation of histones can appear and disappear within minutes in response to particular cell signals.

Table 1-1 Chromatin modifications[14]. me1 (monomethylation); me2 (dimethylation); me3 (trimethylation); me2s (symmetric dimethylation); me2a (asymmetrical dimethylation).

Modification	Nomenclature	Function
acetylation	K-ac	transcription, repair, replication and condensation
methylation (lysine)	K-me1, K-me2, K-me3	transcription and repair
methylation (arginine)	R-me1, R-me2s, R-me2a	transcription
phosphorylation	S-ph, T-ph, Y-ph	transcription,repair and condensation
ubiquitylation	K-ub	transcription and repair,
sumoylation	K-su	transcription
ADP ribosylation	E-ar	transcription
deimination (citrullination)	R to cit	transcription
proline isomerisation	P-cis to P-trans	transcription

Very crudely, histone-modifying enzymes can therefore be classified into modification writers and erasers. Figure 1-1 summarises the main histone modifying enzymes, which are currently known in higher eukaryotes. In addition to their catalytic function many enzymes also possess very specialised reader domains allowing them to bind specific regions in the genome and respond to information conveyed through upstream signaling cascades.

1.1.7. Binders of histone modifications

Histone modifications provide a dynamic binding platform for various factors. Proteins that bind histone modifications contain specialised structured domains that recognise distinct covalent modifications of the nucleosome and other nucleosomal features simultaneously. This combination allows proteins with very similar binding domains to dock at different modified residues or at the same amino acid displaying a different modification state. Figure 1-2 illustrates the main types of domains that bind modified histones with examples of proteins that harbor these specialised binding modules. Domain types recognising lysine methylation show the highest degree of diversity, implying the relative importance of lysine methylation. These domains include PHD fingers and the Tudor 'royal' family of domains, containing chromodomains, Tudor, PWWP and MBT domains[39]. Several domains can recognise the same modified histone lysine. For example, H3K4me3 is bound by the tandem chromodomains within the ATP-dependent chromatin remodeler CHD1[58] and by the double Tudor domains within the histone demethylase JMJD2A (jumonji C domain-containing 2A)[59].

In the early 2000s, the 'histone code' hypothesis was proposed, which defined the combinatorial pattern of histone modifications and the ability of various chromatin interacting proteins to recognise the presence and absence of modifications to translate the code into biological functions[60]. The term 'code' has been controversial, as it implies that it is as predictable as the genetic code, which is not strictly the case for histone modifications. However, ignoring semantics, there are endless examples of how physiological events are not regulated by a single histone modification, but by a combination of several. Many binders and/or protein complexes contain more than one binding module recognising different modifications or different states of modifications at the same time. The transcriptional repressor L3MBTL1 (lethal 3 malignant brain tumor-like 1) is a good example for a multifaceted chromatin binding protein.

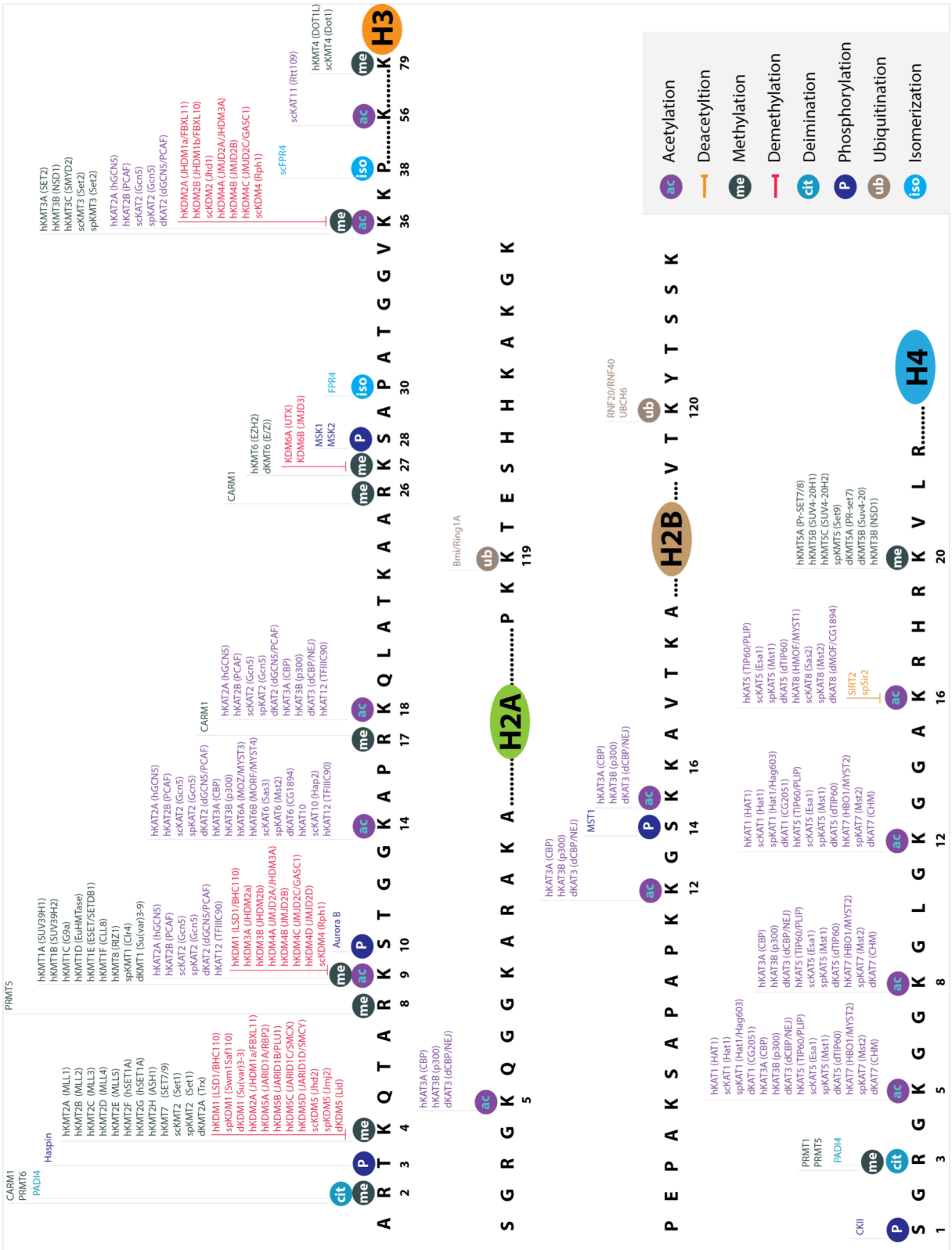


Figure 1-1 Post-translational modifications of histones.

The 4 core histones are schematically illustrated with currently known histone modifications annotated. Listed above each modification are the enzymes known to deposit or erase the modification. (Image was adapted with permission from Dr. Andrew Bannister.)

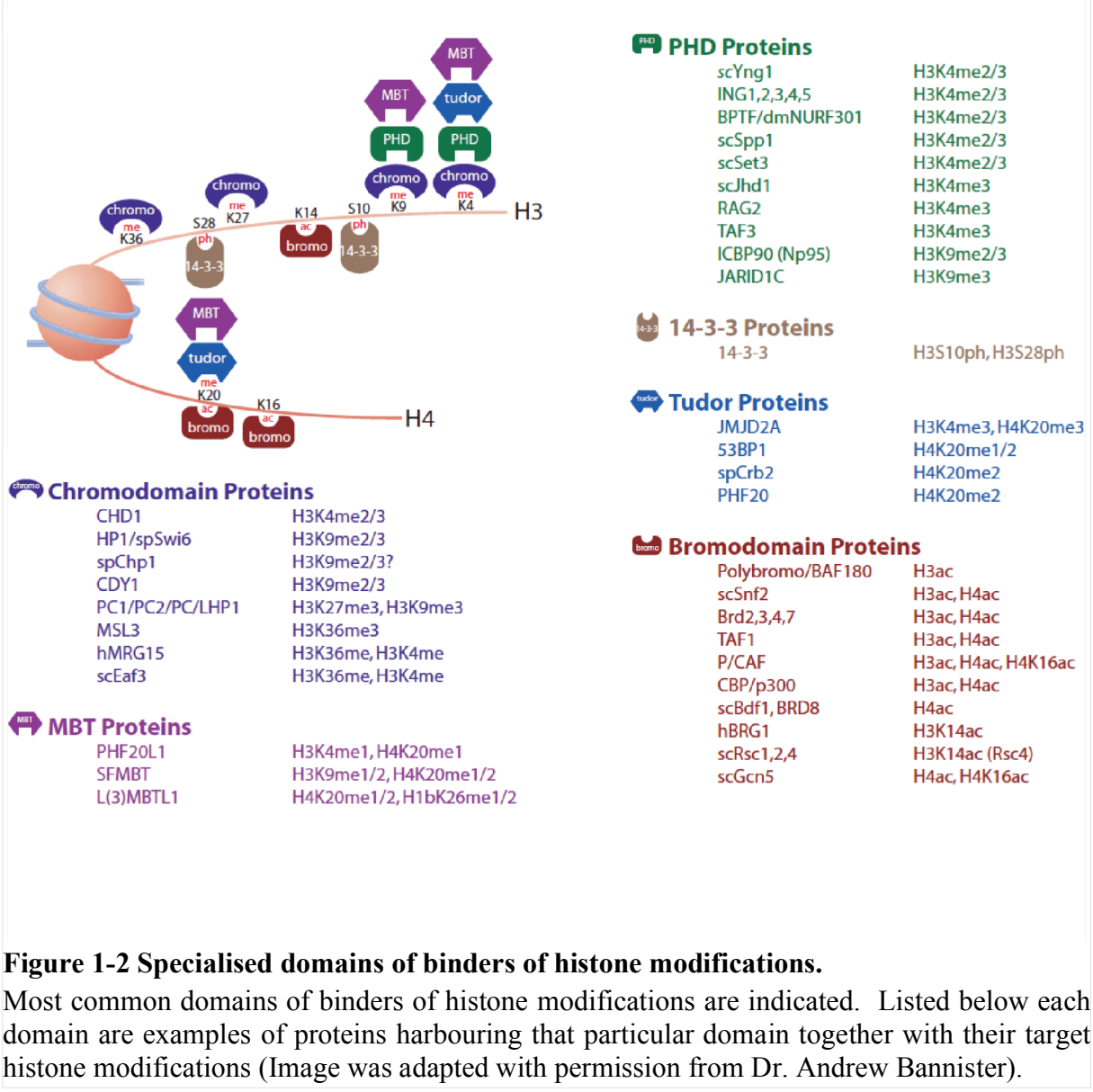


Figure 1-2 Specialised domains of binders of histone modifications.
Most common domains of binders of histone modifications are indicated. Listed below each domain are examples of proteins harbouring that particular domain together with their target histone modifications (Image was adapted with permission from Dr. Andrew Bannister).

It contains three MBT domains recognising H4K20me, as well as H1.4K26me/me2 either on the same nucleosome, on adjacent or on nucleosomes of different chromosomes[61]. In addition, L3MBTL1 interacts with the H3K9me3 interacting protein HP1, thereby linking three repressive chromatin modifications to increase chromatin compaction and transcriptional silencing[61].

1.1.8. Histone methylation

In contrast to phosphorylation and acetylation, methylation does not alter the overall charge of the histone protein. Histones can be methylated to different degrees: lysines may be mono, di

or trimethylated, whilst arginines may be mono, symmetrically or asymmetrically dimethylated[39,62]. The best-characterised sites of histone methylation are lysines. Various histones are methylated on lysine residues either in the histone core or the N-terminal tails. The best-studied lysine methylations are those of the histone tails such as H3K4, H3K9, H3K27, H3K36 and H4K20. Some of these (H3K4, H3K36) are often associated with active genes in euchromatin, whereas others (H3K9, H3K27 and H4K20) are associated with heterochromatic regions of the genome[63]. However, different methylation states on the same residue can be found in different locations in the genome. For example, trimethylation of H3K9 is associated with repressed genes, whilst monomethylation of H3K9 can be detected at active genes[63].

1.1.8.1. Histone lysine methyltransferases (HKMTs)

Many HKMTs have been identified with the vast majority methylating histone tails. All of them contain a SET domain, which possess the methyltransferase enzymatic activity. The only exception to this is hDOT1L, the enzyme that methylates H3K79 in the histone globular core. All HKMTs catalyse the transfer of a methyl group from S-adenosylmethionine (SAM) a lysine's ε-amino group. HKMTs tend to be highly specific enzymes that specifically target certain lysine residues[39].

1.1.8.2. Histone demethylases

The initial hypothesis, that histone methylation is a static modification, has been disproved by the identification of two classes of lysine demethylases. The first class of enzymes demethylate lysines via an amine oxidation reaction with flavin adenine dinucleotide (FAD) as a cofactor. The most prominent example for this group of enzymes, and the first lysine demethylase to be identified is LSD1, which catalyses the removal of H3K4me1/2[64]. However, due to this reaction requiring a protonated nitrogen, amine oxidases only demethylate mono and dimethylated lysine. The second and more expansive class of lysine demethylases is referred to as the Jumonji demethylases. They have a conserved JmjC domain, which functions via an oxidative mechanism and radical attack (involving Fe(II) and α-ketoglutarate). This reaction allows demethylation of trimethylated lysines. The first enzyme identified as a trimethyl lysine demethylase was JMJD2 that demethylates H3K9me3 and H3K36me3[65].

1.1.8.3. H3K36 methylation

In yeast, H3K36 methylation is catalysed by scSet2 (SET domain-containing 2), which performs mono-, di- and trimethylation. The mammalian paralogue of scSet2 is SETD2, however at least eight distinct mammalian enzymes have been described to methylate H3K36, which have varying preferences for H3K36 residues in different methylation states[66] (compare Figure 1-1). H3K36me3 is found in euchromatin. It accumulates at the 3′ end of active genes and is associated with the serine 2 phosphorylated elongating form of RNA pol II. At least in *S. cerevisiae*, it serves as a mark for histone deacetylases (HDACs) to bind and remove any acetylation that was placed in the gene body during the process of transcription. In this way, initiation of inappropriate transcription from cryptic promoters within the gene body is suppressed[67,68]. Evidence suggests this may also be true in humans. However, in humans methylation of H3K9 is used in conjunction with H3K36 to repress aberrant transcription.

1.1.8.4. H3K4 methylation

High levels of H3K4me3 are associated with the 5′ regions of virtually all active genes. Furthermore, there is a strong positive correlation between this modification, transcription rates, active RNA polymerase II occupancy, and histone acetylation[69]. H3K4me1 is a modification associated with enhancers[70]. The first H3K4 methyltransferase to be identified was *S. cerevisiae* Set1, the sole enzyme responsible for H3K4 methylation in yeast. In mammals, the complexity of methyl writers is greater. At least ten known or predicted H3K4 methyltransferases exist (compare Figure 1-1). H3K4me3 is a *bona fide* mark of active genes and often coincides with H3K9ac and H3K14ac at TSSs[69]. The most notable exception of this paradigm is the existence of 'bivalent domains' where H3K4me3 occupies the same promoter as repressive marks, such as H3K27me3. These bivalent domains often mark important developmental genes in stem cells[71]. They are believed to have a role in keeping promoters of developmental genes inactive, yet 'poised' ready to be transcribed[72].

1.1.8.5. H3K9 methylation

While H3K9me1/2 localise specifically to silent domains within euchromatin, H3K9me3 is implicated in the formation of heterochromatin together with HP1. Conversely, H3K9me3

was also described in the gene body of actively transcribed genes, where it is associated with HP1γ[73], indicating that H3K9me3 deposition is not strictly restricted to repressive chromatin and is most probably locus dependent.

Different methylated states of H3K9 are directed by specific histone methyltransferases to mark distinct domains of silent chromatin. The first H3K9 methyltransferase, and the very first HKMT to be identified was SUV39H1[35]. An important paralogue of this enzyme is SUV39H2. Both enzymes catalyse mono-, di-, and trimethylation of H3K9 with a strong preference for H3K9me3. They are crucial for H3K9me3 deposition on pericentromeric heterochromatin[74].

Another important HKMT depositing H3K9me3 is SETDB1 (SET domain bifurcated 1), also called ESET (ERG-associated protein with SET domain)[75,76]. SETDB1 is involved in transcriptional silencing within euchromatin and its specificity is dependent on various interaction partners. It has been shown to bind various proteins, including KAP1 (Kruppel-associated box domain (KRAB)-associated protein), DNMT3A and transcription factors such as OCT4 and SP3[77]. SETDB1's interaction with KAP1 is well characterised. Both proteins form a repressor complexes with components of the NuRD (nucleosome remodeling and histone deacetylation) complex and HP1 to silence coding and non-coding genes in euchromatic regions.

Major depositors of H3K9me1 and H3K9me2 in euchromatin are the HKMTs G9A and GLP, also called EHMT2 and EHMT1 (euchromatic histone-lysine N-methyltransferase 2/1)[78]. Besides H3K9, both enzymes also methylate a number of non-histone substrates, for instance the tumour suppressor protein p53[79].

1.1.9. Histone deimination

1.1.9.1. Peptidylarginine deiminases (PADIs)

Deimination (or citrullination) describes the process of converting peptidyl-arginine to the non-coding amino acid peptidyl-citrulline. This modification results in loss of a positive charge that can change the overall charge of the protein and affect the ability of hydrogen bonding with neighbouring proteins or nucleic acids[80] (Figure 1-3A). The reaction essentially

replaces an imine group with a ketone group and is therefore referred to as a deimination reaction. It is mediated by calcium dependent enzymes called peptidylarginine deiminases (PADIs). In vertebrates, the PADI family consists of 5 tissue-specific isozymes: PADIs 1 - 4 and 6. PADIs 1-4 are enzymatically active, whereas PADI6 appears to be a pseudoenzyme. The enzymatic activity is highly dependent upon the calcium concentration. PADIs 1, 3 and 4 bind five Ca^{2+} ions in their catalytic pocket, whilst PADI2 binds six[81]. X-ray crystal structures of PADI2 and PADI4 enzymes revealed that allosteric binding of Ca^{2+} outside of their catalytic pocket triggers severe structural rearrangements within the pocket[82,83]. In PADI2 this so called 'calcium switch' results in the movement of the 'gate keeper' residue R347 away from the catalytic pocket, thereby facilitating movement of the catalytically important cysteine residue C647 (C645 in PADI4) into the catalytic pocket[83]. Figure 1-3B illustrates conversion from the apo PADI2 to the holo PADI2, which is controlled by 3 Ca^{2+} ions. A very similar mechanism was described for PADI4[83].

All PADI isozymes are associated with the development of different pathological states such as autoimmunity, cancer and neurodegenerative disorders[84]. The enzymatically active members PADIs 1-4 are widely expressed in mammalian tissues. PADI1 is predominantly expressed in the epidermis and the uterus, PADI3 in hair follicles, whilst PADI4 is expressed mainly in neutrophils and eosinophils. PADI2 is the ubiquitous member of the family, expressed for example in skeletal muscles, spleen, brain, secretory glands and other tissues[81].

1.1.9.2. Functions of histone deimination

So far only PADI1 1, 2 and 4 have been reported to localise to the nucleus and regulate gene expression by directly citrullinating histones[48,85,86]. PADI4 citrullinates histone H3 at R2, 8, 17 and 26, as well as histones H2A and H4, both at R3[48,49,87].

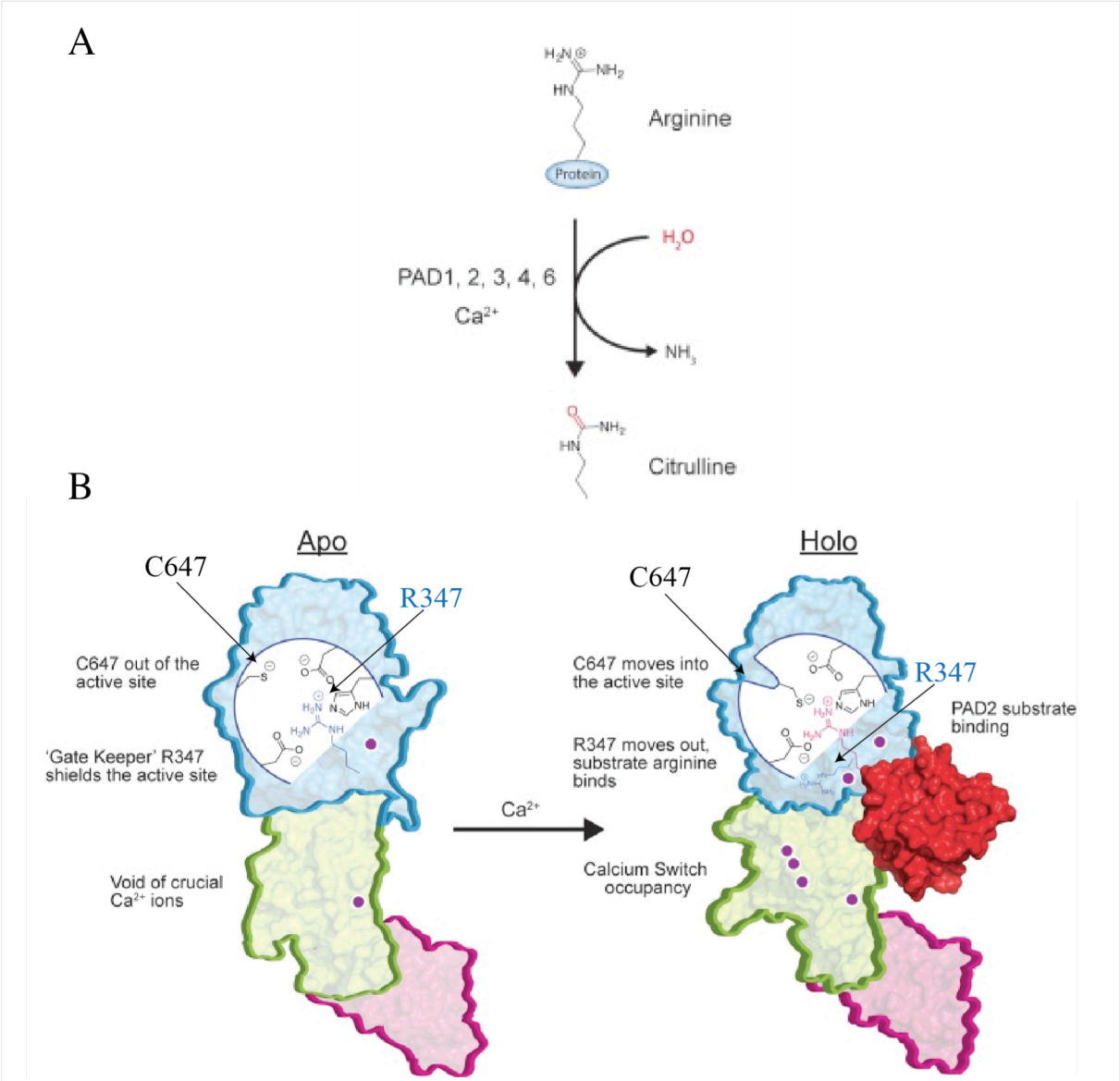


Figure 1-3 Peptidyldeimination is catalysed by PADI enzymes.

A Deimination describes the conversion of peptidylarginine to peptidylcitrulline. Deimination reaction is illustrated.

B Illustration of the X-ray crystal structure of apo and holo form of PADI2. Conversion is triggered by binding of 3 Ca^{2+} ions at the calcium switch, which promotes the relief of 'gate keeper' R374 and proper positioning of the catalytic C647 and subsequent docking of the substrate's arginine. Figure adapted from reference 64.

Citrullination of histones can result in both gene activation and gene repression. For instance, citrullination of H3R17 suppresses estrogen-dependent expression of the *pS2* gene in estrogen receptor α (ERα)-positive MCF7 breast cancer cells, thereby antagonising transcriptional activation mediated via H3R17 methylation by CARM1[48]. In contrast, estrogen-mediated stimulation of PADI4 in the same system was shown to induce citrullination of H3R8 at

promoters of cytokine genes and several human endogenous retroviruses (HERVs). Citrullination of H3R8 at these genomic loci promotes transcription by interfering with H3K9me3 directed binding of HP1α[88]. Also, PADI2 resides within nuclei, where it mediates citrullination of H3 at R2, 8, 17 and 26[86,89,90]. Work from Coonrod's laboratory suggests that PADI2, and not PADI4, exclusively citrullinates H3R26 at ERα target genes in MCF7 cells in response to estrogen stimulation[86]. The authors further propose that H3R26 citrullination leads to local chromatin decondensation, which increases transcriptional activation by enhancing accessibility of ERα to its target genes[90]. In conclusion, these studies show that gene regulation mediated by citrullination of histones is diverse. PADI mediated citrullination interferes with other histone modifying enzymes, binding of chromatin readers or locally affects chromatin structure.

Recently, a role was suggested for nuclear PADI1 in early mouse development[85]. PADI1 expression is induced at the 2-cell stage of mouse pre-implantation development. Knockdown of PADI1 in embryos resulted in loss of citrullination at H4R3 and H3R2, 8 and 17, accompanied by decreased transcriptional activity, and finally complete developmental arrest[85]. It is the first study to demonstrate an enzymatic activity of PADI1 on histones.

Besides epigenetic regulation of gene expression by histone citrullination, PADI4-mediated hypercitrullination of histones is an essential mechanism for the innate immune response[91]. Histone hypercitrullination is critical for neutrophil extracellular trap (NET) formation, a specialised defense mechanism that neutrophils exert against bacterial[92] and fungal pathogens[93]. Rapid chromatin decondensation promotes chromatin extrusion from cells to form web-like structures that capture pathogens. Hypercitrullinated H3 is an important molecular marker for this form of cell death[94].

1.1.10. Constitutive vs. facultative heterochromatin

In contrast to euchromatin that contains active genes, predominantly marked by H3K4me3 and H3K36me3, heterochromatin is generally repressive and contains silenced genes. But not all heterochromatin is the same. In multicellular organisms, two distinct types of heterochromatic regions have been classified: (I.) constitutive heterochromatin and (II.) facultative heterochromatin (Figure 1-4).

(I.) Constitutive heterochromatin predominantly forms at pericentromeric regions and telomeres. Constitutive heterochromatin is characterised by high levels of H3K9me3, HP1α/β and SUV39 enzymes[39]. Telomeres and pericenters are gene-deprived regions and usually made of tandem repeats, called satellites, which vary in size from 5bp up to a few hundreds bp. Pericentromeric repeats in particular show weak sequence conservation across species and do not contain significant functional genic features such as promoter elements. A conserved characteristic is the tandem iteration of DNA motifs, however the mechanisms into how repetitive elements contribute to heterochromatin formation at pericenters repeats remain largely unknown. Interestingly, major satellite repeats are transcribed in sense as well as antisense direction and transcripts were shown to remain chromatin associated and form DNA:RNA hybrids. For both HP1 and Suv39h KMTs interaction with those transcripts is important for their recruitment and retention to pericentric chromatin, suggesting repeat transcription plays an important role in establishing constitutive heterochromatin at these loci[95,96].

(II.) Facultative heterochromatin is more dynamic and can be defined as condensed, transcriptionally silent chromatin, which decondenses and allows transcription in response to developmental cues and/or during cell differentiation[97]. Factors involved in the establishment and/or maintenance of facultative heterochromatin are less well understood. However, it is broadly accepted that combinations of several processes are involved. They include (i.) the incorporation of alternative chromatin components, such as different variants of histone H1, (ii.) interplay of different chromatin modulating factors such as HDACs, HKMTs, DNMTs, chromatin remodelers and/or non-coding RNAs and (iii.) activities of various trans-activating factors such as HP1γ, polycomb group (PcG) proteins or methyl-CpG-binding domain (MBD) containing proteins[97]. A classic example of this type of heterochromatin is the mono-allelic inactivation of the X chromosome in female mammalian organisms in early stages of development. Before gastrulation, the long ncRNA Xist coats the inactive X chromosome (Xi), which is followed by hypo acetylation of the entire Xi chromosome and the appearance of repressive chromatin marks H4K20me1 and H3K27me3 mediated by PR (polycomb repressive) complex I and II[97]. This state is then maintained throughout the whole lifetime of the organism.

1.2. Heterochromatin protein 1 (HP1)

A major breakthrough in our understanding of heterochromatin-mediated gene silencing was achieved in the 1980s, when extensive mutagenesis analyses of modifiers of PEV in *Drosophila* were performed. This resulted in identification of more than 100 genes that either enhanced or suppressed PEV when mutated. Amongst those genes, Heterochromatin protein 1 (HP1), which had previously been described as a non-histone protein that predominantly localises to pericentromeric heterochromatin[98,99], was identified as a dominant suppressor of PEV[100-102].

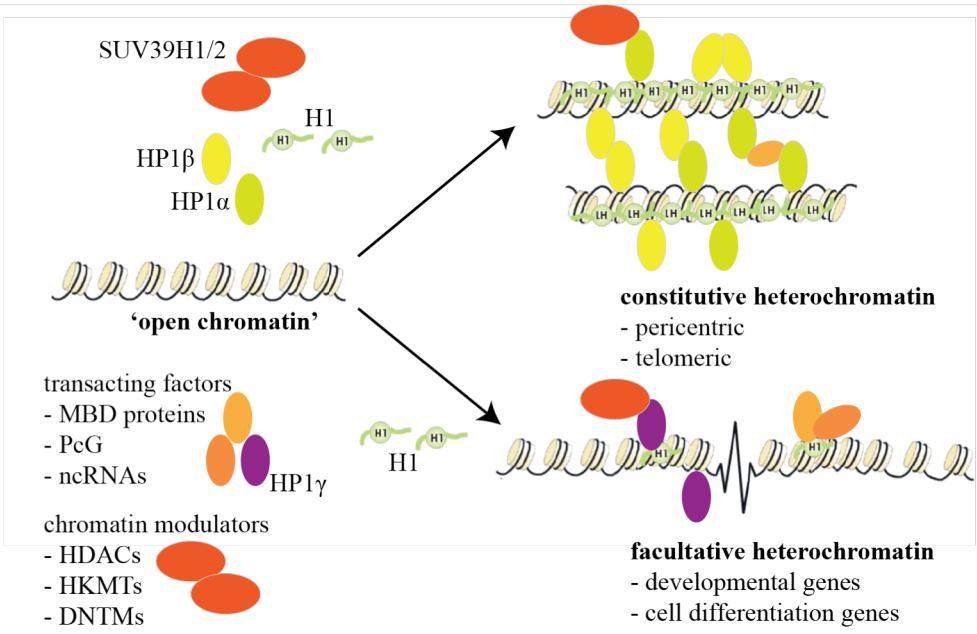


Figure 1-4 Constitutive vs. facultative heterochromatin.
There are two distinct types of heterochromatin: constitutive heterochromatin (CH) and facultative heterochromatin (FH). CH is highly condensed and predominantly contains HP1α/β and SUV39H1/2 enzymes. Condensation of CH is in part mediated by dimerisation of HP1 (hetero- or homo-dimerisation) proteins binding to adjacent histones or interaction with non-HP1 proteins. FH formation is regulated by chromatin modulators (chromatin modifying enzymes) and transacting factors such as HP1γ. FH structure is less condensed. Histone H1 variants are structural proteins that play a role in both CH and FH formation and maintenance. Image adapted from reference 77.

1.2.1. Structure and function of HP1

HP1 is widely conserved throughout the eukaryotic animal kingdom ranging from fission yeast to humans. While fission yeast have only one HP1 paralogue, *Drosophila* (HP1a, HP1b and HP1c) and mammals (HP1α, HP1β and HP1γ) have 3 paralogues. In mammals HP1α (22kDa), HP1β (25kDa) and HP1γ (21kDa) are expressed from the genes *CBX5*, *CBX1*, and *CBX3*, respectively. Members of the HP1 family are characterised by 2 evolutionary conserved domains: an N-terminal chromodomain (CD) and a C-terminal chromo-shadow domain (CSD), which are separated by a variable linker, the hinge region (Figure 1-5A). 3D structures of CD and CSD have been resolved by nuclear magnetic resonance (NMR) spectroscopy[103,104] and X-ray crystallography[105]. Both CD and CSD are globular domains consisting of an anti-parallel, three-stranded β-sheet, against which one (CD) or two α-helices (CSD) are packed. Despite their structural similarity, CD and CSD have distinctive functions. Functions of HP1's CD, hinge and CSD are summarised in Figure 1-6.

In general, HP1s are multifunctional proteins and their association with chromatin is directly or indirectly involved in a myriad of physiological processes such as regulation of gene expression, replication, DNA damage repair, telomere maintenance, cell differentiation and cell cycle progression.

1.2.1.1. The chromodomain (CD) of HP1

CD is known to be the chromatin-binding domain as it specifically interacts with N-terminal tails of H3 that are di- and trimethylated at K9[37,106-108]. Lack of, or mutations in the CD abolish localisation of HP1 at centromeric heterochromatin[37,109]. Figure 1-5C depicts the X-ray crystal structure of msHP1γ CD in complex with a H3K9me3 peptide. In 3-dimensional space, three key amino acids (in msHP1γ: F30, W51 and F53; highlighted in light blue Figure 1-5C) form a 'hydrophobic cage' that specifically recognises di- and trimethylation of H3K9. Other specific sites in the CD interact with the neighboring H3 amino acid residues to stabilize the flexible N-terminal histone tail[106]. Residues forming the 'hydrophobic cage' are highly conserved across HP1 homologues (compare Figure 1-5B). Furthermore, HP1's CD directly interacts with the globular domain of H3.

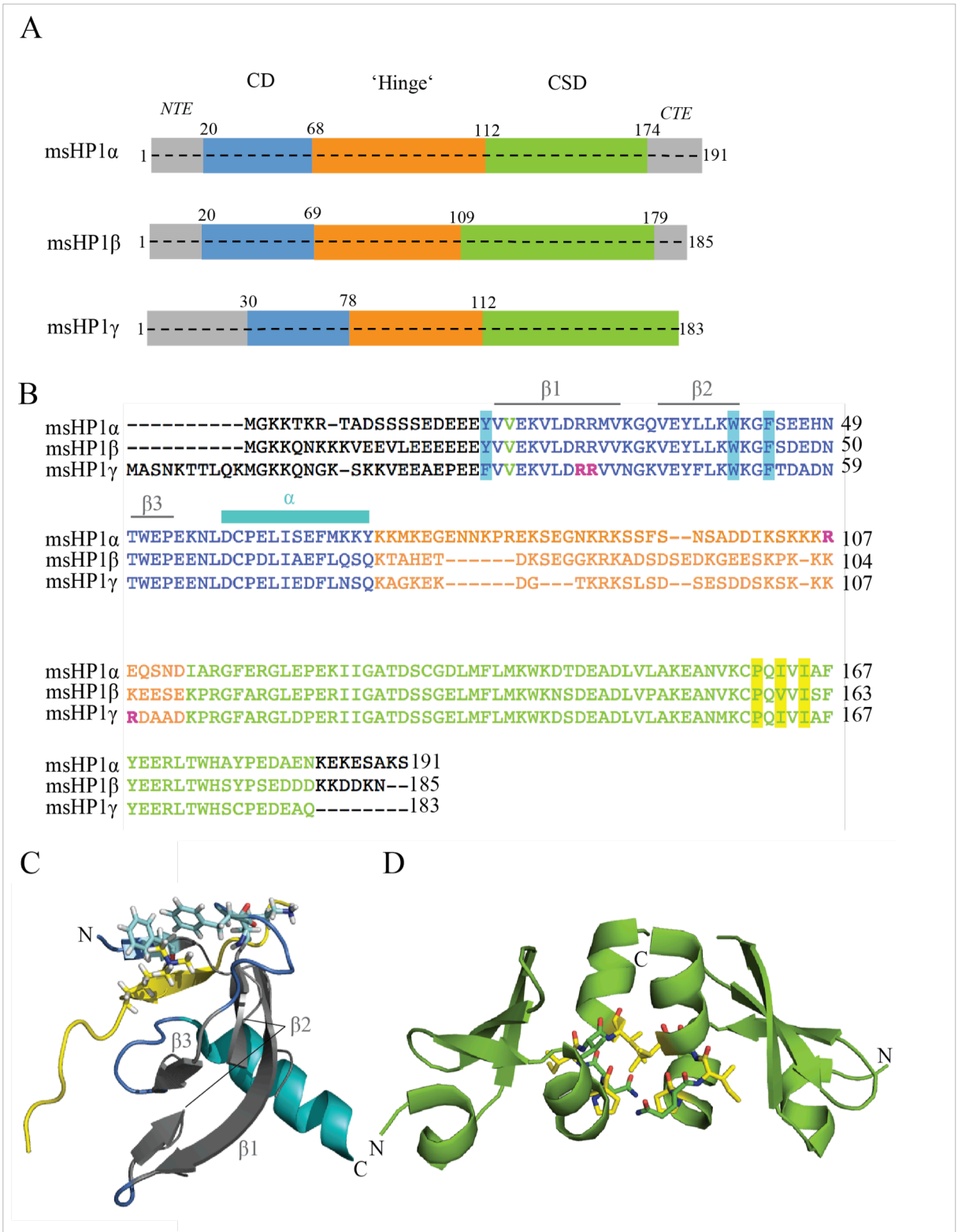


Figure 1-5 Functional domains of HP1.
(Figure legend is located on the next page.)

Figure 1-5 Functional domains of HP1.

A Cartoon of domain structure of mouse HP1 paralogues. Chromodomains (CD) are highlighted in blue, hinge regions in orange, chromo-shadow (CDS) domains in green and NTE (N-terminal extension) and CTE (C-terminal extension) in grey.

B Multiple sequence alignment of mouse HP1 paralogue protein sequences. Multiple sequence alignment was performed using Clustal Omega. CD, hinge region and CSD are colour-coded as indicated in **A**. Beta sheets and alpha helix of CD are indicated. Residues forming the hydrophobic pocket are highlighted in light blue and those of the PxVxL motif in yellow. Citrullinated residues are highlighted in pink. **C** X-ray crystal structure of msHP1γ CD (blue, turquoise, grey) in complex with H3K9me3 peptide (yellow). PDB file '5t1i' was adapted using the PyMOLE software. Beta sheets and alpha helix of CD are colour-coded as indicated in **B**. Residues F30, W51 and F53 forming the hydrophobic cage are highlighted in light blue.

D X-ray crystal structure of msHP1γ CSD dimer. PDB PDB file '3kup' was adapted using the PyMOLE software. Residues of PxVxL motif are highlighted in yellow.

1.2.1.2. The chromo-shadow domain (CSD) of HP1

The CSD is the dimerization domain of HP1[104]. HP1 molecules can interact with each other to form homo – or heterodimers between different homologues. Furthermore, HP1 interacts with several other proteins via its CSD[104,105]. The CSD contains a consensus sequence, the PxVxL motif (where P = proline, V = valine, L = leucine and x is any amino acid). Figure 1-5C depicts the X-ray crystal structure of the CDS dimer of mouse HP1γ, with PxVxL residues highlighted in yellow. HP1 interacts with a myriad of different chromatin associated proteins, many of which contain this consensus motif. This ever-growing list of HP1 interactors includes a number of chromatin components, chromatin remodelers, chromatin modifiers, transcriptional regulators, replication-, cell cycle- and DNA repair-implicated factors, and nuclear structure proteins[110]. A list of examples of HP1 interacting proteins are depicted in Table 1-2.

SUV39 homologues, as important interactors of HP1 also contain a PxVxL motif. In particular at pericentric heterochromatin, their interaction reinforces heterochromatin formation and spreading[110]. Furthermore, HP1 is part of KAP1 co-repressor complexes, where it interacts with KAP1 via its PxVxL motif [111]. Its interaction with HP1 anchors these complexes on chromatin via H3K9me3. When associated with KRAB zinc finger transcription factors, KAP1 co-repressor complexes are the main repressors of transposable elements in eukaryotic cells[112].

Table 1-2 Examples of HP1 interacting proteins. (Adapted from reference [113]).

Protein	Function	Domain
H3	chromatin structure	CD
H1	chromatin structure	CD, hinge
H3K9m2/3	chromatin structure	CD
SUV39H1/2	chromatin modification	CSD
DNMT3a/b	transcription	CSD
KAP1/TIF1β	transcription	CSD
TIF1α	transcription	CSD
TAFII130	transcription	CSD
E2F6	transcription	?
RB	transcription	?
SNF2β/BRG1	transcription	CSD
p150CAF-1	replication	CSD
KU70	DNA repair	CSD
INCENP	centromere structure	hinge
Lamin B receptor	chromatin structure	CSD
Lamin B	chromatin structure	CD
RNA	transcription/chromatin structure	Hinge
FACT	transcription	CSD

1.2.1.3. The hinge region of HP1

CDs and CSDs are connected via a less conserved and less structured hinge region. For example, the CD and CSD of all three mammalian paralogues share approximately 70% sequence identity, while hinge regions show less than 30% identity. The length of the hinge region varies substantially between paralogues. It contains HP1's nuclear localisation sequence and is the site of several post-translational modifications including phosphorylation and SUMOylation[114]. The hinge region is implicated in HP1's interaction with RNA and DNA[115,116]. Because of its weak sequence conservation, the hinge region represents an important source of functional differences between HP1 paralogues. HP1 proteins have different affinities for RNAs *in vitro*[116]. For example, *in vivo* interaction with RNA transcripts originating from pericentric repetitive elements is essential for the initial targeting of HP1α (not HP1β or γ) to pericentric domains and subsequent spreading of heterochromatin[96,116]. However, the molecular basis of this difference is not entirely understood, as the critical charged residues responsible for nucleic acid binding (KRK and KKK) are conserved between all paralogues. Structural differences between the hinge regions of HP1 paralogues are yet to

be identified. Due to its flexible nature, crystallisation of the hinge region is technically extremely challenging.

1.2.2. HP1 and chromatin compaction

The general binding affinity of HP1 to H3K9me3, although it differs between paralogues, is relatively weak and may not solely account for HP1's tight association with heterochromatin[110]. It is therefore likely that all three domains contribute to HP1's association with chromatin and all downstream effects that may result from it. The modular nature in which individual HP1 domains bind a nucleosome raises the possibility that different types of HP1-nucleosome complexes are formed depending upon which HP1 domains are engaged with the nucleosome and which HP1 domains are engaged with other proteins. In condensed chromatin, H3K9me3 within adjacent nucleosomes and therefore HP1 molecules are closer together; increasing the probability of HP1 proteins associated with adjacent histones to interact (Figure 1-4).

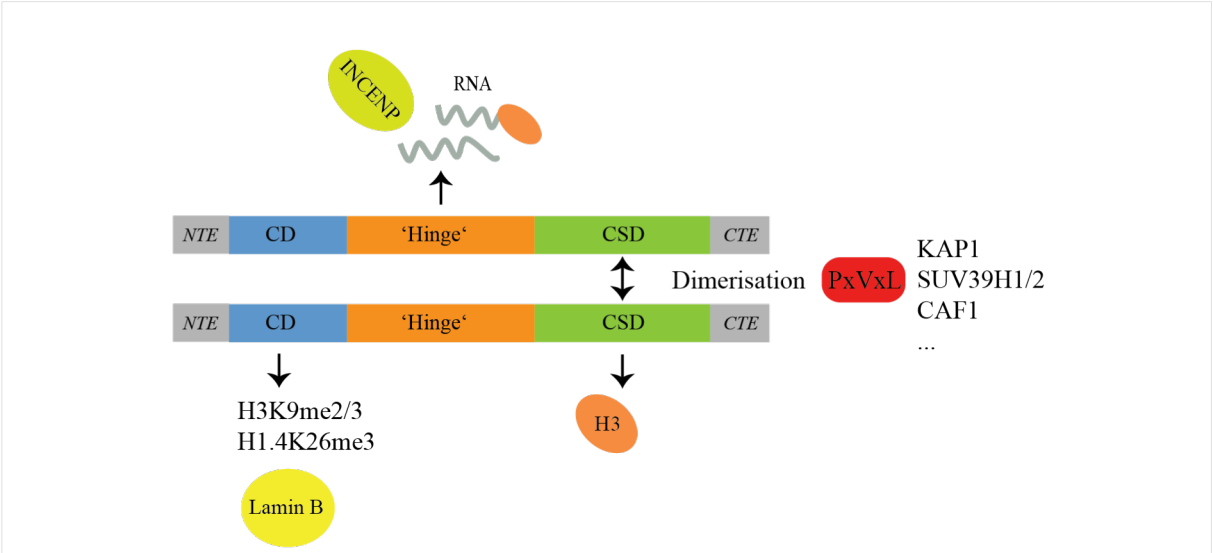


Figure 1-6 Schematic diagram of HP1.
All HP1 paralogues have a conserved N-terminal chromodomain (CD) and a C-terminal chromo-shadow domain (CSD), which are connected via a flexible hinge region. The main functions of each domain and examples of interaction partners are indicated.

For example, *in vitro,* human HP1α's specific recognition of H3K9me3 in tetra nucleosomes depends on both the hinge and the chromo-shadow domain[117]. Furthermore, whilst HP1α

predominantly associates with H3K9me3 on decondensed nucleosomes *in vitro*, HP1γ's binding affinity is specifically enhanced upon addition of histone H1 and Mg^{2+} ions, components that are known to induce compaction of nucleosomes[118]. HP1γ's increased binding affinity to condensed nucleosomes may, therefore, be due to HP1γ's direct interaction with H1. HP1γ is regarded as a trans-acting factor for facultative heterochromatin formation (Figure 1-4).

Condensation of chromatin could also be supported by HP1's interaction with different histone modifications. For example, besides H3K9me3, HP1 also binds methylated H1.4K26[119]. Since linker histone H1.4 is involved in heterochromatin structure, an interesting hypothesis could be that HP1 dimers bind to both modifications simultaneously to reinforce chromatin compaction.

The multipartite organization of HP1 proteins allows several proteins to bind simultaneously and has led to the idea that HP1 can function as a structural adaptor that may be vital for the assembly of macromolecular complexes on chromatin. A good example for this is HP1's role in sequence-specific heterochromatin formation mediated by small RNA pathways. H3K9me3 bound HP1 anchors RNA-induced silencing complexes composed of Argonoute proteins and small guide RNAs to chromatin[120]. These complexes further recruit chromatin modiying enzymes resulting in the spreading of silenced heterochromatin. This mechanism seems to be conserved from yeast to man[120].

1.2.3. Roles for HP1 in DNA damage repair

High compaction of chromatin in heterochromatin is hypothesized to constitute a barrier for DNA repair. DNA double strand breaks (DSBs) that are formed in heterochromatin are processed slower than in euchromatin[121]. Conversely, all HP1 paralogues are recruited to DNA lesions, indicating HP1 proteins play an active role in the DNA damage response[121]. In fact, HP1 depleted cells showed persistent γH2AX foci at late time points after irradiation, which indicates persistent and unrepaired DNA lesions. HP1's recruitment to DSBs is independent of H3K9me3, but depends on interaction with the largest subunit of the histone chaperone chromatin assembly factor 1, CAF1p150 and KAP1[121]. Loss of HP1 paralogues has no effect on non-homologous end joining (NHEJ), however, it affects homologous

recombination (HR). Interestingly, studies using a reporter construct to measure HR efficiency showed that loss of HP1α and β decreases the efficiency of HR, whereas HP1γ depletion results in increased HR efficiency[122,123]. These studies suggest different roles for HP1 paralogues in DNA repair. However, to date it remains unclear how these functional differences are mediated.

1.2.4. Roles for HP1 during cell cycle progression

Localisation of HP1 paralogues during the cell cycle is highly dynamic and paralogue specific functions are still being investigated. What is known is that HP1α predominantly associates with centromeres of metaphase chromosomes, whereas HP1β associates with interphase centromeres but is replaced at the centromere by HP1α as the cell enters metaphase. HP1γ seems to preferentially associate with distinct sites along the chromosomal arms in metaphase[110]. What exactly causes the redistribution of HP1 during cell cycle remains poorly understood. One possible mechanism is post-translational modification of HP1. Indeed, HP1α and γ, but not HP1β were reported to be differentially phosphorylated during the cell cycle. Different kinases such as PIM1 (proviral intergration site 1) and Aurora kinase B phosphorylate HP1 in specific phases of the cell cycle[101]. The latter also phosphorylates H3S10, which leads to eviction of all HP1 paralogues from chromatin during mitosis[124].

Furthermore, it has been proposed that the CD is required for HP1 localisation to chromatin in interphase nuclei, whereas during metaphase the CSD and hinge region that are needed for correct HP1 localisation, indicating that in mitosis HP1's association with chromatin is dependent on other proteins. Indeed, HP1α interacts with centromeric protein INCENP via the hinge region and HP1γ was reported to interact with cohesin during cell division[125].

HP1s are known to associate with proteins involved in chromatin assembly during DNA replication. In late S phase, HP1 interacts with CAF1p150, which transports a tetramer of histone (H3/H4) to the replication fork via interaction with proliferating cell nuclear antigen (PCNA). Recruitment of HP1 to the replication fork maybe important to maintain both heterochromatin integrity (especially at centromeres and telomeres) and possibly the epigenetic status of silenced genes. It suggests that CAF1 somehow promotes incorporation of HP1 into heterochromatin during its replication in late S phase[110].

Distinct localisations of HP1 paralogues during cell cycle progression indicate that they acquire specific functions. These differences seem to be, at least in part, regulated through post-translational modifications and/or interactions with different proteins. Further mechanistic insights are necessary to fully understand the function of each HP1 paralogue in cell cycle progression. Specifics for HP1γ will be discussed in Section 1.2.8.2.

1.2.5. Knockout of *CBX1, CBX5* and *CBX3* in mice

All three HP1 paralogues are ubiquitously expressed and due to high sequence conservation, they display functional redundancy. Nevertheless, it has become evident that each paralogue harbours specific functions with respect to chromatin formation and transcriptional regulation and we are only now starting to understand these functional differences. Knockout mice of HP1 paralogues show distinct phenotypes[126-128] (Table 1-3). Whilst *CBX1* (HP1β) knockout is perinatally lethal, *CBX5* (HP1α)-null mice develop normally. *CBX3* (HP1γ) knockout mice are sterile due to severe defects in spermatogenesis. These knockout phenotypes indicate that HP1 paralogues are not fully redundant.

Table 1-3 Phenotypes of HP1 knockout mice

Gene	Knockout phenotype
CBX5	no phenotype; mice develop normally
CBX1	perinatal lethal due to genomic instability; defects in neuronal development
CBX3	severe defects in spermatogenesis; C57BL/6 knock-out mice exhibit neonatal lethality

HP1 research of the past 2 decades concentrated mainly on HP1α and HP1β. The next sections summarise the main functional and structural differences between HP1 paralogues in mammalian cells with the main focus on HP1γ.

1.2.6. HP1α

HP1α is predominantly found in constitutive heterochromatin, especially in pericentric heterochromatin[110]. In general, binding affinity of HP1α to H3K9me3 is significantly lower than that of HP1β and γ. However, phosphorylation of mouse HP1α at serine residues in the N-terminal extension (NTE) significantly increases its affinity for H3K9me3 peptides and

H3K9me3 within nucleosomes[114,129]. Intriguingly, these NTE serine residues are only present in HP1α, indicating that NTE phosphorylation is a unique way of regulating HP1α's binding to chromatin. As constitutive heterochromatic regions must undergo coordinated assembly and disassembly during developmental transitions and cell cycle, HP1α's binding to chromatin, probably more than any other paralogue, has to be coordinately regulated over large stretches of the genome. Dynamic phosphorylation and dephosphorylation may be a way of regulating HP1α during processes requiring global nuclear rearrangement.

Two very recent studies reported an exclusive role for phosphorylated HP1α in initiating phase separation of heterochromatin *in vitro*[130,131]. These two papers very elegantly demonstrate that intracellular membrane-less compartments can contain heterochromatin, which suggests that heterochromatin-mediated gene silencing occurs in part through spatial organisation of compacted chromatin in phase-separated HP1α droplets. Droplet formation was significantly enhanced by phosphorylation of HP1α at the NTE and the hinge region[131].

HP1α seems to be the only paralogue that is regulated by SUMOylation. Lysine 84 in the hinge region of mouse HP1α is SUMOylated, which promotes its association with noncoding transcripts originating at major satellite repeats. Interaction with these RNAs is essential for HP1α's recruitment to pericentromeric regions and the silencing of repetitive elements located in these genomic loci[96].

Chromodomains of all human HP1s bind to a specific peptide sequence (PGTVAL) within the globular domain of histone H3[132]. HP1α additionally recognises H3Y41, which is located two residues upstream from the PGTVAL motif in H3. Phosphorylation of Y41 by nuclear JAK2 (Janus kinase 2) inhibits this interaction and HP1α's eviction from chromatin correlates with increased transcription of specific genes[133]. In contrast, neither HP1γ nor HP1β seem to require Y41 for binding the PGTVAL region and might therefore be insensitive to Y41 phosphorylation. The molecular nature of this specificity remains unclear.

Other interactors of HP1α include Brm (Protein Brahma Homolog) and Brg1 (BRM/SWI2-related gene 1), the two mutually exclusive catalytic subunits of the human SWI/SNF (hSWI/SNF) chromatin-remodeling complex. This interaction mediates repression of a

reporter construct by a transfected Gal4-HP1α fusion protein, indicating a co-repressive function of HP1α and the SWI/SNF complex[106].

In summary, HP1α is predominantly localised to constitutive heterochromatin. HP1α's recruitment to, and retention on chromatin are regulated through different post-translational modifications, such as phosphorylation and SUMOylation and by specific interaction partners.

1.2.7. HP1β

Similarly to HP1α, HP1β is mainly found in constitutive heterochromatin, but also plays a role in selectively silencing genes in hetero- and euchromatic regions. Very early on, mouse HP1β was reported to cause dose-dependent silencing of transgenes in centromeric regions, but not other genomic regions, indicating a chromosomal context dependent role for HP1β gene silencing[101]. Additionally, HP1β tethering experiments in mammalian cells showed that HP1β localisation leads to transcriptional repression of reporter genes located in repetitive regions of the genome, which is induced by chromatin condensation. Such chromosomal condensation was accompanied by recruitment of SETDB1 and subsequent spreading of H3K9me3[ref.113]. Hence, HP1β can play an active role in promoting local heterochromatin formation.

HP1β is excluded from active promoters, but it is found at promoters of inactive genes located in euchromatic regions. For example, HP1β is present at the repressed cyclin E promoter in a complex with the Rb (retinoblastoma) protein and SUV39H1, which is released upon Rb phosphorylation[134]. Furthermore, HP1β was shown to be involved in transient repression of an inducible viral promoter. In absence of stimulation, HP1β is bound to the promoter via H3K9me3, functioning as a negative regulator of transcription. Upon activation, HP1β is rapidly released, accompanied by removal of H3K9me3, subsequent H3S10 phosphorylation and recruitment of HP1γ[135]. This study nicely demonstrates that HP1 paralogues can acquire different, in this case even antagonistic functions.

CBX1-null mice die before or shortly after birth, while knockouts of other HP1 paralogues are viable. The *CBX1*-null phenotype is caused by genomic instability characterised by severe

chromosomal abnormalities that are indicative of pericentric heterochromatin dysfunction[126]. One hypothesis is that the essential interaction of HP1β with chromatin is not mediated by H3K9me3, but rather by its interaction with the globular domain of H3, which could be directly required for HP1β incorporation into heterochromatic regions[136]. However, binding affinities of all HP1 paralogues to the globular domain of H3 have not been compared yet.

To summarise, HP1β is predominantly found in constitutive heterochromatin, but it is also involved in transcriptional silencing of genes within euchromatin. The dramatic phenotype of *CBX1* knockout mice suggests a general role for HP1β in establishing higher order chromatin structures during development.

1.2.8. HP1γ

In contrast to HP1α and HP1β, HP1γ is less enriched in constitutive heterochromatin and predominantly found in euchromatic regions. The first study to identify mammalian HP1γ occupancy in the gene body of active genes, was published by the group of Gerd Blobel in 2005[73]. They showed for the first time that the repressive H3K9me3 mark can be deposited in the gene body of active genes, which is then bound by HP1γ. Ever since, many groups have attempted to decipher HP1γ's function in actively transcribed genes.

1.2.8.1. Regulation of gene expression

Depending on where HP1γ is located within a gene (promoter, gene body or 3'UTR), it can acquire distinct functions. Figure 1-7 summarises the roles of HP1γ in regulating gene expression that have been proposed within the last decade.

1.2.8.1.1. Promoter bound HP1γ

HP1γ at promoters functions either as a transcriptional repressor or activator. HP1γ mediated silencing of genes within euchromatin has been demonstrated using inducible viral transgenes. For example, HP1γ is part of a co-repressor complex, which is recruited to silence an actively transcribed hormone-regulated viral transgene[137]. This transgene contains a hormone-regulated KRAB repression domain. Upon hormone treatment, KAP1 co-repressor

is recruited to the promoter together with HP1γ and the histone methyltransferase SETDB1 leading to local H3K9me3 deposition and rapid gene silencing[137]. A similar mechanism was demonstrated for endogenous genes. In quiescent cells, together with histone methyltransferases, HP1γ is targeted to E2F response genes mediated by the transcription factor E2F-6, resulting in gene repression[138]. Both of these mechanisms involve local heterochromatin formation at the promoter of HP1γ target genes.

HP1γ can also be displaced from repressed gene promoters upon gene activation. In a breast cancer cell line carrying an integrated single copy of a transgene containing a hormone inducible MMTV-LTR promoter, HP1γ participates in the repression of the viral promoter[139]. Upon treatment, a hormone receptor complex triggers phosphorylation of H3S10, which results in rapid displacement of HP1γ from the promoter. Similarly, HP1γ was found to be part of a repressor complex together with the unliganded progesterone receptor (PR), ncRNA SRA (steroid receptor RNA activator) and several histone modifying enzymes[140]. This complex is rapidly displaced from promoters upon hormone stimulation, which is mediated though the H3K9me3 - H3S10ph switch (compare 1.1.5.3)[140].

In contrast to HP1γ's gene silencing function, it is also recruited to many promoters upon activation. Several different stimuli have been reported to trigger HP1γ's recruitment to chromatin, such as cytokine treatment, heat shock or treatment with drugs such as phorbol ester PMA[132,135,141-144].

In the human colorectal cancer cell line HCT116, upon treatment with DNA damaging agents HP1γ is recruited to the promoter of the *BIRC5* gene, which coincides with HP1β's displacement[142]. The antagonistic relationship between silencing HP1β and activating HP1γ seems to be a common mechanism to regulate inducible genes. However, it remains an open question how HP1γ is binding these active promoters.

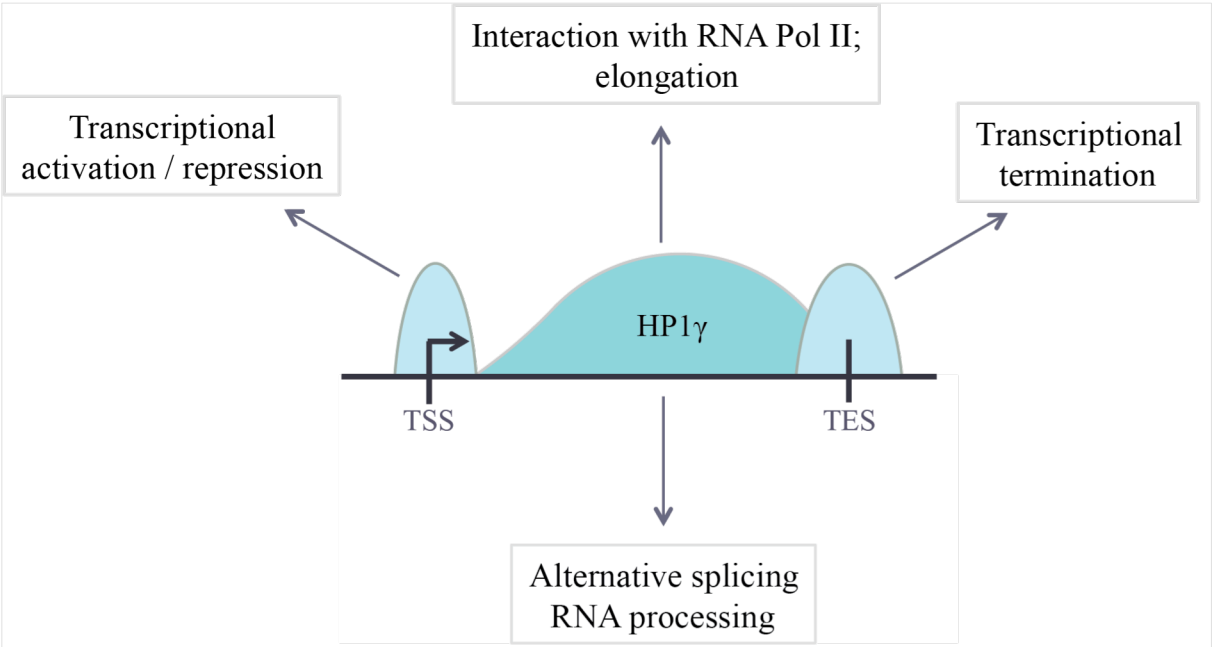


Figure 1-7 Functions of HP1γ in euchromatin.
Schematic illustration of a gene with transcriptional start site (TSS) and termination site (TES), HP1γ binding at TSS, gene body and TES is indicated together with associated functions of the protein.

One possibility is that HP1γ directly interacts with the core of H3. HP1γ was shown to be recruited to the promoter of interferon inducible genes upon interferon alpha (IFNα) treatment and its recruitment is essential for gene activation[132]. Interestingly, this recruitment depends on the chromatin remodeling complex SWI/SNF. As the core of H3 is not directly accessible for the CSD of HP1, the action of SWI/SNF may open up the chromatin enabling HP1γ to bind H3[132].

Another option is that HP1γ specifically interacts with certain histone variants. Evidence supporting this idea comes from the human HSP70 (heat shock 70kDa protein 2) genes. HP1γ is recruited to the HSP70 promoters upon heatshock induced gene activation, which is accompanied by increased incorporation of histone variant H3-3[144]. Both HP1γ and H3.3 physically interact and loss of either protein results in repression of *HSP70,* indicating HP1γ and H3.3 cooperate in activating *HSP70* upon heatshock[144].

In summary, HP1γ bound at promoters can either function as a transcriptional repressor or activator. In both settings it is likely to be part of larger protein complexes.

1.2.8.1.2. HP1γ bound in gene bodies

HP1γ is the only HP1 paralogue that is found within gene bodies of actively transcribed genes. Several functions for HP1γ within active genes and distinct mechanisms of recruitment have been proposed. As mentioned beforehand, HP1γ binding within active gene bodies can coincide with H3K9me3 deposition[73,145]. But not every HP1γ target gene is marked by H3K9me3[143]. What about other genes?

HP1γ directly interacts with RNA Pol II, more precisely with the elongating form of RNA Pol II as determined by immunoprecipitation using antibodies against the carboxy-terminal domain of RNA Pol II phosphorylated at serine 5[ref.73,141,143]. This interaction can result in different transcriptional outputs. HP1γ was proposed to function globally as a transcriptional activator[143]. In HCT116 cells, genes bound by HP1γ within the gene body were found to be predominantly down-regulated upon HP1γ depletion, indicating a positive influence of HP1γ on elongation[143]. Work in drosophila showed that HP1c interacts with the histone chaperone complex FACT and recruits it to active genes, which links FACT to the active form of RNA Pol II[146]. Loss of HP1c impairs recruitment of FACT to heat-shock loci and causes a defect in heat-shock gene expression.

Conversely, HP1γ was also reported to limit RNA Pol II's elongation rate[141,145,147]. For example, HP1γ can function as a transcriptional modulator to 'fine tune' gene expression of immune response genes in mouse embryonic fibroblasts (MEFs)[141]. In unstimulated cells immune response genes are repressed by HP1γ binding to H3K9me3 at their promoters. Upon infection with bacteria, H3K9me3 is removed and H3S10 is phosphorylated by the map kinase MSK1. The same kinase phosphorylates HP1γ at serine 83 (S83) in the hinge region, which they show is a prerequisite for the protein to bind RNA Pol II[ref.141]. Its interaction with elongating polymerase results in moderate transcription. Only upon complete loss of HP1γ do these immune genes become hyper activated in response to stimuli, resulting in a full-blown immune response. The authors propose a model in which phosphorylated HP1γ acts as a transcriptional modulator to fine-tune the expression of immune response genes[141].

Other studies suggest a role for HP1γ in attenuating RNA Pol II elongation rate[145,147], which can result in different effects on gene expression. For example, HP1γ regulates inclusion of

alternate exons within the *CD44* gene via a mechanism involving decreased RNA polymerase II elongation rate. *CD44* is comprised of constant and variant exons. The latter are subject to enhanced alternative splicing. Together with H3K9me3, HP1γ is specifically enriched at variant exons of the gene[145]. Furthermore, it directly binds the nascent pre-mRNA, thereby stabilising the RNA on chromatin. It is therefore regarded as a bridging factor connecting chromatin and pre-mRNA to create structures that slow elongating RNAP II, which in turn facilitates recruitment of the spliceosome. Besides *CD44,* loss of HP1γ results in decreased inclusion of alternate exons of several other genes, indicating HP1γ plays a more general role in alternative splicing[145]. Smallwood *et al.* addressed HP1γ's role in alternative splicing genome wide[143]. In contrast to the previous study, they reported that loss of HP1γ results in a general increase in intron inclusion and loss of splicing and therefore in accumulation of unspliced RNAs, predominantly those originating from genes bound by HP1γ[143]. The same study also showed direct interaction of HP1γ with components of the spliceosome, demonstrating a direct link between the splicing machinery and HP1γ on chromatin. In conclusion, HP1γ favours alternate exon usage by slowing down elongating RNA Pol II and loss of HP1γ either results in reduced usage of alternate exons[145] or a general reduction of splicing[143]. Both studies propose a role for HP1γ in alternative splicing. However, it must be noted that the effects on splicing displayed in both studies are fairly mild and gene specific, which suggests that HP1γ should be regarded as a modulator rather than a main regulator of alternative splicing.

1.2.8.1.3. HP1γ bound to transcriptional termination sites

In the previous section, a functional association between repressive histone marks, HP1γ, Pol II pausing and pre-mRNA processing was discussed[145]. A similar mechanism of RNA Pol II pausing is involved in transcriptional termination. At termination sites G-rich sequences promote R-loop formation, which in turn leads to the synthesis of local dsRNA[147]. As a result dsRNAs recruit the RNAi machinery leading to local heterochromatin formation by deposition of H3K9me2 and recruitment of HP1γ. This local heterochromatin promotes RNA Pol II pausing, a requirement for efficient termination of transcription[147]. HP1γ's recruitment to chromatin within active genes is therefore involved in two very different processes that both require attenuation of RNA Pol II's elongation rate.

To conclude, in contrast to HP1α and β, HP1γ is predominantly found at active and silenced genes within euchromatin, where it is directly associated with gene regulation. Depending on where it is bound within a gene it can acquire distinct functions. At promoters, HP1γ can act as a bivalent regulator of gene expression, as it regulates gene silencing as well as activation. HP1γ occupancy within gene bodies is linked to its interaction with the elongating RNA Pol II. This interaction can either enhance the rate of transcription by of RNA Pol II or attenuate it. HP1γ mediated attenuation of RNA Pol II is associated with alternate splicing, transcriptional termination or 'fine tuning' of transcription of inducible genes.

How HP1γ mediates RNA pol II's elongation rate and whether this interaction is the only mechanism of recruiting HP1γ to gene bodies of active genes remains unknown. Due to the multipartite nature of HP1γ, it is likely that additional, yet unknown interactors are involved in HP1γ mediated gene regulation.

1.2.8.2. Structural features of HP1γ

In contrast to HP1α and β, HP1γ contains 10 extra amino acids at its NTE (Figure 1-5B). Surprisingly little is known about this N-terminal extension. Although the peptide sequence (msHP1γ 1-10: HP1MASNKTTLQK) is not modified in untreated mammalian cells[148], it may influence HP1γ's interaction with histone or non-histone proteins. Furthermore, HP1γ possesses fewer acidic residues in its NTE compared to HP1α and HP1β. Molecular dynamics studies suggest that these differences make the HP1γ CD-H3K9me3 complex more flexible than the HP1β CD-H3K9me3 complex, indicating structural differences between HP1β and HP1γ heterochromatin[149]. Since active transcription requires a flexible chromatin conformation, structural differences between HP1β and HP1γ may be one of the reasons why predominantly HP1γ is recruited to H3K9me3 within active genes.

The hinge regions of HP1 paralogues show weak sequence conservation. HP1γ's hinge region shows lower binding affinity to RNA *in vitro* than that of HP1α[116]. *In vivo*, binding of HP1γ to RNA may be dependent on the correct display of the target within a chromatin structure. Alternatively, HP1γ RNA binding may be more sequence-specific than that of HP1α. In support of this notion is HP1γ's role in modulating splicing of specific transcripts. HP1γ's interaction with nascent pre-mRNAs seems to be gene specific, as not all HP1γ bound genes

show an effect on splicing upon depletion of the protein[143]. Serine residue 83 within the hinge region of HP1γ is differentially phosphorylated[150], whilst the same residue in HP1α is subject to SUMOylation[96]. HP1γ's recruitment to active genes and its interaction with RNA Pol II depends, at least in part on phosphorylation of S83[150,151]. Kinases known to mediate S83 phosphorylation of HP1γ are IKKα (inhibitor of nuclear factor κ-B kinase α)[151], MSK1[141], PKA (protein kinase A)[150], PIM1[152], HIPK2 (homeodomain interacting protein kinase 2)[153] and Aurora kinase B[154]. The latter phosphorylates HP1γS83 specifically during G_2/M phase of the cell cycle. HP1γS83ph then localises to the spindle pole, which is necessary for proper cell division in HeLa cells[154]. Loss of HP1γ or rescue with a S83A mutant in the same cell line results in severe mitotic aberrations. Besides the interaction with RNA Pol II, phosphorylation of HP1γ promotes its interaction with other proteins such as the NHEJ repair protein KU70[150].

Taken together, HP1γ's exclusive functions in euchromatin may partially be explained by structural differences of the protein compared to other HP1 paralogues. Furthermore, HP1γ is regulated by post-translational modifications, which promote interactions with other proteins and affect subnuclear localisation.

1.2.8.3. HP1γ in cancer

HP1 proteins are involved in many different physiological processes such as cell cycle progression, DNA repair, gene regulation or telomere and centromere stabilisation, all of which are directly linked to tumourigenesis. It is therefore not surprising that HP1 paralogues are involved in cancer progression. To date no human diseases have been associated with mutations in the HP1 genes *CBX1*, *CBX5* and *CBX3*. However, changes in the levels of expression have been detected in different cancers.

