Characterization of Nucleosomes Containing Specific Forms of the Histone Variant H2A.Z

MARLEE KAM NG

A DISSERTATION SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN BIOLOGY
YORK UNIVERSITY
TORONTO, ONTARIO

August 2019

Abstract

H2A.Z is a highly conserved histone variant that replaces canonical histone H2A at specific loci to regulate diverse nuclear processes. Amongst these, the role of H2A.Z in transcriptional regulation is of particular interest due to its enrichment at promoters of most genes in yeast and higher eukaryotes. However, its precise role in regulation is complex, as it has been linked to both repression and activation. One possibility is that H2A.Z activity is regulated by post-translational modifications since H2A.Z can be acetylated or monoubiquitylated in mammals. For example, H2A.Z can be multiply acetylated at several lysine residues at its N-terminus, and such modified form is associated with active promoters. In contrast, our lab has previously shown that a fraction of H2A.Z is monoubiquitylated at its C-terminus, and this form is associated with silent chromatin. One aim of this thesis is to characterize monoubiquitylated H2A.Z-nucleosomes in the context of transcriptional regulation. To this end, we devised a biotinylation-based method to enrich for H2A.ZUb1-mononucleosomes, and further characterized their composition and genomic distribution. In the second chapter, I demonstrate that H2A.ZUb1-enriched mononucleosomes are enriched with the histone post-translational modification H3K27me3, but depleted of H3K4 methylation and other modifications associated with transcriptional activity. H2A.ZUb1-eniched mononucleosomes also preferentially co-purify with proteins typically involved in repression, and with CTCF and cohesin. Consistent with these, ChIP-Seq analysis of H2A.ZUb1-nucleosomes identifies non-expressed genes as sites of H2A.ZUb1 enrichment. In addition to post-translational modification, vertebrate H2A.Z is differentiated into non-allelic isoforms H2A.Z-1 and H2A.Z-2. Previously, we used mass-

spectrometry to identify proteins that preferentially associate with H2A.Z-mononucleosomes over H2A-mononucleosomes. In the third chapter, I show that one of these proteins, USP39, is enriched on mononucleosomes containing H2A.Z-1 over those containing H2A.Z-2, and that this selectivity can be mapped to an isoform-specific residue in its C-terminal tail. USP39 is a component of the U4/U6.U5 tri-snRNP, and consistent with a functional link between H2A.Z-1 and USP39, we identify a subset of shared alternative splicing events. Altogether, these data support functional diversification of H2A.Z through monoubiquitylation and isoform-specific amino acid substitution, and collectively, contribute to our understanding of biological pathways converging on H2A.Z.

Acknowledgments

I can't thank Dr. Peter Cheung enough for everything. Thank you especially for doing your best to help me understand that things are seldom black or white and that this is not so bad. You have been a great (and greatly patient) mentor, Peter, and I have learned so much from you.

Thank you also to my committee members, Dr. Emanuel Rosonina and Dr. Mark Bayfield, who both really pulled through to help me complete this thesis.

I would like to dedicate this thesis to my parents, Denny and Johanne, and thank them for everything imaginable. Plus, without them, I'd not have made it to university in the first place, and it was a very good thing that I went, otherwise I'd never have known about all the great stuff there is which I know nothing about. Shout out to Jordan, too, for making sure a sister knows she is never done learning.

I would also like to thank past and present members of the Cheung lab, especially Priscilla Lau, Cindy Law, Ryan Draker, Keyur Adhvaryu, Kashif Khan, Jessica D'Angelo, Myat Ma, and Shahir Morocos for your help. As well as my lab homies: Peter Liuni, Mohamed Salem, and Jelena Brkic. And best homies back home: Kori Novak and Natasha Mills.

This thesis would not have been possible without Dr. Ulrich Braunschweig of the Blencowe lab, who analyzed the high-throughput data. I would really like to thank Dr. Braunschweig for his generosity and expertise, as well as the Donnelly Sequencing Centre for sequencing samples. Finally, I would like to thank Dr. Benjamin Blencowe for his collaboration and helpful suggestions for the research described in Chapter 3 of this thesis.

Table of Contents

List of Figures

List of Abbreviations

5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
AS	Alternative splicing
Avi	Avidin
Avi-HRP	Avidin-conjugated horseradish peroxidase
CE	Cassette exon
ChIP	Chromatin immunoprecipitation
ChIP-Seq	Chromatin immunoprecipitation and sequencing
cPRC1	Canonical Polycomb repressive complex 1
CT	chromosome territory
CTCF	CCCTC-binding factor
DNMT3L	DNA (cytosine-5)-methyltransferase 3-like
FB	Flag-BirA
HDAC	Histone deacetylase
HEK293T	Human embryonic kidney cells (expressing SV40 large T antigen)
IP	Immunoprecipitation
IP-MS	Immunoprecipitation-mass spectrometry
IR	Intron retention
LINE	Long interspersed nuclear element
LSD1	Lysine-specific demethylase
me1	Monomethylation
me2	Dimethylation
me3	Trimethylation
mESC	Mouse embryonic stem cell
MLL	Mixed-lineage leukemia
MNase	Micrococcal nuclease
ncPRC1	Non-canonical Polycomb repressive complex 1
NDR	Nucleosome depleted region
NEM	N-ethylenamine
NLS	Nuclear localization signal
NuRD	Nucleosome remodeling deacetylase
PcG	Polycomb group
PRC2	Polycomb repressive complex 2
PRE	Polycomb response element
PTM	Post-translational modification
qPCR	Quantitative polymerase chain reaction

RT-PCR	Reverse transcription-polymerase chain reaction
SINE	Short interspersed nuclear element
SMC1	Structural maintenance of chromosome protein 1
snRNP	Small nuclear ribonucleoproteins
SR	Serine/arginine
TAD	Topological associated domain
TrxG	Trithorax group
TSS	Transcriptional start site
Ub1	Monoubiquitylation
USP39	Ubiquitin-specific protease 39

Chapter 1

General Introduction

1 General Introduction

1.1 Chromatin structure and function

The human haploid genome comprises roughly 3.0 x 10^9 base pairs (bp) of DNA and encodes around 20,000 – 25,000 protein-coding genes (Venter et al., 2001). In a diploid cell, this naked 6 x 10^9 bp of genetic material would be ~2 meters long if drawn-out end-to-end, but instead, it is intricately packaged within a nucleus only 10 μm in diameter. To accomplish this organizational feat, eukaryotic DNA is bound with histone proteins, in regular and repeating arrays known as nucleosomes, which are the first-order of DNA packaging within the nucleus. The nucleosome is an octomeric particle consisting of two copies each of histones H2A, H2B, H3, and H4, about which 147 bp of DNA is wrapped in 1.67 left-handed superhelical turns. Core particles are connected by stretches of "linker DNA", which can be up to 80 bp long. Linker histones, such as histone H1, bind to the core particle where DNA enters and exits the nucleosome, and to the inter-nucleosomal DNA (Kornberg, 1974, 1977; Olins and Olins, 2003; Oudet et al., 1975) (Fig. 1-1).

In 1884, A. Kossel first purified and coined the term histones for the highly abundant, acid-soluble component of nuclei, which through early studies, were shown to inhibit transcription by virtue of their association with DNA (Kossel, 1883; Olins and Olins, 2003). In 1964, Vincent Allfrey demonstrated by that histones were amenable to post-translational methylation and acetylation, and he proposed that the latter modification served to reduce their inhibitory effect on transcription (Allfrey et al., 1964).

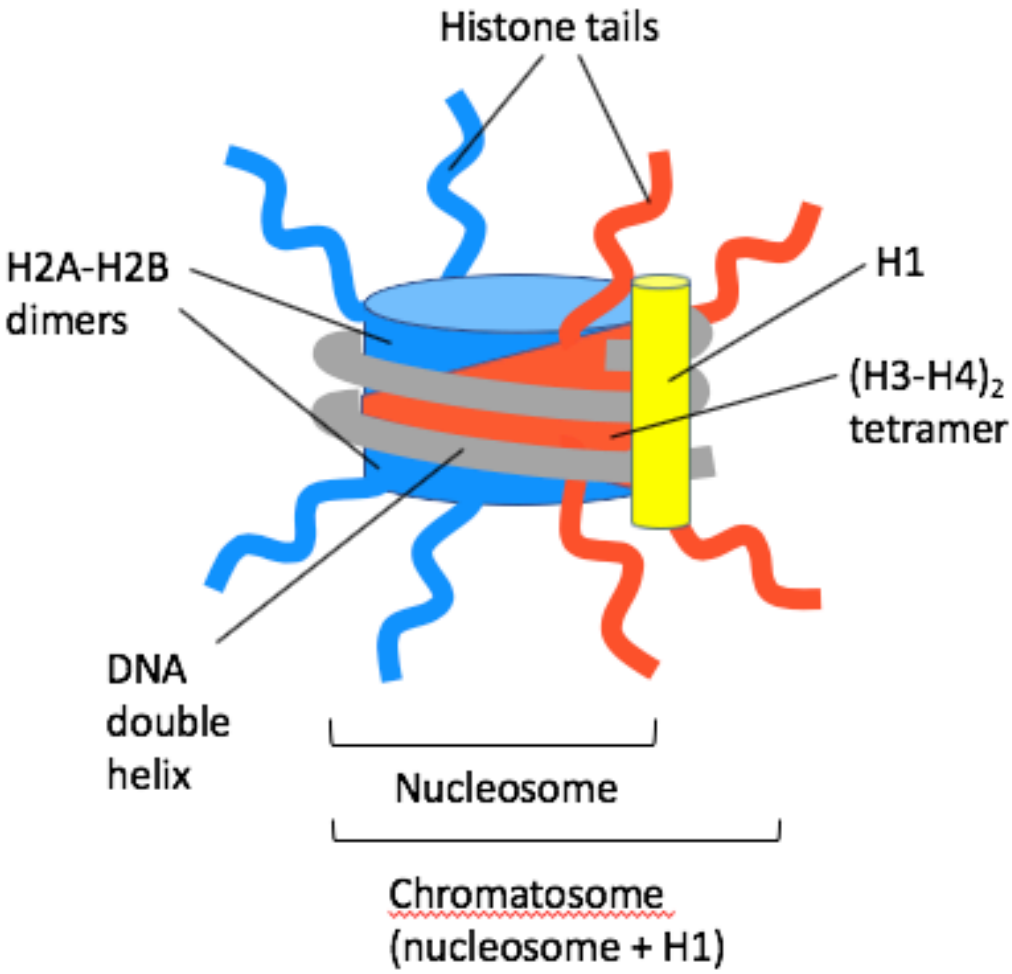


Figure 1-1. Basic structure of the nucleosome. A cartoon version of the basic structure of the nucleosome, showing the two H2A-H2B dimers in blue, and the (H3-H4)2 tetramer in orange, which together comprise the histone octamer. Approximately 150bp of DNA wrap around the histone octamer in a left-handed helix of roughly 1.7 turns. Also shown are the N-terminal tails of the histones, which are unstructured and protrude from the surface of the nucleosome. Histone H1 binds the nucleosome at the DNA entry/exit point.

A major breakthrough in the chromatin field came in the mid-1970s, when the basis for the repeating nature of chromatin as a 200 bp nuclease-resistant fragment was ascribed to an approximately spherical complex of histones, later structurally defined as the “histone octamer”, comprising a central $(H3\text{-}H4)_2$ tetramer flanked by two H2A-H2B dimers (Kornberg, 1974; Olins and Olins, 2003). The discovery of the nucleosome revolutionized the perception of chromatin as a passive packaging medium to a dynamic and organized scaffold, and the nucleosome as the fundamental unit upon which all DNA-templated processes converge (Olins and Olins, 2003).

The high degree of genome compaction achieved by chromatin organization must contend with its accessibility to other factors in order to meet activity demands of the cell. Additionally, while only ~1.5% of the genome is protein-coding, the rest includes non-coding elements such as telomeres, satellite DNA, specialized RNAs, repetitive elements such as LINES and SINES, as well as regulatory sequences such as promoters, enhancers and insulators (de Laat and Duboule, 2013; Lindblad-Toh et al., 2011; Plank and Dean, 2014; Venter et al., 2001). Enhancers are often distal to their cognate promoters and have been shown to function through long-range interactions that loop-out large intervening regions of chromatin (Dekker et al., 2002; Gibcus and Dekker, 2013; de Laat and Duboule, 2013). To satisfy these regulatory constraints, chromatin is non-randomly packaged at multiple layers into spatially segregated regions. Chromatin structure has been shown to be an important regulator of DNA-templated processes at all scales, and hence characterizing the principles that underlie its conformations is essential to understanding the regulation of these activities. Recent advancements in next-generation sequencing technologies have greatly facilitated our

present global view of chromatin topology. These techniques provided mounting evidence that support the widespread existence of folded structures beyond nucleosomes, including loops, topological associating domains (TADs), chromatin domains, and chromosome territories (Gibcus and Dekker, 2013).

1.1.1 Higher-order chromatin structures

During interphase, each chromosome occupies a roughly spherical, spatially distinct "territory", that intermingles with immediately adjacent chromosomes at its peripheries, creating a contiguous body of chromatin. Chromosome territories (CTs) are radially positioned such that gene-rich domains occupy the nuclear interior while gene-poor domains within chromosomes tend to localize at the nuclear periphery and associate with the nuclear lamina (lamina-associated domains; LADs)(Croft et al., 1999; Fritz et al., 2016; Lieberman-Aiden et al., 2009; Tanabe et al., 2002; Zhang et al., 2012). Fluorescence *in situ* hybridization (FISH) has also demonstrated that gene-rich regions tend to localize to the periphery of their CT, which facilitates the intermingling of these loci across different chromosomes (Branco and Pombo, 2006) (Fig. 1-2).

Two general structural states within CTs (termed domains) were initially identified cytologically by how well they stained: *Heterochromatin*, which assumes highly condensed structures characterized by many inter-nucleosomal contacts, stains intensely, whereas *euchromatin*, a state in which chromatin is relatively uncondensed, stains poorly (Bickmore and Sumner, 1989; Craig and Bickmore, 1993; Holmquist, 1992; Holmquist et al., 1982; Passarge, 1979). Many bodies of evidence have since established that the degree of chromatin compaction is inversely correlated with

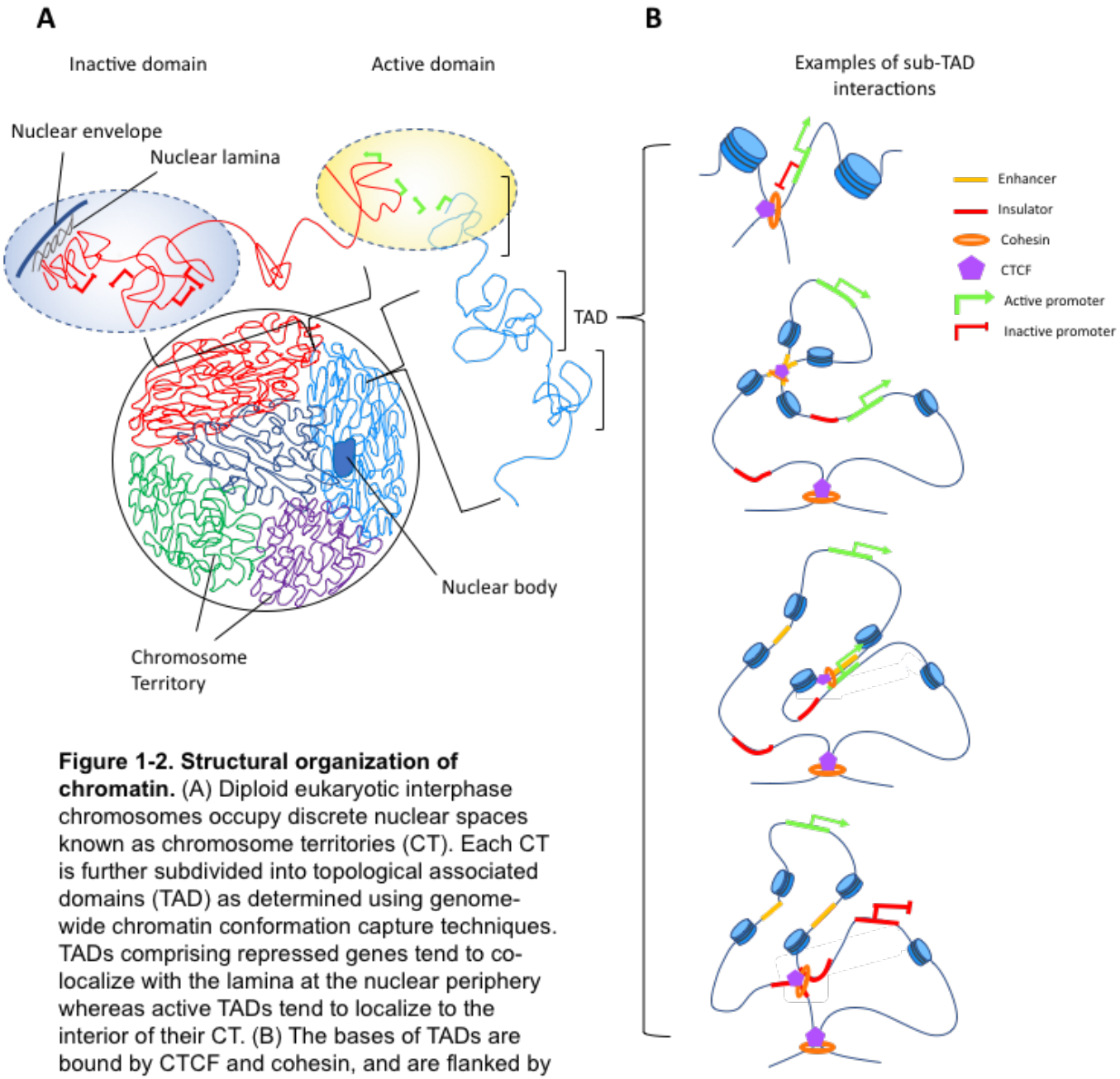


Figure 1-2. Structural organization of chromatin. (A) Diploid eukaryotic interphase chromosomes occupy discrete nuclear spaces known as chromosome territories (CT). Each CT is further subdivided into topological associated domains (TAD) as determined using genome-wide chromatin conformation capture techniques. TADs comprising repressed genes tend to co-localize with the lamina at the nuclear periphery whereas active TADs tend to localize to the interior of their CT. (B) The bases of TADs are bound by CTCF and cohesin, and are flanked by regions of relatively low interaction frequency. Additional intra-TAD interactions occur between *cis* regulator elements mediated by CTCF and cohesin.

