Dissertation zur Erlangung des Doktorgrades

der Naturwissenschaften an der Fakultät für Biologie

der Ludwig-Maximilians-Universität München

Proteomic analysis of specific chromatin domains containing the histone variants H2A.Bbd and macroH2A.1.2

Viola Sansoni

aus

Priverno, Italien

Dissertation eingereicht am: 26-05-2014

Mündliche Prüfung am: 31-07-2014

1. Gutachter: Peter Becker
2. Gutachter: Dirk Eick
3. Gutachter: Anna Friedl
4. Gutachter: Michael Boshart
5. Gutachter: Heinrich Leonhardt
6. Gutachter: Kirsten Jung

Eidesstattliche Erklärung

Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbständig und ohne unerlaubte Hilfe angefertigt ist.

…….	**…….**
München, den	**(Unterschrift)**

Table of contents

1 Abstract

In eukaryotes, chromatin is a highly dynamic and well organized nucleoprotein complex of DNA and histones that influences several important biological processes such as DNA replication, transcription and repair. The fundamental unit of chromatin is the nucleosome consisting of 147 bp of DNA wrapped around an octamer of four histone proteins: H3, H4, H2A and H2B. The structure of nucleosome and chromatin fibres can be modulated in various ways in response to specific biological signals. Together with the presence of specific histone modifications, histone variants are key players in modulating the nucleosome stability and accessibility; thus influencing the overall chromatin structure.

Two particular histone variants, H2A.Bbd and macroH2A.1.2 have been extensively studied in recent years due to their influence on the nucleosome structure and their role in transcription regulation. However not so much is known about the proteomic profile of the chromatin domains that contain H2A.Bbd and macroH2A.1.2.

This PhD project provides a comprehensive analysis of the proteins associated to specific chromatin domains containing the histone variants H2A.Bbd and macroH2A.1.2. Based on the proteomic analysis, it has been found that macroH2A.1.2 is enriched in heterochromatic factors such as the Polycomb Repressive Complexes 1 and 2 (PRC2 and PRC1) and the methyltransferase Suv39h1, whereas H2A.Bbd is mainly present at active chromatin marked by factors involved in splicing and DNA replication.

Moreover, the combination of mass spectrometry and immunofluorescence-based approaches revealed that H2A.Bbd is recruited at sites of DNA synthesis during replication and DNA repair. Interestingly, the cells that overexpress this histone variant show faster replication and higher sensitivity to DNA damage. Furthermore, this study shows for the first time that Hodgkin's Lymphoma (HL) cell lines abnormally express H2A.Bbd and that the expression levels of this histone variant inversely correlate with the HL doubling time.

In summary, this study describes the protein composition of the chromatin containing the histone variants H2A.Bbd and macroH2A.1.2 and reveals new, interesting features of H2A.Bbd and its role in spermatogenesis and cancer.

2 Introduction

2.1 Chromatin

The word ‘Chromatin’ was first used by Flemming at the end of the 19th century to differentiate between a stainable material present in the nuclei and a non-stainable material that he defined as ‘chromatin’ (Flemming, 1882).

Shortly after, at the beginning of the 20th century Heitz observed that various chromatin regions behaved differently during cell replication, already suggesting that chromatin is not uniform in structure. Due to their different staining behavior he defined those regions as ‘euchromatin’ and ‘heterochromatin’. After more than one century from its first description, progresses have been made in understanding chromatin structure and regulation. It has become more and more apparent that chromatin is a dynamic complex of DNA wound around basic histone proteins, and that the higher order chromatin structure greatly influences gene regulation. The classification of chromatin types has developed into a much more complex system than the simple euchromatin- heterochromatin dichotomy, resulting in a color code scheme where chromatin is defined by at least five different classes according to its protein composition (Filion et al., 2010). Chromatin is a highly dynamic structure regulated by several features such as specific histone variants and/or histone modifications. These two specific chromatin features alter the intra- and inter-nucleosomal interactions and recruit specific protein factors that, consequently, modulate gene expression and transcription.

2.2 The Nucleosome

The first evidence that the DNA is present in the nucleus in particles and wound around a histone protein core was found in the early 1970’s (Noll, 1974; Van Holde et al., 1974). During 1974-75 the nucleosome was characterized and defined as the fundamental chromatin unit. It has been shown that the chromatin of various species can be digested by different endonucleases giving rise to fragments of approximately 200 bp and that histone H3 and H4 form the core of these particles (Hewish and Burgoyne, 1973; Kornberg, 1974; Kornberg and Thomas, 1974).

Histone proteins can be classified in core histones (H3, H4, H2A and H2B) and linker histones (H1 family). It was not until 1997 that the atomic structure of the nucleosome was resolved at a resolution of 2.8 Å (Figure 2.1).

Figure 2.1 Nucleosome core particle structure published by Karolin Luger in 1997.
Two different views of the nucleosome particle taken from down the DNA superhelix (left) and perpendicular to it (right). The eight histones proteins forming the octamer are visible: blu-H3; green-H4; yellow-H2A; red-H2B (Figure taken from (Luger et al., 1997).

The nucleosome structure revealed that 146 bp of DNA are wrapped around the histone octamer in 1.65 turns of a flat, left handed superhelix. The central core histone fold domains maintained a high structural similarity during evolution due to their important role in the nucleosome structure. They are formed by three α-helices connected by two loops: L1 and L2. The histone octamer is formed by a tetramer of H3-H4 $(H3\text{-}H4)_2$ and two dimers of H2A-H2B. H3 and H4 form the tetramer via the interaction between H3 and H3' through a 4-helix bundle (4-HB). The interaction of the H2A and H2B pair with the tetramer takes place through a second homologous 4-helix bundle between H2A and H4. The C-terminal docking domain of the histone H2A directs the interaction of the nucleosome with the DNA and it is important for nucleosome stability. These interactions also involve the N-terminal tail of H3 (Luger et al., 1997).

In addition to the core domains every histone contains N-terminal or C-terminal extensions. The histone N-terminal tail protrudes away from the histone octamer structure. In contrast to the histone core domains, the N-terminal histone tails are unstructured and not visible in any of the crystal structures elaborated so far. Nevertheless, the histone tails sequences are highly conserved among species, implying an important role in the regulation of the nucleosome function and the organization of the higher order chromatin structure.

2.3 Higher order chromatin structure

How chromatin is organized in the nucleus in higher ordered structure is still widely debated. The major difficulty scientists encountered in the past was the absence of an experimental tool that allowed for the study of chromatin structure in the presence of an intact nucleus. For this reason all the models obtained so far are based on in vitro experiments where chromatin is assembled and folded, or only folded, outside its natural context.

What is known is that the nucleosome compaction into chromatin structures requires the inter-nucleosomes interactions that involve both the core histone tails and their modifications. The inter-nucleosome interactions promote the formation of the 30 nm fibres from an 11 nm nucleosome array (the beads on a string) (Figure 2.2).

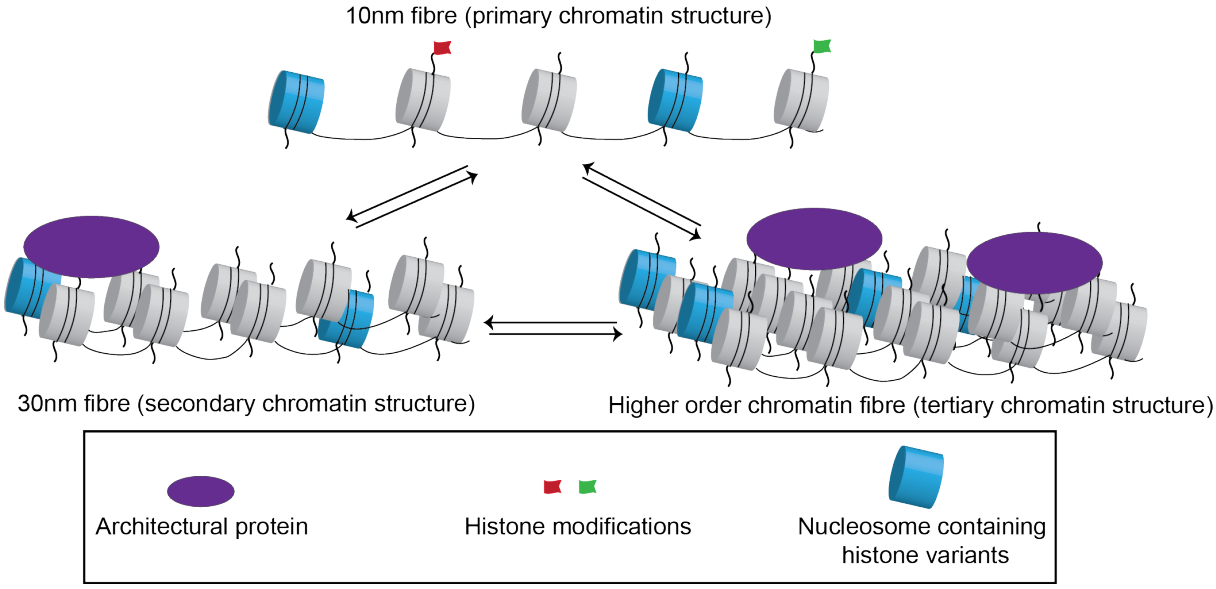

Figure 2.2 Higher order chromatin structure.
The primary structure of chromatin is represented as a nucleosomal array composed either by nucleosome containing the canonical histones (light grey), or nucleosome containing particular histone variants (light blue). In addition, the primary structure is marked by specific histone modifications (green and red flags). The nucleosome-nucleosome interactions form a folded chromatin fibre (secondary structure). The interactions between chromatin fibres result in the higher order chromatin structure (tertiary structure). Secondary and tertiary chromatin structure are stabilized by architectural proteins (purple ellipse). Double arrows indicate the transitions between different structural levels. These transitions can be regulated by changes in the histone modification profiles, by the histone exchange and by the displacement of the architectural proteins (Figure adapted after (Luger et al., 2012)).

The studies that used Electron Microscopy (EM) performed on chromatin isolated from nuclei, suggested two models for the 30 nm fibre. The first one, the 'solenoid' or the 'one-start helix' model, has 6-8 nucleosomes per turn with adjacent nucleosomes connected by a bent linker DNA (Robinson et al., 2006). The second model proposes two rows of interdigitated

nucleosomes resulting in a two-start helix, which is stabilized by the interaction of alternate nucleosomes. In this model the nucleosomes are connected by straight linker DNA (Dorigo, 2004; Schalch et al., 2005).

The use of chromatin isolated from nuclei has the advantage that it most likely resembles the native state of chromatin. However, it is also very heterogeneous with regards to DNA sequence, histone modifications and nucleosome spacing. A good alternative to the native chromatin is the use of homogenous DNA carrying a well-characterized nucleosome positioning sequence for *in vitro* reconstitution of chromatin. With the 601-sequence it was possible to reduce in part the native chromatin complexity, however the structures obtained from these experiments were still largely depending on the experimental conditions. Indeed, when the 601-sequence was used and EM performed on the crystal structure of a tetranucleosome core array lacking the linker histone, a two start crossed-linker model for the 30 nm fibre was proposed (Dorigo, 2004; Schalch et al., 2005). However, the crystallized tetranucleosome had a nucleosomal repeat length (NRL) of only 167 bp, which is shorter than the typical 200 bp found in most of the eukaryotic organisms. In addition, the concentration of divalent cation Mg^{2+}used for the experiments is not common in nature. The use of a longer array (177 to 237 bp), the inclusion of the linker histone and lower divalent cation Mg^{2+} concentration, revealed a one-start interdigitated solenoid structure (Robinson and Rhodes, 2006). Further experiments revealed that the nucleosome repeat length (NRL) and the linker histone play an important role in the chromatin folding (Robinson and Rhodes, 2006). However, the biological role of the linker histone H1 in the higher order chromatin structure is still under debate (Fan et al., 2005).

A recent study using Electron Microscopy-Assisted Nucleosome Capture (EMANIC) in *in vitro* assembled chromatin showed that chromatin might be rather heteromorphic with different types of fibres organization. Indeed, even if the fibres showed predominantly a two-start organization, the structures also contained bend linker DNA typical of the one-start solenoid (Grigoryev et al., 2009).

As previously mentioned, the methods used so far to study the chromatin structure did not allow for its visualization within the nucleus. This problem was recently solved through the use of a combination of Electron Spectroscopic Imaging (ESI) and electron tomography. Surprisingly, the only regular chromatin structure observed in eukaryotic cell nuclei using this

method was the 10 nm fibre. The 10 nm fibre is present at both euchromatic and heterochromatic regions where the level of compaction is higher (Fussner et al., 2012).

Beyond the formation of the 30 nm fibre there are a series of long-range interactions that promote the formation of the chromatin tertiary structure (Figure 2.2). These long-range interactions, or chromatin loops, have been extensively analyzed thanks to the development and improvement of new techniques such as Fluorescence in situ Hybridization (FISH) or Chromosome Conformation Capture (3C) and Hi-C. In this context, regulatory elements such as promoters, enhancers and terminators play an important structural role. Several proteins have also been identified to contribute to the formation of chromatin loops between enhancers and promoters or promoters and terminators.
The higher-order chromatin structure plays a central role in gene regulation and transcription. In the nucleus, chromatin shows different levels of compaction leading to the formation of particular chromatin domains, which differ in protein composition, histone modifications, DNA methylation and histone variants.

2.4 Chromatin domains

In the nucleus, chromatin is organized into domains that share common features such as the same pattern of histone modifications or the presence of specific chromatin binding factors.
One interesting example of chromatin domains is the Lamina-Associated-Domains (LAD). These are large chromatin domains, identified using EM, that are connected to the nuclear lamina and that share common epigenetic feature. In human fibroblasts the majority of the genes present in the LADs is transcriptionally inactive or shows a low expression level. The histone modification H3K27me3 was found to be enriched on the LADs whereas H3K4 dimethylation is rather low at most promoters of the LADs (Guelen et al., 2008). On the other hand, there are large domains associated with the nuclear pore that are actively transcribed (Sood and Brickner, 2014).

Due to the presence of specific chromatin binding proteins, a new classification for different ‘chromatin types’ was recently suggested in *Drosophila melanogaster*. The five chromatin types were defined using a color code system, with a different color for each category. GREEN and BLUE chromatin correspond to the heterochromatin protein 1 (HP1) and the Polycomb group (PcG) binding domains respectively. Together with these well-characterized domains the authors identified the BLACK chromatin, with no transcriptional activity

detected, and the RED and YELLOW chromatin as two distinct types of euchromatin. The euchromatic domains share several chromatin-binding factors, but they differ in the time of replication, with the RED chromatin replicating earlier than the YELLOW, and the marking of different genes categories (Filion et al., 2010). A more recent classification in *Drosophila melanogaster* obtained by the Encyclopedia of DNA Elements (modENCODE) that integrated multiple data types such as histone modifications, chromatin accessibility, short RNAs, non-histone proteins and transcriptional activity, resulted in the identification of 9-state of chromatin (Kharchenko et al., 2011).

There are rare cases in which entire chromosomes are defined by the presence of specific marks that repress or activate their gene expression. This phenomenon happens in the dosage compensation process when transcription from one sex chromosome needs to be either hyper-activated in the heterogametic sex or repressed in the homogametic to balance gene dosage between the two sexes. In mammals, for example, one of the two X chromosomes is randomly inactivated during female development. The inactive X chromosome is characterized by specific epigenetics repressive marks such as H3K27me3, by the presence of the Polycomb Repressive Complexes 1 and 2 (PRC2 and PRC1), and the ubiquitination of the histone H2A. The maintenance of the repressive state involves the replacement of the histone H2A by the histone variant macroH2A (Brockdorff, 2011).
Together with the histone modifications, the histone variants play a crucial role in regulating the gene expression and the structure of specific chromatin domains. This point will be discussed in details below.

2.5 Histone modifications

One system in which the cell modulates nucleosome mobility and turnover is via chemical modification of histones. Histone modifications are important in every cellular process where DNA access is fundamental i.e. replication, transcription and repair. Ever since the acetylation of histones was associated with transcription activation, the interest in histone modifications increased enormously, leading to the discovery of over 100 distinct histone modifications (Dai et al., 2014). All histones can be modified and most of the modifications occur on the histone tails. Among those, two very important and well-characterized modifications are the histone acetylation and methylation (Figure 2.3). Others have only been recently discovered

and their putative function in cellular processes is still largely unknown (Dai et al., 2014; Tan et al., 2011).

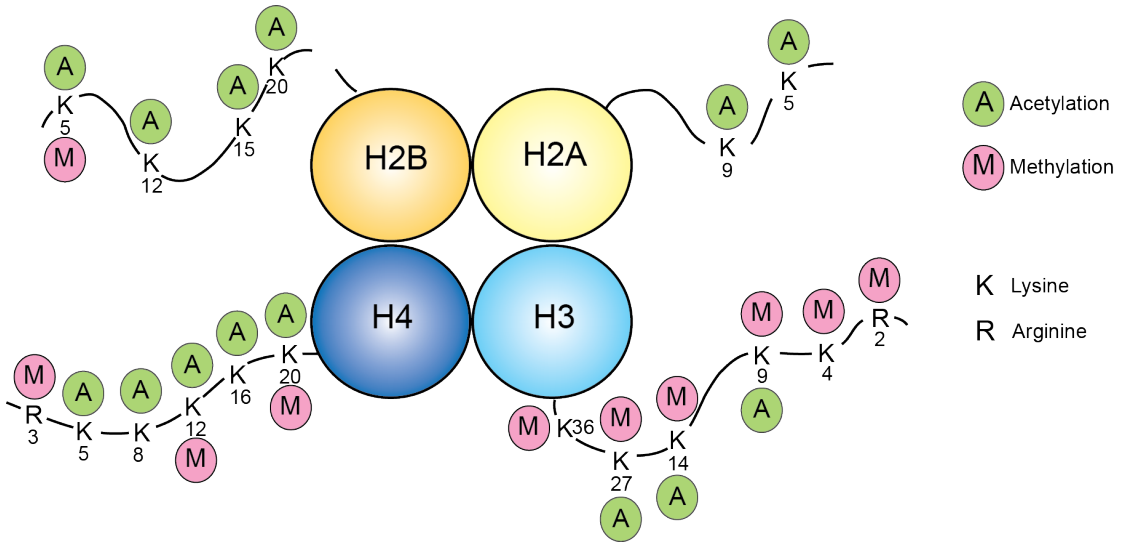

Figure 2.3 Schematic representation of the histone tail methylation and acetylation.
Histone tails are subjected to several post-translational modifications. Two very well characterized modifications are histone methylation and histone acetylation (green and pink circles respectively). By altering the nucleosome structure or by recruiting specific chromatin binding factors, histone methylation and histone acetylation are fundamental for the higher order chromatin structure and for gene expression regulation (Figure adapted after (Marks et al., 2001)).

The first modification discovered in 1961 was histone acetylation and its function was rapidly linked to the gene activation (Allfrey et al., 1964; Pogo et al., 1966). The chemistry of histone acetylation easily explains how this modification could facilitate gene activation. Indeed, the positive charge of lysine residues is neutralized by the acetylation of histones, thus weakening the interactions between the DNA and the nucleosomal histones. This results in an increased accessibility of the DNA to the transcription machinery. Even though acetylation has always been associated with transcription, recent studies show that this modification plays an important role also in replication and DNA repair. This suggests that histone acetylation is important in every process that requires relaxation of the DNA-histone contacts (Unnikrishnan et al., 2010; Xu and Price, 2011). The lysine residues are acetylated by lysine acetyltransferases (HAT) and the acetyl groups are removed by lysine deacetylases (HDAC). A specific protein domain recognizes the acetylated residues: the bromodomain (Tamkun et al., 1992). The bromodomain is found in a broad range of chromatin-associated factors that include the lysine acetyltransferases and also chromatin remodeling factors (Zeng and Zhou, 2002).

A well-known modification, together with the lysine acetylation, is the histone lysine methylation. In general, methylation does not alter the positive charge of the lysine. For this reason, the effect of methylation on nucleosome stability is considered less active than the effect of lysine acetylation. Histone methylation can occur on lysines or arginines, but so far the effect of arginine methylation on the nucleosome dynamics remains largely unclear (Bedford and Clarke, 2009). Interestingly, the histone lysine can be unmodified or gain a mono-, di-, or trimethylated state. However, it is still not clear what influences the methylation state of a given lysine residues. Most likely the accessibility and the exposure duration of a specific residue to its modifying enzyme play an important role in defining the lysine methylation state. Even more interesting is the methylation effect on transcription regulation. Different from lysine acetylation, whose role is mostly related to transcription activation, lysine methylation can be associated to transcription repression and activation. Modifications associated with transcription activation are for example H3K4me3 and H3K36me, whereas methylated residues associated with repression are the lysine 9 (K9) and lysine 27 (K27) on histone H3 (Zentner and Henikoff, 2013).

Even though many modifications have been discovered and characterized, new ones are still emerging. It has been proposed that histone modifications may work singly or in combination thus forming a 'histone code', which could define different expression pattern (Rothbart and Strahl, 2014).

2.6 Histone variants

Histone variants are non-allelic isoforms of canonical histones that differ in primary sequence and expression timing.

While canonical histones are mainly expressed during S-phase, histone variants are transcribed throughout the cell cycle, replacing the canonical histone in a replication-independent way (Wu and Bonner, 1981).

Canonical histones and histone variants differ not only in their time of expression but also in their sequence and, interestingly, in structural characteristics of their mRNA. The canonical histone genes do not contain introns and their mRNA is not polyadenylated but terminates with a stem-loop structure, which is important for the cell cycle regulation of the canonical histones (Busslinger et al., 1979; Kedes et al., 1975a; 1975b; Lüscher et al., 1985). In contrast, histone variants genes encode for longer polyadenylated mRNA containing introns (Brush et al., 1985; Wells and Kedes, 1985). Core histones and histone variants can be defined as

homomorphous or heteromorphous depending on the amino acid sequence difference from the main histone proteins that are expressed during S-phase (West and Bonner, 1980). In general histone proteins that have small sequence differences are classified as homomorphous (i.e. H3.1, H3.2 and H3.3 or H2A.1 and H2A.2), whereas histones that have profound amino acid differences in their sequence that affect large domains are named as heteromorphous (i.e. macroH2A, H2A.Bbd, H2A.X, H2A.Z and CENP-A).

The two prominent core histone variants families belong to H3 and H2A. In contrast, H2B and H4 do not show a high level of sequence variability, this is perhaps due to the role of these two histones in maintaining the histone octamer within the Nucleosome Core Particle (NCP) or promoting higher order chromatin structure (Malik and Henikoff, 2003). There are no H4 variants known thus far, whereas H2B has different tissue-specific variants that are expressed in sperm of vertebrate and invertebrate. In humans there are three H2B variants that have been discovered so far: TH2B, human testis specific H2B (hTSH2B) and H2B family member W testis- specific (H2AFWT) (Churikov et al., 2004; Zalensky et al., 2002)

The incorporation of the histone variants has been associated with several processes involved in cell viability and proliferation such as DNA replication, transcription, recombination and repair. Interestingly, in the last years an abnormal expression of histone variants has been detected in several types of cancer, suggesting a putative role in cancer progression and development (Vardabasso et al., 2013).
The different histone variants belonging to the H2A and H3 families will be described below with particular attention to H2A.Bbd and macroH2A.

2.6.1 H3 histone variants

Five H3 variants have been well characterized thus far. Two of them are the canonical variants that have their peak of expression during S-phase and that are named H3.1 and H3.2, generally referred as H3. The other three replacement variants are expressed outside S-phase. They are H3.3, the centromere-specific variant CenH3 (also named CENP-A in mammals) (Figure 2.4) and the testis-specific histone H3t. Recently, two new H3 histone variants (H3.X and H3.Y) have been identified in primates (Wiedemann et al., 2010).

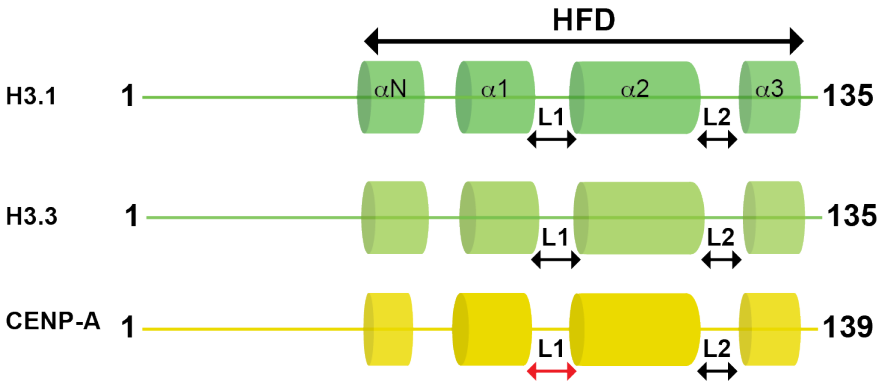

Figure 2.4 Schematic representation of the mammalian H3 histone variants and their functional domains. The H3 histone variants and the canonical histone H3 are represented according to their amino acids length and together with the histone folding domain (HFD) and the histone loops L1 and L2. The sequence divergence from H3.1 is depicted using different colors shades. Significant differences in the L1 loop (L1) are highlighted in red (Figure adapted after (Vardabasso et al., 2013)).

The H3 histone variant H3.3 differs from the canonical H3.1 by only five amino acids, but this small difference is sufficient for a very specific function of H3.3 (Szenker et al., 2011). Through the use of GFP-tagged histones it was shown that H3.3 is incorporated into chromatin outside replication whereas H3.1 incorporation happens only during DNA replication (Ahmad and Henikoff, 2002). Even though the crystal structure of the nucleosome containing H3.3 is very similar to the nucleosome containing H3.1, those with H3.3 show lower stability (Tachiwana et al., 2011). Genome wide analysis of the localization of H3.3 in Drosophila and mammals revealed that H3.3 is enriched at the gene body of actively transcribed genes and at the promoter of active and inactive genes (Szenker et al., 2011). Interestingly, H3.3 incorporation has also been shown at pericentric heterochromatin and telomers in mouse ES cells and Mouse Embryonic Fibroblasts (MEF) (Goldberg et al., 2010; Lewis et al., 2010). The latest discoveries about the H3.3 localization at the silent chromatin and the fact that mutants that lack H3.3 still have a normal gene expression challenged the initial idea that H3.3 enrichment was solely important for gene activation and transcription.

CENP-A histone variant is a rapidly evolving member of the H3 family that replaces H3 at the centromere. Since CENP-A is important for centromere identity and structure, this variant is essential for cell survival. In yeast CENP-A is named as Capping Enzyme Suppressor 4-p (Cse4p) and in *Drosophila* as Centromere Identifier (Cid) (Black and Cleveland, 2011). CENP-A is highly divergent from H3 with only 60% similarity in the HFP domain and a distinct N-terminal tail (Sullivan et al., 1994) (Figure 2.4). The N-terminal sequence varies in

different species in both length and sequence and a histone fold domain that contains the CATD sequence important for targeting CENP-A to the centromere (Black et al., 2004). Another important domain for the CENP-A specific centromere function is the hydrophobic C-terminus (CAC) that mediates the association of CENP-A with a second very important centromere protein CENP-C (Kato et al., 2013). The nucleosomes containing H3 or CENP-A do not show significant differences but CENP-A differs from the canonical H3 in its α-N helix, with at least one helical turn shorter than in H3, and the L1 loop, that has two additional residues (Tachiwana et al., 2011). Remarkably, the loading of CENP-A does not follow the time of its expression. Even though the loading process is still not fully understood it is clear that, whereas CENP-A expression peaks in G2, the loading of this histone variant at centromeres only happens during G1 (Jansen et al., 2007).

2.6.2 H2A histone variants

In the H2A family there are variants for the canonical H2A and replacement variants. The first two variants identified for the canonical H2A are H2A.1 and H2A.2, which differ in one amino acid in position 51 (Franklin and Zweidler, 1977). Several differences between the canonical H2A and its variants are located on the C-terminal tail and the docking domain (Figure 2.5).

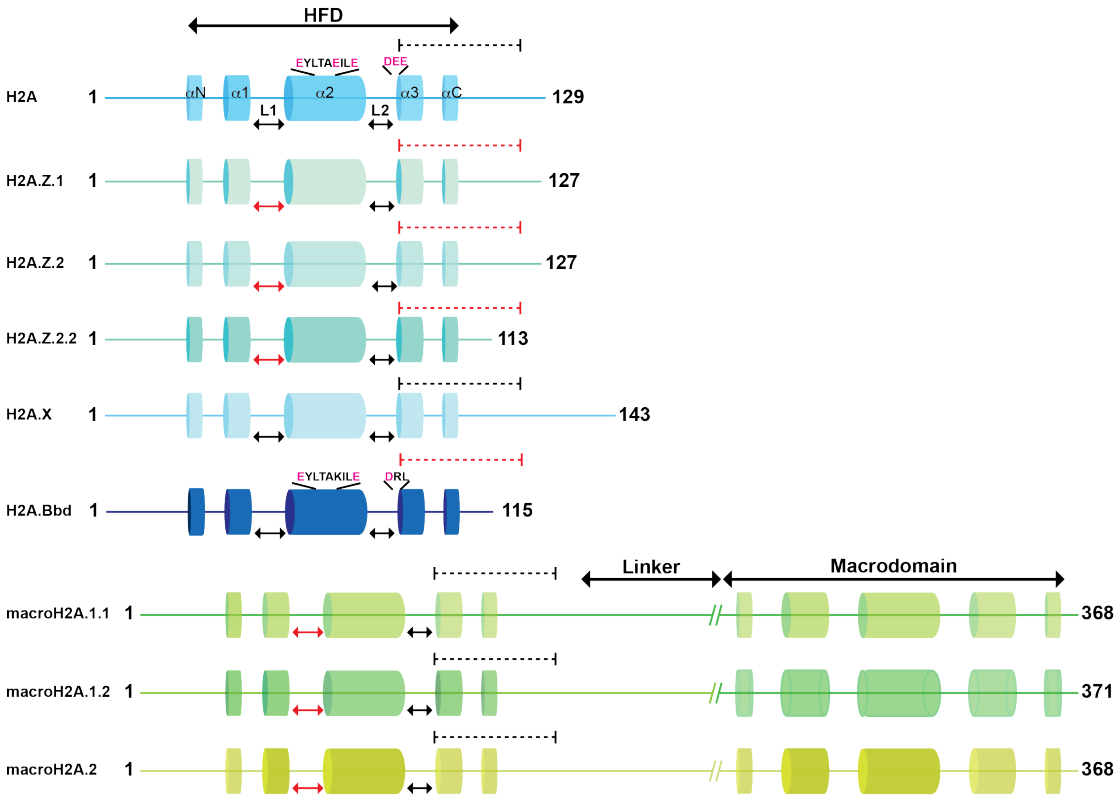

Figure 2.5 Schematic representation of the mammalian H2A family and their functional domains.
The different H2A histone variants and the canonical histone H2A are represented according to their amino acids length and together with the histone folding domain (HFD) and the histone loops L1 and L2. The sequence divergence from the canonical H2A is depicted using different colors shades. Significant differences in the L1 loop (L1) and the docking domain (dashed line) sequences are highlighted in red. The H2A residues that are part of the acidic patch domain are represented in pink (Figure adapted after (Vardabasso et al., 2013)).

Initially, it was shown that the removal of 15 amino acids at the C-terminal tail of the H2A histone decreases the H2A-H2B dimer affinity for the $(H3\text{-}H4)_2$ tetramer, leading to the hypothesis that the carboxyl terminal tail of H2A is important for stabilizing the nucleosome structure (Eickbush et al., 1988). Recently the C-terminal domain of the histone H2A has been shown to interact with the linker histone, and variations in the C-terminal domain could affect internucleosomal interaction and chromatin fibre structure (Vogler et al., 2010). For

this reason, H2A histone variants that carry an altered C-terminal tail, as for example H2A.Bbd, are thought to substantially modify the chromatin fibre structure.

The docking domain is responsible for the interaction between the H2A-H2B dimer and the $(H3\text{-}\ H4)_2$ tetramer and includes three of the seven amino acids that are part of the acidic patch (Luger et al., 1997) (Figure 2.6). As for the C-terminal tail, changes in the docking domain influence the stability of the dimer to tetramer interaction.

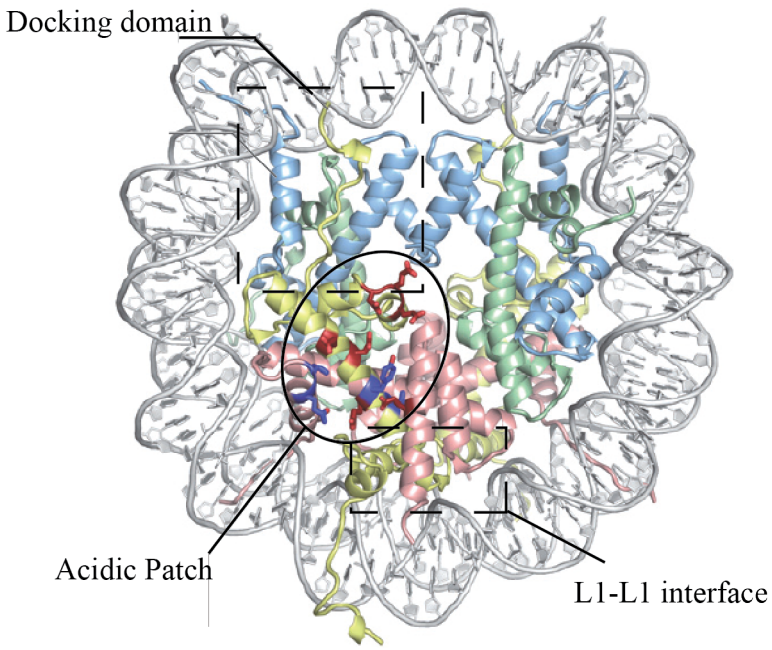

Figure 2.6 H2A relevant domains in the nucleosome structure.
Structural differences between the canonical H2A and its replacement variants are fundamental for the nucleosome stability. The main differences in the H2A family are found in the docking domain, the acidic patch and the L1 loops. All these domains mediate the inter- and intra-nucleosomal interactions. The histones are depicted as follows: H3- light blue; H4- green; H2A-yellow; H2B-red (Figure adapted after (Luger et al., 2012)).

In the nucleosome the L1 loops of the two H2A histones mediate the interaction between the two H2A-H2B dimers (Figure 2.6). This region shows a high diversity among the H2A variants influencing the stability of the dimers within the nucleosome.

H2A.Z is highly conserved among species, but compared to the canonical H2A, H2A.Z shows only 60% sequence identity. In organisms such as *Tetrahymena thermophila*, *Drosophila melanogaster*, *Xenopus laevis* and *Mus musculus,* H2A.Z is essential for the organism survival (Faast et al., 2001; Iouzalen et al., 1996; Liu et al., 1996; van Daal and Elgin, 1992). So far three isoforms have been identified for H2A.Z: H2A.Z.1, H2A.Z.2 and H2A.Z.2.2 (Bönisch et al., 2012; Matsuda et al., 2010) (Figure 2.5).

H2A.Z.1 and H2A.Z.2 differ by only three amino acids at the protein level, but they are encoded by unique nucleotide sequences (Dryhurst et al., 2009). The discovery of a third isoform increased the complexity of the H2A.Z biology. The third isoform, H2A.2.2, has a shorter and specific C-terminus and is loosely associated with chromatin. Nucleosome containing H2A.2.2 are the least stable nucleosome described so far (Bönisch et al., 2012). It is important to specify that what is normally referred to as H2A.Z is the isoform H2A.Z.1. Despite the significant difference between H2A.Z and the canonical H2A, the nucleosome containing H2A.Z is structurally very similar to the nucleosome containing H2A (Suto et al., 2000). The major differences in the nucleosomes containing either H2A or H2A.Z are found in the L1 loop, which is important for the H2A-H2B dimer interaction. Further differences are present in the C-terminal docking domain, which has less than 40% amino acid identity with the canonical H2A. This region is important for the interaction with H3/H4 and it is a putative platform for nucleosome remodeling activities. Modifications in the C-terminal docking domain could have effects on the interaction between nucleosomes, or between nucleosomes and the linker histone, thus altering the higher order chromatin structure (Suto et al., 2000; Thakar et al., 2009). There are conflicting data on the influence of H2A.Z on nucleosome stability. On one hand, several studies propose a destabilization of the nucleosome containing- H2A.Z (Abbott et al., 2001; Suto et al., 2000); on the other hand, other experiments support a rather stable conformation of the nucleosome-containing H2A.Z (Park et al., 2004; Thambirajah et al., 2006). However, these varied outcomes could be the result of technical differences among the performed experiments (Zlatanova and Thakar, 2008).

