# Structure/function analyses of mammalian histone H2A and H3 variants

## Clemens Bönisch

## Dissertation

der Fakultät für Biologie der Ludwig-Maximilians-Universität München

zur Erlangung des akademischen Grades Doctor rerum naturalium (Dr. rer. nat.)

von Clemens Bönisch aus Halle (Saale)

München, den 24. Mai 2012

Eingereicht am 24. Mai 2012

Mündliche Prüfung am 30. Juli 2012

1. Gutachter: Prof. Dr. Peter Becker
2. Gutachter: Prof. Dr. Heinrich Leonhardt/Prof. Dr. Stefan Jentsch
3. Gutachter: Prof. Dr. Dirk Eick
4. Gutachter: Prof. Dr. Thomas Cremer

**Eidesstattliche Erklärung**

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

München, den ..................                ...............................................

(Clemens Bönisch)

# Table of Contents

Table of Contents

# PREFACE

In this cumulative PhD thesis, I combine three original papers comprising my work on histone variants, namely H2A.Z.2.2, macroH2A, and H3.X and H3.Y. In addition, a review article entitled "Chromatin proteomics and epigenetic regulatory circuits", which I partially wrote, constitutes the more general part of the introduction section. In this review, we summarize the vast field of chromatin biology and epigenetics including post-translational modifications of histones, histone variants, DNA methylation, ATP-dependent chromatin remodeling, non-coding RNAs and nuclear architecture as well as their functional interplay.

The centerpiece of this thesis is the first original paper: "H2A.Z.2.2 is an alternatively spliced histone H2A.Z variant that causes severe nucleosome destabilization". In this paper, we identified and characterized both biochemically and biophysically, a novel, putatively primate-specific histone H2A.Z splice variant named H2A.Z.2.2. Our study provides compelling evidence that H2A.Z.2.2 is highly expressed in human brain and strongly destabilizes nucleosomes.

During the process of writing and assembling this PhD thesis, my supervisor Priv.-Doz. Dr. Sandra Hake had the interesting and challenging idea to write the more specific part of the introduction and the discussion of this thesis in form of a review article. Prof. Dr. Peter Becker and Prof. Dr. Heinrich Leonhardt supported the idea from the beginning. Based on my own work on nucleosome stability and mobility presented in the original paper on H2A.Z.2.2, I decided to write a comprehensive review concerning the influence of histone H2A variants on nucleosome as well as chromatin structure and stability. The more specific part of the introduction summarizes the literature on H2A variants and their biological function; the discussion focuses on their influence on nucleosome structure, stability and higher-order chromatin structure. I decided, in accordance with the guide lines for cumulative dissertations, to put the discussion subsequent to the introduction. In order to keep the space between text and references minimal, the references appear after the discussion section and are followed by the results section, constituted by the three original papers mentioned above. Soon after the submission of this PhD thesis, a review article based on both introduction and discussion parts presented here will be submitted to Nucleic Acids Research; whose editor strongly encouraged us to do so.

This thesis does not contain an additional materials and methods section as all relevant information is presented in the here included original papers and corresponding supplementary materials and method sections. Moreover all working protocols and information on plasmids as well as DNA and RNA oligonucleotides will be kept in the laboratory of Priv.-Doz. Dr. Sandra Hake and are available upon request.

# SUMMARY

In the eukaryotic nucleus, DNA is packaged into chromatin, a nucleoprotein complex with the nucleosome as its basic subunit. The nucleosome consists of approximately 150 bp of DNA wrapped around an octamer of core histones with two copies of each H2A, H2B, H3 and H4. Chromatin is highly dynamic and allows DNA compaction to different degrees, thereby permitting or prohibiting accessibility to factors regulating DNA-related processes. Structural alterations of chromatin are strictly controlled and involve various interconnected mechanisms including the exchange of canonical histones with histone variants.

This work comprises three original papers all dealing with histone variants and their biological implications. The centerpiece is the identification and characterization of H2A.Z.2.2 (Z.2.2), a novel, putatively primate-specific splice isoform of the extensively studied histone variant H2A.Z. Z.2.2 mRNA is present in all human tissues analyzed although at lower levels when compared to the major splice isoform, Z.2.1. In the brain however, Z.2.2 mRNA is abundant and constitutes up to 50% of transcripts originating from the *H2A.Z.2* locus. Moreover, characterization of Z.2.2´s mobility *in vivo* and stability *in vitro* demonstrated that this variant strongly destabilizes nucleosomes; to an extent at least as pronounced as H2A.Bbd, the prime example of a destabilizing H2A variant thus far. Z.2.2 is identical to Z.2.1 in most of its sequence but has an altered docking domain/C-terminal sequence, which is shorter and has a stretch of six unique amino acids. Interestingly, by analysis of deletion mutants and chimeric proteins, we could show that Z.2.2´s characteristic properties are dependent on its unique docking domain sequence and structure, and that mere shortening of Z.2.1 to the same length as Z.2.2 does not restore Z.2.2´s high mobility *in vivo*. Taken together, these findings suggest a possible exciting role of Z.2.2 in conferring unique structural properties to chromatin in primate brains, thereby distinguishing them from those of other mammals.

The second paper describes the role of ATRX in regulating macroH2A´s chromatin association. MacroH2A is well characterized as a transcriptional repressor however the mechanisms by which it is targeted to its chromatin sites have remained elusive. We found that ATRX interacts with soluble macroH2A in chromatin free extracts, suggesting a possible role in macroH2A targeting. The findings that ATRX knock down leads to increased macroH2A incorporation at telomeric regions and the $\alpha$-globin gene cluster, concurrent with reduced $\alpha$-globin gene expression, support its function in negatively regulating macroH2A chromatin association. Notably, this study is only the second study pointing towards histone variant targeting by negatively influencing their chromatin association.

In the third paper, we report the identification of H3.X and H3.Y, two novel primate-specific H3 variants. These variants are, in contrast to Z.2.2, only expressed in a subset of cell lines and tissues (including some brain tissues), but H3.Y expression is significantly upregulated under metabolic stress coupled with high cell density in the human osteosarcoma cell line U2OS. Furthermore, knock down of H3.Y impairs cell proliferation by deregulation of cell cycle-specific gene expression. Since H3.Y was expressed only in a subset of neuronal cells in hippocampus, it is possible that this novel variant plays an important role in fine-tuning proliferation of some specialized cell types upon certain stimuli

# ZUSAMMENFASSUNG

DNA liegt in Eukaryoten in Form von Chromatin vor, einem Nukleoproteinkomplex dessen Grundbaustein das Nukleosom darstellt. Das Nukleosom besteht aus ca. 150 bp DNA, die um ein Histonoktamer aus jeweils zwei Kopien der Histone H2A, H2B, H3 und H4 gewunden sind. Die Chromatinstruktur ist sehr dynamisch und ermöglicht eine DNA-Kompaktierung in verschiedenen Ausmaßen, was wiederum die Zugänglichkeit von Faktoren, die in DNA-bezogene Prozessen involviert sind, entweder  erlaubt oder verhindert. Änderungen der Chromatinstruktur sind hochreguliert und werden durch etliche, miteinander verknüpfte Prozesse kontrolliert, wie z.B. durch den Austausch von kanonischen Histonen mit Histonvarianten.

Diese Dissertation enthält drei Originalartikel, die sich alle mit Histonvarianten und deren biologischen Implikationen beschäftigen. Das Herzstück dieser Doktorarbeit stellt die Identifizierung und Charakterisierung von H2A.Z.2.2 (Z.2.2), einer neuen, mutmaßlich primaten-spezifischen Spleißisoform der ausgiebig untersuchten Histonvariante H2A.Z dar. Die mRNA von Z.2.2 kommt in allen untersuchten menschlichen Geweben vor, allerdings in wesentlich geringeren Mengen als die Haupt-Spleißisoform, Z.2.1. Im Gehirn hingegen ist die Z.2.2 mRNA sehr abundant und macht bis zu 50% der Transkripte des H2A.Z.2 Locus aus. Darüberhinaus zeigte die Charakterisierung der Mobilität von Z.2.2 *in vivo* und seiner Stabilität *in vitro*, dass es Nukleosomen signifikant destabilisiert; in einem mindestens genauso großen Ausmaß wie H2A.Bbd, dem bisherigen Paradebeispiel einer destabilisierenden H2A Variante. Z.2.2 ist fast vollständig identisch mit Z.2.1 mit Ausnahme des C-terminus („docking domain"). Dieser ist kürzer und enthält eine Sequenz von sechs charakteristischen Aminosäuren. Interessanterweise konnten wir mittels Deletionsmutanten und chimären Proteinen zeigen, dass die veränderte Aminosäuresequenz und Struktur am C-Terminus von Z.2.2 entscheidend für seine charakteristische Eigenschaften sind; das bloße Verkürzen von Z.2.1 auf die gleiche Länge wie Z.2.2. führte nicht zu einer vergleichbaren Mobilität *in vivo*. Zusammengefasst lassen diese Ergebnisse vermuten, dass Z.2.2 zu einzigartigen strukturelle Eigenschaften von Chromatin im Primatengehirn führt, welche es wiederum von denen anderer Säuger unterscheiden könnte.

Der zweite Artikel beschreibt die Rolle von ATRX in der Regulation der Chromatinassoziation von macroH2A. MacroH2A ist charakterisiert worden als Transkriptionsrepressor, jedoch waren die Mechanismen wie diese Variante gezielt ins Genom lokalisiert wird bisher unklar. Knock-down von ATRX führt zu einem erhöhtem Einbau von macroH2A an Telomeren und am α-Globin Gencluster mit einer gleichzeitig verringerten Transkription dieser Gene. Diese Ergebnisse lassen stark vermuten, dass eine funktionelle Rolle von ATRX darin besteht, die Assoziation von macroH2A mit Chromatin negativ zu beeinflussen. Es ist hervorzuheben, dass dieses erst die zweite Studie ist, die den

Mechanismus der Lokalisation von Histonvarianten durch eine negative Beeinflussung der Chromatinassoziation beschreibt.

Im dritten Artikel beschreiben wir die Identifizierung und Charakterisierung der primaten-spezifischen Histonvarianten H3.X und H3.Y. Diese Varianten sind, anders als Z.2.2, nur in wenigen Geweben und Zelllinen exprimiert, darunter auch einigen speziellen menschlichen Gehirngeweben. Die Expression von H3.Y kann jedoch durch Nährstoffmangel und erhöhte Zelldichte in der humanen Osteosarcomazelllinie U2OS deutlich erhöht werden. Darüberhinaus konnten wir zeigen, dass der Knock-down von H3.Y die Zellproliferation durch eine Deregulation der Expression von Zellzyklus-spezifischen Genen verringert. Da H3.Y nur in einer Subpopulation von neuronalen Zellen im Hippocampus exprimiert wird, ist es denkbar, dass diese neue Histonvariante eine wichtige Rolle in der Feinabstimmung der Proliferation von einigen spezialisierten Zelltypen durch äußere Reize spielt.

# 1.    INTRODUCTION

In eukaryotes, DNA is organized into chromatin to fit into the constrained space of the nucleus. Generally, packaging of DNA into chromatin decreases its accessibility and consequently interferes with biological processes based upon DNA such as transcription, replication and repair but protects it from damage by different kinds of stress. Despite the immense degree of global compaction, access to DNA is achieved by local chromatin decondensation in a highly regulated manner. Chromatin is a very dynamic structure; its regulation involves several interconnected mechanisms (1) such as DNA methylation (2), ATP-dependent chromatin remodeling (3), histone post-translational modifications (PTMs) (4), non-coding RNAs (ncRNAs) (5), arrangement within the three-dimensional nuclear architecture (6) and the replacement of canonical histones by histone variants (7).

Historically, the term chromatin is derived from the Greek *chroma* (color) and was first coined about 130 years ago by Flemming who microscopically analyzed nuclei stained with basophilic dyes. He stated "The word chromatin may stand until its chemical nature is known, and meanwhile stands for that substance in the cell nucleus which is readily stained" (8,9). Although its general chemical nature was revealed, the term still stands. The distinction of chromatin into two states, euchromatin (from the Greek *eu* (good)), the part of the chromatin that is stained lightly with basic dyes, and heterochromatin (from the Greek *hetero* (different, other)), which does not, was established in 1928 by Heitz (10). Since then, a lot of progress in a more detailed and functional understanding of the chromatin structure was made. The identification of DNA as the carrier of the genetic information in 1944 (11), the structure of DNA in 1953 (12), the idea of a regular chromatin subunit, the nucleosome, in the 1970s (13-15) and the high resolution structure of the nucleosome in 1997 were important scientific hallmark events (16). Today, systems biology approaches allow a much more comprehensive analysis of chromatin and lead to fascinating new insights. It becomes more and more apparent that chromatin is more diverse and complex than the original distinction into euchromatin and heterochromatin suggested. Recent studies distinguish five (17) or nine prevalent chromatin states (18), depending on the different parameters analyzed. In the future, the combination between highly correlative approaches providing us with vast quantities of data, together with mechanistic analyses to understand causal relationships, will hopefully help us to understand this fascinating substance more and more.

The introduction section of this PhD thesis contains two parts. The first, more general, part is the review article "Chromatin proteomics and epigenetic regulatory circuits" (1) that summarizes all epigenetic mechanisms mentioned above as well as their functional interconnections. The second, more specific, part is a summary of the literature on histone H2A variants and their biological functions. Here I leave out the roles of H2A variants in manipulating basic nucleosomal properties and higher-order chromatin structure, as this is the topic of the discussion section.

# Chromatin proteomics and epigenetic regulatory circuits

**Clemens Bönisch**\*, Sonja M. Nieratschker\*, Nikos K. Orfanos and Sandra B. Hake
\*These authors contributed equally to this work

# Chromatin proteomics and epigenetic regulatory circuits

EXPERT REVIEWS

Clemens Bönisch*,
Sonja M
Nieratschker*,
Nikos K Orfanos and
Sandra B Hake†

†*Author for correspondence
Adolf-Butenandt-Institute &
Center for Integrated Protein
Science Munich (CIPSM),
Department of Molecular
Biology, Ludwig-Maximilians
University, Schillerstr. 44,
80336 Munich, Germany
Tel.: +49 892 1807 5439
Fax: +49 892 1807 5425
sandra.hake@med.
uni-muenchen.de*

*\*These authors contributed
equally to this work*

Many phenotypic changes of eukaryotic cells due to changes in gene expression depend on alterations in chromatin structure. Processes involved in the alteration of chromatin are diverse and include post-translational modifications of histone proteins, incorporation of specific histone variants, methylation of DNA and ATP-dependent chromatin remodeling. Interconnected with these processes are the localization of chromatin domains within the nuclear architecture and the appearance of various classes of noncoding regulatory RNAs. Recent experiments underscore the role of these processes in influencing diverse biological functions. However, the evidence to date implies the importance of an interplay of all these chromatin-changing functions, generating an epigenetic regulatory circuit that is still not well understood.

Deciphering the genetic code has set the foundation for understanding the biology of life and how it is determined and regulated in every living organism. Although DNA is the hereditary molecule that controls all cellular functions, pioneering findings in the research field of epigenetics have revealed many additional and combinatorial mechanisms that critically contribute to the regulation of gene expression and the subsequent defining of the cellular phenotype.

In the eukaryotic nucleus, DNA does not exist as a naked molecule, but adopts a compact formation by the assembly with basic histone proteins in a nucleoprotein complex, known as chromatin. Four different histone types are known (H2A, H2B, H3 and H4), which organize DNA into basic repeating units of chromatin: the nucleosomes. Nucleosomes consist of a cylindrical octamer core formed from one H3/H4 tetramer and two H2A/H2B dimers, surrounded by 147 bp of DNA in a left-handed superhelix [1]. Further incorporation of the linker histone H1, as well as other nonhistone structural proteins, leads to the compaction of chromatin in higher order structures, generating interphase chromatin and also highly compacted mitotic chromosomes. Thus, chromatin is the template of all DNA-related processes, and the proper regulation of its conformation is essential for the performance of many cellular functions.

Cells have developed several mechanisms to regulate chromatin structure and the subsequent accessibility of DNA, which can be broadly divided into four main categories:

• Post-translational modifications of histones (PTMs)

• Incorporation (or replacement) of histone variants

• DNA methylation

• ATP-dependent chromatin remodeling

Furthermore, recent findings also provide evidence for an important role of noncoding RNAs (ncRNAs), as well as the distinct localization of certain chromatin domains in the context of nuclear architecture in regulating chromatin-related processes. In this review, we introduce the main fields of chromatin research and give examples of novel results that connect these different chromatin pathways, which in turn generate an epigenetic regulatory circuit (FIGURE 1). We conclude with an outlook of future key questions concerning this circuit and the role of epigenetic mechanisms in disease.

## Post-translational modifications of histones

Histones are subject to numerous PTMs, including acetylation and methylation of lysine residues (mono-, di- and trimethylation), methylation of

arginine residues (mono-, asymmetrical and symmetrical dimethylation) and phosphorylation, ubiquitination, sumoylation and ADP ribosylation of other amino acids [2,3]. The number and nature of identified histone PTMs is enormous and a summary of these is reviewed in [4]. Specific enzymes catalyze both the formation and the removal of these PTMs, suggesting that they are dynamic. Histone acetyltransferases (HAT) and deacetylases (HDAC) place and remove acetylation marks, respectively, while kinases and phosphatases catalyze the phosphorylation and dephosphorylation of histones (Figure 2). Methyl marks, established by histone methyltransferases (HMTs) [5], were considered to be stable and irreversible for a long time, until enzymes that demethylate histone tails were reported recently. These enzymes include histone arginine deiminases, such as PAD4 (which catalyze the conversion to citrulline [6,7]), arginine demethylases (e.g., JMJD6 [8]) and lysine-specific demethylases (e.g., LSD1). LSD1, which belongs to the jumonji family, catalyzes the amine oxidation of mono- and dimethylated histone H3K4 [9], while other family members facilitate the removal of di- and trimethyl groups from other H3 lysine residues (Figure 2) [10,11].

The majority of PTMs has been detected on the flexible N-terminal tails of histones, which protrude from the nucleosome core. Histone tails and their PTMs participate in crucial interactions necessary for the proper regulation of the chromatin structure and can be grouped into three main classes: intranucleosomal interactions with the underlying DNA, internucleosomal interactions among adjacent nucleosomes, and interactions with specific DNA/chromatin-related proteins.

Because of their broad interaction potential, PTMs affect a variety of processes such as gene regulation, chromatin condensation, chromosome segregation, DNA replication and repair. The influence of PTMs can be ascribed to either the modification-induced electrostatic changes that influence the affinity of histone tails to the subjacent DNA and/or other nucleosomes, or to the generation of high-affinity sites for the recruitment of specific chromatin modules. Although the majority of PTMs reside on the tails of core histones, recent findings highlight the potential of the few characterized PTMs that occur on the globular core domain of histones (e.g., H3K56ac) in evoking more effective changes to chromatin structure. The drastic nature of these modifications is mainly due to their localization in the nucleosome structure, including its lateral surface, where crucial DNA–histone interactions occur. The latter type of PTMs led to the proposal of a new mechanism for the regulation of nucleosome positioning, which is summarized by the regulated nucleosome mobility model [4,12]. According to this model, the modification(s) of histone residues that interact with DNA changes the chemical affinity between histone and DNA, thereby altering the relative mobility of a nucleosome along DNA. This model also involves ATP-dependent chromatin remodeling, the contribution of which to chromatin structure is described in 'ATP-dependent chromatin remodeling'.

The "histone-code hypothesis" was proposed to explain the influence of PTMs on chromatin-related processes [2,13,14]. It suggests that PTMs act in a combinatorial way that can be read by proteins specifically recognizing certain PTMs or combinations thereof, and consequently influence distinct biological functions. Indeed, many factors, involved in diverse pathways, have been found to bind PTMs on histone tails with specific protein domains (modules), which recognize methyl-groups (chromo-, tudor-, malignant brain tumor (MBT)-domain, WD40 repeats and PHD-finger), acetyl-groups (bromodomain), phospho groups (14-3-3) and unmodified histone residues (SANT-domain and PHD-finger) (Table 1) [15,16].

Histone PTMs play an especially important role in the regulation of transcription and have been characterized as either transcriptionally active or repressive (see later). This classification is not clear cut, because some marks that are linked with gene activation are also found, in some cases, to



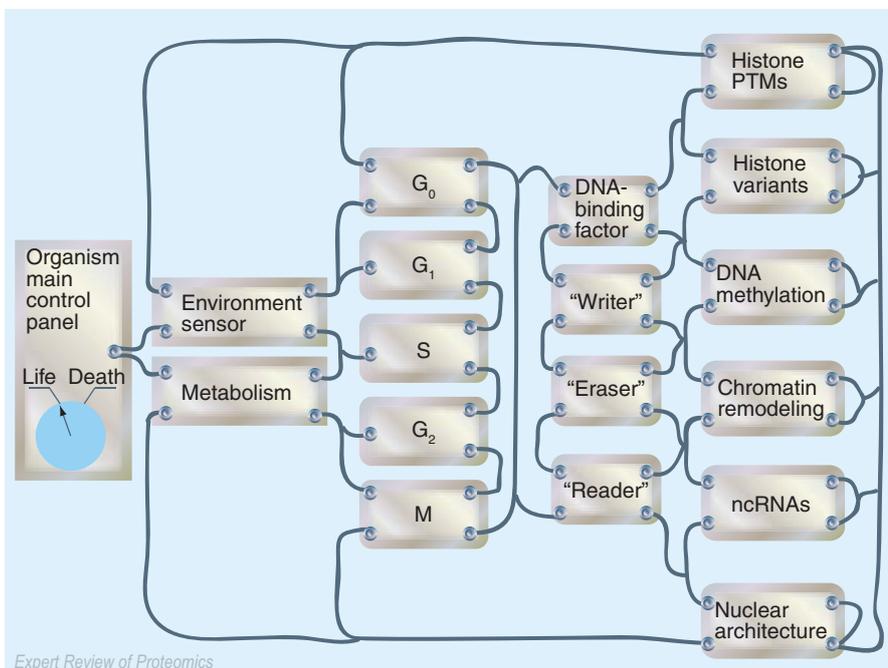**Figure 1. Epigenetic regulatory circuits play a role in the regulation of almost all biological processes in eukaryotes.** They are interconnected and involve histone modifications (PTMs), histone variants, DNA methylation, chromatin remodeling, ncRNAs and the position of the particular chromatin region within the nucleus (nuclear architecture). ncRNA: Noncoding RNA; PTM: Post-translational modification.
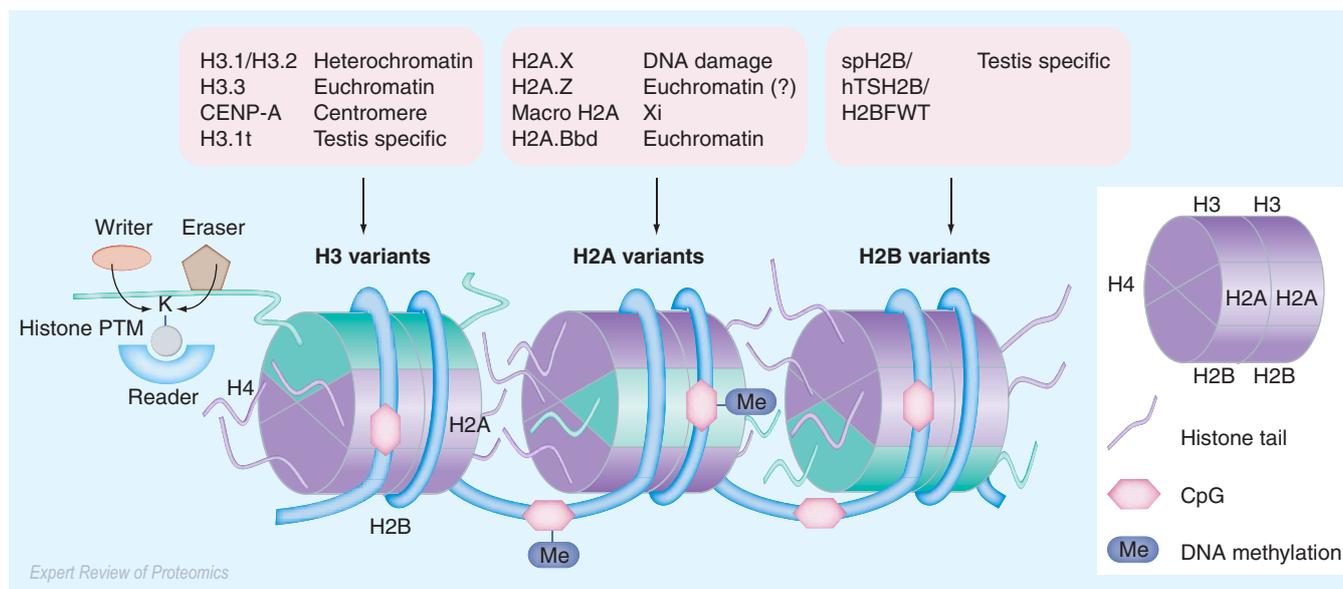
**Figure 2. Constituents of chromatin and variations therein.** Depiction of the components that constitute chromatin; DNA, histones (variants) and their respective PTMs. DNA containing methyl groups on certain CpG islands is wrapped around a nucleosome consisting of an octamer of histone proteins. Variants of the different core histone families are depicted in green in the nucleosome. **(Left)** Histone PTMs, such as lysine (K) methylation, are set by a writer and deleted by an eraser. PTMs are bound by proteins (readers) that contain specific recognition modules. **(Top)** Listing of mammalian core histone variants and their putative function/location in chromatin.

correlate with heterochromatin and vice versa. In general, acetylation of lysines and the phosphorylation of serines and threonines are associated with gene activation, sumoylation appears to repress gene transcription, while methylation and ubiquitination have variable effects, depending on the precise residues and context. On the other hand, ADP ribosylation is considered to be implicated in several chromatin-related functions, such as regulation of DNA repair pathways, chromatin structural alterations during apoptossis, regulation of higher order chromatin structures and imprinting (reviewed in [17]). Besides a single effect of one PTM on biological processes, it has been shown that certain PTMs cross-talk with each other, leading to "trans-histone" pathways. One example is the influence of monoubiquitination of histone H2B (H2Bub) on histone H3 modifications. This modification is necessary for the methylation of H3K79 and H3K4, but not for the methylation of H3K36 [18].

Although single PTMs have clearly been demonstrated to be recognized by proteins containing specialized modules (TABLE 1), studies have recently reported the existence of patterns of co-existing PTMs. For example, the repressive marks H3K27me3 and H2AK119ub have been found to occur together on transcriptionally silent [19], whereas the active marks H3K4me2/3 and H4K16ac have been found together on transcriptionally active homeotic genes, which encode important transcription factors involved in developmental processes [20]. Interestingly, bivalent domains, marked simultaneously by activating (H3K4me3) as well as repressing (H3K27me3) PTMs, have been found at key developmental genes in embryonic stem cells, indicating the plasticity of these genes in pluripotent cells [21].

These and other reports led to the generation of a multivalency-based theory, the so-called "nucleosome code" (reviewed in [22]). Here, the authors suggest that nuclear proteins known to contain numerous effector domains may recognize distinct PTMs on the same or different histones. This would lead to the formation of macromolecular complexes, which would play a role in translating this multivalent PTM code into biological function. However, direct evidence supporting this model is still lacking.

## Histone variants

Beside the canonical histones, many variants of the core histones H2A, H2B, H3 and the linker histone H1 have evolved (FIGURE 2); but no histone H4 variant has been characterized to date. The existence of histone variants was known as early as 1969; however, extensive research on these proteins is a very recent development [23]. Histone variants have the potential to displace the canonical histones from the nucleosomal octamer, altering its nature either by inducing structural changes or by changing its interaction with other factors (e.g., remodeling complexes, transcription machinery or other nucleosomes). Although this exchange procedure is still not thoroughly understood, it is highly dependent on specific proteins (histone chaperones) that deliver histone variants to the DNA. Emerging evidence indicates that there might be specific histone chaperones for different variants [24].

Histone H1 and its variants play an important role in the compaction of chromatin. Although H1 variants are differentially expressed in a developmental and tissue-specific manner,

**Table 1. General histone modifications and their respective protein families that enzymatically catalyze the formation of marks (writers), enzymes that remove PTMs (erasers) and proteins that specifically bind to histone PTMs (readers).**

| Post-translational modifications | "Writer" families | "Eraser" families | "Reader" modules |
|---|---|---|---|
| Lysine methylation | Set and Non-Set proteins | Jumonji | Chromodomain, Tudor, MBT, PHD |
| Arginine methylation | PRMT | Deiminase, Jumonji | WD40 repeat |
| Acetylation | HAT | HDAC | Bromodomain |
| Phosphorylation | Kinase | Phosphatase | 14-3-3 |
| Ubiquitination | E1, E2, E3 ligases | Deubiquitinase | ? |
| SUMOylation | SUMO | Desumoylase (?) | ? |
| ADP ribosylation | PARP | ? | ? |
| Unmodified | | | SANT, PHD |

HAT: Histone acetyltransferases; HDAC: Histone deacetylases; MBT: Malignant brain tumor; PARP: Poly(ADP-ribose) polymerase; PRMT: Protein arginine methyltransferases; SUMO: Small ubiquitin-like modifier.

they appear to be highly redundant (reviewed in [25,26]). Like the core histones, H1 is involved in epigenetic regulation, but as H1 research is such a vast field on its own (reviewed in [25]), we want to focus our attention on the functional differences of core histone variants (Figure 2). In recent years, several connections between histone variants and other epigenetic regulatory mechanisms, particularly variant-specific PTMs, have been discovered. A selection of these will be summarized in the following section.

Five H3 variants have been identified in mammals: CENP-A (centromere specific), H3.1, H3.2, H3.3 and H3.1t (testis specific); the four latter variants differing from each other in only a few amino acids. CENP-A demonstrates great sequence differences to the other H3 variants in its N-terminal tail, which is also highly divergent among species. CENP-A-containing nucleosomes localize subjacent to the active kinetochore and are critical determinants of centromeric chromatin [27]. A CENP-A-specific PTM is its phosphorylation of serine 7 during mitosis, which might play a role in cytokinesis [28].

The histone H3 variants H3.1 and H3.2 differ in only one amino acid from each other, and recent analysis of expression and modification patterns suggest that they might function in different aspects of gene repression [29].

On the other hand, several studies have demonstrated that H3.3 is associated with transcriptionally active regions in different organisms [30,31]. Interestingly, in humans, it was demonstrated that the unique serine residue at position 31 of H3.3 is phosphorylated during metaphase in a subpopulation of H3.3 located directly adjacent to centromeres. The function of this H3.3-specific PTM is presently unknown [32].

In addition to the canonical H2A, the mammalian H2A family contains the variants H2A.X, H2A.Z, MacroH2A (MacroH2A1.1, MacroH2A1.2 and MacroH2A2) and H2A.Bbd (Barr body deficient). The H2A variants differ greatly in their respective functions.

H2A.X contains a unique serine residue (S139 in mammals) that becomes phosphorylated (γ-H2A.X) at foci of DNA double-strand breaks (DSBs) and constitutes a docking site for DNA-repair proteins in many organisms. The mechanisms of recruitment are currently under investigation (see 'ATP-dependent chromatin remodeling') [33].

H2A.Z shows only approximately 60% sequence identity to the canonical H2A. Multiple studies implicated H2A.Z with the formation of heterochromatin (in higher eukaryotes) as well as the prevention of heterochromatin spreading (in yeast), suggesting that H2A.Z might have acquired diverse functions in different species [34].

H2A.Bbd is an unusually short mammalian H2A variant that is only 48% identical to H2A and is found on all chromosomes, except for the female inactive X chromosome (Xi). It co-localizes with acetylated histone H4, suggesting an association with transcriptionally active chromatin [35]. In addition, different groups could show that nucleosomes containing H2A.Bbd rather than canonical H2A are less stable and have a more relaxed conformation, suggesting that the incorporation of H2A.Bbd is another mechanism of unfolding chromatin [36–38].

MacroH2A, the longest and most unusual H2A variant, is vertebrate-specific and consists of a N-terminal H2A-like and a large C-terminal nonhistone region (the macrodomain) [30]. MacroH2A is associated with the Xi and thought to play a role in maintaining the transcriptionally inactive state of this chromosome [39]. Its macrodomain has been shown to constitute a specific-binding module for ADP-ribose and might even display enzymatic activity [40]. Additionally, ubiquitination of MacroH2A seems to be involved in the localization of this variant to the Xi [41]. Besides its role in X inactivation, recent work showed that MacroH2A inhibits transcription. Mechanistically, the macrodomain prevents acetylation of adjacent histone tails and chromatin remodeling [42], (see 'Chromatin remodeling').

MacroH2A incorporation is also influenced by the DNA methylation status of centromeric heterochromatin. The loss of DNA methylation leads to increased incorporation of MacroH2A into pericentric regions and thus to the reorganization of heterochromatin [43].

In mammals, only three bona fide H2B variants (hTSH2B, H2BFWT and spH2B) have been found to date. All of these variants reside in the testis and their functions are largely unknown (reviewed in [30]).

Multiple examples for functional connections between histone variants and chromatin remodeling complexes are known and will be discussed in the following parts.

## DNA methylation

Virtually all vertebrates, many plants and even many invertebrates contain chemical modifications on cytosine bases at distinct positions in the genome that heritably influence the regulation of genetic information (FIGURE 2). DNA methylation, perhaps the best characterized epigenetic modification, results in long-term silencing of the underlying sequence and is essential for many different cellular functions, including cell differentiation, gene regulation, genomic imprinting, X inactivation and aging [44,45]. Two mechanisms contribute to the inhibition of gene transcription:

• DNA methylation inhibits the association of some DNA-binding factors with their recognition sequence [46]

• Methylated CpG dinucleotides recruit specific binding factors involved in gene silencing [47,48]

DNA methylation patterns are highly dynamic and are established during early mammalian development with differences between the paternal and maternal genomes [49,50]. Defects in DNA methylation lead to diverse disorders, ranging from mental retardation and immune deficiencies to malignant transformations [51–54].

Mammals express a number of DNA (cytosine 5) methyltransferase (DNMTs) isoforms, which are responsible for the establishment and maintenance of DNA methylation patterns [55]. These are expressed in cell type- and stage-specific manners due to sex-specific germline promoters [56,57]. In vertebrates, DNA methylation sites are (with minor exceptions) almost exclusively CpG dinucleotides [44], whereas in plants, cytosine methylation also occurs outside of CpGs [58]. The methylation status of the promoters of key transcription factors for pluripotency, Oct-4 and Nanog, correlates well with their expression pattern [59] and tissue-specific demethylation of the Il-2 and Sry promoters leads to transcriptional activation [60,61]. Large-scale analyses (bisulfite sequencing) demonstrated that the DNA methylome differs between embryonic stem cells and differentiated cells, and also between normal and tumor tissue [45].

DNA methylation plays a key role in genomic imprinting of a selected panel of genes where only one of the two parental alleles in mammals and plants is modified. In mammals, imprinted genes are organized in large chromosomal clusters with DNA methylation and gene expression patterns depending on whether they are inherited from the mother or from the father [62]. Genomic imprinting involves both DNA and histone methylation and is controlled by crucial DNA sequence elements called imprinting control regions (ICRs), which are rich in CpG dinucleotides. During female gametogenesis, most ICRs are methylated, whereas during spermatogenesis, methylation occurs only at some ICRs. Following fertilization, methylation of ICRs is maintained resulting in different methylation patterns on paternal and maternal chromosomes in somatic tissues. DNA methylation is only removed by passage through the newly formed germline in the developing embryo.

DNA methylation has been connected to other epigenetic modules in several recent studies within diverse organisms. In many organisms, particularly plants, DNA methylation has been linked to histone PTMs and ncRNAs (see 'Non-coding RNAs'). There is emerging evidence that methyl-CpG binding proteins, such as MeCP2 or methyl-CpG-binding domain protein 2 (MBD2), are found in large protein complexes containing HDACs and chromatin remodeling factors (see 'ATP-dependent chromatin remodeling') [63]. All known DNA methylation in *Neurospora crassa* and some known DNA methylation in *Arabidopsis* depends on methylation of H3K9 [64–66]. In mouse embryonic stem cells, major satellite DNA in pericentric regions, but not minor satellite repeats and endogenous retroviral sequences, show reduced DNA methylation in *Suv39h* (the enzyme responsible for H3K9me3) mutants [67].

Recently, a negative influence of histone PTMs on DNA methylation has been established. It was shown that a DNMT3 complex, containing the enzymatically active subunits DNMT3A2 and DNMT3B as well as the adapter protein DNMT3L, interacts with the N-terminal tail of H3 and in turn methylates the underlying DNA sequence. Interestingly, methylation of H3K4 strongly inhibits binding of this complex, thereby preventing DNA methylation and gene silencing [68]. These data suggest that PTMs and other chromatin modifying activities have strong influences on DNA methylation.

## ATP-dependent chromatin remodeling

In addition to histone PTMs, histone variants and DNA methylation, which are modifications of the constituents of chromatin, the modulation of nucleosomes themselves is another important principle affecting the accessibility of genomic information. This includes ATP-dependent chromatin or nucleosome remodeling, which is achieved by remodeling 'machines', which are most often large multisubunit complexes, using ATP hydrolysis to mobilize histones by weakening histone–DNA contacts. This can result in different alterations of the chromatin substrate, such as relocalization of nucleosomes on the DNA (sliding), changes in the accessibility of the DNA to interacting proteins (e.g., transcription factors) or eviction of histones (reviewed in [69–71]).

In general, remodeling complexes are classified by their ATPase subunit, which belongs to one of the four subfamilies of the Swi2/Snf2 family of DNA helicases/ATPases: ISWI, SWI/SNF, CHD and INO80 [70]. Despite the multitude of complexes and their biological functions, remodelers appear to share the same mechanistic principle: detaching a piece of DNA and subsequent migration of this bulge around the nucleosomal superhelix (bulge propagation) [70,71]. The remodeling reaction can result in both increase or decrease of DNA accessibility, depending on the respective complex [71].

Recruitment of remodelers to their substrate involves genetic as well as epigenetic determinants. Besides recruitment of remodelers by interaction with sequence-specific DNA-binding proteins, histone modifications can target remodelers that contain several modules that are able to interact with either methylated (chromodomain, PHD-finger) or acetylated (bromodomain) lysines (reviewed in [71,72]). Pray-Grant et al. provided the first example of the direct binding of a chromodomain-containing remodeler (Chd1) to a methylated lysine residue (H3K4me2) [73]. Recently, Wysocka et al. coupled the recruitment of the NURF chromatin remodeling complex via interaction of a particular PHD finger in its subunit BPTF to H3K4me3, which leads to gene activation [74]. However, remodeler binding can also be inhibited by specific histone modifications. As Clapier et al. showed, ISWI binding to chromatin is inhibited by PTMs (H4K12ac or H4K16ac, respectively), thereby masking its recognition epitope [75].

Additionally, it has been demonstrated that chromatin remodeling is not only an effect, but also a cause of histone modifications. One example is the cell cycle-regulated expression of the *HO* gene in budding yeast, which encodes an endonuclease participating in mating type switching. The remodeling complex SWI/SNF is required for the recruitment of the HAT complex SAGA to the *HO* gene promoter and the acetylation of nucleosomes residing there, which leads to *HO* gene activation [76,77]. Adding another layer of complexity, remodelers can also be substrates of histone-modifying enzymes, as recently reported for ISWI, which can be acetylated by the HAT Gcn5 [78].

The biological significance of the interaction of remodelers and histone core PTMs can be explained by the regulated nucleosome mobility model (see also 'Histone variants') [4,12]. The model proposes that remodelers (possibly recruited by PTM(s) of the N-terminal tail) increase the histone accessibility for modifying enzymes, which consequently modify residues of the globular histone core domain. This would lead to an altered histone–DNA affinity and thus changes in nucleosome mobility.

Additionally, histone variants and chromatin remodeling influence each other. While it has been shown that MacroH2A can inhibit remodeling by both SWI/SNF and ACF (see 'Histone variants') [39,42], different complexes have been found to only act in concert with certain variants. For example, exchange of H2A/H2B dimers towards Htz1/H2B

(the yeast homolog of H2A.Z) requires the SWR1 remodeling complex in *Saccharomyces cerevisiae* [79–81], while incorporation of H3.3 into chromatin in the decondensing male pronucleus of *Drosophila* embryos requires the remodeler Chd1 [82]. The latter study also highlights the importance of ATP-dependent chromatin remodeling for *de novo* nucleosome assembly *in vivo*.
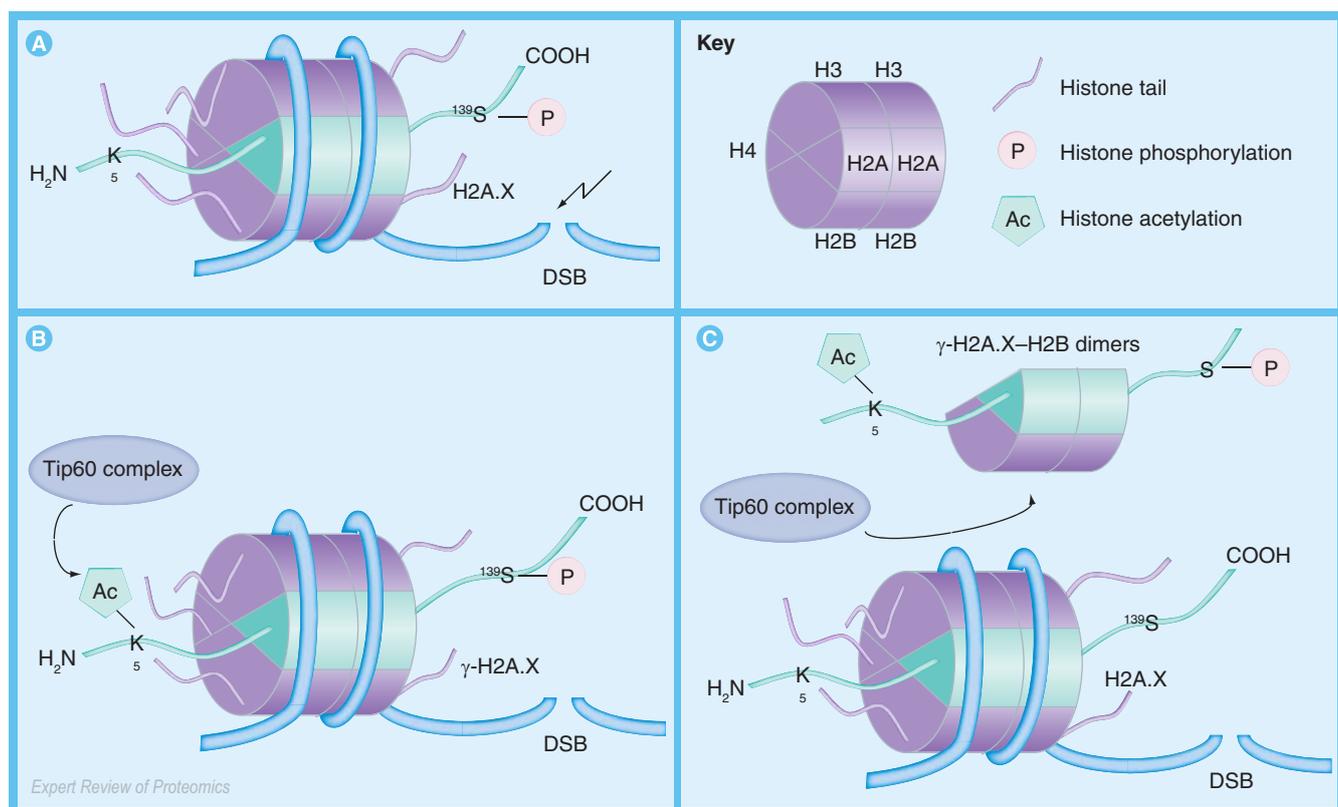
Furthermore, ATP-dependent chromatin remodeling has also been linked to both histone PTMs and histone variants. One example of this circuit can be found in the DSB repair process. The histone variant H2A.X is recruited to sites of DSBs and phosphorylated on serine 139 (γ-H2A.X) (FIGURE 3A). Kusch et al. demonstrated that the *Drosophila* homolog of the Tip60 complex, which contains the HAT Tip60 and the chromatin remodeler Domino, acetylates lysine 5 of the phosphorylated *Drosophila* homolog of H2A.X (γ-H2A.X) shortly after the occurrence of DSBs (FIGURE 3B). This acetylated form of γ-H2A.X is then exchanged with an unmodified H2A.X through the ATPase Domino (FIGURE 3C) [83]. During repair, this potentially creates a new sensor site for DNA damage occurring in the future. Different remodeling complexes, such as RSC, SWI/SNF, INO80 and SWR, have also been implicated in the cellular response to DNA damage, indicating that remodeling activity in connection with other chromatin-related processes, is important for further mechanisms extending gene regulation [84].

Not only histone PTMs can recruit ATP-dependent remodeling complexes, but methylated DNA can also target remodeling activity [69]. One example connecting histone modifications, DNA methylation and chromatin remodeling is the human MeCP1 complex. It consists of the NuRD remodeling complex, which, in addition to the remodeling activity, also contains several histone deacetylases and MBD2 (a methyl-DNA binding protein). Functional analyses of the MeCP1 complex revealed that it preferentially binds to methylated DNA and remodels and deacetylates nucleosomes. These events lead to subsequent gene silencing [85,86].

## Noncoding RNAs

Noncoding RNAs (ncRNAs) are a heterogeneous group of RNAs that are not translated into proteins. They can be categorized based on their regulatory potential. ncRNAs such as rRNAs, tRNAs, snRNAs and snoRNAs, are more or less directly involved in the mechanism of gene expression, but steadily growing evidence illustrates the importance of ncRNAs in the regulation of gene expression. In this section, we want to introduce the main classes of regulatory ncRNAs and discuss their connections with other epigenetic mechanisms. ncRNAs can act on the locus they are derived from (*cis*) or they act on another locus (*trans*).

Since the discovery of RNAi in *Caenorhabditis elegans* in 1998 [87], multiple subgroups of small ncRNAs, 19–31 nucleotides in length, have been identified. These small ncRNAs function in different gene silencing pathways and

**Figure 3. Connection of post-translational modifications, histone variants and remodeling. (A)** The histone variant H2A.X is recruited to sites of DNA DSBs and phosphorylated on serine 139 (γ-H2A.X). **(B)** The histone acetyltransferase Tip60 acetylates lysine 5 (K5) on γ-H2A.X, leading to **(C)** the exchange of γ-H2A.X-H2B dimers with unmodified H2A.X–H2B dimers through the ATP-dependent remodeler Domino within the same complex.
DSB: Double strand break.

include, among others, siRNAs and miRNAs. Both are derived from a long double-stranded precursor RNA processed by Dicer, and incorporated into the RNA-induced silencing complex (RISC), which is responsible for targeting the respective mRNA transcript. The subunits directly interacting with these small ncRNAs belong to the family of Argonaute (Ago) proteins. In general, siRNAs mediate cleavage of target mRNAs, whereas miRNAs repress their translation (reviewed in [88]).

In addition to the post-transcriptional gene silencing described above, small ncRNAs can promote transcriptional silencing by different mechanisms. One such mechanism is the induction of histone modifications, for which the assembly of heterochromatin in *Schizosaccharomyces pombe* is an example. The siRNA component of the RNA-induced initiation of transcriptional gene silencing complex (RITS) is responsible for targeting the complex to homologous sequences, subsequently leading to a distribution of RITS over heterochromatic regions [89–91]. Additionally, the chromo-domain of one of its subunits (Chp1) binds to histone H3 methylated on lysine 9, thus localizing RITS to chromatin [89]. In a later step, H3K9 methylation is spread by the recruitment of the HMT Clr4, leading to the assembly and maintenance of heterochromatin [92].

Interestingly, small ncRNAs have also been connected to DNA methylation by a process called RNA-directed DNA methylation, which is found in plants (reviewed in [93]). These small ncRNAs, many of which originate from transposons and repeats, are targeted to their homologous DNA sequences, to which they recruit DNMTs that, in turn, methylate the underlying DNA. This process is thought to play a role in the regulation of plant development and perhaps stress response.

Only recently, two independent groups demonstrated that small ncRNAs are not only able to repress gene expression, but can also activate it (dsRNA-induced gene activation [RNAa]); a process that combines histone PTMs with ncRNAs [94,95]. Li *et al.* designed double-stranded RNAs against promoter regions of different human genes. Surprisingly, instead of silencing genes, these RNAs led to a long-lasting, sequence-specific transcriptional activation of the respective target gene *in vivo*. The observed effect was Ago2-dependent and occurred with a loss of repressive H3K9 methylation at target sites. By targeting other regions in the respective promoters with the same method, it was still possible to repress transcription of these genes [95]. Although these results clearly show the possibility of activating gene expression by small ncRNAs, to date no evidence has been found that this is an endogenous mechanism.
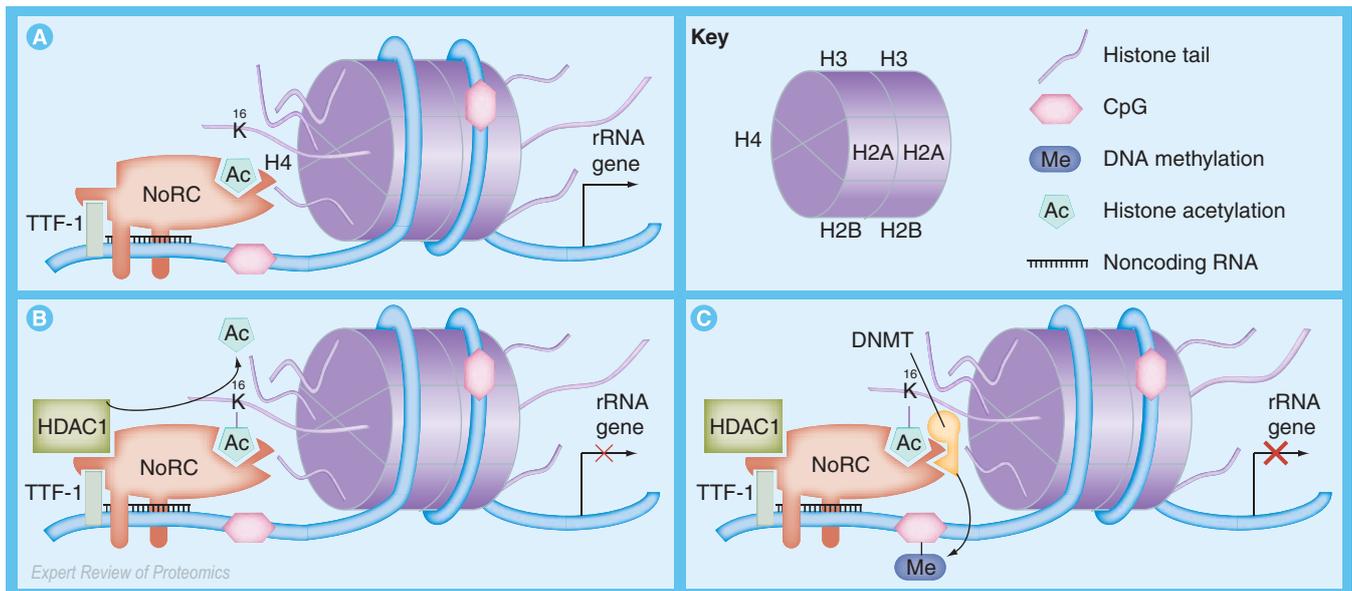
**Figure 4. Connection between post-translational modifications, ATP-dependent remodeling, DNA methylation and noncoding RNAs. (A)** Interactions of the remodeler complex NoRC with TTF-1, H4K16ac (via a bromodomain) and an intergenic ncRNA from the locus lead to its recruitment to the rDNA locus. **(B)** NoRC remodels chromatin (not shown) and recruits HDAC1, which deacetylates histones. **(C)** Binding of a DNA methyltransferase to NoRC leads to CpG methylation and subsequent silencing of the locus.
ncRNA: Noncoding RNA; NoRC: Nucleolar remodeling complex; rRNA: Ribosomal RNA.

In addition to the small ncRNAs described thus far, there are a number of long(er) ncRNAs that are also implicated in the regulation of gene expression. In the following we will provide examples of epigenetic circuits that involve long ncRNAs, histone PTMs, histone variants, DNA methylation and ATP-dependent chromatin remodeling.

One example is the regulation of *HOX* genes during development. *HOX* genes encode transcription factors, which are crucial for establishing morphologies along the head–tail axis in the embryos of various organisms. They are clustered along chromosomes in a co-linear arrangement, with their expression along the body axis [96]. A recent analysis of transcriptional activity at human *HOX* clusters demonstrated a positive correlation between antisense transcription from intergenic regions at the *HOXA* locus and the activity of neighboring *HOXA* genes. *HOXA* gene activation is accompanied by the loss of the polycomb repressive complex PRC2, which methylates H3K27, and the simultaneous increase in histone marks associated with active chromatin. The ncRNAs derived from the *HOXA* cluster therefore act in *cis* to induce and maintain the active state of this cluster [97]. An example for *trans* repression is the long ncRNA HOTAIR (2.2 kb), which is encoded in the *HOXC* cluster. HOTAIR has been identified to act in *trans* to repress the transcription of the *HOXD* locus. HOTAIR was shown to directly recruit PRC2 to the *HOXD* cluster, thus mediating H3K27me3 at this locus [98]. These findings implicate the manifold roles of ncRNAs in the regulation of *HOX* gene transcription.

Another remarkable example for the concerted action of epigenetic mechanisms is the transcriptional silencing of rRNA genes (rDNA) in mammals (FIGURE 4). In this scenario, the nucleolar remodeling complex (NoRC) is targeted to the rDNA locus by specific interactions with the transcription termination factor TTF-1 [99], with H4K16ac (via a bromodomain) [100] and with a long ncRNA derived from the intergenic spacer that separates rRNA genes (FIGURE 4A) [101]. NoRC remodels chromatin and mediates histone deacetylation by recruitment of HDAC1 (FIGURE 4B). Subsequently, NoRC recruits a DNA methyltransferase (DNMT) leading to CpG methylation and gene silencing (FIGURE 4C). Interestingly, this epigenetic state can be inherited by the next generation [102].

One of the best-studied examples for the role of long ncRNAs in gene regulation is found during dosage compensation in flies and mammals. In both cases, females have two X chromosomes while males have only one. In *Drosophila*, dosage compensation is achieved by approximately twofold upregulation of X-linked gene transcription in male flies. Here, X-encoded ncRNAs, together with different proteins, build the dosage compensation complex that has been shown to acetylate H4K16 [103].

By contrast, random X inactivation occurs in the mammalian female embryo (FIGURE 5). The initiation of X inactivation depends on the expression of the 19-kb ncRNA Xi-specific transcript (Xist), which is transcribed from specific X-linked DNA sequences termed the 'X-inactivation center' (Xic). Xist coats the X chromosome and induces silencing in *cis* (FIGURE 5A). At later stages, the repressed state of the Xi is no

longer Xist dependent, but becomes locked by several changes in its chromatin structure (reviewed in [104]). The establishment of X-inactivation starts with the transient localization of PRC2 to the Xi to set the H3K27me3 mark, and subsequent recruitment of PRC1 leading to ubiquitination of H2AK119 (FIGURE 5B). To further lock the inactive state, the histone variant MacroH2A is incorporated into nucleosomes of the Xi and the promoter DNA of X-linked genes is methylated by DNMT (FIGURE 5C). The recruitment of MacroH2A appears to be Xist dependent, however, no direct association between the histone variant and the long ncRNA has been observed to date. Consistent with its heterochromatic state, the Xi is spatially segregated from its active counterpart and associates with the nuclear scaffold protein SAF-A. This interaction helps the formation of a repressive nuclear compartment (reviewed in [105]).

The examples presented above clearly provide evidence for epigenetic regulatory circuits connecting ncRNAs and other components of the chromatin-modifying machinery.

## Nuclear architecture

In the nucleus, chromatin is organized in a nonrandom, highly structured yet dynamic fashion that possibly determines the expression status of genes. In this section, we will first introduce certain features that constitute nuclear architecture and then present some examples that functionally connect nuclear architecture with the other epigenetic circuits.

Chromosomes occupy certain positions, so-called 'chromosome territories' (CTs). Although recent findings revealed that a large proportion of the genome is actually transcribed [106], active and gene-rich (euchromatic) chromosome sections are, generally, located in the interior, silent and gene-poor (heterochromatic) sections are located at the periphery of the nucleus (reviewed in [107,108]). An exception to this rule can be found in yeast. In this organism, some actively transcribed loci are located closely to the nuclear pores, suggesting that facilitating RNA export (gene gating) has an exceptionally positive effect on transcription in the small yeast nucleus (reviewed in [107,109]).
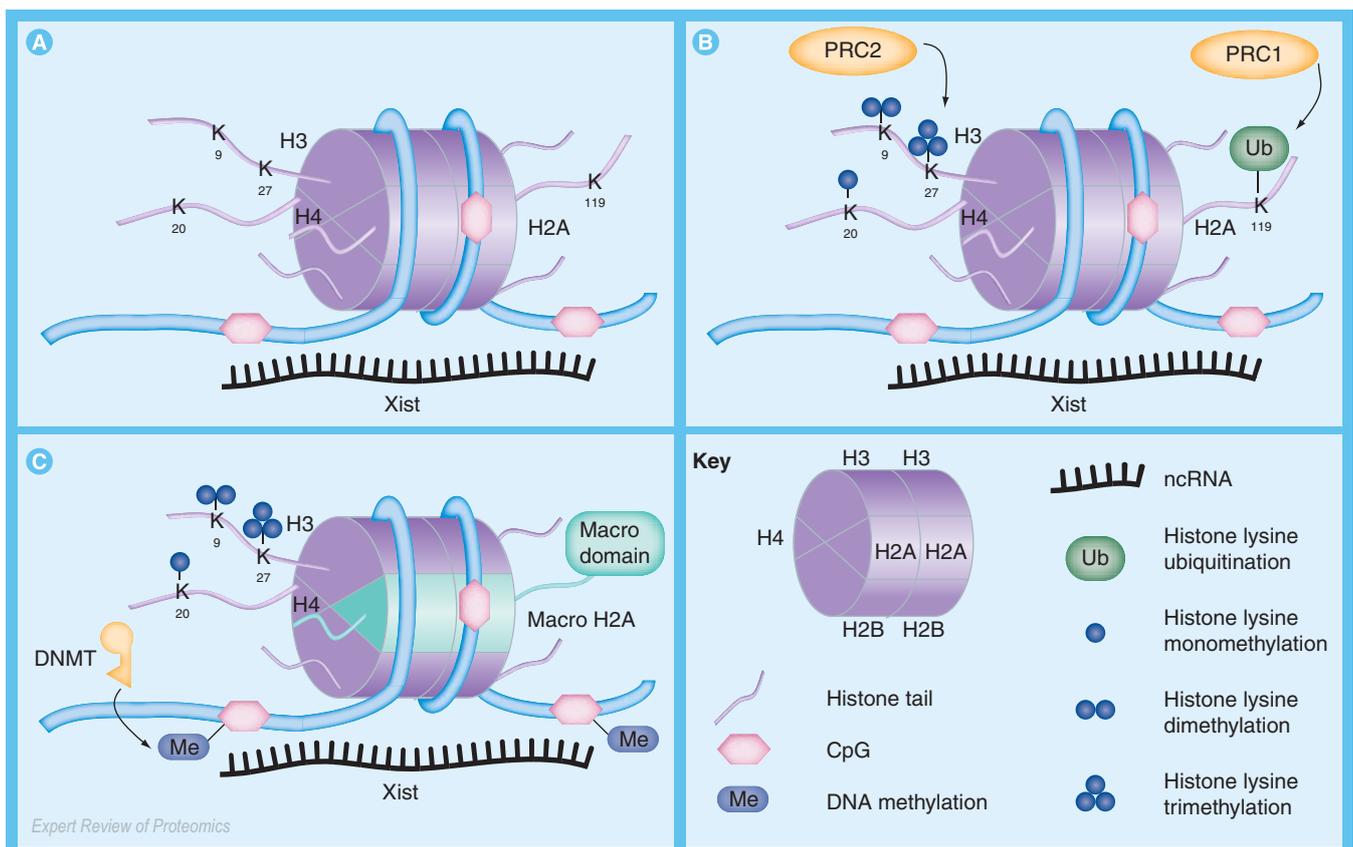


**Figure 5. Connection between post-translational modifications, histone variants, DNA methylation, long noncoding RNAs and nuclear architecture on the Xi. (A)** Initiation of mammalian X chromosome inactivation. The female X chromosome loses its active marks (not shown) and the long ncRNA Xist starts to coat the Xi. **(B)** Establishment of X inactivation. PRC1 and PRC2 are recruited to the Xi and ubiquitinate H2AK119 and methylate H3K27, respectively. **(C)** Maintenance of X inactivation. The promoter sequence of X-linked genes is methylated by a DNMT and the histone variant MacroH2A is incorporated. Eventually, the Xi associates with the nuclear scaffold protein SAF-A, establishing a repressive nuclear compartment (not shown, see also 'Noncoding RNAs').
ncRNA: Noncoding RNA; Ub: Ubiquitination.

In addition to CTs, the nucleus contains several types of structural and functional entities not surrounded by membranes, such as the nuclear lamina (scaffolding function), nucleoli (production and assembly of ribosome components), promyelocytic leukemia (PML) bodies (associated with many biological functions; e.g., transcription, apoptosis and proteolysis), Cajal bodies (assembly and modification of the transcription machinery) and nuclear speckles (enriched in splicing factors); (for review see [109,110]).

Transcription and RNA processing occur in the interchromatin (IC) compartment, which constitutes approximately half of the nucleus without detectable amounts of DNA (reviewed in [108,111]). It is expected that this compartment consists of channels and cavities very different in size, pervading the CTs and thereby leading to an enormous increase of accessible chromatin surface [111]. At the periphery of compact chromatin domains resides the perichromatin region. This region consists of decondensed chromatin looping out of its chromosome territory. Interestingly, these chromatin loops can interact with each other in an event termed 'gene kissing', even if they are located far away on the same chromosome (*cis*) or on different chromosomes (*trans*) [111]. The resulting relocalization of chromosomal loci also contributes to the nuclear architecture, albeit at a smaller scale.

It has been proposed that actively transcribed loci and regulatory elements come together in distinct foci of active transcription, so-called 'transcription factories', where RNA Pol II and associated factors are enriched. Although there is some experimental evidence for their existence [112] (for reviews see [107,113]), location within such foci does not appear to be a requirement for transcription of the respective gene [114]. Consequently, the biological relevance and generality of transcription factories is still of some controversy [114].

An intriguing example for gene activation by interaction of certain loci in concert with DNA methylation is the receptor choice in olfactory sensory neurons in mouse. In these cells, only one out of 1300 odorant receptor (OR) genes of the genome is expressed. Lomvardas *et al.* found that this is caused by the stochastic interaction of a monoallelic functional enhancer sequence with an *OR* gene leading to this gene's expression [115]. Remarkably, only in sensory neurons, one of these enhancer alleles carries a nonclassical CpA methylation, which could be involved in defining the functionality of the enhancer sequence.

In addition, DNA methylation influences the structure of imprinted loci causing their differential expression (for details about imprinting, see 'DNA methylation'). A prominent example is the *Igf2–H19* locus in mammals where only the paternal allele of *Igf2* is expressed. Transcription of the *Igf2* gene requires its interaction with downstream enhancers; both are separated by the ICR *H19*. This ICR, as well as sequences flanking the *Igf2* gene, contain several binding sites for the insulator protein CCCTC-binding factor (CTCF). Insulator proteins act by blocking interactions of genes with enhancers by binding to insulator elements, which are often located in between these sequences. CTCF binding to its recognition site is abolished by CpG methylation, which is present on the paternal, but not on the maternal, allele. Hence, CTCF exclusively interacts with its binding sites within the maternal ICR and *Igf2* flanking sequences leading to the looping of the *Igf2* gene. This event prevents the association with the downstream enhancers and results in transcriptional silencing of the maternal allele. In addition, it was shown by different chromosome conformation capture methods, which allow the detection of loci interactions at a molecular level, that other genomic loci in *cis* and *trans* are regulated by the *H19* ICR via long-range contacts mediated by CTCF (reviewed in [107,116]).

Colocalization of loci can also involve writers of histone PTMs. One example is the gene silencing by Polycomb Group (PcG) proteins in *Drosophila*. PcG proteins specifically recognize DNA sequences, termed 'PcG response elements' (PREs), and lead to heterochromatin formation with subsequent gene silencing (reviewed in [19,117]). This mechanism also involves the colocalization of genomic loci, either of homologous PRE sequences [118] or of independent chromosomal loci [119]. Besides the involvement of chromatin modifiers, this colocalization also requires components of the RNAi machinery, which also co-localizes with PcG bodies [120].

Interestingly, the mechanism of gene silencing at the nuclear envelope functionally connects the position of a gene with epigenetic processes. This mechanism involves interactions of lamin-associated polypeptides (LAPs) with proteins that promote heterochromatin formation ("writers" and "readers" of histone PTMs; for review see [121]). The nuclear envelope protein lamin B receptor has been shown to interact with heterochromatin protein 1 (HP1), and LAP2β was found to bind to HDAC3 (reviewed in [109]). These results demonstrated that specific nuclear compartments are able to recruit different proteins involved in epigenetic processes, which in turn change the structure of chromatin.

Different compaction levels of chromatin (i.e., establishment of either decondensed euchromatin or condensed heterochromatin) influence nuclear architecture and subnuclear localization. Hence, all mechanisms involved in assembly and maintenance of these specific chromatin structures, such as histone PTMs and variants, chromatin remodeling, DNA methylation and ncRNAs, might also contribute to nuclear architecture and vice versa.

## Expert commentary & five-year view

In the coming years, it will be crucial to combine the knowledge researchers have gathered about the different components that constitute and regulate chromatin and its structure into a multidimensional picture that connects all of these aspects. Of course, much work is still needed to obtain a complete picture of all histone PTMs and their functional relationship, to identify all histone variants and their functions, and to understand why so

many ATP-dependent remodelers are so abundant and how their specificity is determined. We also need to enhance our understanding of the functions of different ncRNAs, especially after the finding that almost the complete human genome is transcribed, an observation that changes our view of what constitutes a gene [106]. There are many basic questions to be answered in the next years, for example, how ncRNAs regulate chromatin structure and how the specificity in their regulation is achieved. Furthermore, we still know very little about the relationship between nuclear architecture and epigenetic processes. Does the position of a chromatin domain in the nucleus influence its gene expression or is the specific nuclear location just a consequence of its regulation? What is cause and what is consequence?

Many chromatin-related processes have been shown to play a significant role in the development of various diseases when deregulated. These include histone PTMs, readers and writers of PTMs, DNA methylation and ATP-dependent chromatin remodelers (reviewed in [122,123]). Based on these and numerous other reports, efforts are underway to develop therapeutics that target epigenetic mechanisms. One example is suberoylanilide hydroxamic acid (SAHA), which belongs to the class of HDAC inhibitors, and was recently approved for the treatment of cutaneous T-cell lymphoma [124]. In addition, soon after discovering RNAi as a mechanism to silence genes in mammalian cells, Phase I clinical studies started to evaluate the therapeutic efficacy of this system. The first results obtained from numerous studies in different organisms are promising. However, problems remain, including the delivery of the RNA to the affected tissue, its half-life and potential off-target effects. Although these problems remain to be overcome, the enormous potential of siRNA drugs can already be anticipated from these studies (reviewed in [125]). But not only siRNAs open huge opportunities for therapeutic use, also the recently discovered RNAas have been proposed for this purpose. Regardless of whether RNAa is a biological phenomenon or not, one can imagine RNAas as another useful tool for treating various diseases [94,95].

In addition to diseases such as cancer, epigenetic alterations, such as DNA methylation, have been found to contribute to the process of aging, which in turn might increase the risk of developing cancer. Global hypomethylation and CpG island hypermethylation accumulate with increasing age and might lead to gene deregulation (reviewed in [126]). These and many other observations clearly demonstrate that many different epigenetic mechanisms contribute to the variation in the chromatin polymer and have far-reaching implications for biology and human health.

The next 5 years will hopefully shed light on many open questions regarding the interplay between many chromatin-related processes which we tried to overview here, and the mechanisms by which chromatin-based information is carried from one generation to the next. Only when we learn more about the functional connections of all epigenetic "players" involved, which constitute the epigenetic regulatory circuit, will we gain the crucial insights into how genetic information is regulated and how we can integrate these findings in clinical applications.

## Key issues

- Post-translational histone modifications change chromatin structure directly through charge differences and/or indirectly by serving as a binding platform for 'effector' proteins that translate this coding information into biological readouts.

- Histone variants differ in their amino acid sequence from canonical histones and are found in specific chromatin regions, thereby suggesting that they might have different biological functions.

- DNA methylation predominantly occurs at promoter regions on CpG islands in mammals and is associated with gene silencing. It is inherited without any DNA sequence changes and is therefore considered a true epigenetic mark.

- ATP-dependent chromatin remodeling is performed by multisubunit protein complexes, which slide or evict nucleosomes on specific chromatin sites and modulate chromatin dynamics.

- Noncoding RNAs are involved in the regulation of gene expression either by repressing transcription through influencing chromatin structure, or by guiding RNA cleavage and translation repression.

- Nuclear architecture describes nuclear components as well as the localization of genomic loci in the highly structured nucleus. Correct positioning of chromatin domains to specific nuclear components can lock chromatin in a close and transcriptionally inactive conformation.

## References

Papers of special note have been highlighted as:
- • of interest
- •• of considerable interest

1   Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389(6648), 251–260 (1997).

•• **First study of the nucleosome core particle structure in high resolution, which is key to the understanding of chromatin structure.**

2   Strahl BD, Allis CD. The language of covalent histone modifications. *Nature* 403(6765), 41–45 (2000).

• **Hypothesis about how histone modifications are translated into biological function.**

3   Zhang Y, Reinberg D. Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes Dev.* 15(18), 2343–2360 (2001).

4   Cosgrove MS. Histone proteomics and the epigenetic regulation of nucleosome mobility. *Expert Rev. Proteomics* 4(4), 465–478 (2007).

5   Rea S, Eisenhaber F, O'Carroll D *et al.* Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406(6796), 593–599 (2000).

6   Wang Y, Wysocka J, Sayegh J *et al.* Human PAD4 regulates histone arginine methylation levels via demethylimination. *Science* 306(5694), 279–283 (2004).

7   Cuthbert GL, Daujat S, Snowden AW *et al.* Histone deimination antagonizes arginine methylation. *Cell* 118(5), 545–553 (2004).

8   Chang B, Chen C, Zhao Y, Bruick RK. JMJD6 is a histone arginine demethylase. *Science* 318(5849), 444–447 (2007).

9   Shi Y, Lan F, Matson C *et al.* Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* 119(7), 941–953 (2004).

•• **First report of a histone lysine demethylase. These analyses changed the view that lysine methylation is stable and irreversible.**

10  Klose RJ, Yamane K, Bae Y *et al.* The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. *Nature* 442(7100), 312–316 (2006).

11  Klose RJ, Kallin EM, Zhang Y. JmjC-domain-containing proteins and histone demethylation. *Nat. Rev. Genet.* 7(9), 715–727 (2006).

12  Cosgrove MS, Boeke JD, Wolberger C. Regulated nucleosome mobility and the histone code. *Nat. Struct. Mol. Biol.* 11(11), 1037–1043 (2004).

• **Excellent review on the effects of histone modifications on the mobility of nucleosomes.**

13  Turner BM. Decoding the nucleosome. *Cell* 75(1), 5–8 (1993).

14  Turner BM. Histone acetylation and an epigenetic code. *Bioessays* 22(9), 836–845 (2000).

15  de la Cruz X, Lois S, Sanchez-Molina S, Martinez-Balbas MA. Do protein motifs read the histone code? *Bioessays* 27(2), 164–175 (2005).