transcriptional activity and gene density (Ciabrelli and Cavalli, 2015; Gilbert et al., 2004; Sabo et al., 2004; Weil et al., 2004). Heterochromatin can be further distinguished as *facultative* or *constitutive*. Facultative heterochromatin can interconvert between euchromatic and heterochromatic states depending on needs of the cell, and were first identified as regions that stained differently between cell-types. Such domains often comprise genes that are expressed during differentiation and development, and are subsequently silenced in a cell-type specific manner, and encompass genes that maintain transcriptional tractability (Trojer and Reinberg, 2007). Constitutive heterochromatin, in contrast, maintains an transcriptionally refractive conformation, organizing permanently silenced genes, centromeres, telomeres, and other repetitive DNA elements (Allshire and Madhani, 2018). Structurally alike domains associate through long-range interactions, with heterochromatin domains primarily interacting with other heterochromatin regions on the same chromosome arm, and euchromatin domains interacting with other euchromatin domains within the same chromosome arm, a different chromosome arm, or on other chromosomes (Branco and Pombo, 2006; Simonis et al., 2006; Würtele and Chartrand, 2006; Zhao et al., 2006). This intermingling of domains gives rise to “superdomains” composed of multiple, structurally similar chromosome regions. In some cases, intermingling amongst chromatin domains of similar function gives rise to membrane-less compartments known as “nuclear bodies”. Nuclear bodies are discrete, liquid-like droplets, whose phase-separation is self-driven by specific multivalent protein interactions amongst intrinsically disordered protein regions (IDRs) (Erdel and Rippe, 2018). Nuclear bodies are functionally coordinated macromolecular assemblies involved in specific and diverse cellular

processes. They include, for example: nucleoli, which are involved in ribosome biogenesis; nuclear speckles, which process RNA; transcription factories, which comprise high concentrations of active RNAPII, and conversely, Polycomb bodies, which are depleted of active RNAPII (Cavalli and Misteli, 2013; Erdel and Rippe, 2018; Lamond and Spector, 2003; Misteli, 2005).

Recent technological development of genome-wide chromatin conformation capture techniques (e.g. 4C, HiC, and derivatives) has enabled construction of high-resolution contact frequency maps, revealing chromatin folding within domains. These studies have established the conserved existence of preferentially interacting subdomain “globules”, on scales of tens of kilobases to several megabases (Dekker et al., 2002; Gibcus and Dekker, 2013). These so-called topological associated domains or “TADs” are regions of high local contact frequency containing tens of genes and hundreds of enhancers, and are insulated from other TADs by constrained boundaries across which little contact occurs. TAD boundaries are generally maintained between cell-types; however, their spatial positioning within CTs, as well as the long-range interactions that occur within them are lineage-specific (Dixon et al., 2012; Filippova et al., 2014; Nora et al., 2012). Consistent with these, variability in TAD boundaries and intra-TAD looping events have been shown to regulate specific genetic programs involved in establishing and maintaining cell identity; for example, super-enhancers, which are clusters of active enhancers densely populated by the five master transcription factors (i.e. Oct4, Sox2, Nanog, Klf4, and Esrrb) and which control the expression of lineage-specific genes, are often encompassed within TAD neighbourhoods (Dowen et al., 2014; Huang et al., 2018; Peng and Zhang, 2018).

Boundary regions that insulate chromatin domains, TADs, and intra-TAD looping events are enriched for CCCTC-binding factor (CTCF), an 11-zinc-finger, sequence-specific DNA-binding architectural protein (Dixon et al., 2012; Merkenschlager and Odom, 2013). At the anchored bases of TADs and intra-TAD loops, CTCF co-localizes with the ring-like multiprotein complex, cohesin, which is best known for providing cohesion between sister chromatids during mitosis. Given that CTCF and cohesin are implicated in multiple levels of chromatin folding that can lead to distinct regulatory outcomes – for example, the insulation of regulatory elements through their segregation into adjacent TADs, or the bridging of promoter-enhancer interactions within TADs (Ali et al., 2016; Ong and Corces, 2014) – factors that differentiate these sites are of key interest. At the same time, while the correlation between chromatin folding state and transcriptional activity have been widely documented, the causal effects of these processes, or how they intersect at multiple scales, are poorly understood.

Importantly, research in the past nearly two decades has established that chromatin domains and TADs can be characterized by their epigenetic landscapes (Cain et al., 2011; Gonzalez-Sandoval and Gasser, 2016; Imakaev et al., 2012). These comprise changes in gene expression through mechanisms operating at the nucleosome-level, and include DNA methylation, non-coding RNA, as well as histone variant exchange and histone post-translational modification (PTM) (Gonzalez-Sandoval and Gasser, 2016). The strong correlation between histone modification patterns and specific DNA activities has also been widely observed (Rothbart and Strahl, 2014) and; therefore, understanding how PTMs operate at the nucleosome-level provides a link to

understanding the functional interplay between genomic activity and chromatin structure.

1.1.2 Histone modifications

Nucleosomes are the first order of genome compaction and are dynamically regulated at specific loci by the exchange of canonical histones for variants and by post-translational modification (Jenuwein and Allis, 2001). According to the prevailing model, strings of nucleosomes exist as 10 nm chromatin fibers, which constitute the template for higher-order folding of euchromatin regions (Wolffe, 1998). Heterochromatic nucleosomes, on the other hand, adopt a 30 nm, regular helical structure having a packing density of about 6 to 7 nucleosomes per 11 nm (Song et al., 2014). Histone modifications, such as variant histone incorporation and histone PTM, have been shown to elicit their effects on intra- and inter-nucleosome compaction and by extension, local genome accessibility, through two interrelated mechanisms: intrinsically, by directly altering charge-dependent contacts non-specifically, or extrinsically, by serving as binding platforms for specific chromatin “readers”. Readers of a given histone modification can include architectural proteins, (co-)transcription factors, or DNA (de)methylases. Significantly, readers can also be other histone-exchanging proteins (i.e. histone chaperones and remodelers) or histones modifying enzymes, and these interactions give rise to interconnected feedback loops that reinforce or inhibit the effects and propagation of specific histone modifications (Lee et al., 2010; Torres and Fujimori, 2015; Venne et al., 2014; Zhang et al., 2015).

Achievement of the nucleosome crystal structure significantly advanced our understanding of the interactions stabilizing the nucleosome particle and how these interactions influence the formation of higher order chromatin structures (Luger et al., 1997). Each core histone contains a three-helix core domain known as a “histone fold” motif that directs heterodimerization of H2A with H2B, and H3 with H4, through an extensive protein-protein interface referred to as a “handshake” arrangement. Dimers of H2A/H2B form a region on the nucleosome surface known as the “acidic patch”, comprised of five amino acids from the docking domain of H2A (Glu 56, Glu 61, Glu 64, Asp 90, Glu 91 and Glu 92) and one residue from H2B (Glu 110). *In vitro*, the acidic patch is integral for the compaction of nucleosome arrays through contacts with the basic patch on the H4 tails of neighbouring nucleosomes, and it also mediates interactions with chromatin binding proteins *in vivo* (Fan et al., 2004; Kalashnikova Anna A. et al., 2013; Luger and Richmond, 1998). The H2A-H2B and H3-H4 dimers further associate with each other largely through 4-helical bundles such that an H3-H4 tetramer is formed through an H3:H3 interface and is flanked by H2A-H2B dimers that associate through weaker binding of H2B with H4 (Alberts et al., 2002) (Fig. 1-3). As a result of this structural organization, the H3-H4 tetramer forms a stable core, whereas the two flanking H2A-H2B dimers are more readily displaced (Kulaeva et al., 2010). Histones are highly basic proteins and the histone octamer forms 6 distinct primary contacts with DNA, driven by electrostatic and hydrophobic interactions between the DNA phosphate backbone and the central H3-H4 tetramer (Luger et al., 1997). In addition to these structured regions, 25-30% of the mass of core histones is contained within their unstructured, intrinsically disordered “tail” domains, found at the N-terminal

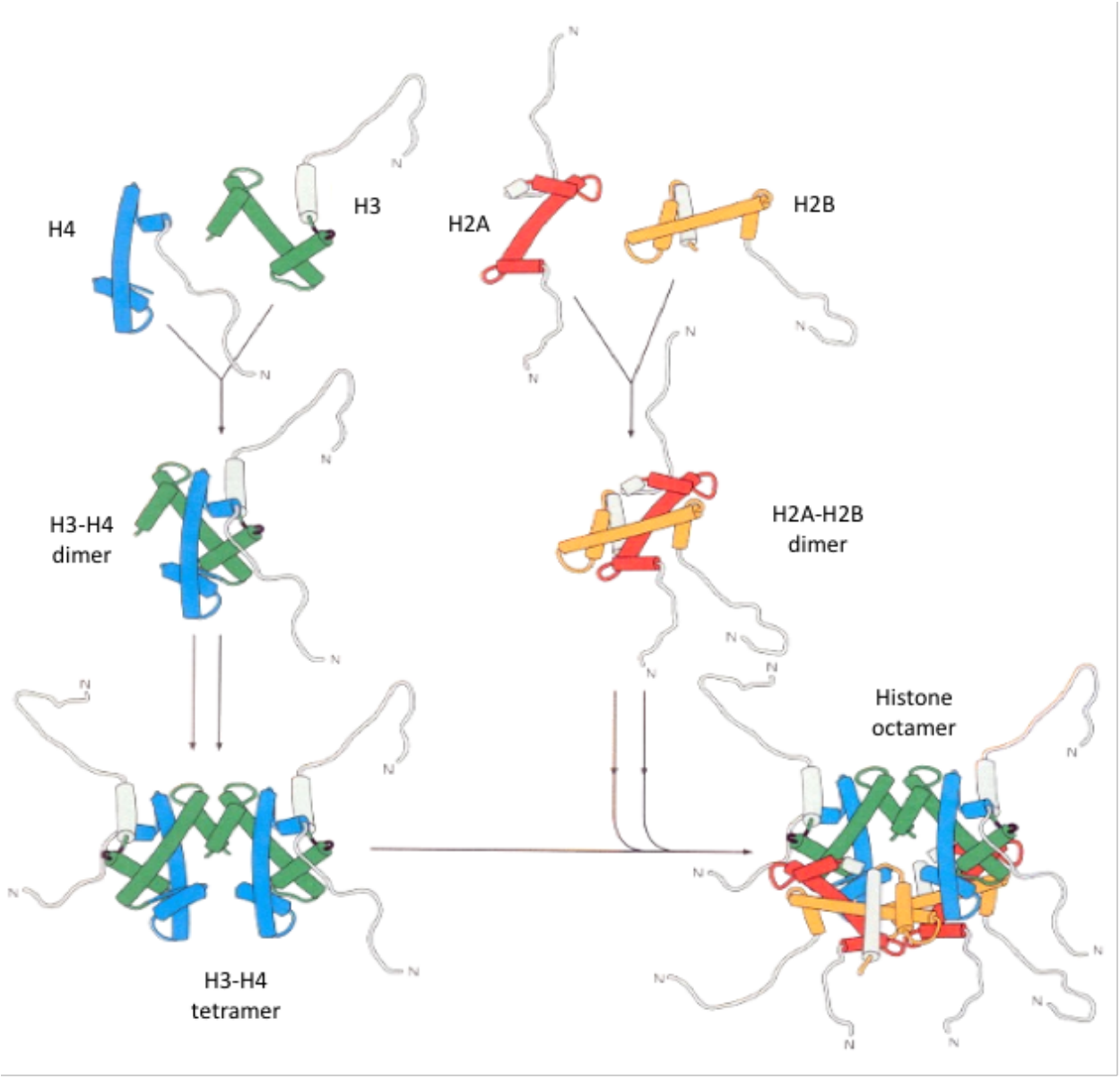


Figure 1-3. Histone-histone interactions that maintain nucleosome structure. Each histone comprises a 'histone fold' domain formed by three alpha-helices connected by loops. H3-H4 dimers interact through a four-helical bundle formed by an H3-H3 interface to yield the H3-H4 tetramer. Two H2A-H2B dimers then interact with the tetramer through a four-helix bundle formed by H2B and H4. This arrangement allows for exchange of H2A-H2B dimers without disrupting the H3-H4 tetramer, whereas H3-H4 replacement requires the removal of H2A-H2B (adapted from Albert et. al 2001).

regions of all four core histones, as well as at the C-terminus of H2A. Histone tails contribute very little to thermostability of the nucleosome core particle, and their primary role instead appears to be in mediating inter-nucleosome interactions. Indeed, experiments *in vitro* have demonstrated that tail domains are essential for compaction of nucleosome arrays into 30 nm fibers (Luger and Richmond, 1998). Histone tails interact with DNA weakly, and also with other intra- and inter-nucleosomal histones, and this binding can be finetuned through their substitution by histone variants or by the enzymatic modification of histones by PTM (Zhou et al., 2019). Binding of reader proteins to histone modifications also influences the chemical composition of nucleosomes, which also contributes to the stability nucleosome stability.

Upwards of 20 types of modifications occur at more than 200 sites of PTM within the tails and lateral surface of canonical and linker histones and their 30 histone variants (Huang et al., 2015; Zhao and Garcia, 2015). Altogether, the expansive combinatorial arrangements of PTMs on different histones impart unique “nucleosome signatures” of distinct chemical compositions that help establish functional regions of the genome through nucleosome specialization.