H2A.Z is expressed during the all cell cycle and incorporated in a replication- independent manner. Even though the incorporation of H2A.Z influences several cellular processes, the best characterized function of H2A.Z is its role in transcriptional regulation. H2A.Z is enriched at gene promoters in yeast and higher eukaryotes (Guillemette et al., 2005; Raisner et al., 2005). Together with its localization at promoter sites, H2A.Z is also found at regulatory regions such as insulator and enhancers (Zlatanova and Thakar, 2008). Notably, it was shown that H2A.Z has active and repressive functions (Marques et al., 2010). The different effects on transcription regulation may suggest that it is not the presence of H2A.Z itself that influences the final readout of gene expression, but rather the combination of different factors such as the presence of specific H2A.Z binding factors and the combination with other histone variants and their post-transcriptional modifications (PTMs).

Of particular interest is the histone variant H2A.X, which is fundamental in preserving genome integrity. H2A.X is highly similar to H2A in its HFD domain, however it is characterized by the presence of a long C-terminal domain containing a serine-glutamine (SQ) motif followed by an acidic and a hydrophobic residue [SQ(E/D)ϕ] (Bönisch and Hake, 2012). H2A.X plays an important role in the DNA damage response (DDR). After DNA double strand break (DSB), H2A.X gets phosphorylated and it can be visualized at specific foci, which in mammals cover several Mb on each DSB side (Rogakou E. P. 1999). In Humans the presence of the serine S139 in the SQ(E/D)ϕ motif is fundamental for the H2A.X specific role in DNA repair. Despite the importance of H2A.X in genome stability, the effects of H2A.X on nucleosome stability are still largely controversial with results pointing in diverging directions (Bönisch and Hake, 2012).

2.6.2.1 H2A.Bbd

H2A.Bbd (H2A. Barr-body-deficient) was first described by Chadwick and Willard as a new histone variant excluded from the inactive X that localizes to H4 hyperacetylated regions (Chadwick and Willard, 2001a). Recently, a new nomenclature was suggested for several histone variants based on historical usage and phylogenetic relationship among histones. According to this new nomenclature, H2A.Bbd should now be named as H2A.B (Talbert et al., 2012). However, in this thesis the prior nomenclature and the name H2A.Bbd will be used. H2A.Bbd is present only in mammals and, among all histones, is the most quickly evolving histone known as of yet (Eirín-López et al., 2008).

Like other histone variants, H2A.Bbd is encoded by a polyadenylated mRNA. Sequence alignment showed that H2A.Bbd is only 50% identical to the canonical H2A. Interestingly, none of the residues that are post-translationally modified in H2A are conserved in H2A.Bbd. Moreover, H2A.Bbd lacks the C-terminal tail and part of the docking domain (Bao et al., 2004; Chadwick and Willard, 2001a), which results in a slightly shorter and less basic protein with a molecular weight of 12.7 kD (Figure 2.5). H2A.Bbd is also characterized by the lack of the acidic patch, which is important for inter-nucleosomal interactions (Zhou et al., 2007).
For this reason H2A.Bbd is also called H2A.Lack of acidic patch 1-4 (H2A.Lap1-4) in mouse (Soboleva et al., 2011).

H2A.Bbd cofractionates and copurifies with core histones after chromatin fractionation and sedimentation (Chadwick and Willard, 2001a). This observation, together with electrophoretic mobility shift assay (EMSA) experiments, confirmed that H2A.Bbd can be incorporated into nucleosomes and efficiently replaces canonical H2A (Angelov et al., 2004). The nucleosomes containing H2A.Bbd are influenced by different perturbations and alterations in the DNase I footprinting pattern (Angelov et al., 2004). Nucleosomes containing H2A.Bbd are not able to refold into histone octamers under high salts conditions. This suggests that the interaction between H2A.Bbd-H2B dimers and $(H3\text{-}H4)_2$ tetramers is impaired resulting in a reduced nucleosome stability (Bao et al., 2004; Eirín-López et al., 2008).
Initially, it was shown that H2A.Bbd containing nucleosomes organize only 118 bp in contrast to the canonical 147 bp (Bao et al., 2004). More recent studies however, in which micrococcal nuclease digestion was combined with atomic force microscopy and electron cryo-microscopy, revealed that the H2A.Bbd nucleosome organizes 130 bp of DNA compared to the canonical NCPs (Doyen et al., 2006).

H2A.Bbd lacks the C-terminal domain and the docking domain. To understand which domain is responsible for the particular nucleosomal conformation of the H2A.Bbd containing NCPs, H2A chimeras containing the H2A.Bbd domains were generated. These experiments showed that the lack of the docking domain and not the C-terminal tail is the reason for the relaxed DNA organization in the nucleosome containing H2A.Bbd. These experiments also pointed to a role of the docking domain in organizing the last 14 base pairs in the nucleosome containing H2A.Bbd (Bao et al., 2004). These results are consistent with the role of the docking domain in stabilizing the nucleosome octamer. Interestingly, when a H2A.Bbd mutant containing the docking domain and the C-terminal tail of the canonical H2A was generated some of the properties of the canonical H2A-NCP were re-established, such as its stability and the length of the organized DNA (Doyen et al., 2006).

Another important contribution to the instability of the nucleosome containing H2A.Bbd is given by the absence of the acidic patch. It was recently shown that the alteration of the acidic patch in the nucleosome-containing H2A.Bbd reduces the magnesium-dependent internucleosomal interactions that guide folding of nucleosomal arrays (Zhou et al., 2007).
In summary, the unfolded state of the chromatin fibres containing H2A.Bbd is very likely a consequence of the differences of both the docking domain and the acidic patch.

In addition to several structural differences, the nucleosome-containing H2A.Bbd has also different functional properties compared to the canonical NCP. Indeed, it was shown that the H2A.Bbd-nucleosome cannot be mobilized by the SWI/SNF and the ACF remodeling machinery, even though the presence of H2A.Bbd does not affect the binding of the remodelers to chromatin (Angelov et al., 2004).

Mammalian H2A.Bbd is expressed at high levels in testis and to a lower degree in the brain. In mice H2A.Bbd is expressed in specific stages of spermatogenesis when chromatin undergoes massive remodeling caused by the histones to protamines exchange (Chadwick and Willard, 2001a; Ishibashi et al., 2010). Recently, it was shown that H2A.Bbd together with the histone variant H2A.Z contribute to create a specific chromatin environment at promoters of active genes during spermatogenesis where H2A.Bbd occupies position -1 and H2A.Z position -2 relative to the transcription start site (TSS) (Soboleva et al., 2011).

2.6.2.2 MacroH2A

MacroH2A is a H2A variant characterized by the presence of a large macro domain that constitutes a binding module for nicotinamide adenine dinucleotide (NAD) metabolites (Kustatscher et al., 2005) (Figure 2.5). The structure of macroH2A consists in an N-terminal histone region connected to an H1-like linker region to a non-histone macro domain. This results in a large histone protein of about 40 kD (Figure 2.5). There are two genes encoding for macroH2A: macroH2A.1/H2AFY and macroH2A.2/H2AFY2. MacroH2A.1 encodes for two alternatively spliced forms, macroH2A.1.1 and macroH2A.1.2. MacroH2A.2 is 68% identical to macroH2A.1. Both, macroH2A.1 and macroH2A.2 localize to the inactive X chromosome (Chadwick and Willard, 2001b). MacroH2A preferentially forms heterotypic nucleosomes where both canonical H2A and macroH2A, are present (Chakravarthy et al., 2005a).

The histone variant macroH2A is enriched on the inactive X chromosome in mammals where it contributes to the maintenance of the repressed state and facilitates dosage compensation in mammals (Costanzi and Pehrson, 1998). Several studies have reported the enrichment of macroH2A on the inactive X chromosome. This particular localization of macroH2A was challenged by using GFP tagged macroH2A and fluorescent recovery after photobleaching (FRAP). This study suggested that the enrichment of macroH2A on the highly compacted

inactive X may reflect the higher chromatin concentration and not a real enrichment of the histone variant (Perche et al., 2000). Subsequently, Chromatin Immunoprecipitation (ChIP) based approaches revealed that the inactive X chromosome in liver cells contains 1.5 to 2 times more MacroH2A in females compared to males (Mietton et al., 2009). Nevertheless, from the first study that showed the localization of macroH2A on the inactive X chromosome, this histone variant has always been associated to transcriptional repression. Several *in vitro* studies of nucleosomes containing macroH2A have given insight on how this variant could contribute to chromatin repression. Indeed, even if the structure of the macroH2A NCPs is similar to the canonical NCPs, there is a four amino acid difference in the LI loop responsible for the interaction of H2A-H2B dimers, which may affect the structure of the macroH2A containing NCPs. In solution the macroH2A NCPs were found to be more stable and less prone to H2A-H2B dimer exchange by the histone chaperone NAP1 (Abbott et al., 2001; Chakravarthy et al., 2005a). Interestingly, it was shown that by introducing the macroH2A L1 loop into the canonical H2A, the H2A chimera was targeted to the inactive X (Nusinow et al., 2007b).

Due to the particular localization of macroH2A, many studies focused on the role of this variant in the X inactivation process (Costanzi and Pehrson, 1998; Costanzi et al., 2000; Hernández-Muñoz et al., 2005; Perche et al., 2000). It is important to note that macroH2A is also present in organisms that do not undergo X inactivation and that macroH2A level expression levels are similar in male and female (Buschbeck and Di Croce, 2010; Rasmussen, 1999). These two fascinating features of macroH2A prompted new investigations in studying its potential role outside X inactivation. In NT2/D1 cells and by using promoter array-based analysis of immunoprecipitated chromatin material it was shown that macroH2A is enriched on promoters of key developmental genes (Buschbeck et al., 2009). However, it seems that macroH2A has different roles in polypotent cells, such as stem and pluripotent cells, when compared to terminally differentiated cells (Buschbeck and Di Croce, 2010). In pluripotent cells, macroH2A is localized to key developmental genes, whereas in liver cells, macroH2A does not show any particular enrichment on specific developmental genes (Changolkar et al., 2010). Moreover, in Namalwa cells, a human B-cell line, macroH2A.1 knockdown caused a dramatic derepression of the IL-8 gene that is normally not expressed in these cells (Agelopoulos and Thanos, 2006).

It is evident that there are several mechanisms that modulate the macroH2A function in different tissues. One type of regulation may be related to different interaction partners of macroH2A. Indeed, many proteins have been identified so far that could specifically interact with macroH2A. Some of those factors show a preferential binding for either macroH2A.1.1 or macroH2A.1.2. An interesting example is the poly-ADP ribose polymerase 1 (PARP1) (Nusinow et al., 2007a). Although the three macroH2A variants are able to inhibit PARP1 via their non-histone region (NHR), only macroH2A.1.1 is able to bind polymeric and monomeric ADP-ribose, O-acetyl-ADP ribose (Kustatscher et al., 2005; Timinszky et al., 2009).

In COS cells, the NHR of macroH2A.1.2 was shown to bind the histone deacetylases 1 and 2 (HDAC1/2) (Chakravarthy et al., 2005b). In a yeast two-hybrid approach macroH2A.1.2 was found to interact with the speckle-type POZ domain protein (SPOP) and later this interaction was found to be important for the deposition of macroH2A.1.2 on the inactive X chromosome (Hernández-Muñoz et al., 2005; Takahashi et al., 2002). Recently, macroH2A has also been shown to localize at Senescence-Associated Heterochromatic Foci (SAHFs), revealing new questions about the role of this variants in chromatin (Zhang et al., 2005).

2.7 Histone variants in cancer

In the last years it became clear that, together with their important role in shaping chromatin structure and gene expression, histone variants play a key role in cancer progression and development. Abnormal expression of different histone variants has been found in several types of cancer. Due to their role in the regulation of key cellular and developmental processes, the deregulation of histone variants expression has been linked to tumor initiation and progression. In some cases, histone variants expression levels correlate with tumor malignancy thus promoting the use of a particular histone variant as a prognostic indicator in cancer (Hua et al., 2008; Sporn and Jung, 2012).

In the H2A family, three variants, H2A.Z, H2A.X and macroH2A, have been identified in playing a role in cancer development. The first evidence comes from microarray studies that detected abnormal expression level of H2A.Z in colorectal, lung, breast and bladder cancers (Dunican et al., 2002; Rhodes et al., 2004; Zucchi et al., 2004). Increased H2A.Z expression is also detected in metastatic melanoma cell lines. Therefore, H2A.Z has been classified as an oncogene but the way it acts to promote cancer progression is still largely unknown in several cancer types. The simple explanation is that H2A.Z modulates gene expression or regulatory

and non-coding regions of genes that are important for cancer development. However, as discussed above, H2A.Z participates in several other cellular processes such as DSBs repair, telomeres integrity, genome stability and chromosome segregation (Greaves et al., 2007; Shia et al., 2006; Xu et al., 2012).

Due to its important role in Double Strand Breaks (DSBs) repair, H2A.X is considered a key factor in tumor progression. In this regards it is surprising that H2A.X null mice are not cancer prone, but they only develop an increased number of T and B lymphomas and solid tumors in the context of p53 deficiency (Bassing et al., 2003; Celeste et al., 2003). Several studies promoted the idea of H2A.X as a 'genome caretaker' and tumor suppressor in defined genetic context (Vardabasso et al., 2013).
As for H2A.X, macroH2A is also considered a tumor suppressor. Its expression is reduced in several types of cancer such as colon, lung and melanoma tumors (Kapoor et al., 2010; Sporn et al., 2009; Sporn and Jung, 2012). In general the down-regulation of macroH2A.1.1 and macroH2A.2 is negatively correlated to cell proliferation (Sporn et al., 2009). As macroH2A is a marker for cellular senescence, which is considered a barrier against tumor progression, it is likely that a low expression of macroH2A allows the cells to bypass senescence thus promoting cancer development. Interestingly, macroH2A is enriched in rodent lung adenomas but it is absent or down regulated in lung carcinomas. Indeed in adenomas there is a high number of senescent cells whereas in carcinomas cells overcome senescence and show a high proliferation rate (Sporn et al., 2009).

Apart from variants of H2A, the H3 variants CENP-A and H3.3 have also been found to play a role in cancer development. There is much evidence that CENP-A is overexpressed in several cancer types. Some examples are the colorectal cancer, lung adenocarcinoma, breast cancer, hepatocellular carcinoma (HCC) and invasive testicular germ cell tumors (Vardabasso et al., 2013). As mentioned above CENP-A is essential for the centromere identity. Thus it is not surprising that the mis-regulation of its expression could promote severe mitotic defects in the cells. Interestingly, it was recently shown that the overexpression of CENP-A in hepatocellular carcinoma (HCC) promotes cell growth and reduce apoptosis by deregulating genes involved in apoptosis and cell proliferation (Li et al., 2011).
H3.3 involvement in cancer development was already reported years ago (Graber et al., 1996). However, H3.3 represents the first case where a mutation in the coding sequence was detected. Two recent studies reported a missense mutation of H3.3 in pediatric Glioblastoma

Multiforme (GBM) and in Diffuse Intrinsic Pontine Gliomas (DIPGs) (Schwartzentruber et al., 2012; Wu et al., 2012). In 31% of GBMs cases, mutations in lysine 27 (K27) or in G34 were reported.

Together with the direct alteration of the histone variants expression or their mutation, in many cancer types there are also changes in their chaperones expression levels or remodeling complexes interacting with specific histone variants (Heaphy et al., 2011; Schwartzentruber et al., 2012; Vardabasso et al., 2013; Wu et al., 2012).

Despite the abundance of data generated in the last years that confirm the involvement of the histone variants in cancer progression, many questions about how they contribute to the disease progression still need to be answered.

3 Aim of the thesis

The core unit of the chromatin, the nucleosome, plays a central role in the regulation of chromatin structure and gene expression regulation. Together with the histone modifications, the canonical histones and their variants are the main regulators of the nucleosome structure and accessibility.

The main aim of this project was to analyze the protein composition of chromatin domains containing the histone variants H2A.Bbd and macroH2A.
H2A.Bbd and macroH2A have been extensively studied for their influence on the nucleosome structure and they opposing nuclear localization. Indeed H2A.Bbd is enriched on H4 hyperacetylated foci and excluded from the Barr body, whereas macroH2A localizes to heterochromatic regions and plays a role in the X inactivation process (Chadwick and Willard, 2001a; Costanzi and Pehrson, 1998).

For this aim the histone variant H2A.Bbd and macroH2A.1.2 were tagged with GFP and expressed in Mouse Embryonic Fibroblasts (MEF). Mass spectrometry analysis was then performed on chromatin fragments containing the GFP-tagged histone variants. Interesting differences were found that reflect the opposing nuclear localization of these two variants confirming that mass spectrometry and label free quantification are valid and powerful tools to study complex samples like chromatin domains.

Interestingly based on the proteomic results it was possible to reveal new feature for the histone variant H2A.Bbd. The mass spectrometry data showed an interesting link between H2A.Bbd and factors involved in DNA replication and repair. Based on this new finding, the mass spectrometry data were validated by using different microscopy techniques that confirmed the presence of GFP-H2A.Bbd at sites of DNA replication and DNA damage.

It was further investigated whether the expression of GFP-H2A.Bbd in MEF cells have an effect on cell cycle or DNA sensitivity and it was found that the MEF cells expressing GFP-H2A.Bbd have shorter S-phase and higher sensitivity to DNA damage.

Due to its specific tissue specific expression in post-mitotic round spermatids it is possible to exclude that H2A.Bbd could play a role in somatic cells DNA replication. However this study

demonstrated that H2A.Bbd is abnormally expressed in Hodgkin's Lymphoma cells (HL) where the level of expression inversely correlates with the doubling time of different HL cell line.

4 Results

4.1 Characterization of mouse embryonic fibroblasts (MEFs) expressing GFP-tagged histone variants

Histone variants H2A.Bbd and macroH2A.1.2 have been found to localize at different chromatin loci. While H2A.Bbd localizes to regions enriched in acetylated H4, macroH2A.1.2 is mostly co-localizing with H3K27me3 (Chadwick and Willard, 2001a; Costanzi and Pehrson, 1998). To better understand the functional properties of H2A.Bbd and macroH2A.1.2 containing chromatin, mass spectrometry analysis was performed on chromatin domains containing GFP-tagged histone H2A.Bbd and macroH2A.1.2. To do this, female MEF cells were used that stably express GFP-tagged version of H2A.Bbd and macroH2A.1.2. The MEF cells expressing GFP- tagged H2A.Bbd and H2A were established by Clemens Boenisch (A.G. Hake), whereas the GFP-macroH2A.1.2 expressing MEFs were generated by me using an expression plasmid provided by Dimitri Thanos. In all experiments, the canonical histone H2A was also included as a control to discriminate specific protein factors that bind H2A.Bbd and macroH2A.1.2 containing chromatin from generic chromatin binding factors.

To obtain a homogenous population of cells expressing the GFP-tagged histone variants, the transfected MEF cells expressing the GFP-tagged histone variants were first selected with the G418 antibiotic and then by isolating single cells clones using a MoFlo sorter. Several clones were obtained and analyzed by FACS to detect the expression levels of the GFP-tagged histone variants (data not shown).

Based on similar expression levels detected by fluorescence-activated cell sorting (FACS) analysis, two clones were chosen expressing GFP-H2A.Bbd and GFP-H2A, respectively. Unfortunately, the expression of GFP-macroH2A.1.2 ceased upon extended passaging of the first three single cell clones selected. Therefore, a population derived from six sorted cells was used for all experiments instead. Within this population, GFP-macroH2A.1.2 expression was maintained in most of the cells even after two month of culturing. The expression of the GFP-tagged histone variants was measured by FACS analysis and compared to non-transfected MEF cells (Figure 4.1 A).

Furthermore, GFP-tagged histone variants expression levels were also determined by Western blotting using antibodies directed against the GFP-tag (Figure 4.1 B). MEF cells transfected

with GFP-H2A.Bbd and GFP-H2A show similar level of expression, whereas GFP-macroH2A.1.2 is expressed at lower levels.

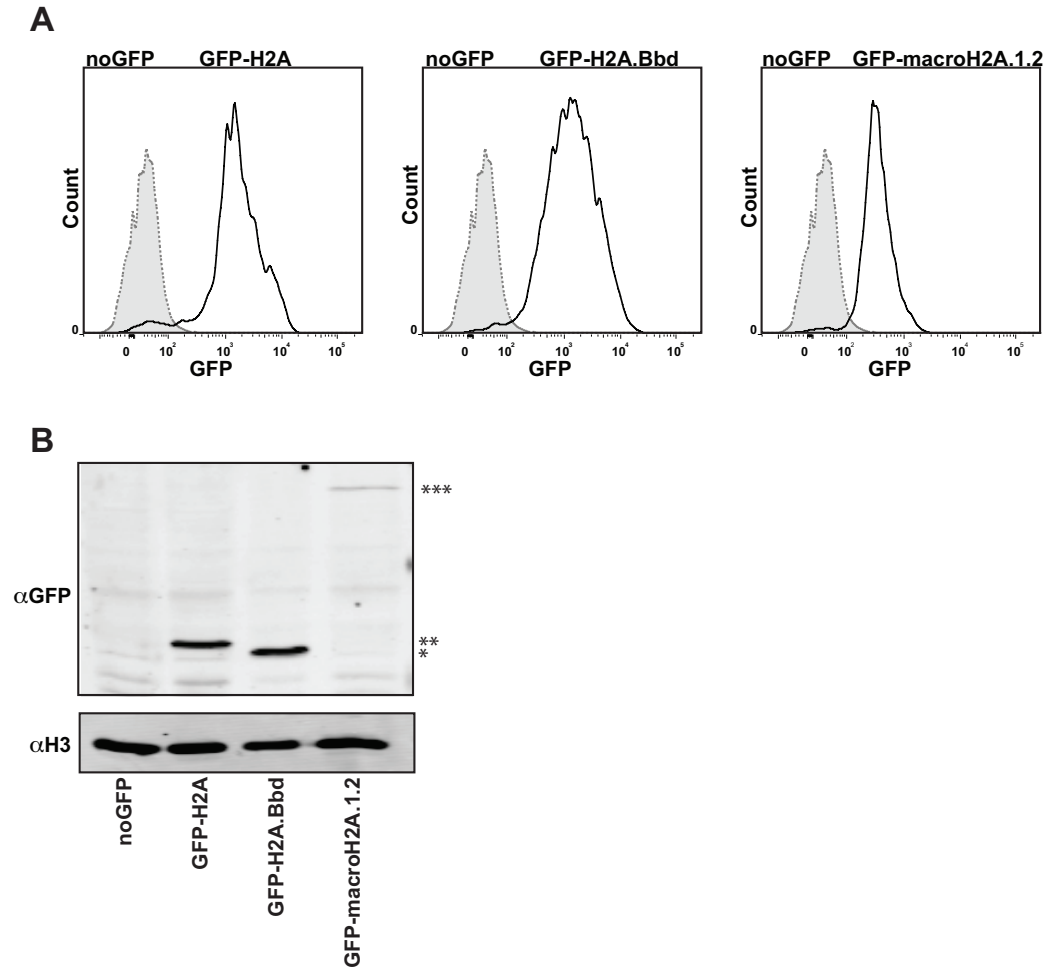

Figure 4.1 Female MEFs stably express the GFP-tagged histone variants.
(A) FACS profile of female MEF cell populations derived from sorted cell clones that express H2A.Bbd, macroH2A.1.2 and H2A tagged with GFP (solid line) in comparison to non-transfected control female MEF cells (dotted and filled histogram). **(B)** Western Blot analysis of GFP-tagged histone variants expression levels with antibody against GFP. Detection of histone H3 served as loading control. The bands corresponding to the GFP-tagged histones are indicated: (*) GFP-H2A.Bbd, (**) GFP-H2A and (***) GFP-macroH2A.1.2.

In order to verify the localization pattern of the GFP-tagged histone variants H2A.Bbd and macroH2A.1.2, MEF cell lines were stained with an anti-H3K27me3 antibody, which marks the inactive X-chromosome forming the Barr body. Consistent with previous reports (Costanzi and Pehrson, 1998), GFP-tagged macroH2A.1.2 was slightly enriched on the Barr body, whereas GFP-H2A.Bbd was excluded (Figure 4.2). This confirmed that in MEF cells the GFP-tagged histone variants indeed mark different chromatin domains.

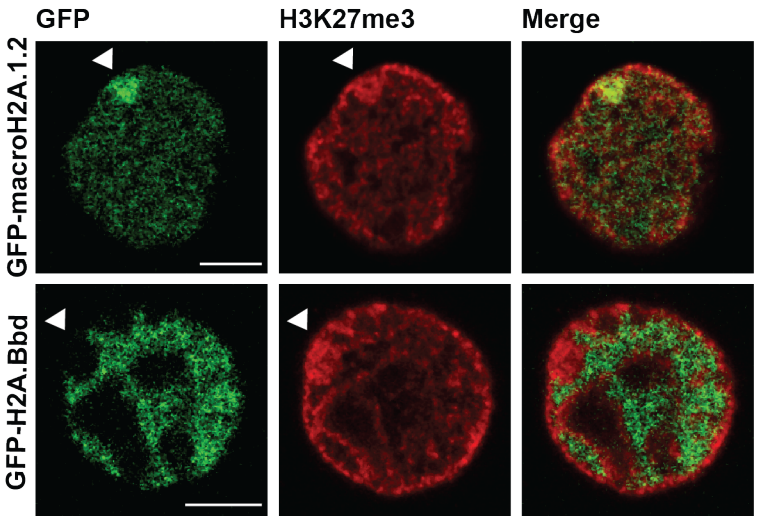

Figure 4.2 GFP-tagged histone H2A.Bbd and macroH2A.1.2 mark different chromatin domains.
Nuclear localization of the histone variants GFP-H2A.Bbd and GFP-macroH2A.1.2 in female MEFs stained with H3K27me3 (red). Arrows indicate the Barr body. Scale bar: 5 µm.

4.2 Purification and accessibility of chromatin containing GFP-tagged histone variants

Chromatin is organized in large domains in which not all the nucleosomes contain the same histone isoforms (Loyola et al., 2006). Therefore, to be able to study entire chromatin domains that contain the GFP-tagged histone variants H2A.Bbd and macroH2A.1.2, chromatin was digested with micrococcal nuclease (MNase) to obtain fragments of different lengths.

After chromatin digestion and centrifugation of the nuclear matrix pellet, the first supernatant was collected (S1) at 150 mM NaCl. The chromatin fragments retained in the nuclear pellet were then extracted using 300 mM NaCl to release the remaining chromatin (second supernatant (S2) (Figure 4.3 A). The chromatin fragments obtained in S1 and S2 differ in size with mononucleosomes preferentially released in the S1 and larger chromatin fragments in the S2 (Figure 4.3 B).

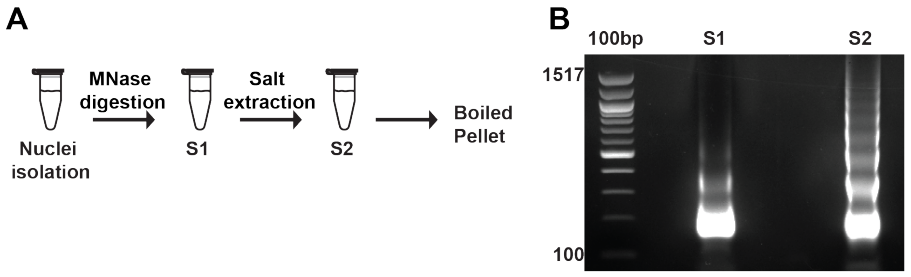

Figure 4.3 Chromatin from the MEF cells expressing the GFP-tagged histone variants is released after MNase digestion.
(A) Chromatin purification protocol. After nuclei isolation the chromatin is digested with MNase. The nuclei are centrifuged and the first supernatant (S1) is collected. The remaining chromatin is extracted with 300mM NaCl and the second supernatant (S2) containing chromatin is collected. **(B)** DNA analysis of the first and second supernatants containing chromatin (S1 and S2, respectively).

The release of the GFP-tagged histone variants was followed by western blotting. The S1, S2 and the boiled nuclear pellet were analyzed for the presence of the tagged histone variants with or without MNase digestion (Figure 4.4). The GFP-tagged histone variants are mainly released in the S1 and S2. Only GFP-H2A.Bbd is still visible in the boiled nuclear pellet. This suggested that the physiochemical properties of the chromatin fibres containing the GFP-histone variants are similar with some differences for GFP-H2A.Bbd. In addition, no significant pools of free histone variants exist in isolated nuclei in the absence of MNase where the GFP-tagged histone variants are not released in the S1 and S2 but are retained in the pellet (Figure 4.4).

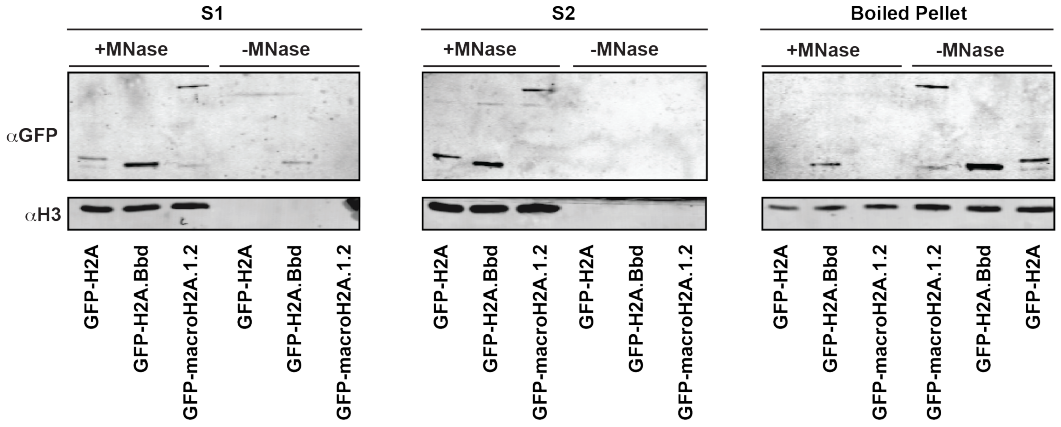

Figure 4.4 The GFP-tagged histone variants are only released after MNase digestion.
Western Blot analysis of the GFP-tagged histone variants release compared to H3 in presence or absence of chromatin MNase digestion.

Chromatin is organized in domains where nucleosomes containing the histone variants are interspersed with nucleosomes containing canonical histones (Loyola et al., 2006). Furthermore, different chromatin regions differ for their solubility and this could also reflect several biological differences as Heinekoff proposed in 2008 (Henikoff et al., 2009). Therefore, the S1 and the S2 fractions were pooled for all further experiments.

It was shown that the incorporation of the histone variants H2A.Bbd and macroH2A into chromatin causes specific modifications on the nucleosome structure and stability (Bönisch and Hake, 2012). In particular the nucleosome containing H2A.Bbd was shown to be rather unstable compared to the canonical H2A-containing nucleosomes. The nucleosomes containing H2A.Bbd have been proven to organize only 130 bp of DNA (Doyen et al., 2006) . For this reason it was asked whether the ectopic expression of H2A.Bbd could lead to a more accessible chromatin structure. To answer this question the chromatin containing the GFP-tagged histone variants was digested with different MNase concentration. The digested DNA was precipitated and separated by agarose gel electrophoresis (Figure 4.5A). The subsequent calculation of the mononucleosomes percentage revealed no major differences in bulk chromatin accessibility in the MEF cells expressing the three different GFP-tagged histone variants compared to the non- transfected cells (Figure 4.5 B).

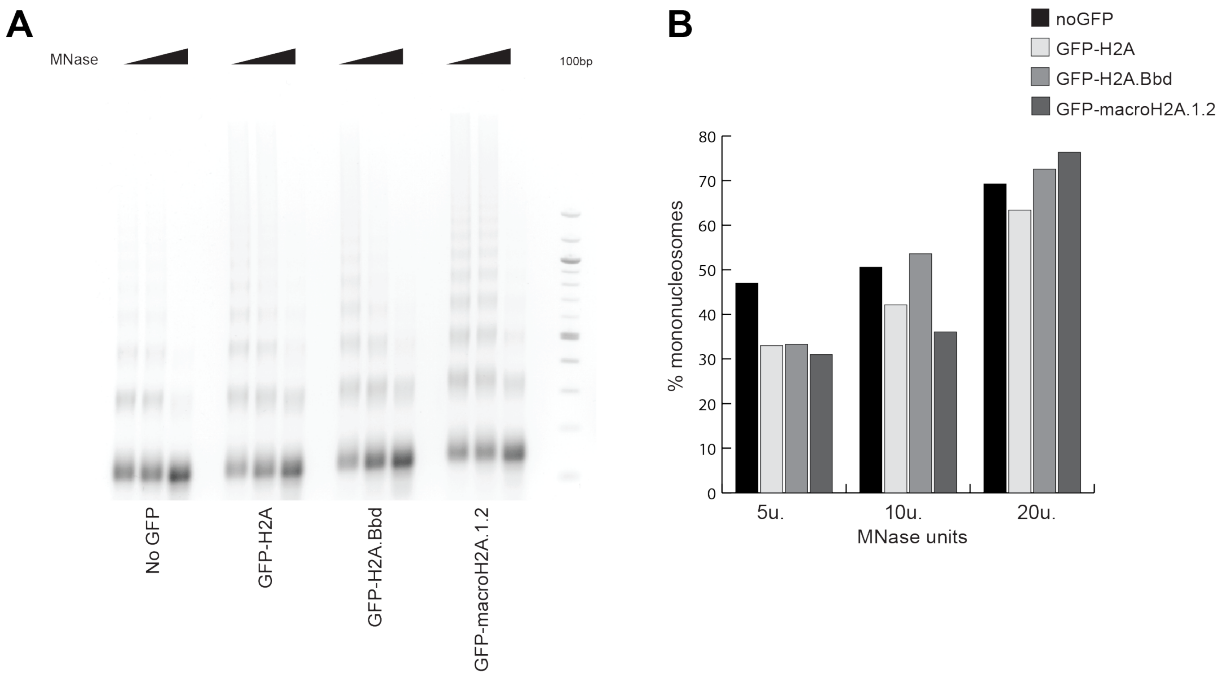

Figure 4.5 GFP-tagged histones expression does not alter bulk chromatin structure in MEF cells.
(A) DNA gel of the chromatin digestion pattern after MNase treatment in cells expressing GFP-H2A, GFP-H2A.Bbd and GFP-macroH2A.1.2. **(B)** Percentage of mononucleosomes after digestion with different concentration of MNase in MEF cells expressing the GFP-tagged histone variant H2A, H2A.Bbd and macroH2A.1.2.

4.3 Proteomic analysis of chromatin containing GFP tagged H2A.Bbd, macroH2A.1.2 and H2A

Using GFP- affinity purification, chromatin-containing the GFP-tagged histone variants was specifically isolated. The representative Coomassie gel shows the purified fractions and the input chromatin (Figure 4.6). The bands corresponding to the three tagged histone variants are clearly visible in the gel. Together with the tagged histone variants several other non-histone proteins bands are detected. In addition, the exclusive presence of chromatin bound histones and the presence of DNA in the GFP-tagged histone purified fraction compared to the non-transfected cells (no-GFP) confirmed that the samples are composed of chromatin without the presence of soluble histones fraction (Figure 4.4 and Figure 4.6 B).