16  Mellor J. It takes a PHD to read the histone code. *Cell* 126(1), 22–24 (2006).

17  Fuchs J, Demidov D, Houben A, Schubert I. Chromosomal histone modification patterns – from conservation to diversity. *Trends Plant Sci.* 11(4), 199–208 (2006).

18  Briggs SD, Xiao T, Sun ZW *et al.* Gene silencing: trans-histone regulatory pathway in chromatin. *Nature* 418(6897), 498 (2002).

19  Schuettengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G. Genome regulation by polycomb and trithorax proteins. *Cell* 128(4), 735–745 (2007).

20  Dou Y, Milne TA, Tackett AJ *et al.* Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. *Cell* 121(6), 873–885 (2005).

21  Bernstein BE, Mikkelsen TS, Xie X *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125(2), 315–326 (2006).

•• **First identification of specific modification pattern, bivalent domains (H3K4me3 and H3K27me3) within regions of key developmental genes in mouse embryonic stem cells. The authors suggest that these bivalent domains play a role in the maintenance of pluripotency.**

22  Ruthenburg AJ, Li H, Patel DJ, Allis CD. Multivalent engagement of chromatin modifications by linked binding modules. *Nature Rev. Mol. Cell. Biol.* 8(12), 983–994 (2007).

23  Pusarla RH, Bhargava P. Histones in functional diversification. Core histone variants. *FEBS J.* 272(20), 5149–5168 (2005).

• **Excellent review of variants of all histone families.**

24  Polo SE, Almouzni G. Chromatin assembly: a basic recipe with various flavours. *Curr. Opin. Genet. Dev.* 16(2), 104–111 (2006).

25  Bustin M, Catez F, Lim JH. The dynamics of histone H1 function in chromatin. *Mol. Cell.* 17(5), 617–620 (2005).

26  Khochbin S. Histone H1 diversity: bridging regulatory signals to linker histone function. *Gene* 271(1), 1–12 (2001).

27  Sullivan KF. A solid foundation: functional specialization of centromeric chromatin. *Curr. Opin. Genet. Dev.* 11(2), 182–188 (2001).

28  Zeitlin SG, Shelby RD, Sullivan KF. CENP-A is phosphorylated by Aurora B kinase and plays an unexpected role in completion of cytokinesis. *J. Cell Biol.* 155(7), 1147–1157 (2001).

29  Hake SB, Garcia BA, Duncan EM *et al.* Expression patterns and post-translational modifications associated with mammalian histone H3 variants. *J. Biol. Chem.* 281(1), 559–568 (2006).

30  Bernstein E, Hake SB. The nucleosome: a little variation goes a long way. *Biochem. Cell Biol.* 84(4), 505–517 (2006).

31  Hake SB, Allis CD. Histone H3 variants and their potential role in indexing mammalian genomes: the "H3 barcode hypothesis". *Proc. Natl Acad. Sci. USA* 103(17), 6428–6435 (2006).

32  Hake SB, Garcia BA, Kauer M *et al.* Serine 31 phosphorylation of histone variant H3.3 is specific to regions bordering centromeres in metaphase chromosomes. *Proc. Natl Acad. Sci. USA* 102(18), 6344–6349 (2005).

33  Tsukuda T, Fleming AB, Nickoloff JA, Osley MA. Chromatin remodelling at a DNA double-strand break site in *Saccharomyces cerevisiae*. *Nature* 438(7066), 379–383 (2005).

34  Jin J, Cai Y, Li B *et al.* In and out: histone variant exchange in chromatin. *Trends Biochem. Sci.* 30(12), 680–687 (2005).

35  Chadwick BP, Willard HF. A novel chromatin protein, distantly related to histone H2A, is largely excluded from the inactive X chromosome. *J. Cell Biol.* 152(2), 375–384 (2001).

36 Bao Y, Konesky K, Park YJ *et al.* Nucleosomes containing the histone variant H2A.Bbd organize only 118 base pairs of DNA. *EMBO J.* 23(16), 3314–3324 (2004).

37 Doyen CM, Montel F, Gautier T *et al.* Dissection of the unusual structural and functional properties of the variant H2A.Bbd nucleosome. *EMBO J.* 25(18), 4234–4244 (2006).

38 Gautier T, Abbott DW, Molla A *et al.* Histone variant H2ABbd confers lower stability to the nucleosome. *EMBO Rep.* 5(7), 715–720 (2004).

39 Angelov D, Molla A, Perche PY *et al.* The histone variant macroH2A interferes with transcription factor binding and SWI/SNF nucleosome remodeling. *Mol. Cell.* 11(4), 1033–1041 (2003).

40 Karras GI, Kustatscher G, Buhecha HR *et al.* The macro domain is an ADP–ribose binding module. *EMBO J.* 24(11), 1911–1920 (2005).

41 Hernandez-Munoz I, Lund AH, van der Stoop P *et al.* 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(21), 7635–7640 (2005).

42 Doyen CM, An W, Angelov D *et al.* Mechanism of polymerase II transcription repression by the histone variant macroH2A. *Mol. Cell. Biol.* 26(3), 1156–1164 (2006).

43 Ma Y, Jacobs SB, Jackson-Grusby L *et al.* DNA CpG hypomethylation induces heterochromatin reorganization involving the histone variant macroH2A. *J. Cell Sci.* 118(Pt 8), 1607–1616 (2005).

44 Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev.* 16(1), 6–21 (2002).

45 Bernstein BE, Meissner A, Lander ES. The mammalian epigenome. *Cell* 128(4), 669–681 (2007).

46 Fuks F. DNA methylation and histone modifications: teaming up to silence genes. *Curr. Opin. Genet. Dev.* 15(5), 490–495 (2005).

47 Boyes J, Bird A. DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell* 64(6), 1123–1134 (1991).

• **Study showing that MeCP1, a methyl-CpG binding protein, represses transcription *in vivo*.**

48 Hendrich B, Bird A. Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol. Cell. Biol.* 18(11), 6538–6547 (1998).

49 Reik W, Dean W, Walter J. Epigenetic reprogramming in mammalian development. *Science* 293(5532), 1089–1093 (2001).

50 Santos F, Hendrich B, Reik W, Dean W. Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev. Biol.* 241(1), 172–182 (2002).

51 Bienvenu T, Chelly J. Molecular genetics of Rett syndrome: when DNA methylation goes unrecognized. *Nat. Rev. Genet.* 7(6), 415–426 (2006).

52 Yasunaga J, Taniguchi Y, Nosaka K *et al.* Identification of aberrantly methylated genes in association with adult T-cell leukemia. *Cancer Res.* 64(17), 6002–6009 (2004).

53 Baylin SB, Esteller M, Rountree MR *et al.* Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer. *Hum. Mol. Genet.* 10(7), 687–692 (2001).

54 Feinberg AP. Phenotypic plasticity and the epigenetics of human disease. *Nature* 447(7143), 433–440 (2007).

55 Goll MG, Bestor TH. Eukaryotic cytosine methyltransferases. *Annu. Rev. Biochem.* 74, 481–514 (2005).

56 Mertineit C, Yoder JA, Taketo T *et al.* Sex-specific exons control DNA methyltransferase in mammalian germ cells. *Development* 125(5), 889–897 (1998).

57 Shovlin TC, Bourc'his D, La Salle S *et al.* Sex-specific promoters regulate Dnmt3L expression in mouse germ cells. *Hum. Reprod.* 22(2), 457–467 (2007).

58 Chan SW, Henderson IR, Jacobsen SE. Gardening the genome: DNA methylation in *Arabidopsis thaliana*. *Nat. Rev. Genet.* 6(5), 351–360 (2005).

59 Blelloch R, Wang Z, Meissner A *et al.* Reprogramming efficiency following somatic cell nuclear transfer is influenced by the differentiation and methylation state of the donor nucleus. *Stem Cells* 24(9), 2007–2013 (2006).

60 Bruniquel D, Schwartz RH. Selective, stable demethylation of the interleukin-2 gene enhances transcription by an active process. *Nat. Immunol.* 4(3), 235–240 (2003).

61 Nishino K, Hattori N, Tanaka S, Shiota K. DNA methylation-mediated control of Sry gene expression in mouse gonadal development. *J. Biol. Chem.* 279(21), 22306–22313 (2004).

62 Feil R, Berger F. Convergent evolution of genomic imprinting in plants and mammals. *Trends Genet.* 23(4), 192–199 (2007).

63 Dobosy JR, Selker EU. Emerging connections between DNA methylation and histone acetylation. *Cell. Mol. Life Sci.* 58(5–6), 721–727 (2001).

64 Tamaru H, Zhang X, McMillen D *et al.* Trimethylated lysine 9 of histone H3 is a mark for DNA methylation in *Neurospora crassa*. *Nat. Genet.* 34(1), 75–79 (2003).

65 Malagnac F, Bartee L, Bender J. An *Arabidopsis* SET domain protein required for maintenance but not establishment of DNA methylation. *EMBO J.* 21(24), 6842–6852 (2002).

66 Jackson JP, Lindroth AM, Cao X, Jacobsen SE. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* 416(6880), 556–560 (2002).

67 Lehnertz B, Ueda Y, Derijck AA *et al.* Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr. Biol.* 13(14), 1192–1200 (2003).

68 Ooi SK, Qiu C, Bernstein E *et al.* DNMT3L connects unmethylated lysine 4 of histone H3 to *de novo* methylation of DNA. *Nature* 448(7154), 714–717 (2007).

• **First report that histone modifications negatively influence DNA methyltransferase binding to chromatin.**

69 Becker PB, Horz W. ATP-dependent nucleosome remodeling. *Annu. Rev. Biochem.* 71, 247–273 (2002).

• **Excellent review on ATP-dependent remodeling machines.**

70 Langst G, Becker PB. Nucleosome remodeling: one mechanism, many phenomena? *Biochim. Biophys. Acta* 1677(1–3), 58–63 (2004).

71 Gangaraju VK, Bartholomew B. Mechanisms of ATP dependent chromatin remodeling. *Mutat. Res.* 618(1–2), 3–17 (2007).

72 Kouzarides T. Chromatin modifications and their function. *Cell* 128(4), 693–705 (2007).

73 Pray-Grant MG, Daniel JA, Schieltz D, Yates JR III, Grant PA. Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. *Nature* 433(7024), 434–438 (2005).

74   Wysocka J, Swigut T, Xiao H *et al.* A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* 442(7098), 86–90 (2006).

75   Clapier CR, Nightingale KP, Becker PB. A critical epitope for substrate recognition by the nucleosome remodeling ATPase ISWI. *Nucleic Acids Res.* 30(3), 649–655 (2002).

76   Cosma MP, Tanaka T, Nasmyth K. Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle- and developmentally regulated promoter. *Cell* 97(3), 299–311 (1999).

77   Krebs JE, Kuo MH, Allis CD, Peterson CL. Cell cycle-regulated histone acetylation required for expression of the yeast HO gene. *Genes Dev.* 13(11), 1412–1421 (1999).

78   Ferreira R, Eberharter A, Bonaldi T *et al.* Site-specific acetylation of ISWI by GCN5. *BMC Mol. Biol.* 8(1), 73 (2007).

79   Krogan NJ, Keogh MC, Datta N *et al.* A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. *Mol. Cell.* 12(6), 1565–1576 (2003).

•• First report of a chromatin remodeling complex that specifically recognizes nucleosomes containing the histone variant Htz1.

80   Mizuguchi G, Shen X, Landry J *et al.* ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science* 303(5656), 343–348 (2004).

81   Kobor MS, Venkatasubrahmanyam S, Meneghini MD *et al.* A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin. *PLoS Biol.* 2(5), E131 (2004).

82   Konev AY, Tribus M, Park SY *et al.* CHD1 motor protein is required for deposition of histone variant H3.3 into chromatin *in vivo*. *Science* 317(5841), 1087–1090 (2007).

• First study demonstrating the function of a remodeler in *de novo* chromatin assembly *in vivo*.

83   Kusch T, Florens L, Macdonald WH *et al.* Acetylation by Tip60 is required for selective histone variant exchange at DNA lesions. *Science* 306(5704), 2084–2087 (2004).

84   Downs JA, Nussenzweig MC, Nussenzweig A. Chromatin dynamics and the preservation of genetic information. *Nature* 447(7147), 951–958 (2007).

85   Zhang Y, Ng HH, Erdjument-Bromage H *et al.* Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes Dev.* 13(15), 1924–1935 (1999).

86   Feng Q, Zhang Y. The MeCP1 complex represses transcription through preferential binding, remodeling, and deacetylating methylated nucleosomes. *Genes Dev.* 15(7), 827–832 (2001).

87   Fire A, Xu S, Montgomery MK *et al.* Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391(6669), 806–811 (1998).

•• First description of the silencing effect of small interfering RNA (RNAi), awarded with the 2006 Nobel Prize in Physiology or Medicine.

88   Chu CY, Rana TM. Small RNAs: regulators and guardians of the genome. *J. Cell. Physiol.* 213(2), 412–419 (2007).

89   Grewal SI, Elgin SC. Transcription and RNA interference in the formation of heterochromatin. *Nature* 447(7143), 399–406 (2007).

90   Verdel A, Jia S, Gerber S *et al.* RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* 303(5658), 672–676 (2004).

• First purification and description of the RNA-induced initiation of the transcriptional gene silencing complex linking siRNAs to heterochromatin assembly in fission yeast.

91   Bernstein E, Allis CD. RNA meets chromatin. *Genes Dev.* 19(14), 1635–1655 (2005).

• Excellent review on noncoding RNAs and their diverse influences on chromatin structure.

92   Irvine DV, Zaratiegui M, Tolia NH *et al.* Argonaute slicing is required for heterochromatic silencing and spreading. *Science* 313(5790), 1134–1137 (2006).

93   Matzke M, Kanno T, Huettel B, Daxinger L, Matzke AJ. Targets of RNA-directed DNA methylation. *Curr. Opin. Plant Biol.* 10(5), 512–519 (2007).

94   Janowski BA, Younger ST, Hardy DB *et al.* Activating gene expression in mammalian cells with promoter-targeted duplex RNAs. *Nat. Chem. Biol.* 3(3), 166–173 (2007).

95   Li LC, Okino ST, Zhao H *et al.* Small dsRNAs induce transcriptional activation in human cells. *Proc. Natl Acad. Sci. USA* 103(46), 17337–17342 (2006).

96   Lemons D, McGinnis W. Genomic evolution of *Hox* gene clusters. *Science* 313(5795), 1918–1922 (2006).

97   Sessa L, Breiling A, Lavorgna G *et al.* Noncoding RNA synthesis and loss of Polycomb group repression accompanies the colinear activation of the human HOXA cluster. *RNA* 13(2), 223–239 (2007).

98   Rinn JL, Kertesz M, Wang JK *et al.* Functional demarcation of active and silent chromatin domains in human *HOX* loci by noncoding RNAs. *Cell* 129(7), 1311–1323 (2007).

• High-resolution mapping of the transcriptional landscape of the four human *HOX* loci led to the identification of noncoding RNAs responsible for gene silencing in *trans*.

99   Nemeth A, Strohner R, Grummt I, Langst G. The chromatin remodeling complex NoRC and TTF-I cooperate in the regulation of the mammalian rRNA genes *in vivo*. *Nucleic Acids Res.* 32(14), 4091–4099 (2004).

100   Zhou Y, Grummt I. The PHD finger/bromodomain of NoRC interacts with acetylated histone H4K16 and is sufficient for rDNA silencing. *Curr. Biol.* 15(15), 1434–1438 (2005).

101   Mayer C, Schmitz KM, Li J, Grummt I, Santoro R. Intergenic transcripts regulate the epigenetic state of rRNA genes. *Mol. Cell.* 22(3), 351–361 (2006).

102   Santoro R. The silence of the ribosomal RNA genes. *Cell. Mol. Life Sci.* 62(18), 2067–2079 (2005).

103   Deng X, Meller VH. Non-coding RNA in fly dosage compensation. *Trends Biochem. Sci.* 31(9), 526–532 (2006).

104   Wutz A. Xist function: bridging chromatin and stem cells. *Trends Genet.* 23(9), 457–464 (2007).

105   Heard E, Disteche CM. Dosage compensation in mammals: fine-tuning the expression of the X chromosome. *Genes Dev.* 20(14), 1848–1867 (2006).

106   Gerstein MB, Bruce C, Rozowsky JS *et al.* What is a gene, post-ENCODE? History and updated definition. *Genome Res.* 17(6), 669–681 (2007).

•• Excellent perspective on the ENCODE consortium project that characterized the functional elements in 1% of the human genome and reported that many noncoding regions are transcribed. These data changed the current definition of a gene.

107 Lanctot C, Cheutin T, Cremer M, Cavalli G, Cremer T. Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. *Nat. Rev. Genet.* 8(2), 104–115 (2007).

• **Excellent review on nuclear architecture.**

108 Foster HA, Bridger JM. The genome and the nucleus: a marriage made by evolution. Genome organisation and nuclear architecture. *Chromosoma* 114(4), 212–229 (2005).

109 Shaklai S, Amariglio N, Rechavi G, Simon AJ. Gene silencing at the nuclear periphery. *FEBS J.* 274(6), 1383–1392 (2007).

110 Handwerger KE, Gall JG. Subnuclear organelles: new insights into form and function. *Trends Cell Biol.* 16(1), 19–26 (2006).

111 Cremer T, Cremer M, Dietzel S *et al.* Chromosome territories – a functional nuclear landscape. *Curr. Opin. Cell Biol.* 18(3), 307–316 (2006).

112 Osborne CS, Chakalova L, Brown KE *et al.* Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat. Genet.* 36(10), 1065–1071 (2004).

113 Faro-Trindade I, Cook PR. Transcription factories: structures conserved during differentiation and evolution. *Biochem. Soc. Trans.* 34(Pt 6), 1133–1137 (2006).

114 de Laat W. Long-range DNA contacts: romance in the nucleus? *Curr. Opin. Cell Biol.* 19(3), 317–320 (2007).

115 Lomvardas S, Barnea G, Pisapia DJ *et al.* Interchromosomal interactions and olfactory receptor choice. *Cell* 126(2), 403–413 (2006).

116 Gaszner M, Felsenfeld G. Insulators: exploiting transcriptional and epigenetic mechanisms. *Nat. Rev. Genet.* 7(9), 703–713 (2006).

117 Grimaud C, Negre N, Cavalli G. From genetics to epigenetics: the tale of Polycomb group and trithorax group genes. *Chromosome Res.* 14(4), 363–375 (2006).

118 Kassis JA. Pairing-sensitive silencing, polycomb group response elements, and transposon homing in *Drosophila*. *Adv. Genet.* 46, 421–438 (2002).

119 Bantignies F, Grimaud C, Lavrov S, Gabut M, Cavalli G. Inheritance of Polycomb-dependent chromosomal interactions in *Drosophila*. *Genes Dev.* 17(19), 2406–2420 (2003).

120 Grimaud C, Bantignies F, Pal-Bhadra M *et al.* RNAi components are required for nuclear clustering of Polycomb group response elements. *Cell* 124(5), 957–971 (2006).

• **Study demonstrating that the RNAi pathway is required for long-distance physical interactions between chromosomes mediated by the Polycomb repressive complex, regulating higher-order nuclear organization.**

121 Schirmer EC, Foisner R. Proteins that associate with lamins: many faces, many functions. *Exp. Cell Res.* 313(10), 2167–2179 (2007).

122 Wang GG, Allis CD, Chi P. Chromatin remodeling and cancer, part I: covalent histone modifications. *Trends Mol. Med.* 13(9), 363–372 (2007).

123 Wang GG, Allis CD, Chi P. Chromatin remodeling and cancer, part II: ATP-dependent chromatin remodeling. *Trends Mol. Med.* 13(9), 373–380 (2007).

124 Marks PA, Breslow R. Dimethyl sulfoxide to vorinostat: development of this histone deacetylase inhibitor as an anticancer drug. *Nat. Biotechnol.* 25(1), 84–90 (2007).

125 Dykxhoorn DM, Palliser D, Lieberman J. The silent treatment: siRNAs as small molecule drugs. *Gene Ther.* 13(6), 541–552 (2006).

126 Fraga MF, Agrelo R, Esteller M. Cross-talk between aging and cancer: the epigenetic language. *Ann. NY Acad. Sci.* 1100, 60–74 (2007).

## Affiliations

• Clemens Bönisch
  Adolf-Butenandt-Institute & Center for Integrated Protein Science Munich (CIPSM), Department of Molecular Biology, Ludwig-Maximilians University, Schillerstr. 44, 80336 Munich, Germany
  Tel.: +49 89 2180 75423
  Fax: +49 89 2180 75425
  clemens.boenisch@med.uni-muenchen.de

• Sonja M Nieratschker
  Adolf-Butenandt-Institute & Center for Integrated Protein Science Munich (CIPSM), Department of Molecular Biology, Ludwig-Maximilians University, Schillerstr. 44, 80336 Munich, Germany
  Tel.: +49 89 2180 75423
  Fax: +49 89 2180 75425
  sonja.nieratschker@med.uni-muenchen.de

• Nikos K Orfanos
  Adolf-Butenandt-Institute & Center for Integrated Protein Science Munich (CIPSM), Department of Molecular Biology, Ludwig-Maximilians University, Schillerstr. 44, 80336 Munich, Germany
  Tel.: +49 89 2180 75423
  Fax: +49 89 2180 75425
  nikolaos.orfanos@med.uni-muenchen.de

• Sandra B Hake
  Adolf-Butenandt-Institute & Center for Integrated Protein Science Munich (CIPSM), Department of Molecular Biology, Ludwig-Maximilians University, Schillerstr. 44, 80336 Munich, Germany
  Tel.: +49 89 2180 75439
  Fax: +49 89 2180 75425
  sandra.hake@med.uni-muenchen.de

## 1.2    THE NUCLEOSOME

The "monomeric building block" of chromatin, the nucleosome, contains about 150 bp of DNA wrapped around a histone octamer consisting of two of each of the core histones H2A, H2B, H3 and H4 in 1.65 left-handed superhelical turns (16). The existence of a chromatin subunit, the nucleosome, was first proposed in 1973/74 based on regular patterns upon nuclease digestion and electron microscopic analyses of chromatin ((13-15), for review see (19,20)). About 25 years later, the nucleosome structure at 2.8 Å resolution revealed its fascinating details (16). The (H3-H4)$_2$ tetramer is built by connecting two H3-H4 dimers at the dyad symmetry axis via a strong 4 helix bundle (4 HB) between the two H3 molecules. Interaction of H2A-H2B dimers with this tetramer is accomplished by a weaker 4 HB between H2B and H4 with additional interactions provided by the C-terminal H2A docking domain with H3 and H4, directing the H3 N-terminal helix to interact with DNA (Figure 1). Furthermore, contacts between the H2A loop 1 (L1) regions of the two H2A-H2B-dimers stabilize their association within the nucleosome (Figure 1). However, the nucleosome is not a static entity but rather flexible and dynamic (see (19-21) and references therein). As reviewed in van Holde et al., evidence for nucleosomes organizing DNA lengths between 100 and 170 bp is abundant, stressing that the crystal structure with 147 bp must rather be viewed as a "snapshot". In addition, recent single molecule analyses using Förster resonance energy transfer (FRET) contributed to the characterization of nucleosome dynamics providing evidence for an alternative, more open nucleosome state (0.2% to 3% under physiological salt conditions *in vitro*) where all histones are bound to DNA but the dimer/tetramer interactions are broken (19,22).

Histone variants are non-allelic isoforms of canonical histones that differ in both, their primary sequence as well as their expression timing (23). Expression of canonical histones is almost completely limited to S-phase, whereas most histone variants are expressed throughout the cell cycle. The S-phase dependent expression of canonical histones is mainly due to their unique mRNA structure (24). In general, canonical histone genes lack introns and their corresponding mRNAs are not polyadenylated but have a unique 3´ stem loop crucial to modulate mRNA stability, transport and translation. In contrast, histone variant mRNAs are polyadenylated and their pre-mRNAs can contain introns (24). To date, only two histone transcripts have been shown to be alternatively spliced, macroH2A1 (25) and H2A.Z.2 (26,27), giving rise to histones with distinct functional and structural properties, respectively.

In the following section I summarize literature on H2A variants biology. Here I do neither discuss their influences on nucleosome structure and stability nor on higher-order chromatin structure, since these are the topics of the discussion section.
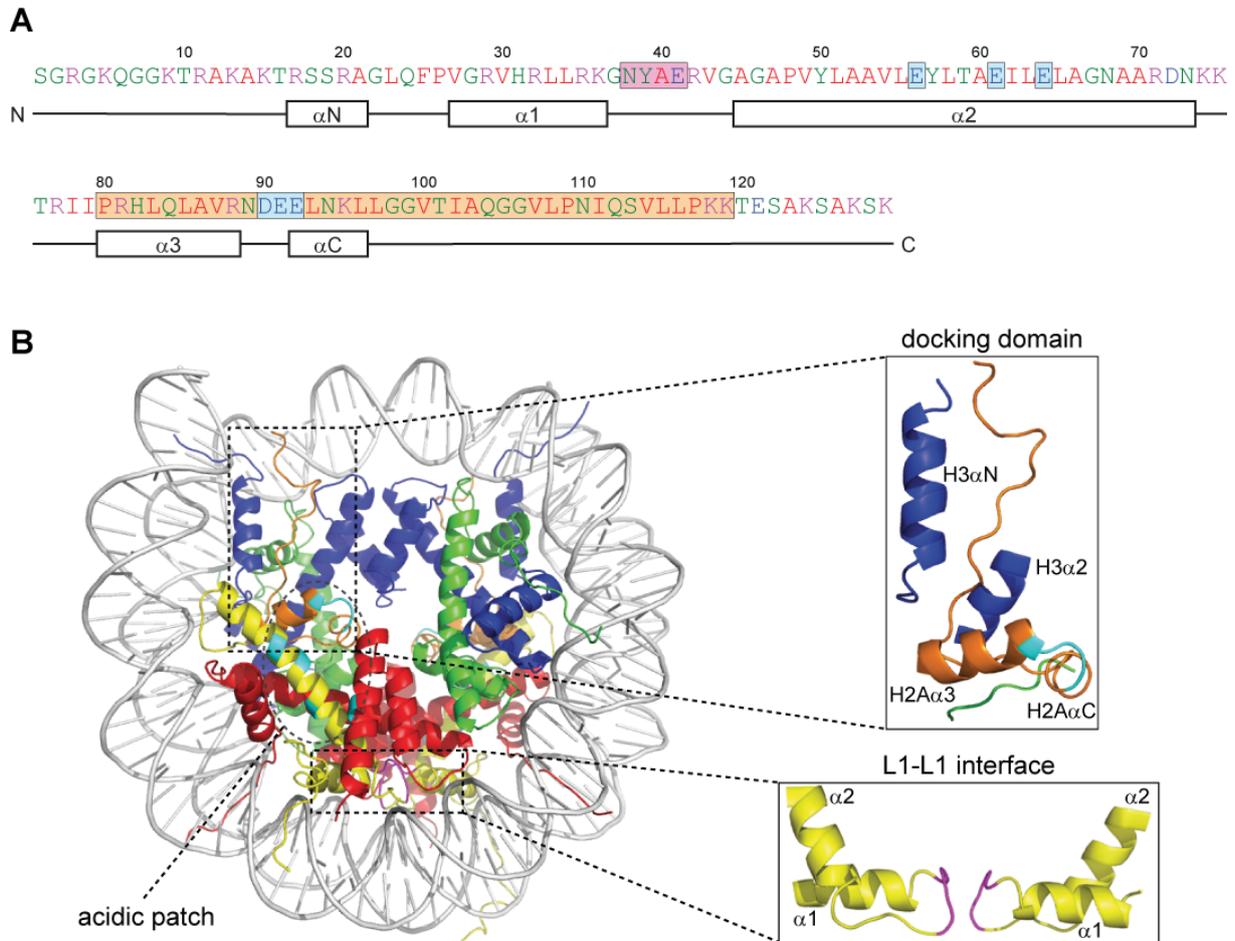
**A**

```
         10        20        30        40        50        60        70
SGRGKQGGKTRAKAKTRSSRAGLQFPVGRVHRLLRKGNYAERVGAGAPVYLAAVLEYLTAEILELAGNAARDNKK
```

N ——————— [αN] ———— [α1] ————————— [α2] ———————

```
         80        90       100       110       120
TRIIPRHLQLAVRNDEELNKLLGGVTIAQGGVLPNIQSVLLPKKTESAKSAKSK
```

——— [α3] —— [αC] ———————————————— C

**B**



**Figure 1: The crystal structure of the nucleosome.** (**A**) Amino acid sequence of histone H2A type 1 from *Xenopus leavis* (NCBI Reference Sequence: NP_001089684.1). α-helices are indicated by boxes below and important structural features are highlighted with colored boxes (L1 loop: magenta, acidic patch: cyan and docking domain: orange). The color code for the amino acids is red: small, hydrophobic (A, V, F, P, M, I, L, W); blue: acidic (D, E); magenta: basic (R, K); green: hydroxyl, sulfhydryl, amine, glycine (S, T, Y, H, C, N, G, Q) (**B**) Nucleosome crystal structure based on ((16), PDB ID: 1AOI). H2A is shown in yellow, H2B in red, H3 in blue, H4 in green and DNA in light gray. L1 loop, acidic patch and docking domain are highlighted and shown in magenta, cyan and orange, respectively. Zoomed images of docking domain and L1-L1 interface are depicted on the right. All pictures were generated using PyMOL (28).

## 1.3 THE HISTONE H2A FAMILY

Due to the specific nucleosomal protein-protein and protein-DNA interactions of each of the core histones, they are subject to different degrees of structural constraint probably resulting in different potentials to evolve variants (7,29). For example, H4 is one of the most slowly evolving eukaryotic proteins (29) with variants only described in tetrahymena (30), trypanosomes (31) and the urochordate, *Oikopleura dioica* (32). The H2A family, on the other hand, contains a plethora of variants with some "universal variants" found in almost all organisms, namely H2A.Z and H2A.X (7). These different degrees of variation might be attributed to the extensive intranucleosomal interactions, in the case of H4, and the location of H2A on the "edges" of the nucleosome. In general, the highest degree of diversification among histone H2A variants is to be found in their C-termini, regarding both length and amino acid sequence ((33) and Figure 2).

**Figure 2: Amino acid sequences of human H2A variants.** Alignments of human H2A type 1 (NCBI Reference Sequence: NP_003501.1) with (**A**) H2A.X (NP_002096.1), (**B**) H2A.Z.1 (NP_002097.1), (**C**) H2A.Bbd (NP_001017990.1) and (**D**) macroH2A1.2 (NP_004884.1). Important structural features are highlighted with colored boxes. For details on color coding see legend of Figure 1. The consensus symbols below the alignment are: an asterisk (*) to indicate fully conserved residues, a colon (:) to indicate conservation between groups of strongly similar properties and a period (.) to indicate conservation between groups of weakly similar properties. (**D**) MacroH2A1.2´s linker region (aa 122-160) and macro domain (aa 161-370) are highlighted with dark gray and light gray boxes, respectively. All alignments were carried out using the ClustalW alignment tool on the EBI homepage (34,35).

H2A´s position within the nucleosome might explain this finding. Its C-terminus is located at the DNA entry/exit site (Figure 1), making variations at this domain a powerful tool to functionally diversify nucleosomes by altering nucleosome stability and dynamics, binding to DNA and/or the linker histone H1 or other interacting factors. Furthermore, the L1 region in the histone fold, where interaction between the two H2A variants takes place, shows a high degree of variation among H2A variants (Figure 2).

In addition to the *bona fide* H2A variants (discussed below), canonical H2A proteins are not completely identical but rather show some sequence variability. In 1977, based on electrophoretic separation and subsequent analysis of amino acid composition, two different H2A isoforms (H2A.1 and H2A.2) that differ in amino acid position 51 (leucine or methionine respectively) were identified in mammals (36). After the human and mouse genomes were sequenced, it became apparent that canonical H2A proteins can differ in many more positions, especially in the C-terminal six amino acids (23). However, thus far no functional specialization of these canonical H2A isoforms has been demonstrated.

Interestingly, in addition to canonical H2A, H2A.X- and H2A.Z-like proteins, plants exhibit a special class of H2A isoforms which have an extended C-terminus comprising SPKK motifs (37,38). This kind of motif (more general T/SPXK) is also present in many subtypes of the linker histone H1 as well as in sea urchin sperm specific H2B, and it is a known target site for phosphorylation (39). This class of H2A proteins has been shown to protect about 16 bp more linker DNA from Micrococcal Nuclease (MNase) digestion than chicken erythrocyte H2A (40). Their property to bind more DNA might help to compact the inactive genome during seed dormancy. Upon germination, the H2A C-terminus is rapidly phosphorylated probably to weaken DNA binding by neutralizing the positive charge of the SPKK motifs (41). A similar situation is found in the sea urchin egg-specific histone variant CS H2A. Here, C-terminal phosphorylation occurs upon fertilization possibly leading to chromatin decondensation, which in turn could facilitate chromatin assembly during replication (39). Recently, the histones of bdelloid rotifers, freshwater invertebrates that are highly resistant to ionizing radiation and desiccation (42), were analyzed (43). Interestingly, the H2A proteins in this organism are very different from all other species with no canonical H2A, H2A.X or H2A.Z present. Instead, bdelloid

H2As have a longer C-terminus that was speculated to play a role in adaption to their environment, especially dealing with DNA damage upon desiccation (7,43).

In the next part, I briefly introduce the major H2A variants known thus far. For more detailed information about histone H2A variants and histone variants in general, see (7,44-46). The influence of H2A variants on nucleosome stability as well as on chromatin folding will not be addressed in this part but in separate sections below.

### 1.3.1 H2A.X

Histone H2A.X was, together with H2A.Z, first described in 1980 (47). H2A.X is defined by its SQ[E/D]Φ motif (where Φ is a hydrophobic amino acid) in the C-terminus. Upon DNA damage, the serine becomes phosphorylated (γH2A.X) and renders H2A.X an important player in preserving genome integrity (see below). Apart from the C-terminus, human H2A and H2A.X differ by just four amino acids in primary sequence; two substitutions in the N-terminal tail, (Q6T and T16S), one in the L1 loop (N38H) and one in the docking domain (K99G) (Figure 2). However, since residue 38 is located in the region where the two H2A-H2B dimers interact with each other (Figure 1), it has been suggested that this substitution might influence the ratio of homotypic (containing two H2A.X) versus heterotypic nucleosomes (containing one H2A and one H2A.X) (48). To my knowledge, it is not known whether H2A.X forms homo- or heterotypic nucleosomes *in vivo*. For a recent review on H2A.X structure and function, see (48). In some organisms, the SQ[E/D]Φ motif is present in other H2A family members hence, no distinct H2A.X exists in these organisms. For example, in yeast and the flagellated protozoan *Giardia lamblia*, the SQ[E/D]Φ motif is present in canonical H2A, whereas in the fly, it is present in the H2A.Z protein called H2AvD.

During S phase, the human and mouse H2A.X transcripts are processed in an identical manner to the canonical histone mRNAs, resulting in a stem loop structure and no polyA tail (49,50). Outside S phase, a longer transcript is produced by using a downstream polyadenylation site. Therefore, *H2A.X* exhibits both characteristics of replication-dependent and replication-independent histone genes.

H2A.X has been shown to be involved in the DNA damage response (DDR). Upon DNA double strand breaks (DSBs), the serine in the SQ[E/D]Φ motif (position 139 in human) is phosphorylated resulting in "γH2A.X" foci which can extend for up to 30 Mb surrounding the damaged site (51). H2A.X phosphorylation is a very early event in DDR leading to structural alterations, at the damaged site, to foster DNA repair. Although there are studies pointing towards a direct destabilization of the nucleosome by γH2A.X (52,53), two studies in yeast, employing serine to glutamate mutants to mimic a phosporylated serine, come to different results. While one study (54) found increased nuclease accessibility, suggesting a more open chromatin structure, another one (55) did not find any evidence for a direct structural influence on chromatin. In contrast to these controversial reports, the importance of ATP-dependent chromatin remodeling during DDR is undisputed. Several studies established the

crucial role of ATP-dependent chromatin remodeling complexes to increase DNA accessibility at the DSB site (reviewed in (56)). Since chromatin decondensation is not severely impaired in H2A.X knock out cells (57,58), it has been suggested that the critical role of γH2A.X is not the primary recruitment of remodeling factors but in retaining them at the repair site to define a "damage neighborhood" and to keep the two DNA strands together for efficient repair (reviewed in (48)).

Recently, two additional phosphorylation sites, in the vicinity of the extensively studied serine 139, have been reported. Firstly, the very C-terminal tyrosine 142 in the SQEY motif can be phosphorylated in vertebrates (59-62). The modification status of this residue, which is absent in yeast (L instead of Y), has been suggested to play a critical role in cell fate decision after DNA damage. If tyrosine 142 is phosphorylated, interaction of H2A.X with the pro-apoptotic JNK1 is increased at the expense of DDR factor recruitment. Hence, its dephosphorylated form facilitates DNA repair whereas its phosphorylated one promotes apoptosis. Secondly, phosphorylation of threonine 136 has been reported (53,63). Although the biological function of this modification is not yet known, it was speculated that together with serine 139 phosphorylation it might alter chromatin structure upon DNA damage (53).

Interestingly, a recent analysis of H2A variant dynamics in pre-implantation embryos suggested a novel role for H2A.X in chromatin remodeling during mouse development (64). The authors found a striking increase in H2A.X chromatin incorporation at the expense of canonical H2A, H2A.Z and macroH2A after fertilization, leading to chromatin containing mostly H2A.X and H2A during the one to four cell stages. Notably, this effect seems to depend primarily on H2A.X's C-terminus but not on serine 139, suggesting an intriguing and not well-understood effect of H2A.X on chromatin structure outside the DDR.

### 1.3.2 H2A.Z

Histone H2A.Z is an almost universal variant which evolved early and only once in evolution (65). H2A.Z is only ~60% identical to canonical H2A within the same species (66), but strikingly more conserved between different species (~80% identity between most organisms) with the most divergent member in trypanosomes (~50-60%) (67). This suggests that it fulfills specific and unique functions that cannot be carried out by other H2A variants. Indeed, H2A.Z has been shown to be essential in many organisms like mouse (68), fly (69), frogs (70) and tetrahymena (71); but not in bakers (72) and fission yeast (73), where knock-out leads to severe growth phenotypes. A very elegant study in drosophila demonstrated that the essential regions for H2A.Z function are located in its C-terminus (M6 and M7, Figure 3) (74). In line with this finding, the M6 region is required for interaction of H2A.Z with the evolutionary conserved SWR1 chromatin remodeling complex important for H2A.Z targeting (discussed below), providing a reasonable explanation for its essential nature (75). Furthermore, this region comprises residues of the acidic patch, important for H2A.Z deposition and function in yeast (76) and chromatin higher-order structure (see respective section below).

**Figure 3: Amino acid sequences of human H2A.Z variants.** Alignment of human H2A.Z.1 with H2A.Z.2.1 and H2A.Z.2.2. α-helices are indicated by open boxes below and structural features that are discussed in the text are highlighted with colored boxes. For details on color coding and consensus symbols see legends of Figures 1 and 2, respectively. Sequence elements required for H2A.Z function (74) are indicated by gray boxes below and sites of PTMs as described in the legend. Alignment was carried out using the ClustalW alignment tool on the homepage of the EMBL-EBI (34,35).

Surprisingly, despite significant sequence divergence, the H2A.Z nucleosome structure (77) revealed, overall, high similarity to the canonical one (16). Striking differences between both structures are found in L1, important for interaction of the two H2A-H2B dimers within the nucleosome. This led to the hypothesis that it is unlikely that H2A and H2A.Z are present within the same nucleosome, proposing the exclusive existence of homotypic nucleosomes for H2A.Z (77). This prediction however, was proven wrong *in vitro* (78) as well as *in vivo,* using tagged H2A variants in HeLa cells, where the main proportion of H2A.Z containing nucleosomes is heterotypic (79). Further differences between the two structures were found in the C-terminal docking domain suggesting a subtle destabilization of the H2A.Z nucleosome and, due to the presence of a metal ion at the nucleosomal surface, a possibly altered interaction site for the linker histone or other factors. Additionally, an increased acidic patch on the nucleosome surface is observed for H2A.Z (Figure 2), suggesting an influence on internucleosomal interactions ((77) and see below).

Like other histones, H2A.Z can be post-translationally modified by acetylation, sumoylation and ubiquitination with different functional outcomes (reviewed in (80)). H2A.Z sumoylation has been implicated in DNA repair in yeast (81), ubiquitination correlates with localization to the inactive X chromosome (Xi) in mammals (82), whereas N-terminal acetylation leads to nucleosome destabilization (83). It was suggested that H2A.Z acetylation works as a switch-like mechanism to modulate H2A.Z nucleosome stability, ascribing repressive functions to the unmodified and activating functions to the acetylated form (83). Furthermore, acetylated H2A.Z was found associated with active genes but its role at these sites is not yet completely understood (reviewed in (80)).

The biological function of H2A.Z has been extensively studied revealing roles in transcription regulation, DNA repair, heterochromatin formation, chromosome segregation and mitosis. Due to

space constraints, I cannot discuss all aspects of H2A.Z biology. Excellent reviews covering the vast amount of literature are available (66,84-89).

Many studies focused on the influence of H2A.Z on transcription (reviewed in (84-86)) revealing that in yeast (90-93) and mammals (94) H2A.Z is enriched at gene promoters. Interestingly, it has been found that H2A.Z can have both, activating as well as repressive influences on transcription (84). Evidence accumulated that H2A.Z affects nucleosome mobility and positioning (84,91,95-99) which could explain the sometimes rather contrasting impact on transcription. As a consequence of such changes, incorporation of H2A.Z could differentially increase or decrease binding of both activating and repressive regulatory factors to their target sequences. Hence, the naïve view of H2A.Z as a transcriptional activator (or repressor), acting merely by structural alterations should be extended by one interpreting H2A.Z as a modulator of nucleosome positioning which consequently influences different biological processes including gene activity by transcription regulation. In addition to gene promoters, H2A.Z is associated with other regulatory regions like enhancers and insulators as well as heterochromatin (reviewed in (66)), consistent with the wide variety of biological processes this variant is implicated in.

The intriguing finding that H2A.Z is non-uniformly localized within the genome leads to the question by which means H2A.Z is enriched at its target sites. Two non-mutually exclusive mechanisms can be envisioned. Firstly, H2A.Z can be actively incorporated at specific sites by targeting factors and secondly, H2A.Z can be randomly incorporated and afterwards (actively) removed from non-target sites (100). Evidence is found for both mechanisms (for recent reviews on H2A.Z deposition see (87,89,100)). The ATP-dependent chromatin remodeling complex SWR-1 is important for H2A.Z deposition, in yeast, by exchanging nucleosomal H2A-H2B dimers (H2A-H2B) for free H2A.Z-H2B dimers (H2A.Z-H2B) (101-103). Here, target sites for H2A.Z incorporation can be defined by at least two different manners. On the one hand, it was suggested that the SWR-1 complex can be recruited by acetylated histones (90,92); on the other hand, the insertion of a certain DNA sequence that harbors elements, common in yeast promoters, into an inactive gene, was shown to be sufficient to induce a typical feature of yeast promoters: a nucleosome-free region (NFR) flanked by two H2A.Z containing nucleosomes (92). Therefore, both genetic as well as epigenetic factors contribute to establish the specific H2A.Z pattern in yeast. In mammals, two SWR-1 related complexes, the p400/NuA4/TIP60 and the SRCAP complex, exist ((89) and references therein). Both can catalyze the exchange of nucleosomal H2A-H2B for free H2A.Z-H2B, but their different compositions implicate functional specialization. More recently, the importance of the ATP-dependent chromatin remodeling complex INO80 (Inositol-requiring protein 80) for H2A.Z localization patterns, in yeast, was established (104). Papamichos-Chronakis et al. could show that INO80 catalyzes the opposite reaction as SWR-1, namely the active exchange of nucleosomal H2A.Z-H2B for free H2A-H2B. Addressing the *in vivo* relevance of this reaction, they found that loss of INO80 leads to mislocalization of unacetylated

H2A.Z, which results in genome instability. Speculating about the underlying mechanism, they hypothesized that impairment of removal of unacetylated H2A.Z might interfere with processes preventing genome stability by altering chromatin structure.

In vertebrates, two H2A.Z genes are present, *H2A.Z.1* (*H2AFZ*) and *H2A.Z.2* (*H2AFV*), which are expressed in a wide variety of tissues (26,105). Both genes contain introns, give rise to polyadenylated mRNAs and to protein products that differ in only three amino acids (106). Both H2A.Z proteins can be acetylated at the same N-terminal lysine residues (105) and show very similar nuclear localization patterns (26,105) and fluorescence recovery after photobleaching (FRAP) mobilities (26). Their promoter structures, however, are different between both genes (105) and knock out of *H2A.Z.2* but not of *H2A.Z.1* leads to BCL6 down regulation and increased apoptosis in chicken DT40 cells, suggesting functional (sub)specialization of the two H2A.Z variants (107).

Recently, we (26) and others (27) showed that the human H2A.Z.2 transcript can be alternatively spliced giving rise to two isoforms, the already known H2A.Z.2.1 (Z.2.1, formerly H2A.Z.2) and the novel H2A.Z.2.2 (Z.2.2). In contrast to the highly conserved major isoform Z.2.1, Z.2.2 is putatively primate-specific and present at much lower levels in most tissues. In brain tissues however, Z.2.2 is significantly enriched with abundances similar to Z.2.1. The two alternatively spliced transcripts differ only in their last exons, resulting in differences only in the C-termini of the encoded proteins. Z.2.2 is the shorter protein, lacking the utmost C-terminal tail and having a unique C-terminus/docking domain but retaining the extended H2A.Z acidic patch completely (Figure 3). Furthermore, we could show that Z.2.2 nucleosomes exhibit striking differences with regards to nucleosome stability both *in vivo* and *in vitro* (discussed in the respective section below). Notably, it has been suggested that alternative splicing can be a significant evolutionary driving force since alternative splicing events are very often, as in the case of H2A.Z.2, associated with exon gain or loss when compared between human, mouse and rat (108,109). Modrek & Lee proposed that alternatively spliced isoforms can serve as "internal paralogs" (108). Initially underrepresented due to weak splice signals, they can be tolerated since they do not interfere with gene function and are not detrimental for the cell. Over time however, they are able to accumulate mutations and become functionally important in a tissue-specific manner where they can represent 30-70% of all transcript isoforms from the respective locus (108). Indeed, more than 90% of human multi-exon genes are alternatively spliced with splicing patterns varying between different tissues (110). These findings are in perfect agreement with data from us (26) and others (27), showing that Z.2.2 mRNA is normally low abundant but constitutes up to 50% of the H2A.Z.2 isoforms in brain tissues. It is tempting to speculate that Z.2.2 is the major H2A.Z.2 isoform in some specialized cell types in the primate brain. There, it might be able to substitute for Z.2.1 and to confer unique structural and functional properties to H2A.Z containing nucleosomes, possibly acting in concert with chaperone complexes containing brain specific subunits (111,112).

### 1.3.3 H2A.Bbd

Histone H2A.Bbd (Barr body deficient) was first described just over one decade ago (113), but identification of the endogenous protein was published only recently (114). As found most often for replacement variants (23), H2A.Bbd is encoded by a polyadenylated mRNA. On the protein level, H2A.Bbd is only about 50% identical to canonical H2A (113) and is the most quickly evolving histone variant known, even exceeding the rate of evolution of the linker histone H1 (114,115). In agreement with its fast evolution, several H2A.Bbd-like proteins (also known as H2AL1-3 (116) or H2A.Lap2-4 (117)) are found in mouse, which are not all present in the human genome (116). Thus far, H2A.Bbd has only been found in mammals (115).

Comparison of histone H2A.Bbd and H2A protein sequences reveals several striking differences (Figure 2). H2A.Bbd is considerably shorter, lacking the C-terminal tail and part of the docking domain. Additionally, it does not contain an acidic patch implicated in internucleosomal contacts and chromatin fiber condensation (see below). Therefore, H2A.Bbd is also called H2A.Lap1 (Lack of acidic patch) in mouse (117). Interestingly, H2A.Bbd contains relatively few lysine residues indicating poor conservation of possible modifications, for example acetylation in the N-terminus. I decided to stick to the more widely used term H2A.Bbd in the following text.

H2A.Bbd is not present in all tissues but strongly expressed in testis (113,114,117) and to a much lesser extent in brain (117), suggesting a tissue-specific function. Indeed, H2A.Bbd plays a role in mouse spermatogenesis (114,117). Soboleva et al. could show that H2A.Bbd is involved in creating a specific chromatin landscape at the promoters of active genes, during spermatogenesis, in a temporally specific manner, where H2A.Z occupies the -2 nucleosome and H2A.Bbd the -1 nucleosome with respect to the transcription start site (TSS) of active genes. Since incorporation of H2A.Bbd hinders chromatin fiber folding to a similar extent like acetylated H3 and H4, the authors suggested that employing H2A.Bbd instead of histone acetylation could be advantageous in the process of rapid chromatin remodeling during spermatogenesis in mouse. H2A.Bbd's association with actively transcribed chromatin is further supported by co-localization of ectopically expressed H2A.Bbd with acetylated H4 (113). To address the mechanism by which H2A.Bbd influences transcription, experiments *in vitro* have been employed. Surprisingly, they found only mild effects on transcription of chromatin reconstituted with H2A.Bbd compared to canonical H2A chromatin (118,119). Angelov et al. found a maximal two-fold increased interaction of a transcription factor (NF-κB) with its binding site for H2A.Bbd versus canonical H2A nucleosomes *in vitro*. Furthermore, they reported a slightly more efficient transcriptional activity when using an H2A.Bbd chromatin template. This effect was dependent on the acetyltransferase p300 and coincident with elevated histone acetylation. Bao et al. also reported slightly more transcription from H2A.Bbd containing chromatin; however, p300 seemed to level out expression from both kinds of templates. Interestingly, and counter-intuitively, H2A.Bbd is much less efficiently remodeled by a variety of ATP-dependent chromatin remodeling complexes

like SWI/SNF (SWItch/Sucrose NonFermentable), ACF (ATP-utilizing Chromatin Assembly and Remodeling Factor) (118) and RSC (Remodels the Structure of Chromatin) (120). However, since H2A.Bbd is expressed more or less testis-specifically, it is possible that it is remodeled by, as yet, unidentified, tissue-specific, molecular machines.

In general, testis-specific variants of other histone families, such as H3t (121) and TSH2B (122), are also known,. Together they contribute to the unique chromatin structure in testis and are speculated to be involved in the process of histone to protamine replacement. However, due to tissue-specific expression of H2A.Bbd and other testis-specific histone variants, findings from studies focusing on them are limited in the generality of their implications on chromatin structure in other tissues. For most tissues and cells types, chromatin structure cannot be influenced by variants like H2A.Bbd simply due to their absence or low expression levels (Z.2.2). Hence, structural alterations must be accomplished by other, more general means. On the other hand, tissue and cell type-specific histone variants (and other chromatin factors) could contribute to specialized chromatin functions only required in certain cell types and tissues like testis (H2A.Bbd) and brain (Z.2.2).

## 1.3.4 MacroH2A

Histone macroH2A is the most diverged H2A variant known to date. It was first described two decades ago (123) and has since fascinated researchers because of its particular domain architecture. MacroH2A has a tripartite structure consisting of an N-terminal histone domain connected via a lysine rich H1-like linker region to a non-histone macro domain (Figure 2) resulting in a protein about three times larger than canonical H2A. MacroH2A is conserved among vertebrates, whereas macro domain-containing non-histone proteins are found in all organisms (124). The highly conserved macro domain is known to be a binding module for NAD metabolites and implicated in diverse biological functions like transcriptional regulation, chromatin remodeling and DNA repair (for a recent review on macro domain proteins see (125)). Two macroH2A genes are present in mammals (*macroH2A1/H2AFY* and *macroH2A2/H2AFY2*), with one of them (*macroH2A1*), known to be alternatively spliced (25), giving rise to two isoforms, macroH2A1.1 and macroH2A1.2. The two splice variants differ only in their macro domains which results in differences in their abilities to interact with NAD metabolites (126,127). MacroH2A1.1 can bind NAD metabolites, including poly(ADP-ribose), whereas macroH2A1.2 cannot. This suggests a unique role for macroH2A1.1 in chromatin remodeling that depends on poly-ADP-ribose polymerase (PARP) activity, which is induced by different biological stimuli such as DNA damage and metabolic stress.

The first insights into macroH2A´s biological function(s) came from immunofluorescence microscopy studies showing an enrichment on the Xi in female mammals (128). Mammalian dosage compensation is accomplished by transcriptional silencing of one of the two X chromosomes in females resulting in the same gene dose as in males (recently reviewed in (129)). The Xi is a *bona fide* model for an epigenetically regulated chromatin state since, once established, it is stable and passed on during

mitosis. These initial findings constituted the basis for the general view of macroH2A as an epigenetic repressor of gene transcription involved in X inactivation. Recently, nuclear transfer experiments in frogs showed that macroH2A inhibits reprogramming and hence contributes to stability and maintenance of differentiated epigenomes (130). Although macroH2A´s role in X inactivation is well established, two findings suggested early on a function(s) outside X inactivation. MacroH2A is present in vertebrates other than mammals that do not undergo X inactivation and it is expressed equally in both male and female mammals (25,124).

Many studies analyzed the influence of macroH2A on gene expression both on the X chromosome and on autosomes. The general view is that macroH2A represses transcription by setting up a repressive chromatin environment ((128,131-133) reviewed in (124,134)). However, some recent studies challenge this view by reporting a positive influence on some macroH2A target genes, thereby also influencing cell differentiation (135-137). How macroH2A mechanistically works on these target genes is not well understood.

In 2005, the structure of the macroH2A containing nucleosome was published showing overall similarity when compared to the canonical one (138). The structure of the macroH2A docking domain, although harboring several substitutions, is not altered and the residues constituting the acidic patch are completely conserved. However, the two structures differ substantially in a four amino acid region in the L1 loop, which constitutes the interaction site of the two H2A-H2B dimers within the nucleosome. Interestingly, macroH2A preferentially forms heterotypic nucleosomes over homotypic ones *in vitro* with an overall similar structure but changes in the L1-L1 interface due to the amino acid sequence differences in the two H2A variants (139). This is in contrast to H2A.Z and H2A.Bbd, which form stochastic mixtures of homo- and heterotypic nucleosomes *in vitro* (78). In addition to affecting nucleosome structure, the L1 region of macroH2A might play a role in targeting to the Xi, since, when inserted into canonical H2A, it is sufficient for Xi enrichment (140). Together with two other regions sufficient for Xi targeting (one in the $\alpha$1 helix and one in the docking domain), the L1 region is located on the outside of the macroH2A-H2B dimer, constituting a possible chaperone-binding site (140).

In contrast to H2A.Z, factors involved in macroH2A targeting are not well characterized; only our recent study provides first insights (141). We showed that macroH2A associates with ATRX ($\underline{\alpha}$-$\underline{t}$halassemia/M$\underline{R}$, $\underline{X}$-linked), although it is not known whether this interaction is direct or not. Importantly, macroH2A-ATRX binding is independent of DAXX (Death-Domain Associated Protein), which acts together with ATRX in H3.3 deposition at telomeres (142-144). These findings demonstrate that two distinct ATRX-containing complexes act together on H3.3 and macroH2A. Interestingly, in contrast to its role in active H3.3 deposition (142,143), ATRX is a negative regulator of macroH2A chromatin association by an, as yet, unknown mechanism. ATRX knock-down leads to increased macroH2A incorporation at telomeres and the $\alpha$-globin gene cluster, concomitant with its

reduced expression. Together with the study by Papamichos-Chronakis et al. (see above), this one contributes to the emerging, thus far underappreciated, regulation of histone variant localization by factors negatively influencing their chromatin association.

Several studies employed *in vitro* experiments to gain insight into the mechanisms by which macroH2A functions to repress transcription. Angelov et al. suggested that macroH2A acts on at least two distinct levels to repress transcription (145). Firstly, by interfering with transcription factor binding if the binding site is close to the nucleosome dyad axis, the part which shows strongest alteration in DNaseI digestion pattern and secondly, by inhibiting ATP-dependent nucleosome remodeling. The authors found that the influence on transcription factor binding is dependent on macroH2A´s non-histone region (NHR; linker and macro domain, amino acids 121-372), whereas the histone domain alone is sufficient to inhibit remodeling. Another study, however, reinvestigated nucleosome remodeling of macroH2A with different results (146). Here, ATP-dependent nucleosome remodeling of macroH2A nucleosomes by SWI/SNF and ACF was not found to be impaired. Using competition experiments, they could show that the activating SWI/SNF complex binds preferentially canonical over macroH2A nucleosomes whereas the ACF complex, mostly involved in gene repression, does not show any preference. Interestingly, all effects were dependent on the NHR in contrast to the study by Angelov et al.

In addition to the mechanisms discussed above, macroH2A also represses transcription more indirectly by reducing histone acetylation via different mechanisms dependent on the NHR. On the one hand, macroH2A inhibits p300-dependent histone acetylation *in vitro* (147); on the other hand it interacts with histone deacetylases resulting in co-precipitation with hypoacetylated chromatin (138).

Taken together, the H2A family has a multitude of different members that differ strikingly with regards to their evolutionary conservation, amino acid sequences and domain architectures, and the biological processes they play roles in. The mechanisms of their functions are often not well understood; open questions remain including how they are targeted to their respective chromatin sites and how specific interaction partners contribute to their biological roles. One very plausible mechanism of function is the alteration of nucleosome and higher-order chromatin structure brought about by H2A variant incorporation. In the next section, H2A variants and the properties they confer to chromatin on different levels, ranging from the nucleosome to higher-order chromatin structure, are discussed.

## 2. DISCUSSION

### 2.1 THE INFLUENCE OF H2A VARIANTS ON NUCLEOSOME STABILITY

#### 2.1.1 The H2A C-terminus influences nucleosome properties such as stability, dynamics, positioning and linker histone binding

Core histones have a common structural architecture as they consist of a histone fold domain (three α helices connected by short loops) and an unstructured N-terminal tail (16). In addition, and in contrast, to the other core histones, H2A also exhibits a flexible tail at the C-terminus. From the crystal structure, it can be seen that the C-terminal part of H2A (amino acids 80-119), including the α3 and αC helices, forms a ladle shaped docking domain that constitutes an important interface for interaction with the $(H3-H4)_2$-tetramer (16). The very C-terminal amino acids protrude from the globular nucleosome structure and interact with DNA, which is illustrated by molecular dynamics (MD) simulations, revealing stable hydrogen bonds between DNA and the lysines 118 and 119 in H2A (148). This is consistent with the recent finding that H2A monoubiquitination destabilizes nucleosomes during repair of UV induced DNA damage (149), possibly by neutralization of the negative charge of the ε-amino group. Interactions of the H2A C-terminus with nucleosomal DNA are modulated by the presence of linker DNA as well as the linker histone H1 (150,151). Moreover, H2A can directly interact with H1, as has been shown by crosslinking experiments (152,153) and, recently, the interaction site was mapped to the last 17 amino acids of H2A further stressing the importance of its C-terminal tail (154).

The question of whether linker histone binding to nucleosomes is affected by H2A variant incorporation has been addressed for all major H2A variants *in vitro*. In general, canonical H2A nucleosomes seem to bind the linker histone most efficiently, in accordance with the direct interaction between the two proteins (152-154). Incorporation of H2A.X into nucleosomes exhibits only mild effects on interaction with H1, but phosphorylation of the H2A.X C-terminus leads to significant impairment of this binding (53). More pronounced reductions of interaction with H1 were reported for H2A.Z (99) and H2A.Bbd nucleosomes (120). In the case of H2A.Bbd, this has been further dissected, showing that reduced interaction with H1 can be attributed to the H2A.Bbd docking domain. In contrast to the other variants, the influence of macroH2A on H1 binding has not been analyzed using *in vitro* assembled nucleosomes. However, fractionation experiments of native chicken chromatin revealed an almost mutually exclusive distribution; chromatin is either associated with linker histone or it contains macroH2A (155). This finding suggests that macroH2A interferes with linker histone binding, probably by its large C-terminal NHR, but it does not address the question whether macroH2A incorporation influences H1 binding as directly as the band shift assay carried out for other

H2A variants (see above). Notably, it has been shown that the H1-like linker domain of macroH2A fosters chromatin folding and compaction in the absence of the macro domain, leading to the speculation that it might fulfill linker histone function after removal of the macro domain by proteolytic cleavage (156). However, no evidence pointing towards relevance of this mechanism *in vivo* is available thus far.

The importance of the H2A C-terminus for protein-protein interactions within the histone octamer was established almost 25 years ago (157). Eickbush et al. found that removal of H2A´s 15 C-terminal amino acids resulted in a significant destabilization of the isolated histone octamer under high salt conditions. (The histone octamer is unstable under physiological salt conditions but can be stabilized by high salt concentrations (158).) Cleaving the peptide bond between valine 114 and leucine 115 destroys a short α helix (Q112-L116) that is present in both the isolated octamer (159) as well as the nucleosome (160) and contributes to complex stability by hydrophobic interactions between H2A and H3 (159). Recently, Vogler et al. analyzed C-terminal truncations of canonical H2A *in vitro* and *in vivo* (154). They reported moderately decreased nucleosome stability due to removal of the C-terminal 15 amino acids. More interestingly, they also found altered nucleosome positioning as well as less H1 binding and decreased susceptibility to ATP-dependent chromatin remodeling consistent with data from others (120). The biological significance of these findings is illustrated by reduced stress resistance of cells expressing H2A truncations at levels of about 10% of endogenous H2A, probably brought about by altered chromatin structure due to insufficient H1 recruitment and erroneous nucleosome positioning. Furthermore, C-terminal truncations of H2A enhance thermal nucleosome mobility, pointing towards the influence of the H2A C-terminus in defining specific and stable nucleosome positions (154,161). The possible role of H2A.Z in creating nucleosomes harboring special properties with regards to nucleosome positioning and mobility (see above) further highlights the importance of H2A variants in defining unique nucleosomal properties.

The evidence for an alternative nucleosome state in which all histones are bound to DNA but where the interface between the $(H3-H4)_2$-tetramer and the H2A-H2B dimer is opened, suggests an intriguing model for the influence of H2A variants on nucleosome stability and dynamics (19,22). H2A variant incorporation can lead to alterations of this particular interface thereby shifting the equilibrium between the closed and the open nucleosome state and consequently confers distinct dynamic properties to variant-containing nucleosomes. Since H2A variants differ significantly in their C-termini that are implicated in these interactions, this could be one mechanism by which they accomplish their distinct biological functions. Consistent with the idea that the H2A-H3 interface is sensitive to changes on both sides and critical for nucleosome properties, mutations of residues within the H3 αN helix (I51A or Q55A), involved in interactions with the H2A C-terminus, greatly increase nucleosome thermal mobility, H2A-H2B dimer exchange and abolish octamer formation under high

salt conditions *in vitro* (161). The influence of the different H2A variants on nucleosome stability is discussed below.

### 2.1.2    H2A.X

In general, the biochemical and biophysical properties of H2A.X have not been studied as extensively as for the other major H2A variants. This might be due to its high similarity to canonical H2A. However, a recent study analyzed the stability of H2A.X as well as γH2A.X containing nucleosomes by analytical ultracentrifugation (53). Surprisingly, they found striking nucleosome destabilization by H2A.X, further enhanced by C-terminal phosphorylation. Indeed, H2A.X harbors two substitutions in comparison to H2A (N38H and K99G) which were suggested to influence nucleosome stability as they are located in regions important for protein-protein interactions, within the nucleosome (48). Li et al. state that the observed destabilization of H2A.X nucleosomes is similar to observations on yeast nucleosomes which also exhibit decreased salt stability (162). However, this comparison is difficult to draw even though yeast H2A can be seen as an ortholog of H2A.X. Slight alterations in amino acid sequences are present in all yeast histones and distributed throughout the whole nucleosome structure (163), thereby making it hard to evaluate the influence of yeast H2A on nucleosome stability in an isolated manner. It is tempting to speculate that the extended C-terminal tail present in H2A.X might be involved in the changes discussed above, since nucleosome stability is further reduced upon C-terminal phosphorylation. Future studies will hopefully reveal which changes in H2A.X´s primary structure are relevant for the observed destabilization.

### 2.1.3    H2A.Z

The stability of the H2A.Z containing nucleosome has been intensively studied with contrasting results (reviewed in (66)). Some studies found stabilization (83,164,165), whereas others found destabilization of the nucleosome upon H2A.Z incorporation (77,90,166). Some FRET measurements detected only subtle effects on stability *in vitro* (26,167), consistent with two studies measuring H2A.Z mobility *in vivo* using FRAP (26,168). The reported differences can have a multitude of reasons, for example the use of H2A.Z from different organisms (note: 80% identity means 20% divergence), different experimental setups and different sources of chromatin (recombinant vs. native chromatin). Comparison of the available studies is further complicated by the fact that recombinant chromatin consists of homotypic, whereas native chromatin consists mostly of heterotypic nucleosomes (79) that can also be post-translationally modified (80). Moreover, comparing studies with *in vitro* assembled chromatin is complicated by the different DNA sequences used (167). Two examples nicely illustrate these problems; Zhang et al. (90) found destabilization by analyzing native chromatin fibers prepared from yeast, whereas Park et al. (165) found stabilization by performing FRET analyses of *in vitro* reconstituted nucleosomes from Xenopus histones, produced in E. coli, on 5S rDNA. What one could hypothesize from these studies is that H2A.Z is probably not the sole

determinant of nucleosome stability but might modulate it, integrating influences like DNA sequence, post-translational modifications and nucleosome composition.

In line with this, the Felsenfeld lab reported that H2A.Z severely destabilizes nucleosomes if present with H3.3 in the same particle (169). Nucleosomes prepared from native chromatin containing both H2A.Z and H3.3, are highly salt sensitive and are disrupted in the presence of minimal (80 mM) NaCl. In a second paper (170), they analyzed the genome wide distribution of H3.3/H2A.Z-containing, double variant, nucleosomes and found that they mark the NFRs of active promoters, enhancers and insulator regions. These nucleosomes are highly unstable and can therefore be more easily replaced by other DNA binding proteins, such as transcription factors. Unfortunately, all experiments employed ectopically expressed H3 variants with the tag located at the C-terminus, close to the H3-H3-dimerization interface. Taking into account the dynamic nature of the nucleosome (19,20), this could, potentially, have differential influence on H3.3/H2A.Z-containing double variant nucleosomes compared to those containing H3/H2A.Z ones *in vivo*. Surprisingly, another study (99), employing nucleosomes reconstituted, *in vitro,* from human histones purified from E. coli, did not find any drastic changes for H3.3/H2A.Z-containing double variant nucleosomes. As stated above, technical differences in these studies can explain the different outcomes and hamper the drawing of final conclusions.

In the last two years, the importance of the H2A.Z C-terminus for nucleosome stability and chromatin association has been explored in yeast as well as in human (26,27,171). Two studies in yeast revealed that C-terminal deletions, depending on the extent of truncation, decrease or completely abolish chromatin association (27,171). As expected, loss of chromatin association leads to phenotypes similar to the complete knock-out of the H2A.Z gene in yeast, such as reduced resistance to genotoxic stress and spreading of heterochromatin into euchromatic regions, indicating that chromatin association is essential for H2A.Z function. Interestingly, by analyzing chimeric proteins, both groups found that the C-terminus of canonical H2A can completely restore chromatin association and rescue the H2A.Z knock-out phenotype, consistent with the idea that the primary function of the H2A.Z C-terminus in yeast is anchoring the protein to chromatin.

The recent discovery of Z.2.2, an alternatively spliced H2A.Z isoform, provided fascinating new insights into the role of H2A.Z´s C-terminus (26,27). Alternative splicing of the H2A.Z.2 gene gives rise to two isoforms, Z.2.1 and Z.2.2 (Figure 3). The novel isoform 2 is different from isoform 1 in two respects: it is 14 amino acids shorter and has a stretch of six unique amino acids in its very C-terminus. We and others have found identical properties with respect to chromatin association and nucleosome stability of Z.2.2 (26,27). In contrast to isoform 1, the cellular pool of the shorter isoform 2 is not completely associated with chromatin but exhibits a major soluble pool. Moreover, the chromatin-bound fraction is less tightly incorporated into nucleosomes, both *in vitro* and *in vivo*, further establishing the importance of H2A.Z´s C-terminus in providing stable chromatin

incorporation. To further break down which of the two distinguishing properties of Z.2.2´s C-terminus, its shortened length or unique amino acid sequence, are critical for its decreased extent and stability of chromatin incorporation, we analyzed deletion mutants and chimeric proteins. Surprisingly, mere shortening of Z.2.1 to the same length as Z.2.2 does not dramatically alter chromatin incorporation *in vivo*. In contrast, transferring Z.2.2´s unique docking domain to the respective site of H2A results in a protein with chromatin incorporation virtually identical to Z.2.2. These results demonstrate that the specific sequence within Z.2.2´s docking domain and not just its shortened length is the critical determinant for the unique properties of Z.2.2 with respect to its incorporation into chromatin. To gain insight into the underlying structural alterations in Z.2.2 nucleosomes, we performed MD simulations that point towards a more flexible C-terminus of Z.2.2 which is also more distant to the H3 αN helix, thereby reducing interactions with the $(H3-H4)_2$-tetramer in a sequence-specific manner. These MD simulations are further supported by Z.2.2´s inability to form stable histone octamers under high salt conditions, which is in contrast to Z.2.1 or canonical H2A. From the results obtained *in silico* and *in vitro*, one can hypothesize that the changed interaction interface with a striking increase in C-terminal flexibility leads to less stable DNA organization but increased DNA breathing instead, which is confirmed by decreased resistance to MNase digest. Taken together, Z.2.2 is an intriguing protein that, by specific changes in its C-terminus, drastically alters basic H2A.Z properties possibly leading to a shift in H2A.Z function in certain tissues of high Z.2.2 abundance, e. g. brain tissues (26,27).

### 2.1.4 H2A.Bbd

Because of its shorter length and highly divergent amino acid sequence (about 50% identical to H2A (113)), H2A.Bbd was expected to alter nucleosome structure and organization of DNA significantly. Indeed, several studies investigated H2A.Bbd nucleosome properties mostly using *in vitro* assays; all of which consistently revealed an open structure of H2A.Bbd containing chromatin. H2A.Bbd organizes DNA less tightly, leading to a more relaxed and elongated structure with almost 180° between the DNA entry/exit sites in contrast to the V-shaped canonical nucleosomes (119,172). These differences in nucleosomal DNA constraint are concomitant with less resistance to digestion by MNase (119,172). Notably, no H2A.Bbd crystal structure is available thus far, compatible with global structural alterations leading to a more dynamic particle that prevents formation of well diffracting crystals (19). This is in line with findings that DNaseI footprinting experiments showed significant changes of DNA organization in the H2A.Bbd containing nucleosome (118,120,172). Analysis of H2A.Bbd nucleosome stability showed that it does not refold into histone octamers under high salt conditions (26,119), indicating weaker interaction of H2A.Bbd-H2B dimers with the $(H3-H4)_2$-tetramer ultimately resulting in reduced nucleosome stability (26,115,172,173). In accordance with *in vitro* studies discussed above, determination of H2A.Bbd mobility *in vivo* using FRAP showed a much faster exchange than canonical H2A (26,174).