1.1.2.1 Histone post-translational modification

Histones are post-translationally modified by the covalent, reversible addition of chemical groups to specific residues. These modifications include acetylation, phosphorylation, methylation, ubiquitylation, sumoylation, ADP-ribosylation, and deamination, as well as lesser-abundant, more recently discovered PTMs such as

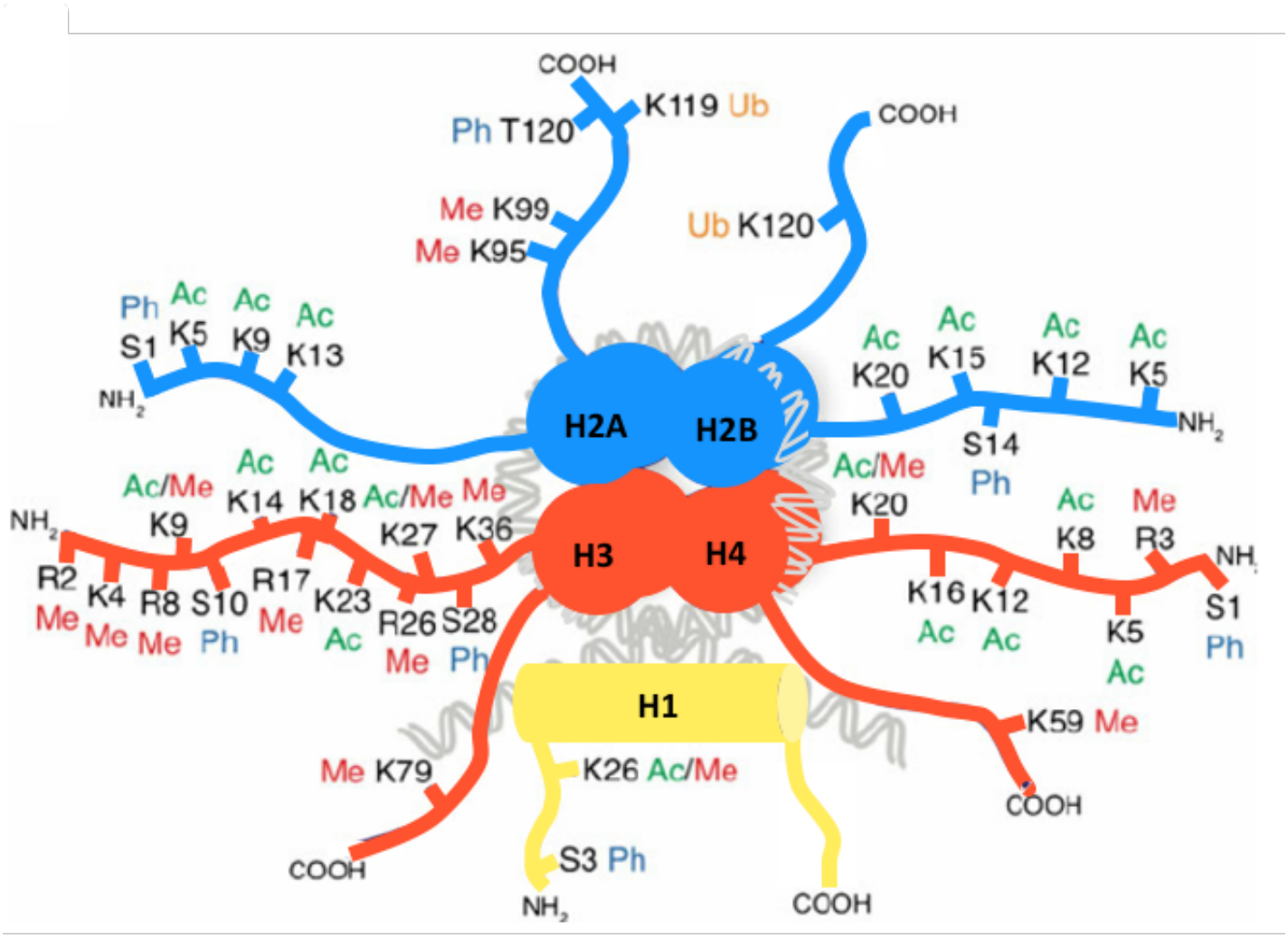


Figure 1-4. Histone tail post-translational modifications. Common PTMs occurring on the N- and C-terminal tails of the four core histones (H3, H4, H2A, H2B) and linker histone H1. Me=methylation; Ac= acetylation; Ub=ubiquitination; Ph=phosphorylation (adapted from Tollervey and Lunyak 2012).

butyrylation, propionylation, crotonylation, and serotonylation (Farrelly et al., 2019; Zhao and Garcia, 2015). Histone PTMs are mainly found within the intrinsically disordered N- and C-terminal tails that protrude from the core particle, and to a lesser extent, on the structured (globular) surface of the core particle itself (Tollervey and Lunyak, 2012; Zhao and Garcia, 2015) (Fig. 1-4).

In general, active chromatin domains are characterized by distinct compositions of histone PTMs. For example, acetylation of histone H3 at lysine 27 (H3K27ac) and monomethylation of histone H3 at lysine 4 (H3K4me1) are associated with active enhancers, and H3K4me3 as well as H3 and H4 acetylation levels at promoters strongly correlate with their transcriptional activity. Ubiquitylation of histone H2B at lysine 120 (H2BK120ub1), H3K79me3, and H3K36me3 are also linked to active transcription. In contrast, H3K27me3 and H2AK119ub1, as well as H3K9me3 are associated with transcriptional repression (Zhang et al., 2015). Although combinations of histone modifications associate with distinct chromatin states, the interplay between PTMs and PTM readers poses a challenge to assigning strict causality. The functional and biological significance of histone PTMs, especially in the context of epigenetic regulation, have been a matter of debate in the past decade. If causally linked, histone modifications may be deposited first to regulate transcription or alternatively, they can be deposited as a consequence of transcriptional activity. It is also possible that both *de novo* histone modifications and gene activity are the consequence of sequence-specific transcription factors, and that once established, the primary function of histone modifications is to act as a form of cellular memory within larger regulatory systems (Henikoff and Greally, 2016).

1.1.2.1.1.1 Intrinsic and extrinsic dynamics

Histone PTMs can have direct or intrinsic effects on nucleosome dynamics. This typically occurs when they are deposited at key residues residing at the interface of histone-histone intra- or inter-nucleosome contacts, or histone-DNA contacts. At the same time, histone PTMs (including those having an intrinsic effect) can act as signals for the recruitment of chromatin binding proteins which in turn mediate or elicit chromatin function.

When first discovered, the mechanism by which histone PTMs influence chromatin dynamics was thought to occur through their disruption of charge-dependent contacts between histones and DNA (Allfrey et al., 1964). Incidentally, a large proportion of PTMs occurring on the globular domains of the core particle, including on the lateral surface of octamer (which is in direct contact with DNA) are lysine acetylation, and serine or threonine phosphorylation events - PTMs that alter electrostatic potential of the nucleosome surface by neutralizing a charge or adding a negative charge, respectively. Examples of PTMs occurring on the globular domains are limited however, and most documented histone PTMs are observed on the flexible N- and C-terminal tail domains (Lawrence et al., 2016; Zhao and Garcia, 2015). Post-translational modification within tail domains can also have direct effects on chromatin structure, and are important determinants of inter-nucleosome interactions. For example, histone hyperacetylation, which is associated with open, active euchromatin regions, only modestly affects the stability of individual nucleosomes, while a single acetylation event at lysine 16 within the tail of H4 (H4K16ac) dramatically reduces the

propensity of oligonucleosome arrays to self-associate *in vitro* (Shogren-Knaak et al., 2006). This effect of H4K16ac has been attributed to its disruption of a key interaction between H4K16 on the tail of one nucleosome, with the acidic patch on an adjacent nucleosome. Notably, such interaction provides an example whereby intrinsic properties of the nucleosome interface with its extrinsic capacities, as PTM status of the H4 tail is in a position to competitively regulate accessibility to the acidic patch by reader binding domains, and vice-versa.

Another type of dynamic interplay between histone modifications is “crosstalk”, which is the bridging of modifications by reader complexes that possess multiple activities and can promote or antagonize a given functional state (Lee et al., 2010; Torres and Fujimori, 2015; Venne et al., 2014; Zhang et al., 2015). For example, one of the first described instances of crosstalk accounts for the duality of effects H3 serine 10 phosphorylation (H3S10p) imparts on chromatin structure - its decompaction during transcriptional activation of immediate early (IE) genes during mitogen stimulation, as well as its condensation during mitosis. During the former event, H3S10p has been shown to promote acetylation of the adjacent K14 and K9 residues on the same histone molecule to activate transcription of IE genes such as *c-fos* and *c-jun* (Cheung et al., 2000; Clayton et al., 2000). At another target gene, *FOSL1*, H3S10p results in the recruitment of the MOF acetyltransferase, which then acetylates H4K16. Within this context, H4K16ac has been shown to recruit Brd4 through its acetyl-binding bromodomain, which then recruits the positive elongation factor b (p-TEFb) to phosphorylate serine 2 of paused RNA polymerase II (Zippo et al., 2009). In the mitosis context, H3S10p occurs concomitantly with trimethylation of K9 on H3 (H3K9me3)

during M-phase, where it has been shown to repel the binding of HP1 (heterochromatin binding protein 1) to H3K9me3 (presumably facilitating access to chromatin by factors involved in condensation) (Fischle et al., 2005). Histone PTM crosstalk is also proposed to act during *de novo* deposition and maintenance of H3K4me3, a PTM predominately associated with genes and strongly correlated with their active transcription (Santos-Rosa et al., 2002; Strahl et al., 1999). H3K4me3 levels peak around the TSS of genes and this striking feature, conserved from yeast to humans, correlates with both antisense and sense transcription (Bornelöv et al., 2015). In mammals, H3K4me3 can be targeted to nonmethylated CpG islands by a zinc-finger-CXXC domain in the H3K4-methylases MLL1 or MLL2. MLL complexes also contain PHD finger domains, which may enable it to bind its own H3K4me3 mark and positively reinforce it by directing H3K4-methylase activity towards adjacent nucleosomes (Shi et al., 2007; Wang et al., 2010). H3K4me3 can also be maintained at promoters by monoubiquitylation of H2BK120 (H2BK120Ub1), through an activity-dependent crosstalk pathway purportedly conserved from yeast to humans. This pathway was first discovered in *S. cerevisiae*, where loss of H2Bub1 (either through deletion of Rad6 ubiquitin-conjugating enzyme, or mutation of the H2B ubiquitylation site, K123) resulted in genome-wide loss of H3K4-methylation (Sun and Allis, 2002). In mammalian cells, H2BUb1 is catalyzed by the RNF20 and RNF40 ubiquitin ligases which associate with elongating RNAPII and are further activated by additional co-transcription factors (Osley, 2006). Significantly, H2BUb1 has been shown to allosterically stimulate activity of the MLL complexes by binding to its core subunit, ASH2L (Wu et al., 2013; Zhang et al., 2015).

The importance of histone PTM activity as a function of its overall chromatin environment – both within its cognate nucleosome and adjacent nucleosomes – is further demonstrated by evidence suggesting that epitope recognition by a single reader can occur across multiple histones and engage or sense multiple surfaces of the nucleosome (Ng and Cheung, 2015; Su and Denu, 2016). Examples of individual PTM-binding domains include PHD fingers that bind methylated lysines, bromodomains that bind acetylated lysines, PWWP repeats, HEAT, WD40, MBT, ankyrin, and chromodomains that bind methylated lysines/arginines, and 14-3-3, BIR, and BRCT domains that bind phosphorylated serine/threonines (Yun et al., 2011). One of the first examples of such multivalent engagement at the nucleosome level is the binding of BPTF, a subunit of the Nucleosome Remodeling Factor (NURF), to active chromatin through adjacent PHD finger and bromodomain modules. The PHD finger of BPTF was first identified as a motif that specifically engages H3K4me3 (Wysocka et al., 2006). It was noted that a bromodomain is located adjacent to this PHD finger, and this tandem arrangement, termed the PHD-Bromo cassette, was predicted to bind combinatorial methyl/acetyl marks on histones. Subsequent characterization of the BPTF bromodomain by *in vitro* peptide binding found that it prefers to bind H4K12ac, H4K16ac, and H4K20ac modified peptides with comparable affinities (Ruthenburg et al., 2011). Strikingly, when Ruthenburg et al. tested the combinatorial PTM binding preference of the BFTF PHD-Bromo cassette with semisynthetic nucleosomes bearing H3K4me3, and H4 acetylated at K12, K16, or K20, they found enhanced binding only with nucleosomes containing the H3K4me3-H4K16ac combination, but not when H3K4me3 is paired with other acetylation sites (Ruthenburg et al., 2011). These findings

highlight the importance of context (e.g. nucleosome versus peptide) that determine binding of PTM-binding domains to their cognate epitopes. Another example of multivalent binding at the nucleosome level is seen in the catalytic subunit of the NuRD complex, CHD4, which comprises tandem PHD fingers connected by a short linker (Musselman et al., 2012). These PHD fingers have been shown to mediate concomitant engagement of two H3 tails within a single reconstituted tetrasome (H3/H4 tetramer wrapped by 80 bp of DNA), which was used as a nucleosome substitute. Binding of these tandem PHD fingers was further shown to be enhanced by H3K9ac or H3K9me3, and weakened by H3K4me, as determined by NMR and pull-down approaches, suggesting that Lys9 hydrophobicity is a determinant of CHD4 PHD module association whereas successive methylation of Lys4 has an abrogative effect (Musselman et al., 2012). A similar mode of multivalent recognition was also demonstrated for the tandem PHD finger protein CHD5, which has been shown to simultaneously engage two unmodified N-terminal H3 tails, and several PTMs within this region have been found to disrupt this high–affinity binding (Oliver et al., 2012). Finally, the PHD-bromodomain protein p300 has been identified as a multivalent reader *in vitro*. Using a DNA-barcode nucleosome library and a streamlined method for producing semisynthetic modified nucleosomes, the Muir lab found that concomitant hyperacetylation of nucleosomal H3 and H4 resulted in a dramatically increased binding affinity of p300, albeit through hitherto undefined contacts (Nguyen et al., 2014).

1.1.2.1.2 Polycomb silenced chromatin

Polycomb silencing has long been considered a paradigmatic model for the epigenetic maintenance of gene transcription programs (Aranda et al., 2015). Links between epigenetic-mediated repression and chromatin structure began 70 years ago with identification of *Polycomb* (*Pc*) in *Drosophila melanogaster*. The eponymous Polycomb gene was named after its striking dominant phenotype – the presence of multiple sex combs on all three pairs of legs in *Pc* mutant flies, while normally appearing exclusively on the most posterior pair. It was later surmised that transformation of embryonic body segments to resemble more posterior ones in response to *Pc* mutation was caused by the ectopic expression of homeotic (Hox) genes, which are master regulators of biaxial body pattern (Jürgens, 1985; Lewis, 2004). Genetic screens subsequently identified additional proteins whose loss-of-function mutations resulted in a similar phenotype, and are collectively referred to as Polycomb group (PcG) proteins (Ringrose and Paro, 2004). Shortly after identifying PcG proteins, the first of several so-called Trithorax Group (TrxG) genes was discovered and found to antagonize activity of PcG, causing posterior body segments to display anterior traits upon mutation (e.g. loss of sex combs) (Klymenko and Müller, 2004; Poux et al., 2002). The additional early observation that PcG and TrxG proteins maintain Hox gene expression patterns after germline transcription factors that established their expression have long been diluted from the embryo led to hypothesize that antagonistic PcG and TrxG activities act as long-term cellular memory systems (Ringrose and Paro, 2004). Indeed, research has since extended these discoveries in mammals and established PcG and TrxG proteins as key regulators of developmental processes, including X chromosome inactivation,

genomic imprinting, cell-cycle control, stem cell plasticity, and dynamic response to developmental and environmental cues (Aranda et al., 2015; Bajusz et al., 2018).

Traditionally, PcG proteins are primarily divided amongst two large complexes: PRC1 and PRC2 (Fig. 1-5). PRC2 consists of three components: EZH1/2, SUZ12, and EED. It catalyzes trimethylation of H3K27 (H3K27me3) through its SET domain-containing methyltransferase EZH1/2, and this complex can also subsequently bind H3K27me3 through a WD40-repeat domain contained within its EED subunit. Binding of PRC2 to methylated H3 allosterically enhances activity of the SET domain of EZH1/2 and results in the spreading of H3K27me3 (Aranda et al., 2015). Research in the past decade has recognised that PRC1 is further divided into canonical (cPRC1) and non-canonical/variant (ncPRC1) complexes, and the latter is thought to have appeared earlier in evolution (Bajusz et al., 2018; Gao et al., 2012; Tavares et al., 2012). Both mammalian PRC1 complexes share a core comprising RING1 proteins (RING1A or RING1B), which possess ubiquitin ligase activities that catalyze monoubiquitylation of H2A at lysine K119 (H2AUb1), as well as one of six Polycomb group ring-finger domains (PCGF1-6). Canonical PRC1 complexes are only assembled around PCGF2 or PCGF4 and in contrast to ncPRC1, contain a chromobox protein (CBX2, CBX4 or CBX6/8) that binds to H3K27me3. Canonical PRC1 complexes are further typified by the presence of a Polyhomeotic homologous protein domain (PHC1 through 3) comprising a sterile alpha (SAM) motif. Non-canonical PRC1, in contrast, can assemble with PCGF1 through PCGF6 to form the core of six discrete complexes respectively named ncPRC1.1 – ncPCR1.6, and these are further characterized by their complement of co-purifying ancillary factors (Bajusz et al., 2018; Di Croce and Helin, 2013; Junco et

al., 2013). Non-canonical PRC1 complexes lack a chromodomain and instead contain an RYBP (zinc-finger domain and YY1 binding protein) domain or its paralog YAF2, which are enhancers RING1 ubiquitin ligase activity of both *in vitro* and *in vivo* (Rose et al., 2016).