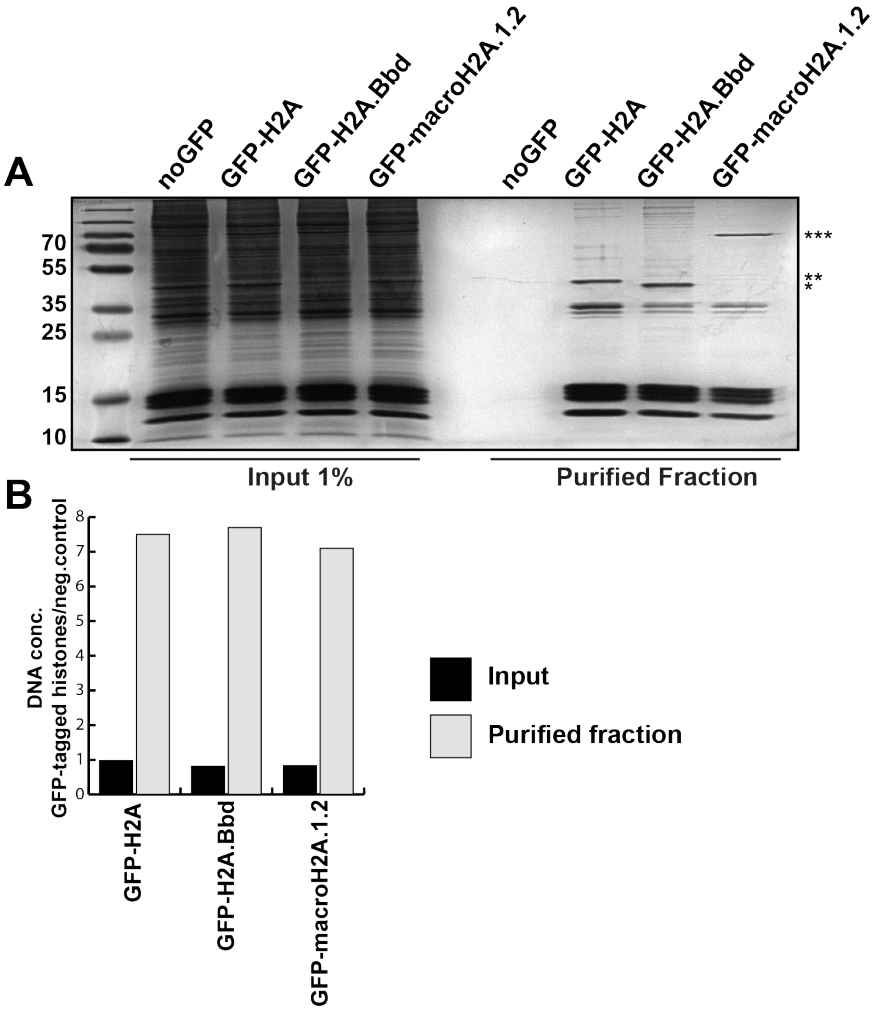

Figure 4.6 The GFP purified fraction is enriched in chromatin containing the GFP-tagged histone variants. **(A)** Coomassie blue staining of SDS-PAGE gel of chromatin fibres purified by GFP-nanotrap beads. Chromatin input and purified fractions are shown. The bands corresponding to the GFP-tagged histones are indicated: (*) GFP-H2A.Bbd, (**) GFP-H2A and (***) GFP-macroH2A.1.2. **(B)** Isolated DNA after chromatin fibres purification with the GFP-nanotrap beads. The ratio between the DNA concentration in the chromatin containing GFP-tagged histones and the non-transfected cells is shown before and after the affinity purification (black and grey bars respectively).

As expected, the four canonical histones are also present in the purified fraction (Figure 4.6 A). This is consistent with previous reports that suggested that chromatin domains are composed by canonical nucleosomes and nucleosome containing the histone variants (Loyola et al., 2006).

To identify non-histone proteins present in the sample, in-gel trypsin digestion was performed using the entire lane cut into 8 pieces. The eluted peptides were analyzed by Mass spectrometry. A total number of 1428 proteins were identified using the Andromeda algorithm of the MaxQuant protein analysis package (Cox et al., 2011). The overlap among replicates for every GFP-histone pull down was calculated based on protein identification. Using these criteria more than the 50% of the identified proteins are present in at least two replicates. The overlap varies among the GFP-tagged histone samples: 53.4% for GFP-H2A.Bbd, 51.6% for GFP-H2A and 53.1% for GFP-macroH2A.1.2. (Figure 4.7).

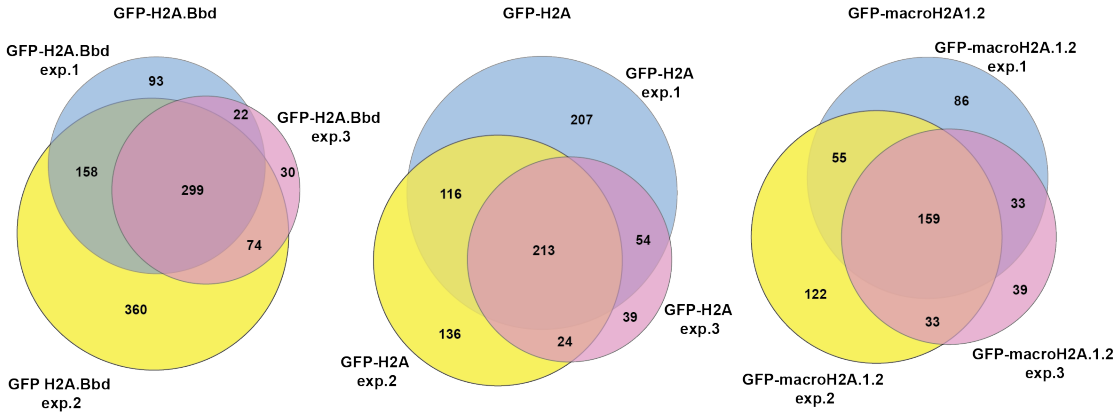

Figure 4.7 Identified protein overlap among the replicates.
Protein factors identified after the GFP-H2A variants chromatin purification using the GFP nanotrap beads are displayed in Venn diagrams showing the identified proteins overlap in three independent biological replicates (e.g. H2A.Bbd exp.#1, #2 and #3).

In order to quantify the amount of every single protein in the pull down sample and compare the chromatin fragments that contain different histone variants, the Intensity Based Absolute Quantification (iBAQ) values were used (Schwanhäusser et al., 2011). The iBAQ values are calculated using the sum of all the tryptic peptides intensities divided by the number of theoretically observable peptides of a specific protein and they have been shown to correlate with the absolute protein amount in the samples (Schwanhäusser et al., 2011). Due to the different expression of GFP-tagged H2A variants, all replicates were normalized using largely unmodified peptides of histone H4. A protein was considered to be chromatin-bound when

its average log2 transformed iBAQ value showed a difference of at least 3.5 between the GFP-H2A variant sample and the background control (noGFP). The samples were further filtered excluding all the proteins that were identified only in one experiment. The enrichment factors were used to evaluate how the proteome of the chromatin containing GFP-tagged histone variants varies among the pull downs. This was obtained by calculating the log2 enrichment of the ratio between the enrichment factors of different GFP-tagged histone variants pull downs (Figure 4.8 A). The values obtained for every protein were displayed on a scatter plot (Figure 4.8 B).

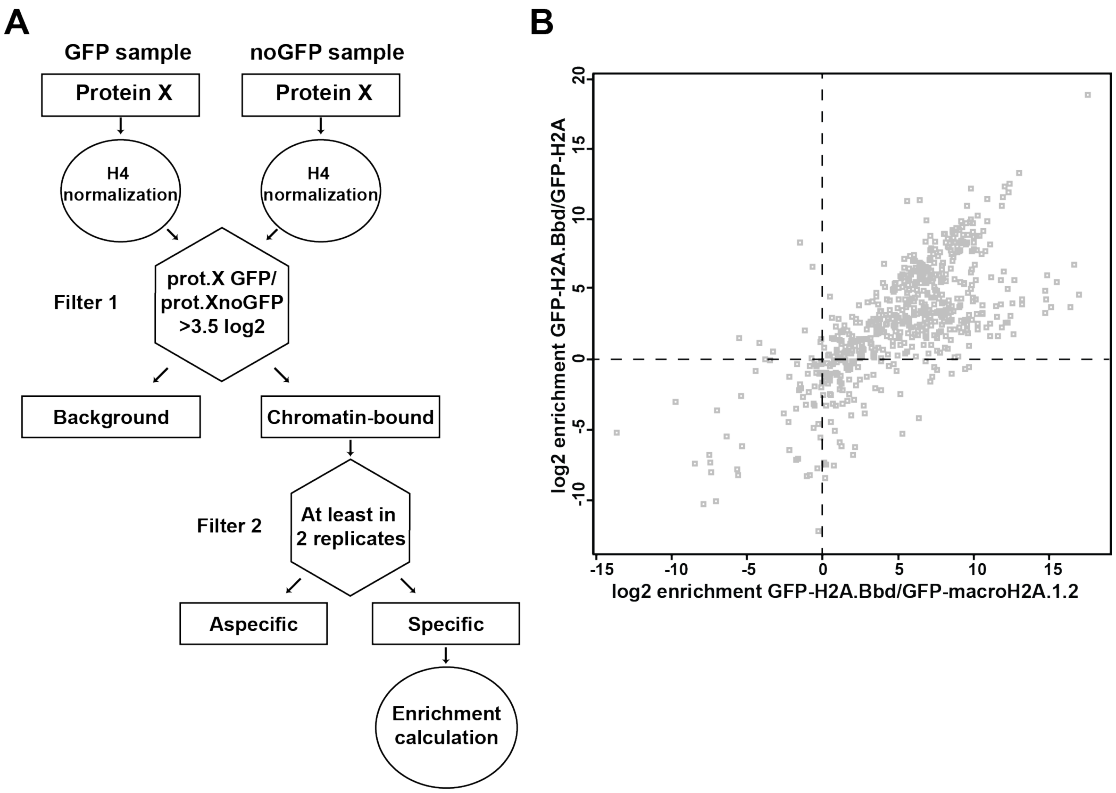

Figure 4.8 Chromatin containing the GFP-tagged histone variants shows different protein composition.
(A) Schematic representation of the mass spectrometry data analysis workflow. **(B)** Proteomic analysis of GFP-H2A.Bbd, GFP-macroH2A.1.2 and GFP-H2A containing chromatin. Specific proteins enriched in the GFP-H2A.Bbd pull down over the proteins in the GFP-macroH2A.1.2 (x-axis) and GFP-H2A (y-axis) pull downs are displayed in a scatter plot according to their log2 enrichment values.

Consistent with the observation that H2A.Bbd and macroH2A.1.2 show a very different localization pattern, the proteomic profile of the chromatin containing their GFP-tagged version turned out to be very divergent.

As there is not so much knowledge about the effects of H2A.Bbd on chromatin structure and its molecular function, the subsequent analysis focused on the protein factors enriched in the GFP-H2A.Bbd sample. Due to the large number of protein enriched in the chromatin

containing GFP-H2A.Bbd, Gene Ontology (GO) analysis was performed to understand whether there are particular functional categories overrepresented. The GO analysis of the proteins enriched in the GFP-H2A.Bbd compared to the GFP-macroH2A.1.2 containing chromatin revealed the presence of several factors involved in splicing regulation and mRNA processing (Appendix Table I). This is consistent with the recent discovery that H2A.Bbd might play a role in RNA splicing (Tolstorukov et al., 2012).

In addition, several proteins involved in the replication process were enriched in the GFP-H2A.Bbd compared to the GFP-macroH2A.1.2 and GFP-H2A containing chromatin (Appendix Table I). Among those some revealed a very high enrichment factors, such as the proliferating cell nuclear antigens (PCNA), the facilitates chromatin transcription complex (FACT), the chromatin assembly complex 1 (CAF1) and some components of the mini chromosome maintenance complex (MCM) (Table 4.1). Some of these proteins are also enriched when comparing GFP-H2A.Bbd chromatin with chromatin isolated from cells expressing GFP-H2A, suggesting that they are specific for GFP-H2A.Bbd rather than excluded from chromatin that contains GFP-macroH2A.1.2 (Appendix Table II). Interestingly, there is one protein factor that showed a very high and significant enrichment in the chromatin containing GFP-H2A.Bbd. Its name is Protein Phosphatase 1G (Ppm1G) and it was shown to be a putative chaperone for the H2A-H2B dimers (Kimura et al., 2006) (Table 4.1).

Protein names	log2 enrich. bbd/h2a samples	p-value	log2 enrich. bbd/macro samples	p-value
Fin13;Ppm1g	18.9	0.0004	17.53	0.0005
Pcna	9.26	0.006	9.03	0.004
Gins1	7.42	0.02	5.8	0.03
Mcm3	5.44	0.29	4.84	0.27
Cdc21;Mcm4	-1.04	0.48	0.09	0.96
Cdc46;Mcm5	0.25	0.97	6.21	0.17
Cdc47;Mcm7	1.17	0.75	2.63	0.27
Mcm6	0.86	0.92	3.21	0.47
Ssrp1	4.31	0.18	10.13	0.01
Fact140	5.35	0.15	9.86	0.02
Chaf1b	7.02	0.09	5.98	0.10
Rbbp4	2.56	0.61	3.46	0.36

Table 4.1 Replication associated factors are enriched in the chromatin containing GFP-H2A.Bbd.
Enriched replication factors in the GFP-H2A.Bbd containing chromatin are displayed according to their enrichment factors and p-values calculated in the GFP-H2A.Bbd over GFP-H2A and in the GFP-H2A.Bbd over GFP-macroH2A.1.2 samples.

In contrast to GFP-H2A.Bbd, the chromatin containing GFP-macroH2A.1.2 showed an enrichment of factors involved in transcriptional repression and heterochromatin formation (Table 4.2). So are, for example, two components of the Polycomb Repressive Complex 2 (PRC2), EED and EZH2 significantly enriched in the GFP-macroH2A.1.2 purified chromatin. Interestingly, also the components of the Polycomb Repressive Complex 1 (PRC1), CBX8 and BMI1 and the methyltransferase SUV39 were identified and enriched in the GFP-macroH2A.1.2 chromatin.

Protein names	log2 enrich. bbd/h2a samples	p-value	log2 enrich. bbd/macro samples	p-value
Eed	-7.37	0.00494	-8.45	0.003
Ezh2	-10.26	0.00095	-7.86	0.002
Cbx8	-6.14	0.58	-5.31	0.548
Suv39h1	-5.49	0.51	-6.36	0.334
BmI1	0.59	0.86	-3.31	0.108

Table 4.2 GFP-macroH2A.1.2 containing chromatin is enriched in heterochromatic factors.
Enriched factors in the GFP-macroH2A.1.2 containing chromatin are displayed according to their enrichment factors and p-values calculated in the GFP-H2A.Bbd/GFP-macroH2A.1.2 and in the GFP-H2A.Bbd/GFP-H2A samples.

Taken together the mass spectrometry results confirmed that the histone variants H2A.Bbd and macroH2A.1.2 are incorporated at different chromatin domains marked by specific protein factors and complexes. Consistent with the current literature, showing that H2A.Bbd is detected preferably at sites of alternative splicing, the analysis showed a strong enrichment of factors that play a role in RNA metabolism. As expected from its nuclear localization and in contrast to GFP-H2A.Bbd, GFP-macroH2A.1.2 associates with repressive chromatin.
Moreover, the proteomic analysis of the chromatin containing GFP-H2A.Bbd revealed an interesting connection of this variant with factors involved in the replication process.

4.4 Histone modification analysis of chromatin containing GFP-tagged histone H2A.Bbd and macroH2A.1.2

As the proteomic analysis of the GFP-H2A.Bbd and GFP-macroH2A.1.2 containing chromatin showed significant differences in the protein composition, it was further hypothesized that the histone modifications patters of the purified chromatin fibres could also reveal interesting differences.

MacroH2A is incorporated at sites where the repressive mark H3K27me3 is highly enriched (Costanzi and Pehrson, 1998). In contrast, H2A.Bbd was first discovered to be present at H4 hyperacetylated loci (Chadwick and Willard, 2001a). To test whether the chromatin fibres containing GFP-H2A.Bbd and GFP-macroH2A.1.2 displayed distinct histone modifications patterns, the histone H3 and H4 bands were cut out of the Coomassie gel after chromatin GFP-affinity purification and processed for histone modifications analysis.

To understand whether particular modifications could be specifically enriched in the chromatin containing GFP-H2A.Bbd and GFP-macroH2A.1.2, the histone H3 and H4 were also analyzed for the chromatin containing GFP-H2A. The analysis was performed in collaboration with Teresa Barth.

The histone modifications found in the chromatin purified fractions showed different enrichment levels from the background histone fractions represented by the histones that bind the GFP-trap beads, providing further evidences that the samples are constituted by specific chromatin domains associated to the tagged histone variants.

Particular attention was given to the peptide derived from the amino acids 27 to 40 and 9 to 17 of H3 that contain the repressive marks H3K27me3 and H3K9me3. According to the literature and to the proteomic data that showed an enrichment of the PRC2 complex and SUV39 enzyme, these two modifications were expected to be enriched on the chromatin fibres containing GFP-macroH2A.1.2. However, no particular enrichment for H3K27me3 and H3K9me3 was observed in the GFP-macroH2A.1.2 compared to GFP-H2A.Bbd containing chromatin (Figure 4.9 A-B).

H2A.Bbd was previously shown to colocalize with H4 hyperacetylated loci (Costanzi and Pehrson, 1998). However, the analysis of the modification of H4 present in the chromatin containing GFP-H2A.Bbd did not show any particular enrichment of acetylation compared to the H4 present in the GFP-macroH2A.1.2 and GFP-H2A containing chromatin (Figure 4.9 C).

The proteomic analysis of the chromatin containing specific GFP tagged histone variants revealed an interesting enrichment of factors involved in the replication process associated with GFP-H2A.Bbd indicating the presence of replicating chromatin in the sample (Table 2.9). To determine whether the enrichment of replicating chromatin in the GFP-H2A.Bbd sample could also result in a different H3.1/H3.3 ratio, the levels of H3.1 and H3.3 were further quantified. However, no differences were detected in the H3.1/H3.3 levels in the purified chromatin containing the three GFP-tagged variants (Figure 4.9 D).

In conclusion, despite the different nuclear localization and mutual chromosomal exclusion of H2A.Bbd and macroH2A, chromatin fibres containing the GFP tagged histone variants showed similar distribution of the histone modifications. However, the analyzed histones seem to be specific for the purified chromatin as they show clear differences when compared to the histones bound to the GFP-trap beads.

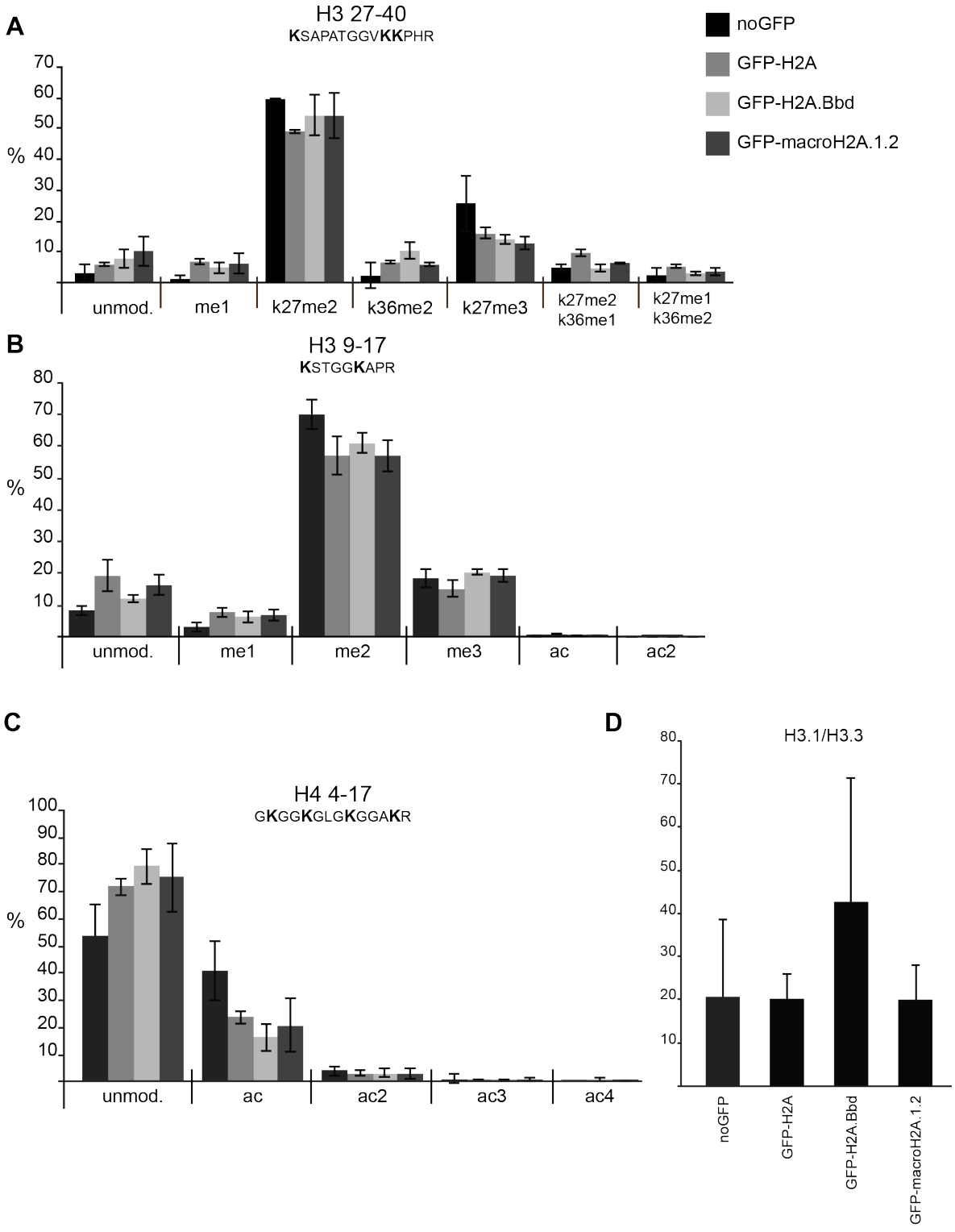

Figure 4.9 Chromatin containing the GFP-tagged histone variants H2A.Bbd, macroH2A.1.2 or H2A does not show any enrichment for particular histone variants.
A) Modifications percentage of H3 peptide 27 to 40 **B)** Modifications percentage for the histone H3 peptide 9 to 17 **C)** Modification percentage for the histone H4 peptide 4 to 17 **D)** Ratio between the level of H3.1 and H3.3 after GFP chromatin purification.

4.5 H2A.Bbd localizes at replication foci during S-phase

Ishibashi and colleagues described a particular localization pattern in a mouse embryonic cell line that expressed tagged-H2A.Bbd. In those cells, they detected a 'dotty pattern' where H2A.Bbd was co-localizing with DNA dense regions (Ishibashi et al., 2010). Interestingly, a fraction of MEF cells expressing GFP-H2A.Bbd generated for this project showed the same localization pattern described above (Figure 4.10).

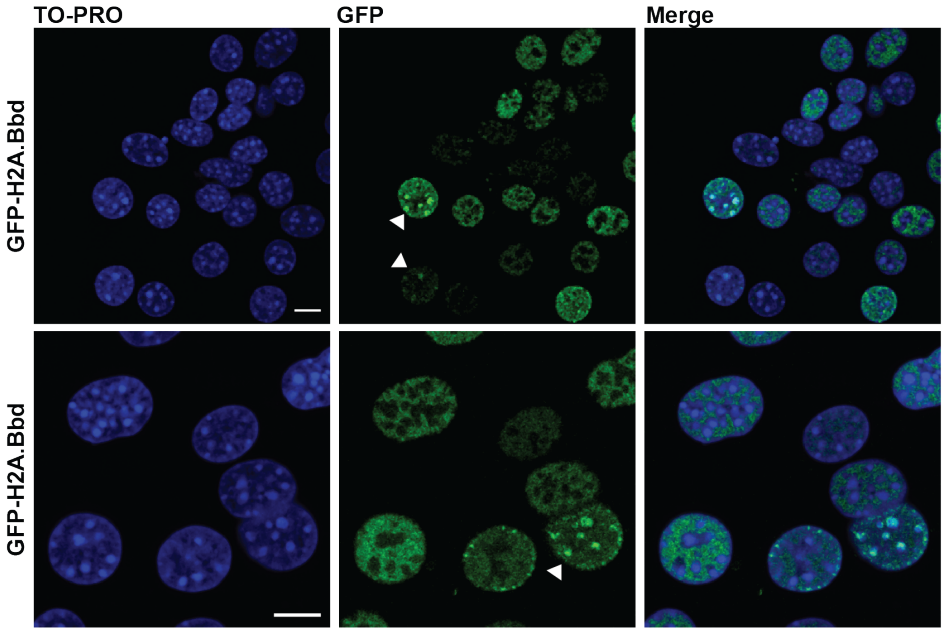

Figure 4.10 H2A.Bbd co-localizes with DNA-dense regions in a sub-population of cells expressing GFP-H2A.Bbd.
Single confocal slice of asynchronous population of female MEF cells that express GFP-H2A.Bbd. To-PRO: DNA (blue); GFP: GFP-H2A.Bbd (green). The arrow indicates GFP-H2A.Bbd co-localizing with DNA dense regions. The two panels represent different areas of acquisition. Scale bar upper panel: 5 µm; lower panel: 10 µm.

Together with the mass spectrometry results, that showed an enrichment of replication-associated factors in chromatin containing GFP-H2A.Bbd, the co-localization with DAPI dense regions suggested a possible replication coupled incorporation of GFP-H2A.Bbd. To test whether the MEF cells that show the co-localization of GFP-H2A.Bbd with the DAPI dense regions are in S-phase, an asynchronous population of cells was labeled with 5-ethynyl-2'-deoxyuridine (EdU) for 30 minutes to visualize sites of DNA synthesis in proliferating cells (Salic and Mitchison, 2008). This experiment revealed that the cells that showed the 'dotty pattern' were cells in S-phase (Figure 4.11).

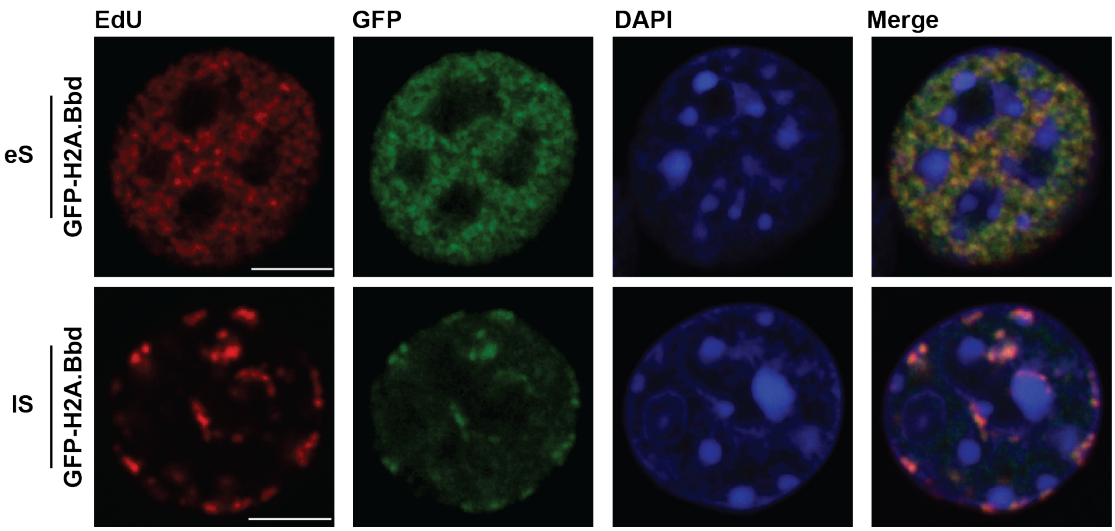

Figure 4.11 H2A.Bbd localizes at sites of replication during all stages of the S-phase.
Single confocal slice of female MEFs expressing H2A.Bbd labeled with EdU and stained with anti-GFP at different time of replication. EdU: replication foci (red); GFP-H2A.Bbd: GFP (green); DAPI: DNA (blue). Early S-phase: eS; late S-phase: lS. Scale bar: 5 µm.

Co-localization of GFP-H2A.Bbd and EdU is observed throughout S phase exhibiting the expected distribution pattern in early replicating (eS) and late replicating chromatin (lS) described for mammalian cells (Figure 4.11) (Chakalova et al., 2005). Outside S-phase the GFP-H2A.Bbd, as well as the GFP-macroH2A.1.2 and the GFP-H2A signals are distributed all over the nucleus without any particular pattern (data not shown).

To better understand the dynamics of the GFP-H2A.Bbd incorporation at sites of replication, an asynchronous MEF cells expressing GFP-tagged H2A.Bbd, macroH2A.1.2 and H2A were labeled for 30 minutes with EdU. First the cells were fixed immediately after EdU labeling (Figure 4.12 A-B).

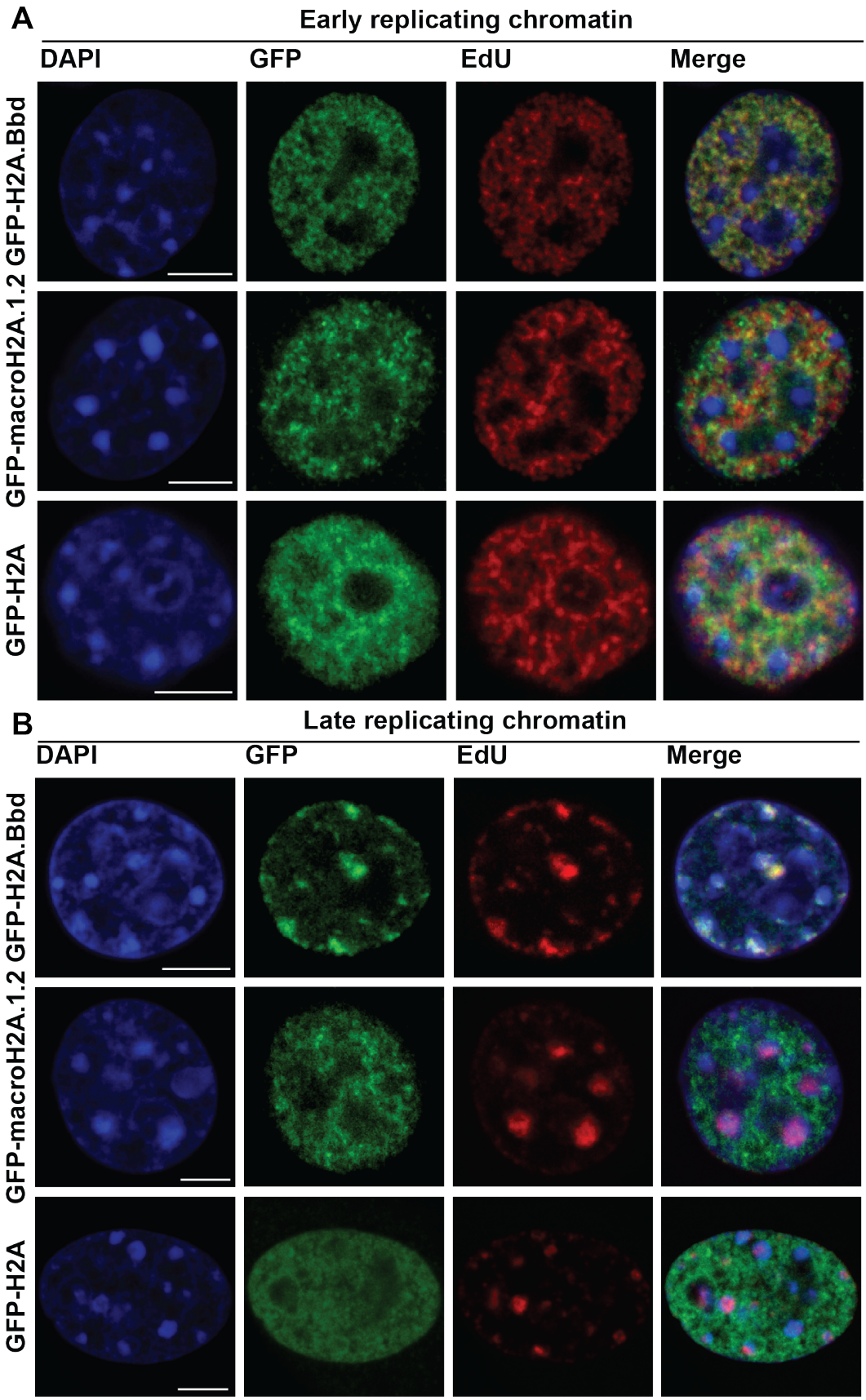

Figure 4.12 H2A.Bbd localizes at sites of DNA synthesis during replication.
(A) Single confocal slice of female early- replicating MEFs expressing the GFP-tagged histone variants pulsed labeled with EdU. **(B)** Single confocal slice of late-replicating MEFs expressing the GFP-tagged histone variants pulsed labeled with EdU. DAPI: DNA (blue); GFP: GFP-tagged histone variant (green); EdU: replication foci (red). Scale bar: 5 µm.

The analysis of the channel intensities revealed that only GFP-H2A.Bbd shows a very strong co-localization with the EdU labeled sites suggesting that this histone variant is recruited to sites of DNA replication (Figure 4.13 A-B).

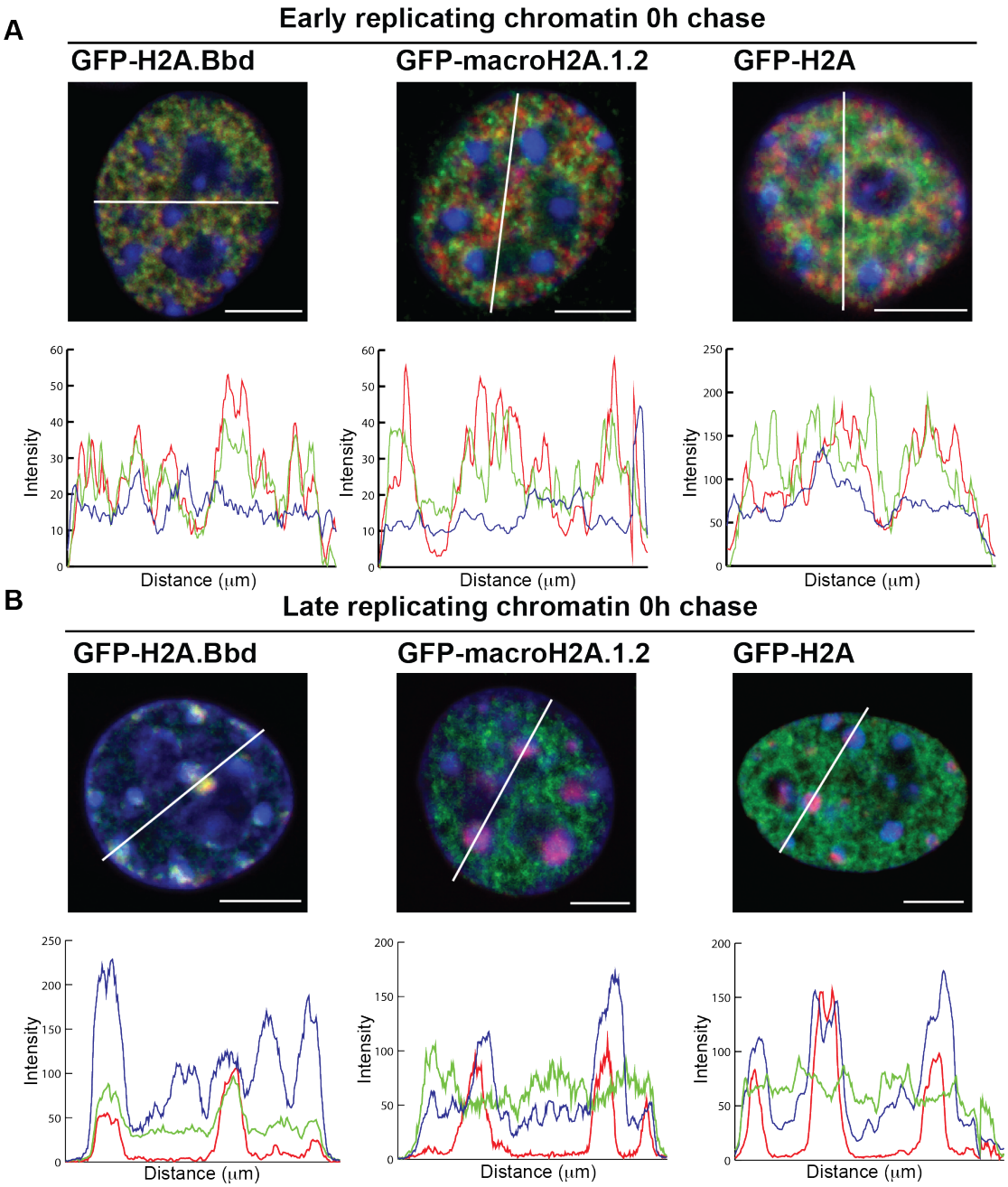

Figure 4.13 GFP-H2A.Bbd, but not GFP-macroH2A.1.2 or GFP-H2A, co-localizes with DNA replication foci during S-phase.
(A) The channels intensity recorded in early replicating cells along the white line is plotted on a diagram. Blue line: DNA; green line: GFP-tagged histone variants; red line: replication foci. **(B)** The channels intensity recorded in late- replicating cells along the white line is plotted on a diagram. Blue line: DNA; green line: GFP-tagged histone variants; red line: replication foci. Scale bar: 5 μm.