Several studies investigated the role of H2A.Bbd´s C-terminus for the observed changes in structure and stability. As apparent for Z.2.2, the C-terminus of H2A.Bbd differs from canonical H2A in both length and amino acid composition. Hence, the question was whether the shortened length or amino acid sequence is the main determinant for H2A.Bbd´s unique properties. Bao et al. analyzed canonical H2A truncations *in vitro*. They found that mere shortening of the C-terminal tail neither impairs histone octamer assembly under high salt conditions nor significantly alters DNA organization, indicating that H2A.Bbd´s shortened length cannot be the sole determinant for its characteristic properties (119). In contrast, chimeric proteins consisting of H2A.Bbd´s C-terminus/docking domain fused to the N-terminal part of canonical H2A exhibit properties characteristic for H2A.Bbd. They do not refold into histone octamers under high salt conditions and bind DNA less tightly with a H2A.Bbd-like geometry, pointing towards an essential role of H2A.Bbd´s docking domain in defining interactions with the $(H3-H4)_2$-tetramer as well as DNA (119,120,172). This role is further underlined by the finding that the C-terminus of canonical H2A fused to the H2A.Bbd histone fold is sufficient to organize DNA comparably to canonical H2A nucleosomes and partly restores the normal V-shaped geometry (172). The analysis of H2AL2, an H2A.Bbd-like protein present in mouse (116,117), revealed striking similarities to H2A.Bbd, as H2AL2 nucleosomes arrange nucleosomal DNA in a more open structure as canonical ones (175). In conclusion, H2A.Bbd incorporation results in reduction of nucleosome stability and structural constraint of nucleosomal DNA in a manner highly dependent on its docking domain, consistent with its presence at active genes during spermatogenesis (117).

### 2.1.5    MacroH2A

The crystal structure of the macroH2A nucleosome provided important clues about alterations upon incorporation of this variant (138). Despite the overall high structural similarity to the canonical particle, a four amino acid sequence in L1, which is implicated in interactions between the two H2A-H2B dimers within the nucleosome (Figure 1 and 2), showed noticeable differences. This finding led to the suggestion of increased stability for macroH2A-containing nucleosomes due to stronger interactions between the two macroH2A-H2B dimers. Whether this is indeed the case was addressed by analysis of the macroH2A-containing histone octamer in the absence of DNA. Interestingly, Chakravarthy et al. found that the macroH2A-containing octamer is less reliant on high salt stabilization than the canonical one (139). Canonical octamers dissociate if salt concentration is lowered to 1.1 M NaCl whereas macroH2A octamers are still completely stable under these conditions. Importantly, by mutational studies, the authors could show that the four amino acid substitutions in the L1 region are solely responsible for the changes observed in octamer stability, pointing towards the importance of the L1 region in defining interactions within the nucleosome. Consistent with these findings, from studies on *in vitro* reconstituted nucleosomes, analysis of native chromatin from chicken cells showed an increased stability of macroH2A chromatin incorporation as well (155). Taken together macroH2A increases nucleosome stability by alterations within a four

amino acid stretch in L1, which is in strong contrast to Z.2.2 or H2A.Bbd that lead to a significant decrease in nucleosome stability mediated by their characteristic docking domains.

Further evidence supporting macroH2A´s role in constituting nucleosomes that are more stable and static is provided by the finding that chaperone-assisted H2A(variant)-H2B dimer exchange is inhibited by macroH2A containing nucleosomes (139). Interestingly, the L1 region and the docking domain are not sufficient to transfer this property to canonical H2A, thereby indicating the importance of other regions for macroH2A´s static nature. The authors state that the best explanation for these findings is reduction of macroH2A´s relative affinity to the chaperone used (yNAP1) compared to canonical H2A. This points towards a question neglected in most *in vitro* studies, namely the influence of the relative affinity of histone variants to factors other than the nucleosome such as chaperones and remodeling complexes. In principle, the affinity of a histone variant to soluble protein complexes promoting its absence from chromatin must also be considered, since these factors are abundant and contribute significantly to the equilibrium between soluble and chromatin-bound histone variant in the cell. However, it is complicated to exhaustively analyze these protein complexes *in vitro* due to their immense diversity in the living cell. On the other hand, *in vivo* assays such as FRAP can provide valuable insights; unfortunately, these data are hard to dissect due to the complexity of the experimental system, i.e. the cell. Thorough analyses should therefore follow a complementary approach employing both *in vitro* and *in vivo* analyses to compensate for the limitations of each of them.

## 2.2    THE INFLUENCE OF H2A VARIANTS ON CHROMATIN STRUCTURE

The work discussed in the last section focused primarily on the influence of H2A variants on the nucleosome and its basic properties. To understand how H2A variants can alter the "monomeric building block" of chromatin provided us with plenty of insight into the mechanisms of their biological functions. In the cell however, chromatin is not present in a linear "beads-on-a-string" conformation but adopts higher-order structures impacted by the complex interplay of DNA, core and linker histones and other chromatin architectural proteins (reviewed in (176-179)). Short-range (intramolecular) interactions within a linear chromatin strand lead to a more compact secondary structure, the 30 nm fiber. In addition (or instead), long-range (intermolecular) interactions between distinct chromatin fibers lead to large oligomeric tertiary complexes.

To understand the intricate relationship of structure and function in chromatin, the influence of chromatin components on secondary and tertiary chromatin structures must be taken into account. In this section, I discuss the influence of histone variants on chromatin structure with a special focus on the crucial role of the H2A acidic patch.

### 2.2.1    The H2A acidic patch is a key regulator of higher-order chromatin structure

H2A and H2B form an obligate dimer under physiological conditions (180), hence it is the structure of the H2A-H2B dimer that must be considered to be involved in biologically relevant protein-protein interactions (181). Despite the overall basic nature of histones, the nucleosome crystal structure revealed the presence of an acidic patch on the surface of the H2A-H2B dimer which is mainly comprised by H2A (six out of seven amino acids) (16). Interestingly, in this structure, the H4 N-terminal tail (K16-N25) contacts the acidic patch on the adjacent nucleosome and this contact is required for crystallization (16). In addition to the H4 tail, interactions with at least five more non-histone proteins make the acidic patch an important binding site in chromatin with the potential to differentially contribute to diverse biological processes by its alteration due to H2A variant incorporation (181). In support of this notion, Interleukin-33 interacts with the acidic patch of H2A or H2A.Z but binding to H2A.Bbd, which lacks an acidic patch, is strongly decreased (182).

The importance of the H4 tail for the establishment of proper secondary and tertiary chromatin structure has been established and depends on its charge and PTMs (reviewed in (176)). Richmond and coworkers could show that the H4 tail has a critical role beyond the other histone tails for both intra- and intermolecular interactions (183). Furthermore, by using mutant proteins, they could crosslink the H4 tail (H4-V21C) to the acidic patch of H2A (H2A-E64C) upon array folding, thereby providing evidence for the direct interaction in solution when a more compact secondary structure, the 30 nm fiber, is formed (184). More recently, intermolecular crosslinks between H4-V21C and H2A-E64C have been reported as well (185). However, additional intermolecular contacts between chromatin fibers must be very important since arrays containing only $(H3-H4)_2$ tetramers can

oligomerize just as nucleosomal arrays (186) and binding of the H4 tail to DNA is another important mechanism for the establishment of intermolecular interactions (187).

The modulation of chromatin folding by PTMs is of particular importance as histone PTMs are abundant and implicated in a multitude of biological processes (4). Two H4 tail modifications have been studied with regards to their influence on chromatin folding, acetylation of lysine 16 and trimethylation of lysine 20. These two modifications appear to have opposing biological functions as H4K16ac is associated with euchromatin and active transcription, whereas H4K20me3 plays a role in heterochromatin formation (reviewed in (188)). In accordance with the opposing biological functions, H4K16ac inhibits both intra- and intermolecular interactions of the H4 tail and consequently promotes an open chromatin structure (189). Contrariwise, H4K20me3 leads to more efficient intramolecular folding which results in a more compact secondary chromatin structure without influencing intermolecular folding ((190), reviewed in (178)). Taken together, the interaction of the H4 tail with the acidic patch of H2A is important for both short-range (intramolecular) and long-range (intermolecular) chromatin interactions and can be modulated by PTMs of the H4 tail as well as incorporation of H2A variants (see below).

**2.2.2    Alterations of the acidic patch due to H2A variant incorporation influence higher-order chromatin structure**

Two H2A variants, H2A.Z and H2A.Bbd, have been studied with regards to the influence of their acidic patch on secondary and tertiary chromatin structure. These studies contributed significantly to our understanding of the importance of the acidic patch on higher-order chromatin structure. Compared to canonical H2A, H2A.Z has an extended acidic patch whereas H2A.Bbd virtually lacks it (Figure 2). In H2A.X and macroH2A, the residues constituting the acidic patch are completely conserved.

Ten years ago, Tremethick and coworkers investigated the influence of H2A.Z on chromatin folding *in vitro* (95). They reported that arrays assembled with H2A.Z exhibit increased intramolecular folding and therefore a more compact secondary structure than canonical arrays. Interestingly, H2A.Z arrays impair intermolecular contacts and therefore array oligomerization. Two years later, the authors extended their studies by employing acidic patch mutants and H4 tail deletions to mechanistically understand H2A.Z´s influence on chromatin folding (191). They found that the extended acidic patch of H2A.Z leads to increased intramolecular folding and decreased intermolecular oligomerization. Furthermore, intramolecular folding also requires the H4 tail suggesting that the same mechanism of folding seen for canonical arrays also applies to H2A.Z variant-containing chromatin fibers (183), but with a higher affinity due to the larger acidic patch of H2A.Z which allows stronger electrostatic interactions. In addition, they analyzed HP1$\alpha$ (heterochromatin binding protein 1 $\alpha$) binding to arrays containing either canonical H2A or H2A.Z. The authors found that HP1$\alpha$ binds highly folded

chromatin (hence H2A.Z containing chromatin) to further enhance intramolecular folding but it does not bridge chromatin fibers. These findings could be relevant for structure and function of H2A.Z containing chromatin at centromeres *in vivo* (192). Greaves et al. reported the association of H2A.Z with both major satellite repeats in pericentric heterochromatin and minor satellite repeats in CENP-A-containing centric chromatin. Moreover, they suggest that H2A.Z could be important for the assembly of H3K4me2 euchromatin into compact chromatin structures within the CENP-A-containing centric chromatin regions, required for centromere function on chromosomes relatively deficient in pericentric heterochromatin.

To explain opposing influences of H2A.Z´s extended acidic patch on chromatin compaction and oligomerization, they propose a model based on competition between intra- and intermolecular interaction partners for the H4 tail (Figure 4). The H4 tail can either interact with the acidic patch within one fiber and lead to a more compact secondary structure or it can participate in other contacts that promote array oligomerization, for example with DNA of another chromatin fiber (187). Hence, the interaction of the H4 tail with the acidic patch in an intramolecular manner inhibits any other (intermolecular) interactions and therefore inhibits oligomerization. Therefore, the stronger the interaction of the acidic patch with the H4 tail, the more favored the compact secondary structure and the less favored array oligomerization and *vice versa*.

To test whether the competition model applies more generally, the Tremethick lab analyzed arrays assembled *in vitro* with H2A.Bbd, canonical H2A and acidic patch mutants of both (193). In line with the proposed model, H2A.Bbd inhibits intramolecular folding but fosters internucleosomal oligomerization. Moreover, H2A.Bbd mutants with a restored acidic patch increase the tendency to form compact secondary structures depending on the H4 tail and the extent of the acidic patch restoration, whereas mutants of canonical H2A that lack the acidic patch form compact secondary structures less efficiently than wild type but oligomerize chromatin fibers more efficiently. Taken together, these analyses support the competition model and further strengthen the view of the acidic patch as a key regulator of chromatin structure. On the one hand, the extended acidic patch of H2A.Z results in more efficient formation of compact secondary structures while inhibiting oligomerization of chromatin fibers; on the other hand the smaller acidic patch of H2A.Bbd has contrasting effects by fostering oligomerization at expense of secondary structure formation (Figure 4).
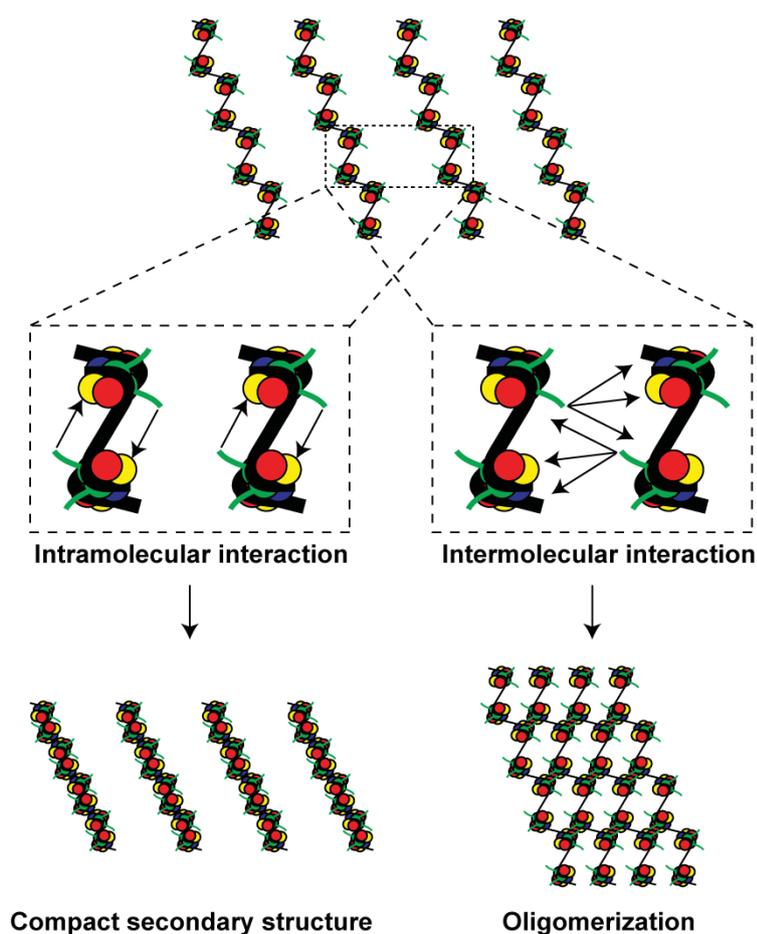
**Figure 4: The acidic patch regulates chromatin structure by interaction with the H4 tail.** The H4 tail can engage in intramolecular interactions with the acidic patch of neighboring nucleosomes within the same chromatin fiber to form more compact secondary structures (left). Alternatively, it can form different intermolecular interactions with DNA and histones of other chromatin fibers to form large tertiary oligomeric complexes (right). Which interactions are preferred is influenced by acidic patch alterations in H2A variants. The extended acidic patch of H2A.Z fosters compact secondary structure formation (right) whereas the reduced acidic patch of H2A.Bbd leads to preferred oligomerization (left). DNA is shown in black, H2A in yellow, H2B in red, H3 in blue and H4 in green. Flexible histone tails for histones other than H4 are omitted for clarity.

The lack of acidic patch in H2A.Bbd could also explain why no crystal structure is available thus far. Luger et al. reported that the contact of the H4 tail with the acidic patch, which cannot be formed with H2A.Bbd nucleosomes, is required for crystallization (16). Whether this is the only reason for the inability of H2A.Bbd nucleosomes to be crystallized could be easily tested by using the reported H2A.Bbd mutant with a restored acidic patch (193).

Further evidence of how sensitive chromatin folding responds to alterations of the acidic patch was provided by the experimental comparison of human H2A.Bbd with its mouse homologue (117). A single amino acid substitution from human to mouse (T100D) increases the acidic patch of mouse H2A.Bbd that in turn is able to partially fold chromatin into more compact secondary structures. Mutation of this residue back to threonine, as found in the human protein, disables higher-order chromatin folding indicating its functional importance.

Interestingly, Z.2.2 combines features of H2A.Bbd as well as H2A.Z. One the one hand, it significantly destabilizes nucleosomes, similarly to H2A.Bbd; on the other hand it completely retains the extended acidic patch of H2A.Z (Figure 2). Thus far, no analysis of chromatin folding and oligomerization of Z.2.2 containing arrays are available but from the literature on the acidic patch, one would expect that Z.2.2, although severely destabilizing nucleosomes, behaves like H2A.Z and allows the formation of compact secondary chromatin structures due to the key role of the acidic patch. This

would give Z.2.2 an intriguing role by promoting a compact chromatin structure of unstable nucleosomes. Future studies will ascertain whether this is indeed the case or not.

In addition to their structural analyses, Zhou et al. also functionally investigated the influence of the acidic patch and its impact on secondary and tertiary chromatin structure on transcription (193). Surprisingly, efficient transcription can occur within large oligomeric chromatin structures but is only impaired by the formation of compact secondary chromatin structures by H2A.Bbd mutants that restore the acidic patch. This is consistent with H2A.Bbd´s euchromatic localization and role in gene activation by inhibiting the formation of highly compact chromatin structures (117). For H2A.Z, these findings would suggest a repressive role in gene transcription if present in large domain. However, the localization of H2A.Z on gene promoters does not appear to build large domains but to rather position individual nucleosomes (see above) and is therefore hard to compare with *in vitro* experiments employing nucleosomal arrays.

Another important player involved in establishment of higher-order chromatin structure is the linker histone H1 (176,178,194) that facilitates formation of chromatin higher-order structures by neutralization of the negatively charged DNA. As discussed above, both H2A.Bbd and H2A.Z mononucleosomes bind H1 less efficiently than canonical H2A (99,120). To our knowledge, linker histone binding has not been analyzed on H2A variant-containing chromatin fibers, which might influence this interaction by adoption of secondary and tertiary structures. However, if H2A.Z chromatin indeed binds H1 less efficiently than canonical H2A *in vivo*, this might compensate for the higher tendency of H2A.Z to form compact secondary structures and even out structural differences between H2A- and H2A.Z-containing chromatin. More interestingly, it is tempting to speculate that H2A.Z and H2A form structurally and functionally different chromatin due to their different inherent properties to engage in secondary and tertiary chromatin folding as well as in recruitment of H1 and other chromatin factors.

It is important to mention that another group found contrasting influences of H2A.Z incorporation on chromatin folding (166). However, the experiments are hard to compare since here, monovalent cations have been used to enhance chromatin folding whereas the Tremethick lab used divalent cations and the sources of histones were different in the studies by the two labs.

In conclusion, incorporation of H2A variants into chromatin can alter its secondary and tertiary structure. The key regulators for these alterations are the H2A acidic patch and the H4 tail, which together define different kinds of interactions with distinct structural and functional outcomes. Moreover, the different affinities of H2A variant-containing nucleosomes to the linker histone H1 could also play an important role in specifying distinct chromosomal domains. Although a lot of progress has been made in the last decade, the complex composition of chromatin *in vivo* makes it hard to set up suitable models *in vitro*.

## 2.3    CONCLUDING REMARKS

The importance of H2A variants in a multitude of biological processes is well established; however the mechanisms by which they function are not yet completely understood. From the literature available today, one can envision that histone variants, in general, and H2A variants, in particular, function by conferring characteristic properties to chromatin both on the nucleosomal and higher-order structural level. Our mechanistic understanding of chromatin structure alterations and histone domains involved revealed the functional significance of different regions in H2A, like the L1 loop, the docking domain and C-terminal tail as well as the acidic patch. The recent discovery of the H2A.Z splice isoform Z.2.2 questions the completeness of our knowledge on existing histone variants.

Although a lot of progress was made in the last decade, we are far from understanding the structure/function interplay of H2A variants. Notably, the importance of H2A variants in a tissue specific manner was shown for H2A.Bbd (117) and was suggested for Z.2.2 (26,27). In future studies it will be of particular interest to analyze histone variants with respect to their tissue specific influences on chromatin structure and function.

# REFERENCES

1. Bonisch, C., Nieratschker, S.M., Orfanos, N.K. and Hake, S.B. (2008) Chromatin proteomics and epigenetic regulatory circuits. *Expert review of proteomics*, **5**, 105-119.

2. Bird, A. (2002) DNA methylation patterns and epigenetic memory. *Genes & development*, **16**, 6-21.

3. Becker, P.B. and Horz, W. (2002) ATP-dependent nucleosome remodeling. *Annual review of biochemistry*, **71**, 247-273.

4. Bannister, A.J. and Kouzarides, T. (2011) Regulation of chromatin by histone modifications. *Cell research*, **21**, 381-395.

5. Zhang, R., Zhang, L. and Yu, W. (2012) Genome-wide expression of non-coding RNA and global chromatin modification. *Acta biochimica et biophysica Sinica*, **44**, 40-47.

6. Lanctot, C., Cheutin, T., Cremer, M., Cavalli, G. and Cremer, T. (2007) Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. *Nature reviews. Genetics*, **8**, 104-115.

7. Talbert, P.B. and Henikoff, S. (2010) Histone variants--ancient wrap artists of the epigenome. *Nature reviews. Molecular cell biology*, **11**, 264-275.

8. Flemming, W. (1882) *Zellsubstanz, Kern und Zelltheilung*. Vogel, Leipzig.

9. Olins, D.E. and Olins, A.L. (2003) Chromatin history: our view from the bridge. *Nature reviews. Molecular cell biology*, **4**, 809-814.

10. Heitz, E. (1928) Das Heterochromatin der Moose. *Jb. Wiss. Bot.*, **69**, 728–818.

11. Avery, O.T., Macleod, C.M. and McCarty, M. (1944) Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types : Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type Iii. *The Journal of experimental medicine*, **79**, 137-158.

12. Watson, J.D. and Crick, F.H. (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, **171**, 737-738.

13. Hewish, D.R. and Burgoyne, L.A. (1973) Chromatin sub-structure. The digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease. *Biochemical and biophysical research communications*, **52**, 504-510.

14. Olins, A.L. and Olins, D.E. (1974) Spheroid chromatin units (v bodies). *Science*, **183**, 330-332.

15. Woodcock, C.L., Safer, J.P. and Stanchfield, J.E. (1976) Structural repeating units in chromatin. I. Evidence for their general occurrence. *Experimental cell research*, **97**, 101-110.

16. Luger, K., Mader, 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.

17. 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.

18. 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.

19. Andrews, A.J. and Luger, K. (2011) Nucleosome structure(s) and stability: variations on a theme. *Annual review of biophysics*, **40**, 99-117.

20. Zlatanova, J., Bishop, T.C., Victor, J.M., Jackson, V. and van Holde, K. (2009) The nucleosome family: dynamic and growing. *Structure*, **17**, 160-171.

21. van Holde, K. and Zlatanova, J. (1999) The nucleosome core particle: does it have structural and physiologic relevance? *BioEssays : news and reviews in molecular, cellular and developmental biology*, **21**, 776-780.

22. Bohm, V., Hieb, A.R., Andrews, A.J., Gansen, A., Rocker, A., Toth, K., Luger, K. and Langowski, J. (2011) Nucleosome accessibility governed by the dimer/tetramer interface. *Nucleic acids research*, **39**, 3093-3102.

23. Marzluff, W.F., Gongidi, P., Woods, K.R., Jin, J. and Maltais, L.J. (2002) The human and mouse replication-dependent histone genes. *Genomics*, **80**, 487-498.

24. Marzluff, W.F., Wagner, E.J. and Duronio, R.J. (2008) Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nature reviews. Genetics*, **9**, 843-854.

25. Rasmussen, T.P., Huang, T., Mastrangelo, M.A., Loring, J., Panning, B. and Jaenisch, R. (1999) Messenger RNAs encoding mouse histone macroH2A1 isoforms are expressed at similar levels in male and female cells and result from alternative splicing. *Nucleic acids research*, **27**, 3685-3689.

26. Bonisch, C., Schneider, K., Punzeler, 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 research*.

27. Wratting, D., Thistlethwaite, A., Harris, M., Zeef, L.A. and Millar, C.B. (2012) A conserved function for the H2A.Z C-terminus. *The Journal of biological chemistry*.

28. The PyMOL Molecular Graphics System, V.r.p., Schrödinger, LLC.

29. Malik, H.S. and Henikoff, S. (2003) Phylogenomics of the nucleosome. *Nature structural biology*, **10**, 882-891.

30. Hayashi, H., Nomoto, M. and Iwai, K. (1984) Tetrahymena histone H4. Complete amino acid sequences of two variants. *Journal of biochemistry*, **96**, 1449-1456.

31. Siegel, T.N., Hekstra, D.R., Kemp, L.E., Figueiredo, L.M., Lowell, J.E., Fenyo, D., Wang, X., Dewell, S. and Cross, G.A. (2009) Four histone variants mark the boundaries of polycistronic transcription units in Trypanosoma brucei. *Genes & development*, **23**, 1063-1076.

32. Moosmann, A., Campsteijn, C., Jansen, P.W., Nasrallah, C., Raasholm, M., Stunnenberg, H.G. and Thompson, E.M. (2011) Histone variant innovation in a rapidly evolving chordate lineage. *BMC evolutionary biology*, **11**, 208.

33. Ausio, J. and Abbott, D.W. (2002) The many tales of a tail: carboxyl-terminal tail heterogeneity specializes histone H2A variants for defined chromatin function. *Biochemistry*, **41**, 5945-5949.

34. Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R. *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947-2948.

35. Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J. and Lopez, R. (2010) A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic acids research*, **38**, W695-699.

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

37. Rodrigues Jde, A., Brandt, W.F. and von Holt, C. (1985) The amino acid sequence of wheat histone H2A(1). A core histone with a C-terminal extension. *European journal of biochemistry / FEBS*, **150**, 499-505.

38. Yi, H., Sardesai, N., Fujinuma, T., Chan, C.W., Veena and Gelvin, S.B. (2006) Constitutive expression exposes functional redundancy between the Arabidopsis histone H2A gene HTA1 and other H2A gene family members. *The Plant cell*, **18**, 1575-1589.

39.     Green, G.R. (2001) Phosphorylation of histone variant regions in chromatin: unlocking the linker? *Biochemistry and cell biology = Biochimie et biologie cellulaire*, **79**, 275-287.

40.     Lindsey, G.G., Orgeig, S., Thompson, P., Davies, N. and Maeder, D.L. (1991) Extended C-terminal tail of wheat histone H2A interacts with DNA of the "linker" region. *Journal of molecular biology*, **218**, 805-813.

41.     Green, G.R., Gustavsen, L.C. and Poccia, D.L. (1990) Phosphorylation of Plant H2A Histones. *Plant physiology*, **93**, 1241-1245.

42.     Gladyshev, E. and Meselson, M. (2008) Extreme resistance of bdelloid rotifers to ionizing radiation. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 5139-5144.

43.     Van Doninck, K., Mandigo, M.L., Hur, J.H., Wang, P., Guglielmini, J., Milinkovitch, M.C., Lane, W.S. and Meselson, M. (2009) Phylogenomics of unusual histone H2A Variants in Bdelloid rotifers. *PLoS genetics*, **5**, e1000401.

44.     Boulard, M., Bouvet, P., Kundu, T.K. and Dimitrov, S. (2007) Histone variant nucleosomes: structure, function and implication in disease. *Sub-cellular biochemistry*, **41**, 71-89.

45.     Banaszynski, L.A., Allis, C.D. and Lewis, P.W. (2010) Histone variants in metazoan development. *Developmental cell*, **19**, 662-674.

46.     Izzo, A., Kamieniarz, K. and Schneider, R. (2008) The histone H1 family: specific members, specific functions? *Biological chemistry*, **389**, 333-343.

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

48.     Pinto, D.M. and Flaus, A. (2010) Structure and function of histone H2AX. *Sub-cellular biochemistry*, **50**, 55-78.

49.     Mannironi, C., Bonner, W.M. and Hatch, C.L. (1989) H2A.X. a histone isoprotein with a conserved C-terminal sequence, is encoded by a novel mRNA with both DNA replication type and polyA 3' processing signals. *Nucleic acids research*, **17**, 9113-9126.

50.     Nagata, T., Kato, T., Morita, T., Nozaki, M., Kubota, H., Yagi, H. and Matsushiro, A. (1991) Polyadenylated and 3' processed mRNAs are transcribed from the mouse histone H2A.X gene. *Nucleic acids research*, **19**, 2441-2447.

51.     Rogakou, E.P., Boon, C., Redon, C. and Bonner, W.M. (1999) Megabase chromatin domains involved in DNA double-strand breaks in vivo. *The Journal of cell biology*, **146**, 905-916.

52.     Heo, K., Kim, H., Choi, S.H., Choi, J., Kim, K., Gu, J., Lieber, M.R., Yang, A.S. and An, W. (2008) FACT-mediated exchange of histone variant H2AX regulated by phosphorylation of H2AX and ADP-ribosylation of Spt16. *Molecular cell*, **30**, 86-97.

53.     Li, A., Yu, Y., Lee, S.C., Ishibashi, T., Lees-Miller, S.P. and Ausio, J. (2010) Phosphorylation of histone H2A.X by DNA-dependent protein kinase is not affected by core histone acetylation, but it alters nucleosome stability and histone H1 binding. *The Journal of biological chemistry*, **285**, 17778-17788.

54.     Downs, J.A., Lowndes, N.F. and Jackson, S.P. (2000) A role for Saccharomyces cerevisiae histone H2A in DNA repair. *Nature*, **408**, 1001-1004.

55.     Fink, M., Imholz, D. and Thoma, F. (2007) Contribution of the serine 129 of histone H2A to chromatin structure. *Molecular and cellular biology*, **27**, 3589-3600.

56.     Downs, J.A., Nussenzweig, M.C. and Nussenzweig, A. (2007) Chromatin dynamics and the preservation of genetic information. *Nature*, **447**, 951-958.

57.     Celeste, A., Fernandez-Capetillo, O., Kruhlak, M.J., Pilch, D.R., Staudt, D.W., Lee, A., Bonner, R.F., Bonner, W.M. and Nussenzweig, A. (2003) Histone H2AX phosphorylation is dispensable for the initial recognition of DNA breaks. *Nature cell biology*, **5**, 675-679.

58. Kruhlak, M.J., Celeste, A., Dellaire, G., Fernandez-Capetillo, O., Muller, W.G., McNally, J.G., Bazett-Jones, D.P. and Nussenzweig, A. (2006) Changes in chromatin structure and mobility in living cells at sites of DNA double-strand breaks. *The Journal of cell biology*, **172**, 823-834.

59. Xiao, A., Li, H., Shechter, D., Ahn, S.H., Fabrizio, L.A., Erdjument-Bromage, H., Ishibe-Murakami, S., Wang, B., Tempst, P., Hofmann, K. *et al.* (2009) WSTF regulates the H2A.X DNA damage response via a novel tyrosine kinase activity. *Nature*, **457**, 57-62.

60. Krishnan, N., Jeong, D.G., Jung, S.K., Ryu, S.E., Xiao, A., Allis, C.D., Kim, S.J. and Tonks, N.K. (2009) Dephosphorylation of the C-terminal tyrosyl residue of the DNA damage-related histone H2A.X is mediated by the protein phosphatase eyes absent. *The Journal of biological chemistry*, **284**, 16066-16070.

61. Singh, R.K. and Gunjan, A. (2011) Histone tyrosine phosphorylation comes of age. *Epigenetics : official journal of the DNA Methylation Society*, **6**, 153-160.

62. Cook, P.J., Ju, B.G., Telese, F., Wang, X., Glass, C.K. and Rosenfeld, M.G. (2009) Tyrosine dephosphorylation of H2AX modulates apoptosis and survival decisions. *Nature*, **458**, 591-596.

63. Bennetzen, M.V., Larsen, D.H., Bunkenborg, J., Bartek, J., Lukas, J. and Andersen, J.S. (2010) Site-specific phosphorylation dynamics of the nuclear proteome during the DNA damage response. *Molecular & cellular proteomics : MCP*, **9**, 1314-1323.

64. Nashun, B., Yukawa, M., Liu, H., Akiyama, T. and Aoki, F. (2010) Changes in the nuclear deposition of histone H2A variants during pre-implantation development in mice. *Development*, **137**, 3785-3794.

65. Thatcher, T.H. and Gorovsky, M.A. (1994) Phylogenetic analysis of the core histones H2A, H2B, H3, and H4. *Nucleic acids research*, **22**, 174-179.

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

67. Lowell, J.E., Kaiser, F., Janzen, C.J. and Cross, G.A. (2005) Histone H2AZ dimerizes with a novel variant H2B and is enriched at repetitive DNA in Trypanosoma brucei. *Journal of cell science*, **118**, 5721-5730.

68. 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. *Current biology : CB*, **11**, 1183-1187.

69. van Daal, A. and Elgin, S.C. (1992) A histone variant, H2AvD, is essential in Drosophila melanogaster. *Molecular biology of the cell*, **3**, 593-602.

70. Iouzalen, N., Moreau, J. and Mechali, M. (1996) H2A.ZI, a new variant histone expressed during Xenopus early development exhibits several distinct features from the core histone H2A. *Nucleic acids research*, **24**, 3947-3952.

71. Liu, X., Li, B. and GorovskyMa. (1996) Essential and nonessential histone H2A variants in Tetrahymena thermophila. *Molecular and cellular biology*, **16**, 4305-4311.

72. Jackson, J.D. and Gorovsky, M.A. (2000) Histone H2A.Z has a conserved function that is distinct from that of the major H2A sequence variants. *Nucleic acids research*, **28**, 3811-3816.

73. Carr, A.M., Dorrington, S.M., Hindley, J., Phear, G.A., Aves, S.J. and Nurse, P. (1994) Analysis of a histone H2A variant from fission yeast: evidence for a role in chromosome stability. *Molecular & general genetics : MGG*, **245**, 628-635.

74. Clarkson, M.J., Wells, J.R., Gibson, F., Saint, R. and Tremethick, D.J. (1999) Regions of variant histone His2AvD required for Drosophila development. *Nature*, **399**, 694-697.

75. Wu, W.H., Alami, S., Luk, E., Wu, C.H., Sen, S., Mizuguchi, G., Wei, D. and Wu, C. (2005) Swc2 is a widely conserved H2AZ-binding module essential for ATP-dependent histone exchange. *Nature structural & molecular biology*, **12**, 1064-1071.

76.     Jensen, K., Santisteban, M.S., Urekar, C. and Smith, M.M. (2011) Histone H2A.Z acid patch residues required for deposition and function. *Molecular genetics and genomics : MGG*, **285**, 287-296.

77.     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. *Nature structural biology*, **7**, 1121-1124.

78.     Chakravarthy, S., Bao, Y., Roberts, V.A., Tremethick, D. and Luger, K. (2004) Structural characterization of histone H2A variants. *Cold Spring Harbor symposia on quantitative biology*, **69**, 227-234.

79.     Viens, A., Mechold, U., Brouillard, F., Gilbert, C., Leclerc, P. and Ogryzko, V. (2006) Analysis of human histone H2AZ deposition in vivo argues against its direct role in epigenetic templating mechanisms. *Molecular and cellular biology*, **26**, 5325-5335.

80.     Thambirajah, A.A., Li, A., Ishibashi, T. and Ausio, J. (2009) New developments in post-translational modifications and functions of histone H2A variants. *Biochemistry and cell biology = Biochimie et biologie cellulaire*, **87**, 7-17.

81.     Kalocsay, M., Hiller, N.J. and Jentsch, S. (2009) Chromosome-wide Rad51 spreading and SUMO-H2A.Z-dependent chromosome fixation in response to a persistent DNA double-strand break. *Molecular cell*, **33**, 335-343.

82.     Sarcinella, E., Zuzarte, P.C., Lau, P.N., Draker, R. and Cheung, P. (2007) Monoubiquitylation of H2A.Z distinguishes its association with euchromatin or facultative heterochromatin. *Molecular and cellular biology*, **27**, 6457-6468.

83.     Thambirajah, A.A., Dryhurst, D., Ishibashi, T., Li, A., Maffey, A.H. and Ausio, J. (2006) H2A.Z stabilizes chromatin in a way that is dependent on core histone acetylation. *The Journal of biological chemistry*, **281**, 20036-20044.

84.     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.

85.     Svotelis, A., Gevry, N. and Gaudreau, L. (2009) Regulation of gene expression and cellular proliferation by histone H2A.Z. *Biochemistry and cell biology = Biochimie et biologie cellulaire*, **87**, 179-188.

86.     Draker, R. and Cheung, P. (2009) Transcriptional and epigenetic functions of histone variant H2A.Z. *Biochemistry and cell biology = Biochimie et biologie cellulaire*, **87**, 19-25.

87.     Altaf, M., Auger, A., Covic, M. and Cote, J. (2009) Connection between histone H2A variants and chromatin remodeling complexes. *Biochemistry and cell biology = Biochimie et biologie cellulaire*, **87**, 35-50.

88.     Guillemette, B. and Gaudreau, L. (2006) Reuniting the contrasting functions of H2A.Z. *Biochemistry and cell biology = Biochimie et biologie cellulaire*, **84**, 528-535.

89.     Billon, P. and Cote, J. (2012) Precise deposition of histone H2A.Z in chromatin for genome expression and maintenance. *Biochimica et biophysica acta*, **1819**, 290-302.

90.     Zhang, H., Roberts, D.N. and Cairns, B.R. (2005) Genome-wide dynamics of Htz1, a histone H2A variant that poises repressed/basal promoters for activation through histone loss. *Cell*, **123**, 219-231.

91.     Guillemette, B., Bataille, A.R., Gevry, 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 biology*, **3**, e384.

92.     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.

93.	Li, J., Krichevsky, A., Vaidya, M., Tzfira, T. and Citovsky, V. (2005) Uncoupling of the functions of the Arabidopsis VIP1 protein in transient and stable plant genetic transformation by Agrobacterium. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 5733-5738.

94.	Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I. and Zhao, K. (2007) High-resolution profiling of histone methylations in the human genome. *Cell*, **129**, 823-837.

95.	Fan, J.Y., Gordon, F., Luger, K., Hansen, J.C. and Tremethick, D.J. (2002) The essential histone variant H2A.Z regulates the equilibrium between different chromatin conformational states. *Nature structural biology*, **9**, 172-176.

96.	Flaus, A., Rencurel, C., Ferreira, H., Wiechens, N. and Owen-Hughes, T. (2004) Sin mutations alter inherent nucleosome mobility. *The EMBO journal*, **23**, 343-353.

97.	Albert, I., Mavrich, T.N., Tomsho, L.P., Qi, J., Zanton, S.J., Schuster, S.C. and Pugh, B.F. (2007) Translational and rotational settings of H2A.Z nucleosomes across the Saccharomyces cerevisiae genome. *Nature*, **446**, 572-576.

98.	Schones, D.E., Cui, K., Cuddapah, S., Roh, T.Y., Barski, A., Wang, Z., Wei, G. and Zhao, K. (2008) Dynamic regulation of nucleosome positioning in the human genome. *Cell*, **132**, 887-898.

99.	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.

100.	Hardy, S. and Robert, F. (2010) Random deposition of histone variants: A cellular mistake or a novel regulatory mechanism? *Epigenetics : official journal of the DNA Methylation Society*, **5**, 368-372.

101.	Krogan, N.J., Keogh, M.C., Datta, N., Sawa, C., Ryan, O.W., Ding, H., Haw, R.A., Pootoolal, J., Tong, A., Canadien, V. *et al.* (2003) A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. *Molecular cell*, **12**, 1565-1576.

102.	Mizuguchi, G., Shen, X., Landry, J., Wu, W.H., Sen, S. and Wu, C. (2004) ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science*, **303**, 343-348.

103.	Kobor, M.S., Venkatasubrahmanyam, S., Meneghini, M.D., Gin, J.W., Jennings, J.L., Link, A.J., Madhani, H.D. and Rine, J. (2004) A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin. *PLoS biology*, **2**, 587-599.

104.	Papamichos-Chronakis, M., Watanabe, S., Rando, O.J. and Peterson, C.L. (2011) Global regulation of H2A.Z localization by the INO80 chromatin-remodeling enzyme is essential for genome integrity. *Cell*, **144**, 200-213.

105.	Dryhurst, D., Ishibashi, T., Rose, K.L., Eirin-Lopez, 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.

106.	Eirin-Lopez, J.M., Gonzalez-Romero, R., Dryhurst, D., Ishibashi, T. and Ausio, J. (2009) The evolutionary differentiation of two histone H2A.Z variants in chordates (H2A.Z-1 and H2A.Z-2) is mediated by a stepwise mutation process that affects three amino acid residues. *BMC evolutionary biology*, **9**, 31.

107.	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 research*, **38**, 4263-4273.

108. Modrek, B. and Lee, C.J. (2003) Alternative splicing in the human, mouse and rat genomes is associated with an increased frequency of exon creation and/or loss. *Nature genetics*, **34**, 177-180.

109. Kalsotra, A. and Cooper, T.A. (2011) Functional consequences of developmentally regulated alternative splicing. *Nature reviews. Genetics*, **12**, 715-729.

110. Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P. and Burge, C.B. (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature*, **456**, 470-476.

111. Harata, M., Mochizuki, R. and Mizuno, S. (1999) Two isoforms of a human actin-related protein show nuclear localization and mutually selective expression between brain and other tissues. *Bioscience, biotechnology, and biochemistry*, **63**, 917-923.

112. Kuroda, Y., Oma, Y., Nishimori, K., Ohta, T. and Harata, M. (2002) Brain-specific expression of the nuclear actin-related protein ArpNalpha and its involvement in mammalian SWI/SNF chromatin remodeling complex. *Biochemical and biophysical research communications*, **299**, 328-334.

113. Chadwick, B.P. and Willard, H.F. (2001) A novel chromatin protein, distantly related to histone H2A, is largely excluded from the inactive X chromosome. *The Journal of cell biology*, **152**, 375-384.

114. Ishibashi, T., Li, A., Eirin-Lopez, J.M., Zhao, M., Missiaen, K., Abbott, D.W., Meistrich, M., Hendzel, M.J. and Ausio, J. (2010) H2A.Bbd: an X-chromosome-encoded histone involved in mammalian spermiogenesis. *Nucleic acids research*, **38**, 1780-1789.

115. Eirin-Lopez, J.M., Ishibashi, T. and Ausio, J. (2008) H2A.Bbd: a quickly evolving hypervariable mammalian histone that destabilizes nucleosomes in an acetylation-independent way. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, **22**, 316-326.

116. Govin, J., Escoffier, E., Rousseaux, S., Kuhn, L., Ferro, M., Thevenon, J., Catena, R., Davidson, I., Garin, J., Khochbin, S. *et al.* (2007) Pericentric heterochromatin reprogramming by new histone variants during mouse spermiogenesis. *The Journal of cell biology*, **176**, 283-294.

117. Soboleva, T.A., Nekrasov, M., Pahwa, A., Williams, R., Huttley, G.A. and Tremethick, D.J. (2012) A unique H2A histone variant occupies the transcriptional start site of active genes. *Nature structural & molecular biology*, **19**, 25-30.

118. 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. *The EMBO journal*, **23**, 3815-3824.

119. 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. *The EMBO journal*, **23**, 3314-3324.

120. Shukla, M.S., Syed, S.H., Goutte-Gattat, D., Richard, J.L., Montel, F., Hamiche, A., Travers, A., Faivre-Moskalenko, C., Bednar, J., Hayes, J.J. *et al.* (2011) The docking domain of histone H2A is required for H1 binding and RSC-mediated nucleosome remodeling. *Nucleic acids research*, **39**, 2559-2570.

121. Witt, O., Albig, W. and Doenecke, D. (1996) Testis-specific expression of a novel human H3 histone gene. *Experimental cell research*, **229**, 301-306.

122. 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. *The Journal of biological chemistry*, **277**, 43474-43480.

123.    Pehrson, J.R. and Fried, V.A. (1992) MacroH2A, a core histone containing a large nonhistone region. *Science*, **257**, 1398-1400.

124.    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.

125.    Han, W., Li, X. and Fu, X. (2011) The macro domain protein family: structure, functions, and their potential therapeutic implications. *Mutation research*, **727**, 86-103.

126.    Kustatscher, G., Hothorn, M., Pugieux, C., Scheffzek, K. and Ladurner, A.G. (2005) Splicing regulates NAD metabolite binding to histone macroH2A. *Nature structural & molecular biology*, **12**, 624-625.

127.    Timinszky, G., Till, S., Hassa, P.O., Hothorn, M., Kustatscher, G., Nijmeijer, B., Colombelli, J., Altmeyer, M., Stelzer, E.H., Scheffzek, K. *et al.* (2009) A macrodomain-containing histone rearranges chromatin upon sensing PARP1 activation. *Nature structural & molecular biology*, **16**, 923-929.

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

129.    Lee, J.T. (2011) Gracefully ageing at 50, X-chromosome inactivation becomes a paradigm for RNA and chromatin control. *Nature reviews. Molecular cell biology*, **12**, 815-826.

130.    Pasque, V., Gillich, A., Garrett, N. and Gurdon, J.B. (2011) Histone variant macroH2A confers resistance to nuclear reprogramming. *The EMBO journal*, **30**, 2373-2387.

131.    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.

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

133.    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.

134.    Gamble, M.J. and Kraus, W.L. (2010) Multiple facets of the unique histone variant macroH2A: from genomics to cell biology. *Cell cycle*, **9**, 2568-2574.

135.    Buschbeck, M., Uribesalgo, I., Wibowo, I., Rue, P., Martin, D., Gutierrez, A., Morey, L., Guigo, R., Lopez-Schier, H. and Di Croce, L. (2009) The histone variant macroH2A is an epigenetic regulator of key developmental genes. *Nature structural & molecular biology*, **16**, 1074-1079.

136.    Gamble, M.J., Frizzell, K.M., Yang, C., Krishnakumar, R. and Kraus, W.L. (2010) The histone variant macroH2A1 marks repressed autosomal chromatin, but protects a subset of its target genes from silencing. *Genes & development*, **24**, 21-32.

137.    Creppe, C., Janich, P., Cantarino, N., Noguera, M., Valero, V., Musulen, E., Douet, J., Posavec, M., Martin-Caballero, J., Sumoy, L. *et al.* (2012) MacroH2A1 Regulates the Balance between Self-Renewal and Differentiation Commitment in Embryonic and Adult Stem Cells. *Molecular and cellular biology*, **32**, 1442-1452.

138.    Chakravarthy, S., Gundimella, S.K., Caron, C., Perche, P.Y., Pehrson, J.R., Khochbin, S. and Luger, K. (2005) Structural characterization of the histone variant macroH2A. *Molecular and cellular biology*, **25**, 7616-7624.

139.    Chakravarthy, S. and Luger, K. (2006) The histone variant macro-H2A preferentially forms "hybrid nucleosomes". *The Journal of biological chemistry*, **281**, 25522-25531.

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

141. Ratnakumar, K., Duarte, L.F., LeRoy, G., Hasson, D., Smeets, D., Vardabasso, C., Bonisch, C., Zeng, T., Xiang, B., Zhang, D.Y. *et al.* (2012) ATRX-mediated chromatin association of histone variant macroH2A1 regulates alpha-globin expression. *Genes & development*, **26**, 433-438.

142. 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.

143. 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. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 14075-14080.

144. Drane, P., Ouararhni, K., Depaux, A., Shuaib, M. and Hamiche, A. (2010) The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. *Genes & development*, **24**, 1253-1265.

145. Angelov, D., Molla, A., Perche, P.Y., Hans, F., Cote, J., Khochbin, S., Bouvet, P. and Dimitrov, S. (2003) The histone variant macroH2A interferes with transcription factor binding and SWI/SNF nucleosome remodeling. *Molecular cell*, **11**, 1033-1041.

146. Chang, E.Y., Ferreira, H., Somers, J., Nusinow, D.A., Owen-Hughes, T. and Narlikar, G.J. (2008) MacroH2A allows ATP-dependent chromatin remodeling by SWI/SNF and ACF complexes but specifically reduces recruitment of SWI/SNF. *Biochemistry*, **47**, 13726-13732.

147. Doyen, C.M., An, W., Angelov, D., Bondarenko, V., Mietton, F., Studitsky, V.M., Hamiche, A., Roeder, R.G., Bouvet, P. and Dimitrov, S. (2006) Mechanism of polymerase II transcription repression by the histone variant macroH2A. *Molecular and cellular biology*, **26**, 1156-1164.

148. Biswas, M., Voltz, K., Smith, J.C. and Langowski, J. (2011) Role of histone tails in structural stability of the nucleosome. *PLoS computational biology*, **7**, e1002279.

149. Lan, L., Nakajima, S., Kapetanaki, M.G., Hsieh, C.L., Fagerburg, M., Thickman, K., Rodriguez-Collazo, P., Leuba, S.H., Levine, A.S. and Rapic-Otrin, V. (2012) Monoubiquitinated Histone H2A Destabilizes Photolesion-containing Nucleosomes with Concomitant Release of UV-damaged DNA-binding Protein E3 Ligase. *The Journal of biological chemistry*, **287**, 12036-12049.

150. Usachenko, S.I., Bavykin, S.G., Gavin, I.M. and Bradbury, E.M. (1994) Rearrangement of the histone H2A C-terminal domain in the nucleosome. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 6845-6849.

151. Lee, K.M. and Hayes, J.J. (1998) Linker DNA and H1-dependent reorganization of histone-DNA interactions within the nucleosome. *Biochemistry*, **37**, 8622-8628.

152. Bonner, W.M. and Stedman, J.D. (1979) Histone 1 is proximal to histone 2A and to A24. *Proceedings of the National Academy of Sciences of the United States of America*, **76**, 2190-2194.

153. Boulikas, T., Wiseman, J.M. and Garrard, W.T. (1980) Points of contact between histone H1 and the histone octamer. *Proceedings of the National Academy of Sciences of the United States of America*, **77**, 127-131.

154. 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 genetics*, **6**, e1001234.

155. Abbott, D.W., Chadwick, B.P., Thambirajah, A.A. and Ausio, J. (2005) Beyond the Xi: macroH2A chromatin distribution and post-translational modification in an avian system. *The Journal of biological chemistry*, **280**, 16437-16445.

156. Muthurajan, U.M., McBryant, S.J., Lu, X., Hansen, J.C. and Luger, K. (2011) The linker region of macroH2A promotes self-association of nucleosomal arrays. *The Journal of biological chemistry*, **286**, 23852-23864.

157. 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. *The Journal of biological chemistry*, **263**, 18972-18978.

158. Thomas, J.O. and Kornberg, R.D. (1975) An octamer of histones in chromatin and free in solution. *Proceedings of the National Academy of Sciences of the United States of America*, **72**, 2626-2630.

159. Wood, C.M., Nicholson, J.M., Lambert, S.J., Chantalat, L., Reynolds, C.D. and Baldwin, J.P. (2005) High-resolution structure of the native histone octamer. *Acta crystallographica. Section F, Structural biology and crystallization communications*, **61**, 541-545.

160. Davey, C.A., Sargent, D.F., Luger, K., Maeder, A.W. and Richmond, T.J. (2002) Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. *Journal of molecular biology*, **319**, 1097-1113.

161. Ferreira, H., Somers, J., Webster, R., Flaus, A. and Owen-Hughes, T. (2007) Histone tails and the H3 alphaN helix regulate nucleosome mobility and stability. *Molecular and cellular biology*, **27**, 4037-4048.

162. Pineiro, M., Puerta, C. and Palacian, E. (1991) Yeast nucleosomal particles: structural and transcriptional properties. *Biochemistry*, **30**, 5805-5810.

163. White, C.L., Suto, R.K. and Luger, K. (2001) Structure of the yeast nucleosome core particle reveals fundamental changes in internucleosome interactions. *The EMBO journal*, **20**, 5207-5218.

164. Li, W., Nagaraja, S., Delcuve, G.P., Hendzel, M.J. and Davie, J.R. (1993) Effects of histone acetylation, ubiquitination and variants on nucleosome stability. *The Biochemical journal*, **296 ( Pt 3)**, 737-744.

165. 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. *The Journal of biological chemistry*, **279**, 24274-24282.

166. Abbott, D.W., Ivanova, V.S., Wang, X., Bonner, W.M. and Ausio, J. (2001) Characterization of the stability and folding of H2A.Z chromatin particles: implications for transcriptional activation. *The Journal of biological chemistry*, **276**, 41945-41949.

167. Hoch, D.A., Stratton, J.J. and Gloss, L.M. (2007) Protein-protein Forster resonance energy transfer analysis of nucleosome core particles containing H2A and H2A.Z. *Journal of molecular biology*, **371**, 971-988.

168. Higashi, T., Matsunaga, S., Isobe, K., Morimoto, A., Shimada, T., Kataoka, S., Watanabe, W., Uchiyama, S., Itoh, K. and Fukui, K. (2007) Histone H2A mobility is regulated by its tails and acetylation of core histone tails. *Biochemical and biophysical research communications*, **357**, 627-632.

169. Jin, C. and Felsenfeld, G. (2007) Nucleosome stability mediated by histone variants H3.3 and H2A.Z. *Genes & development*, **21**, 1519-1529.

170. Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K. and Felsenfeld, G. (2009) H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nature genetics*, **41**, 941-945.

171.    Wang, A.Y., Aristizabal, M.J., Ryan, C., Krogan, N.J. and Kobor, M.S. (2011) Key functional regions in the histone variant H2A.Z C-terminal docking domain. *Molecular and cellular biology*, **31**, 3871-3884.

172.    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. *The EMBO journal*, **25**, 4234-4244.

173.    Okuwaki, M., Kato, K., Shimahara, H., Tate, S. and Nagata, K. (2005) Assembly and disassembly of nucleosome core particles containing histone variants by human nucleosome assembly protein I. *Molecular and cellular biology*, **25**, 10639-10651.

174.    Gautier, T., Abbott, D.W., Molla, A., Verdel, A., Ausio, J. and Dimitrov, S. (2004) Histone variant H2ABbd confers lower stability to the nucleosome. *EMBO reports*, **5**, 715-720.

175.    Syed, S.H., Boulard, M., Shukla, M.S., Gautier, T., Travers, A., Bednar, J., Faivre-Moskalenko, C., Dimitrov, S. and Angelov, D. (2009) The incorporation of the novel histone variant H2AL2 confers unusual structural and functional properties of the nucleosome. *Nucleic acids research*, **37**, 4684-4695.

176.    Li, G. and Reinberg, D. (2011) Chromatin higher-order structures and gene regulation. *Current opinion in genetics & development*, **21**, 175-186.

177.    Hansen, J.C. (2002) Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions. *Annual review of biophysics and biomolecular structure*, **31**, 361-392.

178.    Schlick, T., Hayes, J. and Grigoryev, S. (2012) Toward convergence of experimental studies and theoretical modeling of the chromatin fiber. *The Journal of biological chemistry*, **287**, 5183-5191.

179.    McBryant, S.J., Adams, V.H. and Hansen, J.C. (2006) Chromatin architectural proteins. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology*, **14**, 39-51.

180.    Kelley, R.I. (1973) Isolation of a histone IIb1-IIb2 complex. *Biochemical and biophysical research communications*, **54**, 1588-1594.

181.    Wyrick, J.J., Kyriss, M.N. and Davis, W.B. (2012) Ascending the nucleosome face: Recognition and function of structured domains in the histone H2A-H2B dimer. *Biochimica et biophysica acta*.

182.    Roussel, L., Erard, M., Cayrol, C. and Girard, J.P. (2008) Molecular mimicry between IL-33 and KSHV for attachment to chromatin through the H2A-H2B acidic pocket. *EMBO reports*, **9**, 1006-1012.

183.    Dorigo, B., Schalch, T., Bystricky, K. and Richmond, T.J. (2003) Chromatin fiber folding: requirement for the histone H4 N-terminal tail. *Journal of molecular biology*, **327**, 85-96.

184.    Dorigo, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R.R. and Richmond, T.J. (2004) Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science*, **306**, 1571-1573.

185.    Sinha, D. and Shogren-Knaak, M.A. (2010) Role of direct interactions between the histone H4 Tail and the H2A core in long range nucleosome contacts. *The Journal of biological chemistry*, **285**, 16572-16581.

186.    Schwarz, P.M., Felthauser, A., Fletcher, T.M. and Hansen, J.C. (1996) Reversible oligonucleosome self-association: dependence on divalent cations and core histone tail domains. *Biochemistry*, **35**, 4009-4015.

187.    Kan, P.Y., Caterino, T.L. and Hayes, J.J. (2009) The H4 tail domain participates in intra- and internucleosome interactions with protein and DNA during folding and oligomerization of nucleosome arrays. *Molecular and cellular biology*, **29**, 538-546.

188. Peterson, C.L. and Laniel, M.A. (2004) Histones and histone modifications. *Current biology : CB*, **14**, R546-551.

189. Shogren-Knaak, M., Ishii, H., Sun, J.M., Pazin, M.J., Davie, J.R. and Peterson, C.L. (2006) Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science*, **311**, 844-847.

190. Lu, X., Simon, M.D., Chodaparambil, J.V., Hansen, J.C., Shokat, K.M. and Luger, K. (2008) The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure. *Nature structural & molecular biology*, **15**, 1122-1124.

191. Fan, J.Y., Rangasamy, D., Luger, K. and Tremethick, D.J. (2004) H2A.Z alters the nucleosome surface to promote HP1alpha-mediated chromatin fiber folding. *Molecular cell*, **16**, 655-661.

192. Greaves, I.K., Rangasamy, D., Ridgway, P. and Tremethick, D.J. (2007) H2A.Z contributes to the unique 3D structure of the centromere. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 525-530.

193. Zhou, J., Fan, J.Y., Rangasamy, D. and Tremethick, D.J. (2007) The nucleosome surface regulates chromatin compaction and couples it with transcriptional repression. *Nature structural & molecular biology*, **14**, 1070-1076.

194. Caterino, T.L. and Hayes, J.J. (2011) Structure of the H1 C-terminal domain and function in chromatin condensation. *Biochemistry and cell biology = Biochimie et biologie cellulaire*, **89**, 35-44.

# H2A.Z.2.2 is an alternatively spliced histone H2A.Z variant that causes severe nucleosome destabilization

**Clemens Bönisch**, Katrin Schneider, Sebastian Pünzeler, Sonja M. Wiedemann, Christina Bielmeier, Marco Bocola, H. Christian Eberl, Wolfgang Kuegel, Jürgen Neumann, Elisabeth Kremmer, Heinrich Leonhardt, Matthias Mann, Jens Michaelis, Lothar Schermelleh and Sandra B. Hake

# H2A.Z.2.2 is an alternatively spliced histone H2A.Z variant that causes severe nucleosome destabilization

Clemens Bönisch[1], Katrin Schneider[2], Sebastian Pünzeler[1], Sonja M. Wiedemann[1], Christina Bielmeier[2], Marco Bocola[3], H. Christian Eberl[4], Wolfgang Kuegel[5], Jürgen Neumann[2], Elisabeth Kremmer[6], Heinrich Leonhardt[2,7], Matthias Mann[4,7], Jens Michaelis[4,7,8], Lothar Schermelleh[2,*] and Sandra B. Hake[1,7,*]

[1]Department of Molecular Biology, Adolf-Butenandt-Institute, Ludwig-Maximilians-University Munich, 80336 Munich, [2]Department of Biology, Biozentrum, Ludwig-Maximilians-University Munich, 82152 Planegg-Martinsried, [3]Department of Biochemistry II, University Regensburg, 93053 Regensburg, [4]Department of Proteomics and Signal Transduction, Max-Planck-Institute of Biochemistry, 82152 Martinsried, [5]Department of Chemistry, Ludwig-Maximilians-University Munich, [6]Institute of Molecular Immunology, Helmholtz Center Munich, German Research Center for Environmental Health, [7]Center for Integrated Protein Science Munich (CIPSM), 81377 Munich and [8]Department of Physics, Ulm University, 89081 Ulm, Germany

## ABSTRACT

**The histone variant H2A.Z has been implicated in many biological processes, such as gene regulation and genome stability. Here, we present the identification of H2A.Z.2.2 (Z.2.2), a novel alternatively spliced variant of histone H2A.Z and provide a comprehensive characterization of its expression and chromatin incorporation properties. Z.2.2 mRNA is found in all human cell lines and tissues with highest levels in brain. We show the proper splicing and *in vivo* existence of this variant protein in humans. Furthermore, we demonstrate the binding of Z.2.2 to H2A.Z-specific TIP60 and SRCAP chaperone complexes and its active replication-independent deposition into chromatin. Strikingly, various independent *in vivo* and *in vitro* analyses, such as biochemical fractionation, comparative FRAP studies of GFP-tagged H2A variants, size exclusion chromatography and single molecule FRET, in combination with *in silico* molecular dynamics simulations, consistently demonstrate that Z.2.2 causes major structural changes and significantly destabilizes nucleosomes. Analyses of deletion mutants and chimeric proteins pinpoint this property to its unique C-terminus. Our findings enrich the list of known human variants by an unusual protein belonging to the H2A.Z family that leads to the least stable nucleosome known to date.**

## INTRODUCTION

In the eukaryotic nucleus, DNA is packaged into chromatin. The fundamental unit of this structure is the nucleosome consisting of a histone octamer (two of each H2A, H2B, H3 and H4) that organizes ~147 bp of DNA (1). In order to allow or prevent nuclear regulatory proteins access to the DNA, the chromatin structure has to be flexible and dynamic. Several mechanisms ensure controlled chromatin changes, one being the incorporation of specialized histone variants (2,3).

Variants of the histone H2A family are the most diverse in sequence and exhibit distinct functions (4,5), comprising DNA damage repair, transcriptional regulation, cell cycle control and chromatin condensation, though the exact mechanisms of action are not fully understood yet. Interestingly, the highest sequence variation among H2A variants is found in the C-terminus, suggesting that differences in structure and biological function might be

*To whom correspondence should be addressed. Tel: +49 89 2180 75435; Fax: +49 89 2180 75425; Email: sandra.hake@med.uni-muenchen.de
Correspondence may also be addressed to Lothar Schermelleh. Tel: +44 1865 613264; Fax: +49 89 2180 74236;
Email: lothar.schermelleh@bioch.ox.ac.uk
Present address:
Lothar Schermelleh, Department of Biochemistry, University of Oxford, South Park Road, Oxford OX1 3QU, UK.

primarily attributed to this domain (6–9). One of the best investigated and highly conserved but also functionally enigmatic histone variant is H2A.Z. This variant is essential in most eukaryotes and possesses unique functions (10,11). H2A.Z is involved in transcriptional regulation, chromosome segregation and mitosis, acting in an organism- and differentiation-dependent manner (12,13). Furthermore, H2A.Z has been implicated in regulating epigenetic memory (14) and in inhibiting read-through antisense transcription (15). In higher eukaryotes, H2A.Z might play a role in heterochromatin organization (16), genome stability and chromosome segregation (17). Despite many efforts to elucidate the exact biological functions of H2A.Z, its roles have been and remain controversial (18). Furthermore, deregulation of H2A.Z expression or localization seems to be connected to the development of several neoplasias (19–23). Interestingly, in vertebrates two non-allelic genes coding for two highly similar H2A.Z proteins, H2A.Z.1 and H2A.Z.2, exist (24) (previously named H2A.Z-1 and H2A.Z-2, prefixes were changed due to a new histone variant nomenclature; Talbert P.B., manuscript in preparation). They have a common origin in early chordate evolution, are both acetylated on the same N-terminal lysines (25–27) and might be ubiquitinated on either one of the two C-terminal lysines (28).

Here, we report the identification and structural characterization of H2A.Z.2.2 (Z.2.2), an unusual alternative splice form of H2A.Z. We show that Z.2.2 mRNA is expressed to different degrees in all human cell lines and tissues examined, with highest levels found in brain. Cell biological and biochemical analyses consistently reveal the presence of two distinct Z.2.2 populations within the cell. The majority of Z.2.2 is freely dispersed in the nucleus, whereas only a minority is stably incorporated into chromatin, most likely through the H2A.Z-specific p400/NuA4/TIP60 (TIP60) and SRCAP chaperone complexes. *In vivo* and *in vitro* analyses, in agreement with molecular dynamic (MD) simulations, demonstrate that due to its unique docking domain Z.2.2 chromatin incorporation leads to severely unstable nucleosomes. Our data provide compelling evidence that a novel H2A.Z variant exists in humans that plays a distinct and novel role in chromatin structure regulation.

## MATERIALS AND METHODS

See Supplementary Materials and Methods section for detailed protocols.

### Cell culture, transfection, FACS and cloning

Cell lines were grown in DMEM medium (PAA) supplemented with 10% FCS (Sigma) and 1% penicillin/streptomycin at 37°C and 5% $CO_2$. Cells were transfected using FuGene HD (Roche Applied Science) according to the manufacturer's instructions. For details on cell selection, FACS and cloning of expression plasmids see Supplementary Materials and Methods section.

### RNA expression analysis

RNA isolation and cDNA generation were performed as previously described (29). Data were analyzed with the advanced relative quantification tool of the Lightcycler 480 (Roche) software including normalization to HPRT1 and HMBS levels. Statistical evaluation was done using *t*-test (two-tailed distribution, heteroscedastic). Total RNA from different human tissues was commercially acquired from: Applied Biosystems: normal lung, breast and tumor breast, lung and ovary; Biochain: tumor lung, breast, thyroid and bone, normal testis, cerebellum, cerebral cortex, hippocampus, thalamus and total fetal brain; amsbio: frontal lobe.

### Histone extraction, RP–HPLC purification, sucrose gradient, cellular fractionation and salt stability experiments

Acid extraction of histones was done as previously described (30). Histones were separated by RP–HPLC as previously described (29). Fractions were dried under vacuum and stored at −20°C.

Details on MNase digest and sucrose gradient fractionation can be found in Supplementary Materials and Methods section.

Fractionation and salt stability experiments were carried out as described previously (31–33) with minor changes. For details on these methods see Supplementary Materials and Methods section.

### Antibodies

For the generation of a Z.2.2-specific antibody (αZ.2.2), a peptide spanning the last C-terminal amino acids GGEKRRCS of Z.2.2 was synthesized (Peptide Specialty Laboratories GmbH) and coupled to BSA and OVA, respectively. Development of Z.2.2-specific monoclonal antibodies in rats was done as previously described (29). The αZ.2.2 clone 1H11-11 of rat IgG1 subclass was applied in this study. Rabbit αZ.2.2 antibody (rabbit 2, bleed 3) was generated by the Pineda-Antikörper-Service company using the identical peptide epitope followed by affinity purification. Following other primary antibodies were used: αGAPDH (sc-25778, Santa Cruz), αGFP (Roche Applied Science), αH2A (ab 13923, abcam), αH3 (ab1791, abcam) and αH2A.Z (C-terminus: ab4174, abcam; N-terminus: ab18263, abcam). Following secondary antibodies and detection kits were used in immunoblots: GFP-Z.2.2 and GFP-Bbd histones (αGFP) and endogenous Z.2.2 (αZ.2.2) were detected using HRP-conjugated secondary antibodies (Amersham) with ECL advance (Amersham), all other proteins were detected using ECL (Amersham). Detection of recombinant proteins to evaluate histone stoichiometry of *in vitro* assembled nucleosomes was carried out using IRDye-labeled secondary antibodies (LI-COR).

### Fluorescence microscopy of cells and chromosomes

Preparation of cells and chromosome spreads for fluorescence microscopy was done as previously reported (34). Wide-field fluorescence imaging was performed on

a PersonalDV microscope system (Applied Precision) equipped with a 60×/1.42 PlanApo oil objective (Olympus), CoolSNAP ES2 interline CCD camera (Photometrics), Xenon illumination and appropriate filtersets. Iterative 3D deconvolution of image *z*-stacks was performed with the SoftWoRx 3.7 imaging software package (Applied Precision).

### FRAP and exponential fitting

For details see Supplementary Materials and Methods section.

### Stable isotope labeling with amino acids in cell culture (SILAC) and mass spectrometric identification of H2A.Z-specific chaperone complexes

HeLa cells expressing GFP-Z.2.1 or GFP-Z.2.2 were SILAC labeled and nuclear extracts were prepared as described before (35,36). High-resolution LC MS/MS analysis was performed on an Orbitrap platform: details on the experimental procedure are found in Supplementary Materials and Methods section. Mass spectrometric (MS) operation and raw data analysis (37) are described in Supplementary Materials and Methods section. A complete list of all proteins identified is found in Supplementary Table S1.

### Immunofluorescence microscopy of cell cycle-dependent GFP-Z.2.1 and GFP-Z.2.2 chromatin incorporation

Details on the experimental labeling (38) and microscopy procedures are found in Supplementary Materials and Methods section.

### Expression of recombinant human histone proteins in *Escherichia coli*, *in vitro* octamer and nucleosome reconstitution

Histones were expressed, purified and assembled into octamers as described (39) and mononucleosomes were assembled on DNA containing the 601-positioning sequence (40) according to (39,41). For details on *in vitro* octamer and nucleosome reconstitution, see Supplementary Materials and Methods section.

### Single molecule Förster resonance energy transfer

Single molecule Förster resonance energy transfer (smFRET) single molecule burst analysis followed by the removal of multi-molecular events (42–45) are described in detail in the Supplementary Materials and Methods section.

### Molecular modeling and MD simulations

The molecular modeling suite YASARA-structure version 9.10.29 was employed, utilizing the AMBER03 force field (46) for the protein and the general amber force field (GAFF) (47) throughout this study. The partial charges were computed using the AM1/BCC procedure (48) as implemented in YASARA structure (49). The starting point for molecular modeling was the crystal structure of a nucleosome core particle containing the histone variant H2A.Z (PDB 1F66) (50). Missing side chain atoms were added (Glu E 634). The missing N-terminal and C-terminal residues were not modeled, although they might interact with the neighboring DNA, e.g. in the case of missing C-terminal residues in H2A.Z (119–128; GKKGQQKTV). All structures were solvated in a water box with 0.9% NaCl and neutralized (51). The structures were initially minimized using steepest descent and simulating annealing procedures. All deletions and mutations were introduced sequentially using YASARA structure. MD simulations were carried out at 300 K over 2.5 ns in an NPT ensemble using PME. All simulations were performed four times using various starting geometries. The 2.5 ns MD trajectories were sampled every 25 ps, resulting in 100 simulation frames per run, which were evaluated after an equilibration phase of 500 ps to derive statistical averages and properties of the corresponding variant. Finally, the interaction energy of H2A and H3 was calculated from a simulation of the solvated octamer and the isolated $(H3–H4)_2$ tetramer or the isolated respective H2A.Z–H2B dimer. The interaction energy is calculated as energy difference of the solvated octamer minus the solvated $(H3–H4)_2$ tetramer and H2A.Z–H2B dimer.

## RESULTS

### Alternative splicing of H2A.Z.2 occurs *in vivo*

Two non-allelic intron-containing genes with divergent promoter sequences that code for H2A.Z variants exist in vertebrates (24,27). In humans, the H2A.Z.2 (H2AFV) primary transcript is predicted to be alternatively spliced thereby generating five different gene products (Supplementary Figure S1A). Using PCR and confirmed by sequencing we detected not only H2A.Z.2.1 (Z.2.1) but also H2A.Z.2.2 (Z.2.2) mRNA, though none of the other splice variants in human cells (Supplementary Figure S1B) showing that the H2A.Z.2 primary transcript is indeed alternatively spliced *in vivo*. Interestingly, database searches found Z.2.2 mRNA to be predicted in chimpanzee (*Pan troglodytes*) and Northern white-cheeked gibbon (*Nomascus leucogenys*) as well. In addition, the coding sequence of the unique exon 6 was present downstream of the *H2AFV* locus of several other primate genomes, such as gorilla (*Gorilla gorilla gorilla*), macaque (*Macaca mulatta*), orangutan (*Pongo abelii*) and white-tufted-ear marmoset (*Callithrix jacchus*) (data not shown). In all of these primates, with the exception of marmoset, the resulting protein sequence, if translated, is 100% identical to the unique human Z.2.2 peptide. Further searches revealed that the genomes of horse, and to a certain extent also rabbit and panda bear, contain sequences downstream of their *H2AFV* loci that could, if translated, lead to proteins with some similarities to human Z.2.2, although they are much more divergent and even longer (rabbit, panda bear). Due to these differences, it is highly likely that those species do not express a Z.2.2 protein homolog. Surprisingly, we could not detect Z.2.2-specific sequences in mouse, rat or other eukaryotic genomes, suggesting that Z.2.2 might be primate specific.