Many studies have demonstrated that in differentiated cells, absence of PRC1 and PRC2 interplay can result in the erosion of repressive Polycomb domains rendering target genes susceptible to inappropriate expression signals (Mills, 2010; Sashida and Iwama, 2017). Silencing by PcG must therefore be dynamically responsive yet robust. The relationships between cPRC1, ncPRC1 and PRC2 are intricate, and mechanistically, how these complexes are targeted to chromatin and how they orchestrate gene repression programs, is incompletely understood.

One source of elusiveness comes from the fact that much of what we know about Polycomb silencing derives from studies of PRC2 and cPRC1, the functional homologue of *Drosophila,* and this has led to a simplistic hierarchal recruitment model in which PRC2-mediated H3K27me3 recruits PRC1 through its chromodomain, and PRC1-catalyzed monoubiquitylation of H2A ultimately inhibits FACT (Facilitates Chromatin Transcription) through an unknown mechanism (Aranda et al., 2015). This model, however, does not explain how PRC2 itself is targeted to chromatin, nor does it account for sites of PRC1-binding that lack H3K27me3. Unidirectional recruitment of PcG complexes is also inconsistent with observations that H2AUb1 can also localize PRC2 activity to a subset of genes. Studies in *Drosophila* have established that Hox gene silencing is accomplished through long-range looping events that link *cis* Polycomb response elements (PREs), and have demonstrated that exclusion of Hox genes from

these topological assemblies leads ectopic transcription and acquisition of H3K4me3 (Montavon et al., 2011; Noordermeer et al., 2011). Although mammals appear to lack consensus PREs, PRC2, and to a lesser extent PRC1, has been shown to localize to unmethylated CpG islands, and cPRC1 can form topological structures through auto-

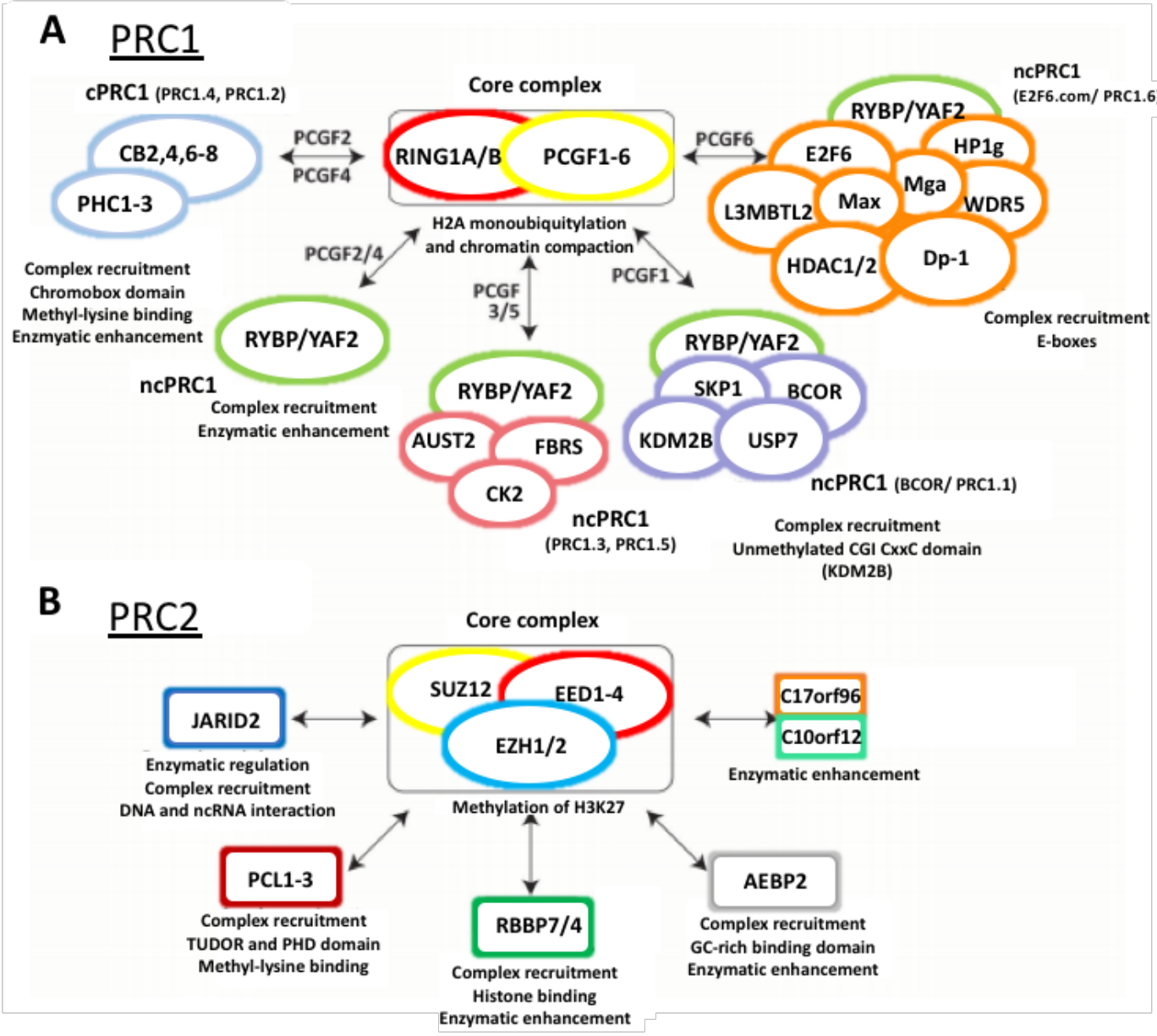


Figure 1-5. Mammalian Polycomb complexes. PcG proteins are divided amongst two main complexes: (A) PRC1 and (B) PRC2. Both complexes are typified by their core components, which are shared by their respective subcomplexes. The differential interactions of core components with accessory proteins further defines the composition of each subcomplex, and can modulate the recruitment or enzymatic activity of the core complex. (A) PRC1 is divided into cPRC1 and ncPRC1. These subcomplexes associate with different Pcgf proteins. cPRC1 complexes include subcomplexes incorporating Pcgf2 or Pcgf4 (cPCR1.2 and cPRC1.4, respectively). Pcgf1 and Pcgf4 can also interact with the ncPRC1-associated Rybp or YAF proteins. Pcgf3 and Pcgf5 are present in ncPRC1.3 and ncPRC1.5, respectively. Pcgf1 is present in ncPRC1.1 (also known as BCOR), and Pcgf6 is present in ncPRC1.6, known also as E2F6.com. (B) The PRC2 core complex can associate with different accessory proteins that interact with PRC2 non-exclusively (adapted from Aranda et. al, 2015).

polymerization of its SAM domains in a RING1B-independent manner (Isono et al., 2013). At the same time, ncPRC1 complexes have been shown to possess greater H2A-monoubiquitylation activity than cPRC1, but do not contain SAM domains, and hence how they mediate Polycomb silencing, in addition to recruiting or retaining of H3K27me3, remains obscure. Nevertheless, enzymatic activity of PRC1 is clearly a key component of PcG silencing at a subset of genes; *PAX3,* for example, is almost completely derepressed in RING1B mutants (Blackledge et al., 2019; Endoh et al., 2012; Stoop et al., 2008). Although the discovery of ncPRC1 in addition to cPRC1 may hint at possible divergent functions and non-redundant targets amongst PRC1 complexes, the activities of both PRC1 complexes could also converge within a more complex system of mutual reinforcement in which the interplay of cPRC1 and ncPRC1 is required for silencing fidelity of the same transcriptional programs.

1.1.2.2 Histone variants

Histone variants further contribute to and expand the intrinsic and extrinsic properties of the nucleosome. These are non-allelic isoforms that differ from canonical histones by one to a few dozen amino acids. Unlike canonical histones, whose expression are restricted to S-phase to package newly synthesized DNA, histone variants are expressed throughout the cell-cycle and are substituted into specific sites of the genome through the activity of histone variant-specific chaperones and shared ATP-driven nucleosome remodeling enzymes (Hamiche and Shuaib, 2012). Functional divergency of histone variants from canonical histones is first apparent by the

organization of their respective genes, as core histones are encoded by multiple, clustered copies of genes, and their mRNAs lack introns and poly-(A) tails. In contrast, variant histones are typically present as single-copy genes, and their mRNAs contain introns that may be alternatively spliced and are modified by poly-(A) tails (Osley, 2006). Although representing a only small proportion of the total cellular histone pool, histone variants are uniquely linked to key cellular and developmental processes. Evolutionarily, H4 and H2B are more constrained, while greater diversity is displayed amongst H3 and H2A (Henikoff 2013). For example, H2A has eight variants - H2A.X, H2A.Z-1, H2A.Z-2, H2A.Z-2.2 (an alternatively spliced form of H2A.Z-2 causing severe nucleosome destabilization and found most abundantly in brain (Bönisch et al., 2012)), H2A Barr body deficient (H2A.Bbd; also known as H2A.B), macroH2A1.1, macroH2A1.2 (which are splice variants), and macroH2A2 – while H3 has six variants – H3.3, histone H3-like centromeric protein A (CENP-A), H3.1T, H3.5, H3.X (also known as H3.Y.2), and H3.Y (also known as H3.Y.1) (Buschbeck and Hake, 2017). Additionally, two testis-specific variants of H2B (H2BFWT or H2B.W, and TSHH2B or H2B type A), and one variant of H4 (H4G) have been identified (Buschbeck and Hake, 2017; Long et al.).

Some histone variants have well-defined effects on chromatin structure. H2A.B for example, has a C-terminus that is 19-amino acids shorter than canonical H2A and this has several effects on nucleosome structure. The portion of the H2A.B C-terminal tail that is absent encompasses the histone “docking domain” present on H2A and other family members that interfaces and interacts with the H3-H4 tetramer, resulting in nucleosome instability and a core particle that protects ~30 bp less DNA from micrococcal nuclease digestion. H2A.B also lacks a surface acidic patch and consistent

with this, arrays of H2A.B-containing nucleosomes cannot form compact chromatin fibers *in vitro* (Bao et al., 2004; Doyen et al., 2006; Zhou et al., 2007). Another H2A variant imparting distinct structural variation is mH2A (formerly macroH2A). mH2A possesses a large, non-histone, 'macro' domain of ~30 kDa at its C-terminal end, connected to its histone fold by an unstructured linker, resulting in a histone approximately three times the size of its canonical counterpart (Karras et al., 2005; Kustatscher et al., 2005). Structurally, incorporation of mH2A stabilizes the mH2A-H2B dimer with the H3-H4 tetramer (Chakravarthy et al., 2005) and consistent with this, mH2A is primarily associated with heterochromatin (Costanzi and Pehrson, 1998; Grigoryev et al., 2004; Zhang et al., 2005). The macro domain of mH2A is related to a family of proteins that includes a class of ADP-ribose processing enzymes and NAD+ metabolite-binding proteins, while the linker region of mH2A has been shown to bind and stabilize the DNA entry/exit site of the nucleosome in a manner reminiscent of histone H1, enhancing compaction of nucleosome arrays and fiber-fiber interactions *in vitro* (Kozlowski et al., 2018). Histone variant H3.3 incorporation also influences chromatin stability, and it differs from canonical H3 by only four amino acids. Three of these amino acids reside within the core particle interior (residues 87-90) and one is solvent-exposed (Ser31). Interestingly, all four residues promote chromatin array de-compaction but have no effect on the stability of mononucleosomes, *in vitro* (Chen et al., 2013). Histone H3.3 has numerous and context-specific functions, and *in vivo,* its observed effects on chromatin compaction are likely linked to its turnover rate through the activity of H3.3-specific histone chaperones, HIRA and ATRX/DAXX. HIRA is linked to deposition of H3.3-H4 tetramers at the promoter and within the body of actively

transcribing genes, which exhibit high nucleosome-turnover, whereas ATRX/DAXX regulates incorporation of H3.3 at telomeres and pericentric regions, where nucleosome turnover is low (Kraushaar et al., 2013; Ricketts and Marmorstein, 2017; Szenker et al., 2011). The differential recognition of histone variants by chaperones and its consequences for the nucleosome assembly pathway is one way in which histone variants can regulate chromatin. Additionally, nucleosome readers may be sensitive to histone variant status. For example lysine 36 methylation in the context of H3.3 (H3.3K36me3) is specifically recognized by the putative tumour suppressor BS69/ZYMND11 through its Bromo-Zinc-PWWP cassette, and this interaction is antagonized by the phosphorylation of H3.3-specific residue Ser31 (Guo et al., 2014; Wen et al., 2014b). Another histone that undergoes variant-specific PTM is H2A.X, which is phosphorylated at Ser139 (known as γ-H2A.X) in response to DNA damage, and this PTM is engaged by MDC1 (mediator of DNA damage checkpoint protein 1) through its BRCT (breast cancer associated carboxy-terminal) domain (Sawicka and Seiser, 2014).

1.1.2.2.1 Histone variant H2A.Z

The histone variant H2A.Z accounts for ~15% of total H2A in mammals. H2A.Z appears to have arisen once in early eukaryotic evolution and shares ~60% sequence identity with canonical H2A (Weber and Henikoff, 2014) (Fig. 1-6). The importance of H2A.Z is demonstrated by its requisite for viability of *Tetrahymena thermophila, D. melanogaster, Xenopus leavis,* and mice (van Daal and Elgin, 1992; Faast et al., 2001; Iouzalen et al., 1996; Liu et al., 1996). H2A.Z is highly conserved and ~90% sequence identity is

shared amongst the protozoan parasite *Plasmodium falciparia, Saccharomyces cerevisiae,* mice, and humans, and is postulated to likewise be an essential factor in human development (Zlatanova and Thakar, 2008). Consistent with this, H2A.Z has been implicated in diverse biological processes including transcription activation and repression, chromosome segregation, heterochromatin silencing and boundary formation, cell cycle progression, and more recently, splicing regulation (Keogh et al., 2006; Krogan et al., 2004; Meneghini et al., 2003; Mizuguchi et al., 2004; Neves et al., 2017; Nissen et al., 2017; Zofall et al., 2009). H2A.Z is non-randomly distributed throughout the genome in both euchromatic and heterochromatic domains; however, the mechanisms by which it is localized by H2A.Z chaperones (which so far include ANP32E in metazoans) and remodeling complexes (SRCAP and p400/Tip60) are poorly understood (Cai et al., 2005; Choi et al., 2009; Ikura et al., 2000; Obri et al., 2014; Ruhl et al., 2006). While the mechanisms underlying the multiplicity of its functions are also incompletely delineated, modifications of H2A.Z and their modularity have been shown to contribute to the diversity of H2A.Z-containing nucleosomes. H2A.Z can be post-translationally modified at specific lysine residues by acetylation, ubiquitylation, sumoylation, and methylation, and in vertebrates, is present as non-allelic paralogs, H2A.Z-1 and H2A.Z-2 (Eirín-López et al., 2009; Sevilla and Binda, 2014). Nucleosomes can also contain one copy of H2A.Z (heterotypic nucleosomes) or two copies (homotypic), and this distinction may also have functional consequences for chromatin.