Histone variants have been already shown to localize at sites of DNA synthesis sites (Polo et al., 2006). However, such a strong co-localization with DNA replication has not been observed so far. Moreover, the incorporation of GFP-H2A.Bbd is visible during different stages of the S-phase suggesting that the histone variant follows the replication process.

The next question was whether the incorporation of GFP-H2A.Bbd at replication foci is stably maintained after the replication machinery moves on or whether GFP-H2A.Bbd is removed from sites of DNA replication. To address this point an asynchronous population of MEF cells was pulse for 30 minutes with EdU and then chased for 1h before fixation. After 1h chase the GFP-H2A.Bbd expressing MEF cells did not show the co-localization of the histone variant with the newly replicated DNA anymore (Figure 4.14 A-B).

The analysis of the channels intensity confirmed that GFP-H2A.Bbd is no longer co-localizing with the EdU signal after 1h chase but is rather neighboring the EdU labeled DNA (Figure 4.15 A-B). This is particularly evident in the late- replicating cells (Figure 4.15 B), suggesting that GFP-H2A.Bbd incorporation follows the replication process.

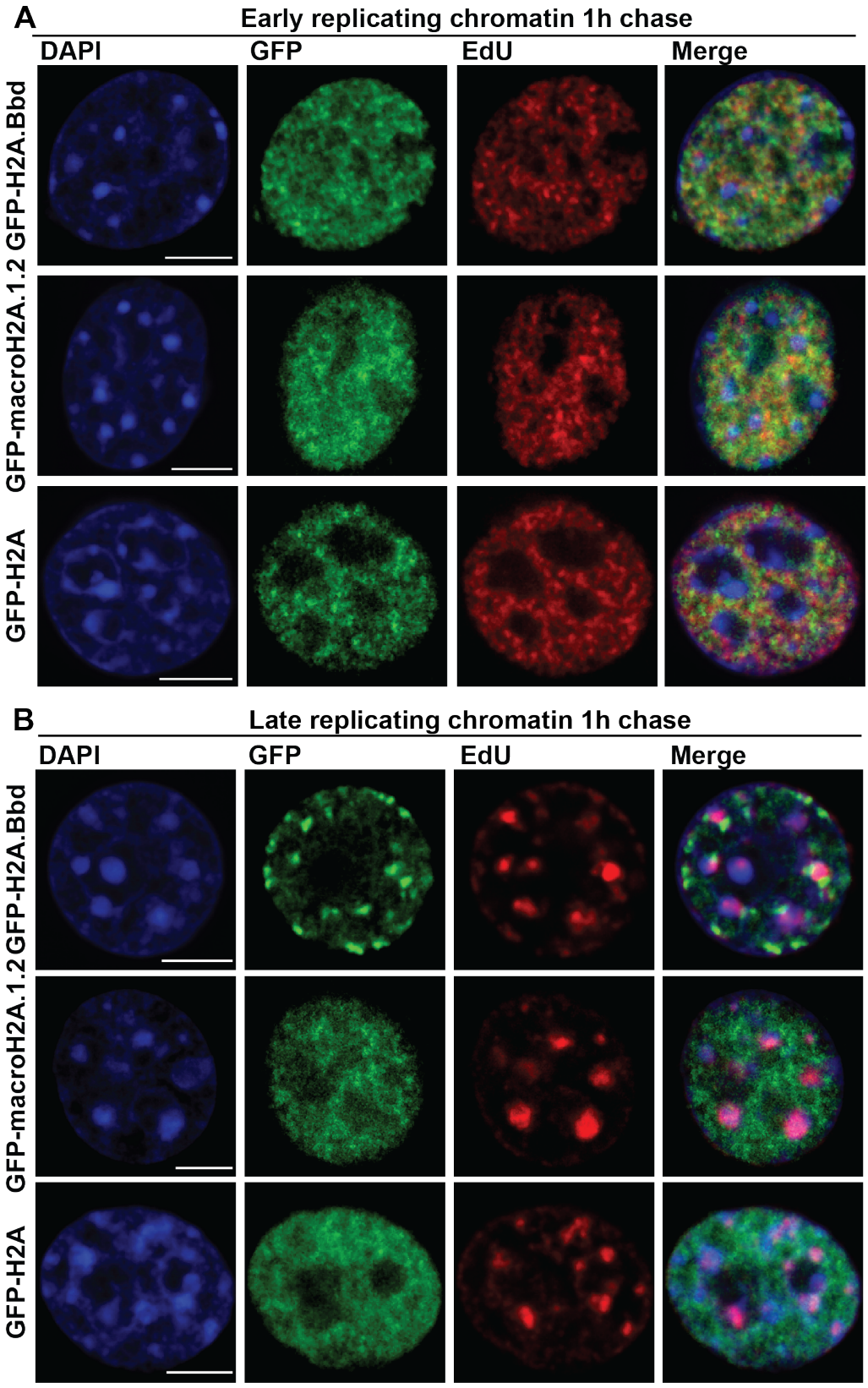

Figure 4.14 H2A.Bbd localizes at sites of DNA synthesis during replication and is removed within 1h. **(A)** Single confocal slice of female early- replicating MEFs expressing the GFP-tagged histone variants pulsed labeled with EdU and chased for 1h. **(B)** Single confocal slice of late-replicating MEFs expressing the GFP-tagged histone variants pulsed labeled with EdU and chased for 1h. DAPI: DNA (blue); GFP: GFP-tagged histone variant (green); EdU: replication foci (red). Scale bar: 5 µm.

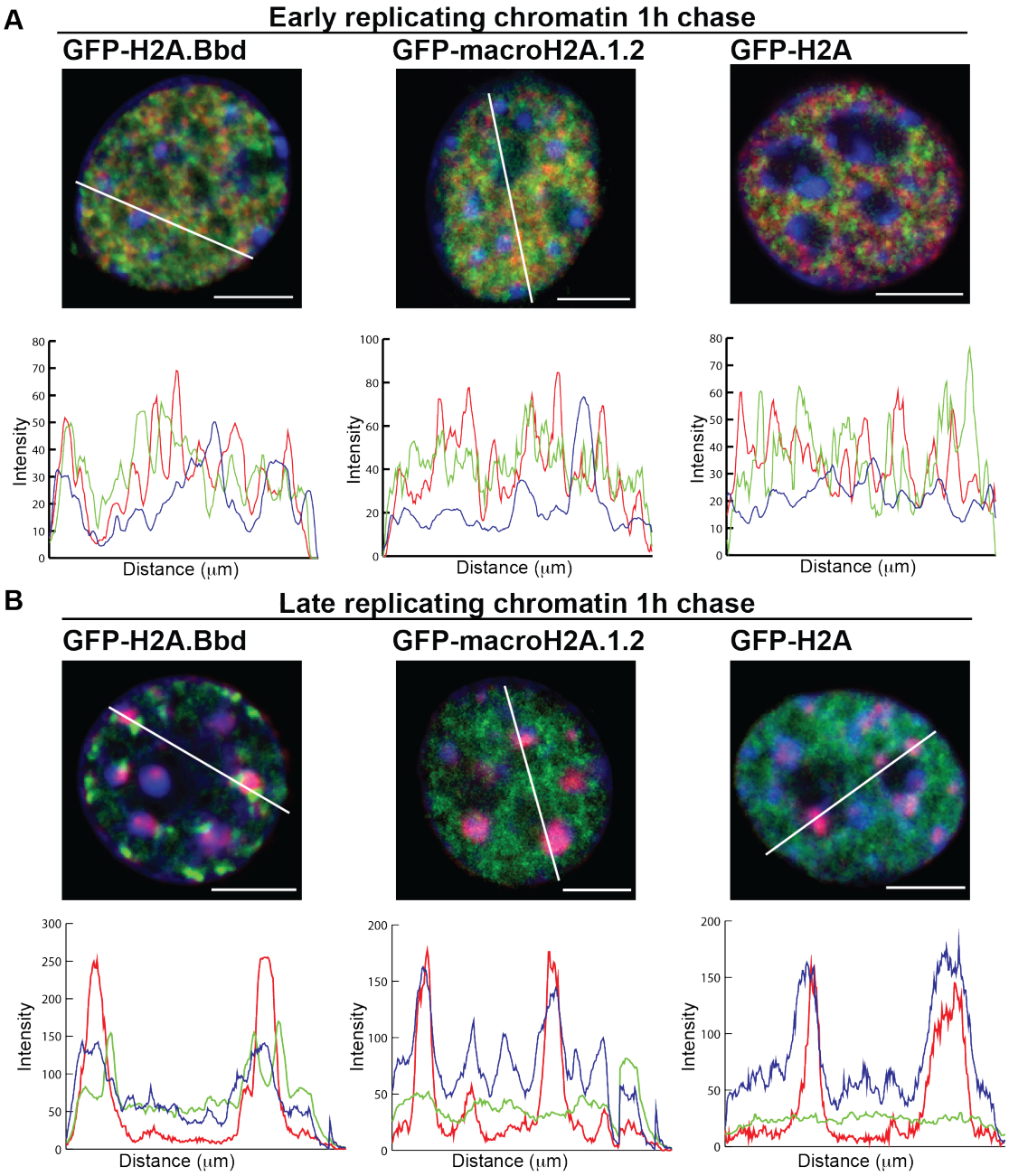

Figure 4.15 GFP-H2A.Bbd is removed from sites of DNA synthesis within 1h during replication.
(A) The channels intensity recorded in early replicating cells along the white line is plotted on a diagram. Blue line: DNA; green line: GFP-tagged histone variants; red line: replication foci. **(B)** The channels intensity recorded in late- replicating cells along the white line is plotted on a diagram. Blue line: DNA; green line: GFP-tagged histone variants; red line: replication foci. Scale bar: 5 µm.

All together these results show that GFP-H2A.Bbd is incorporated into chromatin during replication. The incorporation of GFP-H2A.Bbd during replication is not stable, but the tagged histone variant is removed within 1h.

4.6 GFP-H2A.Bbd follows PCNA at replication foci

One interesting protein identified and enriched in the GFP-H2A.Bbd chromatin affinity purification was the Proliferating Cell Nuclear Antigen (PCNA).

PCNA is a DNA polymerase cofactor during the replication process. It adopts a ring-shape structure and encircles the replicated DNA strand. The presence of PCNA at the replication foci is the key event for the association of the replicative polymerases to the DNA strands. Moreover, PCNA enhances the processivity of the polymerases enzymes and is a binding platform for many proteins involved in the replication process (Moldovan et al., 2007).

To investigate whether H2A.Bbd is following the replication fork we looked at live cell imaging using MEFs expressing GFP-H2A.Bbd co-transfected with an mRFP-PCNA. The following experiment was performed in collaboration with Dr. Corella Casas Delucchi. By following the mRFP-PCNA during replication it is clear that GFP-H2A.Bbd is recruited after PCNA at the replication foci (Figure 4.16).

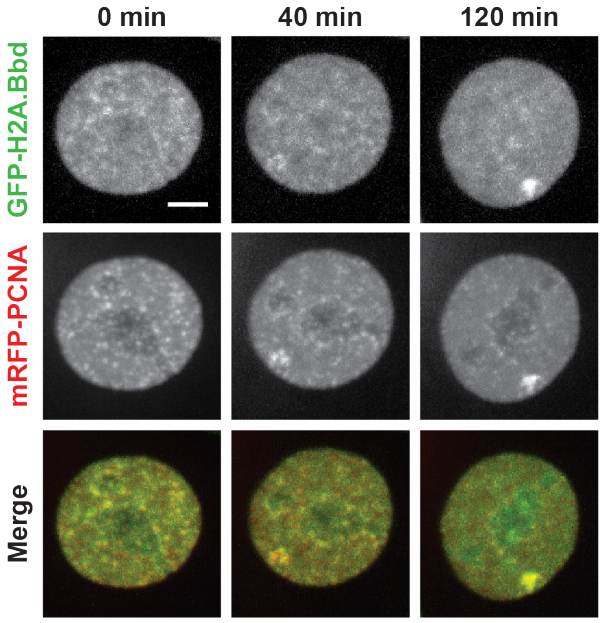

Figure 4.16 H2A.Bbd follows PCNA at replication foci.
Selected frames of time-lapse imaging of MEF cells stably expressing GFP-H2A.Bbd (green) transfected with mRFP-PCNA (red). Z-stacks were collected every 20 min over night using a spinning disk microscope equipped with temperature, humidity and CO_2 control. Shown are maximum intensity projections.

4.7 H2A.Bbd is recruited at DNA damage sites after microirradiation

PCNA is not only involved in DNA replication, but also plays an important role in DNA repair (Moldovan et al., 2007). The recruitment of GFP-H2A.Bbd to replication foci following PCNA suggested that GFP-H2A.Bbd could be also recruited to sites of DNA damage. To test this hypothesis, microirradiation experiments were performed in collaboration with Dr. Corella Casas Delucchi.

MEF cells expressing GFP-H2A.Bbd were microirradiated and GFP-H2A.Bbd recruitment was monitored in live imaging cells. Different frames of a live cell imaging movie were selected and clearly show that GFP-H2A.Bbd is recruited at sites of DNA damage after microirradiation (Figure 4.17). It is interesting to notice that when recruited at DNA damage sites, GFP-H2A.Bbd signal is still detectable after 110 minutes post-irradiation in contrast to its fast replacement during replication.

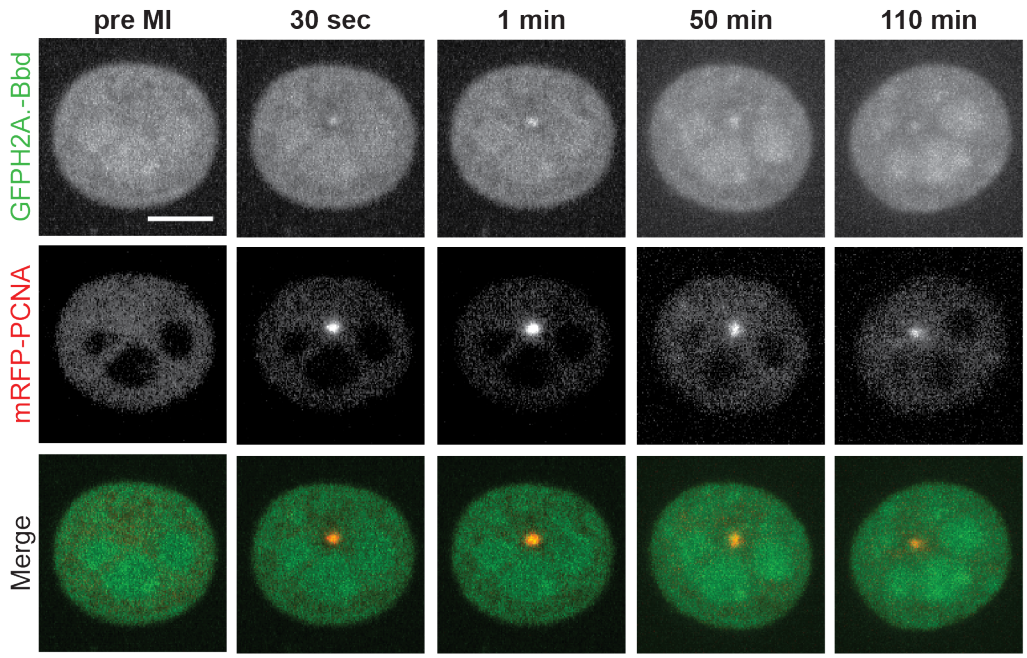

Figure 4.17 H2A.Bbd accumulates at DNA damage sites following PCNA.
Time-lapse imaging of microirradiated MEF cells stably expressing GFP-H2A.Bbd (green) and transfected with mRFP-PCNA (red) at 405 nm for 1200 ms. Images were collected before microirradiation and 5 seconds, 10/50 and 110 minutes after microirradiation. Scale bars: 5 µm.

4.8 MEF cells expressing GFP-H2A.Bbd have a shorter S-phase and are more sensitive to DNA damage

The association of GFP-H2A.Bbd to replication and DNA damage questioned whether its recruitment could have an effect on the S-phase length or DNA damage sensitivity. To test whether the localization of GFP-H2A.Bbd at replication and DNA damage loci has an effect on replication and sensitivity to damage, the doubling time and the S-phase time were measured in MEF cells expressing GFP-H2A.Bbd and GFP-H2A. These measurements were obtained in collaboration with Malini Rajan. This analysis demonstrated that MEF cells expressing GFP-H2A.Bbd have a shorter doubling time and a shorter S-phase compared to the MEF cells expressing GFP-H2A (Figure 4.18 A).

Moreover, the analysis of the Cyclobutane Pyrimidine Dimers (CPD), which indicates the level of DNA damage in UV treated cells, shows a higher sensitivity to DNA damage in GFP-H2A.Bbd compared to GFP-H2A expressing MEF cells 30 minutes after irradiation (Figure 4.18 B).

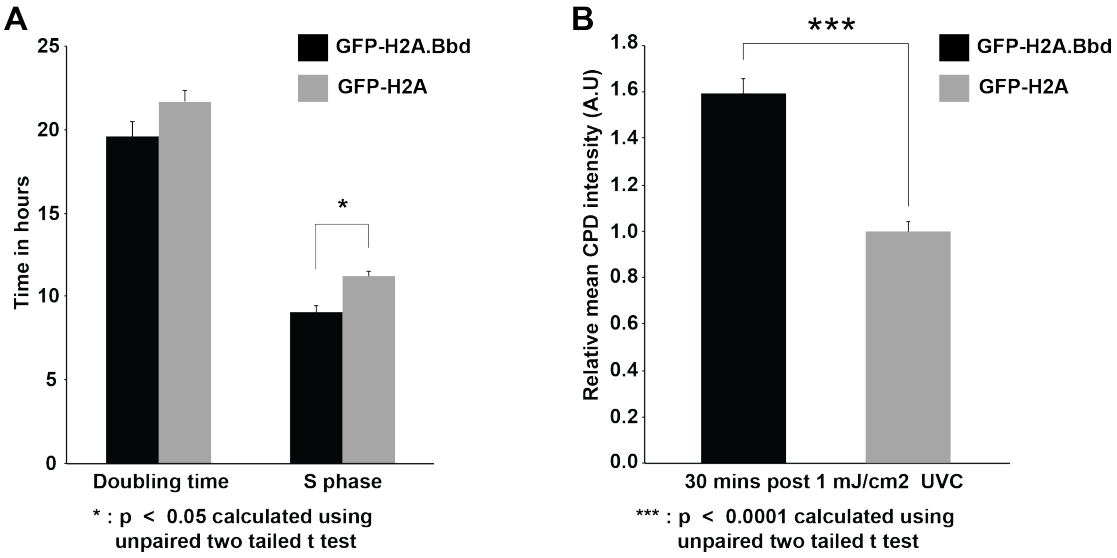

Figure 4.18 MEF cells expressing H2A.Bbd replicate faster and are more sensitive to DNA damage.
(A) Doubling time and S-phase length in MEF cells stably expressing GFP-H2A.Bbd and GFP-H2A. **(B)** Mean intensity of nuclear CPD signal 10 and 30 minutes post 1mJ UVC irradiation. Error bars represent the 95% confidence interval. N>30.

4.9 H2A.Bbd expression in Hodgkin's Lymphoma cell lines

Under physiological conditions H2A.Bbd is only expressed in testis and to a lower extent in brain (Chadwick and Willard, 2001a; Ishibashi et al., 2010). This particular expression pattern suggested a specific role of H2A.Bbd during spermatogenesis. Interestingly, there is a special class of genes that are expressed in testis and silent in other tissues, which become re-activated in different cancer types. These genes are classified as Cancer Testes Antigens (CTA) (Cheng et al., 2011). Recently, a microarray based screening for new CTAs in Hodgkin's Lymphoma cells revealed that H2A.Bbd could indeed be a new potential CTA gene (Winkler et al., 2012). In collaboration with Dr. Martin Staege the expression levels of H2A.Bbd were measured in five Hodgkin's Lymphoma cell lines by qPCR. H2A.Bbd is indeed expressed in these cell lines with different expression levels (Figure 4.19 A). The H2A.Bbd protein levels were measured by Western Blotting using nuclear extracts from five different Hodgkin's Lymphoma cell lines. As control EBV-immortalized B-cell lines, LCL2 and EBV1.11, were used. Interestingly, H2A.Bbd protein can be detected in at least four out of five Hodgkin's Lymphoma cell lines with different expression levels confirming the results obtained in the mRNA analysis (Figure 4.19 B).

When the doubling time of the Hodgkin's Lymphoma cells lines reported in (Drexler et al., 2010) was compared to the level of H2A.Bbd expression in Figure 4.19 A, an interesting correlation was found between the time of replication and the expression levels. In summary, the cells that express the higher levels of H2A.Bbd, such as the L428 or the L1236, replicate faster whereas the cell line that show lower level of expression, i.e. the L540 and the HDLM-2, are slower in doubling (Figure 4.19 C).

Taken together these results show for the first time that H2A.Bbd protein is abnormally expressed in Hodgkin's Lymphoma cell lines and that the level of expression correlates with their replication time.

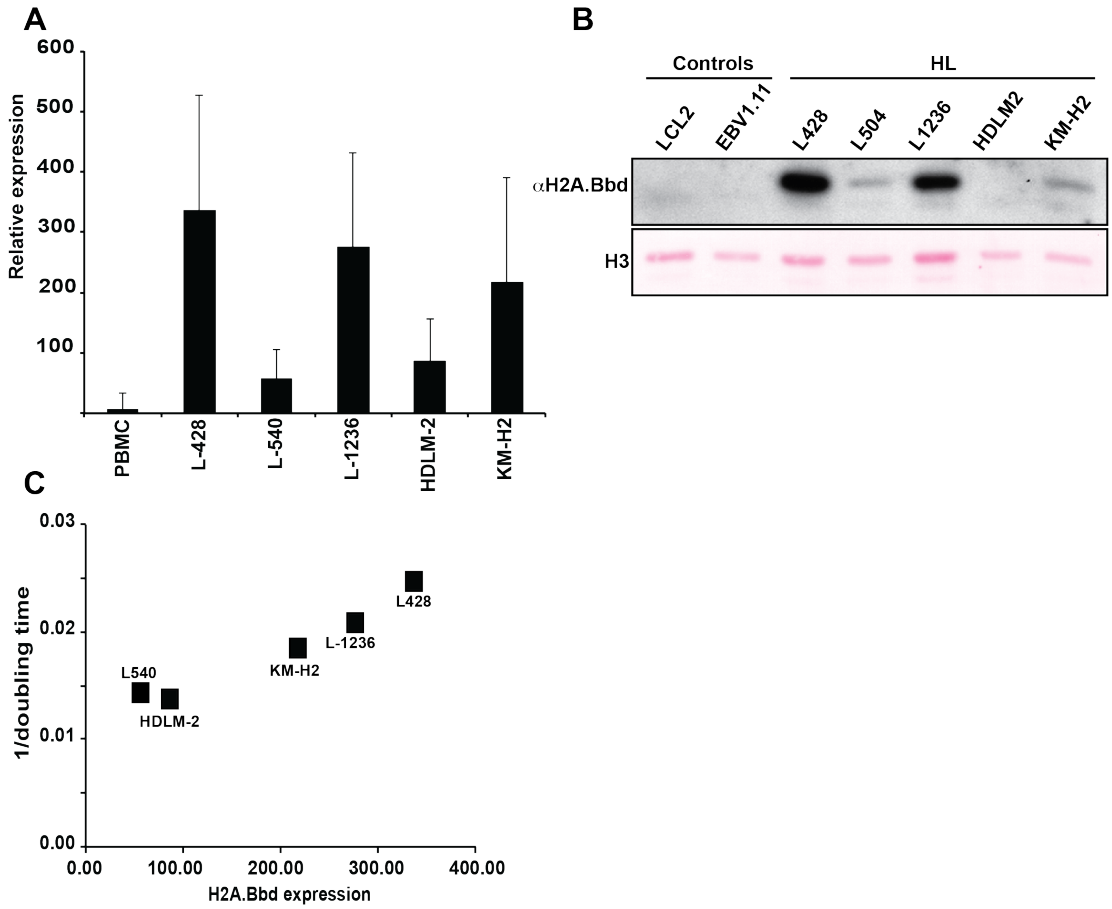

Figure 4.19 H2A.Bbd is expressed in Hodgkin's Lymphoma cell lines.
(A) rRT-PCR expression analysis of H2A.Bbd in HL cell and control cell lines. cDNA from HL cell line (2 independent experiments and normal PBMC (21 donors) was used for quantitative PCR. For calculation of relative expression values, actin beta was used as housekeeping control and the median expression in normal PBMC was set as 1. **(B)** Western blot analysis of H2A.Bbd in Hodgkin's lymphoma cells (HL) and control cell lines. Ponceau staining of the histone H3 is used as loading control. **(C)** Dot plot showing the correlation in Hodgkin's Lymphoma cell lines between the doubling time derived from (REF) (y-axis) and the expression levels of H2A.Bbd (x-axis_ measured by qRT-PCR). Every dot represents a specific Hodgkin's Lymphoma cell lines.

5 Discussion

Chromatin is regulated in distinct ways including the addition or removal of specific histone modifications and the recruitment of specific protein factors that read and erase the histone modifications. In addition, the histone variants play a significant role in the regulation of chromatin accessibility by modulating the nucleosome structure and consequently the higher order chromatin structure. However, in recent years, the role of the histone variants in different cellular processes has been addressed mainly by looking at their effects on the nucleosome stability or their association to specific DNA sequences.

The main aim of this work was to better define the function of the histone variants H2A.Bbd and macroH2A in the light of the protein environment they are in. To achieve this, specific chromatin domains containing the GFP-tagged histone H2A.Bbd and macroH2A.1.2 were purified and their proteomic profile was analyzed using mass spectrometry.
After analyzing the possible reason why it was not possible to achieve comparable levels of expression for the three GFP-tagged histone variants, the discussion will mainly focus on the results obtained from the proteomic analysis of the GFP-tagged histone containing chromatin. These experiments revealed remarkable differences in the protein composition of the two 'types' of chromatin containing the tagged histone variants. The presence of specific heterochromatic or euchromatic factors and their putative link with the histone variant H2A.Bbd and macroH2A.1.2 will be further explored in this section.

Moreover, this study also revealed new, interesting features of H2A.Bbd. Indeed it was shown for the first time that GFP-H2A.Bbd ectopically expressed in MEF cells localizes to sites of DNA synthesis during replication and DNA repair. Due to its tissue-specific expression, it is unlikely that H2A.Bbd could play a role in somatic cell during DNA replication and repair, but this new finding is particularly relevant and will be further discussed in the light of the H2A.Bbd expression detected in Hodgkin's Lymphoma cell lines. The effects on DNA damage sensitivity observed in MEF overexpressing GFP-H2A.Bbd will be also used to decipher the possible role that H2A.Bbd could have in testis, where it is normally expressed during male germ cells differentiation.
In summary, this study not only provided new evidences about the importance of the histone variants in chromatin regulation but also opened new questions about the role of H2A.Bbd in cancer and spermatogenesis.

5.1 Potential negative effect of macroH2A.1.2 expression in MEF cells

To study the proteomic composition of the chromatin containing the histone variants H2A.Bbd and macroH2A.1.2, MEF cells were transfected with GFP tagged H2A.Bbd and macroH2A.1.2. Thereafter, single cell clones expressing the GFP-tagged histones were isolated from the transfected MEFs by FACS sorting.
Interestingly, the first three single cell clones selected for GFP-macroH2A.1.2 lost the expression of the GFP tagged variant after some consecutive weeks in culture. The population derived from six sorted cells maintained the expression of GFP-macroH2A.1.2 in most of the cells resulting in a low expression level compared to the MEF cells expressing GFP-H2A.Bbd and GFP-H2A (Figure 4.1 B). It is not possible to exclude that the repression of the exogenous gene is the result of a technical problem such as the CpG content of the construct.

However, one intriguing hypothesis is that the MEF cells expressing GFP-macroH2A.1.2 do not allow high levels of the exogenous gene expression. This would explain the failure in achieving the same expression levels for all the GFP-tagged histone variants and the repression observed in the first three single cell clones. In support of this hypothesis, increasing evidence points towards a repressive role of macroH2A in cancer progression (Cantariño et al., 2013). In fact, macroH2A.1.1 was found up-regulated in tumors that show a low mitotic index, whereas tumors with higher mitotic indices displayed low expression of this histone variant. This suggests the role of macroH2A.1.1 as a tumor-suppressor (Sporn et al., 2009). MacroH2A.1.2 function in tumorigenesis in contrast appears to be more context-dependent. Up-regulation in melanoma cells, for example, leads to a reduced metastatic potential (Kapoor et al., 2010), whereas in breast cancer, high levels of macroH2A.1.2 promoted cell growth, invasion and migration (Dardenne et al., 2012; Li et al., 2012).
Viewed together, these findings suggest that macroH2A.1.2 overexpression decreases the proliferative potential of MEF cells. Extended passaging of GFP-macroH2A.1.2 transfected MEFs may therefore enrich for cells with lower expression level.

5.2 GFP tagged H2A.Bbd and macroH2A.1.2-containing chromatin domains show differences in their proteomic profile

After chromatin affinity purification, the identification of the chromatin-bound proteins revealed an obvious difference in the number of proteins present in the respective H2A variant specific chromatin. However, also the biological replicates only showed a partial overlap (Figure 4.7). It is important to mention that the samples were measured in the mass spectrometer at different times. This could in part explain the variability observed among the replicates. In fact, many studies have been dedicated to increase the reproducibility of mass spectrometry datasets acquired on affinity purification samples. The most recent studies point to the importance of reducing the sample variability by using the same sample treatment, machines and conditions (Dazard et al., 2012; Varjosalo et al., 2013). Remarkably, increased sample variability may have also resulted from the use of native/unfixed chromatin because the abundance of proteins that displayed weak or transient interactions with the bait might be sometimes below the sensitivity of detection. Considering these technical issues, if only the proteins present in all the three replicates are considered as specific chromatin interactors, novel, interesting interactions might be discarded. Therefore, the quantification was performed using proteins present in at least two biological replicates. Using these criteria, the overlap reached an average of 50% in every GFP-H2A variant chromatin purification.

As mentioned above, the samples revealed an obvious difference in the number of identified proteins, with the GFP-H2A.Bbd containing chromatin having a higher number of entries compared to the GFP-macroH2A.1.2 chromatin. One of the possible explanations for this marked difference could be the clear expression level discrepancy observed in the MEF cell expressing GFP-H2A.Bbd and GFP-macroH2A.1.2 (Figure 4.1 B). To avoid that the expression levels differences could affect the quantification of the protein present in the purified chromatin, all the samples were normalized to the level of H4.

After statistical analysis, the two samples that showed relevant differences were the GFP-H2A.Bbd and GFP-macroH2A.1.2 purified chromatin samples. This result reflects the specific and mutual exclusive nuclear localization of the two variants ((Chadwick and Willard, 2001a; Costanzi and Pehrson, 1998) and Figure 4.2)

It is worth noting that the different chromatin composition can not be explained by major differences in the canonical histone composition, which do not show particular variation, but rather in the presence of the histone variants H2A.Bbd and macroH2A.1.2 thus supporting the important role that the histone variants play in chromatin regulation and gene expression.

5.3 Proteomic profile of GFP-macroH2A.1.2

MacroH2A is a large H2A variant characterized by a non-histone macro domain that serves as a binding module for NAD metabolites (Kustatscher et al., 2005). MacroH2A has been extensively studied in recent decades for its role in the late stages of the X chromosome inactivation that pointed to a repressive role for macroH2A in transcription regulation (Costanzi and Pehrson, 1998). However, the presence of macroH2A in species that do not undergo X inactivation and, in addition, the similar expression of macroH2A in male and female, suggested that this histone variant could be implicated in the regulation of processes outside X inactivation (Buschbeck and Di Croce, 2010; Rasmussen, 1999).
Several studies attempted to elucidate the role of macroH2A in gene expression but the more evidence obtained the more difficult it became to give macroH2A one unique function in transcription regulation. It is now clear that macroH2A functions are basically context-dependent and that it is important to discriminate among the two macroH2A genes, macroH2A.1 and macroH2A.2, and isoforms, macroH2A.1.1 and macroH2A.1.2. In this context, the identification of putative chromatin binding factors that surround the macroH2A containing chromatin has become relevant.

When comparing the proteins present in the GFP-macroH2A.1.2 chromatin together with those present in the GFP-H2A.Bbd samples, a striking enrichment of factors associated with repression of transcription was found (Table 4.2).
In particular, some of the Polycomb Repressive Complex 2 (PRC2) factors showed significant enrichment in the GFP-macroH2A.1.2 containing chromatin. This result confirmed previous data that showed an overlap between macroH2A and PRC2 binding regions (Buschbeck et al., 2009). Despite the striking enrichment of the PRC2 complex in the GFP-macroH2A.1.2 containing chromatin, the analysis of the histone modifications did not reveal any particular enrichment for the H3K27me3, which is generally known to be placed by the PRC2 complex (Margueron and Reinberg, 2011). The presence of the PRC2 complex and the absence of the H3K27me3 enrichment would suggest that not all the nucleosomes bound by the Polycomb carry the H3K27me3 modification. This would involve the presence of an alternative form of the PRC2 complex at specific loci not engaged in setting this modification. Interestingly, recent studies demonstrate the presence of the PRC2 complex at sites that are not enriched for H3K27me3. In *Xenopus* it has been shown that the presence of the PRC2 complex does not correlate with the presence of H3K27me3 at defined stages during development (van

Heeringen et al., 2014). In mouse embryonic stem cells (ESC), the PRC2 complex bound to TSSs can be repressed by nascent transcript and this results in the absence of H3K27me3 (Kaneko et al., 2013). In essence, the data support the overlap between the PRC2 complex and the histone variant macroH2A.1.2 already observed by others but has raised new questions about the PRC2 complex and its role in modifying the histone H3 on lysine 27.