Next, we wanted to determine to what degree all three H2A.Z mRNAs are expressed in different human cell lines and tissues and performed quantitative PCR (qPCR).

Z.2.2 mRNA was present to different degrees in all human cell lines and tissues tested, though less abundant than Z.1 and Z.2.1 mRNAs that are expressed in similar amounts (Supplementary Figure S1C and D). Z.2.2 constituted between 5% and 15% of total Z.2 transcripts in all cell lines and tissues, with the exception of brain, where it was statistically significant upregulated ($p = 1.7 \times 10^{-4}$; Figure 1A). In some regions of this particular organ Z.2.2 accounted for up to 50% of all Z.2 transcripts pointing toward an exciting brain-specific function of this novel variant.

Encouraged by our findings we next investigated whether the endogenous protein is present *in vivo*. The distinctive feature of Z.2.2 is its C-terminus that is 14 amino acids shorter and contains six amino acids differences compared to Z.2.1 (Figure 1B). Due to this shortened C-terminal sequence, ubiquitination sites at positions K120 and K121 (28) and part of the H3/H4 docking domain (50) are lost in Z.2.2. We generated antibodies against Z.2.2's unique C-terminal amino acids (αZ.2.2) in rats and rabbits and confirmed their specificity in immunoblots (IB) with recombinant Z.2.1 and Z.2.2

proteins (Supplementary Figure S1E and data not shown). We extracted histones from several human and mouse cell lines, purified them by reversed phase–high performance liquid chromatography (RP–HPLC) and analyzed obtained fractions by IB (Figure 1C). Using αZ.2.2 (polyclonal rabbit), we observed a signal of the calculated weight of Z.2.2 that elutes shortly before Z.1- and Z.2.1-containing fractions in all human samples. Similar results were obtained with a monoclonal αZ.2.2 rat antibody (data not shown). In agreement with the finding that Z.2.2-specific exon 6 sequences are mainly restricted to primate genomes, we could detect Z.2.2 protein in human but not in mouse cells (Figure 1C). In summary, our data show that Z.2.2 protein indeed exists *in vivo*, albeit at a low expression level.

### GFP-Z.2.2 is partially incorporated into chromatin

Having demonstrated the existence of this novel variant *in vivo*, we next sought to clarify whether Z.2.2 constitutes a bona fide histone by being part of the chromatin structure. Due to high background of all our αZ.2.2 antibodies in IB with cell extracts (data not shown), we generated
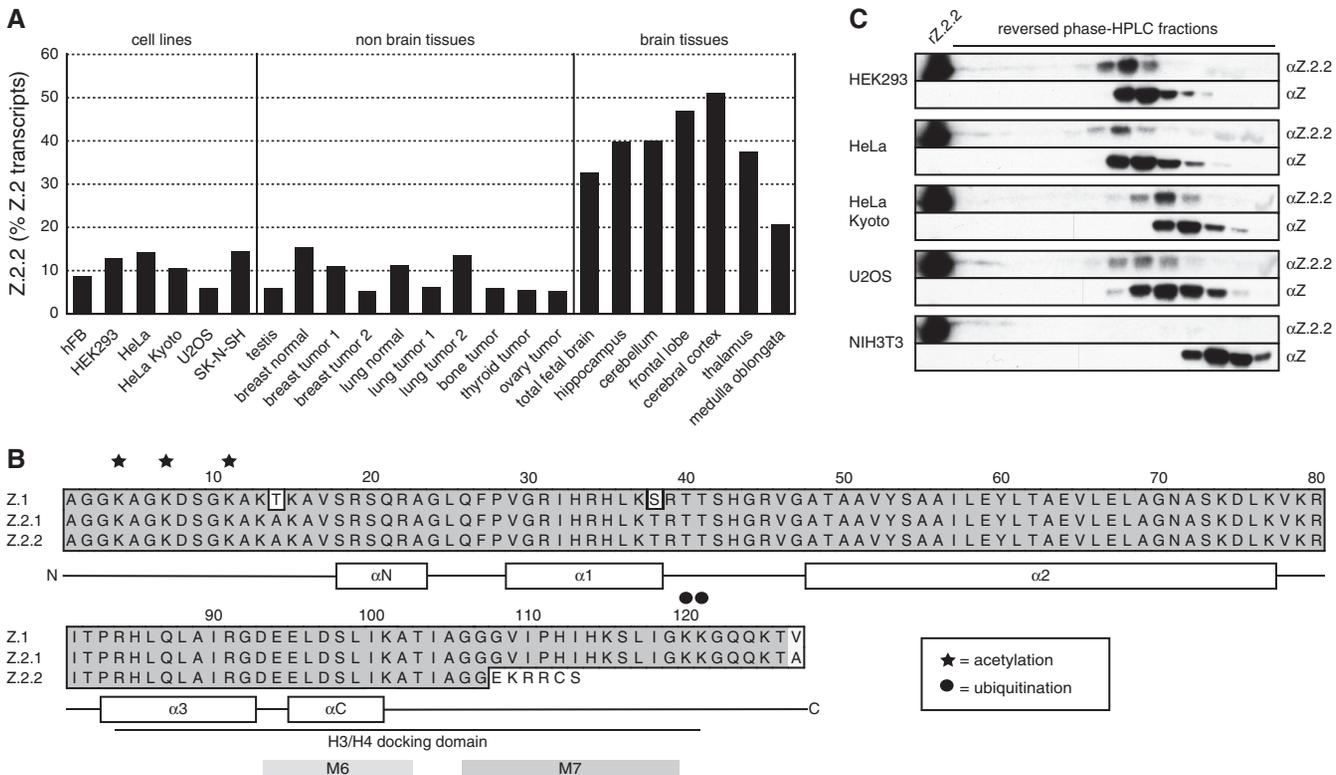


**Figure 1.** Identification of Z.2.2. (**A**) qPCR with cDNA from different human cell lines and tissues using primers specific for Z.2.1 and Z.2.2. Data were normalized to HPRT1 and HMBS expression levels. Controls generated without reverse transcriptase (no RT) were used to assess amplification threshold. Shown are the levels of Z.2.2 mRNA as percentages of total Z.2 transcripts (Z.2.1 + Z.2.2). For an evaluation of absolute expression levels see Supplementary Figure S1C and D. (**B**) Amino acid alignment of human Z.1, Z.2.1 and Z.2.2 proteins using ClustalW Alignment (MacVector 10.0.2). Identical amino acids are highlighted in dark gray, similar amino acids in light gray and changes are set apart on white background. Known acetylation sites are depicted with stars and ubiquitination sites with circles. A schematic representation of the secondary structure of Z.1 and Z.2.1 is shown below the alignment, including depiction of the H3/H4 docking domain (50). M6 and M7 boxes indicate regions important for H2A.Z-specific biological functions in *D. melanogaster* (60). (**C**) IB analyses of RP–HPLC purified fractions from different human (HEK293, HeLa, HeLa Kyoto and U2OS) and mouse (NIH3T3) cell lines using a polyclonal rabbit αZ.2.2 and αH2A.Z (αZ, C-terminal) antibodies. Recombinant Z.2.2 protein (rZ.2.2) was loaded in the first lane as positive control for αZ.2.2 antibody. Similar results were obtained when using a monoclonal rat αZ.2.2 antibody (data not shown).

HeLa Kyoto cell lines stably expressing GFP-tagged H2A variants (HK-GFP cells) for subsequent analyses. Expression levels of GFP-tagged histone variants were determined by FACS (Supplementary Figure S2A) and by comparing expression levels of GFP-tagged variants with endogenous H2A.Z proteins in IB (Supplementary Figure S2B). GFP-Z.1 and -Z.2.1 were expressed in similar amounts as the endogenous H2A.Z protein, and GFP-Z.2.2 expression levels were considerably lower than those of other GFP-tagged H2A variants, with the exception of GFP-H2A.Bbd (Barr body deficient; Bbd). These data show that all GFP-H2A variants were not expressed in abnormal amounts in cell clones used for further analyses.

In fluorescence microscopy, GFP-Z.2.2 exhibited a sole but rather diffuse nuclear distribution similar to GFP-Bbd, suggesting that both variants might have similar properties (Figure 2A). Additionally, GFP-Z.2.2 was detected in condensed mitotic chromosomes, with a faint residual staining in the surrounding area (Figure 2B), suggesting that it is incorporated into chromatin, although to a lesser extent than other GFP-H2A variants. To discriminate between a potential non-specific DNA binding and nucleosomal incorporation of Z.2.2 we purified mononucleosomes by sucrose gradient centrifugation. GFP-Z.2.2 was detected by IB in fractions containing mononucleosomes (Figure 2C), suggesting that Z.2.2 is indeed a nucleosomal constituent.

To analyze the extent of Z.2.2 chromatin incorporation in more detail, we isolated soluble (*sol*) and chromatin (*chr*) fractions from HK-GFP cells. IB analyses revealed, as expected, that similar to GFP-Bbd, GFP-Z.2.2 is predominantly nuclear soluble, with only minor amounts present in chromatin (Figure 3A). Based on fractionation and fluorescence imaging results, we hypothesized that this novel variant behaves in a different manner as compared to other H2A variants with regard to chromatin exchange mobility *in vivo*. To test this prediction, we performed fluorescence recovery after photobleaching (FRAP) experiments with HK-GFP cells. Using spinning disk confocal microscopy we monitored the kinetic behavior of H2A variants with variable intervals over 2 min (short-term) up to several hours (long-term) after bleaching a $5\,\mu m \times 5\,\mu m$ square nuclear region (Figure 3B and Supplementary Figure S3). As expected, GFP alone showed the highest mobility. In contrast, GFP-H2A, -Z.1 and -Z.2.1 showed a slow recovery, which is in agreement with a previous report (52). GFP-Bbd has been described to exhibit low nucleosomal stability and a fast FRAP kinetic (53), which we also observed in our experiment. Interestingly, GFP-Z.2.2 showed an even faster recovery than GFP-Bbd, with $\sim$80% of initial fluorescence reached after 1 min. Careful assessment and bi-exponential fitting of FRAP data allowed us to also calculate ratios of fractions with fast, intermediate and slow recovery and their respective half-time of recovery ($t_{1/2}$) as an indication of exchange rate thereby revealing quantitative differences between Z.2.2 and other H2A variants (Figure 3D, Supplementary Figure S3C and E). For Z.2.2 as well as for Bbd, we identified a fast fraction of unbound or very transiently
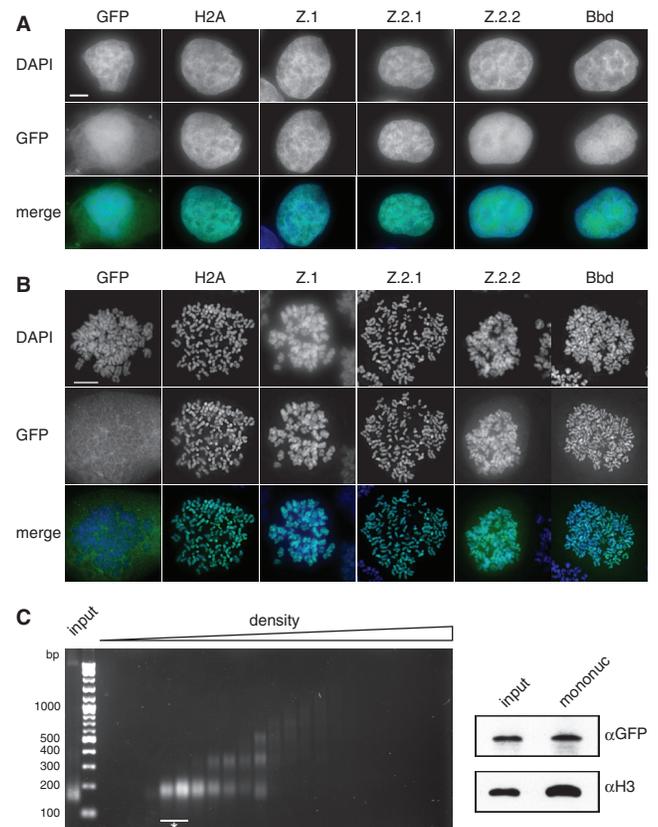


**Figure 2.** Z.2.2 localizes to the nucleus and is partially incorporated into chromatin. (**A**) Fluorescence imaging of stably transfected HeLa Kyoto cells shows nuclear localization of all GFP-H2A variants (middle). DNA was counterstained with DAPI (top). Overlay of both channels in color is shown at the bottom (Merge; GFP: green, DAPI: blue). Scale bar = 5 μm. (**B**) Deconvolved images of metaphase spreads of HeLa Kyoto cells stably expressing GFP-H2A variants (middle). Merged images in color are shown below (GFP: green; DAPI: blue). Scale bar = 10 μm. (**C**) Chromatin from HeLa Kyoto cells stably expressing GFP-Z.2.2 was digested with MNase followed by a purification of mononucleosomes using sucrose gradient centrifugation. Isolated DNA from subsequent sucrose gradient fractions was analyzed by agarose gel electrophoresis (left). Fractions containing pure mononucleosomes (marked with asterisk) were combined and analyzed by IB (right) using αGFP antibody for the presence of GFP-Z.2.2 (top), and αH3 (bottom).

interacting molecules (78%, $t_{1/2} \sim 1.1$ s and 52%, $t_{1/2} \sim 2.5$ s, respectively; for comparison GFP $t_{1/2} \sim 0.4$ s) and a substantially slower fraction with a $t_{1/2}$ in the range of 7–9 min. In contrast, GFP-H2A, -Z.1 and -Z.2.1 showed no fast mobile fraction but intermediate slow fractions with $t_{1/2}$ in the range of 8–17 min and a second even slower class exchanging with a $t_{1/2}$ of a few hours. For comparison, we measured the linker histone H1.0 (54–57) and the histone binding protein HP1α (58,59), both DNA-associated proteins, and found that HP1α shows an overall much faster recovery than all H2A variants. In contrast to Z.2.2 and Bbd, no unbound fraction of H1.0 was detected. More importantly, with regards to the bound Z.2.2 and Bbd fractions overall H1.0 showed a faster recovery, arguing against an unspecific DNA-association of Z.2.2 and Bbd. In agreement with cell biological and biochemical analyses,

these data clearly demonstrate that a large fraction of the splice variant Z.2.2 is very rapidly exchanged or chromatin unbound, and a minor population is incorporated into chromatin.

### Z.2.2's unique docking domain, but not its shortened length, weakens chromatin association

The functional importance of specific C-terminal domains of H2A.Z has previously been demonstrated by nucleosomal structure analyses (7,50) and in rescue experiments in flies (60). Since the C-terminus of Z.2.2 is shorter and has a distinct sequence when compared to Z.1 and Z.2.1, it is not clear which of these features determines Z.2.2's unusual chromatin-association.

Therefore, we generated deletion and domain-swap constructs (Supplementary Figure S3D) for FRAP experiments (short-term: Figure 3C and long-term: Supplementary Figure S3B). Surprisingly, C-terminal deletions of GFP-H2A ($H2A^{111}$) and GFP-Z.2.1 ($Z.2.1^{113}$) to mimic the shortened length of Z.2.2 did not affect their original mobility in short-term and only modestly in long-term FRAP. Hence, the mere shortening of the C-terminus is not sufficient to weaken stable chromatin association.

To investigate whether the unique six C-terminal amino acids of Z.2.2 are sufficient to generate highly mobile proteins, we created a further C-terminally truncated GFP-H2A construct ($H2A^{105}$) and added the Z.2.2 specific C-terminal six amino acids ($H2A^{105}+CZ.2.2$). Although both mutant constructs are slightly more mobile than $H2A^{111}$, their indistinguishable recovery kinetics demonstrate that the unique six C-terminal amino acids of Z.2.2 alone are not sufficient to cause its extreme mobility *in vivo*.

To explore whether the complete Z.2.2 docking domain is able to induce high-protein mobility, we transferred the respective domain of either Z.2.1 (amino acids 91–127) or Z.2.2 (amino acids 91–113) onto a C-terminally truncated H2A ($H2A^{88}+CZ.2.1$ and $H2A^{88}+CZ.2.2$, respectively). Interestingly, only the docking domain of Z.2.2, but not the one of Z.2.1, confers high mobility. In conclusion, the six unique C-terminal amino acids of Z.2.2 prevent chromatin-association of a large proportion of this protein, but only when present in the context of the preceding H2A.Z-specific docking domain sequence.

### Z.2.2 interacts with H2A.Z-specific TIP60 and SRCAP chaperone complexes and is deposited into chromatin outside of S-phase

Our so far obtained data strongly imply that at least a minor amount of the cellular Z.2.2 protein is incorporated into nucleosomes. Since previous studies have shown that evolutionary conserved Swr1-related ATP-dependent chromatin remodelers specifically exchange canonical H2A–H2B with H2A.Z–H2B dimers within nucleosomes (10,61), we wondered if such complexes are also able to actively deposit Z.2.2 into chromatin. HK cells and HK cells stably expressing GFP-Z.2.1 or -Z.2.2 were SILAC labeled, soluble nuclear proteins isolated, GFP-tagged
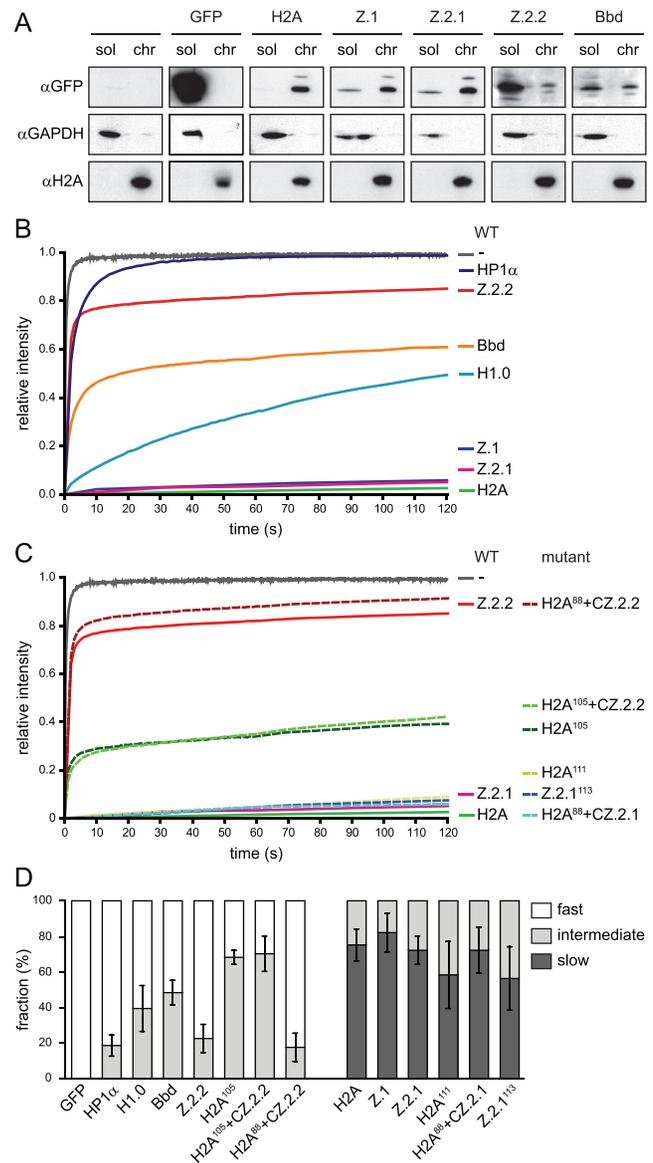


**Figure 3.** The majority of Z.2.2 protein is nuclear soluble and highly mobile in a sequence-dependent manner. (**A**) HK-GFP cells were subjected to biochemical fractionation. Fractions *sol* and *chr* of identical cell equivalents were probed in IB with αGFP (top), αH2A (middle) and αGAPDH (bottom). (**B**) FRAP quantification curves of average GFP signal relative to fluorescence intensity prior to bleaching are depicted for GFP, GFP-tagged wild-type H2A variants, linker histone H1.0 and heterochromatin protein 1α (HP1α). Mean curves of 10–29 cells are shown for each construct. Error bars are omitted for clarity. (**C**) FRAP quantification curves similar to (**B**) are depicted for GFP, GFP-tagged wild-type H2A, Z.2.1, Z.2.2 and mutant constructs. (**D**) Quantitative evaluation of FRAP curves. Plot shows calculated mobility fraction sizes of different wild-type and mutant H2A variant constructs, as well as H1.0 and HP1α, based on bi-exponential fitting of FRAP data. Error bars indicate SD (see Supplementary Figure S3 for long-term FRAP and for numerical values).

Z.2.1 and Z.2.2-associated proteins precipitated using GFP nanotrap beads and identified by quantitative mass spectrometry (Figure 4 and Supplementary Table S1 for a complete list of all identified proteins). Whereas the majority of proteins are background binders clustering

**A**  Z.2.1

**B**  Z.2.2

**C**

| protein name | ratio H/L Z.2.1 | ratio L/H Z.2.2 (ratio H/L) | complex |
|---|---|---|---|
| EP400 | 25,43 | 81,32 (0,012) | TIP60 |
| EPC1 | 25,18 | 49,73 (0,020) | |
| TIP60 | 12,86 | 21,30 (0,047) | |
| BRD8 | 12,75 | 18,42 (0,054) | |
| ING3 | 12,34 | 30,86 (0,032) | |
| PAF400 | 11,10 | 28,32 (0,035) | |
| MRGBP | 4,95 | 3,34 (0,299) | |
| MEAF6 | 4,68 | 5,16 (0,194) | |
| *MRG15* | *2,79* | *2,05 (0,488)* | |
| DMAP1 | 22,85 | 65,40 (0,015) | TIP60 / SRCAP |
| TIP49B | 22,05 | 57,00 (0,018) | |
| TIP49A | 21,42 | 55,18 (0,018) | |
| GAS41 | 15,24 | 54,06 (0,018) | |
| YL1 | 14,24 | 31,44 (0,032) | |
| BAF53A | 10,10 | 20,98 (0,048) | |
| ARP6 | 27,52 | 71,25 (0,014) | SRCAP |
| SRCAP | 25,51 | 59,24 (0,017) | |
| ZNHIT1 | NA | 19,93 (0,050) | |
| GFP | 23,82 | 108,88 (0,009) | |

**Figure 4.** Z.2.2 associates with H2A.Z-specific SRCAP and TIP60 chaperone complexes. GFP-pull-downs for H2A.Z-specific chaperone complexes are shown. HK cells stably expressing GFP-Z.2.1 (**A**) and GFP-Z.2.2 (**B**) were SILAC-labeled and subjected to single-step affinity purifications of soluble nuclear proteins in a 'forward' (GFP-Z.2.1) or 'reverse' (GFP-Z.2.2) pull-down using GFP nanotrap beads. In

around 0, specific interactors can be found on the right side having a high ratio H/L or ratio L/H for Z.2.1 and Z.2.2, respectively. In accordance with previous studies (62–65), we found GFP-Z.2.1 to be part of two major complexes, the SRCAP and the p400/NuA4/TIP60 (TIP60) complexes (Figure 4A), as we were able to detect all of their thus far identified members, with the exception of actin, as significant outliers. Interestingly, GFP-Z.2.2 also associated with both SRCAP and TIP60 complexes (Figure 4B), showing an almost identical binding composition as GFP-Z.2.1 (Figure 4C). These results strongly imply that Z.2.2 is, similar to other H2A.Z variants, actively deposited into chromatin through specific chaperone complexes.

Based on these results, we predicted that Z.2.1 and Z.2.2 should be incorporated into chromatin in a highly similar spatial manner. Since both SRCAP and TIP60 chaperone complexes are evolutionary conserved between different species, we tested mouse C127 cells that do not express endogenous Z.2.2 for their ability to deposit GFP-Z.2.2. Hereby we should be able to distinguish whether SRCAP and TIP60 complexes are sufficient for deposition, or if other potential primate-specific factors are needed. GFP-Z.2.1 and -Z.2.2 were transiently expressed in C127 cells, S-phase stages highlighted by EdU-incorporation and co-localization patterns visualized by fluorescence microscopy (Figure 5). GFP-Z.2.1 and -Z.2.2 showed an almost identical chromatin localization and deposition pattern, suggesting that Z.2.2 is, like Z.2.1, deposited through SRCAP and TIP60 complexes. In accordance with a recent study, we observed an enrichment of both H2A.Z variants in facultative heterochromatin regions in interphase nuclei (66). Surprisingly, although H2A.Z is expressed in all cell cycle phases (67), and GFP-Z.2.1 and -Z.2.2 expression is driven by a constitutive active promoter, chromatin deposition of both proteins is underrepresented at replication foci. This result underlines our findings that Z.2.2 interacts with all members of both TIP60 and SRCAP complexes and is actively and not passively deposited, as would have been the case during S-phase when nucleosomes are highly exchanged.

### Structural changes in Z.2.2's C-terminus prevent histone octamer folding and enhance DNA breathing on structurally destabilized nucleosomes

Our findings thus far imply that Z.2.2 is incorporated into nucleosomes and most likely targeted by TIP60 and

**Figure 4.** Continued
each panel the ratio of the identified proteins after MS is plotted. Proteins known to interact with H2A.Z are indicated in the following way: members of the SRCAP complex in red, members of the TIP60 complex in blue and shared subunits in purple. Potential novel H2A.Z-interacting proteins are shown as green dots ('other outliers') and are distinguished from background binders (gray dots) and contaminants (yellow dots). See also Supplementary Table S1 for a list of all identified proteins. (**C**) List of the SRCAP and TIP60 complex members and their normalized binding intensity to Z.2.1 or Z.2.2. Note that for comparison reasons the obtained H/L ratios of GFP-Z.2.2 binders (numbers in brackets) were calculated in the corresponding L/H ratios. See also Supplementary Table S1 for a list of all identified proteins and their normalized H/L ratios.
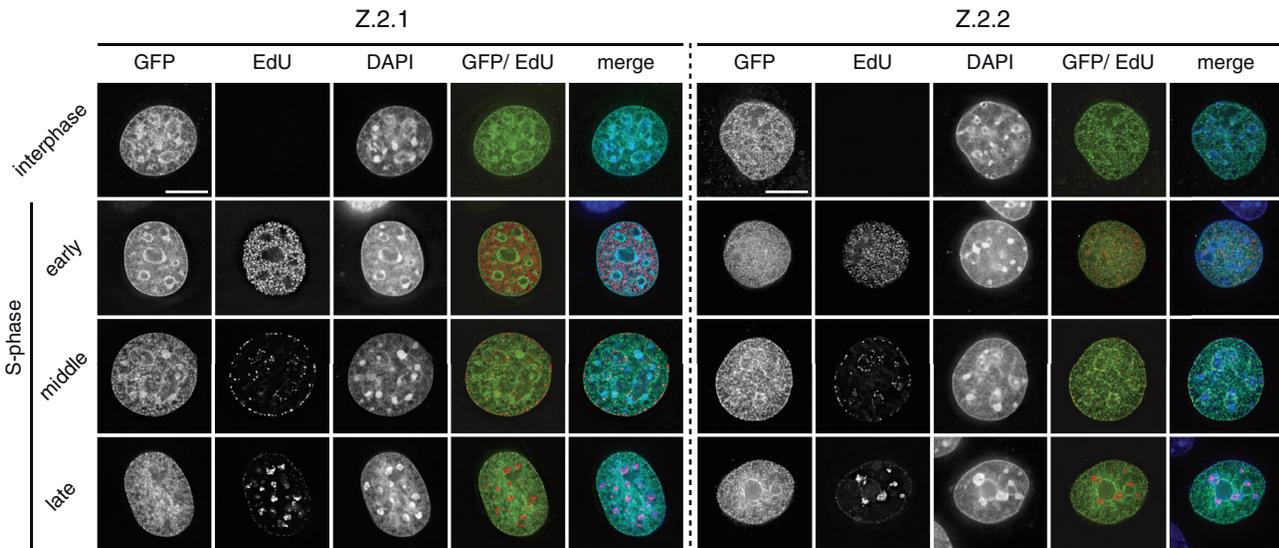
**Figure 5.** Z.2.1 and Z.2.2 are actively deposited into chromatin and are under-represented at replication foci. C127 cells transiently expressing GFP-Z.2.1 (left) and -Z.2.2 (right) were pulse labeled with EdU to visualize replication foci and to identify S-phase stages. DNA was counterstained with DAPI and analyzed by wide-field deconvolution microscopy. To remove the unbound fraction in GFP-Z.2.2 expressing cells, an *in situ* extraction was performed prior to fixation. Cells in early, middle and late S-phases were distinguished due to their characteristic differential EdU replication labeling patterns of eu- and heterochromatic regions. Merged images in color are shown alongside (GFP: green; EdU: red; DAPI: blue). Scale bar = 5 μm.
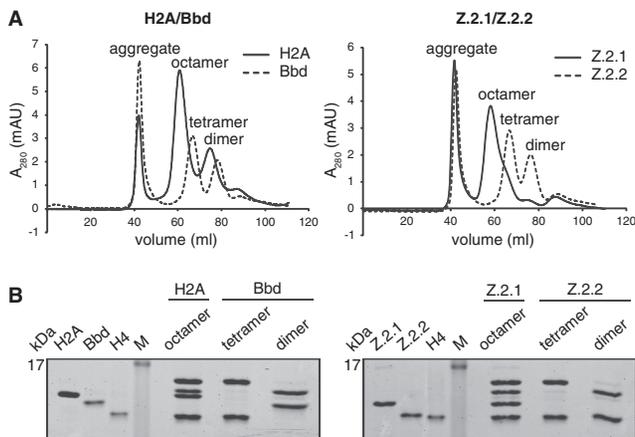


**Figure 6.** Z.2.2 does not constitute stable histone octamers with H2B, H3 and H4 *in vitro*. (**A**) Size exclusion chromatography of refolding reactions using recombinant human H3, H4 and H2B proteins together with either H2A (solid line) or Bbd (dashed line) (left overlay) or with either Z.2.1 (solid line) or Z.2.2 (dashed line) (right overlay). Peaks corresponding to aggregates, histone octamers, tetramers or dimers are labeled respectively. (**B**) Fractions corresponding to H2A-containing octamers, Bbd-containing tetramers and dimers (left) or Z.2.1-containing octamers and Z.2.2-containing tetramers and dimers (right) were analyzed by 18% SDS–PAGE and stained with Coomassie brilliant blue.

SRCAP complexes. Then why does a large fraction of the cellular Z.2.2 protein pool shows a high mobility and is freely dispersed in the nucleus? One plausible possibility is that Z.2.2 severely destabilizes nucleosomes due to its divergent C-terminal docking domain and is hence rapidly exchanged. To test this hypothesis, we used an *in vitro* reconstitution system. Recombinant human H2A

variants together with H3, H2B and H4 (Supplementary Figure S4A) were refolded by dialysis, and formed complexes purified by size exclusion chromatography. As expected, both H2A and Z.2.1 containing samples readily formed histone octamers (Figure 6A, solid lines). Bbd served as a negative control, because it has been demonstrated to not form octamers under these conditions (41), a result we also observed (Figure 6A left, dotted line). Interestingly, in accordance with our FRAP data, Z.2.2 behaved like Bbd in that it only formed Z.2.2–H2B dimers, but did not complex together with (H3–H4)$_2$ tetramers to generate octamers (Figure 6A right, dotted line), which was further confirmed by SDS–PAGE analyses of the separate fractions (Figure 6B). Thus, like for Bbd the incorporation of Z.2.2 destabilizes the interface between Z.2.2–H2B dimers and (H3–H4)$_2$ tetramers in a C-terminal sequence dependent manner (Supplementary Figure S4B and C). In conclusion, the Z.2.2 docking domain is sufficient to prevent octamer formation.

Although no Z.2.2 containing histone octamers could be generated *in vitro*, our results using GFP-Z.2.2 strongly suggest that Z.2.2 can be part of nucleosomes. To test this *in vitro* and to evaluate the effect of Z.2.2 on nucleosome stability, we reconstituted mononucleosomes by mixing Z.2.2–H2B dimers, (H3–H4)$_2$ tetramers and DNA containing a 'Widom 601' DNA positioning sequence in a 2:1:1 ratio. As controls, we reconstituted H2A or Z.2.1 containing nucleosomes by mixing octamers and DNA in a 1:1 ratio. As expected, analysis of all nucleosomes by native PAGE showed a single band before and after heat shift (Figure 7A), indicating a unique position on the 'Widom 601' DNA template. Purification of nucleosomes from a native gel and analysis of the protein content by
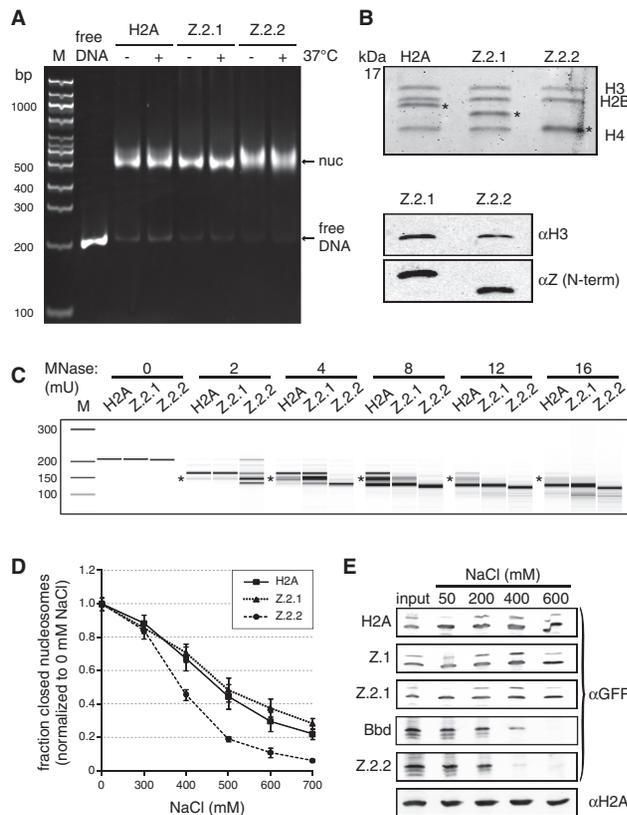
**Figure 7.** Z.2.2-containing nucleosomes are less resistant to MNase digestion and increased ionic strength. (**A**) H2A, Z.2.1 or Z.2.2 containing nucleosomes were assembled on DNA by salt gradient deposition, incubated at 4°C or 37°C to evaluate DNA positioning and separated by a native 5% PAGE gel. (**B**) Agarose-gel-electro-eluted material from (A) was analyzed by 18% SDS–PAGE and Coomassie stained to evaluate stoichiometry of histones after nucleosome assembly (top). Stars indicate H2A variants that were used for assembly. Further evaluation of histone stoichiometry after nucleosome assembly was done by IB using a LI-COR instrument (bottom). Assembled nucleosomes containing Z.2.1 or Z.2.2 were immunoblotted and the amount of histones was visualized using an αH3 antibody (top) and an N-terminal αZ antibody (recognizes all H2A.Z variants, bottom). (**C**) Mononucleosomes containing either H2A, Z.2.1 or Z.2.2 were digested with increasing concentrations of MNase and extracted DNA was separated using Bioanalyzer. Stars indicate DNA length of 146 bp. For detailed electropherogram analyses of fragment lengths in each sample see Supplementary Figure S5. (**D**) Mononucleosomes containing either H2A, Z.2.1 or Z.2.2 histones together with double dye labeled DNA were incubated with increasing amounts of salt. smFRET measurement values of each salt concentration were normalized to 0 mM NaCl. Error bars represent SEM of six measurements. (**E**) Chromatin from HK-GFP cells was isolated and incubated with increasing amounts of salt. Chromatin-bound histones were precipitated and detected by IB using αGFP antibody. Staining with αH2A was used as loading control.

SDS–PAGE (Coomassie staining and immunoblot) showed that Z.2.2 was indeed incorporated into nucleosomes (Figure 7B). All nucleosomes were further evaluated for their resistance to MNase cleavage as an indicator of stably organized nucleosomes and to determine nucleosomal DNA length (Figure 7C and Supplementary Figure S5). We observed fragments corresponding to protected nucleosomal DNA with the length of

146 bp for all variant nucleosomes tested. The appearance of smaller, subnucleosomal fragments indicates that DNA breathing occurred (68). Interestingly, DNA of Z.2.2 nucleosomes is less protected, since subnucleosomal fragments were obtained at lower MNase concentrations than with H2A or Z.2.1 nucleosomes. Additionally, at higher MNase concentrations a stable DNA fragment of about 120 bp was most abundant for Z.2.2 nucleosomes (Supplementary Figure S5), indicating that this might be the preferred DNA length wrapped around this octamer. These data suggest that increased DNA breathing occurs in Z.2.2 nucleosomes, which as a result might be less stable. To quantify nucleosome stability *in vitro* we measured salt-dependent changes in nucleosome structure using smFRET (69). In line with the results presented above, Z.2.2 containing recombinant nucleosomes lost their compact structure at lower salt concentrations than Z.2.1 or H2A-containing ones (Figure 7D). To investigate whether the observed Z.2.2-dependent nucleosome destabilization is true in the context of chromatin, we isolated chromatin from HK cells expressing GFP-H2A variants and incubated it with buffer containing increasing amounts of salt. Histones that remained stable chromatin components were precipitated and detected by IB (Figure 7E). As observed with FRET techniques, Z.2.2-containing nucleosomes disintegrated between 200 and 400 mM NaCl, and were therefore even less stable than Bbd-containing ones. In summary, incorporation of Z.2.2 leads to a severely reduced nucleosome stability due to C-terminal sequence dependent changes in its docking domain and subsequent loss of its interaction with histone H3.

Our FRAP data suggest that the Z.2.2 C-terminal amino acids might have a direct influence on the nucleosomal structure by affecting interactions with DNA and/or adjacent histones. Based on the existing structural data (50), we performed MD simulations of nucleosomes containing Z.1 (Supplementary Figure S7) or Z.2.2. In addition, we also included the deletion mutant Z.2.1[113], which did not show any change in short-term FRAP (Figure 3C), but some increase in mobility in long-term FRAP (Supplementary Figure S3B) in our assay. These *in silico* models revealed that changes in the C-terminus of H2A.Z strongly affect its protein structure (Figure 8A). Strikingly, different statistical descriptors over the MD-trajectory like distance and mobility (B-factor) show in contrast to Z.1 and Z.2.1[113] unique properties for the Z.2.2 tail. Only Z.2.2 leads to a substantial structural change in the C-terminus resulting in an increased distance to histone H3, which in turn makes a hydrogen bond interaction between peptide backbone NH of Cys112 in Z.2.2 and the oxygen in the Gln55 side chain in H3 impossible (Figure 8B). Additionally, an increase in the B-factor for Z.2.2 indicates a substantially enhanced mobility of Z.2.2's C-terminus (Figure 8C). We also calculated the Z.2.2–H3 interaction energy and observed a switch from negative to positive values in the case of Z.2.2 suggesting that this histone variant destabilizes the nucleosome (Figure 8D). In summary, these data suggest that the C-terminal sequence of Z.2.2 leads to a more dynamic structure that in turn loses binding to histone
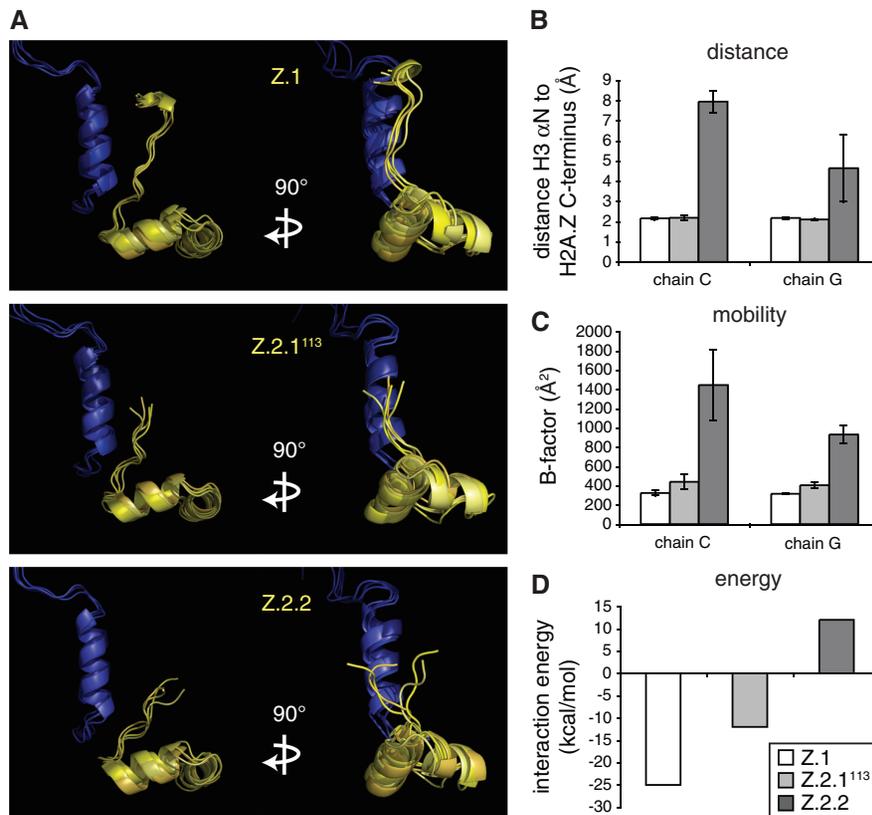
**Figure 8.** Unique Z.2.2 C-terminal amino acids cause significant changes in protein and nucleosome structure. (**A**) *In silico* models of Z.1, Z.2.1[113] and Z.2.2 C-terminal C-chains (yellow; from amino acids 84 to C-terminus, including the complete docking domain) together with the E-chain of histone H3 (blue; amino acids 33–60, including αN-helix). Side (left) and frontal views (right) of four MD simulations are shown respectively. See Supplementary Figure S7 for complete *in silico* model of H2A.Z-containing nucleosome. (**B**) Simulated distances between peptide backbone NH of amino acids 112 in H2A.Z (His or Cys, respectively) variants and the oxygen in the Gln55 sidechain in H3 based on *in silico* nucleosome models containing either Z.1 (white), Z.2.1[113] (light gray) or Z.2.2 (dark gray) proteins. Error bars represent SD of four independent simulations. (**C**) Simulated mobility measuring B-factor values between amino acids 108 and 113 in respective H2A.Z variant C-termini. Error bars represent SD of four independent simulations. (**D**) Simulated interaction energy between tetramer versus respective H2A.Z variant-containing dimers.

H3 and destabilizes the nucleosomal structure, providing a reasonable explanation for the observed *in vivo* and *in vitro* data.

## DISCUSSION

In this work, we have identified a previously unknown histone H2A.Z variant and provide a comprehensive characterization of its nucleosomal properties. This alternatively spliced variant, Z.2.2, is present to different degrees in all human cell lines and tissues investigated, with a significant enrichment in brain. Z.2.2 contains a shortened and in six amino acids divergent C-terminus compared to Z.1 and Z.2.1 that is necessary, but not sufficient, to weaken chromatin association. Only in the context of the unique Z.2.2 docking domain does the C-terminal sequence negatively affect nucleosome stability *in vitro* and *in vivo*. To our knowledge, Z.2.2 has the strongest destabilizing effect on nucleosomal structure compared to other histone H2A variants reported to date.

Only one other histone variant, macroH2A, has been shown thus far to be alternatively spliced (70). Here, like

our observation with H2A.Z, two independent genes *mH2A1* and *mH2A2* exist in mammals, with only mH2A1 being alternatively spliced resulting in functional different proteins (71). In our study, we demonstrate that the human H2A.Z.2 (H2AFV) primary transcript is alternatively spliced generating Z.2.1 and Z.2.2 mRNAs and proteins. These observations suggest that Z.2.2 is tightly regulated in a tissue-specific manner through alternative splicing and/or RNA stability. Our findings now raise the intriguing possibility that alternative splicing of histone variants might not be rare but more common than previously thought. If true, it will be of interest to reevaluate other intron-containing histone variant genes with regard to their possible alternative transcripts and protein products.

Bioinformatic genome analyses revealed the existence of Z.2.2-specific sequences only in humans, old and new world primates and to some extend in other mammals, with the exclusion of mouse, rat and even lower eukaryotes. It remains to be seen, whether Z.2.2's evolution is indeed limited to primates only. Primate-specific gene products have been often identified in human brain and reproductive tissues (72), supporting the notion that

their RNAs and proteins might be essential to adaptive changes leading to human development and further speculates that primate-specific genes might be important in reproductive function and disease. Since we have found Z.2.2 transcripts to be strongly enriched in brain samples of higher brain function in comparison to other tissues and cell types, it will be of great interest to determine in future studies, if this novel variant might play an important functional role in this particular organ. These observations also raise the interesting question of how alternative splicing and/or differential stability of H2AFV transcripts are tissue specifically regulated.

Another intriguing feature of Z.2.2 is its influence on nucleosome stability. Although Z.2.2 localizes exclusively to the nucleus, only a minor proportion is stably incorporated into chromatin. The only other exception in humans known thus far is Bbd, which has previously been demonstrated to destabilize the nucleosome structure (41,53,73). Bbd, similar to Z.2.2, is a shorter H2A variant with an unusual C-terminus and a considerable different primary histone fold sequence that might explain its ability to destabilize nucleosomes. In agreement, a recent study demonstrated that the incomplete C-terminal docking domain of Bbd results in structural alterations in nucleosomes and that those are in turn associated with an inability of the chromatin remodeler RSC to both remodel and mobilize nucleosomes (8). Z.2.2, on the other hand, is identical to Z.2.1, except that its C-terminus is 14 amino acids shorter and in six amino acids altered. How can this small change in Z.2.2's primary sequence lead to such drastic effects on chromatin association?

We show that Z.2.2 can be part of a bona fide nucleosome and that it interacts with the H2A.Z-specific TIP60 and SRCAP chaperone complexes. These complexes have been shown to catalyze the exchange of H2A–H2B dimers with H2A.Z–H2B dimers in nucleosomes and our finding therefore suggests that both complexes are also involved in an active chromatin incorporation of Z.2.2. Supporting this idea is the observation that both Z.2.1 and Z.2.2 are incorporated into chromatin in a replication-independent manner, even in mouse cells that do not express endogenous Z.2.2. Both H2A.Z variants are not primarily deposited at replication foci, not even in middle S-phase when facultative heterochromatin is replicated, where the majority of the H2A.Z protein pool is found in interphase cells (66). Our findings are in agreement with a model proposed by Hardy and Robert, in which H2A.Z variants are randomly deposited into chromatin by specific chaperone complexes in a replication-independent manner coupled to a subsequent targeted H2A.Z depletion (74). As a consequence, an enrichment of H2A.Z at non-transcribed genes and heterochromatin regions over several cell generations can be observed (74). It might be possible that INO80 facilitates this eviction function, as it has been shown to exchange nucleosomal H2A.Z–H2B dimers with free H2A–H2B dimers (75). It will be of interest in future studies to determine whether Z.2.2 exchange is subjected to a similar mechanism. Taken together, our findings strongly imply that Z.2.2 is actively deposited into chromatin through the interaction

with evolutionary conserved chaperone complexes. Nevertheless, a large fraction of Z.2.2 protein is not chromatin bound and we have mapped the region crucial for high FRAP mobility to its docking domain. In addition to Z.2.2's unique C-terminal amino acids this region spans the highly conserved acidic patch responsible for deposition (76), the M6 region that is functionally essential in fly H2A.Z (60) and required for the interaction with the SWR1 complex in yeast (77). Strikingly, *in silico* simulation of Z.2.2 predicted dynamic structural changes that in turn weaken interaction with histone H3 and destabilize the nucleosome structure. Such a gross structural alteration explains why Z.2.2 is not able to form stable octamers *in vitro* and leads to enhanced DNA breathing in a nucleosomal context. Hence, Z.2.2 incorporation into chromatin disrupts nucleosomes more easily and supports a model in which Z.2.2 is more rapidly exchanged than Z.2.1.

What functional outcome might Z.2.2 cause when incorporated into chromatin? And how is the variant composition of Z.2.2 containing nucleosomes? It has been shown that a special class of nucleosomes containing both H2A.Z and H3.3 variants exists in humans (78). These nucleosomes are enriched at promoters, enhancers and insulator region and promote the accessibility of transcription factors to these DNA regions (78), most likely due to their extreme sensitivity to disruption (79). Since these studies nicely demonstrate that differential nucleosome stabilities due to the incorporation of different histone variants influence transcriptional regulation, it is tempting to speculate that Z.2.2 might also affect chromatin-related processes. Future experiments will shed light on Z.2.2 function(s), especially with regard to its increased expression in human brain areas, and explain why and where nucleosomal destabilization is needed. This is of particular interest, since Bbd that also leads to nucleosomal destabilization is almost exclusively present in testis (80–82) in contrast to the apparently ubiquitously expressed Z.2.2, possibly pointing toward distinct roles of both destabilizing H2A variants in different tissues. Our data suggest that additional interesting, yet unidentified, histone variants may exist and await their discovery.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Table S1, Supplementary Figures S1–S7 and Supplementary Materials and Methods.

manuscript. The authors thank especially the Hake lab and members of the Adolf-Butenandt Institute for constructive discussions.

## REFERENCES

1. van Holde,K.E. (1988) *Chromatin.* Springer, New York.
2. Bonisch,C., Nieratschker,S.M., Orfanos,N.K. and Hake,S.B. (2008) Chromatin proteomics and epigenetic regulatory circuits. *Exp. Rev. Proteomics*, **5**, 105–119.
3. Bernstein,E. and Hake,S.B. (2006) The nucleosome: a little variation goes a long way. *Biochem. Cell Biol.*, **84**, 505–517.
4. Boulard,M., Bouvet,P., Kundu,T.K. and Dimitrov,S. (2007) Histone variant nucleosomes: structure, function and implication in disease. *Subcell. Biochem.*, **41**, 71–89.
5. Redon,C., Pilch,D., Rogakou,E., Sedelnikova,O., Newrock,K. and Bonner,W. (2002) Histone H2A variants H2AX and H2AZ. *Curr. Opin. Genet. Dev.*, **12**, 162–169.
6. Ausio,J. and Abbott,D.W. (2002) The many tales of a tail: carboxyl-terminal tail heterogeneity specializes histone H2A variants for defined chromatin function. *Biochemistry*, **41**, 5945–5949.
7. Wang,A.Y., Aristizabal,M.J., Ryan,C., Krogan,N.J. and Kobor,M.S. (2011) Key functional regions in the histone variant H2A.Z C-terminal docking domain. *Mol. Cell. Biol.*, **31**, 3871–3884.
8. Shukla,M.S., Syed,S.H., Goutte-Gattat,D., Richard,J.L., Montel,F., Hamiche,A., Travers,A., Faivre-Moskalenko,C., Bednar,J., Hayes,J.J. *et al.* (2011) The docking domain of histone H2A is required for H1 binding and RSC-mediated nucleosome remodeling. *Nucleic Acids Res.*, **39**, 2559–2570.
9. 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.
10. Billon,P. and Cote,J. (2011) Precise deposition of histone H2A.Z in chromatin for genome expression and maintenance. *Biochim Biophys Acta*, **1819**, 290–302.
11. 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*, **5**, 267–272.
12. Draker,R. and Cheung,P. (2009) Transcriptional and epigenetic functions of histone variant H2A.Z. *Biochem. Cell Biol.*, **87**, 19–25.
13. Svotelis,A., Gevry,N. and Gaudreau,L. (2009) Regulation of gene expression and cellular proliferation by histone H2A.Z. *Biochem. Cell Biol.*, **87**, 179–188.
14. Brickner,D.G., Cajigas,I., Fondufe-Mittendorf,Y., Ahmed,S., Lee,P.C., Widom,J. and Brickner,J.H. (2007) H2A.Z-mediated localization of genes at the nuclear periphery confers epigenetic memory of previous transcriptional state. *PLoS Biol.*, **5**, e81.
15. Zofall,M., Fischer,T., Zhang,K., Zhou,M., Cui,B., Veenstra,T.D. and Grewal,S.I. (2009) Histone H2A.Z cooperates with RNAi and heterochromatin factors to suppress antisense RNAs. *Nature*, **461**, 419–422.
16. Rangasamy,D., Berven,L., Ridgway,P. and Tremethick,D.J. (2003) Pericentric heterochromatin becomes enriched with H2A.Z during early mammalian development. *EMBO J.*, **22**, 1599–1607.
17. Rangasamy,D., Greaves,I. and Tremethick,D.J. (2004) RNA interference demonstrates a novel role for H2A.Z in chromosome segregation. *Nat. Struct. Mol. Biol.*, **11**, 650–655.
18. Zlatanova,J. and Thakar,A. (2008) H2A.Z: view from the top. *Structure*, **16**, 166–179.
19. Dryhurst,D., McMullen,B., Fazli,L., Rennie,P.S. and Ausio,J. (2012) Histone H2A.Z prepares the prostate specific antigen (PSA) gene for androgen receptor-mediated transcription and is upregulated in a model of prostate cancer progression. *Cancer Lett.*, **315**, 38–47.
20. Valdes-Mora,F., Song,J.Z., Statham,A.L., Strbenac,D., Robinson,M.D., Nair,S.S., Patterson,K.I., Tremethick,D.J., Stirzaker,C. and Clark,S.J. (2012) Acetylation of H2A.Z is a key epigenetic modification associated with gene deregulation and epigenetic remodeling in cancer. *Genome Res.*, **22**, 307–321.
21. Conerly,M.L., Teves,S.S., Diolaiti,D., Ulrich,M., Eisenman,R.N. and Henikoff,S. (2010) Changes in H2A.Z occupancy and DNA methylation during B-cell lymphomagenesis. *Genome Res.*, **20**, 1383–1390.
22. Svotelis,A., Gevry,N., Grondin,G. and Gaudreau,L. (2010) H2A.Z overexpression promotes cellular proliferation of breast cancer cells. *Cell Cycle*, **9**, 364–370.
23. 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.
24. Eirin-Lopez,J.M., Gonzalez-Romero,R., Dryhurst,D., Ishibashi,T. and Ausio,J. (2009) The evolutionary differentiation of two histone H2A.Z variants in chordates (H2A.Z-1 and H2A.Z-2) is mediated by a stepwise mutation process that affects three amino acid residues. *BMC Evol. Biol.*, **9**, 31.
25. Beck,H.C., Nielsen,E.C., Matthiesen,R., Jensen,L.H., Sehested,M., Finn,P., Grauslund,M., Hansen,A.M. and Jensen,O.N. (2006) Quantitative proteomic analysis of post-translational modifications of human histones. *Mol. Cell Proteomics*, **5**, 1314–1325.
26. Bonenfant,D., Coulot,M., Towbin,H., Schindler,P. and van Oostrum,J. (2006) Characterization of histone H2A and H2B variants and their post-translational modifications by mass spectrometry. *Mol. Cell Proteomics*, **5**, 541–552.
27. Dryhurst,D., Ishibashi,T., Rose,K.L., Eirin-Lopez,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 Biol.*, **7**, 86.
28. Sarcinella,E., Zuzarte,P.C., Lau,P.N., Draker,R. and Cheung,P. (2007) Monoubiquitylation of H2A.Z distinguishes its association with euchromatin or facultative heterochromatin. *Mol. Cell. Biol.*, **27**, 6457–6468.
29. Wiedemann,S.M., Mildner,S.N., Bonisch,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.
30. Shechter,D., Dormann,H.L., Allis,C.D. and Hake,S.B. (2007) Extraction, purification and analysis of histones. *Nat. Protoc.*, **2**, 1445–1457.
31. Wysocka,J., Reilly,P.T. and Herr,W. (2001) Loss of HCF-1-chromatin association precedes temperature-induced growth arrest of tsBN67 cells. *Mol. Cell. Biol.*, **21**, 3820–3829.
32. Wessel,D. and Flugge,U.I. (1984) A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal. Biochem.*, **138**, 141–143.
33. Zhang,H., Roberts,D.N. and Cairns,B.R. (2005) Genome-wide dynamics of Htz1, a histone H2A variant that poises repressed/basal promoters for activation through histone loss. *Cell*, **123**, 219–231.

34. Hake,S.B., Garcia,B.A., Kauer,M., Baker,S.P., Shabanowitz,J., Hunt,D.F. and Allis,C.D. (2005) Serine 31 phosphorylation of histone variant H3.3 is specific to regions bordering centromeres in metaphase chromosomes. *Proc. Natl Acad. Sci. USA*, **102**, 6344–6349.

35. Vermeulen,M., Eberl,H.C., Matarese,F., Marks,H., Denissov,S., Butter,F., Lee,K.K., Olsen,J.V., Hyman,A.A., Stunnenberg,H.G. *et al.* (2010) Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell*, **142**, 967–980.

36. Vermeulen,M., Mulder,K.W., Denissov,S., Pijnappel,W.W., van Schaik,F.M., Varier,R.A., Baltissen,M.P., Stunnenberg,H.G., Mann,M. and Timmers,H.T. (2007) Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell*, **131**, 58–69.

37. Cox,J. and Mann,M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.*, **26**, 1367–1372.

38. 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.

39. Dyer,P.N., Edayathumangalam,R.S., White,C.L., Bao,Y., Chakravarthy,S., Muthurajan,U.M. and Luger,K. (2004) Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol.*, **375**, 23–44.

40. Thastrom,A., Lowary,P.T., Widlund,H.R., Cao,H., Kubista,M. and Widom,J. (1999) Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences. *J. Mol. Biol.*, **288**, 213–229.

41. 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.

42. Gansen,A., Valeri,A., Hauger,F., Felekyan,S., Kalinin,S., Toth,K., Langowski,J. and Seidel,C.A. (2009) Nucleosome disassembly intermediates characterized by single-molecule FRET. *Proc. Natl Acad. Sci. USA*, **106**, 15308–15313.

43. Muller,B.K., Zaychikov,E., Brauchle,C. and Lamb,D.C. (2005) Pulsed interleaved excitation. *Biophys. J.*, **89**, 3508–3522.

44. Nir,E., Michalet,X., Hamadani,K.M., Laurence,T.A., Neuhauser,D., Kovchegov,Y. and Weiss,S. (2006) Shot-noise limited single-molecule FRET histograms: comparison between theory and experiments. *J. Phys. Chem. B*, **110**, 22103–22124.

45. Lee,N.K., Kapanidis,A.N., Wang,Y., Michalet,X., Mukhopadhyay,J., Ebright,R.H. and Weiss,S. (2005) Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation. *Biophys. J.*, **88**, 2939–2953.

46. Duan,Y., Wu,C., Chowdhury,S., Lee,M.C., Xiong,G., Zhang,W., Yang,R., Cieplak,P., Luo,R., Lee,T. *et al.* (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.*, **24**, 1999–2012.

47. Wang,J., Wolf,R.M., Caldwell,J.W., Kollman,P.A. and Case,D.A. (2004) Development and testing of a general amber force field. *J. Comput. Chem.*, **25**, 1157–1174.

48. Jakalian,A., Jack,D.B. and Bayly,C.I. (2002) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.*, **23**, 1623–1641.

49. Krieger,E., Darden,T., Nabuurs,S.B., Finkelstein,A. and Vriend,G. (2004) Making optimal use of empirical energy functions: force-field parameterization in crystal space. *Proteins*, **57**, 678–683.

50. 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.

51. Krieger,E., Nielsen,J.E., Spronk,C.A. and Vriend,G. (2006) Fast empirical pKa prediction by Ewald summation. *J. Mol. Graph Model.*, **25**, 481–486.

52. Higashi,T., Matsunaga,S., Isobe,K., Morimoto,A., Shimada,T., Kataoka,S., Watanabe,W., Uchiyama,S., Itoh,K. and Fukui,K. (2007) Histone H2A mobility is regulated by its tails and acetylation of core histone tails. *Biochem. Biophys. Res. Commun.*, **357**, 627–632.

53. Gautier,T., Abbott,D.W., Molla,A., Verdel,A., Ausio,J. and Dimitrov,S. (2004) Histone variant H2ABbd confers lower stability to the nucleosome. *EMBO Rep.*, **5**, 715–720.

54. Lever,M.A., Th'ng,J.P., Sun,X. and Hendzel,M.J. (2000) Rapid exchange of histone H1.1 on chromatin in living human cells. *Nature*, **408**, 873–876.

55. Th'ng,J.P., Sung,R., Ye,M. and Hendzel,M.J. (2005) H1 family histones in the nucleus. Control of binding and localization by the C-terminal domain. *J. Biol. Chem.*, **280**, 27809–27814.

56. Misteli,T., Gunjan,A., Hock,R., Bustin,M. and Brown,D.T. (2000) Dynamic binding of histone H1 to chromatin in living cells. *Nature*, **408**, 877–881.

57. Stasevich,T.J., Mueller,F., Brown,D.T. and McNally,J.G. (2010) Dissecting the binding mechanism of the linker histone in live cells: an integrated FRAP analysis. *EMBO J.*, **29**, 1225–1234.

58. Schmiedeberg,L., Weisshart,K., Diekmann,S., Meyer Zu Hoerste,G. and Hemmerich,P. (2004) High- and low-mobility populations of HP1 in heterochromatin of mammalian cells. *Mol. Biol. Cell.*, **15**, 2819–2833.

59. Cheutin,T., McNairn,A.J., Jenuwein,T., Gilbert,D.M., Singh,P.B. and Misteli,T. (2003) Maintenance of stable heterochromatin domains by dynamic HP1 binding. *Science*, **299**, 721–725.

60. Clarkson,M.J., Wells,J.R., Gibson,F., Saint,R. and Tremethick,D.J. (1999) Regions of variant histone His2AvD required for Drosophila development. *Nature*, **399**, 694–697.

61. Lu,P.Y., Levesque,N. and Kobor,M.S. (2009) NuA4 and SWR1-C: two chromatin-modifying complexes with overlapping functions and components. *Biochem. Cell Biol.*, **87**, 799–815.

62. Cai,Y., Jin,J., Florens,L., Swanson,S.K., Kusch,T., Li,B., Workman,J.L., Washburn,M.P., Conaway,R.C. and Conaway,J.W. (2005) The mammalian YL1 protein is a shared subunit of the TRRAP/TIP60 histone acetyltransferase and SRCAP complexes. *J. Biol. Chem.*, **280**, 13665–13670.

63. Ruhl,D.D., Jin,J., Cai,Y., Swanson,S., Florens,L., Washburn,M.P., Conaway,R.C., Conaway,J.W. and Chrivia,J.C. (2006) Purification of a human SRCAP complex that remodels chromatin by incorporating the histone variant H2A.Z into nucleosomes. *Biochemistry*, **45**, 5671–5677.

64. Doyon,Y., Selleck,W., Lane,W.S., Tan,S. and Cote,J. (2004) Structural and functional conservation of the NuA4 histone acetyltransferase complex from yeast to humans. *Mol. Cell. Biol.*, **24**, 1884–1896.

65. Auger,A., Galarneau,L., Altaf,M., Nourani,A., Doyon,Y., Utley,R.T., Cronier,D., Allard,S. and Cote,J. (2008) Eaf1 is the platform for NuA4 molecular assembly that evolutionarily links chromatin acetylation to ATP-dependent exchange of histone H2A variants. *Mol. Cell. Biol.*, **28**, 2257–2270.

66. Hardy,S., Jacques,P.E., Gevry,N., Forest,A., Fortin,M.E., Laflamme,L., Gaudreau,L. and Robert,F. (2009) The euchromatic and heterochromatic landscapes are shaped by antagonizing effects of transcription on H2A.Z deposition. *PLoS Genet.*, **5**, e1000687.

67. Wu,R.S., Tsai,S. and Bonner,W.M. (1982) Patterns of histone variant synthesis can distinguish G0 from G1 cells. *Cell*, **31**, 367–374.

68. Ramaswamy,A., Bahar,I. and Ioshikhes,I. (2005) Structural dynamics of nucleosome core particle: comparison with nucleosomes containing histone variants. *Proteins*, **58**, 683–696.

69. Ha,T., Enderle,T., Ogletree,D.F., Chemla,D.S., Selvin,P.R. and Weiss,S. (1996) Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc. Natl Acad. Sci. USA*, **93**, 6264–6268.

70. Rasmussen,T.P., Huang,T., Mastrangelo,M.A., Loring,J., Panning,B. and Jaenisch,R. (1999) Messenger RNAs encoding mouse histone macroH2A1 isoforms are expressed at similar levels in male and female cells and result from alternative splicing. *Nucleic Acids Res.*, **27**, 3685–3689.

71. 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.

72. Tay,S.K., Blythe,J. and Lipovich,L. (2009) Global discovery of primate-specific genes in the human genome. *Proc. Natl Acad. Sci. USA*, **106**, 12019–12024.

73. 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.

74. Hardy,S. and Robert,F. (2010) Random deposition of histone variants: a cellular mistake or a novel regulatory mechanism? *Epigenetics*, **5**, 368–372.

75. Papamichos-Chronakis,M., Watanabe,S., Rando,O.J. and Peterson,C.L. (2011) Global regulation of H2A.Z localization by the INO80 chromatin-remodeling enzyme is essential for genome integrity. *Cell*, **144**, 200–213.

76. Jensen,K., Santisteban,M.S., Urekar,C. and Smith,M.M. (2011) Histone H2A.Z acid patch residues required for deposition and function. *Mol. Genet. Genomics.*, **285**, 287–296.

77. Wu,W.H., Alami,S., Luk,E., Wu,C.H., Sen,S., Mizuguchi,G., Wei,D. and Wu,C. (2005) Swc2 is a widely conserved H2AZ-binding module essential for ATP-dependent histone exchange. *Nat. Struct. Mol. Biol.*, **12**, 1064–1071.

78. Jin,C., Zang,C., Wei,G., Cui,K., Peng,W., Zhao,K. and Felsenfeld,G. (2009) H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nat. Genet.*, **41**, 941–945.

79. Jin,C. and Felsenfeld,G. (2007) Nucleosome stability mediated by histone variants H3.3 and H2A.Z. *Genes Dev.*, **21**, 1519–1529.

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

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

82. 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.*, **19**, 25–30.