Homotypic and heterotypic H2A.Z-nucleosomes have been crystallized and while their overall structures are very similar to those containing H2A, several differences can be captured (Horikoshi et al., 2016; Suto et al., 2000). Firstly, substitution of Gln104 in

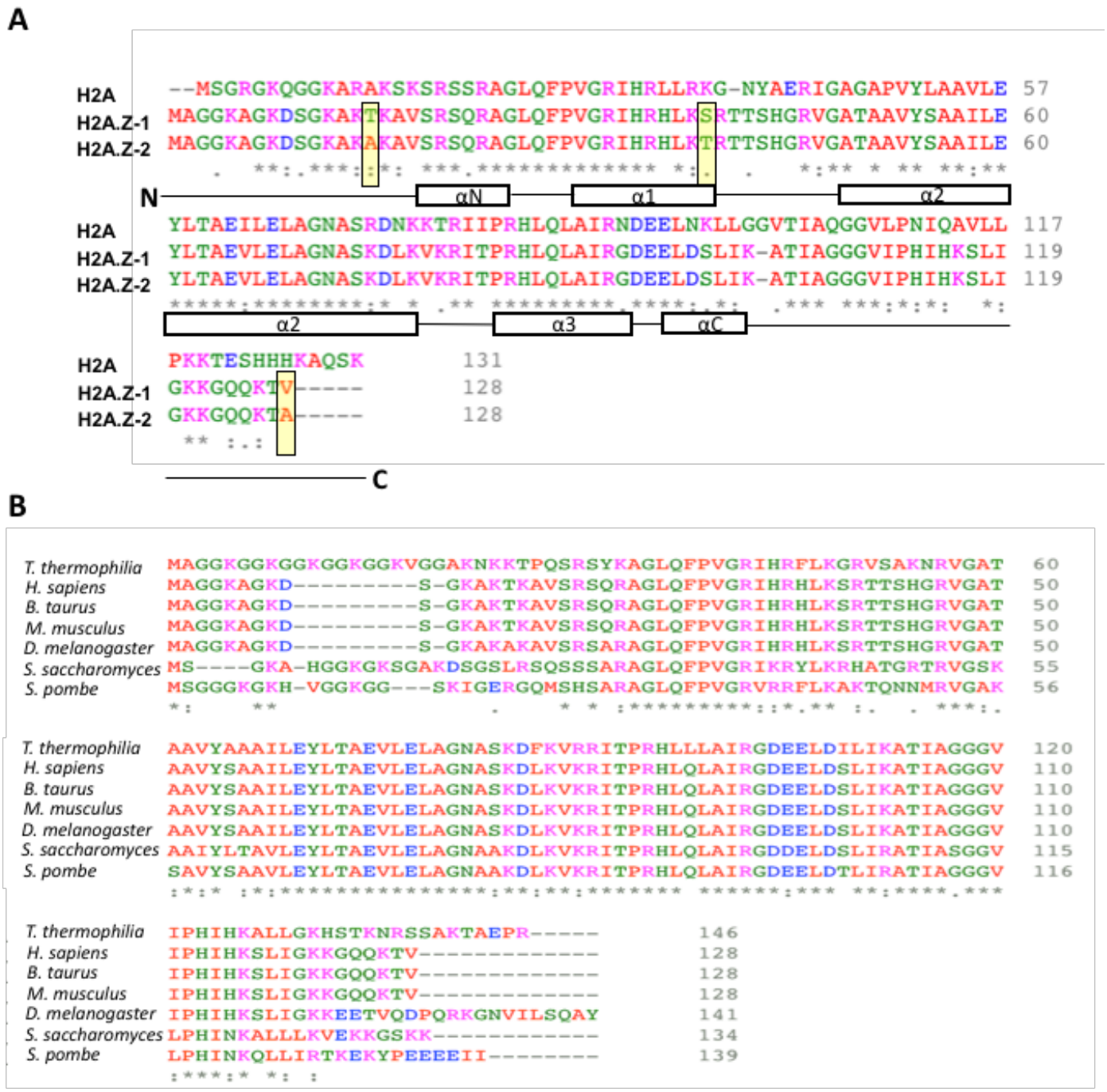


Figure 1-6. Amino acid sequence alignment of H2A.Z. (A) Alignment of human H2A with isoforms H2A.Z-1 and H2A.Z-2 (non-conserved residues between isoforms are highlighted in yellow boxes. White boxes represent alpha helices; the histone fold domain encompasses the region between α1 to α3, inclusive (B) Multiple sequence alignment of H2A.Z(-1) in different organisms.

H2A by Gly106 in H2A.Z results in the loss of 3 hydrogen bonds between H2A.Z and the H3-H4 tetramer and is predicted to result in slight destabilization of the H2A.Z-nucleosome; an expectation that appears to hold *in vivo* for acetylated H2A.Z-nucleosomes or those also paired with H3.3 (Jin and Felsenfeld, 2007; Thambirajah et al., 2006). Secondly, dimers of H2A.Z-H2B display an extended acidic patch by replacement of Asn94 in H2A with an aspartate (H2A.Z Asp97), and H2A Lys95 with a serine (H2A.Z Ser98). The increased negative charge of this functional domain is predicted to provide a unique epitope for a distinct set of chromatin readers and enhance the propensity of H2A.Z-nucleosomes to form compact fibers through its increased affinity with neighboring H4 tails. Interestingly, with respect to the latter observation, arrays of H2A.Z-oligonucleosomes have been found to favor intra-fiber (local) interactions with the H4 tail, over inter-fiber (global) contact *in vitro*, resulting in unique chromatin domains refractory to highly condensed structures, and hence poised for de-compaction (Fan et al., 2002). At the same time, in specific contexts, highly dense intra-molecular folding of H2A.Z-arrays can stimulate binding by HP1α (heterochromatin protein 1α), which further promotes local chromatin compaction and a transcriptionally-refractive state. Indeed, this interaction is thought to contribute to compacted folding of pericentromeric and telomeric regions where H2A.Z and HP1α have been found to co-localize (Fan et al., 2004; Rangasamy et al., 2003). Finally, crystal structures predict increased thermostability of heterotypic versus homotypic H2A.Z-nucleosomes, as the L1 loop of H2A.Z is displaced in homotypic nucleosomes in comparison to those containing only H2A (Horikoshi et al., 2016).

In both yeast and vertebrates, H2A.Z is enriched within gene promoters as well at enhancers and insulators (Barski et al., 2007; Jin et al., 2009; Obri et al., 2014; Raisner et al., 2005; Weber and Henikoff, 2014). These regulatory regions are characterized by an accessible, nucleosome depleted region (NDR) flanked by strongly-positioned nucleosomes (Hughes and Rando, 2014; Lee et al., 2004; Thurman et al., 2012). Across Eukarya, H2A.Z is specifically localized at the promoter within the +1 nucleosome located immediately downstream of the TSS (which is preceded a NDR), and a few positioned nucleosomes further downstream, but is relatively depleted over gene bodies (Barski et al., 2007; Lantermann et al., 2010; Mavrich et al., 2008; Zilberman et al., 2008). The +1 nucleosome plays an important role in impeding RNAPII progression and its barrier can be lowered through incorporation of H2A.Z, and is marked by H3K4me3 during transcription (Bönisch and Hake, 2012; Jin et al., 2009; Weber et al., 2014a). In *S. cerevisiae* and mammals, H2A.Z is also enriched at the -1 nucleosome upstream of the TSS NDR, though this enrichment is not seen in *Drosophila*, *Arabidopsis thaliana* or *Schizosaccharomyces pombe* (Bagchi and Iyer, 2016). H2A.Z is purportedly localized within the well-positioned -1 nucleosome in silenced genes, but is specifically depleted from this nucleosome in an RNAPII-dependent manner upon activation (Schones et al., 2008). The -1 nucleosome is also an important regulator of antisense transcription from bidirectional promoters (Bagchi and Iyer, 2016). It is interesting to note that *S.pombe* cells lacking H2A.Z display increased antisense transcription (Zofall et al., 2009), while alternatively, in *S.cerevisiae,* incorporation of H2A.Z at the 3' end of genes bodies has been found to promote overlapping antisense transcription (Bagchi and Iyer, 2016). Together, these

suggest that altering H2A.Z incorporation at the +1 and −1 nucleosomes may modulate the permissibility of nucleosomes bordering the NDR to RNAPII, and that H2A.Z may be able to inhibit transcription in both directions from the TSS through yet undefined mechanisms, which might include species-specific functions or differential PTM (Bagchi and Iyer, 2016).

H2A.Z can be acetylated at specific lysines in its N-terminus and this form of the variant is involved in transcription activation. In vertebrates, H2A.Z is mainly acetylated at K4, K7, and K11, though it can also be acetylated at K13 and K15. In yeast, acetylated H2A.Z (acH2A.Z) is correlated with transcription activity, where it is found at the promoters of actively transcribing genes, and is required for galactose-dependent gene induction (Halley et al., 2010; Millar et al., 2006). Alternatively, non-acetylated H2A.Z is associated with inducible but silenced genes (Millar et al., 2006), and ectopically expressed unacetylatable H2A.Z is found at heterochromatin boundaries in yeast (Babiarz et al., 2006). However, in the latter juxtaposition, mutant H2A.Z, in contrast to its wild-type counterpart, is unable to prevent the spread of silent heterochromatin marks from telomeres into adjacent euchromatin regions (Babiarz et al., 2006; Meneghini et al., 2003). In chicken erythroblast cells, acH2A.Z is similarly enriched at the 5’ end of transcriptionally active genes, but is depleted in inactive genes (Bruce et al., 2005). More recently, acH2A.Z has also been implicated in myogenesis, as ectopic expression of unacetylatable H2A.Z mutants has been shown to reduce chromatin accessibility at the MyoD promoter, inhibiting MyoD expression and resulting in inhibition of the myogenic differentiation process (Law and Cheung, 2015). Studies in prostate cancer cells have also demonstrated the

presence of acH2A.Z at the promoter of active genes regulated by the androgen receptor and at the prostate specific antigen (PSA) enhancer, prior to rapid cycles of transcription in response to androgen induction (Dryhurst et al., 2012; Valdes-Mora et al., 2012), and is anti-correlated with promoter DNA methylation and H3K27me3 (Valdes-Mora et al., 2012). It is yet unknown how acH2A.Z is mechanistically linked to transcriptional activation; however, studies of *Drosophila* H2A.Z (H2AvD) localization during DNA repair have suggested that this PTM can be coupled with its exchange for the unmodified form of H2A.Z (Kusch et al., 2004). Acetylation of H2A.Z has been proposed to de-stabilize the nucleosome as determined by *in vitro* salt-dependent dissociation assays (Thambirajah et al., 2006), and it is possible that acetylation of H2A.Z is coupled to its removal during active transcription. In an non-mutually exclusive manner, the acetylated tail of H2A.Z may act to bind or repel chromatin reader proteins which then regulate transcription. For example, our lab has shown that the bromodomain containing transcriptional activator Brd2 preferentially associates with H2A.Z-nucleosomes and is required for androgen receptor-mediated gene activation (Draker et al., 2012). Bromodomain-containing proteins recognize acetylated lysine residues and acH2A.Z is present at androgen-dependent promoters (Valdes-Mora et al., 2012). It is thus tempting to speculate that Brd2 is mechanistically linked to functional outcomes through acH2A.Z as well.

Our lab has found that in addition to acetylation, H2A.Z can also be monoubiquitylated at either K120, K121, or K125, and that this PTM is catalyzed by the PRC1 E3 ligase Ring1b (Sarcinella et al., 2007). Work in our lab has previously demonstrated that monoubiquitylated H2A.Z (H2A.ZUb1) associates with the

transcriptionally inactive X-chromosome, and that H2A.Z de-ubiquitylation precedes activation of androgen receptor-mediated genes (Draker et al., 2011). Monoubiquitylated H2A.Z reportedly co-localizes with acH2A.Z at bivalent promoters in mouse embryonic stem cells (mESCs), which are characterized by nucleosomes modified with both 'repressive' H3K27me3 and 'active' H3K4me3 PTMs (Ku et al., 2012). Genes marked by bivalent promoters encode the majority of developmental regulators in mESCs and are transcriptionally silent, but poised for rapid activation in response to developmental cues (Ku et al., 2012). Recently, H2A.ZUb1 has been implicated as an important repressor of bivalent genes, akin to previous reports of H2AK119Ub1 in mESCs (Endoh et al., 2012; Surface et al., 2016). In this context, H2A.ZUb1 has been shown to antagonize the binding of Brd2, and absence of H2A.ZUb1 leads to faulty lineage commitment (Surface et al., 2016).

Finally, vertebrates possess two non-allelic isoforms of H2A.Z, H2A.Z-1 (*H2AFZ*) and H2A.Z-2 (*H2AFV*), which differ by three amino acids: residues 14 (Thr in H2A.Z-1 and Ala in H2A.Z-2), 38 (Ser in H2A.Z-1 and Thr in H2A.Z-2), and 127 (Val in H2A.Z-1 and Ala in H2A.Z-2) (Eirín-López et al., 2009). Residues 14 and 127 are located within the unstructured N- and C-terminal tails, respectively, while residue 38 is located within the histone-fold domain. Most previous studies involving H2A.Z have not distinguished between these isoforms; however, several studies have suggested that they could possess distinct functions in transcriptional regulation and modulate non-overlapping sets of genes (Dunn et al., 2017; Faast et al., 2001; Matsuda et al., 2010; Vardabasso et al., 2015). *In vitro*, salt-dissociation assays reveal that nucleosomes containing H2A.Z-1 or H2A.Z-2 do not differ in their stability or overall

nucleosome structure, while FRAP studies indicate that H2A.Z-1 is exchanged more rapidly than H2A.Z-2 (Horikoshi et al., 2016).

1.2 Alternative splicing

Introns are removed from mRNA precursors (pre-mRNA) and exons are ligated to form mature RNA through a process called splicing. Most Pre-mRNA splicing takes place within the major spliceosome, a large complex comprising five ribonucleoproteins (RNPs) containing the small nuclear RNAs (snRNAs) U1, U2, U4, U5, and U6, and as many as 150 other proteins. The spliceosome is assembled *de novo* for each splicing event, and recognizes exons and introns through multiple *cis*-acting signals, which promote networks of interactions that result in exon definition and intron definition, respectively. Four core splice signals demarcate exon-intron boundaries: the 5' and 3' splice sites (5'SS and 3'SS), which are upstream and downstream exon-intron junctions, respectively, the branchpoint site, and the polypyrimidine tract located upstream of the 3'SS (Fig. 1-6A). Human genes contain an average of 10 introns and nearly all transcripts are subject to alternative splicing (AS) in one or more cell types, providing a major source of transcriptomic and proteomic diversity (Pan et al., 2008; Wang et al., 2008). AS is the process whereby one gene produces a variety of isoforms through differential selection of splice sites (Fig. 1-7). Cassette exon skipping occurs when an intervening exon between two exons can be either included or skipped depending on context (e.g. spatiotemporally). Alternatively, exons that are always included are termed constitutive exons (Cui et al., 2017). In order to achieve AS, the splicing machinery must

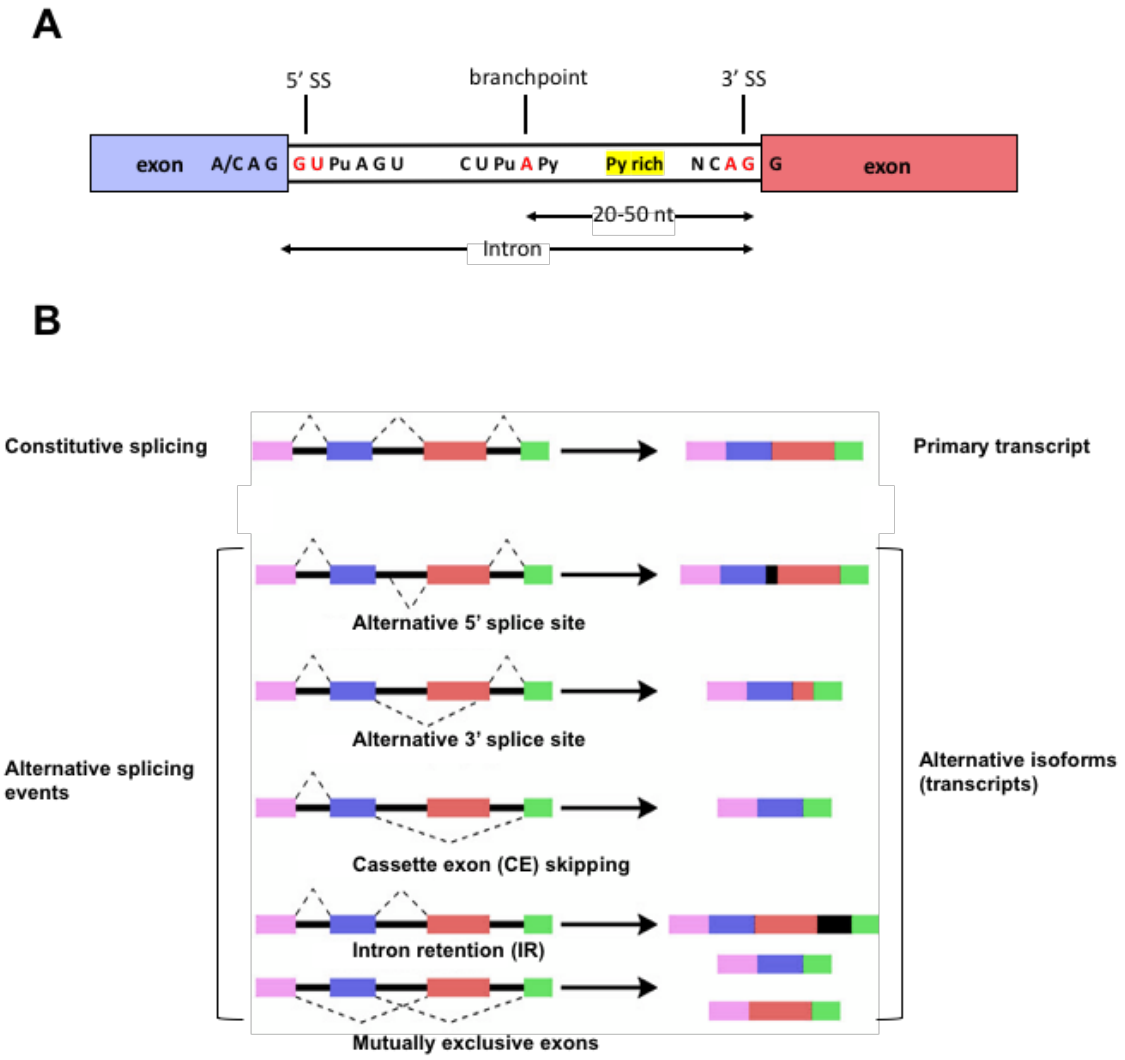


Figure 1-7. Alternative splicing events. (A) Most metazoan introns begin with **GU** and end in **AG** (in the 5' to 3' direction); these are referred to as splice donor (5'SS) and splice acceptor (3'SS), respectively. These sequences alone are not sufficient to signal presence of an intron. Another important signal is the branchpoint site located 20-50 nt upstream of the 3'SS. The consensus branchpoint sequence is CU(A/G)**A**(C/U) amongst higher eukaryotes. Pu = A or G; Py = C or U. (B) Several types of alternative splicing events are shown. Coloured rectangles represent exons whereas black lines denote introns. Connected dashed lines indicate splice site usage.

discriminate between splice- sites in a context-dependent manner by integrating multiple *cis* and *trans* signals in addition to those of the basal core. The main types of AS events include exon skipping (cassette exons), where an exon is spliced in or out of the transcript; alternative 5'SS and 3'SS usage, which involves the recognition of one or more adjacent splice-sites in introns or exons; and intron retention, where an intron may be retained in otherwise mature mRNA (Iñiguez and Hernández, 2017). In vertebrates and invertebrates, exon skipping is the most prevalent form of AS, accounting for ~30-40% of all events. Recent work has shown that intron retention is also more widespread than previously thought, affecting transcripts from approximately two-thirds of human genes (Braunschweig et al., 2014).