High protein levels of HP1γ were linked to enhanced cell proliferation and oncogenesis in different cancers such as colorectal [155,156], non-small cell lung [157], prostate [158], cervical and breast[159,160]. Mechanistic insights into how HP1γ may be involved in tumourigenesis, however, remain poorly investigated. Given HP1γ's versatile role in regulating gene expression, overexpression of HP1γ in cancer cells is likely to be linked to aberrant gene regulation. First insights into how HP1γ may be involved in deregulating genes linked to

malignant growth were provided in colorectal cancer[155,156]. In human colorectal cancer HCT116 cells, elevated levels of HP1γ repress gene expression of the cell cycle regulator p21(Waf1/Cip1) resulting in uncontrolled cell proliferation[155]. Consequently, knockout of HP1γ in HCT116 cells leads to cell cycle arrest and impaired proliferation *in vitro* and *in vivo*. Survival of mouse with xenografts of HP1γ depleted HCT116 is significantly increased when compared to control mice. The same study identified a tumour-suppressive microRNA (miR-30a) that is commonly down-regulated in colorectal cancer to target HP1γ[155].

Another recent study demonstrated a role for HP1γ mediated gene regulation in controlling the tumour immune environment[161]. In CD8 positive T cells HP1γ suppresses the expression of effector T cell factors such as perforin, granzyme B and the inflammatory cytokine interferon γ[161]. These factors are all necessary for CD8 positive effector T cells to kill target cells. HP1γ deficient CD8 positive T cells express high levels of these factors and therefore harbour increased tumour-killing capacity. In fact, the tumour burden of mice treated with HP1γ deficient CD8 positive T cells is reduced, indicating that targeting HP1γ in CD8 positive T cells may be a novel therapeutic approach to control growth of solid tumours[161]. HP1γ mediated gene regulation and their association to cancer progression may therefore be highly cell type specific.

Given the number of putative HP1 interactors it is likely that HP1γ participates with, or acts downstream of other oncogenic transcriptional regulators. For example, HP1γ is part of repressive steroid hormone receptor complexes (compare Section 1.2.8.1.1), which are deregulated in different solid and blood cancers[162]. Furthermore, several H3K9 methyltransferases are known to be deregulated in cancer[163]. Thus, HP1γ (and other HP1s) could act downstream of these enzymes to reinforce their malignant behaviour.

Besides regulating gene expression, HP1γ may contribute to tumourigenesis by directly regulating physiological processes. Loss of HP1γ causes chromosome segregation defects, while over-expression of the protein enhances cell proliferation[154]. HP1γ interacts with cohesin at the spindle poles and is thought to be physically involved in chromosome segregation. Furthermore, HP1γ interacts with the shelterin complex at human telomeres and is therefore involved in sister telomere cohesion and maintaining telomere lengths by

mediating telomerase activity[164]. Telomerase and telomere length are often deregulated in human cancers[165].

HP1γ plays a role in DNA damage repair. It seems to be the only paralogue that suppresses homologous recombination (see Section 1.2.3), suggesting that enhanced levels of HP1γ could severely affect genomic instability - a key feature of cancer progression[165]. Additionally, HP1γ interacts with different proteins directly involved in the repair of DSBs[150,153,166,167].

In summary, elevated levels of HP1γ protein are associated with enhanced cell proliferation and tumourigenesis. HP1γ is a versatile transcriptional regulator and its over-expression results in aberrant gene regulation linked to malignant cell growth. Given its function in cell cycle regulation, DNA repair and telomere regulation, HP1γ is a protein with high oncogenic potential. However, to date, only a few studies report mechanistic insights into how HP1γ may be involved in cancer progression.

1.2.8.4. HP1γ in stem cell differentiation

Mouse embryonic stem cells (mESCs) are derived from the inner cell mass of the blastocyst approximately 4 days after fertilisation. They are pluripotent stem cells with the unique ability to differentiate into every cell type of the body and to self-renew. These characteristics correlate with a distinct nuclear architecture, epigenetic signatures enriched for active chromatin marks and hyperdynamic binding of structural chromatin proteins. The process of differentiation is accompanied by major chromatin remodeling, which predominantly requires formation of facultative heterochromatin to establish cell type specific gene expression patterns. HP1 and other structural chromatin factors such as histone H1 play important roles in this process[97,168]. It is therefore not surprising that the subnuclear localisation of HP1 changes dramatically during differentiation. For all three paralogues, increased heterochromatic foci formation was observed upon differentiation of mESCs by immunofluorescence staining[169],

Overexpression (OE) of HP1γ in P19 embryonal carcinoma (EC) cells, which are pluripotent cells that share many features of mESCs, affects spontaneous cell differentiation[170]. Undifferentiated HP1γ OE cells exhibit no phenotypic differences to P19 parental cells and

continue to express pluripotency markers. Interestingly, after embryoid body formation, HP1γ OE cells spontaneously differentiate equally into cells of all three germ layers (meso, endo and ectoderm), indicating that HP1γ may play a role in the decision of cell fate[170]. In agreement with these results is the recent observation that HP1γ regulates neuronal differentiation[171]. Loss of HP1γ significantly impairs differentiation of mESCs to neural progenitor cells (NPCs), while up-regulating genes involved in mesodermal lineage decisions. In mESCs, HP1γ binds to promoters of differentiation-related genes, suggesting it directly regulates neural and mesodermal gene expression programs during cell fate decision[171]. In addition, loss of HP1γ enhances reprogramming by de-repressing expression of the pluripotency gene *Nanog*[172]. Importantly, the same study identified a switch in the location of HP1γ between pluripotent and non-pluripotent cells, confirming its dynamic role in gene regulation during stem cell differentiation. In summary, these studies highlight an important function for HP1γ in maintaining the integrity of pluripotent stem cells and regulating the establishment of cell type specific gene expression during differentiation.

1.2.9. Dynamics of HP1 molecules

The maintenance of stable heterochromatin domains in living cells involves the transient binding and dynamic exchange of HP1 from chromatin. Fluorescence after photo-bleaching (FRAP) experiments using living cells with GFP-tagged HP1 variants showed relatively rapid exchange of HP1 molecules in both heterochromatin and euchromatin. This implies that heterochromatin is not a static entity and not completely inaccessible to other molecules[109,173,174]. Although HP1 molecules are highly dynamic, the mobility of all three HP1 paralogues is significantly lower in heterochromatin when compared to euchromatin. For example, for euchromatic HP1 full recovery after photo-bleaching is reached after 5s, whilst it takes 60s to recover when HP1 is associated with heterochromatin[109]. This difference in mobility or the amount of the 'immobile' fraction in these two regions may be due to differences in the density of HP1 binding sites in euchromatin versus heterochromatin.

With fluorescence correlation spectroscopy (FCS), a technique to determine the diffusion coefficient by monitoring the Brownian movement of individual proteins, different populations of molecules within euchromatin and heterochromatin with distinct diffusion modes were discovered[174,175]. This analysis revealed at least two differently mobile HP1

populations in euchromatin and three in heterochromatin of mammalian cells, the latter including a fraction of very slow molecules in constitutive heterochromatin. There is a highly mobile population (fast) of all HP1 isoforms with diffusion coefficients (DC) ranging between 0.6 and 0.7±0.03$\mu m^2 s^{-1}$ and a second (slow), much less mobile population with diffusion coefficients ranging from 0.07 to 0.04±0.02$\mu m^2 s^{-1}$ref.174,175. In heterochromatin, an additional population (DC<<0.04$\mu m^2 s^{-1}$) of immobile molecules was identified. In euchromatin, about 70% of molecules show fast and 30% slow diffusion, while in heterochromatin the opposite is the case with 30% of molecules showing fast and 60% slow diffusion, while the immobile population contains about 10% of all molecules[174,175]. Taken together, these analyses indicate three distinct binding modes of HP1 molecules on chromatin. In heterochromatin the majority of molecules move relatively slowly, whilst in euchromatin the majority of molecules diffuse relatively fast. These studies did not detect any significant differences between HP1 paralogues.

Studying the plasticity of HP1 molecules during stem cell differentiation can provide valuable insights into how HP1 paralogues orchestrate local and global chromatin condensation. mESCs show highly diffused staining of HP1, while differentiation to NPCs is accompanied by the appearance of defined heterochromatic foci[176]. FRAP experiments revealed that a global reduction of dynamic binding of HP1 precedes formation of distinct heterochromatin foci, suggesting that immobilisation of HP1 is an early step in differentiation-dependent chromatin remodeling and silencing process[176]. To date, no differences in subnuclear dynamics of HP1 paralogues during differentiation have been identified.

In summary, the maintenance and *de novo* formation of heterochromatin during physiological processes such as cell differentiation involves the transient binding and dynamic exchange of HP1 molecules on chromatin. In general, within mammalian nuclei the fraction of HP1 molecules with fast diffusion is higher in euchromatic regions than in heterochromatic regions.

1.3. Aims of PhD thesis

The overall purpose of this PhD thesis is to discover novel functions of HP1γ in different mammalian cell systems. This work has the following aims:

I. Determine the role of residues R38 and R39 located within the CD in HP1γ's association with chromatin and characterise a novel post-translational modification of HP1γ at these very residues *in vitro* and *in vivo* using mouse embryonic stem cells (mESCs).

II. Understand whether HP1γ regulates growth of breast cancer cells by characterising the phenotype of HP1γ depletion in different human breast cancer cell models. Furthermore, I asked the question whether HP1γ is a transcriptional regulator in breast cancer cells using whole transcriptome analysis in cells depleted of HP1γ, cross-referenced the with genomic localisation of HP1γ.

Chapter 2

Materials and methods

2.1. General materials

2.1.1. Standard solutions

The composition of bacterial media and stock solutions used in this thesis are shown in Table 2-1. All modifying and restriction enzymes were purchased from Roche or New England Biolabs.

Table 2-1 Standard solutions.

Name	Composition per litre
LB	5g NaCl, 5g yeast extract, 10g bacto-tryptone
SOB	20g bacto-tryptone, 5g bacto-yeast, 0.58g NaCl, 0.19g KCl,10mM MgCl2, 10mM MgSO4 and 0.4% glucose
SOC	20g bacto-tryptone, 5g bacto-yeast, 0.58g NaCl, 0.19g KCl, 10mM MgCl2 and 10mM MgSO4
TE	10mM Tris, 1mM EDTA, pH 8.0
TBE	90mM Tris-borate, 2mM EDTA, pH 6.5
TBS	150mM NaCl, 20mM Tris-HCl pH 7.6
PBS	150mM NaCl, 2.5mM KCl, 10mM Na2HPO4, 2mM K2HPO4
IPH	150mM NaCl, 50mM Tris-HCl pH 8.0, 5mM EDTA, 0.5% NP-40
RIPA	10mM Tris (pH 8.0), 1mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, 140mM NaCl

2.1.2. Vectors

The plasmids used for the various applications detailed below are summarised in Table 2-2. The cDNA for human PADI4 was a kind gift from Maria Christophorou (MRC Human Genetics Unit, University of Edinburgh). The vector pSB7 was a kind gift from Srinjan Basu (Department of Biochemistry, University of Cambridge). It was generated from the pEF/*myc*/ER backbone

(Invitrogen), where the *myc*/ER regulatory elements were replaced with the sequence encoding mEeos3.2-HaloTag. The remaining plasmids were present as stocks in the Kouzarides laboratory (Gurdon Institute).

Table 2-2 Vectors.

Vector	Insert	Application
pcDNA3.1	Human HA-PADI4	Transfection into mammalian cells
pcDNA3.1	Mouse HA-HP1γ	Transfection into mammalian cells
pSB7	Mouse HA-HP1β/γ WT and mutants	Transfection into mammalian cells
pCMV6-SPORT	Human PADI2	Transfection into mammalian cells
pLKO.1	shHP1γ_1, shHP1γ_2, shCtrl (Sigma)	Lentiviral infection of mammalian cells
lentiCRISPR v2	sgRNAs against human *CBX3*	Lentiviral infection of mammalian cells
PGEX-2T	Mouse HP1β/γ WT and mutants	Recombinant protein expression from *E. coli*
PGEX-6P	Human PADI4	Recombinant protein expression from *E. coli*

2.2. Bacterial strains and transformation

2.2.1. *E.coli* strains

XL10-Gold: *Tet^{r}D(mcrA)183 D(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte [F′ proAB lacI^{q}ZDM15 Tn10 (Tet^r) Amy Cam^r]*

BL21(DE3): *F$^-$ ompT gal dcm lon hsdS$_B$(r_B^- m_B^-) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])*

DH5a: *F$^-$ endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(r_K^- m_K^+), λ–*

2.2.2. Growth of bacterial cultures

Plated cultures were grown at 37°C overnight on LB plates (LB with 20g bacto-agar per litre). Bacterial suspension cultures were grown in LB at 37°C overnight. To allow the selection of bacteria transformed with plasmid DNA expressing the gene of interest, ampicillin (100μg/ml) or kanamycin (100μg/ml) were included in the media.

2.2.3. Transformation of *E. Coli* by heat shock

For the transformation of *E. Coli*, 0.1μg plasmid DNA was added to 100μl of competent bacteria and incubated on ice for 30 min. Samples were heat shocked for 45 seconds at 42°C, returned to ice for 3 min following which 1ml of SOC was added. This was incubated at 37°C for 1 hour. Bacteria were plated immediately.

2.3. Preparation of DNA

2.3.1. Large scale (maxi prep) plasmid DNA preparation from *E. coli*

The Qiagen Maxiprep kit was used according to the manufacturer's instructions. 300ml LB and antibiotic was inoculated with a single colony of the appropriate transformant and grown overnight at 37°C. Bacteria were harvested by centrifugation at 3500rpm for 10 minutes. The bacterial pellet was resuspended in 10ml P1 solution (50mM Tris-HCL pH 8.0, 10mM EDTA, 100ug RNase A) and then lysed by the addition of 10ml P2 solution (200mM NaOH, 1%SDS). Following 5min incubation at room temperature, 10ml ice-cold P3 solution (3.0M potassium acetate pH 5.5) was added, the lysate was gently mixed by inversion and incubated on ice for 20 minutes. The lysate was cleared by centrifugation for 30 minutes at 4000rpm, 4°C. The cleared lysate was applied through muslin to a Qiagen 500 (anion-exchange) column, previously equilibrated with 10ml QBT buffer (750nM NaCl, 50mM MOPS pH 7.0, 15% isopropanol, 0.15% Triton-X100). The column was washed two times with 30ml QC buffer (1.0M NaCl, 50mM MOPS pH 7.0, 15% isopropanol). Plasmid DNA was eluted with 15ml buffer QF (1.25M NaCl, 50mM Tris pH 8.5, 15% isopropanol) and precipitated by the addition of 10.5ml isopropanol. The DNA precipitate was centrifuged for 30 minutes at 4000rpm, 4°C. The DNA

pellet was washed with 70% ethanol and dried prior to being dissolved in TE or nuclease-free H_2O. DNA concentration was determined by measuring the absorbance at 260nm using a NanoDrop® spectrophotometer.

2.3.2. Small scale (mini prep) plasmid DNA preparation from E. Coli

The QIAprep Miniprep kit (Qiagen) was used according to the manufacturer's instructions. 3ml LB and antibiotic was inoculated with a single bacterial colony and grown at 37°C overnight. Bacteria were pelleted by centrifugation at 3500rpm for 5 minutes. The bacterial pellet was resuspended in 250µl P1 buffer (50mM Tris-HCl pH 8.0, 10mM EDTA, 100µg/ml RNase A) and lysed by the addition of 250µl of P2 buffer (200mM NaOH, 1% SDS). Following incubation at room temperature for 5 minutes, 350µl ice cold N3 buffer was added, the solution mixed by inversion and incubated on ice for 5 minutes. The debris was pelleted by centrifugation and the supernatant transferred to a Qiagen plasmid spin column and centrifuged again for 1 minute. The column was sequentially washed in 500µl PB buffer, 750µl PE buffer and centrifuged for 1 minute. After the final wash, the column was spun a second time to remove any residual buffer and then the DNA eluted in 50µl EB buffer (10mM Tris-HCl, pH8.5) by micro-centrifugation for 1 minute.

2.4. DNA cloning

Several precautions were taken to maximise the generation of recombinant plasmids. (i) Both vector and insert fragments were gel-purified to exclude the presence of undigested plasmids in the ligation reaction. (ii) Recircularisation of vector fragments was minimised in two ways: (a) Where possible DNA fragments were inserted into vectors cut with two enzymes generating incompatible ends. (b) If compatible ends were generated, the 5' terminal phosphate was removed from vector DNA with alkaline phosphatase. Ligation by T4 DNA ligase requires the presence of 5' phosphate groups on only one of the two DNA fragments.

2.4.1. DNA restriction digests and dephosphorylation

0.5-1µg DNA was digested with 5-10 units of restriction endonuclease for 1 hour at 37°C. The enzyme buffer and incubation temperature were followed according to the manufacturer's instructions (NEB). 10 units of calf intestinal alkaline phosphatase (Roche) was added directly to the completed restriction digest of the vector DNA, and incubated for a further 15 minutes at 37°C. 6x agarose gel loading buffer (0.25% bromophenol blue, 30% glycerol in H2O) was added to the digestions before gel electrophoresis.

2.4.2. Agarose gel electrophoresis

DNA fragments were resolved using horizontal agarose gels. 0.8-2.0% agarose gels were made by dissolving agarose in 150ml TBE and adding ethidiumbromide to a final concentration of 0.5µg/ml. The gel was poured into a tray and allowed to set. Once solid, the gel was mounted in an electrophoresis tank and covered with TBE. DNA samples were loaded onto the gel in 1x agarose gel loading buffer, and electrophoresed at 140V. 1kb ladder (Invitrogen) was run alongside the samples as a marker.

2.4.3. Purification of DNA from agarose gels

DNA that was stained with ethidiumbromide was visualised in an agarose gel by low wavelength UV light. The relevant fragments were cut from the gel with a scalpel. DNA was extracted from the agarose using the QIAquick Gel Extraction kit (Qiagen), according to the manufacturer's instructions. Three volumes of QG buffer were added to the agarose slice and the mixture incubated at 50°C for 10 minutes with shaking. 1 gel volume of isopropanol was added to the dissolved agarose, the mix was then applied it to the Qiagen spin-column and micro-centrifugation. The DNA was retained in the column. The column was then washed with 750µl of PE buffer and micro centrifuged for 1 minute. Any residual buffer was removed by a second spin. DNA was eluted in 30µl of EB buffer.

2.4.4. DNA ligations

Purified, linearized vector and insert DNA were ligated with T4 DNA ligase, using the Rapid DNA ligation Kit (NEB) according to the manufacturer's instructions. A molar ratio of 1:5 or 1:7 (vector : insert) DNA and 5 units of T4 DNA ligase were used as appropriate. The ligation reaction was carried out for 5 minutes at room temperature before transformation into competent *E. coli.*

2.4.5. Polymerase chain reaction (PCR)

PCRs were carried out in PCR Eppendorf Mastercycler Nexus Gradient Thermal Cycler. PFU (Stratagene) was used according to manufacturer's instructions. See Table 2-3 for reaction mixtures and protocol. After the PCR, 5µl of the PCR product was analysed by agarose gel electrophoresis. Table 2-4 summarises primers used to amplify cDNA constructs for restriction cloning.

Table 2-3 PCR protocols.

	Standard PCR	Mutagenesis PCR
Mix:		
Made up to 50µl in ddH_2O	5 µl each primer (10uM) 4µl of 2.5mM dNTP mix 5µl of 10x reaction buffer 3ul of $MgCl_2$ 100ng template DNA	125ng each primer 1µl 10uM dNTP mix 5µl 10x reaction buffer (including 20mM $MgSO_4$) 2.5µl DMSO 50ng DNA template
Enzyme:		
	0.25ul Amplitaq Gold	Pfu (2 units)
Protocol		
	95°C, 5 minutes 95°C, 1 minute 58°C, 1 minute 72°C, 2 minutes per kb Back to 2 (35 cycles) 72°C for 10 minutes	95°C, 1minute 30 seconds 95°C, 50 seconds 60°C, 50 seconds 68°C, 2 minutes per kb Back to 2 (18 cycles)

2.4.6. PCR product purification

PCR products were purified using the QIAquick PCR purification kit (Qiagen), according to the manufacturer's instructions. Five volumes of PB buffer were added to the PCR sample. The sample was applied to the QIAquick spin-column and micro-centrifuged (13,000rpm, 1min). The DNA was retained in the column. The column was then washed with 750µl of PE buffer and micro-centrifuged. Any residual buffer was removed by a second spin. DNA was eluted in 30µl of EB buffer by micro-centrifugation.

Table 2-4 Cloning primers.

Construct	PCR primer
pcDNA3.1 HA-/ pGEX-2TK -HP1γ	F_EcoR1:GGAGAATTCATGGGAAAGAAACAAAATGGAAA R_BamH1:GGAGGATCCTTATTGTGCTTCATCTTCAGGAC
pGEX-2TK -HP1β	F_EcoR1:GGAGAATTCATGGGGAAAAAGCAAAACAAGAAG R_BamH1: GGAGGATCCCTAATTCTTGTCGTCTTTTTTGTC
pSB7-Halo-mEos3.2-HP1β	F_Xho1: GGACTCGAGATGGGGAAAAAGCAAAACAAGAAG R_Xba1: GGATCTAGACTAATTCTTGTCGTCTTTTTTGTC
pEF-Halo-mEos3.2- HP1γ	F_Xho1: GGACTCGAGATGGGAAAGAAACAAAATGGAAA R_Xba1: GGATCTAGATTATTGTGCTTCATCTTCAGGAC

2.4.7. Site-directed mutagenesis

Mutagenesis was carried out using the Quickchange Site-directed Mutagenesis kit (Stratagene) according to the manufacturers instructions. The parental vector was PCR extended with two complementary mutagenic primers using PFU DNA polymerase. See Table 2-3 for reaction mix and protocol and Table 2-5 for mutagenesis primers used in this thesis. The PCR product was then treated with 10 units of the restriction endonuclease DpnI (specific for methylated and hemi-methylated DNA) and incubated for 3 hour at 37ºC. This digests the parental vector strand (dam methylated from *E. coli*) and selects for mutated synthesised DNA. 5µl of the DpnI digested PCR product was transformed into DH5α *E. coli*. After plating, single colonies were grown overnight at 37ºC. The mini prep kit (see Section 2.3.2) used to isolate DNA and subsequently sequenced.

Table 2-5 Mutagenesis PCR primers.

Construct generated	Primer sequence
msHP1γ R38/9A	Fwd: AAAGTACTGGACGCGGCGGTAGTGAATGGGAAG
	Rev: CTTCCCATTCACTACCGCCGCGTCCAGTACTTT
msHP1γ R38/9K	Fwd: AAAGTACTGGACAAGAAGGTAGTGAATGGGAAG
	Rev: CTTCCCATTCACTACCTTCTTGTCCAGTACTTT
msHP1γ R39A	Fwd: AAAGTACTGGACCGTGCTGTAGTGAATGGGAAG
	Rev: CTTCCCATTCACTACAGCACGGTCCAGTACTTT
msHP1γ R38A	Fwd: AAAGTACTGGACGCTCGTGTAGTGAATGGGAAG
	Rev: CTTCCCATTCACTACACGAGCGTCCAGTACTTT
msHP1γ V32M	Fwd: CCTGAAGAATTTGTGATGGAAAAAGTACTG
	Rev: CAGTACTTTTTCCATCACAAATTCTTCAGG
msHP1γ I165K	Fwd: TGTCCTCAGATTGTCAAGGCCTTCTACGAGGAG
	Rev: CTCCTCGTAGAAGGCCTTGACAATCTGAGGACA
msHP1β I161M	Fwd: TGCCCACAGGTTGTCATGCCCTTCTATGAGGAA
	Rev: TTCCTCATAGAAGGGCATGACAACCTGTGGGCA

2.4.8. Oligonucleotides and DNA sequencing

All oligonucleotides were synthesised by Sigma Genosys and IDT. The automated DNA sequencing service at the Department of Biochemsitry performed all sequencing.

2.5. RNA purification and mRNA quantification

2.5.1. RNA purification

RNA isolation was performed using the RNeasy Plus RNA isolation kit (Qiagen) according to manufacturer's instructions. With this method all RNA molecules longer than 200bp are purified, therefore RNA is enriched for mRNA molecules. Up to 10^6 cells were pelleted and lysed in 350µl RLT Plus lysis buffer supplemented with β-mercaptoethanol at 0.143M concentration. To remove genomic DNA, the homogenised lysate was transferred to a gDNA eliminator spin column and centrifuged for 30 s at 10000rpm. One volume of 70% ethanol was added to the flow-through and mixed by pipetting. The mix was then transferred to a RNeasy spin column to bind total RNA

and centrifuged for 15 s at 10000rpm. Flow-through was discarded, column was washed with 700µl of RW1 wash buffer and centrifuged for 15s at 10000rpm. Column was washed another 2 times with 500µl RPE wash buffer and centrifuged for 15s and 2min respectively. Column was then placed in a fresh Eppendorf tube and RNA was eluted in 50µl nuclease-free H_2O. RNA concentration was measured using a Nanodrop spectrophotometer at 260nm.

2.5.2. First strand cDNA synthesis

The Superscript III Reverse Transcriptase kit (Invitrogen) was used according to manufacturers instructions to synthesise first strand cDNA. RNA purified as described (Section 2.5.1) was mixed with 1µl of DNase I for 30 minutes at room temperature. The enzyme was inactivated by heating for 10 minutes at 37°C. 500ng of RNA was mixed with 1µl hexanucleotide primers (100ng/µl), 1µl of oligo (dT) primers (50µM), 1µl dNTP (10mM) and 12µl H_2O. Samples were heated for 5 minutes at 65°C then mixed with 4µl 5xfirst strand buffer, 1µl 0.1M DTT, 1µl of RNAse out (a recombinant ribonuclease inhibitor at 40U/µl) and 1µl Superscript III RT for 5 minutes at 25°C. cDNA was synthesised by incubating samples at 50°C for 30-60 minutes. Reactions were inactivated for 15 minutes at 70°C. RNA complimentary to cDNA was removed by the addition of 1µl RNase H for 15 minutes at 37°C.

2.5.3. Quantitative real-time RT-PCR

cDNA purified as described (Section 2.5.2) was used to quantify mRNA levels. Alternatively genomic DNA isolated from cell lines, primary human samples or chromatin immunoprecipitation analysis was used to quantify DNA. Quantitative real-time PCR was performed with the ABI Prism 7300 apparatus (Applied Biosystems) sequence detection system using SYBRGreen as the detector. A 10µl reaction was used on 96 well plates (SYBRGreen master-mix 5µl, forward primer (200 µM) 0.05µl, reverse primer (200 µM) 0.05 µl, cDNA 1 µl, RNase-free water 3.9 µl.) A standard curve was included for each gene. PCR amplification was performed with an initial step of 10 minutes at 95°C, followed by 40 cycles of 15 seconds at 95°C and 1 minute at 55-60°C. Primes are summarised in Table 2-6.

Table 2-6 RT-PCR primers. (next page)

Gene	Primer sequence
hsHP1γ exon 3-4	Fwd:TTCCTGAAGTGGAAGGGATTT
	Rev: TTTTCTTTGCCAGCTTTCTGA
hsHP1γ exon 4-5	Fwd: GCTGGCAAAGAAAAAGATGG
	Rev: TCTGGCAAATCCTCTTGGTT
hsB2M	Fwd: GAGTATGCCTGCCGTGTG
	Rev: AATCCAAATGCGGCATCT
hsTUBB exon	Fwd:GTCCTGGATGTGGTACGGAA
	Rev: GCTGATAAGGAGAGTGCCCA
hsTUBB exon	Fwd: GGTGGCAAATATGTTCCTCGT
	Rev:CCAAAAGGACCTGAGCGAAC
hsZHX1	Fwd: TGTTGGTACTGCACAGCCTA
	Rev: CTGAAGGACACGAAGCTGC
hsZNF229	Fwd: GCCTCAGCCATTTCCCAAG
	Rev: GCTCCTCCTCAGTGAAGACC
hsZNFX1	Fwd:CTGGCTATTATTCAAGGACCT
	Rev: GGGGAACTTCTGGAGGCTAA
hsZNF778	Fwd: CCAGCACTTCGGCAGGATA
	Rev: GAAAGCTTCTCCACACTGCG
hsZBTB43	Fwd: TCCCAAAGCCCAAGAACTGA
	Rev: TCATGCTCTGTGGACGAGTT
hsZMAT1	Fwd: GCCAGAGGACTAGAGCCAA
	Rev:TCTGGAAAGTCTCACATGGGA
hsZNF124	Fwd: GGAATGCGGAAAGAAGCCAT
	Rev: ACTGTGTCTCTGTGAACCCT
hsPIM1	Fwd: CGCGACATCAAGGACGAAAA
	Rev: GAGGGCTATACACTCGGGTC
hsIFIT3	Fwd: CAAACAGATGTCCTCCGCAG
	Rev: CCTTGTAGCAGCACCCAATC
hsIFNL2/3	Fwd: CTGGGACCTGAGGCAGCT
	Rev: TATGGTGCAGGGTGTGAAGG
hsIFNB1	Fwd: TGCTCTCCTGTTGTGCTTCT
	Rev: AAGCCTCCCATTCAATTGCC
hsSTAT1	Fwd: TGCCACCATCCGTTTTCATG
	Rev: GACATCTGGATTGGGTCTTCC
hsCXCL8	Fwd: CAGTTTTGCCAAGGAGTGCT
	Rev: TGGTCCACTCTCAATCACTCT
hsXAF1	Fwd: CTGCCAACTTCACCCTCCAT
	Rev: ACACATCGTACACCCAACCT
hsBIK	Fwd: GAGATGGACGTGAGCCTCAG
	Rev: TGATGTCCTCAGTCTGGTCG
hsBID	Fwd: CGCTTGGGAAGAATAGAGGC
	Rev: GGAGGGATGCTACGGTCC
hsTNF	Fwd: GTCAACCTCCTCTCTGCCAT
	Rev: CTGAGTCGGTCACCCTTCTC

hsDDX58	Fwd:CGTCAGTGATAGCAACAGAA Rev: TCTCTGTGTCCCTCATCAGC

2.6. Protein purification, detection and analysis

2.6.1. SDS polyacrylamide gel electrophoresis (PAGE)

Proteins were separated in 10-20% acrylamide gels according to size, by the standard Laemmli procedure of one-dimensional gel electrophoresis under denaturing conditions. The mini-protean II gel apparatus (BioRad) was assembled according to the manufacturers instructions with 0.75mm spacers and combs. To prepare a 10% separating gel, 2.5ml 30% acrylamide (37.5:1 acrylamide:bisacrylamide), 2.8ml 1M Tris pH8.8, 37.5ml 20% SDS and 2.18ml ddH_2O were mixed. For preparation of gels with lower or higher acrylamide concentration the volumes of acrylamide and ddH_2O used were adjusted accordingly. 20µl 20% APS and 10µl TEMED were added to initiate polymerisation of the acrylamide. The gel solution was applied to the sandwiched glass until the height of the solution was 2cm from the top of the small plate. ddH_2O-saturated butanol was then overlaid and the gel left to polymerise. After polymerisation, the overlaid butanol was removed. The stacking gel was prepared by mixing 0.85ml 30% acrylamide (37.5:1 acrylamide:bisacrylamide), 0.625ml 1M Tris pH6.8, 25ml 20% SDS, 3.5ml ddH_2O, 20µl 20% APS and 5µl TEMED. The solution was applied into the glass sandwich, a comb inserted into the stacking solution and the gel was left to polymerise. Protein samples were denatured by boiling in 1xSDS loading buffer (200mM Tris-HCl pH6.8, 20% (v/v) β-mercaptoethanol, 2% (v/v) SDS, 0.1% (w/v) bromophenol blue, 40% (v/v) glycerol) for 5 minutes, and then loaded onto the gel. The gel was electrophoresed in Tris-glycine running buffer (25mM Tris-HCl, 250mM glycine, 0.1% (v/v) SDS) at 150-200V. Pre-stained high molecular weight protein markers were run alongside the samples (Bio-Rad).

2.6.2. Coomassie-blue staining of protein gels

Following electrophoresis, proteins were visualised by staining the gel with a staining solution (0.5% (w/v) Coomassie Brilliant Blue 40%, (v/v) methanol, 10% (v/v) acetic acid). Gels were incubated at room temperature for 60 minutes with gentle agitation. The staining solution was

then replaced with de-stain solution (20% (v/v) methanol, 10% (v/v) acetic acid) and shaken gently until blue bands containing protein were visible against a clear background within the gel.

2.6.3. Preparation of mammalian cell samples for Western blotting

Approximately $1x10^6$ cells washed twice in 1xPBS buffer to remove residual culture medium components and pelleted. Pellets were lysed by the addition of 500µl 1x SDS loading buffer. The lysed cells were transferred to an Eppendorf tube and sonicated at the high setting (Utrasound waterbath) for 5 minutes with 30 seconds off and 30 seconds on. The samples were then boiled for 5 minutes. 10-30µl of the whole cell extracts (lysate from ~ 20000-60000 cells) were analysed by SDS-PAGE and immunoblotting.

2.6.4. Western blotting

Proteins were transferred from polyacrylamide gels to nitrocellulose membranes using the mini-trans blot cell (BioRad). Following electrophoresis, gels were transferred into transfer buffer. Tris-glycine transfer buffer (192mM glycine, 25mM Tris-HCl) was used for proteins > 50kDa and for proteins <50 kDa phosphate transfer buffer (50x: 500mL of 1M Na2HPO4, 425ml 1M NaH2PO4) was used. The gel was sandwiched next to nitrocellulose (Amersham Protram), between 2 pieces of Whatman 3MM paper and between 2 sponge pads. The sandwich was inserted into the transfer chamber filled with 1x phosphate transfer buffer such that the nitrocellulose was facing the positive anode. Proteins were transferred to the nitrocellulose by applying a constant current (0.4A) for 1.5h, using an ice-block to keep the buffer cool. After the transfer, the nitrocellulose membrane was stained with ponceau S to ensure equal transfer and then placed either in milk blocking buffer (TBS with 0.5% (v/v) Tween-20, 5% (w/v) non-fat milk powder) or BSA blocking buffer (TBS with 0.5% (v/v) Tween-20, 5% BSA), for a minimum of 1 hour at room temperature. The filter was then sealed in a bag containing primary antibody diluted in blocking buffer, and incubated at room temperature on a rocking table for 2 hours or overnight at 4°C. The filter was washed three times for 5min in TBS with 0.1% (v/v) Tween-20 and gentle agitation. The filter was then incubated with secondary antibody (conjugated to horse-radish peroxidase) diluted in blocking buffer. This was incubated at room

temperature for 1 hour and the filter washed twice for 20 minutes in TBS with 0.1% (v/v) Tween-20. Detection of protein bands was carried out using the ECL system (Promega) according to the manufacturers instructions. ECL reagents were mixed 1:1 and incubated with nitrocellulose for required time. After incubation, excess fluid was drained off and the filter was exposed to X-ray film (Fuji) or analysed using a Chemidoc™ imaging system (BioRad).

2.6.5. Dot blot assays

Dot blot assays were performed by spotting synthetic peptide onto pre-wet PVDF membranes. Membranes were blocked in BSA blocking buffer for 1 h at RT. Blocked membranes were then probed with primary antibodies at 4°C o/n. Membranes were washed 3 times for 30s in TBS with 0.1% (v/v) Tween-20 followed by another 10m wash in the same buffer to remove residual antibodies. Membranes were then incubated in appropriate secondary antibodies and developed as described above (Section 2.6.4).

2.6.6. Antibody purification from crude sera

Principle of antibody purification is illustrated in Figure 3-12. Sulfhydryl-groups of peptides (see Table 2-8) were reduced using Bond-Breaker™ TCEP Solution (Thermofisher Scientific). Reactions were all performed at RT. 200µg of peptide were mixed with coupling buffer (50mM Tris HCl pH 8.5, 5mM EDTA in H_2O) and 25mM TCEP to a final volume of 2ml and incubated at RT for 30min. To couple reduced peptides to SulfoLink™ Coupling Resin (Thermofisher Scientific), 200µl of pure resin (400µl slurry) was washed in coupling buffer at RT and then incubated with 2ml of reduced peptide solution. The mix was incubated at RT on a rotating wheel for 15min and then for another 30min, this time standing upright without rotation. Resin was then spun down gently at 800rpm for 2min ad washed 3x in coupling buffer. To block unspecific binding, resin was incubated in quenching reagent (50mM L-cysteine in coupling buffer) at RT on a rotating wheel for 15min and then for another 30min, this time standing upright without rotation. Resin was then washed 3x in wash buffer (1M NaCl in H_2O) and 2x in storage buffer (0.05%(w/v) NaN_3 in 1xTBS). Resin could now be stored at 4°C until usage. Before usage resin was washed 3x in 1xTBS. 5ml of crude sera were spun down at full speed for 10min and

supernatant was moved into fresh 15ml falcon. Resin was then added and the coupling reaction was incubated at RT on a rotating wheel for 1h. Resin was then washed 4x in 1xTBS. To elute the antibodies off the resin 200µl of elution buffer (0.2M glycine HCl pH 2.5) were added and tube was quickly flicked and directly microcentrifuged at 800rpm for 1min. Supernatant was moved to a fresh tube and placed on ice. Immediately 10µl of ice-cold neutralisation buffer (1M Tris HCl ph 8.8) were added and mixed thoroughly to neutralise the pH to 7. Elution step was repeated 2-3 times. Beads were restored by washing 6x in binding/wash buffer and stored in storage buffer at 4°C.

2.6.7. Antibodies

The following antibodies were used in the listed concentrations for western blot, immunoprecipitation and immunofluorescence assays (Table 2-7).

Table 2-7 Primary antibodies.

Antibody	Description	Concentration		
		WB	IP	ChIP
HP1β	Rabbit polyclonal, Abcam	1:1000		
HP1γ	Rabbit polyclonal, Abcam ab1080	1:2000		
HP1γ	Mouse monoclonal, Millipore clone42s2	1:1000	5µg	5µg
GAPDH	Rabbit polyclonal Abcam ab9485	1:1000		
β-Tubulin	Rabbit Polyclonal, Abcam ab6046	1:1000		
H3	Rabbit polyclonal, Abcam; ab1791	1:1000		
H3S10ph	Rabbit Polyclonal, Abcam ab5176	1:1000		
H3K4me3	Rabbit Polyclonal, Abcam ab8580			2µg
H3K36me3	Rabbit Polyclonal, Abcam ab9050			2µg
H3R2-Cit	Rabbit polyclonal, Abcam ab176843	1:1000		
Citrulline	Rabbit Millipore clone 07-377	1:1000		
HA	Rabbit Polyclonal, Abcam ab9110	1:1000		
GAPDH	Rabbit Polyclonal, Abcam ab8245	1:1000		
Cleaved PARP1	Rabbit polyclonal, Abcam ab3251	1:5000		
GFP	Rabbit polyclonal, Abcam ab290			5µg

Table 2-8 Secondary antibodies.

Antibody	Description	Concentration
Rabbit IgG H&L	Goat anti-rabbit HRP conjugated, Abcam ab6721	1:20000
Mouse IgG H&L	Rabbit anti-mouse HRP conjugated, Abcam ab6728	1:5000
Sheep IgG H&L	Rabbit anti-sheep HRP conjugated, Abcam ab6747	1:10000

2.6.8. Immunoprecipitation from mammalian cell extracts (IP)

Approximately $10x10^6$ cells were washed twice in 1xPBS. Cells were pelleted and resuspended in ice-cold IPH lysis buffer (see Table 2-1) or ice RIPA lysis buffer with protease inhibitor cocktail (Complete^{TM}, Roche) and incubated on ice for 20 minutes and vortexed every 5min. IPH lysis buffer was used for co-immunoprecipitation assays and RIPA lysis buffer for standard immunoprecipitation assays. Extracts were cleared of debris by centrifugation (15 minutes, 13000rpm, 4°C). Extracts were then incubated with 3-5µg of antibody, on a rotating wheel at 4°C for a minimum of 2 hours or o/n. 20µl protein A/protein G sepharose beads (1:1 mix) were added and the incubation continued for 4h. The sepharose beads were pelleted in a cooled micro-centrifuge (4000rpm, 1 min, 4°C) and washed 3 times with ice-cold IPH. After the final centrifugation, the beads were resuspended in 1xSDS loading buffer and resolved by SDS-PAGE and immunoblotting.

2.6.9. Fractionation of mammalian cells

Approximately $5x10^6$ cells were washed twice in 1xPBS. Cells were pelleted and resuspended in ice-cold IPH lysis buffer (see Table 2-1) with protease inhibitor cocktail (Complete^{TM}, Roche) and incubated on ice for 20 minutes and vortexed every 5min. Extracts were centrifuged for 15min at 13000rpm and 4°C. The supernatant contains the nucleo- and cytoplasmic fraction and the insoluble pellet is composed primarily of chromatin and associated proteins.

2.6.10. Production of recombinant fusion proteins from *E. coli*

Protein coding sequences for each of the HP1β and γ domains were cloned into the relevant vector so that the region was in frame with the fusion. These constructs were transformed into the relevant *E. coli* as described (Section 2.2.1). Table 2-4 details the primers used to clone each of the HP1β and γ variants into pGEX-2T using the restriction enzyme sites for BamH1 and EcoRI. For GST fusion proteins, 10ml overnight cultures were set up in LB/antibiotic and then diluted 1:10 into 100ml LB/antibiotic the following morning. Cultures were grown at 37°C until OD_{600} reached 0.6-0.8 (log-phase). Protein expression was induced at log-phase by the addition of IPTG to 1mM final concentration and cultures grown for a further 4 hours at 37°C. The bacteria were pelleted by centrifugation and resuspended in ice-cold 5ml TBS containing 1% (v/v) Triton-X-100 and protease inhibitor cocktail. The cell suspension was sonicated with a probe sonicator for 3min at high settings with 30 seconds on and 30 seconds off (Bioruptor, Diagenode) and centrifuged (5000 rpm, 15 minutes, 4°C). The supernatant was poured into a 15ml Falcon tube containing 300µl of Glutathione linked sepharose beads slurry (Amersham), and rocked for 1 hour at 4°C. The beads were washed twice with 1xTBS, 1% (v/v) Triton-X-100 and twice with 1xTBS. The fusion proteins were either stored frozen on the beads in an equal volume 1xTBS with 20% (v/v) glycerol or eluted from the beads using 300µl glutathione solution (25mM reduced glutathione, 50mM Tris pH 8.0).

2.7. *In vitro* assays

2.7.1. *In vitro* binding assays

The GST-fusion proteins and biotin-conjugated peptides (Table 2-9) were incubated with glutathione agarose beads or streptavadin sepharose beads respectively in binding buffer (150mM NaCl, 50mM Tris-HCl pH 8.0, 5mM EDTA and 0.2% NP-40) for 1 hour at room temperature. After washing three times in binding buffer the beads were incubated with their potential binding proteins for 1 hour at room temperature. The beads were then washed 4 times with binding buffer, following which the bound protein was eluted with hot 2x SDS loading buffer and resolved by SDS-PAGE for western blot analysis or coomassie-blue staining.

2.7.2. *In vitro* citrullination assays

Recombinant human PADI4-GST was expressed from pGEX6p constructs in LB media, induced with 0.1mM IPTG at 25°C, purified using glutathione-sepharose resin, eluted using a 25mM glutathione solution. PADI4-GST was used fresh. *In vitro* deimination of HP1γ wildtype and HP1γ R38/9A mutant protein was carried out in deimination buffer [50mM HEPES, pH 7.5; 2mM DTT; in the presence of absence of 10mM $CaCl_2$] at 37°C for 1h, using the active enzyme. Samples were eluted with hot 2xSDS loading buffer and analysed by immunoblotting.

Table 2-9 Peptide sequences.

Name	Peptide sequence
H3 (1-16)	Biotin-ARTKQTARKSTGGKAPGGC
H3 (1-16) K9me3	Biotin-ARTKQTARK(me3)STGGKAPGGC
HP1γ (34-44)	KVLDRRVVNGKC
HP1γ (34-44) R38-Cit	KVLD(Cit)RVVNGKC
HP1γ (34-44) R39-Cit	KVLDR(Cit)VVNGKC
HP1γ (34-44) R38/9-Cit	KVLD(Cit)(Cit)VVNGKC
HP1γ (104-111)	SKKKRDAADKC
HP1γ (104-111) R108-Cit	SKKK(Cit)DAADKC
HP1α (103-112)	SKKKREQSNDC
HP1α (103-112) R107-Cit	SKKK(Cit)EQSNDC

2.8. Mammalian cell culture

2.8.1. Cell culture media, reagents and maintenance

All tissue culture manipulations were carried out in sterile conditions in a standard laminar flow hood. Human cell lines (see Table 2-10, row 1-4) were maintained in 10cm culture dishes in DMEM (Gibco), 10% FCS, 1% penicillin/streptomycin, 2mM L-glutamine. MCF10A were cultured in DMEM/F12, 5% horse serum, 20ng/ml human EGF, 0.5mg/ml hydrocortisone, 100ng/ml cholera toxin, 10μg/ml human insulin and 1% penicillin/streptomycin. All cell cultures were stored in incubators at 37°C in 5% CO_2 (100% humidity) and passaged between 1:3 and 1:5 every 2-3 days. Cells were passaged using conventional cell culture techniques.

Mouse embryonic stem (ES) cell lines were cultured in standard serum and mouse leukemia inhibitory factor (mLIF) conditions. DMEM (Gibco) supplemented with 15% FCS, 0.1mM non-essential amino acids (Gibco), 1% penicillin/streptomycin, 2mM L-glutamine (Gibco), 1mM sodium pyruvate (Gibco), 0.1mM beta-mercaptoethanol (Life tech) and 10^6 units/L mLIF (ESGRO, Millipore). E14 ES cells were stored in incubators at 37°C in 7.5% CO_2 (100% humidity) and maintained in 0.1% gelatine coated 10cm culture dishes with 1:10 passaging every 2 days. Media was changed on a daily basis.

Table 2-10 Cell lines.

Name	Cell type	Details
MCF7	Human mammary adenocarcinoma	ER+/PR+/Her2-, p53 wildtype, Luminal A type breast cancer cells
MDA-MB-231	Human mammary adenocarcinoma	p53 mutant chemotherapy resistant, metastatic breast cancer cells
HEK293T	Human embryonic kidney cells	Used for lentiviral packaging
MCF10A	Human mammary gland	Immortalised, non-transformed
mESC	Mouse embryonic stem cells	E14 cell clone

2.8.2. Transient transfection of MCF7 using FuGENE®

MCF7 cells were seeded into 6-well dishes 24h prior to transfection. Dishes contained approximately $4x10^5$ cells or were 80% confluent on the day of transfection. Media was replaced with 2ml of antibiotic-free DMEM (Gibco) media with 10% FCS. 6µl of FuGENE® was mixed with 72µl of Opti-MEM™ (Gibco). 2µg of DNA was mixed with 40µl of Opti-MEM™ (Gibco), then added to the mix and incubated for 20min at RT. The transfection mix was dropped onto the cells and 6-16 hours later 2ml of normal media were added.

2.8.3. Transient transfection of mESCs using Lipofectamine 2000®

5µl of Lipofectamine 2000® was mixed with 245µl of Opti-MEM™ (Gibco). 2µg of DNA was mixed with 250µl of Opti-MEM™ (Gibco), then added to the mix and incubated for 20min at RT. The transfection mix was mixed with $5x10^5$ mESCs in 1.5ml antibiotic-free ESC media. The

mix was seeded into 0.1% gelatin coated 6-well dishes. Media was replaced carefully with 2ml of antibiotic-free ESC media 4h post transfection.

2.8.4. Transient transfection with siRNAs using Dharmafect® (reverse)

2µl of Dharmafect® or DharmafectDuo® was mixed with 198µl of Opti-MEM™ (Gibco). siRNAs were used at a final concentration of 100nM.10µl of siRNAs (20µM stock, Dharmafect) were mixed with 190µl of Opti-MEM™ (Gibco), then added to the mix and incubated for 20min at RT. The transfection mix was mixed with $5x10^5$ MCF7 or MDA-MB-231 cells in 1.6ml antibiotic-free DMEM media with 10% (v/v) FCS. 12-16h later 2ml of normal media were added. siRNAs were ordered from Dharmacon. A list of siRNAs used in this study is summarised in Table 2-11.

Table 2-11 siRNA sequences.

Name	Sequence	Target
siHP1γ_1	AUCUGACAGUGAAUCUGAU	hsCBX3, CDS
siHP1γ_2	GAGGCAGAGCCUGAAGAAU	hsCBX3, CDS
siHP1γ_3	UCAGAAAGCUGGCAAAGAA	hsCBX3, CDS
siCtrl	UCUCGACCCUGAACGAAUAAU	Non-targeting control (Dharmacon)

2.8.5. Production of lentiviral particles

5x10^6 HEK293T cells were seeded onto L-polylysine (1µg/ml) coated 10cm dishes 24h prior to transfection in antibiotic-free DMEM media with 10% (v/v) FCS. Cells were 80% confluent on the day of transfection. For production of lentiviral particles, the 3rd generation packaging system was used. 50µl Lipofectamine 2000® was added to 450µl of Opti-MEM™ (Gibco). 4µg of pCMV-VSVG (Addgene#8454) encoding viral envelope proteins, 12µg of psPAX2 (Addgene#12260) encoding viral gag, pol, rev and tat proteins, 8µg of pLKO.1_shRNA vectors or lenti_CRISPR_V2 vectors were mixed with Opti-MEM™ (Gibco) to a final volume of 500µl. The DNA mix was added to the Lipofectamine 2000® mix and incubated for 20min at RT. The

transfection mix was dropped onto the cells and 12-14h later 6ml of normal media were added. Cells were incubated and handled under Cat.2 safety conditions. 24h after the media change cellular supernatant containing lentiviral particles was harvested and sterile filtered using 0.45µm syringe filters (Millipore). Aliquots of lentiviral supernatants were stored at -80°C. Titer of lentiviral particles was determined using MCF7 cells. 5x10^5 cells were seeded into 6-well dishes 24h prior to infection. Media was removed and replaced by 5 serial dilutions of viral supernatant (SN) in a total volume of 1ml starting from undiluted SN (1000, 100, 10, 1, 0.1 µl of viral supernatants). Cells were incubated at 37°C. As the pLKO.1 shRNA constructs contain a puromycin resistance gene cassette, puromycin was used for positive selection of transduced cells at 1µg/ml. Puromycin supplemented media was replaced every 3 days. After 8 days and the number of colonies was determined. Therefore, cells were washed once with 1xPBS and stained and fixed with a crystal violet solution (0.05% (w/v) crystal violet, 20% (v/v) EtOH in ddH_2O) for 1h in the dark at RT. Cells were then washed twice in ddH_2O. Plates were dried over night and colonies were counted. Titers of viral SN were calculated according to the following formula:

$$\text{Titer SN}\left[{}^{\text{TU}}/_{\text{ml}}\right] = \frac{2 \cdot (number\ cells\ seeded\ DAY1)}{\frac{1000\ \mu l}{Volume\ (viral\ SN)[\mu l]}} \cdot number\ colonies$$

2.8.6. Lentiviral infection

MCF7, MCF10A or MDA-MB-231 cells were seeded into 6-well dishes 24h prior to infection. Dishes contained approximately $4x10^5$ cells or were 80% confluent on the day of infection. pLKO.1 shRNA containing lentiviral particles were produced as described in Section 2.8.5. A list of shRNA sequences used in this study is summarised in Table 2-12. Lentiviral supernatants were added to cells at an estimated MOI of 10 together with 8µg/ml of transduction enhancer polybrene (Sigma Aldrich). 12h later 2ml of normal media were added. 24h post infection cells were cultured in puromycin selection media (MCF7: 1µg/ml, MDA-MB-231: 0.5µg/ml, MCF10A: 2µg/ml).

Table 2-12 shRNA and guideRNA (sgRNA) sequences.

Name	Sequence	Target
shHP1γ_1	GCGTTTCTTAACTCTCAGAAA	hsCBX3, CDS
shHP1γ_2	CTGGCGAAAGAGGCAAATATG	hsCBX3, CDS
shCtrl	CAACAAGATGAAGAGCACCAA	Non-targeting (Sigma)
sgCBX3 #79	CACCGGGCCTCCAACAAAACTACAT	hsCBX3 exon 2
sgCBX3 #81	CACCGAGAGCCTGAAGAATTTGTCG	hsCBX3 exon 3
sgCBX3 #83	CACCGAGAAACGCTTCAATCAATTC	hsCBX3 exon 4
sgCtrl	CACCGGCGAGGTATTCGGCTCCGCG	Non-targeting

2.8.7. Proliferation assay

Proliferation of human cell lines was assessed using the CyQUANT® proliferation assay (ThermoFisher Scientific) according to the manufacturer's instructions. At day 0, cells were seeded at a density of 500 cells per well into wells of a 96 well plate. Per condition and per time point 18 wells were seeded. Proliferation was measured at 5 time points (day 0 (=12h post seeding), 2, 4, 6, 8). Cell were prepared for analysis by removing media, washing the wells twice with 1xPBS and freezing the plates at -80°C until analysis. 200µl of 1x cell lysis buffer supplemented with 2x QUANT® GR DNA intercalating dye were added per well and incubated in the dark until measurement. Fluorescence was measure using a Clariostar® plate reader at 520nm. Values were blank corrected and both the three highest and lowest fluorescence values were regarded as outliers and not included into further data analysis. In order to correct for seeding variability, all fluorescence values were normalised to the average of all fluorescence values of day 0 of the respective condition and plotted using GraphPad Prism software (GraphPad Software Inc. (La Jolla, United States of America)).

2.8.8. Cell cycle analysis by flow cytometry

MCF7 and MDA-MB-231 cells were seeded into wells of 6 well plate 2 days after infection with shRNA vectors. At the day of analysis, cell at approximately 80% confluency were trypsinised and washed twice in 1xPBS. After the last wash cell pellet was resuspended in 100µl of 1xPBS. 200µl of staining solution were added (6% (v/v) NP-40, 50 µg/ml propidium iodide, 0.1 mg/ml

RNase in 1xPBS) and incubated at RT in the dark for 15min. The propidium iodide stained cells were then analysed by flow cytometry on a Becton Dickinson FACS CaliburTM system.

Flow cytometry analysis was performed through a custom R script. Briefly, raw FCS acquisition files were imported in R taking advantage of the package flowCore[177]. Acquired events were manually gated by the physical parameters Forward Scatter (FSC) and Side Scatter (SSC) in order to filter out and minimise the cellular debris, characterised by low intensity values. The G0/G1 and G2/M populations were fitted by two Gaussian distributions by first finding the local maxima of the FL2-A intensity distribution (Filter: 564-606nm) closest to the expected values of 200 and 400, respectively; then, relative weights and the standard deviations of each distribution were calculated from their FWHMs (Full widths at half maximum). The fraction of each cell cycle state was computed as the fraction of the area under the curve of the corresponding population (see Figure 4-6C: as an example plot of the analysis). Values of S population were derived by subtracting the fitted areas of G0/G1 and G2/M populations from the global intensity distribution.

2.8.9. Annexin V and propidium iodide staining

MCF7 and MDA-MB-231 cells were seeded into wells of 6 well plate 2 days after infection with shRNA vectors. At the day of analysis at approximately 80% confluency, detached cells were collected and attached cells trypsinsed. Detached and trypsinsed cells were recombined and washed twice in 1xPBS. To measure the proportion of viable and apoptotic cells AnnexinV-FITC Apoptosis staining/detection kit (Abcam ab14085) was used at manufacturer's instructions. Cells were then resuspended in 300 µl of the staining solution containing 5µl Annexin V-FITC detection reagent and 5µl of propidium iodide (PI) (final concentration: 50µg/ml) and incubated for 5min at RT in the dark. PI staining was used to distinguish between live and dead cells. Annexin V-FITC as a marker for apoptotic cells. Cells were analysed by flow cytometry on a Becton Dickinson FACS CaliburTM system. Properties of fluorophores are: FITC (ex = 488nm, em = 530nm), PI (ex = 488nm, em = 617nm).

2.8.10. Transwell migration assay

MDA-MB-231 or MCF7 cells were transfected with siRNAs (Section 2.8.4, Table 2-10) and incubated for 44h (MDA-MB-231) or 24h (MCF7) at 37°C, after which cells were prepared for the transwell migration assay. 2h before the assay the bottom side of 24-well Corning® BioCoat™ Control Inserts (8.0µm pores) were incubated in 1µg/ml human fibronectin solution and washed in 1xPBS right before cells were added. Cells were trypsinsed and washed twice in 1xPBS. To obtain a single cell solution, the cell solution was strained through a 40µm nylon filter mesh cell strainer. Cell numbers were determined by bromophenol blue staining using a Countess II FL cell counter (Thermofisher scientific). The migration chamber was assembled by placing the fibronectin coated transwell inserts into dishes of a 24-well plate containing 750µl of DMEM media supplemented with 15%(v/v) FBS. 50.000 cells in 500µl FBS free DMEM were then added drop-wise on top of the transwell. Migration assays was performed for for 4h (MDA-MB-231) or 24h (MCF7) at 37°C. Transwells were the washed carefully 2x in 1xPBS from both sides and cells in the upper side of the transwell were removed with cotton buds. Migrated cells on the bottom side of the transwell were stained and fixed in a crystal violet staining solution (5mg/ml crystal violet, 25%(v/v) EtOH in H_2O_{dest}) for 60min at RT in the dark. Cells were then washed 2x in $H_2O_{dest.}$ Cells were counted using a Brightfield microscope. At 20x magnification, cells within 15-20 fields of view were counted manually.

2.8.11. Wound healing ‘scratch assay’ using Incucyte®ZOOM system

MDA-MB-231 cells were transfected with siRNAs (Section 2.8.4, Table 2-10) and incubated for 24h at 37°C, after which cells were seeded into wells of a 96-well plates specialised for the IncuCyteZOOM® system (Essenbioscience) at a density of 30.000cells/well, 32 wells per condition. 12h later a wound of defined size was introduced into the cell monolayer using the Essenbioscience 2D WoundMaker™ tool. Cells were washed 2x in completed DMEM media to remove cell debris and were then placed into IncuCyteZOOM® incubator system. Cells were imaged every 45min for 48h. Images were then used to calculate the relative wound healing density (RWD in%) for each timepoint using the IncuCyteZOOM® software. The algorithm is based on measuring the spatial cell density in the wound area relative to the spatial cell density

outside of the wound area at every time point. It is designed to be zero at t=0, and 100% when the cell density inside the wound is the same as the cell density outside the initial wound, which was achieved for control cells after 48h of incubation.

2.9. Single molecule localisation microscopy (SMLM)

2.9.1. mES cell line generation

ES cells expressing mouse HP1 constructs tagged at the N- or C-terminus with the mEos3.2-HaloTag were generated by transfecting a plasmid expressing the tagged protein, followed by selection in 600µg/ml geneticin (Life tech, cat. 10131019). After 2 weeks of geneticin selection, cells were sorted using a Sony SH800 cell sorter to ensure that they were labeled with the mEos3.2 fluorophore (excitation at 488nm, emission at 515nm).

2.9.2. Microscope setup

An IX73 Olympus inverted microscope was used with circularly polarized laser beams aligned and focused at the back aperture of an Olympus 1.40 NA 100× oil objective (Universal Plan Super Apochromat, 100×, NA 1.40, UPLSAPO100XO/1.4). Continuous wavelength diode laser light sources used include a 561nm (Cobolt, Jive 200, 200mW) and a 405nm laser (Stradus, Toptica, 405-100, 100mW). Total internal reflection was achieved by aligning the laser off axis such that the emergent beam at the sample interface was near-collimated and incident at an angle greater than the critical angle $\theta_c \sim 67°$ for a glass/water interface for TIRF imaging and slightly less than θ_c for oblique-angle illumination imaging. This generated a ~50µm diameter excitation footprint. For TIRF, the power density at the coverslip for the 561 nm laser was in the range 0.5-100 W/cm^2. For oblique-angle illumination, the power of the collimated beams at the back aperture of the microscope was 10 kW/cm^2 and 10-100 W/cm^2 for the 561nm and 405 nm laser beams respectively. The lasers were reflected by adichroic mirrors, which also separated collected fluorescence emission from the TIR beam (Semrock, Di01- R405/488/561/635). The fluorescence emission was collected through the same objective and then further filtered using a combination of long-pass and band-pass filters (BLP01-561R and FF01-587/35 for 561 nm

excitation). The emission signal was projected onto an EMCCD (Photometrics, Evolve 512 Delta) with an electron multiplication gain of 250 ADU/photon operating in a frame transfer mode. The instrument was automated using the open-source software micro-manager (https://www.micro-manager.org) and the data displayed using the ImageJ software.

2.9.3. Mammalian live-cell single-molecule imaging

ES cells expressing mEos3.2-HaloTag-tagged HP1 were passaged two days before imaging onto 35 mm glass bottom dishes No 1.0 (MatTek Corporation P35G-1.0-14-C Case) in phenol red-free serum and LIF conditions. Just before single-particle tracking (SPT) imaging experiments, cells were labeled with 0.5nM HaloTag-JF_{549} ligand for at least 15 minutes, followed by two washes in PBS and a 30 minute incubation at 37 °C in media, before imaging the cells in fresh phenol red-free serum and LIF conditions. Cells were under-labeled to prevent overlap of fluorophores during SPT. Using 561 nm excitation, fluorescence images were collected as movies of 10,000 frames at 15ms or 1,000 frames at 500ms exposure. Mammalian cell image processing and analysis

Live-cell single-molecule traces were analyzed using software that detects single-molecule trajectories from SPT movies[178]. Briefly, the code detected puncta positions in each image frame using a brightness-weighted centroid after applying a band-pass filter to remove low-frequency background and high-frequency noise from the image data. To track the HP1 single-molecules, only fluorescent puncta smaller than 3 pixels and with a signal-to-noise greater than 8 were analyzed. Fluorescent puncta were considered to be the same molecule if they were within 6 pixels (936 nm) between frames.

2.10. Chromatin immunoprecipitation

2.10.1. ChIP and RT-PCR analysis

2x10^7 MCF7 cells (~ confluent 15cm dishes) were cross-linked on the dish with 1% [v/v] formaldehyde for 10 minutes at room temperature and cross-linking stopped by the addition of 0.125M glycine for 5min. To pre-extract the chromatin fraction, cells were then lysed in ice-cold

IPH chromatin extraction buffer supplemented with EDTA-free protease inhibitors (Roche) and incubated for 20min on ice intermitted by vortexing every 5min. Lysates were micro centrifuged at full speed for 20min. Supernatant was discarded and chromatin pellet was then lysed in ChIP lysis buffer (1% [v/v] SDS, 10mM EDTA, 50mM Tris-HCL pH8.0 and EDTA-free protease inhibitors). Cells were sonicated in a QSonica sonicator for 1.5h 5s ON/ 5s OFF to achieve a mean DNA fragment size of 500bp. An equal volume of protein A and G agarose beads (GE Healthcare), equilibrated in modified RIPA buffer (1% [v/v] Triton-X-100, 0.1% deoxycholate, 0.1% SDS, 90mM NaCl, 10mM Tris-HCL pH8.0 and EDTA-free protease inhibitors) and was used to pre-clear chromatin for 1h prior to immunoprecipitation. After pre-clearing 2% of chromatin was stored as input material. Immunoprecipitation was performed for a minimum of 12 hours at 4°C in modified RIPA buffer. An equal volume of protein A and G agarose beads, equilibrated in modified RIPA buffer were used to bind the antibody and associated chromatin for a minimum of 2 hours. The beads were then washed twice with wash buffer (0.1% [v/v] SDS, 1% [v/v] Triton-X-100, 2mM EDTA, 20mM Tris pH8.0, 150mM NaCl) and once with final wash buffer (0.1% [v/v] SDS, 1% [v/v] Triton-X-100, 2mM EDTA, 20mM Tris pH8.0, 500mM NaCl). Beads and input samples were then mixed with 200µl of elution buffer (1% [v/v] SDS, 100mM NaHCO3) and 5µl RNase (DNase free [Roche]) for 1h at 37°C. To reverse the cross link, the same mix was incubated at 65°C for at least 4h or o/n. Reverse crosslinking of DNA was followed DNA purification using the ChIP DNA Clean & Concentrator™ kit (Zymo Research) according to manufacturer's instructions. DNA was eluted in 100µl nuclease free H_2O when analysed by RT-PCR and 40µl when analysed by ChIP sequencing. Immunoprecipitated DNA was analysed on an ABI 7300 real-time PCR machine (as described in Section 2.5.3). Primers used for RT-PCR analysis are summarised in Table 2-13.

2.10.2. ChIP-sequencing library preparation

ChIP libraries used for parallel sequencing were prepared using the NEXTflex Illumina ChIP-Seq library prep kit (Bioo Scientific) designed for ultra-low DNA input Illumina® sequencing. Following chromatin immunoprecipitation and DNA purification, DNA was quantified using Qubit fluorimetric quantification technology.10-50ng of DNA (including 1% of total input) was

used to perform end repair of the ChIP DNA. Libraries were generated exactly according to manufacturer's instrusctions. Size selection of library samples was performed with AMPure® XP beads (Agencourt) as dictated in the NEXTflex Illumina ChIP-Seq library prep kit protocol. 2µl of NEXTflex™ ChIP-Seq Barcodes (#1-24, #514122) were used per library, they were ligated onto individual libraries in the adapter ligation step of the protocol. PCR conditions for the final library amplifications are summarised in Table 2-14. The completed library was then tested for both size selection and quantity of DNA using a D1000 DNA tape analysed on the Tapestation® system. DNA was additionally quantified by Qubit fluorimetric quantification technology. Libraries were multiplexed at equal molarities. The final ChIP-seq library, with an average size of 300-450bp (at 10-15nM), was submitted for 50bp, paired-end sequencing on an Illumina HiSeq 1500®.

Table 2-13 ChIP pPCR primers. All primers are designed against human genes.

Target gene	Primer sequence
ZNF37A 3'UTR	Fwd: TGAGAAGACACACAGGGGAG
	Rev: TTAGGGCTGACTTCTGACGG
ZNF221 3'UTR	Fwd: TGAGAAGACACACAGGGGAG
	Rev: TTAGGGCTGACTTCTGACGG
ZNF594 3'UTR	Fwd: TTCTCCGACGTTGAATAAGG
	Rev: CACCTTGTCACACACCAGAA
ZNF709 3'UTR	Fwd: CTCCAGGAATGTCTGGTCAC
	Rev: CTTGATGCTGGGTCTCATCT
ZNF788 3'UTR	Fwd: GCACTCGGGACAGATCTCTT
	Rev: CACGGCTTCCATTTCCATGT
ZNF75A 3'UTR	Fwd: CATCAGCAGGCGTCATATCA
	Rev: GGTGCCATTTTCCCACTCTG
ZNF174 3'UTR	Fwd: CAGCTCCCCAAATGCTCAAA
	Rev: TCTCTCAGTGGGTGGCTTTT
ZSCAN32 3'UTR	Fwd: AACCTGGATGTTGTGCTTGC
	Rev: CTTTGTAACCCAGCCACCAC
α-SAT 3'UTR	Fwd: ATCGAATGGAAATGAAAGGAGTCA
	Rev: GACCATTGGATGATTGCAGTCA
ZNF3 3'UTR	Fwd: TGTGGATTCTCTGGTGCTGA
	Rev: TACGCCTGCAATGAATGTGG
ZNFX1 TSS	Fwd: CGGCCTGGACAACTACTAGA
	Rev: CGAAACGCTCTCTTTCCCG
ZHX1 TSS	Fwd: CCTTGACGTTCACTCCCCA
	Rev: GGATTAATGTGGCTCAGCGG
ZBTB43 TSS	Fwd: GCAACCACGCTCAGCTAATT

	Rwd: GAGGCTAGGAGTTCGAGACC
ZNF229 TSS	Fwd: AGAAGCAATCAAGGTGACGC
	Rwd: GGAGGGTTCTGGGAAAGGTA
ZMAT1 TSS	Fwd: TGCGATCTCACTCACTGCTA
	Rev: TGCCTGTAATCCCAGCTACT
ZNF124 TSS	Fwd:ACTGTGTCTCTGTGAACCCT
	Rev:GGAATGCGGAAAGAAGCCAT

Table 2-14 PCR program for final ChIP-seq library amplification.

98°C, 2 min	
98°C, 30 sec	Repeat 11 cycles*
65°C, 30 sec	
72°C, 1 min	
72°C for 4 min	

2.10.3. Analysis of ChIP-Seq data

After adaptor trimming with trim galore, unique reads were mapped against the reference genome sequence (GRCh38) using bwa (v0.7.12)[179]. Biological replicates were then combined and used for peak calling with MACS2[180] with a p-value cutoff of 0.01 (peak calling against Input sample). The resulting peaks were then filtered to only retain those with a q-value less than 0.05 and all peaks overlapping an IgG peak were discarded. We then used bedtools to identify Ensembl genes with a peaks in the gene body or in the promoter (defined as TSS+/-2kb).

To produce metagene coverage profile and heatmaps the bwa BAM files were first converted to bedGraph format using bedtools[181] (genomeCoverageBed, v2.26.0, scaling each sample proportionally to the mapped library depth) and subsequently to bigWig using the UCSC Genome Browser tool bedGraphToBigWig[182]. The bigWig files were then used to produce signal coverage matrices using deepTools computeMatrix (v2.5.1)[183] using as reference transcripts all Ensembl transcripts overlapping a peak and scaled to a uniform size of 5kb (+2kb upstream of the TSS and downstream of the TES). Heatmaps and profile plots were rendered in R using ggplot2. Open access to ChIP-data presented in this thesis is currently being prepared.

2.11. RNA-sequencing

2.11.1. Depletion of ribosomal RNA

Whole transcriptome analysis was performed using RNA-sequencing (RNA-seq) technology. As input material for RNA-seq analysis total RNA depleted from ribosomal RNA (rRNA) was used. Total RNA was extracted (see Section 2.5.1). rRNA using the Ribo-Zero® rRNA Removal Kit (Illumina) according to manufacturer's instructions with the following adjustments: Due to a generally low efficiency of this kit, 500ng of total RNA was used at conditions suggested for 2.5-5µg of input RNA (10µl of Removal solution per sample). The RNA Clean & Concentartor™-5 kit (Zymo Research) was used to clean up the rRNA depleted RNA sample. RNA was eluted in 40µl nuclease free H2O. Efficiency of rRNA removal was confirmed RNA ScreenTape assay (Agilent) analysed on an Agilent 2200 Tapestation® system. In total RNA, peaks for small (18S) and large (28S) ribosomal subunits were detected, reflected in a RNA integrity number (RIN) >8. RIN is a measure of RNA integrity ranging from 1-10, with 10 describing the most intact RNA profile. The RIN algorithm is based on the ratio of 18S and 28S RNA fragments detected in a given RNA sample[184]. Accordingly, upon ribo-depletion of the same sample, the RIN was reduced to <3. 1-10ng of RNA were used for RNA-seq library preparation.