# Supplementary information for:

# H2A.Z.2.2 is an alternatively spliced histone H2A.Z variant that causes severe nucleosome destabilization

Nucleic Acids Research, 2012, 1–14; doi:10.1093/nar/gks267
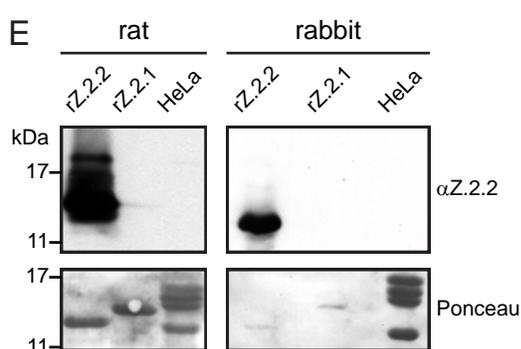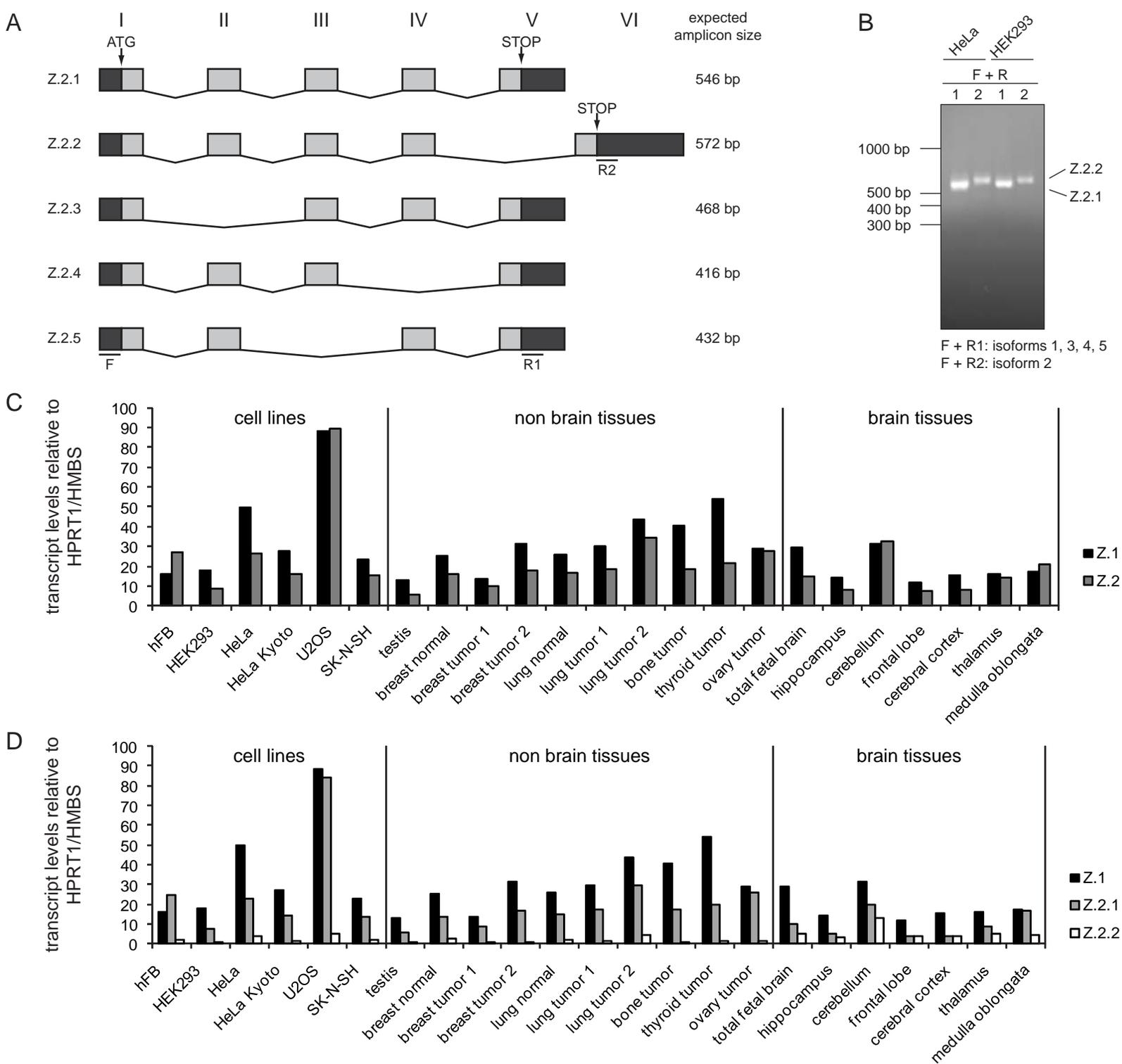
**Supplementary Figure S1: Identification of Z.2.2**.
**(A)** Schematic representation of exon/intron structure of the predicted five alternative human Z.2 (H2AFV) splice forms (http://www.ncbi.nlm.nih.gov/gene/94239). Untranslated sequences (UTR) are depicted in dark gray, coding regions in light gray and introns as kinked lines. Sizes of exons/introns are not drawn to scale. Primers used to amplify specific splice forms by PCR (see Supplementary Figure S1B) are shown as F (hybridizes to 5'UTR of all spliced mRNAs), R1 (specific for exon V of splice forms 1, 3, 4 and 5) and R2 (specific for exon VI of H2A.Z.2.2). Expected PCR amplicon sizes of each splice form are indicated on the right. **(B)** Agarose gel to visualize amplified DNA after PCR with cDNA generated from HeLa and HEK293 total RNA, respectively, using F+R1 (A, lanes 1 and 3) or F+R2 (B, lane 2 and 4) primer pairs. Expression levels of Z.1 and Z.2 (Z.2.1 + Z.2.2), and **(C)** of Z.1, Z.2.1 and Z.2.2 **(D)** mRNAs measured by qPCR and normalized to HPRT1 and HMBS. Controls generated without reverse transcriptase (no RT) were used to assess amplification threshold. **(E)** Immunoblot with recombinant Z.2.1, recombinant Z.2.2 and acid extracted HeLa histones to validate specificity of both, rat monoclonal and rabbit polyclonal antibodies (top). Equal loading was ensured by Ponceau staining of the membrane before antibody incubation (bottom).
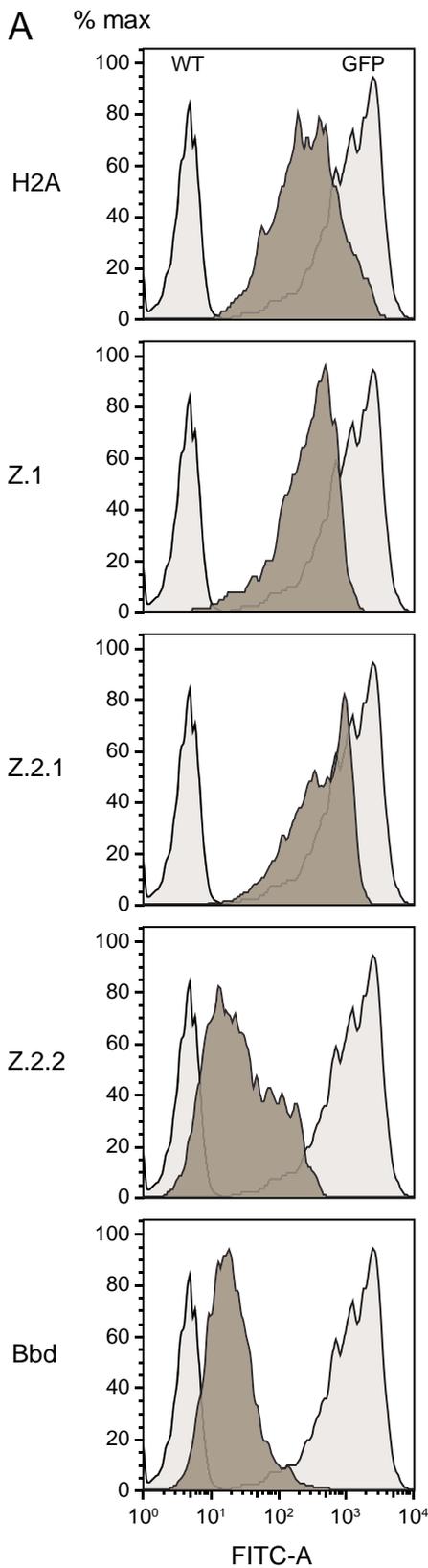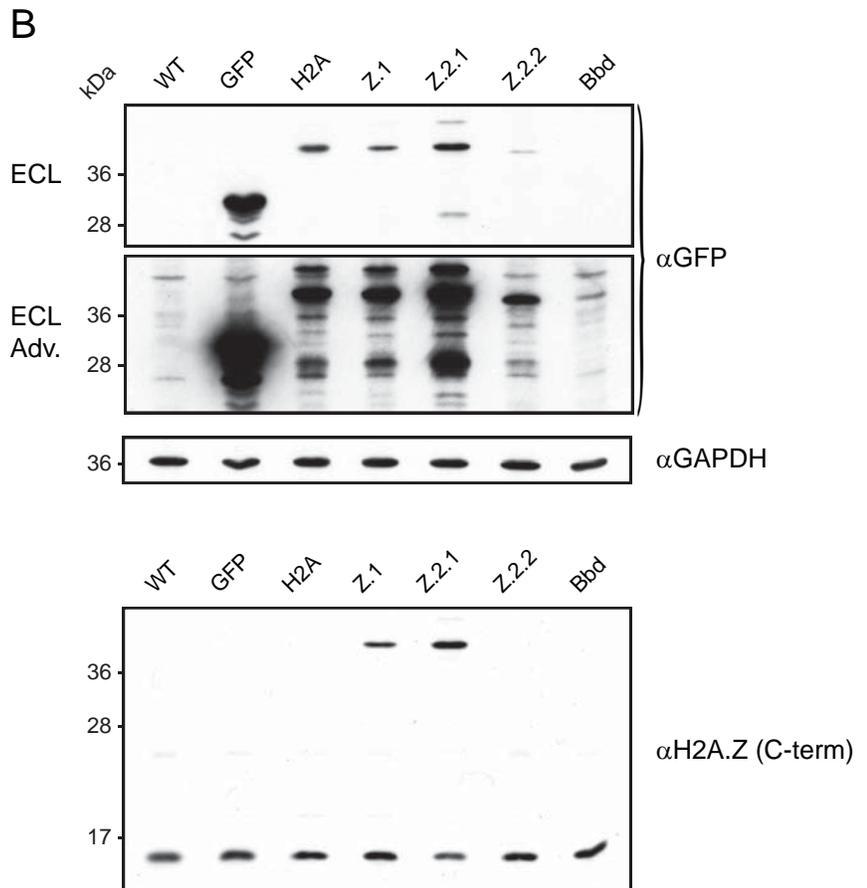
**Supplementary Figure S2: Characterization of HeLa Kyoto cell lines stably expressing GFP-H2A variants.**
**(A)** FACS profiles of representative HeLa Kyoto cell lines (single cell clones, dark gray) used in this study. WT and GFP controls (light gray) are shown for comparison. Table of mean fluorescence intensity (MFI) is shown on the right. **(B)** Immunoblot of whole cell extracts prepared from HeLa Kyoto cell lines stably expressing GFP-H2A variants. Immunoblot with αGFP antibody (top) to compare expression levels of different GFP-H2A variants. Expression of GFP-Z.2.2 and GFP-Bbd is much lower than for the other variants and GFP-Bbd is only detectable using ECL Advance. Equal loading was ensured by αGAPDH antibody. Immunoblot with antibody against H2A.Z C-terminus (bottom) shows similar expression of GFP tagged Z.1 and Z.2.1 compared with endogenous protein.

A



B

C



D



E

| | N | $t_{1/2}$ | | | fraction [%] | | |
|---|---|---|---|---|---|---|---|
| | | fast (s) | intermediate (min) | slow (h) | fast | intermediate | slow |
| GFP | 17 | 0.4 ± 0.2 | - | - | 100 | - | - |
| HP1$\alpha$ | 14 | 1.6 ± 0.3 | 0.3 ± 0.1 | - | 82 ± 6 | 18 ± 6 | - |
| H1.0 | 15 | 61.8 ± 23.5 | 9.9 ± 4.7 | - | 61 ± 13 | 39 ± 13 | - |
| Bbd | 14 | 2.5 ± 0.8 | 9.0 ± 2.5 | - | 52 ± 7 | 48 ± 7 | - |
| Z.2.2 | 29 | 1.1 ± 0.3 | 7.3 ± 4.1 | - | 78 ± 8 | 22 ± 8 | - |
| H2A[105] | 10 | 1.4 ± 0.3 | 14.5 ± 2.7 | - | 32 ± 4 | 68 ± 4 | - |
| H2A[105]+CZ.2.2 | 14 | 2.1 ± 1.0 | 10.6 ± 5.2 | - | 30 ± 10 | 70 ± 10 | - |
| H2A[88]+CZ.2.2 | 15 | 1.1 ± 0.2 | 2.7 ± 0.6 | - | 83 ± 8 | 17 ± 8 | - |
| | | | | | | | |
| H2A | 24 | - | 17.3 ± 6.8 | 4.0 ± 1.4 | - | 25 ± 9 | 75 ± 9 |
| Z.1 | 13 | - | 8.4 ± 7.7 | 1.6 ± 0.7 | - | 18 ± 11 | 82 ± 11 |
| Z.2.1 | 15 | - | 16.8 ± 9.5 | 4.9 ± 2.5 | - | 28 ± 8 | 72 ± 8 |
| H2A[111] | 14 | - | 10.7 ± 8.0 | 1.5 ± 1.2 | - | 42 ± 19 | 58 ± 19 |
| H2A[88]+CZ.2.1 | 25 | - | 16.7 ± 10.9 | 4.2 ± 2.7 | - | 28 ± 13 | 72 ± 13 |
| Z.2.1[113] | 13 | - | 13.5 ± 9.1 | 1.9 ± 1.1 | - | 44 ± 18 | 56 ± 18 |

**Supplementary Figure S3: FRAP analysis of GFP-H2A variants.**
**(A)** FRAP experiment to evaluate nucleosomal stability of wild type and mutant H2A variants using spinning disk confocal microscopy. A small nuclear area (box) of HeLa Kyoto cells expressing GFP-tagged H2A variants, H1.0, HP1$\alpha$ or GFP alone was photobleached and the recovery of the fluorescent signal was monitored for up to 5 hours. For each construct, selected time points of one exemplary time series are shown. Scale bar = 5 µm. **(B)** Long-term (60 min) FRAP quantification curves of average GFP signal recovery after photobleaching relative to flourescence intensity prior to bleaching are depicted for wild type (solid lines) and mutant (dashed lines) GFP-tagged H2A variant constructs, H1.0, HP1$\alpha$ and GFP. Mean curves of 10 to 29 individual cells are shown for each construct. For clarity error bars are omitted. (See also Figure 3B and 3C for short-term FRAP curves and Supplementary Figure S2E for numerical values and standard deviations). **(C)** Quantitative evaluation of FRAP curves. Plot shows calculated half-time of recovery ($t_{1/2}$) values of different control constructs, wild type and mutant H2A variant constructs based on bi-exponential fitting of FRAP data (± standard deviation (SD), see Fig. S2E for numerical values) **(D)** Schematic representation of GFP-tagged (green box) H2A variants deletion and domain swap constructs used in FRAP studies. H2A is depicted in yellow, Bbd in purple, Z.1/.2.1/.2.2 in orange with Z.2.2's unique C-terminal amino acids highlighted in red. **(E)** Table of mean half-time of recoveries ($t_{1/2}$) and respective mobility fraction sizes (± SD) as determined from exponential fitting. While a single exponential function was sufficient to fit GFP alone, all other GFP fusion proteins were fitted with a bi-exponential function identifying either a fast ($t_{1/2}$ < 62 s) and an intermediate fraction ($t_{1/2}$ 0.3-20 min), or an intermediate and a slow fraction ($t_{1/2}$ > 1.5 h).

A



B **H2A$^{88}$+CZ.2.1/H2A$^{88}$+CZ.2.2**
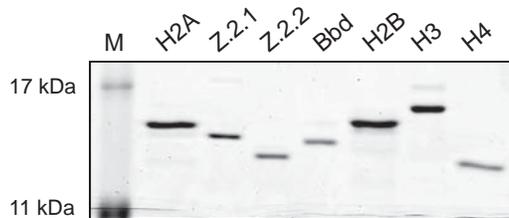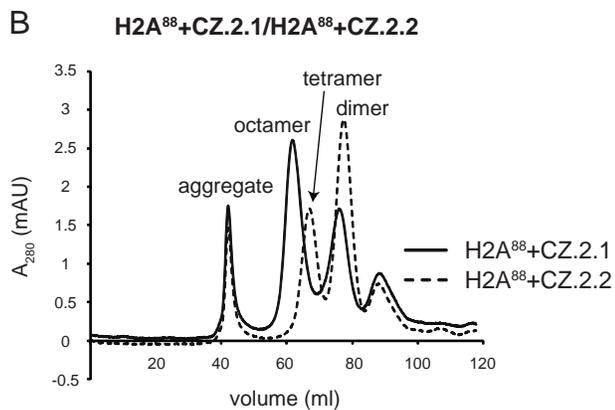


C
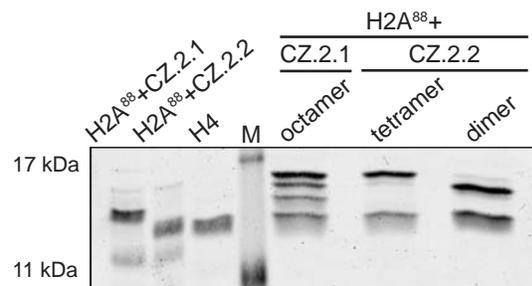


**Supplementary Figure S4: Z.2.2 destabilizes histones octamers in a sequence-dependent manner.**
**(A)** Evaluation of purity of recombinant human histone proteins. Human core histones (H3, H4, H2B, H2A) and H2A variants Bbd, Z.2.1 and Z.2.2 were expressed in E. coli. After purification, recombinant proteins were separated on 18% SDS-PAGE and stained with Coomassie brilliant blue. **(B)** Size exclusion chromatography of refolding reactions using recombinant histones, as described in Figure 6, but with the use of H2A$^{88}$+CZ.2.1 (solid line) and H2A$^{88}$+CZ.2.2 (dashed line) proteins instead of H2A. **(C)** Fractions corresponding to octamers, tetramers and dimers depicted in Supplementary Figure S4B were analyzed by 18% SDS-PAGE and stained with Coomassie brilliant blue.
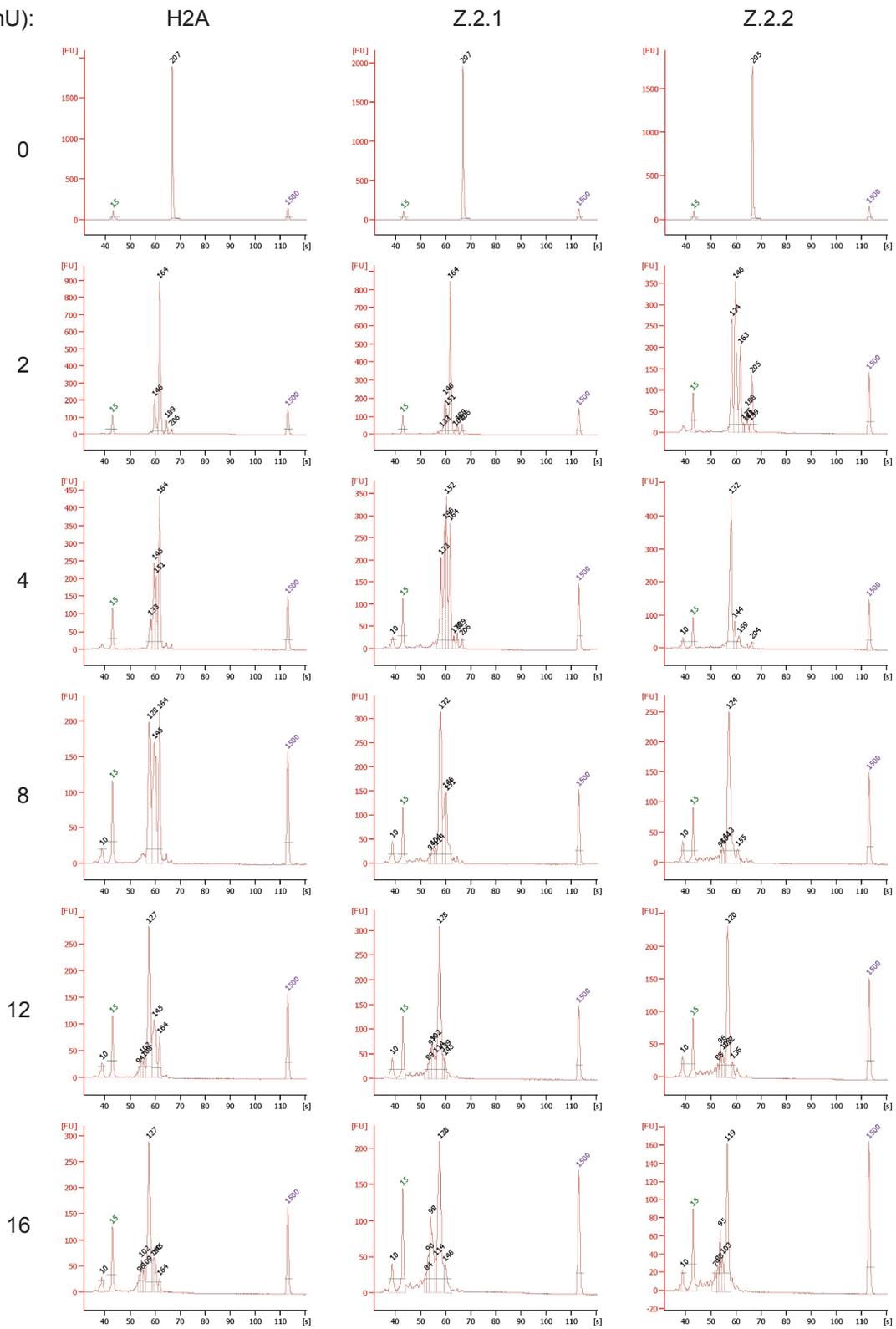
**Supplementary Figure S5: MNase digest of mononucleosomes containing different H2A variants.**
Electropherograms of deproteinized DNA fragments separated on a Bioanalyzer after digestion of mononucleosomes containing either H2A, Z.2.1 or Z.2.2 with increasing amounts of MNase (see also Figure 6C). Numbers indicate respective DNA fragment length in bp.
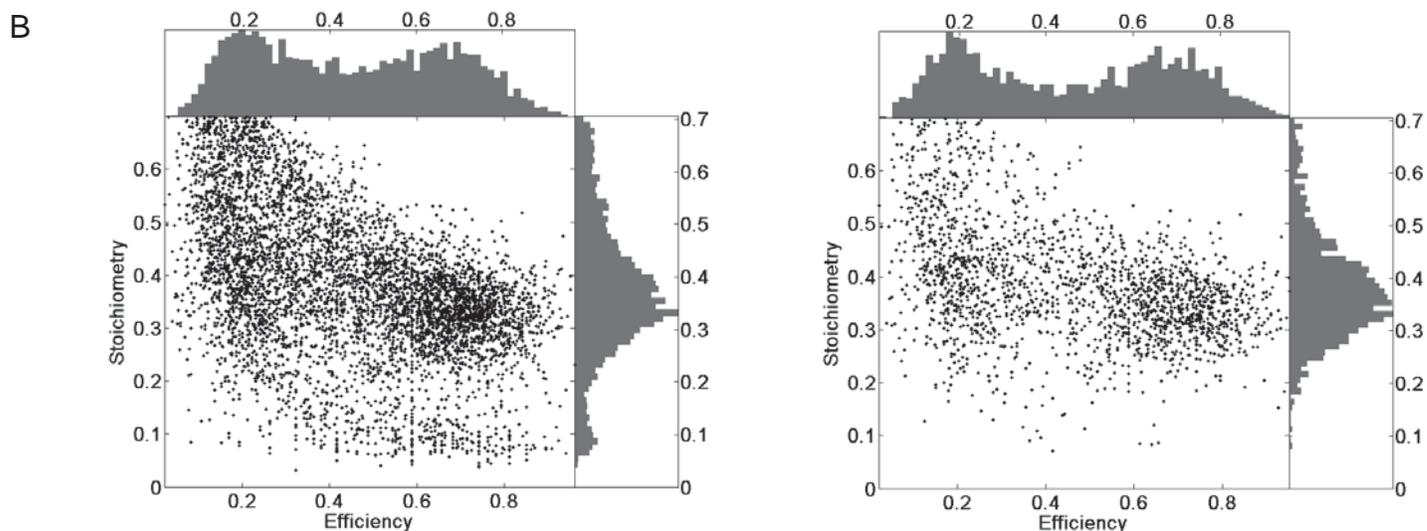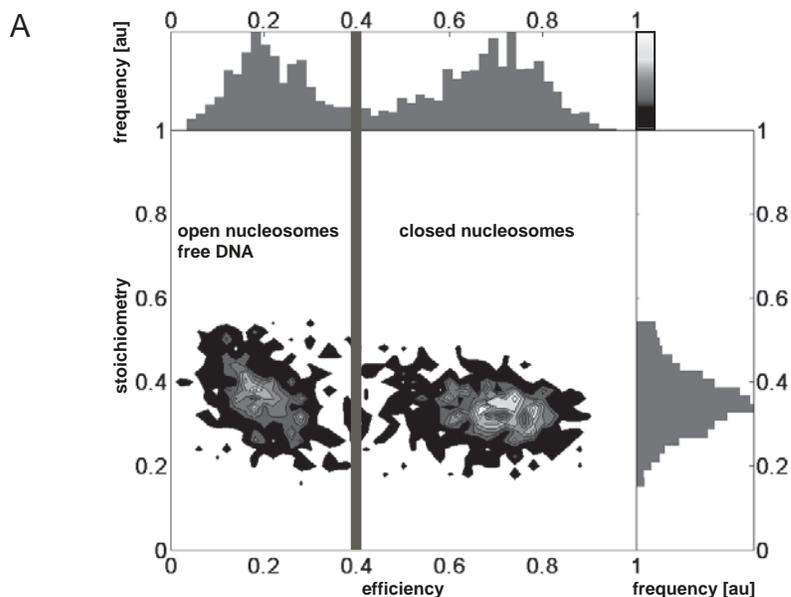
**Supplementary Figure S6: smFRET analysis of mononucleosomes containing different H2A variants.**
**(A)** Exemplary two-dimensional FRET efficiency versus stoichiometry histogram showing H2A Nucleosomes at 0 mM NaCl. Data is filtered using TDS < 1 and TDSred-PIE < 0.6 and a stoichiometry threshold (S = 0.15 – 0.55). The separation between low and high FRET at E = 0.4 (40% FRET efficiency) used to analyze the data presented here is marked by a dark grey line. **(B)** Composite of raw data, TDS and additional dynamic filter. The two dimensional plot of stoichiometry versus FRET efficiency recorded with a mononucleosome sample containing Z.2.1 (TE pH 7.6), together with the one-dimensional projections. Besides dual labeled nucleosomes, the sample contained also impurities of donor and acceptor only complexes. Only bursts within a stoichiometry range of S = 0 – 0.7 are shown. The high concentration of molecules in the sample combined with the significant amount of single labeled impurities cause strong multi-molecular trailing (left). A significant improvement of the data quality is reached by removing the multi-molecule events using TDS < 0.6 and TDSred-PIE < 0.6 (right).
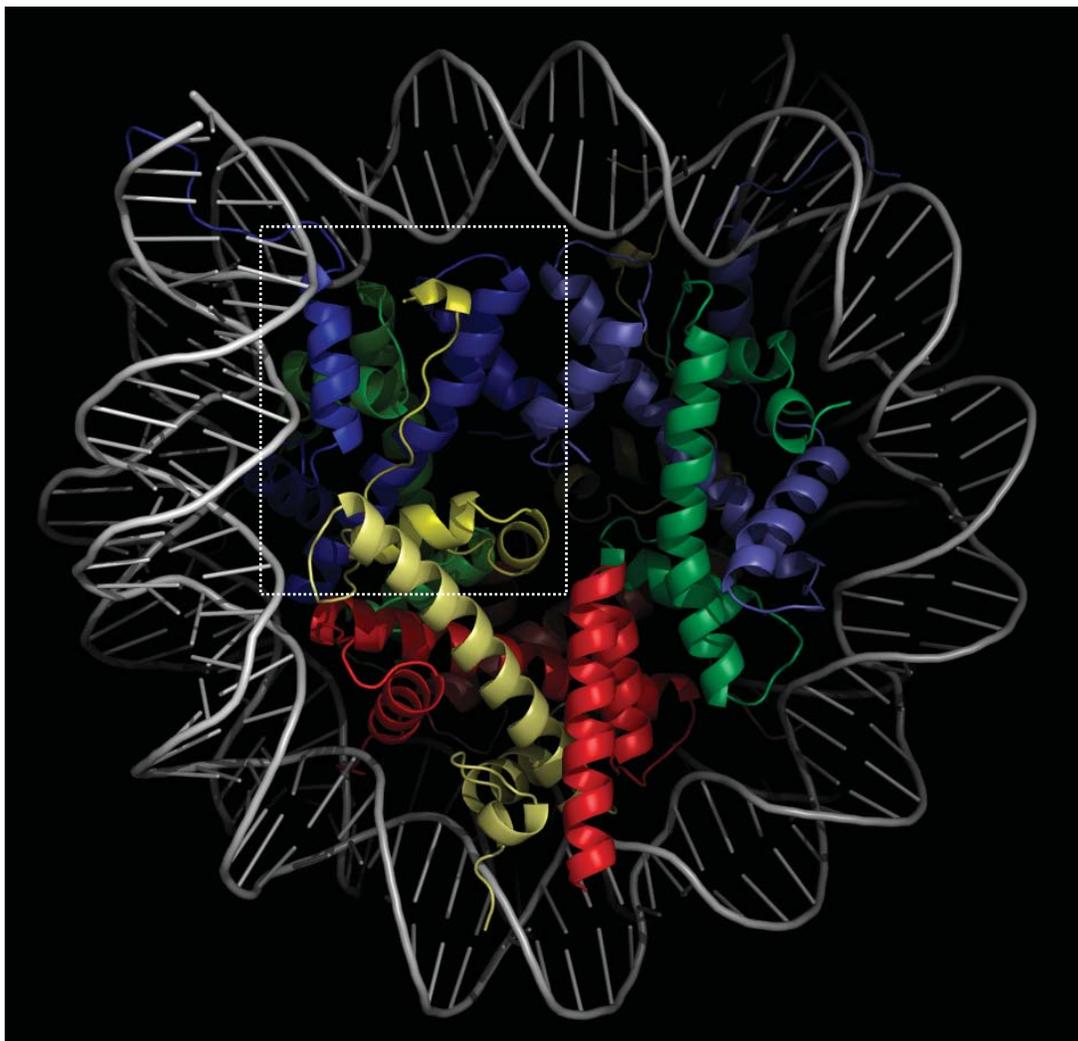
**Supplementary Figure S7: Simulation of Z.1-containing nucleosome.**
*In silico* model of nucleosome based on H2A.Z published crystal structure data (50) containing Z.1 (yellow), H2B (red), H3 (blue), H4 (green) and DNA (gray). Box highlights region depicted in Figure 8A.

**"H2A.Z.2.2 is an alternatively spliced histone H2A.Z variant that causes severe nucleosome destabilization" by Bönisch et al.**


## SUPPLEMENTARY MATERIALS AND METHODS

### Cell culture, transfection, FACS analysis and cloning

Cell lines were grown in DMEM medium (PAA) supplemented with 10% FCS (Sigma) and 1% penicillin/streptomycin or 50 µg/ml gentamicin (C127 cells) at 37°C and 5% $CO_2$. The following human cell lines were used in this study: HEK293 (embryonic kidney), HeLa (cervix carcinoma), HeLa Kyoto (cervix carcinoma), U2OS (osteosarcoma), hFB (fibroblasts), SK-N-SH (neuroblastoma), and the following mouse cell lines were used: NIH3T3 and C127. Human cell lines were transfected using FuGene HD (Roche Applied Science) and mouse C127 cells were transiently transfected using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. Z.2.1 and Z.2.2 were cloned from HeLa cDNA into pT7blue3 (Novagen). For expression in human cells, Z.2.1, Z.2.2, Bbd and deletion or domain swap mutants were cloned into the pEGFP-C1 vector (Clontech) to generate N-terminally tagged proteins. For simplicity eGFP-tagged constructs are referred to as GFP-tagged throughout the text. Plasmids coding for GFP-H2A (H2A type 1, NP_003501.1) and GFP-Bbd (H2A.Bbd type 2/3, NP_542451.1) were kindly provided by Emily Bernstein. The plasmid coding for GFP-Z.1 was a gift from Sachihiro Matsunaga, the H1.0-GFP construct was kindly given by M.J. Hendzel (1) and the GFP-HP1α construct was provided by T. Misteli (2). Stable cell lines were selected with 600 µg/ml G418 (PAA) and individual cell clones sorted by using a FACSAria machine (Becton Dickinson). Expression levels of GFP-proteins were quantified by using a FACSCanto machine (Becton Dickinson). For expression in *E. coli*, Z.2.1, Z.2.2, Bbd and domain swap mutants were cloned into the pET-21a(+) vector (Novagen). Plasmids for expression of recombinant human H2A, H2B, H3 and H4 were kindly provided by Robert Schneider. Cloning and PCR amplification accuracy was verified by sequencing (MWG).

### Histone extraction, RP-HPLC purification, cellular fractionation, MNase digestion, sucrose gradient fractionation and salt stability experiments

Acid extraction of histones was done as previously described (3). Histones from HEK293 cells were separated by RP-HPLC as previously described (4). Fractions were dried under vacuum and stored at -20°C.

Fractionation experiments were carried out as described previously (5) with minor changes. Briefly, $2 \times 10^7$ cells were resuspended in 1 ml buffer A supplemented with 0.1% NP40, incubated for 10 min and collected by centrifugation. The pellet was washed once with buffer A and incubated in buffer B for 30 min. The resulting chromatin pellet was washed once with buffer B and resuspended in SDS loading buffer (chromatin fraction). After sonication (Diagenode Bioruptor) and denaturation, nucleic acids were degraded by benzonase (VWR) treatment. All centrifugations were performed at 6.500 *g* for 5 min at 4°C except the final one (20.000 *g* for 20 min at 4°C). The soluble fraction was obtained by combining the supernatants of all centrifugations (incl. washing steps). Proteins were pelleted as described (6) and resuspended in SDS loading buffer. Identical cell equivalents of soluble and chromatin fractions were analyzed by immunoblotting.

For MNase digestion, chromatin was prepared from 5 x $10^7$ HK cells stably expressing GFP-Z.2.2 as described above, resuspended in 500 μl EX100 (10 mM HEPES pH 7.6, 100 mM NaCl, 1.5 mM $MgCl_2$, 0.5 mM EGTA, 10 % (v/v) glycerol, 1 mM DTT, 1 x Roche protease inhibitors) and digested with 1.5 U MNase (Sigma) for 20 min at 26°C. The reaction was stopped by addition of EGTA to a final concentration of 10 mM and centrifuged (3.200 *g* for 20 min at 4°C). A gradient of 5%-35% sucrose in EX100 was prepared using a Gradient Master (Biocomp), 400 μl of the supernatant after MNase digest were loaded on top and ultracentrifuged (30.000 rpm for 18 h at 4°C using a Beckmann SW41 rotor). Afterwards, 500 μl fractions were manually taken from top and analyzed for DNA content after RNase A and Proteinase K digest by agarose gel electrophoresis. Purity of mononucleosome containing fractions was verified using DNA 1000 reagents (Agilent Technologies) with the 2100 Bioanalyzer (Agilent Technologies). Pure mononucleosome fractions were combined, concentrated by TCA precipitation and analyzed by immunoblot.

For salt stability assays (7), chromatin was prepared in the same manner as in fractionation experiments. After washing with buffer B, chromatin was incubated with incubation buffer (10 mM TrisHCl pH 7.5, 1 mM DTT, 1 mM EDTA, 1x protease inhibitors (Roche Applied Science), 0.1% Triton X-100) containing different salt concentrations ranging from 50 mM to 600 mM NaCl for 1 h at room temperature. Chromatin was pelleted, solubilized in the same manner as in fractionation experiments and analyzed by immunoblotting.

**FRAP and exponential fitting**

FRAP experiments were performed using an UltraVIEW VoX spinning disk microscope system (PerkinElmer) equipped with a heated environmental chamber and $CO_2$ perfusion as previously described (4) with the following changes. To determine short-term recovery kinetics, 2D time series were recorded for 2 min to 1 h with time intervals between 0.1 s and 1 min depending on the recovery kinetics of the construct. For long-term recovery kinetics, image z-stacks were recorded with intervals between 1 min and 5 min. The central 3-5 image planes were average projected for quantitative evaluation. To correct for cell-to-cell differences in bleaching depth, the normalized mean intensity values of the first postbleach values were linearly interpolated to determine an estimated value for the time point t = 0. This value was subtracted from all mean fluorescence values after previous double normalization to correct for potential gain or loss of total fluorescence, e.g. by import and bleaching-by-acquisition.

The normalized FRAP curves were further evaluated and quantified by a commonly used fitting procedure (8): A sum of exponential time dependencies can be used to describe the intensity I(t) of the fluorescence recovery.

$$I(t) = I_0 * (1 - \sum_i A_i * e^{-\alpha_i t})$$

To discriminate at least two different species within our sample, we restricted our self to the bi-exponential case, where $A_1 + A_2 = 1$. Therefore the parameters $\alpha_1$, $\alpha_2$ and A had to be determined from the individual recovery curves by a least-square optimization algorithm. This was carried out automatically by a self-made python (www.python.org) script, which applied the *leastsq* function from the *scipy.optimize* package (www.scipy.org). Only curves with a reasonable set of resulting parameters were taken into account for the final summary, wherein the exponents $\alpha$ are transformed into their corresponding, more intuitive half-times of recovery.

$$T_1 = \frac{\ln 2}{\alpha_1}$$

Most of the recovery data were described adequately by this bi-exponential characteristic, manifested in a close approximation by the fitted curve. An evaluation of the goodness of the fits was therefore not necessary. However, in some experiments where GFP alone was expressed, a single exponential curve already allowed a sufficiently close approximation. Applying the bi-exponential model to these data sets leads to two almost identical exponentials, showing the invalidity of the more complex model. This categorization allowed us therefore to distinguish between two types of complexity: One type with only a single mobile species and on the other hand the case, where at least two mobile species are apparent. Extension to three mobile species did typically not result in significant improvement of the fits.

**Stable isotope labeling with amino acids in cell culture (SILAC) and mass spectrometry (MS) identification of H2A.Z-specific chaperone complexes**
HeLa cells expressing GFP-Z.2.1 or GFP-Z.2.2 were SILAC labeled and nuclear extracts were prepared as described before (9). Nuclear extracts were diluted in incubation buffer (150 mM NaCl, 50 mM Tris-HCl pH 8.0, 0.25% NP40) supplemented with complete protease inhibitors w/o EDTA (Roche) and 0.5 mM DTT to a concentration of 1.5 µg/µl. 400 µl solution were incubated with GFP trap (Chromotek) for 3 h at 4°C. Beads were washed 3 times with 1 ml incubation buffer, combined and eluted by boiling in loading buffer. Samples were loaded on 1D NuPAGE gels (Invitrogen), lanes cut into 8 slices, in gel digested with trypsin and desalted by stage tipping. Mass spectrometry (MS) was performed on an LTQ Orbitrap mass spectrometer as described before (10). Peptides were separated online to the mass spectrometer by using an easy nano-LC system (Proxeon Biosystems) with a 15-cm fused silica emitter with an inner diameter of 75 µm packed in house with RP ReproSil-Pur C18-AQ 3 µm resin (Dr. Maisch). Peptides were eluted with a segmented gradient from 5% to 60% B with a constant flow of 0.25 ml/min (solvent B: 80% acetonitrile, 0.5% acetic acid) over 110 min.
The MS was operated in data dependent mode. A full scan MS (m/z 300 - 1650) was acquired in the Orbitrap cell with a resolution of 60,000 at a theoretical m/z ratio of 400 after accumulation of 1,000,000 ions in the C-trap (maximum filling time of 1000 ms); the lock mass option was enabled to improve mass accuracy. The 5 most intense ions from the preview survey scan were isolated (target value of 5,000 ions at a maximum filling time of 150 ms) fragmented by collision-induced dissociation (collision energy 35 %) and measured in the ion trap concurrently to full scan acquisition in the Orbitrap. Precursor ion charge state screening was enabled, and all unassigned charge states as well as singly charged peptides were rejected. The dynamic exclusion list was set to a maximum of 500 entries with a maximum retention period of 90 s and a relative mass window of 5 ppm.
Raw data were analyzed using the MaxQuant software suite (11) (version 1.2.2.7) with the integrated Andromeda search engine (11) at default parameters using the IPI human database version 3.68 concatenated with a database containing common contaminants. For further analysis we removed contaminants and defined a ratio cutoff of 4 (Z.2.1 pull-down) and 0.25 (Z.2.2 pull-down). A complete list of all proteins identified is found in Supplementary Table S1.

**EdU replication labeling, *in situ* extraction and fluorescence microscopy to assay cell cycle dependent GFP-Z.2.2 chromatin incorporation**

C127 cells were pulse labeled for 25 min with 10 µM 5-ethynyl-2′-deoxyuridine (EdU, Baseclick) 48 h after transfection with GFP-Z.2.1 or -Z.2.2 plasmids, and immediately subjected to *in situ* extraction or fixation. For *in situ* extraction, GFP-Z.2.2 transfected cells were washed with PBS and incubated 15 sec in permeabilization buffer (0.1% Triton X-100, 150 mM NaCl in PBS). Thereafter, cells were fixed for 10 min at room temperature with 2% formaldehyde (Sigma) in PBS containing 0.1% Triton X-100. Cells expressing GFP-Z.2.1 were not *in situ* extracted and were fixed with 2% formaldehyde in PBS for 10 min at room temperature. All washing steps after fixation were performed with 0.02% Tween20 in PBS (PBST). After permeabilization with 0.2% Triton X-100 in PBS for 10 min, cells were blocked for at least 1 h in blocking buffer (2% BSA in PBST). Before EdU-detection, cells were incubated 1 h with GFP-booster (Chromotek) diluted in blocking buffer. Incorporated EdU was detected by incubating cells 30 min in 100 mM Tris-HCl pH 7, 4 mM $CuSO_4$, 20 µM azide dye Alexa Fluor 594 (Invitrogen) and 50 mM sodium ascorbate (adapted from (12)). Cells were counterstained with 200 ng/ml DAPI in PBST for 10 min and mounted on microscope slides in Vectashield mounting medium (Vector Laboratories).

Wide-field imaging was performed on a PersonalDV microscope system (Applied Precision) equipped with a 60x/1.42 PlanApo oil objective (Olympus), CoolSNAP ES2 interline CCD camera (Photometrics), Xenon illumination and appropriate filter sets. Image stacks were recorded with a z-distance of 200 nm and subjected to a constrained iterative deconvolution (enhanced ratio, 10 cycles, medium noise filtering, SoftWoRX, 3.7. imaging software package, Applied Precision).

**Expression and purification of recombinant human histone proteins in *E.coli*, *in vitro* octamer and mononucleosome reconstitution and MNase digestion of recombinant mononucleosomes**

Histones were expressed, purified and assembled into octamers as described (13). DNA for mononucleosome assembly was obtained from a pUC18 plasmid containing 25 repeats of the 601 nucleosome positioning sequence (14) kindly donated by Daniela Rhodes. After AvaI digestion, monomeric DNA was purified by gel electrophoresis and electroeluted using the Elutrap system (Whatman). Assembly of nucleosomes was performed by salt gradient deposition (13,15). For H2A and Z.2.1 nucleosomes, respective octamers and DNA were mixed in a 1:1 ratio; for Z.2.2 nucleosomes, Z.2.2-H2B dimers, $(H3-H4)_2$ tetramers and DNA were mixed in a 2:1:1 ratio. Assembly of histones on DNA was evaluated by EMSA using 5% native PAGE or native 1.5% agarose gels. Incubation of mononucleosomes for 1 h at 37°C (15) did not change position as evaluated by 5% native PAGE (data not shown). To analyze the histone content of nucleosomes after assembly, the corresponding band was excised from native 1.1% agarose gels, nucleosomes were electroeluted using the Elutrap system and protein content was analyzed by 18% SDS-PAGE after Benzonase treatment by Coomassie staining or immunoblot.

Equal amounts of nucleosomes (1 µg) were digested with different amounts of MNase (Sigma) for 10 min at 37 °C in MNase digest buffer (13.85 mM TrisHCl ph 7.5, 67 mM KCl, 10.75% Glycerol, 1 mM DTT, 5 mM $CaCl_2$). The reaction was stopped by addition of nine volumes of 5 mM EGTA. DNA was deproteinized by phenol/chloroform extraction, ethanol precipitated and analyzed using DNA 1000 reagents (Agilent Technologies) with the 2100 Bioanalyzer (Agilent Technologies).

**Single molecule Förster resonance energy transfer (smFRET)**
*Single molecule burst analysis:*
To gain information on salt dependent nucleosome stability single-molecule Förster resonance energy transfer (smFRET) measurements of dual labeled nucleosomes freely diffusing through the focal volume of a confocal microscope were performed. To this end a dual labeled 159 bp DNA was prepared using dye-labeled primers (IBA), a DNA template containing the 601 (14) sequence and six additional bases on each side together with dNTPs (Finnzymes) and the Phusion DNA-polymerase (Finnzymes) in a PCR reaction. Dye labels were at position 65 (donor dye, Tamra) and position 20 on the reverse strand (acceptor dye, Alexa647). A mixture of this labeled DNA and unlabeled DNA (molar ratio of 1:50) was used for mononucleosome assembly (see above).

It is well known, that ultra-low concentrations of nucleosomes are prone to become instable in typical experimental geometries (16). To minimize such effects measurements were performed using a 1:250 mixture of double labeled to unlabeled nucleosomes at a total concentration of 25 nM in commercially available TE buffer (Sigma-Aldrich, pH 7.6) containing 10 mM DTT and 0 mM, 300 mM, 400 mM, 500 mM, 600 mM and 700 mM NaCl, respectively. The samples were incubated at the respective salt concentration for 1 hour at 21°C before a drop of 20 μl was put onto cover slips (Marienfeld) for data collection. The cover slips were cleaned with 2% Hellmanex III (Hellma) and water prior to silanization for 15 min with 2% (3-aminopropyl)triethoxysilane (Sigma-Aldrich) in Acetone, and coating with 40 mg/100 μl polyethylene glycol (mPEG-SVA MW 5000, Laysan Bio Inc.) in ddH$_2$O for 1 h.

The confocal measurements were performed on a custom built experimental setup using pulsed interleaved excitation (PIE) (17) with lasers at 532 nm (Pico-TA-Picoquant, power before the objective 80μW) and 640 nm (LDH-D-C-640, Picoquant, power before the objective 80 μW) at a repetition rate of 26.66 MHz. The fluorescence was separated for polarization and color and detected on four avalanche-photo-diodes (green channel AQRH-14, red channel AQR-16, Perkin Elmer). Photon arrival times were recorded using four single-photon-counting-modules (Becker&Hickl SPC-150) and data was processed using custom software written in MATLAB (MathWorks). Since the anisotropy of the molecules was not of importance for this study, photons of identical wavelength but different polarization were merged into one detector channel. Data were collected for 10 min and the collected photons were sorted into three different channels, namely donor detection after direct excitation (green), acceptor detection after direct excitation (red) and acceptor detection after donor excitation (fret). Labeled complexes diffusing through the focal volume of the microscope resulted in bursts of detected photons. An all photons burst search (APBS) was applied with the criteria of detecting at least 3 photons within 500 μs with a total of at least 60 photons per burst (18). From the photon bursts, the Stoichiometry (*S*) and FRET Efficiency (*E*) were calculated (including the predetermined correction factor) according to:

$$E = \frac{GR}{GR + \gamma * GG} \text{ , and } S = \frac{GR + \gamma * GG}{GR + \gamma * GG + RR}$$

Where *GR* are the red photons after green excitation, *RR* are the red photons after red excitation, *GG* are the green photons after green excitation and *y* is a factor correcting for the different efficiencies of the red and green detection channels (19). Multi-molecular events were removed from the data as described below using *TDS < 1* and *TDS$_{red-PIE}$* < 0.6 for all complexes not showing molecular dynamics. Remaining

donor-only and acceptor-only bursts were removed using a stoichiometry threshold (S = 0.15 – S = 0.55).

Due to the chosen labeling positions on the nucleosome, closed nucleosomes have the donor and acceptor dyes positioned adjacent to each other leading to a high FRET state with an efficiency of ~80% while open or incomplete nucleosomes show a very low FRET signal. The fraction of closed nucleosomes was quantified for each salt concentration by analyzing how many of the detected fluorescence bursts have $E >$ 40% (Supplementary Figure S6A). The data was normalized to the fraction of closed molecules at 0 mM NaCl to allow for a comparison of the salt dependence for the three investigated samples (H2A, Z.2.1, Z.2.2).

*Removal of multi-molecular events*

Nucleosomes are prone to become instable at low concentrations as well as when interacting with surfaces. Thus in order to avoid artifacts the duration of the experiment has to be minimized. To address this difficulty relatively high nucleosome concentrations were used in the experiments to ensure that the occurrence of multi-complex bursts is not negligible. For a homogeneous population with only a single FRET species this is not a problem, however if several FRET states exist, multi-molecule events of different species will alter the determined FRET values. Moreover, impurities such as complexes labeled with only donor or acceptor observed simultaneously with double-labeled complexes will also lead to changes in FRET efficiencies. However, as two independently diffusing complexes do not enter and exit the excitation volume exactly at the same time it is possible to differentiate these multi-molecular events from single molecule events and to exclude them from further analysis. Independently diffusing molecules involved in a multi-molecule event will yield different values for the mean-macro-time (i.e. the time where 50% of the respective photons have arrived) for all photons of a burst, as compared to that for the photons of one color. We therefore calculated the characteristic Time-Deviation-Signal (TDS) defined as

$$TDS = \left( (D_{total} - D_{green}) + \left| T_{total} - T_{green} \right| \right) * (1 - P) + \left( (D_{total} - D_{fret}) + \left| T_{total} - T_{fret} \right| \right) * P * \gamma$$
**(1)**

where $D_x$ is the burst duration, $T_x$ is the mean-macro-time, $y$ is a factor correcting for the different efficiencies of the red and green detection channels and $P$ is the proximity ratio given by the number of photons in the burst as $P = GR/(GR+GG)$.

In eq. 1 we compute the TDS of the green and red channels simultaneously and adjust the relative value to the percentage of photons detected. In addition, multi-molecule events containing low (0%) FRET and donor only complexes (which cannot be found using eq. 1) can also be determined using a PIE setup and defining:

$$TDS_{red-PIE} = (D_{total} - D_{red}) + \left| T_{total} - T_{red} \right| \quad \textbf{(2)}$$

To demonstrate the capabilities of this analysis scheme a sample containing Z.2.1 nucleosomes (TE pH 7.6, 0 mM NaCl) as well as impurities of donor only and acceptor only complexes was measured. In order to stress the discussed effects for this control experiment, the concentration of labeled complexes was increased to ~150 pM. By using thresholds of *TDS* < 1 and *TDS_{red-PIE}* < 0.6 in the TDS parameter space we are able to remove most of the observed trailing (i.e. events with high *S* and medium to low *E*) caused by multi-molecule events as well as photo-physics and receive a distribution showing populations of distinct FRET efficiencies and stoichiometries (Supplementary Figure S6B).

**REFERENCES**

1. Th'ng, J.P., Sung, R., Ye, M. and Hendzel, M.J. (2005) H1 family histones in the nucleus. Control of binding and localization by the C-terminal domain. *J Biol Chem*, **280**, 27809-27814.
2. Cheutin, T., McNairn, A.J., Jenuwein, T., Gilbert, D.M., Singh, P.B. and Misteli, T. (2003) Maintenance of stable heterochromatin domains by dynamic HP1 binding. *Science*, **299**, 721-725.
3. Shechter, D., Dormann, H.L., Allis, C.D. and Hake, S.B. (2007) Extraction, purification and analysis of histones. *Nat Protoc*, **2**, 1445-1457.
4. Wiedemann, S.M., Mildner, S.N., Bonisch, 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.
5. Wysocka, J., Reilly, P.T. and Herr, W. (2001) Loss of HCF-1-chromatin association precedes temperature-induced growth arrest of tsBN67 cells. *Mol Cell Biol*, **21**, 3820-3829.
6. Wessel, D. and Flugge, U.I. (1984) A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal Biochem*, **138**, 141-143.
7. Zhang, H., Roberts, D.N. and Cairns, B.R. (2005) Genome-wide dynamics of Htz1, a histone H2A variant that poises repressed/basal promoters for activation through histone loss. *Cell*, **123**, 219-231.
8. Phair, R.D., Gorski, S.A. and Misteli, T. (2004) Measurement of dynamic protein binding to chromatin in vivo, using photobleaching microscopy. *Methods Enzymol*, **375**, 393-414.
9. Vermeulen, M., Eberl, H.C., Matarese, F., Marks, H., Denissov, S., Butter, F., Lee, K.K., Olsen, J.V., Hyman, A.A., Stunnenberg, H.G. *et al.* (2010) Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell*, **142**, 967-980.
10. Vermeulen, M., Mulder, K.W., Denissov, S., Pijnappel, W.W., van Schaik, F.M., Varier, R.A., Baltissen, M.P., Stunnenberg, H.G., Mann, M. and Timmers, H.T. (2007) Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell*, **131**, 58-69.
11. Cox, J. and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol*, **26**, 1367-1372.
12. Salic, A. and Mitchison, T.J. (2008) A chemical method for fast and sensitive detection of DNA synthesis in vivo. *Proc Natl Acad Sci U S A*, **105**, 2415-2420.
13. Dyer, P.N., Edayathumangalam, R.S., White, C.L., Bao, Y., Chakravarthy, S., Muthurajan, U.M. and Luger, K. (2004) Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol*, **375**, 23-44.
14. Thastrom, A., Lowary, P.T., Widlund, H.R., Cao, H., Kubista, M. and Widom, J. (1999) Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences. *J Mol Biol*, **288**, 213-229.
15. 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.

16.    Gansen, A., Valeri, A., Hauger, F., Felekyan, S., Kalinin, S., Toth, K., Langowski, J. and Seidel, C.A. (2009) Nucleosome disassembly intermediates characterized by single-molecule FRET. *Proc Natl Acad Sci U S A*, **106**, 15308-15313.

17.    Muller, B.K., Zaychikov, E., Brauchle, C. and Lamb, D.C. (2005) Pulsed interleaved excitation. *Biophys J*, **89**, 3508-3522.

18.    Nir, E., Michalet, X., Hamadani, K.M., Laurence, T.A., Neuhauser, D., Kovchegov, Y. and Weiss, S. (2006) Shot-noise limited single-molecule FRET histograms: comparison between theory and experiments. *J Phys Chem B*, **110**, 22103-22124.

19.    Lee, N.K., Kapanidis, A.N., Wang, Y., Michalet, X., Mukhopadhyay, J., Ebright, R.H. and Weiss, S. (2005) Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation. *Biophys J*, **88**, 2939-2953.

# ATRX-mediated chromatinassociation of histone variant macroH2A1 regulates α-globin expression

Kajan Ratnakumar, Luis F. Duarte, Gary LeRoy, Dan Hasson,
Daniel Smeets, Chiara Vardabasso, **Clemens Bönisch**, Tianying Zeng,
Bin Xiang, David Y. Zhang, Haitao Li, Xiaowo Wang, Sandra B. Hake,
Lothar Schermelleh, Benjamin A. Garcia, and Emily Bernstein

RESEARCH COMMUNICATION

# ATRX-mediated chromatin association of histone variant macroH2A1 regulates α-globin expression

Kajan Ratnakumar,[1,2] Luis F. Duarte,[1,2] Gary LeRoy,[3,10] Dan Hasson,[1,4,10] Daniel Smeets,[5,10] Chiara Vardabasso,[1,2] Clemens Bönisch,[6] Tianying Zeng,[7] Bin Xiang,[8] David Y. Zhang,[9] Haitao Li,[8] Xiaowo Wang,[7] Sandra B. Hake,[6] Lothar Schermelleh,[5,11] Benjamin A. Garcia,[3] and Emily Bernstein[1,2,12]

[1]Department of Oncological Sciences, [2]Department of Dermatology, Mount Sinai School of Medicine, New York, New York 10029, USA; [3]Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544, USA; [4]Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, New York 10029, USA; [5]Department of Biology II, Ludwig Maximilians University Munich, 82152 Martinsried, Germany; [6]Munich Center for Integrated Protein Science, Adolf-Butenandt Institute, Ludwig-Maximilians University, 80336 Munich, Germany; [7]MOE Key Laboratory of Bioinformatics, Bioinformatics Division, TNLIST, Department of Automation, Tsinghua University, Beijing 100084, China; [8]Center for Structural Biology, School of Life Sciences, School of Medicine, Tsinghua University, Beijing 100084, China; [9]Department of Medicine, Mount Sinai School of Medicine, New York, New York 10029, USA

The histone variant macroH2A generally associates with transcriptionally inert chromatin; however, the factors that regulate its chromatin incorporation remain elusive. Here, we identify the SWI/SNF helicase ATRX (α-thalassemia/MR, X-linked) as a novel macroH2A-interacting protein. Unlike its role in assisting H3.3 chromatin deposition, ATRX acts as a negative regulator of macroH2A's chromatin association. In human erythroleukemic cells deficient for ATRX, macroH2A accumulates at the *HBA* gene cluster on the subtelomere of chromosome 16, coinciding with the loss of α-globin expression. Collectively, our results implicate deregulation of macroH2A's distribution as a contributing factor to the α-thalassemia phenotype of ATRX syndrome.

Supplemental material is available for this article.

The replacement of canonical histones with histone variants contributes to the dynamic nature of chromatin.

Due to amino acid differences and, in turn, unique post-translational modifications, histone variants can alter nucleosome structure, stability, and binding of effector proteins. Histone variants have unique genomic localization patterns, and thus specialized roles such as regulating gene expression or chromosome segregation during cell division (Banaszynski et al. 2010). Therefore, the differential genomic incorporation of histone variants directly impacts critical cellular functions.

The histone variant macroH2A (mH2A) is a vertebrate-specific member of the H2A family and is unusual due to the presence of a C-terminal macro domain (Pehrson and Fried 1992). Two different genes encode mH2A1 and mH2A2 (*H2AFY* and *H2AFY2*, respectively), and two splice forms of mH2A1 exist: mH2A1.1 and mH2A1.2 (Costanzi and Pehrson 2001). mH2A is abundant in heterochromatin, including senescence-associated heterochromatic foci (SAHF) and the inactivated X chromosome (Xi) (Costanzi and Pehrson 1998; Zhang et al. 2005). In vitro studies suggest that the macro domain sterically hinders access of transcription factors to DNA, while mH2A's L1 loop produces inflexible nucleosomes (Angelov et al. 2003; Chakravarthy et al. 2005).

Our group has recently demonstrated a role for mH2A isoforms in suppressing melanoma progression, and others have linked mH2A expression or its splice patterns to breast and lung cancer (Sporn et al. 2009; Kapoor et al. 2010; Novikov et al. 2011). However, the factors that regulate the association of mH2A with chromatin remain obscure. Therefore, identifying regulators of the incorporation of histone variants at distinct genomic loci is key to understanding how chromatin domains are established and maintained and how these may go awry in disease.

A second group of factors contributing to chromatin dynamics are ATP-dependent chromatin remodeling complexes that rearrange or mobilize nucleosomes. Deregulation of members of the SWI/SNF family is implicated in various cancers and mental retardation (MR) syndromes, including ATRX (α-thalassemia/MR, X-linked), (Wilson and Roberts 2011). Mutations in ATRX, predominantly found in the H3K9me3-binding ADD (ATRX–DNMT3–DNMT3L) and/or helicase domains, are associated with ATRX syndrome (Higgs et al. 2005; Iwase et al. 2011). This syndrome is characterized by MR and α-thalassemia—a loss of α-globin gene production (Higgs et al. 2005). However, the mechanisms by which *HBA* (hemoglobin α) gene repression occurs are unknown (Higgs et al. 2005).

In addition to its role in regulating gene expression, ATRX acts in concert with Daxx to deposit the H3 variant H3.3 specifically at telomeres (Drane et al. 2010; Goldberg et al. 2010; Lewis et al. 2010), and ATRX deficiency results in loss of telomere integrity (Goldberg et al. 2010; Wong et al. 2010; Heaphy et al. 2011). However, it remains unclear how loss of functional ATRX protein affects the global chromatin landscape of ATRX patients, which may have tissue-specific effects (Berube 2011).

Here, we sought to discover factors involved in regulating mH2A's chromatin association. By isolating mH2A in its chromatin-free state, we identified ATRX as a novel mH2A partner. Unlike H3.3, mH2A does not interact with Daxx in chromatin-free extracts, suggesting that these two variants interact with unique ATRX complexes. As such, we observed a mutual exclusion between mH2A1.2 and

Ratnakumar et al.

H3.3 in the nucleosome. We further demonstrate that ATRX negatively regulates mH2A1 chromatin incorporation. Loss of ATRX results in increased mH2A1 levels at telomeres, as well as at the α-globin locus in erythroleukemic cells, concomitant with reduced transcription of the *HBA* genes. These data implicate dysregulation of mH2A's chromatin incorporation as a novel facet of the α-thalassemia phenotype of ATRX syndrome.

## Results and Discussion

### mH2A interacts with ATRX in a chromatin-free cellular fraction

To identify factors involved in the regulation of mH2A chromatin association, we reasoned they would associate in the soluble nuclear and/or cytoplasmic fractions ("chromatin-free"). Due to their different cellular localization patterns (Fig. 1A; Costanzi and Pehrson 1998) and differential mobility in chromatin, as assayed by fluorescence recovery after photobleaching (FRAP) analysis (Supplemental Fig. 1), we hypothesized that mH2A and H2A have unique regulatory factors. To this end, we employed a large-scale biochemical strategy to purify chromatin-free GFP-H2A or mH2A1.2-GFP, similar to that used for the identification of the CENPA chaperone HJURP (Fig. 1B; Foltz et al. 2009). Confirmation of immunoprecipitated histones was performed by immunoblotting (Fig. 1C). Following extensive washing and TCA precipitation of the entire immunoprecipitated material, proteins were resolved on a gradient gel, both lanes were excised (10



**Figure 1.** Identification of mH2A1.2 chromatin-free interacting factors. (*A*) Fluorescence microscopy of HEK293 cells stably expressing GFP-H2A and mH2A1.2-GFP. Arrowhead indicates Xi. (*B*) Procedure used to isolate chromatin-free H2A- and mH2A1.2-interacting factors. (*C*) αGFP immunoblot confirms expression and immunoprecipitation of histones in stable cell lines. (*D*) Immunoblots of Imp9, Npm, and NAP1 association with GFP-H2A or mH2A-GFP isoforms. Arrows on all αGFP blots indicate GFP-H2A (*bottom*) and mH2A-GFP (*top*), and asterisks indicate degradation products.

slices per lane), and mass spectrometry (MS) analysis on all gel slices was performed (Fig. 1B; Supplemental Fig. 2A).

From the proteins retrieved (Supplemental Table 1), we focused on factors that regulate chromatin association of histones, including nuclear import factors (Supplemental Fig. 2B). For example, kap114p mediates the nuclear import of H2A/H2B in *Saccharomyces cerevisiae* (Mosammaparast et al. 2005), and accordingly, we identified the mammalian homolog Importin 9 (Imp9) as an H2A import factor. We validated Imp9 by immunoblots and independent MS experiments where specific bands were excised (Fig. 1D; Supplemental Fig. 2C,D). These data suggest that the mechanism of histone import is evolutionarily conserved and, importantly, validate our technical approach.

As we were interested in factors that directly regulate chromatin association of mH2A1, we further focused on histone chaperones (De Koning et al. 2007; Park and Luger 2008). We identified peptides from NAP1, SET/TAF-I, nucleolin (Ncl), and nucleophosmin (Npm) in our MS analysis (Supplemental Figs. 2A, 3A). Some of these factors, such as Ncl and Npm, were present in both the GFP-H2A and mH2A1.2-GFP immunoprecipitations, and we reasoned they were general histone-interacting proteins. Indeed, this is the case (Dunleavy et al. 2009; Gaume et al. 2011), as confirmed by our immunoblots (Fig. 1D). We further confirmed specificity of NAP1 for H2A by immunoblot (Fig. 1D), as reported (Park and Luger 2008). Due to the lack of specificity for mH2A, these factors were unlikely candidates for regulating its chromatin association.

Of the potential histone chaperones identified by MS, the SWI/SNF chromatin remodeling protein ATRX interacted uniquely with mH2A1.2 (Supplemental Fig. 3A). Nineteen peptides spanning the entire ATRX protein were identified (Supplemental Fig. 3B). We confirmed this interaction using the immunoprecipitation protocol as performed for MS analysis (Fig. 2A) and via an alternative chromatin-free approach with similar results (Supplemental Fig. 3C; Mendez and Stillman 2000). Of note, we did not detect H3 in the immunoprecipitation from either protocol, suggesting that the tagged histones in our extracts were indeed chromatin-free (nonnucleosomal) and that the mH2A–ATRX interaction is independent of H3 binding (Fig. 2A). H2B peptides were detected via MS, suggesting the presence of (m)H2A–H2B dimers.

We next examined the interaction between ATRX and mH2A isoforms in chromatin-free extracts. ATRX interacts with all mH2A isoforms (Fig. 2B), suggesting the interaction occurs through the highly conserved H2A domain. As expected, Imp9 preferentially associated with H2A, and Parp1 uniquely interacted with mH2A1.1, as previously described (Timinszky et al. 2009). We then performed reverse coimmunoprecipitation (co-IP) experiments in HEK293 cells transfected with GFP-tagged ATRX fragments (Supplemental Fig. 4). Using whole-cell extracts, we narrowed down the region required for mH2A1 binding to amino acids 1–841 of ATRX (Fig. 2C). As positive controls, this fragment also bound H3, while a construct spanning the middle region of ATRX (800–1670) bound Daxx (Tang et al. 2004). We next validated this interaction in chromatin-free immunoprecipitations (Fig. 2D). Interestingly, the N terminus of ATRX contains an ADD domain, an HP1-binding "PxVxL" motif, an α-helical region, and an acid-rich motif (B Xiang and H Li, unpubl.). It will be key to decipher whether mH2A–ATRX binding is direct and, if so, which domains mediate this interaction.
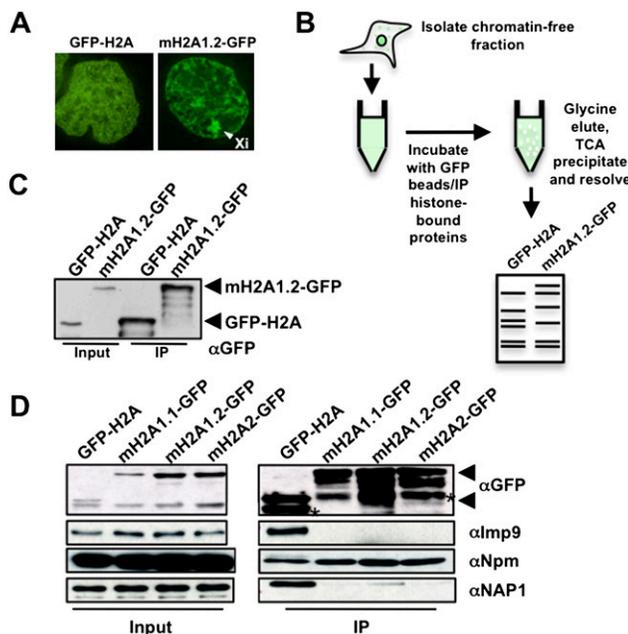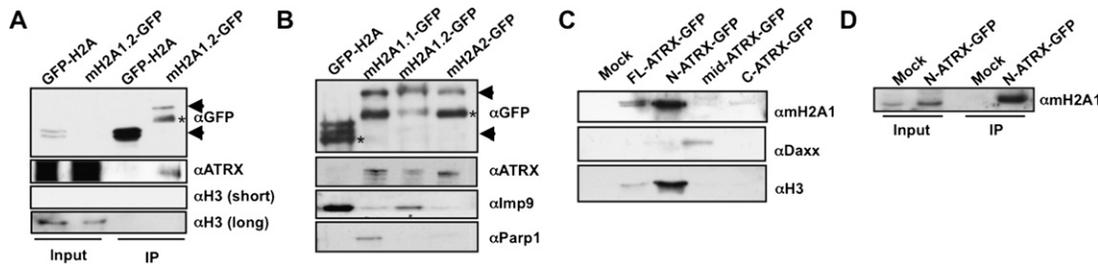
**Figure 2.** ATRX interacts with mH2A isoforms in chromatin-free extracts. (*A,B*) Immunoblots of GFP and ATRX from chromatin-free extracts. The absence of H3 in immunoprecipitations confirms chromatin-free interactions; see the long exposure. (*B*) Co-IP of GFP-H2A and all mH2A isoforms in chromatin-free extracts for ATRX. Immunoblots for Imp9 and Parp1 demonstrate chromatin-free interactions. (*C*) Whole-cell immunoprecipitations of GFP-tagged ATRX constructs (full-length, N-terminal [1–841], middle region [800–1670], and C-terminal [1670–2492]), followed by mH2A1, H3, and Daxx immunoblots. (*D*) Chromatin-free co-IP of N-ATRX-GFP with mH2A1.

### mH2A and H3.3 are in distinct ATRX complexes

While ATRX deposits H3.3 at telomeres via its interaction with Daxx (Drane et al. 2010; Goldberg et al. 2010; Lewis et al. 2010), we were unable to detect an interaction between mH2A isoforms and Daxx via immunoblot (Fig. 3A), and Daxx peptides were not detected in our MS analysis (Supplemental Table 1). However, Daxx indeed interacts specifically with H3.3 in chromatin-free extracts (Fig. 3B). These results suggest that H3.3 and mH2A are in distinct ATRX complexes, distinguished by the presence of Daxx.

Based on the above, we hypothesized that mH2A1 and H3.3 exist in mutually exclusive nucleosomes. To test this directly, we immunoprecipitated mononucleosomes from H2A- and mH2A1.2-Flag-tagged HeLa cells (Fig. 3C) and performed MS analysis to determine their H3 variant composition. While H2A-containing nucleosomes contain >40% H3.3, those of mH2A1.2 contain ~4% (Fig. 3D). This suggests that distinct factors or complexes regulate chromatin association of mH2A1 and H3.3, and we hypothesized that while ATRX–Daxx deposits H3.3 into chromatin, ATRX inhibits mH2A chromatin incorporation.

### Loss of ATRX results in altered levels of mH2A1 in chromatin

To test this hypothesis, we probed the effects of ATRX depletion on mH2A chromatin association. We engineered HEK293 cells to stably express shRNAs targeting ATRX or luciferase (control). We selected two shRNA lines that induced significant knockdown (sh90 and sh92) (Fig. 4A; Supplemental Fig. 5A) for further experiments and ensured that mH2A levels were unaffected (Fig. 4A). This knockdown may mimic ATRX syndrome, as patients with ATRX mutations have reduced protein levels or decreased enzymatic activity of the helicase domain (Berube 2011; Mitson et al. 2011).

We next inquired whether loss of ATRX altered mH2A chromatin association. Loss of ATRX resulted in a global increase of mH2A1 in chromatin, while total cellular levels remained constant (Fig. 4B), implicating ATRX as a negative regulator of mH2A's chromatin incorporation. Quantitative MS (qMS) analysis on histones extracted from chromatin (Plazas-Mayorca et al. 2009) of shluc and sh92 HEK293 cells revealed that loss of ATRX results in ~30% more mH2A1 in chromatin (Fig. 4C; Supplemental Fig. 6). To examine the dynamics of mH2A1.2-GFP in the absence of ATRX in vivo, we performed FRAP using stable

shRNA lines generated in HeLa1.2.11 cells (Supplemental Fig. 5B,C). FRAP studies revealed a decrease in fluorescence recovery of mH2A1.2 in sh92 cells, suggesting a more stable association of mH2A1.2 with chromatin upon loss of ATRX (Fig. 4D).

### ATRX regulates mH2A1 incorporation at telomeres and the α-globin cluster

As mH2A1 chromatin association increased in ATRX-depleted cells, we inquired which genomic regions are enriched in mH2A1. Because ATRX localizes to telomeres (Goldberg et al. 2010; Wong et al. 2010), we hypothesized that global increase of mH2A1 might, in part, be a result of telomeric accumulation. Chromatin immunoprecipitation (ChIP) followed by Southern blot demonstrated increased association of mH2A1 with telomeres in ATRX knockdown lines of HEK293 and erythroleukemic K562 cells, which express α-globin (Fig 5B; Supplemental Fig. 7; see below). The weaker of the two ATRX knockdowns
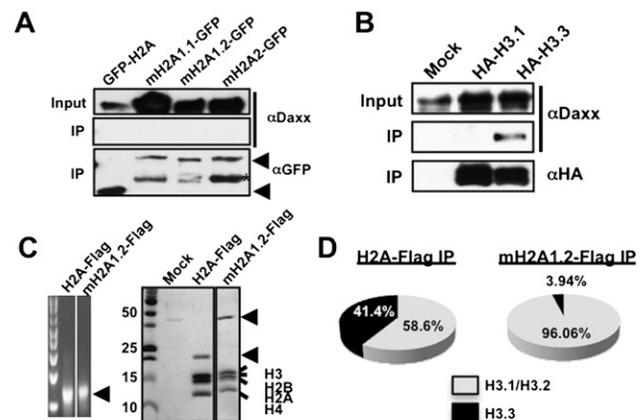


**Figure 3.** mH2A and H3.3 are in distinct ATRX complexes. Co-IPs of chromatin-free association of H2A and mH2A isoforms (*A*) or H3 variants (*B*) with Daxx. Immunoblots detected the presence of Daxx specifically with H3.3. (*C*) Ethidium bromide-stained (*left*) and Coomassie-stained (*right*) mononucleosomes from H2A-Flag and mH2A1.2-Flag immunoprecipitations. (*D*) H3 composition of H2A-Flag and mH2A1.2-Flag mononucleosomes as analyzed by MS. The pie chart depicts the abundance of H3.1/H3.2 (gray) and H3.3 (black) in immunoprecipitated nucleosomes.
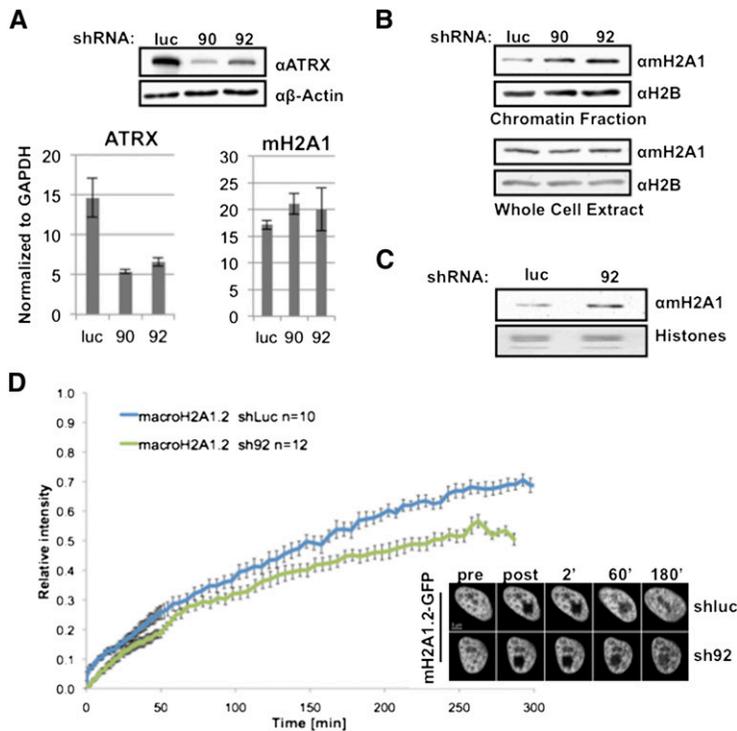
Ratnakumar et al.



**Figure 4.** ATRX knockdown results in increased levels and stability of mH2A1 in chromatin. (*A*) shRNA-mediated knockdown of ATRX (sh90 and sh92) in HEK293 cells results in the loss of ATRX protein and mRNA, compared with shluc, without affecting mH2A1 mRNA. β-Actin was used for loading. (*B*) Loss of ATRX results in accumulation of mH2A1 in chromatin (*top* panel), and the whole-cell content remains unaffected (*bottom* panel). H2B was used for loading. (*C*) mH2A1 immunoblot of chromatin-extracted histones analyzed by qMS. (*D*, *left*) Quantitation of FRAP experiments indicates slower recovery of mH2A1.2-GFP in HeLa1.2.11 cells depleted of ATRX (sh92; green line) compared with control (shluc; blue line). (*Right*) Representative images of FRAP time series pre- and post-bleach are shown.