One way AS is accomplished is through the action of *cis* regulatory sequences within the exons and introns of pre-mRNA referred to as enhancers and silencers. These recruit positive- or negative-acting splicing factors which then facilitate or inhibit assembly of the spliceosome at proximal splice-sites. These factors include the SR family of proteins, which contain one or two RNA recognition motifs (RRMs) and a C-terminal RS-domain that is rich in alternating Arg and Ser residues. SR proteins typically promote recruitment of multiple factors throughout the spliceosome assembly pathway by recognizing exonic splicing enhancers (ESEs) and are required for formation of the catalytically competent core of the spliceosome. Members of the heterogenous nuclear ribonucleoprotein (hnRNP) family, in contrast, are structurally diverse and often antagonize SR protein activity, and include factors that bind exonic splicing silencers (ESSs), such as hnRNPA1, and PTBP1 or hnRNP I, which often binds to splicing silencers within introns (ISSs).

Splicing and transcription are intimately coupled both spatially and temporally, with introns often being removed as soon as they emerge from RNAPII. RNAPII can regulate splicing through two mechanisms. First, the C-terminal domain (CTD) of RNAPII, which consists of 52 YSPTSPS heptad repeats, can be phosphorylated at two sites (Ser2 and Ser5) in coordination with the transcription cycle and can act as a "landing pad" for recruitment of splicing factors, in differentially phosphorylated states. Secondly, RNAPII elongation rate influences the proportion of transcript available for spliceosome recognition, such that in general, faster elongation rates expose longer stretches of nascent mRNA and can favour the use of stronger, more distal splice-sites. Accordingly, slower elongation limits the RNA sequence presented and can favour usage of weaker, proximal splice-sites, resulting in the splicing of exons with suboptimal splice sites. Elongation rate can also affect the way pre-mRNA folds, and these structures are additional important determinants of spliceosome activity.

1.2.1 Connections between chromatin and alternative splicing

Chromatin can regulate splicing decisions by modulating both the rate of RNAPII and the recruitment of splicing factors. For example, nucleosomes are enriched on GC-rich exons in comparison to introns, and can reduce RNAPII progression. While *in vitro,* the rate of RNAPII elongation is unaffected by the presence of nucleosomes, *in vivo,* RNAPII has been shown to pause preferentially at the DNA entry point and 45 bases into the nucleosome, where DNA contacts the H3/H4 tetramer. Although the signals that elicit RNAPII pausing have not been definitely established, its widespread occurrence is supported by ChIP-Seq and NET-Seq (native elongating transcript sequencing) data

demonstrating that RNAPII occupancy is greater and elongation rate is slower over exons than in introns (Mylonas and Tessarz, 2019; Saldi et al., 2016). Histone chaperones and remodelers that co-transcriptionally target H2A/H2B dimers influence the progression of RNAPII by promoting the assembly or disassembly of nucleosomes, and may indirectly modulate pre-mRNA processing through the deposition of histone variants which can further influence splicing outcomes in complex ways (Venkatesh and Workman, 2015). In yeast, for example, H2A.Z has recently been shown to both stimulate elongation rate and favour the usage of weak splice-sites that result in intron retention (Neves et al., 2017; Nissen et al., 2017).

In addition to their potential effects on RNAPII kinetics, histone PTMs have been shown to regulate AS by recruiting splicing factors through chromatin binding intermediary or ‘adaptor’ proteins. Exons and introns differ in their profiles of histone PTMs, though in most cases, their effects on splicing have not yet been characterized. In general, H3K27me1/2/3, H3K36me3, H3K79me1, H4K20me1, and H2BK5me1 are enriched on exons, while H3K79me1/2, H2BK5me1, H3K4me1/2, H3K9me1, H3K23ac, and H2BUb1 are relatively enriched within introns (Kolasinska-Zwierz et al., 2009; Saldi et al., 2016; Spies et al., 2009). Of exonic PTMs, H3K36me3 is the most enriched and is present in actively transcribed gene bodies and over exons compared to flanking intronic sequence (Spies et al., 2009). H3K36me2/3 has been shown to recruit MRG15, a multifunctional chromodomain-containing protein that is a component of several histone-modifying complexes (Luco et al., 2010). In turn, MRG15 is thought to recruit the negative splicing factor PTBP1 to ISS elements to suppress exon inclusion, particularly when the ISS elements are suboptimal for PTBP1 binding (Luco et al.,

2010). This interaction has been demonstrated at the *FGFR2* gene, which is alternatively spliced into two mutually-exclusive and tissue-specific isoforms containing either exon IIIb or exon IIIc. In human mesenchymal stem cells, exon IIIb is enriched with H3K36me2/3 and its inclusion is suppressed in favour of exon IIIc. Alternatively, exon IIIb is included in epithelial PNT2 cells where *FGFR2* contains lower levels of H3K36me2/3 (Luco et al., 2010). Another liaison between H3K36me3 and AS involves Psip1, which, through its PWWP domain, recruits SRSF1 to modulate exon inclusion or inclusion (Pradeepa et al., 2012). Trimethylation of K36 has also been linked to intron retention events within the context of H3.3, where it interacts with BS69/ZMYND11. BS69 binds directly to the U5 snRNP component EFTUD2 and thus possibly promotes IR by inhibiting formation of the active spliceosome (Guo et al., 2014). Another PTM involved in splicing regulation is H3K4me3, which can recruit the U2 snRNP subcomplex Sf3a through its interaction with CHD1 (Sims et al., 2007). Additionally, unmodified H3, H3K9ac, H3K9me, and H3K14ac have been shown to be important for tethering the SR proteins SRSF3 and ASF2/SF2 to interphase chromatin, and they are released from chromatin in response to H3S10 hyperphosphorylation during mitosis (Loomis et al., 2009). Finally, the histone variant H2A.B has recently been implicated in AS, and can bind both splicing factors and RNA (Soboleva et al., 2017). It has been proposed that H2A.B is able to sequester splicing factors that are competitively released in favour of nascent pre-mRNA which it then anchors, promoting spliceosome interaction and activity resulting in exon inclusion (Soboleva et al., 2017).

1.3 Overview of thesis goals

A comprehensive description of H2A.Z requires an understanding of the activities different forms of H2A.Z carry-out within the nucleosome context. To this end, one objective of this thesis is to provide insight into the nucleosomal context and genome-wide localization of H2A.Z-mononucleosomes post-translationally modified with monoubiquitin (described in detail in Chapter 3 of this thesis). Previously, we used a proteomics approach to identify proteins that selectively engage H2A.Z-nucleosomes and have identified splicing factors amongst the most abundant H2A.Z-enriched proteins. Therefore, a second aim of this thesis is to interrogate a potentially isoform-specific role for H2A.Z in alternative splicing. In particular, we test a possible the link between H2A.Z-1 and the U4/U6.U5 tri-snRNP component USP39 (described in Chapter 4 of this thesis). Completion of these objectives not only involved development of new experimental approaches and methods, but also led to new and unexpected discoveries that raise further questions and ideas for future studies.

Chapter 2

Characterization of Mononucleosomes Enriched for Monoubiquitylated H2A.Z

2 Characterization of mononucleosomes enriched for monoubiquitylated H2A.Z

2.1 Introduction

Histone H2A.Z is a highly conserved histone variant that replaces canonical histone H2A at specific parts of the genome to regulate diverse nuclear processes. Many functional roles have been ascribed to H2A.Z, including transcriptional regulation , heterochromatin boundary formation (Meneghini et al., 2003), DNA repair (Kalocsay et al., 2009), maintenance of chromosome stability and segregation (Ahmed et al., 2007; Hou et al., 2010; Rangasamy et al., 2004), and more recently, the efficient splicing of pre-mRNA (Neves et al., 2017; Nissen et al., 2017). As suggested by the enrichment of H2A.Z at the promoters of most genes in yeast and higher eukaryotes, transcriptional regulation may be a key activity of this variant (Barski et al., 2007; Raisner et al., 2005). However, the precise function of H2A.Z in this process is still controversial, as nucleosomes containing H2A.Z are reported to have roles in both gene activation and repression (Guillemette and Gaudreau, 2006). Moreover, incorporation of this histone variant into chromatin also appears to have disparate structural effects, and studies have reported either a stabilizing or de-stabilizing effect on nucleosomes (Abbott et al., 2001; Suto et al., 2000; Thambirajah et al., 2006). These discrepancies may be in part explained by differential histone post-translational modifications, since H2A.Z is amenable to both acetylation and monoubiquitylation in mammalian cells. For example, H2A.Z can be multiply acetylated at lysine residues proximal to its N-terminus (K4, K7, K11) (Beck et al., 2006; Bonenfant et al., 2007), and such modified forms of the variant is associated with the promoters of actively transcribing genes (Bruce et al., 2005; Millar

et al., 2006). In contrast, we and others have found that a fraction of H2A.Z is monoubiquitylated at its C-terminus, which predominantly occurs on K120, but is occasionally found on K121 or K125 in low frequencies (Ku et al., 2012; Sarcinella et al., 2007). This form of H2A.Z associates with transcriptionally silent facultative heterochromatin, and with the repressed state of the prostate-specific antigen (PSA) gene prior to androgen receptor-mediated transcriptional activation (Draker et al., 2011; Sarcinella et al., 2007). More recently, it has been demonstrated using indirect methods that monoubiquitylated H2A.Z functionally antagonizes binding of the BET bromodomain family member Brd2 at the promoters of bivalent, developmentally-poised, pluripotency genes in mouse embryonic stem cells (mESC) (Surface et al., 2016). Altogether, these have suggested that monoubiquitylated H2A.Z is linked to transcriptional silencing, although the mechanisms by which it elicits this function is unknown.

One well-established mechanism by which histone PTMs mediate their regulatory function within the nucleosome is by recruiting or retaining specific effector proteins. Combinations of histone PTMs can serve as multivalent docking sites which stabilize contacts between chromatin and the recruited protein or complex (Flanagan et al., 2005; Kikuchi et al., 2009; Wysocka et al., 2006). Another critically revealing facet of histone PTM signaling is that many chromatin binding proteins exist within large, multi-subunit complexes (Ikura et al., 2000; Zippo et al., 2009). Importantly, the machineries that catalyze the deposition of PTMs often contain domains that bind the same, or different PTM, resulting in PTM crosstalk, which is the reinforcement of modifications through a positive feedback loop or inhibition of their activity by the deposition of

antagonizing marks on the same or different histone tail (eg. Hunter, 2007). Elucidating the combinatorial patterns of histone PTMs and how they cross-regulate one another is hence essential to understanding how they establish and maintain functional chromatin states.

A major obstacle to fully characterizing H2A.Z ubiquitylation is the lack of an antibody that specifically recognizes the monoubiquitylated form of this variant. Accordingly, in order to directly study monoubiquitylated H2A.Z (H2A.ZUb1) within the chromatin environment, we have developed an affinity purification system to isolate H2A.ZUb1-containing nucleosomes and bypass the previous reagent limitations. This method takes advantage of the specific biotinylation of proximal AviTag sequences (a unique 15 amino acid sequence) by the *Escherichia coli* BirA biotin ligase. When expressed as a fusion protein with BirA, H2A.Z-nucleosomes that are modified with ubiquitin containing an AviTag sequence will automatically be biotinylated by the BirA fusion, and thus can be specifically captured by streptavidin-conjugated beads (Fig. 2-1). Using this strategy, we investigated the PTM status and genome-wide occupancy of H2A.ZUb1-enriched nucleosomes. Importantly, we observe that H2A.ZUb1-nucleosomes are hypomethylated at H3K4, hypoacetylated at H2A.Z, H3K27 and H4, and hypermethylated at H3K27me3. In addition, by examining a subset of interacting partners that co-purify with H2A.ZUb1, we provide insight into the possible mechanisms by which it could function in chromatin repression. Consistent with these findings, ChIP-Seq experiments reveal that H2A.ZUb1 is significantly enriched at the promoters of repressed genes, depleted at active enhancers, and is enriched at developmentally-regulated genes. Collectively, these findings provide the first evidence for a genome-

wide function of monoubiquitylation of H2A.Z in transcriptional repression, and further suggest previously unknown links between H2A.ZUb1 and other key chromatin processes.

2.2 Methods

2.2.1 Cell culture, transfection, plasmids and antibodies

HEK293T cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum (DMEM; Wisent). All transfections were carried out using polyethylenimine (PEI; Polysciences). All expression constructs used were based on the pcDNA 3.1 (+) (Invitrogen) backbone with the Flag-BirA cloned in-frame to the C-terminus of H2A.Z-1(hereafter termed H2A.Z-FB) or H2A.Z-K3R3 where all known sites of ubiquitylation on H2A.Z is mutated to arginines. In addition, the AviTag was cloned in-frame to the N-terminus of ubiquitin. To generate H2A.Z-FB-K3R3, K120/121/125 were mutated to R120/121/125. To generate a non-biotinylatable AviTag, K10 of the AviTag sequence (GLNDIFEAQ**K**IEWHE) was converted to R10. In co-transfection experiments, the ratio of H2A.Z-Flag-BirA to AviTag-Ub plasmids was 3:1. The commercial antibodies used were: H3 (Abcam ab1791), Flag (Sigma F7425), Anti-Avidin (Genscript A00674), Avi-HRP (Sigma A3151), H3K4me1 (Diagenode C15410194), H3K4me2 (Upstate 07-030), H3K4me3 (Active Motif AM39159), H3K27me3 (Millipore 07-449), H3K27Ac (Abcam ab4729), H3K9me2 (Upstate 07-444), H3K9me3 (Millipore UBI 07-442), H2A.ZK4/7/11Ac (Abcam ab18262), H2A.ZK7 (Diagenode C15210012), H4K5/8/12/16Ac (Millipore 06-946), Brd2 (Abcam ab3718),

LSD1 (Abcam 17721), DNMT3L (Abcam ab3493), SMC1 (Bethyl A300-055A), Rad21 (Abcam ab154769), CTCF (Millipore 07-729).