2.11.2. RNA-seq library preparation

10 ng of ribo-depleted RNA was subjected to RNA seq library preparation using the NEXTflex® Rapid Illumina Directional RNA-Seq Library Prep Kit (Bioo Scientific) according to manufacturer's instructions. This kit uses the deoxy-UTP (dUTP) strand-marking methodology in order to preserve the information on a transcript's directionality[185]. Figure 2-1 depicts a schematic illustration of the library preparation protocol using this method. In brief, cDNA first strand synthesis is performed using random hexamers. For second strand synthesis, dUTP is used instead of dTTP. In the next steps, adaptors (containing individual barcodes) are ligated to both ends of the double stranded (ds) cDNA. The adaptor sequences have a Y-shaped structure, which retain directionality of the cDNA molecules. Next Uracil-DNA glycosylase (UDG) is added to degrade the second strand. All remaining first strand cDNAs, ligated to Y-shaped adapter sequences, allow directional sequencing of all molecules while retaining the strand information of

each RNA molecule[185]. NEXTflex® RNA-Seq Barcodes (#1-24 NOVA-512913) were ligated onto individual libraries during adapter ligation step. PCR conditions for final library amplification are summarised in Table 2-14.

Table 2-14 PCR program for final RNA-seq library amplification.

37°C, 30min	
98°C, 2 min	
98°C, 30 sec 65°C, 30 sec 72°C, 1 min	Repeat 15 cycles*
72°C for 4 min	

The average size of final libraries was measured using a DNA ScreenTape (Agilent) assay analysed on an Agilent 2200 Tapestation® system and DNA concentrations were determined using the Qbit fluorimetric quantitation system®. Libraries were multiplexed at equal molarities. The final multiplexed library, with an average size of 300-450bp (at 15-20nM) was submitted for 125bp, paired-end sequencing on an Illumina HiSeq 4000® instrument.

2.11.3. Analysis of RNA-seq data

Fastq files from technical replicates were merged and then mapped against the reference genome sequence (GRCh38) using STAR (v2.5.0)[186] with indexes built with the Ensembl[187] reference transcriptome GTF file. The number of reads mapping to each Ensembl annotated gene were counted with STAR and then imported into DESeq2[188] for differential expression analysis. The DESeq2 workflow was run with default parameters using a one-factor design representing the treatment condition. The resulting p-values were corrected using the Benjamini-Hochberg method as implemented in DESeq2.

GO and KEGG enrichment analysis were realised using the BioConductor GAGE package[189]. The input gene list to GAGE was ordered by log2 fold change in the comparison of interest, and Entrez IDs for each gene were retrieved from the corresponding ENSEMBL IDs using biomaRt. Open access to RNA-data presented in this thesis is currently being prepared.

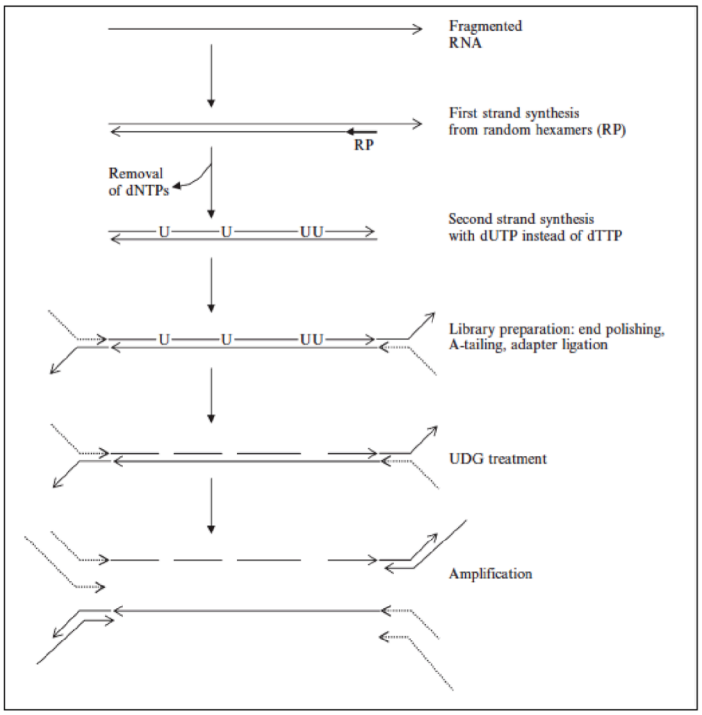


Figure 2-1 Deoxy-UTP (dUTP) strand-marking methodology for RNA-seq library preparation. Image was adapted from reference 165. For description, see main text.

Chapter 3

Citrullination – a novel post-translational modification of HP1γ

3.1. Introduction

Citrullination describes the conversion of peptidyl-arginine to the non-coding amino acid peptidyl-citrulline. This modification results in loss of a positive charge that can change the overall charge of a protein and affect the ability of hydrogen bonding with neighbouring proteins or nucleic acids[80]. Reactions are catalysed by calcium-dependent enzymes called peptidylarginine deiminases (PADIs). PADI1, 2 and 4 have been reported to be nuclear and regulate gene expression by directly citrullinating histones[48,85,86] (compare Section 1.1.8).

3.1.1. A role for citrullination in pluripotency

Work from our laboratory recently identified a role for PADI4 during reprogramming and ground state pluripotency[190]. It showed that expression and enzymatic activity of PADI4 are induced during reprogramming and that knockdown of the protein reduces reprogramming efficiency. Active PADI4 binds to regulatory elements of key stem cell genes and activates their transcription. To understand the molecular mechanism of gene regulation by PADI4, stable isotope labeling of amino acids in cell culture (SILAC) was used to find PADI4 substrates in mouse embryonic stem cells (mESCs). The experimental setup is illustrated in Figure 3-1. In brief, mESCs overexpressing PADI4 were labeled with $^{13}C_6$ L-lysine (heavy), whereas cells expressing an empty control vector were labeled with $^{12}C_6$ L-lysine (light). The chromatin fraction was then analysed by mass spectrometry. As citrullination results in a mass change of ~ 1Da it can be reliably detected using this method[191]. This approach identified linker histone H1

variant 2 (H1.2) as a main PADI4 target, and that it was citrullinated within its DNA binding domain. H1 isoforms are regarded as important drivers of facultative chromatin formation[97]. They can be incorporated in a localised manner, facilitating local chromatin condensation and transcriptional repression. PADI4 mediated citrullination impairs H1.2 binding to DNA in pluripotent stem cells which leads to H1 eviction from chromatin, resulting in extensive chromatin decondensation. The 'loosened' chromatin structure, in return, facilitates gene expression of stem cell genes[190]. Besides citrullination of linker histone H1.2, this paper identified more than 700 non-histone citrullinated proteins, 20% of which were directly associated with chromatin. The list of citrullinated proteins included heterochromatin proteins HP1α and HP1γ. Citrullination of HP1α was detected at residue R107 within the hinge region. HP1γ was found to be citrullinated at residue R108 (the equivalent site to R107 in HP1α) again in the hinge region and additionally at residues R38 and R39 located within its chromodomain (CD) (Figure 1-4).

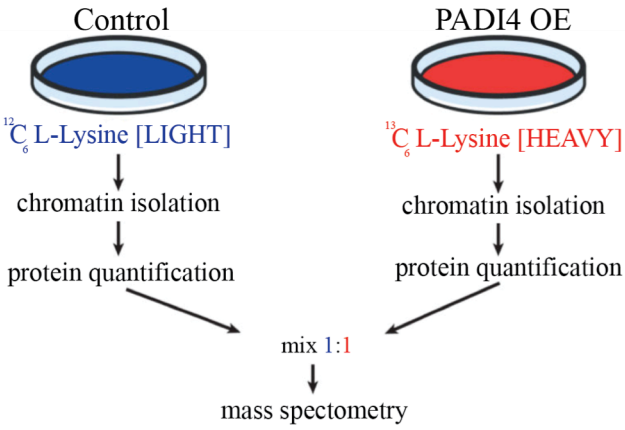


Figure 3-1 Proteomic approach to identify citrullinated proteins in mESCs.

3.1.2. Dynamics of HP1

The maintenance and *de novo* formation of heterochromatin involves the transient binding and dynamic exchange of HP1 molecules on chromatin. The dynamics of HP1 have mainly been studied using Fluorescence after photo-bleaching (FRAP) and fluorescence correlation spectroscopy (FCS) technologies. Both analyses identify that the mobility of all three HP1

paralogues is significantly lower in heterochromatin when compared to euchromatin. Diffusion measurements in mammalian cells imply that in euchromatin and heterochromatin at least two HP1 populations with different diffusion modes exist - one slow moving population, bound to chromatin and one freely diffusing population. In this chapter single molecule localisation microscopy (SMLM) is used to study HP1 dynamics in mESCs.

3.1.2.1. The principle of SMLM in live cells

Two variations of SMLM are photo-activated localisation microscopy (PALM) and stochastic optical reconstruction microscopy (STORM). PALM uses photo-activatable fluorescent proteins, whereas STORM uses organic dyes to label the protein of interest. Both techniques rely on photo-switchable fluorophores that modulate their fluorescence between an 'on' and 'off' state to separate fluorophores temporally so that their point spread functions (PSF) do not overlap (Figure 3-2)[192]. Each fluorophor can be identified and its centre localised with high precision. Single molecules can then be tracked over several consequent image frames (~movie). Compared to conventional microscopy techniques, SMLM is therefore not dependent on the diffraction limit of the microscope, but on stochastic activation of single molecules, their density and the size of the fluorophore.

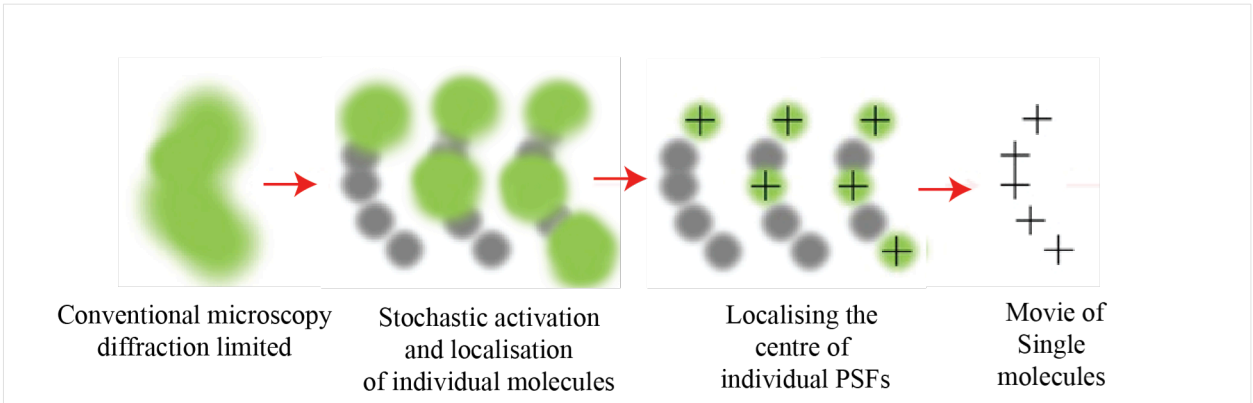


Figure 3-2 General principle of single molecule localisation microscopy.
PSF: point spread function.

SMLM allows super resolved imaging of proteins within mammalian nuclei at a spatial resolution of ~10-20nm[193]. Live-cell SMLM approaches allow single-particle tracking (SPT) of proteins on chromatin to obtain information on diffusion coefficients of proteins and their residence time on DNA. By adapting a total internal reflection fluorescence (TIRF) microscope to achieve oblique-angle illumination, a 2D plane through a cell's nucleus is illuminated, enabling detection of all molecules within that plane per image frame.

This chapter aims to characterise, *in vitro* and *in vivo,* the citrullinated sites within the CD of HP1γ. To further explore the function of HP1γ residues R38 and R39, I will present my results from studying the dynamics of HP1γ molecules during mESC differentiation using SMLM. I performed a comparative analysis of HP1γ wildtype and R38/9A mutant proteins in mESCs in order to dissect the role of residues R38/9 on protein dynamics during differentiation.

3.2. Aims of this chapter

This chapter has the following aims related to characterising citrullination of HP1γ at residues R38 and R39

1. To determine whether HP1γ residues R38 and R39 are important for HP1γ binding to H3K9me3 *in vitro*

2. To identify HP1γ citrullinated at residues R38 and R39 *in vitro*

3. To determine whether HP1γ is citrullinated in mESCs

4. To determine the effects of mutations in HP1's chromo- and chromo-shadow domains upon the dynamics of HP1 molecules in mESCs

5. To measure the dynamics of HP1γ molecules in mESCs upon differentiation

3.3. HP1γ CD residues R38 and R39 are citrullinated

Whilst sharing only ~ 70% protein sequence homology, the CDs of all three HP1 paralogues are structurally highly conserved. Residues R38/9, located within the β1-sheet of the CD, are absolutely conserved between all three homologues (Figure 1-5). Citrullination of these residues, however, was exclusively detected in HP1γ. In 3-dimensional space, R38/9 are in close proximity to amino acids F30, W51 and F53, which form the 'hydrophobic cage' that specifically recognises N-methyl groups of H3K9 in H3 tails[106]. The X-ray crystal structure of HP1γ's CD in complex with H3K9me3 illustrates that R38/9 are actually on the opposite side to the H3K9me3 binding pocket, and facing away from it, suggesting that these residues are not in direct contact with the histone tail (Figure 3-3). Nevertheless, neutralising the positive charge of R38/9 by citrullination may affect the β1-sheet structure and allosterically hinder binding of HP1γ to H3K9me2/3.

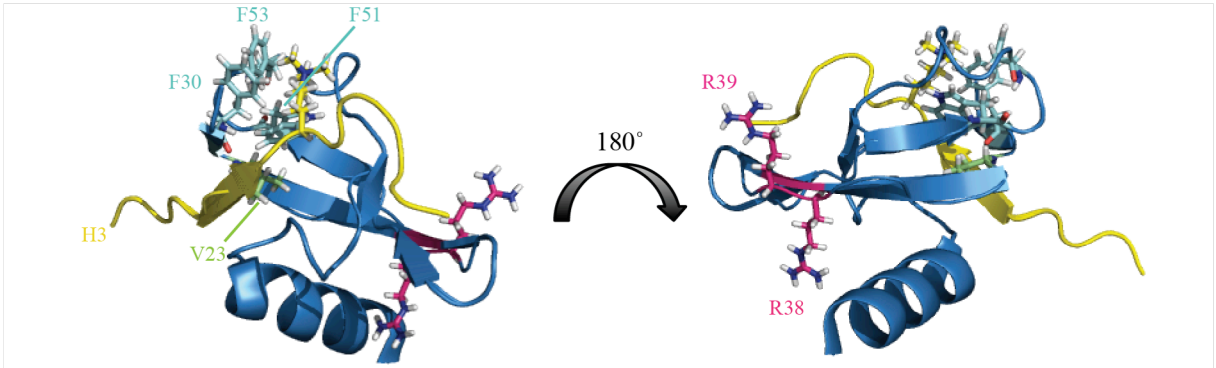


Figure 3-3 X-ray crystal structure of the CD of HP1γ in complex with H3K9me3.
PDB file '5t1i' was adapted using Pymole software. Residues forming the hydrophobic pocket are highlighted in light blue, V23 residue in green, citrullinated residues in pink and H3K9me3 tail in yellow.

3.4. Bacterial protein expression of HP1 isoforms

In order to study the contribution of R38/9 to binding H3K9me2/3 *in vitro*, both arginines were mutated to alanine. Like citrullination, this removes the positive charge associated with these residues and I will refer to the mutant as R38/9A. Single mutations of either R38 or R39 to alanine were also generated. As controls, HP1β wildtype (WT), HP1β V23M mutant and the corresponding mutation in HP1γ, V32M, were also analysed. HP1β V23M is a well-characterised mutation that abolishes binding to H3K9me2/3 *in vitro*[37] and *in vivo*[194]. A summary of the mutants is illustrated in Figure 3-4.

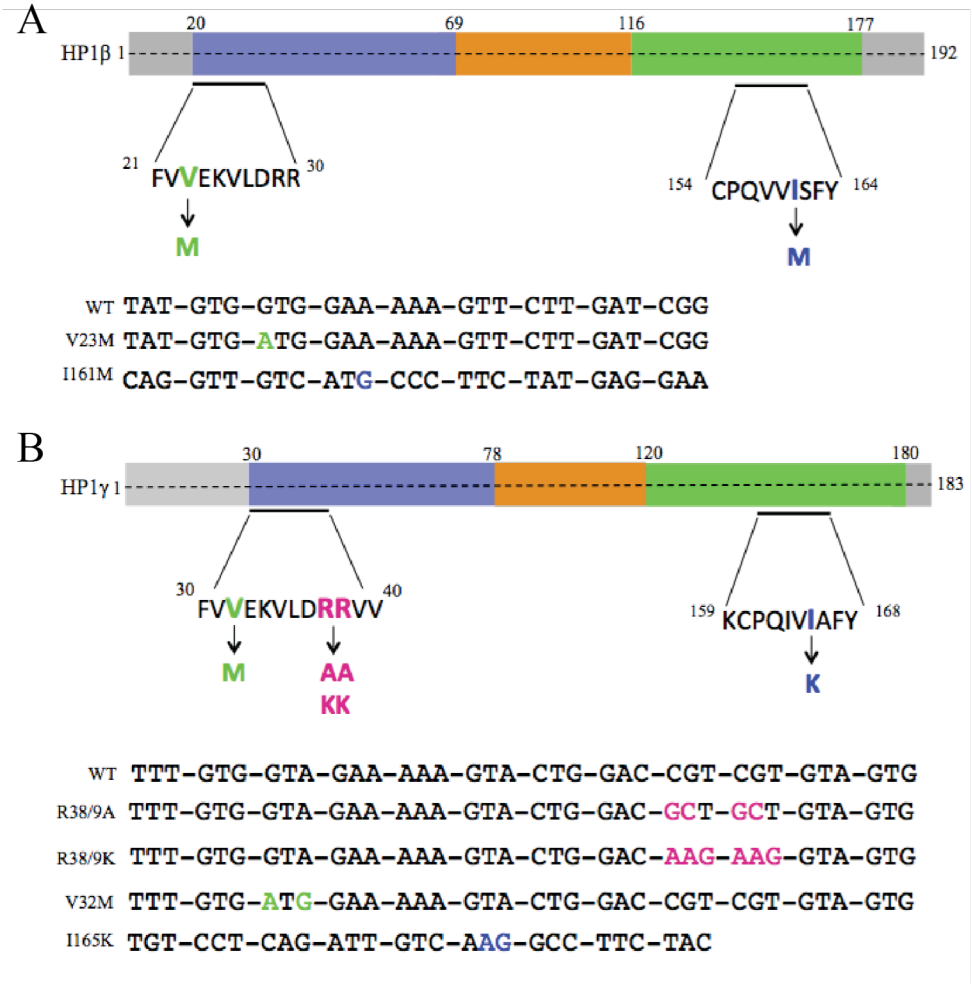


Figure 3-4 Schematic illustration of HP1β and γ mutants.
Changes in amino acid and DNA sequence are indicated.
A Summary of HP1β mutants used in this study.
B Summary of HP1γ mutants used in this study.

Full-length WT HP1γ and WT HP1β, and the CD mutants, were cloned into the bacterial expression vector pGEX-2TK. This generates a fusion protein with GST at the N-terminus. E.*coli* were transformed with plasmids encoding GST-HP1γ (WT, R38/9A, R38A, R39A and V32M) and GST-HP1β (WT and V23M). Proteins were affinity purified using glutathione sepharose beads and analysed by SDS-PAGE and Coomassie blue staining (Figure 3-5B-D). All proteins were expressed at equivalent and high levels. In order to estimate protein amounts, Coomassie stained GST-fusion proteins were compared to a BSA protein standard of known concentration (Figure 3-5A).

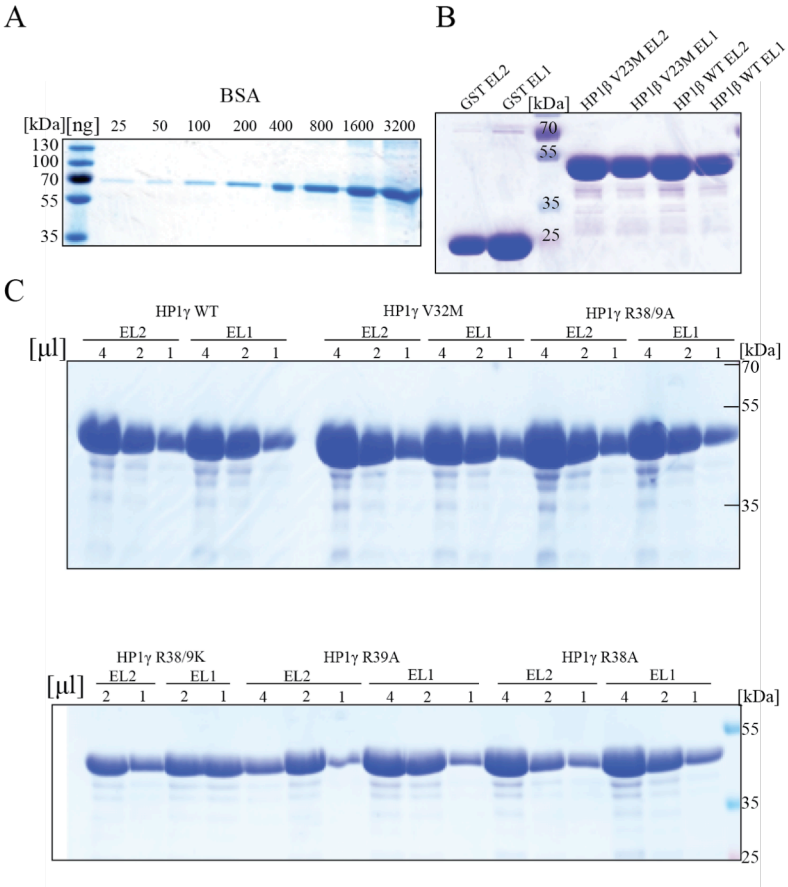


Figure 3-5 HP1β and γ - GST fusion protein expression.
Proteins were resolved by SDS-PAGE and gels stained with Coomassie blue. Eluate 1 and 2 of each GST-protein preparation were loaded.
A BSA protein standard ranging from 25-3200ng was used as a reference.
B Wildtype and mutant HP1β proteins and GST were loaded at 1µl per well.
C Wildtype and mutant HP1γ wildtype, loading as indicated.

3.5. HP1γ R38/9A has impaired binding to H3K9me3

To test binding of HP1γ and its various mutants to H3K9me3, an *in vitro* peptide pulldown assay was used. C-terminally biotinylated H3 (1-16) peptides, methylated or not at K9 (Figure 3-6A) were immobilized on streptavidin sepharose beads and incubated with 1µg of GST-HP1 fusion proteins. Equal amounts of input proteins were determined by comparing various dilutions of the GST-tagged proteins by SDS-PAGE and Coomassie blue staining to a BSA standard (Figure 3-5A). As a proof of principle, the assay was initially performed with HP1β WT and the V23M mutant. Pulldowns were performed at physiological salt conditions (150mM NaCl). As expected, the wildtype protein does not bind to unmethylated H3 peptides, but it specifically binds to the methylated peptides (Figure 3-6B). The V23M mutation abolished binding completely. The same result was obtained for HP1γ WT and the V32M mutant respectively (Figure 3-6C).

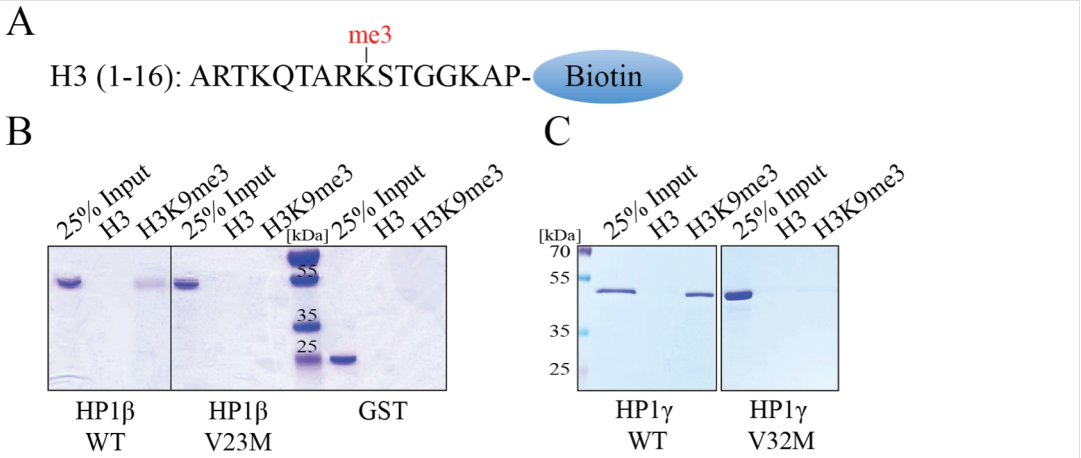


Figure 3-6 Proof of principle H3K9me3 peptide pulldown assay with HP1β and HP1γ. 1µg of GST or the indicated GST-HP1 fusion proteins were incubated with methylated or unmethylated H3 (1-16) peptides immobilised on streptavidin beads in buffer containing 150mM NaCl. Bound proteins were eluted off with 2x SDS loading buffer. Proteins were resolved by SDS-PAGE and gels stained with Coomassie blue staining solution. Images of representative pulldown assays are illustrated.
A Sequence of biotin coupled H3 (1-16) peptide trimethylated at lysine 9.
B Result of HP1β wildtype (WT) and V23M mutant pulldown.
C Result of HP1γ WT and V32M mutant pulldown.

These proof of principle experiments, using known CD mutants V23M/V32M, indicate effective assay conditions for both HP1β and HP1γ (Figure 3-6). Next, the HP1γ double R38/9A mutant was analysed using the same methodology and compared to WT HP1γ

protein. At physiological salt conditions (150mM NaCl) the R38/9A mutant showed approximately 50% less binding to the methylated H3 peptide compared to WT HP1γ protein. At 300mM NaCl, 70% less R38/9A mutant protein was bound compared to WT HP1γ protein (Figure 3-7). These results indicate that the R38/9A mutation impairs binding of HP1γ to methylated H3K9 peptides. Binding of the R38/9A mutant is more sensitive to increased salt concentration than that of the WT protein.

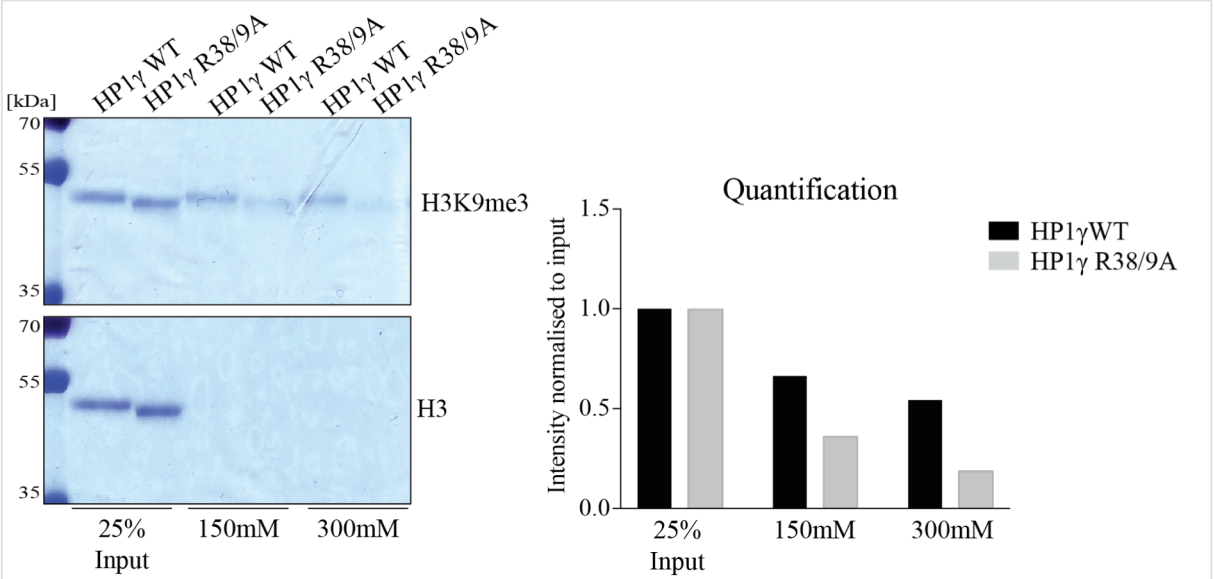


Figure 3-7 H3K9me3 peptide pulldown assay of HP1γ R38/9A mutant protein.
1μg of indicated GST-HP1γ fusion proteins were incubated with methylated or unmethylated H3 (1-16) peptides immobilised on streptavidin beads in buffer containing 150mM or 300mM NaCl. Bound proteins were eluted off with 2x SDS loading buffer. Proteins were resolved by SDS-PAGE and gels stained with Coomassie blue staining solution. Images of representative pulldown assays are illustrated (left panel). Intensities were quantified using ImageJ and values were normalised to input (25%) (right panel).

In order to understand whether both R38 and R39 residues are equally important for HP1γ to bind methylated H3K9 peptides, the experiment was repeated with single mutants R38A and R39A (here at physiological salt conditions). Whilst binding to H3K9me3 peptides is equally reduced in both R38A and R39A when compared to the wildtype protein, binding is significantly reduced in the double 38/39 alanine substituted HP1γ (Figure 3-8A). Importantly however, mutating both R38 and R39 to lysine (R38/9K) restored binding back to WT levels, suggesting that the positive charges at positions 38 and 39 are required to maintain the local structure of the chromodomain and/or the electrostatic interaction(s) between H3K9me3 and the chromodomain of HP1γ (Figure 3-8B).

In summary, my results suggest that a positive charge at either position 38 or 39 within HP1γ's CD is important for it to interact with methylated histone H3 tails *in vitro*. Mutating both residues to alanine reduces binding significantly, whereas the R38/9K mutant, which maintains a positive charge at both sites, has full binding capacity. My data indicate that the charge loss accompanied by citrullination of residues R38/9 may interfere with HP1γ's binding to chromatin.

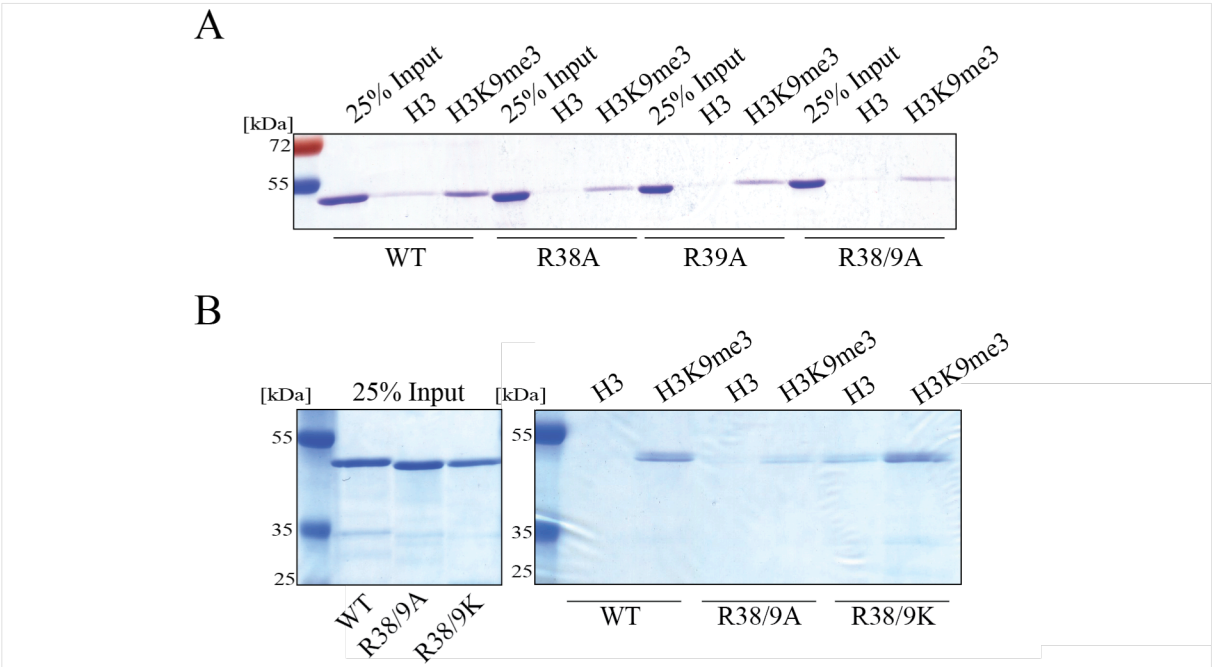


Figure 3-8 H3K9me3 peptide pulldown assay of HP1γ mutant protein.
1µg of indicated GST-HP1γ fusion proteins were incubated with methylated or unmethylated H3 (1-16) peptides immobilised on streptavidin beads in buffer containing 150mM NaCl. Bound proteins were eluted off with 2x SDS loading buffer. Proteins were resolved by SDS-PAGE and gels stained with Coomassie blue staining solution. Images of representative pulldown assays are illustrated.
A Peptide pulldown performed with GST-HP1γ wildtype (WT) and R38A, R39A and R38/9A mutant proteins.
B Peptide pulldown performed with GST-HP1γ WT, R38/9A and R38/9K mutant proteins.

3.6. HP1γ citrullination *in vitro*

In order to investigate if HP1γ is a direct target of PADI4, an *in vitro* citrullination assay using recombinant human PADI4 was used. Full length human PADI4 was cloned into the bacterial expression vector pGEX-6p to generate a N-terminal tagged GST-PADI4 fusion construct. Recombinant proteins were affinity purified, resolved by SDS-PAGE and

visualised by Coomassie blue staining (Figure 3-9A). *In vitro* citrullination assays were performed as previously described[190]. In brief, 100ng GST-PADI4 enzyme was incubated with an equal amount of substrate in deimination buffer in the presence or absence of 10mM $CaCl_2$ for 1h at 37°C. Citrullination was assessed by immunoblotting using an α-citrulline antibody. As a positive control, human recombinant histone H3.1 (NEB) (Figure 3-9B) - a known substrate of PADI4[49] was used. Citrullination of H3.1 was assessed using an α-H3R2-cit antibody. As expected, citrullination of H3.1 was exclusively detected in the presence of Ca^{2+} indicating that the enzymatic activity of PADI4 depends on Ca^{2+}.

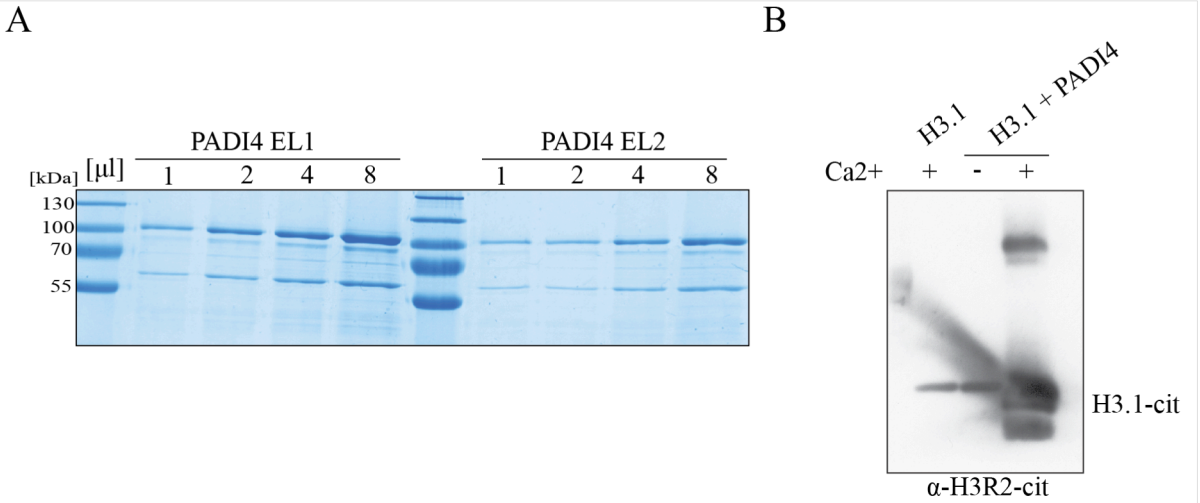


Figure 3-9 *In vitro* citrullination assay of H3.1 with PADI4.
In vitro citrullination assays were performed with a recombinant PADI4-GST fusion protein for 1h at 37°C in the presence and absence of Ca^{2+}. Reactions were resolved by SDS-PAGE and analysed by immunoblotting.
A SDS-PAGE of GST-PADI4 fusion proteins, stained by Coomassie blue solution. Eluate 1 (EL1) and 2 (EL2) of each GST-PADI4 preparation were loaded at the indicated amounts.
B Immunoblotting of *in vitro* citrullination reaction with 1µg of H3.1 as a substrate and 100ng PADI4 with a α-H3R2-cit antibody.

To optimise citrullination assays for GST-HP1γ, titrations of either enzyme or substrate were performed (Figure 3-10). Equivalent input protein amounts, as judged from Coomassie blue stained SDS gels, were confirmed by Ponceau S staining of transferred proteins (data not shown). First, the amount of HP1γ was kept constant (1µg) and increasing amounts (50-400ng) of PADI4 were added (Figure 3-10A). Citrullinated HP1γ was detected only in the presence of PADI4 and Ca^{2+}. The strongest signal relative to background was obtained when assaying 100ng of PADI4. Interestingly, adding more than 200ng of PADI4 appeared to alter

its substrate specificity as citrullination of HP1γ diminished, whilst PADI4 autocitrullination increased.

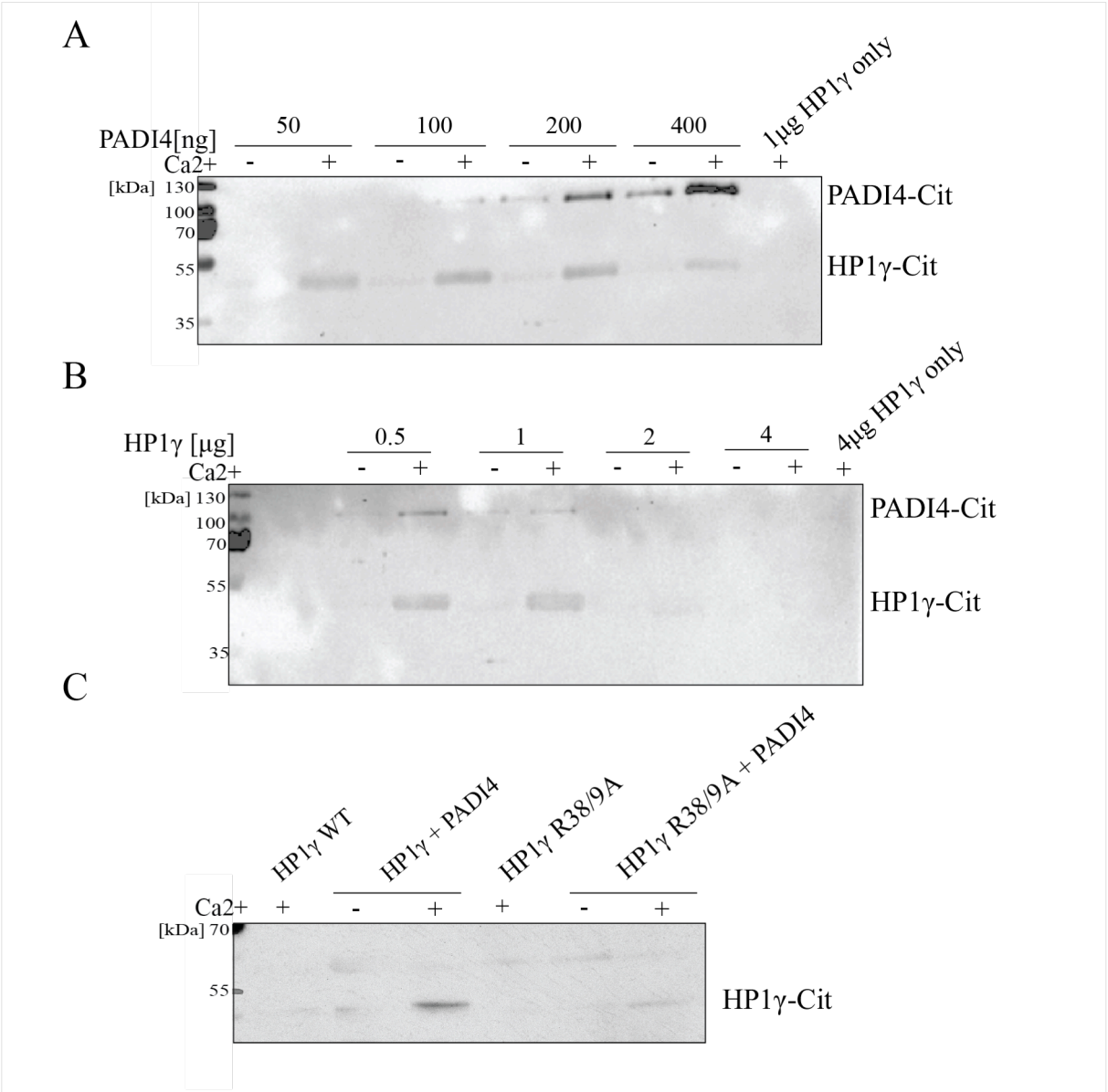


Figure 3-10 *In vitro* citrullination of HP1γ with PADI4.
In vitro citrullination assays of HP1γ were performed with a recombinant GST-PADI4 fusion protein for 1h at 37°C in the presence or absence of Ca^{2+}. Reactions were resolved by SDS-PAGE and analysed by immunoblotting with an α-citrulline antibody.
A Titration of PADI4 ranging from 50-400ng with a constant amount of HP1γ WT (1µg).
B Titration of HP1γ WT ranging from 0.5-4µg with a constant amount of PADI4 (200ng).
C Reaction of 1µg HP1γ WT and R38/9A mutant with 200ng PADI4.

In the next step, the amount of enzyme was kept constant (200ng) and increasing amounts (0.5-4μg) of HP1γ were added (Figure 3-10B). The strongest signal was achieved with 1μg HP1γ in the presence of PADI4 and Ca^{2+}. Increasing HP1γ protein input above 2μg completely inhibited the citrullination reaction, indicating a substrate-mediated inhibition of the enzyme.

In order to investigate whether PADI4 catalyses citrullination of HP1γ residues R38/9, *in vitro* citrullination assays were performed using WT HP1γ compared to the R38/9A mutant protein. 200ng of PADI4 was incubated with 1μg of WT HP1γ and R38/9A mutant protein respectively. Compared to WT HP1γ, the signal for R38/9A, in presence of PADI4 and Ca^{2+}, is considerably reduced, suggesting that *in vitro* citrullination of residues R38/9 is mediated by PADI4 (Figure 3-10C). The residual signal observed for the reaction of R38/9A mutant protein and PADI4 could be due to the citrullination of residue R108 or other, yet undetected citrullination sites within HP1γ.

My results show that HP1γ is citrullinated by recombinant PADI4 *in vitro* in a Ca^{2+} dependent manner and that mutating R38/9 to alanine notably reduces citrullination of HP1γ. Furthermore, the data show that, as with many enzymes, the activity of PADI4 *in vitro* strongly depends on concentration of substrate and enzyme.

3.7. Generation of site-specific antibodies against HP1γ R38/9-citrulline

In the previous section, I showed that mutation of R38/9 to alanine leads to significantly reduced citrullination of HP1γ in *in vitro* citrullination assays. This suggests very strongly that PADI4 citrullinates HP1γ at residues R38/9, at least *in vitro*. Therefore, in order to study this modification in more physiologically relevant settings, I decided to raise antibodies that would specifically recognise HP1γ citrullinated at R38/9. A peptide immunogen was designed based upon the mouse HP1γ protein sequence, with the citrullinated residues 38/9 in the centre, flanked by 4-5 amino acids each side (HP1γ 34-44) and with a C-terminal cysteine (Figure 3-11A). Peptides were covalently bound to the carrier protein keyhole limpet hemocyanin (KLH) via their C-terminal cysteine. KLH is highly immunogenic and is therefore used to help initiate an immune response in the animal. 4mg of immunogen,

homogenized in complete Freund's adjuvant, were injected subcutaneously into one sheep. The complete immunisation schedule is illustrated in (Figure 3-11B).

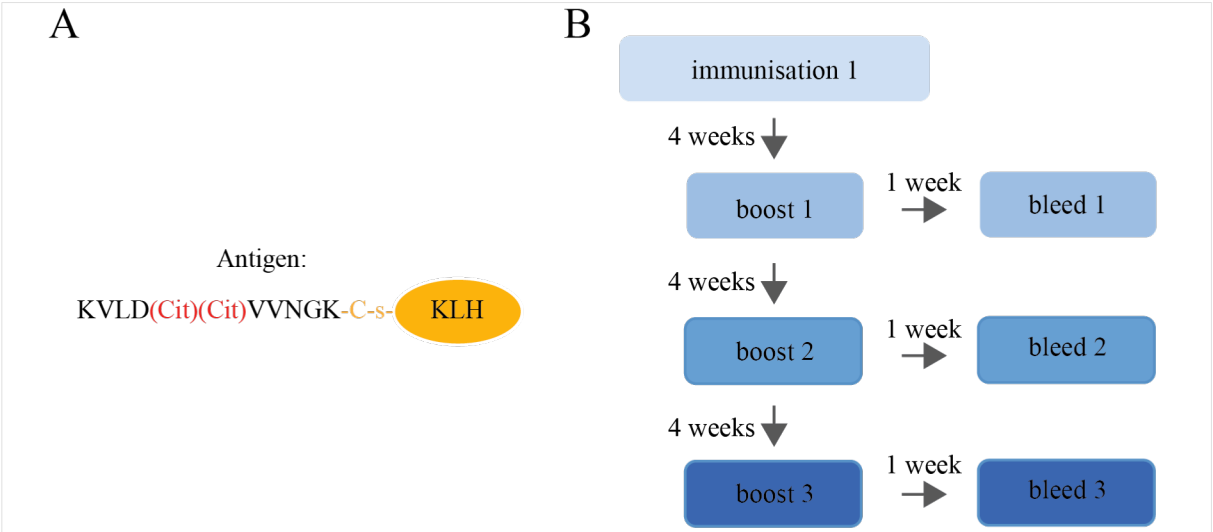


Figure 3-11 Sheep immunisation to produce HP1γ R38/9-citrulline antibodies.
A Sequence of HP1γ (33-44) R38/9-citrulline antigen coupled to KLH used for sheep immunisation.
B Immunisation schedule.

A total number of 4 sheep were immunised with the same immunogen. Sera were collected after 4 weeks (1st bleed), 8 weeks (2nd bleed) and 12 weeks (3rd bleed and final bleed). Crude sera of all three bleeds were initially screened in dot blots against mouse HP1γ unmodified (34-44) and the double citrulline R38/9-cit peptides (data not shown). Bleeds showing the most specific recognition of citrullinated peptides were subjected to affinity purification as illustrated in Figure 3-12 (compare also Section 2.6.6). Purification was performed with R38/9-cit peptides and antigen-specific antibodies were serially eluted using low pH (0.1M glycine pH 2.5) and neutralized with ice-cold Tris HCl pH8.0 buffer. A representative methylene blue staining of purified antibodies (here S780D) pre-dialysis indicates the relative amount of protein obtained in each elution step (Figure 3-13A). After dialysis the antibody was resolved in a SDS gel and Coomassie blue stained (Figure 3-13B). Bands for both heavy and light chains of purified antibodies were detected in each elution step.

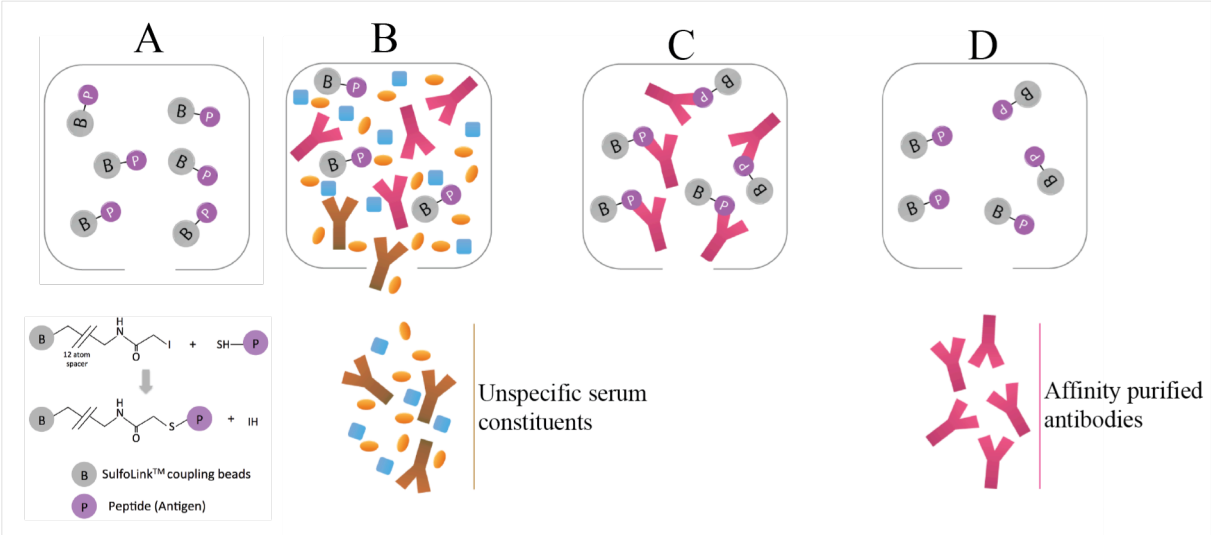


Figure 3-12 Scheme of antibody affinity purification strategy.
Compare Section 2.6.6 for further technical details.
A Peptides (**Figure 3-11**A) are coupled to Sulfolink™ beads via disulfide bridge.
B Peptides are incubated with crude serum.
C Serial washes to remove unbound serum constituents.
D Affinity purification using use low pH (0.1M glycine pH 2.5). Purified antibodies were neutralized using ice-cold Tris HCl pH8.0.

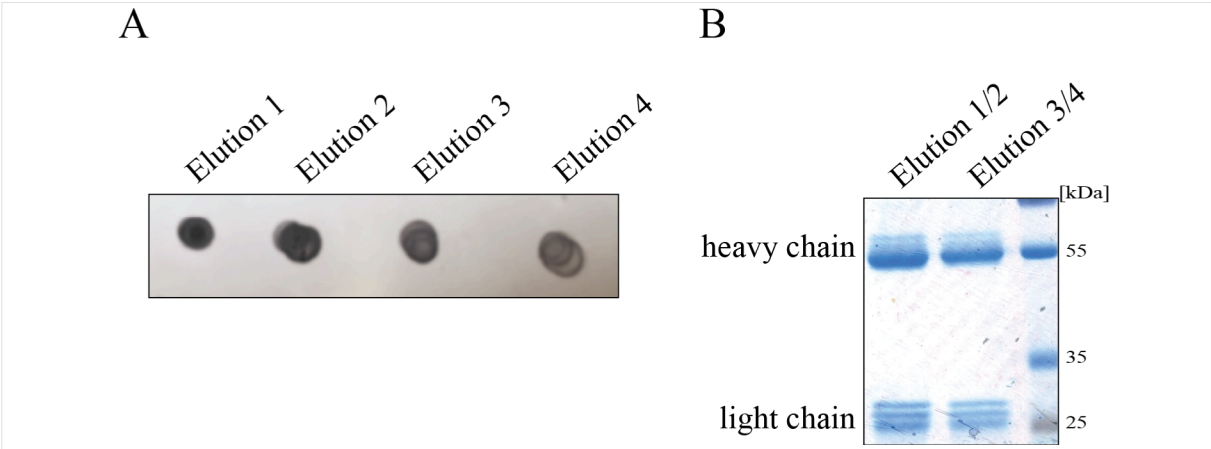


Figure 3-13 Quality control of purified HP1γ R38/9-citrulline antibodies.
A 2µl of purified antibody eluates were immobilized on humidified PVDF membrane, dried and stained with 2% [w/v] methylene blue solution.
B SDS-PAGE of purified antibody eluates (elution 1-2 and 3-4 were combined). 2µl were loaded per well. Proteins were stained by coomassie staining solution. Heavy and light chains of antibodies are indicated.

3.8. Dot blot analysis of HP1γ R38/9-citrulline antibodies

Four affinity-purified polyclonal antibodies against HP1γ R38/9-citrulline, referred to as S780D, S781D, S923D and S924D, were further tested in dot blots. They were tested against increasing amounts of various HP1γ peptides. Peptides screened were HP1γ (34-44) unmodified peptide (HP1γ UM), double citrulline R38/9-cit (HP1γ-R38/9-cit), single citrulline R38-cit (HP1γ-R38-cit), single citrulline R39-cit (HP1γ-R39-cit), R108-cit (HP1γ-R108-cit) and R107-cit (HP1α-R107-cit) (Figure 3-14). Peptide sequences are summarised in Table 2-9 (Section 2.7.2). All antibodies specifically recognise the HP1γ-R38/9-cit peptides with very low binding to unmodified peptides. S780D was the most specific antibody with no specificity against any other peptides tested (Figure 3-14A). S781D showed moderate specificity against the R39-cit peptide (Figure 3-14B).

S923D recognised the R39-cit to the same extend as the R38/9-cit peptide, suggesting that this polyclonal antibody contains an equal amount of reactive antibodies recognising R39-cit or R38/9-cit (Figure 3-14C). At this point, I decided to exclude this antibody from further characterisation, although it may prove useful in the future if the laboratory decides to specifically study citrullination of HP1γ R39-cit.

S924D showed high specificity against R38/9-cit with low to moderate sensitivity to both R38-cit and R39-cit (Figure 3-14D). To enrich S924D for clones recognising only the doubly modified R38/9-cit peptide, the antibody preparation was sequentially passed over R38-cit and R39-cit peptide columns to remove antibodies that recognise each single modification. In each case, the column flow through was recovered. Dot blot results using the antibody after these additional purification steps confirmed that S924D was successfully depleted of antibodies specifically recognizing R38-cit and R39-cit, whilst it still recognised R38/9-cit (Figure 3-14E). This additionally purified antibody, S924D, together with antibodies S870D and S871D were used for further characterisation.

In summary, my results demonstrate that I successfully generated three polyclonal antibodies with high specificity for the modified HP1γ R38/9-cit peptide.

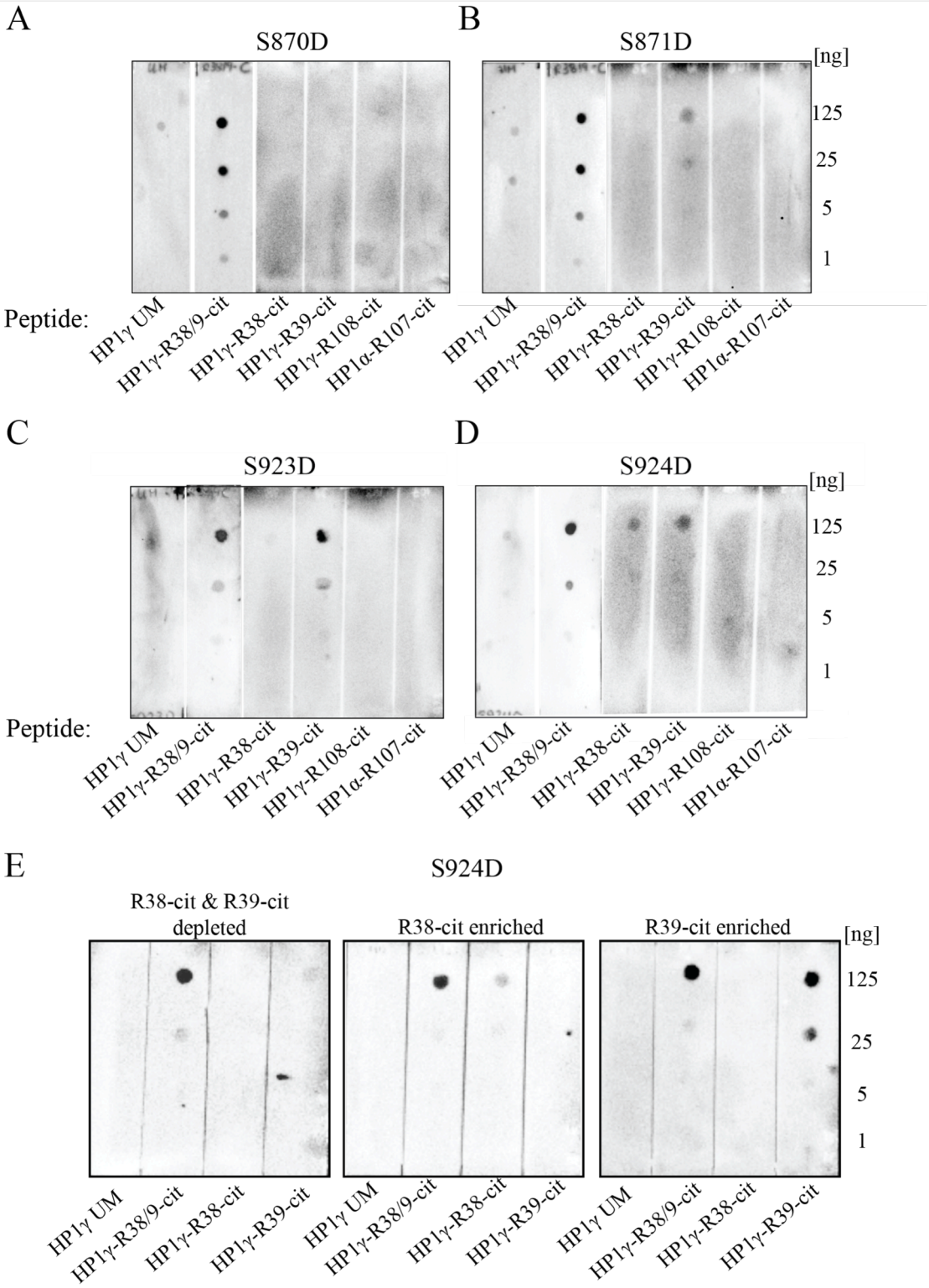


Figure 3-14 Dot blot analysis of HP1γ R38/9-citrulline antibodies.
(Figure legend is located on the next page).

Figure 3-14 Dot blot analysis of HP1γ R38/9-citrulline antibodies.
HP1γ-R38/9-cit antibodies were tested against various peptides: HP1γ (34-44) unmodified peptide (HP1γ UM), double cit R38/9-cit (HP1γ-R38/9-cit), single cit R38-cit (HP1γ-R38-cit), single citR39-cit (HP1γ-R39-cit), single cit R108-cit (HP1γ(104-111)-R108-cit) and single cit R107-cit (HP1γ(103-112)-R107-cit). Peptides were immobilised on PVDF membranes at indicated amounts (1-125ng) and incubated with purified HP1γ-R38/9-cit antibodies at 1:500 dilution.
A S870D, **B** S871D, **C** S923D, **D** S924D
E S924D antibody was repurified against single peptides HP1γ-R38-cit and R39-cit.

3.9. HP1γ R38/9-citrulline antibodies tested in *in vitro* citrullination assays

In order to test whether the antibodies recognise the epitope within full-length HP1γ, I performed *in vitro* citrullination assays (see Section 3.6) using 1μg of HP1γ WT and R38/9A mutant proteins. Equivalent input protein amounts, as judged from Coomassie blue stained SDS gels compared to a BSA standard (Figure 3-5) were used and confirmed by Ponceau S staining of transferred proteins (data not shown). S871D recognised both HP1γ WT and R38/9A in the absence of PADI4 enzyme (Figure 3-15A). This signal did not increase upon addition of PADI4, indicating the antibody recognised the backbone of HP1γ, with no specificity for citrullinated residues R38/9.

Two antibodies S870D (Figure 3-15B) and S924D ((Figure 3-15C), did not recognise WT HP1γ in the absence of PADI4 but did so when PADI4 was included in the reaction. This indicates that the antibodies do indeed specifically recognise citrullinated HP1γ. Intriguingly though, both recognised the HP1γ R38/9A mutant protein even in the absence of PADI4, though S924D did less so than S870D. However, unlike with WT HP1γ, the antibodies' recognition of HP1γ R38/9A was unaffected by addition of PADI4 to the reaction (Figure 3-15B/C). Based on these results, I concluded that both S870D and S924D specifically recognise an epitope surrounding R38/9, which they bind to only when the two arginines (and/or the associated positive charges) are replaced by other amino acids (either citrulline or alanine). In other words, the fact that the antibodies showed no specificity against HP1γ WT indicates that they exclusively recognise their epitope only once the arginines (and their positive charges) are lost either through citrullination or mutating the residues to alanines.

To further test whether loss of charge within the region is required to allow antibody recognition, I repeated the experiment using a HP1γ R38/9K mutant protein. The results for both antibodies, S870D (top panel) and S924D (bottom panel), are summarized in Figure 3-16. Indeed, antibody recognition of R38/9K in absence of PADI4 was significantly weaker compared to recognition of HP1γ R38/9A protein. Importantly, addition of PADI4 had no further influence on the signal which is consistent with my hypothesis that the antibodies cannot recognise their epitopes if positive charges are present.

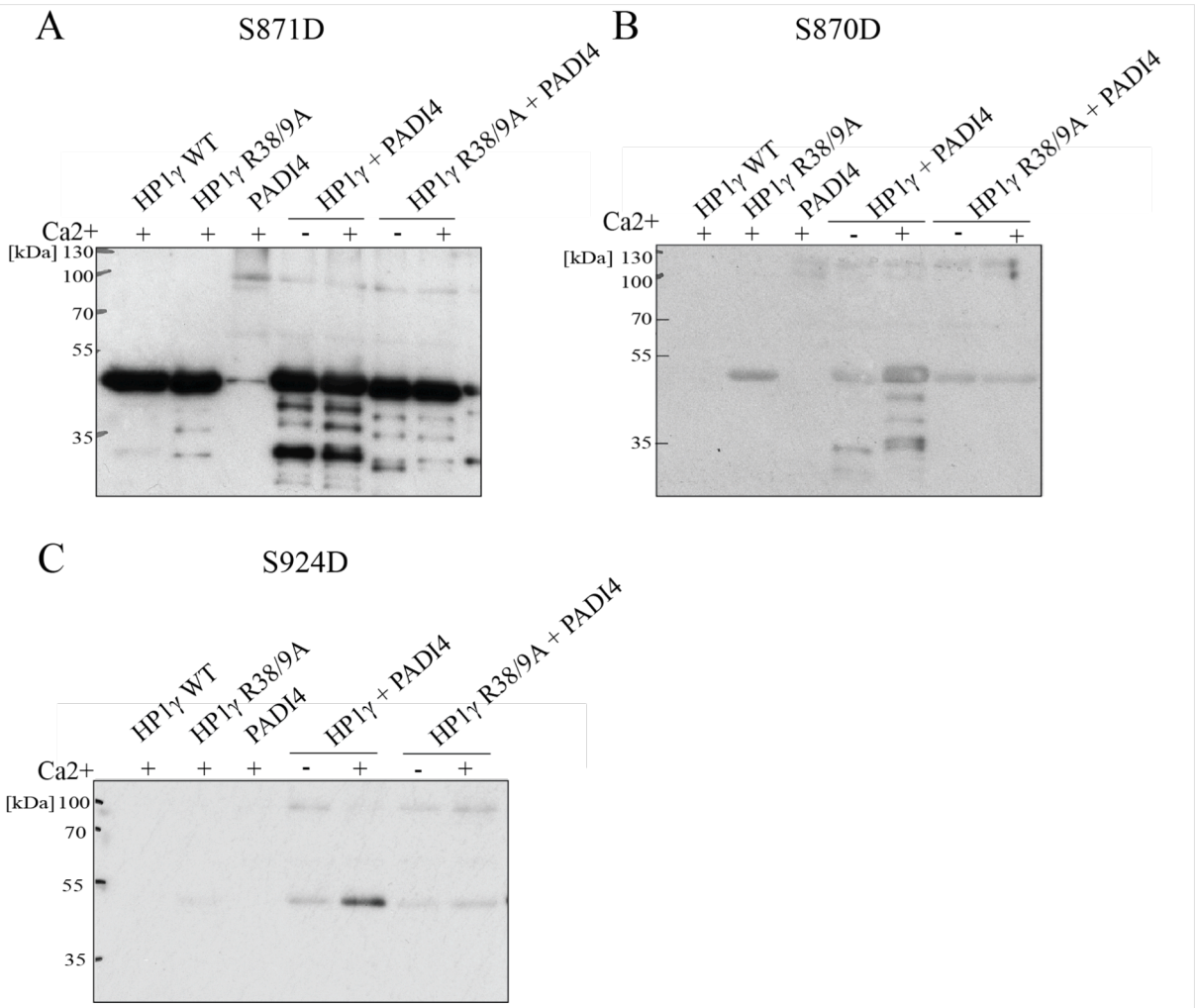


Figure 3-15 Test of HP1γ-R38/9-cit antibodies using *in vitro* citrullination of HP1γ.
In vitro citrullination assays with 1µg of HP1γ WT or R38/9A mutant protein were performed with 200ng of GST-PADI4 fusion protein for 1h at 37°C in the presence or absence of Ca^{2+}. Reactions were resolved by SDS-PAGE and analysed by immunoblotting with indicated HP1γ-R38/9-cit antibodies.
A: S871D**, B:** S780D**, C:** S924D

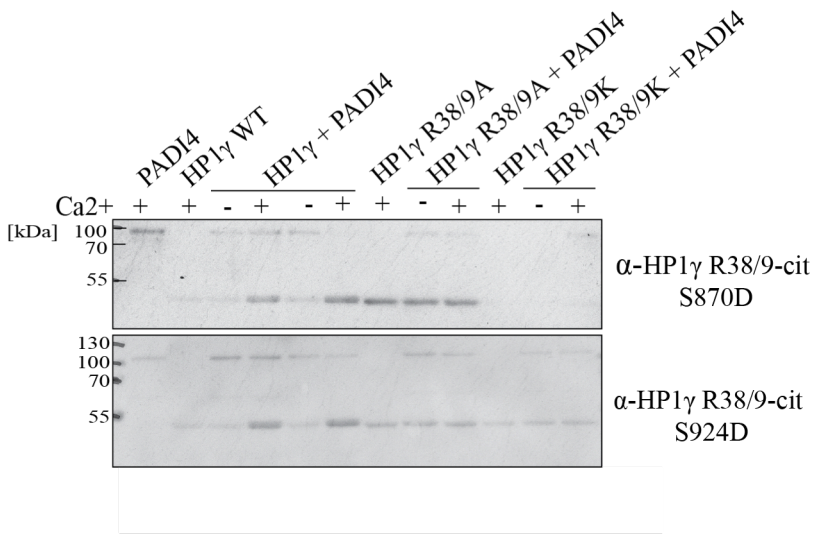


Figure 3-16 Test of HP1γ-R38/9-cit antibodies S870D and S924D using *in vitro* citrullination of HP1γ.

In vitro citrullination assays of 1µg of HP1γ WT, R38/9A or R38/9K mutant proteins were performed with 200ng of GST-PADI4 fusion protein for 1h at 37°C in the presence or absence of Ca^{2+}. Reactions were resolved by SDS-PAGE and analysed by immunoblotting with HP1γ-R38/9-cit S870D (upper panel) and S924D (lower panel) antibodies. HP1γ WT reaction was performed in duplicate.

In conclusion, *in vitro* citrullination data demonstrate that both antibodies, S870D and S924D, specifically recognize their epitopes within HP1γ when arginines 38 and 39 are converted to citrulline or alanine. Crucially, a robust PADI4 dependent increase in signal is observed with HP1γ WT protein, indicating that the antibody can be used to detect citrullination of HP1γ at R38 and R39.

3.10. Detection of HP1γ R38/9-citrulline in mESCs lysates

Dot blot and *in vitro* citrullination data suggest that antibodies S780D and S924D specifically recognise PADI4 dependent citrullination of HP1γ residues R38/9. To understand whether this modification occurs *in vivo,* I transiently overexpressed HA-PADI4 in mESCs. As a comparison, I also included PADI2, another nuclear PADI enzyme[86]. Extracts were then analysed by immunoblotting using S780D and S924D antibodies. In order to enrich for HP1γ protein, chromatin was extracted using IPH chromatin extraction buffer. Protein expression of HA-PADI4 and PADI2 was observed by immunoblotting using a α-HA or PADI2 antibody respectively (Figure 3-17A).