(sh92) revealed >10-fold mH2A1 enrichment in K562 cells (Fig. 5B), suggesting partial loss of function may have significant consequences, possibly akin to the syndrome (Berube 2011; Mitson et al. 2011). While currently unclear, ATRX-mediated regulation of mH2A levels at telomeres may help to ensure telomeric integrity.

Next, we queried whether loss of ATRX affects expression of the subtelomeric α-globin gene cluster on human chromosome 16 in erythroid cells. Upon ATRX knockdown in K562 cells (Fig. 5A; Supplemental Fig. 8A), HBA mRNA and protein levels were dramatically reduced (Fig. 5C). Other genes in this region, including the ATRX target *NME4* (Law et al. 2010), were also transcriptionally decreased, while *CDK8* on chromosome 13 was unaffected (Supplemental Fig. 8B; Kapoor et al. 2010).

As mH2A is generally transcriptionally repressive, we hypothesized that deregulation of mH2A nucleosome occupancy represses the *HBA* genes, which are silenced in ATRX patients by undefined mechanisms (Higgs et al. 2005; Berube 2011). To examine mH2A1 distribution across the α-globin cluster, we performed native ChIP-seq in shluc and sh92 K562 cells (Supplemental Fig. 9A). We obtained 56,540,184 reads for shluc, 67,219,237 for sh92, and 148,165,330 for input DNA using Illumina Hi-Seq (Supplemental Fig. 9B). Analyses were performed on normalized alignments (to the total number of alignments) to

account for the different number of reads between samples. We found mH2A1 to be generally (1) excluded from transcriptional start sites (TSSs) and (2) in broad domains both upstream of and downstream from the TSS, particularly at genes transcribed at low levels. This supports its role as a repressive variant and is consistent with ChIP–chip studies (Supplemental Fig. 10A; Buschbeck et al. 2009; Gamble et al. 2010).

While K562 cells express α-globin, levels are lower than primary erythroblasts (D Higgs, pers. comm.). In accordance, we observed a distinct mH2A1 domain at the α-globin cluster in K562 cells, however, with more significant peaks of enrichment in sh92 cells (Fig. 5D; Supplemental Fig. 11A). These data strongly suggest that mH2A1 is enriched at this gene cluster in the absence of ATRX. Globally, the total number of base pairs covered by mH2A1 significant peaks is 20% higher in sh92 cells than shluc cells. This is likely due to mH2A1 redistribution, as only ~45% of the peaks are shared between shluc and sh92 (Supplemental Fig. 9B,C). Interestingly, we observed a global anti-correlation of mH2A1 domains and ATRX peaks, which are generally concentrated around TSSs (Fig. 5D; Supplemental Fig 10B,C; Law et al. 2010), suggesting that ATRX prevents mH2A1 chromatin incorporation. Finally, by quantitative PCR (qPCR) analysis of native ChIP DNA and cross-linked ChIP DNA, we observed marked increase at regions in the α-globin cluster previously reported to be mH2A1-enriched (Supplemental Fig. 11B,C; Gamble et al. 2010). Taken together, these data suggest that mH2A1 is specifically deposited at the α-globin gene cluster in an ATRX-mediated fashion.

Here, we took an unbiased approach to identify factors that specifically associate with mH2A in its chromatin-free state. We identify ATRX as a negative regulator of mH2A1 chromatin incorporation, particularly at telomeres and the α-globin locus. Such regulation of histone incorporation via inhibitory factors remains relatively unexplored. A recent study identified INO80, also an ATP-dependent chromatin remodeling enzyme, as a negative regulator of H2A.Z nucleosomal incorporation (Papamichos-Chronakis et al. 2011). In the absence of INO80, genomic distribution of H2A.Z is perturbed, resulting in a reduced response to transcriptional changes. Whether ATRX directly interacts with mH2A to evict this histone variant from chromatin or inhibit its deposition remains unclear. If the interaction is direct, dissecting the surfaces that mediate binding will be important. In addition, it remains formally possible that alternatively spliced or modified forms of ATRX differentially interact with mH2A and H3.3 (Berube et al. 2000; Garrick et al. 2004). We look forward to future studies addressing the positive and negative regulation of histone variants within the chromatin template, the factors involved, and the underlying mechanisms.

Finally, our data point toward a novel mechanism by which the histone variant mH2A1 is involved in the α-thalassemia phenotype of ATRX patients. While the ATRX–Daxx complex has been shown to deposit H3.3, the genomic localization and function of H3.3 have yet to be explored in the context of ATRX syndrome. Here, we
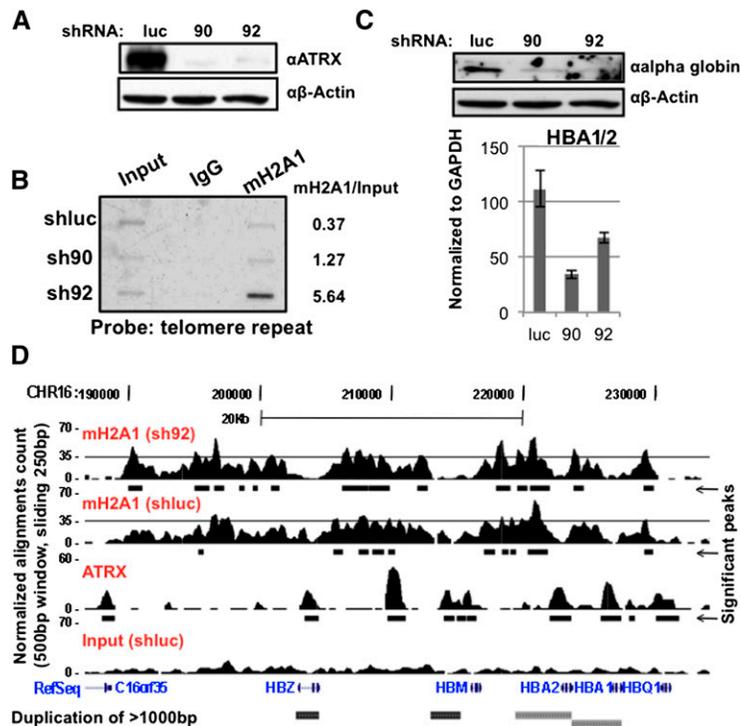
**Figure 5.** ATRX loss results in increased association of mH2A1 at telomeres and the α-globin cluster, concomitant with loss of α-globin expression. (A) shRNA-mediated knockdown of ATRX (sh90 and sh92) in K562 cells. β-Actin was used for loading. (B) ChIP reveals increased association of mH2A1 with telomeres in the absence of ATRX. (Right) Densitometry quantitation. One of two biological replicates is shown. (C) Loss of ATRX results in decreased α-globin protein and HBA mRNA. (D) Capture of the University of California at Santa Cruz Genome Browser showing an ~50-kb region around the α-globin locus. ChIP-seq-enriched peaks are shown for mH2A1 (sh92 and shluc), ATRX (Law et al. 2010), and input (shluc). Significant peaks (MACS) are shown below each panel as black bars. (Bottom) RefSeq annotated genes. The threshold line was set at 35 to facilitate visualization.

implicate ATRX in nucleosomal association of mH2A1, which may be important for establishing and/or maintaining chromatin states.

## Materials and methods

### Cell culture, plasmids, and shRNA

HEK293 and HeLa cells were grown in DMEM, and K562 cells were suspension-cultured in RPMI (10% FBS, 1% penicillin/streptomycin). GFP-tagged H2A or mH2A isoforms were expressed in HEK293 cells (pEGFP-C1 or N1, respectively), HA-tagged H3.1 and H3.3 were expressed in HeLa cells (Wiedemann et al. 2010), and Flag-tagged H2A or mH2A1.2 were infected into HeLa cells (pQXCIP). Selection was carried out in either 800 μg/mL neomycin or 1 μg/mL puromycin. For shRNA studies, HEK293, HeLa1.2.11, and K562 cells were infected with lentiviral plasmids encoding ATRX shRNAs (Open Biosystems, RHS4533-NM_000489) or luciferase shRNA by standard procedures and grown in 1 μg/mL puromycin. GFP-tagged plasmids of ATRX (N, mid, C, and full-length; gift of D. Picketts) were transiently transfected into HEK293 cells.

### Chromatin-free immunoprecipitation

Chromatin-free fractions were isolated essentially as described (Foltz et al. 2009). Material from large-scale chromatin-free immunoprecipitations was glycine-eluted, TCA-precipitated, and resolved by 4%–12% gel (NuPAGE, Invitrogen). Immunoprecipitations were carried out with αGFP beads (Vector Laboratories or Chromotek) for 3 h (at 250–375 mM salt). For immunoblots,

beads were washed and boiled in Laemmli loading buffer, and proteins were resolved by PAGE.

### Chromatin fractionation, histone acid extraction, and immunoblots

Chromatin fractionation was performed as described (Mendez and Stillman 2000). Histone acid extraction was performed as described (Kapoor et al. 2010), with the exception that isolated chromatin was treated with $H_2SO_4$. Immunoblots were performed with the following antibodies: αGFP (Roche, 11 814 460 001), αATRX (Santa Cruz Biotechnology, sc15408), αIMP9 (Abcam, ab52605), αNap1L1 (Abcam, ab33076), αNpm (Chemicon, MAB4500), αH3 (Millipore, 05-928), αDaxx (Cell Signaling, 25C12), αParp1 (Active Motif, 39561), αβActin (Sigma, A5316), αH2B (Millipore, 07-371), αHemoglobin α (Santa Cruz Biotechnology, sc-21005), and αmH2A1 (Abcam, ab37264).

### LC-MS/MS and qMS

Protein identification via LC-MS/MS was carried out essentially as described (Kaneko et al. 2010). qMS of mH2A1 was performed as described (Kapoor et al. 2010).

### Mononucleosome immunoprecipitation and MS quantification of H3 variants

Mononucleosome immunoprecipitation and qMS of the histone peptides were carried out essentially as described (Viens et al. 2006; Kapoor et al. 2010). In order to quantify the ratio of H3.3 to H3.1 and H3.2, residue 31 was used (serine in place of an alanine, respectively). All methylated and acetylated forms of the peptides were considered.

### cDNA isolation and qPCR

qPCR and mRNA analysis were carried out as described (Kapoor et al. 2010). cDNA expression was normalized to GAPDH levels. Primer sequences are provided in the Supplemental Material.

### FRAP

Live-cell imaging and long-term FRAP experiments were carried out essentially as reported (Wiedemann et al. 2010). For visualization of the results, single-cell measurements were averaged and plotted together with the respective standard error for every time point.

### Native ChIP-seq

Native mH2A1 ChIP (Abcam, ab37264) and Input DNA were prepared from K562 cells, and subsequent sequencing was performed using Illumina Hi-Seq. See the Supplemental Material for full details.

### Telomere Southern blot

DNA was isolated following mH2A1 ChIP, and telomere Southern blot was carried out as described (Goldberg et al. 2010), with the exception that the probe used was $T_2AG_3$ from pSty-11.

## References

Angelov D, Molla A, Perche PY, Hans F, Cote J, Khochbin S, Bouvet P, Dimitrov S. 2003. The histone variant macroH2A interferes with transcription factor binding and SWI/SNF nucleosome remodeling. *Mol Cell* **11:** 1033–1041.

Banaszynski LA, Allis CD, Lewis PW. 2010. Histone variants in metazoan development. *Dev Cell* **19:** 662–674.

Berube NG. 2011. ATRX in chromatin assembly and genome architecture during development and disease. *Biochem Cell Biol* **89:** 435–444.

Berube NG, Smeenk CA, Picketts DJ. 2000. Cell cycle-dependent phosphorylation of the ATRX protein correlates with changes in nuclear matrix and chromatin association. *Hum Mol Genet* **9:** 539–547.

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

Chakravarthy S, Gundimella SK, Caron C, Perche PY, Pehrson JR, Khochbin S, Luger K. 2005. Structural characterization of the histone variant macroH2A. *Mol Cell Biol* **25:** 7616–7624.

Costanzi C, Pehrson JR. 1998. Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature* **393:** 599–601.

Costanzi C, Pehrson JR. 2001. MACROH2A2, a new member of the MACROH2A core histone family. *J Biol Chem* **276:** 21776–21784.

De Koning L, Corpet A, Haber JE, Almouzni G. 2007. Histone chaperones: An escort network regulating histone traffic. *Nat Struct Mol Biol* **14:** 997–1007.

Drane P, Ouararhni K, Depaux A, Shuaib M, Hamiche A. 2010. The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. *Genes Dev* **24:** 1253–1265.

Dunleavy EM, Roche D, Tagami H, Lacoste N, Ray-Gallet D, Nakamura Y, Daigo Y, Nakatani Y, Almouzni-Pettinotti G. 2009. HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. *Cell* **137:** 485–497.

Foltz DR, Jansen LE, Bailey AO, Yates JR 3rd, Bassett EA, Wood S, Black BE, Cleveland DW. 2009. Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. *Cell* **137:** 472–484.

Gamble MJ, Frizzell KM, Yang C, Krishnakumar R, Kraus WL. 2010. The histone variant macroH2A1 marks repressed autosomal chromatin, but protects a subset of its target genes from silencing. *Genes Dev* **24:** 21–32.

Garrick D, Samara V, McDowell TL, Smith AJ, Dobbie L, Higgs DR, Gibbons RJ. 2004. A conserved truncated isoform of the ATR-X syndrome protein lacking the SWI/SNF-homology domain. *Gene* **326:** 23–34.

Gaume X, Monier K, Argoul F, Mongelard F, Bouvet P. 2011. In vivo study of the histone chaperone activity of nucleolin by FRAP. *Biochem Res Int* **2011:** 187624. doi: 10.1155/2011/187624.

Goldberg AD, Banaszynski LA, Noh KM, Lewis PW, Elsaesser SJ, 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.

Heaphy CM, de Wilde RF, Jiao Y, Klein AP, Edil BH, Shi C, Bettegowda C, Rodriguez FJ, Eberhart CG, Hebbar S, et al. 2011. Altered telomeres in tumors with ATRX and DAXX mutations. *Science* **333:** 425.

Higgs DR, Garrick D, Anguita E, De Gobbi M, Hughes J, Muers M, Vernimmen D, Lower K, Law M, Argentaro A, et al. 2005. Understanding α-globin gene regulation: Aiming to improve the management of thalassemia. *Ann N Y Acad Sci* **1054:** 92–102.

Iwase S, Xiang B, Ghosh S, Ren T, Lewis PW, Cochrane JC, Allis CD, Picketts DJ, Patel DJ, Li H, et al. 2011. ATRX ADD domain links an atypical histone methylation recognition mechanism to human mental-retardation syndrome. *Nat Struct Mol Biol* **18:** 769–776.

Kaneko S, Li G, Son J, Xu CF, Margueron R, Neubert TA, Reinberg D. 2010. Phosphorylation of the PRC2 component Ezh2 is cell cycle-

regulated and up-regulates its binding to ncRNA. *Genes Dev* **24:** 2615–2620.

Kapoor A, Goldberg MS, Cumberland LK, Ratnakumar K, Segura MF, Emanuel PO, Menendez S, Vardabasso C, Leroy G, Vidal CI, et al. 2010. The histone variant macroH2A suppresses melanoma progression through regulation of CDK8. *Nature* **468:** 1105–1109.

Law MJ, Lower KM, Voon HP, Hughes JR, Garrick D, Viprakasit V, Mitson M, De Gobbi M, Marra M, Morris A, et al. 2010. ATR-X syndrome protein targets tandem repeats and influences allele-specific expression in a size-dependent manner. *Cell* **143:** 367–378.

Lewis PW, Elsaesser SJ, Noh KM, Stadler SC, Allis CD. 2010. Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. *Proc Natl Acad Sci* **107:** 14075–14080.

Mendez J, Stillman B. 2000. Chromatin association of human origin recognition complex, cdc6, and minichromosome maintenance proteins during the cell cycle: Assembly of prereplication complexes in late mitosis. *Mol Cell Biol* **20:** 8602–8612.

Mitson M, Kelley LA, Sternberg MJ, Higgs DR, Gibbons RJ. 2011. Functional significance of mutations in the Snf2 domain of ATRX. *Hum Mol Genet* **20:** 2603–2610.

Mosammaparast N, Del Rosario BC, Pemberton LF. 2005. Modulation of histone deposition by the karyopherin kap114. *Mol Cell Biol* **25:** 1764–1778.

Novikov L, Park JW, Chen H, Klerman H, Jalloh AS, Gamble MJ. 2011. Qki-mediated alternative splicing of the histone variant Macroh2a1 regulates cancer cell proliferation. *Mol Cell Biol* **31:** 4244–4255.

Papamichos-Chronakis M, Watanabe S, Rando OJ, Peterson CL. 2011. Global regulation of H2A.Z localization by the INO80 chromatin-remodeling enzyme is essential for genome integrity. *Cell* **144:** 200–213.

Park YJ, Luger K. 2008. Histone chaperones in nucleosome eviction and histone exchange. *Curr Opin Struct Biol* **18:** 282–289.

Pehrson JR, Fried VA. 1992. MacroH2A, a core histone containing a large nonhistone region. *Science* **257:** 1398–1400.

Plazas-Mayorca MD, Zee BM, Young NL, Fingerman IM, LeRoy G, Briggs SD, Garcia BA. 2009. One-pot shotgun quantitative mass spectrometry characterization of histones. *J Proteome Res* **8:** 5367–5374.

Sporn JC, Kustatscher G, Hothorn T, Collado M, Serrano M, Muley T, Schnabel P, Ladurner AG. 2009. Histone macroH2A isoforms predict the risk of lung cancer recurrence. *Oncogene* **28:** 3423–3428.

Tang J, Wu S, Liu H, Stratt R, Barak OG, Shiekhattar R, Picketts DJ, Yang X. 2004. A novel transcription regulatory complex containing death domain-associated protein and the ATR-X syndrome protein. *J Biol Chem* **279:** 20369–20377.

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

Viens A, Mechold U, Brouillard F, Gilbert C, Leclerc P, Ogryzko V. 2006. Analysis of human histone H2AZ deposition in vivo argues against its direct role in epigenetic templating mechanisms. *Mol Cell Biol* **26:** 5325–5335.

Wiedemann SM, Mildner SN, Bonisch 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.

Wilson BG, Roberts CW. 2011. SWI/SNF nucleosome remodellers and cancer. *Nat Rev Cancer* **11:** 481–492.

Wong LH, McGhie JD, Sim M, Anderson MA, Ahn S, Hannan RD, George AJ, Morgan KA, Mann JR, Choo KH. 2010. ATRX interacts with H3.3 in maintaining telomere structural integrity in pluripotent embryonic stem cells. *Genome Res* **20:** 351–360.

Zhang R, Poustovoitov MV, Ye X, Santos HA, Chen W, Daganzo SM, Erzberger JP, Serebriiskii IG, Canutescu AA, Dunbrack RL, et al. 2005. Formation of MacroH2A-containing senescence-associated heterochromatin foci and senescence driven by ASF1a and HIRA. *Dev Cell* **8:** 19–30.
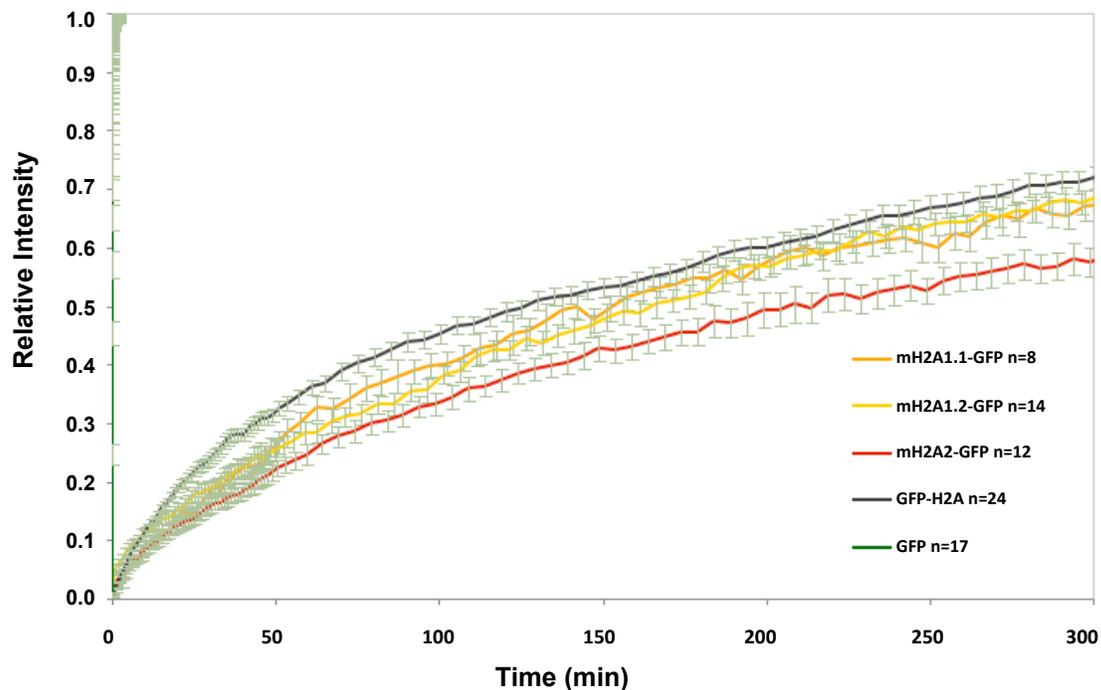
# Supplementary information for:

# ATRX-mediated chromatin association of histone variant macroH2A1 regulates α-globin expression
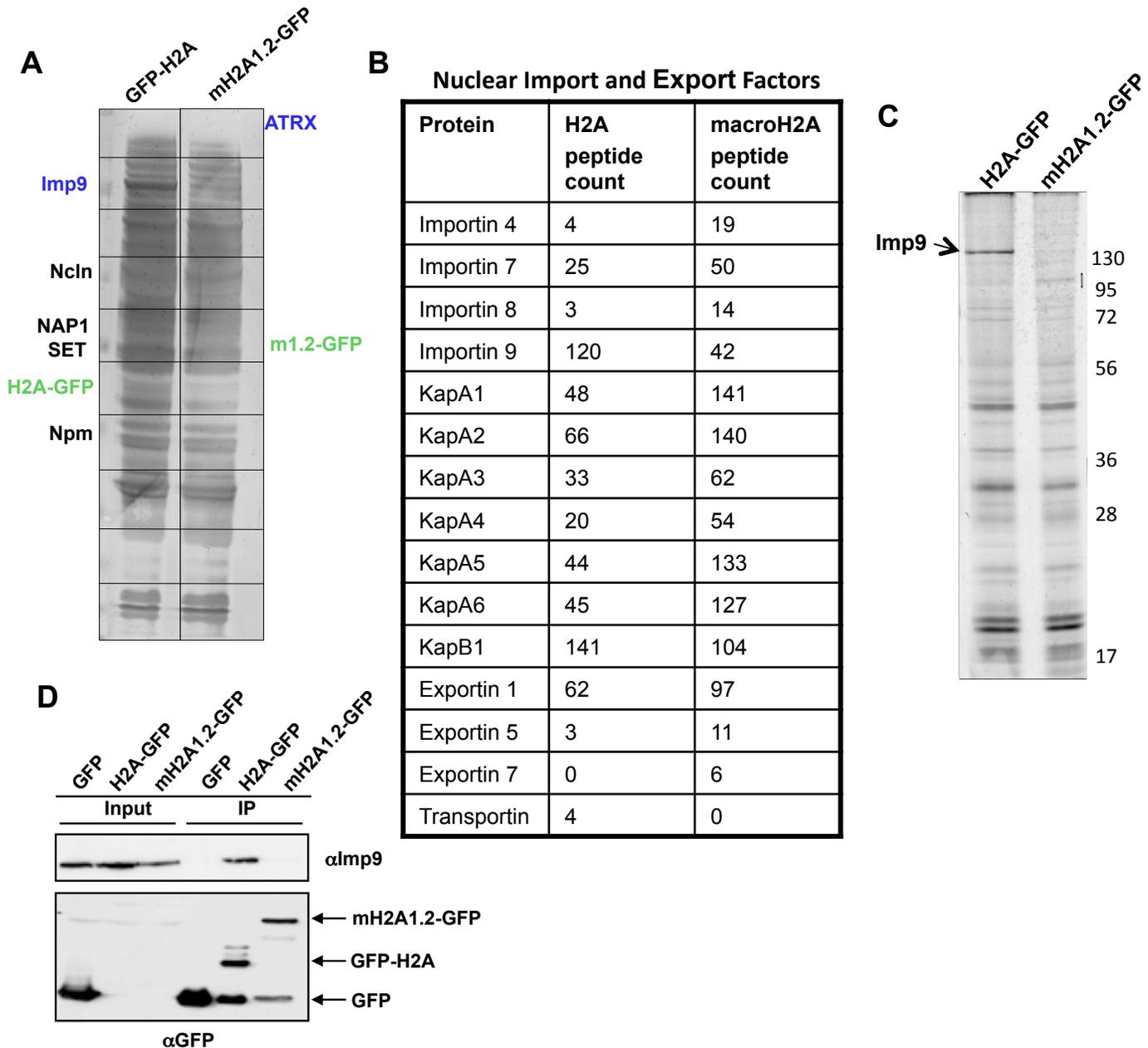
**A**



**B**
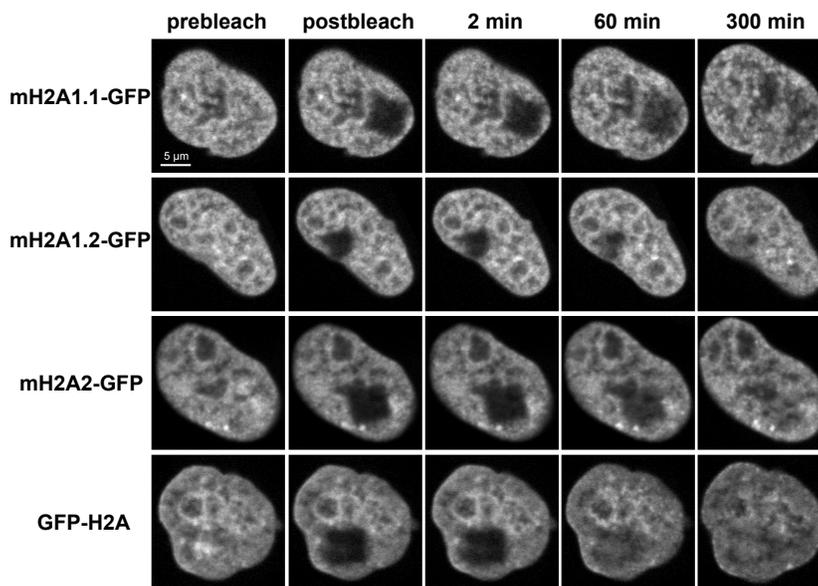


**Supplemental Figure 1. FRAP analysis of canonical H2A and mH2A isoforms.** (*A*) Representative images from long-term FRAP time series of HeLa cells transiently transfected with GFP-H2A or mH2A-GFP isoforms. Squares of 5 μm x 5 μm within the nucleus were photobleached and fluorescence recovery followed for at least 5 hours. Prebleach, postbleach, and 2min, 60min, 300min postbleach are shown. (*B*) Quantitative FRAP evaluation. mH2A1 isoforms (orange and yellow) display a decreased recovery compared to canonical H2A (blue). mH2A2 (red) displays decreased recovery compared with canonical H2A and mH2A1 isoforms. In contrast, GFP alone shows full recovery within less than a minute (green).

**A**

**B** Nuclear Import and Export Factors

| Protein | H2A peptide count | macroH2A peptide count |
|---|---|---|
| Importin 4 | 4 | 19 |
| Importin 7 | 25 | 50 |
| Importin 8 | 3 | 14 |
| Importin 9 | 120 | 42 |
| KapA1 | 48 | 141 |
| KapA2 | 66 | 140 |
| KapA3 | 33 | 62 |
| KapA4 | 20 | 54 |
| KapA5 | 44 | 133 |
| KapA6 | 45 | 127 |
| KapB1 | 141 | 104 |
| Exportin 1 | 62 | 97 |
| Exportin 5 | 3 | 11 |
| Exportin 7 | 0 | 6 |
| Transportin | 4 | 0 |

**C**

**D**

**Supplemental Figure 2. H2A and mH2A1.2 associate with nuclear-cytoplasmic shuttling factors and chromatin-associated factors in chromatin-free extracts.** (*A*) Proteins interacting with either GFP-H2A or mH2A1.2-GFP in chromatin-free extracts were resolved on 4-12% NuPAGE and silver stained. Each lane was cut into ten slices and MS analyzed. Chromatin-associated factors identified by MS are presented alongside the gel (see Fig. 1D for peptide counts). Blue characters = H2A or mH2A1.2-specific factors; black characters = chromatin-associated factors which interact with both H2A and mH2A1.2; green = GFP histones. (*B*) Table of nuclear-cytoplasmic shuttling factors that interact with H2A or mH2A1.2 in chromatin free extracts. These data were collected from MS analysis of the gel presented in (*A*) and Supplemental Table 1. (*C*) Chromatin-free IP and silver staining of proteins interacting with GFP-H2A or mH2A1.2-GFP resolved on 4-12% NuPAGE. Unique bands were excised and MS analyzed. Arrow indicates Imp9; molecular weight marker shown on right. (*D*) GFP, GFP-H2A or mH2A1.2-GFP IPs were resolved on 12% PAGE and IB'd for Imp9 (upper panel) or GFP (lower panel).

**A**

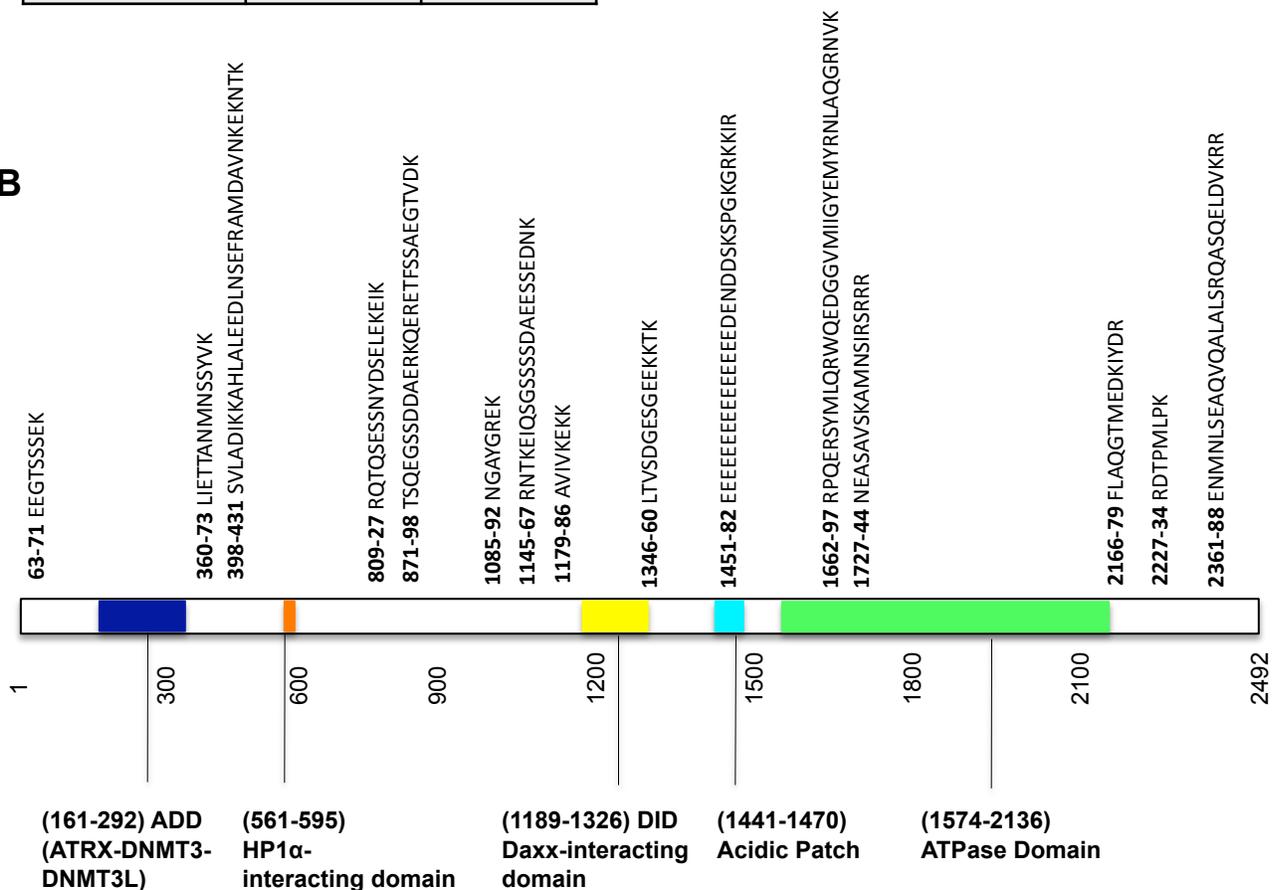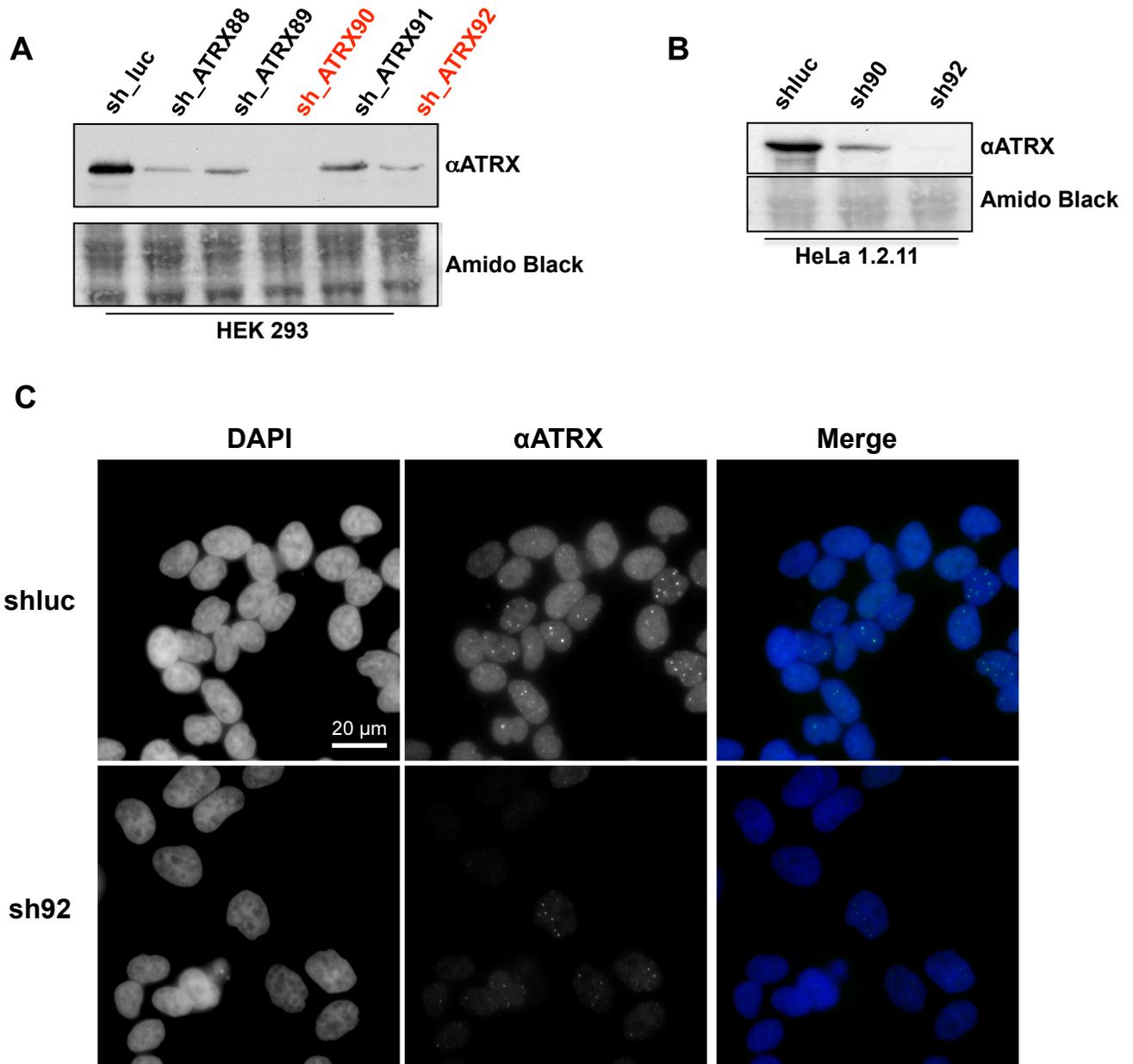| Protein | H2A Peptide counts | mH2A1.2 Peptide counts |
|---|---|---|
| NAP1 | 32 | 72 |
| SET | 2 | 19 |
| Nucleolin | 79 | 81 |
| Nucleophosmin | 7 | 42 |
| **ATRX** | **0** | **19** |
| Importin 9 | 120 | 42 |

**C**



**B**
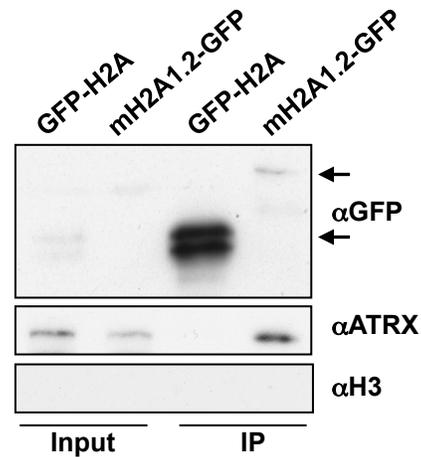


**Supplemental Figure 3. ATRX interacts with mH2A in chromatin-free extracts.** (*A*) Peptide counts of chaperones identified by MS for H2A and mH2A1.2 (See corresponding gel in Supplemental Fig. 2A). (*B*) ATRX peptides identified by MS analysis in mH2A1.2 IPs from chromatin-free extracts. Peptides presented next to a schematic of ATRX protein with its various domains shown. Nineteen peptides in total were identified (15 are presented here as 4 peptides were overlapping). (*C*) IBs of GFP and ATRX from chromatin-free extracts by independent chromatin fractionation method (Mendez and Stillman 2000).

ATRX fragments

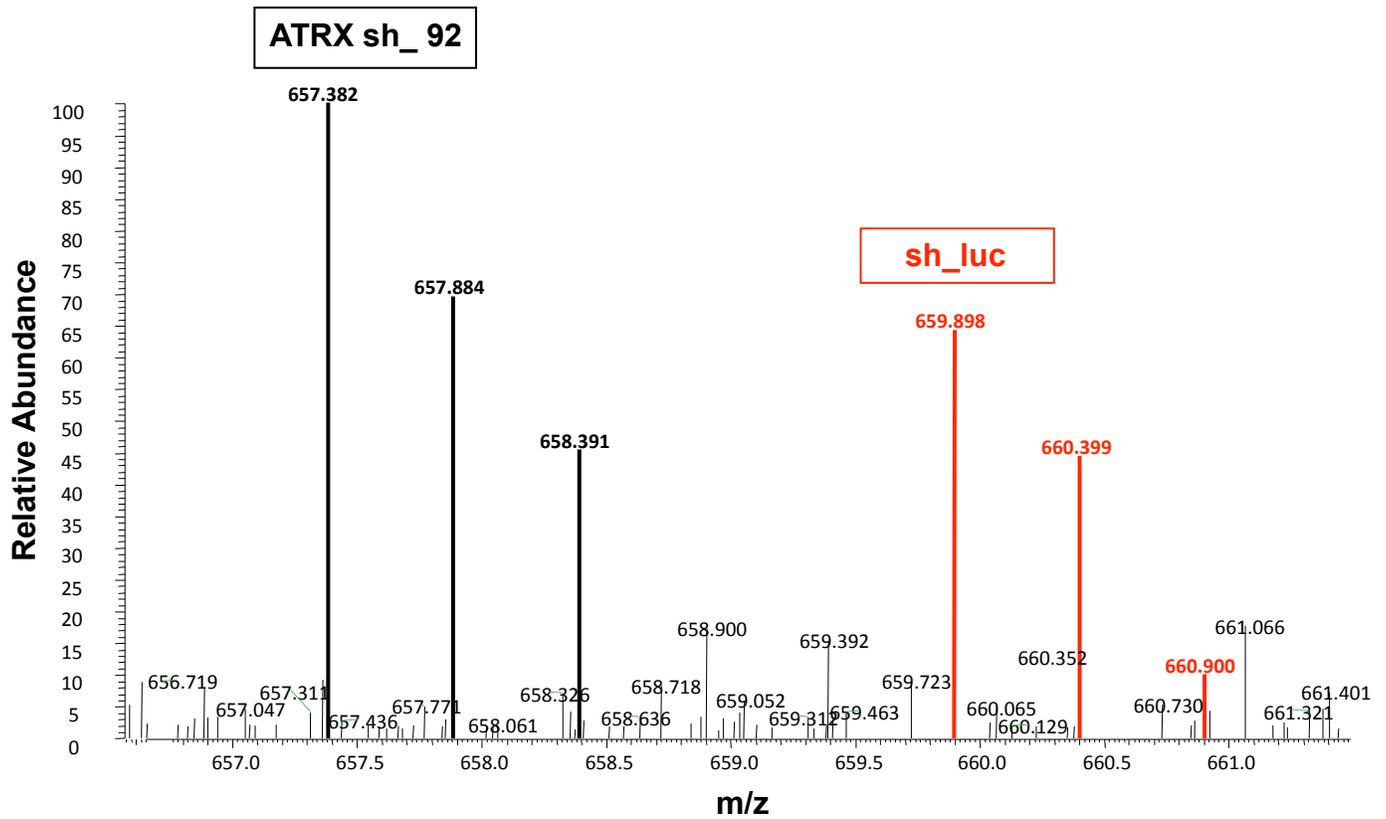| Full length ATRX | N-term (1-841) | Middle (800-1670) | C-term (1670-2942) |
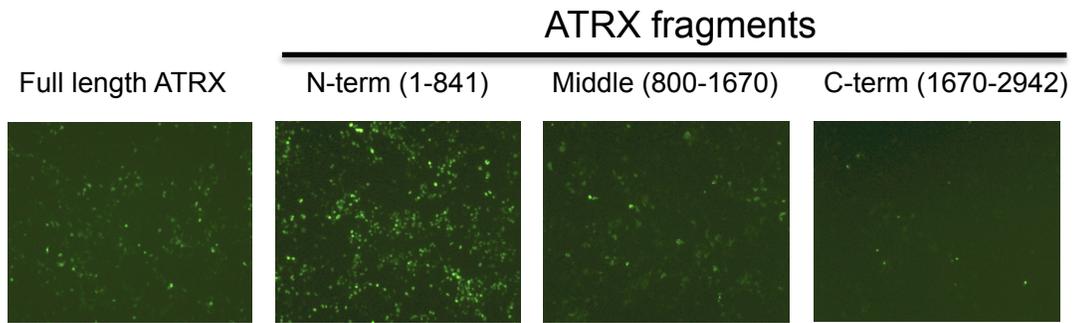


**Supplemental Figure 4. Expression of ATRX deletion constructs.** Fluorescent images of GFP-tagged ATRX constructs transiently expressed in HEK293 cells. While the N-terminal construct expressed strongly, enrichment of mH2A1 signal observed in the N-terminal IP (Fig. 2C, D) suggests that this region is sufficient for binding to mH2A1.
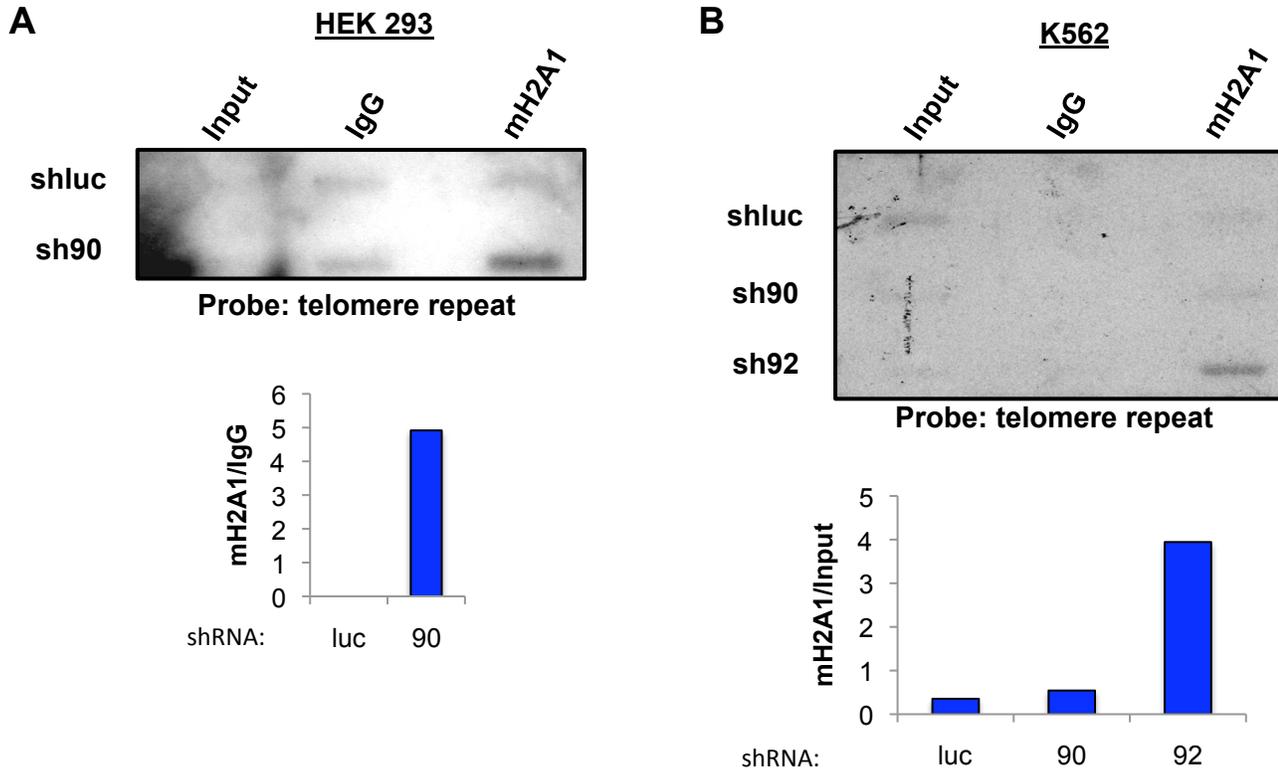
**Supplemental Figure 5. shRNA-mediated depletion of ATRX in 293 and HeLa1.2.11 cells.** (*A*) αATRX IB of HEK 293 cells expressing five lentivirally introduced ATRX shRNAs, and control shluc. sh90 and sh92, used for subsequent studies, are highlighted in red; Amido black of histones shown for loading. (*B, C*) ATRX depletion in HeLa1.2.11 cells shown by IB (*B*) and immunofluorescence (*C*). Note presence of ATRX at telomeres.
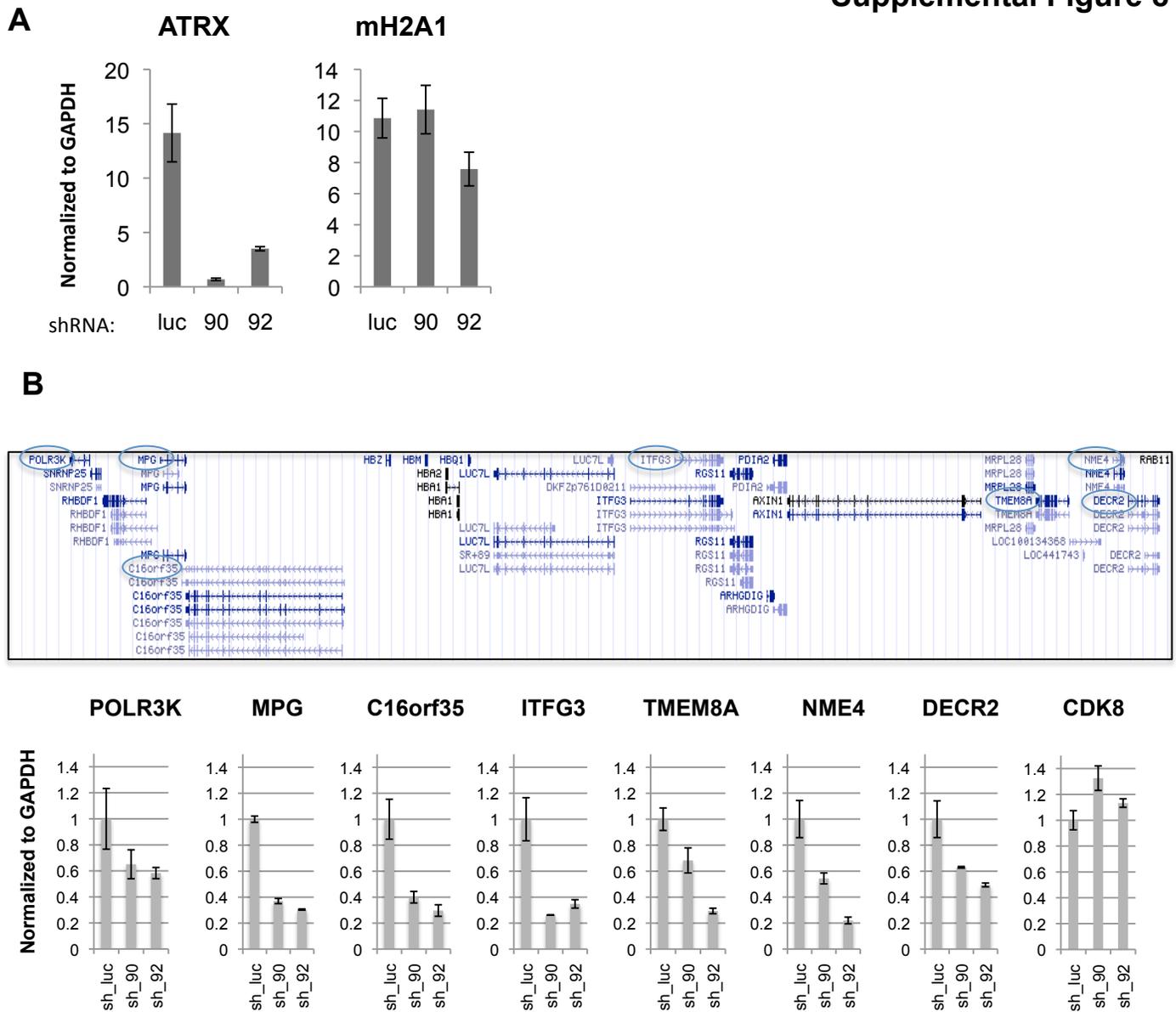
**Supplemental Figure 6. ATRX knockdown results in mH2A1 accumulation in chromatin.** Comparison of an averaged full mass spectrum for the +2 charged mH2A1 peptide (SAKAGVIFPVGR) from sh92 ATRX knockdown ($d_0$-labeled) and shluc control ($d_5$-labeled) histones extracted from chromatin.

**A** HEK 293

Input   IgG   mH2A1

shluc

sh90

**Probe: telomere repeat**

mH2A1/IgG

6
5
4
3
2
1
0

shRNA:   luc   90

**B** K562

Input   IgG   mH2A1

shluc

sh90

sh92

**Probe: telomere repeat**

mH2A1/Input

5
4
3
2
1
0

shRNA:   luc   90   92

**Supplemental Figure 7. Loss of ATRX results in telomeric accumulation of mH2A1.** *(A)* One of two representative ChIP-telomere Southern blots for HEK293 cells shows the increased association of mH2A1 with telomeric chromatin in the absence of ATRX. Only sh90 was used as it induced the most efficient knockdown in this cell line (see Supplemental Fig. 4). *(B)* Biological replicate of telomere Southern blot in K562 cell line. mH2A1's presence at telomeres of K562 cells is increased in the absence of ATRX in both sh90 and sh92 lines. Densitometry quantitation presented below each graph.

**A**



**B**



**Supplemental Figure 8. ATRX knockdown results in reduced RNA levels of sub-telomeric chromosome 16 genes.** (*A*) shRNA-mediated knockdown of ATRX (sh90 and sh92) in K562 cells results in the loss of ATRX mRNA, compared to shluc, without affecting mH2A1 mRNA. (*B*) Knock down of ATRX results in decreased mRNA levels of genes found in the α globin cluster of sub-telomeric chromosome 16. Starting with the most telomere proximal, genes assayed include POLR3K, MPG, C16orf35, ITFG3, TMEM8A, NME4 and DECR2. Genes assayed are circled in blue in the UCSC browser above. CDK8 transcription (chromosome 13) is unaffected by ATRX knockdown, similar to mH2A1 (chromosome 5) in (A). Expression was measured relative to GAPDH and to the control shluc, whose expression values were arbitrarily set as 1.
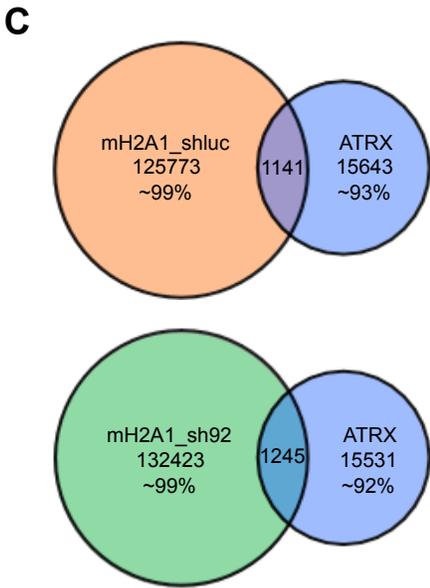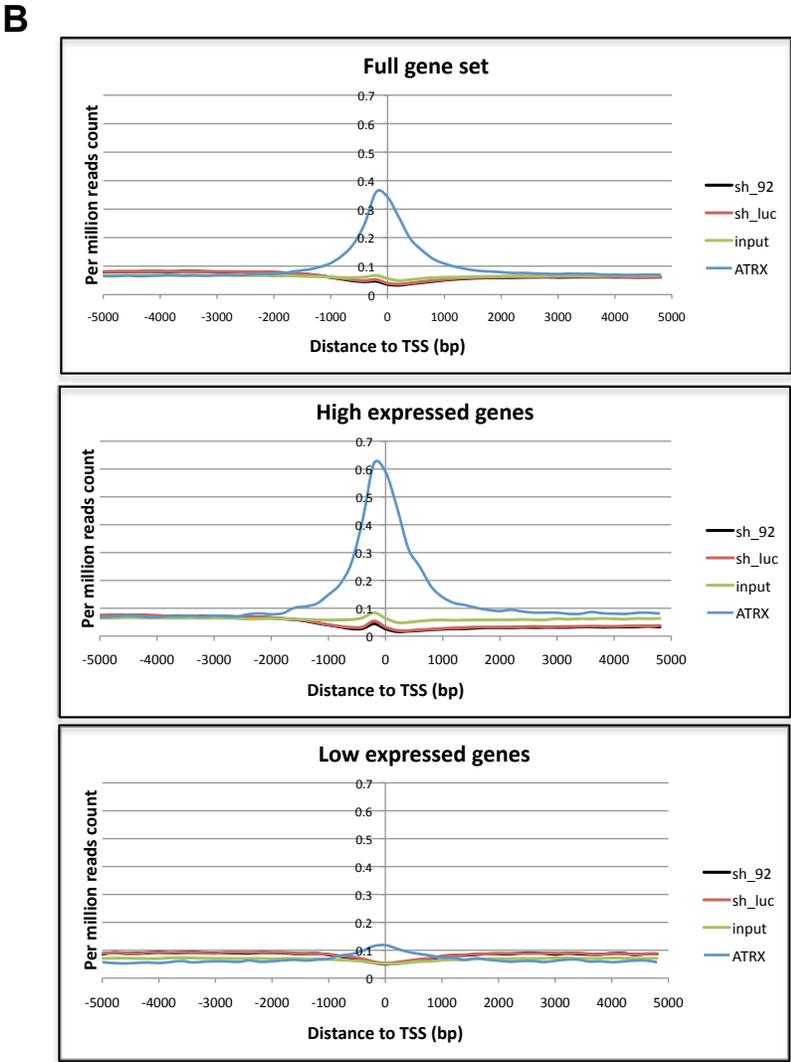
**A**
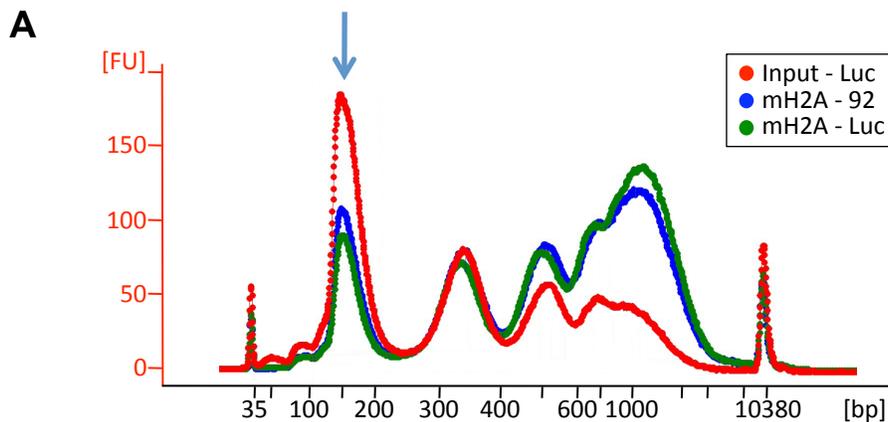


**B**

|  | sh92 | shluc | Input (shluc) | ATRX | Input (ATRX) |
|---|---|---|---|---|---|
| Raw reads | 67219237 | 56540184 | 148165330 | 9139129 | 18232783 |
| Bowtie alignments (wiggle) | 69827565 | 59143870 | 145176252 | 5563975 | 18468175 |
| Alignments analyzed (MACS) | 51200836 | 45308710 | 63296750 | 4654137 | 17770258 |
| Total peaks count (MACS) | 133763 | 127000 | - | 16848 | - |
| Total peak length (bp) (MACS) | 113477382 | 94296714 | - | 12914783 | - |

**C**



shluc 65413 ~54%    55397    sh92 74145 ~57%
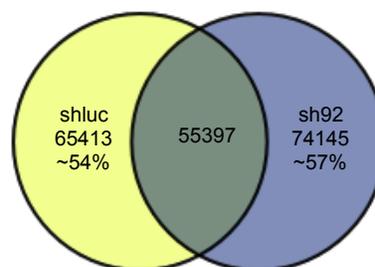
**Supplemental Figure 9. ChIP-sequencing of mH2A in shluc and sh92 K562 cells.** (*A*) Bioanalyzer traces of ChIP (mH2A1) and Input DNA from MNase digested chromatin (High sensitivity DNA chip; Agilent Technologies). Isolated mononucleosomal DNA (arrow) was size selected for ChIP-Seq library preparation. Input (red), sh92 (blue) and shluc (green). Peaks at 35bp and 10380bp are internal size markers. (*B*) Raw number of reads obtained by Illumina Hi-Seq, total number of alignments (Bowtie: -m 20 -k 20 -n 2 -l 50), and alignments used for peak calling (MACS) for sh92, shluc and Input shown. Raw reads for the ATRX ChIP-Seq were downloaded from GEO (GSE22162). Also shown are the total number of peaks for sh92 and shluc (p value cutoff = 1.00e-4) and total number of base pairs covered by significant peaks (MACS). (*C*) Overlap of significant mH2A1 peaks (MACS) from shluc and sh92 ChIP-seq data sets (HOMER software, Heinz et al. 2010).

**A**



**C**

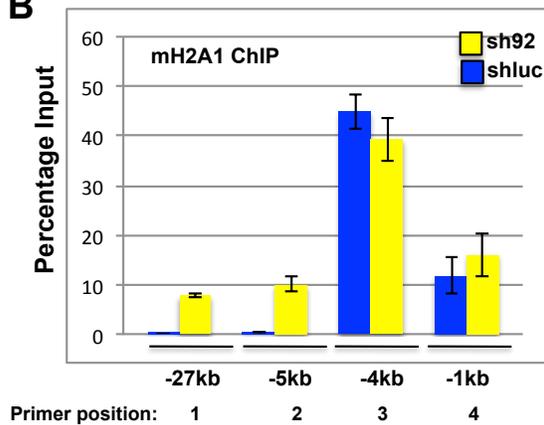**Supplemental Figure 10**



**B**



**Supplemental Figure 10. TSS/TES analysis and peak overlap of mH2A1 and ATRX.** (*A*) Read counts (200bp window) normalized to total number of reads (counts per million reads), plotted against the distance (-5Kb, +5Kb), from the nearest annotated Transcription Start Site (TSS, left), and Transcription End Site (TES, right). Genes were grouped by expression levels to Full (all annotated genes), High and Low categories based on K562 RNA-Seq data from the ENCODE project. sh92 (black), shluc (red), and Input (green). (*B*) Same TSS analysis as in (A) also containing ATRX (blue). Scale is different than in (*A*). (*C*) Overlap of significant mH2A1 and ATRX peaks (MACS) in shluc and sh92 samples (Law et al. 20120; Heinz et al. 2010).
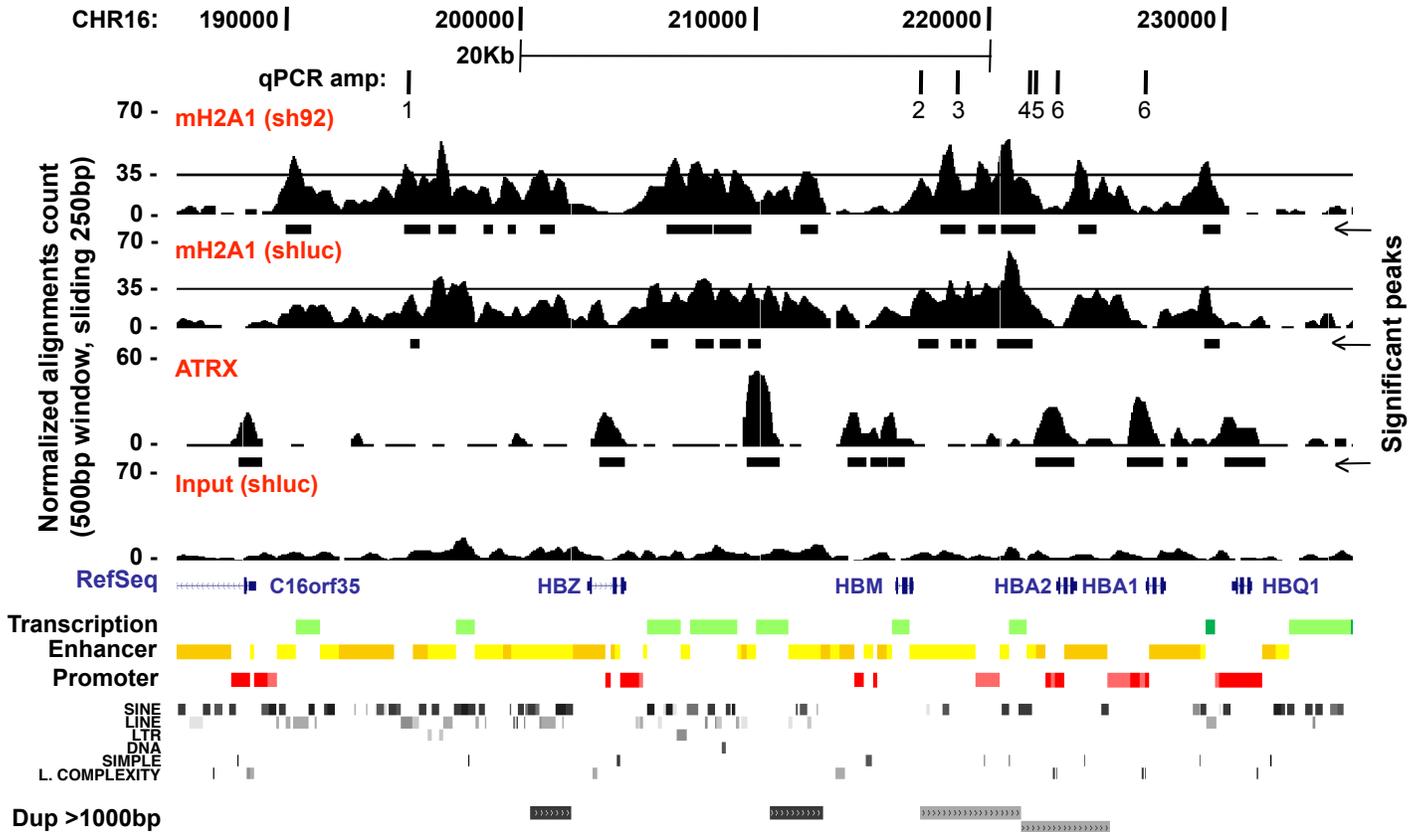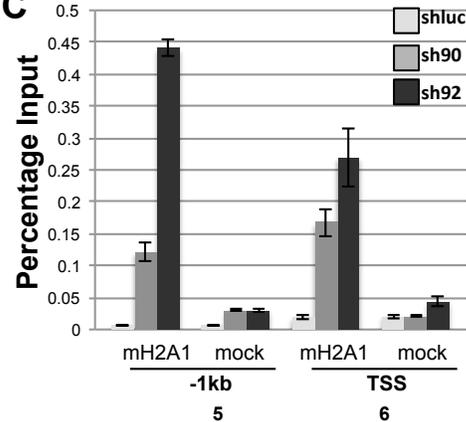
**Supplemental Figure 11. mH2A1 is enriched at the α globin cluster**. (A) Capture of UCSC genome browser showing ~50kb region around α globin locus. ChIP –seq analysis of mH2A1 in K562 cells. Loss of ATRX (sh92) results in redistribution of mH2A1 compared to control cells (shluc) as shown in Fig. 5D (shown on the X-axis genomic position in Kb; Y-axis alignment counts in 500bp window sliding 250bp. Window counts are normalized to total number of alignments and scaled by 10^7 for mH2A1 and 10^6 for ATRX). ChIP-seq data for mH2A1 was compared to published ATRX ChIP-seq data (Law et al. 2010). Shluc Input enrichment is shown for reference. Threshold line set at 35 to facilitate visualization. Regions of significant enrichment are indicated by black bars below the respective enrichment plots. UCSC custom tracks are shown, RefSeq gene annotation, K562 Chromatin State Segmentation by HMM from ENCODE/Broad, Repeat elements by RepeatMasker, and Duplications of >1000 Bases of Non-RepeatMasked Sequence are shown, respectively. Primer pairs used in B (native ChIP) and C (fixed ChIP) shown on top. As the HBA genes are duplicated, primer pair 6 at the TSS is presented twice. (B) Validation by qPCR of the ChIP-sequencing library indicates enrichment of mH2A1 at regions upstream of the HBA1/2 genes. (C) ChIP via standard formaldehyde cross-linking and sonication demonstrates increased mH2A1 levels 1kb upstream and at TSS of the HBA1/2 genes when ATRX is depleted. Mock = no antibody. One of three replicates (two biological and one technical) shown. The stronger enrichment of mH2A1 signal in native ChIP (B) compared with formaldehyde-fixed ChIP (C) is likely due to greater antibody affinity for mH2A1 in the native ChIP protocol.

Ratnakumar et al.

**Supplemental Materials and Methods**

**Immunofluorescence.** Cells were fixed with 2% formaldehyde/PBS (Sigma Aldrich) for 10 min at room temperature (RT) and subsequently permeabilized with 0.5% Triton X-100/PBS for 10 min. After incubation in 2% BSA/PBST blocking solution (bovine serum albumin in PBS with 0.02% Tween-20) for 1 h at RT, cells were incubated with rabbit αATRX antibody (Santa Cruz, sc15408) for 1 h at RT, and after washing, with goat αrabbit Alexa-488 (Invitrogen) for 30 min. Cells were post-fixed with 4% formaldehyde/PBST for 10 min. Counterstaining of DNA was performed with 4',6-diamidino-2-phenylindole (DAPI, 2.5 µg/ml; Sigma-Aldrich) and samples were mounted in Vectashield Anti-fade Mounting Medium (Vector Laboratories). Images were taken using a pDV wide-field epifluorescence microscope (Applied Precision Imaging, Inc.) equipped with a 60x/1.42 NA Plan-Apochromat oil objective lens (Olympus) and a CoolSnap HQ$^2$ CCD camera (Photometrics).

**ChIP-qPCR and ChIP-Southern blot.** ChIP in K562 cells was performed using the Magna ChIP Kit (Millipore) with modifications (described by the Myers Lab online protocol at Stanford University). Briefly, 2 x 10^7 cells were lysed (Magna ChIP Cell Lysis Buffer) and nuclei sonicated in 1ml RIPA buffer (1XPBS/1% NP-40/0.5% sodium deoxycholate/0.1% SDS), 4 minutes (30sec on, 30 sec off - Diagenode Bioruptor) to yield 500-800bp DNA fragments. Sonicated chromatin from 2x10^6 cells was immunoprecipitated overnight in Magna ChIP Kit Dilution Buffer at 4C with 5ug of αmH2A1 (Abcam, ab37264 or Millipore, 07-219), αIgG (Millipore 12-370), or no antibody control. Capture was performed with Millipore Magna ChIP Protein A beads for two hours at 4C. Beads were washed and reverse cross-linked as described by Magna ChIP Kit protocol. DNA was isolated using QiaQUICK PCR Purification kit.

**ChIP-sequencing and peak calling.** Chromatin from K562 cells was digested with MNase and used for ChIP with $\alpha$mH2A1, essentially as described (Alonso et al. 2003; Hori et al. 2008). Mononucleosomal DNA from Input and ChIP samples were processed following Illumina standard library protocol with modifications to allow multiplexing. High throughput sequencing was performed using Illumina Hi-Seq obtaining a total of 67,219,237 reads for sh92, 56,540,184 for shluc and 148,165,330 for Input. 100bp reads were aligned to the hg19 genome assembly (NCBI build 37) using Bowtie short read aligner (v 0.12.7) (Langmead et al. 2009), with the following parameters: seed of 50bp, maximum 2 mismatches, suppression (m) = 20, and reported alignments (k) = 20. Wiggle files (HAFEZ, unpublished pipeline D.H.), were generated using a 500bp window sliding 250bp, counting the number of aligned reads (5' end of each aligned read), for both ChIP and Input samples. The number of alignments from each window was normalized to the total number of alignments and scaled by factor of $10^7$, to allow comparison between different samples. The MACS software (v 1.4.1) (4) was used to identify peaks (pvalue cutoff = 1.00e-4, shiftsize = 75) (Zhang et al. 2008). For the ATRX (Law et al. 2010) analysis, raw reads were downloaded from GEO (GSE22162) and analyzed as described above (pvalue cutoff = 1.00e-5). The University of California at Santa Cruz (UCSC) genome browser was used to examine the distribution of sequenced tags (wiggle), on specific genomic regions.

**TSS/TES analysis.** The relative positions of the aligned reads to the TSS/TES were generated using RefSeq gene annotations downloaded from the UCSC genome browser. Analysis of reads distribution performed with a sliding window (200bp). Reads were normalized to total number of alignments and presented per $10^6$ reads. Gene expression classification was done based on the K562 RNA-Seq (RPKM) from ENCODE annotations

(http://h.gdownload.cse.ucsc.edu/goldenPath/hg18/encodeDCC/wgEncodeCaltechRnaSeq/wgEncodeC altechRnaSeqRawData5Rep2K562CellLongpolyaErng32x75.rpkm.gz). Gene clusters show all genes

Ratnakumar et al.

(full), top and bottom 20% of genes (high and low clusters respectively). For the ATRX (Law et al. 2010) analysis, raw reads were downloaded from GEO (GSE22162).