Together with the PRC2, the PRC1 complex is also enriched in the GFP-macroH2A.1.2 containing chromatin (Table 4.2). In this study, one of the subunits found particularly enriched in the sample is the Chromobox homolog 8 (CBX8). The presence of CBX8 in the chromatin containing GFP-macroH2A.1.2 is very intriguing and it could be related to the putative role that macroH2A plays in cell senescence. CBX8 is a chromodomain containing protein, which was first characterized as a component of the PRC1 complex that associates with BMI1 repressing transcription (Bárdos et al., 2000). Further investigations showed that CBX8 repressed the INC4A/ARF locus regulating the proliferation of mouse and human fibroblasts and the senescence process (Dietrich et al., 2007). Interestingly, macroH2A has been found to repress Cdkn2a/p16 (INK4A) in colorectal cancer cell line leading to an enhanced proliferation (Barzily-Rokni et al., 2011). In summary, the data from the GFP-macroH2A.1.2 chromatin purification provide new evidences that these two proteins could interplay in the regulation of particular genes that are targeted by the CBX8 containing PRC1 complex.

Another factor associated with repression of transcription, which is present in the macroH2A.1.2 containing chromatin, is the methyltransferase SUV39h1 (Table 4.2). This suggests that pericentric heterochromatin is also recovered in these purifications. The presence of macroH2A at pericentric heterochromatin has already been reported, but only upon chromatin reorganization induced by nuclear transplantation or DNA hypomethylation (Pasuq v. et al., 2011; Ma Y et al., 2005). It would be possible that only a small fraction of the histone macroH2A is present at pericentric heterochromatin. This would also explain why there is no evident enrichment of H3K9me3 in the GFP-macroH2A.1.2 containing chromatin.

5.4 Proteomic profile of GFP-H2A.Bbd

Mass spectrometry analysis showed particular differences between the chromatin factors associated with GFP-H2A.Bbd and GFP-macroH2A.1.2. As of yet there is little information about putative direct or indirect interaction partners of H2A.Bbd. Recently, it was shown that in HeLa cells GFP-H2A.Bbd is incorporated in chromatin enriched for splicing factors suggesting a putative role of H2A.Bbd in the regulation of the splicing process (Tolstorukov et al., 2012). In keeping with this, the analysis of the chromatin containing GFP-H2A.Bbd showed an enrichment of factors involved in splicing regulation and mRNA processing, partially overlapping with the previous report (Appendix Table I and II).
In addition, DNA replication was also found as an overrepresented category in the GFP-H2A.Bbd containing chromatin compared to the GFP-macroH2A.1.2 containing chromatin (Appendix Table I)). This category comprised members of the DNA replication initiation complex such as the MCM proteins and the GINS, and factors involved in the replication process such as PCNA (Table 4.1). However, even if several factors involved in the replication process showed a high enrichment in the GFP-H2A.Bbd chromatin, the p-values associated with some of them are not significant. The high p-values could be explained in different ways. First of all the chromatin purification was performed using an asynchronous population of MEF cells expressing the GFP-tagged histone variant, with only a small fraction of cells that are actively replicating at the time of harvesting. Secondly, the immunofluorescence experiments revealed that the localization of GFP-H2A.Bbd at sites of active replication is only transient and disappears within 1h. As already mentioned all the experiments were performed using native chromatin, which is not the best tool to study transient interactions and this can influence the reproducibility of the data.

H2A.Bbd was shown to localize at sites of H4 hyperacetylation (Chadwick and Willard, 2001a). However, the histone modification analysis did not reveal any particular acetylation enrichment on H4 peptides compared to the GFP-macroH2A.1.2 containing chromatin. This apparent discrepancy could be related to the presence of large chromatin fragments that do not all contain the histone variant H2A.Bbd. The histone modification analysis performed on mononucleosomes would, perhaps, give better information about the modifications related to the histone variant H2A.Bbd.

In summary, the mass spectrometry analysis of the chromatin containing GFP-H2A.Bbd in part confirmed previous reports that showed a role of H2A.Bbd in RNA splicing, but also revealed new, interesting features and possible new interactors of this histone variant.

5.5 GFP-H2A.Bbd recruitment at DNA synthesis sites during DNA replication and DNA repair

GFP-H2A.Bbd is removed within 1h from the replication foci, whereas it is retained longer at DNA damage site (Figure 4.15 and Figure 4.17). One possibility is that the incorporation of GFP-H2A.Bbd at replication foci is generating unstable nucleosomes. In support of this idea there are several studies that showed an evident destabilization of the nucleosome containing H2A.Bbd compared to the canonical nucleosome (Bao et al., 2004; Eirín-López et al., 2008; González-Romero et al., 2008). Notably, a recent article showed that synchronized HeLa cells overexpressing H2A.Bbd were more sensitive during S-phase to MNase digestion compared to the cell overexpressing H2A (Goshima et al., 2014). This experiment is particularly relevant for two reasons: first, it explains why it was not possible to see a relevant variation of the chromatin compaction in non-synchronized cells after MNase digestion (Figure 4.5 A-B); second because, that H2A.Bbd needs to be removed rapidly from sites of DNA synthesis during replication to ensure the correct chromatin compaction.

Goshima also showed that the overexpression of H2A.Bbd leads to cell cycle arrest and apoptosis (Goshima et al., 2014). In contrast here it was showed that GFP-H2A.Bbd overexpressing cells are replicating faster than the GFP-H2A cells. The apparent discrepancy of phenotypes obtained in the two studies could be explained either by the use of different cell types or by a dosage dependent effect of H2A.Bbd expression. In this regards it would be interesting to express different levels of H2A.Bbd and observe whether this brings to different effects on the cell cycle.

The fast removal of GFP-H2A.Bbd from DNA replication sites could also be the reason why it was not possible to detect a specific enrichment of H3.1 in the chromatin containing GFP-H2A.Bbd (Figure 4.9).

As mentioned before, GFP-H2A.Bbd is rapidly removed from DNA synthesis sites during replication. However, after DNA microirradiation, GFP-H2A.Bbd is recruited at DNA damage sites following PCNA and it is still detectable after 110 min. at repair foci (Figure 4.17). One possibility is that GFP-H2A.Bbd is simply following the dynamic of PCNA at

DNA damage sites. It was already shown that PCNA is engaged at least two times longer at sites of DNA damage than at replication foci (Essers et al., 2005). However, it would be interesting to test whether the presence of GFP-H2A.Bbd alters the residence time of PCNA at DNA damage sites compared to non-transfected cells or to the GFP-H2A expressing cells.

5.6 H2A.Bbd and its putative role in spermatogenesis

In this study, GFP-H2A.Bbd was found to localize at sites of DNA synthesis during DNA replication and DNA repair (Figure 4.11 and 4.17). However, H2A.Bbd is known to be a testis specific histone variant (Ishibashi et al., 2010) and its expression is detected at the time of the elongating spermatids just prior to the replacement of the canonical histone when H4 is highly acetylated (Ishibashi et al., 2010). The tissue and the time-specific expression of H2A.Bbd excludes the possibility that this histone variant could play a role in DNA replication and repair in somatic cells and mitotic germ cells. However, the ability of H2A.Bbd to be recruited at DNA synthesis sites during replication and DNA damage suggests that this histone variant could play a role in the replacement of the histone with protamines during spermatogenesis (Figure 5.1).

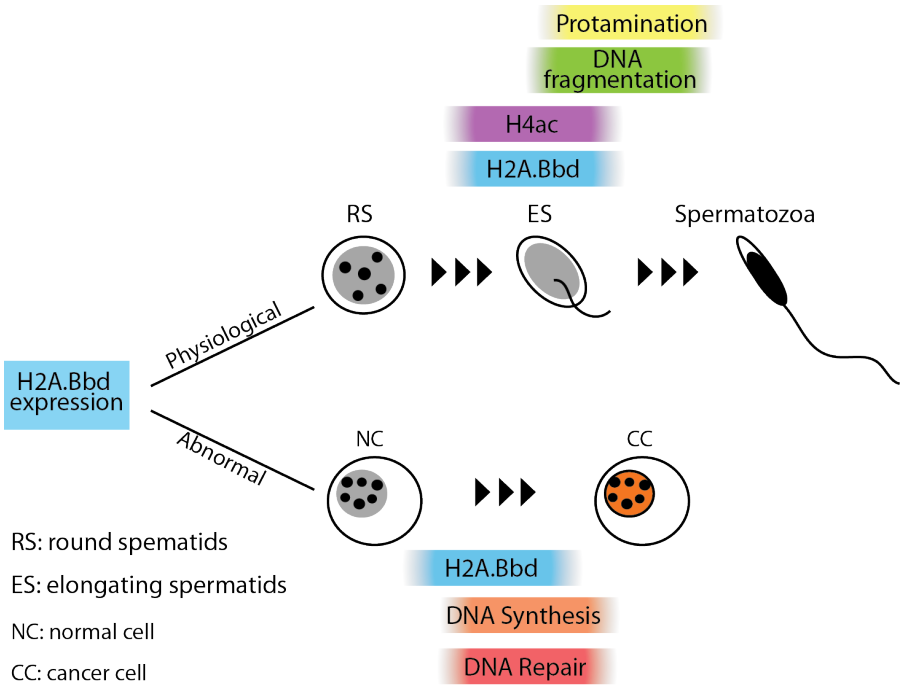

Figure 5.1 H2A.Bbd in spermatogenesis and cancer.
The histone variant H2A.Bbd is expressed in the elongating spermatids (ES) during spermatogenesis. The expression of H2A.Bbd precedes the acetylation of H4 and the DNA fragmentation, which are fundamental processes for the histone to protamine exchange (upper panel). The abnormal expression of H2A.Bbd in somatic cells may promote cancer development and progression by influencing DNA synthesis during DNA replication and repair (lower panel).

During mammalian spermatogenesis chromatin undergoes massive remodeling that leads to the formation of a very compacted chromatin structure required for proper spermatocytes formation and function (Govin et al., 2004). A key process in the context of spermatogenesis is the replacement of the canonical histones with protamine that ensure the proper level of chromatin compaction. Replacement of canonical histones by the protamines is not direct but involves a transition phase where several testis specific histone variants are recruited into chromatin (Govin et al., 2004). Immediately before replacement takes place, chromatin undergoes a massive wave of H4 acetylation (Meistrich et al., 1992). This event is considered to be the first signal for the subsequent chromatin remodeling. Moreover, during the latest stages of the spermatogenesis, a large number of DNA breaks are needed to ensure the massive chromatin remodeling and the transition from histones to protamines observed in the round spermatids (Marchetti and Wyrobek, 2008).
Due to the fact that GFP-H2A.Bbd is incorporated into chromatin after DNA damage (Figure 4.17) it is tempting to speculate that the DNA repair machinery in male germ cells could facilitate the incorporation of H2A.Bbd into chromatin thus promoting DNA recombination and protamine exchange.

Unfortunately, there are no studies available that explore the role H2A.Bbd in its natural context during spermatogenesis. A few studies focused on its putative homologs, the four mouse H2A.Bbd- like proteins called H2A.Lap (lack of acidic patch) 1-4 (Soboleva et al., 2011). Both, H2A.Bbd and its homologous in mouse H2A.Lap1, were shown to destabilize the nucleosome by only protecting 118 bp of DNA. In the context of spermatogenesis, the ability of the human H2A.Bbd and the mouse H2A.Lap1 to make the nucleosome less stable could be essential for the histone to protamine replacement. Indeed, H2A.Bbd and mouse H2A.Lap1 incorporation could favor chromatin de-compaction during the histone to protamine replacement. Moreover, H2A.Lap1 is enriched at the Transcription Start Sites (TSS) of active genes in spermatogenesis during meiosis and pot-meiotic stages (Nekrasov et al., 2013; Soboleva et al., 2011). The enrichment of H2A-Lap1 at the transcription start sites would suggest a potential role of this variant and, theoretically of H2A.Bbd, in activating specific genes during male germ cell line differentiation (Soboleva et al., 2011). However, it is important to mention that H2A.Lap1 and H2A.Bbd can not be totally functionally related as H2A.Lap1 gained one amino acid located in the H2A C-terminal tail α–helix responsible for the formation of the 30 nm fibres (Soboleva et al., 2011).

In this study, the most striking factor enriched in the GFP-H2A.Bbd containing chromatin is the Protein Phosphatase 1G (Ppm1G), also called Protein Phosphatase 2C isoform gamma (PP2C-gamma) or Fibroblast growth factor-inducible protein 13 (Fin13) (Table 4.1). Ppm1g was shown to have a high expression level in testis where H2A.Bbd is endogenously expressed and where the histone variant potentially plays a role (Guthridge et al., 1997). In addition, it was shown that Ppm1g regulates H2A/H2B exchange and there is evidence that this phosphatase could be involved in the regulation of the splicing process (Allemand et al., 2007; Kimura et al., 2006). Overall, these features make Ppm1g an interesting candidate for further analysis to understand its relation with H2A.Bbd.

Finally, the expression regulation of the testis-specific histone variants is essential for proper male germ cells differentiation. Here, it was shown that MEF cells overexpressing GFP-H2A.Bbd showed higher sensitivity to DNA damage (Figure 4.18). One possible hypothesis is that the overexpression of H2A.Bbd in germ cell could lead to higher DNA damage sensitivity or defect in chromatin compaction. A similar effect was already detected in sub-fertile patients where variable amounts of the testis-specific H2B histone variant (TSH2B) correlated with different level of chromatin compaction. Interestingly, the sub-fertile patients with higher levels of TSH2A showed s stronger Chromomycin A3 (CMA3) staining, which marks spermatozoa with defecting packaging (Singleton et al., 2007).
In summary, H2A.Bbd could be essential for normal germ cell differentiation and it would be interesting to understand whether the overexpression of H2A.Bbd in testis could have similar effects to what was observed here in MEF cells.

5.7 Abnormal expression of H2A.Bbd could have effects in cancer development

Due to their important role in influencing the nucleosome structure and chromatin accessibility, it is intuitive to hypothesize that the de-regulation of the histone variants expression could affect the cellular expression profile and interfere with proper cell cycle regulation.
In this respect, abnormal expression of several histone variants has been detected in different types of cancer. H2A.Z plays an important role in many cellular processes, such as DSBs repair, telomere integrity, genome stability and chromosome segregation. Its expression is de-regulated in colorectal, lung and bladder (Dunican et al., 2002; Rhodes et al., 2004; Zucchi et

al., 2004). H2A.X has been defined as a 'genome caretaker' and tumor suppressor in defined genetic contexts (Vardabasso et al., 2013). The role of macroH2A has already been discussed in this section. All these studies underline the importance of regulating histone variant expression. This PhD project focused on the effects of H2A.Bbd overexpression in MEF cells and Hodgkin's Lymphoma (HL) cells. Interestingly, the Hodgkin's Lymphoma cell line tested for the expression of H2A.Bbd showed an inverse correlation between the levels of expression and the doubling time. In other words, HL cells that express high levels of H2A.Bbd replicate faster (Figure 4.19).
Moreover, the MEF cells transfected with GFP-H2A.Bbd also showed higher sensitivity to DNA damage (Figure 4.18) and this correlates with the high sensitivity of the Hodgkin's lymphoma cells to radio- and chemotherapy that has led to highly efficient cancer treatments (Yeoh and Mikhaeel, 2011). All these findings suggest that the abnormal presence of H2A.Bbd in somatic cells could be fundamental for cancer progression and development.

It was previously discussed that one putative chaperone for H2A.Bbd could be the protein phosphatase PPM1G. This phosphatase is expressed in all tissues, but with higher levels in testis and it was shown to be a putative chaperone for H2A-H2B dimers (Kimura et al., 2006). PPM1g was found as the strongest interactor of the chromatin containing GFP-H2A.Bbd (Table 4.1 and Appendix Table III). It may well be that these two proteins interact in testis under physiological conditions and it is tempting to speculate that H2A.Bbd abnormally expressed in somatic cells is able to compete with H2A for the binding to PPM1g, leading to the preferential incorporation of H2A.Bbd-H2B dimers into chromatin.
In summary, this study revealed new properties of the testis specific histone variants H2A.Bbd and opened new questions about its role in spermatogenesis and cancer development (Figure 5.1). Future experiments are required to investigate the way in which H2A.Bbd is incorporated into chromatin, and to better understand the consequences of its incorporation in chromatin assembly and maturation after DNA replication and repair.

6 Materials and Methods

6.1 Materials

6.1.1 Technical devices

Description	Supplier
MoFlo cell sorter	Beckman Coulter
FACSCanto	Beckton Dickinson
Ultimate 3000 HPLC	LC Pachings Dionex
LTQ Orbitrap mass spectrometer	Thermo Fisher Scientific
LMS 510 confocal microscope	Zeiss
TCS SP5 confocal microscope	Leica

6.1.2 Antibodies

Description	Dilution	Supplier
a) Western Blot primary antibodies		
Mouse anti-GFP	1:1000	Roche
Rabbit anti-H3	1:5000	Abcam
Rabbit anti-H2A.Bbd	1:1000	Millipore
b) Western Blot secondary antibodies		
HRP-couples secondary antibodies	1:5000	
Licor anti-rabbit	1:10000	Licor
Licor anti-mouse	1:10000	Licor
c) Immunofluorescence primary antibodies		
anti-H3K27me3	1:1000	Gift of Gunnar Schotta
anti- GFP	1:200	Roche
d) Immunofluorescence secondary antibodies		
donkey anti-mouse Alexa 488	1:1000	Jackson Immuno Research
goat anti- rabbit Cy3	1:300	Millipore

6.1.3 DNA staining

Topro	1:1000	Invitrogen
DAPI		Vectashield

6.1.4 Kits and enzymes

Descriprion	**Supplier**
MNase	Sigma
Trypsin	Sigma
ProteinaseK	Gennaxxon
RNaseA	Roche
Click-it kit	Invitrogen
Benzonase	VWR

6.1.5 Cell culture media

DMEM	Sigma
RPMI 1640/GlutaMAX	Gibco
Penicillin/streptomicyn	PAA
Fetal calf serum	Sigma

6.1.6 Mass spectrometry material

0.2 ml tubes, strips of 8 (low binding)	Nunc
0.5 ml tubes (low binding)	Eppendorf LoBind
Acetonitrile	Sigma
Formic acid	Sigma
H_2O HPLC grade	Merk
Propionic acid	Merk
TFA	Merk

6.2 Methods

6.2.1 Cell biology methods

6.2.1.1 Cultivation of mammalian cells

Female Mouse Embryonic Fibroblasts cells (MEFs) were grown under standard conditions in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% Fetal Bovine Serum (Sigma) and 1x penicillin/streptomycin (PAA). Human Hodgkin's Lymphoma cell lines L-1236, L-540, L-428, HDLM-2, and KM-H2 were obtained from the Deutsche Sammlung für Mikroorganismen und Zellkulturen (Braunschweig, Germany). HL cell lines and control Epstein-Barr virus-immortalised cell lines (LCL) were cultured in RPMI 1640/GlutaMAX (Gibco) supplemented with 10% Fetal Bovin Serum (Sigma), betamercaptoethanol, and 1x penicillin/streptomycin (PAA). Peripheral blood mononuclear cells (PBMC) from healthy donors and patients with HL were isolated as described with informed consent and approval by the ethics committee of the Medical Faculty of the Martin Luther University Halle-Wittenberg.

6.2.1.2 Transfection

MEF cells were transfected with H2A.Bbd-GFP and GFP-H2A expression plasmids using FuGENE® HD transfection reagent (Roche) according to the manufacturer's instructions. Stable cell lines were selected using G418 (PAA) at a concentration of 600 µg/ml. The GFP-macroH2A expression plasmid was transfected using the TransIT LT1 reagent (Mirus) according to the manufacturer's instructions. For live cell imaging MEF cells stably expressing GFP-H2A.Bbd were transfected with mRFP-PCNA using Amaxa nucleofection as previously described.

6.2.1.3 FACS sorting and FACS analysis

GFP single cell sorting was carried out using a MoFlo cell sorter (Beckman Coulter). Single cell clone suspensions of GFP-H2A.Bbd, GFP-macroH2A and GFP-H2A MEFs were analysed using FACS Canto. The GFP signal was detected in the FITC channel and the data analysis to quantify the GFP positive cells was performed using FlowJo software.

6.2.1.4 Immunofluorescence

Cells grown on glass coverslips were fixed in PBS/3.7% FA (Roth:7398.1) 10 minutes at room temperature. After fixation the cells were washed in PBS and the permealized in 3.3

mM sodium citrate and 0.1% TRITON X-100, for 5 min at room temperature. The cells were then washed 2 times with PBS and 2 times with washing solution I (PBS-Tween (0.1%) plus 0.25% BSA). Blocking was performed with blocking solution I (5 ml washing solution I/2.5% BSA) for 30 min at room temperature. Primaries antibodies were diluted in the blocking solution I. Incubation with the primary antibody was performed overnight at 4°C. The day after the cells were washed with washing solution I and incubated with the seconday antibodies diluted in blocking solution II (blocking solution I plus 10% normal goat serum) for 1 hour at room temperature. Slides were mounted with Vectashield containing DAPI (Vectore Laboratories) and sealed with nail polish. Images were acquired using a Zeiss LSM 510 META confocal microscope equipped with a 63x oil immersion objective NA=1.4 or a Leica TCS SP5 confocal microscope with a 63x glycerol immersion objective NA = 1.3. Digital image processing, light linear adjustment of brightness and contrast was done using ImageJ.

The primary antibodies used were anti-H3K27me3 (kind gift of Gunnar Schotta) and anti-GFP (mixture of mouse monoclonal antibodies from Roche). For detection the cells were stained with donkey anti-mouse Alexa 488 (Jackson Immuno Research); goat anti-rabbit Cy3 (Millipore); TO-Pro-3 (Invitrogen). For live cell imaging MEF transfected cells were plated on a glass bottom p35 dish and grown under standard conditions. 3D stacks were acquired at 20-min intervals overnight. Time lapse imaging was carried out on a UltraVIEW VoX spinning disc confocal system (PerkinElmer, UK) in a closed live cell microscopy chamber (ACU control, Olympus, Japan) heated to 37°C, with 5% CO_2 and 60% air humidity control, mounted on a Nikon Ti microscope (Nikon, Japan), using a 60x/1.45 NA Planapochromat oil immersion objective lens. Images were obtained with a cooled 14-bit EMCCD camera (C9100-50, CamLink). Maximum intensity Z-projections were assembled into videos and annotated using ImageJ (http://rsb.info.nih.gov/ij/). For CPD analysis cells were grown on glass coverslips. Fixation of the cells was done with 3.7% formaldehyde for 10 min at room temperature followed by a combined permeabilization and DNA denaturation step with ice cold 0.7% Triton X-100/ 0.4 N HCL/ PBS for 10 min. For blocking step, the cells were incubated with 4% BSA for 30 min. Immunostaining for CPD was performed by incubating the cells with monoclonal anti-CPD (Kamiya Biomedical Company, Cat no: MC-062) at a dilution of 1:200 for 1 hour at room temperature. For detection, cells were incubated with polyclonal anti mouse IgG Cy3 (Jackson Laboratories, Cat no: 715-165-151) at a dilution of 1:330 for 1 hour at room temperature. The DNA content was stained with 4',6-diamidino-2-phenylindole (DAPI) and finally mounted in Mowiol 4-88 (Sigma-Aldrich Chemie).

6.2.1.5 In situ replication labeling

To visualize newly synthesized DNA, cells grown on coverslips were pulse labeled with either 10 µM EdU (Fig. 4) or 100 µM 5-bromo-2′-deoxyuridine (BrdU, Sigma-Aldrich) (Fig. 5) and either fixed immediately or chased for the indicated time in pre-warmed medium supplemented with 200 µM thymidine. Cells were fixed with 3.7 % formaldehyde for 10 minutes and permeabilized with 0.5 % Triton X-100 for 20 minutes. BrdU detection was performed with rat-anti-BrdU (Gentaur, Cat no: OBT0030CX, dilution 1:100) in the presence of DNaseI as described (Casas-Delucchi et al.,2012). Click reaction was performed using the click-it kit (Invitrogen). Alexa Fluor 647 azide was added in a final concentration of 2 µM. The click reaction was performed for 30 minutes at room temperature. The cells were then stained with anti-GFP as described above.

6.2.1.6 Microirradiation experiments

Microirradiation experiments were carried out in Nikon Ti spinning disc microscope (see live cell imaging) with a 405 nm diode laser set to 100% transmission. Approximately 1 µm spots were selected inside the nucleus and microirradiated with 1 mJ. Pre and post bleach images were acquired as indicated.

6.2.1.7 UV-C irradiation experiments

For UV-C irradiation experiments, the cells were irradiated with a UV-C source containing 5.8 W short wave lamps (Amersham Biosciences) at a wavelength of 254 nm. For damage induction the media was removed from the cells and were globally irradiated for 10 seconds with a final dose of 1 mJ/cm^2. Unirradiated cells were used as control.

6.2.1.8 CPD Quantification

The levels of DNA damage were calculated from confocal 3D images of CPD stainings and DNA counterstained with DAPI using the image analysis platform Priithon (http://code.google.com/p/ priithon/). The DAPI images were used to create 3D nuclear masks by applying a median filter (sigma = 500 nm) and segmenting the filtered images using automatic thresholding. The integrated CPD signal within the nuclear mask was divided by the mask volume to obtain the mean CPD intensity in each nucleus.

6.2.2 Molecular biology methods

6.2.2.1 Quantitative reverse transcription-polymerase chain reaction

RNA from HL cell lines and PBMC was isolated using Trizol reagent (Invitrogen, Karlsruhe, Germany) following the manufacturer's protocol. After reverse transcription of 2 µg of RNA, quantitative reverse transcription-polymerase chain reaction (RT-PCR) was performed essentially as described (Neumann et al., 2010). The following primer combinations have been used: actin beta (ACTB): 5'-ggc atc gtg atg gac tcc g-3' and 5'-gct gga agg tgg aca gcg a-3'; H2A.Bbd: 5'-tcg ttt tca gtg agc cag gt-3' and 5'-cag aat taa tga agg ccc aag-3'. Quantification of gene expression was performed using the $2^{-\Delta\Delta Ct}$ method [H] and for comparative analysis, the median Dct value from PBMC (21 independent donors) was set as one.

6.2.3 Biochemical methods

6.2.3.1 Chromatin purification

Adherently growing mouse embryonic fibroblasts (MEFs) were detached by trypsinisation followed by centifugation at 200 g. Cells were washed twice with PBS and nuclei were isolated by the addition of ice-cold 0.3% Triton-X dissolved in PBS. After a 10 min incubation step at 4°C nuclei were pelleted for 10 min at 1000 xg, washed once with PBS and resuspended in EX-100 (100 mM Hepes pH 7.6); 100 mM NaCl; 1.5 mM $MgCl_2$; 0.5 mM EGTA; 10% v/v Glycerol) freshly supplemented with proteinase inhibitors. Chromatin was solubilized by 20 min Microccocal Nuclease (MNase) digestion at 26°C in the presence of 2 mM $CaCl_2$. The reaction was stopped by adding 1/50 th volume of 0.5 M EGTA. All further steps were carried out at 4°C. The first supernatant (S1) was recovered after centrifugation at 1000xg and the pellet was resuspended in resuspension buffer (1x PBS; 150 mM NaCl; 2 mM EDTA; 0.1% Triton) and over-head rotated over night (o/n) to further extract solubilized chromatin fragments. A second supernatant (S2) was recovered by a 30 min centrifugation, pooled with the first supernatant to yield the purified chromatin. A chromatin aliquot was used to check the MNase digestion by isolating the DNA fragments obtained.

6.2.3.2 GFP affinity purification

MNase digested chromatin was precleared with 20 µl of slurry agarose beads for 1 h at 4°C. The supernatant was collected after 2 minutes of centrigugation at 2000xg, 4°C. 20 µl of

slurry GFP binders (Chromotek) were equilibrated 3 times with EX-100 (100 mM Hepes pH 7.6; 100 mM NaCl; 1.5 mM $MgCl_2$; 0.5 mM EGTA; 10% v/v Glycerol) and then added to the supernatant. Immunoaffinity purification was performed for 2 h at 4°C. The beads were washed 3 times with Resuspension Buffer (1x PBS; 150 mM NaCl; 2 mM EDTA; 0.1% Triton) plus proteinase inhibitors and centrifuged at 4°C at 2000xg. 15 µl of 2X Laemmli buffer were added to the beads and boiled 15 minutes at 95°C. Samples were run on a 15% SDS gel for 2 cm and stainded with Coomassie Brilliant Blue G250 for 20 minutes. After destaining with 10% acetic acid every lane was cut into 8 slices and every slice was collected into a tube.

6.2.3.3 Trypsin in-gel digestion and mass spectrometry

Gel pieces were washed once with milliQ water and two times with NH_4HCO_3 20 mM. Destain of the gel pieces was preformed in 50% ACN 50% NH_4HCO_3 at 37°C for 1 hour. Samples were washed two times with 20 mM NH_4HCO_3 and dehydrated with ACN ultrapure. Reduction was performed in 10 mM DTT diluted in 20 mM NH_4HCO_3, 1 hour at RT. Alkylation of the gel pieces was done in 50 mM IAA diluted in NH_4HCO_3, 30 minutes at RT in the dark. Before tryptic digestion the samples were washed once with 20 mM NH_4HCO_3 and dehydrated with 100% ACN. 30 to 50 µl of 25 ng/µl Trypsin were added to the gel pieces in ice till the gel pieces hydrated completely. Non absorbed trypsin was removed and 50 µl of 20 mM NH_4HCO_3 were added. The samples were incubated overnight at 37°C. The day after the peptides were collected and placed in a fresh tube and acid extraction was performed as follow. The gel pieces were incubated twice in 50% ACN- 0.25% TFA, 10 minutes at room temperature followed by two times 100% ACN, 10 minutes at room temperature. The peptides were collected and pooled together with the first supernatant. The samples were dried and then resuspended in 15 µl 0.1% FA. 50% of the sample was injected in an Ultimate 3000 HPLC system (LC Packings Dionex). Samples were desalted on-line by a C18 micro column (300 mm i.d. 5 mm, packed with C18 PepMapTM, 5 mm, 100 Å by LC Packings), and peptides were separated with a gradient from 5% to 60% acetonitrile in 0.1% formic acid over 40 min at 300 nl/min on a C18 analytical column (75 mm i.d.x10 cm, packed with C18 PepMapTM, 3 mm, 100 Å by LC Packings). The effluent from the HPLC was directly electrosprayed into the LTQ Orbitrap mass spectrometer (Thermo Fisher Scientific). The MS instrument was operated in the data-dependent mode to automatically switch between full scan MS and MS/MS acquisition. Survey full scan MS spectra (*m/z* 580-830) were acquired in the Orbitrap with resolution R= 60 000 at *m/z* 400. The six most intense peptide ions with

charge states between two and five were sequentially isolated (window = 350-2000 *m/z*) to a target value of 10000 and fragmented in the linear ion trap by collision-induced dissociation (CID). Fragment ion spectra were recorded in the linear trap of the instrument. For all measurements with the Orbitrap detector, three lock-mass ions from ambient air (m/z=371.10123, 445.12002, 519.13882) were used for internal calibration as described. Typical mass spectrometric conditions were: spray voltage, 1.4 kV; no sheath and auxiliary gas flow; heated capillary temperature, 200°C; normalized collision energy, 35% for CID in linear ion trap. The ion selection threshold was 10000 counts for MS2. An activation q=0.25 and activation time of 30 ms were used.

6.2.3.4 Mass spectrometry data analysis

For protein identification the raw data were analysed with the Andromeda algorithm of the MaxQuant protein analysis package (version 1.2.2.5) against the IPI mouse v.3.68.Fasta including reverse sequences and contaminants. The Trypsin/P enzyme was selected, allowing for maximally 2 missed cleavages. Carbamidomethylation of C was set as fixed modification whereas methionine oxidation and protein N-acetylation were included as variable modifications. The mass tolerance of the initial search was 20 ppm; after recalibration, a 6 ppm mass error was applied for the main search. Fragment ions were searched with a mass offset of 0.5 Da using the 6 most intense signals per 100 Da. Searching for secondary peptide hits within already assigned MS/MS spectra was enabled. The search results were filtered with a peptide and protein false discovery rate of 0.01 with a minimum peptide length of 6 amino acids. Protein identifications with at at least one unique or razor peptide were accepted. For the quantification the Intensity Based Absolute Quantification (iBAQ) values were calculated from peptide intensities and the protein sequence information (Schwanhäusser et al., 2011) of unmodified, M/oxidated and acetylated peptide species with a minimum of 2 peptides per protein.

6.2.3.5 Statistical data analysis

Statistical analysis of protein enrichment was performed from the three affinity purification-MS replicates of the three GFP-H2A variants. First, all replicates were normalized using the median of abundant histone H4 peptides, including only peptides without modifications, less than 2 missed cleavages and excluding tryptic cleavage after Pro, as normalization factor. The control samples (NoGFP) data were not normalized, since detected H4 levels were to low. Secondly, normalized iBAQ values were log2 transformed and missing values were imputed

from a gaussian distribution which was shifted by a difference of about 2.5 and had a width of 0.3 using the Perseus software tool (Cox and Mann, 2012) The average difference of the GFP-H2A variants over the control samples (NoGFP) was calculated and only proteins having a difference of more than 3.5 and quantiative information in at least two experiments were kept. To determine enrichment factors and p-values for enrichment, ANOVA was perfromed in DanteR and resulting p-values were subsequently corrected for multiple hypothesis testing using the Benjamini-Hochberg method in DanteR (Benjamini and Hochberg, 2008). Data were plotted with Perseus software. The Gene Ontology was performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) version 6.7. (Huang et al., 2009a; 2009b). Functional Annotation Chart were created using default settings.

6.2.3.6 Histone preparation for mass spectrometry

For histone modifications analysis, histones were separated by 15% SDS–PAGE and stained with Coomassie blue for 20 minutes at RT. After destaining with 10% acetic acid, the histone bands for H3 and H4 were excised from the gel and placed in a 0.5ml low binding tube (Eppendorf LoBind). The pieces were washed once with 200 µl milliQ water (Merk) and once with 200 µl 10 mM NH_4HCO_3 for 5 min. at 37°C. The destain was obtained incubating the gel pieces in 200 µl of 50%ACN/50 mM NH_4HCO_3 for 30-60 min. at 37°C. In case of partial destaining, the step was repeated again once. At this point the protocol can be stopped and the gel pieces can be stored in milliQ water at 4°C overnight. Histone lysine acylation was perfomed adding 1 µl of propionic anhydride in the sample tubed followed by 10 µl of 0.1 M NH_4HCO_3. Due to the production of CO_2 the tubes were opened and closed several time after the reaction started. The reaction was performed adding 1 M NH_4HCO_3 to cover the gel pieces. The reaction was perfomed at 37°C for 30-60 min. To prevent the aspecific acrylation the pH of the reaction was monitored and kept between 7-8. After removing the supernatant the gel pieces were washed three times with 0.1M NH_4HCO_3 for 10 min. at room temperature followed by one time with milliQ water and one time with 50% ACN for 15 min. at room temperature. To avoid aspecific N-terminal acylation the washing steps were never shorter than the time requested. The dehydration of the gel pieces is perfomed by washing the gel pieces two times with 100% ACN on ice. Trypsin digestion was perfomed using 0.2 ng of trypsin. The trypsin absorbtion in the gel pieces was obtained using 10 µl of 50 mM NH_4HCO_3 in ice for 5 min. After trypsin absorbion the gel pieces were incubated in 30 µl 50 mM NH_4HCO_3 overnight at 37°C. The day after the peptides released in the supernatant were collected in a fresh tubes and 1/3 of the sample was used for mass spectrometry

measurament. The method applied for the LC-MS/MS is the same as described in the proteomic analysis of the purified chromatin.

6.2.3.7 Histone modifications data analysis

To determine all modifications occurring on histones the spectra were processed and analyzed with the XCalibur Qual Browser software (Thermo Fisher Scientific). Doubly and triply charged peptides were quantified by using the masses of the corresponding extracted ion chromatograms (XICs). The sum of the area from all peaks derived from a single peptide was defined as 100% and the relative percentage for a single modification was calculated accordingly. All the calculations were performed manually in Excel.