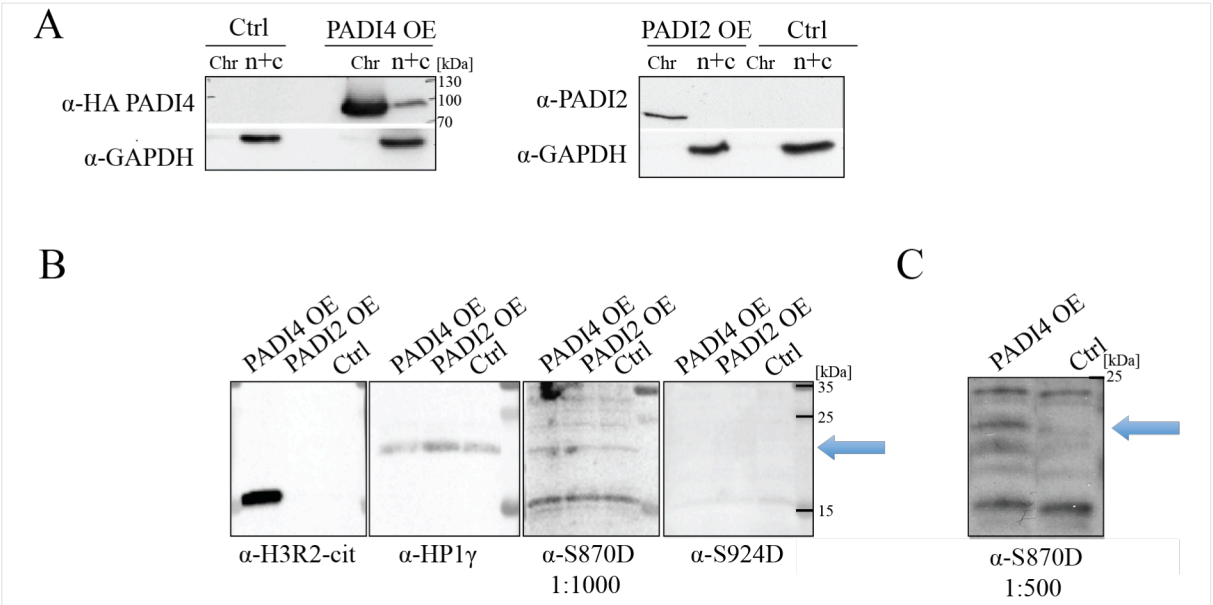


Figure 3-17 Detection of HP1γ R38/9-cit in mESCs upon PADI4 overexpression.
mESCs were transfected with pCDNA3.1-HA-PADI4 or pCMV6-PADI2 constructs and harvested 48h post transfection. Cellular fractionation of mESCs was performed using IPH chromatin extraction buffer. Immunoblotting was performed with the indicated antibodies.
A Equivalent protein lysates of each fraction (from 40000 cells) were loaded per well. Chr: chromatin fraction; n+c: nucleo- and cytoplasmic fraction.
B Chromatin fractions (from 40000 cells) from cells overexpressing PADI4, PADI2 or control cells were resolved in a 12.5%(w/v) SDS PAG. Immunoblotting was performed with antibodies as indicated (S870D was used at 1:1000 dilution). Blue arrow indicates molecular weight of HP1γ.
C Chromatin fractions (from 80000 cells) from cells overexpressing PADI4 or control cells were dissolved in a 17.5%(w/v) SDS PAG. Immunoblotting was performed with S870D antibody at 1:500 dilution. Blue arrow indicates molecular weight of HP1γ.

Both enzymes were detected within the chromatin fraction. The enzymatic activity of both enzymes was tested by detection of citrullinated H3 by immunoblotting using a commercially available H3R2-cit antibody (Figure 3-17B). H3R2-cit was only detected in PADI4 and not in PADI2 overexpressing cells, which indicates that overexpressed PADI2 was not enzymatically active (at least towards H3). Notably, a weak signal at the molecular weight of HP1γ was observed using the S870D antibody, only upon overexpression of PADI4 (Figure 3-17B). Immunoblotting with antibody S924D showed no significant signal.

To enhance the HP1γ R38/9-cit signal detected with the S870D antibody, I repeated the immunoblotting experiment with double the amount of protein loaded per well, increased the polyacrylamide concentration of the gel and increased the antibody concentration. The

previous HP1γ R38/9-cit signal observed after overexpression of PADI4 (Figure 3-17B) was significantly enhanced (Figure 3-17C). By resolving the proteins on a 17.5% SDS gel a second band at a molecular weight approximately 2-3kDa smaller than that of HP1γ was detected.

To conclude, immunoblotting experiments using the site-specific HP1γ R38/9-cit antibody S870D suggest HP1γ is citrullinated at residues R38/9 upon overexpression of PADI4 in mESCs. Although the S924D antibody showed very promising results in *in vitro* assays, no specific signal was detected in mESC lysates.

3.11. Single molecule localisation microscopy (SMLM) of HP1

To further explore the role of residues R38 and R39 on HP1γ's function in mESCs, I studied the dynamics of HP1γ wildtype and R38/9A mutant molecules using 2D single molecule localisation microscopy (SMLM) in live mESCs during differentiation. I performed a comparative analysis in undifferentiated mESCs of the dynamics of HP1γ and HP1β wildtype and different CD and CSD mutants. This part of the project is an ongoing collaboration with Dr. Srinjan Basu (Prof. Ernest Laue laboratory, Department of Biochemistry, University of Cambridge).

In order to measure the diffusion of single HP1 molecules, 2D imaging at 15ms time resolution was performed using high laser excitation powers. To measure the residence time of proteins bound to DNA, imaging was performed at lower laser excitation powers at 500ms time resolution because longer exposures lead to motion blurring of freely diffusing molecules as previously described[195]. Fluorescence images were collected as movies of 10,000 frames at 15ms or 1,000 frames at 500ms exposure. Single-molecule traces were analysed using software that detects single-molecule trajectories from SPT movies. High excitation laser powers are necessary to achieve high-precision localisation at 15ms exposures and thereby accurately determine diffusion coefficients of single molecules. However, high excitation laser powers can lead to photobleaching, therefore low excitation powers with longer exposures are used to ensure long track lengths that are ideal for determining the residence times of proteins bound to chromatin. A detailed description of the data analysis is given in Section 2.9.4.

3.11.1. Generation of mESC lines expressing mEos3.2-HaloTag -HP1 proteins

In order to measure both fast diffusion and residence times of HP1 proteins on chromatin, mammalian expression vectors were used encoding HP1 proteins N-terminally tagged to a fusion protein consisting of mEos3.2[ref.86] and a HaloTag[196] (Figure 3-18A). mEos3.2 is a widely used photo-convertible fluorophore. In the fluorescent 'off' state mEos3.2 is a green fluorescent protein that can be excited at 488nm, emitting light at 516nm. It can be activated to a fluorescent 'on' state with a 405nm laser triggering a conformational change of the protein. In the fluorescent 'on' state mEos3.2 is a red fluorescent protein, which can be excited at 561nm and emits light at 587nm[86]. mEos3.2 has limited photostability as it can only be activated once or twice before molecules are irreversibly bleached.

HaloTag is a bacterial enzyme used to label proteins with organic dyes that are introduced into cells. HaloTag is a hydrolase with a genetically modified active site, which binds specifically to reactive chloroalkane ligand forming a covalent alkyl-enzyme adduct[197]. The ligand is coupled to the fluorophore of choice. In this study a HaloTag chloroalkane ligand coupled to Janelia Fluor (JF_{549}) was used[198]. In order to avoid overlaps of PSFs (point spread functions) of single molecules, molecules were labeled using low concentrations of HaloTag-JF_{549} ligand (0.5nM). JF_{549} (excitation maximum: 549nm; emission maximum: 571nm) is a photostable dye enabling imaging of molecules for longer, before irreversible photobleaching. In this study, the HaloTag-JF_{549} ligand labeling was used to measure both fast diffusion and residence time of HP1 molecules on chromatin.

mESCs were transfected with pSB7 mammalian expression vectors containing mEos3.2-HaloTag in frame with mouse HP1β or γ cDNA sequences (Figure 3-18A). Both HP1β or γ wildtype proteins were compared to CD mutants (HP1β V23M; HP1γ V32M and R38/9A) and CSD mutants (HP1β I161M; HP1γ I165K). As described before, mutations in the CD abolish binding to H3K9me3 of HP1β and γ proteins.

Isoleucine residues in the chromo-shadow domain are conserved residues located within the PxVxL motif of HP1β and γ. Their mutation abolishes homo- and hetero dimerisation of HP1 and interaction with other proteins[118,199]. A summary of all HP1β and γ mutants is illustrated in Figure 3-4.

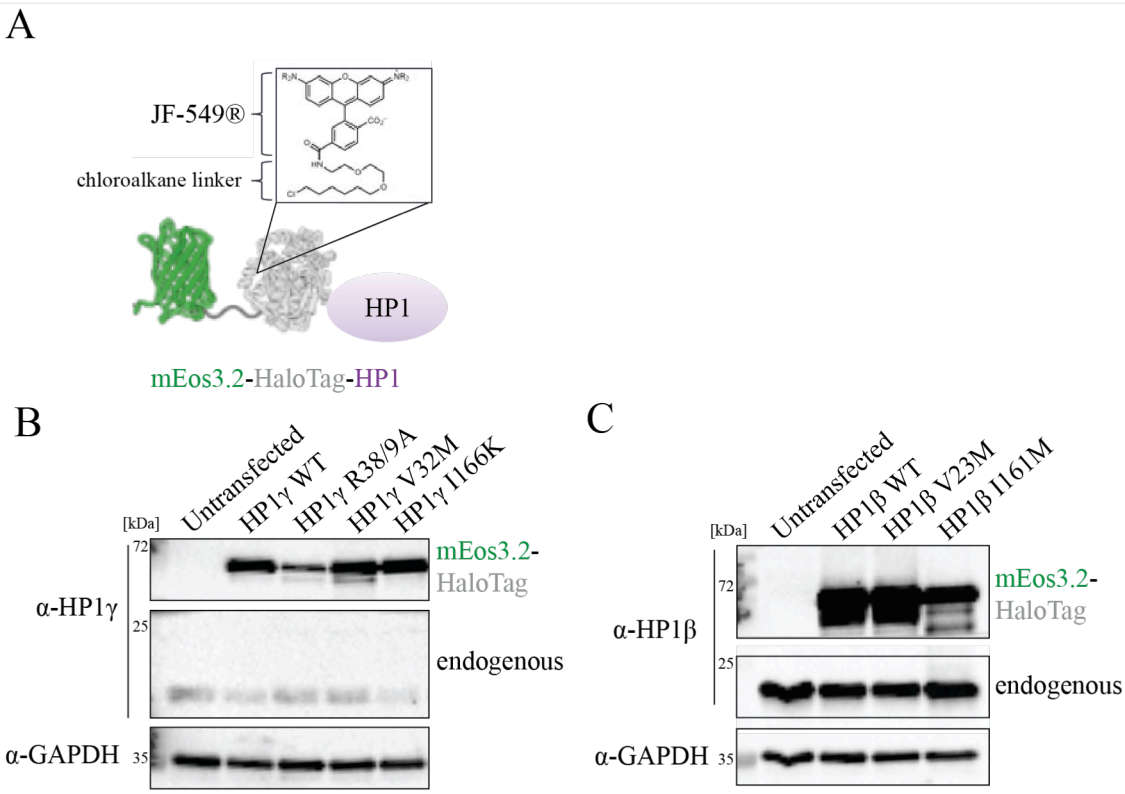


Figure 3-18 Stable expression of mEos3.2-HaloTag - HP1 constructs in mESCs.
mESCs were transfected with pSB7 mammalian expression vectors encoding mEos3.2-HaloTag-HP1 fusion proteins. After 14 days of geneticin selection, mEos3.2 positive cells were enriched by cell sorting. Immunoblotting using the indicated antibodies was performed 7 days after cell sorting. Whole cell lysates from 50000 cells were loaded per well.
A Schematic illustration of mEos3.2-HaloTag-HP1 fusion proteins. X-ray crystal structure of green EosFP (PDB identifier 1ZUX) and HaloTag (PDB identifier 5UXZ) were used to generate this image. Chemical structure of JF_{549} dye coupled to HaloTag-chloroalkane ligand and its approximate binding pocket is illustrated.
B Immunoblotting, with the indicated antibodies, of HP1γ WT and mutant fusion proteins.
C Immunoblotting, with the indicated antibodies, of HP1β WT and mutant fusion proteins.

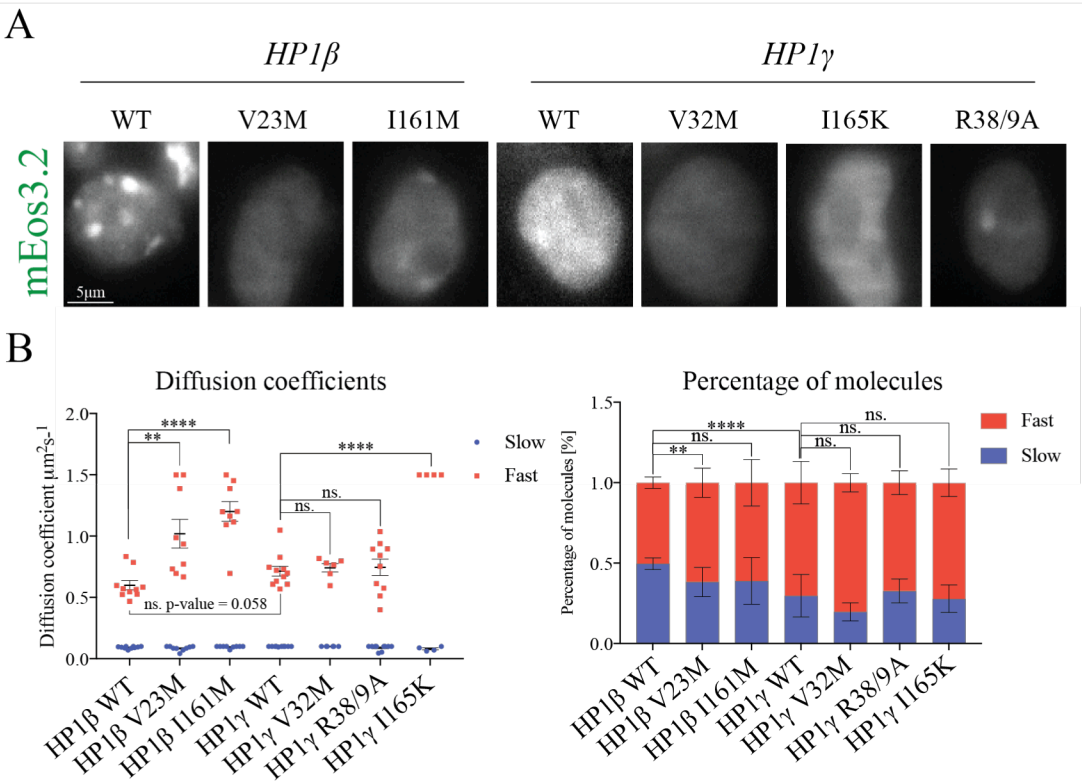
pSB7 vectors contain a neomycin resistance cassette to generate stable cell lines. After 48h, cells were cultured in medium supplemented with geneticin. 14 days later mEos3.2 positive cells (excitation: 488nm; emission: 516nm) were sorted in bulk. Protein expression of HP1β and γ fusion proteins in stable cell lines was confirmed by immunoblotting (Figure 3-18B/C). HP1 fusion proteins have a molecular weight of approximately 70kDa. mEos3.2-HaloTag-HP1γ wildtype, R38/9A, V32M and I165K mutant proteins were expressed at higher levels than endogenous HP1γ protein (Figure 3-18B). HP1γ R38/9A fusion proteins were expressed at approximately 50% lower levels than HP1γ wildtype fusion proteins. Also mEos3.2-

HaloTag-HP1β wildtype, V23M and I161M mutant proteins were expressed at higher levels than endogenous HP1β protein levels (Figure 3-18C).

3.11.2. Diffusion of HP1 molecules in mESCs

mESCs stably expressing mEos3.2-HaloTag-HP1β and γ fusion proteins were subjected to live-cell SMLM as described in Section 2.9.3. mEos3.2 green fluorescence images confirmed nuclear localisation of HP1 fusion proteins in each cell line (Figure 3-19A). Clear differences were observed in the sub-nuclear localisation of HP1β and γ constructs. Whilst wildtype HP1β proteins form clear heterochromatic foci, the signal for both HP1β I161M and V23M was more diffuse, indicating that molecules were less tightly associated with chromatin (Figure 3-19A, left panel). HP1γ wildtype proteins also formed foci. However, the signal was more diffuse than that of HP1β wildtype. No foci formation was observed with HP1γ mutant proteins (Figure 3-19A, right panel).

In order to quantify diffusion of HP1 proteins, diffusion coefficients (DCs) of single molecules and the percentage of molecules with a specific diffusion mode were determined. Therefore, 2D imaging at 15ms time resolution with 67 frames per second (fps) was performed. Fluorescence images were collected as movies of 10,000 frames per cell and single-molecule traces were analysed using software that detects single-molecule trajectories from SPT movies (see Section 2.9.3). For each cell line, an average of 300 trajectories per movie were recorded. Up to 10 cells were imaged per cell line. Jump-distance analysis was performed as previously described[178] and showed that HP1 proteins have two major DCs (compare Section 1.2.9). The data represent an average of all molecule trajectories within a 2D plane, independent of their spatial distribution within the nucleus, as we did not distinguish between euchromatic and heterochromatic regions. As expected HP1 molecules exhibit one slow fraction (0.1 $\mu m^2 s^{-1}$) and one fast diffusing fraction (> 0.5 $\mu m^2 s^{-1}$) (Figure 3-19B, left panel). DCs of the slow fraction were similar across samples. As slow molecules show similar diffusion to that observed for histone H2B[200] they are regarded as stably bound to chromatin. The fast diffusing HP1β wildtype 50.2% fraction diffused at 0.6 ± 0.12 $\mu m^2 s^{-1}$, whereas both V23M (61.7%; 1.01 ± 0.34 $\mu m^2 s^{-1}$, p-value = 0.0016) and I161M (61.7%; 1.24 ± 0.16 $\mu m^2 s^{-1}$, p-value = 0.0001) moved significantly more quickly (Figure 3-19B/C).



C

	Diffusion coefficient [µm²s⁻¹]		Percentage [%]	
	Fast	Slow	Fast	Slow
HP1β WT	0.60 ± 0.12	0.09 ± 0.01	50.2	49.8 ± 3.6
HP1β V23M	1.02 ± 0.34	0.08 ± 0.02	61.7	38.3 ± 9.0
HP1β I161M	1.24 ± 0.16	0.10 ± 0.01	61.1	38.9 ± 14.4
HP1γ WT	0.78 ± 0.26	0.10 ± 0.01	70.2	29.7 ± 13.1
HP1γ V32M	0.74 ± 0.08	0.10 ± 0.001	80.3	19.7 ± 5.6
HP1γ R38/9A	0.74 ± 0.21	0.09 ± 0.02	67.3	32.7 ± 7.4
HP1γ I165K	1.50 ± 0.01	0.08 ± 0.02	72.1	27.8 ± 8.6

Figure 3-19 Diffusion of HP1 molecules in mESCs.

Stable mESC lines expressing the indicated mEos3.2-HaloTag-HP1β and γ proteins were labeled with 0.5nM HaloTag-JF_{549} ligand and subjected to live-cell SMLM. Using 561nm excitation, fluorescence images were collected as movies of 10,000 frames at 15ms exposure times. Approximately 10 cells were imaged per cell line.

A Representative fluorescent images of mESC nuclei expressing mEos3.2-HaloTag-HP1 proteins as indicated. Signals of mEos3.2 (em: 515nm) green fluorescence are shown.

B Diffusion coefficients of indicated mEos3.2-HaloTag- HP1β and γ molecules [$\mu m^2 s^{-1}$] are illustrated (left panel). Each data point represents an average of ~300 molecules tracked per cell. Percentages of molecules in fast and slow fraction are indicated (right panel). SEM are indicated. SEM of single cell data is indicated. Significance of diffusion data was assessed using Students T-test. Significance of percentages was assessed using two-way ANOVA tests (****: p-value =0.0001, **: p-value < 0.0023).

C Tabulated summary of results shown in **B**.

These results indicate that fast diffusion of HP1β molecules is affected when mutating the chromo and chromo-shadow domains. Compared to HP1β, HP1γ wildtype molecules diffused quicker (70.2%; 0.78±0.26 $\mu m^2 s^{-1}$, p-value = 0.0577). For HP1γ, 20% more molecules were detected in the fast diffusing fraction compared to HP1β (p-value = 0.0001), indicating that HP1γ molecules are less tightly associated with chromatin. Interestingly, diffusion of both HP1γ CD mutants was similar to that of wildtype molecules (p-value < 0.7), suggesting that fast diffusion of HP1γ is less dependent on their ability to bind H3K9me3. Similar to HP1β, HP1γ I165K mutant molecules moved significantly faster than wildtype molecules (1.5±0.01 $\mu m^2 s^{-1}$, p-value = 0.0001). The clear differences in diffusion of wildtype compared to chromo-shadow domain mutants for both HP1 paralogues, may indicate that HP1 moves within bigger complexes inside the nucleus.

In summary, single molecule particle tracking of mEos3.2-HaloTag-HP1β and γ fusion proteins revealed that HP1 molecules show two major diffusion modes, which are characteristic for HP1 dynamics. The slow DC refers to molecules stably bound to chromatin and the fast DC to freely diffusing molecules. HP1γ wildtype molecules in the fast diffusing fraction appear to move significantly quicker than HP1β wildtype molecules. Furthermore, diffusion of HP1β molecules seems to be more dependent on its ability to bind H3K9me3 than that of HP1γ molecules. However, mutating the PxVxL domain significantly increases fast diffusion of both paralogues.

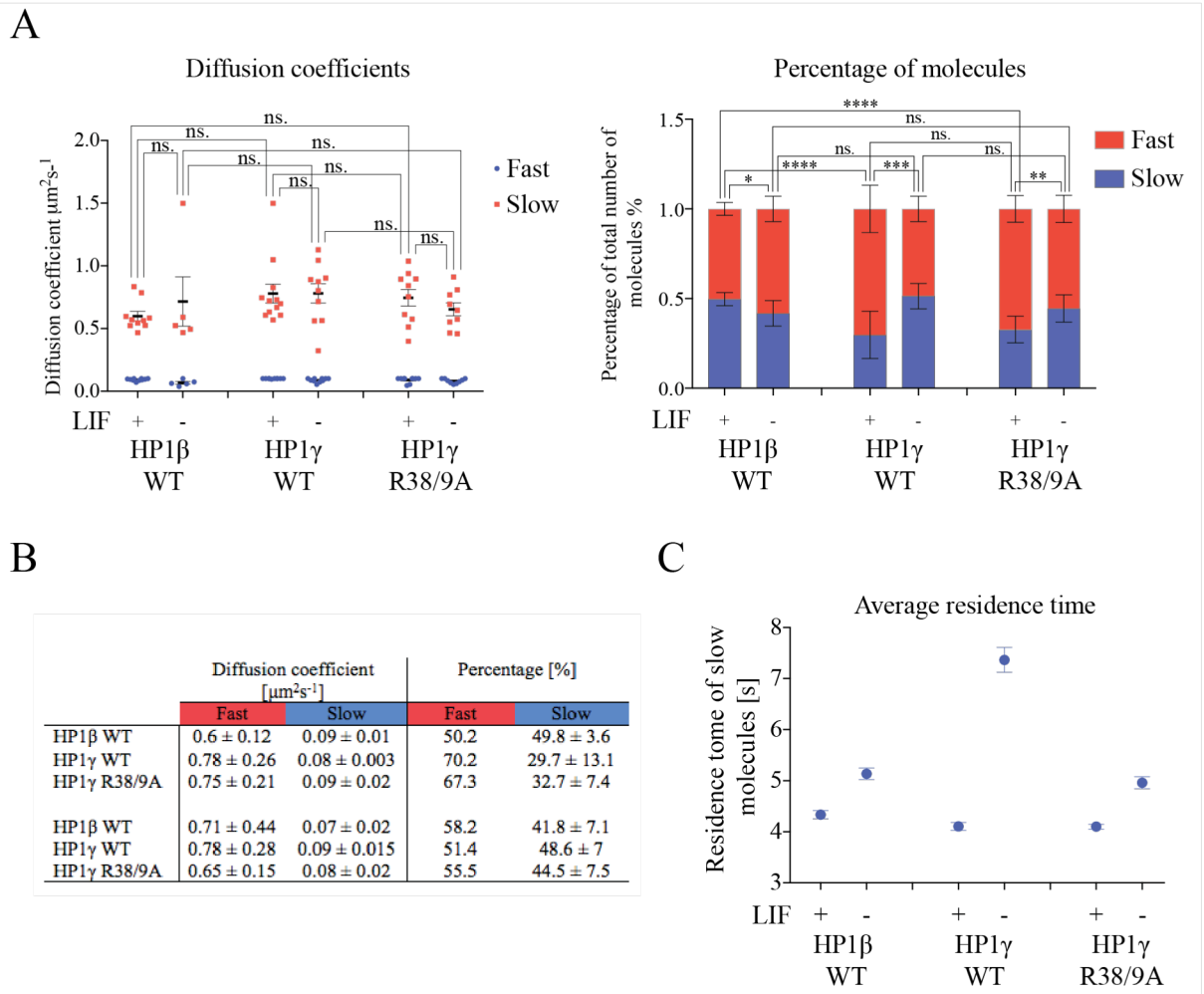
3.11.3. Dynamics of HP1 during mESC differentiation

I previously demonstrated that HP1γ R38/9A mutant proteins show impaired binding to H3K9me3 peptides *in vitro* (see Section 3.5) and provided evidence that HP1γ is citrullinated by PADI4 at residues R38/9 (see Section 3.9 and 3.10). During reprogramming PADI4 enzymes are activated and citrullinate linker histone H1.2, which results in their eviction from chromatin and global chromatin decondensation[190]. Conversely, stem cell differentiation is accompanied by facultative heterochromatin formation and global chromatin condensation, which involves the association of structural chromatin proteins such as histone H1 and HP1 with chromatin[176,201]. Given the independent functions of both citrullination and HP1 during stem cell homeostasis (compare Section 1.2.8.4), I decided to extend the analysis to characterise the dynamics of HP1γ R38/9A mutant compared to HP1γ and HP1β wildtype

proteins during mESCs differentiation. mES cell lines stably expressing mEos3.2-HaloTag-HP1β and γ fusion proteins (see Section 3.11.1) were subjected to live cell SMLM before and after cellular differentiation, which was induced by withdrawing LIF for 72h. The same protocol was previously shown to induce differentiation in mESCs at high efficiency[176].

Diffusion coefficients (DCs) of single molecules with the percentages in each fraction were determined (compare Section 3.11.2). Again, two major DCs were detected for HP1 molecules. In general, the diffusion coefficients of HP1 molecules remained unchanged upon mESC differentiation. However, percentages of the slow fraction of HP1γ both wildtype (48.6 ± 7.0%; 0.09 ± 0.014 $\mu m^2 s^{-1}$, p-value = 0.0004) and R38/9A mutant (44.5 ± 7.6%; 0.08 ± 0.017 $\mu m^2 s^{-1}$, p-value = 0.002) molecules increased upon differentiation by approximately 20% and 15%, respectively. Conversely, the percentages of slow HP1β even decreased by approximately 8% upon differentiation (41.8 ± 7.0 %; 0.07 ± 0.02 $\mu m^2 s^{-1}$, p-value = 0.0143) (Figure 3-20A/B). Whilst in undifferentiated cells both HP1γ wildtype (p-value = 0.0001) and R38/9A mutant (p-value = 0.0001) molecules showed significantly more molecules in the fast moving population when compared to HP1β. These differences decreased upon differentiation for both HP1γ wildtype (p-value = 0.174) and R38/9A mutant (p-value = 0.0619) molecules. These results indicate that HP1γ showed a striking redistribution towards the slow moving fraction upon differentiation, whereas percentage of HP1β slow moving molecules even decreased, which would suggest that HP1γ is more critical in establishing heterochromatin during differentiation in this system. However, a similar increase of stably bound molecules was observed for HP1γ R38/9A mutant molecules.

To measure the residence time of proteins on chromatin in differentiated versus undifferentiated cells, imaging was performed at 500ms time resolution with 2 fps. Fluorescence images were collected as movies of 1000 frames per cell and single-molecule traces were analysed using software that detects single-molecule trajectories from SPT movies (HP1γ). For each cell line, an average of ~200 trajectories per movie were recorded. Up to 10 cells were imaged per cell line. Figure 3-20C depicts the average residence times of HP1 molecules before and after differentiation. Single particle tracking at 500ms time resolution only detects movement of molecules bound to chromatin, as faster moving molecules cannot be detected in more than one frame. The average residence times therefore refer to molecules stably bound to chromatin.



B

	Diffusion coefficient [µm²s⁻¹]		Percentage [%]	
	Fast	Slow	Fast	Slow
HP1β WT	0.6 ± 0.12	0.09 ± 0.01	50.2	49.8 ± 3.6
HP1γ WT	0.78 ± 0.26	0.08 ± 0.003	70.2	29.7 ± 13.1
HP1γ R38/9A	0.75 ± 0.21	0.09 ± 0.02	67.3	32.7 ± 7.4
HP1β WT	0.71 ± 0.44	0.07 ± 0.02	58.2	41.8 ± 7.1
HP1γ WT	0.78 ± 0.28	0.09 ± 0.015	51.4	48.6 ± 7
HP1γ R38/9A	0.65 ± 0.15	0.08 ± 0.02	55.5	44.5 ± 7.5

Figure 3-20 Dynamics of HP1 upon mESC differentiation.

Stable mESC lines expressing the indicated mEos3.2-HaloTag-HP1β and γ proteins were differentiated by 72h LIF withdrawal. Differentiated and undifferentiated cells were labeled with 0.5nM HaloTag-JF_{549} ligand and subjected to live-cell SMLM. Using 561nm excitation, fluorescence images were collected as movies of 10,000 frames at 15ms for diffusion and 1000 frames at 500ms for residence data.

A Diffusion coefficients of indicated mEos3.2-HaloTag- HP1β and γ molecules [$\mu m^2 s^{-1}$] are illustrated (left panel). Each data point represents an average of ~300 molecules tracked per cell. Percentages of molecules in fast and slow fraction are indicated (right panel). Approximately 10 cells were imaged per cell line. SEM of single cell data is indicated. Significance of diffusion data was assessed using Students T-test. Significance of percentages was assessed using two-way ANOVA tests (****p-value = 0.0001, ***p-value = 0.0004, **p-value = 0.002, *p-value = 0.0143).

B Tabulated summary of results depicted in **A**.

C Average residence time of indicated mEos3.2-HaloTag- HP1β and γ molecules [s] are illustrated. Each data point represents an average of ~200 molecules tracked per cell. Approximately 10 cells were imaged per cell line. SEM of exponential fit are indicated.

In agreement with the increase of molecules in the slow moving population in differentiated cells (Figure 3-20A/B), HP1γ WT molecules show an increased average residence time (t_{res}) of approximately 3s upon differentiation (t_{res}[+LIF] = 4.10±0.05s; t_{res}[-LIF] = 7.37±0.24), while the residence time of HP1β molecules only increased slightly (t_{res}[+LIF] = 4.33±0.08s; t_{res}[-LIF] = 5.13±0.11s).

Intriguingly, although the percentage of immobilised HP1γ R38/9A mutant molecules increased in differentiated cells (Figure 3-20A/B), their average residence time did not change significantly (t_{res} [+LIF] = 4.1±0.07s; t_{res} [-LIF] = 4.96±0.11s). Hence, upon differentiation, the overall percentage of slow moving and chromatin associated HP1γ R38/9A molecules increases similar to WT molecules. However, the average time these molecules remain stably bound to chromatin is lower than that of WT HP1γ molecules.

In summary, upon differentiation the diffusion of fast and slow moving HP1β or γ molecules remains unchanged, however the percentage of slow moving molecules significantly increases for HP1γ (WT and R38/9A), but not for HP1β molecules. These results suggest that HP1γ's contribution to local and global chromatin condensation during differentiation is higher than that of HP1β. While the percentages of slow moving molecules of HP1γ WT and R38/9A are similar in differentiated cells, their average residence on chromatin is lower, which suggests that the residues R38 and R39 are not necessarily important for HP1γ's recruitment to chromatin during differentiation, but influence the stable binding of HP1γ molecules on chromatin.

3.12. Discussion

Mass spectrometry analysis in mESCs identified citrullination of HP1γ in the hinge region at residue R108 and in the CD at residues R38/9. My PhD project focused on citrullination of residues R38/9 since these were within a domain of well-characterised function. This chapter documents my efforts to characterise these modifications using *in vitro* and *in vivo* assays.

3.12.1. R38/9A mutant shows impaired binding to H3K9me3 peptides

Residues R38/9 are located within the β1-sheet of the CD (Figure 1-5). As they are on the opposite side and facing away from the H3K9me2/3 binding pocket they probably do not

directly contact the H3 tail. However, correct folding of the β1-sheet is important for H3 recognition[106] and losing the positive charge of two adjacent arginines due to citrullination, may affect CD structure and hence the affinity of HP1 to H3K9me2/3. To test this hypothesis, I performed *in vitro* pulldown assays with H3K9me3 peptides using recombinant HP1γ wildtype and R38/9A mutant proteins (see Section 3.5). The R38/9A mutant showed impaired binding to H3K9me3 peptides compared to wildtype, suggesting positive charges at residues R38/9 are important for HP1γ to efficiently bind methylated H3 tails. The fact that the R38/9K mutant restores binding indicates that positive charges at position 38 and 39 are critical for HP1γ's CD to interact with H3K9me3. When comparing different pulldown experiments, it became apparent that the efficiency varied from experiment to experiment. This might be due to experimental variants such as GST-preparations, peptide batches or temperature fluctuations in the lab. Despite this variability the results presented here were qualitatively highly reproducible.

The R38/9A mutant was designed to mimic the loss of charge accompanied with citrullination. Previous work from our lab showed that linker histone H1.2 R54A mutant phenocopies citrullinated R54 in *in vitro* nucleosome binding assays[190]. However, the effect of mutating two adjacent, relatively long arginine residues to short alanine tails may interfere with the protein structure of HP1γ more than citrullination of these residues. Here, nuclear magnetic resonance (NMR) spectroscopy of wildtype and mutant CDs could provide more insight into the impact of R38/9A mutations on the overall doain structure.

In conclusion, my results show that maintaining positive charges at residues R38/9 is important for HP1γ to bind H3K9me3 *in vitro* and suggest that citrullination may influence this interaction.

3.12.2. HP1γ R38/9-citrulline antibodies and potential applications

In order to characterise these modifications *in vivo*, I generated site-specific antibodies against HP1γ R38/9-citrulline. Antibodies were validated using *in vitro* assays (see Section 3.8 and 3.9) and finally tested on mESCs whole cell extracts (see Section 3.10). Out of four antibodies tested, S870D and S924D appeared the most specific for HP1γ R38/9-cit. *In vitro* citrullination assays revealed binding of S870D or S924D to R38/9A, but not to the R38/9K

mutant. This suggests that antibody recognition of its epitope is inhibited by positive charges at positions 38 and 39 within HP1γ's CD, although steric hindrance due to the lysine side-chain may also be involved in blocking antibody access to its epitope. Most importantly though, the signal on wildtype HP1γ notably increased upon addition of PADI4, confirming that citrullination of R38/9 promotes epitope recognition. Indeed, recognition of the citrullinated form of HP1γ is even stronger than the HP1γ R38/9A mutant indicating that both antibodies prefer citrulline at these positions for full binding activity.

The S870D antibody detected an increase in HP1γ R38/9-cit in mESCs upon overexpression of PADI4, suggesting that the modification can exist *in vivo*. To verify that the signal is specific for HP1γ, it is crucial to test the antibody on lysates obtained from cells depleted from endogenous HP1γ in the presence or absence of HA-PADI4 over-expression. Signal reduction in HP1γ knockdown cells over-expressing PADI4 would confirm that S780D specifically recognises PADI4 dependent citrullination of HP1γ at residues R38/9.

The signal for HP1γ R38/9-cit was weak and was only detected upon PADI4 overexpression indicating that the overall levels of this modification are low in untreated mESCs. To gain further insight into the role of HP1γ citrullination it will be very important to determine where citrullinated HP1γ is located in the cell. Immunostainings or cellular fractionations analysed by immunoblotting would be useful in this regard.

Besides HP1, other chromatin associated proteins have been described to be citrullinated[202,203]. Hormone induced nuclear receptors (NRs) regulate expression of target genes by association with coactivator proteins to remodel chromatin structure and assemble the transcription machinery[204]. The important coactivators of NRs, histone acetyltransferases CBP/p300[202], have been shown to be regulated by citrullination. PADI4 citrullinates p300 on R2142, which in turn enables the protein to interact with the coactivator complex that activates estrogen receptor-mediated transcription[205]. Similarly, citrullination of HP1γ could generate new binding sites for proteins affecting HP1γ localisation or function. Since HP1γ seems to be the only paralogue which is part of NR complexes[139,140], it would be very interesting to test whether this interaction is regulated by citrullination. Furthermore, *in vitro* pulldown assays with unmodified or R38/9-cit modified HP1γ peptides using mESC whole cell lysates, analysed by mass spectrometry could be a way of identifying readers of these

modifications. The HP1γ R38/9-cit antibody would be a valuable tool in helping to dissect the role of potential interaction partners.

Citrullinated H1.2 is evicted from chromatin in mESCs[190]. My *in vitro* data indicate that binding of HP1γ R38/9A to H3K9me3 is significantly reduced compared to WT HP1γ suggesting HP1γ's binding to chromatin may also be regulated by citrullination. In mESCs, the amount of citrullinated HP1γ seems to be low, suggesting that citrullination only affects a small subset of HP1γ molecules per cell. ChIP experiments using the HP1γ R38/9-cit antibody would be a crucial experiment to determine whether citrullination affects HP1γ's binding to chromatin and, if so, which genomic loci are affected. To gain more insight into the role of PADI4 mediated HP1γ citrullination, these data could be cross-compared with ChIP data from cells overexpressing PADI4, using HP1γ and PADI4 antibodies. An interesting starting point would be to study the potential interplay of HP1γ and PADI4 in regulating expression of the stem cell gene *Nanog*. In differentiated cells, HP1γ binds to the *Nanog* promoter via H3K9me3 and represses gene expression. Furthermore, loss of HP1γ increases reprogramming efficiency by de-repressing *Nanog* expression, indicating a role for HP1γ in regulating cell fate decisions[172]. During reprogramming, expression of PADI4 is induced, which leads to activation of stem cell genes, including *Nanog*[190]. It would therefore be very interesting to understand whether HP1γ's displacement from the *Nanog* promoter during reprogramming is regulated by PADI4.

3.12.3. Single molecule localisation microscopy of HP1

SMLM was used to determine the dynamics of HP1β and γ wildtype and mutant proteins in mESCs. For that purpose, stable cell lines expressing mEos3.2-HaloTag-HP1 proteins were generated. All constructs were expressed at levels higher than those of the respective endogenous proteins (Figure 3-18B). However, we deliberately chose cells with lower expression to decrease the likelihood of overexpression artefacts (Figure 3-19A). Furthermore, the analysis to determine diffusion parameters for single molecules is corrected for the overall fluorescence intensity. Therefore, individual expression levels of fusion proteins should not influence the measurements dramatically.

SPT of mEos3.2-HaloTag-HP1β and γ fusion proteins detected two characteristic diffusion modes of HP1 molecules, which confirms results from previous studies using complementary technologies (FRAP and FCS)[109,173-175]. FCS is based on the observation of single molecules or several molecules within one diffraction-limited spot. Also FRAP measures kinetics of molecule diffusion within one region of interest. Data acquisition is therefore limited to single regions within a nucleus. On the other hand, 2D SMLM allows one to detect every molecule within a 2D plane within a nucleus. SPT is a much more quantitative technique, as diffusion of hundreds of molecules can be determined simultaneously.

The two major diffusion modes of HP1 relate to molecules that are either stably bound to chromatin (slow) or freely diffusing molecules (fast). In contrast to published data stating that HP1 paralogues share similar dynamics[109,168,174], we observed clear differences between HP1β and γ. HP1γ wildtype molecules in the fast diffusing fraction appear to move significantly quicker than HP1β wildtype molecules. Furthermore, diffusion of HP1β molecules seems to be more dependent on its ability to bind H3K9me3 than that of HP1γ molecules. This observation is interesting in the sense that HP1γ is the only paralogue predominantly found in H3K9me3 deprived regions of the genome[143]. Mutations within the CD were previously shown to result in complete loss of HP1 from heterochromatin, as mutations completely abolish heterochromatic foci formation[109]. Conversely, our data suggest that a significant proportion of CD mutant molecules is still stably bound to chromatin. It is likely that HP1 recruitment to chromatin is not entirely regulated by its ability to bind H3K9me3 and/or that the molecules we detect interact with chromatin via their CSD or hinge region. Mutating the PxVxL domain significantly increases fast diffusion of both HP1β and HP1γ, which suggests that HP1 molecules move in dimers or even in bigger complexes. The mechanisms of HP1 recruitment to chromatin are largely unknown. Given HP1's multipartite nature, recruitment of HP1 via association with larger protein complexes seems highly likely.

Both HP1γ[170-172] and PADI4[ref.190] are associated with stem cell homeostasis. To understand whether the residues R38 and R39 of HP1γ influence the proteins dynamics during stem cell differentiation, we compared the dynamics of HP1γ R38/9A mutant to WT HP1γ and HP1β molecules in mESCs before and after LIF withdrawal. The first interesting observation from this experiment is that, upon differentiation the percentage of slow moving molecules significantly increases for HP1γ (WT and R38/9A), but not for HP1β molecules. This

suggests that, at least in our system, HP1γ's contribution to chromatin remodeling during differentiation is greater than that of HP1β. Conversely, data derived from FRAP experiments reported no apparent difference between HP1 paralogues[169,176]. SPT is a more quantitative method and more sensitive than FRAP. It is possible that functional differences between HP1 paralogues have not been detected yet, due to limitations of other technologies.

Secondly, we observed that although the percentages of slow moving molecules of HP1γ WT and R38/9A are similar in differentiated cells, their average residence time on chromatin is lower. These data suggest that residues R38 and R39 are not necessarily critical for HP1γ's recruitment to chromatin during differentiation, but they influence the stability of HP1γ binding on chromatin. This hypothesis is supported by *in vitro* data, which showed that interaction of HP1γ R38/9A with methylated H3K9 peptides is less stable than that of WT proteins. Binding of HP1γ R38/9A is significantly impaired under physiological salt conditions and seems to be more sensitive to increasing salt concentrations when compared to WT proteins, indicating that the interaction of the mutant protein with H3K9me3 peptides is weaker.

The diffusion data presented here represent an average over all molecule trajectories, independent of their spatial distribution within the nucleus. Therefore, our analyses do not distinguish between euchromatic and heterochromatic regions. Data derived from FRAP and FCS experiments suggest differential dynamics of HP1 in heterochromatin and euchromatin. To understand the plasticity of HP1γ and the R38/9A mutant during differentiation, it will be crucial to determine the diffusion and residence time of molecules in euchromatin versus heterochromatin. It is also possible to determine association rates of single molecules in different regions of the nucleus. This work is currently ongoing. In the future, we are planning to analyse DNA residence time data of HP1 in 3D to understand how the bound molecules are moving. In particular, it will be possible to determine whether molecules are sliding along the DNA or are bound to specific sites. This information will be very valuable to understand how HP1 molecules are interacting with chromatin and which domains or residues are necessary.

In this chapter, I presented my results characterising a novel post-translational modification of HP1γ. I showed that HP1γ is citrullinated by PADI4 *in vitro* and *in vivo* at residues R38/9. The mutant R38/9A, which is designed to mimic the loss of charge accompanied with

citrullination, affects HP1γ's binding to H3K9me3 peptides and reduces the residence time of HP1γ molecules on chromatin in differentiated cells. This suggests a role for citrullination in regulating HP1γ binding to chromatin during differentiation.

Chapter 4

Functional characterisation of HP1γ in breast cancer cells

4.1. Introduction

Gene regulation is tightly controlled at different levels to ensure correct transcriptional programming of a cell. Deregulation of these processes may result in many phenotypic changes including aberrant cell growth and oncogenesis. It is well documented that cancer progression is a multistep process involving many changes in chromatin structure and gene expression[165]. Given their widespread role of telomere and centromere stability and gene regulation, it is not surprising that HP1 proteins have links to cancer progression[160,206]. High protein levels of HP1γ have been associated with enhanced cell proliferation and oncogenesis in various cancers such as colorectal[155,156], non-small cell lung[157], prostate[158], cervical and breast[159,160] (compare Section 1.2.8.3). However, the involvement of HP1γ in tumorigenesis remains unclear.

In this chapter, I will present data on the physiological role of HP1γ in breast cancer cells. I have taken advantage of two breast cancer cell line models, one with less invasive characteristics (MCF7) and one with a more invasive phenotype (MDA-MB-231). This chapter describes the changes to the phenotype of these cells upon HP1γ depletion using assays that assess cell growth, cell cycle progression, cell viability and motility.

4.1.1. The MCF7 breast cancer cell model

The MCF7 (Michigan cancer foundation 7) cell line is a widely used human breast cancer cell line generated from a 69- year old Caucasian woman in 1973[ref.207] with an invasive ductal breast carcinoma. The cells express estrogen receptor α (ERα) and progesterone receptor (PR), and are receptor tyrosine kinase ERBB2 (HER2) negative. Therefore, they are regarded as a model for luminal A type breast cancer. Luminal A type breast cancer is the most common type of breast cancer with the best prognosis[208].

MCF7 cells are adherent cells that have retained characteristics of the mammary epithelium, and they therefore display an epithelial-like morphology. In culture, MCF7 cells grow as a monolayer, are 20-25μm in diameter (Figure 4-1) and have a doubling time of 30-40h.

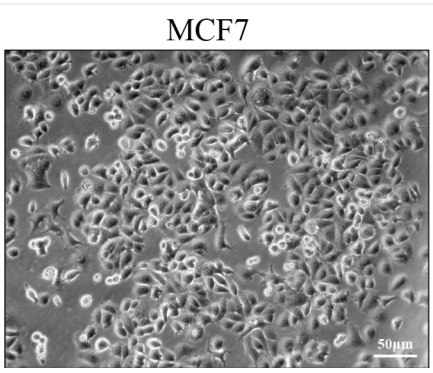


Figure 4-1 Morphology of MCF7 cells.
Brightfield microscopy image of adherent MCF7 cells at 80% confluency, captured with a Leica EC3 digital camera at 10x magnification. Scale bar indicated (50μm).

4.1.2. The MDA-MB-231 breast cancer cell model

The MDA-MB-231 cell line is a human breast cancer cell line generated from a 51- year old Caucasian woman in 1973[209] with a poorly differentiated invasive ductal breast carcinoma. This cell line is an example of triple negative breast cancer due to lack of the molecular markers ERα, PR and HER2. It further lacks expression of tight junction proteins including claudin 3 and E-cadherin, but it shows high expression of mesenchymal markers and therefore classifies as a model for the Claudin-low (CL) breast cancer subtype[210]. CL tumors are poorly differentiated and are frequently found in residual mammary tumour tissue after hormone therapy or chemotherapy, and consequently patients overall have a poor prognosis[211].

MDA-MB-231 cells are adherent cells that form loose networks of round or short, spindle-shaped cells, they are 15-20µm in diameter (Figure 4-2) and have a doubling time of 35-40h. The cells are used as model for chemotherapy resistant metastatic breast cancer cells.

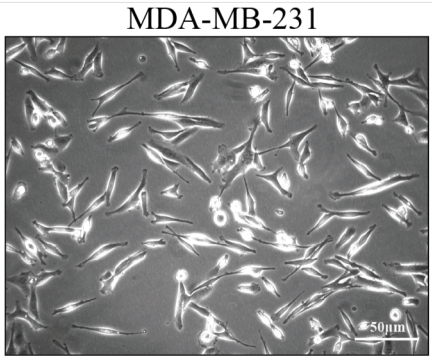


Figure 4-2 Morphology of MDA-MB-231 cells.
Brightfield microscopy image of adherent MDA-MB-231 cells at 60% confluency, captured with a Leica EC3 digital camera at 10x magnification. Scale is bar indicated (50µm).

4.2. Aims of this chapter

This chapter has the following aims related to characterising the phenotype of HP1γ depletion in the two breast cancer models MCF7 and MDA-MB-231.

1. To determine the effects of HP1γ depletion on cell proliferation and cell cycle progression
2. To determine the effects of HP1γ depletion on cell viability
3. To investigate whether loss of HP1γ affects cellular mobility

4.3. Phenotype changes upon HP1γ depletion in MCF7 cells

4.3.1. Expression of HP1γ in MCF7 cells

HP1γ is a ubiquitously expressed nuclear protein associated with chromatin. In order to determine HP1γ protein levels and subcellular localisation in MCF7 cells, cellular fractionation using a 0.5% (v/v) NP-40 lysis buffer (IPH buffer) was performed to separate chromatin from the nucleo- and cytoplasmic fraction. Cellular equivalents from each fraction were resolved by SDS PAGE and analysed by immunoblotting using antibodies against HP1γ, the cytoplasmic marker GAPDH and the chromatin marker histone H3 (Figure 4-3). GAPDH was absent from the chromatin fraction and only a small amount of H3 was detected in the nucleo-cytoplasmic fraction, indicating cellular fractionations were performed successfully. As expected, the strongest signal for HP1γ was observed in the chromatin fraction. However, when compared to total HP1γ as shown in the whole cell lysate, approximately 33%, of the protein was detected in the nucleo- and cytoplasmic fraction. My results suggest that, in MCF7 cells, the majority of HP1γ is associated with chromatin but a significant fraction is also present within the nucleoplasm and/or cytoplasm.

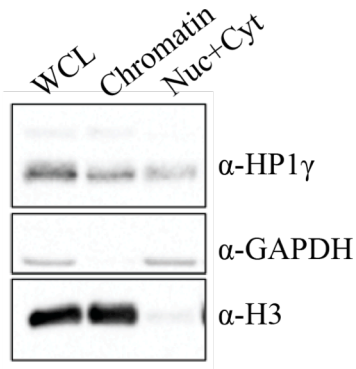


Figure 4-3 Cellular fractionation of MCF7 cells.
Immunoblotting of MCF7 lysates using indicated antibodies. MCF7 cells were lysed directly in 2x Laemmli buffer to prepare whole cell lysates (WCL). Cellular fractionation was performed using IPH chromatin extraction buffer. Nuc+Cyt: nucleo- and cytoplasmic fraction. Cellular equivalents (representing 40000 cells) of each fraction were loaded per well.

4.3.2. shRNA mediated knockdown of HP1γ

To analyse HP1γ function in MCF7 cells, I generated stable knockdown cell lines using lentiviral vectors that express small hairpin RNAs (shRNAs) targeting the CDS of the human HP1γ gene *CBX3*. shRNAs against HP1γ and a non-targeting control shRNA were cloned into lentiviral vector pLKO.1 and lentiviral particles were produced in 293T cells using the 3rd generation packaging system[212]. MCF7 cells were then infected with lentiviral supernatant. 24h post infection cells were cultured in puromycin selection media. HP1γ knockdown efficiency was determined by measuring HP1γ mRNA and protein levels for two shRNAs: shHP1γ_1 and 2 versus the control shCtrl over a 6 day time course. HP1γ mRNA levels were measured by RT-PCR and protein level by immunoblotting using a HP1γ antibody (Figure 4-4). Protein levels were quantified using ImageJ and signal intensities normalised to the loading control GAPDH. 2 days post infection, a 95% reduction of HP1γ mRNA was observed with shHP1γ_1 and a 50% reduction with shHP1γ_2, as compared to shCtrl control levels (Figure 4-4A). At the same time point, protein levels were reduced by 75% with shHP1γ_1 and remained unchanged with shHP1γ_2. After 6 days, the most efficient knockdown of HP1γ protein was observed for both shRNAs. shHP1γ_1 resulted in nearly complete protein depletion and shHP1γ_2 reduced the protein level by 80% (Figure 4-4B). Based on these results, day 4 and 6 time points were chosen for functional assays.

4.3.3. HP1γ depletion affects proliferation of MCF7 cells

Previous studies using different cancer cell lines showed that loss of HP1γ inhibits tumour cell growth[155,159]. To understand whether loss of HP1γ affects cell growth of MCF7 cells, I performed shRNA-mediated knockdown of HP1γ and assessed cell proliferation using the CyQUANT® cell proliferation assay. The CyQUANT® assay uses a fluorescence dye that intercalates into DNA and RNA. The fluorescence signal is directly proportional to the number of cells and is therefore used as an estimate for cell proliferation. 4 days after infection with lentiviral vectors encoding shRNAs (shCtrl, shHP1γ_1 and shHP1γ_2), cells were seeded at equal densities into wells of a 96-well plate ($t = d_0$) and fluorescence was measured every 2 days for a total of 8 days. HP1γ protein knockdown at d_0 was confirmed by immunoblotting using a HP1γ antibody (Figure 4-5A).

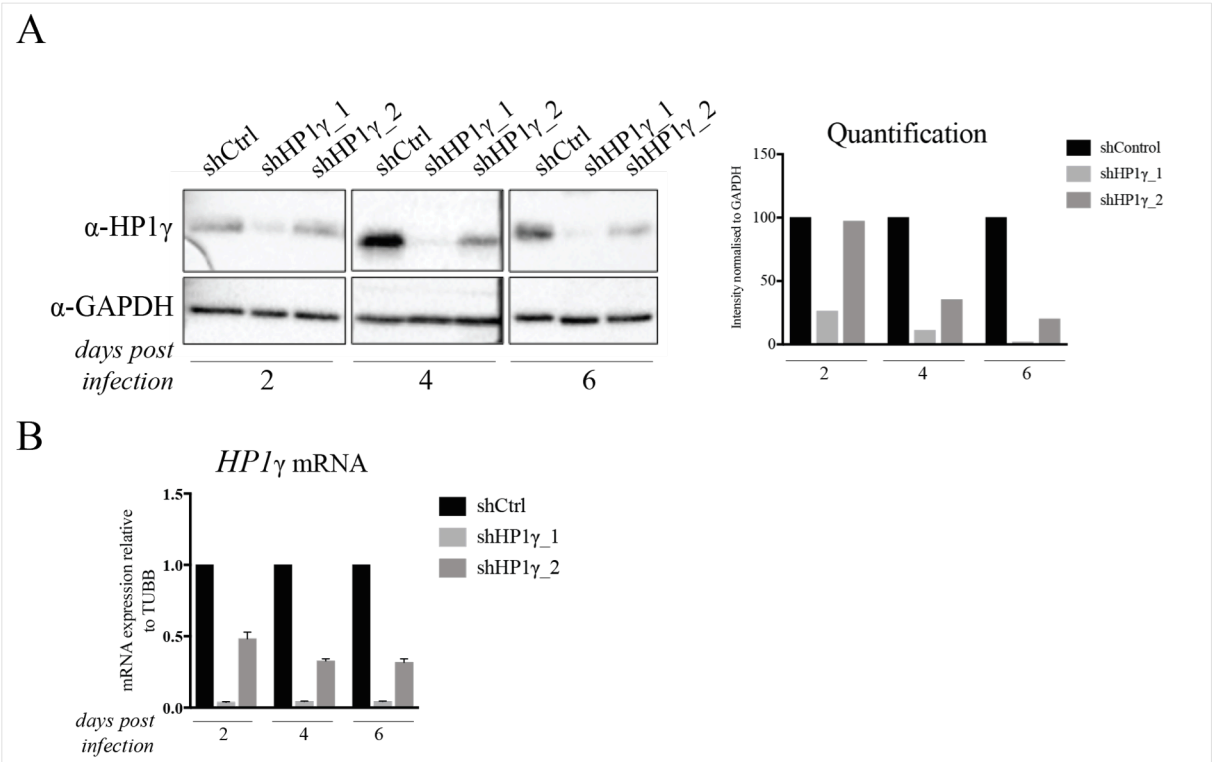


Figure 4-4 Time course for knockdown of HP1γ in MCF7 cells.
MCF7 cells were infected with shCtrl and shHP1γ_1 and 2 shRNA vectors and processed 2, 4 and 6 days post infection.
A Immunoblot of MCF7 whole cell lysates, using indicated antibodies (left panel). Signal was quantified using ImageJ and normalised to GAPDH (right panel). Protein lysates representing 40000 cells were loaded per well.
B HP1γ mRNA level assessed by RT-PCR relative to shCtrl samples. Graph shows relative mRNA expression after normalisation to *TUBB* expression. The fold change over shCtrl samples is shown. An average of duplicates is shown. SEM are indicated.

With both shRNAs the fluorescence signal was significantly lower compared to shCtrl cells, indicating cell proliferation was impaired upon HP1γ depletion (Figure 4-5B). To independently confirm the CyQUANT® assay results, I performed a single counting experiment for one shRNA (shHP1γ_1) (Figure 4-5C). Cells infected with shCtrl vector grew exponentially, whereas HP1γ depleted cells merely doubled over the time course of 8 days. Hence, a cell counting assay confirmed the results obtained from CyQUANT® assays. To further confirm that this effect is dependent on HP1γ protein levels, a rescue experiment was performed. MCF7 cell lines expressing HA-HP1γ or an empty control construct were generated using the lentiviral vector pHIV-Zsgreen, which expresses HA-HP1γ and GFP linked via an IRES sequence. GFP positive cells were sorted 5 days post infection and propagated to generate a cell line. HA-HP1γ overexpressing and control cells were infected

with shCtrl and shHP1γ_1 vectors and 4 days after infection cells were subjected to one CyQUANT® proliferation assay as described above. Expression of endogenous and exogenous (HA) HP1c protein levels were confirmed by immunoblotting (Figure 4-5D). Preliminary results showed that proliferation of MCF7 was partially rescued in cells expressing HA-HP1γ (Figure 4-5D, green line).

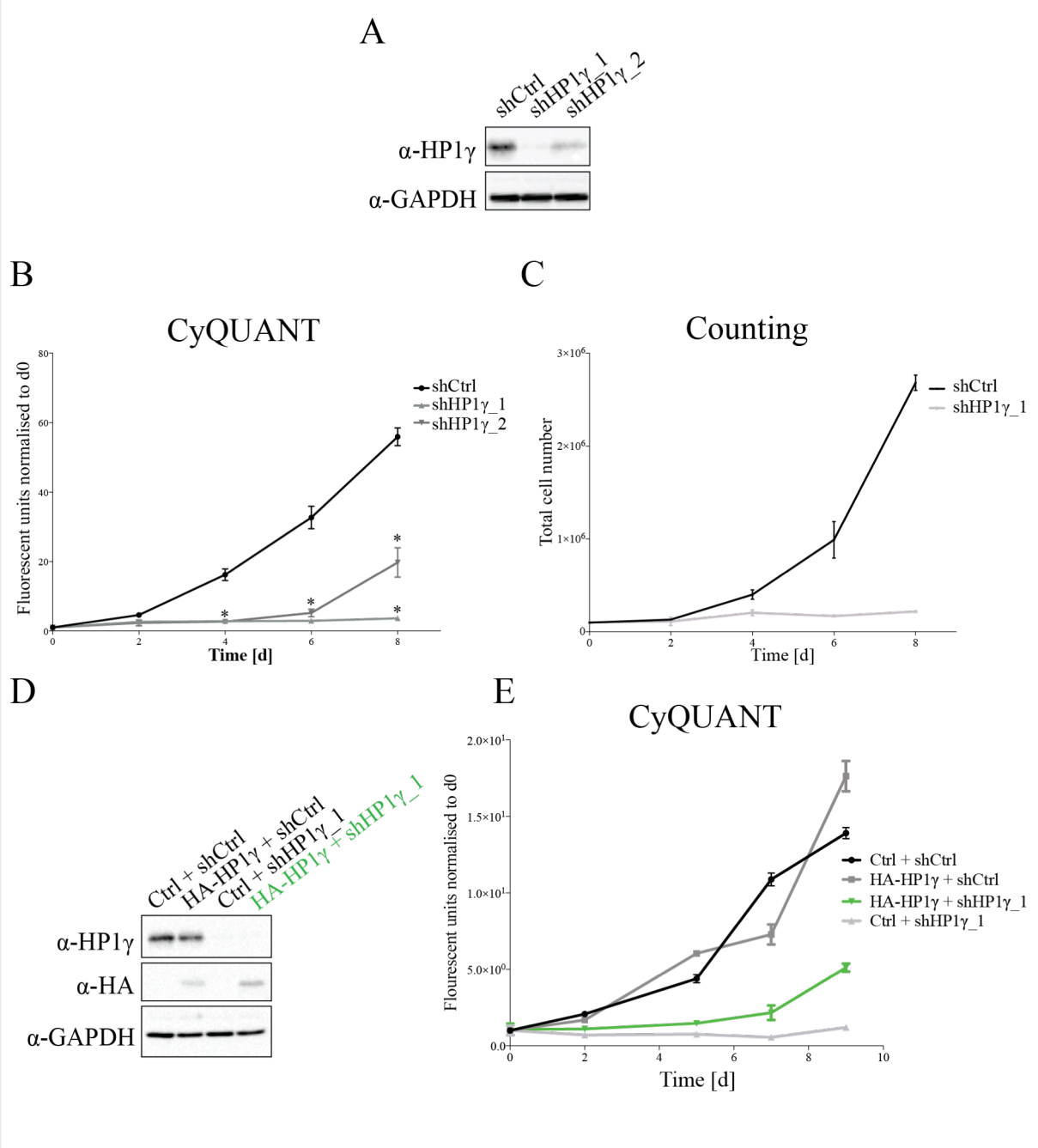


Figure 4-5 legend is located on the following page.

Figure 4-5 Proliferation of MCF7 cells.
MCF7 cells were infected with shCtrl and shHP1γ_1 and 2 shRNA vectors and processed 4 days post infection (t = d0).

A Immunoblot of MCF7 whole cell lysates at d_0 using indicated antibodies. Protein lysates from 40000 cells loaded per well.

B CyQUANT® proliferation assays over an 8 day time course. Cells were seeded at 500 cells/well in a 96 well plate at d_0. Fluorescence was measured at 520 nm in a Clariostar® microplate reader at the indicated time points. Fluorescent units were normalised to signals measured at d_0. Graph includes results from 3 independent experiments, each performed with 12 wells per condition. Significance was assesses using one-way ANOVA test (* p-value = 0.0001). SEMs are indicated.

C Cell proliferation assessed by cell counting. 1.0*10^5 cells were seeded in wells of a 12-well dish at d_0. Cells were trypsinsed and counted over the course of 8 days. Graph includes data from one experiment with 2 replicates per condition. SEMs are indicated.

D Immunoblot of MCF7 stably expressing HA-msHP1γ or empty vector (Ctrl) whole cell lysates at d_0 using indicated antibodies. Protein lysates from 40000 cells loaded per well.

E CyQUANT® proliferation assays over an 8 day time course. Cells were seeded at 500 cells/well in a 96 well plate at d_0. Fluorescence was measured at 520 nm in a Clariostar® microplate reader at the indicated time points. Fluorescent units were normalised to signals measured at d_0. Graph includes results from 1 experiment performed with 12 wells per condition. SEMs of technical replicates are indicated.

To test whether impaired proliferation upon HP1γ depletion is a general cellular response, or more specific for cancer cells, I assessed cell proliferation in the non-tumorigenic breast epithelial cell line MCF10A - a model cell line for normal mammary epithelial cells. 4 days post infection with shRNA vectors (t = d_0), HP1γ protein knockdown was confirmed by immunoblotting (Figure 4-6A). CyQUANT® results summarised in Figure 4-6B indicated that proliferation of MCF10A was not affected in HP1γ depleted cells. For both shRNAs, the measured fluorescent signals were similar to those of the control.

In summary, loss of HP1γ significantly affects MCF7 cell proliferation, whereas proliferation of the non-tumorigenic mammary epithelial cell line MCF10A was unaffected. My results indicate that proliferation of breast cancer cells is dependent upon HP1γ, whereas proliferation of normal breast epithelia cells is not.

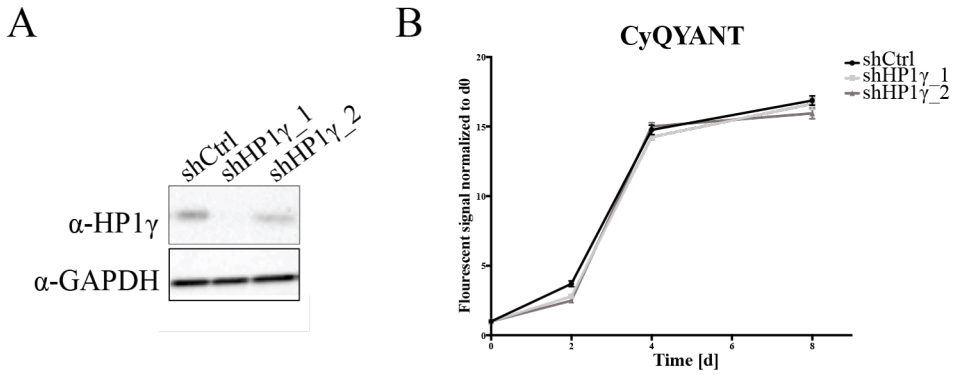


Figure 4-6 Proliferation of MCF10A cells.
MCF10A cells were infected with shCtrl and shHP1γ_1 and 2 shRNA vectors and processed 4 days post infection ($t = d_0$).

A Immunoblot of MCF10A whole cell lysates at d_0 using indicated antibodies. Protein lysates from 30000 cells were loaded per well.

B CyQUANT® proliferation assays over an 8 day time course. Cells were seeded at 500 cells/well in a 96 well plate at d_0. Fluorescence was measured at 520 nm in Clariostar® microplate reader at the indicated time points. Fluorescent units were normalised to signals measured at d_0. Graph includes results from 2 independent experiments. SEMs are indicated.

4.3.4. Cell cycle analysis of MCF7 cells

To help understand whether the impaired cell proliferation observed in HP1γ depleted MCF7 cells is due to cell cycle progression defects, I performed cell cycle analysis in HP1γ knockdown cells. Knockdown was confirmed by immunoblotting (Figure 4-7A). Data from studies in colorectal cancer cells suggest that HP1γ regulates expression of the cyclin-dependent kinase inhibitor p21 by directly repressing the promoter activity of the *CDKN1A* gene[155]. Consequently, HP1γ depletion causes an increase of p21 expression in colorectal cancer cells, which then accumulate in G1 phase of the cell cycle[155].

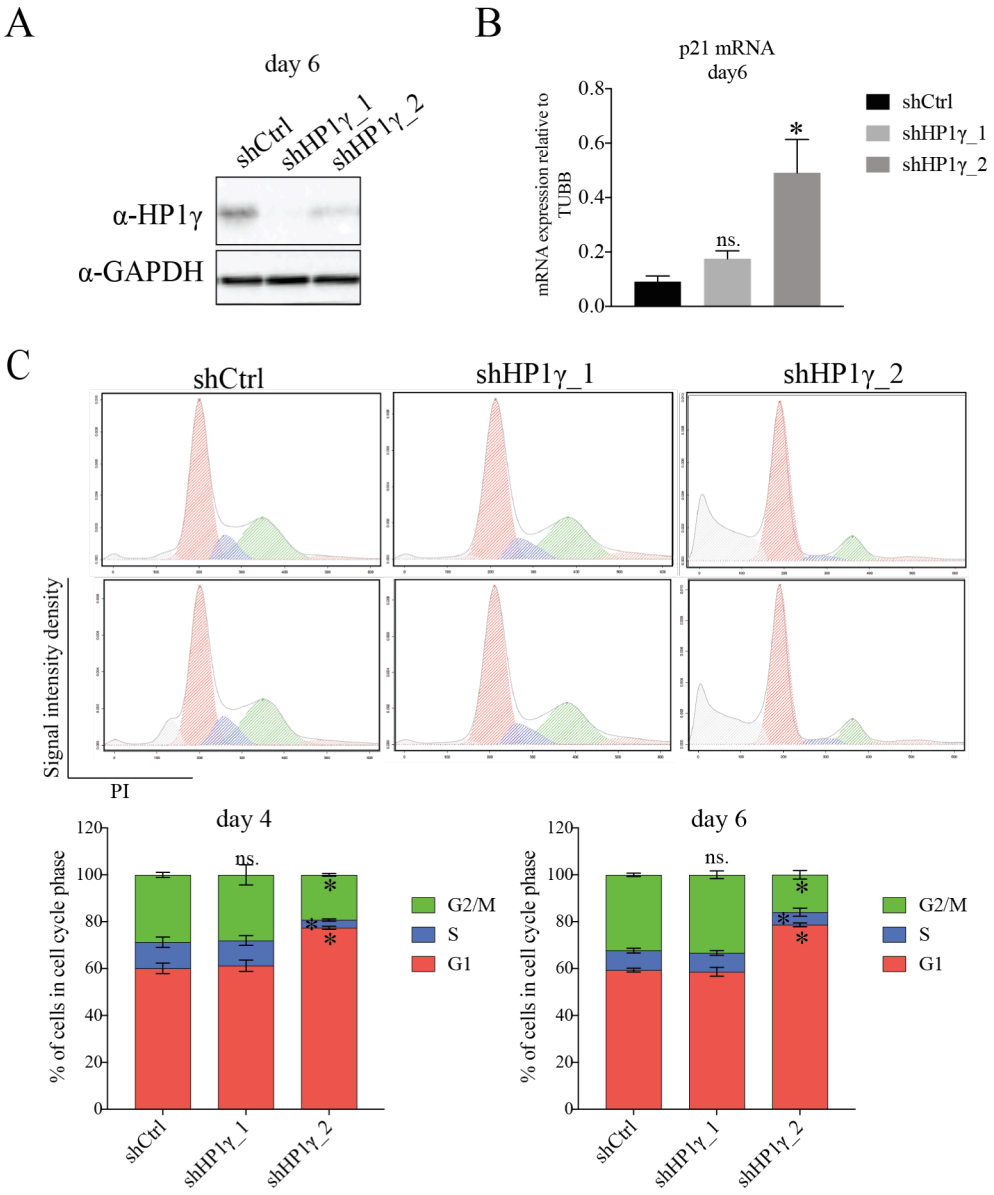


Figure 4-7 Cell cycle analysis of MCF7 cells.

MCF7 cells were infected with shCtrl and shHP1γ_1 and 2 shRNA lentiviral vectors and processed 4 or 6 days post infection.

A Immunoblot of MCF7 whole cell lysates at day 6 using indicated antibodies. Protein lysates from 40000 cells were loaded per well.

B mRNA level of CDKN1A encoding p21 at 6 days post infection, assessed by RT-PCR. Graph contains data from 3 experiments. Data normalised to TUBB. Significance was assessed using one-way ANOVA tests (* p-value <0.002; ns. p-value <0.095). SEM are indicated

C Cell cycle analysis of MCF7 cells by flow cytometry. Unsynchronised cells were stained with propidium iodide (PI) to assess DNA content. A total of 30000 cells were recorded per sample. Representative image of histograms are shown; curves were fitted using a custom R script (top panel). Percentages of cells in each cell cycle phase are shown (bottom panel). Plots contain results from 2 independent experiments performed in triplicate. Significance was assessed using one-way ANOVA tests (* p-value <0.008, ns. p-value <0.882). SEM are indicated.