2.2.2 Mononucleosome affinity purification

Generation of mononucleosomes was performed as described previously (Draker et al., 2012). In brief, HEK293T cells were grown in 15 cm-diameter plates and were transfected with various constructs according to the experiments. Cells were trypsinized, counted, and washed in 1X PBS, 48 hrs following transfection. Cellular pellets were resuspended in buffer A (20mM HEPES, pH 7.5, 10mM KCl, 1.5mM $MgCl_2$, 0.34M sucrose, 10% glycerol, 1mM dithiothreitol, 5mM sodium butyrate, 10mM NEM, and protease inhibitors), pelleted and then resuspended in buffer A containing 0.2% Triton X-100 and incubated on ice for 5 min. The nuclear suspension was centrifuged at 600 x g; nuclei were then washed once in buffer A, then resuspended in cutting buffer (15 mM NaCl, 60 mM KCl, 10 mM Tris pH 7.5, 5mM sodium butyrate, 10mM NEM, and protease inhibitors) plus 2mM $CaCl_2$. Microccocal nuclease (MNase; Worthington) was added at a concentration of 10 units/1.0 x 10^7 cells then incubated at 37°C for 30 min. The reaction was stopped by the addition of 20mM EGTA (one twenty-fifth of the reaction volume) and immediate gentle mixing by inversion. The MNase-digested nuclei were centrifuged at 1300 x g. The resulting supernatant (S1) was saved and kept on ice. The digested nuclear pellet was subjected to hypotonic lysis by resuspension in TE buffer (10mM Tris-HCL, pH 8.0, 1mM EDTA). Samples were incubated on ice for 1 hr, with occasional mixing by pipette. The suspension was then centrifuged at 16 000 x g and the supernatant (S2) was transferred to a new tube. Salt was adjusted in S1 to 150mM NaCl by adding 2X buffer D (30 mM Tris pH 7.5, 225 mM NaCl, 3 mM $MgCl_2$,

20% glycerol, 0.4% Triton-X 100, 5mM sodium butyrate, 10mM NEM, and protease inhibitors) drop-wise, with constant mixing on a vortex set to low speed. S2 was also titrated to 150mM NaCl by the drop-wise addition of 3X buffer E (60 mM HEPES pH 7.5, 450 mM NaCl, 4.5 mM $MgCl_2$, 0.6 mM EGTA, 0.6 % Triton-X 100, 30% glycerol, 5mM sodium butyrate, 10mM NEM, and protease inhibitors). Insoluble material was pelleted via centrifugation. The clarified supernatants were combined and then used for affinity purification. Streptavidin-agarose (Sigma) or Flag M2-agarose beads (Sigma) were added and incubated overnight at 4°C on an end- over-end rotator. Beads were washed 4 times in 1X Buffer D, followed by 3 washes in 1X Buffer D containing 0.5% Triton X-100. Proteins were eluted from the beads by resuspension in 2X SDS sample buffer and boiled for 10min. For Western blot analysis, samples were run on SDS-polyacrylamide electrophoresis gels according to standard practices.

2.2.3 ChIP-Seq (analysis performed by Ulrich Braunschweig through collaboration with Dr. Benjamin Blencowe at U of Toronto)

Mononucleosome affinity purification was performed in duplicate as described above using streptavidin-agarose, Flag M2-agarose, or H3 antibody (pulled-down using protein G-coupled Dynabeads; Invitrogen) and eluted in buffer D containing 1% SDS by end-over-end rotation at room temperature for 2 X 10 min. DNA was treated with RNase A and proteinase K, purified by phenol-chloroform extraction, and then re-precipitated with ethanol and resuspended in water. DNA was converted to libraries by the Donnelly Sequencing Centre using Illumina TruSeq ChIP-Seq and sequenced on an Illumina NextSeq500 in single end mode. Reads were converted to FASTQ format and mapped to the human hg19 genome using Bowtie. Duplicate reads were removed. Peaks were

called using MACS2 using H3 as a control. Normalized peaks from duplicate experiments were then averaged and matched H3 samples were subtracted to yield fragments per million reads (FPM). The IDR procedure (Li et al., 2011) employed by ENCODE was then used to generate a merged peak set for both streptavidin and Flag with all raw peaks as input. At FDR < 0.05, 143k peaks (from 453k and 552k in streptavidin or Flag samples, respectively) were recovered in both. Enhancers were identified as co-localization of H3K4me1, H3K27ac, and DNaseI hypersensitivity sites (HSS) in either HeLa or several other cells and tissues.

2.2.4 ChIP-qPCR

Affinity-purified mononucleosomes were eluted in buffer D containing 1% SDS as described above. DNA was treated with RNase A and proteinase K, phenol-chloroform extracted from mononucleosomes, re-precipitated with ethanol, and then resuspended in water. Quantitative polymerase chain reactions (qPCR) were assembled in triplicate using PerfeCta SYBR Green SuperMix (Quanta Biosciences) and gene-specific primers. Reactions were run on an Optocon 2 thermocyler (Biorad). Primers used are listed in table S1 of Appendix.

2.3 Results

2.3.1 Isolation of H2A.ZUb1

In order to isolate H2A.ZUb1-containing nucleosomes, we developed an affinity purification technique that harnesses the specificity of *E.coli* BirA for the AviTag (acceptor peptide) sequence (Fig. 2-1). BirA is a biotin ligase that catalyzes the

proximity-dependent and site-specific biotinylation of the AviTag. Importantly, the bacterial BirA enzyme does not biotinylate any endogenous mammalian proteins, and similarly, mammalian biotin-ligases do not recognize the AviTag (Barker and Campbell, 1981; Beckett et al., 1999; Cull and Schatz, 2000; Schatz, 1993). We generated an H2A.Z-Flag-BirA fusion protein (H2A.Z-FB) and an AviTag-ubiquitin (AviTag-Ub) construct.

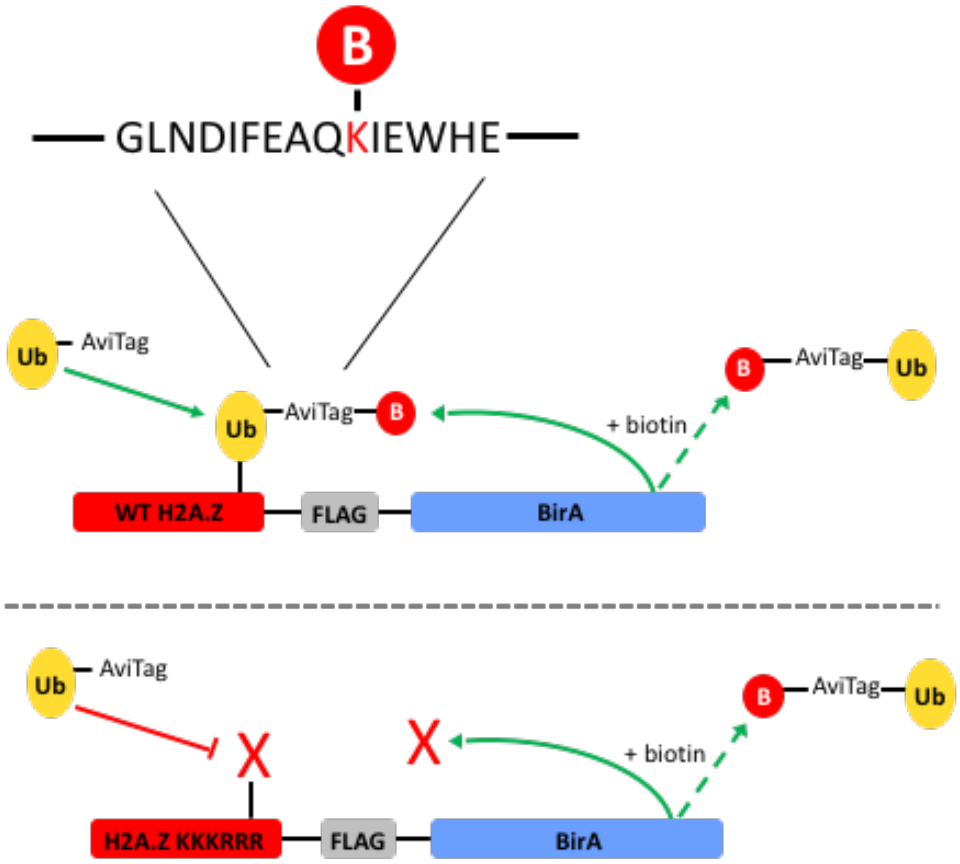


Figure 2-1. Schematic of H2A.ZUb1 affinity purification system. (Top) Monoubiquitylation of wildtype H2A.Z-Flag-BirA with AviTag-ubiquitin results in biotinylation of the AviTag by BirA. Biotinylated H2A.Z-Flag-BirA-AviTag-Ub1 (H2A.ZUb1) can then be isolated by streptavidin-conjugated beads. (Bottom) H2A.Z-Flag-BirA in which K120R, K121R, and K125R mutations have been introduced to H2A.Z cannot be monoubiquitylated and hence do not become modified by AviTag-ubiquitin and cannot be captured by streptavidin.

We predicted that co-expression of these constructs in mammalian cells would result in preferential biotinylation of AviTag-ubiquitin incorporated on H2A.Z-FB, and in this manner, H2A.Z-FB-AviTag-Ub1-nucleosomes could be affinity-purified using streptavidin-conjugated beads.

We first assessed the ability of H2A.Z-FB to specifically biotinylate itself in the presence of AviTag-Ub (Fig. 2-2).

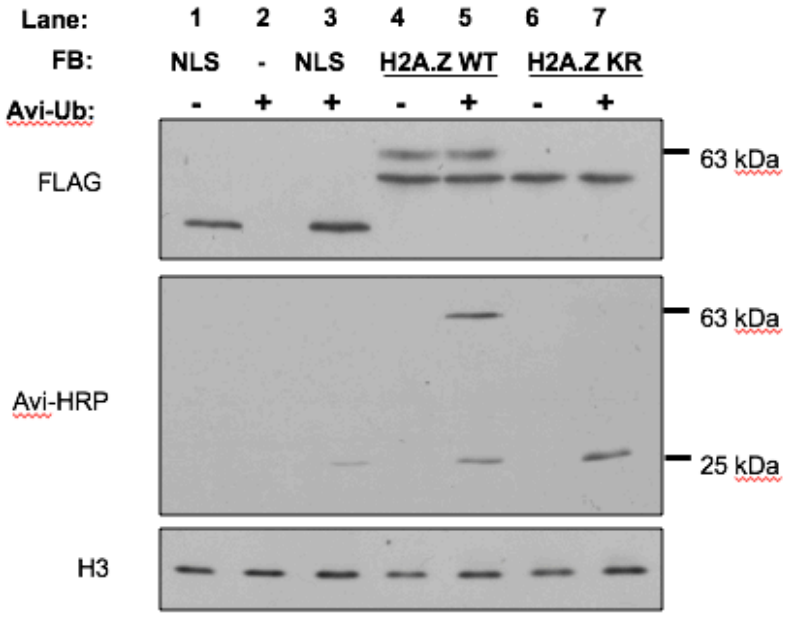


Figure 2-2. H2A.ZUb1 is the main product of co-expressing H2A.Z-Flag-BirA and AviTag-ubiquitin in HEK293T cells. Flag-BirA fused to a nuclear localization signal (NLS), H2A.Z-Flag-BirA (H2A.Z WT), or H2A.Z-K3R3-Flag-BirA (H2A.Z KR) were transiently transfected with AviTag-ubiquitin (+) or singly expressed without (-). Western blots of nuclear lysates are shown and are probed as indicated. Molecular weights indicated align with migration of H2A.Z-Flag-BirA-AviTag-ubiquitin (top band) or endogenous H2A/H2A.Z-AviTag-ubiquitin (bottom band).

For comparison, we included a construct in which Flag-BirA alone is fused to a nuclear localization signal (NLS), as well as an H2A.Z-FB construct in which the three sites of H2A.Z monoubiquitylation are converted to arginine and hence rendered non-ubiquitylatable (H2A.Z-K3R3-FB). As predicted, when co-expressed with AviTag-ubiquitin, the predominant band detected by Western blot when nuclear lysates were probed with avidin-conjugated horseradish peroxidase (Avi-HRP) is approximately 63 kDa, which roughly corresponds to the molecular weight of H2A.Z-FB modified with biotinylated AviTag-Ub (Fig. 2-2, lane 5). The presence of this band is dependent on the co-transfection of AviTag-ubiquitin and also on the capacity of H2A.Z-FB to be monoubiquitylated (compare lane 5 to lanes 3 and 7 in Fig. 2-2). Additionally, a second, weaker signal is detected by Avi-HRP that is roughly 25 kDa. This minor band is similarly dependent on co-expression of AviTag-ubiquitin; however, it is also present when AviTag-ubiquitin is co-expressed with FB-NLS, albeit in lesser-abundance than when co-expressed with H2A.Z-FB or H2A.Z-K3R3-FB. The estimated molecular weight of this minor band approximates those of endogenous ubiquitylated H2A.Z, or H2A and we surmise it is likely due to H2A.Z-FB reaching and biotinylating Avi-ubiquitylated endogenous H2A.Z or H2A that co-exists with H2A.Z-FB within the nucleosome context. Nevertheless, as our H2A.ZUb1 construct is the most readily detectable biotinylated band, and presence of the minor band is more dependent on the fusion of Flag-BirA and

H2A.Z, we concluded that intramolecular biotinylation of monoubiquitinated H2A.Z-FB is a sufficiently selective reaction.

We next used mononucleosomes prepared from nuclear lysates as input for affinity purification of H2A.ZUb1 using streptavidin- or Flag-conjugated beads. Here we included the co-transfection of H2A.Z-FB with a non-biotinylatable AviTag-ubiquitin construct (AviTag-K10R-Ub) in order to assess specificity of the streptavidin beads in our pull-down assays. Coomassie staining and visualization of affinity-purified samples confirmed assembly of H2A.Z-FB into nucleosomes based on the stoichiometric co-

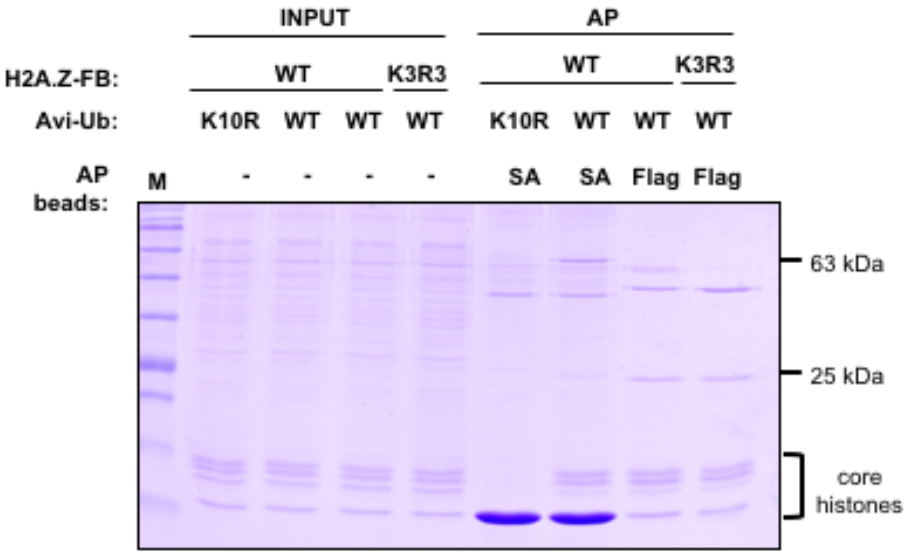


Figure 2-3. H2A.ZUb1 is incorporated into nucleosomes and can be enriched using streptavidin-coupled beads. Mononucleosomes were prepared from HEK293Ts co-expressing H2A.Z-Flag-BirA (H2A.Z-FB) wildtype (WT) or non-ubiquitylatable mutant H2A.Z (K3R3) fused to Flag-BirA with either biotinylatable AviTag-ubiquitin (WT) or non-biotinylatable AviTag-ubiquitin (K10R), as indicated, and used as input for affinity purification (AP) using streptavidin beads (SA) or Flag beads (Flag). Samples are normalized for nucleosome content and separated by SDS-polyacrylamide gel electrophoresis stained with Coomassie Brilliant Blue. Note that streptavidin runs at the same molecular weight as histone H4. M indicates molecular marker.

precipitation of additional core histones, and also confirmed the enrichment of H2A.ZUb1 using streptavidin beads (Fig. 2-3).