6.2.3.8 Nuclear extract for western blot analysis

GFP-histone variants transfected MEFs and Hodgkin's Lymphoma cell lines were incubated with 0.1% Triton in PBS supplemented with proteinase inhibitor for 10 minutes at 4°C. After centrifugation at 3000 rpm the nuclei pellet was resuspended in 1X PBS supplemented with proteinase inhibitors and the DNA was digested with Benzonase (VWR) for 15 minutes at 37°C. 2X Laemmli buffer was added to the nuclei suspension and boiled 15 minutes at 95°C

6.2.3.9 Western Blot antibodies

anti-GFP antibodies: a mixture of two mouse anti-GFP monoclonal antibodies (Roche) was used at a dilution of 1:1000, the H3 antibody (Abcam ab1719) was used at a concentration of 1:5000, H2A.Bbd 1:1000 (Millipore polyclonal). HRP-coupled secondary antibodies were used at a concentration of 1:5000 and detection was performed using ECL chemiluminescence.

7 Abbreviations

3C	chromosome conformation capture
4-HB	4-helix bundle
BMI1	Polycomb Complex Protein 1
Bp	basepair
CAF1	Chromatin Assembly Factor 1
CBX8	Chromobox protein homolog 8
CENP-A	Centromeric Protein A
ChIP	chromatin immunoprecipitation
Cid	Centromeric Identifier
CPD	cyclobutane pyrimidine dimer
CTA	Cancer Testis Antigens
DAPI	4',6-diamidino-2-phenylindole
DIPG	diffuse intrinsic pontine glioma
DNA	deoxyribonucleic acid
DSB	double strand breaks
EBV	Epstein-barr virus
EdU	5-ethynyl-2′-deoxyuridine
EM	electron microscopy
EMANIC	electron microscopy-assisted nucleosome capture
EMSA	electrophoretic mobility shift assay
FACS	fluorescence-activated cell sorting
FACT	facilitated chromatin transcription complex
FISH	fluorescence in situ hybridization
FRAP	fluorescent recovery after photobleaching
GBM	glioblastoma multiforme
GFP	Green Fluorescent Protein
GO	Gene Ontology

h	hour
H2A.Bbd	H2A Barr body deficient
H3/H4/H2A/H2B/H1	histones
HCC	hepatocellular carcinoma
HDAC1/2	histone deacetylase 1/2
HL	Hodgkin's Lymphoma
HP1	heterochromatin protein 1
HZH2	Enhancer of zeste homolog 2
iBAQ	intensity based absolute quantification
K	lysine
kDa	kilodalton
L1	loop1
L2	loop2
LAD	lamina-associated-domains
LCL	lymphoblastoid cell lines
Mb	megabase
MCM	mini chromosome maintenance
me	methylation
MEFs	Mouse Embryonic Fibroblasts
Mg	magnesium
mM	millimolar
mRNA	messenger RNA
NaCl	Sodium Chloride
NAP1	Nucleosome assembly protein 1
NCP	nucleosome core particle
NHR	non-histone region
nm	nanometer
PARP1	poly-ADP polymerase 1
PcG	Polycomb group

PCNA	proliferating cell nuclear antigens
PRC1/2	Polycomb repressive complex 1/2
qPCR	quantitative polymerase chain reaction
S-phase	Synthesis phase
SPOP	speckle-type POZ domain protein
SUV39	suppressor of variegation 3-9
TH2B	testis specific H2B
TSS	transcription start site

8 Acknowledgement

First of all I would like to thank Prof. Dr. Axel Imhof for giving me the opportunity to work in his group and for his scientific supervision and support.

I thank Prof. Dr. Peter Becker for providing and excellent and motivating scientific environment at the Adolf Butenandt Institute.

A special thank goes to all the member of my group and in particular to Dr. Andreas Thomae (Postdoc!) for being always there in case of a 'scientific emergency' and for being always willing to share his great knowledge with me.

I also thank Irene Vetter, for her great technical support and for her tutorials about the Bavarian culture.

I would like to truly thank Teresa, Simo and Miri, for sharing the last five years of PhD together in 'good and bad times', but most of all for their friendship and constant help in understanding the German world.

Ringrazio Alessandra e Dario, per aver portato un po' di casa ogni giorno in laboratorio tra risate, battute e caffe,' anche nei giorni piu' freddi e cupi.

Ringrazio i miei amici lontani, in particolare Antonia ed Elisabetta, per essermi rimaste vicine e per aver continuato a condividere la propria vita con me quotidianamente, magari durante una improbabile cena virtuale.

Infine, un ringraziamento speciale va ad Enza, Sandro e Serena, per avermi incoraggiato a seguire i miei sogni e le mie aspirazioni nonostante queste mi abbiano portato lontano da casa. Grazie davvero!

9 Curriculum vitae

Personal information

Name	Viola Sansoni
Date of birth	28/02/1983
City of birth	Priverno (Italy)
Nationality	Italian

Education

04/2009- 07/2014	**PhD student**
	Adolf Butenandt Institute, LMU, Munich, Germany
	Prof. Dr. Axel Imhof
09/2006- 10/2008	**Master Degree in Molecular Biology**
	University La Sapienza, Rome, Italy
09/2001- 07/2006	**Bachelor Degree in Biology**
	University La Sapienza, Rome, Italy
09/1996- 07/2001	**Liceo Classico** 'Dante Alighieri', Anagni, Italy

Publications

Viola Sansoni, Corella S. Casas-Delucchi, Malini Rajan, Andreas Schmidt,Clemens Boenisch, Andreas W. Thomae, Martin S. Staege, Sandra B. Hake, M.Cristina Cardoso and Axel Imhof The histone variant H2A.Bbd is enriched at sites of DNA synthesis, *Nucleic Acid Research*, 2014

Iacobini C, Menini S, Ricci C, Scipioni A, Sansoni V, Cordone S, Taurino M, Serino M, Marano G, Federici M, Pricci F, Pugliese G., Arteriosclerosis Thrombosis Vascular Biology. Accelerated lipid-induced atherogenesis in galectin-3-deficient mice: role of lipoxidation via receptor-mediated mechanisms, *Atherosclerosis Thrombosis Vascular Biology*, 2009

Iacobini C, Menini S, Ricci C, Scipioni A, Sansoni V, Mazzitelli G, Cordone S, Pesce C, Pugliese F, Pricci F, Pugliese G. Advanced lipoxidation enr-products mediate lipid- induced glomerular injury: role of receptor-mediated mechanisms, *Journal of Pathology*, 2009

10 Appendix

10.1 Table I. Gene Ontology analysis of proteins enriched in the GFP-H2A.Bbd over the GFP-macroH2A.1.2 containing chromatin

Term	Count	%	PValue	Bonferroni	Benjamini	FDR
RNA processing	100	17.5	3.32E-59	3.71E-56	3.71E-56	5.33E-56
translation	75	13.2	1.35E-44	1.51E-41	7.55E-42	2.17E-41
mRNA processing	65	11.4	5.53E-40	6.17E-37	2.06E-37	8.86E-37
ribosome biogenesis	45	7.9	1.22E-37	1.36E-34	3.40E-35	1.95E-34
RNA splicing	56	9.8	3.24E-37	3.62E-34	7.23E-35	5.19E-34
mRNA metabolic process	66	11.6	5.59E-37	6.24E-34	1.04E-34	8.96E-34
ribonucleoprotein complex biogenesis	47	8.2	1.10E-35	1.23E-32	1.76E-33	1.76E-32
rRNA processing	29	5.1	9.06E-24	1.01E-20	1.26E-21	1.45E-20
rRNA metabolic process	29	5.1	1.40E-23	1.57E-20	1.74E-21	2.25E-20
chromatin organization	47	8.2	5.21E-19	5.82E-16	5.82E-17	8.35E-16
chromosome organization	53	9.3	6.67E-19	7.44E-16	6.77E-17	1.07E-15
ncRNA processing	31	5.4	7.42E-16	8.67E-13	7.23E-14	1.24E-12
ncRNA metabolic process	31	5.4	6.50E-13	7.25E-10	5.58E-11	1.04E-09
DNA replication	27	4.7	8.31E-13	9.28E-10	6.63E-11	1.33E-09
chromatin modification	33	5.8	1.38E-12	1.54E-09	1.03E-10	2.22E-09
DNA metabolic process	44	7.7	3.61E-12	4.03E-09	2.52E-10	5.78E-09
covalent chromatin modification	20	3.5	7.62E-11	8.50E-08	5.00E-09	1.22E-07
histone modification	19	3.3	2.95E-10	3.30E-07	1.83E-08	4.73E-07
chromatin assembly or disassembly	20	3.5	7.51E-10	8.38E-07	4.41E-08	1.20E-06
cellular macromolecular complex assembly	25	4.4	5.55E-08	6.20E-05	3.10E-06	8.89E-05
histone acetylation	11	1.9	6.14E-08	6.86E-05	3.27E-06	9.84E-05
protein amino acid acetylation	11	1.9	8.28E-08	9.24E-05	4.20E-06	1.33E-04
cellular macromolecular complex subunit organization	25	4.4	5.41E-07	6.04E-04	2.63E-05	8.67E-04
transcription	91	16.0	6.46E-07	7.21E-04	3.00E-05	0.001
protein amino acid acylation	11	1.9	1.53E-06	0.002	6.84E-05	0.002
histone H4 acetylation	8	1.4	1.95E-06	0.002	8.38E-05	0.003

10.2 Table II. Gene Ontology analysis of proteins enriched in the GFP-H2A.Bbd over the GFP-H2A containing chromatin

Term	Count	%	PValue	Bonferroni	Benjamini	FDR
RNA processing	90	18.6	1.61E-55	1.66E-52	1.66E-52	2.56E-52
translation	73	15.1	7.94E-48	8.15E-45	4.07E-45	1.26E-44
mRNA processing	60	12.4	5.30E-39	5.43E-36	1.81E-36	8.40E-36
RNA splicing	53	11.0	1.18E-37	1.21E-34	3.02E-35	1.87E-34
mRNA metabolic process	61	12.6	2.38E-36	2.44E-33	4.88E-34	3.77E-33
ribosome biogenesis	40	8.3	6.15E-34	6.31E-31	1.05E-31	9.75E-31
ribonucleoprotein complex biogenesis	41	8.5	2.40E-31	2.46E-28	3.52E-29	3.80E-28
rRNA processing	24	5.0	3.70E-19	3.80E-16	4.75E-17	5.87E-16
rRNA metabolic process	24	5.0	5.21E-19	5.35E-16	5.94E-17	8.26E-16
ncRNA processing	26	5.4	3.25E-13	3.34E-10	3.34E-11	5.16E-10
ncRNA metabolic process	26	5.4	8.46E-11	8.68E-08	7.89E-09	1.34E-07
chromatin organization	31	6.4	7.43E-10	7.62E-07	6.35E-08	1.18E-06
chromosome organization	34	7.0	5.25E-09	5.39E-06	4.14E-07	8.32E-06
DNA replication	20	4.1	1.39E-08	1.42E-05	1.02E-06	2.20E-05
DNA metabolic process	31	6.4	5.34E-07	5.48E-04	3.66E-05	8.47E-04
chromatin modification	22	4.5	8.78E-07	9.01E-04	5.63E-05	0.001
cellular macromolecular complex assembly	21	4.3	9.26E-07	9.50E-04	5.59E-05	0.001

10.3 Table III. List of proteins enriched in the GFP-H2A.Bbd compared to the GFP-macroH2A.1.2 and GFP-H2A containing chromatin.

Gene name	IPI number	Uniprot	bbd vs macro	pvalue	log2 bbd vs H2A	pvalue
Fin13;Ppm1g	IPI00117072	Q61074	17.53	0.00048	18.90	0.00045
Rpl19	IPI00122426	P84099	16.95	0.03664	4.65	0.62298
Rpl36	IPI00869475	Q5M9L1	16.64	0.00319	6.79	0.07355
Rpl35	IPI00263879	Q6ZWV7	16.42	0.04882	3.72	0.73830
Rplp2	IPI00139795	P99027	15.51	0.06783	5.54	0.59351
Csda;Msy4	IPI00274739	Q9JKB3-2	15.10	0.00668	3.72	0.39704
Rps24	IPI00465568	P62849-1	14.86	0.00790	5.98	0.17831
Rpl22	IPI00222546	P67984	14.78	0.00278	4.27	0.17057
Rpl34	IPI00466153	Q9D1R9	14.76	0.00429	3.36	0.36442
Rplp1	IPI00113377	P47955	14.73	0.07042	4.92	0.62751
Rps19	IPI00875584	Q5M9P3	13.22	0.02496	4.31	0.47935
Rpl24	IPI00323806	Q3UW40	13.20	0.07783	3.98	0.69359
Ddx49	IPI00354271	Q4FZF3	13.02	0.00013	13.34	0.00014
Purb	IPI00128867	O35295	12.69	0.00790	5.68	0.13861
Top1	IPI00109764	Q04750	12.63	0.01648	1.80	0.77044
1810009A15Rik	IPI00112004	Q9CSS6	12.54	0.00649	4.01	0.24924
Npm1	IPI00127415	Q61937	12.44	0.12993	3.14	0.80998
Brix;Brix1	IPI00130436	Q9DCA5	12.40	0.00964	12.51	0.01255
Ebna1bp2	IPI00111829	Q9D903	12.33	0.04950	11.95	0.07369
Phf14	IPI00415328	Q9D4H9-2	12.23	0.00631	5.06	0.12589
Caper	IPI00223371	Q8VH51-2	12.22	0.02665	2.65	0.68362
Nol5	IPI00463468	Q6DFW4	12.11	0.14087	2.63	0.84404
Cenpa	IPI00128526	O35216	12.05	0.00011	12.31	0.00009
Kiaa1820	IPI00890011	Q61818-2	11.98	0.00017	6.32	0.00304
Eif6	IPI00857575	A6PWZ2	11.94	0.01036	11.56	0.01480
Unnamed protein	IPI00626366		11.89	0.04653	10.93	0.08102
Ccdc86	IPI00402914	Q9JJ89	11.84	0.00362	2.28	0.43004
Gnl3	IPI00222461	Q8CI11-1	11.80	2.71E-09	5.64	1.13E-07
Top2b	IPI00135443	Q64511	11.77	0.00470	3.92	0.18894
Nol1;Nop2	IPI00311453	Q922K7	11.63	0.00668	1.66	0.67876
Rsl1d1	IPI00226149	Q3TAJ5	11.56	0.14181	3.07	0.80492
Llrep3;Rps2	IPI00880400	P25444	11.48	0.22453	3.82	0.79884
Mbd2	IPI00131088	Q9Z2E1-1	11.14	0.02496	5.89	0.21179
Llph	IPI00112043	Q9D945-1	11.12	0.02191	8.24	0.07369
Ddx50	IPI00117771	Q99MJ9	11.11	0.01057	2.56	0.52428
C7orf50 homolog	IPI00110456	Q9CXL3	11.07	0.01591	1.70	0.74887
Igf2bp3	IPI00331315	Q9CPN8	10.91	0.09223	9.87	0.17057
Polr2g	IPI00263106	P62488	10.89	0.00019	11.48	0.00019
Atrx;Hp1bp2	IPI00322707	Q61687	10.84	0.00429	2.27	0.41933
Ssr1	IPI00110852	Q52PE3	10.80	0.06411	7.73	0.21252
Polr2h	IPI00124284	Q923G2	10.62	0.08410	9.20	0.17594

Rpl14	IPI00473728	Q9CWK0	10.61	0.13268	3.85	0.69893
Nhp2	IPI00133550	Q9CRB2	10.58	0.23076	3.23	0.82172
P198	IPI00223217	P19253	10.57	0.05072	3.31	0.61120
C3orf26 homolog	IPI00113109	Q9CZT6	10.53	0.06089	7.55	0.20667
Polr2d	IPI00268215	Q9D7M8	10.38	0.00261	8.82	0.00585
Dnaj2	IPI00132208	P63037	10.29	0.00012	10.26	0.00013
Polr2c	IPI00875500	Q99M46	10.28	0.08328	2.59	0.75624
Rpl36a	IPI00225066	P83882	10.23	0.09990	2.13	0.82373
Ddx18	IPI00459381	Q8K363	10.22	0.00221	3.25	0.09561
Igf2bp1	IPI00131056	O88477	10.21	0.02968	8.81	0.06492
Zan75	IPI00314507	O88291-1	10.20	0.07873	7.83	0.21252
Ssrp1	IPI00407571	Q08943-2	10.13	0.00919	4.31	0.17831
Zfr	IPI00131810	O88532	10.13	0.03028	5.64	0.21252
Mrt4	IPI00132578	Q9D0I8	10.05	0.18469	6.79	0.47935
Ddx6	IPI00109932	P54823	10.00	0.00278	8.38	0.00660
Bxdc1	IPI00121159	Q9JJ80	9.99	0.09235	7.88	0.23169
MNCb-2643	IPI00318725	Q9CYH6	9.98	0.03183	8.78	0.06558
Ssr3	IPI00120826	Q9DCF9-1	9.98	0.00820	9.72	0.01255
D19Bwg1357e	IPI00222675	Q8BKS9	9.96	0.01036	4.45	0.17594
Snrnp70	IPI00625105	Q62376-1	9.96	0.02496	9.13	0.04460
Pura	IPI00118447	P42669	9.95	0.09372	4.17	0.58358
A301	IPI00123005	Q810D6	9.86	0.00002	7.70	0.00007
Mta1l1	IPI00128230	Q9R190	9.86	0.01288	2.87	0.43117
Fact140	IPI00120344	Q920B9	9.86	0.01665	5.35	0.14979
D13Wsu177e	IPI00458704	Q9CPT5	9.85	0.06089	6.43	0.24924
Ssr4	IPI00953773	Q9D6F7	9.83	0.09842	10.00	0.13273
Rpl10a	IPI00849927	Q3U561	9.82	0.21551	5.03	0.64107
Ddx47	IPI00187240	Q9CWX9	9.81	0.13911	6.21	0.44546
Npm3	IPI00131725	Q9CPP0	9.81	0.03466	12.23	0.02281
Crfg;Gtpbp4	IPI00117642	Q99ME9	9.75	0.00877	9.79	0.01209
Mnar;Pelp1	IPI00321597	Q9DBD5	9.70	0.02431	0.97	0.86489
Meaf6	IPI00110388	Q2VPQ9-2	9.69	0.02298	-0.52	0.93456
Mta1	IPI00853911	Q8K4B0	9.67	0.00746	3.20	0.25882
Spbp	IPI00407458	Q9EPQ8-1	9.65	0.00558	4.81	0.06831
Exosc6;Mtr3	IPI00466593	Q8BTW3	9.61	0.00790	8.92	0.01359
Rpl8	IPI00137787	P62918	9.60	0.16309	1.86	0.86613
Nhp2l1;Ssfa1	IPI00621272	Q9D0T1	9.59	0.19692	2.86	0.81243
Rpl35a	IPI00115902	O55142	9.58	0.15344	2.51	0.81407
H1fx	IPI00118590	Q80ZM5	9.56	0.04263	2.11	0.73542
Srpk1	IPI00387234	O70551	9.55	0.00406	8.32	0.00879
Rpl22l1	IPI00110724	Q9D7S7-1	9.54	0.01288	10.96	0.01160
Ddx1	IPI00127172	Q91VR5	9.52	0.03119	9.75	0.03934
Rfc4	IPI00653266	Q3UI84	9.51	0.07626	8.25	0.15511
Ck2n;Csnk2b	IPI00126762	P67871	9.43	0.10561	2.84	0.73840
Glyr1;Np60	IPI00817029	Q922P9	9.43	0.00548	-0.64	0.85691
Pabp1	IPI00124287	P29341	9.39	0.03061	2.63	0.59164

Rpl5	IPI00308706	P47962	9.35	0.39845	5.48	0.73740
Polr2b;rpB2	IPI00320034	Q8CFI7	9.35	0.07344	5.37	0.35136
Gclm;Glclr	IPI00114329	O09172	9.33	0.01036	8.89	0.01578
Hmg20a;Ibraf	IPI00120110	Q9DC33-1	9.29	0.00048	8.43	0.00095
Abcf2	IPI00116825	Q99LE6	9.22	0.00744	9.43	0.00948
Kiaa0461	IPI00227539	Q8BZH4	9.21	0.00261	1.98	0.29768
Nip7	IPI00133555	Q9CXK8	9.20	0.24776	2.77	0.82784
Dpy30	IPI00117168	Q99LT0	9.19	0.03263	8.29	0.06486
Las1l	IPI00462502	A2BE28	9.16	0.02365	8.47	0.03974
Yeats4	IPI00132946	Q9CR11	9.16	0.00820	9.87	0.00948
Arf1	IPI00221613	P84078	9.15	0.19811	4.40	0.65766
Rpl18a	IPI00162790	P62717	9.13	0.15286	3.85	0.66583
Csnk2a2	IPI00118795	O54833	9.12	0.04619	7.77	0.10184
2700060E02Rik	IPI00132456	Q9CQE8	9.11	0.11903	5.52	0.43494
Nhn1;Zc3h18	IPI00673693	Q0P678-1	9.08	0.00011	3.74	0.00409
Farp1	IPI00757373	C4IXU2	9.06	0.00099	9.23	0.00107
mCG_1675;Sfrs6	IPI00310880	Q3TWW8	9.06	0.09853	2.30	0.77513
Znhit1	IPI00857945	Q8R331	9.05	0.00048	7.74	0.00107
Pcna	IPI00113870	P17918	9.03	0.00416	9.26	0.00555
Krip1;Trim28	IPI00312128	Q62318-1	8.96	0.01128	0.23	0.96168
Exosc4;Rrp41	IPI00310857	Q921I9	8.95	0.00559	7.67	0.01176
Rpl32	IPI00230623	P62911	8.94	0.18326	4.34	0.63200
Bcas2	IPI00316184	Q9D287	8.90	0.00231	8.06	0.00409
Sec61b	IPI00755226	Q9CQS8	8.83	0.07926	9.37	0.09263
Csl4;Exosc1	IPI00118360	Q9DAA6-1	8.81	0.02421	8.98	0.02976
Luc7l	IPI00410804	Q9CYI4-1	8.78	0.22783	8.32	0.33777
Fusip1	IPI00314984	Q9R0U0-1	8.76	0.20605	0.19	0.98192
Igf2bp2	IPI00761863	Q5SF07-1	8.75	0.00057	8.32	0.00095
Ing4	IPI00515302	Q8C0D7-5	8.75	0.00295	2.17	0.25379
Adnp	IPI00672180	Q9Z103	8.74	0.07049	5.72	0.26984
Mbd3	IPI00131067	Q9Z2D8-1	8.68	0.26312	3.45	0.77037
Arf4;mCG_5889	IPI00876029	P61750	8.65	0.00261	7.31	0.00585
Crop;Luc7l3	IPI00122418	Q5SUF2-2	8.64	0.00668	8.72	0.00888
D3Wsu161e;Larp7	IPI00340860	Q05CL8-1	8.60	0.05838	7.04	0.14243
Hrb2;Krr1	IPI00223570	Q8BGA5	8.59	0.00695	9.09	0.00866
Luc7l2	IPI00380309	Q7TNC4-1	8.58	0.01370	9.26	0.01373
C2f;Emg1	IPI00127554	O35130	8.56	0.03154	7.54	0.06506
0610010K14Rik	IPI00265101	Q9DCT6-1	8.55	0.16309	2.83	0.75624
Ehmt1	IPI00555042	A2AIS3	8.52	0.01629	-0.99	0.82194
Rpl13	IPI00224505	P47963	8.52	0.07882	3.79	0.51330
Rpl15	IPI00273803	Q9CZM2	8.50	0.06147	3.41	0.51494
Nol9	IPI00625464	Q3TZX8-1	8.49	0.00012	1.42	0.08777
Exosc2;Rrp4	IPI00121437	Q8VBV3	8.48	0.07873	6.51	0.21252
Surf6	IPI00137612	P70279	8.46	0.04941	7.83	0.08481
Imp3	IPI00122383	Q921Y2	8.44	0.10808	2.17	0.78077
Slc25a11	IPI00230754	Q9CR62	8.41	0.00254	8.25	0.00325

Prpf6	IPI00130409	Q91YR7-1	8.40	0.00208	2.74	0.08371
Rpl31	IPI00123007	P62900	8.38	0.16728	4.20	0.60516
Hnrpq	IPI00406117	Q7TMK9-1	8.37	0.11770	3.98	0.56208
Tcfl1	IPI00126141	Q62481	8.37	0.01201	7.99	0.01847
Hnrnpul2	IPI00222208	Q00PI9	8.35	0.10783	3.03	0.67183
Pmx;Pmx1	IPI00230653	P63013-2	8.31	0.06938	8.05	0.10301
Actr6	IPI00110922	Q9D864	8.28	0.00161	8.61	0.00157
Tcfubf	IPI00114869	P25976-1	8.26	0.02431	0.81	0.86613
Rbms2	IPI00121520	Q8VC70	8.19	0.00778	5.61	0.03681
Hp1bp3	IPI00342766	Q3TEA8-1	8.18	0.08398	1.71	0.81243
Gar1;Nola1	IPI00110931	Q9CY66-1	8.18	0.16350	3.52	0.66764
Prpf4	IPI00458908	Q9DAW6	8.15	0.10051	6.08	0.27507
C1qbp	IPI00132799	Q8R5L1	8.14	0.19534	8.33	0.25704
Rbm25	IPI00421119	B2RY56-1	8.11	0.01629	8.95	0.01480
Rpl23a	IPI00461456	P62751	8.08	0.04263	3.17	0.47740
Rpl29	IPI00849550	P47915	8.08	0.11067	3.45	0.60516
Ddx21	IPI00652987	Q3TVJ3	8.08	0.02534	0.82	0.86489
Dkc1	IPI00113635	Q9ESX5	8.05	0.41005	4.09	0.77786
Snrnp40	IPI00461621	Q6PE01	8.04	0.00099	2.72	0.04662
Rps9	IPI00420726	Q6ZWN5	8.03	0.02496	3.45	0.31544
Arf6	IPI00221616	P62331	8.02	0.00295	8.97	0.00263
Rps26	IPI00377441	P62855	8.00	0.32031	3.32	0.79211
Zfp593	IPI00321357	Q9DB42	7.98	0.09775	6.87	0.20443
Myl6	IPI00848866	Q60605-2	7.96	0.17703	5.79	0.42270
Snrpb2	IPI00132576	Q9CQI7	7.91	0.12605	2.49	0.74578
Pc4	IPI00225633	P11031	7.90	0.09009	6.72	0.19610
Rpl18	IPI00555113	P35980	7.90	0.02496	3.09	0.36530
Rpl10	IPI00915054	Q3THJ6	7.90	0.04915	4.01	0.34408
Tcof1	IPI00115660	O08784	7.90	0.01551	6.39	0.04137
Cktsf1b1	IPI00280124	O70326	7.89	0.12475	4.62	0.47081
Polr2a	IPI00136207	P08775	7.84	0.21507	5.51	0.49208
Kiaa1823	IPI00137302	Q9D4J7-1	7.83	0.16863	1.99	0.82373
Pop101	IPI00605037	Q52KI8-1	7.82	0.04211	5.67	0.14834
Rpl17	IPI00626233	Q9CPR4	7.82	0.08481	3.17	0.58358
Cpsf5;Nudt21	IPI00132473	Q9CQF3	7.76	0.12254	0.53	0.95010
Unnamed protein	IPI00551538		7.76	0.50606	7.69	0.60516
Ncl;Nuc	IPI00317794	P09405	7.74	0.01442	3.47	0.21179
Gnl2	IPI00938448	B1ASC3	7.73	0.00778	5.98	0.02335
Utp15	IPI00226889	Q8C7V3	7.73	0.23076	4.36	0.61525
Nol5a;Nop56	IPI00318048	Q9D6Z1	7.68	0.26632	1.95	0.86489
Rras2	IPI00323822	P62071	7.67	0.03240	7.41	0.04844
Rpl28	IPI00222547	P41105	7.65	0.07488	3.32	0.51477
Usp39	IPI00457815	Q3TIX9	7.62	0.06119	3.33	0.47935
Ascc3l1	IPI00420329	Q69ZZ3	7.60	0.02431	1.56	0.69025
Ckap4	IPI00223047	Q8BMK4	7.58	0.02980	-0.74	0.88076
Rpl27	IPI00122421	P61358	7.57	0.22205	3.39	0.70631

Ctcf	IPI00320741	Q61164	7.57	0.00649	0.63	0.82673
Rps6	IPI00113655	P62754	7.56	0.26632	1.61	0.89697
Rps16	IPI00469918	Q5CZY9	7.56	0.25721	4.89	0.58824
Rig;Rps15	IPI00319231	P62843	7.53	0.54307	4.74	0.79211
Dmap1;Mmtr	IPI00317722	Q9JI44-1	7.50	0.27429	6.11	0.49036
Rps15a	IPI00230660	P62245	7.48	0.36126	4.99	0.65723
C130057N11Rik	IPI00828741	A2AU63	7.47	0.17581	2.74	0.74108
Ddx31	IPI00421196	Q6NZQ2	7.44	0.01805	7.88	0.01913
Tceb2	IPI00131224	P62869	7.44	0.00287	8.43	0.00248
L27a;mCG_2315	IPI00626628	Q9CQ16	7.42	0.00278	3.97	0.02633
Prp31;Prpf31	IPI00331424	Q8CCF0-1	7.40	0.15286	5.22	0.39704
Rfc2;mCG_16697	IPI00124744	Q9WUK4	7.39	0.06119	6.89	0.10285
Ing5	IPI00114689	Q9D8Y8-1	7.39	0.06775	-0.55	0.93854
Kiaa0017;Sf3b3	IPI00122011	Q921M3-1	7.38	0.10315	1.72	0.80679
Sf3b2;mCG_19856	IPI00349401	Q3UAI4	7.35	0.01381	4.41	0.09561
D7Wsu180e	IPI00133532	Q9CRA8	7.33	0.02496	5.82	0.07004
Mrpl22	IPI00225318	Q8BU88	7.31	0.00341	7.55	0.00420
Senp3;Smt3ip	IPI00109326	Q9EP97	7.29	0.01952	5.59	0.06244
Rps8;mCG_14435	IPI00466820	P62242	7.26	0.04619	3.31	0.39287
Adprt2;Adprtl2	IPI00131935	O88554	7.23	0.02496	-1.21	0.75624
Arbp;Rplp0	IPI00314950	P14869	7.23	0.26312	4.45	0.61110
Rbm28	IPI00229472	Q8CGC6	7.22	0.07626	6.28	0.15460
Csnk2a1	IPI00408176	Q61177	7.20	0.27269	6.34	0.44546
Rps23	IPI00131357	P62267	7.20	0.10496	2.44	0.69415
Rfc5	IPI00132481	Q9D0F6	7.18	0.37655	1.03	0.94248
Pes;Pes1	IPI00620167	Q9EQ61	7.18	0.01172	7.92	0.01176
Hnrnpr;Hnrpr	IPI00128441	Q3U8W9	7.16	0.14181	2.69	0.69652
Rps17;mCG_15301	IPI00755495	P63276	7.15	0.03758	5.93	0.09263
Sr140	IPI00830853	Q6NV83-1	7.14	0.02496	5.82	0.06537
Btf3;mCG_114993	IPI00515257	Q64152-1	7.13	0.08481	7.26	0.11281
Arl10c;Arl8b	IPI00133218	Q9CQW2	7.13	0.00766	7.82	0.00845
Rrp1b;mKIAA0179	IPI00130246	Q91YK2	7.13	0.00406	-1.59	0.36122
Ilf3	IPI00130591	Q9Z1X4-3	7.12	0.09816	8.80	0.07429
Rpl26	IPI00132460	P61255	7.12	0.06622	3.09	0.49463
Rrbp1	IPI00121149	Q99PL5-1	7.07	0.00820	7.11	0.01160
Rps7;Rps7A	IPI00136984	P62082	7.06	0.10604	3.52	0.51477
Rpl9;mCG_10266	IPI00122413	P51410	7.03	0.06119	3.57	0.38950
4933421E11Rik	IPI00228791	Q8CDD9-1	7.00	0.03552	0.42	0.93854
Cdc73	IPI00170345	Q8JZM7	6.98	0.00508	6.53	0.00879
Hdgfrp2	IPI00116442	Q3UMU9-3	6.97	0.22299	6.38	0.35138
Exosc8;Rrp43	IPI00109639	Q9D753	6.95	0.29741	4.79	0.59351
Hdac2;Yy1bp	IPI00137668	P70288	6.95	0.26632	2.63	0.78610
Ddx10	IPI00896604	Q80Y44	6.92	0.00665	4.63	0.03238
Gatad2b	IPI00128615	Q8VHR5-1	6.89	0.01579	9.92	0.00688
Kiaa0007;Wdr43	IPI00938465	Q6ZQL4	6.88	0.06119	1.90	0.69442
Thrap3;Trap150	IPI00556768	Q569Z6	6.88	0.35734	5.46	0.57917

Sfrs2;Pr264	IPI00474430	Q8C671	6.87	0.48861	5.54	0.68061
Eftud2	IPI00649950	A2AH85	6.86	0.18326	2.30	0.76616
Aly;Ref1;Refbp1	IPI00114407	O08583-1	6.86	0.56008	1.53	0.93854
mCG_132913	IPI00225634	Q6ZWZ6	6.86	0.31242	6.27	0.47034
Fam98a	IPI00457756	Q3TJZ6	6.85	0.14181	6.44	0.22985
Glut1;Glut-1	IPI00308691	P17809	6.84	0.20510	-0.11	0.98307
Rpl7	IPI00311236	P14148	6.84	0.58313	6.39	0.69652
Eif2a;Eif2s1	IPI00474446	Q6ZWX6	6.83	0.51967	3.75	0.81552
Aporp1;Arp1	IPI00117234	P43135	6.81	0.00295	8.41	0.00167
Lyar	IPI00113232	Q08288	6.81	0.42882	1.69	0.92156
Zc3hav1	IPI00136572	Q3UPF5-1	6.81	0.02122	6.62	0.03035
Rnu1a-1;Snrpa	IPI00122350	Q62189	6.80	0.42055	5.66	0.61239
Rps4;Rps4x	IPI00331092	P62702	6.80	0.12475	3.57	0.51477
Rps14;rps14	IPI00322562	P62264	6.80	0.02595	3.25	0.26517
Top2;Top-2	IPI00122223	Q01320	6.79	0.10505	3.34	0.51477
Rpl21	IPI00315548	Q4VA28	6.78	0.02552	3.28	0.25898
Rps11	IPI00117569	Q9DB79	6.77	0.04467	3.31	0.34762
Baz2a;Kiaa0314	IPI00944144	Q91YE5-3	6.75	0.16309	3.97	0.51477
Ilf2;Nf45	IPI00318550	Q9CXY6	6.73	0.50733	2.90	0.85691
Aatf;Che1;Trb	IPI00123565	Q9JKX4-1	6.72	0.02662	3.69	0.20660
Cpsf6	IPI00421085	Q6NVF9	6.69	0.15846	1.04	0.90436
Rpl23	IPI00139780	P62830	6.68	0.55008	4.94	0.75504
Gpx4;GPx4	IPI00660262	O70325-2	6.67	0.00790	6.13	0.01372
Mybbp1a	IPI00331361	Q7TPV4	6.64	0.03240	2.49	0.45568
Pgam5	IPI00848970	Q8BX10-1	6.63	0.25692	6.07	0.39287
Rpl12	IPI00849793	P35979	6.62	0.02264	3.36	0.21252
Pop1	IPI00169797	Q8K205	6.61	0.09223	6.78	0.12037
Exosc7	IPI00387489	Q4VBW5	6.61	0.05135	5.98	0.09561
Stau1	IPI00749596	A2A5S3	6.59	0.00470	6.52	0.00688
Rps5	IPI00944141	Q91V55	6.55	0.04838	4.05	0.23514
Rpl7a;Surf3	IPI00330363	P12970	6.53	0.27094	2.09	0.82596
Impnb;Kpnb1	IPI00323881	P70168	6.50	0.40686	3.46	0.76077
Eif2s2	IPI00116302	Q99L45	6.50	0.26632	5.45	0.46509
Gatad2a	IPI00625995	Q8CHY6	6.47	0.21847	6.82	0.26611
Nle1	IPI00930808	B1ARD5	6.47	0.05507	5.99	0.09561
Nxf1;Tap	IPI00115588	Q99JX7	6.45	0.42010	2.50	0.84416
Ant2;Slc25a5	IPI00127841	P51881	6.43	0.36541	11.38	0.16183
G22p2;Xrcc5	IPI00321154	P27641	6.43	0.03240	6.88	0.03681
Unnamed protein	IPI00849047		6.42	0.04938	6.63	0.06244
Tex10	IPI00458057	Q3URQ0	6.42	0.13522	5.88	0.23169
Rnasep2;Rpp30	IPI00134346	O88796	6.39	0.05251	6.33	0.07369
Pdlim2	IPI00153375	Q8R1G6	6.38	0.00143	6.35	0.00162
Gltscr2	IPI00122471	Q3USL3	6.37	0.03263	6.40	0.04460
Ki-67;Mki67	IPI00124959	Q61769	6.36	0.04467	-4.19	0.20024
Rpl4;mCG_12784	IPI00111412	Q9D8E6	6.32	0.30451	2.00	0.84416
Cebpb	IPI00116613	P28033	6.29	0.01271	4.63	0.04728