I assessed p21 mRNA level by RT-PCR in MCF7 cells 6 days after infection with HP1γ shRNA vectors. For both shRNAs, an increase in p21 expression was observed compared to shCtrl cells (Figure 4-7B). However, the results varied significantly between the two shRNAs. Notably, the most potent shRNA, shHP1γ_1, which results in almost complete depletion of HP1γ protein (Figure 4-7A), resulted in less than a 2-fold increase in p21 mRNA compared to shCtrl, whereas shHP1γ_2 that less efficiently depletes of HP1γ protein (Figure 4-7A), increased p21 mRNA levels approximately 4-fold compared to shCtrl cells. These results indicate that *p21* mRNA levels are increased in HP1γ knockdown cells, however, *p21* expression does not directly correlate with HP1γ protein levels. To analyse cell cycle progression, unsynchronised MCF7 cells were stained with propidium iodide and DNA content per cell was assessed by flow cytometry. Cell cycle analysis was performed 4 and 6 days after infection with shRNA vectors. Results are summarised in Figure 4-7C. The results differed significantly between HP1γ shRNAs. At both time points, shHP1γ_1 showed no effect on cell cycle progression (p-value >0.882), whereas shHP1γ_2 significantly increased the number of cells in G1 phase by approximately 20% (p-value <0.002) with a concomitant decrease of cells in G2/M and S phase.

Cell cycle analysis showed that shHP1γ_2 treated cells, which express higher levels of p21, arrest in the G1 phase of the cell cycle, whilst shHP1γ_1 treated cells remain unaffected. However, effects on cell cycle could not be directly correlated with HP1γ protein levels, as cells treated with the more effective shRNA cycled normally.

4.3.5. Morphology of HP1γ depleted MCF7 cells

MCF7 cells undergo morphological changes upon HP1γ knockdown. Representative brightfield images of cells 6 days after infection with shRNA vectors are depicted in Figure 4-8. ShCtrl cells showed normal MCF7 morphology (compare to Figure 4.1), whereas both knockdown cell lines showed signs of apoptosis such as membrane blebbing, cell condensation and detachment (highlighted by red arrows), indicating cell viability was affected in HP1γ depleted cells. However, morphological differences were evident between cells infected with shHP1γ_1 or shHP1γ_2 vectors. Cells infected with shHP1γ_1 were rounded (middle panel), whereas shHP1γ_2 cells were flat and showed condensation of cellular organelles (bottom panel).

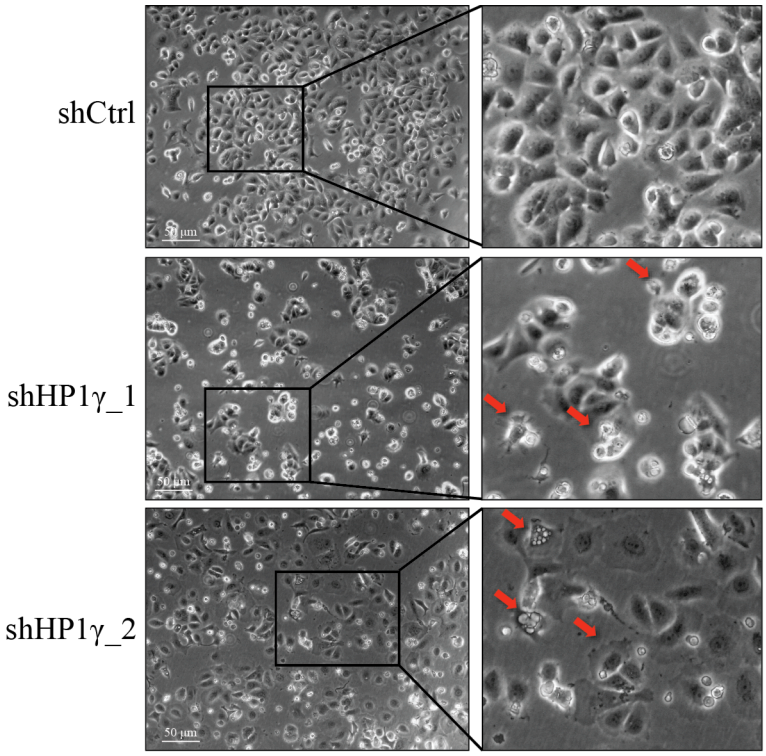


Figure 4-8 Morphology of HP1γ depleted MCF7 cells.
Brightfield microscopy images of MCF7 cells 6 days post infection with lentiviral shRNA vectors as indicated. Images were captured with a Leica EC3 digital camera at 10x magnification. Scale bar indicated (50 µm). Regions of interest are 3-fold magnified and red arrows highlight cell aberrations.

4.3.6. HP1γ affects cell viability of MCF7 cells

To measure viability of MCF7 cells upon HP1γ depletion, I performed annexin V/propidium iodide (PI) staining, analysed by flow cytometry. Annexin V is a phospholipid-binding protein with high affinity for the phospholipid phosphatidylserine (PS). In viable cells PS is located in the inner leaflet of the plasma membrane. Upon early stages of apoptosis, PS translocates to the outer leaflet of the plasma membrane, where it can be detected by annexin V coupled to a fluorophor (here FITC was used)[213]. Whilst annexin V is used as a marker for apoptotic cells, PI staining distinguishes between viable and dead cells. Assay conditions were optimised using MCF7 cells treated with 5 mM hydrogen peroxide (H_2O_2) as a positive control.

The excessive exposure of cells to reactive oxygen species (ROS) is well known to induce apoptotic pathways[214]. Flow cytometry results for annexin/PI stainings are summarised in Figure 4-9A. Gatings were adjusted manually using control (shCtrl) cells as a reference. A representative plot showing cells detected in FITC and PI channels, for H_2O_2 treated versus untreated MCF7 cells is depicted in Figure 4-9. As expected, H_2O_2 treatment clearly increased the number of annexin V positive and annexin V/PI double positive cells, indicating an increase in apoptotic and dead cells. Figure 4-9B shows the results of one representative experiment using MCF7 cells 4 days after infection with shCtrl and shHP1γ_1 and 2 shRNA lentiviral vectors.

With shRNA_1, a significant 20% decrease of viable cells compared to shCtrl cells was observed (p-value <0.0001), indicating that cell viability was affected (Figure 4-9C). Furthermore, an approximate 10% increase of dead cells (p-value <0.0001) and 5% increase of apoptotic cells (p-value <0.06) were observed with shHP1γ_1. The number of PI positive, necrotic cells remained unchanged. With shRNA_2 a modest non-significant increase of 5% of dead cells (p-value >0.224) and 5% decrease of live cells (p-value >0.231) were observed. In summary, these results demonstrate that compared to control cells, shHP1γ_1 cells show increased cell death with an increase in apoptosis as determined by annexin V staining. The effects on cell viability were less robust with shHP1γ_2, which depletes HP1γ less efficiently than shHP1γ_1.

4.3.7. Detection of apoptosis in MCF7 cells

The majority of annexin V positive cells detected in HP1γ knockdown cells, co-stained with PI, indicating cells were in late stages of apoptosis or already dead. To further investigate the apoptotic process in my system, I analysed the levels of cleaved poly (ADP-ribose) polymerase-1 (PARP1). PARP1 is a known substrate of cysteinyl-aspartate proteases (caspases)[215]. Caspases are important mediators of signaling cascades involved in apoptosis. They initiate cell death by cleaving and activating other, so called, effector caspases, which drive the process of apoptosis. Effector caspases then activate DNA nucleases resulting in DNA fragmentation, another key feature of apoptosis[216].

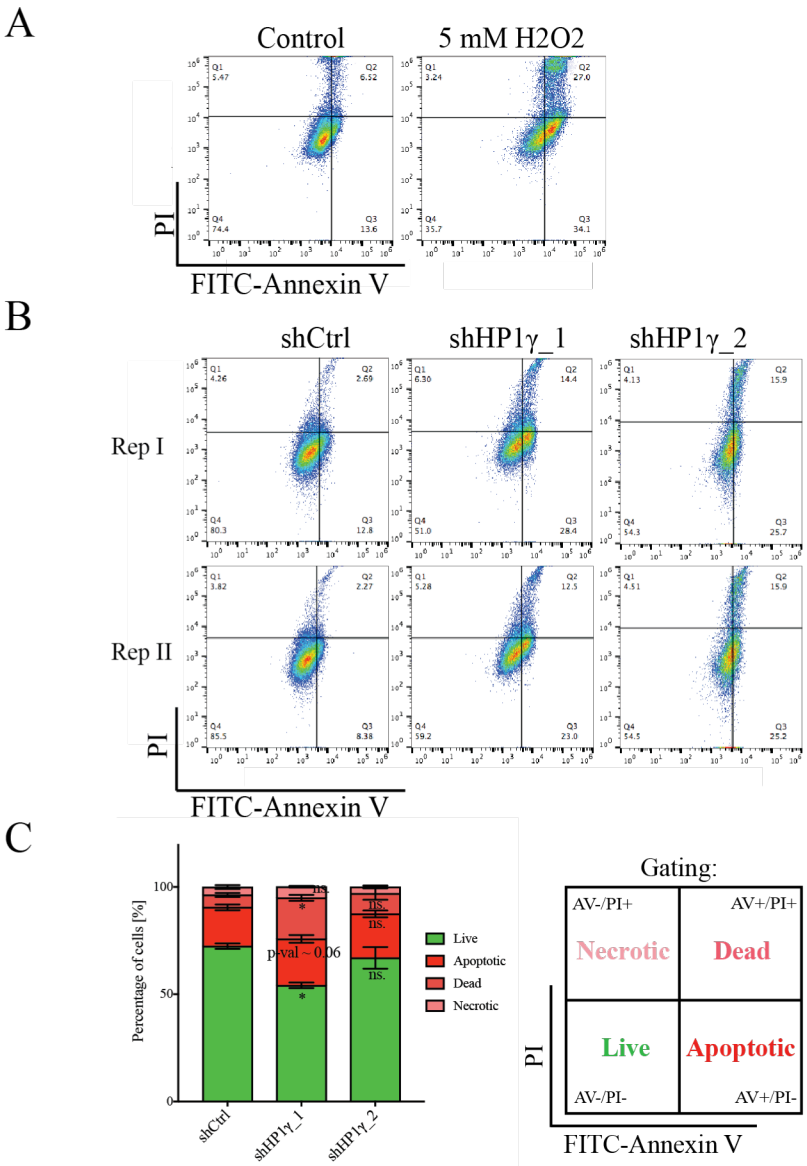


Figure 4-9 Annexin V-FITC/PI staining of MCF7 cells.

MCF7 cells were infected with shCtrl and shHP1γ_1 and 2 shRNA lentiviral vectors and processed 4 days post infection. Cells were trypsinsed and stained with annexin V-FITC and propidium iodide (PI) and analysed by flow cytometry. Gating was adjusted manually using shCtrl cells as a reference. A total of 30000 cells were recorded per sample.

A Flow cytometry results of cells treated with 5mM H_2O_2 for 30min compared to untreated cells.

B Flow cytometry results of HP1γ knockdown cells compared to control cells. Representative plots of one experiment are shown.

C Summary of flow cytometry results (left panel). Gating of cell populations and their definition are summarised in the right panel. Plots contain results from 3 independent experiments performed in duplicate. Significance was assessed using one-way ANOVA tests (* p-value <0.0001, ns. p-value >0.224). SEM are indicated.

Caspase-mediated PARP1 cleavage is considered a hallmark of apoptosis. It results in 2 very specific fragments of 24 and 89kDa in size[215,217]. To detect PARP1 cleavage upon HP1γ depletion in MCF7 cells a specific antibody against the 24kDa fragment of PARP1 was used. Figure 4-10A summarises the results of a time course monitoring the level of cleaved PARP-1 after 2, 4 and 6 days after infection with shRNA vectors. An increase of cleaved PARP1 protein was observed 4 days after infection in both HP1γ knockdown cell lines compared to shCtrl cells. The most effective knockdown of sHP1γ was observed at day 6 with both shRNAs, which correlated with a further increase of cleaved PARP1 levels, indicating that apoptotic signaling cascades were activated upon HP1γ depletion. Higher levels of PARP1 were detected for cells treated with the more effective shRNA shHP1γ_1. Notably, a slight increase in cleaved PARP1 levels was also observed in shCtrl cells, which may be due to puromycin selection mediated stress.

In summary, elevated levels of cleaved PARP-1 were observed in HP1γ depleted MCF7 cells compared to control cells. This indicated that cells activate caspase dependent apoptotic pathways.

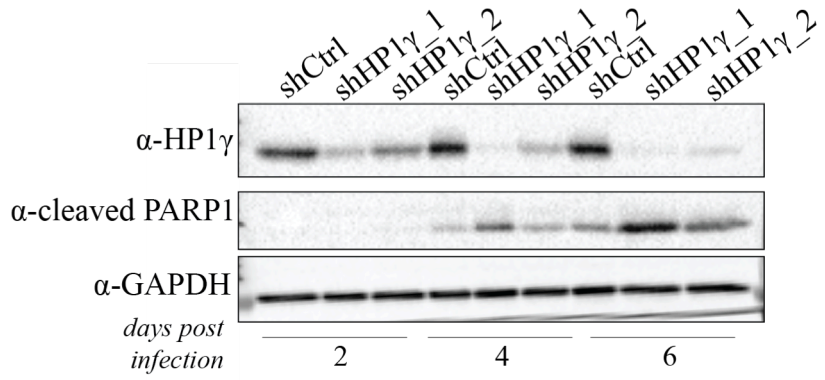


Figure 4-10 Apoptosis detection in MCF7 cells.
MCF7 cells were infected with shCtrl and shHP1γ_1 and 2 shRNA lentiviral vectors and processed 2, 4 or 6 days post infection. Immunoblotting of whole cell lysates using indicated antibodies. Protein concentrations were determined using a BCA protein quantification assay.10 µg of protein were loaded per well.

4.3.8. Migration of MCF7 cells

To study the migratory response of cells towards a serum gradient, a standard membrane-based transwell assay was used for MCF7 cells upon siRNA-mediated knockdown of HP1γ (principle of transwell assay is explained in Section 2.8.11).

The duration of migration was optimised for MCF7 cells so that approximately 100 cells were counted per window at 20x magnification. As MCF7 cells are epithelial cells with low intrinsic motility, an optimal migration period of 24h was determined. 24h after transfection with siRNAs, MCF7 cells were seeded into transwell chambers and incubated for 24h. HP1γ knockdown with 2 independent siRNAs was confirmed by immunoblotting (Figure 4-11A) and RT-PCR (Figure 4-11B). On average, 120 migrated cells were counted for cells transfected with control siRNA and siRNA_3, whereas cells transfected with siHP1γ_1 migrated approximately 40% less effectively (Figure 4-11C). However, this result was not statistically significant (p-value >0.08). To conclude, these results indicated that HP1γ depletion induced no consistent effects upon MCF7 cell migration.

4.4. Phenotype changes upon HP1γ depletion in MDA-MB-231 cells

4.4.1. Expression of HP1 in MDA-MB-231 cells

To determine HP1γ protein levels in MDA-MB-231 cells, cellular fractionation was performed using IPH chromatin extraction buffer (compare Section 4.3.1). Cellular equivalents from each fraction were resolved by SDS PAGE and analysed by immunoblotting using antibodies against HP1γ, GAPDH and histone H3 (Figure 4-12). No GAPDH was detected in the chromatin fraction and only a small amount of H3 was present in the nucleo-cytoplasmic fraction, indicating cellular fractionations were performed successfully. The highest signal for HP1γ was observed in the chromatin fraction. Similarly to my results in MCF7 cells (Figure 4-3) approximately 33% of the protein was detected in the nucleo- and cytoplasmic fraction, indicating that, also in MDA-MB-231 cells, a significant fraction of HP1γ is not tightly associated with chromatin.

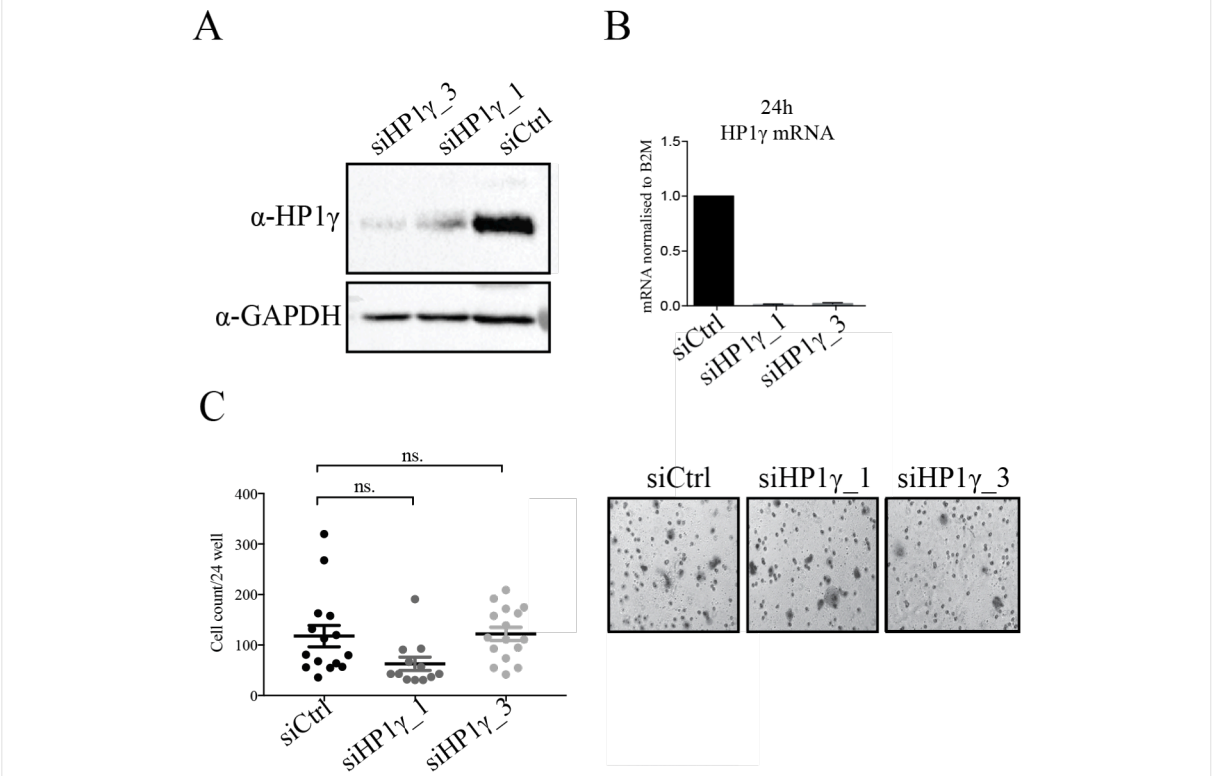


Figure 4-11 Migration of MCF7 cells.

MCF7 cells were transfected with siRNA vectors (siHP1γ_1, 3 and siCtrl). 24h post transfection, 5*10^4 cells were seeded into transwell chambers, which were then placed into 24 wells containing 15% (v/v) FBS culture medium. Cells were incubated for 24h and migrated cells were fixed and stained with crystal violet. Cells were counted manually at 20x magnification with the transwell area divided into 15 windows. Each dot represents an average of cells counted per window.

A HP1γ protein level 48h post transfection with siRNAs assessed by immunoblotting using indicated antibodies. Whole cell lysates from 40000 cells were loaded per well.

B HP1γ mRNA levels 24h post transfection with the indicated siRNAs assessed by RT-PCR, normalised to B2M. Results of PCR replicates are shown.

C Quantification of migration assays (left panel). Results of 5 independent experiments are shown. Significance was assessed using one-way ANOVA tests. Results were not significant (p-value >0.08). SEM are indicated. Representative images of migrated cells in each condition are shown (right panel).

4.4.2. shRNA mediated knockdown of HP1 γ in MDA-MB-231 cells

Stable knockdown of HP1γ in MDA-MB-231 cells was achieved using lentiviral vectors expressing shRNA shHP1γ_1 or shCtrl. HP1γ knockdown efficiency was determined by measuring HP1γ mRNA and protein levels over the course of 6 days (Figure 4-13). HP1γ mRNA levels were more than 95% down-regulated after two days in shHP1γ_1 cells

compared to shCtrl cells (Figure 4-13B). 4 days post infection, HP1γ protein level was reduced by 85% (Figure 4-13A) and almost complete depletion was observed after 6 days.

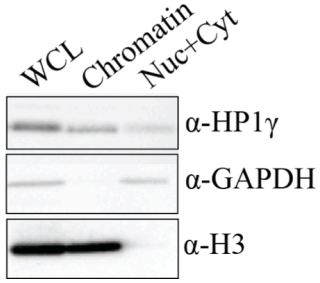


Figure 4-12 Cellular fractionation in MDA-MB-231 cells.
Immunoblotting of MDA-MB-231 lysates using indicated antibodies. MDA-MB-231 cells were lysed directly in 2x Laemmli buffer to prepare whole cell lysates (WCL). Cellular fractionation was performed using IPH chromatin extraction buffer. Nuc+Cyt: nucleo- and cytoplasmic fraction. Cellular equivalents (representing 50000 cells) of each fraction were loaded per well.

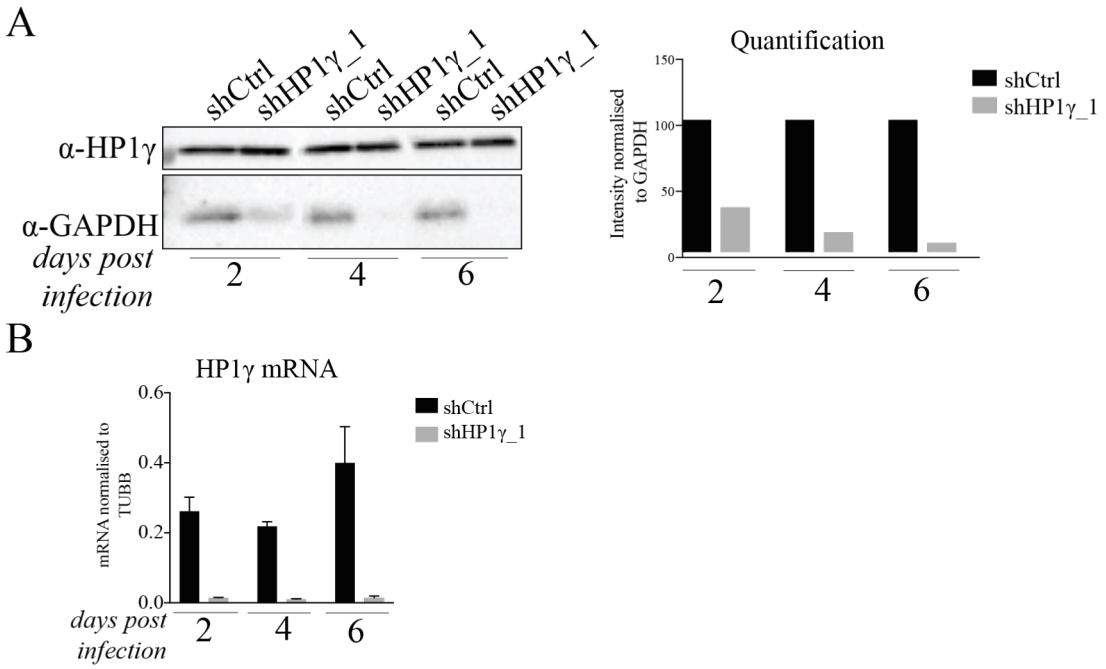


Figure 4-13 Time course of HP1γ knockdown in MDA-MB-231 cells.
MDA-MB-231 cells were infected with shCtrl and shHP1γ_1 shRNA vectors and processed 2, 4 and 6 days post infection.

A Immunoblot of MDA-MB-231 whole cell lysates, using indicated antibodies (left panel). Signal was quantified using ImageJ and normalised to GAPDH (right panel). Protein lysates representing 50000 cells were loaded per well.

B HP1γ mRNA levels assessed by RT-PCR relative to shCtrl samples. Graph shows relative mRNA expression after normalisation to *TUBB* expression. The fold change over shCtrl samples is shown. An average of duplicates is shown. SEM are indicated.

4.4.3. HP1γ depletion affects proliferation of MDA-MB-231 cells

Previous results showed that loss of HP1γ significantly affects cell proliferation of MCF7 cells, but not of normal breast epithelial cells (Section 4.3.3). To study the effect of HP1γ depletion on MDA-MB-231 cell growth, CyQUANT® cell proliferation assays were performed with HP1γ knockdown cells. 4 days after infection with shRNA vectors, cells were seeded into 96-well plates at equal densities ($t = d_0$). HP1γ protein knockdown at d_0 was confirmed by immunoblotting (Figure 4-14A). The fluorescence signal was significantly reduced in HP1γ knockdown cells compared to shCtrl cells, indicating MDA-MB-231 cell proliferation was impaired upon HP1γ depletion (Figure 4-14B). The average signal detected at day 8 in HP1γ knockdown cells was approximately 50% lower than at day 0, indicating that initial cell numbers were reduced.

In conclusion, similarly to MCF7 cells, depletion of HP1γ from MDA-MB-231 cells significantly impairs their proliferation and viability (initial number reduced).

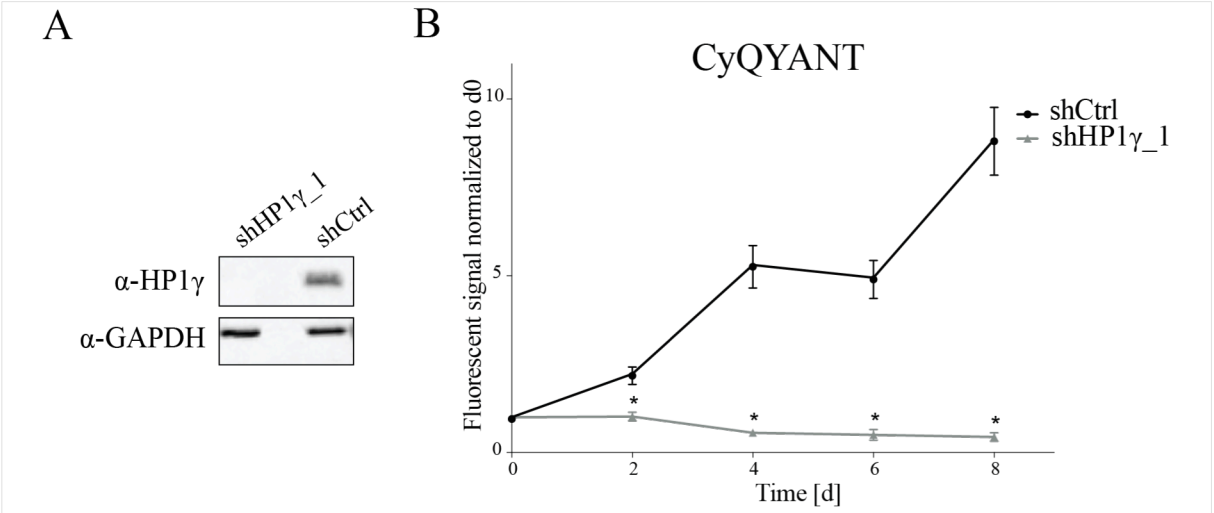


Figure 4-14 Proliferation of MDA-MB-231 cells.
MDA-MB-231 cells were infected with shCtrl and shHP1γ_1 vectors and processed 4 days post infection (t = d0).
A Immunoblot of MDA-MB-231 whole cell lysates at d_0 using indicated antibodies. Protein lysates from 40000 cells were loaded per well.
B CyQUANT® proliferation assays over an 8 day time course. Cells were seeded at 500 cells/well in a 96-well plate at d_0. Fluorescence was measured at 520 nm in a Clariostar® microplate reader at indicated times. Fluorescent units were normalised to signals measured at d_0. Graph includes results of 2 independent experiments, each performed with 12 wells per condition. Significance was assesses using one-way ANOVA test (* p-value > 0.0001). SEMs are indicated.

4.4.4. Cell cycle analysis in MDA-MB-231 cells

To analyse cell cycle progression, unsynchronised MDA-MB-231 cells were stained with PI and DNA content per cell was assessed by flow cytometry. Cell cycle was analysed after 4 days of infection with shCtrl and shHP1γ_1 vectors. HP1γ protein knockdown was confirmed by immunoblotting (Figure 4-15A). Cell cycle analysis revealed no significant difference between HP1γ knockdown and shCtrl cells (Figure 4-15C). In addition to analysing cell cycle profiles, I measured the levels of phosphorylated serine 10 of histone H3 (H3S10ph), a well-established histone modification of mitotic chromosomes catalysed by Aurora kinase B[218]. Immunoblotting of whole cell lysates of MDA-MB-231 cells, 4 days after infection with shHP1γ_1 and shCtrl vectors was performed using an α-H3S10ph and α-H3 antibodies (Figure 4-15B). H3S10ph levels were normalised to total H3 levels for each sample. On average a reduction of H3S10ph signal of approximately 40% was observed in HP1γ knockdown samples.

To summarise, cell cycle profiles of MDA-MB-231 cells depleted of HP1γ by its knockdown did not differ from control cells, indicating that cells were not arrested in a particular phase of the cell cycle. The reduction of H3S10ph in HP1γ knockdown cells suggests that either fewer cells were in M-phase or that cells simply had less of the modification as they progressed through this phase of the cell cycle.

4.4.5. Morphology of HP1γ depleted MDA-MB-231 cells

Loss of HP1γ also causes MDA-MB-231 cells to undergo morphological changes. 4 days after infection with shRNA vectors, HP1γ knockdown cells were aggregated into clusters and began to round up and detach from the culture dish. In contrast, shCtrl cells showed normal MDA-MB-231 morphology (Figure 4-16). After 6 days, more than 80% of knockdown cells were detached (not shown).

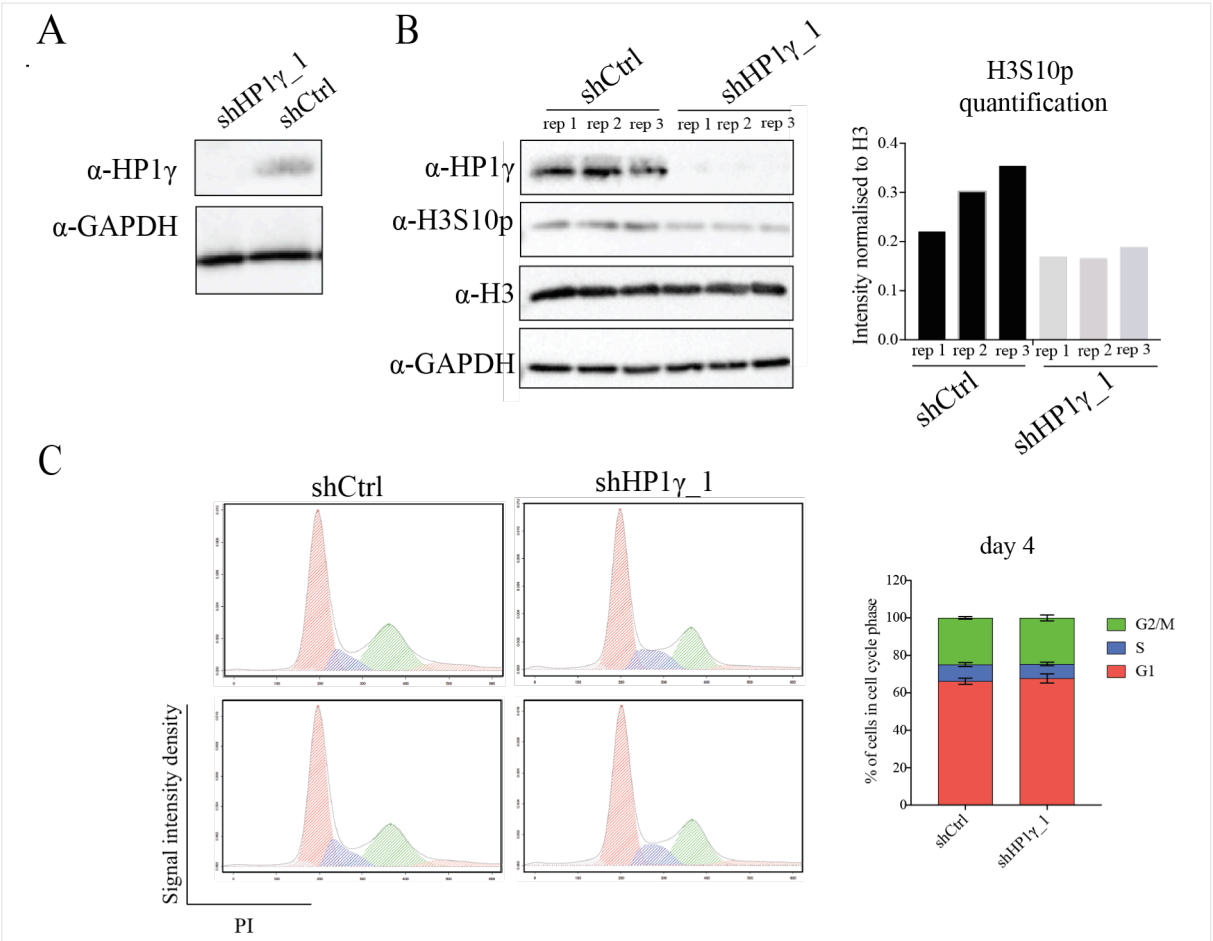


Figure 4-15 Cell cycle analysis of MDA-MB-231 cells.
MDA-MB-231 cells were infected with shCtrl and shHP1γ_1 shRNA lentiviral vectors and processed 4 post infection.

A Immunoblot of MDA-MB-231 whole cell lysates at day 4 using the indicated antibodies. Protein lysates from 40000 cells loaded per well.

B Immunoblot of MDA-MB-231 whole cell lysates at day 4 in triplicates per condition, using indicated antibodies. Protein lysates from 40000 cells were loaded per well. H3S10-ph blot was stripped and re-blotted with an α-H3 antibody (left panel). H3S10ph signal was quantified using ImageJ and normalised to H3 (right panel).

C Cell cycle analysis of MDA-MB-231 cells by flow cytometry. Unsynchronised cells were stained with PI to assess DNA content. A total of 30000 cells were recorded per sample. Representative images of histograms are shown, curves were fitted using a custom R script (left panel). Percentages of cells in each cell cycle phase are summarised in the right hand panel. Plots contain results from 3 independent experiments performed in triplicates. SEM are indicated.

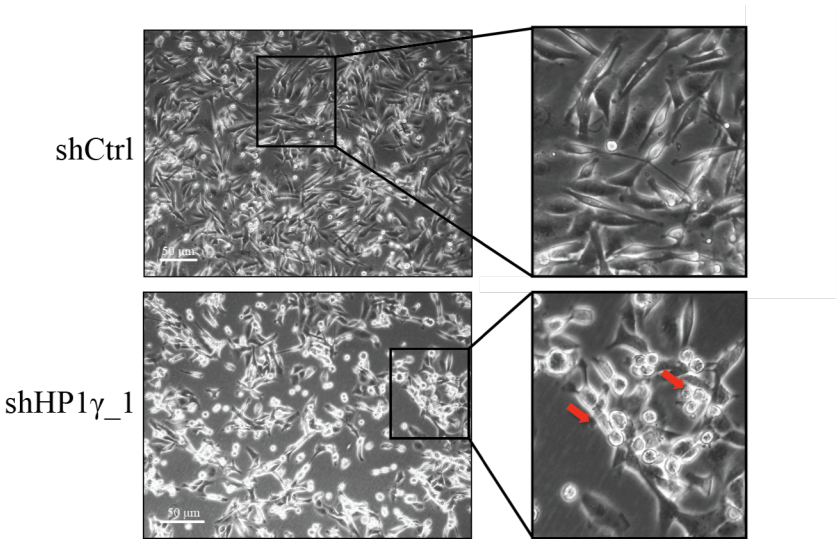


Figure 4-16 Morphology of HP1γ depleted MDA-MB-231 cells.
Brightfield microscopy images of MDA-MB-231 cells 4 days post infection with lentiviral shRNA vectors (as indicated). Images were captured with a Leica EC3 digital camera at 10x magnification. Scale bar indicated (50 µm). Regions of interest are 3-fold magnified and red arrows highlight cell aberrations.

4.4.6. HP1γ depletion affects cell viability of MDA-MB-231 cells

To determine cell viability of HP1γ depleted MDA-MB-231 cells, annexin V/PI staining was performed. Cells were analysed by flow cytometry. Results from two experiments are summarised in (Figure 4-17). Upon HP1γ knockdown, an increase of approximately 14% of dead and apoptotic cells together with a 10% decrease in viable cells was observed compared to shCtrl cells. However, these results were not statistically significant (p-value > 0.15), which suggests that the experiment should be repeated.

4.4.7. Detection of apoptosis in MDA-MB-231 cells

To understand whether cells were detaching due to apoptosis, I divided HP1γ depleted cells into attached and detached cells and tested the levels of cleaved PARP1 by immunoblotting. Both fractions were depleted of HP1γ (Figure 4-17A). Interestingly, cleaved PARP1 was only detected in detached and not attached HP1γ knockdown cells suggesting that caspase dependent apoptotic pathways were activated after cells detached (Figure 4-17B).

In conclusion, MDA-MB-231 cells detached from culture dishes upon HP1γ depletion. Increased levels of PARP1 were observed in detached cells, suggesting that detached cells were less viable than attached MDA-MB-231 cells. Assessment of cell viability and apoptosis in MDA-MB-231 cells indicated an increase of apoptotic cells together with a decrease in the number of viable cells.

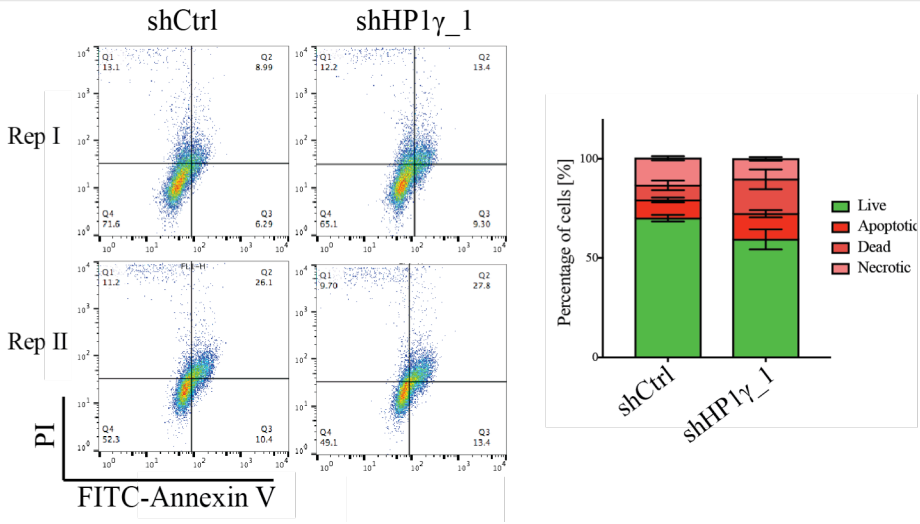


Figure 4-17 Viability and apoptosis assessment in MDA-MB-231 cells.
MDA-MB-231 cells were infected with shCtrl and shHP1γ_1 lentiviral vectors and processed 4 days post infection. Cells were stained with annexin V-FITC/PI and analysed by flow cytometry. A total of 30000 cells were recorded per sample. Gating was adjusted manually using shCtrl cells as a reference. Flow cytometry results are shown of HP1γ knockdown compared to shCtrl cells (left panel). Representative plots for one experiment are shown. Summary of flow cytometry results (right panel). Refer to **Figure 4-9**C for gating key. Plots contain results from 2 independent experiments performed in duplicate. Significance was assessed using Student T-tests. Results were not significant (p-value > 0.15). SEM are indicated.

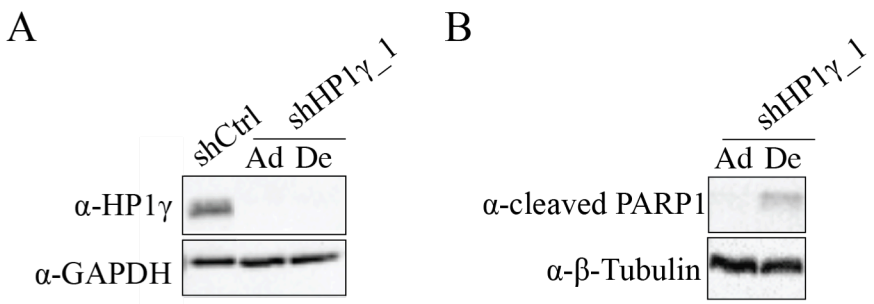


Figure 4-18 Apoptosis assessment in MDA-MB-231 cells.
MDA-MB-231 cells were infected with shCtrl and shHP1γ_1 lentiviral vectors and processed 4 days post infection. Immunoblotting of whole cell lysates using the indicated antibodies was performed. shHP1γ_1 cells were divided into attached (Ad) and detached (De) cells, and adherent shCtrl cells were used as a control. Protein lysates from 40000 cells were loaded per well.
A Immunoblot of shHP1γ_1 and shCtrl cell lysates with indicated antibodies.
B Immunoblot of shHP1γ_1 attached and detached cell lysates with indicated antibodies.

4.4.8. Migration of MDA-MB-231 cells

MDA-MB-231 cells are of metastatic origin and therefore have intrinsically high cellular motility. An optimal migration period of 4h was established for this cell line. Migration assays were set up 44h after transfection with siRNAs. HP1γ knockdown with 2 independent siRNAs was confirmed by immunoblotting (Figure 4-19A). Intriguingly, both knockdown cell lines migrated 80% less compared to control cells (p-value <0.0001), indicating that depletion of HP1γ influences cell motility of MDA-MB-231 cells (Figure 4-19B).

To further confirm that this effect is dependent on HP1γ, a rescue experiment was performed. Together with the siRNAs, MDA-MB-231 cells were transfected with mammalian expression vector pcDNA3.1 containing DNA encoding mouse HP1γ N-terminally fused to a HA-tag. An empty pcDNA3.1-HA vector was used as control. Protein expression was observed by immunoblotting using an α-HA antibody (Figure 4-19C). HA-HP1γ protein levels were expressed at approximately 50% of the endogenous HP1γ levels as assessed by immunoblotting with an α-HP1γ antibody (compare lane siHP1γ_3 + HA-HP1γ to lane siCtrl + HA-Ctrl). An average of 75 control cells migrated through the membrane, whereas an average of 22 HP1γ knockdown cells (siHP1γ_3 + HA-Ctrl) did so (p-value <0.0001). For HP1γ rescue cells (siHP1γ_3 + HA-HP1γ) an average of 42 migrated cells were counted,

indicating an increase of approximately 50% compared to knockdown cells (p-value <0.031). These preliminary results indicate that the migratory phenotype of HP1γ knockdown cells was partially rescued in cells expressing HA-HP1γ.

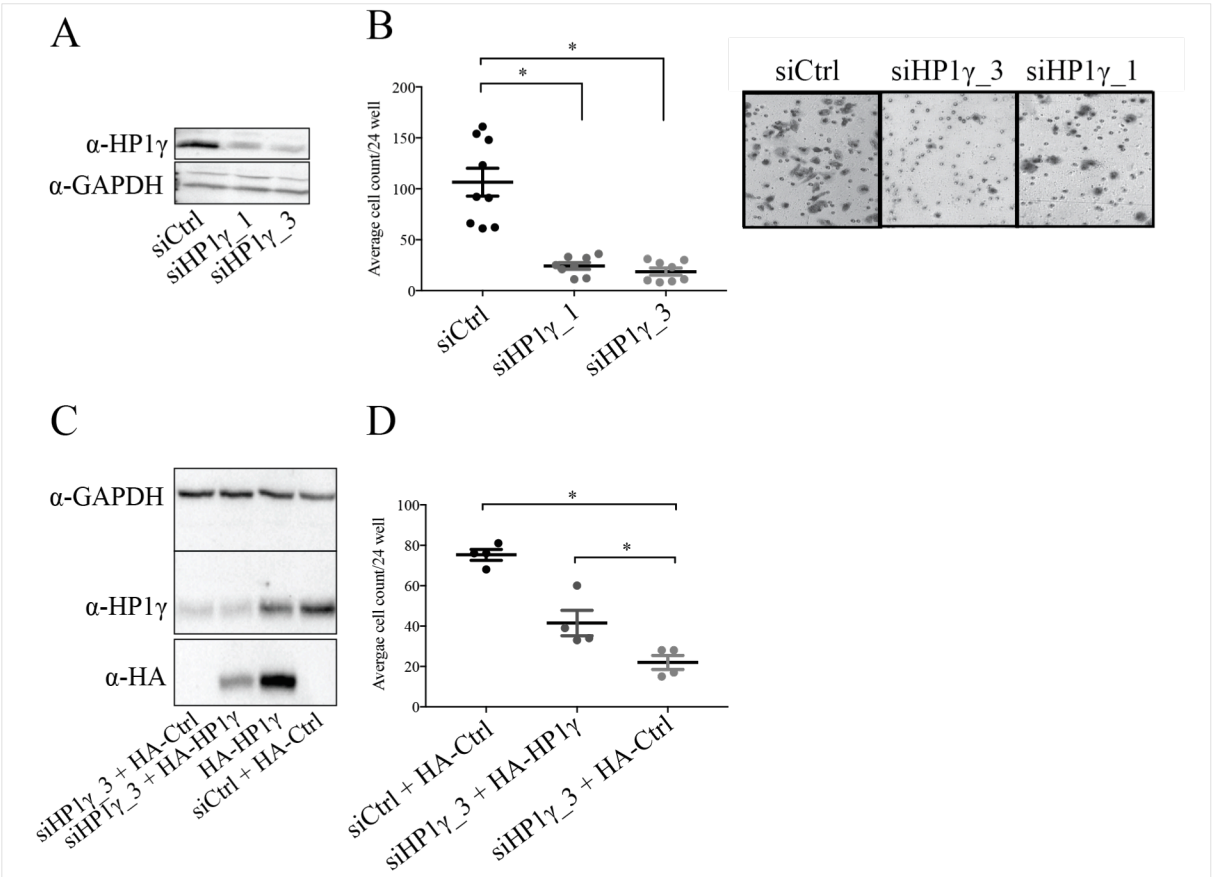


Figure 4-19 Migration of MDA-MB-231 cells.

MDA-MB-231 cells were transfected with siRNA vectors (siHP1γ_1, 3 and siCtrl). 44h post transfection 5*10^4 cells were seeded into transwell chambers, which were placed into 24 wells containing 15% (v/v) FBS culture medium. Cells were incubated for 4h and migrated cells were fixed and stained with crystal violet. Cells were counted manually at 20x magnification, where the transwell area was divided into 15 windows. Each dot represents an average of cells counted in each window.

A HP1γ protein levels 44h post transfection with the indicated siRNAs, assessed by immunoblotting using indicated antibodies. Whole cell lysates from 40000 cells loaded per well.

B Quantification of migration assays (left panel). Results of 3 independent experiments are shown. Significance was assessed using one-way ANOVA tests (p-value <0.0001). SEM are indicated. Representative images of migrated cells in each condition are shown (right panel).

C HP1γ protein level 44h post transfection with the indicated siRNA and pcDNA3.1-HA constructs assessed by immunoblotting using indicated antibodies.

D Preliminary results from one migration assay. Significance was assessed using one-way ANOVA tests (p-value < 0.031). SEM are indicated.

In summary, migration assays showed that HP1γ knockdown significantly affects the migratory response of MDA-MB-231 cells towards a serum gradient. This phenotype was partially rescued by reintroducing HP1γ expression, suggesting that migration of MDA-MB-231 cells depends on HP1γ.

4.4.9. Wound healing of MDA-MB-231 cells

To study the effect of HP1γ depletion on MDA-MB-231 cell motility further, I performed 2D wound healing (scratch) assays with MDA-MB-231 cells upon siRNA mediated HP1γ knockdown. In contrast to transwell migration assays, which acquire endpoint measurements, wound healing assays measure speed and efficiency with which cells migrate to close a gap in a monolayer over time. Wound healing assays were performed using the Essenbioscience IncuCyteZOOM® system. MDA-MB-231 cells were seeded into 96-well plates and transfected with siRNAs.

24h later a wound of known area was introduced into the cell monolayer of each well. Images were taken every 45min over a total of 34h to follow the cell movement over time. The images were then used to calculate the relative wound healing density (RWD in %), which was calculated using an algorithm based on the density of both the cell region and the wound region. After just 2h, a significant 2-fold decrease in the RWD of both knockdown cell lines compared to siCtrl cells was observed (p-value <0.0001). The difference between knockdown and control cells increased further over the course of 34h, indicating that cell movement of MDA-MB-231 cells was impaired upon HP1γ knockdown (Figure 4-20A). Representative lightfield images of wound healing taken at different time points illustrate the differences of the migration fronts in knockdown and control cells (Figure 4-20B).

To conclude, both transwell migration and wound healing assays indicated that MDA-MB-231 cells have impaired cellular migration upon HP1γ depletion.

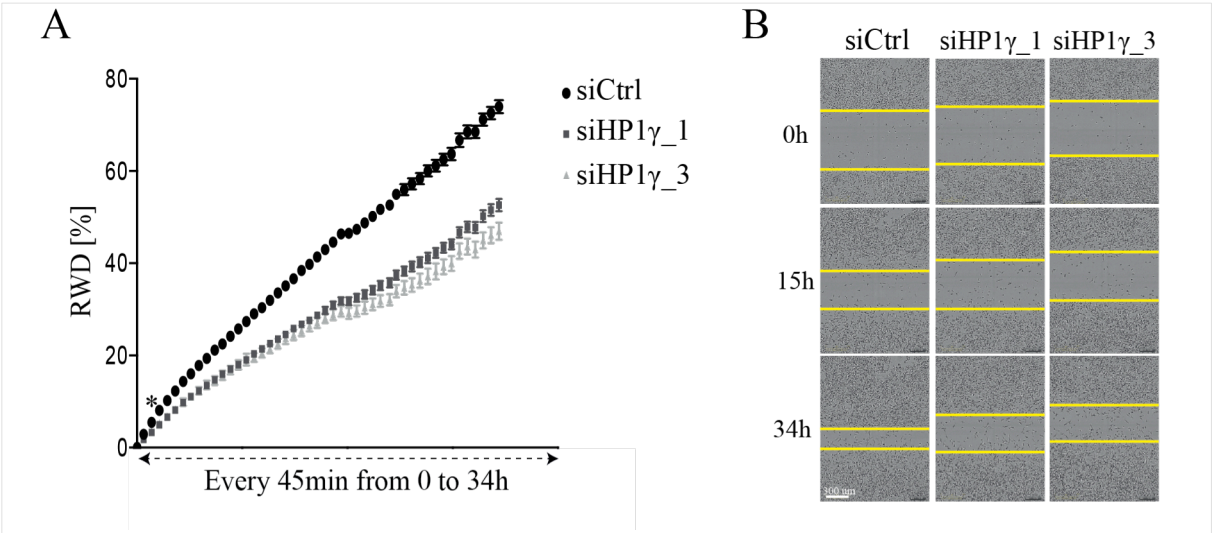


Figure 4-20 Wound healing assay of MDA-MB-231 cells.

Wound healing assays were performed using the Essenbioscience IncuCyteZOOM® system. MDA-MB-231 cells were seeded in to a 96-well plate and transfected with siRNAs (as indicated). After 24h a wound of known area was introduced and cells were imaged every 45min over the course of 34h.

A Relative wound healing density (RWD) for each time point. Graph contains results of two experiments performed with 16 wells per condition. Significance was assessed using one-way ANOVA tests (p-value <0.0001 from indicated time point (t = 2h) onwards). SEM are indicated.

B Representative lightfield images of wound healing at time points t = 0, 15 and 34h. Representative scale bar is indicated (300 µm). Migration fronts are highlighted by yellow lines.

4.5. Discussion

In this chapter, I presented functional data on the role of HP1γ in two the breast cancer cell lines, MCF7 and MDA-MB-231. The phenotype of HP1γ depletion was characterised in both cell lines using assays to assess cell growth, cell cycle progression, cell viability and motility. My data indicate that loss of HP1γ affects proliferation and cell viability of both cancer cell lines.

4.5.1. Cell growth

In both MCF7 and MDA-MB-231 cells lines loss of HP1γ impairs cell proliferation, whereas proliferation of non-tumorigenic epithelial breast cells remains unaffected, indicating a cancer cell specific function for HP1γ. In contrast to the phenotype described in the colon cancer cell line HCT116[ref.155,156], cell cycle analysis provided no sufficient evidence that reduced proliferation is directly associated with a defect in cell cycle progression. In MDA-MB-231 loss of HP1γ had no effect on cell cycle.

MCF7 cells infected with shHP1γ_2 accumulated in G1 phase of the cell cycle, whilst knockdown cell line shHP1γ_1 showed no difference to control cells. Differences between these two shRNAs cannot be directly correlated with HP1γ protein levels, as cells treated with the more effective shRNA progressed through the cell cycle normally. At this stage it is unclear why knockdown of HP1γ with both shRNAs results in such different results. In order to exclude any off-target effects from either shRNA, it will be important to repeat the experiments with additional independent shRNAs.

It is possible that cells are not arresting in a particular phase of the cell cycle, but rather that the length of each cycle is generally increased. To determine cell cycle duration one could use live cell imaging of synchronised cells to measure the time it takes for cells to progress through a cell cycle. Alternatively, one could use dual pulse labeling of cells with different nucleosides such as BrdU (5-bromo-2'-deoxyuridine) and EdU *(5*-ethynyl-*2′*-deoxyuridine), which incorporate into the DNA of growing cells. Sequential pulses of BrdU and EdU are useful in determining cell cycle kinetics and they provide an estimate of total cell cycle length[219].

Of cause, the cell cycle may not be affected at all. Reduced proliferation rates may simply reflect increased cell death.

4.5.2. Cell viability and apoptosis

Annexin V/PI staining detected an increase of dead cells in both cell lines, indicating that HP1γ depletion affects cell viability. In MCF7 cells, a significant decrease of viable cells together with an increase of dead and apoptotic cells was observed, whilst the phenotype in MDA-MB-231 cells was less well defined. Furthermore, both cell lines showed elevated levels of cleaved PARP-1 upon HP1γ knockdown, together with an increased number of cells with fragmented DNA, indicating induction of apoptosis.

MCF7 cells showed classical morphological signs of apoptosis such as membrane blebbing, condensation and cell shrinkage. However, these signs were not observed in MDA-MB-231 cells. In contrast, upon loss of HP1γ, MDA-MB-231 cells rounded up and detached from the culture dish, with up to 80% of cells becoming detached 6 days after infection with shRNA vectors. Control cells remained attached. The fact that both cell lines show such different morphologies upon loss of HP1γ might suggest that apoptotic pathways are activated in different ways. Indeed, cleaved PARP-1 was only detected in detached MDA-MB-231 cells and not in the attached cell population, suggesting that caspases were only activated after cells detached.

Apoptosis can be triggered via two major pathways, the intrinsic and the extrinsic pathway. In the intrinsic pathway, caspase activation occurs as a consequence of mitochondrial permeabilisation. This is regulated by the pro-apoptotic BH3-only protein family, which control formation of pores in the outer mitochondrial membrane (OMM), releasing proapoptotic factors such as cytochrome *c*. The release of cytochrome *c* in turn activates caspases. The extrinsic pathway is initiated by ligation of death receptors on the cell surface, such as TNF receptor or Fas, resulting in the assembly of a death-inducing signaling complex (DISC). DISC then recruits FADD (Fas-associated protein with death domain) and caspase 8, resulting in activation of the latter. Both pathways ultimately result in activation of the major effector caspase 3, which triggers the execution of apoptosis leading to protein and DNA degradation[220].

When comparing the activation of apoptosis in MDA-MB-231 and MCF7 cells their genetic backgrounds have to be considered. MDA-MB-231 cells express high levels of caspases, whereas MCF7 cells are known to lack expression of caspase 3 and are therefore dependent on other caspases to trigger cell destruction[221]. Besides caspases, MDA-MB-231 cells express high levels of mutant p53[222], whereas MCF7 express low levels of wildtype p53[223]. p53 is an important activator of pro-apoptotic BH3-only proteins, hence a potent activator of the intrinsic pathway[224]. It would be interesting to understand whether HP1γ's role in mediating apoptosis, either directly or indirectly, is cell type specific and whether it is linked to p53.

Finally, it is possible that in MDA-MB-231 cells induction of apoptotic pathways is not the predominant consequence of HP1γ depletion. Loss of HP1γ could trigger a change of cell surface molecules necessary for MDA-MB-231 cells to attach and/or to interact with each other. A process called 'Anoikis' describes the phenomenon of activating apoptosis in anchorage-dependent cells in response to inappropriate cell-to-cell or extra cellular matrix (ECM) interactions. Integrins, as major components of adherens junctions linking the cytoskeleton to components of the ECM play an important role in mediating survival signaling in many cell types[225]. It would be interesting to understand why MDM-MB-231 cells lose contact and whether, for example HP1γ depletion changes expression of genes encoding cell adhesion molecules. Furthermore, HP1γ could indirectly affect localisation of adherens junction or cytoskeleton components. Immunofluorescence stainings for proteins involved in these processes such as integrins, focal adhesion complex members or cytoskeletal components could provide insights into the mechanism by which MDA-MB-231 detachment is mediated upon HP1γ depletion.

4.5.3. Cell migration

Cell migration is a multistep process and a fundamental component of many biological and pathological processes such as tissue re-organization, angiogenesis, immune cell trafficking, wound healing and tumour metastasis. To understand whether HP1γ depletion affects breast cancer cell motility, I performed transwell migration assays in MCF7 and MDA-MB-231 cells upon HP1γ knockdown. MCF7 cells showed no significant effect on cell migration upon HP1γ knockdown, whereas migration of MDA-MB-231 cells was significantly impaired. The effect on MDA-MB-231 cell mobility was confirmed by wound healing assays. Impairment

of cellular movement could be directly associated with the morphological changes observed in HP1γ depleted MDA-MB-231 cells. For cells to migrate they need to be attached to the matrix and since part of the phenotype of HP1γ depleted cells I observed is detachment from the surface this is likely, at least in part, to result in their reduced mobility.

Both MCF7 and MDA-MB-231 cell lines originated from invasive ductal breast carcinomas. However, MCF7 cells are differentiated and retain an epithelial-like morphology, whereas MDA-MB-231 cells are poorly differentiated and have lost any morphological resemblance to epithelial cells. Furthermore, MDA-MB-231 cells are a model for Claudin-low (CL) metastatic breast cancer and express high levels of epithelia-to-mesenchymal transition markers such as vimentin. There are several ways how HP1γ could be involved in the regulation of cell motility. Firstly, HP1γ could be directly or indirectly involved in regulating genes encoding proteins involved in cell adhesion, cell-to-cell interactions or cell communication. Important factors associated with breast cancer metastasis include cadherins, selectins and integrin subunits, which are major components of adherens junctions linking the cytoskeleton to components of the extra cellular matrix. Intercellular adhesion molecules or components of tight junctions such as claudins are further factors linked to breast cancer metastasis[226].

Cell migration is initiated by a stimulus that activates a set of signaling pathways leading to cellular polarization and a rapid reorganization of actin filaments and microtubules. HP1γ could be either directly involved in signaling cascades or affect chromatin condensation, an interesting feature of cell migration. Using wound healing assays, Gerlitz *el al.* showed that upon induction of cell migration, levels of markers for facultative heterochromatin such as H3K27me3 or H4K20me1 are increased in cells at the migratory front[227]. Furthermore, interference with heterochromatin formation using the HDAC inhibitor TSA and the methyltransferase inhibitor chaetocin reduced cellular migration[227]. HP1γ itself is classified as a marker for facultative heterochromatin[97] and so it may be involved in migration mediated chromatin condensation. The nuclear lamina is directly connected to the cytoskeleton through components of the so-called LINC complex, which is integrated into the nuclear membrane[228]. Interestingly, interference with the lamina or the structure of the LINC complex can inhibit cell migration both *in vivo* and *in vitro*[229]. As HP1 proteins directly interact with the nuclear lamina[230], they may be involved in LINC mediated cell migration.

My findings clearly indicate that loss of HP1γ specifically affects cellular motility of more aggressive breast cancer cells, but not of less aggressive breast cancer cells. From a clinical point of view, these results are very interesting, as they suggest a role for HP1γ in breast cancer invasiveness. Importantly, my work highlights HP1γ as a good candidate target for therapeutic intervention in aggressive breast cancer. To dissect the functional differences between aggressive and less aggressive breast cancer cells in more detail, soft agar colony formation assays could be used to monitor the effect of HP1γ depletion on anchorage-independent growth.

To summarise this chapter, loss of HP1γ results in impaired proliferation and reduces cellular viability of the two breast cancer cell lines MCF7 and MDA-MB-231. In both cell lines an increase in apoptotic cells was observed. While MCF7 cells show classical signs of apoptosis, the phenotype in MDA-MB-231 cells is less defined. The analysis of cell motility revealed a significant impairment of cell migration in the more aggressive breast cancer cell line MDA-MB-231, whereas less aggressive MCF7 cells remained unaffected, suggesting a role for HP1γ in breast cancer invasiveness.

The following chapter includes transcriptome analysis in MCF7 before and after HP1γ depletion. The analysis identified pathways that may underlie the aforementioned phenotypic changes.

Chapter 5

Transcriptomic analysis of HP1γ in breast cancer cells

5.1. Introduction

HP1γ overexpression has been linked to increased cell proliferation and oncogenesis in various cancers such as colorectal[155,156], non-small cell lung[157], prostate[158], cervical and breast[159]. First mechanistic insights into how HP1γ might be involved in cancer progression were described for colorectal cancer. Liu *et al.* showed that in human colorectal cancer HCT116 cells elevated levels of HP1γ repress gene expression of the cell cycle regulator p21(Waf1/Cip1) resulting in uncontrolled cell proliferation[155]. Consequently, knockout of HP1γ in HCT116 cells leads to cell cycle arrest and impaired proliferation *in vitro* and *in vivo*. In addition, in the same cell line HP1γ regulates a feedback mechanism by repressing expression of cyclin dependent kinase Cdk6, an activator of p21(Waf1/Cip1)[156]. Similar results were observed in primordial germ cells. HP1γ knockout mice show a dramatic reduction in the number of primordial germ cells due to cell cycle arrest[127]. Cell cycle defects were again linked to HP1γ mediated regulation of p21(Waf1/Cip1) expression.

In the previous chapter, I showed that HP1γ depletion impairs proliferation of the two breast cancer cell lines MCF7 and MDA-MB-231. In contrast to previous studies, a drecrease in proliferation could not directly be linked to cell cycle progression defects. However, in both cell lines an increase in apoptotic cells was observed, indicating that impaired proliferation may be a result of reduced cell viability.

To find pathways that may underlie these phenotypic changes, I performed whole transcriptome analysis in MCF7 cells before and after HP1γ depletion. In this chapter I explore the function of HP1γ in transcriptional regulation and link the findings to the phenotype of HP1γ depletion.

5.2. Aims of this chapter

This chapter has the following aims related to changes in gene expression in MCF7 upon HP1γ depletion.

1. To detect differentially expressed genes in MCF7 cells depleted of HP1γ

2. To identify differentially regulated pathways upon HP1γ depletion

3. To investigate whether gene expression changes are linked to phenotypic changes observed in HP1γ depleted cells

5.3. RNA-seq in MCF7 cells upon HP1γ knockdown

5.3.1. Preparation of RNA-seq libraries

Whole transcriptome analysis in MCF7 cells upon HP1γ knockdown was performed using RNA-sequencing (RNA-seq) technology. Knockdown efficiency was confirmed by immunoblotting using a HP1γ antibody (Figure 5-1A) and RT-PCR (Figure 5-1B), 6 days after infection with shRNA vectors containing shHP1γ_1 or shCtrl sequences. The experiment was performed with three technical replicates for each condition. Total RNA was depleted from ribosomal RNA (see Section 2.11.1) and RNA-seq libraries were prepared as described in Section 2.11.2. Libraries were multiplexed at equal molarities and were submitted for 125bp, paired-end sequencing on an Illumina HiSeq 4000® instrument.

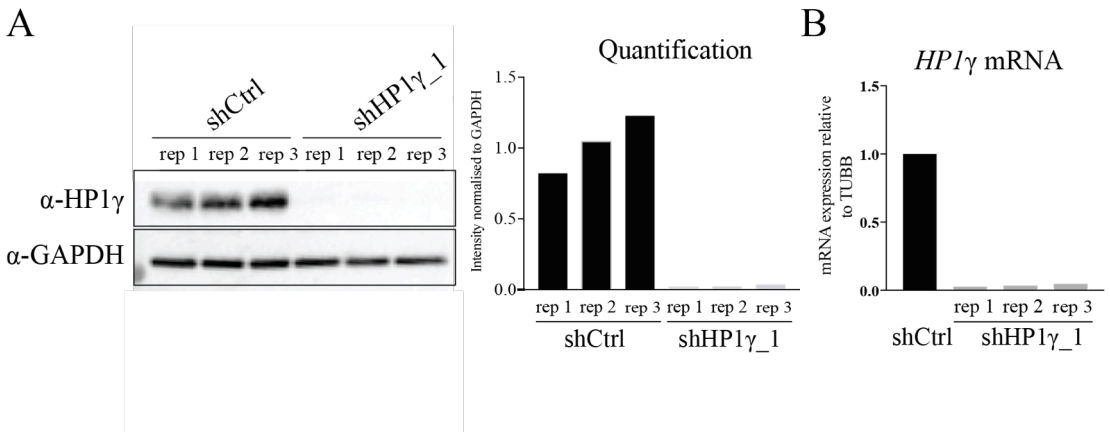


Figure 5-1 Preparation of RNA-seq samples from MCF7 cells following HP1γ knockdown.

MCF7 cells were infected with shCtrl and shHP1γ_1 shRNA vectors and processed 6 days post infection.

A Immunoblot using the indicated antibodies of MCF7 whole cell lysates expressing either control or HP1γ shRNAs as indicated (left panel). Signals were quantified using ImageJ and normalised to GAPDH (right panel).

B HP1γ mRNA levels were assessed by RT-PCR. Data were normalised to *TUBB* and the fold change expression over the average of 3 shCtrl samples is shown.

5.3.2. RNA-seq analysis

5.3.2.1. Quality control of RNA-seq data

The sequencing data was mapped to the reference genome assembly and the total number of mapped reads per replicate was determined (compare Section 2.11.3). The total read counts, summarised in Table 5-1, show that multiplexed samples were well balanced, with no over-represented sample in the pool. The similarity between replicates was determined by hierarchical clustering of normalised reads. Overall replicates clustered together, indicating a high degree of similarity within each sample group (shCtrl and shHP1γ_1) (Appendix I, Figure 1B). The gene-wise dispersion estimates were obtained with DESeq 2[ref.188], and the dispersion-mean curve displayed the expected profile, with shrinkage of the individual gene-wise estimates towards the fitted value (Appendix I, Figure 1A).

In conclusion, quality control assessment detected low level of variability within the data and high degree of similarity across technical replicates within sample groups.

Table 5-1 Total read counts of MCF7 RNA-seq data. Number of uniquely mapped reads per replicate is depicted for each sample.

	rep 1	rep 2	rep 3
shCtrl	38837750	46670329	37493876
shHP1γ_1	36191074	35948714	33089260

5.3.2.2. Differential expression analysis

Differential expression analysis comparing control (shCtrl) versus HP1γ knockdown (shHP1γ_1) samples was performed using the Bioconductor package DESeq2. Results are visualised as a volcano plot (Figure 5-2A). Here, the log2 of expression fold change for each gene on the x-axis is plotted against log10 of the adjusted p-value on the y-axis. Each dot represents a single gene with significant fold change in red (adjusted p-value <0.05) and non-significant (adjusted p-value <0.05) in black. This analysis identified a total number of 4673 genes significantly deregulated upon HP1γ knockdown including 65% of genes with greater than log2 fold change. Differentially expressed genes were both up- (52.4%) and down-

regulated (47.6%). The most significantly up-regulated genes with adjusted p-values <1E-75 were *IFI27, PARP10 DDX60L, FBX032* and *RASD1*. The most significantly down-regulated and the only gene with adjusted p-values <1E-75, was *CBX3* itself. As an example, UCSC browser snapshots of the most up-regulated gene *IFI27* and the most down gene *CBX3* are shown in Figure 5-2B.

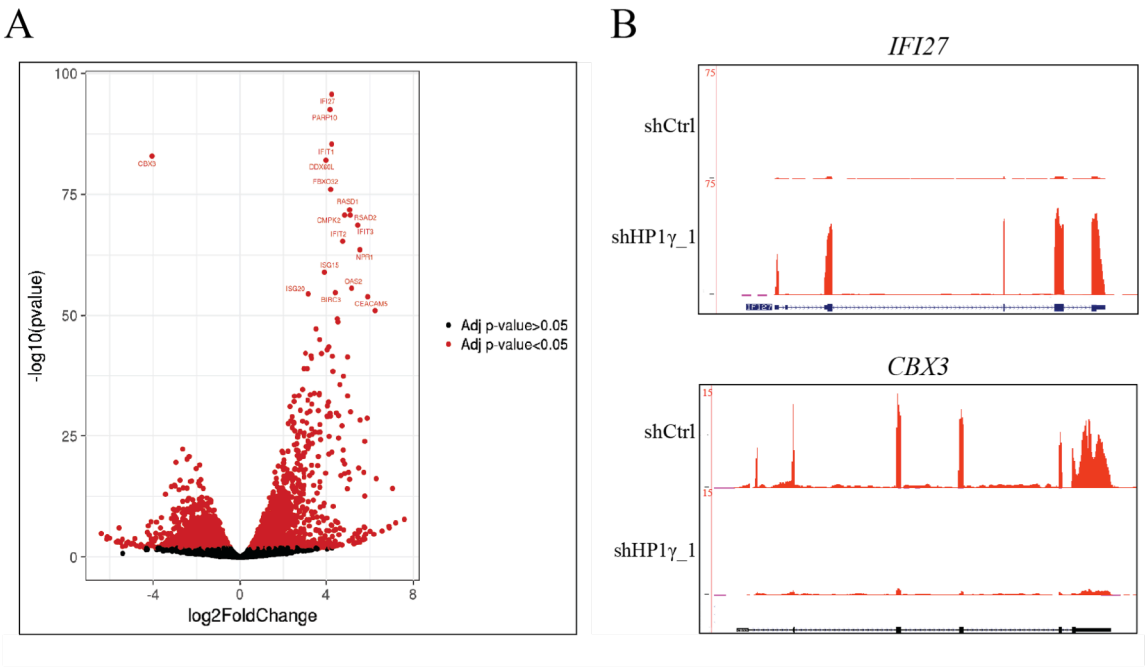


Figure 5-2 Differential expression analysis of MCF7 RNA-seq.
Differential expression analysis performed using Bioconductor package DESeq2.
A Volcano plot of differentially expressed genes. $\log_2$ fold change on X-axis plotted against log10 of the adjusted p-value on Y-axis. Dots represent single genes with significant fold change (red: p-value >0.05) and non-significant (black: p-value <0.05).
B UCSC genome browser snapshots of normalised RNA-seq reads of most up-regulated gene *IFI27* (top panel) and most down gene *CBX3* (bottom panel). Equal scaling is used for shCtrl and shHP1γ _1 tracks.