**Analysis of overlapping regions.** Significant peaks from sh92, shluc and ATRX samples generated using MACS were compared to find overlapping and differentially bounds peaks using Homer (v3.8.2, 1-06-2012, Heinz et al. 2010), with mergePeaks (-d given, -venn). Venn diagrams were generated based on this analysis.

**Data deposition.** All ChIP-Seq data sets deposited to NCBI's Gene Expression Omnibus with the deposition number GSE35339.

**ChIP and RT-PCR Primers.**

**Native ChIP Primers**

-27kb Fwd : TCGTAGCCTGGGTCAGAACT, Rev : ACCCAAATGTCCATCACTGG

-5kb Fwd : CCCTTTTTGCAGAGAGATGC, Rev : GAAGGGGAAAGGAACAGGTC

-4kb Fwd : AGAAATAAGCTGCCGTGGTG, Rev : CATAGCCAGCAGTGTGGTTG

-1kb Fwd : CTGGCCAAACCATCACTTTT, Rev : AGGAAGGAAGGGGTGGACT

**Fixed ChIP Primers:**

**HBA1**

-1kb Fwd : TCCCCTCACCTACATTCTGC, Rev : GCCTAGAGGTCGTGGTTCAC

Ratnakumar et al.

TSS : HBA1/2 primers (Law et al. 2010)

**RT-PCR primers:**

ATRX: Fwd: GCAACCTTGGTCGAAAGGAGT; Rev : GGCTCTGGGTGACAAATGTAG (Primer Bank)

mH2A1: Previously described (Kapoor et al. 2010)

HBA1: Previously described (Privitera et al. 1995)

POLR3K, MPG, C16orf35, ITFG3, TMEM8A, NME4, DECR2: Previously described (Law et al. 2010)

Ratnakumar et al.

**Supplemental References**

Alonso A, Mahmood R, Li S, Cheung F, Yoda K, Warburton PE. 2003. Genomic microarray analysis reveals distinct locations for the CENP-A binding domains in three human chromosome 13q32 neocentromeres. *Hum Mol Genet* **12**: 2711-2721.

Foltz DR, Jansen LE, Bailey AO, Yates JR, 3rd, Bassett EA, Wood S, Black BE, Cleveland DW. 2009. Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. *Cell* **137**: 472-484.

Hori T, Amano M, Suzuki A, Backer CB, Welburn JP, Dong Y, McEwen BF, Shang WH, Suzuki E, Okawa K et al. 2008. CCAN makes multiple contacts with centromeric DNA to provide distinct pathways to the outer kinetochore. *Cell* **135**: 1039-1052.

Kapoor A, Goldberg MS, Cumberland LK, Ratnakumar K, Segura MF, Emanuel PO, Menendez S, Vardabasso C, Leroy G, Vidal CI et al. 2010. The histone variant macroH2A suppresses melanoma progression through regulation of CDK8. *Nature* **468**: 1105-1109.

Heinz S, Benner C, Spann N, Bertolino E et al. Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. Mol Cell 2010 May 28;38(4):576-589

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**: R25.

Law MJ, Lower KM, Voon HP, Hughes JR, Garrick D, Viprakasit V, Mitson M, De Gobbi M, Marra M, Morris A et al. 2010. ATR-X syndrome protein targets tandem repeats and influences allele-specific expression in a size-dependent manner. *Cell* **143**: 367-378.

Mendez J, Stillman B. 2000. Chromatin association of human origin recognition complex, cdc6, and minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late mitosis. *Mol Cell Biol* **20**: 8602-8612.

Privitera E, Schiro R, Longoni D, Ronchi A, Rambaldi A, Bernasconi S, Ottolenghi S, Masera G, Biondi A. 1995. Constitutive expression of GATA-1, EPOR, alpha-globin, and gamma-globin genes in myeloid clonogenic cells from juvenile chronic myelocytic leukemia. *Blood* **86**: 323-328.

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W et al. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**: R137.

# Identification and characterization of two novel primate-specific histone H3 variants, H3.X and H3.Y

Sonja M. Wiedemann, Silke N. Mildner, **Clemens Bönisch**, Lars Israel, Andreas Maiser, Sarah Matheisl, Tobias Straub, Rainer Merkl, Heinrich Leonhardt, Elisabeth Kremmer, Lothar Schermelleh and Sandra B. Hake

This article has original data in the JCB Data Viewer

http://jcb-dataviewer.rupress.org/jcb/browse/2505

JCB: Article

# Identification and characterization of two novel primate-specific histone H3 variants, H3.X and H3.Y

Sonja M. Wiedemann,[1] Silke N. Mildner,[1] Clemens Bönisch,[1] Lars Israel,[2] Andreas Maiser,[3] Sarah Matheisl,[1] Tobias Straub,[1] Rainer Merkl,[4] Heinrich Leonhardt,[3,6] Elisabeth Kremmer,[5] Lothar Schermelleh,[3] and Sandra B. Hake[1,6]

[1]Adolf-Butenandt-Institute, Department of Molecular Biology, and [2]Protein Analysis Unit, Ludwig Maximilians University of Munich, 80336 Munich, Germany
[3]LMU Biocenter, Department of Biology, Ludwig Maximilians University of Munich, 82152 Planegg-Martinsried, Germany
[4]Department of Biochemistry II, University Regensburg, 93053 Regensburg, Germany
[5]Institute of Molecular Immunology, Helmholtz Center Munich, German Research Center for Environmental Health, 81377 Munich, Germany
[6]Center for Integrated Protein Science Munich 5, 81377 Munich, Germany

Nucleosomal incorporation of specialized histone variants is an important mechanism to generate different functional chromatin states. Here, we describe the identification and characterization of two novel primate-specific histone H3 variants, H3.X and H3.Y. Their messenger RNAs are found in certain human cell lines, in addition to several normal and malignant human tissues. In keeping with their primate specificity, H3.X and H3.Y are detected in different brain regions. Transgenic H3.X and H3.Y proteins are stably incorporated into chromatin in a similar fashion to the known H3 variants. Importantly, we demonstrate biochemically and by mass spectrometry that endogenous H3.Y protein exists in vivo, and that stress stimuli, such as starvation and cellular density, increase the abundance of H3.Y-expressing cells. Global transcriptome analysis revealed that knockdown of H3.Y affects cell growth and leads to changes in the expression of many genes involved in cell cycle control. Thus, H3.Y is a novel histone variant involved in the regulation of cellular responses to outside stimuli.

## Introduction

Chromatin, the storage and regulatory form of genetic information in eukaryotes, consists of nucleosomes that are composed of DNA and octamers of the core histones H2A, H2B, H3, and H4 (van Holde, 1988). To allow changes in chromatin structure, which are necessary to promote different biological functions, several interconnected mechanisms have evolved (for review see Bönisch et al., 2008). Among others, these include the sliding or eviction of nucleosomes by ATP-dependent chromatin remodeling machines (for review see Clapier and Cairns, 2009), posttranslational modifications (PTMs) of histone proteins (Strahl and Allis, 2000), and the exchange of canonical histones with specialized histone variants (for reviews see Pusarla and Bhargava, 2005; Bernstein and Hake, 2006). Histone variants

differ in sequence and expression timing from their canonical counterparts and are enriched in chromatin of specific functional states, ranging from DNA repair and centromere determination to the regulation of gene expression. In mammals, variants of the H3, H2A, and H2B families of histones have been identified whose incorporation results in nucleosomes with novel functional and structural properties (Suto et al., 2000; Abbott et al., 2001; Angelov et al., 2003; Bao et al., 2004; Gautier et al., 2004).

To date, five different H3 variants have been found in mammals: H3.1, H3.2, H3.3, H3.1t (tH3), and CENP-A. The centromeric H3 variant CENP-A causes changes to the nucleosomal structure (Black et al., 2004) and is crucial for proper chromosome segregation (for review see Allshire and Karpen, 2008). tH3 is a testis-specific histone variant with a putative

The Rockefeller University Press   $30.00
J. Cell Biol. Vol. 190 No. 5   777–791
www.jcb.org/cgi/doi/10.1083/jcb.201002043

JCB   777

THE JOURNAL OF CELL BIOLOGY

function in chromatin reorganization during spermatogenesis (Witt et al., 1996). H3.1 and H3.2 sequences are distinguishable by just one amino acid. Although expression of both is replication dependent (Ahmad and Henikoff, 2002a), they differ in their cell type expression levels as well as their enrichment of PTMs (Hake et al., 2006). Furthermore, H3.1 has been implicated in DNA damage response pathways (Polo et al., 2006) and is deposited into chromatin by the chaperone complex CAF-1 (Tagami et al., 2004), whereas H3.3 is expressed and incorporated into chromatin in a replication-independent manner by HIRA (Tagami et al., 2004). The latter variant is highly decorated with modifications associated with gene transcription (McKittrick et al., 2004; Hake et al., 2006) and is thought to be involved in activating gene expression (Ahmad and Henikoff, 2002a) and epigenetic reprogramming (for review see Santenard and Torres-Padilla, 2009).

Here, we describe the identification of two novel primate-specific histone H3 variants (*H3.X* and *H3.Y*) that are transcribed at low levels in some human cell lines and in normal and malignant tissues. Their amino acid sequences are highly similar to each other, but differ in several functionally important residues from other H3 variants, e.g., S10 and S28. Transgenic H3.X and H3.Y proteins localize to the nucleus and are stably incorporated into chromatin. We demonstrate that endogenous H3.Y protein is highly expressed in some osteosarcoma cells. Interestingly, the number of cells expressing H3.Y is increased by particular stress stimuli. Knockdown of H3.Y leads to diminished growth and changes in the expression of many genes controlling cell cycle progression. The observed presence of primate-specific histone H3 variants in specific human brain areas opens up fascinating questions about their roles in human cell determination and differentiation.

## Results

### Identification of novel H3 variant genes

Using the nucleotide sequence of human histone *H3.1f* (HIST1H3I; available from GenBank/EMBL/DDBJ under accession no. NM_003533), we searched the public database (National Center for Biotechnology Information) and identified two highly similar genes initially annotated as pseudogenes (Fig. S1 A). These two intron-free genes, which we named *H3.X* (GenBank/EMBL/DDBJ accession no. LOC340096) and *H3.Y* (GenBank/EMBL/DDBJ accession no. LOC391769), are located on human chromosome 5 (5p15.1). Further database searches revealed the existence of similar genes in primate genomes (*Pan troglodytes* H3.X, GenBank/EMBL/DDBJ accession no. LOC471464; and H3.Y, GenBank/EMBL/DDBJ accession no. LOC471473; *Macaca mulatta* H3.X, GenBank/EMBL/DDBJ accession no. LOC718189; and H3.Y, GenBank/EMBL/DDBJ accession no. LOC718280; Fig. S1 B). Searches for these genes in other mammalian genomes yielded no positive hits, which suggests that they evolved in evolutionarily younger terms and might constitute primate-specific histones. Both human genes contain a sequence matching the translation initiation start site consensus (underline) for vertebrates (GCC-GCCACCAUGGCG; Kozak, 1991; Nakagawa et al., 2008), and

depending on the search program used ("polyadq" or PolyA_SVM program), *H3.X* and *H3.Y* 3′ genomic sequences of human and primate origins were predicted to include a conserved poly-A site (Tabaska and Zhang, 1999; Cheng et al., 2006). Alignment of human and primate *H3.X* and *H3.Y* coding sequences with respective human *H3* variant sequences revealed a higher similarity to *H3.3* than to *H3.2* and *H3.1* (Fig. S1 C).

*H3.X* and *H3.Y* genes are predicted to encode proteins of 146 and 135 amino acids, respectively (Fig. 1 A). Both putative variant proteins are highly similar to each other, with differences of only four amino acids in their overlapping region (89.7% identity). H3.X has an unusual long C-terminal tail with no sequence homology to other proteins (Fig. 1 A). H3.X and H3.Y display interesting changes in amino acids that are known to be posttranslationally modified in H3.1, H3.2, and H3.3: serine 10 and 28, which are phosphorylated during mitosis (Hendzel et al., 1997) and immediate early gene induction (Clayton and Mahadevan, 2003), are altered to alanine and arginine, respectively. Regions surrounding lysine 14, usually acetylated and found in actively transcribed genes (Yan and Boyd, 2006), and lysine 79, methylated in transcriptionally active regions (Im et al., 2003) and upon DNA damage (Huyen et al., 2004), are completely altered (Fig. 1 A). In several instances, single amino acid changes between novel and known H3 variants have occurred, thereby generating or eliminating potential PTM sites. The region surrounding aa 87–90 is important for the chaperone-dependent chromatin incorporation of histone variants (Ahmad and Henikoff, 2002b). CAF-1 deposits H3.1, whereas HIRA catalyzes the incorporation of H3.3 (Tagami et al., 2004). The residues of H3.X and H3.Y in this region are identical to H3.3 (Fig. 1 A), leading to the hypothesis that these putative novel variants might be incorporated by HIRA in a replication-independent manner.

To assess whether both genes are expressed in human cells, we analyzed their RNA expression levels in a variety of cell lines by quantitative PCR (qPCR; Fig. 1 B). Because the nucleotide sequences of *H3.X* and *H3.Y* are almost identical (Fig. S1 A), we generated one primer set detecting both H3.X and H3.Y mRNAs and one additional primer set that hybridizes with H3.X-specific 3′ sequences. We demonstrate that both *H3.X* and *H3.Y* genes are transcribed at low levels in the human osteosarcoma cell line U2OS, with *H3.Y* in higher amounts than *H3.X*, but ∼16-fold lower than the histone variant CENP-A (Fig. S1 D). Minimal expression of *H3.Y* was detected in human HEK293 cells, whereas no expression was seen in HeLa as well as in mouse cells (negative control; Fig. 1 B). These data suggest that *H3.X* and *H3.Y* genes are transcribed in some human cell lines, albeit at low total levels.

Our observation that H3.X and H3.Y mRNAs are expressed in U2OS cells inspired us to examine RNAs from different human tissues by qPCR. Of particular interest were human malignant tissues, as U2OS cells are derived from a moderately differentiated sarcoma of the tibia. We also chose to investigate several regions of the human brain and testis due to the primate-specific appearance of *H3.X* and *H3.Y*. Interestingly, we found that some bone, breast, lung, and ovary tumors express low but significant levels of H3.X/Y mRNA (Fig. 1 C). Analyzing RNA from human brain areas and testis, H3.X/Y expression was
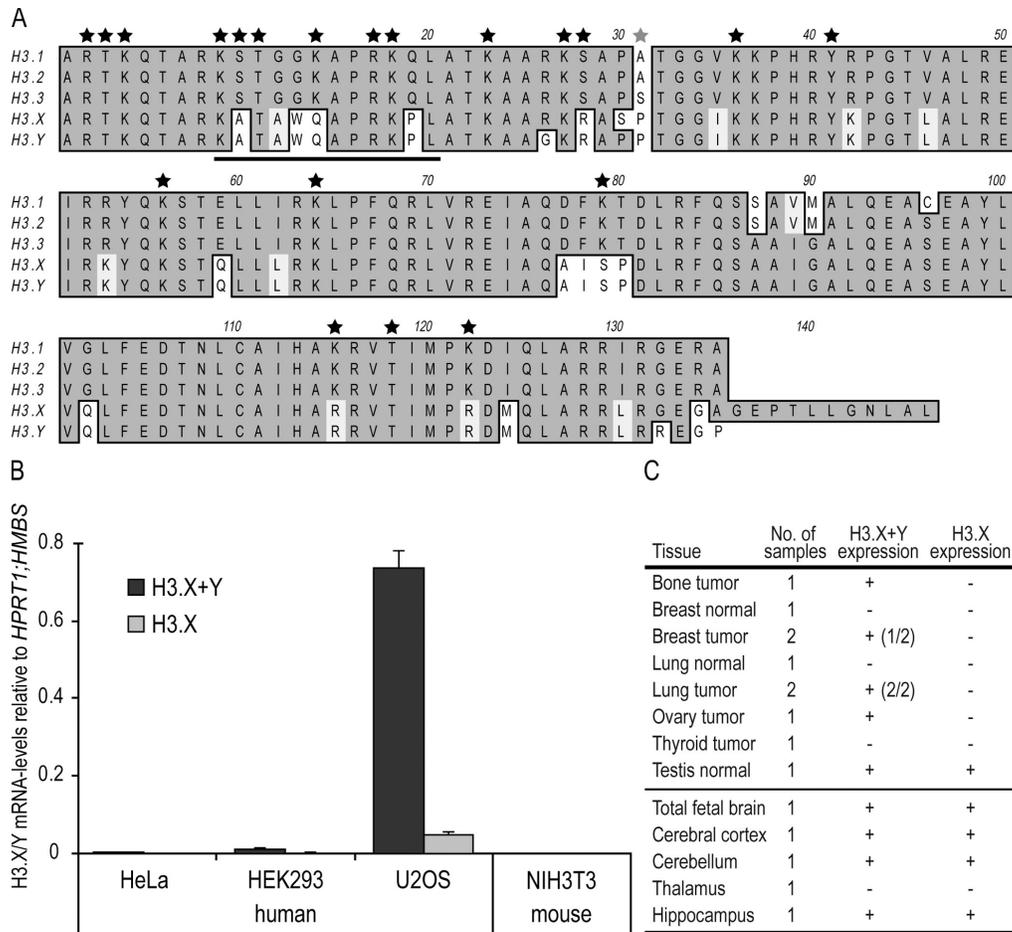
## A



## B



## C



Figure 1. **Sequence and mRNA expression analysis of novel H3 variants H3.X and H3.Y.** (A) Amino acid sequence alignment of human histone variants H3.1, H3.2, and H3.3 with novel variants H3.X and H3.Y. Alignments were made with ClustalW Alignment (MacVector 10.0.2). Identical amino acids are highlighted in dark gray, similar amino acids are highlighted in light gray, and changes are set apart on a white background. The black bar indicates the peptide sequence used for antibody generation. Black stars mark known PTM sites in H3.1, H3.2, and H3.3. The gray star indicates an H3.3-specific modification site. Amino acid numbers are indicated on top. (B) qPCR analysis with cDNA from different human cell lines shows expression of H3.Y and to a lesser extent H3.X mRNA in U2OS cells. Primer pair H3.X+Y (dark gray) specifically recognizes H3.X and H3.Y nucleotide sequences, whereas another primer pair is H3.X specific (H3.X, light gray). NIH3T3 mouse cDNA was used as a negative control. Data were normalized to HPRT1 and HMBS expression levels. Controls generated without reverse transcription were used to assess amplification threshold. Error bars represent SEM of two independent biological experiments. (C) Expression of H3.X and H3.Y mRNA in normal and malignant human tissues. Commercially available total RNA from human tissues was analyzed for H3.X+Y and H3.X expression in qPCR experiments and compared with results obtained with controls generated without reverse transcription. The number of samples that were positive for H3.X+Y and H3.X expression (+) are indicated in brackets.

observed in all samples except for the thalamus. In contrast to all other tissues tested, H3.X mRNA could also be detected in the brain and testis samples (Fig. 1 C).

In conclusion, mRNAs of the novel primate-specific H3 variants H3.X and H3.Y are not only present in human cell lines, but are also detected in primary cells from different human organs.

### Tagged H3.X and H3.Y proteins are similar to canonical histone H3 in their nuclear localization, nucleosome structure, and exchange mobility kinetics

The detection of H3.X and H3.Y mRNA in human cells prompted us to investigate their protein products. We amplified their coding sequences from U2OS cDNA and cloned them under the constitutively active cytomegalovirus promoter in frame with an N-terminal HA or GFP tag. These constructs

were transfected into HeLa cells and, in the case of HA-tagged variants, at least two independent stable cell populations were selected. In both HA-H3.X– and HA-H3.Y–expressing HeLa cells, nuclear HA signals in interphase and staining of condensed chromatin in mitotic cells were observed by immunofluorescence (IF) microscopy (Fig. 2 A). Signals were also present on metaphase chromosome spreads (Fig. 2 B), which suggests that these novel tagged H3 variants are stably incorporated into chromatin and constitute core nucleosomal components. Next, we asked if the amino acid differences between H3.X, H3.Y, and other H3 variants have any impact on nucleosome structure. Thus, we calculated in silico models of H3.X and H3.Y protein structures, using the published structure of *Xenopus laevis* H3.2 as a template (Davey et al., 2002). With these, we computationally assembled a nucleosome containing H3.X or H3.Y. Our data show only few differences to the published structure of H3.2, with an exception being the unusual C terminus of H3.X
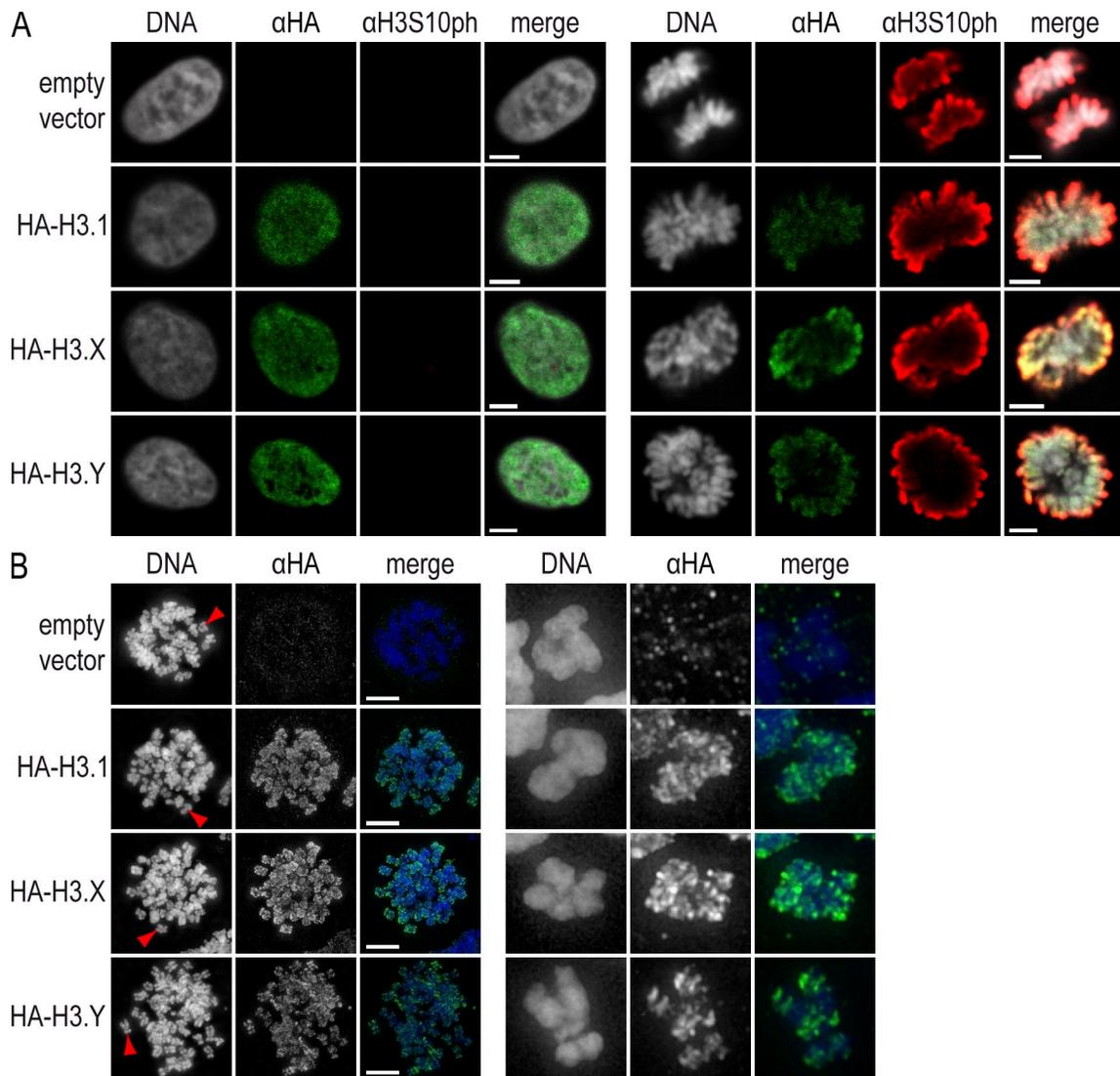
Figure 2. **Subcellular localization of HA-H3.X and -H3.Y histone proteins.** (A) Confocal imaging of stably transfected HeLa cells shows nuclear localization of HA-H3.1, -H3.X, and -H3.Y proteins. Cells were costained with TO-PRO3 (DNA, gray, left), α-HA (green, middle left), and α-H3S10ph (red, middle right). Overlay is shown on the right (merge). The left side shows interphase and the right side shows mitotic cells, as indicated by α-H3S10ph staining. (B) Metaphase spreads of mitotically arrested HeLa cells transfected with empty vector and HA-H3.1, -H3.X, and -H3.Y. Deconvolved wide-field images of chromosomes costained with DAPI (DNA, blue, left) and α-HA (green, middle). Sections containing one chromosome (red arrowhead) stained with α-HA are depicted on the far right. All HA-H3 variants are incorporated into chromosomes. Bars: (A) 5 μm; (B) 10 μm.

(Fig. 3 A). Surprisingly, this extended C-terminal tail of H3.X is proposed to fold into an additional α-helix that fits into the free space of the nucleosomal core structure (Fig. 3 B, left).

Because our in silico models proposed no dramatic changes in nucleosomal structure for H3.Y, but some differences for H3.X, we asked whether H3.X and/or H3.Y are incorporated into nucleosomes in a cellular context and whether they exhibit a similar stability compared with nucleosomes containing H3.1, H3.2, or H3.3. Immunoprecipitation (IP) of mononucleosomes generated from HeLa cells expressing HA-tagged H3 variants using an α-HA antibody confirmed the incorporation of HA-H3.X and -H3.Y into chromatin (Figs. 3 C and S2 A). Additionally, silver staining and immunoblots revealed that HA-H3.X and -H3.Y are present in nucleosomes containing all four core histones. Interestingly, HA-H3.X and -H3.Y containing mononucleosomes also contained endogenous H3 (Fig. 3 C, right), which suggests

that, unlike H3.1 and H3.3 (Tagami et al., 2004), these tagged novel variants form heterotypic nucleosomes with regard to their variant composition. To assess the incorporation stability of these novel variants, we performed FRAP experiments with GFP-tagged H3 variants transiently transfected into HeLa Kyoto cells, which were chosen because of their slow cell motility. These experiments demonstrate similar dissociation kinetics for the novel GFP-tagged H3 variants compared with GFP-H3.1 and -H3.3 (Fig. 3, D and E; and Fig. S2, B–D). GFP fluorescence recovery rates of all H3 variants were very slow (>8 h; Fig. S2, B–D) in comparison to GFP alone (a few seconds; Fig. 3, D and E), which is indicative of stable chromatin-bound immobile populations.

In summary, tagged H3.X and H3.Y protein containing nucleosomes are similar in their predicted structure and indistinguishable from nucleosomes containing H3.1 or H3.3 in their FRAP mobility and chromatin-association behavior.
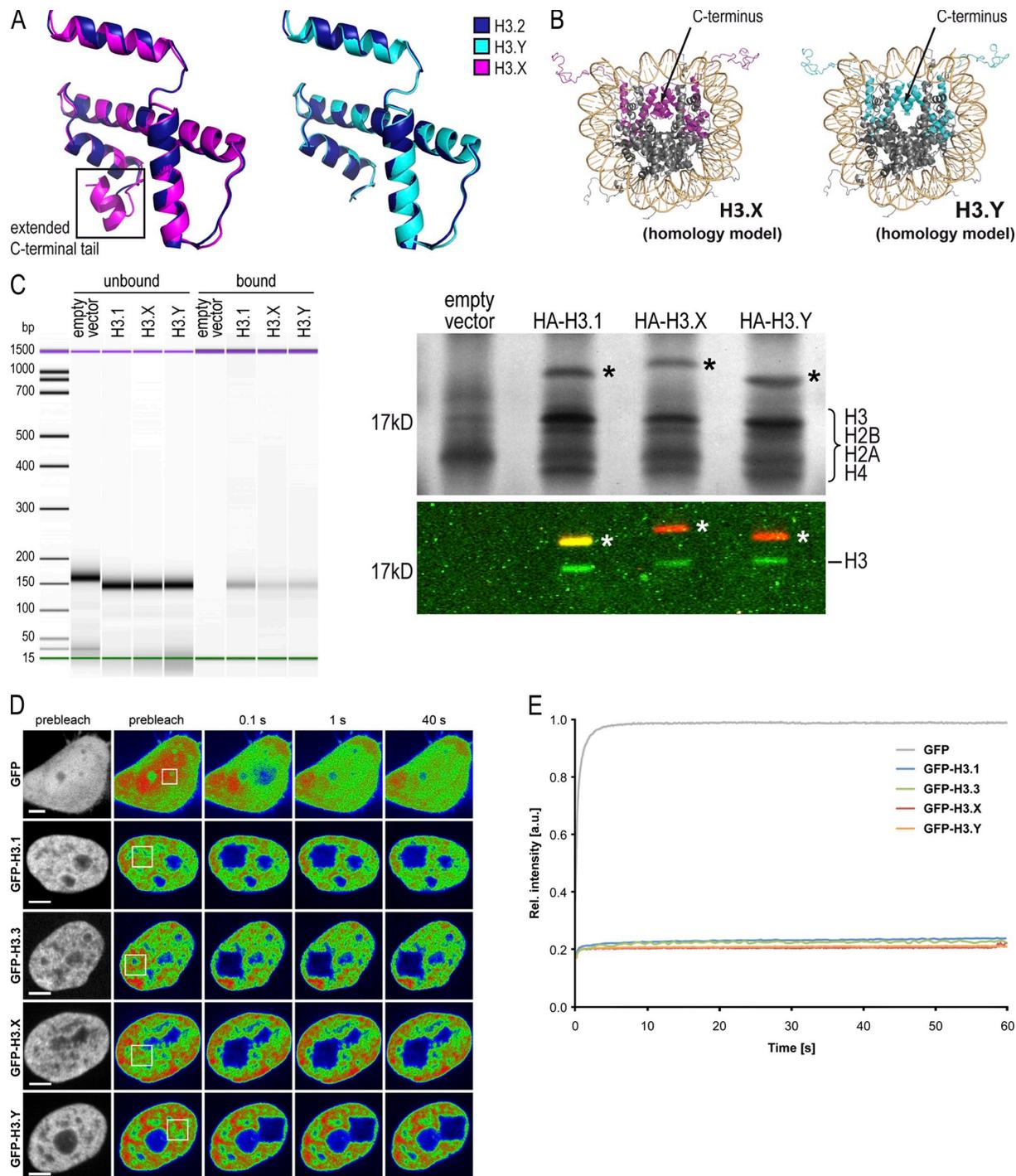
Figure 3. **Structure and stability of H3.X- and H3.Y-containing nucleosomes.** (A) In silico homology model of H3.X (purple, left) and H3.Y (light blue, right) protein structures in overlay with the crystal structure of H3.2 (dark blue). (B) Crystal structure of nucleosome with H3.2 exchanged by in silico homology models of H3.X (purple, left) and H3.Y (light blue, right), respectively. (C) IP of mononucleosomes generated from HeLa cells transfected with empty vector, HA-H3.1, -H3.X, and -H3.Y shows incorporation of novel H3 variants into nucleosomes. Bioanalyzer evaluation of purified DNA after IP of MNase-treated chromatin (unbound and bound material) shows digestion of chromatin to mononucleosomes and their successful precipitation (left; see also Fig. S2 A for DNA size and quality). Silver stain of 15% SDS-PAGE with α-HA IPs of mononucleosomes revealed successful binding of HA-tagged H3 variants (asterisks) and pull-down of core histones (top, right). Immunoblot of immunoprecipitates with α-HA (red) and α-H3 C-terminal (green) antibodies visualized by the Odyssey infrared imaging system (bottom, right). Notice that endogenous H3 is coimmunoprecipitated with all H3 variants analyzed. (D) FRAP experiments to evaluate nucleosomal stability of novel H3 variants using spinning disk confocal microscopy. HeLa Kyoto cells were transiently transfected with GFP, GFP-H3.1, -H3.3, -H3.X, and -H3.Y constructs. A small nuclear area was photobleached (box) and the recovery of the fluorescent signal was monitored over 1 min and up to 8 h (see Fig. S2, B–D, for long-term FRAP). Depicted is a short-term FRAP series (selected time points are shown) of GFP-tagged H3 variants compared with GFP alone. Bar, 5 µm. (E) Quantification of short-term FRAP experiment. Mean curves of 10–20 individual cells are shown. Standard deviations were very small (in the range of ± 0.02) and were omitted for clarity (for details see Fig. S2 D). All GFP-H3 variants show almost no recovery within the first 60 s after bleaching, which indicates that all expressed fusion protein was stably incorporated into nucleosomes. In contrast, GFP alone recovers to almost 100% within 5 s.

### Endogenous H3.X/Y proteins are expressed in U2OS cells and localize to the nucleus

To characterize endogenous H3.X/Y proteins, we generated a monoclonal antibody against these variants (α-H3.X/Y; see Fig. 1 A, black line, for peptide sequence). α-H3.X/Y specificity was tested in diverse experiments using HeLa cells stably expressing HA-tagged H3 variants. Immunoblot analysis of acid-extracted histones from these cell lines revealed that α-H3.X/Y specifically recognizes HA-H3.X and -H3.Y, but not HA-H3.1, -H3.2, or -H3.3 (Fig. 4 A). In IF microscopy analyses with this antibody, HA-H3.X and -H3.Y showed an exclusive nuclear staining (Fig. S2 E), which could be confirmed to be specific for H3.X/Y by peptide competition assays (Fig. S2 F).

To analyze the expression of endogenous H3.X and H3.Y proteins, we isolated total histones from different cell lines. As a negative control, we used mouse NIH3T3 cells and, as a positive control, HeLa cells expressing HA-H3.X. Immunoblotting of these histones with α-H3.X/Y showed a faint band in the lane containing U2OS histones corresponding to the size of H3, which would be expected for H3.Y (Fig. 4 B). These data are consistent with our previous finding that U2OS cells express mostly H3.Y mRNA (Fig. 1 B). Next, we addressed the subcellular localization of endogenous H3.X and/or H3.Y in U2OS cells by confocal IF microscopy. α-H3.X/Y shows a faint and dotted staining associated with the nuclear rim in all cells, which might be due to cross-reactivity with nuclear pores. Interestingly, we also noticed that few U2OS cells showed a strong nuclear α-H3.X/Y staining, colocalizing with DNA (Fig. 4 C, arrowhead), which was not observed in cell lines of mouse or rat origin (Fig. 4 D). Some few mitotic U2OS cells, as well as metaphase chromosomes, also stained positive with α-H3.X/Y (Fig. 4, E and F, respectively), which suggested that the observed α-H3.X/Y signal marks a stable chromatin component. To shed light on the subnuclear localization of endogenous H3.X/Y in U2OS cells, we performed super-resolution (i.e., below the diffraction limit of conventional optical microscopy) imaging with 3D structured illumination microscopy (3D-SIM; Gustafsson et al., 2008; Schermelleh et al., 2008). As depicted in Fig. 4 G, H3.X/Y is predominantly located outside of DAPI-dense regions, arguing for an association of endogenous H3.X/Y with less condensed, more euchromatic regions.

### Nutritional- and growth-associated stress stimuli increase the number of H3.X/Y-expressing cells

Because only ∼0.1% of U2OS cells showed a general nuclear staining with α-H3.X/Y, we wondered what cellular features distinguished these from other cells. We hypothesized that specific stress stimuli induce the nuclear α-H3.X/Y staining. To test this assumption, we induced DNA damage with different treatments (UV, etoposide, and hydroxyurea) but were unable to detect a significant increase in the percentage of cells positive for α-H3.X/Y staining (unpublished data). Next, we addressed the question of whether cell growth and nutrition could have an impact on H3.X/Y expression. Interestingly, induction of starvation and overgrowth (SO) in U2OS cells for 8 d, but not overgrowth alone (for details, see Materials and methods), led

to an increase in the percentage of cells with nuclear α-H3.X/Y staining (unpublished data). To evaluate if starvation alone, or in combination with high cell density, caused this increase in α-H3.X/Y–positive cell numbers, we either subjected U2OS cells to starvation or cultivated them under SO conditions. In three independent biological experiments, we quantified (1) the percentage of α-H3.X/Y nuclear-stained cells by IF microscopy, (2) H3.X/Y mRNA levels by qPCR, and (3) H3.X/Y protein levels by immunoblot analyses. Interestingly, we observed an approximately sixfold increase in the number of α-H3.X/Y nuclear-stained cells under SO conditions, whereas starvation alone only showed a minor effect (threefold increase; Fig. 5 A). In addition, we could reproducibly show, by qPCR (Fig. 5 B) and immunoblots (Fig. 5 C), that both H3.X/Y RNA and protein are up-regulated upon SO treatment compared with normally growing cells. We observed a stronger increase of H3.X/Y- than H3.X-specific amplicons (Fig. 5 B), which suggests that H3.Y mRNA expression is specifically affected by this particular stress condition. This was further validated by a "primer-walk" experiment with several different primer pairs that span the entire H3.X or H3.Y coding sequence (Fig. S3, A–C). Additionally, the induced protein band observed in immunoblots corresponds to a protein of the size of recombinant H3.Y but not H3.X (Fig. 5 C). To assess the question of which factors in the growth medium cause the observed effect, we also analyzed α-H3.X/Y nuclear stained cells after serum starvation and amino acid depletion. However, we did not observe any significant difference caused by both treatments (unpublished data).

These data show a direct correlation between α-H3.X/Y antibody staining and the increased mRNA and protein levels due to SO treatment. Furthermore, they suggest that H3.Y expression, in particular, is positively affected by these growth conditions.

### Endogenous H3.Y protein is expressed after SO treatment

Our experiments indicate that H3.Y expression is mostly, if not exclusively, induced in U2OS cells under SO conditions. To confirm this observation, we purified histones from HEK293, NIH3T3, and SO-treated and normally grown U2OS cells by reversed-phase HPLC (RP-HPLC; Fig. S4 A). Fractions surrounding and including peaks corresponding to H3.2 and H3.3 (peak I) and H3.1 (peak II; Fig. 6 A) were analyzed by immunoblots with α-H3.X/Y. Specific signals were observed in fractions eluting at the beginning of peak I (region A bands) and between peak I and II (region B bands; Fig. 6 B). Interestingly, region A bands correspond to proteins of the molecular weight of histone H3, the predicted size of H3.Y. As expected for H3.Y, these bands are faintly visible in normally grown U2OS cells and strongly enhanced in those cultured under SO conditions. In contrast, no such bands are present in NIH3T3 and HEK293 cells. Region B bands are indicative of proteins of ∼35 kD and are clearly visible without any change in intensity in all human cell lines, but not in mouse NIH3T3 cells. Fractions corresponding to region A and B bands from SO treated cells from two independent experiments and normally grown U2OS cells were analyzed by liquid chromatography tandem mass spectrometry (LC-MS/MS). For region B fractions, no conclusive information
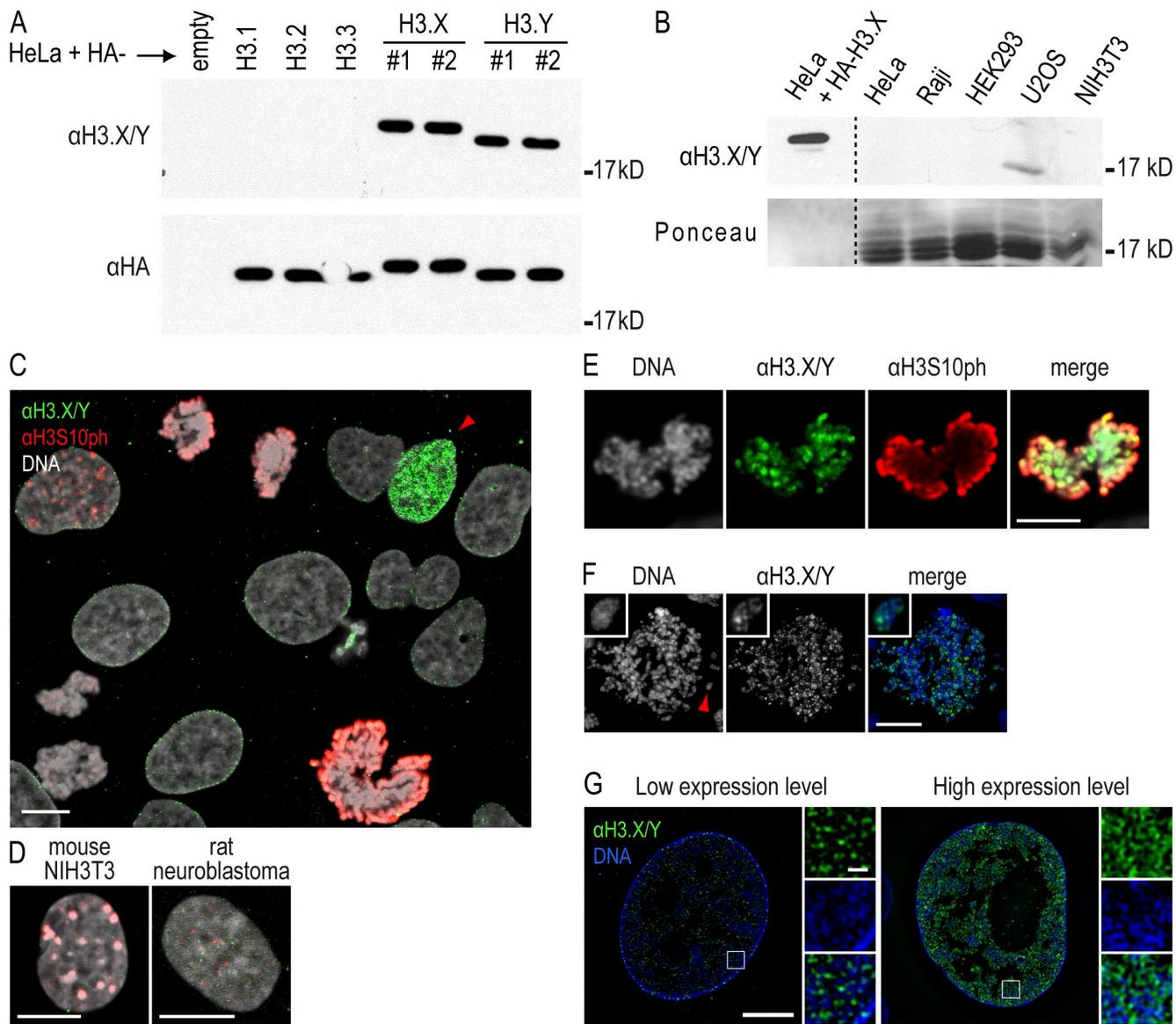
Figure 4.  **Detection of endogenous H3.X and/or H3.Y proteins.** (A) A monoclonal antibody against H3.X/Y (α-H3.X/Y) was generated by immunizing rats with an N-terminal peptide specific for H3.X and H3.Y (aa 9–20, see also black line in Fig. 1 A). Immunoblots of acid-extracted histones from HeLa cells stably transfected with empty vector, or vectors containing HA-H3.1, -H3.2, -H3.3, -H3.X, and -H3.Y. Histones from two independently selected HeLa cell populations expressing HA-H3.X and -H3.Y were used (#1 and #2). (A, top) Staining of the membrane with α-H3.X/Y antibody shows only signals in lanes loaded with histones from HeLa cells expressing HA-H3.X and -H3.Y, but not in lanes containing general HA-H3 variants, which demonstrates the specificity of the antibody toward the novel variants in immunoblotting. (A, bottom) Equal loading was controlled by α-HA staining. Note that all bands run slower, as expected, because of the HA tag. HA-H3.X runs even slower than all other HA-H3 variants because of its extended C-terminal tail. (B) Immunoblot analysis of acid-extracted histones from different cell lines with α-H3.X/Y antibody. (B, top) Histones from HeLa cells expressing HA-H3.X served as positive control, and histones from mouse NIH3T3 cells served as a negative control. A faint signal in the lane containing U2OS histones can be seen. (B, bottom) The identical membrane was stained with Ponceau S solution before antibody incubation to control for protein loading. Dotted lines indicate that intervening lanes have been spliced out. (C) Confocal IF analysis of U2OS cells costained with α-H3.X/Y (green), α-H3S10ph (mitosis-specific, red), and TO-PRO3 (DNA, gray). Confocal midsections are shown. (D) Confocal IF analysis of mouse (left) and rat (right) cells costained with α-H3.X/Y (green), α-H3S10ph (red), and TO-PRO3 (DNA, gray) as negative controls. (E) Confocal IF analysis of mitotic U2OS cells costained with α-H3.X/Y (green), α-H3S10ph (mitosis-specific, red), and TO-PRO3 (DNA, gray). (F) Costaining of metaphase chromosomes derived from mitotically arrested U2OS cells with α-H3.X/Y (green) and DAPI (DNA, blue). The boxed inlet shows enlargement of one chromosome. (G) Super-resolution 3D-SIM imaging of U2OS cells costained with α-H3.X/Y (green) and DAPI (DNA, blue). Depicted are cells expressing low (left) and high (right) levels of H3.X/Y. Bars: (C–F) 10 µm; (G, left) 5 µm; (G, right) 0.5 µm.

about the protein identity could be obtained. Only one single peptide corresponding to H3.X/Y (aa 53–63) could be observed in B-type fractions, rendering the analysis of these particular fractions inconclusive. To eliminate the possibility that region B bands constitute a dimeric aggregate caused by formation of disulfide bonds between cysteines, we treated these fractions

with increasing amounts of DTT (Fig. S4 B). In contrast to H3.1, H3.2, and H3.3, we did not observe a change in molecular weight for region B bands, which suggests that these particular proteins are not dimers of H3.X and/or H3.Y.

Interestingly, in region A fractions, we were able to identify two peptides specific for H3.X/Y (aa 53–63 and 73–83), in
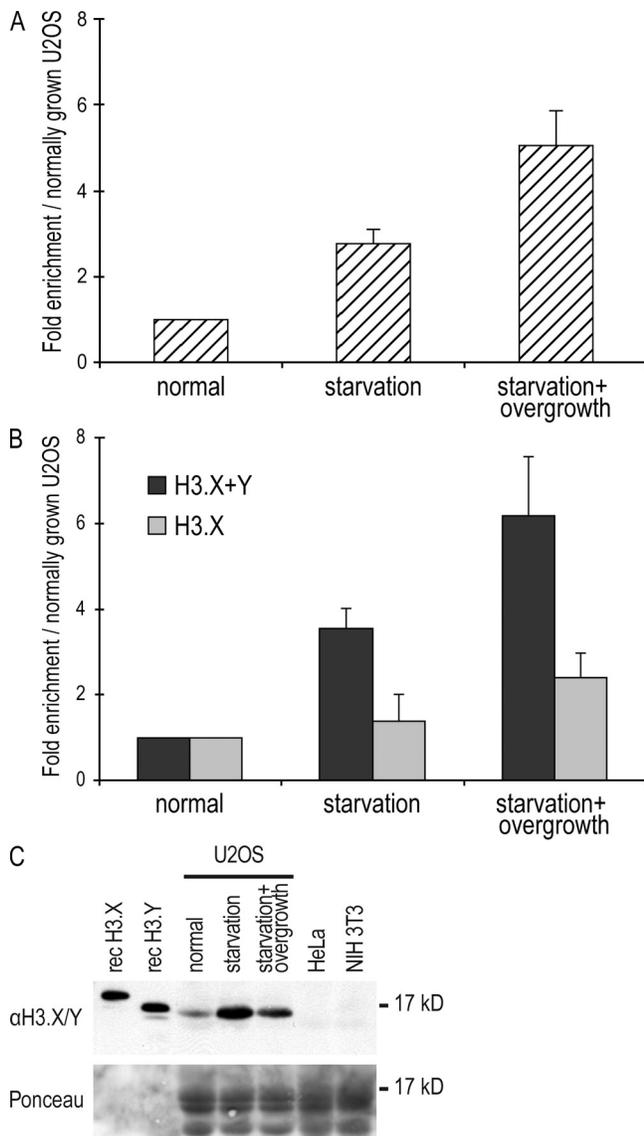
Figure 5. **The number of cells expressing H3.X and/or H3.Y is increased by nutritional and proliferative stress.** (A) Quantification of the percentage of U2OS cells positive for α-H3.X/Y nuclear staining by IF analyses under different growth conditions. U2OS cells were stained with α-H3.X/Y after 8 d of growth under different conditions: normal, starvation, or SO. The percentage of cells positive for α-H3.X/Y nuclear IF staining was determined (see Materials and methods for details) and plotted. A clear increase in the number of cells containing α-H3.X/Y–positive nuclei can be seen after growth under SO conditions. Error bars represent SEM of three independent biological experiments. (B) qPCR analysis of H3.X and H3.X+Y mRNA expression levels from U2OS cells grown under different conditions. From the same plates described in Fig. 5 A, cells were harvested, RNA was isolated, and cDNA was generated. A clear increase in H3.X+Y (dark gray) but not H3.X mRNA (light gray) under SO conditions can be detected, which is similar to the data obtained with IF analyses (Fig. 5 A). Error bars represent SEM of three independent biological experiments. (C) Immunoblot analysis of H3.X and/or H3.Y proteins isolated from U2OS cells grown under different conditions. From the same plates described in Fig. 5 A, cells were harvested, and histones were acid extracted and immunoblotted with α-H3.X/Y antibody (top). A clear increase in a 17-kD signal can be seen in U2OS cells grown under different stress conditions. Recombinant H3.X and H3.Y proteins serve as positive controls and histones from human HeLa and mouse NIH3T3 cells serve as negative controls. Staining of the same membrane with Ponceau S solution before antibody incubation was performed to ensure similar loading (bottom). One representative blot from three independent biological experiments is shown.

addition to the expected general H3 peptides (Figs. 6 C and S4, C and D). Furthermore, one peptide exclusively present in H3.Y but not H3.X (aa 18–28) was repeatedly detected (Figs. 6 C and S4 E), leading to a combined sequence coverage of 33% for H3.Y. Interestingly, in conjunction with unmodified H3.Y peptide, we were also able to identify acetylation of lysines 18, 23, and 27 (Fig. S4, F and G). These PTMs are also present in the known H3 variants, indicating that H3.Y was a likely part of a nucleosome and present in chromatin fibers.

In summary, these data show that posttranslationally modified H3.Y protein is expressed in vivo, and that under stress-inducing conditions, more H3.Y-expressing cells are present.

### Knockdown of H3.Y affects expression of cell cycle–related genes and cell growth

Because we could clearly demonstrate the presence of endogenous H3.Y protein in human cells, we wondered about the biological consequences of its expression. We hypothesized that the natural substitution of amino acids S10, K14, S28, and K79 in H3.Y and its preferential exclusion from DAPI-dense regions (Fig. 4 G) might result in gene expression changes at the sites of chromatin incorporation. Therefore, we used global expression arrays to examine a potential impact on the regulation of gene expression by H3.Y. We used RNAi to significantly reduce H3.Y and H3.X+H3.Y mRNAs and proteins. siRNAs directed against H3.Y, H3.X+Y, and luciferase, as control, were designed and their specificity verified by transfection of HeLa cells expressing HA-tagged H3.X and H3.Y, followed by IF analysis (Fig. 7 A). Next, we sought to quantitatively compare global RNA expression profiles after RNAi in SO-treated U2OS cells using Affymetrix Human Gene 1.0 ST Arrays (performed twice with independent transfections on different days). We verified the proper knockdown of H3.Y (and H3.X) mRNAs by qPCR (Fig. 7 B) and observed a modest but significant deregulation of genes when compared with control luciferase RNAi in global microarray analyses (Fig. 7, C and D; and Tables S1 and S2). Cells transfected with H3.Y-specific siRNAs showed 293 genes up-regulated and 974 genes down-regulated. Cells treated with H3.X+Y-specific siRNAs had 1,106 genes up-regulated and 1,249 genes down-regulated (local false discovery rate cutoff of 0.2). To eliminate off-target and sole H3.X-specific effects and to focus only on H3.Y-specific responses, Venn diagrams of overlapping deregulated genes between H3.Y- and H3.X+Y-specific siRNA changes were used. They revealed up-regulation of 73 and down-regulation of 229 shared genes (Fig. 7 C). Gene ontology (GO) enrichment analysis of the deregulated genes classified shared down-regulated genes to mainly belong to cell cycle-controlling pathways, but also chromatin organization and metabolic pathways, whereas only few GO groups of shared up-regulated genes could be determined (Tables S1 and S2). Because the majority of genes affected by H3.Y RNAi play a role in cell cycle control (mitosis), as shown by a simplified GO enrichment analysis (Fig. 7 D), and we previously noticed a reduction in cell number after H3.Y and H3.X+Y knockdown, we quantitatively determined cell growth of U2OS cells with reduced H3.Y mRNA and protein levels. Therefore, we transfected U2OS cells with different variant-specific siRNAs and quantitatively monitored cell growth for 96 h using the xCELLigence system (Roche). Interestingly, and in
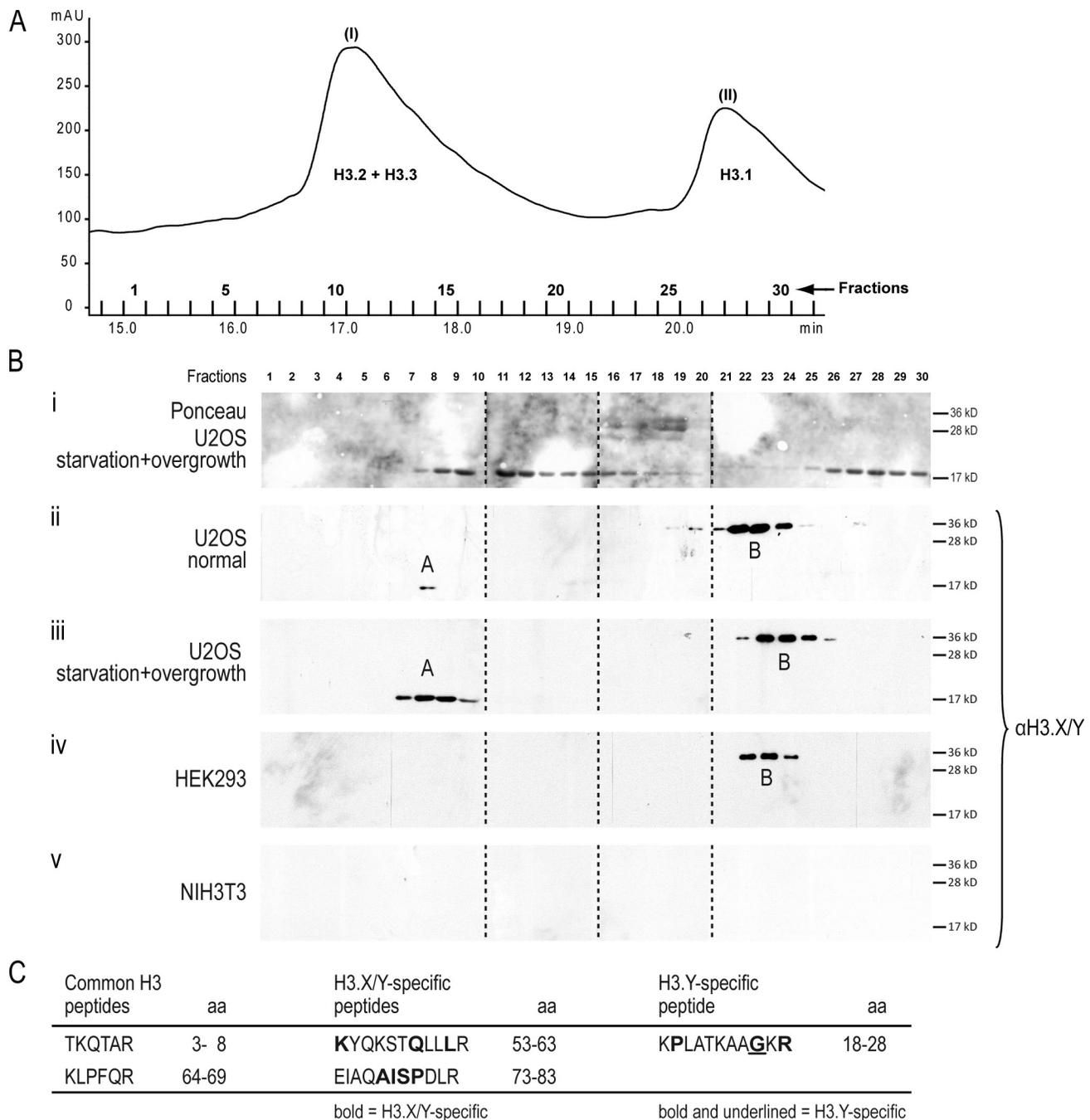
Figure 6. **Purification and identification of endogenous H3.Y variant protein.** (A) RP-HPLC section showing histone H3 peaks (see Fig. S4 A for complete RP-HPLC profile). Acid-extracted histones from starved and overgrown U2OS cells (Fig. 5) were separated by RP-HPLC, and histone H3 peaks (peak I, H3.2+H3.3; peak II, H3.1) spanning fractions 1–31 are shown. (B) Immunoblotting analyses of RP-HPLC fractions 1–30 spanning histone H3 peaks I to II from different cell lines under distinct growth conditions with α-H3.X/Y antibody. (i) Ponceau S staining of membrane containing histone fractions from starved and overgrown U2OS cells to detect H3.1, H3.2, and H3.3 proteins. Immunoblots incubated with α-H3.X/Y from RP-HPLC fractions from U2OS cells (ii), starved and overgrown U2OS cells (iii), HEK293 cells (iv), and mouse NIH3T3 cells (v). Dotted lines indicate that intervening lanes have been spliced out. The two anti-H3.X/Y–positive fractions are indicated with A and B. Proteins of both fractions were independently subjected to MS/MS analyses. (C) List of H3-, H3.X/Y-, and H3.Y-specific peptides identified by LC-MS/MS from combined band A–corresponding fractions from U2OS cells (normal and SO treated; see iii). Amino acids highlighted in bold are specific for H3.X and H3.Y; a bold and underlined amino acid is found only in H3.Y.

accordance with our global transcriptome analyses, knockdown of H3.Y and H3.X+Y, but not of control RNAi (luciferase), resulted in a significant reduction of cell growth (Fig. 7 E).

These results suggest that loss of H3.Y especially affects the expression of genes involved in cell cycle control, leading to diminished cell growth.

## Neuronal cell subpopulations in human hippocampus express novel H3 variant proteins

Encouraged by our findings that endogenous H3.Y protein exists in vivo and plays an essential role in cell growth and gene regulation, we wondered if H3.Y and/or H3.X proteins
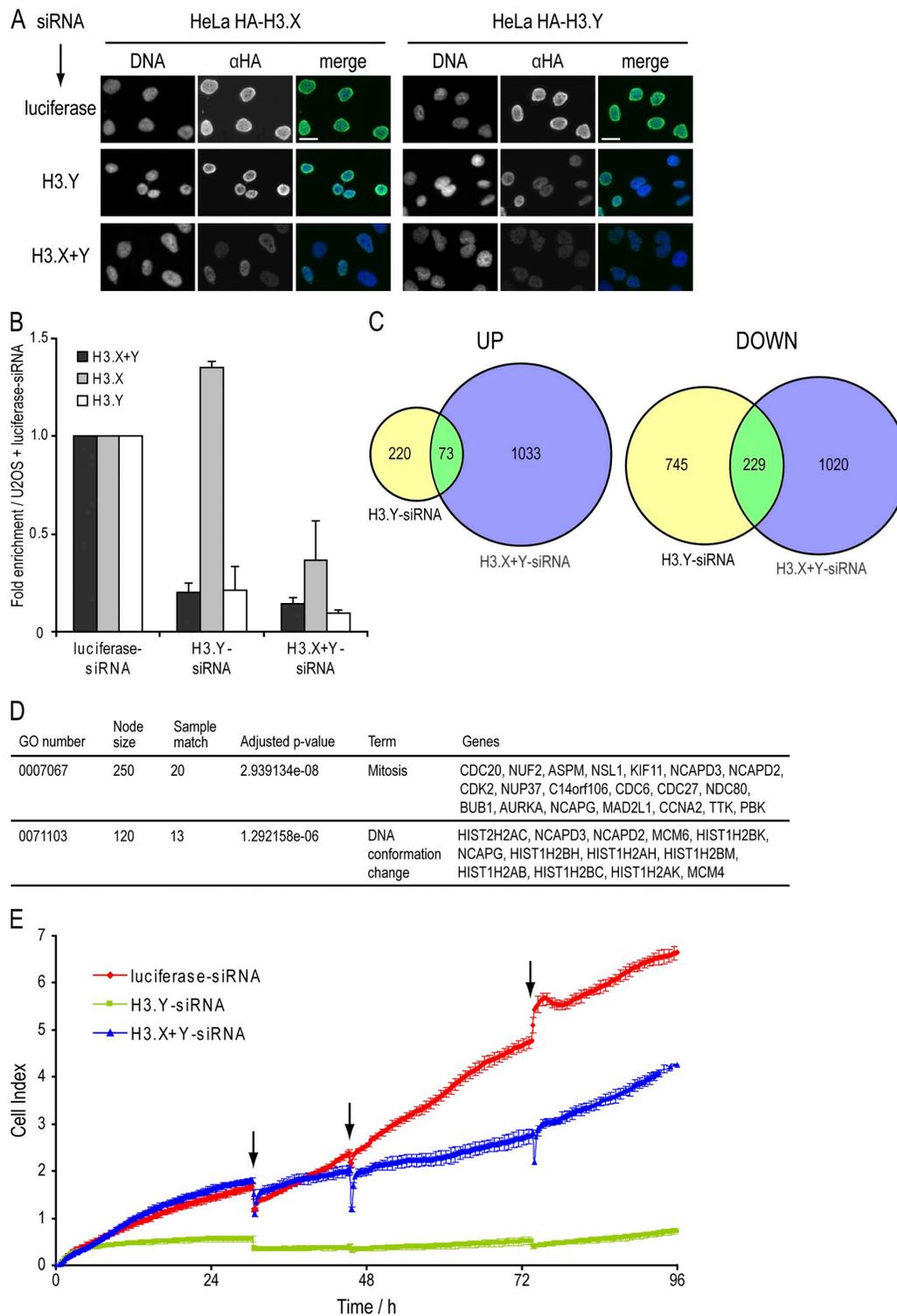
Figure 7. **Influence of H3.Y expression on global gene regulation and cell growth.** (A) Specificity determination of siRNAs against novel variants. IF micros-copy using α-HA (green) and DAPI (blue) staining of HeLa cells expressing HA-H3.X and -H3.Y 4 d after RNAi treatment with indicated siRNAs. Bar, 20 µm. (B) qPCR analysis to verify efficient H3.X and H3.Y RNAi knockdown before global transcriptome analysis. Primer pair H3.X+Y (dark gray) specifically recognizes H3.X and H3.Y nucleotide sequences, whereas two other primer pairs are H3.X- (light gray) or H3.Y-specific (white). Data were normalized to HPRT1 and HMBS expression levels and depict fold enrichment of expression in comparison to luciferase control RNAi. Controls generated without reverse transcriptase were used to assess amplification threshold. Error bars represent SEM of two independent biological experiments. (C) Venn diagrams of genes deregulated after H3.X+Y (blue) and H3.Y (yellow) RNAi in SO-treated U2OS cells, as identified by microarray analyses of two independent experi-ments when compared with luciferase control knockdown. Digits indicate numbers of genes significantly up- (left) or down-regulated (right) in comparison to luciferase control knockdown. (D) Simplified GO analysis of overlapping genes after H3.X+Y and H3.Y knockdown. Detailed GO lists are shown in Tables S1 and S2. Node size = total number of genes analyzed in this node (GO term/group). (E) Growth curve of U2OS cells after RNAi (red, luciferase control siRNA; blue, H3.X+Y siRNA; green, H3.Y siRNA). Arrows mark changes of growth medium.
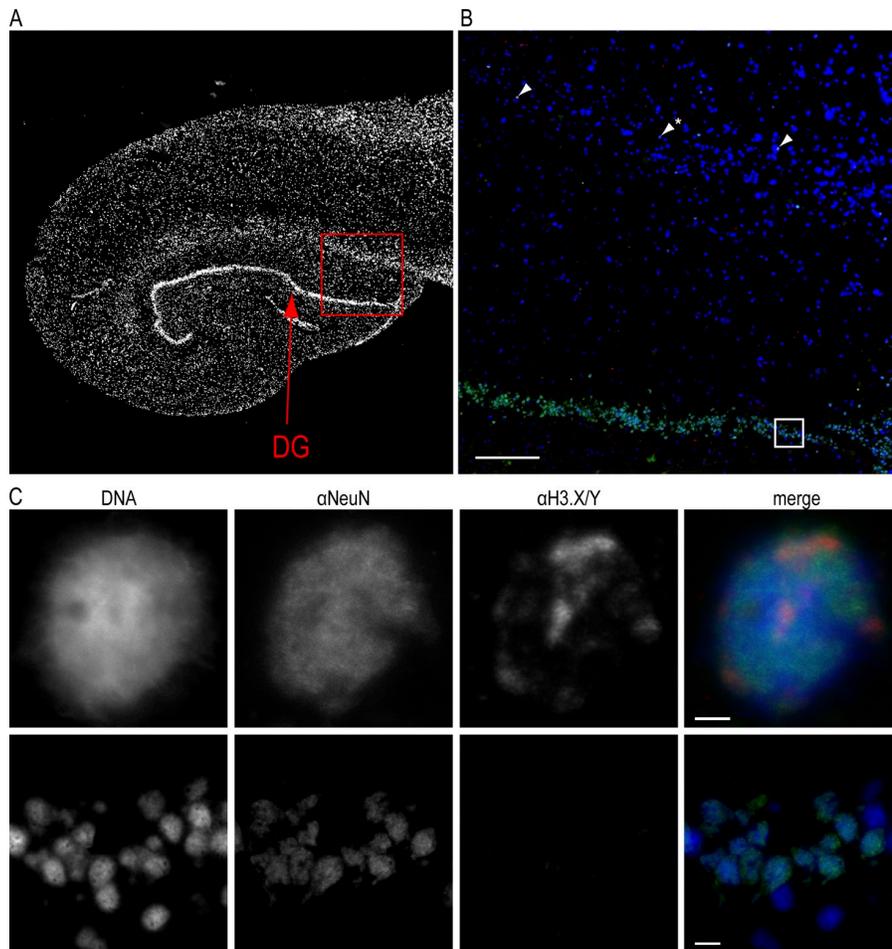
Figure 8. **H3.X/Y protein expression in human brain.** (A) Overview IF picture of commercially available human hippocampus section stained with DAPI (DNA, gray). (B) Human hippocampus sections were costained with α-H3.X/Y (red), α-NeuN (neuronal marker, green), and DAPI (DNA, blue). The boxed section from A is shown. Arrows indicate neuronal cells with positive α-H3.X/Y staining in the region above DG. One out of three representative stainings is shown. (C) The boxed section and the α-H3.X/Y–positive cell marked with an asterisk in B are shown in higher resolution. Costainings with astrocyte marker antibody (α-GFAP) are shown in Fig. S5. Bars: (A and B) 200 µm; (C, top) 2 µm; (C, bottom) 10 µm.

are expressed in human tissue. Based on our qPCR results (Fig. 1 C), and the observation that many primate-specific genes are expressed in the brain and reproductive organs (Tay et al., 2009), we chose to analyze sections of human hippocampus in IF (Fig. 8 A).

Interestingly, some few cells in the region above the dentate gyrus (DG) of the hippocampus stained positive with α-H3.X/Y (Fig. 8 B). We determined their cellular origin by costaining with antibodies against neurons (α neuronal nuclei [α-NeuN] and astrocytes (α glial fibrillary acidic protein [α-GFAP]; Fig. S5). H3.X/Y proteins were identified in a subpopulation of neurons outside of the DG (Fig. 8, B and C). This finding is in accordance with our observation that only few U2OS cells express large amounts of H3.Y. Surprisingly, α-H3.X/Y staining of neurons displayed an enrichment in certain chromatin areas, different from that observed in U2OS cells (Fig. 4, C and G).

In conclusion, H3.X and/or H3.Y proteins are expressed in a subpopulation of neurons in the human hippocampus, highlighting the possibility that these novel histone variants might have cell type–specific functions.

Collectively, our data demonstrate that the novel histone variant H3.Y is expressed in human cells and specialized tissues, is incorporated into chromatin, is posttranslationally modified, and impacts the regulation of many genes implicated in cell cycle progression.

## Discussion

More than 20 yr ago, the mammalian histone H3 variants CENP-A (Palmer et al., 1987) and tH3 (Trostle-Weige et al., 1984) were found. Since then several variants of the mammalian H2A family, such as macroH2A (Pehrson and Fried, 1992) and H2A.Bbd (Chadwick and Willard, 2001), have been described, but it was long thought that all members of the histone H3 family had been identified. To better understand the role of H3 variants in chromatin-related processes, we set out to investigate if there might be even more, yet unknown, H3 variants present in mammals.

### H3.X and H3.Y are primate-specific histone genes

Using the DNA sequence of human histone *H3.1* in genomic searches, we have now identified two novel intron-free histone H3 variant genes on human chromosome 5, which we named *H3.X* and *H3.Y*. Sequences for *H3.X* and *H3.Y* genes were, in addition to humans, only found in chimpanzee and macaque primate genomes; sequence information of other primate genomes is unfortunately not yet available. Additional Basic Local Alignment Search Tool (BLAST) searches in mouse, rat, and other genomes did not yield any positive hits, which suggests that *H3.X* and *H3.Y* genes are present in primates, but not in other mammals or even lower eukaryotes. Both genes might be duplications

from one ancestral gene, as their sequences are highly conserved even in their 5′ and 3′ untranslated regions, with only some differences in their putative promoter sequences. Phylogenetic analyses of *H3.X* and *H3.Y* coding and 3′ genomic sequences revealed a higher sequence homology to *H3.3* genes than to other *H3* variants, leading to the speculation that *H3.X* and *H3.Y* might be evolutionary derivatives of *H3.3*, although they do not contain any introns. To our knowledge, these are the first histone genes that are primate specific outside of testes and likely confer specialized chromatin functions unique to these higher vertebrates.

### H3.X and H3.Y genes are both transcribed, but only H3.Y protein can be detected in vivo

So far we have obtained evidence that H3.X and especially H3.Y mRNAs are present at low but significant amounts in some human bone, breast, lung, and ovary tumor tissues, as well as in testis and certain areas of the brain. Therefore, their expression is not only found in transformed cell lines (U2OS) but also in primary human tissues. Surprisingly, we were able to verify only the existence of H3.Y but not H3.X protein in human cell lines thus far. One possibility is that alternative transcription of H3.X may occur. Two previous annotations of the *H3.X* locus in the NCBI database predicted several introns and different splice sites (Fig. S3 D); however, our primer walk experiment and negative qPCR data using primers specific for these splice forms (unpublished data) provided no evidence that alternative H3.X transcripts exist. Based on our results, *H3.X* is transcribed only at low levels in U2OS cells, and because no protein could be detected thus far, it is feasible that this particular variant is a pseudogene without any functional consequences. Alternatively and more interestingly, it is also possible that its mRNA, although it is present only at low levels in some cells, may have unknown regulatory functions and may even be influencing H3.Y gene expression. Future studies will hopefully shed more light onto a putative function of H3.X mRNA.

### The number of H3.Y-expressing U2OS cells is increased by stress conditions

We showed that H3.Y mRNA and protein is present in some U2OS cells (∼0.1%) and that the number of cells expressing high levels of this novel histone variant can be increased by a stress response that involves nutritional starvation in combination with high cellular density. It is interesting to note that when calculating the amount of endogenous H3.Y mRNA per cell, we observe an ∼40-fold higher expression compared with a cell expressing HA-H3.Y mRNA. This suggests that once a cell receives a specific signal to start transcribing endogenous H3.Y or reaches a certain threshold of an unknown factor, this variant is expressed in extremely high amounts. It will be crucial to identify the physiological trigger of H3.Y expression in future studies to learn more about H3.Y's cellular function.