Big;Big3	IPI00139957	P61965	6.28	0.07719	1.74	0.72038
Anc1;Ant1	IPI00115564	P48962	6.27	0.26632	4.99	0.49133
Rpl7l1	IPI00111695	Q9D8M4	6.27	0.43184	2.36	0.85713
Nfib	IPI00130129	P97863-1	6.22	0.03466	6.30	0.04662
Rps3	IPI00134599	P62908	6.22	0.07061	3.24	0.39704
Cdc46;Mcm5	IPI00309398	P49718	6.21	0.16728	0.25	0.97349
Cbfb;Pebp2b	IPI00229487	Q08024-2	6.18	0.02167	6.56	0.02231
Cbp80;Ncbp1	IPI00458056	Q3UYV9	6.15	0.01607	5.53	0.03080
Chd3	IPI00675483	B1AR17	6.11	0.00820	7.01	0.00842
Ppp1a;Ppp1ca	IPI00130185	P62137	6.10	0.44103	1.79	0.90033
Nolc1	IPI00720058	Q6ZQK6	6.09	0.03504	4.46	0.12390
Rpl6;mCG_16023	IPI00313222	P47911	6.08	0.58160	5.41	0.71242
Ssb;Ss-b	IPI00134300	P32067	6.06	0.46912	3.70	0.75672
Sfrs2ip	IPI00874991	Q6ZPE9	6.05	0.01551	6.37	0.01620
Kiaa0731;Larp	IPI00929786	Q6ZQ58-1	6.04	0.06449	7.43	0.04662
Rps13	IPI00125901	P62301	6.03	0.08090	3.36	0.39002
Ngd;Ngdn	IPI00119201	Q9DB96	6.03	0.33024	4.17	0.61336
Rps3a;mCG_1128	IPI00331345	P97351	6.02	0.69084	3.20	0.88241
Chaf1b	IPI00132770	Q9D0N7	5.98	0.10042	7.02	0.09077
Snrpd2;mCG_4862	IPI00119220	P62317	5.97	0.42010	5.05	0.60516
Baz1a;Cbp146	IPI00461396	O88379	5.93	0.03281	1.59	0.61525
Rab5c	IPI00404579	Q8C266	5.92	0.05829	0.89	0.85070
Np95;Uhrf1	IPI00130200	Q8VDF2	5.90	0.54269	4.35	0.75118
Sfrs4	IPI00606760	A2A837	5.86	0.03154	6.06	0.03904
Msy1;Msy-1	IPI00120886	P62960	5.86	0.35608	2.28	0.82156
Jund;Jun-d	IPI00126223	P15066	5.84	0.03504	6.24	0.03974
Cdc5l;Kiaa0432	IPI00284444	Q6A068	5.81	0.24776	4.32	0.50407
Gins1;Psf1	IPI00122249	Q9CZ15	5.80	0.03130	7.42	0.01802
Gzf1;Zfp336	IPI00756220	Q4VBD9	5.76	0.00875	5.67	0.01255
Fxr2;Fxr2h	IPI00652944	Q3TA75	5.76	0.41331	4.99	0.58744
Gpsn2;Tecr	IPI00875068	Q52L67	5.73	0.01629	6.52	0.01373
Slit3;Slit1	IPI00667854	B1ATW4	5.73	0.00231	5.68	0.00263
Hnrnpu;Hnrpu	IPI00458583	Q8VEK3	5.72	0.02496	2.45	0.31811
Prp8;Prpf8	IPI00121596	Q99PV0	5.70	0.05848	2.81	0.39547
Chd4	IPI00857771	Q5DTP7	5.69	0.04977	1.87	0.59164
Haf;Sart1	IPI00323674	Q9Z315	5.69	0.21653	3.93	0.50371
Fmr1;Fmr-1	IPI00227005	P35922-3	5.67	0.59095	2.88	0.85713
Kiaa0162	IPI00454050	Q62383	5.67	0.03240	5.98	0.03793
Mphosph6	IPI00134044	Q9D1Q1	5.64	0.05838	5.81	0.07355
Slit2	IPI00128415	Q9R1B9	5.63	0.00649	5.35	0.01015
D10Wsu52e	IPI00116850	Q99LF4	5.62	0.10937	11.34	0.01718
Rab18	IPI00116770	P35293	5.62	0.36932	4.75	0.55231
Rpl11;mCG_5487	IPI00331461	Q9CXW4	5.59	0.02294	3.00	0.19155
Ddx17	IPI00653307	Q3U741	5.57	0.11208	0.58	0.92734
Nup205	IPI00929832	B0LAE1	5.56	0.06119	3.97	0.21094
Unnamed protein	IPI00330406		5.51	0.08107	3.76	0.27622

Kpna2;Rch1	IPI00124973	P52293	5.47	0.48889	0.47	0.97171
Matr3	IPI00453826	Q8K310	5.46	0.46262	4.28	0.67423
Rps25	IPI00137735	P62852	5.45	0.07061	3.15	0.33817
Arf5	IPI00221615	P84084	5.43	0.05445	5.61	0.06819
Fbp11;Fnbp3	IPI00284213	Q9R1C7-1	5.43	0.17563	2.72	0.61284
Hrs;Sfrs5	IPI00314709	O35326	5.42	0.70030	7.17	0.66583
Ipo7;Ranbp7	IPI00331444	Q9EPL8	5.40	0.09816	6.57	0.07752
Nipbl	IPI00421052	Q6KCD5-1	5.38	0.29787	1.19	0.90033
Snrpd1	IPI00322749	P62315	5.31	0.68866	5.22	0.75504
U2af2;U2af65	IPI00113746	P26369	5.30	0.21827	3.54	0.51477
Kif20a;Rab6kifl	IPI00123947	P97329	5.30	0.08398	-5.25	0.12037
Tubb2c;Tubb2c1	IPI00169463	P68372	5.30	0.57140	6.08	0.59351
Sfrs9	IPI00132340	Q9D0B0	5.29	0.29319	2.57	0.73740
Sin3a;mKIAA4126	IPI00117932	Q60520-1	5.21	0.03263	5.81	0.03241
Hnrnpa0;Hnrpa0	IPI00109813	Q9CX86	5.17	0.59637	1.94	0.90612
Aprin;As3	IPI00845638	Q4VA53-3	5.16	0.16228	5.93	0.16326
Srfbp1	IPI00395196	Q9CZ91	5.16	0.42010	1.08	0.92734
Wdr18	IPI00136252	Q4VBE8	5.15	0.47078	-2.43	0.82373
Hnrnpc;Hnrpc	IPI00223443	Q9Z204-2	5.08	0.10654	1.58	0.73231
FAM120A	IPI00830478	Q6A0A9	5.05	0.03466	5.11	0.04654
Eif2s3x	IPI00230415	Q9Z0N1	5.04	0.55141	2.81	0.82367
Kiaa1341;Mef2	IPI00226891	Q8C854-4	5.04	0.42871	2.99	0.75188
Rpl3;mCG_11520	IPI00755309	Q3T9U9	5.03	0.50877	1.66	0.89906
Ddx15;Deah9	IPI00128818	O35286	5.01	0.21827	4.27	0.38369
Fbl	IPI00119581	P35550	5.00	0.72286	1.51	0.94248
Wdr61;mCG_9684	IPI00112320	Q9ERF3	4.99	0.42640	3.66	0.67391
Wiz	IPI00263016	O88286-1	4.98	0.15013	0.28	0.96448
Pcbp1	IPI00128904	P60335	4.88	0.41215	3.15	0.70631
Alg4;Kiaa0185	IPI00551454	Q6NS46	4.87	0.25432	-0.79	0.92617
Supt5h	IPI00656221	O55201-1	4.87	0.22322	2.01	0.74059
Mcm3;Mcmd	IPI00108338	P25206	4.84	0.27094	5.44	0.29432
Pak1ip1	IPI00120790	Q9DCE5	4.81	0.43038	2.51	0.78637
Smn;Snrpn	IPI00128699	P63163	4.80	0.12428	2.27	0.57922
Mem3;Vps35	IPI00111181	Q9EQH3	4.80	0.03077	5.01	0.03681
Srcap	IPI00620743	Q8BKT0	4.79	0.42055	5.37	0.47034
Nol14;Nop14	IPI00353010	Q8R3N1	4.78	0.23026	2.58	0.63695
Ddx5;Tnz2	IPI00420363	Q61656	4.77	0.24082	3.22	0.53949
Lamr1;P40-8;Rps	IPI00123604	P14206	4.73	0.09223	2.42	0.47887
Rps18;mCG_23000	IPI00317590	P62270	4.72	0.10051	3.22	0.32317
Lig3;mCG_8187	IPI00124272	P97386-1	4.71	0.09329	5.04	0.10960
Eef1d;mCG_22130	IPI00944009	P57776-3	4.70	0.50884	-0.34	0.97611
Epb7.2;Epb72	IPI00323748	P54116	4.69	0.21847	4.81	0.27744
Ppih	IPI00848926	Q9D868-2	4.67	0.57140	1.37	0.92734
Tuba1;Tuba1a	IPI00110753	P68369	4.66	0.58037	1.01	0.94248
Rps10	IPI00112448	P63325	4.66	0.20056	2.95	0.52428
Slc25a1	IPI00276926	Q3TDH6	4.66	0.52588	4.95	0.59314

Ddx9;Dhx9	IPI00339468	O70133-2	4.49	0.49433	1.64	0.87635
Hdac1	IPI00114232	O09106	4.44	0.09770	4.25	0.15511
Mrps26	IPI00108194	Q80ZS3	4.42	0.12307	3.50	0.28132
Crp1;Csrp	IPI00123891	P97315	4.39	0.27429	-0.43	0.95157
Chd2	IPI00845761	Q3V0V7	4.34	0.19811	0.39	0.94583
Hnrnpm;Hnrpm	IPI00132443	Q9D0E1-1	4.31	0.49585	2.96	0.74401
Polr2e	IPI00337955	Q80UW8	4.26	0.04619	1.65	0.49133
Snrp1c;Snrpc	IPI00123376	Q62241	4.26	0.37374	1.66	0.82373
Csnk1a1	IPI00330729	Q8BK63-1	4.26	0.01201	4.32	0.01480
Rpl30;mCG_20799	IPI00222549	P62889	4.23	0.38719	2.55	0.71841
Rbbp5	IPI00226384	Q3TBL4	4.22	0.07052	2.52	0.32084
Trim56	IPI00453527	Q80VI1	4.21	0.00412	3.20	0.01255
Toe1;mCG_19662	IPI00134747	Q9D2E2-1	4.16	0.29612	5.19	0.26517
C5orf24 homolog	IPI00330644	Q80X32	4.08	0.13233	1.97	0.58033
Puf60	IPI00652919	Q3UEB3-1	4.07	0.68102	6.81	0.51947
Rbmx2	IPI00261255	Q8R0F5	4.01	0.39406	0.71	0.93441
Plrg1	IPI00331172	Q922V4	4.00	0.36459	5.50	0.27507
Xrcc1;Xrcc-1	IPI00118139	Q60596	3.98	0.36226	1.97	0.75830
Hnrnpf;Hnrpf	IPI00226073	Q9Z2X1-1	3.92	0.47078	2.14	0.79252
Ehmt2	IPI00756765	A2CG77	3.91	0.11903	0.12	0.97611
Dimt1;Dimt1l	IPI00132415	Q9D0D4	3.89	0.69084	-0.14	0.98509
A170;Sqstm1	IPI00133374	Q64337-1	3.88	0.34273	-2.14	0.71870
Sap155	IPI00623284	Q99NB9	3.88	0.58399	2.18	0.83647
Alg2;Pdcd6	IPI00121736	P12815	3.86	0.68547	3.48	0.77581
Mecp2	IPI00775806	Q9Z2D6-2	3.83	0.38340	3.46	0.52656
Sfrs3	IPI00129323	P84104-1	3.78	0.05789	2.95	0.16060
Ruvbl2	IPI00123557	Q9WTM5	3.78	0.29428	3.22	0.48403
Hmgts;Tfam	IPI00112822	P40630-1	3.77	0.72175	2.39	0.86613
Afg3l2	IPI00170357	Q8JZQ2	3.74	0.10319	2.73	0.29213
Drim;Utp20	IPI00850851	Q5XG71	3.72	0.00668	2.92	0.01847
Chd1;Chd-1	IPI00107999	P40201	3.70	0.00406	3.95	0.00456
Snrpd3	IPI00119224	P62320	3.69	0.26703	1.87	0.69652
Mphosph8	IPI00121617	A6H600	3.68	0.52274	0.83	0.93441
Cdk9;mCG_18574	IPI00114953	Q99J95-1	3.62	0.42871	2.37	0.71870
Dnaja2	IPI00136251	Q9QYJ0	3.61	0.72197	2.91	0.82855
Actg;Actg1	IPI00874482	P63260	3.57	0.72723	4.47	0.71962
Adnp2;Kiaa0863	IPI00875894	Q8CHC8	3.55	0.04155	4.80	0.02066
Pairbp1;Serbp1	IPI00471476	Q9CY58-2	3.55	0.75116	1.80	0.92156
Rent1;Upf1	IPI00420949	Q9EPU0-1	3.55	0.43068	2.84	0.63715
Fip1l1	IPI00881462	Q9D824-1	3.55	0.66647	2.22	0.84702
Ddb1	IPI00316740	Q3U1J4	3.49	0.26632	-2.03	0.63695
Bclaf1	IPI00620639	Q05C67	3.48	0.77755	0.75	0.96867
MNCb-1706	IPI00283671	Q9CY57-1	3.48	0.71676	0.71	0.96168
mCG_12245;Rbbp4	IPI00828412	A2A875	3.46	0.35786	2.56	0.60753
Cbx;Cbx1	IPI00129466	P83917	3.41	0.12307	0.14	0.97171
Dnajc9	IPI00128268	Q91WN1	3.41	0.71949	0.39	0.97611

Spin;Spin1	IPI00653709	Q3UWM7	3.38	0.41982	0.44	0.94836
Cbx3;mCG_119115	IPI00677454	Q32P00	3.35	0.12586	0.25	0.94620
Fxr1;Fxr1h	IPI00122521	Q61584-1	3.34	0.77755	-1.09	0.94583
Rbm14;p16	IPI00404707	Q8C2Q3-1	3.34	0.60846	2.90	0.74401
Skp1;Skp1a	IPI00331163	Q9WTX5	3.28	0.67209	2.49	0.81407
Prkcbp1	IPI00108978	A2A484	3.28	0.27696	4.15	0.24266
Mcm6;Mcmd6	IPI00123881	P97311	3.21	0.47078	0.86	0.92156
Ddx48;Eif4a3	IPI00126716	Q91VC3	3.20	0.72286	1.13	0.93784
Cpsf100;Cpsf2	IPI00314302	O35218	3.20	0.22252	1.18	0.76756
Ruvbl1;Tip49	IPI00133985	P60122	3.18	0.46938	2.69	0.64054
Rbap46;Rbbp7	IPI00122698	Q60973	3.14	0.39812	1.55	0.77800
Nap1l1	IPI00929813	P28656	3.07	0.69652	2.61	0.80679
Kiaa0324	IPI00785240	Q8BTI8-1	3.01	0.35776	1.44	0.76616
H2a.x;H2afx	IPI00230264	P27661	2.97	0.66894	4.11	0.60516
Cenpl	IPI00830718	Q14A61	2.92	0.52274	2.18	0.73740
Kiaa1707;Senp7	IPI00403228	Q8BUH8	2.89	0.54661	-1.20	0.86995
Rfc3	IPI00665571	Q8R323	2.87	0.84512	1.49	0.93854
Orc4l	IPI00752477	Q8BN98	2.78	0.75497	-3.85	0.71242
Actl6;Actl6a	IPI00323660	Q9Z2N8	2.77	0.37415	1.45	0.75504
Gapd;Gapdh	IPI00874964	P16858	2.75	#VALUE!	0.74	0.27845
Kiaa1584;Suhw3	IPI00169658	Q6P3Y5-1	2.70	0.72175	-3.30	0.71962
Rnu3ip2;Rrp9	IPI00128256	Q91WM3	2.68	0.71670	0.70	0.94583
Kpna3;Qip2	IPI00230429	O35344	2.64	0.74740	0.75	0.94677
Cdc47;Mcm7	IPI00126396	Q61881	2.63	0.27269	1.17	0.75118
Pno1	IPI00131909	Q9CPS7	2.60	0.74740	0.30	0.97734
Baf170;Smarcc2	IPI00459742	Q6PDG5-1	2.59	0.10148	1.12	0.58744
Hnrnph1;Hnrph1	IPI00224729	Q8C2Q7	2.59	0.67620	1.17	0.90272
D1Pas1-rs2;Ddx3	IPI00230035	Q62167	2.58	0.40717	0.70	0.89889
Kiaa0650;Smchd1	IPI00137433	Q6P5D8	2.57	0.40894	0.25	0.96317
Nsa2;Tinp1	IPI00468437	Q9CR47	2.49	0.58015	0.45	0.94836
Prp19;Prpf19	IPI00222760	Q99KP6-2	2.46	0.80291	2.67	0.82596
Hbo1;Myst2	IPI00515403	Q5SVQ0-1	2.44	0.24062	-1.43	0.61120
Tim50;Timm50	IPI00111045	Q9D880	2.44	0.67570	0.13	0.98192
Baf190a;Brg1	IPI00460668	Q3TUD7	2.40	0.54559	0.66	0.92734
Ccdc55	IPI00345960	Q5NCR9	2.39	0.00002	3.90	0.00000
Sdpr	IPI00135660	Q63918	2.38	0.77755	1.48	0.90501
Cbx5;Hp1a	IPI00123755	Q61686	2.36	0.37655	-0.52	0.92156
Hsc70;Hsc73	IPI00323357	P63017	2.27	0.35118	1.31	0.70631
Fug1;Rangap1	IPI00467338	P46061	2.24	0.42010	-1.80	0.61840
Col8a1	IPI00399464	Q00780	2.21	0.78852	2.65	0.79211
Fn1	IPI00352163	B7ZNJ1	2.20	0.75750	2.94	0.73343
Orc1;Orc1l	IPI00130290	Q9Z1N2	2.18	0.68126	-6.25	0.23169
Nsd3;Whsc1l1	IPI00663331	Q6P2L6-1	2.16	0.77217	0.98	0.92868
Brd1;mKIAA4191	IPI00330331	Q571F6	2.10	0.62300	-0.12	0.98192
cbp37;Gtl6	IPI00623570	O88380	2.06	0.66408	-2.89	0.59314
Tpx2	IPI00420481	A2APB8	2.02	0.84263	-6.81	0.47740

Ase1;Cd3eap	IPI00169700	Q76KJ5	1.99	0.75111	1.69	0.83647
Tra2a	IPI00377298	Q6PFR5	1.98	0.86442	-0.24	0.98192
Cdw5;Wdr82_	IPI00221414	Q8BFQ4	1.97	0.59637	-0.10	0.98192
Hbxap;Rsf1	IPI00122845	Q3UJ04	1.97	0.22256	0.67	0.79211
Kpna4;Qip1	IPI00129792	O35343	1.93	0.81499	3.02	0.74887
Grp78;Hspa5	IPI00319992	P20029	1.92	0.42871	1.28	0.71242
Fn1	IPI00652813	B9EHT6	1.92	0.23324	1.33	0.51477
Hnrnpl	IPI00620362	Q8R081	1.90	0.59637	0.45	0.93854
Poldip3	IPI00229721	Q8BG81	1.90	0.58327	-3.98	0.27845
Unnamed protein	IPI00118632		1.84	0.68163	0.47	0.94415
Prss3	IPI00130391	B9EJ35	1.83	0.86688	1.10	0.93854
Erh;mCG_7617	IPI00403589	P84089	1.82	0.75256	-0.65	0.93854
H1f1;Hist1h1a	IPI00228616	P43275	1.82	0.64893	-0.44	0.94248
Smarca5;Snf2h	IPI00396739	Q91ZW3	1.80	0.38340	-1.44	0.59320
Ddx5	IPI00648763	B1ARB9	1.78	0.79981	0.92	0.92734
Baz1b;Wbscr9	IPI00923656	Q9Z277-1	1.78	0.42871	-2.47	0.34425
U2af1;mCG_14259	IPI00318548	Q9D883	1.76	0.91103	1.82	0.92734
Cnbp;Cnbp1	IPI00454151	P53996-3	1.73	0.51967	-0.73	0.86318
Hmga1	IPI00954313	Q3TE85	1.63	0.62738	0.51	0.92734
D4Cole1e;Nmnat	IPI00109667	Q9EPA7	1.62	0.45784	2.11	0.41933
mKIAA0138;Safb2	IPI00605227	Q6A0C0	1.62	0.77217	-2.93	0.62410
Smc1a	IPI00944715	A0JLM6	1.60	0.82764	0.53	0.95994
Rrp15	IPI00458958	Q9CYX7	1.59	0.89696	-0.57	0.97154
Cct1;Ccta	IPI00459493	P11983-1	1.53	0.84956	-0.17	0.98192
Cct8;Cctq	IPI00469268	P42932	1.53	0.67543	0.91	0.85963
Insrr;mCG_3075	IPI00896646	B2RQC5	1.49	0.08398	-0.02	0.98307
Jun	IPI00121829	P05627	1.47	0.74719	0.31	0.96317
Copb2	IPI00115097	O55029	1.45	0.37796	0.49	0.85691
Cse1l;Xpo2	IPI00112414	Q9ERK4	1.42	0.82301	2.42	0.73478
Nrap	IPI00135182	Q80XB4-1	1.40	0.79335	2.06	0.74401
Adprp;Adprt	IPI00112473	P11103-1	1.35	0.77217	-3.11	0.51115
Hmga2	IPI00752993	Q6NSP9	1.34	0.73919	-2.21	0.61509
Gnas;Gnas1	IPI00416906	Q6R0H7-1	1.32	0.74182	1.93	0.67439
Men1;mCG_5034	IPI00918121	O88559	1.27	0.87420	-6.14	0.36388
Ledgf;Psip1	IPI00115257	Q99JF8-1	1.26	0.86442	-1.72	0.84416
H1f3;Hist1h1d	IPI00331597	P43277	1.23	0.88704	1.11	0.92303
Hmga1;Hmgi	IPI00624711	P17095-1	1.23	0.77755	1.32	0.81407
Tubb5	IPI00117352	P99024	1.21	0.90369	2.04	0.85691
Ddx51	IPI00396728	Q6P9R1	1.20	0.72286	-0.01	0.99682
Lmn1;Lmna	IPI00620256	P48678-1	1.16	0.72197	-1.11	0.79211
Grp75;Hsp74	IPI00133903	P38647	1.16	0.82439	1.39	0.82373
Orc5;Orc5l	IPI00125261	Q9WUV0	1.15	0.91061	-5.92	0.49977
Hmga2	IPI00331612	P52927	1.14	0.66524	0.41	0.92734
Rrp7a	IPI00133594	Q9D1C9	1.13	0.89602	-2.10	0.82172
Hnrnpa2b1	IPI00853914	O88569-1	1.10	0.89367	-1.54	0.86613
Aros;Rps19bp1	IPI00226227	Q8C6B9	1.06	0.80745	2.94	0.49036

Gnai3	IPI00338854	Q9DC51	1.04	0.76088	0.06	0.98307
Hist1h2bk	IPI00875277	Q8CGP1	1.03	0.92860	-1.00	0.94468
Rab21	IPI00337980	P35282	0.97	0.84670	1.70	0.75624
Ddx52;Rok1	IPI00336965	Q8K301	0.96	0.51525	-0.20	0.93854
Atp5h	IPI00230507	Q9DCX2	0.96	0.87323	-1.08	0.88076
H3.1-221	IPI00553538	P68433	0.96	0.95503	0.74	0.97154
D2Wsu81e	IPI00224127	Q3UHX9-1	0.93	0.79335	0.55	0.92156
H3.3a;H3.3b	IPI00785343	P84244	0.89	0.96150	2.36	0.92303
DinG;Ring1b	IPI00133880	Q9CQJ4	0.85	0.94907	-0.96	0.94836
mCG_10339;Tmpo	IPI00896574	B2RUB9	0.84	0.92743	-3.29	0.69966
Mafg;mCG_2153	IPI00755238	O54790	0.80	0.92743	-5.04	0.47922
Ftsj3	IPI00119632	Q9DBE9	0.77	0.60568	1.45	0.36882
Naa40;Nat11	IPI00137271	Q8VE10-1	0.77	0.77755	-7.56	0.01435
Hmga1;Hmgi	IPI00314240	P17095-2	0.73	0.88704	-0.15	0.97957
D1Wsu40e;Imp4	IPI00395166	Q8VHZ7	0.71	0.89131	0.22	0.97171
H1f2;Hist1h1c	IPI00223713	P15864	0.69	0.82439	-0.34	0.93854
Aldh18a1	IPI00944043	Q3TWN8	0.68	0.92743	-3.14	0.63567
Msi2;Msi2h	IPI00120924	Q920Q6-1	0.63	0.92860	1.54	0.84416
Sdha	IPI00230351	Q8K2B3	0.61	0.92860	2.97	0.64043
Mtf2;Pcl2	IPI00467892	Q02395-1	0.57	0.86113	-4.46	0.10233
D1Ertd578e	IPI00222122	Q8BIG4	0.53	0.92743	-2.55	0.63148
;Smarcd2;Baf60b	IPI00649685	B1ARJ6	0.52	0.95823	-2.53	0.81243
Hp1bp3	IPI00896020	Q3TEA8-2	0.52	0.89359	-0.46	0.92734
Pb1;Pbrm1;Baf18	IPI00461676	Q8R134	0.51	0.91106	4.52	0.21167
H1f0;H1fv	IPI00467914	P10922	0.50	0.91226	-1.50	0.75188
Ran;Rasl2-8	IPI00134621	P62827	0.50	0.90329	-2.47	0.47935
Kiaa1403;Rpap1	IPI00377618	Q80TE0-1	0.49	0.96263	-1.11	0.93441
Kiaa1709;Nat10	IPI00276866	Q8K224	0.47	0.91335	-1.02	0.82721
Ccdc137	IPI00153212	Q8R0K4	0.44	0.91335	-1.26	0.76228
H1f4;Hist1h1e	IPI00223714	P43274	0.39	0.92860	0.37	0.94415
Hmgn2	IPI00918033	B7ZCQ2	0.38	0.97767	-1.69	0.92303
Imp9;Ipo9	IPI00130158	Q91YE6	0.36	0.96150	0.28	0.97331
Hist1h2af	IPI00229543	Q8CGP5	0.28	0.95503	-1.18	0.81243
Kcnc3	IPI00467430	Q63959-1	0.26	0.98236	-0.22	0.98192
Hmg14	IPI00338745	P18608	0.23	0.98412	-3.28	0.85109
Ark2;Aurkb	IPI00268655	O70126	0.23	0.94414	-7.45	0.01495
H1f5	IPI00230133	P43276	0.23	0.96718	-0.90	0.90218
Kiaa0938	REV__IPI001	Q80TN7	0.19	0.95503	-8.45	0.01160
Sap18	IPI00915480	O55128	0.12	0.98672	-1.20	0.92734
Orc3;Orc3l	IPI00311034	Q9JK30-1	0.12	0.98297	-7.35	0.17057
Cdc21;Mcm4	IPI00117016	P49717	0.09	0.96122	-1.04	0.47740
H2afy2;H2afy3	IPI00652934	Q8CCK0	0.08	0.98236	0.84	0.86489
D11Ertd530e	IPI00396676	Q80U70	0.02	0.99279	-2.53	0.38950
H2afv	IPI00555055	Q3THW5	0.00	0.99908	-0.90	0.51330
Hmgn3	IPI00120653	Q9DCB1-1	-0.03	0.99279	-1.60	0.76228
H3.2	IPI00282848	P84228	-0.08	0.89423	-0.23	0.69025

Rnps1;RNPS1	IPI00122227	Q62150	-0.09	0.98412	-0.34	0.96867
Ahnak2	IPI00850843	Q3UUE0	-0.10	0.96471	1.45	0.44451
Bp75;Brd7	IPI00133099	O88665	-0.10	0.97922	-1.94	0.56475
H4-12	IPI00623776	P62806	-0.12	0.92743	-0.14	0.92734
Sfrs10	IPI00139259	P62996-1	-0.13	0.98814	-0.93	0.95761
Ptrf	IPI00117689	O54724	-0.18	0.98410	1.26	0.92734
Rrp8	IPI00653834	Q3U4B0	-0.18	0.97922	-5.54	0.30595
Arglu1	IPI00652831	Q3UL36-1	-0.20	0.94414	-1.03	0.71870
Hist2h2ab	IPI00623951	Q64522	-0.25	0.96571	-12.24	0.02638
Scaf1;Sfrs19	IPI00344430	Q5U4C3-1	-0.27	0.92743	-4.55	0.05920
Ctdspl2	IPI00454047	Q8BG15-1	-0.34	0.96571	-7.75	0.23784
Rps27a	IPI00470152	P62983	-0.39	0.92844	-0.47	0.92960
Plec1	IPI00626385	Q9QXS1-2	-0.41	0.93800	-0.91	0.88258
Hist1h2bh	IPI00876550	Q64478	-0.54	0.14181	-0.32	0.47935
Parp1	IPI00139168	Q3TF18	-0.54	0.71934	-3.26	0.03941
Phc2	IPI00894922	B2KFW6	-0.55	0.94414	-1.89	0.82338
Phgdh	IPI00225961	Q61753	-0.63	0.95503	-3.13	0.78091
Ndrp;Phip	IPI00311490	Q8VDD9	-0.63	0.92860	-4.88	0.42149
Dek	IPI00227720	Q7TNV0	-0.64	0.70862	6.64	0.00645
Fus;Taf15	IPI00117063	P56959	-0.73	0.79145	0.07	0.98192
Sf3a1	IPI00408796	Q8K4Z5	-0.86	0.93833	-0.49	0.97171
Coq6	IPI00222526	Q8R1S0	-0.88	0.82439	-8.22	0.03457
Fosl2;Fra2	IPI00338178	P47930	-0.94	0.92860	-2.33	0.85070
Hist2h2aa1	IPI00330000	Q6GSS7	-0.99	0.51286	-1.02	0.59164
Orc2;Orc2l	IPI00121509	Q60862	-1.06	0.82439	-8.29	0.05898
Rbmxrt;Hnrnpg	IPI00663587	Q91VM5	-1.13	0.89367	-3.21	0.69166
Actbl2	IPI00221528	Q8BFZ3	-1.16	0.94414	2.11	0.92303
Smu1	IPI00331342	Q3UKJ7-1	-1.28	0.72175	-2.67	0.47935
Numa1	IPI00263048	Q3TH77	-1.46	0.88071	-2.02	0.85691
Csrp2;Dlp1	IPI00470178	P97314	-1.49	0.74292	8.36	0.06413
Cct4;Cctd	IPI00116277	P80315	-1.51	0.75256	-1.79	0.75624
Kiaa0170;Mdc1	IPI00753701	Q5PSV9	-1.56	0.89367	-2.19	0.86489
Kif22	IPI00116757	Q3V300	-1.59	0.64893	-7.09	0.05761
Nfix	IPI00331515	Q5CZY4	-1.67	0.85785	-0.32	0.97621
Sfpq	IPI00129430	Q8VIJ6	-1.73	0.86688	-3.49	0.75504
Baf47;Ini1	IPI00129145	Q9Z0H3-1	-1.76	0.84956	-7.14	0.38267
Lrwd1	IPI00225459	Q8BUI3-1	-2.20	0.80190	-6.41	0.45568
Wdr76	IPI00856332	A6PWY4-1	-2.22	0.84198	-1.18	0.93854
Nono	IPI00320016	Q99K48-1	-2.24	0.89934	-4.43	0.82048
Rcc1;Chc1	IPI00123762	Q3U6D2	-2.60	0.58930	-3.80	0.49629
Bmi1;Bmi-1	IPI00114786	P25916	-3.31	0.10817	0.59	0.85963
Pab2;Pabp2	IPI00136169	Q8CCS6-1	-3.47	0.69084	-0.03	0.99578
Kif20b;Mphosph1	IPI00874587	Q80WE4-1	-3.78	0.07783	0.03	0.98509
Lap2	IPI00320399	Q61029-1	-4.20	0.03552	1.18	0.61120
H2afy	IPI00378480	Q9QZQ8-1	-4.46	0.00375	-0.81	0.47740
Cbx8;Pc3	IPI00135606	Q9QXV1	-5.31	0.54780	-6.14	0.57713

Cav;Cav1	IPI00117829	P49817-1	-5.37	0.10465	-2.58	0.52656
Ppp2r1a	IPI00310091	Q76MZ3	-5.49	0.09816	1.53	0.75214
H2afy	IPI00137852	Q9QZQ8-2	-5.58	0.65424	-8.19	0.56221
Tpm4	IPI00421223	Q6IRU2	-5.63	0.09223	-7.79	0.04844
Suv39h1	IPI00261633	O54864-2	-6.36	0.33417	-5.49	0.51330
Khdrbs1	IPI00458765	Q60749	-6.99	0.16309	-3.62	0.58824
Ahctf1;Elys	IPI00122594	Q8CJF7	-7.07	0.14321	-10.07	0.07576
Unnamed protein	IPI00850532		-7.36	0.00983	-8.05	0.01088
Baf57;Smarce1	IPI00119892	O54941	-7.42	0.02826	-7.32	0.03974
Kiaa1470;Rcc2	IPI00222509	Q8BK67	-7.50	0.07736	-6.75	0.14259
Enx1h;Ezh2	IPI00468525	Q61188-1	-7.86	0.00228	-10.26	0.00095
Eed	IPI00900431	Q921E6-2	-8.45	0.00254	-7.37	0.00494
Lpl	IPI00319188	P11152	-9.75	0.03803	-2.97	0.58824
Rgs18	IPI00121020	Q99PG4	-13.61	0.01919	-5.20	0.33318

11 References

Abbott, D.W., Ivanova, V.S., Wang, X., Bonner, W.M., and Ausió, J. (2001). Characterization of the stability and folding of H2A.Z chromatin particles: implications for transcriptional activation. J Biol Chem *276*, 41945–41949.

Agelopoulos, M., and Thanos, D. (2006). Epigenetic determination of a cell-specific gene expression program by ATF-2 and the histone variant macroH2A. Embo J *25*, 4843–4853.

Ahmad, K., and Henikoff, S. (2002). The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. Mol Cell *9*, 1191–1200.

Allemand, E., Hastings, M.L., Murray, M.V., Myers, M.P., and Krainer, A.R. (2007). Alternative splicing regulation by interaction of phosphatase PP2Cgamma with nucleic acid-binding protein YB-1. Nat Struct Mol Biol *14*, 630–638.

Allfrey, v.g., faulkner, r., and mirsky, a.e. (1964). acetylation and methylation of histones and their possible role in the regulation of rna synthesis. Proc Natl Acad Sci USA *51*, 786–794.