5.3.2.3. Gene ontology analysis

In order to understand which biological processes were affected by HP1γ depletion, gene ontology enrichment analysis was performed using the generally applicable gene-set enrichment for pathway analysis (GAGE) Bioconductor package. The results are summarised in Figure 5-3. GO enrichments are ranked according to the p-value corrected for multiple testing (q-value) ranging from red (significant) to black (non-significant). GO terms for up-

regulated genes are depicted on the right and down genes on the left hand side. GO terms for up-regulated genes were predominantly linked to angiogenesis and vascular development, followed by regulation of immune system, metabolic processes and transcriptional regulation. On the other hand, among down-regulated genes, GO terms involved in transcriptional elongation, translation, RNA processing and DNA repair were highlighted.

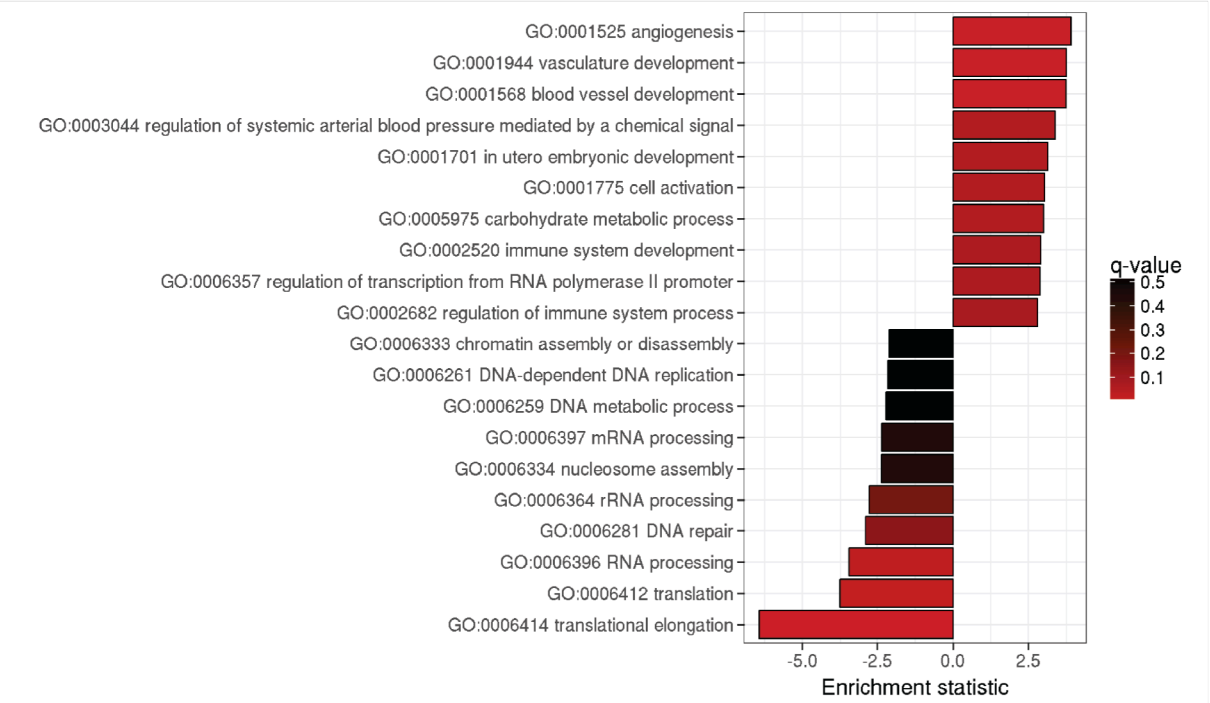


Figure 5-3 Gene ontology analysis of MCF7 RNA-seq data.
GO enrichment statistics are plotted and ranked according to the q-value ranging from red (significant) to black (less significant). Both GO terms for up-regulated genes (right) and are down genes (left) are depicted.

5.3.2.4. KEGG pathway analysis

Next, KEGG pathway analysis was used to understand which biological pathways were affected. Here, KEGG pathway maps covering all known biological pathways ranging from metabolism, cellular processes to human disease pathways are scored according to their enrichment in differentially expressed genes. This analysis allows to identify biological networks and signaling pathways significantly affected in HP1γ depleted cells. Results are summarised in Figure 5-4. Here, the statistical mean for each enrichment group is illustrated. As the majority of genes were up-regulated, more pathways (16 in total) were identified for those genes (Figure 5-4A). For down-regulated genes only 3 pathways were found (Figure

5-4B). Intriguingly, among up-regulated genes several pathways indicating activation of an interferon mediated immune response, reacting to nucleic acid sensing, such as JAK-STAT, RIG-I-like receptor, cytosolic DNA-sensing, NOD-like receptor, Toll-like receptor and apoptosis signaling were identified. This analysis implies that upon knockdown of HP1γ, MCF7 cells activate pathways that, under normal circumstances, should only be triggered in response to acute viral or bacterial infections. Besides immune activation, pathways suggesting regulation of cell adhesion molecules and cell migration were identified. Among down-regulated genes, pathways involving down-regulation of physiological functions, such as ribosomal and mitochondrial processes were found (Figure 5-4B). Figure 5-5 illustrates two examples of the pathways RIG-I-like receptor signaling (upper panel) and ribosomal function (lower panel). Pathway maps for JAK-STAT, cytosolic DNA-sensing, NOD-like receptor, Toll-like receptor and apoptosis signaling are illustrated in Appendix I (Figures 2 to 6). Gene expression changes of individual genes are colour-coded, where red indicates up-regulation and green down-regulation. It illustrates clearly that many genes within the intracellular RNA sensing RIG-I-like pathway are up-regulated, whereas the majority of ribosomal genes are collectively down-regulated.

Taken together, the KEGG pathway analysis suggests that the cells are under physiological stress and are in the process of shutting down biological processes that are important for cell homeostasis.

5.3.2.5. Determination of differential exon usage

Two independent reports proposed a role for HP1γ in alternative splicing[143,145]. By attenuating RNA Pol II, the protein was shown to favor alternate exon usage of all variable exons (8 in total) within the human *CD44* gene[145]. Loss of HP1γ either results in reduced usage of alternate exons within the *CD44* gene[145] or a general reduction of splicing[143]. To determine whether HP1γ depletion affects differential exon usage in MCF7 cells, RNA-seq data was analysed using the Bioconductor tool *DEXSeq*[231]. For a p-value threshold of 0.01 a total number of 882 genes with at least one differentially used exon were identified. As an example, the result for *CD44* is depicted in Appendix I (Figure 6A). Each of the 18 exons (10 constant (C) and 8 variable (V) exons) are plotted on the x-axis against the exon usage on the y-axis. This value is normalised to the expression level of the gene. The only significant

differential exon usage for this gene was detected for exon V2, which has higher relative abundance in transcripts of HP1γ knockdown cells compared to controls (log2 fold change = 1.2, adjusted p-value = 0.049). This result would suggest that loss of HP1γ causes increased exon usage of exon V2, which is the opposite result to published data[145]. However, it has to be noted that in general this analysis revealed highly inconclusive results. Figure 6B of the Appendix I depicts the result for the histone gene *HISTH4H*, the gene with the most significant changes. Virtually every exon is either significantly included (exon 1-4) or excluded (exon 5-6) (adjusted p-value < 0.009). For many genes similar, biologically rather unlikely results were observed.

DEXSeq is a powerful method to identify differential exon usage, however, to produce biologically meaningful results, RNA-seq data of great sequencing depth should be used. Our rather inconclusive results are probably due to limited sequencing depth and the small number of biological replicates used in this experiment (n=3). The effect of HP1γ depletion on alternative splicing in MCF7 cells was not pursued any further.

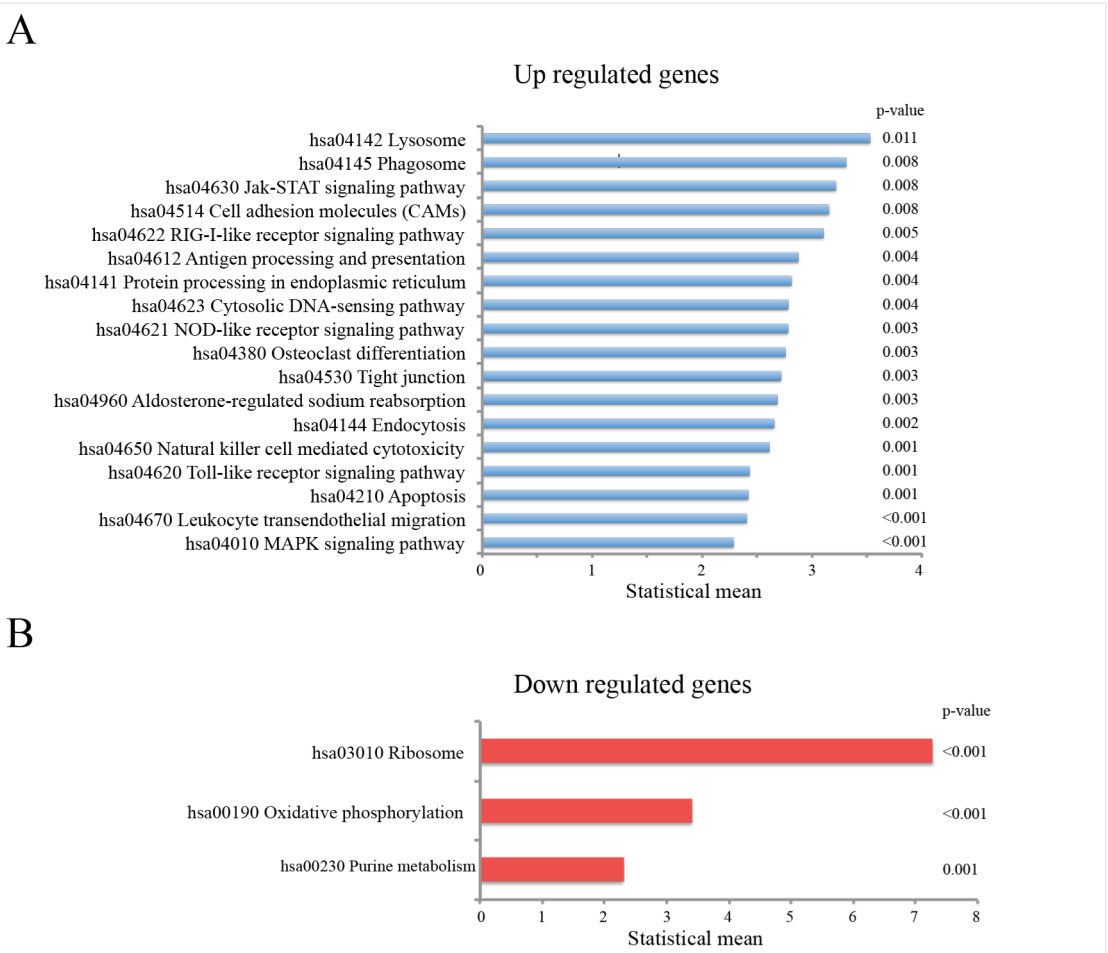
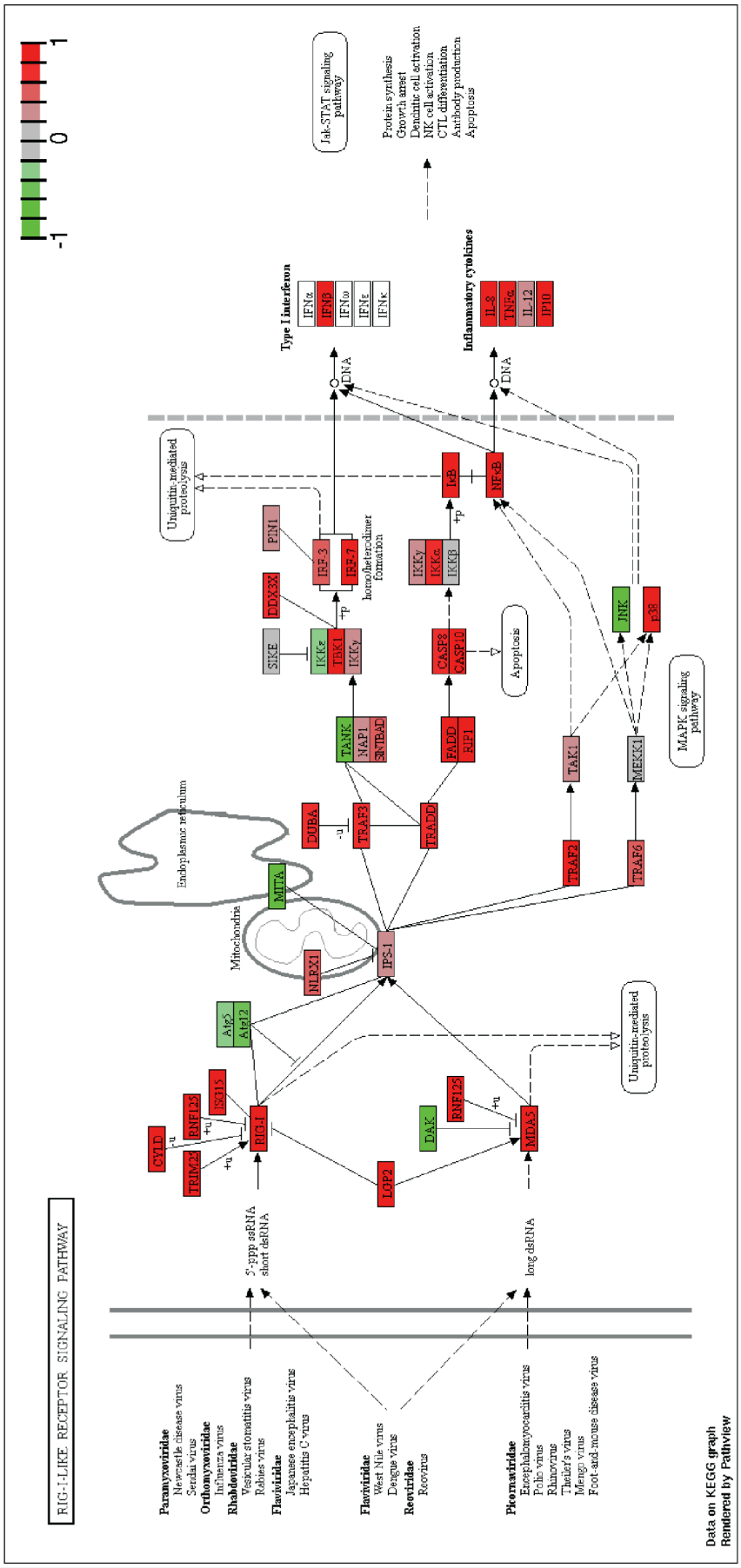


Figure 5-4 KEGG pathway analysis of MCF7 RNA-seq data.

KEGG pathway analysis performed using Bioconductor package for KEGG enrichment analysis. Input data: all differentially expressed genes (p-value <0.05). Statistical mean for each biological pathway class plotted, p-values indicated.

A Up-regulated genes. **B** Down-regulated genes

A


RIG-I-LIKE RECEPTOR SIGNALING PATHWAY
-1
0
1
Paramyxoviridae
Newcastle disease virus
Sendai virus
Orthomyxoviridae
Influenza virus
Rhabdoviridae
Vesicular stomatitis virus
Rabies virus
Flaviviridae
Japanese encephalitis virus
Hepatitis C virus
Flaviviridae
West Nile virus
Dengue virus
Reoviridae
Reovirus
Picornaviridae
Encephalomyocarditis virus
Polio virus
Rhinovirus
Theiler's virus
Mengo virus
Foot-and-mouse disease virus
5'-ppp ssRNA
short dsRNA
long dsRNA
CYLD
TRIM25
RNF125
ISG15
RIG-I
Atg5
Atg12
LGP2
DAK
RNF125
MDA5
NLRX1
IPS-1
Mitochondria
MITA
Endoplasmic reticulum
DUBA
TRAF3
TRADD
TRAF2
TRAF6
TANK
NAP1
SINTBAD
FADD
RIP1
TAK1
MEKK1
JNK
p38
MAPK signaling pathway
CASP8
CASP10
Apoptosis
SIKE
IKKε
TBK1
IKKγ
DDX3X
PIN1
IRF-3
IRF-7
homodimerization
formation
IKKγ
IKKα
IKKβ
IκB
NF-κB
+p
Ubiquitin-mediated proteolysis
Ubiquitin-mediated proteolysis
DNA
DNA
Type I interferon
IFNα
IFNβ
IFNω
IFNε
IFNκ
Inflammatory cytokines
IL-8
TNFα
IL-12
IP10
Jak-STAT signaling pathway
Protein synthesis
Growth arrest
Dendritic cell activation
NK cell activation
CTL differentiation
Antibody production
Apoptosis
Data on KEGG graph
Rendered by Pathview

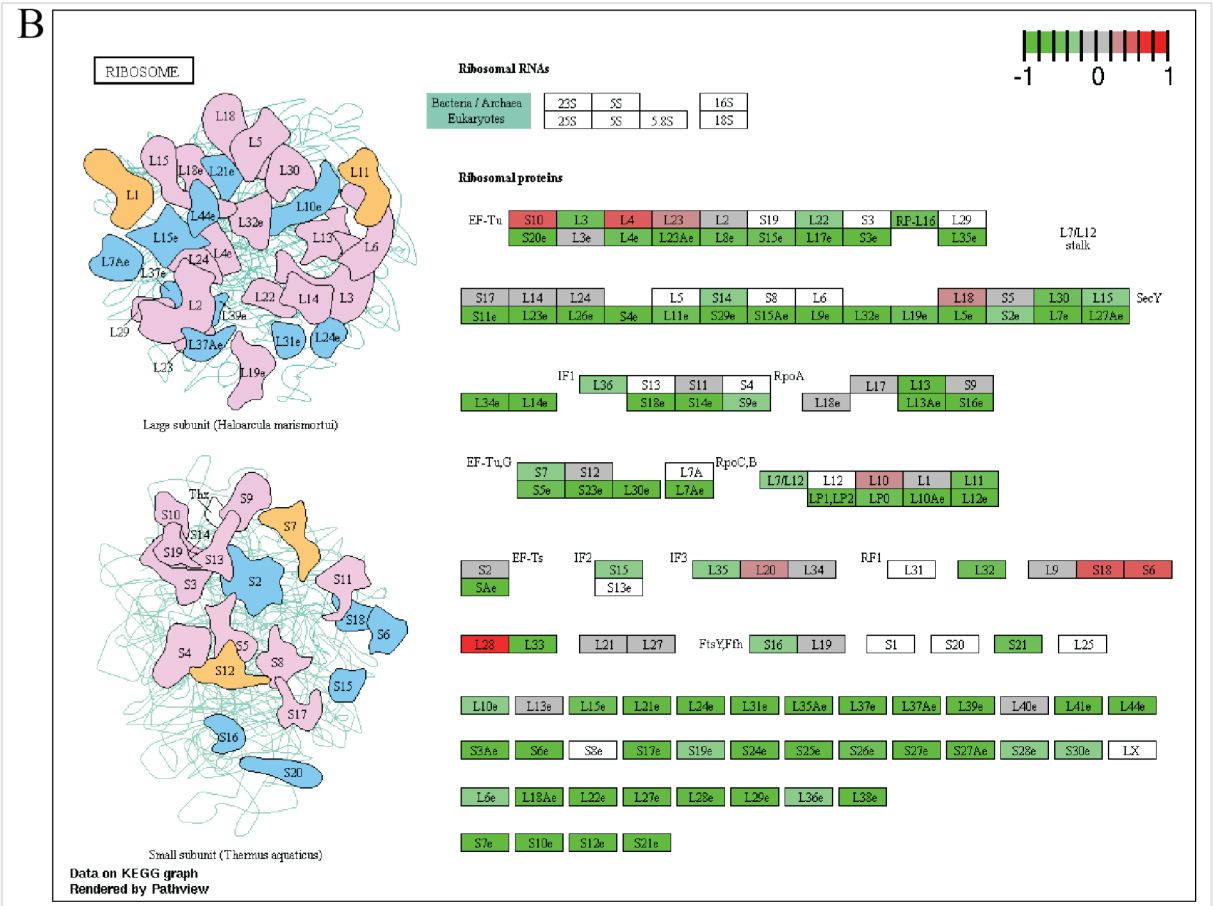


Figure 5-5 KEGG pathway maps.

Examples of KEGG pathway maps. Differentially expressed genes are highlighted: Up-regulated genes (red) and down-regulated genes (green).

A hsa04622: RIG-I-like receptor signaling

B hsa03010: Ribosome

5.3.3. Validation of RNA-seq in MCF7 cells

5.3.3.1. Activation of interferon response pathways

KEGG pathway analysis identified the most up-regulated genes to be enriched for pathways indicating activation of an interferon mediated immune response, such as JAK-STAT, RIG-I-like receptor, cytosolic DNA-sensing, NOD-like receptor and Toll-like receptor signaling. All of these pathways are functionally linked. Normally they are activated in response to acute viral or bacterial infection. Components of RIG-I-like receptor, NOD-like receptor and Toll-like receptor signaling pathways directly bind double stranded RNA or DNA and rapidly induce signaling cascades resulting in release of pro-inflammatory cytokines such as type I interferons (IFNs), interleukins and TNF ligands. These cytokines, in particular type I IFNs

(IFNα, β), in turn act as activators for the same pathways. Cells use these positive feedback mechanisms to assure immediate reaction to infections.

The JAK-STAT signaling is downstream of the above-mentioned pathways and directly activated by type I IFNs. STATs are monomeric transcription factors that reside within the cytoplasm. Upon IFN binding to its receptor, receptor-associated Janus kinases (JAKs) are activated and recruit monomeric STATs. JAK mediated phosphorylation of STATs facilitates their dimerisation and localisation to the nucleus where they act as transcription factors[232]. JAK-STAT signaling results in activation of genes inducing growth arrest and apoptotic pathways - phenotypes that relate to the physiological changes observed in HP1γ depleted MCF7 cells. Apart from JAK-STAT signaling all of these pathways result in the release of IFNβ. Other common cytokines induced by these pathways include TNFα, interleukin 6, 8, 10 (IL6, 8, 10), as well as other interleukins. To confirm activation of pathways involved in the interferon mediated immune response in MCF7 cells, mRNA levels of IFNβ (*IFNB1*), IL8 (*CXCL8*), TNFα (*TNF*) and IL28A/B (*IFNL2/3*) were analysed by RT-PCR over a time course of 6 days after infection with HP1γ shRNA vectors. Significant increase in mRNA expression of these genes was detected after 4 days with further increase at day 6 for both shRNAs targeting HP1γ compared to shCtrl cells (p-values <0.05, Figure 5-6A, Figure 5-6B). Effects were more pronounced for shHP1γ_1. The fact that expression of these cytokines was only induced after 4 days indicates that the effects are not directly due to infection with lentiviral shRNA vectors. In fact, gene expression changes can be correlated with down-regulation of HP1γ protein as assessed by immunoblotting (Figure 4-10A).

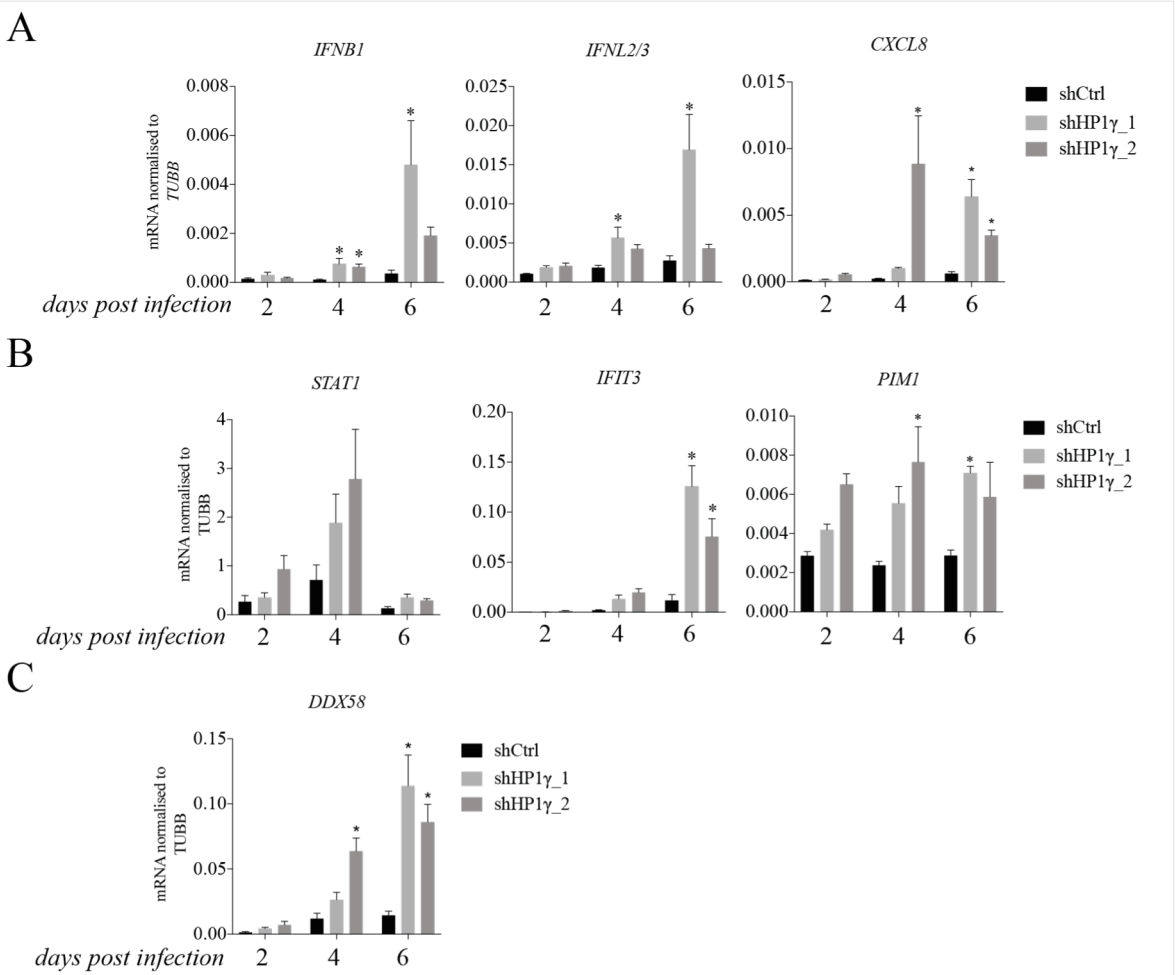


Figure 5-6 Expression of genes linked to interferon immune response in MCF7 cells. MCF7 cells were infected with shCtrl and shHP1γ_1 and 2 shRNA vectors and processed 2, 4 and 6 days post infection. mRNA levels of indicated genes were assessed at the 3 indicated time points by RT-PCR. Graphs show relative mRNA expression after normalisation to *TUBB* expression and include data from 3 independent experiments. Significance was assessed using one-way ANOVA tests, highest p-values per panel are indicated. SEM are indicated.

A Genes encoding regulatory immunity cytokines (* p-value <0.05).

B JAK-STAT target genes (* p-value <0.05).

C Gene encoding dsRNA sensing DEAD-box helicase DDX58 (* p-value <0.05).

The activity of the JAK-STAT signaling pathway was assessed by measuring expression of selected STAT target genes such as *PIM1* and *IFIT3*. Both *PIM1* and *IFIT3* show similar expression patterns to genes encoding cytokines (Figure 5-6B). mRNA expression was induced 4 days after infection with shRNA vectors in both knockdown cell lines (p-values <0.05). *IFIT3* encodes an interferon response protein. Differential expression analysis (Section 5.5.3) identified more than 10 such interferon response genes that were significantly up-regulated upon HP1γ knockdown (log2 >2, p-value <0.0001). Besides STAT target genes, up-regulation of several *STAT* transcription factor genes *(STAT1*, *2, 3* and *5A,* log2 >1.5, p-value <1E-6) was identified. I validated the mRNA levels for *STAT1* (Figure 5-6B). Interestingly, mRNA expression of *STAT1* mRNA peaked at day 4 and decreased significantly at day 6, indicating dynamic regulation.

Differential expression analysis identified up-regulation of genes encoding dsRNA sensing proteins, which are important components of the interferon immune response. Genes such as *DDX58 (*also called *RIG-I)* and *DDX60* both encoding DEAD box helicases recognising viral dsRNA were detected *(*log2 >2, p-value <1E-26). Furthermore, genes encoding viral dsRNA sensing oligoadenylate synthetases *OAS1, 2, 3* (log2 > 2, p-value <1E-19) or genes encoding members of the toll like receptor family *TLR1, 2, 3 (*log2 >2, p-value <0.03) were also detected. TLR1 and 2 are located in the cytoplasma membrane, predominantly detecting bacteria in the extracellular space, whereas TLR3 located in endosomal membranes detecting dsRNA. mRNA expression was validated for *DDX58* (Figure 5-6C), which was induced after 4-6 days in both knockdown cell lines (p-values <0.027).

Stimulation of an interferon immune response in MCF7 upon loss of HP1γ was further confirmed by immunoblotting using antibodies against phosphorylated STAT1. With both HP1γ targeting shRNAs an increase of STAT1 phosphorylation was observed 4 days after infection, whilst total STAT1 protein levels remained unchanged (Figure 5-7). Phosphorylation of STAT1 correlates with a decrease in HP1γ protein levels.

In conclusion, KEGG pathway analysis indicated that MCF7 cells activate signaling pathways linked to an interferon-mediated antiviral response upon HP1γ depletion. I confirmed the activation of these pathways by measuring mRNA expression of the cytokine genes *IFNB1*, *TNF*, *IL8* and *IL28A/B*. Furthermore, expression for a subset of JAK-STAT targets was

assessed, confirming the activation of JAK-STAT signaling, which is downstream of the IFN response. Induction of IFN response genes 4-6 days after infection with HP1γ shRNA vectors correlates with decreased HP1γ protein level.

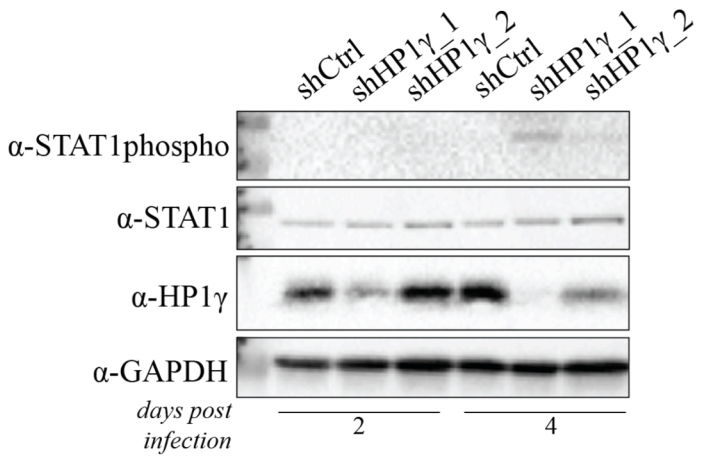


Figure 5-7 STAT1 phosphorylation in MCF7 cells upon HP1γ knockdown.
MCF7 cells were infected with shCtrl and shHP1γ_1 and 2 shRNA vectors and processed 2 and 4 days post infection. Immunoblot of MCF7 whole cell lysates, using indicated antibodies. Protein lysates representing 40000 cells were loaded per well.

5.3.3.2. Up-regulation of pro-apoptotic genes

Interferon mediated activation of apoptotic pathways is a well known self defense mechanism to prevent viral replication in virus-infected cells[233]. Given the clear activation of interferon-mediated pathways, it is not surprising that up-regulation of numerous pro-apoptotic genes was identified in HP1γ depleted MCF7 cells. Such genes include those encoding members of the pro-apoptotic BH3-only protein family (*BIK, BID* and *PUMA),* apoptosis inducing members of tumor necrosis factor (TNF) ligand family (*TNFSF10, TNF)* and their receptors *(TNFRSF10A/C),* a negative regulator of the 'inhibitor of apoptosis protein' IAP family (*XAF1*) family and genes encoding caspases (*CASP4, 10, 7)*. Furthermore, down-regulation of *BCL2,* encoding a member of the Bcl-2 protein family of anti-apoptotic proteins, was also observed. Both *XAF1* and *TNFSF10* (also called *TRIAL*) are direct targets of IFNβ[234,235]. mRNA levels of a subset of pro-apoptotic genes were followed over a time course of 6 days by RT-PCR using shHP1γ_1 and shHP1γ_2 mediated knockdown of HP1γ. Results are summarised in Figure 5-. Up-regulation of mRNAs encoding pro-apoptotic proteins *BID, BIK* and *PUMA* was observed after 2 days of knockdown with both shRNAs compared to shCtrl samples, with a significant increase after 4 days (Figure 5-A). A non-significant down-

regulation of the anti-apoptotic gene *BCL2* was observed on day 6 only with shHP1γ_1 (Figure 5-C). Expression of *XAF1*, *TNF* and *TNFSF10* mRNA was induced at day 4 with both shRNAs (Figure 5-B/C). In comparison, reduction of HP1γ protein levels were detected after 2 days with shHP1γ_1 and after 4 days with shHP1γ_2 (Figure 4-10A) indicating the observed gene expression changes can be correlated with a reduction of HP1γ protein levels. Notably, increased expression of pro-apoptotic genes was also observed over time in shCtrl cells, which may be due to puromycin selection mediated stress. To avoid selection stress, gene expression changes were further validated using siRNA-mediated knockdown of HP1γ in MCF7s cells (Figure 7, Appendix II). Gene expression changes resulted in an increase in protein expression for a subset of apoptosis factors. Immunoblotting of MCF7 whole cell lysates 4 days after infection with shRNA vectors confirmed an increase of XAF1, BID and PUMA protein levels (Figure 5-9).

In summary, upon HP1γ depletion, MCF7 cells showed clear up-regulation of pro-apoptotic genes. Induction of apoptotic genes could be a direct result of interferon mediated signaling.

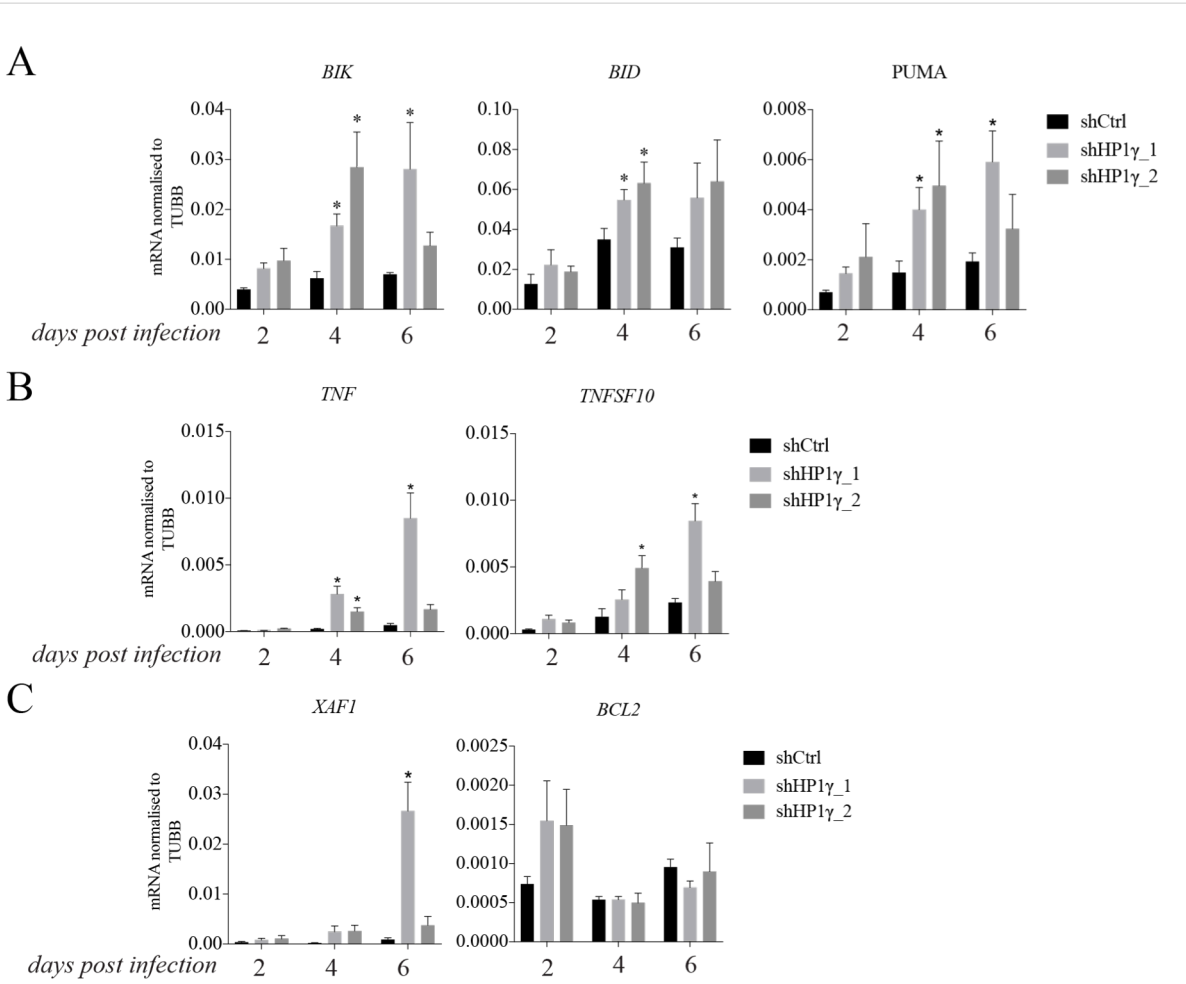


Figure 5-8 mRNA expression of apoptotic genes in MCF7 cells.
MCF7 cells were infected with shCtrl and shHP1γ_1 and 2 shRNA vectors and processed 2, 4 and 6 days post infection. mRNA levels of indicated genes were assessed at the 3 indicated time points by RT-PCR. Graphs show relative mRNA expression after normalisation to *TUBB* levels and include data from 3 independent experiments. Significance was assessed using one-way ANOVA tests, highest p-values per panel are indicated. SEM are indicated.
A Genes encoding members of pro-apoptotic BH3-only protein family (*p-value <0.04).
B Genes encoding members of family of tumor necrosis factor ligands (*p-value <0.04).
C Genes encoding pro-apoptotic factor XAF1 and anti-apoptotic factor BCL2 (*p-value <0.04).

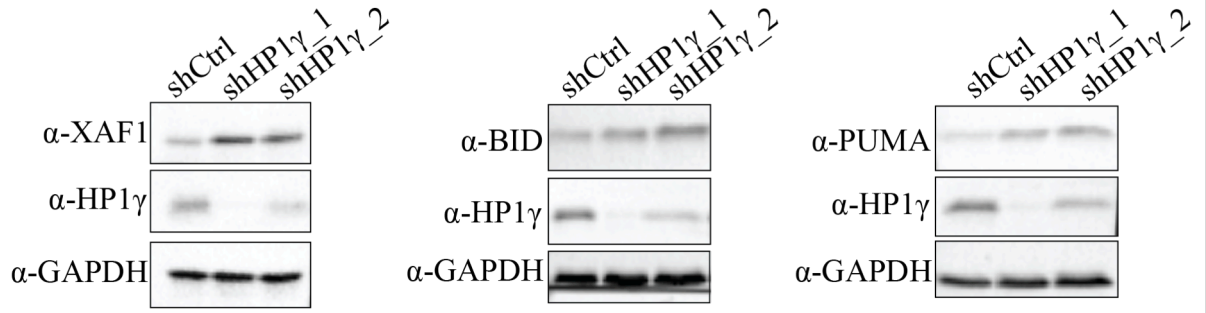


Figure 5-9 Protein expression of apoptotic factors in MCF7 cells.
MCF7 cells were infected with shCtrl and shHP1γ_1 and 2 shRNA vectors and processed 4 days post infection. Immunoblot of MCF7 whole cell lysates, using indicated antibodies. Protein lysates representing 40000 cells were loaded per well.

5.4. Discussion

In this chapter, I presented whole transcriptome analysis in MCF7 cells upon HP1γ depletion. Differential expression analysis identified 4673 de-regulated genes with an equal split between up- and down-regulation. Subsets of genes linked to phenotypic changes determined in Chapter 5 were validated by RT-PCR.

5.4.1. Interferon response

To confirm activation of pathways involved in the interferon mediated immune response in MCF7 cells, induction of genes encoding IFNβ, IL8, TNFα and IL28A/B was analysed. Gene expression was induced 4-6 days after infection with shRNA vectors, which can be correlated with a reduction of HP1γ protein levels observed at these time points. I also confirmed activation of a subset of JAK-STAT target genes that are downstream of interferon signaling pathways. Upon stimulus, STAT transcription factors are phosphorylated and activated by receptor-associated JAK[232]. Activation of the JAK-STAT pathway was confirmed by by immunoblotting with antibodies recognising phosphorylated STAT proteins 4 days after infection with shRNA vectors. Besides JAK-STAT target genes, induction of genes encoding STAT transcription factors was detected. In response to acute viral infection JAK-STAT signaling was shown to activate expression of *STAT1* and *3* to reinforce the immune response[236]. This positive feedback mechanism could be an explanation for elevated *STAT1* expression and further emphasizes that MCF7 cells fully activated the interferon mediated immune response.

HP1γ has been associated with regulating expression of interferon response genes in mouse embryonic fibroblasts (MEFs)[132,141]. HP1γ regulates the transcription rate of immune response genes, such as *CXCL8,* in response to bacterial infection or phorbol ester PMA treatment, by directly interacting with RNA Pol II[ref.141]. Only upon loss of HP1γ do these genes become hyper activated in response to stimuli, resulting in a full-blown immune response. Interestingly, knockdown of HP1γ alone induced selected immune related genes, however, did not lead to full activation of the interferon immune response. Similar results were observed in HeLa cells depleted of HP1γ[132]. My data suggests that in un-stimulated MCF7 cells, loss of HP1γ is sufficient to induce an immune response normally activated in response to viral infection.

Another recent report demonstrated a role for HP1γ in regulating the tumor immune environment[161]. In CD8 positive T cells HP1γ regulates the expression of effector T cell factors such as perforin, granzyme B and the inflammatory cytokine interferon γ, which are all necessary to kill target cells[161]. HP1γ deficient CD8 positive T cells express high levels of these factors and therefore harbor increased tumor-killing capacity. In fact tumour burden of mice treated with HP1γ deficient CD8 positive T cells was reduced, indicating that targeting HP1γ in CD8 positive T cells may be a novel therapeutic approach to control growth of solid tumours[161].

Both of these studies clearly link HP1γ to the regulation of immune response genes. The question whether HP1γ directly regulates expression of immune response genes in MCF7 cells will be addressed in Chapter 6, which includes a genomic analysis of HP1γ in MCF7 cells.

5.4.2. Activation of apoptosis

Depletion of HP1γ impairs proliferation and reduces cell viability of MCF7 cells (compare Section 4.2.). Besides morphological changes, MCF7 show an increase of apoptotic markers such as cleaved PARP1 and elevated levels of DNA fragmentation (Figure 4-10). These results indicate that MCF7 cells activate classical apoptotic pathways upon loss of HP1γ. In fact, KEGG pathway analysis identified enrichment of up-regulated genes associated with pro-apoptotic pathways (Figure 5-4). I validated increased expression of genes encoding

members of the pro-apoptotic BH3-only protein family *BIK, BID* and *PUMA* at mRNA. In each case expression was induced 4-6 days after infection with shRNA vectors, which correlates with reduced HP1γ protein levels. Similar results were obtained for the activation of the tumour necrosis factor ligand *TNFSF10* (TRAIL). Based on this analysis it is difficult to conclude which apoptotic pathway is activated. Up-regulation of TRIAL indicates activation of the extrinsic, while up-regulation of BH3-only proteins indicates activation of the intrinsic apoptotic pathway. Given the plethora of signaling cascades activated in HP1γ depleted MCF7 cells, it is likely that both apoptotic pathways are triggered simultaneously.

Up-regulation of pro-apoptotic genes is in agreement with the phenotypic changes observed in MCF7 cells upon HP1γ depletion. Both phenotypic changes and effects on gene expression were less robust with shHP1γ_2, which depletes HP1γ less efficiently than shHP1γ_1. Differences between the two shRNAs could suggest a dose dependent effect of HP1γ depletion. To ultimately confirm that phenotypic changes are HP1γ dependent, it will be crucial to perform HP1γ rescue experiments. In addition to rescuing with wildtype HP1γ protein, rescue experiments with HP1γ mutant proteins would provide interesting insights into which domain of the protein is required for its function.

Loss of HP1γ induces apoptosis in non-small cell lung cancer (NSCLC) cell lines[237]. Phenotypic changes in this study were correlated with increased protein levels of growth arrest and DNA-damage-inducible protein alpha (GADD45α) and BAX, another member of the BH3-only protein family. The same genes were only slightly up-regulated in MCF7 cells (log2=0.8 p-value <0.015, log2=0.6 p-value <0.015, respectively), indicating different apoptotic pathways are activated in HP1γ depleted MCF7 cells. Besides, a correlation with GADD45α and BAX expression data in NSCLC cell lines did not reveal any mechanistic insights into how HP1γ depletion induces apoptosis[237].

Given the direct link between the interferon response and activation of apoptosis, it is possible that MCF7 cells undergo apoptosis as a consequence of triggering interferon signaling. In fact, upon loss of HP1γ, I observed a clear activation of *XAF1* and *TNFSF10* expression both of which are direct targets of IFNβ[234,235]. In support of this hypothesis is a recent study analysing the influence of IFNβ, produced by adipose tissue-derived mesenchymal stem cells (ASCs), on MCF7 cell physiology[238]. It showed that ASCs release IFNβ into the culture

medium. Culturing MCF7 cells in this conditioned media resulted in impaired cell proliferation and induction of apoptosis[238]. This study demonstrates that IFNβ stimulation is sufficient to trigger the same phenotypic changes I detect in MCF7 cells upon HP1γ depletion. In order to understand whether apoptosis is activated via IFNβ signaling, HP1γ depleted cells could be treated with neutralizing IFNβ antibodies or JAK (JAK1 or JAK2) inhibitors to inhibit JAK-STAT downstream signaling. Alternatively, time course experiments with more time points between 2 and 6 days after infection with shRNA vectors could help to dissect the order of gene activation events in more detail.

5.4.3. Proliferation

Depletion of HP1γ severely impaired proliferation of MCF7 cells, however no significant effect on cell cycle progression was detected. Control of proliferation usually occurs during G1 phase of the cell cycle. Several different extrinsic and intrinsic cues influence the cell's decision to enter S phase. Progression through G1 phase is controlled by members of the retinoblastoma (pRB) proteins, which function to repress E2F-dependent transcription. Phosphorylation of pRB proteins causes them to dissociate from E2F at promoters, thereby de-repressing E2F transcription factors and allowing them to activate expression of various genes whose products stimulate entry into S phase. pRB phosphorylation is regulated by three CDKs: cyclin-D-dependent CDK4 and CDK6 and cyclin-E-dependent CDK2[239]. These cyclin-CDK complexes are tightly controlled by CDK inhibitors (CDKIs) such as p21, p27 or members of the INK4 family of CDKIs (p15, p16, p18 or p19). Transcriptional regulation of G1 cyclin (D and E) genes adds another level of complexity, which is mediated through signaling cascades induced by mitogens, cytokines or other signaling molecules[239].

Differential expression analyses did not provide any clear insight into how cell proliferation is affected in MCF7 cells. Gene ontology analysis identified down-regulation of genes involved in DNA-replication, indicating that DNA duplication in S phase may be affected. In contrast, KEGG pathway analysis identified activation of MAPK signaling, which suggests an activation of pro-proliferative signaling cascades. Manual inspection of the data identified up-regulation of genes encoding CDKIs such as *CDKN1A* (p21; log2=1 p-value<1E-4), *CDKN1B* (p27; log2=0.7 p-value< 0.001) and *CDKN2D* (p19; log2=1.6 p-value<0.001), all of which repress G1 to S phase transition. On the contrary, up-regulation of cyclin D and E

genes *CCND3* (log2=1.6 p-value <1E-5) and *CCNE1* (log2=0.9 p-value<0.001) was observed, which should have the opposite effect. Based on this differential expression analysis in MCF7 cells, it is unclear how cell proliferation is affected. Deregulation of genes could be due to feedback mechanisms and/or directly linked to the strong immune response that was detected in MCF7 cells upon HP1γ depletion. Overall, given these results, it seems likely that impaired cell proliferation is predominantly due to induction of apoptosis.

5.4.4. Cell adhesion

KEGG pathway analysis identified up-regulation of genes associated with cell adhesion molecules. Amongst these were genes encoding cell adhesion molecules such as ICAM1 and 4; CEACAM1, 5 and 6; L1CAM and BCAM. Several genes encoding focal adherens components such as integrin subunits (ITGA3, 5, 9 and ITGB1, 6, 8), cadherins (CDH15, 18 and 5) and selectin (SELE) were also significantly up-regulated. Furthermore up-regulation of genes encoding tight junction components such as claudins (CDLN 1, 3, 4, 23, 7 and 9) and occludin (OCLN1) were identified. This list of up-regulated genes strongly suggests that cell morphology and/or cell mobility of MCF7 cells is affected upon HP1γ depletion. However, migration assays with MCF7 cells showed that loss of HP1γ had no effect on their cell migration. As MCF7 cells have low intrinsic mobility, an effect on cell migration will be difficult to measure. Intriguingly, HP1γ depleted MDA-MB-231 cells, which are of metastatic origin, showed a clear effect on cell migration. As most of the above mentioned genes have direct links to cancer metastasis, it would be very interesting to determine expression of cell adhesion genes, identified in MCF7 cells, in MDA-MB-231 cells. Beyond this candidate approach, a comprehensive comparison of whole transcriptome analysis in MDA-MB-231 and MCF7 cells upon HP1γ depletion will be very important to identify cell type specific roles of HP1γ and may elucidate a function for HP1γ in breast cancer invasiveness. RNA-seq analysis in MDA-MB-231 cells before and after HP1γ depletion is currently ongoing.

To summarise this Chapter, whole transcriptome analysis in MCF7 cells upon HP1γ depletion identified differentially expressed genes. Gene enrichment analysis identified clear activation of IFNβ mediated signaling pathways and activation of apoptosis. Together with a down-regulation of genes enriched for ribosomal and mitochondrial functions, this analysis indicates that upon HP1γ depletion, MCF7 cells are under extreme physiological stress and shut down

biological processes that are important for cell homeostasis. Gene expression changes are in agreement with phenotypic changes observed in HP1γ depleted MCF7 cells.

In Chapter 6, I will present the results from HP1γ chromatin immunoprecipitation (ChIP) analysis linked to next generation sequencing (ChIP-seq) using samples isolated from MCF7 cells. ChIP-seq data will be cross-compared with whole transcriptome data discussed in this chapter to understand whether gene expression changes are directly mediated by HP1γ.

Chapter 6

Genomic analysis of HP1γ in breast cancer cells

6.1. Introduction

HP1γ is a versatile transcriptional regulator that can operate either as a repressor or a transcriptional activator. In contrast to HP1α and β, HP1γ is predominantly associated with euchromatin and it localises to sites of active transcription[73,135,143,145]. HP1γ regulates alternative splicing of specific gene transcripts by directly interacting with components of the splicing machinery[135,143]. Furthermore, it interacts with and mediates RNA Pol II action resulting in different transcriptional outcomes. It was proposed that HP1γ enhances RNA Pol II's transcription rate, suggesting it functions as a transcriptional activator[143]. Other studies, however, associate HP1γ with 'slowing down' transcription via its interaction with the elongating form of RNA Pol II, which either facilitates alternate exon inclusion of nascent transcripts[145], facilitates transcriptional termination[147] or 'fine tunes' transcription of inducible immune response genes[141] (see Section 1.2.8.1.2). The means by which HP1γ regulates RNA Pol II activity mechanistically remains, however, an unanswered question.

In this chapter I explore the function of HP1γ in regulating gene expression in MCF7 cells. Whole transcriptome data presented in Chapter 5 will be cross-referenced with the genomic localisation of HP1γ as determined here by ChIP-sequencing.

6.2. Aims of this chapter

This chapter has the following aims related to the genomic localisation of HP1γ and its role in regulating gene expression.

1. To determine the genomic locations of HP1γ in MCF7 cells

2. To determine differentially expressed genes directly bound by HP1γ

3. To investigate whether HP1γ is involved in regulating ZNF genes

6.3. Using ChIP-seq to determine HP1γ bound loci in MCF7 cells

6.3.1. ChIP protocol for HP1γ

In Chapter 4, I showed that approximately 33% of HP1γ is located in the nucleo-cytoplasmic fraction of MCF7 cells indicating a significant proportion of HP1γ is not tightly associated with chromatin (Section 4.3.1). Furthermore, my analysis on the dynamics and subcellular localisation of HP1 molecules in mESCs revealed that HP1γ molecules are more dynamic and less tightly bound to chromatin than other HP1 isoforms (presented in Section 3.10). Therefore, to avoid immunoprecipitation (IP) of freely diffusing HP1γ proteins and to enrich for chromatin bound HP1γ, I extracted chromatin before pursuing the standard ChIP protocol. I cross-linked adherent cells with 2% (v/v) formaldehyde for 10 min. After neutralisation with 125mM glycine, cells were lysed in IPH chromatin extraction buffer supplemented with 0.1% (w/v) SDS. Chromatin pellets were then separated from the nucleo- and cytoplasmic fraction by centrifugation. Chromatin was then resuspended in ChIP lysis buffer and subjected to sonication. Sonication conditions were optimized for MCF7 cells. To test sonication efficiency, sheared DNA was extracted and resolved in a 1% (w/v) agarose gel. DNA was visualised following ethidium bromide staining of the gel. I tested different sonication times ranging from 30min to 1.5h (data not shown). The best sonication result was obtained with 1.5h of total sonication time (at 5sec ON/ 5sec OFF and 100% amplitude; Qsonica™ sonicator). A representative image of sonicated DNA using these conditions is depicted in Figure 6-1A. Compared to unsonicated, genomic DNA of high molecular weight, the sonicated DNA size ranged from 200 to 800bp, indicating a resolution ranging from mono- to tetra-nucleosomes. Sheared chromatin was then subjected to immunoprecipitation (IP) using a HP1γ specific antibody. The antibody was tested for IP under ChIP buffer conditions. The IP efficiency was confirmed by immunoblotting using the same antibody (Figure 6-1B). IP with the HP1γ antibody showed an enrichment of HP1γ protein over input. A mouse IgG isotype control was used as a negative control. The ChIP protocol using this antibody was then followed as described in Section 2.10.1.

6.3.2. ChIP-seq library preparation

For the ChIP-sequencing experiment, ChIP was performed in untreated MCF7 cells with antibodies against HP1γ, H3K4me3, H3K36me3 and GFP, the latter being an IgG isotype

control. Histone modifications H3K4me3 and H3K36me3 were selected in order to identify actively transcribed genes. Samples were prepared in technical replicates and 1% of total input was used. Library preparation was performed as described in Section 2.10.3. The average size of final libraries and their DNA concentrations were determined and a total of 10 libraries were multiplexed at equal molarities. Libraries were analysed by 50bp, paired-end sequencing on an Illumina HiSeq 1500® system.

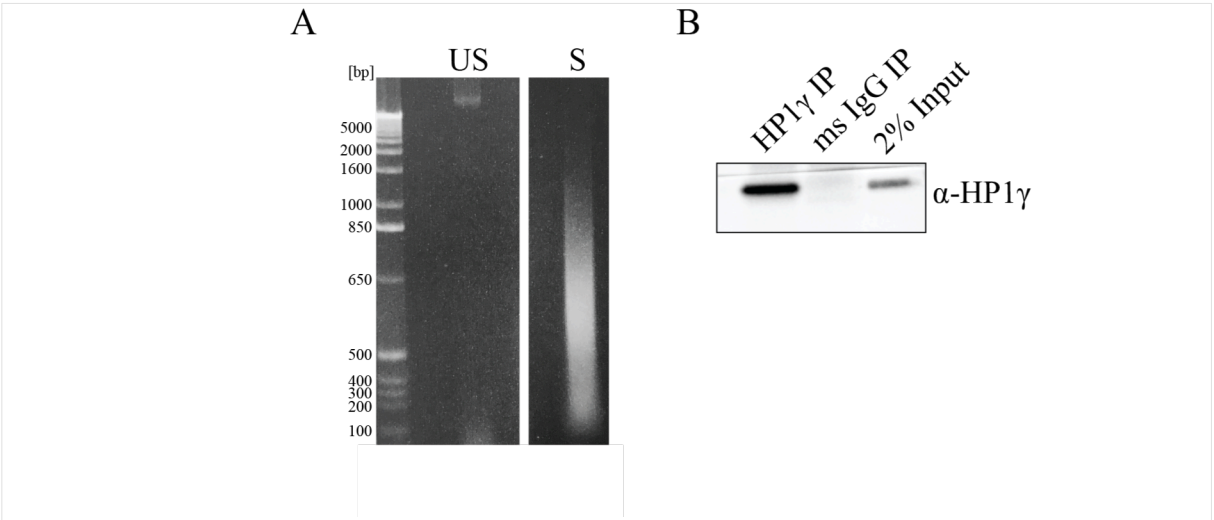


Figure 6-1 ChIP protocol adapted to HP1γ.
2*10^7 MCF7 cells were cross-linked using 1% (v/v) formadehyde (10 min), neutralised with 125 mM glycine and lysed in IPH buffer (+ 0.1% (w/v) SDS). Chromatin was separated from nucleo-cytoplasmic fraction by centrifugation. Chromatin was then resuspended in ChIP lysis buffer and sonicated for 1.5 h of total sonication time (at 5 sec ON/ 5 sec OFF and 100% amplitude; Qsonica sonicator).
A Sonicated DNA was phenol-chloroform extracted and resolved on 2%(w/v) agarose gel stained with ethidium bromide. DNA ladder is indicated [bp].
B Immunoblot using HP1γ antibodies of immunoprecipitated material from HP1γ and mouse IgG isotype antibody immunoprecipitation (IP) performed under ChIP conditions. 2% of total input lysate was analysed alongside IP samples.

6.3.3. ChIP-seq analysis

Read count of uniquely mapped reads per sample are summarised in Table 6-1. Processed reads for each replicate were mapped against the human reference genome and peaks were called on combined replicates against input samples (compare Section 2.10.3). With further filtering, peaks were cleared from those overlapping with an IgG peak and enriched for significant peaks (q-value <0.05) only. A total of 7637 peaks before and 7596 after filtering were identified for HP1γ. In order to estimate the similarity between replicates, pairwise

correlation analysis was performed (Appendix II, Figure 8). Pearson correlation coefficients (R) are indicated for each sample group. For HP1γ a correlation of R=0.8 and for both histone modifications correlations of R>0.9 were obtained, indicating that replicates showed high degrees of similarity.

To understand HP1γ's role in direct regulation of gene expression, as a primary approach ChIP-seq analysis was focussed on all HP1γ promoter bound genes (n=3859). The analysis of genes with HP1γ binding exclusively within gene bodies is currently on going and will not be discussed in this chapter.

Table 6-1 Total read count of ChIP-seq data. Number of uniquely mapped reads per replicate are depicted for each sample.

Antibody	Rep 1	Rep 2
HP1γ	28940004	17711924
H3K4me3	23395080	14518886
H3K36me3	21830521	15089423
IgG	2501319	3580514
Input (1%)	22795178	24171690

6.3.4. Distribution of ChIP-seq peaks

Distributions of HP1γ, H3K4me3 and H3K36me3 ChIP-seq signal over all genes with a HP1γ peak 2kb up- and/or downstream of the TSS (n = 3859) were visualised as metagene coverage profile plots and heatmaps (Figure 6-2). Each sample is scaled proportionally to the mapped library size and all transcripts overlapping a peak are scaled to a uniform size of 5kb (plus 2kb upstream of the TSS and downstream of the TES). Both metagene plot (Figure 6-2A) and heatmap (Figure 6-2B) show similar results. As expected, the signal for H3K4me3, a marker for active genes, peaks at the TSS. Also H3K36me3, a histone mark found in the gene body of actively transcribed genes, shows the expected profile of being absent at the TSS, but increasing over the gene body. For HP1γ the signal peaks at the TSS and, at a lower level, continues over the gene body. The heatmap clearly illustrates that for a subset of genes HP1γ binding is also enriched throughout the gene body (Figure 6-2B).

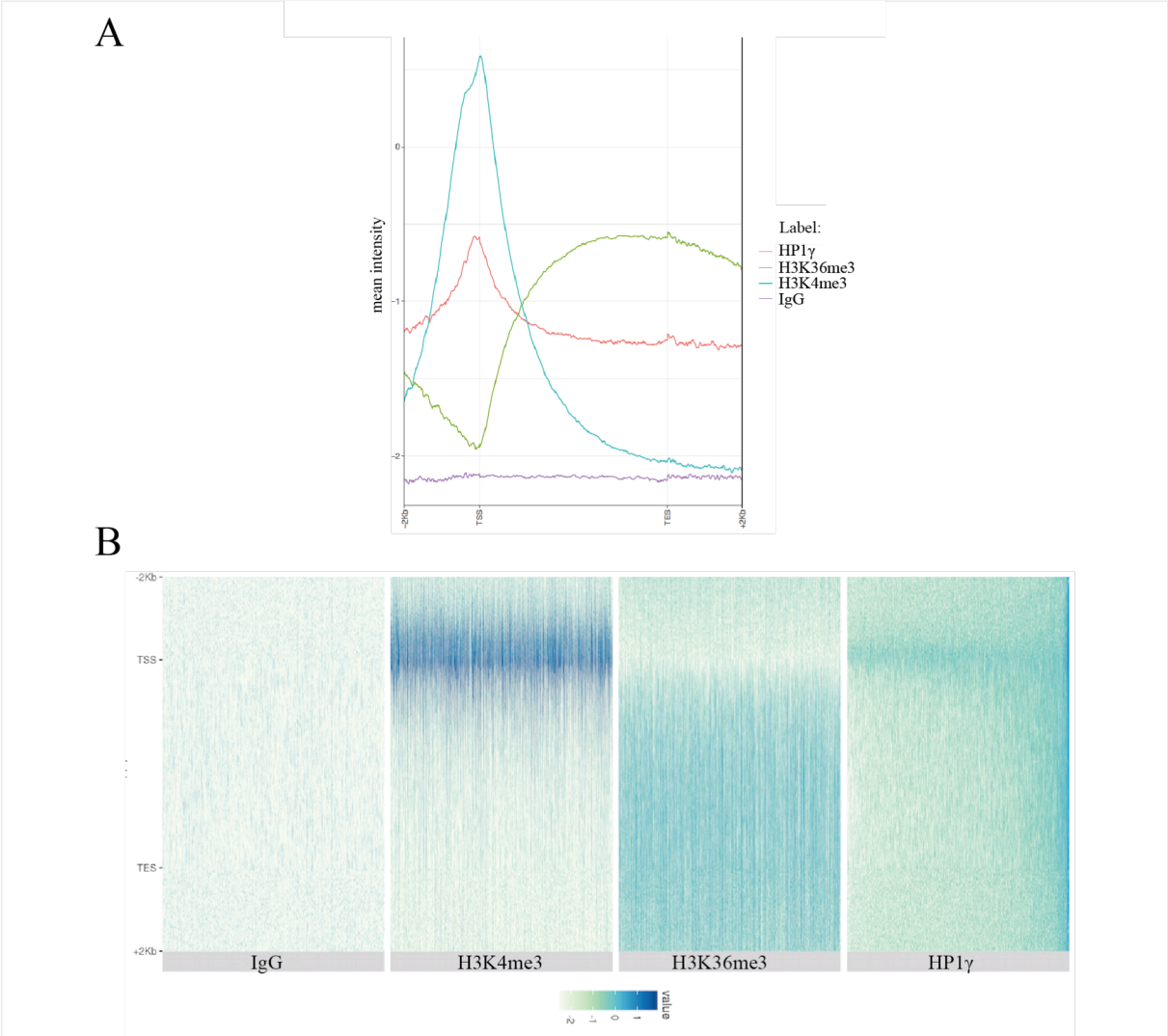


Figure 6-2 Distribution analysis of ChIP-seq peaks of HP1γ bound genes.
Metagene coverage analysis of HP1γ, H3K4me3 and H3K36me3 ChIP-seq data restricted to genes with a HP1γ peak 2kb up- and/or downstream of the TSS (n=3859). The signal for individual samples was scaled proportionally to mapped library size and log10 transformed. All transcripts overlapping a peak are scaled to a uniform size of 5kb (plus 2 kb upstream of the TSS and downstream of the TES).
A Metagene coverage plot illustrating the mean distribution of all peaks.
B Visualisation of all transcripts overlapping a peak as a heatmap. Colour code represents normalised ChIP-seq signal.

These data clearly show that the majority of genes bound by HP1γ, are also marked by H3K4me3 and H3K36me3, indicating HP1γ is predominantly marking actively transcribed genes. This observation is in agreement with published data sets for HP1γ using the same antibody but in different cell types[143].

6.3.5. Gene ontology analysis of HP1γ bound genes

To understand whether HP1γ bound genes are functionally related, gene ontology (GO) enrichment analysis was applied using the topGO Bioconductor package. Results are summarised in Figure 6-3. For each *Molecular Function* GO term, the dots indicate its enrichment score (x-axis), statistical significance (colour) as well as the absolute number of HP1γ bound genes in the category (size). Among the most significant GO terms this analysis identified enrichment in 'DNA binding' (GO: (GO:0003677, adjusted p-value<10^{-30}), 'metal ion binding' (GO:0046872, adjusted p-value<10^{-30}) and 'transcription factor activity' (GO:0003700, adjusted p-value<10^{-30}). Intriguingly, the majority of genes within these enrichment classes encode zinc finger (ZNF) proteins. Indeed, a total number of 290 ZNF genes were bound by HP1γ. Therefore, I decided to study this subgroup of HP1γ bound genes in more detail.

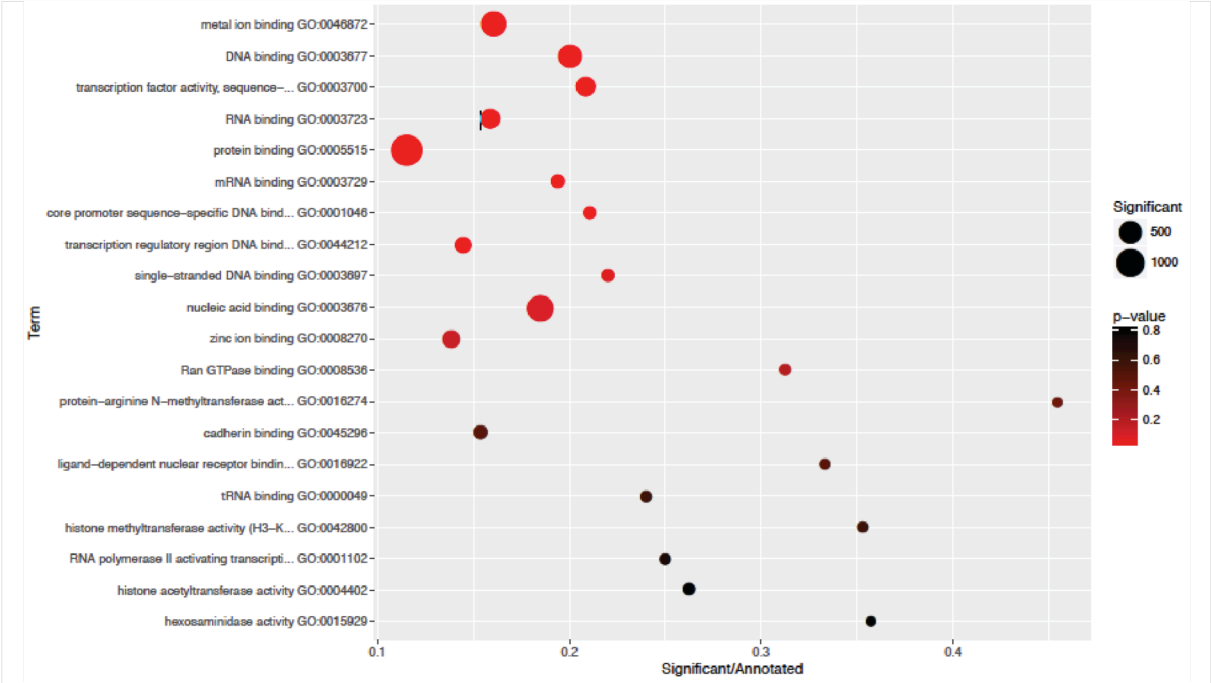


Figure 6-3 Gene ontology analysis of HP1γ bound genes.
The plot reports the top 20 Molecular Function GO terms with the strongest enrichment for HP1γ promoter bound genes (n=3859) obtained with the topGO Bioconductor package. Each dot represents a different GO term (y-axis) with its enrichment score (x-axis). Size of dots indicates the absolute number of gene targets that are annotated within the given class. Colour coding represents adjusted p-value: significant (red), less significant (black).

6.3.6. HP1γ is enriched at ZNF gene clusters

Out of 937 ZNF genes annotated in the human genome, I identified 290 (31% of total) bound by HP1γ. ZNF genes encode the biggest group of transcription factors with almost half of them (>400) encoding Krüppel-associated box (KRAB)-domain containing ZNF proteins. Amongst the 290 genes, ZNF genes of different classes such as C2H2- and KRAB ZNF genes were identified. Metagene analysis of HP1γ's distribution over all annotated ZNF genes (n=937) showed that besides enrichment at the TSS, HP1γ signal is enriched throughout the gene body with a significant increase towards the 3'-exon. (Figure 6-4A). A closer inspection of the data visualised in the UCSC genome browser revealed that HP1γ is enriched across entire ZNF gene clusters (Figure 6-4B).