### Loss of H3.Y expression impacts transcriptional regulation and cell growth

To get a first glimpse on H3.Y's potential impact on gene regulation, we knocked down H3.Y in U2OS cells using RNAi.

Our global microarray data imply that H3.Y influences the transcription of several genes, either directly or indirectly. Interestingly, loss of H3.Y led to more genes being down- rather than up-regulated, which suggests that H3.Y might be involved in transcriptional activation of some genes. This is in accordance with our finding that H3.Y mainly localizes to euchromatic regions (Fig. 4 G). The majority of genes affected by H3.Y reduction are components of cell cycle– and chromatin structure–regulating pathways, leading to a significant impairment of cell growth. It is difficult to unequivocally assign a gene or group of genes as direct targets of H3.Y depletion, which would require the identification of genomic regions containing H3.Y on a global scale by chromatin IP followed by chip hybridization or sequencing. Unfortunately, IP of endogenous H3.Y protein with α-H3.X/Y antibody turned out to be insufficient with standard protocols, and therefore more time and testing will be required to work out a functional ChIP protocol. It was surprising to us that the whole U2OS cell population was equally impaired in its cell growth after H3.Y knockdown although only few cells appeared to strongly express H3.Y protein, as shown by IF quantification (Figs. 4 C and 5 A). It is therefore possible that all U2OS cells express a minimal amount of H3.Y RNA and protein, which is beyond our detection limit, but nevertheless crucial for the cell to traverse through the cell cycle. To our knowledge, only one other histone variant showed a similar effect on the expression of cell cycle–related genes after RNAi depletion: knockdown of the linker histone variant H1.2 altered the expression of ∼2% of genes genome wide (most of them being repressed), including a relevant proportion of cell cycle–related genes (Sancho et al., 2008). It is therefore not completely unlikely that H3.Y plays an important role in the regulation of genes involved in cell cycle control. It is tempting to speculate that the stress-related appearance of more U2OS cells expressing H3.Y is caused by continued cell proliferation or the prevention of cell death. One interesting observation was that knockdown of H3.Y led to stronger cell growth defects than combined H3.X+H3.Y RNAi. Because we repeatedly saw an increase of H3.X mRNA after H3.Y depletion, one speculative idea would be that H3.X and H3.Y are connected by some sort of regulatory feedback mechanism. It will be of great interest in future studies to shed light on the regulatory expression pathways of these novel variants, particularly in the context of cellular stress and their existence in neuronal subpopulations of the human hippocampus.

## Materials and methods

### Cell culture, cloning of H3.X and H3.Y expression constructs, and transfection

Human HeLa, HeLa Kyoto, HEK293, U2OS, mouse NIH3T3, and rat neuroblastoma cells were grown in DME medium (PAA) supplemented with 10% FCS (Sigma-Aldrich) and 1% penicillin/streptomycin (PAA) at 37°C and 5% $CO_2$. Human Raji cells were cultured in RPMI-1640 medium (Invitrogen) supplemented with 10% FCS (Sigma-Aldrich) and 1% penicillin/streptomycin (PAA) at 37°C and 5% $CO_2$. H3.1, H3.2, H3.3, H3.Y, and H3.X cDNA were cloned into pEGFP-C1 (Takara Bio Inc.) and pIRES-neo vector (Takara Bio Inc.) containing an HA tag (a gift from H. Dormann, Rockefeller University, New York, NY) to generate N-terminally tagged H3 variants after transfection into cell lines. For simplicity, we will refer to the EGFP-tagged constructs as GFP-tagged throughout the text. All constructs were sequenced (MWG-Biotech AG) to verify cloning and amplification accuracy.

Transfections were performed with FuGene HD (Roche) according to the manufacturer's instructions. In brief, cells were incubated with the transfection complex for 1 d, after which selection medium containing 600 µg/ml G-418 sulfate (PAA) was added. Stable cell lines were maintained in medium containing 400 µg/ml G-418 sulfate.

### Expression of recombinant H3.X and H3.Y proteins in *Escherichia coli*
BL21-CodonPlus (DE3)-RIL bacteria (Agilent Technologies) were transformed with pET-21a(+) plasmids (EMD) containing H3.X and H3.Y cDNA. Expression of recombinant proteins was induced by incubation with 0.1% IPTG (Carl Roth) overnight and controlled through SDS-PAGE analysis of boiled bacteria (Coomassie and immunoblot).

### RNA expression analysis
Total RNA was isolated using the RNeasy mini kit (QIAGEN) according to the manufacturer's instructions. Genomic DNA contaminations were removed by performing an on-column DNase I digest (RNase-free DNase; QIAGEN). cDNA was synthesized using the ProtoScript First Strand cDNA Synthesis kit (New England Biolabs, Inc.), priming with random nonamers. qPCR analysis was performed on a LightCycler 480 (Roche) using the Light-Cycler 480 SYBR green I mastermix (Roche). Data analysis was performed with the advanced relative quantification tool of the LightCycler 480 software; results were normalized to HPRT1 and HMBS levels. H3.X and H3.Y primer sequences (Sigma-Aldrich) are listed in Fig. S3 C. Total RNA from different human tissues was commercially acquired from Applied Biosystems (normal lung, breast and tumor breast, and lung and ovary) and BioChain (tumor lung, breast, thyroid and bone, normal testis, cerebellum, cerebral cortex, hippocampus, thalamus, and total fetal brain).

### Histone extraction and mononucleosome IP
Histones were acid extracted as described previously (Shechter et al., 2007). In brief, nuclei were isolated by hypotonic lysis and extracted using 0.4 M sulfuric acid. Soluble histones were precipitated with trichloroacetic acid and resuspended in water.

Mononucleosomes for IP experiments were generated as described previously (Wysocka et al., 2001), with the following changes: For cell lysis, 0.1% NP-40 was used instead of Triton X-100. All centrifugation steps before MNase treatment were performed at 3,200 *g*. Mononucleosomes were generated by digestion of chromatin with 0.25 U MNase (Sigma-Aldrich) for 15 min in buffer A (10 mM Hepes, pH 7.9, 10 mM KCl, 1.5 mM MgCl$_2$, 0.34 M sucrose, 10% glycerol [vol/vol], 1 mM DTT, and protease inhibitor cocktail [Roche] plus 1 mM CaCl$_2$) and stopped by the addition of EGTA (final concentration of 2 mM). Centrifugation was performed at 20,000 g for 20 min. Supernatants of four MNase digests were combined, and salt concentration was adjusted to 150 mM KCl. Magnetic beads (Invitrogen) were washed four times with buffer C (20 mM Hepes/KOH, pH 7.9, 20% [vol/vol] glycerol, 0.2 mM EDTA, 0.2% [vol/vol] Triton X-100, 300 mM KCl, and protease inhibitor cocktail [Roche]). One part was evaluated for DNA size and quality using DNA 1000 reagents (Agilent Technologies) with the 2100 Bioanalyzer (Agilent Technologies), another part was analyzed by immunoblotting using the Odyssey infrared imaging system (LI-COR Biosciences) and evaluated with Odyssey Software Version 2.1 (LI-COR Biosciences), and the last part was analyzed by silver staining.

### Antibodies
To generate an antibody against H3.X and H3.Y (α-H3.X/Y), a peptide spanning amino acids KATAWQAPRKLP of histone H3.X and H3.Y was synthesized (Peptide Specialty Laboratories GmbH) and coupled to BSA and ovalbumin, respectively. Rats were immunized subcutaneously and intraperitoneally with a mixture of 50 µg peptide-ovalbumin, 5 nmol CPG oligonucleotide (TIB MOLBIOL GmbH), 500 µl PBS, and 500 µl of incomplete Freund's adjuvant. A boost without adjuvant was given 6 wk after the primary injection. Fusion was performed using standard procedures (Köhler and Milstein, 1975). Supernatants were tested by differential ELISA with the histone peptide coupled to BSA and an irrelevant peptide coupled to the same carrier. Monoclonal antibodies that reacted specifically with the peptide were further analyzed in immunoblot and IF studies. α-H3.X/Y clone 8H6-2111 of rat IgG2a subclass was deployed in this study.

The following primary antibodies were used in this study: α-HA (clone 12CA5; Roche), α-H3S10ph (Millipore), α-H3 (C terminus; Abcam), α-NeuN (Millipore), and α-GFAP (Dako).

The following secondary antibodies were used: α-rat HRP and α-rabbit HRP (GE Healthcare); α-mouse–IRDye700DX and α-rabbit–IRDye800DX (Rockland Immunochemicals, Inc.); α-rat–Alexa Fluor 488, α-mouse–Alexa Fluor 488, and streptavidin-Alexa 555 (Invitrogen); and

α-rabbit–Rhodamine red X and α-rabbit–Cy5 (Dianova, Inc.); and α-rat–biotin (Vector Laboratories).

### IF microscopy
Adherent mammalian cells were grown on coverslips, washed, fixed in 1% or 3.7% formaldehyde–PBS solution, permeabilized with 0.1% or 0.5% Triton X-100–PBS solution, and blocked with 1% BSA in 0.1% or 0.5% Triton X-100–PBS solution. After stepwise incubation with primary and then secondary fluorescent antibody, cells were stained with DAPI (Invitrogen) and mounted with a ProLong Antifade kit (Invitrogen). Samples for 3D-SIM were prepared on precision cover glass No. 1.5 (thickness 0.170 ± 0.005 mm; Carl Roth) and embedded in Vectashield mounting medium (Vector Laboratories). Chromosome spreads were generated as described previously (Hake et al., 2005). For confocal microscopy, samples were stained with To-Pro3 (Invitrogen) and mounted with Vectashield mounting medium (Vector Laboratories).

### FRAP
For FRAP experiments, human HeLa Kyoto cells were transiently transfected with GFP-tagged histone H3 constructs. After 24 h, cells were seeded in Lab-Tek chamber slides (Thermo Fisher Scientific) and incubated for several hours or overnight before performing FRAP experiments. FRAP was performed on a spinning disc microscope (UltraVIEW VoX; PerkinElmer) with an integrated FRAP PhotoKinesis accessory (PerkinElmer) assembled to an Axio Observer D1 stand (Carl Zeiss, Inc.). The microscope was equipped with a heated environmental chamber set to 37°C and CO$_2$ perfusion set to 5%. For photobleaching experiments, several square bleach regions with a size of 5 × 5 µm were positioned on selected cell nuclei within the field of view. Photobleaching was performed using two iterations with the acousto-optical tunable filter of the 488-nm and the 514-nm laser line set to 100% transmission. To determine long-term recovery kinetics, 3D image stacks of 8-µm height and a z distance between image planes of 1 µm were recorded with an initial speed of 1 frame/min for the first 10 frames, followed by intervals of 10 min.

Quantitative evaluations were performed with ImageJ (http://rsb .info.nih.gov/ij/). Lateral and rotational movements of the cell nucleus were corrected by image registration using the StackReg plug-in of ImageJ. Mean intensities over time were extracted from the total nuclear area (*T*). The background ROI outside of the cell was defined manually from the initial field of view. The mean gray values over time were measured, background subtracted, and normalized to the respective means of the last prebleach values. The resulting postbleach *B* values were then divided by the respective *T* value to correct for the superimposed gain or loss of total fluorescence during postbleach acquisition, potentially caused by newly synthesized GFP-histones, bleaching-by-acquisition, and flux of residual fluorescence from above and below the recorded optical plane.

### Peptide competition
α-H3.X/Y antibody was incubated for 4 h at 4°C with the following peptides (1 µg/ml) prior to addition to the fixed cells: biotin-coupled H3.3 unmodified, aa 22–41 (Proteomics Resource Center of the Rockefeller University); and H3.X/Y unmodified, aa 9–20 (Peptide Specialty Laboratories GmbH).

### Immunohistochemistry
Commercially available frozen sections of human hippocampus (Biochain) were thawed and blocked with 0.5% BSA in 0.5% Triton X-100–PBS solution. After stepwise incubation with primary and then secondary fluorescent antibodies, cells were stained with DAPI (Invitrogen) and mounted with Aqua-Poly/Mount (Polysciences, Inc.). To enhance α-H3.X/Y signals, samples were incubated with a biotin-coupled α-rat antibody (Vector Laboratories), followed by streptavidin coupled to an Alexa 555 fluorophore (Invitrogen).

### Microscopes
Unless stated otherwise, samples were kept at room temperature during image acquisition. Stained cells were analyzed on a microscope (Axiovert 200M) with an EC Plan-Neofluar 40×/1.3 NA oil Ph3 (differential interference contrast [DIC] III) objective (both from Carl Zeiss, Inc.). Images were processed with AxioVision software (Carl Zeiss, Inc.). Confocal imaging was performed with a confocal microscope (LSM 510 META; Carl Zeiss, Inc.) equipped with an argon-ion and two helium-ion lasers using a Plan-Apochromat 63×/1.4 NA oil DIC or a Plan-Neofluar 40×/1.3 NA oil DIC objective lens (both from Carl Zeiss, Inc.). Images were processed with LSM 510 META software (Carl Zeiss, Inc.). Images of metaphase chromosome spreads were acquired using a personalDV wide-field epifluorescence microscope (Applied Precision) equipped with a 60×/1.42 NA Plan-Apochromat oil objective lens (Olympus) and a CoolSNAP HQ$^2$

charge-coupled device (CCD) camera (Photometrics). Typically, stacks with a z distance of 200 nm were recorded and then subjected to a constrained iterative deconvolution (enhanced ratio, 10 cycles, medium noise filtering) using the SoftWoRx 3.7 imaging software package (Applied Precision). For a description about the microscopic setup used for FRAP, see the description for in the FRAP section. Imaging with 3D-SIM was performed with an DeltaVision OMX v3 prototype (Applied Precision) equipped with high-power diode lasers with wavelengths of 405, 488 and 592.2 nm; a 100×/1.4 NA Plan-Apochromat oil objective lens (Olympus); and Cascade II:512 EM CCD cameras (Photometrics). Laser light is directed through a movable optical grating to generate a fine-striped illumination pattern on the sample plane. Image stacks with 15 images per plane (five phases, three angles) and a z distance of 125 nm were acquired and subjected to a computational reconstruction to obtain a high-resolution 3D dataset with a two-fold enhanced optical resolution in all three special directions compared with conventional light microscopy (Gustafsson et al., 2008; Schermelleh et al., 2008). Immunohistochemistry samples were analyzed using a confocal LSM 710 microscope (Carl Zeiss, Inc.) equipped with three argon, one diode-pumped solid-state, one helium, and one diode laser. Images were acquired using an EC Plan-Neofluar 10×/0.3 NA M27, an LD C-Apochromat 40×/1.1 W Korr M27 (water), or a LCI Plan-Neofluar 63×/1.3-mm Korr Ph3 M27 (water) objective lens (all from Carl Zeiss, Inc.), and Image processing was performed with ZEN 2009 (Carl Zeiss, Inc.) software.

For downstream image processing of all microscope images, Photoshop (Adobe) software packages were used.

### In silico modeling

To generate homology models from protein sequences, we used the I-TASSER server (Zhang, 2008). I-TASSER combines threading, assembly, and refinement to generate 3D models. For sequences of both H3.X and H3.Y, the server automatically selected as a template the chain A (histone H3.2), PDB accession no. 1kx5. This structure was obtained using *X. laevis* histones assembled into nucleosomes with 147 bp of DNA from *Homo sapiens* (Davey et al., 2002). The model for H3.X had a C score of 0.07, and the model for H3.Y had a C score of 1.50. In both cases, the second-best model had a considerably lower C score. Datasets of the models were downloaded and visualized by means of the Swiss-PdbViewer (Guex and Peitsch, 1997) and PyMOL (DeLano Scientific LLC; http://www.pymol.org).

### RP-HPLC and MS/MS

Histones from U2OS, HEK293, and NIH3T3 cells were separated by RP-HPLC on a C4 column (250 × 4.6-mm Jupiter, 10 µm, 300 Å; Phenomenex) using a linear gradient of 35–53% solvent B (solvent A, 0.1% trifluoroacetic acid; solvent B, 99.92% acetonitrile and 0.08% trifluoroacetic acid) over 30 min at 1.0 ml/min on a Biotech Ettan microLC (GE Healthcare). The H3-containing fractions were dried under vacuum and stored at −20°C. For mass spectrometry analysis, RP-HPLC fractions containing histones were resolved in 0.1 M ammonium bicarbonate and digested. Proteolytic digestions were performed overnight at 37°C with sequencing grade trypsin (Promega). In this case, histones were chemically modified beforehand by treatment with propionic anhydride (Merck) to convert free amino groups of lysine residues to propionic amides. After digestion, samples were directly loaded onto a nano electrospray ionization (ESI) LC-MS/MS for protein identification. Each sample was first separated on a C18 reversed phase column via a linear acetonitrile (Sigma-Aldrich) gradient (UltiMate 3000 system [Dionex] and an LC Packings column; 75 µm inner diameter × 15 cm, packed with C18 PepMap, 3 µm, 100 Å) before MS and MS/MS spectra were recorded on an Orbitrap mass spectrometer (Thermo Fisher Scientific). The resulting data were analyzed via the Mascot Software (Matrix Science) using a home-made database that contains the diverse H3 sequences. Fragment spectra were also interpreted manually.

### RNAi and growth curve

siRNAs were designed using siDESIGN Center (http://www.dharmacon .com/designcenter/designcenterpage.aspx; Thermo Fisher Scientific) and synthesized (MWG-Biotech AG). siRNAs have been prevalidated to confirm their targeting specificity to H3.Y or H3.X+H3.Y and to reduce the chance of off-target effects by BLAST searches (NCBI). The following double-stranded siRNAs, which had differences of at least three nucleotides from other targets, were used. Luciferase, 5′-CUUACGCUGAGUACUUCGA-3′; H3.Y, 5′-CCGCAGAGAGGGUCCUUAA-3′; H3.X+Y, 5′-GCGGGAAAU-CAGAAAGUAC-3′. Cells were transfected twice with siRNAs using oligofectamine (Invitrogen) according to the manufacturer's instructions. 4 d after transfection, cells were used for various assays. Determination of cell growth after RNAi was performed using xCELLigence system (Roche). Cells were seeded in E plates 1 d after the second siRNA transfection. Cell growth was

monitored over the next 4 d. To eliminate side effects caused by the lack of nutrients or the accumulation of waste products, growth medium was exchanged every day.

### Microarray hybridization

Total RNA preparations were further purified with the RNeasy MinElute Cleanup kit (QIAGEN). 100 ng of RNA were used as starting material for all target preparations. RNA amplification, labeling, and hybridization to Human Gene 1.0 ST Arrays (Affymetrix) were performed according to the Human Gene 1.0 ST Array kit protocol (Affymetrix). The raw microarray data were processed in R/Bioconductor (http://www.bioconductor .org/) as follows: gene-based expression values were calculated using the robust multichip average (RMA) method provided by the oligo package. Genes that had a log2 expression value of at least 4 in at least one of the treatment conditions were kept for downstream analyses. Differential expression estimation was based on a moderated t statistic (limma package) with subsequent calculation of the local false discovery rate (lfdr; locfdr package). Genes were classified as responders by an lfdr cutoff of 0.2. GO enrichment analysis was performed using a hypergeometric distribution test and subsequent Bonferroni correction as supplied by the GOHyperGALL script (http://faculty.ucr.edu/~tgirke/Documents/R_BioCond/My_R_Scripts/ GOHyperGAll.txt). We reduced term redundancy by applying the GOHyperGAll_Simplify function with a P-value cutoff of 0.001.

### Online supplemental material

Fig. S1 shows H3.X and H3.Y alignments, evolutionary origin, and expression level determination. Fig. S2 shows analysis of H3.X- and H3.Y-containing nucleosomes, and α-H3.X/Y antibody specificity determination. Fig. S3 shows evaluation of human H3.X sequences and inducible endogenous H3.X and H3.Y gene expression. Fig. S4 shows purification and mass spectrometrical identification of endogenous H3.Y protein. Fig. S5 shows H3.X/Y expression in the human hippocampus. Table S1 shows a GO list with shared up-regulated genes after H3.Y knockdown. Table S2 shows a GO list with shared down-regulated genes after H3.Y knockdown. Online supplemental material is available at http://www.jcb .org/cgi/content/full/jcb.201002043/DC1.

## References

Abbott, D.W., V.S. Ivanova, X. Wang, W.M. Bonner, and J. Ausió. 2001. Characterization of the stability and folding of H2A.Z chromatin particles: implications for transcriptional activation. *J. Biol. Chem.* 276:41945–41949. doi:10.1074/jbc.M108217200

Ahmad, K., and S. Henikoff. 2002a. Histone H3 variants specify modes of chromatin assembly. *Proc. Natl. Acad. Sci. USA*. 99:16477–16484. doi:10.1073/ pnas.172403699

Ahmad, K., and S. Henikoff. 2002b. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol. Cell.* 9:1191–1200. doi:10.1016/S1097-2765(02)00542-7

Allshire, R.C., and G.H. Karpen. 2008. Epigenetic regulation of centromeric chromatin: old dogs, new tricks? *Nat. Rev. Genet.* 9:923–937. doi:10.1038/nrg2466

Angelov, D., A. Molla, P.Y. Perche, F. Hans, J. Côté, S. Khochbin, P. Bouvet, and S. Dimitrov. 2003. The histone variant macroH2A interferes with transcription factor binding and SWI/SNF nucleosome remodeling. *Mol. Cell.* 11:1033–1041. doi:10.1016/S1097-2765(03)00100-X

Bao, Y., K. Konesky, Y.J. Park, S. Rosu, P.N. Dyer, D. Rangasamy, D.J. Tremethick, P.J. Laybourn, and K. Luger. 2004. Nucleosomes containing the histone variant H2A.Bbd organize only 118 base pairs of DNA. *EMBO J.* 23:3314–3324. doi:10.1038/sj.emboj.7600316

Bernstein, E., and S.B. Hake. 2006. The nucleosome: a little variation goes a long way. *Biochem. Cell Biol.* 84:505–517. doi:10.1139/O06-085

Black, B.E., D.R. Foltz, S. Chakravarthy, K. Luger, V.L. Woods Jr., and D.W. Cleveland. 2004. Structural determinants for generating centromeric chromatin. *Nature.* 430:578–582. doi:10.1038/nature02766

Bönisch, C., S.M. Nieratschker, N.K. Orfanos, and S.B. Hake. 2008. Chromatin proteomics and epigenetic regulatory circuits. *Expert Rev. Proteomics.* 5:105–119. doi:10.1586/14789450.5.1.105

Chadwick, B.P., and H.F. Willard. 2001. A novel chromatin protein, distantly related to histone H2A, is largely excluded from the inactive X chromosome. *J. Cell Biol.* 152:375–384. doi:10.1083/jcb.152.2.375

Cheng, Y., R.M. Miura, and B. Tian. 2006. Prediction of mRNA polyadenylation sites by support vector machine. *Bioinformatics.* 22:2320–2325. doi:10.1093/bioinformatics/btl394

Clapier, C.R., and B.R. Cairns. 2009. The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.* 78:273–304. doi:10.1146/annurev.biochem.77.062706.153223

Clayton, A.L., and L.C. Mahadevan. 2003. MAP kinase-mediated phospho-acetylation of histone H3 and inducible gene regulation. *FEBS Lett.* 546:51–58. doi:10.1016/S0014-5793(03)00451-4

Davey, C.A., D.F. Sargent, K. Luger, A.W. Maeder, and T.J. Richmond. 2002. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. *J. Mol. Biol.* 319:1097–1113. doi:10.1016/S0022-2836(02)00386-8

Gautier, T., D.W. Abbott, A. Molla, A. Verdel, J. Ausio, and S. Dimitrov. 2004. Histone variant H2ABbd confers lower stability to the nucleosome. *EMBO Rep.* 5:715–720. doi:10.1038/sj.embor.7400182

Guex, N., and M.C. Peitsch. 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis.* 18:2714–2723. doi:10.1002/elps.1150181505

Gustafsson, M.G., L. Shao, P.M. Carlton, C.J. Wang, I.N. Golubovskaya, W.Z. Cande, D.A. Agard, and J.W. Sedat. 2008. Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination. *Biophys. J.* 94:4957–4970. doi:10.1529/biophysj.107.120345

Hake, S.B., B.A. Garcia, M. Kauer, S.P. Baker, J. Shabanowitz, D.F. Hunt, and C.D. Allis. 2005. Serine 31 phosphorylation of histone variant H3.3 is specific to regions bordering centromeres in metaphase chromosomes. *Proc. Natl. Acad. Sci. USA.* 102:6344–6349. doi:10.1073/pnas.0502413102

Hake, S.B., B.A. Garcia, E.M. Duncan, M. Kauer, G. Dellaire, J. Shabanowitz, D.P. Bazett-Jones, C.D. Allis, and D.F. Hunt. 2006. Expression patterns and post-translational modifications associated with mammalian histone H3 variants. *J. Biol. Chem.* 281:559–568. doi:10.1074/jbc.M509266200

Hendzel, M.J., Y. Wei, M.A. Mancini, A. Van Hooser, T. Ranalli, B.R. Brinkley, D.P. Bazett-Jones, and C.D. Allis. 1997. Mitosis-specific phosphorylation of histone H3 initiates primarily within pericentromeric heterochromatin during G2 and spreads in an ordered fashion coincident with mitotic chromosome condensation. *Chromosoma.* 106:348–360. doi:10.1007/s004120050256

Huyen, Y., O. Zgheib, R.A. Ditullio Jr., V.G. Gorgoulis, P. Zacharatos, T.J. Petty, E.A. Sheston, H.S. Mellert, E.S. Stavridi, and T.D. Halazonetis. 2004. Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks. *Nature.* 432:406–411. doi:10.1038/nature03114

Im, H., C. Park, Q. Feng, K.D. Johnson, C.M. Kiekhaefer, K. Choi, Y. Zhang, and E.H. Bresnick. 2003. Dynamic regulation of histone H3 methylated at lysine 79 within a tissue-specific chromatin domain. *J. Biol. Chem.* 278:18346–18352. doi:10.1074/jbc.M300890200

Köhler, G., and C. Milstein. 1975. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature.* 256:495–497. doi:10.1038/256495a0

Kozak, M. 1991. An analysis of vertebrate mRNA sequences: intimations of translational control. *J. Cell Biol.* 115:887–903. doi:10.1083/jcb.115.4.887

McKittrick, E., P.R. Gafken, K. Ahmad, and S. Henikoff. 2004. Histone H3.3 is enriched in covalent modifications associated with active chromatin. *Proc. Natl. Acad. Sci. USA.* 101:1525–1530. doi:10.1073/pnas.0308092100

Nakagawa, S., Y. Niimura, T. Gojobori, H. Tanaka, and K. Miura. 2008. Diversity of preferred nucleotide sequences around the translation initiation codon in eukaryote genomes. *Nucleic Acids Res.* 36:861–871. doi:10.1093/nar/gkm1102

Palmer, D.K., K. O'Day, M.H. Wener, B.S. Andrews, and R.L. Margolis. 1987. A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones. *J. Cell Biol.* 104:805–815. doi:10.1083/jcb.104.4.805

Pehrson, J.R., and V.A. Fried. 1992. MacroH2A, a core histone containing a large nonhistone region. *Science.* 257:1398–1400. doi:10.1126/science.1529340

Polo, S.E., D. Roche, and G. Almouzni. 2006. New histone incorporation marks sites of UV repair in human cells. *Cell.* 127:481–493. doi:10.1016/j.cell.2006.08.049

Pusarla, R.H., and P. Bhargava. 2005. Histones in functional diversification. Core histone variants. *FEBS J.* 272:5149–5168. doi:10.1111/j.1742-4658.2005.04930.x

Sancho, M., E. Diani, M. Beato, and A. Jordan. 2008. Depletion of human histone H1 variants uncovers specific roles in gene expression and cell growth. *PLoS Genet.* 4:e1000227. doi:10.1371/journal.pgen.1000227

Santenard, A., and M.E. Torres-Padilla. 2009. Epigenetic reprogramming in mammalian reproduction: contribution from histone variants. *Epigenetics.* 4:80–84. doi:10.4161/epi.4.2.7838

Schermelleh, L., P.M. Carlton, S. Haase, L. Shao, L. Winoto, P. Kner, B. Burke, M.C. Cardoso, D.A. Agard, M.G. Gustafsson, et al. 2008. Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy. *Science.* 320:1332–1336. doi:10.1126/science.1156947

Shechter, D., H.L. Dormann, C.D. Allis, and S.B. Hake. 2007. Extraction, purification and analysis of histones. *Nat. Protoc.* 2:1445–1457. doi:10.1038/nprot.2007.202

Strahl, B.D., and C.D. Allis. 2000. The language of covalent histone modifications. *Nature.* 403:41–45. doi:10.1038/47412

Suto, R.K., M.J. Clarkson, D.J. Tremethick, and K. Luger. 2000. Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. *Nat. Struct. Biol.* 7:1121–1124. doi:10.1038/81971

Tabaska, J.E., and M.Q. Zhang. 1999. Detection of polyadenylation signals in human DNA sequences. *Gene.* 231:77–86. doi:10.1016/S0378-1119(99)00104-3

Tagami, H., D. Ray-Gallet, G. Almouzni, and Y. Nakatani. 2004. Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell.* 116:51–61. doi:10.1016/S0092-8674(03)01064-X

Tay, S.K., J. Blythe, and L. Lipovich. 2009. Global discovery of primate-specific genes in the human genome. *Proc. Natl. Acad. Sci. USA.* 106:12019–12024. doi:10.1073/pnas.0904569106

Trostle-Weige, P.K., M.L. Meistrich, W.A. Brock, and K. Nishioka. 1984. Isolation and characterization of TH3, a germ cell-specific variant of histone 3 in rat testis. *J. Biol. Chem.* 259:8769–8776.

van Holde, K.E. 1988. Chromatin. Springer, New York. 497 pp.

Witt, O., W. Albig, and D. Doenecke. 1996. Testis-specific expression of a novel human H3 histone gene. *Exp. Cell Res.* 229:301–306. doi:10.1006/excr.1996.0375

Wysocka, J., P.T. Reilly, and W. Herr. 2001. Loss of HCF-1-chromatin association precedes temperature-induced growth arrest of tsBN67 cells. *Mol. Cell. Biol.* 21:3820–3829. doi:10.1128/MCB.21.11.3820-3829.2001

Yan, C., and D.D. Boyd. 2006. Histone H3 acetylation and H3 K4 methylation define distinct chromatin regions permissive for transgene expression. *Mol. Cell. Biol.* 26:6357–6371. doi:10.1128/MCB.00311-06

Zhang, Y. 2008. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics.* 9:40. doi:10.1186/1471-2105-9-40

# Supplementary information for:

# Identification and characterization of two novel primate-specific histone H3 variants, H3.X and H3.Y

THE JOURNAL OF CELL BIOLOGY
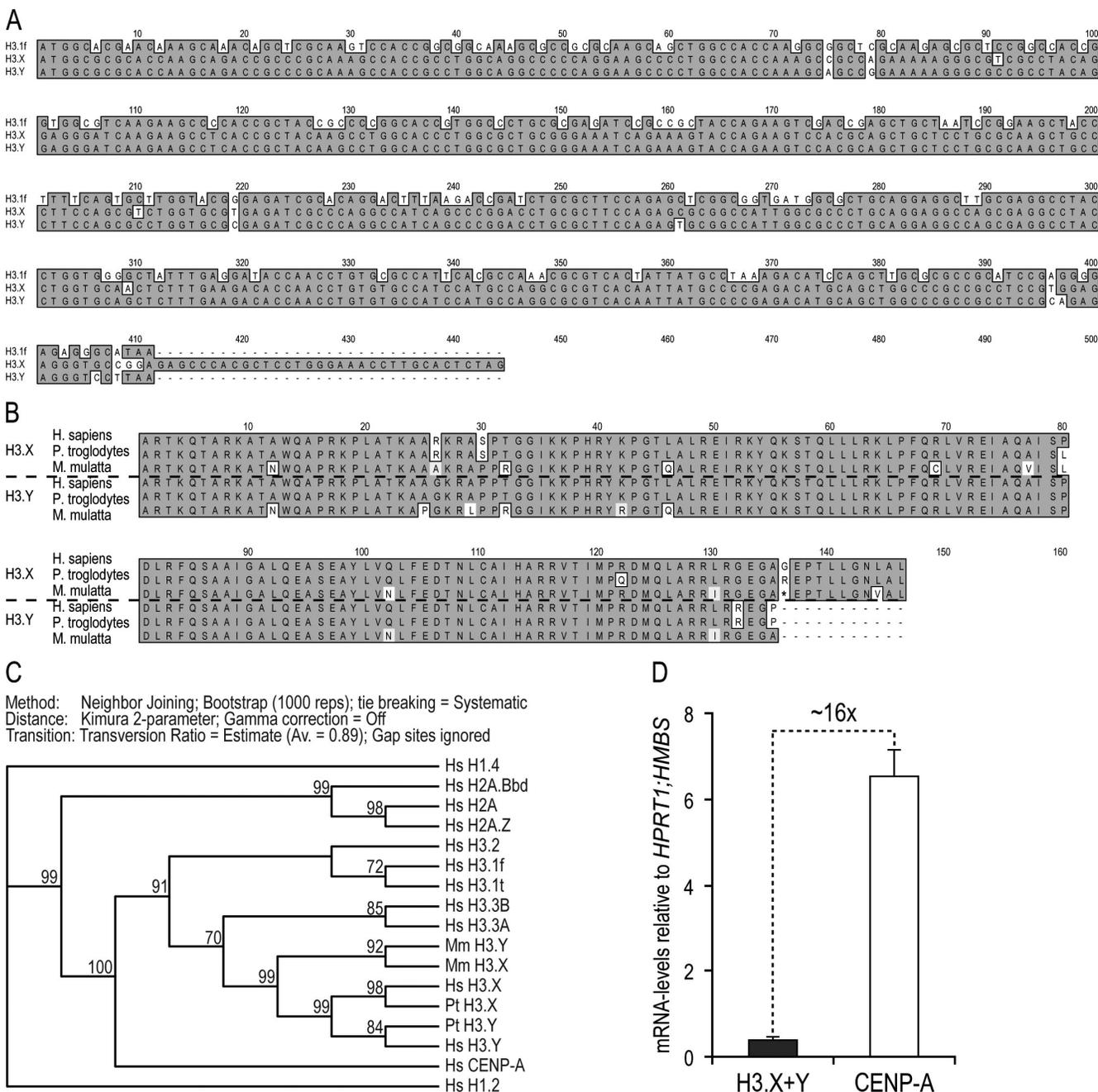


**Figure S1.  H3.X and H3.Y alignments, evolutionary origin, and expression level determination.** Alignments were made with ClustalW Alignment (Mac-Vector 10.0.2). (A) Nucleotide alignment of human *H3.1f* (available from GenBank/EMBL/DDBJ under accession no. NM_003533), *H3.X* (accession no. LOC340096), and *H3.Y* (accession no. LOC391769) genes. Identical nucleotides are highlighted in dark gray and changes are set apart on a white background. Nucleotide numbers are indicated on top. (B) Amino acid alignment of human (*H. sapiens*), chimpanzee (*P. troglodytes*), and macaque (*M. mulatta*) H3.X and H3.Y proteins. Note that *M. mulatta* H3.X protein contains a stop (asterisk) at position 136, but the following nucleotides would code for an H3.X-specific C-terminal tail. Identical amino acids are highlighted in dark gray, similar amino acids are highlighted in light gray, and changes are set apart on a white background. Amino acid numbers are indicated on top. The dotted line separates the results of H3.X and H3.Y proteins. (C) Phylogenetic tree of H3.X and H3.Y from different species. Coding nucleotide sequences of different histones from human (Hs), chimpanzee (Pt), and macaque (Mm) were aligned using ClustalW Alignment (MacVector 10.0.2). Evolutionary relationships were determined using the neighbor-joining tree-building method based on the Kimura 2 parameter with gamma correction off. Numbers indicate bootstrap probabilities. (D) qPCR analysis with cDNA from U2OS cells. Primer pairs for H3.X+Y (dark gray; specifically recognizes H3.X and H3.Y nucleotide sequences) and CENP-A (white) were used. Data were normalized to HPRT1 and HMBS expression levels. Controls generated without reverse transcription were used to assess amplification threshold. Error bars represent SEM of two independent biological experiments.
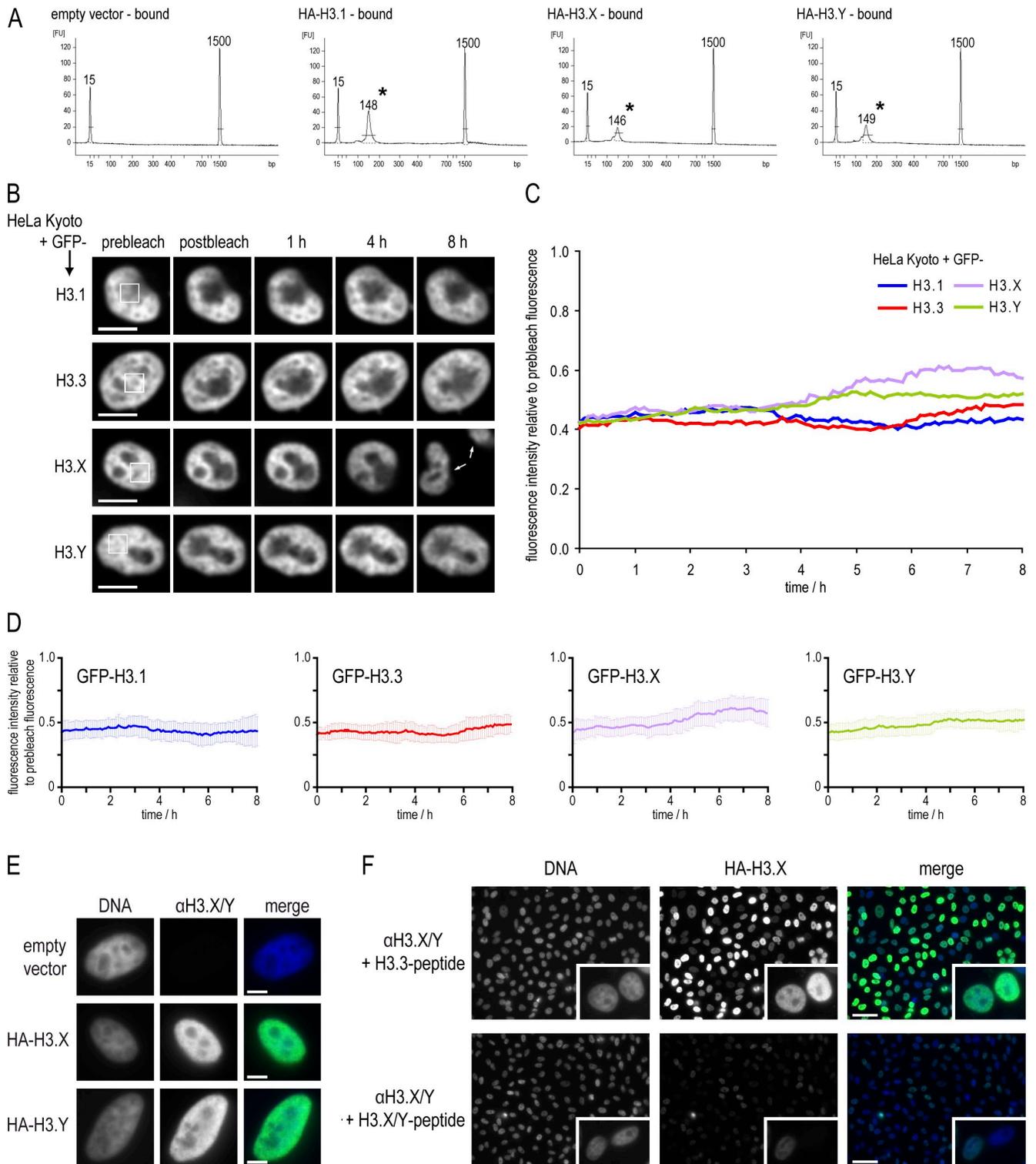
Figure S2.    **Analysis of H3.X- and H3.Y-containing nucleosomes and α-H3.X/Y antibody specificity determination.** (A) Evaluation of DNA size and quality after MNase IP (bound material, see also Fig. 3 C). Purified DNA after MNase IP was separated and analyzed using a 2100 Bioanalyzer (Agilent Technologies). Shown are the electropherograms of bound DNA material. Peaks left and right (15 and 1,500) indicate marker DNA measured in base pairs. Peaks labeled with asterisks show base pair length of precipitated DNA after MNase IP. (B) FRAP experiment to evaluate nucleosomal stability of novel H3 variants in vivo. HeLa Kyoto cells were transiently transfected with GFP and GFP-H3.1, -H3.3, -H3.X, and -H3.Y constructs. Live cell imaging was performed with confocal spinning disk microscopy. A small nuclear area was photobleached (box) with a focused laser beam, and the recovery of the fluorescent signal was monitored over 8 h. Mean projection of the 3–4 central image planes is shown. Many cells underwent cell division during or at the end of the time series (e.g., arrows pointing to daughter cells), which indicates the general viability of the cells. (C) Quantitative evaluation of mean GFP signal recovery after photobleaching relative to fluorescence intensity before bleaching. Mean curves of 13–25 individual cells are depicted for each construct. For clarity, error bars are omitted here (see D for details). Note that the FRAP curves start at a higher initial postbleach value than the short-term FRAP curves depicted in Fig. 3 E as a consequence of the mean projections of image stacks at each time point. (D) Quantification of mean GFP signal

recovery after photobleaching relative to fluorescence intensity before bleaching is shown. Depicted are the mean curves for GFP-H3.1 (dark blue, *n* = 13), GFP-H3.3 (red, *n* = 15), GFP-H3.X (purple, *n* = 18), and GFP-H3.Y (green, *n* = 25). Error bars indicate respective standard deviations. Note that subtle differences in recovery curves are within the dynamic range and that no statistically significant differences in FRAP recovery rates between novel and known H3 variants could be detected. (E) IF microscopy of HeLa cells stably expressing HA-H3.X and -H3.Y costained with DAPI (DNA, blue, left) and α-H3.X/Y (green, middle). Merged pictures are shown on the right. (F) Peptide competition experiment to evaluate epitope specificity of α-H3.X/Y antibody. α-H3.X/Y was preincubated with peptides (1 µg/ml) corresponding to aa 9–20 of H3.X/Y (bottom) or a control peptide corresponding to aa 22–41 of H3.3 (top), then used for IF microscopy with HeLa cells stably expressing HA-H3.X (green, middle). DAPI (blue, left) stains DNA, and merged pictures with antibody stainings are shown on the right. Insets show enlargements of two cells, highlighting the H3.X/Y sequence–specific loss of antibody signal (bottom). Bars: (B) 10 µm; (E) 5 µm; (F) 50 µm.
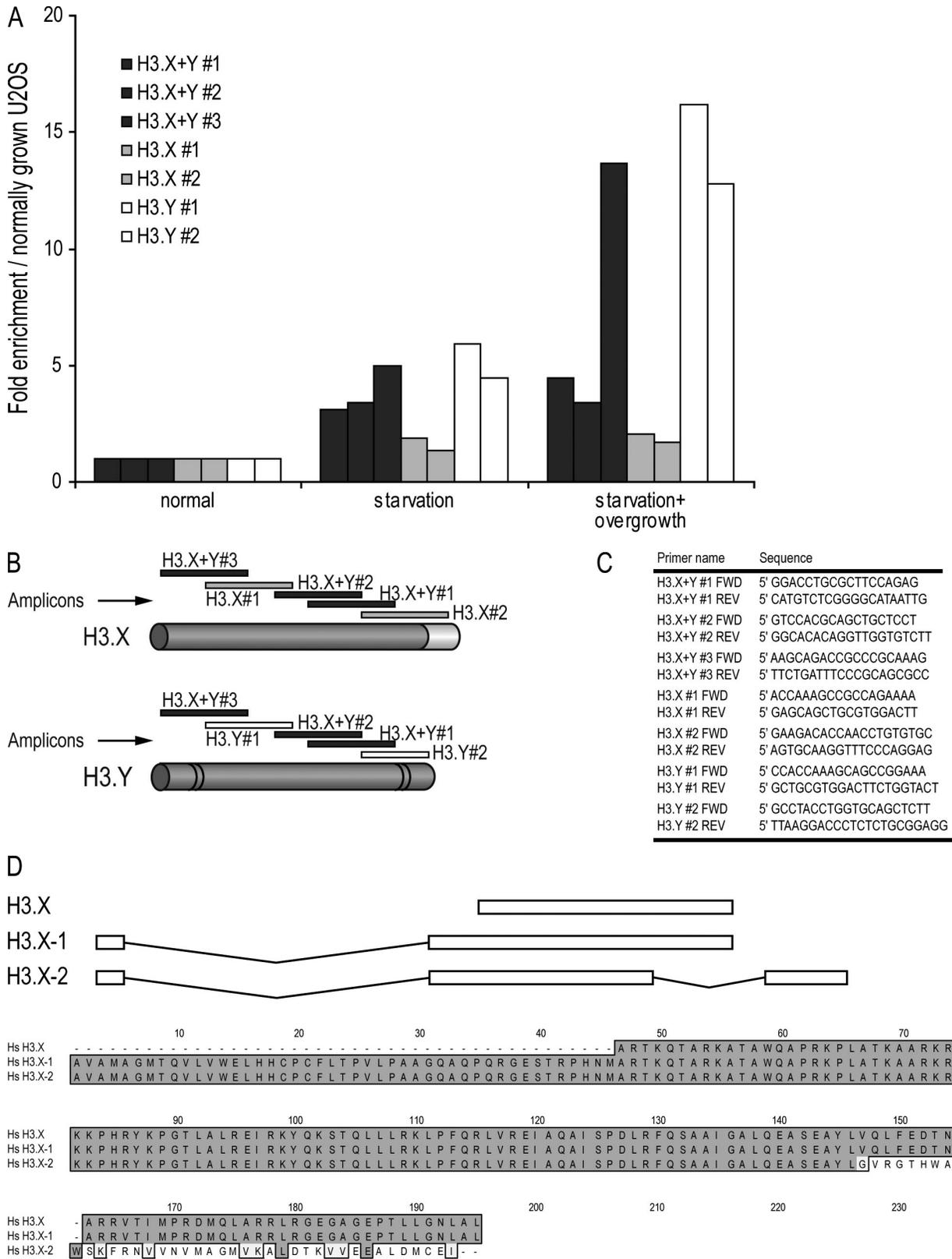
Figure S3. **Evaluation of human H3.X sequences and inducible endogenous H3.X and H3.Y gene expression.** (A) qPCR analysis of H3.X and H3.Y mRNA expression in U2OS cells under different stress conditions using different primer pairs ("primer-walk"). Three primer pairs specific for H3.X+H3.Y (dark gray) and two primer pairs specific for H3.X (light gray) and H3.Y (white) were used. Shown is the result from one of our three biological replicates (Fig. 5). (B) qPCR generated amplicons of each primer pair. (C) List of H3.X+Y-, H3.X-, and H3.Y-specific primer pair sequences. (D, top) Scheme depicting the different transcripts that have been annotated in the NCBI database for the human H3.X locus over the last years. (D, bottom) Alignment of NCBI predicted H3.X protein sequences. Identical amino acids are highlighted in dark gray, similar amino acids are highlighted in light gray, and changes are set apart on a white background. Amino acid numbers are indicated on top.

# A

# B



αH3.X/Y   αH3

# C   H3.X/Y-specific: KYQKSTQLLLR (aa 53-63)

| aa | B | Y |
|---|---|---|
| K | 185.1284545 | - |
| Y | 348.191783 | 1305.752512 |
| Q | 476.2503605 | 1142.689184 |
| K | 660.3715385 | 1014.630606 |
| S | 747.4035669 | 830.5094284 |
| T | 848.4512453 | 743.4774 |
| Q | 976.5098228 | 642.4297215 |
| L | 1089.593887 | 514.371144 |
| L | 1202.677951 | 401.2870801 |
| L | 1315.762015 | 288.2030161 |
| R | - | 175.1189522 |

# D   H3.X/Y-specific: EIAQAISPDLR (aa 73-83)

| aa | B | Y |
|---|---|---|
| E | 130.0498695 | - |
| I | 243.1339335 | 1083.615684 |
| A | 314.1710473 | 970.5316204 |
| Q | 442.2296248 | 899.4945066 |
| I | 513.2667386 | 771.4359291 |
| S | 626.3508025 | 700.3988153 |
| P | 713.3828309 | 587.3147514 |
| D | 810.4355948 | 500.282723 |
| L | 925.4625378 | 403.2299591 |
| L | 1038.546602 | 288.2030161 |
| R | - | 175.1189522 |

# E   H3.Y-specific: KPLATKAAGKR (aa 18-28)

| aa | B | Y |
|---|---|---|
| K | 185.1284545 | - |
| P | 282.1812183 | 1124.678619 |
| L | 395.2652823 | 1027.625856 |
| A | 466.302396 | 914.5417917 |
| T | 567.3500745 | 843.5046779 |
| K | 751.4712525 | 742.4569994 |
| A | 822.5083663 | 558.3358214 |
| A | 893.5454801 | 487.2987077 |
| G | 950.5669438 | 416.2615939 |
| K | 1134.688122 | 359.2401302 |
| R | - | 175.1189522 |

# F   H3.Y-specific: K(Ac)PLATKAAGK(Ac)R (aa 18-28)

| aa | B | Y |
|---|---|---|
| K(Ac) | 171.1128365 | - |
| P | 268.1656003 | 1110.663001 |
| L | 381.2496643 | 1013.610238 |
| A | 452.2867780 | 900.5261737 |
| T | 553.3344565 | 829.4890599 |
| K | 737.4556345 | 728.4413814 |
| A | 808.4927483 | 544.3202034 |
| A | 879.5298621 | 473.2830897 |
| G | 936.5513258 | 402.2459759 |
| K(Ac) | 1106.656886 | 345.2245122 |
| R | - | 175.1189522 |

# G   H3.Y-specific: KPLATK(Ac)AAGKR (aa 18-28)

| aa | B | Y |
|---|---|---|
| K | 185.1284545 | - |
| P | 282.1812183 | 1110.663001 |
| L | 395.2652823 | 1013.610238 |
| A | 466.3023960 | 900.5261737 |
| T | 567.3500745 | 829.4890599 |
| K(Ac) | 737.4556345 | 728.4413814 |
| A | 808.4927483 | 558.3358214 |
| A | 879.5298621 | 487.2987077 |
| G | 936.5513258 | 416.2615939 |
| K | 1120.672504 | 359.2401302 |
| R | - | 175.1189522 |

**Figure S4.   Purification and identification of endogenous H3.Y protein.** (A) RP-HPLC profile of acid-extracted total histones from starved and overgrown U2OS cells. Histones corresponding to different peaks are indicated. The box marks the section of H3 peaks shown in Fig. 6 A. (B) Evaluation of α-H3.X/Y region B bands (Fig. 6 B) by immunoblotting under reducing conditions. To test the possibility that the α-H3.X/Y region B band protein is a dimer of H3.Y and/or H3.X, proteins of these fractions were treated with increasing amounts of DTT and analyzed in immunoblots with α-H3.X/Y. Pooled peak I and II H3 fractions contained H3 dimers that were reduced to monomers after treatment with 250 and 500 mM DTT, whereas the α-H3.X/Y band B proteins did not change running behavior after the same treatment. (C–G) MS/MS spectra of peptides specific for novel H3 variants. NanoLC ESI MS/MS analysis of propionylated and trypsin-digested proteins found in RP-HPLC fractions corresponding to region A bands from U2OS SO experiment No. 1 (Fig. 6 B, iii). MS/MS spectra of the doubly charged precursor ions at m/z 745.44 (C), 606.83 (D), 654.90 (E), 640.88 (F), and 647.90 (G). Boxes on the right show the specific b and y ions after fragmentation.

DNA    αNeuN

αH3.X/Y    αGFAP

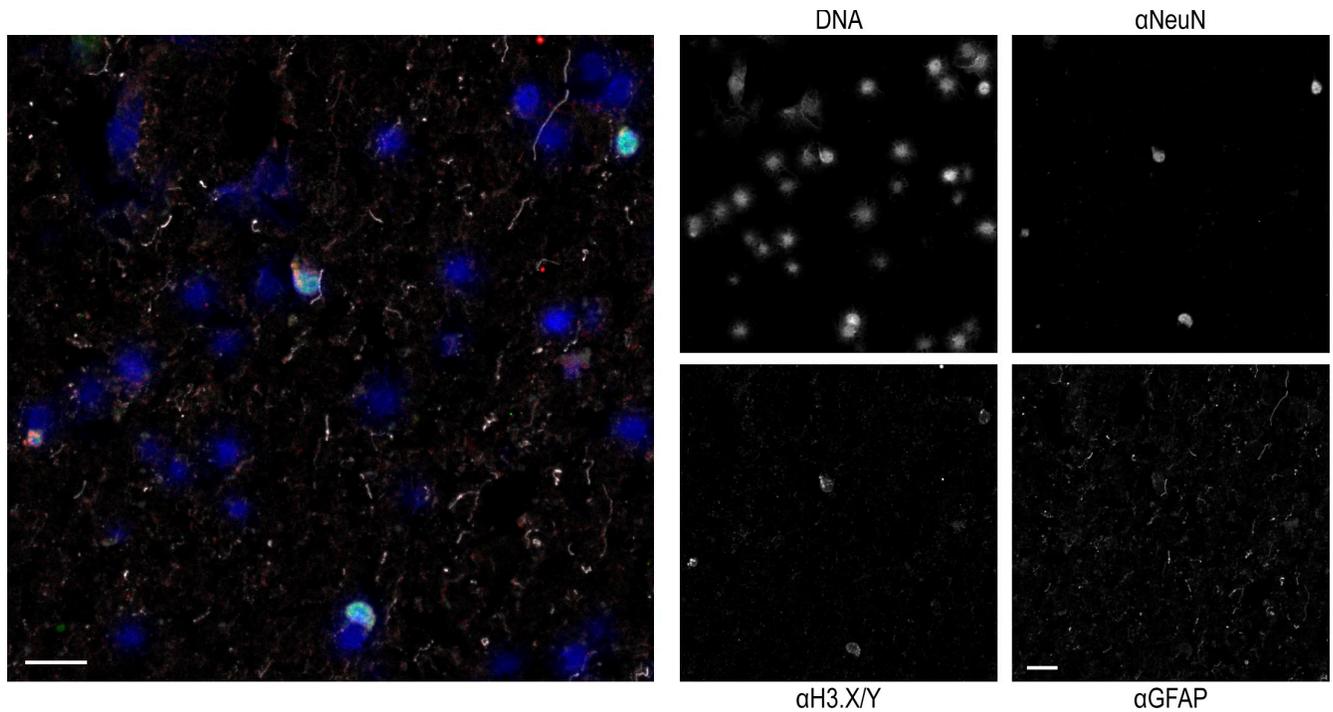Figure S5.  **H3.X/Y expression in the human brain.** Commercially available human hippocampus sections were costained with α-H3.X/Y (red), α-NeuN (neuronal marker, green), α-GFAP (astrocyte marker, white), and DAPI (DNA, blue). One out of three representative stainings is shown. For further costainings of human hippocampus, see Fig. 8. Bars, 20 µm.

Table S1.  **GO list with shared up-regulated genes after H3.Y knockdown**

| GOBPID | P-value | Odds ratio | Exp count | Count | Size | Term |
|---|---|---|---|---|---|---|
| GO:0032274 | 0.000 | 166.667 | 0 | 2 | 5 | Gonadotropin secretion |
| GO:0042523 | 0.000 | 166.667 | 0 | 2 | 5 | Positive regulation of tyrosine phosphorylation of Stat5 protein |
| GO:0046884 | 0.000 | 166.667 | 0 | 2 | 5 | Follicle-stimulating hormone secretion |
| GO:0042522 | 0.000 | 99.983 | 0 | 2 | 7 | Regulation of tyrosine phosphorylation of Stat5 protein |
| GO:0042506 | 0.000 | 83.312 | 0 | 2 | 8 | Tyrosine phosphorylation of Stat5 protein |
| GO:0060986 | 0.001 | 62.473 | 0 | 2 | 10 | Endocrine hormone secretion |

Comprehensive GO list of combined genes significantly up-regulated upon H3.Y and H3.X+Y RNAi in comparison to luciferase control RNAi (see the Venn diagram in Fig. 7 C for the number of overlapping genes). Count, actual number of responders found in the node; Exp Count, number of responders expected (entering the term by chance); GOBPID, GO identified in the biological process domain (numbers are identifiers based on the GO database released by the Gene Ontology Consortium; http://www.geneontology.org/); Size, total number of genes analyzed in this node (GO term/group).

**Table S2.  GO list with shared down-regulated genes after H3.Y knockdown**

| GOBPID | P-value | Odds Ratio | Exp Count | Count | Size | Term |
|---|---|---|---|---|---|---|
| GO:0000280 | 0.000 | 6.951 | 3 | 20 | 248 | Nuclear division |
| GO:0007067 | 0.000 | 6.951 | 3 | 20 | 248 | Mitosis |
| GO:0048285 | 0.000 | 6.710 | 4 | 20 | 256 | Organelle fission |
| GO:0000087 | 0.000 | 6.653 | 4 | 20 | 258 | M phase of mitotic cell cycle |
| GO:0051301 | 0.000 | 5.937 | 4 | 22 | 318 | Cell division |
| GO:0000279 | 0.000 | 5.515 | 5 | 22 | 340 | M phase |
| GO:0022403 | 0.000 | 4.767 | 6 | 24 | 428 | Cell cycle phase |
| GO:0007049 | 0.000 | 3.538 | 12 | 35 | 864 | Cell cycle |
| GO:0000278 | 0.000 | 4.671 | 6 | 24 | 436 | Mitotic cell cycle |
| GO:0071103 | 0.000 | 9.106 | 2 | 13 | 122 | DNA conformation change |
| GO:0022402 | 0.000 | 4.019 | 8 | 27 | 571 | Cell cycle process |
| GO:0006323 | 0.000 | 9.123 | 1 | 11 | 102 | DNA packaging |
| GO:0007059 | 0.000 | 10.287 | 1 | 10 | 83 | Chromosome segregation |
| GO:0006261 | 0.000 | 11.797 | 1 | 9 | 66 | DNA-dependent DNA replication |
| GO:0065004 | 0.000 | 9.327 | 1 | 9 | 81 | Protein–DNA complex assembly |
| GO:0000075 | 0.000 | 8.952 | 1 | 9 | 84 | Cell cycle checkpoint |
| GO:0006334 | 0.000 | 9.422 | 1 | 8 | 71 | Nucleosome assembly |
| GO:0051276 | 0.000 | 3.453 | 7 | 20 | 470 | Chromosome organization |
| GO:0006260 | 0.000 | 4.924 | 3 | 13 | 213 | DNA replication |
| GO:0031497 | 0.000 | 8.856 | 1 | 8 | 75 | Chromatin assembly |
| GO:0009066 | 0.000 | 21.496 | 0 | 5 | 22 | Aspartate family amino acid metabolic process |
| GO:0034728 | 0.000 | 8.475 | 1 | 8 | 78 | Nucleosome organization |
| GO:0006270 | 0.000 | 20.300 | 0 | 5 | 23 | DNA replication initiation |
| GO:0009067 | 0.000 | 29.070 | 0 | 4 | 14 | Aspartate family amino acid biosynthetic process |
| GO:0051726 | 0.000 | 3.772 | 4 | 14 | 295 | Regulation of cell cycle |
| GO:0006259 | 0.000 | 2.988 | 7 | 19 | 508 | DNA metabolic process |
| GO:0000070 | 0.000 | 12.588 | 0 | 5 | 34 | Mitotic sister chromatid segregation |
| GO:0006996 | 0.000 | 2.203 | 18 | 35 | 1,311 | Organelle organization |
| GO:0000819 | 0.000 | 12.168 | 0 | 5 | 35 | Sister chromatid segregation |
| GO:0051318 | 0.000 | 19.372 | 0 | 4 | 19 | G1 phase |
| GO:0009086 | 0.000 | 43.353 | 0 | 3 | 8 | Methionine biosynthetic process |
| GO:0006333 | 0.000 | 5.579 | 2 | 8 | 114 | Chromatin assembly or disassembly |
| GO:0006082 | 0.000 | 2.841 | 7 | 18 | 502 | Organic acid metabolic process |
| GO:0007091 | 0.000 | 16.139 | 0 | 4 | 22 | Mitotic metaphase/anaphase transition |
| GO:0007093 | 0.000 | 10.425 | 1 | 5 | 40 | Mitotic cell cycle checkpoint |
| GO:0007094 | 0.000 | 30.961 | 0 | 3 | 10 | Mitotic cell cycle spindle assembly checkpoint |
| GO:0045841 | 0.000 | 30.961 | 0 | 3 | 10 | Negative regulation of mitotic metaphase/anaphase transition |
| GO:0008652 | 0.000 | 9.599 | 1 | 5 | 43 | Cellular amino acid biosynthetic process |
| GO:0031570 | 0.000 | 9.118 | 1 | 5 | 45 | DNA integrity checkpoint |
| GO:0007076 | 0.000 | 27.089 | 0 | 3 | 11 | Mitotic chromosome condensation |
| GO:0071174 | 0.000 | 27.089 | 0 | 3 | 11 | Mitotic cell cycle spindle checkpoint |
| GO:0009309 | 0.000 | 6.867 | 1 | 6 | 70 | Amine biosynthetic process |
| GO:0044283 | 0.000 | 3.415 | 4 | 12 | 275 | Small molecule biosynthetic process |
| GO:0016043 | 0.001 | 1.813 | 33 | 51 | 2,371 | Cellular component organization |
| GO:0006555 | 0.001 | 24.077 | 0 | 3 | 12 | Methionine metabolic process |
| GO:0071173 | 0.001 | 24.077 | 0 | 3 | 12 | Spindle assembly checkpoint |
| GO:0019752 | 0.001 | 2.682 | 7 | 17 | 498 | Carboxylic acid metabolic process |
| GO:0043436 | 0.001 | 2.682 | 7 | 17 | 498 | Oxoacid metabolic process |
| GO:0010948 | 0.001 | 12.098 | 0 | 4 | 28 | Negative regulation of cell cycle process |
| GO:0007346 | 0.001 | 4.575 | 2 | 8 | 137 | Regulation of mitotic cell cycle |
| GO:0031577 | 0.001 | 21.667 | 0 | 3 | 13 | Spindle checkpoint |
| GO:0045839 | 0.001 | 21.667 | 0 | 3 | 13 | Negative regulation of mitosis |
| GO:0051784 | 0.001 | 21.667 | 0 | 3 | 13 | Negative regulation of nuclear division |
| GO:0042180 | 0.001 | 2.614 | 7 | 17 | 510 | Cellular ketone metabolic process |
| GO:0000097 | 0.001 | 19.696 | 0 | 3 | 14 | Sulfur amino acid biosynthetic process |

Comprehensive GO list of combined genes significantly down-regulated upon H3.Y and H3.X+Y RNAi in comparison to luciferase control RNAi (see the Venn diagram in Fig. 7 C for the number of overlapping genes). Count, actual number of responders found in the node; Exp Count, number of responders expected (entering the term by chance); GOBPID, GO identified in the biological process domain (numbers are identifiers based on the GO database released by the Gene Ontology Consortium; http://www.geneontology.org/); Size, total number of genes analyzed in this node (GO term/group).

# APPENDIX

## CONTRIBUTIONS

Declaration of contributions to **"Chromatin proteomics and epigenetic regulatory circuits"**

This project was conceived by Sandra B. Hake. I wrote the sections on ATP-dependent chromatin remodeling and nuclear architecture. In addition, I put together the individual section and set up the final manuscript.

Declaration of contributions to **"H2A.Z.2.2 is an alternatively spliced histone H2A.Z variant that causes severe nucleosome destabilization"**

I conceived the project, cloned H2A.Z.2.2 initially from cDNA and all other constructs with a few exceptions that are mentioned in the manuscript. I established all cell lines used, designed and performed all experiments with a few exception. Figure 1C: I prepared the acid extracted histones and carried out the western blot, Sonja M. Wiedemann purified the acid extracted histones via HPLC; Figure 3B-D and Supplementary Figure S3: Katrin Schneider and Christina Bielmeier; Figure 4: H. Christian Eberl; Figure 5: Christina Bielmeier; Figure 7D and Supplementary Figure S6: I prepared labeled nucleosomes and Wolfgang Kuegel performed and analyzed FRET measurements; Figure 8 and Supplementary Figure S7: Marco Bocola. I made all the figures. Sandra B. Hake wrote the first draft of the manuscript and we developed the final manuscript together.

Declaration of contributions to **"ATRX-mediated chromatin association of histone variant macroH2A1 regulates α-globin expression"**

This project was performed in collaboration with the lab of Emily Bernstein. I cloned HA-tagged H3 constructs and established and characterized stable HeLa and HEK293 cell lines. In addition, I performed size exclusion chromatography with chromatin free extracts to confirm the absence of chromatin from these extracts and to analyze the fractionation behavior of ATRX, macroH2A, H3.3 and DAXX (data not shown).

Declaration of contributions to **"Identification and characterization of two novel primate-specific histone H3 variants, H3.X and H3.Y"**

This project was conceived by Sandra B. Hake. I cloned all GFP-tagged H3 constructs and all HA-tagged H3 constructs with the exception of HA-H3.Y. In addition, I established stable HeLa cell lines with the exception of the one expressing HA-H3.Y. Furthermore, I established the protocol to perform MNase digestion of chromatin from HeLa cells and performed the IP as well as the subsequent experiments shown in Figure 3C.

## CURRICULUM VITAE

### Personal Information

Name                    Clemens Bönisch

Date of birth           November 04th 1981

Place of birth          Halle (Saale)

Nationality             German

### Education

04/07 – 07/12          **Graduate student (Dr. rer. nat.) at LMU München** in Priv.-Doz. Dr. Sandra
                       Hake´s group (Adolf-Butenandt-Institute, chair Prof. Dr. Peter Becker)

                       Project: "Structure/function analyses of mammalian histone H2A and H3
                       variants"

09/06 – 03/07          **Scientific coworker at MLU Halle-Wittenberg** in Prof. Dr. Elmar Wahle´s
                       group, (Institute for Biochemistry), successfully finishing and publishing
                       project started in diploma thesis

10/01 – 07/06          **Study of biochemistry at MLU Halle-Wittenberg**

                       11/05 – 07/06: Diploma thesis in Prof. Dr. Elmar Wahle´s group,
                       (Institute for Biochemistry)

                       Project: "Degradation of Hsp70 mRNA in *Drosophila melanogaster*"

10/00 – 08/01          **Alternative civil service (Zivildienst)**

                       Information center for the elderly (Seniorenberatungsstelle) of the city of
                       Halle (Saale) as household assistant (Hauswirtschaftshilfe)

08/92 – 07/00          **Secondary school** "Georg-Cantor Gymansium", Halle (Saale)
                       Abitur

## Publications

**Bönisch C**, Schneider K, Pünzeler S, Wiedemann SM, Bielmeier C, Bocola M, Eberl HC, Kuegel W, Neumann J, Kremmer E, Leonhardt H, Mann M, Michaelis J, Schermelleh L and Hake SB. H2A.Z.2.2 is an alternatively spliced histone H2A.Z variant that causes severe nucleosome destabilization. *Nucleic Acids Res.* 2012 Mar 29. doi:10.1093/nar/gks267

Ratnakumar K, Duarte LF, LeRoy G, Hasson D, Smeets D, Vardabasso C, **Bönisch C**, Zeng T, Bin X, Zhang DY, Li H, Wang X, Hake SB, Schermelleh L, Garcia BA and Bernstein E. ATRX-mediated chromatin association of histone variant macroH2A1 regulates a-globin expression. *Genes Dev.* 2012 Mar; 26 (5): 433-8.

Müller-Taubenberger A, **Bönisch C**, Fürbringer M, Wittek F and Hake SB. The histone methyltransferase Dot1 is required for DNA damage repair and proper development in Dictyostelium. *Biochem Biophys Res Commun*. 2011 Jan; 404 (4): 1016-22

Wiedemann SM, Mildner SN, **Bönisch C**, Chioda M, Straub T, Israel L, Maiser A, Merkl R, Leonhardt H, Kremmer E, Schermelleh L, Hake SB. Identification and characterization of the two novel primate-specific histone H3 variants H3.X and H3.Y. *J Cell Biol.* 2010 Sep; 190 (5): 777-791

**Bönisch C**\*, Nieratschker SM\*, Orfanos NK, Hake SB. Chromatin proteomics and epigenetic regulatory circuits. *Expert Rev Proteomics*. 2008 Feb; 5 (1): 105-19 (\* equal contribution)

**Bönisch C**\*, Temme C\*, Moritz B, Wahle E. Degradation of hsp70 and other mRNAs in Drosophila via the 5' 3' pathway and its regulation by heat shock. *J Biol Chem.* 2007 Jul; 282 (30): 21818-28 (\* equal contribution)

**ACKNOWLEDGEMENT**

First I want to thank Priv.-Doz. Dr. Sandra Hake for giving me the opportunity to work as a PhD student in her group. I am very grateful that you always found time to discuss my work, for providing valuable input and for just being a nice boss who understands that work is not everything in life. I always appreciated your attitude to always discuss on eye level without ever being bossy, especially when things were not going well. I thank you for always supporting me to develop my own ideas. Your style of supervision was exactly what I needed and helped me a lot to develop both as a scientist and as a person.

I want to thank Prof. Dr. Peter Becker, Prof. Dr. Axel Imhof and Dr. Tobias Straub for being members of my thesis advisory committee and giving me honest and helpful input on my work whenever I needed it.

I am very grateful for having had the opportunity to work in outstanding collaborations with excellent scientists. I want to especially thank all co-authors of the H2A.Z.2.2 paper: Katrin Schneider, Sebastian Pünzeler, Sonja M. Wiedemann, Christina Bielmeier, Marco Bocola, H. Christian Eberl, Wolfgang Kuegel, Jürgen Neumann, Elisabeth Kremmer, Heinrich Leonhardt, Matthias Mann, Jens Michaelis, Lothar Schermelleh and Sandra B. Hake. Without your help and your contributions this paper would not have been published!

I am glad that I could work in the motivating and scientifically very inspiring environment of the Adolf-Butenandt-Institute. Especially, I want to thank all present and former members of the Hake and Korber group ("the north lab") to create a great work environment and constantly filling the cookie drawer. I want to point out that I do not thank Sebastian Sommer to challenge me in speed and efficiency to empty it again. In addition, I want to thank Dr. Tobias Straub for his time to discuss all scientific mysteries ranging from experimental design to special western blot applications as well as instant help with all kinds of computer problems. I am deeply obliged to Felix Müller-Planitz, Henrike Klinker and Silvia Dambacher for introducing me in the fascinating and sometimes surprisingly challenging world of chromatin assembly *in vitro*. In addition, I want to thank Silvia Dambacher for providing me with recombinant histones which made experiments faster and life easier. Moreover, I want to thank Torsten Fauth, Christian Wippo and Jochen Abel for paper discussions in the "GAP" and the latter two and Tobias Straub for sharing numerous thrilling and highly competitive "Kicker Manager Game Pro" seasons.

Last but not least, I want to thank all the people in my life outside the lab, my family and my friends. Especially, I want to thank Antonia (a.k.a. Frau Jack) – my girlfriend – for becoming much more than a colleague. You always helped me to keep a good work/life balance and showed me that there is more in life than unsuccessful cloning. Vielen Dank!