Angelov, D., Verdel, A., An, W., Bondarenko, V., Hans, F., Doyen, C.-M., Studitsky, V.M., Hamiche, A., Roeder, R.G., Bouvet, P., et al. (2004). SWI/SNF remodeling and p300-dependent transcription of histone variant H2ABbd nucleosomal arrays. Embo J *23*, 3815–3824.

Bao, Y., Konesky, K., Park, Y.-J., Rosu, S., Dyer, P.N., Rangasamy, D., Tremethick, D.J., Laybourn, P.J., and Luger, K. (2004). Nucleosomes containing the histone variant H2A.Bbd organize only 118 base pairs of DNA. Embo J *23*, 3314–3324.

Barzily-Rokni, M., Friedman, N., Ron-Bigger, S., Isaac, S., Michlin, D., and Eden, A. (2011). Synergism between DNA methylation and macroH2A1 occupancy in epigenetic silencing of the tumor suppressor gene p16(CDKN2A). Nucleic Acids Res *39*, 1326–1335.

Bassing, C.H., Suh, H., Ferguson, D.O., Chua, K.F., Manis, J., Eckersdorff, M., Gleason, M., Bronson, R., Lee, C., and Alt, F.W. (2003). Histone H2AX: a dosage-dependent suppressor of oncogenic translocations and tumors. Cell *114*, 359–370.

Bárdos, J.I., Saurin, A.J., Tissot, C., Duprez, E., and Freemont, P.S. (2000). HPC3 is a new

human polycomb orthologue that interacts and associates with RING1 and Bmi1 and has transcriptional repression properties. J Biol Chem *275*, 28785–28792.

Bedford, M.T., and Clarke, S.G. (2009). Protein arginine methylation in mammals: who, what, and why. Mol Cell *33*, 1–13.

Benjamini, Y., and Hochberg, Y. (2008). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. 1–13.

Black, B.E., and Cleveland, D.W. (2011). Epigenetic Centromere Propagation and the Nature of CENP-A Nucleosomes. Cell *144*, 471–479.

Black, B.E., Foltz, D.R., Chakravarthy, S., Luger, K., Woods, V.L., and Cleveland, D.W. (2004). Structural determinants for generating centromeric chromatin. Nature *430*, 578–582.

Bönisch, C., and Hake, S.B. (2012). Histone H2A variants in nucleosomes and chromatin: more or less stable? Nucleic Acids Res *40*, 10719–10741.

Bönisch, C., Schneider, K., Pünzeler, S., Wiedemann, S.M., Bielmeier, C., Bocola, M., Eberl, H.C., Kuegel, W., Neumann, J., Kremmer, E., et al. (2012). H2A.Z.2.2 is an alternatively spliced histone H2A.Z variant that causes severe nucleosome destabilization. Nucleic Acids Res *40*, 5951–5964.

Brockdorff, N. (2011). Chromosome silencing mechanisms in X-chromosome inactivation: unknown unknowns. Development *138*, 5057–5065.

Brush, D., Dodgson, J.B., Choi, O.R., Stevens, P.W., and Engel, J.D. (1985). Replacement variant histone genes contain intervening sequences. Mol Cell Biol *5*, 1307–1317.

Buschbeck, M., and Di Croce, L. (2010). Approaching the molecular and physiological function of macroH2A variants. Epigenetics : Official Journal of the DNA Methylation Society *5*, 118–123.

Buschbeck, M., Uribesalgo, I., Wibowo, I., Rué, P., Martin, D., Gutierrez, A., Morey, L., Guigó, R., López-Schier, H., and Di Croce, L. (2009). The histone variant macroH2A is an epigenetic regulator of key developmental genes. Nat Struct Mol Biol *16*, 1074–1079.

Busslinger, M., Portmann, R., and Birnsteil, M.L. (1979). A regulatory sequence near the 3'

end of sea urchin histone genes. Nucleic Acids Res *6*, 2997–3008.

Casas-Delucchi,C.S., van Bemmel,J.G., Haase,S., Herce,H.D.,Nowak,D.,Meilinger,D., Stear,J.H., Leonhardt,H. and Cardoso,M.C. (2012) Histone hypoacetylation is required tomaintain late replication timing of constitutive heterochromatin. *Nucleic Acids Res.*, **40**, 159–169

Cantariño, N., Douet, J., and Buschbeck, M. (2013). MacroH2A--an epigenetic regulator of cancer. Cancer Lett *336*, 247–252.

Celeste, A., Difilippantonio, S., Difilippantonio, M.J., Fernandez-Capetillo, O., Pilch, D.R., Sedelnikova, O.A., Eckhaus, M., Ried, T., Bonner, W.M., and Nussenzweig, A. (2003). H2AX haploinsufficiency modifies genomic stability and tumor susceptibility. Cell *114*, 371–383.

Chadwick, B.P., and Willard, H.F. (2001a). A novel chromatin protein, distantly related to histone H2A, is largely excluded from the inactive X chromosome. J Cell Biol *152*, 375–384.

Chadwick, B.P., and Willard, H.F. (2001b). Histone H2A variants and the inactive X chromosome: identification of a second macroH2A variant. Hum Mol Genet *10*, 1101–1113.

Chakalova, L., Debrand, E., Mitchell, J.A., Osborne, C.S., and Fraser, P. (2005). Replication and transcription: shaping the landscape of the genome. Nature Reviews Genetics *6*, 669–677.

Chakravarthy, S., Gundimella, S.K.Y., Caron, C., Perche, P.-Y., Pehrson, J.R., Khochbin, S., and Luger, K. (2005a). Structural characterization of the histone variant macroH2A. Mol Cell Biol *25*, 7616–7624.

Chakravarthy, S., Gundimella, S.K.Y., Caron, C., Perche, P.-Y., Pehrson, J.R., Khochbin, S., and Luger, K. (2005b). Structural characterization of the histone variant macroH2A. Mol Cell Biol *25*, 7616–7624.

Changolkar, L.N., Singh, G., Cui, K., Berletch, J.B., Zhao, K., Disteche, C.M., and Pehrson, J.R. (2010). Genome-wide distribution of macroH2A1 histone variants in mouse liver chromatin. Mol Cell Biol *30*, 5473–5483.

Cheng, Y.-H., Wong, E.W., and Cheng, C.Y. (2011). Cancer/testis (CT) antigens, carcinogenesis and spermatogenesis. Spermatogenesis *1*, 209–220.

Churikov, D., Siino, J., Svetlova, M., Zhang, K., Gineitis, A., Bradbury, E.M., and Zalensky, A. (2004). Novel human testis-specific histone H2B encoded by the interrupted gene on the X chromosome. Genomics *84*, 745–756.

Costanzi, C., and Pehrson, J.R. (1998). Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. Nature *393*, 599–601.

Costanzi, C., Stein, P., Worrad, D.M., Schultz, R.M., and Pehrson, J.R. (2000). Histone macroH2A1 is concentrated in the inactive X chromosome of female preimplantation mouse embryos. Development *127*, 2283–2289.

Cox, J., and Mann, M. (2012). 1471-2105-13-S16-S12. BMC Bioinformatics *13*, S12.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J.V., and Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. J Proteome Res *10*, 1794–1805.

Dai, L., Peng, C., Montellier, E., Lu, Z., Chen, Y., Ishii, H., Debernardi, A., Buchou, T., Rousseaux, S., Jin, F., et al. (2014). Lysine 2-hydroxyisobutyrylation is a widely distributed active histone mark. Nat Chem Biol *10*, 365–370.

Dardenne, E., Pierredon, S., Driouch, K., Gratadou, L., Lacroix-Triki, M., Espinoza, M.P., Zonta, E., Germann, S., Mortada, H., Villemin, J.-P., et al. (2012). Splicing switch of an epigenetic regulator by RNA helicases promotes tumor-cell invasiveness. Nat Struct Mol Biol *19*, 1139–1146.

Dazard, J.-E., Saha, S., and Ewing, R.M. (2012). ROCS: a reproducibility index and confidence score for interaction proteomics studies. BMC Bioinformatics *13*, 128.

Dietrich, N., Bracken, A.P., Trinh, E., Schjerling, C.K., Koseki, H., Rappsilber, J., Helin, K., and Hansen, K.H. (2007). Bypass of senescence by the polycomb group protein CBX8 through direct binding to the INK4A-ARF locus. Embo J *26*, 1637–1648.

Dorigo, B. (2004). Nucleosome Arrays Reveal the Two-Start Organization of the Chromatin Fiber. Science *306*, 1571–1573.

Doyen, C.-M., Montel, F., Gautier, T., Menoni, H., Claudet, C., Delacour-Larose, M., Angelov, D., Hamiche, A., Bednar, J., Faivre-Moskalenko, C., et al. (2006). Dissection of the

unusual structural and functional properties of the variant H2A.Bbd nucleosome. Embo J *25*, 4234–4244.

Drexler,H.G., Gaedicke,G., Lok,M.S., Diehl,V. and Minowada,J. (1986) Hodgkin's disease derived cell lines HDLM-2 and L-428: comparison of morphology, immunological and isoenzyme profiles.*Leuk. Res.*, **10**, 487–500.

Dryhurst, D., Ishibashi, T., Rose, K.L., Eirín-López, J.M., Mcdonald, D., Silva-Moreno, B., Veldhoen, N., Helbing, C.C., Hendzel, M.J., Shabanowitz, J., et al. (2009). Characterization of the histone H2A.Z-1 and H2A.Z-2 isoforms in vertebrates. BMC Biology *7*, 86.

Dunican, D.S., McWilliam, P., Tighe, O., Parle-McDermott, A., and Croke, D.T. (2002). Gene expression differences between the microsatellite instability (MIN) and chromosomal instability (CIN) phenotypes in colorectal cancer revealed by high-density cDNA array hybridization. Oncogene *21*, 3253–3257.

Eickbush, T.H., Godfrey, J.E., Elia, M.C., and Moudrianakis, E.N. (1988). H2a-specific proteolysis as a unique probe in the analysis of the histone octamer. J Biol Chem *263*, 18972–18978.

Eirín-López, J.M., Ishibashi, T., and Ausió, J. (2008). H2A.Bbd: a quickly evolving hypervariable mammalian histone that destabilizes nucleosomes in an acetylation-independent way. Faseb J *22*, 316–326.

Essers, J., Theil, A.F., Baldeyron, C., van Cappellen, W.A., Houtsmuller, A.B., Kanaar, R., and Vermeulen, W. (2005). Nuclear Dynamics of PCNA in DNA Replication and Repair. Mol Cell Biol *25*, 9350–9359.

Faast, R., Thonglairoam, V., Schulz, T.C., Beall, J., Wells, J.R., Taylor, H., Matthaei, K., Rathjen, P.D., Tremethick, D.J., and Lyons, I. (2001). Histone variant H2A.Z is required for early mammalian development. Curr Biol *11*, 1183–1187.

Fan, Y., Nikitina, T., Zhao, J., Fleury, T.J., Bhattacharyya, R., Bouhassira, E.E., Stein, A., Woodcock, C.L., and Skoultchi, A.I. (2005). Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation. Cell *123*, 1199–1212.

Filion, G.J., van Bemmel, J.G., Braunschweig, U., Talhout, W., Kind, J., Ward, L.D.,

Brugman, W., de Castro, I.J., Kerkhoven, R.M., Bussemaker, H.J., et al. (2010). Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. Cell *143*, 212–224.

Flemming W. (1882) Zellsubstanz, Kern un Zelltheilung. Vogel, Leipzig..

Franklin, S.G., and Zweidler, A. (1977). Non-allelic variants of histones 2a, 2b and 3 in mammals. Nature *266*, 273–275.

Fussner, E., Strauss, M., Djuric, U., Li, R., Ahmed, K., Hart, M., Ellis, J., and Bazett-Jones, D.P. (2012). Open and closed domains in the mouse genome are configured as 10-nm chromatin fibres. EMBO Rep *13*, 992–996.

Goldberg, A.D., Banaszynski, L.A., Noh, K.-M., Lewis, P.W., Elsaesser, S.J., Stadler, S., Dewell, S., Law, M., Guo, X., Li, X., et al. (2010). Distinct factors control histone variant H3.3 localization at specific genomic regions. Cell *140*, 678–691.

González-Romero, R., Méndez, J., Ausió, J., and Eirín-López, J.M. (2008). Quickly evolving histones, nucleosome stability and chromatin folding: all about histone H2A.Bbd. Gene *413*, 1–7.

Goshima, T., Shimada, M., Sharif, J., Matsuo, H., Misaki, T., Johmura, Y., Murata, K., Koseki, H., and Nakanishi, M. (2014). Mammal-specific H2A Variant, H2ABbd, Is Involved in Apoptotic Induction via Activation of NF- B Signaling Pathway. J Biol Chem *289*, 11656–11666.

Govin, J., Caron, C., Lestrat, C., Rousseaux, S., and Khochbin, S. (2004). The role of histones in chromatin remodelling during mammalian spermiogenesis. Eur J Biochem *271*, 3459–3469.

Graber, M.W., Schweinfest, C.W., Reed, C.E., Papas, T.S., and Baron, P.L. (1996). Isolation of differentially expressed genes in carcinoma of the esophagus. Ann Surg Oncol *3*, 192–197.

Greaves, I.K., Rangasamy, D., Ridgway, P., and Tremethick, D.J. (2007). H2A.Z contributes to the unique 3D structure of the centromere. Proc Natl Acad Sci USA *104*, 525–530.

Grigoryev, S.A., Arya, G., Correll, S., Woodcock, C.L., and Schlick, T. (2009). Evidence for heteromorphic chromatin fibers from analysis of nucleosome interactions. Proc Natl Acad Sci USA *106*, 13317–13322.

Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W., et al. (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature *453*, 948–951.

Guillemette, B., Bataille, A.R., Gévry, N., Adam, M., Blanchette, M., Robert, F., and Gaudreau, L. (2005). Variant histone H2A.Z is globally localized to the promoters of inactive yeast genes and regulates nucleosome positioning. PLoS Biol *3*, e384.

Guthridge, M.A., Bellosta, P., Tavoloni, N., and Basilico, C. (1997). FIN13, a novel growth factor-inducible serine-threonine phosphatase which can inhibit cell cycle progression. Mol Cell Biol *17*, 5485–5498.

Heaphy, C.M., de Wilde, R.F., Jiao, Y., Klein, A.P., Edil, B.H., Shi, C., Bettegowda, C., Rodriguez, F.J., Eberhart, C.G., Hebbar, S., et al. (2011). Altered telomeres in tumors with ATRX and DAXX mutations. Science *333*, 425.

Henikoff, S., Henikoff, J.G., Sakai, A., Loeb, G.B., and Ahmad, K. (2009). Genome-wide profiling of salt fractions maps physical properties of chromatin. Genome Research *19*, 460–469.

Hernández-Muñoz, I., Lund, A.H., van der Stoop, P., Boutsma, E., Muijrers, I., Verhoeven, E., Nusinow, D.A., Panning, B., Marahrens, Y., and Van Lohuizen, M. (2005). Stable X chromosome inactivation involves the PRC1 Polycomb complex and requires histone MACROH2A1 and the CULLIN3/SPOP ubiquitin E3 ligase. Proc Natl Acad Sci USA *102*, 7635–7640.

Hewish, D.R., and Burgoyne, L.A. (1973). Chromatin sub-structure. The digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease. Biochem Biophys Res Commun *52*, 504–510.

Hua, S., Kallen, C.B., Dhar, R., Baquero, M.T., Mason, C.E., Russell, B.A., Shah, P.K., Liu, J., Khramtsov, A., Tretiakova, M.S., et al. (2008). Genomic analysis of estrogen cascade reveals histone variant H2A.Z associated with breast cancer progression. Mol Syst Biol *4*, 188.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res *37*, 1–13.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44–57.

Iouzalen, N., Moreau, J., and Méchali, M. (1996). H2A.ZI, a new variant histone expressed during Xenopus early development exhibits several distinct features from the core histone H2A. Nucleic Acids Res *24*, 3947–3952.

Ishibashi, T., Li, A., Eirín-López, J.M., Zhao, M., Missiaen, K., Abbott, D.W., Meistrich, M., Hendzel, M.J., and Ausió, J. (2010). H2A.Bbd: an X-chromosome-encoded histone involved in mammalian spermiogenesis. Nucleic Acids Res *38*, 1780–1789.

Jansen, L.E.T., Black, B.E., Foltz, D.R., and Cleveland, D.W. (2007). Propagation of centromeric chromatin requires exit from mitosis. J Cell Biol *176*, 795–805.

Kaneko, S., Son, J., Shen, S.S., Reinberg, D., and Bonasio, R. (2013). PRC2 binds active promoters and contacts nascent RNAs in embryonic stem cells. Nat Struct Mol Biol *20*, 1258–1264.

Kapoor, A., Goldberg, M.S., Cumberland, L.K., Ratnakumar, K., Segura, M.F., Emanuel, P.O., Menendez, S., Vardabasso, C., LeRoy, G., Vidal, C.I., et al. (2010). The histone variant macroH2A suppresses melanoma progression through regulation of CDK8. Nature *468*, 1105–1109.

Kato, H., Jiang, J., Zhou, B.-R., Rozendaal, M., Feng, H., Ghirlando, R., Xiao, T.S., Straight, A.F., and Bai, Y. (2013). A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. Science *340*, 1110–1113.

Kedes, L.H., Chang, A.C., Houseman, D., and Cohen, S.N. (1975a). Isolation of histone genes from unfractionated sea urchin DNA by subculture cloning in E. coli. Nature *255*, 533–538.

Kedes, L.H., Cohn, R.H., Lowry, J.C., Chang, A.C., and Cohen, S.N. (1975b). The organization of sea urchin histone genes. Cell *6*, 359–369.

Kharchenko, P.V., Alekseyenko, A.A., Schwartz, Y.B., Minoda, A., Riddle, N.C., Ernst, J., Sabo, P.J., Larschan, E., Gorchakov, A.A., Gu, T., et al. (2011). Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. Nature *471*, 480–485.

Kimura, H., Takizawa, N., Allemand, E., Hori, T., Iborra, F.J., Nozaki, N., Muraki, M., Hagiwara, M., Krainer, A.R., Fukagawa, T., et al. (2006). A novel histone exchange factor, protein phosphatase 2Cgamma, mediates the exchange and dephosphorylation of H2A-H2B. J Cell Biol *175*, 389–400.

Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. Science *184*, 868–871.

Kornberg, R.D., and Thomas, J.O. (1974). Chromatin structure; oligomers of the histones. Science *184*, 865–868.

Kustatscher, G., Hothorn, M., Pugieux, C., Scheffzek, K., and Ladurner, A.G. (2005). Splicing regulates NAD metabolite binding to histone macroH2A. Nat Struct Mol Biol *12*, 624–625.

Lewis, P.W., Elsaesser, S.J., Noh, K.-M., Stadler, S.C., and Allis, C.D. (2010). Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. Proc Natl Acad Sci USA *107*, 14075–14080.

Li, X., Kuang, J., Shen, Y., Majer, M.M., Nelson, C.C., Parsawar, K., Heichman, K.A., and Kuwada, S.K. (2012). The atypical histone macroH2A1.2 interacts with HER-2 protein in cancer cells. J Biol Chem *287*, 23171–23183.

Li, Y., Zhu, Z., Zhang, S., Yu, D., Yu, H., Liu, L., Cao, X., Wang, L., Gao, H., and Zhu, M. (2011). ShRNA-targeted centromere protein A inhibits hepatocellular carcinoma growth. PLoS ONE *6*, e17794.

Liu, X., Li, B., and GorovskyMA (1996). Essential and nonessential histone H2A variants in Tetrahymena thermophila. Mol Cell Biol *16*, 4305–4311.

Loyola, A., Bonaldi, T., Roche, D., Imhof, A., and Almouzni, G. (2006). PTMs on H3 variants before chromatin assembly potentiate their final epigenetic state. Mol Cell *24*, 309–316.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251–260.

Luger, K., Dechassa, M.L., and Tremethick, D.J. (2012). New insights into nucleosome and

chromatin structure: an ordered state or a disordered affair? Nature Reviews Molecular Cell Biology *13*, 436–447.

Lüscher, B., Stauber, C., Schindler, R., and Schümperli, D. (1985). Faithful cell-cycle regulation of a recombinant mouse histone H4 gene is controlled by sequences in the 3'-terminal part of the gene. Proc Natl Acad Sci USA *82*, 4389–4393.

Malik, H.S., and Henikoff, S. (2003). Phylogenomics of the nucleosome. Nat Struct Biol *10*, 882–891.

Marchetti, F., and Wyrobek, A.J. (2008). DNA repair decline during mouse spermiogenesis results in the accumulation of heritable DNA damage. DNA Repair (Amst) *7*, 572–581.

Margueron, R., and Reinberg, D. (2011). The Polycomb complex PRC2 and its mark in life. Nature *469*, 343–349.

Marks, P., Rifkind, R.A., Richon, V.M., Breslow, R., Miller, T., and Kelly, W.K. (2001). Histone deacetylases and cancer: causes and therapies. Nat Rev Cancer *1*, 194–202.

Marques, M., Laflamme, L., Gervais, A.L., and Gaudreau, L. (2010). Reconciling the positive and negative roles of histone H2A.Z in gene transcription. Epigenetics : Official Journal of the DNA Methylation Society *5*, 267–272.

Matsuda, R., Hori, T., Kitamura, H., Takeuchi, K., Fukagawa, T., and Harata, M. (2010). Identification and characterization of the two isoforms of the vertebrate H2A.Z histone variant. Nucleic Acids Res *38*, 4263–4273.

Meistrich, M.L., Trostle-Weige, P.K., Lin, R., Bhatnagar, Y.M., and Allis, C.D. (1992). Highly acetylated H4 is associated with histone displacement in rat spermatids. Mol Reprod Dev *31*, 170–181.

Mietton, F., Sengupta, A.K., Molla, A., Picchi, G., Barral, S., Heliot, L., Grange, T., Wutz, A., and Dimitrov, S. (2009). Weak but uniform enrichment of the histone variant macroH2A1 along the inactive X chromosome. Mol Cell Biol *29*, 150–156.

Moldovan, G.-L., Pfander, B., and Jentsch, S. (2007). PCNA, the maestro of the replication fork. Cell *129*, 665–679.

Nekrasov, M., Soboleva, T.A., Jack, C., and Tremethick, D.J. (2013). Histone variant selectivity at the transcription start site: H2A.Z or H2A.Lap1. Nucleus *4*, 431–438.

Neumann,I., Foell,J.L., Bremer,M., Volkmer,I., Korholz,D., Burdach,S. and Staege,M.S. (2010). Retinoic acid enhances sensitivityof neuroblastoma cells for imatinib mesylate. *Pediatr. Blood Cancer*, **55**, 464–470.

Noll, M. (1974). Internal structure of the chromatin subunit. Nucleic Acids Res *1*, 1573–1578.

Nusinow, D.A., Hernández-Muñoz, I., Fazzio, T.G., Shah, G.M., Kraus, W.L., and Panning, B. (2007a). Poly(ADP-ribose) polymerase 1 is inhibited by a histone H2A variant, MacroH2A, and contributes to silencing of the inactive X chromosome. J Biol Chem *282*, 12851–12859.

Nusinow, D.A., Sharp, J.A., Morris, A., Salas, S., Plath, K., and Panning, B. (2007b). The histone domain of macroH2A1 contains several dispersed elements that are each sufficient to direct enrichment on the inactive X chromosome. J Mol Biol *371*, 11–18.

Park, Y.-J., Dyer, P.N., Tremethick, D.J., and Luger, K. (2004). A new fluorescence resonance energy transfer approach demonstrates that the histone variant H2AZ stabilizes the histone octamer within the nucleosome. J Biol Chem *279*, 24274–24282.

Perche, P.Y., Vourc'h, C., Konecny, L., Souchier, C., Robert-Nicoud, M., Dimitrov, S., and Khochbin, S. (2000). Higher concentrations of histone macroH2A in the Barr body are correlated with higher nucleosome density. Curr Biol *10*, 1531–1534.

Pogo, B.G., ALLFREY, V.G., and MIRSKY, A.E. (1966). RNA synthesis and histone acetylation during the course of gene activation in lymphocytes. Proc Natl Acad Sci USA *55*, 805–812.

Polo, S.E., Roche, D., and Almouzni, G. (2006). New histone incorporation marks sites of UV repair in human cells. Cell *127*, 481–493.

Raisner, R.M., Hartley, P.D., Meneghini, M.D., Bao, M.Z., Liu, C.L., Schreiber, S.L., Rando, O.J., and Madhani, H.D. (2005). Histone variant H2A.Z marks the 5' ends of both active and inactive genes in euchromatin. Cell *123*, 233–248.

Rasmussen, T.P. (1999). Messenger RNAs encoding mouse histone macroH2A1 female cells and result from alternative splicing. 1–5.

Rhodes, D.R., Yu, J., Shanker, K., Deshpande, N., Varambally, R., Ghosh, D., Barrette, T., Pandey, A., and Chinnaiyan, A.M. (2004). Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. Proc Natl Acad Sci USA *101*, 9309–9314.

Robinson, P.J.J., and Rhodes, D. (2006). Structure of the "30 nm" chromatin fibre: a key role for the linker histone. Curr Opin Struct Biol *16*, 336–343.

Robinson, P.J.J., Fairall, L., Huynh, V.A.T., and Rhodes, D. (2006). EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure. Proc Natl Acad Sci USA *103*, 6506–6511.

Rothbart, S.B., and Strahl, B.D. (2014). Interpreting the language of histone and DNA modifications. BBA - Gene Regulatory Mechanisms 1–17.

Salic, A., and Mitchison, T.J. (2008). A chemical method for fast and sensitive detection of DNA synthesis in vivo. Proc Natl Acad Sci USA *105*, 2415–2420.

Schalch, T., Duda, S., Sargent, D.F., and Richmond, T.J. (2005). X-ray structure of a tetranucleosome and its implications for the chromatin fibre. Nature *436*, 138–141.

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature *473*, 337–342.

Schwartzentruber, J., Korshunov, A., Liu, X.-Y., Jones, D.T.W., Pfaff, E., Jacob, K., Sturm, D., Fontebasso, A.M., Quang, D.-A.K., Tönjes, M., et al. (2012). Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. Nature *482*, 226–231.

Shia, W.-J., Li, B., and Workman, J.L. (2006). SAS-mediated acetylation of histone H4 Lys 16 is required for H2A.Z incorporation at subtelomeric regions in Saccharomyces cerevisiae. Genes Dev *20*, 2507–2512.

Singleton, S., Zalensky, A., Doncel, G.F., Morshedi, M., and Zalenskaya, I.A. (2007). Testis/sperm-specific histone 2B in the sperm of donors and subfertile patients: variability and relation to chromatin packaging. Hum Reprod *22*, 743–750.

Soboleva, T.A., Nekrasov, M., Pahwa, A., Williams, R., Huttley, G.A., and Tremethick, D.J.

(2011). A unique H2A histone variant occupies the transcriptional start site of active genes. Nat Struct Mol Biol.

Sood, V., and Brickner, J.H. (2014). Nuclear pore interactions with the genome. Curr Opin Genet Dev *25*, 43–49.

Sporn, J.C., Kustatscher, G., Hothorn, T., Collado, M., Serrano, M., Muley, T., Schnabel, P., and Ladurner, A.G. (2009). Histone macroH2A isoforms predict the risk of lung cancer recurrence. Oncogene *28*, 3423–3428.

Sporn, J.C., and Jung, B. (2012). Differential regulation and predictive potential of MacroH2A1 isoforms in colon cancer. Am J Pathol *180*, 2516–2526.

Sullivan, K.F., Hechenberger, M., and Masri, K. (1994). Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere. J Cell Biol *127*, 581–592.

Suto, R.K., Clarkson, M.J., Tremethick, D.J., and Luger, K. (2000). Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. Nat Struct Biol *7*, 1121–1124.

Szenker, E., Ray-Gallet, D., and Almouzni, G. (2011). The double face of the histone variant H3.3. Cell Res *21*, 421–434.

Tachiwana, H., Osakabe, A., Shiga, T., Miya, Y., Kimura, H., Kagawa, W., and Kurumizaka, H. (2011). Structures of human nucleosomes containing major histone H3 variants. Acta Crystallogr D Biol Crystallogr *67*, 578–583.

Takahashi, I., Kameoka, Y., and Hashimoto, K. (2002). MacroH2A1.2 binds the nuclear protein Spop. Biochim Biophys Acta *1591*, 63–68.

Talbert, P.B., Ahmad, K., Almouzni, G., Ausió, J., Berger, F., Bhalla, P.L., Bonner, W.M., Cande, W.Z., Chadwick, B.P., Chan, S.W.L., et al. (2012). A unified phylogeny-based nomenclature for histone variants. Epigenetics Chromatin *5*, 7.

Tamkun, J.W., Deuring, R., Scott, M.P., Kissinger, M., Pattatucci, A.M., Kaufman, T.C., and Kennison, J.A. (1992). brahma: a regulator of Drosophila homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. Cell *68*, 561–572.

Tan, M., Luo, H., Lee, S., Jin, F., Yang, J.S., Montellier, E., Buchou, T., Cheng, Z., Rousseaux, S., Rajagopal, N., et al. (2011). Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. Cell *146*, 1016–1028.

Thakar, A., Gupta, P., Ishibashi, T., Finn, R., Silva-Moreno, B., Uchiyama, S., Fukui, K., Tomschik, M., Ausio, J., and Zlatanova, J. (2009). H2A.Z and H3.3 histone variants affect nucleosome structure: biochemical and biophysical studies. Biochemistry *48*, 10852–10857.

Thambirajah, A.A., Dryhurst, D., Ishibashi, T., Li, A., Maffey, A.H., and Ausió, J. (2006). H2A.Z stabilizes chromatin in a way that is dependent on core histone acetylation. J Biol Chem *281*, 20036–20044.

Timinszky, G., Till, S., Hassa, P.O., Hothorn, M., Kustatscher, G., Nijmeijer, B., Colombelli, J., Altmeyer, M., Stelzer, E.H.K., Scheffzek, K., et al. (2009). A macrodomain-containing histone rearranges chromatin upon sensing PARP1 activation. Nat Struct Mol Biol *16*, 923–929.

Tolstorukov, M.Y., Goldman, J.A., Gilbert, C., Ogryzko, V., Kingston, R.E., and Park, P.J. (2012). Histone variant H2A.Bbd is associated with active transcription and mRNA processing in human cells. Mol Cell *47*, 596–607.

Unnikrishnan, A., Gafken, P.R., and Tsukiyama, T. (2010). Dynamic changes in histone acetylation regulate origins of DNA replication. Nat Struct Mol Biol *17*, 430–437.

van Daal, A., and Elgin, S.C. (1992). A histone variant, H2AvD, is essential in Drosophila melanogaster. Mol Biol Cell *3*, 593–602.

van Heeringen, S.J., Akkers, R.C., van Kruijsbergen, I., Arif, M.A., Hanssen, L.L.P., Sharifi, N., and Veenstra, G.J.C. (2014). Principles of nucleation of H3K27 methylation during embryonic development. Genome Research *24*, 401–410.

Van Holde, K.E., Sahasrabuddhe, C.G., and Shaw, B.R. (1974). A model for particulate structure in chromatin. Nucleic Acids Res *1*, 1579–1586.

Vardabasso, C., Hasson, D., Ratnakumar, K., Chung, C.-Y., Duarte, L.F., and Bernstein, E. (2013). Histone variants: emerging players in cancer biology. Cell Mol Life Sci.

Varjosalo, M., Sacco, R., Stukalov, A., Van Drogen, A., Planyavsky, M., Hauri, S., Aebersold,

R., Bennett, K.L., Colinge, J., Gstaiger, M., et al. (2013). Interlaboratory reproducibility of large-scale human protein-complex analysis by standardized AP-MS. Nat Methods *10*, 307–314.

Vogler, C., Huber, C., Waldmann, T., Ettig, R., Braun, L., Izzo, A., Daujat, S., Chassignet, I., Lopez-Contreras, A.J., Fernandez-Capetillo, O., et al. (2010). Histone H2A C-terminus regulates chromatin dynamics, remodeling, and histone H1 binding. PLoS Genet *6*, e1001234.

Wells, D., and Kedes, L. (1985). Structure of a human histone cDNA: evidence that basally expressed histone genes have intervening sequences and encode polyadenylylated mRNAs. Proc Natl Acad Sci USA *82*, 2834–2838.

West, M.H., and Bonner, W.M. (1980). Histone 2A, a heteromorphous family of eight protein species. Biochemistry *19*, 3238–3245.

Wiedemann, S.M., Mildner, S.N., Bönisch, C., Israel, L., Maiser, A., Matheisl, S., Straub, T., Merkl, R., Leonhardt, H., Kremmer, E., et al. (2010). Identification and characterization of two novel primate-specific histone H3 variants, H3.X and H3.Y. J Cell Biol *190*, 777–791.

Winkler, C., Steingrube, D.S., Altermann, W., Schlaf, G., Max, D., Kewitz, S., Emmer, A., Kornhuber, M., Banning-Eichenseer, U., and Staege, M.S. (2012). Hodgkin's lymphoma RNA-transfected dendritic cells induce cancer/testis antigen-specific immune responses. Cancer Immunol Immunother *61*, 1769–1779.

Wu, G., Broniscer, A., McEachron, T.A., Lu, C., Paugh, B.S., Becksfort, J., Qu, C., Ding, L., Huether, R., Parker, M., et al. (2012). Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. Nat Genet *44*, 251–253.

Wu, R.S., and Bonner, W.M. (1981). Separation of basal histone synthesis from S-phase histone synthesis in dividing cells. Cell *27*, 321–330.

Xu, Y., and Price, B.D. (2011). Chromatin dynamics and the repair of DNA double strand breaks. Cell Cycle *10*, 261–267.

Xu, Y., Ayrapetov, M.K., Xu, C., Gursoy-Yuzugullu, O., Hu, Y., and Price, B.D. (2012). Histone H2A.Z controls a critical chromatin remodeling step required for DNA double-strand break repair. Mol Cell *48*, 723–733.

Yeoh, K.-W., and Mikhaeel, N.G. (2011). Role of Radiotherapy in Modern Treatment of Hodgkin's Lymphoma. Advances in Hematology *2011*, 1–6.

Zalensky, A.O., Siino, J.S., Gineitis, A.A., Zalenskaya, I.A., Tomilin, N.V., Yau, P., and Bradbury, E.M. (2002). Human testis/sperm-specific histone H2B (hTSH2B). Molecular cloning and characterization. J Biol Chem *277*, 43474–43480.

Zeng, L., and Zhou, M.M. (2002). Bromodomain: an acetyl-lysine binding domain. FEBS Letters *513*, 124–128.

Zentner, G.E., and Henikoff, S. (2013). Regulation of nucleosome dynamics by histone modifications. Nat Struct Mol Biol *20*, 259–266.

Zhang, R., Poustovoitov, M.V., Ye, X., Santos, H.A., Chen, W., Daganzo, S.M., Erzberger, J.P., Serebriiskii, I.G., Canutescu, A.A., Dunbrack, R.L., et al. (2005). Formation of MacroH2A-containing senescence-associated heterochromatin foci and senescence driven by ASF1a and HIRA. Developmental Cell *8*, 19–30.

Zhou, J., Fan, J.Y., Rangasamy, D., and Tremethick, D.J. (2007). The nucleosome surface regulates chromatin compaction and couples it with transcriptional repression. Nat Struct Mol Biol *14*, 1070–1076.

Zlatanova, J., and Thakar, A. (2008). H2A.Z: view from the top. Structure *16*, 166–179.

Zucchi, I., Mento, E., Kuznetsov, V.A., Scotti, M., Valsecchi, V., Simionati, B., Vicinanza, E., Valle, G., Pilotti, S., Reinbold, R., et al. (2004). Gene expression profiles of epithelial cells microscopically isolated from a breast-invasive ductal carcinoma and a nodal metastasis. Proc Natl Acad Sci USA *101*, 18147–18152.