An example of the UCSC tracks over a ZNF gene cluster located on the p-arm of chromosome 19 is shown in Figure 6-4B. This figure also reports Encode ChIP-seq data for the HP1γ interactor KAP1[240], which was also described to bind ZNF genes in a very similar fashion[241]. Whereas HP1γ peaks are broad and extended over the whole locus, KAP1 peaks are more defined. Figure 6-4B illustrates that the signal for both factors clearly defines the borders of this particular locus. Both KAP1 and HP1γ predominantly bind to 3'-exons of ZNF genes, with additional peaks at the TSS[241]. Figure 6-4C depicts UCSC tracks zoomed in on two divergent genes located in a ZNF gene cluster on chromosome 16 (*ZSCAN32* and *ZNF174)* demonstrating enrichment of both factors over the 3'-exons. However, only HP1γ shows additional enrichment at the TSS spanning the promoter regions of both genes.

HP1γ ChIP-seq results were validated by RT-PCR using specific primers spanning HP1γ bound regions. Validation experiments were performed in wildtype MCF7 cells compared to HP1γ CRISPR knockout cells. HP1γ knockout cells were generated by Katharina Lindner - a master's student whom I supervised during my PhD. CRISPR knockout was achieved using the lentiviral CRISPR_V2 vector system with two independent guide RNAs targeting exon 2 or exon 4 of *CBX3*. Clonal cell lines for each guide RNA were generated from single cells. A summary of the cell line derivation is depicted in Appendix II (Figure 9). As a control, a non-targeting guide RNA was used. Hereafter I refer to control cells as 'wildtype cells' (WT). Knockout of HP1γ was confirmed by immunoblotting using a HP1γ antibody.

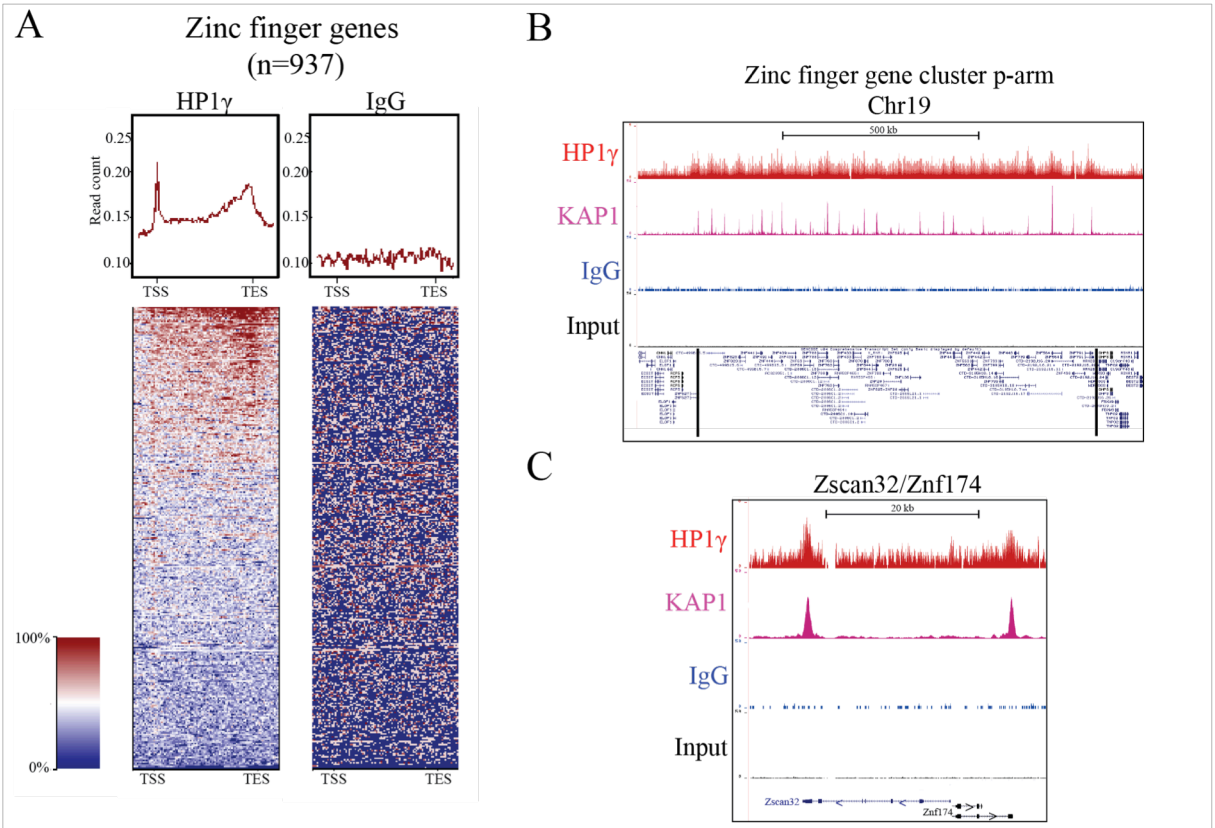


Figure 6-4 HP1γ - enrichment over ZNF gene clusters.

A Metagene coverage plots of HP1γ reads over all annotated ZNF genes (n = 937). Individual samples are scaled proportionally to mapped library size. All transcripts overlapping a peak are scaled to a uniform size of 5kb (plus 2kb upstream of the TSS and downstream of the TES). Colour scheme: ranging from 100% enrichment (red) to 0% enrichment (blue).

B UCSC genome browser snapshot of normalised HP1γ, KAP1 (Encode data set), IgG peaks over ZNF cluster on p-arm of chromosome 19. ChIP-seq peaks of combined replicates are shown. Boundaries of ZNF gene cluster are indicated by black lines.

C UCSC genome browser snapshot of HP1γ, KAP1 (Encode data set), IgG peaks over example genes *ZSCAN32* and *ZNF174* located on chromosome 16. ChIP-seq peaks of combined replicates are shown.

Figure 6-5A shows the results from two wildtype clones (WT_1 and 2) and one knockout clone for each guide RNA (KO_1 and 2), all of which were used for validation experiments. RT-PCR primers were designed against the 3'-exon of ZNF genes. Experiments were performed with two biological replicates. RT-PCR results for one representative experiment including 9 different ZNF genes are summarised in Figure 6-5. For each gene, the percentage binding relative to input is plotted for HP1γ and an IgG control. As a positive control HP1γ enrichment at pericentric α-satellite repeats was determined (Figure 6-5B). ZNF genes were separated in to 3 different classes: (i.) located within Chr19 ZNF cluster (Figure 6-5C), (ii.) Chr16 ZNF cluster (Figure 6-5D) or (iii.) located outside of ZNF gene clusters (Figure 6-5E).

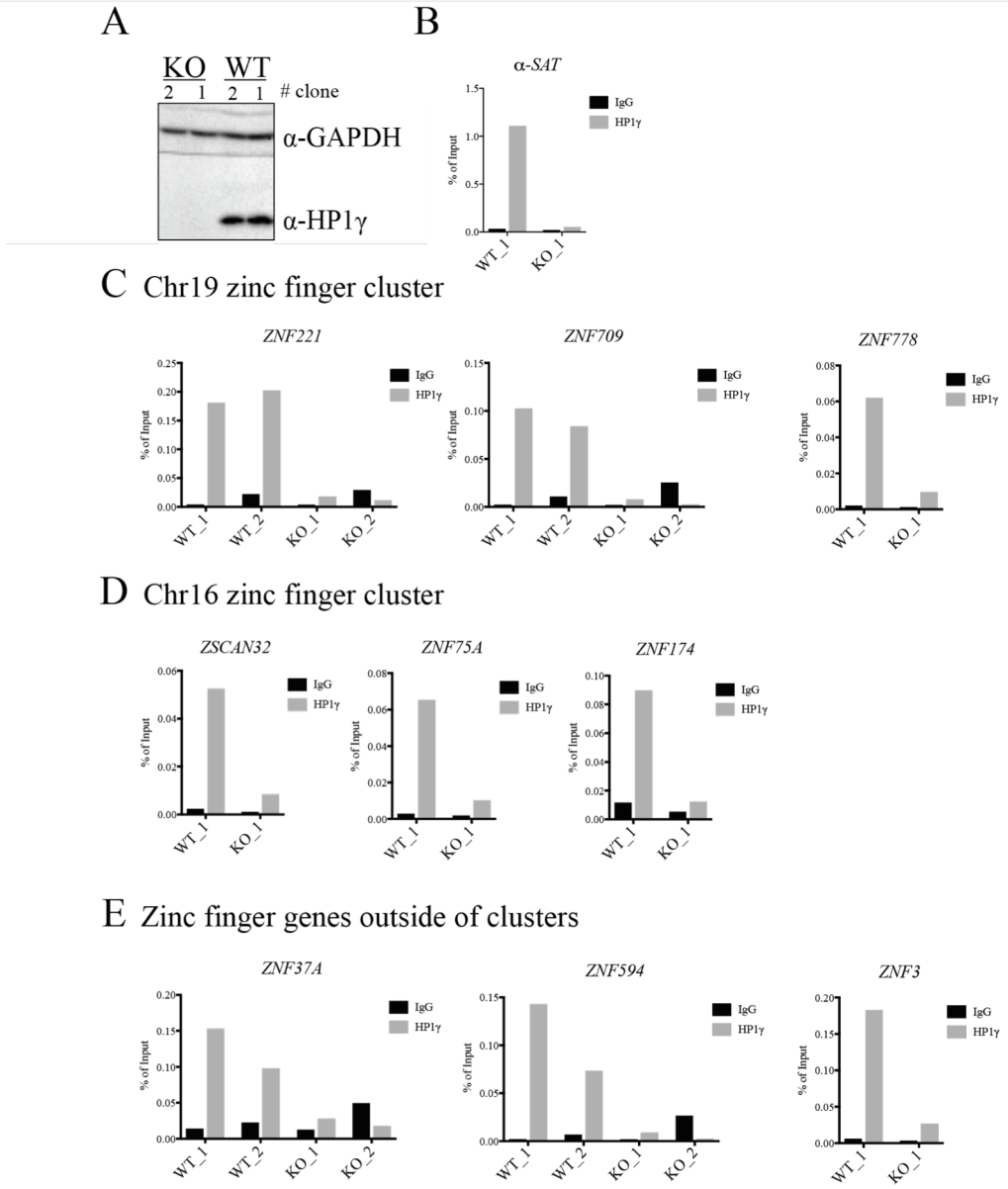


Figure 6-5 HP1γ ChIP-seq validation at 3'-exon of ZNF genes in MCF7 cells.

HP1γ CRISPR knockout MCF7 cells were generated using the lentiviral CRISPR_V2 vector system with 2 independent guide RNAs targeting exon 2 (knockout 1 (KO_1)) or exon 4 (knockout 2 (KO_2)) of *CBX3*. Cells were used for ChIP experiments using HP1γ and IgG antibodies and results were analysed by RT-PCR. RT-PCRs were performed in duplicates; the average of Ct values was used for further calculations. ChIP results are plotted as percentage of bound HP1γ over input.

A Whole cell lysates of two wildtype (WT) and two knockout clones analysed by immunoblotting using the indicated antibodies.

B ChIP RT-PCR of alpha-satellite repeat region as a positive control for HP1γ binding.

C - D ChIP RT-PCR of representative genes located in ZNF gene cluster of Chr19 (**C**), Chr16 (**D**) and outside of ZNF gene clusters (**E**).

The overall ChIP efficiency was relatively low with 1% being the highest enrichment detected for HP1γ at α-satellite repeats. However, in wildtype samples I detected 5-10 fold enrichment of HP1γ compared to IgG for all genes tested. Importantly, HP1γ signal was notably reduced in knockout cells. For most genes, HP1γ signal was at IgG level in knockout cells indicating that the enrichment of HP1γ measured in wildtype cells was specific.

In conclusion, ChIP-seq analysis of HP1γ in MCF7 cells identified HP1γ peaks in approximately 30% of all ZNF genes annotated in the human genome. Unbiased metagene analysis of HP1γ's distribution over all ZNF genes (n=937) showed that the signal for HP1γ peaks at the TSS and is further distributed over the whole gene body with a significant increase towards the 3'-exon. Interestingly, HP1γ seems to cover entire gene clusters of ZNF genes. To validate these findings I performed ChIP experiments analysed by RT-PCR in MCF7 wildtype cells compared to HP1γ knockout cells. RT-PCR primers were designed to span part of the 3'-exon of different ZNF genes. My results confirmed the ChIP-seq analysis for ZNF genes located within and outside of gene clusters. The majority of ZNF genes are not bound by HP1γ, suggesting a specific role for HP1γ at only a subset of ZNF genes. In order to further investigate the function of HP1γ binding to ZNF genes further, I cross-compared my RNA-seq data with the HP1γ ChIP-seq data. Results are summarised in the next section.

6.4. HP1γ regulates ZNF gene expression

To understand whether HP1γ regulates gene expression of ZNF genes, I cross-compared the list of all differentially regulated genes (n=4673, adjusted p-value<0.05) with all genes with a HP1γ peak 2kb up- and/or down-stream of the TSS (n=3859) as identified by ChIP-seq analysis. The results are visualized as a Venn diagram in Figure 6-6. It shows that only 14% of differentially expressed genes (n=665) are also bound by HP1γ, suggesting that the majority of genes might not be directly regulated by HP1γ at the expression level. Of those genes, 52% were up-regulated and 48% down-regulated upon HP1γ knockdown. Surprisingly, these genes were not enriched for those linked to IFNβ signaling and apoptosis pathways (as discussed in Chapter 5), indicating that loss of HP1γ most likely indirectly triggered activation of these signaling cascades. The only genes identified by this analysis were *STAT1* and *STAT2*.

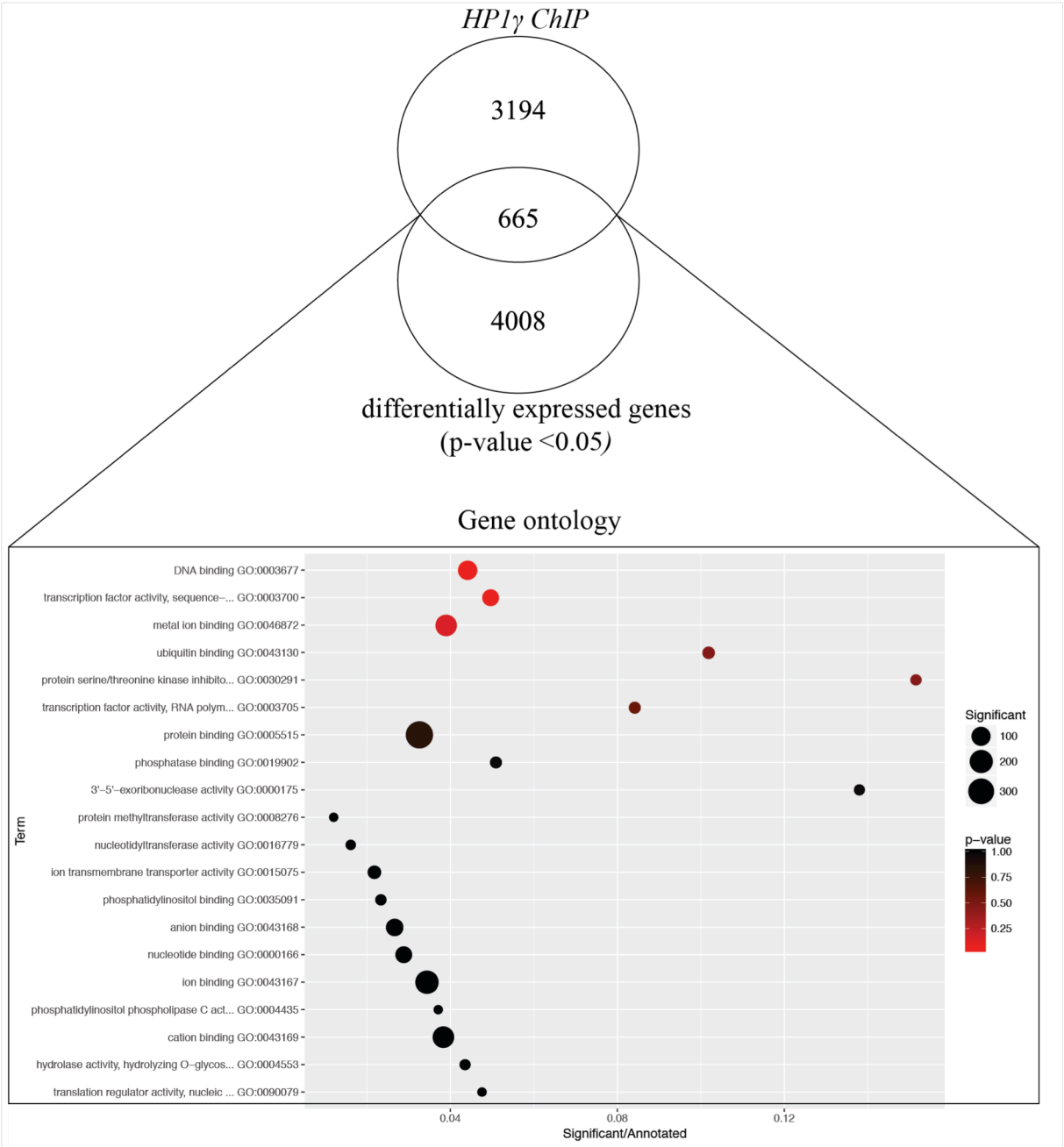


Figure 6-6 Cross-comparison of HP1γ bound and differentially expressed genes upon HP1γ knockdown in MCF7 cells.

List of differentially expressed genes upon HP1γ knockdown (p-value <0.05) (n=4673) was compared to genes identified by HP1γ ChIP-seq (n=3859). Results were visualised as a Venn diagram (top panel). GO enrichment analysis of genes contained within the intersection was performed (bottom panel) with the topGO Bioconductor package. Each dot represents a different GO term (y-axis), and the value (x-axis) indicates the fraction of genes in the given term, both differentially expressed and bound by HP1γ. Size of dots indicates the absolute number of gene targets that are annotated within the given class. Colour coding represents adjusted p-value: significant (red), less significant (black).

The JAK-STAT pathway is activated downstream of the interferon response, therefore it seems unlikely that up-regulation of STAT transcription factors is sufficient to activate the immune response observed in HP1γ depleted MCF7 cells.

Gene ontology analysis identified significant enrichment of genes associated with DNA binding, transcription factor binding and metal ion binding (Figure 6-6, bottom panel). A list of genes identified in the top GO enrichment class 'DNA binding' plotted against their log2 fold change is illustrated (Figure 6-7). 40 out of 108 genes (37%) within this list are ZNF genes (highlighted by green box). This analysis highlights that HP1γ may be directly involved in regulating gene expression of a subset of ZNF genes. The majority of ZNF genes within this list were up-regulated with only 25% identified being down-regulated.

As an example, Figure 6-8A, shows a UCSC genome browser snapshot of RNA-seq and ChIP-seq tracks for the gene *ZNF778*. It clearly demonstrates mRNA up-regulation of *ZNF778* upon HP1γ depletion. HP1γ binding at the 3'-exon of this gene was validated using HP1γ knockout cells Figure 6-5B. I validated gene expression changes by RT-PCR using an additional shRNA (shHP1γ_2) for a subset of genes. Figure 6-8B summarises representative results for 5 up-regulated genes. I furthermore validated HP1γ binding to the TSS of *ZNFX1* and *ZBTB43* genes by ChIP RT-PCR using HP1γ knockout cells (Figure 6-8D). HP1γ knockdown using shHP1γ_1 significantly (p-value <0.001) increased mRNA levels of all genes tested. A similar result was obtained for shHP1γ_2. However, the effect was less robust, which could be due to incomplete knockdown of HP1γ by this shRNA (compare to Figure 4-2). I also measured mRNA levels of the two most down-regulated ZNF genes (Figure 6-8C). However, due to the generally low expression levels of these genes, I was unable to validate the down-regulation of their mRNA with either of the shRNAs.

In summary, I showed that 14% of genes deregulated upon HP1γ depletion are actually bound by HP1γ. Intriguingly, gene ontology analysis of this gene subset revealed a significant enrichment for ZNF genes. Hence, HP1γ might directly regulate ZNF gene expression. To further confirm this hypothesis, I validated the expression changes of several ZNF genes by RT-PCR using two independent shRNAs. HP1γ binding to a subset of genes was validated using HP1γ knockout cells.

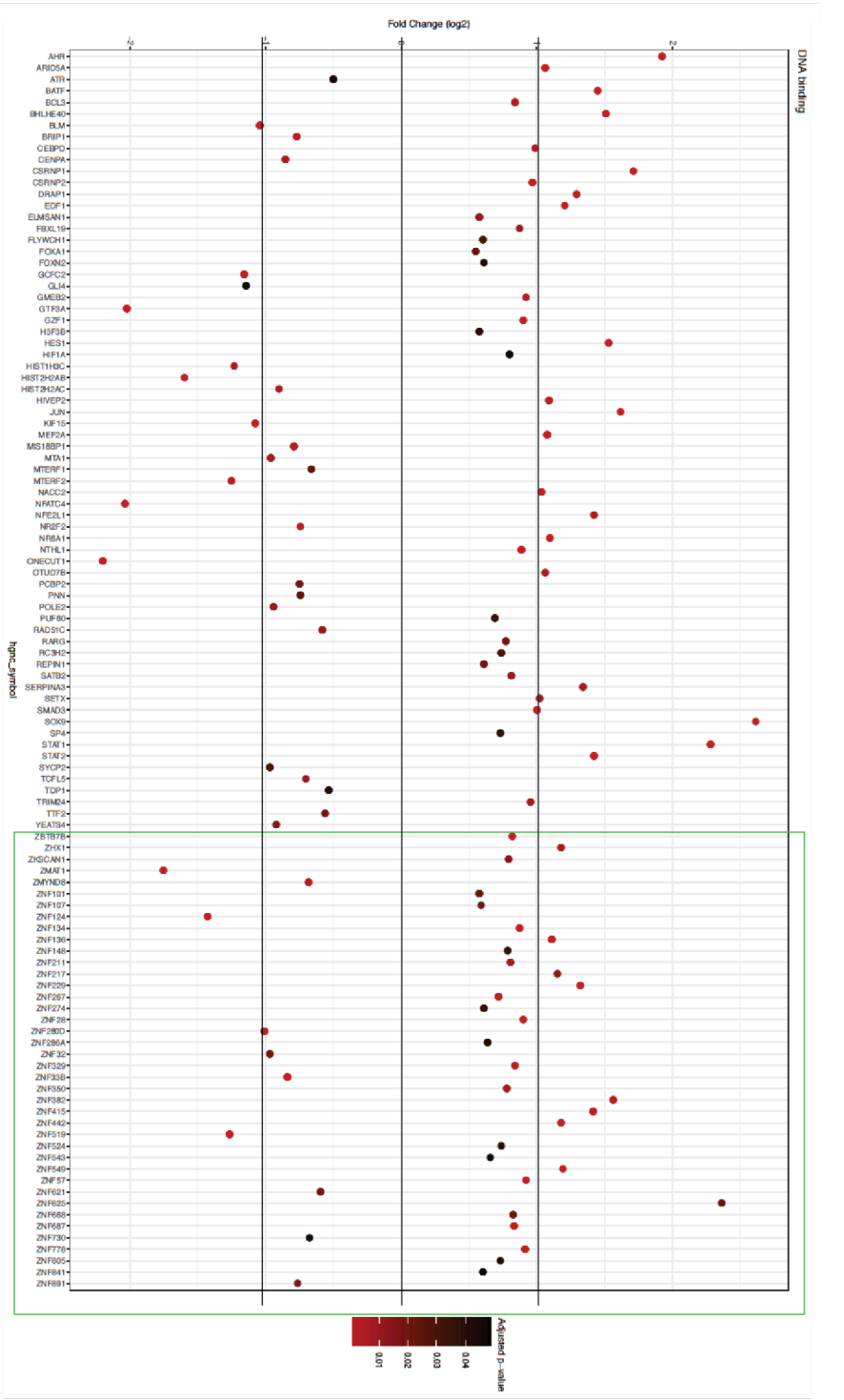


Figure 6-7 Genes identified in 'DNA binding' GO class. Genes plotted against their log2 fold change in HP1γ knockdown vs. control cells. The colour scale reports the adjusted p-value of differential expression statistics. Colour coding: red (adjusted p-value <0.01), black (adjusted p-value <0.05). ZNF genes are highlighted by green box.

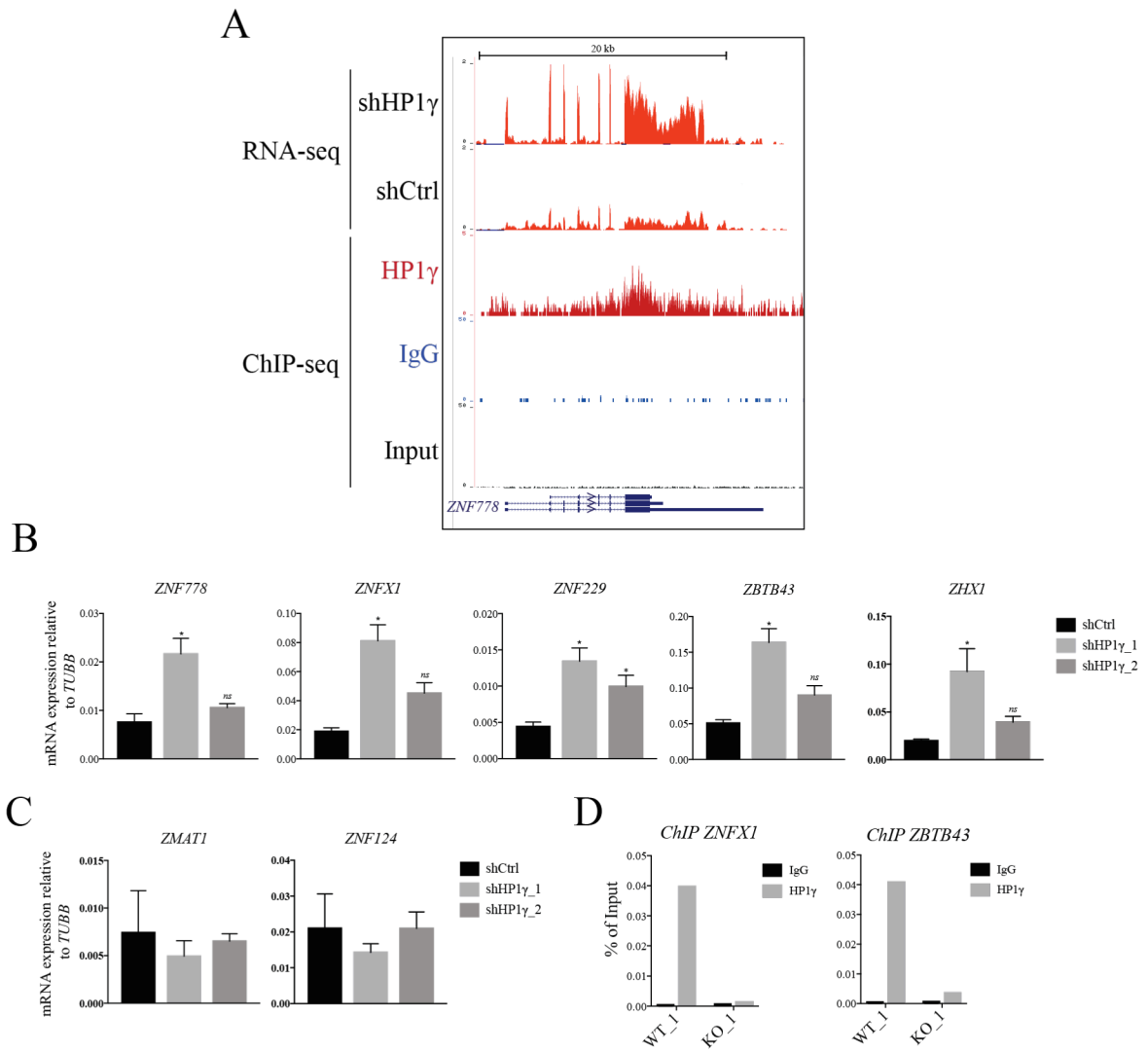


Figure 6-8 ZNF genes are deregulated upon HP1γ knockdown.

A UCSC genome browser tracks of normalised RNA-seq and HP1γ ChIP-seq data for *ZNF778* gene.

B/**C** mRNA level of indicated genes was assessed by RT-PCR after 6 days of knockdown using two independent shRNAs. Relative mRNA expression normalised to *TUBB* is shown. Graphs include data from 3 independent experiments. Significance was assessed using one-way ANOVA tests (*p-value < 0.001). SEM are indicated. **B** Up-regulated genes. **C** Down-regulated genes.

D HP1γ CRISPR knockout and wildtype MCF7 cells were used for ChIP experiments using HP1γ and IgG antibodies and data were analysed by RT-PCR. RT-PCRs were performed in duplicate; the average of Ct values was used for calculations. ChIP results are plotted as percentage of input.

6.5. Discussion

HP1 proteins are chromatin proteins functionally linked to various aspects of gene regulation. Whereas HP1α and HP1β are mainly involved in chromatin compaction and gene repression, HP1γ's role in gene and chromatin regulation is less well defined. In this chapter I studied the role of HP1γ in transcriptional regulation in breast cancer cells. I cross-referenced whole transcriptome analysis in HP1γ depleted MCF7 cells with the genomic localisation of HP1γ.

6.5.1. HP1γ ChIP in MCF7 cells

Cellular fractionation in MCF7 (Figure 4-2) and MDA-MB-231 cells (Figure 4-12) showed that a significant amount of HP1γ was not directly associated with chromatin. Furthermore, single molecule localisation microscopy of HP1γ in mESCs indicated that HP1γ molecules have higher diffusion coefficients and are less tightly bound to chromatin compared to HP1β molecules (Figure 3-19). HP1γ ChIP experiments were performed on isolated chromatin after cross-linking to avoid IP of freely diffusing HP1γ molecules. ChIP-seq validation by RT-PCR showed that HP1γ enrichment was in general relatively low, indicating that ChIP efficiency was not optimal. However, HP1γ signal in ChIP RT-PCRs was significantly reduced in HP1γ CRISPR knockout cells, demonstrating specific binding of HP1γ to all analysed loci. ChIP efficiency could be further optimised, for example by using different cross-linking reagents. All ChIP experiments described in this chapter used formaldehyde, which predominantly cross-links protein to DNA. In order to stabilise HP1γ molecules that are not in direct contact to chromatin, interacting with chromatin via other proteins, an additional cross-linking step using protein-protein cross-linking reagents such as disuccinimidyl glutarate (DSG) or dimethyl adipimidate (DMA) could be used[242]. These cross-linkers have a much higher effective cross-linking length than formaldehyde. If a significant proportion of HP1γ molecules are located further away from chromatin, changing the cross-linking conditions could increase the efficiency of HP1γ ChIP.

Nevertheless, ChIP-seq analysis of HP1γ in MCF7 cells identified 7596 significant HP1γ specific peaks. Further analysis showed that 3859 genes have a HP1γ peak 2kb up- and/or downstream of their TSS, indicating HP1γ is significantly enriched at promoters.

6.5.2. Regulation of gene expression by HP1γ in MCF7 cells

6.5.2.1. IFNβ response genes are not bound by HP1γ

In Section 6.3.5, I cross-compared HP1γ promoter bound genes in MCF7 cells with differentially expressed genes upon HP1γ depletion. Intriguingly, only a small subset (14%) of total differentially expressed genes was bound by HP1γ, indicating that the majority of gene expression changes may not be directly mediated by HP1γ. In Chapter 5, I showed that IFNβ signaling and pro-apoptotic pathways were activated in HP1γ depleted cells (Figure 5-6 and Figure 5-). The majority of genes involved in these pathways were not directly bound by HP1γ, suggesting signaling cascades were triggered downstream of HP1γ depletion. This analysis identified *STAT1* and *STAT2* as direct targets of HP1γ. As the JAK-STAT pathway is activated downstream of the interferon response, it seems unlikely to be the trigger that activates the immune response observed in MCF7 cells. As discussed in Section 5.4.1, *STAT1/3* can activate their own expression during acute viral infection[236]. It is therefore possible that up-regulation of STAT genes due to loss of HP1γ, reinforces the immune reaction and sensitizes MCF7 cells to activation of IFNβ signaling. It remains, however, unclear what triggered the immune reaction in the first place.

Cross-comparison of ChIP-seq and RNA-seq data did not include genes bound by HP1γ exclusively within the gene body. As HP1γ was reported to directly regulate genes involved in the immune response through either mediating RNA Pol II[ref.132,141] or repressing gene promoters[161], I manually inspected HP1γ ChIP-seq reads in the UCSC browser of genes involved in IFNβ immune regulation. However, I detected no significant enrichment of HP1γ within gene bodies. A comprehensive comparison of differentially expressed genes with genes harbouring a HP1γ peak exclusively in the gene body is currently ongoing.

ChIP-seq analysis of HP1γ, therefore, revealed no significant evidence that HP1γ directly regulates gene expression of genes that could trigger an immune response in MCF7 cells.

6.5.3. HP1γ marks actively transcribed genes

The first study to analyse HP1γ genome wide was published by the Ren laboratory in 2012[ref.143]. The authors performed ChIP-seq for HP1γ in the colorectal cancer cell line HCT-

116 and cross-compared it with RNA-seq data upon HP1γ knockdown. The paper describes a positive correlation between HP1γ binding and down-regulation of genes. They further confirmed that HP1γ interacts with elongating RNA Pol II[ref.73]. These results led them to define HP1γ as a transcriptional activator regulating transcriptional elongation. My ChIP-seq analysis in MCF7 cells indicated that HP1γ predominantly marks actively transcribed genes (Figure 6-2), which is consistent with these previous studies[73,141,143,145]. Notwithstanding the above, my results in MCF7 cells also show that 52% of genes bound by HP1γ are significantly up-regulated upon its depletion, indicating that HP1γ also plays a prominent role in transcriptional repression in MCF7 cells.

6.5.3.1. Understanding HP1γ's function in euchromatin

How HP1γ regulates expression of genes located predominantly in euchromatin remains poorly understood. Several studies describe recruitment of HP1γ to specific genes resulting in local 'heterochromatin' formation and gene repression. HP1γ can be part of repressor complexes containing histone methyltransferases that are recruited to actively transcribed genes in euchromatic regions. Recruitment then results in H3K9me3 deposition and ultimately gene repression[137,138]. However, H3K9me3 is not exclusively found in heterochromatin, as H3K9me3 within bodies of active genes is bound by HP1γ[73].

Of course, HP1γ may affect gene expression via post-transcriptional mechanisms. Indeed, the protein interacts with nascent transcripts[145], so it may play a role in stabilising or destabilising transcripts in addition to regulating alternative splicing. Although HP1γ was described as a general RNA binding protein[243], it may selectively and perhaps only temporarily bind to its target gene mRNAs. RNA binding by HP1 proteins is mediated by their hinge region[116]. An interesting experiment would be to test whether mutations within, or complete deletions of the hinge region affect binding of HP1γ to its target genes and whether expression of these mutants in place of wildtype HP1γ affect transcription in the same way as HP1γ depletion. Alternatively, HP1γ ChIP before and after treatment with RNase could be used to test the dependence of RNA on HP1γ's association to chromatin.

Several studies suggest HP1γ is localised to chromatin via its interaction with RNA Pol II, which can result in different transcriptional outputs. On the one hand, HP1γ was proposed to

enhance RNA Pol II transcription rate, suggesting HP1γ functions as a transcriptional activator[143]. On the other hand, a role for HP1γ in retarding transcription through slowing down the elongating RNA Pol II was described[141,145,147]. How HP1 regulates RNA Pol II activity mechanistically remains an open question in the field. It would be important to follow RNA Pol II progression at a given gene, in order to understand if its rate of transcription is immediately affected when released from a given promoter. This could for example affect loading of the next polymerase. Alternatively, the rate may be affected later during the elongation process. RNA Pol II progression can be measured by quantifying the amount of nascent transcripts with technologies such as global-run-on (GRO) sequencing[244] or native elongating transcript (NET) sequencing[245]. With these techniques it would be interesting to follow gene activation of inducible genes in the presence and absence of HP1γ.

HP1γ has been described quite extensively to relocate to inducible genes upon stimulation by cytokines[143], bacterial infection[141,151] or phorbol ester treatment[141,145]. I thought it would be interesting to study gene induction and RNA Pol II progression upon rapid inducible HP1γ degradation. To achieve this, I generated homozygous knock-in HCT-116 cells introducing an auxin-inducible degron (AID) tag at the *CBX3* locus using the CRISPR Cas9 system (as described in[246].) HCT116 cells are deficient in NHEJ repair mechanisms and therefore they are the preferred cell line for CRISPR knock-in experiments[246]. In theory, treatment of cells stably expressing auxin perceptive F-box protein TIR1 (OsTIR1), which forms a functional SCF (Skp1–Cullin–F-box) ubiquitin ligase, should rapidly (within minutes) induce degradation of proteins fused to AID following addition of auxin to the culture medium. Unfortunately, I detected reduced HP1γ protein levels in several clones only after 24h of auxin treatment, which could be due to the high abundance of HP1γ and/or low protein turn-over (data not shown). 24h is too long to detect immediate effects on transcription initiation and/or elongation, as I wished to do. Therefore, I decided not to pursue this investigation of HP1γ's role in regulating RNA Pol II function any further.

As both enhancing and attenuating RNA Pol II progression were proposed it would be interesting to investigate whether other proteins are part of the elongation machinery together with HP1γ. A technique to identify protein complexes on chromatin was published recently[247]. Using this technique with different antibodies against phosphorylated RNA Pol II to distinguish between initiation (RNA Pol II Ser2 phosphorylation) and elongation (RNA Pol

II Ser5 phosphorylation) complexes would be very useful to determine how complexes change upon HP1γ depletion. In turn, this could identify novel co-mediators of HP1γ.

6.5.4. HP1γ - a novel regulator of ZNF genes?

ChIP-seq analysis of HP1γ in MCF7 cells indicated that HP1γ target genes are highly enriched in ZNF genes. Approximately 30% (n = 290) of all ZNF genes annotated within the human genome are bound by HP1γ. Metagene coverage analysis of HP1γ signal over all ZNF genes (n = 937) showed that HP1γ signal is enriched at the TSS as well as the 3'-exon (Figure 6-4). HP1γ binding to the 3'-exon of ZNF genes was validated for a subset of genes (Figure 6-5), as well as binding to the TSS for another subset of genes (Figure 6-8). Cross-comparison of HP1γ bound genes with differentially expressed genes upon HP1γ knockdown showed that 17% (n = 48) of all ZNF genes bound by HP1γ are significantly deregulated, with approximately 33% (n =14) down- and 67% (n = 34) up-regulated (Table 6-2). The majority of ZNF genes bound by HP1γ (n = 242), does not change expression upon loss of HP1γ. Furthermore, this analysis identifies a group of genes, which is differentially expressed, but not bound by HP1γ, indicating loss of HP1γ perturbs ZNF gene expression in multiple ways.

Table 6-2 Summary of ZNF gene regulation by HP1γ. ZNF genes bound by HP1γ (n=290) are cross-compared with differentially expressed ZNF genes upon HP1γ knockdown (n=139). NC: not changing expression, UP/DOWN: up- and down-regulated (p-value <0.05).

	NC	UP	DOWN
Bound	242	34	14
Not bound	x	46	45

HP1γ might not be the only HP1 paralogue present at ZNF gene clusters. A similar observation was made for HP1β in the same cell line using the DamID method[248]. Furthermore, HP1γ interactor KAP1 binds to ZNF genes in a very similar fashion to HP1γ[241]. Remarkably however, loss of neither HP1β[248] nor KAP1[241] affects expression of its ZNF target genes, indicating an unique role for HP1γ in regulating ZNF gene expression.

6.5.4.1. HP1β binding at ZNF genes

Similarly to the binding profile of HP1γ at ZNF genes, HP1β binding was found strongly enriched at the 3'-exon[248] with additional peaks at the TSS. Vogel *et al*. demonstrate that

knockdown of HP1β and HP1α (HP1γ knockdown was not performed) had little influence on gene expression of ZNF genes in MCF7 cells. Overall they observed that HP1β target genes, ZNF and non-ZNF genes showed slightly lower expression levels than non-target genes[248]. My results indicate that in contrast to HP1β, HP1γ might regulate ZNF gene expression. But what makes HP1γ exclusive?

ZNF genes have a particularly interesting chromatin environment; genes are marked both by H3K9me3 and H3K36me3, indicating they are located within a state of 'heterochromatin' that is permissive for transcription[241,249]. A comparison of HP1γ with H3K36me3 and H3K4me3 ChIP-seq data for all ZNF genes (n = 937), shows that the majority of ZNF genes, including HP1γ target and non-target genes, are transcriptionally active (metagene plots are summarised in Appendix III Figure 8). RNA-seq data from the same cell line confirms these results. As discussed above, HP1γ is the only HP1 paralogue that directly mediates RNA-Pol II elongation rate and might be involved in regulating loading of the next polymerase during transcriptional initiation. HP1γ's direct association with transcriptional regulation could be an explanation why only loss of HP1γ, and not of HP1β, affects transcription of ZNF genes.

ZNF gene clusters are marked by H3K9me3[249,250] and this provides a logical explanation for HP1 binding at these loci. Knockdown of histone methyltransferases known to deposit H3K9me3 at ZNF gene loci, such as SETDB1[75,251] or Suv39H1[248], to reduce H3K9me3 would be useful in helping to understand how HP1γ binds to these loci. Furthermore, it is well documented that HP1 molecules homo- and heterodimerise. To test whether localisation of HP1γ depends on HP1β or vice versa, knockdown of either isoform followed by ChIP analysis for the other could be performed. Interestingly, the promoter region of ZNF genes seems to be depleted of H3K9me3, whereas 3'-exons are highly enriched[241], suggesting that recruitment of HP1 paralogues to ZNF gene promoters is independent of their chromodomain.

6.5.4.2. KAP1 binding at ZNF genes

It is likely that HP1γ is part of large protein complexes regulating gene expression. KAP1 is a large scaffold protein known to form large repressor complexes with histone methyltransferase SETDB1 or components of nucleosome remodeling and histone deacetylation (NuRD) complexes[75,252]. As KAP1 itself lacks a DNA binding domain it

associates with a variety of transcription factors including a large subset of KRAB ZNF proteins. This interaction is mediated through its N-terminal RBCC domain[111]. The predominant role of this repressor complex is the suppression of transposable elements[112].

Also, KAP1 predominantly binds 3'-exons of ZNF genes with additional peaks at the TSS[241]. ChIP experiments by Iyengar *et al.* demonstrated that KAP1 binding to the 3'-exon of ZNF genes depends on its interaction with KRAB ZNF proteins, as mutating the RBCC domain completely abolished KAP1's binding to 3'-exons[241]. Interestingly, mutating the RBCC domain did not affect KAP1's binding to promoter regions of ZNF genes, indicating a different mechanism of recruitment. Furthermore, the authors mutated the PxVxL motif of KAP1 to abolish its interaction with HP1 and demonstrated that the protein is still able to localise to promoter regions, but less efficiently[241]. These results suggest that KAP1 recruitment is largely independent of HP1, but HP1 may affect its retention time.

Whether binding of HP1γ to ZNF genes is dependent on KAP1 remains unclear. To determine how HP1γ is recruited to its ZNF target genes, ChIP experiments using different mutants of HP1γ that either abolish binding to H3K9me3 (V32M point mutation in chromodomain), abolish the interaction with co-factors (PxVxL mutant I166K in chromo-shadow domain) or its interaction with RNA (mutation or deletion of hinge region) will be critical to understand which domains of HP1γ are important for targeting it to ZNF gene loci.

6.5.4.3. Is HP1γ suppressing homologous recombination of ZNF genes?

Expression of the vast majority of HP1γ bound ZNF genes (83%) remained unaffected by HP1γ knockdown. ZNF genes have arisen from genomic duplications, hence they are highly homologous[253]. Their 3′-exons encode tandemly arranged highly repetitive ZNF domains and it was proposed that heterochromatinisation of these 3'-exons suppresses recombination-mediated gene deletions[241,248]. Intriguingly, both KAP1 and HP1γ, but not HP1α and β, were described to suppress homologous recombination in synthetic reporter assays[122,123], suggesting both proteins could be involved in suppressing recombination of ZNF gene 3'-exons.

In conclusion, I showed that in MCF7 cells, 14% of genes deregulated upon HP1γ depletion are actually bound by HP1γ. HP1γ target genes are both down- and up-regulated indicating that HP1γ can act as a transcriptional activator and repressor. Differentially expressed genes associated with the phenotype of HP1γ depletion in MCF7 cells were not directly bound by HP1γ, indicating their gene expression was affected downstream. Intriguingly, differentially expressed genes bound by HP1γ are significantly enriched for ZNF genes. HP1γ binding to ZNF genes seems to coincide with binding of HP1β and KAP1. Work by other groups, however, shows that loss of neither HP1β nor KAP1 affects expression of ZNF genes. My results suggest a novel role for HP1γ in regulating ZNF gene expression. To understand the means by which HP1γ is involved regulating these genes, it will be critical to determine how HP1γ is recruited to these genomic loci.

Chapter 7

Final conclusions

The three mammalian HP1 proteins: HP1α, HP1β and HP1γ are multifunctional proteins and their association with chromatin is involved in a myriad of physiological processes. HP1 paralogues are ubiquitously expressed and due to their high sequence conservation they exhibit functional redundancy. Nevertheless, it has become increasingly evident that each HP1 paralogue possesses specific functions with respect to chromatin formation and transcriptional regulation, which are likely to be cell type specific.

My PhD work focused on HP1γ, which in contrast to HP1α and HP1β is predominantly associated with euchromatic regions and plays a role in transcriptional elongation. I identified novel functions of HP1γ in mESCs and breast cancer cells. I characterised a new post-translational modification of HP1γ called peptidyl-citrullination using *in vitro* and *in vivo* assays. This work suggests a function for citrullination of HP1γ in stem cell homeostasis. Furthermore, I studied the phenotype of HP1γ depletion in two human breast cancer models and found that HP1γ is important for cell viability of cancer, but not of normal epithelial cells. Whole transcriptome analysis in breast cancer cells depleted of HP1γ, cross-referenced with its genomic localisation, identified HP1γ as a novel regulator of zinc finger (ZNF) genes.

7.1. A role for HP1γ citrullination in pluripotency

Citrullination describes the conversion of peptidyl-arginine to peptidyl-citrulline resulting in loss of a positive charge. In mammalian nuclei, citrullination is catalysed by PADI1, 2 and 4. Mass spectrometry analysis in mESCs identified citrullination of HP1γ in the CD at residues R38/9. My PhD work shows that HP1γ is citrullinated by PADI4 *in vitro* and *in vivo* at residues R38/9. The mutant R38/9A, designed to mimic the loss of charge accompanied with

citrullination, affects HP1γ's binding to H3K9me3 peptides *in vitro* and reduces the residence time of HP1γ molecules on chromatin in living differentiated cells, as determined by SPT.

Pluripotent mESCs require a distinct nuclear architecture, characterised by epigenetic signatures enriched with active chromatin marks and hyperdynamic binding of structural chromatin proteins. My data show that in these cells the diffusion and residence times of HP1γ WT and R38/9A mutant molecules are highly similar, indicating that the time molecules spend 'ON' and 'OFF' chromatin is similar (Figure 7-1). The process of stem cell differentiation is accompanied by major chromatin remodeling, which predominantly requires formation of facultative heterochromatin to establish cell type specific gene expression. HP1γ and other structural chromatin factors such as histone H1 play important roles in this process[97,168].

Upon differentiation the percentage of molecules associated with chromatin increases for both HP1γ WT and the R38/9A mutant. However, the time that mutant molecules remain stably bound to chromatin ('ON') is reduced compared to WT proteins, indicating that mutant molecules are spending less time 'ON' and more time 'OFF' chromatin (Figure 7-1). My results would suggest that residues R38 and R39 are not necessary for HP1γ's recruitment to chromatin during differentiation, but influence the stable binding of HP1γ molecules once on chromatin. These data are currently being analysed to determine association rates of single molecules in different regions of the nucleus.

Our laboratory described a role for PADI4 in pluripotency and showed that activity of PADI4 enzymes in differentiated cells is significantly decreased compared to pluripotent stem cells[190]. A subsequent decrease in HP1γ citrullination would fit a model in which stable interaction of HP1γ with chromatin is necessary for heterochromatin formation during stem cell differentiation. Conversely in pluripotent cells, where a relatively open chromatin conformation must be maintained, the regulation of HP1γ by citrullination may be important to prevent stable interaction of the protein with chromatin, either to regulate expression of specific genes or to prevent chromatin condensation more globally. In the context of recent publications describing a role for HP1γ in regulating cell fate decisions and stem cell differentiation[170-172], the findings of my PhD are compatible with a novel potential function for HP1γ citrullination in the maintenance of stem cell homoeostasis.

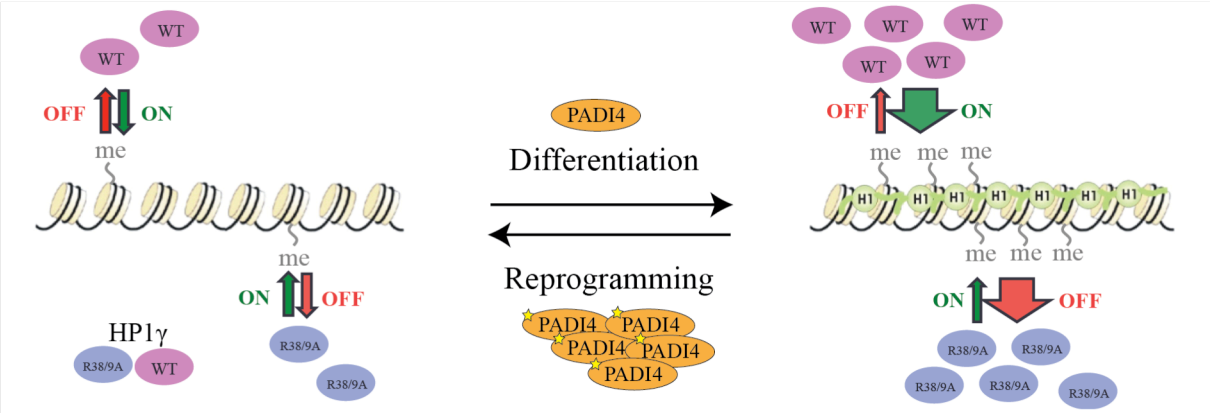


Figure 7-1 Model I: Role for HP1γ residues R38/39 during stem cell differentiation. Chromatin of pluripotent stem cells has a open and relaxed conformation. In these cells dynamics of HP1γ WT and R38/9A mutant molecules are similar. Differentiation is accompanied by increased chromatin condensation mediated by structural proteins such as HP1 and H1. Although the overall percentage of HP1γ molecules on chromatin is similar, residence time (ON) R38/9A mutants is reduced. PADI4 activity is decreased during differentiation, while expression and activity (star) are enhanced during reprogramming[170].

7.2. HP1γ - a novel regulator of zinc finger genes

The biggest group of ZNF genes encodes KRAB zinc finger transcription factor proteins, which are characterised by an N-terminal KRAB domain and a C-terminal region containing tandemly arranged arrays of C2H2 zinc finger modules[112]. Together with their co-factor KAP1 they participate in repression of transposable elements (TEs). About 50% of the whole human genome comprises TEs and the biggest group of TEs are retroelements[112]. Although less than 1 out of 10000 human TEs are still capable of transposition[254], they can cause disruption of genes or affect gene expression, as they contain transcription factor binding sites such as promoters, enhancers or insulators. Moreover, TEs consist of highly repetitive sequences making them potential hot spots of homologous recombination events that could cause various genomic aberrations[112]. Reactivating TEs or altering their chromatin environment can have fatal consequences for a cell. Many genetic disorders and different types of cancer are associated with novel TE insertions or other types of TE deregulation[254,255]. The correct regulation of TEs is therefore very important.

My PhD work identifies a novel function for HP1γ in regulating ZNF genes in MCF7 breast cancer cells. HP1γ is highly enriched at ZNF gene clusters. A similar observation was made for KAP1 and HP1β[241,248]. Intriguingly, loss of HP1γ but not HP1β or KAP1, affects

expression of ZNF target genes, indicating a unique role for HP1γ in regulating ZNF gene expression. However, the majority of HP1γ bound ZNF genes does not change expression. Furthermore, my analysis identifies another group of ZNF genes that is differentially expressed, but not bound by HP1γ (compare Table 6-2). These results suggest that loss of HP1γ in MCF7 cells perturbs zinc finger gene expression in various ways. There are several possible scenarios of how HP1γ could be involved in regulating ZNF genes (Figure 7-2).

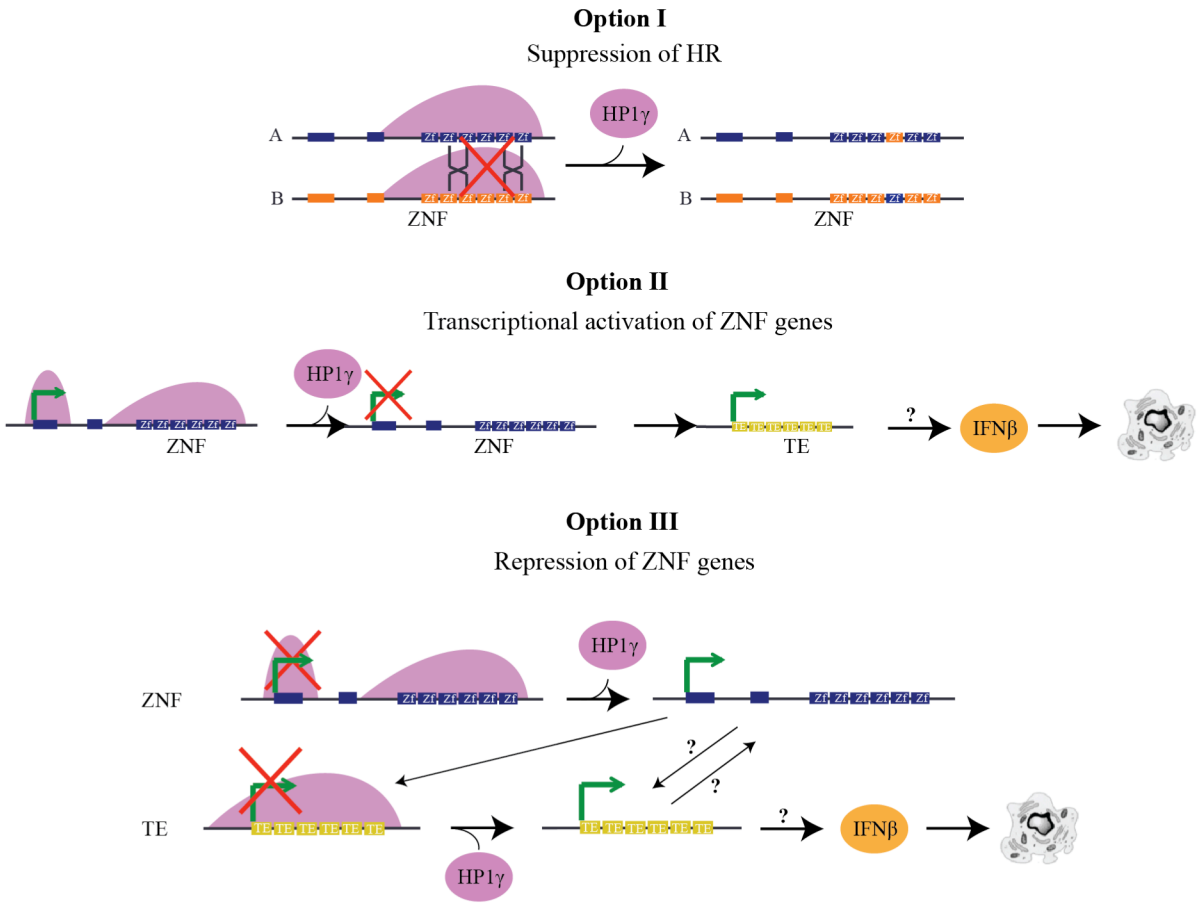


Figure 7-2 Model II: HP1γ regulates ZNF genes in MCF7 cells.
Potential roles for HP1γ in regulating zinc finger (ZNF) genes. **Option I:** HP1γ suppresses homologous recombination (HR) and loss of HP1γ results in aberrant recombination of C2H2 zinc finger modules within 3'-exons of ZNF genes. **Option II:** HP1γ acts as a transcriptional activator of ZNF genes and HP1γ depletion leads to down-regulation of expression and consequent de-repression of transposable elements (TEs). **Option III:** HP1γ acts as a transcriptional repressor of ZNF genes and/or TEs. Loss of HP1γ results in enhanced expression of ZNF genes and TEs. Possible feedback loops between ZNF gene and TEs expression are indicated by black arrows. Reactivation of TEs could be a trigger for activating the INFβ/antiviral response and an increase in apoptosis, observed in HP1γ depleted MCF7 cells.

Upon loss of HP1γ, the majority of HP1γ bound ZNF genes (>80%) are not differentially expressed. At those genomic loci, the main function of HP1γ may be to suppress homologous recombination (HR) (Option I. Figure 7-2). HP1γ is a potent suppressor of HR, while HP1α and HP1β are not[122,123]. C2H2 zinc finger modules of zinc finger proteins are encoded within the 3'-exon of ZNF genes. Since ZNF genes derived from gene duplications, these zinc finger modules are highly homologous and therefore hot spots of HR[112]. Zinc finger modules dictate the transcription factor's DNA binding specificity. In order to prevent *de novo* targeting of potent transcription factors and triggering gene expression changes that could be detrimental to the cell, HR at 3'-exons of ZNF genes has to be strictly regulated. In order to test my hypothesis, I have cultured HP1γ knockout MCF7 cell clones over a long period of time and I will sequence their genomic DNA to study aberrant recombination events within the 3'-exons of HP1γ target ZNF genes.

HP1γ can function as a transcriptional activator[143]. Upon HP1γ depletion, a subset of ZNF genes is down-regulated, which suggests a role for HP1γ in positively regulating expression of these genes (Option II. Figure 7-2). The main role for KRAB zinc finger transcription factors is to repress TEs as part of the KAP1 co-repressor complex[112]. Hence, a direct consequence of down-regulation of KRAB zinc finger proteins may be de-repression of TEs. We are currently analysing the RNA-seq data in MCF7s cells to detect differential expression of transcripts originating from TEs. As TEs consist of highly repetitive sequences that are distributed throughout the human genome, standard mapping algorithms are not suitable to detect differential expression of these transcripts.

The last hypothesis refers to up-regulated ZNF genes, where HP1γ's role(s) may be more complex (Option III. Figure 7-2). Firstly, HP1γ could function as a repressor of ZNF gene expression by directly binding to ZNF gene promoters. Loss of HP1γ consequently results in up-regulation of the gene. Besides KRAB zinc finger proteins, HP1s are also essential components of the KAP1 co-repressor complex and are directly linked to repression of TEs[240]. Therefore, secondly, loss of HP1γ may directly result in de-repression of TEs. Transcripts originating from TEs could in turn trigger up-regulation of certain ZNF genes. In order to assure TE repression, one could imagine that mammalian cells have developed such 'safe-lock' feedback mechanisms. In support of this hypothesis is the group of differentially expressed ZNF genes in HP1γ depleted MCF7 cells, which is not directly bound by HP1γ.

Taken together, HP1γ may perturb the regulation of ZNF genes in several different ways and none of the aforementioned options are necessarily mutually exclusive. Given the central role of KRAB zinc finger transcription factors in suppressing TEs, deregulation of ZNF genes may cause or be an indirect result of deregulation of TEs.

The main phenotype of HP1γ depletion in MCF7 cells is a reduction of cell viability and increased apoptosis. Gene expression analysis upon loss of HP1γ identifies a clear activation of INFβ response, which in theory should only be activated in response to acute viral infection. Conversely, the majority of genes involved in these pathways are not bound by HP1γ, indicating pathway activation occurred further downstream. But what triggers the INFβ response in MCF7 cells?

Increased expression of the IFN response/antiviral defense genes in tumor cells can result from enhanced transcription of genomic repeats such as retroelements[256-259]. This fits a model in which loss of HP1γ in MCF7 cells either directly or indirectly (via ZNF genes) results in de-repression of TEs, resulting in activation of the INFβ response and activation of apoptosis (Figure 7-2). In strong support of this hypothesis is a recent study demonstrating that survival of drug-tolerant cancer cells requires increased levels of H3K9me3 over LINE-1 (long interspersed nuclear element 1) retroelements[258]. Loss of regulators of H3K9me3-mediated heterochromatin formation, including HP1γ, in drug-tolerant cancer cells results in re-expression of LINE-1 elements, activation of IFN/antiviral defense mechanisms and increased cell death[258].

In summary, my PhD thesis identifies a novel role for HP1γ in regulating ZNF genes, which may directly or indirectly influence the expression of TEs. Activation of the INFβ response in HP1γ depleted breast cancer cells establishes an exciting link between HP1γ mediated regulation of the ZNF - TE axis and the viability of breast cancer cells. Overall, my work highlights HP1γ as good candidate target for therapeutic intervention in breast cancer.

Appendix I

RNA-seq in MCF7 cells

Quality control

The dispersion plot depicted in Figure 1A shows the relationship between the variance in the data and the mean. The mean of the normalised read counts across all samples on the x-axis is plotted against the dispersion score on the y-axis. Dispersion scores are a measure of the square of the coefficient of biological variation. The higher the dispersion, the more varied the data are between replicates. Variance here describes a combination of biological noise and technical noise. The fitted line indicates the estimated dispersion value based on modelling the values across all genes. The black points are the values estimated from the RNA-seq data. The final dispersion scores, accounting for the sample-to-sample variation, are illustrated as blue circles. To test the similarity between replicates, hierarchical clustering algorithms were applied. Results are visualised as a heatmap (Figure 1B). The tree-building algorithm calculates the similarity between each sample and all other samples at every iteration step. Samples are then connected to the most similar sample at a height that is proportional to the similarity. The heatmap illustrates the Pearson correlation coefficient (R) for each pairwise comparison.

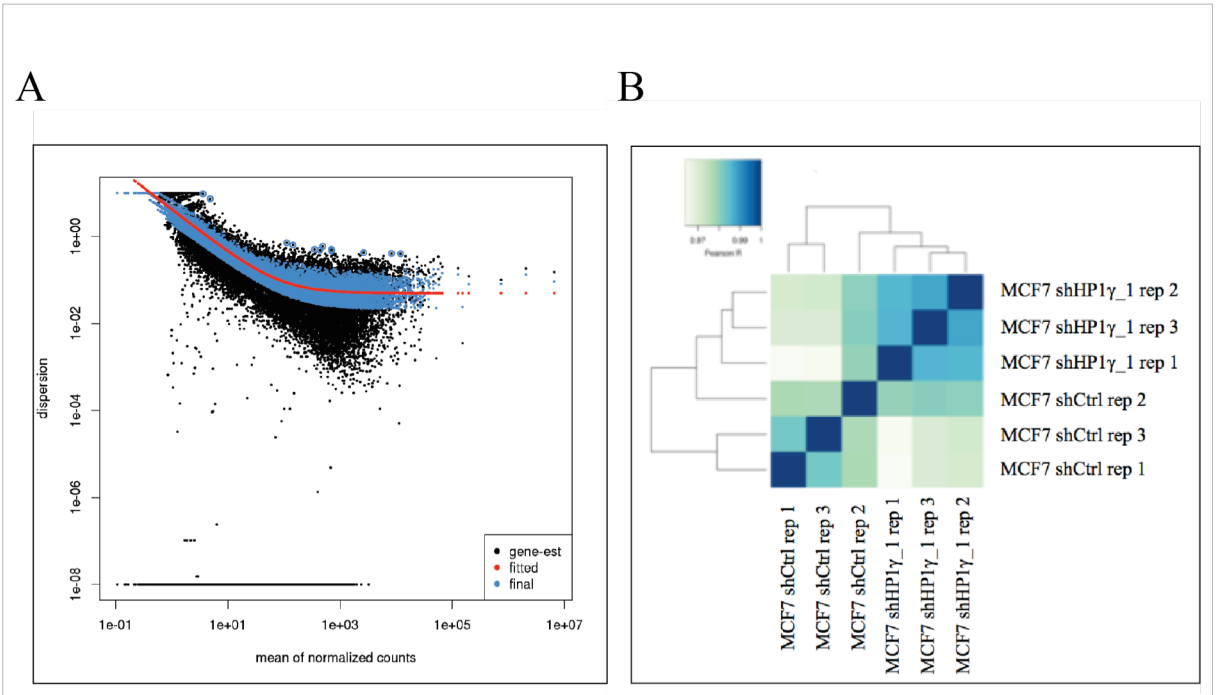


Figure 1 Quality control of MCF7 RNA-seq data.

A Dispersion plot: Mean of normalised read counts across samples on the x-axis plotted against the dispersion score on y-axis. Red: Fitted line indicating estimated dispersion value; Black dots: values estimated from the RNA-seq data; Blue circles: final dispersion scores.

B Hierarchical clustering of technical replicates of sample groups: shCtrl and shHP1γ_1. Heatmap illustrates Pearson correlation scores (R) for pairwise comparisons.

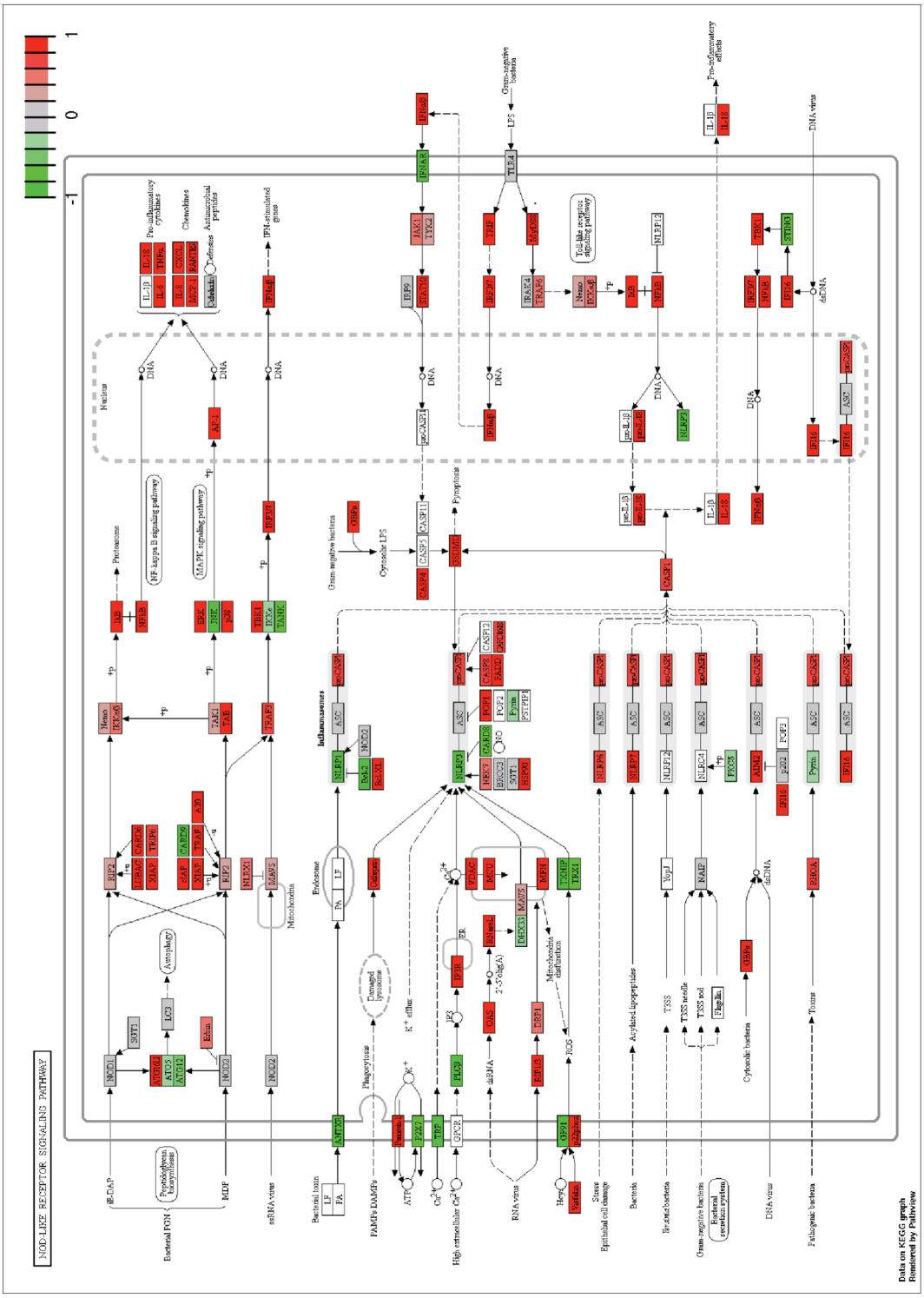


Figure 2 NOD like receptor signaling (hsa04621).

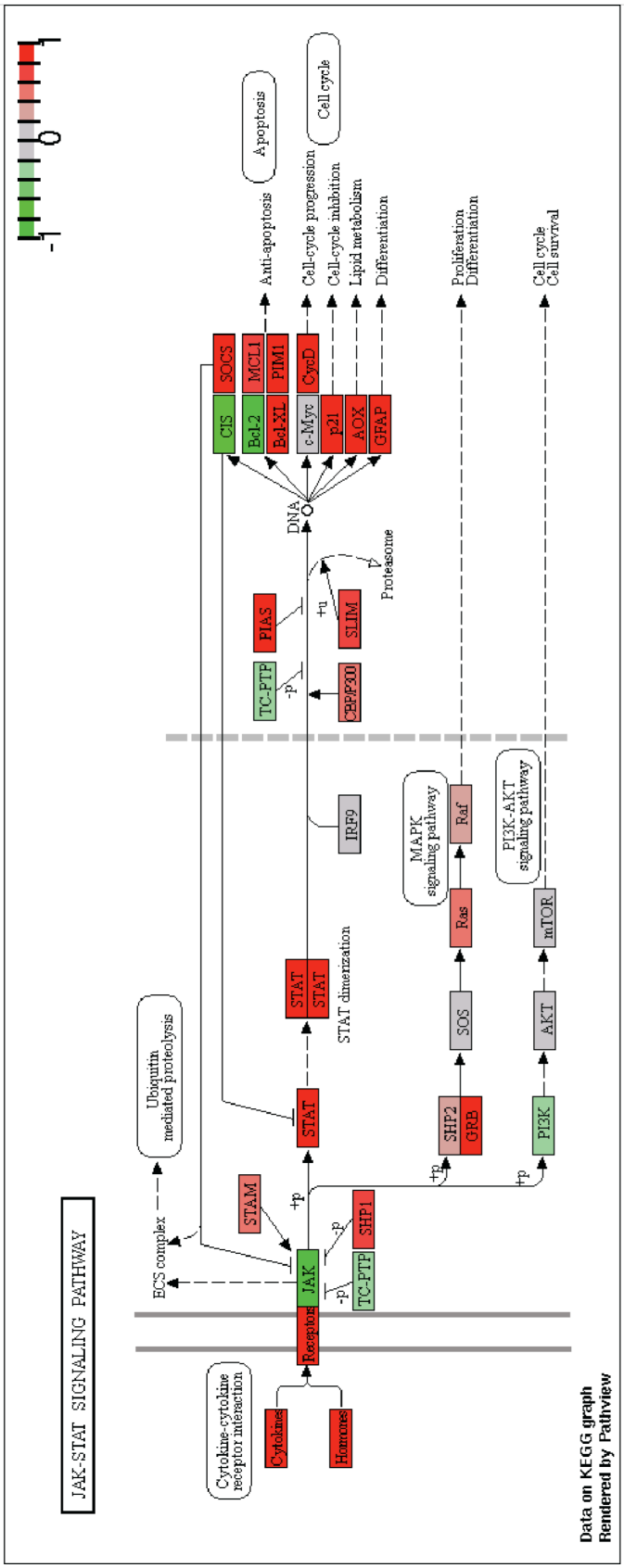


Figure 3 JAK STAT signaling (hsa04630).

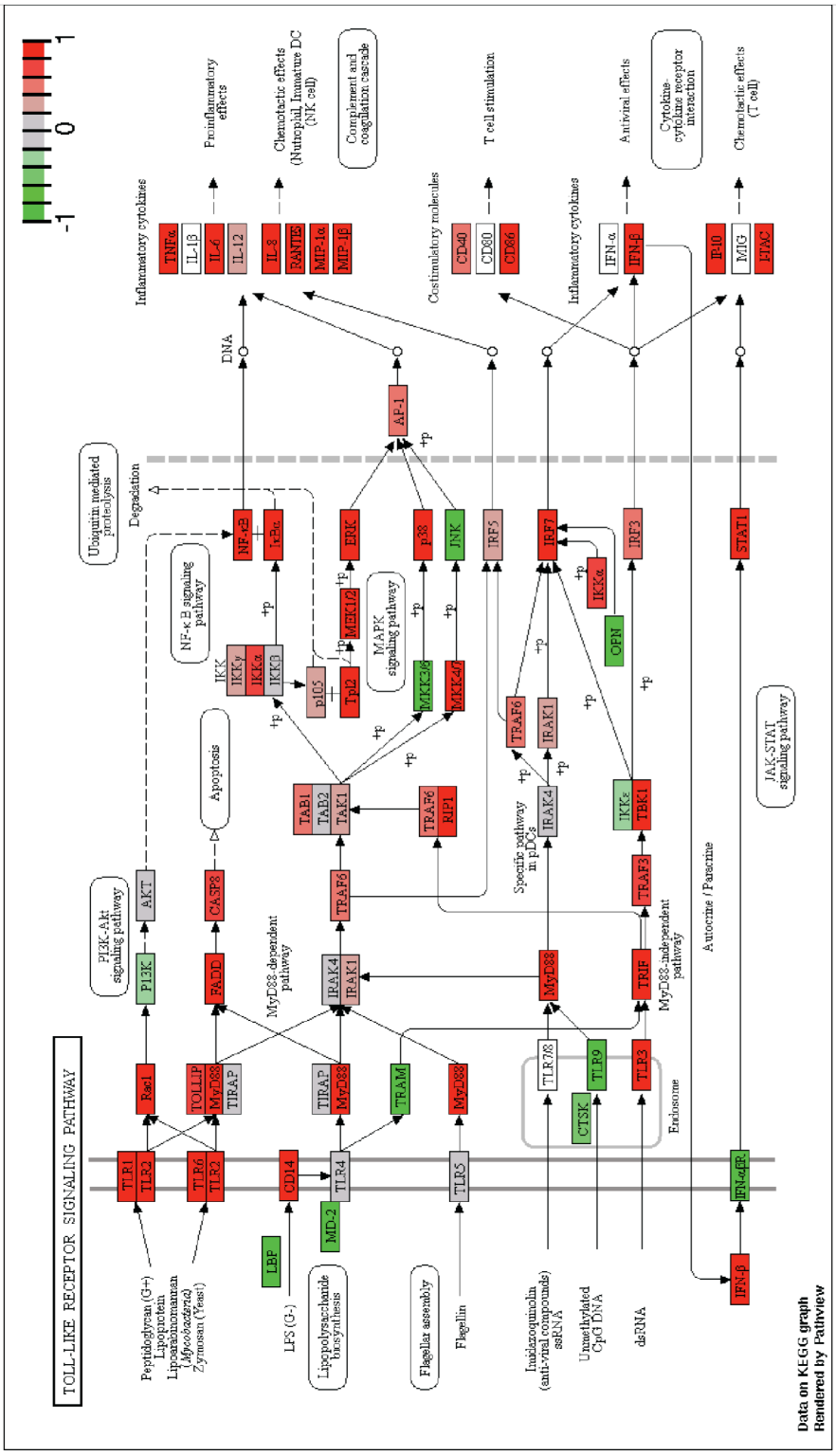


Figure 4 Toll like receptor signaling (hsa044620).

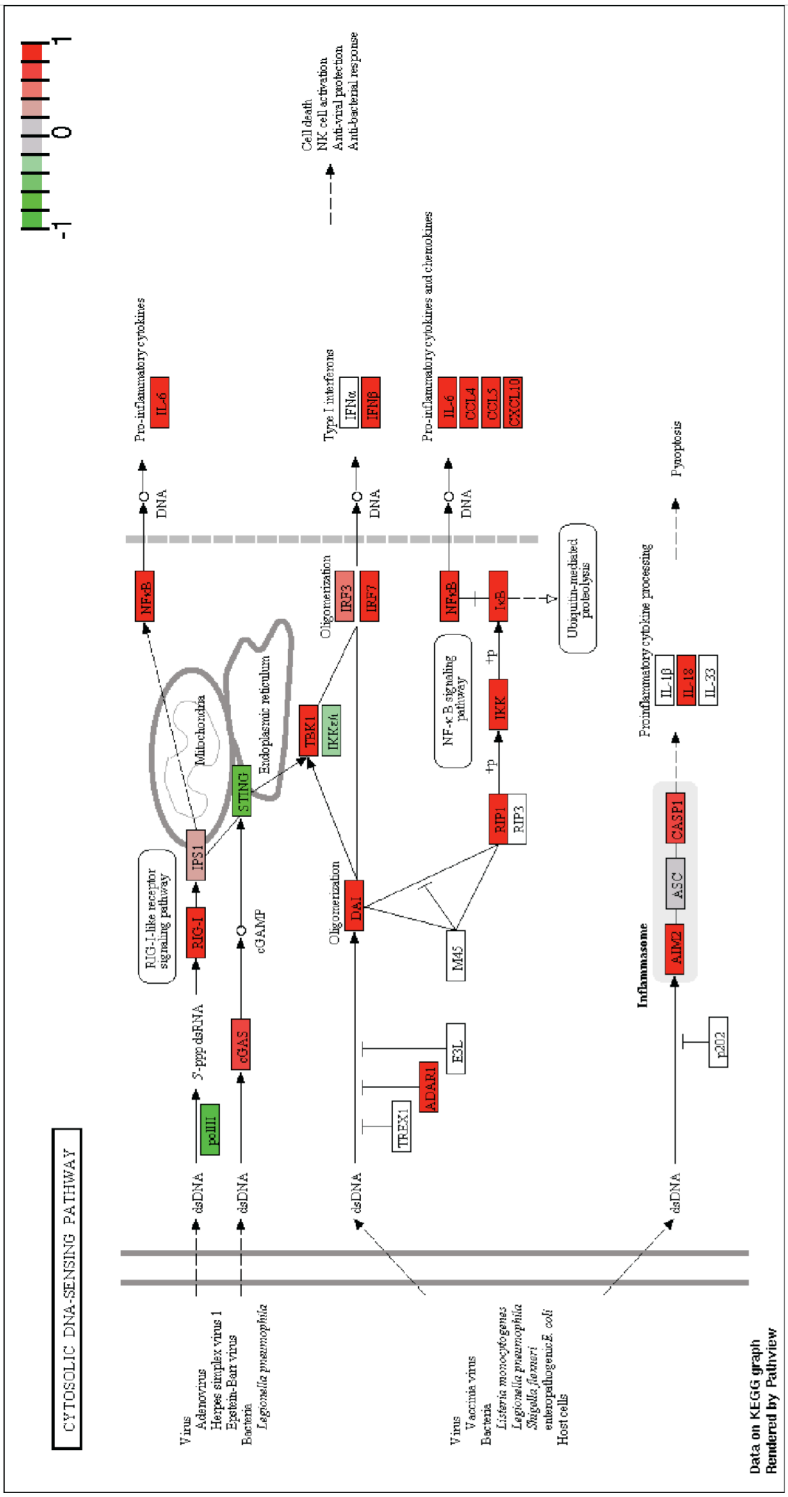
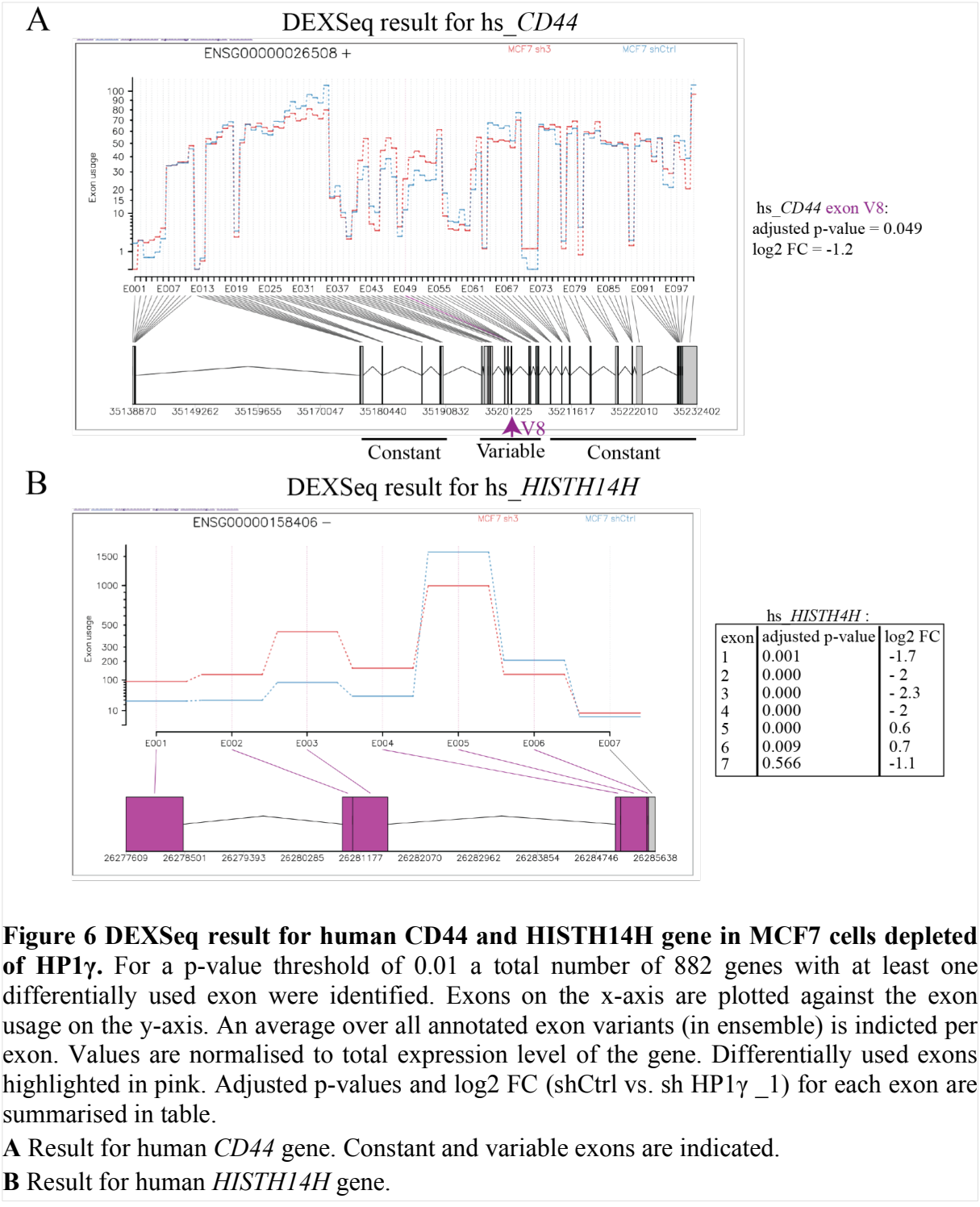


Figure 5 Cytosolic DNA sensing (hsa04623).



hs_*HISTH4H* :

exon	adjusted p-value	log2 FC
1	0.001	-1.7
2	0.000	- 2
3	0.000	- 2.3
4	0.000	- 2
5	0.000	0.6
6	0.009	0.7
7	0.566	-1.1

Figure 6 DEXSeq result for human CD44 and HISTH14H gene in MCF7 cells depleted of HP1γ. For a p-value threshold of 0.01 a total number of 882 genes with at least one differentially used exon were identified. Exons on the x-axis are plotted against the exon usage on the y-axis. An average over all annotated exon variants (in ensemble) is indicted per exon. Values are normalised to total expression level of the gene. Differentially used exons highlighted in pink. Adjusted p-values and log2 FC (shCtrl vs. sh HP1γ _1) for each exon are summarised in table.

A Result for human *CD44* gene. Constant and variable exons are indicated.

B Result for human *HISTH14H* gene.

Validation of RNA-seq by siRNA mediated knockdown

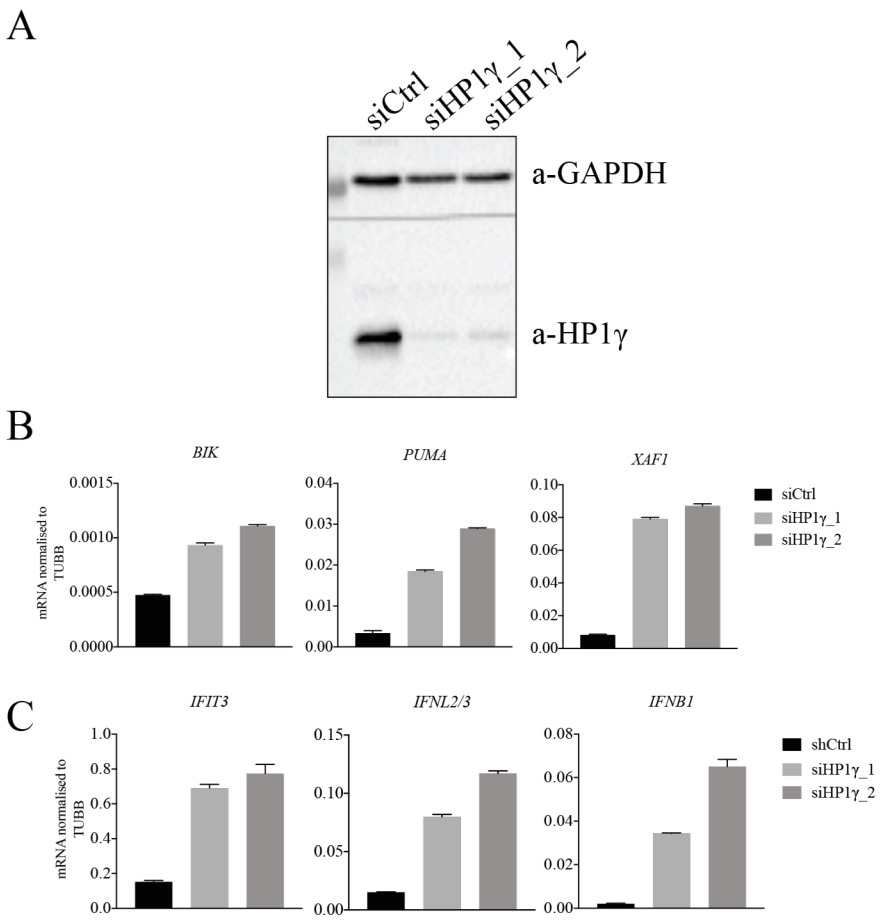


Figure 7 Validation of RNA-seq targets by siRNA knockdown of HP1 γ.

MCF7 cells were transfected with siRNAs (as indicated) and re-transfected after 48h.

A Immunoblotting of MCF7 lysates after 4 days of siRNA mediated knockdown using antibodies as indicated. Lysates from 40000 cells were loaded per well.

B mRNA levels of indicated genes were assessed 4 days of siRNA mediated knockdown by RT-PCR. Graphs show relative mRNA expression after normalisation to *TUBB* expression. SEM of technical replicates are indicated.

Appendix II

ChIP-seq MCF7

Quality control

Shown in are distribution plots of the peak coverage for each sample group (HP1γ, H3K4me3 and H3K36me3). Peaks for each replicate were plotted as histograms and pairwise correlation plots of both replicates are shown. Each point represents a single peak from the joint peak calling.

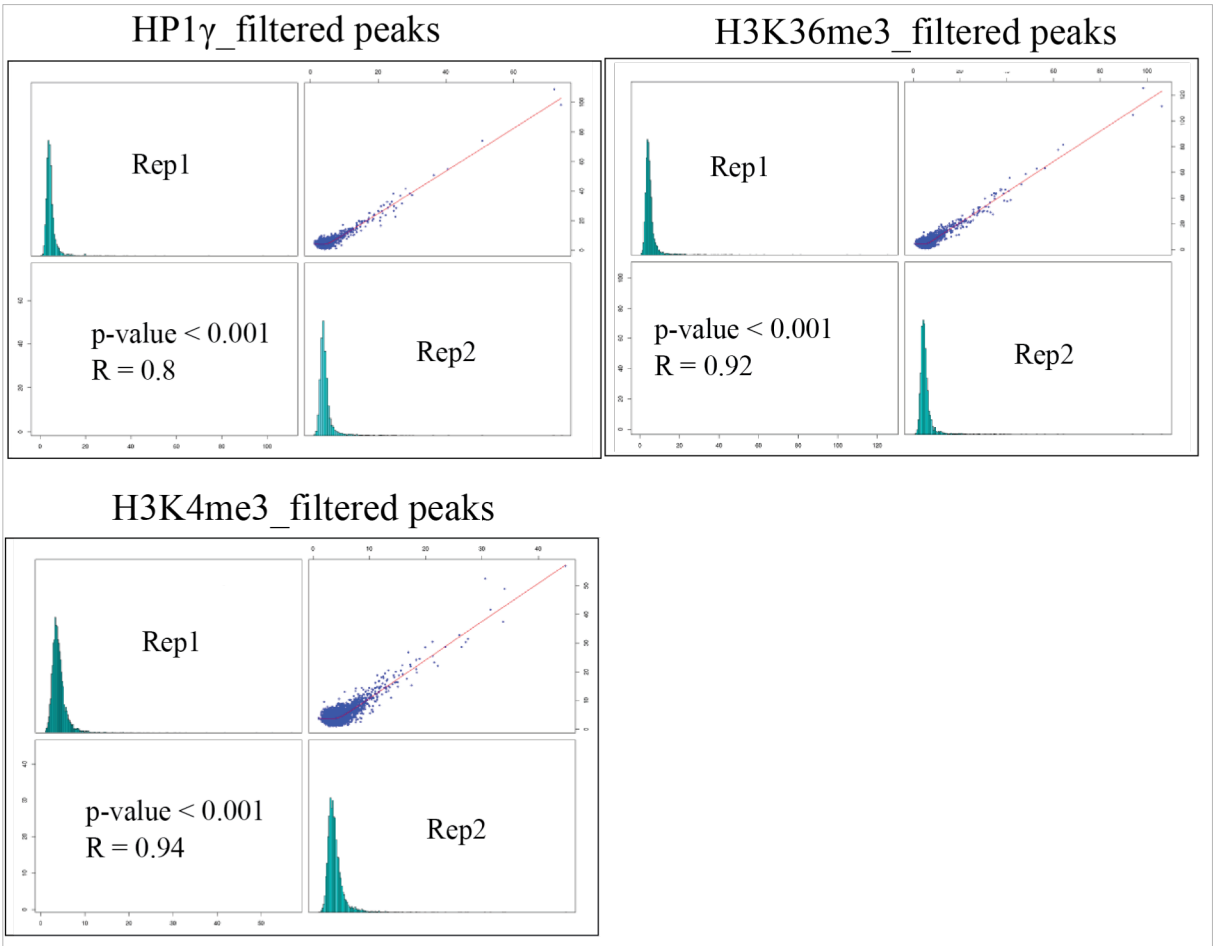
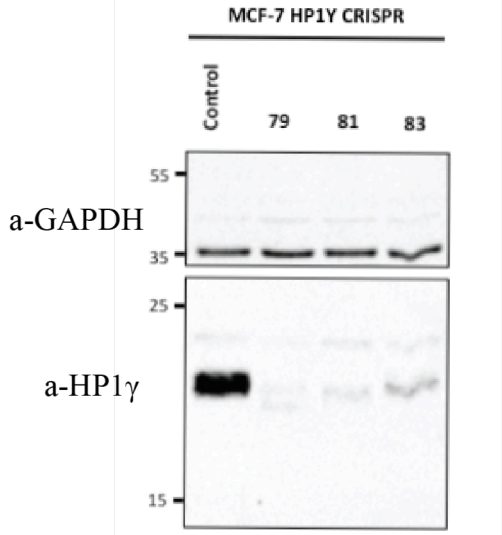


Figure 8 Pairwise correlation of ChIP-seq replicates.
Pairwise correlation of filtered peaks of technical replicates for sample groups HP1γ, H3K4me3 and H3K36me3. Distribution of peak coverage plotted as histograms. Pairwise correlation plots indicated with fitted line (red) indicating correlation. p-value and Pearson correlation coefficients (R) are indicated for each pair.

Generation of MCF7 HP1γ CRISPR knockout clones

A



B

Cell line	sgRNA	Clone ID	Knockout status
MCF-7	Control	2	-
		3	-
		4	-
		8	-
		9	-
		10	-
		11	-
		14	-
	#79	2	KO
		3	KO
		7	Truncated
		11	Truncated
		14	Truncated
		15	Truncated
	#81	16	KO/Truncated
		22	Truncated
	#83	5	KO
		9	KO

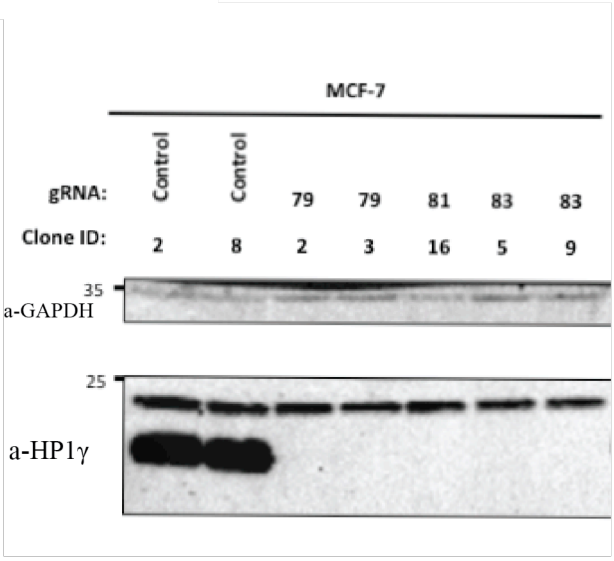


Figure 9 Generation of MCF7 HP1γ CRISPR knockout clones.

CRISPR knockout was achieved using the lentiviral CRISPR_V2 vector system with three independent guide RNAs targeting *CBX3*. MCF7 cells were infected with lentiviral vectors containing different sgRNA sequences (#79, #81, #83) (see Table 2-11). After selection with puromycin for 5 days HP1γ protein levels of bulk cell populations were assessed by immunoblotting using indicated antibodies **(A)**. From each bulk cell population single cell clones were generated. Number of cell clones generated and their HP1γ status are summarised in **(B).** HP1γ knockout of single cell clones was confirmed by immunoblotting using indicated antibodies **(C).** Control clone #2 (= WT_1) and #8 (= WT_2); Knockout clone #79_2 (KO_1) and 83_5 (= KO_2) were used for ChIP RT-PCR validation experiments (Section 6.3 and Section 6.4). Cell lines and figures were generated by Katharina Lindner.

HP1γ bound ZNF genes are actively transcribed

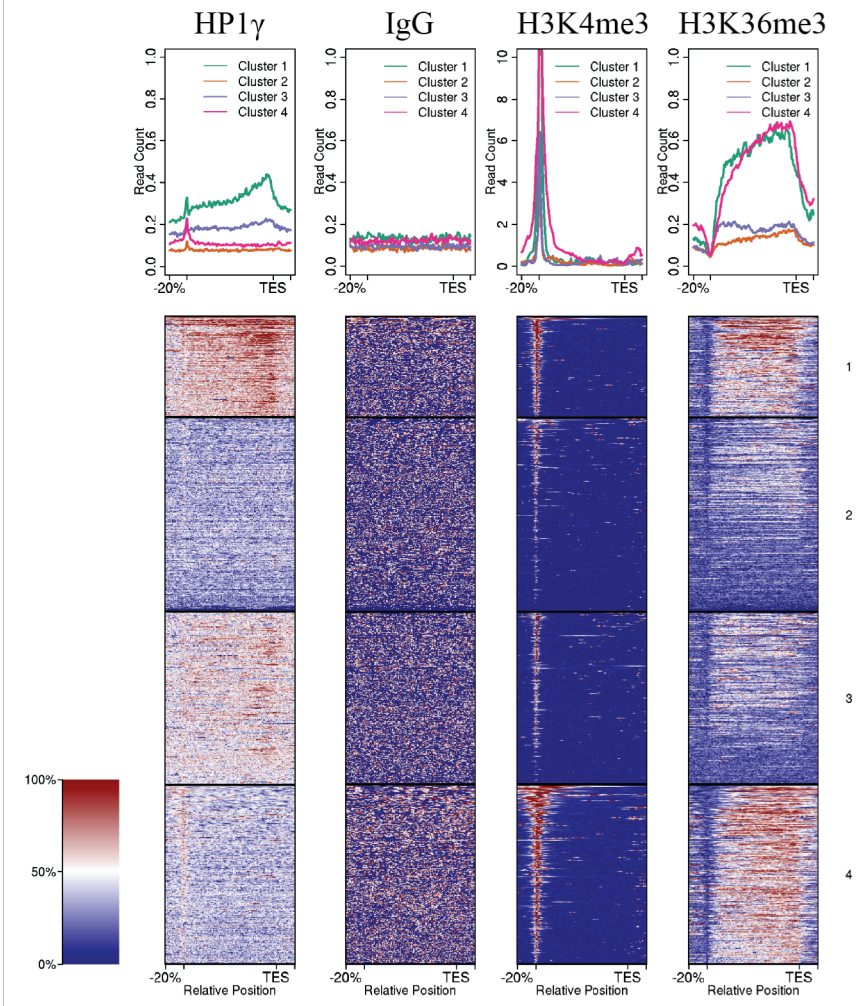


Figure 10 Metagene plots of HP1γ, H3K4me3 and H3K36me3 ChIP-seq data over ZNF genes.

Metagene coverage plots of HP1γ, IgG, H3K4me3 and H3K36me3 reads over all annotated ZNF genes (n=937). Individual samples are scaled proportionally to mapped library size. All transcripts overlapping a peak are scaled to a uniform size of 5kb (plus 2kb upstream of the TSS and downstream of the TES). Colour scheme: ranging from 100% enrichment (red) to 0% enrichment (blue). Genes were clustered into different groups (1-4). Signals for each sample were scaled individually (top panel).

Bibliography

1. Flemming, W. Beitraege zur Kenntniss der Zelle und ihrer Lebenserscheinungen. *Arch. Mikroskop. Anat.* **16**, 34 (1878).

2. Boveri, T. *Ergebnisse über die Konstitution der chromatischen Substanz des Zellkerns*, iv, 130 p (Gustav Fischer, Jena, 1904).

3. Heitz, E. Das Heterochromatin der Moose. *Jahrb. Wiss. Bot.* **69**, 56 (1928).

4. Avery, O.T., Macleod, C.M. & McCarty, M. Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types : Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type Iii. *J Exp Med* **79**, 137-58 (1944).

5. Watson, J.D. & Crick, F.H. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171**, 737-8 (1953).

6. Jacob, F. & Monod, J. Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol* **3**, 318-56 (1961).

7. Kornberg, R.D. Chromatin structure: a repeating unit of histones and DNA. *Science* **184**, 868-71 (1974).

8. Kornberg, R.D. & Thomas, J.O. Chromatin structure; oligomers of the histones. *Science* **184**, 865-8 (1974).

9. Olins, A.L. & Olins, D.E. Spheroid chromatin units (v bodies). *Science* **183**, 330-2 (1974).

10. Oudet, P., Gross-Bellard, M. & Chambon, P. Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. *Cell* **4**, 281-300 (1975).

11. Waddington, C.H. The Epigenotype. *International Journal of Epidemiology* **41**, 10-13 (2012).

12. Berger, S.L., Kouzarides, T., Shiekhattar, R. & Shilatifard, A. An operational definition of epigenetics. *Genes Dev* **23**, 781-3 (2009).

13. Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F. & Richmond, T.J. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**, 251-60 (1997).

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

15. Kornberg, R.D. & Lorch, Y. Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell* **98**, 285-94 (1999).

16. Woodcock, C.L. Chromatin architecture. *Curr Opin Struct Biol* **16**, 213-20 (2006).

17. Allis, C.D. *et al.* New nomenclature for chromatin-modifying enzymes. *Cell* **131**, 633-6 (2007).

18. Segal, E. & Widom, J. What controls nucleosome positions? *Trends Genet* **25**, 335-43 (2009).

19. Widom, J. Role of DNA sequence in nucleosome stability and dynamics. *Q Rev Biophys* **34**, 269-324 (2001).

20. Pennings, S., Allan, J. & Davey, C.S. DNA methylation, nucleosome formation and positioning. *Brief Funct Genomic Proteomic* **3**, 351-61 (2005).

21. Gurard-Levin, Z.A., Quivy, J.P. & Almouzni, G. Histone chaperones: assisting histone traffic and nucleosome dynamics. *Annu Rev Biochem* **83**, 487-517 (2014).

22. Clapier, C.R., Iwasa, J., Cairns, B.R. & Peterson, C.L. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat Rev Mol Cell Biol* **18**, 407-422 (2017).

23. Fuks, F. DNA methylation and histone modifications: teaming up to silence genes. *Curr Opin Genet Dev* **15**, 490-5 (2005).

24. Heard, E. Delving into the diversity of facultative heterochromatin: the epigenetics of the inactive X chromosome. *Curr Opin Genet Dev* **15**, 482-9 (2005).

25. Robertson, K.D. DNA methylation and human disease. *Nat Rev Genet* **6**, 597-610 (2005).

26. Sandoval, J. & Esteller, M. Cancer epigenomics: beyond genomics. *Curr Opin Genet Dev* **22**, 50-5 (2012).

27. Guibert, S. & Weber, M. Functions of DNA methylation and hydroxymethylation in mammalian development. *Curr Top Dev Biol* **104**, 47-83 (2013).

28. Jones, P.A. & Liang, G. Rethinking how DNA methylation patterns are maintained. *Nat Rev Genet* **10**, 805-11 (2009).

29. Okano, M., Bell, D.W., Haber, D.A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247-57 (1999).

30. Baylin, S.B. & Jones, P.A. A decade of exploring the cancer epigenome - biological and translational implications. *Nat Rev Cancer* **11**, 726-34 (2011).

31. Tahiliani, M. *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930-5 (2009).

32. Wu, X. & Zhang, Y. TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat Rev Genet* **18**, 517-534 (2017).

33. Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev* **16**, 6-21 (2002).

34. Fuks, F. *et al.* The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. *J Biol Chem* **278**, 4035-40 (2003).

35. Rea, S. *et al.* Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406**, 593-9 (2000).

36. Lehnertz, B. *et al.* Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol* **13**, 1192-200 (2003).

37. Bannister, A.J. *et al.* Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120-4 (2001).

38. Allfrey, V.G., Faulkner, R. & Mirsky, A.E. Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. *Proc Natl Acad Sci U S A* **51**, 786-94 (1964).

39. Bannister, A.J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res* **21**, 381-95 (2011).

40. Tessarz, P. & Kouzarides, T. Histone core modifications regulating nucleosome structure and dynamics. *Nat Rev Mol Cell Biol* **15**, 703-8 (2014).

41. Sterner, D.E. & Berger, S.L. Acetylation of histones and transcription-related factors. *Microbiol Mol Biol Rev* **64**, 435-59 (2000).

42. Bannister, A.J., Schneider, R. & Kouzarides, T. Histone methylation: dynamic or static? *Cell* **109**, 801-6 (2002).

43. Tessarz, P. *et al.* Glutamine methylation in histone H2A is an RNA-polymerase-I-dedicated modification. *Nature* **505**, 564-8 (2014).

44. Nowak, S.J. & Corces, V.G. Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. *Trends Genet* **20**, 214-20 (2004).

45. Shilatifard, A. Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annu Rev Biochem* **75**, 243-69 (2006).

46. Hassa, P.O., Haenni, S.S., Elser, M. & Hottiger, M.O. Nuclear ADP-ribosylation reactions in mammalian cells: where are we today and where are we going? *Microbiol Mol Biol Rev* **70**, 789-829 (2006).

47. Cubenas-Potts, C. & Matunis, M.J. SUMO: a multifaceted modifier of chromatin structure and function. *Dev Cell* **24**, 1-12 (2013).

48. Cuthbert, G.L. *et al.* Histone deimination antagonizes arginine methylation. *Cell* **118**, 545-53 (2004).

49. Wang, Y. *et al.* Human PAD4 regulates histone arginine methylation levels via demethylimination. *Science* **306**, 279-83 (2004).

50. Nelson, C.J., Santos-Rosa, H. & Kouzarides, T. Proline isomerization of histone H3 regulates lysine methylation and gene expression. *Cell* **126**, 905-16 (2006).

51. Hyllus, D. *et al.* PRMT6-mediated methylation of R2 in histone H3 antagonizes H3 K4 trimethylation. *Genes Dev* **21**, 3369-80 (2007).

52. Vaquero, A. *et al.* SIRT1 regulates the histone methyl-transferase SUV39H1 during heterochromatin formation. *Nature* **450**, 440-4 (2007).

53. Talbert, P.B. & Henikoff, S. Histone variants on the move: substrates for chromatin dynamics. *Nat Rev Mol Cell Biol* **18**, 115-126 (2017).

54. Biterge, B. & Schneider, R. Histone variants: key players of chromatin. *Cell Tissue Res* **356**, 457-66 (2014).

55. Elsaesser, S.J. & Allis, C.D. HIRA and Daxx constitute two independent histone H3.3-containing predeposition complexes. *Cold Spring Harb Symp Quant Biol* **75**, 27-34 (2010).

56. Tagami, H., Ray-Gallet, D., Almouzni, G. & Nakatani, Y. Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* **116**, 51-61 (2004).

57. Rogakou, E.P., Pilch, D.R., Orr, A.H., Ivanova, V.S. & Bonner, W.M. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J Biol Chem* **273**, 5858-68 (1998).

58. Sims, R.J., 3rd *et al.* Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. *J Biol Chem* **280**, 41789-92 (2005).

59. Huang, Y., Fang, J., Bedford, M.T., Zhang, Y. & Xu, R.M. Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A. *Science* **312**, 748-51 (2006).

60. Strahl, B.D. & Allis, C.D. The language of covalent histone modifications. *Nature* **403**, 41-5 (2000).

61. Trojer, P. *et al.* L3MBTL1, a histone-methylation-dependent chromatin lock. *Cell* **129**, 915-28 (2007).

62. Lan, F. & Shi, Y. Epigenetic regulation: methylation of histone and non-histone proteins. *Sci China C Life Sci* **52**, 311-22 (2009).

63. Barski, A. *et al.* High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823-37 (2007).

64. Mosammaparast, N. & Shi, Y. Reversal of histone methylation: biochemical and molecular mechanisms of histone demethylases. *Annu Rev Biochem* **79**, 155-79 (2010).

65. Whetstine, J.R. *et al.* Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. *Cell* **125**, 467-81 (2006).

66. Wagner, E.J. & Carpenter, P.B. Understanding the language of Lys36 methylation at histone H3. *Nat Rev Mol Cell Biol* **13**, 115-26 (2012).

67. Carrozza, M.J. *et al.* Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123**, 581-92 (2005).

68. Joshi, A.A. & Struhl, K. Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. *Mol Cell* **20**, 971-8 (2005).

69. Ruthenburg, A.J., Allis, C.D. & Wysocka, J. Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. *Mol Cell* **25**, 15-30 (2007).

70. Heintzman, N.D. *et al.* Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108-12 (2009).

71. Bernstein, B.E. *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315-26 (2006).

72. Voigt, P., Tee, W.W. & Reinberg, D. A double take on bivalent promoters. *Genes Dev* **27**, 1318-38 (2013).

73. Vakoc, C.R., Mandat, S.A., Olenchock, B.A. & Blobel, G.A. Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. *Mol Cell* **19**, 381-91 (2005).

74. Rice, J.C. *et al.* Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. *Mol Cell* **12**, 1591-8 (2003).

75. Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G. & Rauscher, F.J., 3rd. SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev* **16**, 919-32 (2002).

76. Yang, L. *et al.* Molecular cloning of ESET, a novel histone H3-specific methyltransferase that interacts with ERG transcription factor. *Oncogene* **21**, 148-52 (2002).

77. Kang, Y.K. SETDB1 in Early Embryos and Embryonic Stem Cells. *Curr Issues Mol Biol* **17**, 1-10 (2015).

78. Shinkai, Y. & Tachibana, M. H3K9 methyltransferase G9a and the related molecule GLP. *Genes Dev* **25**, 781-8 (2011).

79. Huang, J. *et al.* G9a and Glp methylate lysine 373 in the tumor suppressor p53. *J Biol Chem* **285**, 9636-41 (2010).

80. Tarcsa, E. *et al.* Protein unfolding by peptidylarginine deiminase. Substrate specificity and structural relationships of the natural substrates trichohyalin and filaggrin. *J Biol Chem* **271**, 30709-16 (1996).

81. Vossenaar, E.R., Zendman, A.J., van Venrooij, W.J. & Pruijn, G.J. PAD, a growing family of citrullinating enzymes: genes, features and involvement in disease. *Bioessays* **25**, 1106-18 (2003).

82. Slade, D.J. *et al.* Protein arginine deiminase 2 binds calcium in an ordered fashion: implications for inhibitor design. *ACS Chem Biol* **10**, 1043-53 (2015).

83. Arita, K. *et al.* Structural basis for Ca(2+)-induced activation of human PAD4. *Nat Struct Mol Biol* **11**, 777-83 (2004).

84. Wang, S. & Wang, Y. Peptidylarginine deiminases in citrullination, gene regulation, health and pathogenesis. *Biochim Biophys Acta* **1829**, 1126-35 (2013).

85. Zhang, X. *et al.* Peptidylarginine deiminase 1-catalyzed histone citrullination is essential for early embryo development. *Sci Rep* **6**, 38727 (2016).

86. Zhang, X. *et al.* Peptidylarginine deiminase 2-catalyzed histone H3 arginine 26 citrullination facilitates estrogen receptor alpha target gene activation. *Proc Natl Acad Sci U S A* **109**, 13331-6 (2012).

87. Hagiwara, T., Hidaka, Y. & Yamada, M. Deimination of histone H2A and H4 at arginine 3 in HL-60 granulocytes. *Biochemistry* **44**, 5827-34 (2005).

88. Sharma, P. *et al.* Citrullination of histone H3 interferes with HP1-mediated transcriptional repression. *PLoS Genet* **8**, e1002934 (2012).

89. Cherrington, B.D., Morency, E., Struble, A.M., Coonrod, S.A. & Wakshlag, J.J. Potential role for peptidylarginine deiminase 2 (PAD2) in citrullination of canine mammary epithelial cell histones. *PLoS One* **5**, e11768 (2010).

90. Guertin, M.J. *et al.* Targeted H3R26 deimination specifically facilitates estrogen receptor binding by modifying nucleosome structure. *PLoS Genet* **10**, e1004613 (2014).

91. Li, P. *et al.* PAD4 is essential for antibacterial innate immunity mediated by neutrophil extracellular traps. *J Exp Med* **207**, 1853-62 (2010).

92. Brinkmann, V. *et al.* Neutrophil extracellular traps kill bacteria. *Science* **303**, 1532-5 (2004).

93. Urban, C.F., Reichard, U., Brinkmann, V. & Zychlinsky, A. Neutrophil extracellular traps capture and kill Candida albicans yeast and hyphal forms. *Cell Microbiol* **8**, 668-76 (2006).

94. Wang, Y. *et al.* Histone hypercitrullination mediates chromatin decondensation and neutrophil extracellular trap formation. *J Cell Biol* **184**, 205-13 (2009).

95. Velazquez Camacho, O. *et al.* Major satellite repeat RNA stabilize heterochromatin retention of Suv39h enzymes by RNA-nucleosome association and RNA:DNA hybrid formation. *Elife* **6**(2017).

96. Maison, C. *et al.* SUMOylation promotes de novo targeting of HP1alpha to pericentric heterochromatin. *Nat Genet* **43**, 220-7 (2011).

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

98. James, T.C. *et al.* Distribution patterns of HP1, a heterochromatin-associated nonhistone chromosomal protein of Drosophila. *Eur J Cell Biol* **50**, 170-80 (1989).

99. James, T.C. & Elgin, S.C. Identification of a nonhistone chromosomal protein associated with heterochromatin in Drosophila melanogaster and its gene. *Mol Cell Biol* **6**, 3862-72 (1986).

100. Eissenberg, J.C. *et al.* Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in Drosophila melanogaster. *Proc Natl Acad Sci U S A* **87**, 9923-7 (1990).

101. Festenstein, R. *et al.* Heterochromatin protein 1 modifies mammalian PEV in a dose- and chromosomal-context-dependent manner. *Nat Genet* **23**, 457-61 (1999).

102. Eissenberg, J.C., Morris, G.D., Reuter, G. & Hartnett, T. The heterochromatin-associated protein HP-1 is an essential protein in Drosophila with dosage-dependent effects on position-effect variegation. *Genetics* **131**, 345-52 (1992).

103. Ball, L.J. *et al.* Structure of the chromatin binding (chromo) domain from mouse modifier protein 1. *EMBO J* **16**, 2473-81 (1997).

104. Brasher, S.V. *et al.* The structure of mouse HP1 suggests a unique mode of single peptide recognition by the shadow chromo domain dimer. *EMBO J* **19**, 1587-97 (2000).

105. Cowieson, N.P., Partridge, J.F., Allshire, R.C. & McLaughlin, P.J. Dimerisation of a chromo shadow domain and distinctions from the chromodomain as revealed by structural analysis. *Curr Biol* **10**, 517-25 (2000).

106. Nielsen, P.R. *et al.* Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. *Nature* **416**, 103-7 (2002).

107. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. & Jenuwein, T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116-20 (2001).

108. Jacobs, S.A. & Khorasanizadeh, S. Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science* **295**, 2080-3 (2002).

109. Cheutin, T. *et al.* Maintenance of stable heterochromatin domains by dynamic HP1 binding. *Science* **299**, 721-5 (2003).

110. Hiragami, K. & Festenstein, R. Heterochromatin protein 1: a pervasive controlling influence. *Cell Mol Life Sci* **62**, 2711-26 (2005).

111. Iyengar, S. & Farnham, P.J. KAP1 protein: an enigmatic master regulator of the genome. *J Biol Chem* **286**, 26267-76 (2011).

112. Ecco, G., Imbeault, M. & Trono, D. KRAB zinc finger proteins. *Development* **144**, 2719-2729 (2017).

113. Hediger, F. & Gasser, S.M. Heterochromatin protein 1: don't judge the book by its cover! *Curr Opin Genet Dev* **16**, 143-50 (2006).

114. Canzio, D., Larson, A. & Narlikar, G.J. Mechanisms of functional promiscuity by HP1 proteins. *Trends Cell Biol* **24**, 377-86 (2014).

115. Meehan, R.R., Kao, C.F. & Pennings, S. HP1 binding to native chromatin in vitro is determined by the hinge region and not by the chromodomain. *EMBO J* **22**, 3164-74 (2003).

116. Muchardt, C. *et al.* Coordinated methyl and RNA binding is required for heterochromatin localization of mammalian HP1alpha. *EMBO Rep* **3**, 975-81 (2002).

117. Mishima, Y. *et al.* Hinge and chromoshadow of HP1alpha participate in recognition of K9 methylated histone H3 in nucleosomes. *J Mol Biol* **425**, 54-70 (2013).

118. Mishima, Y. *et al.* Nucleosome compaction facilitates HP1gamma binding to methylated H3K9. *Nucleic Acids Res* **43**, 10200-12 (2015).

119. Daujat, S., Zeissler, U., Waldmann, T., Happel, N. & Schneider, R. HP1 binds specifically to Lys26-methylated histone H1.4, whereas simultaneous Ser27 phosphorylation blocks HP1 binding. *J Biol Chem* **280**, 38090-5 (2005).

120. Holoch, D. & Moazed, D. RNA-mediated epigenetic regulation of gene expression. *Nat Rev Genet* **16**, 71-84 (2015).

121. Lemaitre, C. & Soutoglou, E. Double strand break (DSB) repair in heterochromatin and heterochromatin proteins in DSB repair. *DNA Repair (Amst)* **19**, 163-8 (2014).

122. Kalousi, A. *et al.* The nuclear oncogene SET controls DNA repair by KAP1 and HP1 retention to chromatin. *Cell Rep* **11**, 149-63 (2015).

123. Soria, G. & Almouzni, G. Differential contribution of HP1 proteins to DNA end resection and homology-directed repair. *Cell Cycle* **12**, 422-9 (2013).

124. Fischle, W. *et al.* Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* **438**, 1116-22 (2005).

125. Zeng, W., Ball, A.R., Jr. & Yokomori, K. HP1: heterochromatin binding proteins working the genome. *Epigenetics* **5**, 287-92 (2010).

126. Singh, P.B. HP1 proteins--what is the essential interaction? *Genetika* **46**, 1424-9 (2010).

127. Abe, K. *et al.* Loss of heterochromatin protein 1 gamma reduces the number of primordial germ cells via impaired cell cycle progression in mice. *Biol Reprod* **85**, 1013-24 (2011).

128. Takada, Y. *et al.* HP1gamma links histone methylation marks to meiotic synapsis in mice. *Development* **138**, 4207-17 (2011).

129. Hiragami-Hamada, K. *et al.* N-terminal phosphorylation of HP1{alpha} promotes its chromatin binding. *Mol Cell Biol* **31**, 1186-200 (2011).

130. Strom, A.R. *et al.* Phase separation drives heterochromatin domain formation. *Nature* **547**, 241-245 (2017).

131. Larson, A.G. *et al.* Liquid droplet formation by HP1alpha suggests a role for phase separation in heterochromatin. *Nature* **547**, 236-240 (2017).

132. Lavigne, M. *et al.* Interaction of HP1 and Brg1/Brm with the globular domain of histone H3 is required for HP1-mediated repression. *PLoS Genet* **5**, e1000769 (2009).

133. Dawson, M.A. *et al.* JAK2 phosphorylates histone H3Y41 and excludes HP1alpha from chromatin. *Nature* **461**, 819-22 (2009).

134. Nielsen, S.J. *et al.* Rb targets histone H3 methylation and HP1 to promoters. *Nature* **412**, 561-5 (2001).

135. Mateescu, B., Bourachot, B., Rachez, C., Ogryzko, V. & Muchardt, C. Regulation of an inducible promoter by an HP1beta-HP1gamma switch. *EMBO Rep* **9**, 267-72 (2008).

136. Billur, M., Bartunik, H.D. & Singh, P.B. The essential function of HP1 beta: a case of the tail wagging the dog? *Trends Biochem Sci* **35**, 115-23 (2010).

137. Ayyanathan, K. *et al.* Regulated recruitment of HP1 to a euchromatic gene induces mitotically heritable, epigenetic gene silencing: a mammalian cell culture model of gene variegation. *Genes Dev* **17**, 1855-69 (2003).

138. Ogawa, H., Ishiguro, K., Gaubatz, S., Livingston, D.M. & Nakatani, Y. A complex with chromatin modifiers that occupies E2F- and Myc-responsive genes in G0 cells. *Science* **296**, 1132-6 (2002).

139. Vicent, G.P. *et al.* Induction of progesterone target genes requires activation of Erk and Msk kinases and phosphorylation of histone H3. *Mol Cell* **24**, 367-81 (2006).

140. Vicent, G.P. *et al.* Unliganded progesterone receptor-mediated targeting of an RNA-containing repressive complex silences a subset of hormone-inducible genes. *Genes Dev* **27**, 1179-97 (2013).

141. Harouz, H. *et al.* Shigella flexneri targets the HP1gamma subcode through the phosphothreonine lyase OspF. *EMBO J* **33**, 2606-22 (2014).

142. Smallwood, A., Esteve, P.O., Pradhan, S. & Carey, M. Functional cooperation between HP1 and DNMT1 mediates gene silencing. *Genes Dev* **21**, 1169-78 (2007).

143. Smallwood, A. *et al.* CBX3 regulates efficient RNA processing genome-wide. *Genome Res* **22**, 1426-36 (2012).

144. Kim, H., Heo, K., Choi, J., Kim, K. & An, W. Histone variant H3.3 stimulates HSP70 transcription through cooperation with HP1gamma. *Nucleic Acids Res* **39**, 8329-41 (2011).

145. Saint-Andre, V., Batsche, E., Rachez, C. & Muchardt, C. Histone H3 lysine 9 trimethylation and HP1gamma favor inclusion of alternative exons. *Nat Struct Mol Biol* **18**, 337-44 (2011).

146. Kwon, S.H. *et al.* Heterochromatin protein 1 (HP1) connects the FACT histone chaperone complex to the phosphorylated CTD of RNA polymerase II. *Genes Dev* **24**, 2133-45 (2010).

147. Skourti-Stathaki, K., Kamieniarz-Gdula, K. & Proudfoot, N.J. R-loops induce repressive chromatin marks over mammalian gene terminators. *Nature* **516**, 436-9 (2014).

148. LeRoy, G. *et al.* Heterochromatin protein 1 is extensively decorated with histone code-like post-translational modifications. *Mol Cell Proteomics* **8**, 2432-42 (2009).

149. Machado, M.R., Dans, P.D. & Pantano, S. Isoform-specific determinants in the HP1 binding to histone 3: insights from molecular simulations. *Amino Acids* **38**, 1571-81 (2010).

150. Lomberk, G., Bensi, D., Fernandez-Zapico, M.E. & Urrutia, R. Evidence for the existence of an HP1-mediated subcode within the histone code. *Nat Cell Biol* **8**, 407-15 (2006).

151. Thorne, J.L., Ouboussad, L. & Lefevre, P.F. Heterochromatin protein 1 gamma and IkappaB kinase alpha interdependence during tumour necrosis factor gene transcription elongation in activated macrophages. *Nucleic Acids Res* **40**, 7676-89 (2012).

152. Koike, N., Maita, H., Taira, T., Ariga, H. & Iguchi-Ariga, S.M. 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 Lett* **467**, 17-21 (2000).

153. Akaike, Y. *et al.* Homeodomain-interacting protein kinase 2 regulates DNA damage response through interacting with heterochromatin protein 1gamma. *Oncogene* **34**, 3463-73 (2015).

154. Grzenda, A. *et al.* Functional impact of Aurora A-mediated phosphorylation of HP1gamma at serine 83 during cell cycle progression. *Epigenetics Chromatin* **6**, 21 (2013).

155. Liu, M. *et al.* Heterochromatin protein HP1gamma promotes colorectal cancer progression and is regulated by miR-30a. *Cancer Res* **75**, 4593-604 (2015).

156. Fan, Y., Li, H., Liang, X. & Xiang, Z. CBX3 promotes colon cancer cell proliferation by CDK6 kinase-independent function during cell cycle. *Oncotarget* **8**, 19934-19946 (2017).

157. Han, S.S. *et al.* RNA sequencing identifies novel markers of non-small cell lung cancer. *Lung Cancer* **84**, 229-35 (2014).

158. Slezak, J., Truong, M., Huang, W. & Jarrard, D. HP1gamma expression is elevated in prostate cancer and is superior to Gleason score as a predictor of biochemical recurrence after radical prostatectomy. *BMC Cancer* **13**, 148 (2013).

159. Takanashi, M. *et al.* Heterochromatin protein 1gamma epigenetically regulates cell differentiation and exhibits potential as a therapeutic target for various types of cancers. *Am J Pathol* **174**, 309-16 (2009).

160. Liang, Y.K., Lin, H.Y., Chen, C.F. & Zeng. Prognostic values of distinct CBX family members in breast cancer. *Oncotarget* **8**, 92375-92387 (2017).

161. Sun, M. *et al.* Cbx3/HP1gamma deficiency confers enhanced tumor-killing capacity on CD8+ T cells. *Sci Rep* **7**, 42888 (2017).

162. Ahmad, N. & Kumar, R. Steroid hormone receptors in cancer development: a target for cancer therapeutics. *Cancer Lett* **300**, 1-9 (2011).

163. Dawson, M.A. & Kouzarides, T. Cancer epigenetics: from mechanism to therapy. *Cell* **150**, 12-27 (2012).

164. Canudas, S. *et al.* A role for heterochromatin protein 1gamma at human telomeres. *Genes Dev* **25**, 1807-19 (2011).

165. Hanahan, D. & Weinberg, R.A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-74 (2011).

166. Choi, J.D., Park, M.A. & Lee, J.S. Suppression and recovery of BRCA1-mediated transcription by HP1gamma via modulation of promoter occupancy. *Nucleic Acids Res* **40**, 11321-38 (2012).

167. Bot, C. *et al.* Independent mechanisms recruit the cohesin loader protein NIPBL to sites of DNA damage. *J Cell Sci* **130**, 1134-1146 (2017).

168. Meshorer, E. & Misteli, T. Chromatin in pluripotent embryonic stem cells and differentiation. *Nat Rev Mol Cell Biol* **7**, 540-6 (2006).

169. Dialynas, G.K. *et al.* Plasticity of HP1 proteins in mammalian cells. *J Cell Sci* **120**, 3415-24 (2007).

170. Morikawa, K., Ikeda, N., Hisatome, I. & Shirayoshi, Y. Heterochromatin protein 1gamma overexpression in P19 embryonal carcinoma cells elicits spontaneous differentiation into the three germ layers. *Biochem Biophys Res Commun* **431**, 225-31 (2013).

171. Huang, C. *et al.* Cbx3 maintains lineage specificity during neural differentiation. *Genes Dev* **31**, 241-246 (2017).

172. Sridharan, R. *et al.* Proteomic and genomic approaches reveal critical functions of H3K9 methylation and heterochromatin protein-1gamma in reprogramming to pluripotency. *Nat Cell Biol* **15**, 872-82 (2013).

173. Festenstein, R. *et al.* Modulation of heterochromatin protein 1 dynamics in primary Mammalian cells. *Science* **299**, 719-21 (2003).

174. Schmiedeberg, L., Weisshart, K., Diekmann, S., Meyer Zu Hoerste, G. & Hemmerich, P. High- and low-mobility populations of HP1 in heterochromatin of mammalian cells. *Mol Biol Cell* **15**, 2819-33 (2004).

175. Muller, K.P. *et al.* Multiscale analysis of dynamics and interactions of heterochromatin protein 1 by fluorescence fluctuation microscopy. *Biophys J* **97**, 2876-85 (2009).

176. Meshorer, E. *et al.* Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev Cell* **10**, 105-16 (2006).

177. Hahne, F. *et al.* flowCore: a Bioconductor package for high throughput flow cytometry. *BMC Bioinformatics* **10**, 106 (2009).

178. Weimann, L. *et al.* A quantitative comparison of single-dye tracking analysis tools using Monte Carlo simulations. *PLoS One* **8**, e64287 (2013).

179. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-60 (2009).

180. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).

181. Quinlan, A.R. & Hall, I.M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-2 (2010).

182. Kent, W.J., Zweig, A.S., Barber, G., Hinrichs, A.S. & Karolchik, D. BigWig and BigBed: enabling browsing of large distributed datasets. *Bioinformatics* **26**, 2204-7 (2010).

183. Ramirez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**, W160-5 (2016).

184. Schroeder, A. *et al.* The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol* **7**, 3 (2006).

185. Parkhomchuk, D. *et al.* Transcriptome analysis by strand-specific sequencing of complementary DNA. *Nucleic Acids Res* **37**, e123 (2009).

186. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).

187. Aken, B.L. *et al.* Ensembl 2017. *Nucleic Acids Res* **45**, D635-D642 (2017).

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

189. Luo, W., Friedman, M.S., Shedden, K., Hankenson, K.D. & Woolf, P.J. GAGE: generally applicable gene set enrichment for pathway analysis. *BMC Bioinformatics* **10**, 161 (2009).

190. Christophorou, M.A. *et al.* Citrullination regulates pluripotency and histone H1 binding to chromatin. *Nature* **507**, 104-8 (2014).

191. De Ceuleneer, M., Van Steendam, K., Dhaenens, M. & Deforce, D. In vivo relevance of citrullinated proteins and the challenges in their detection. *Proteomics* **12**, 752-60 (2012).

192. Basu, S., Tan, Y.L., Taylor, E.J., Laue, E.D. & Lee, S.F. Studying the Dynamics of Chromatin-Binding Proteins in Mammalian Cells Using Single-Molecule Localisation Microscopy. *Methods Mol Biol* **1431**, 235-63 (2016).

193. Ricci, M.A., Cosma, M.P. & Lakadamyali, M. Super resolution imaging of chromatin in pluripotency, differentiation, and reprogramming. *Curr Opin Genet Dev* **46**, 186-193 (2017).

194. Platero, J.S., Hartnett, T. & Eissenberg, J.C. Functional analysis of the chromo domain of HP1. *EMBO J* **14**, 3977-86 (1995).

195. Etheridge, T.J. *et al.* Quantification of DNA-associated proteins inside eukaryotic cells using single-molecule localization microscopy. *Nucleic Acids Res* **42**, e146 (2014).

196. Chen, J. *et al.* Single-molecule dynamics of enhanceosome assembly in embryonic stem cells. *Cell* **156**, 1274-85 (2014).

197. Los, G.V. *et al.* HaloTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS Chem Biol* **3**, 373-82 (2008).

198. Grimm, J.B. *et al.* A general method to improve fluorophores for live-cell and single-molecule microscopy. *Nat Methods* **12**, 244-50, 3 p following 250 (2015).

199. Hiragami-Hamada, K. *et al.* Dynamic and flexible H3K9me3 bridging via HP1beta dimerization establishes a plastic state of condensed chromatin. *Nat Commun* **7**, 11310 (2016).

200. Izeddin, I. *et al.* Single-molecule tracking in live cells reveals distinct target-search strategies of transcription factors in the nucleus. *Elife* **3**(2014).

201. Gaspar-Maia, A., Alajem, A., Meshorer, E. & Ramalho-Santos, M. Open chromatin in pluripotency and reprogramming. *Nat Rev Mol Cell Biol* **12**, 36-47 (2011).

202. Chen, H., Lin, R.J., Xie, W., Wilpitz, D. & Evans, R.M. Regulation of hormone-induced histone hyperacetylation and gene activation via acetylation of an acetylase. *Cell* **98**, 675-86 (1999).

203. Deplus, R. *et al.* Citrullination of DNMT3A by PADI4 regulates its stability and controls DNA methylation. *Nucleic Acids Res* **42**, 8285-96 (2014).

204. Tsai, M.J. & O'Malley, B.W. Molecular mechanisms of action of steroid/thyroid receptor superfamily members. *Annu Rev Biochem* **63**, 451-86 (1994).

205. Lee, Y.H., Coonrod, S.A., Kraus, W.L., Jelinek, M.A. & Stallcup, M.R. Regulation of coactivator complex assembly and function by protein arginine methylation and demethylimination. *Proc Natl Acad Sci U S A* **102**, 3611-6 (2005).

206. Dialynas, G.K., Vitalini, M.W. & Wallrath, L.L. Linking Heterochromatin Protein 1 (HP1) to cancer progression. *Mutat Res* **647**, 13-20 (2008).

207. Soule, H.D., Vazguez, J., Long, A., Albert, S. & Brennan, M. A human cell line from a pleural effusion derived from a breast carcinoma. *J Natl Cancer Inst* **51**, 1409-16 (1973).

208. Brenton, J.D., Carey, L.A., Ahmed, A.A. & Caldas, C. Molecular classification and molecular forecasting of breast cancer: ready for clinical application? *J Clin Oncol* **23**, 7350-60 (2005).

209. Cailleau, R., Young, R., Olive, M. & Reeves, W.J., Jr. Breast tumor cell lines from pleural effusions. *J Natl Cancer Inst* **53**, 661-74 (1974).

210. Prat, A. *et al.* Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res* **12**, R68 (2010).

211. Creighton, C.J. *et al.* Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features. *Proc Natl Acad Sci U S A* **106**, 13820-5 (2009).

212. Dull, T. *et al.* A third-generation lentivirus vector with a conditional packaging system. *J Virol* **72**, 8463-71 (1998).

213. Koopman, G. *et al.* Annexin V for flow cytometric detection of phosphatidylserine expression on B cells undergoing apoptosis. *Blood* **84**, 1415-20 (1994).

214. Whittemore, E.R., Loo, D.T., Watt, J.A. & Cotman, C.W. A detailed analysis of hydrogen peroxide-induced cell death in primary neuronal culture. *Neuroscience* **67**, 921-32 (1995).

215. Lazebnik, Y.A., Kaufmann, S.H., Desnoyers, S., Poirier, G.G. & Earnshaw, W.C. Cleavage of poly(ADP-ribose) polymerase by a proteinase with properties like ICE. *Nature* **371**, 346-7 (1994).

216. Li, J. & Yuan, J. Caspases in apoptosis and beyond. *Oncogene* **27**, 6194-206 (2008).

217. Kaufmann, S.H., Desnoyers, S., Ottaviano, Y., Davidson, N.E. & Poirier, G.G. Specific proteolytic cleavage of poly(ADP-ribose) polymerase: an early marker of chemotherapy-induced apoptosis. *Cancer Res* **53**, 3976-85 (1993).

218. Adams, R.R., Carmena, M. & Earnshaw, W.C. Chromosomal passengers and the (aurora) ABCs of mitosis. *Trends Cell Biol* **11**, 49-54 (2001).

219. Bradford, J.A. & Clarke, S.T. Dual-pulse labeling using 5-ethynyl-2'-deoxyuridine (EdU) and 5-bromo-2'-deoxyuridine (BrdU) in flow cytometry. *Curr Protoc Cytom* **Chapter 7**, Unit 7 38 (2011).

220. Taylor, R.C., Cullen, S.P. & Martin, S.J. Apoptosis: controlled demolition at the cellular level. *Nat Rev Mol Cell Biol* **9**, 231-41 (2008).

221. Liang, Y., Yan, C. & Schor, N.F. Apoptosis in the absence of caspase 3. *Oncogene* **20**, 6570-8 (2001).

222. Olivier, M. *et al.* The IARC TP53 database: new online mutation analysis and recommendations to users. *Hum Mutat* **19**, 607-14 (2002).

223. Lu, X., Errington, J., Curtin, N.J., Lunec, J. & Newell, D.R. The impact of p53 status on cellular sensitivity to antifolate drugs. *Clin Cancer Res* **7**, 2114-23 (2001).

224. Chipuk, J.E. *et al.* Direct activation of Bax by p53 mediates mitochondrial membrane permeabilization and apoptosis. *Science* **303**, 1010-4 (2004).

225. Gilmore, A.P. Anoikis. *Cell Death Differ* **12 Suppl 2**, 1473-7 (2005).

226. Li, D.M. & Feng, Y.M. Signaling mechanism of cell adhesion molecules in breast cancer metastasis: potential therapeutic targets. *Breast Cancer Res Treat* **128**, 7-21 (2011).

227. Gerlitz, G. & Bustin, M. Efficient cell migration requires global chromatin condensation. *J Cell Sci* **123**, 2207-17 (2010).

228. Crisp, M. *et al.* Coupling of the nucleus and cytoplasm: role of the LINC complex. *J Cell Biol* **172**, 41-53 (2006).

229. Gerlitz, G. & Bustin, M. The role of chromatin structure in cell migration. *Trends Cell Biol* **21**, 6-11 (2011).

230. Ye, Q. & Worman, H.J. Interaction between an integral protein of the nuclear envelope inner membrane and human chromodomain proteins homologous to Drosophila HP1. *J Biol Chem* **271**, 14653-6 (1996).

231. Anders, S., Reyes, A. & Huber, W. Detecting differential usage of exons from RNA-seq data. *Genome Res* **22**, 2008-17 (2012).

232. Schindler, C., Levy, D.E. & Decker, T. JAK-STAT signaling: from interferons to cytokines. *J Biol Chem* **282**, 20059-63 (2007).

233. Chawla-Sarkar, M. *et al.* Apoptosis and interferons: role of interferon-stimulated genes as mediators of apoptosis. *Apoptosis* **8**, 237-49 (2003).

234. Chawla-Sarkar, M., Leaman, D.W. & Borden, E.C. Preferential induction of apoptosis by interferon (IFN)-beta compared with IFN-alpha2: correlation with TRAIL/Apo2L induction in melanoma cell lines. *Clin Cancer Res* **7**, 1821-31 (2001).

235. Leaman, D.W. *et al.* Identification of X-linked inhibitor of apoptosis-associated factor-1 as an interferon-stimulated gene that augments TRAIL Apo2L-induced apoptosis. *J Biol Chem* **277**, 28504-11 (2002).

236. He, F. *et al.* A positive autoregulatory loop of Jak-STAT signaling controls the onset of astrogliogenesis. *Nat Neurosci* **8**, 616-25 (2005).

237. Zhou, J. *et al.* Overexpression of HP1gamma is associated with poor prognosis in non-small cell lung cancer cell through promoting cell survival. *Tumour Biol* **35**, 9777-85 (2014).

238. Ryu, H. *et al.* Adipose tissue-derived mesenchymal stem cells cultured at high density express IFN-beta and suppress the growth of MCF-7 human breast cancer cells. *Cancer Lett* **352**, 220-7 (2014).

239. Duronio, R.J. & Xiong, Y. Signaling pathways that control cell proliferation. *Cold Spring Harb Perspect Biol* **5**, a008904 (2013).

240. Ryan, R.F. *et al.* KAP-1 corepressor protein interacts and colocalizes with heterochromatic and euchromatic HP1 proteins: a potential role for Kruppel-associated box-zinc finger proteins in heterochromatin-mediated gene silencing. *Mol Cell Biol* **19**, 4366-78 (1999).

241. Iyengar, S., Ivanov, A.V., Jin, V.X., Rauscher, F.J., 3rd & Farnham, P.J. Functional analysis of KAP1 genomic recruitment. *Mol Cell Biol* **31**, 1833-47 (2011).

242. Tian, B., Yang, J. & Brasier, A.R. Two-step cross-linking for analysis of protein-chromatin interactions. *Methods Mol Biol* **809**, 105-20 (2012).

243. D, G.H., Kelley, D.R., Tenen, D., Bernstein, B. & Rinn, J.L. Widespread RNA binding by chromatin-associated proteins. *Genome Biol* **17**, 28 (2016).

244. Core, L.J., Waterfall, J.J. & Lis, J.T. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322**, 1845-8 (2008).

245. Churchman, L.S. & Weissman, J.S. Native elongating transcript sequencing (NET-seq). *Curr Protoc Mol Biol* **Chapter 4**, Unit 4 14 1-17 (2012).

246. Natsume, T., Kiyomitsu, T., Saga, Y. & Kanemaki, M.T. Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors. *Cell Rep* **15**, 210-8 (2016).

247. Mohammed, H. *et al.* Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes. *Nat Protoc* **11**, 316-26 (2016).

248. Vogel, M.J. *et al.* Human heterochromatin proteins form large domains containing KRAB-ZNF genes. *Genome Res* **16**, 1493-504 (2006).

249. Blahnik, K.R. *et al.* Characterization of the contradictory chromatin signatures at the 3' exons of zinc finger genes. *PLoS One* **6**, e17121 (2011).

250. O'Geen, H. *et al.* Genome-wide analysis of KAP1 binding suggests autoregulation of KRAB-ZNFs. *PLoS Genet* **3**, e89 (2007).

251. Tchasovnikarova, I.A. *et al.* GENE SILENCING. Epigenetic silencing by the HUSH complex mediates position-effect variegation in human cells. *Science* **348**, 1481-1485 (2015).

252. Schultz, D.C., Friedman, J.R. & Rauscher, F.J., 3rd. Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2alpha subunit of NuRD. *Genes Dev* **15**, 428-43 (2001).

253. Emerson, R.O. & Thomas, J.H. Adaptive evolution in zinc finger transcription factors. *PLoS Genet* **5**, e1000325 (2009).

254. Hancks, D.C. & Kazazian, H.H., Jr. Roles for retrotransposon insertions in human disease. *Mob DNA* **7**, 9 (2016).

255. Ayarpadikannan, S., Lee, H.E., Han, K. & Kim, H.S. Transposable element-driven transcript diversification and its relevance to genetic disorders. *Gene* **558**, 187-94 (2015).

256. Chiappinelli, K.B. *et al.* Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses. *Cell* **162**, 974-86 (2015).

257. Roulois, D. *et al.* DNA-Demethylating Agents Target Colorectal Cancer Cells by Inducing Viral Mimicry by Endogenous Transcripts. *Cell* **162**, 961-73 (2015).

258. Guler, G.D. *et al.* Repression of Stress-Induced LINE-1 Expression Protects Cancer Cell Subpopulations from Lethal Drug Exposure. *Cancer Cell* **32**, 221-237 e13 (2017).

259. Cossetti, C. *et al.* Extracellular vesicles from neural stem cells transfer IFN-gamma via Ifngr1 to activate Stat1 signaling in target cells. *Mol Cell* **56**, 193-204 (2014).