Using greater resolution gel electrophoresis and Western blot, we were able to separate monoubiquitylated H2A.Z-FB into two bands in pre-purified (input) lysates, corresponding to forms of H2A.Z-FB modified by either endogenous ubiquitin or a smaller fraction of AviTag-ubiquitin (see cartoon depiction on Fig. 2-4). Two bands of monoubiquitylation were also discernable in the input lysates of cells co-transfected with AviTag-K10R-Ub (Fig. 2-4). Importantly, streptavidin pull-downs from lysates co-expressing H2A.Z-FB and AviTag-Ub specifically enrich for mononucleosomes containing H2A.Z-FB modified by biotinylated AviTag-Ub (i.e. H2A.ZUb1). In contrast, pull-downs from the same lysates using Flag antibody-coupled beads yields a mixture of mainly non-ubiquitylated H2A.Z (unmodified) and a smaller fraction of monoubiquitylated H2A.Z, and is considered a pool of bulk or “total” H2A.Z in our system. The distinct forms of H2A.Z pulled-down by the respective beads (i.e. streptavidin versus Flag) was most evident when equal amounts of affinity-purified nucleosomes were loaded on the same gel (normalized for H3 content) for comparison. As shown in Figure 2-4, we find that nucleosomes pulled-down with streptavidin exclusively contain the monoubiquitylated form of H2A.Z-FB (as detected by the Flag antibody) whereas in the Flag pull-down, the main form of H2A.Z-FB purified was non-ubiquitylated (marked by the triple asterisk in the figure) and only a very small amount of ubiquitylated H2A.Z-FB (shifted band with double-asterisk) was co-purified. These results confirm the selective biotinylation of the Avi-ubiquitylated H2A.Z-FB and high

degree of purification achieved by streptavidin beads. Therefore, along with affinity-purified mononucleosomes enriched with non-ubiquitylatable H2A.Z-K3R3-FB, we next

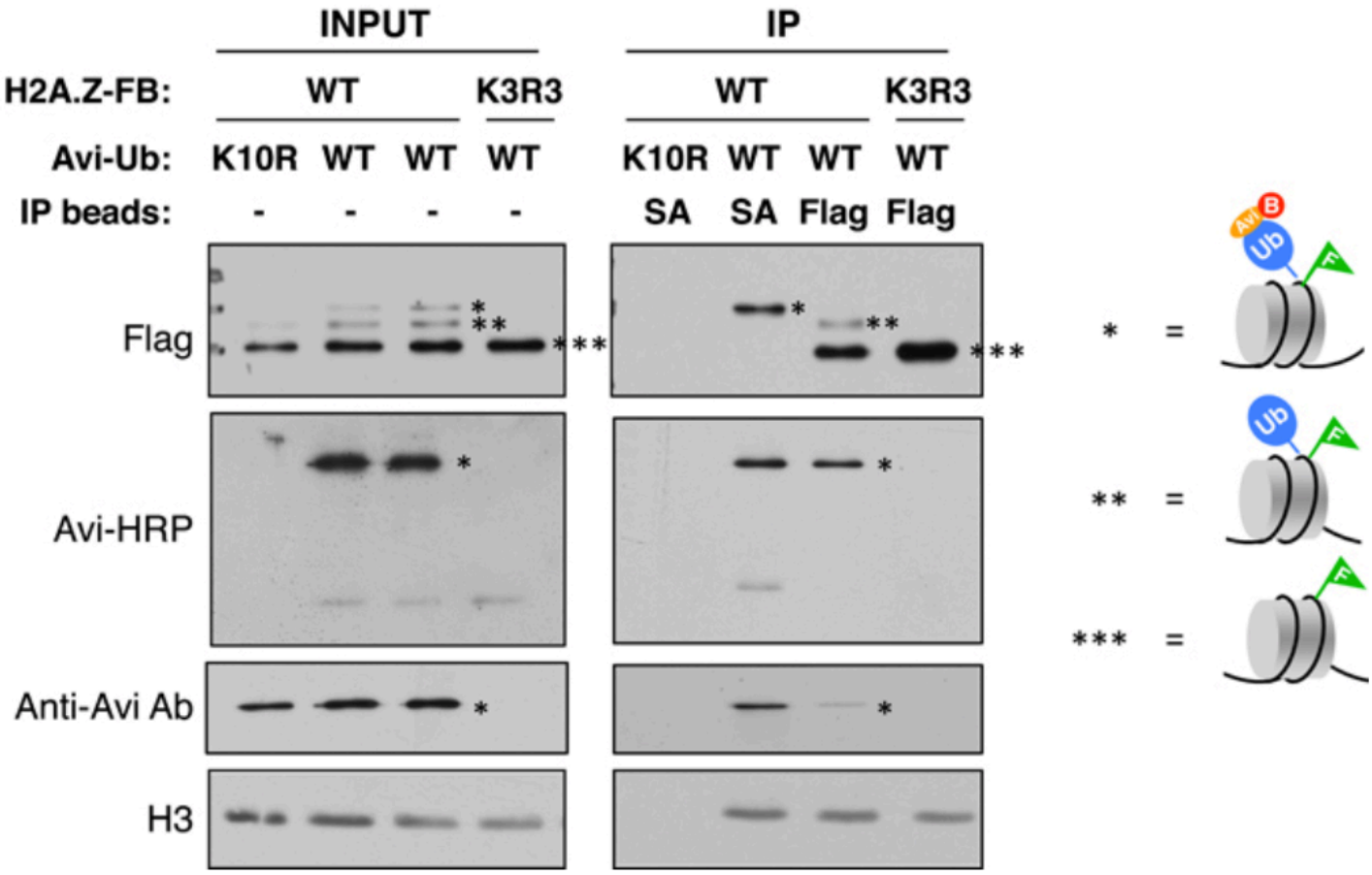


Figure 2-4. H2A.ZUb1 can be differentially enriched using streptavidin- and Flag-coupled beads. Mononucleosomes were prepared from HEK293Ts co-expressing H2A.Z-Flag-BirA (H2A.Z-FB) wildtype (WT) or non-ubiquitylatable mutant H2A.Z (K3R3) fused to Flag-BirA with either biotinylatable AviTag-ubiquitin (WT) or non-biotinylatable AviTag-ubiquitin (K10R), as indicated, and used as input for affinity purification (AP) using streptavidin beads (SA) or Flag beads (Flag). Samples are normalized for nucleosome content probed with antibodies as indicated.

used these samples to assess the effect of H2A.ZUb1 on nucleosome composition, binding partners, and genome-wide localization.

2.3.2 H2A.ZUb1 nucleosomes possess a distinct composition of histone post-translational modifications

Using the affinity purification scheme described, we first compared the associations of the differentially ubiquitylated populations of H2A.Z in HEK293T cells with a subset of well-characterized histone PTMs in order to garner insight into potential histone crosstalk involving H2A.ZUb1-nucleosomes (Fig. 2-5). We observe that H2A.ZUb1-enriched mononucleosomes are enriched for H3K27me3, consistent with our earlier finding that H2A.Z can be monoubiquitylated by the Polycomb repressive complex 1 (PRC1), and the well-established interplay between PRC1 and the H3K27-methyltransferase complex, PRC2 (Blackledge et al., 2014; Cooper et al., 2014; Fischle et al., 2003; Holoch and Margueron, 2017; Min et al., 2003). In contrast, we did not observe any obvious differences amongst differentially enriched samples for H3K9me2/me3, histone modifications linked to constitutive repression (Bannister et al., 2001; Lachner et al., 2001; Nakayama et al., 2001; Rea et al., 2000). We also observe that H2A.ZUb1-enriched mononucleosomes are depleted of H3K4-methylation marks (H3K4me1/2/3). Previously, we reported an enrichment of the promoter activity-linked modification H3K4me3 on total H2A.Z-nucleosomes in comparison to those containing H2A (Sarcinella et al., 2007). However, the preferential co-occurrence of these marks appears to be biased towards unmodified H2A.Z-mononucleosomes. In addition to

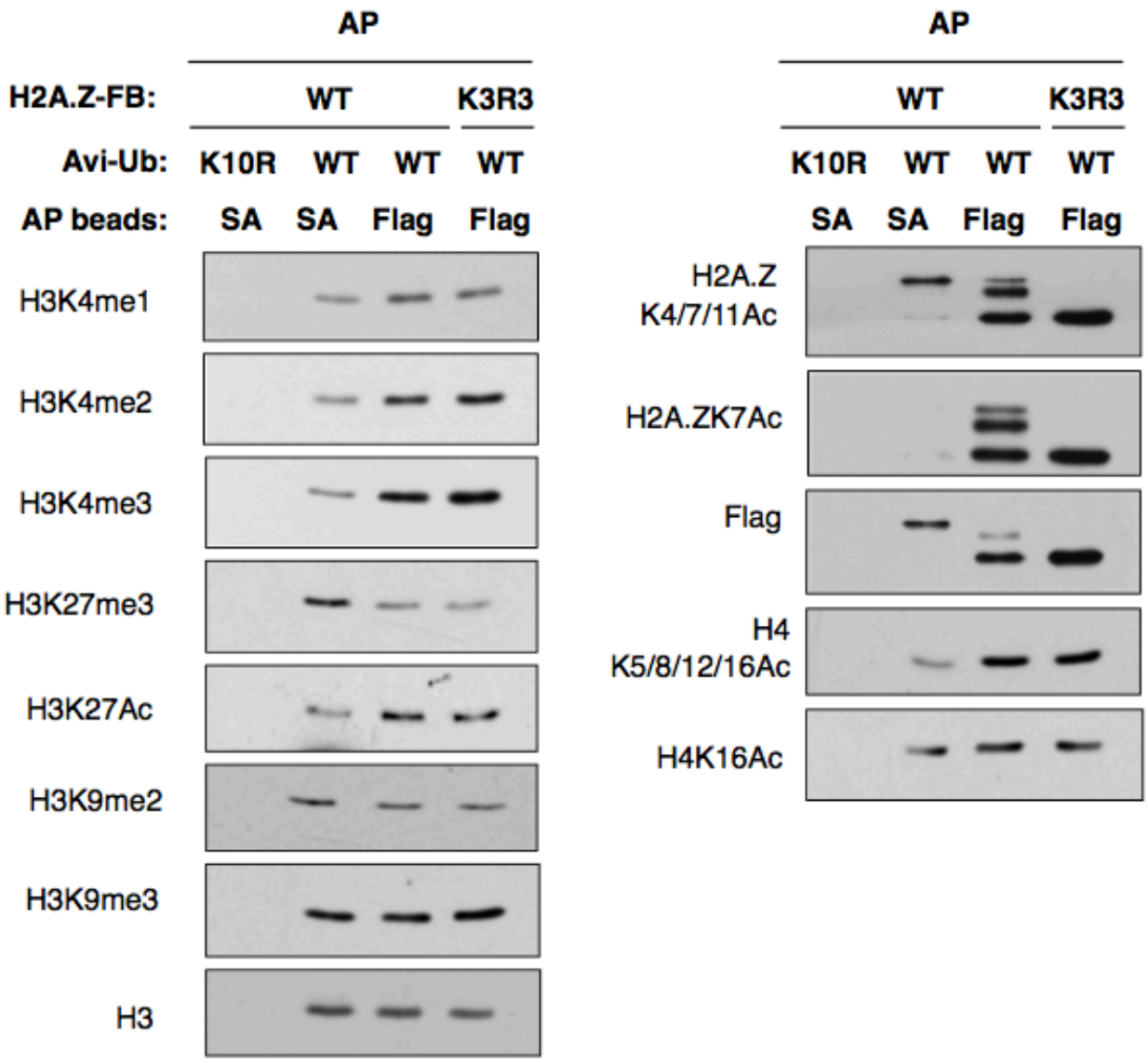


Figure 2-5. H2A.ZUb1 nucleosomes are depleted for histone PTMs associated with transcriptional activity and enriched for H3K27me3, a mark of transcriptional repression. HEK293Ts were co-transfected with H2A.Z-FB (WT or K3R3) and Avi-Ub (WT or K10R). Mononucleosome preparations were immunoprecipitated with either streptavidin-agarose beads (SA) or Flag-agarose beads (F). Immunoprecipitates were normalized for nucleosome content using H3 and probed with antibodies indicated.

demarcating nucleosome depleted regions (NDR) around the transcriptional start site (TSS) of promoters, H2A.Z in combination with H3.3 also occupies active enhancer regions (Brunelle et al., 2015; Gévry et al., 2009; Jin et al., 2009). Enhancers can exist in 4 states: decommissioned (no histone PTM), poised (H3K4me1 and H3K27me3), primed (H3K4me1), or active (H3K4me1 and H3K27ac**)** (Creyghton et al., 2010; Rada-Iglesias et al., 2011; Zentner et al., 2011). We find that H2A.ZUb1-mononucleosomes are depleted of both H3K4me1 and H3K27ac, suggesting that H2A.ZUb1-containing nucleosomes, in contrast to earlier reports studying total H2A.Z, are relatively depleted at active enhancers.

In addition to the depletion of H3K27ac, we observe that H2A.ZUb1-mononucleosomes are hypo-acetylated at other sites as well. Specifically, we observe a dramatic depletion of acetylation of H2A.Z at lysine 7 (H2A.ZK7ac) on H2A.ZUb1-enriched mononucleosomes. The contrasting presence of H2A.ZK7ac in Flag-immunoprecipitates indicates, however, that our H2A.ZUb1-constructs can also be acetylated at K7. This could suggest differences in the composition of affinity-purified H2A.ZUb1 between differentially enriched samples. For example, if our H2A.ZUb1-enriched population also enriched for homotypic H2A.ZUb1-mononucleosomes (where two copies of H2A.Z are monoubiquitylated) then it could be this state that precludes the concomitant acetylation of H2A.ZK7. Alternatively, we find that H2A.Z acetylation detected using an antibody that recognizes acetylation at K4, K7, and K11 does not discernably vary, raising the possibility that acetylation/ deacetylation of H2A.ZUb1 at K7 is a selective reaction.

We have also previously found that bulk H2A.Z-containing nucleosomes are hyperacetylated at H4 in comparison to nucleosomes containing H2A (Draker et al., 2012; Sarcinella et al., 2007). However, in the case of H2A.ZUb1, by probing the pulled-down nucleosomes with an antibody recognizing H4 acetylation at K5/8/12, or K16, we observe that monoubiquitylation is associated with H4 acetylation depletion. Interestingly, we also find that levels of H4K16ac do not vary with H2A.Z monoubiquitylation status, indicating that H4 acetylation dynamics are affected by H2A.ZUb1 in a site-specific manner.

In summary, we have found H2A.ZUb1-enriched nucleosomes to be depleted of modifications associated with active promoter and enhancer function, and instead to be enriched for the repressive mark H3K27me3.

2.3.3 Chromatin binding proteins differentially associate with H2A.ZUb1 nucleosomes

We and others previously identified Brd2 as a chromatin binding protein with preferential enrichment on H2A.Z-nucleosomes, hinting at a mechanism for H2A.Z-mediated transcriptional regulation (Draker et al., 2012; Kim et al., 2013; Vardabasso et al., 2015). More recently, indirect approaches have suggested that monoubiquitylation of H2A.Z antagonizes the binding of Brd2 at bivalent promoters (i.e. those marked by nucleosomes bearing both H3K4me3 and H3K27me3) in mESCs (Surface et al., 2016). Given these findings, we first assessed the binding of Brd2 to H2A.ZUb1-mononucleosomes and find that they are dramatically depleted of this interaction. This is consistent with the observation of antagonism in mESCs, and indicates that

antagonism between Brd2 and H2A.ZUb1 is a general phenomenon that occurs in differentiated cells as well as ESCs (Fig. 6).

We next screened H2A.ZUb1-mononucleosomes for interactions with several possible effector proteins that are functionally linked to histone PTMs for which we observe differential enrichment on H2A.ZUb1 (Fig. 2-6). Consistent with the depletion of H3K4 methylation on H2A.ZUb1, H2A.ZUb1-mononucleosomes are enriched for KDM1A/ LSD1 (lysine-specific demethylase-1). LSD1 possesses H3 lysine 4 demethylase activity as part of the NuRD (nucleosome remodeling and histone deacetylase) repressor complex, which couples histone demethylation to histone deacetylation (Wang et al., 2009). In contrast we find that menin, a co-factor of the MLL1/2 (mixed-lineage leukemia 1/2) H3K4 methyltransferases, does not differentially associate with H2A.ZUb1, and is co-purified with H2A.ZUb1- and H2A.Z-K3R3-nucleosomes at similar levels. Though seemingly inconsistent given the observed hypomethylation of H3K4 in H2A.ZUb1-nucleosomes, it has previously been reported that H2A119KUb1 and MLL can regulate - and are regulated - allosterically, respectively (Wu et al., 2013; Yuan et al., 2013), and hence it might be the activity, and not the targeting, of MLL that is affected by presence of H2A.ZUb1.

We also observe a clear, preferential enrichment of DNMT3L [DNA (cytosine-5)-Methyltransferase 3-Like] on H2A.ZUb1-mononucleosomes (Fig. 2-6). DNMT3L is an enzyme that binds to, and stimulates the activity of, the *de novo* DNA methyltransferases, DNMT3A and DNMT3B (Chédin et al., 2002; Jia et al., 2007; Suetake et al., 2004). Like DNMT3A and DNMT3B, DNMT3L is able to sense the methylation status of H3K4, and preferentially interacts with unmethylated H3K4

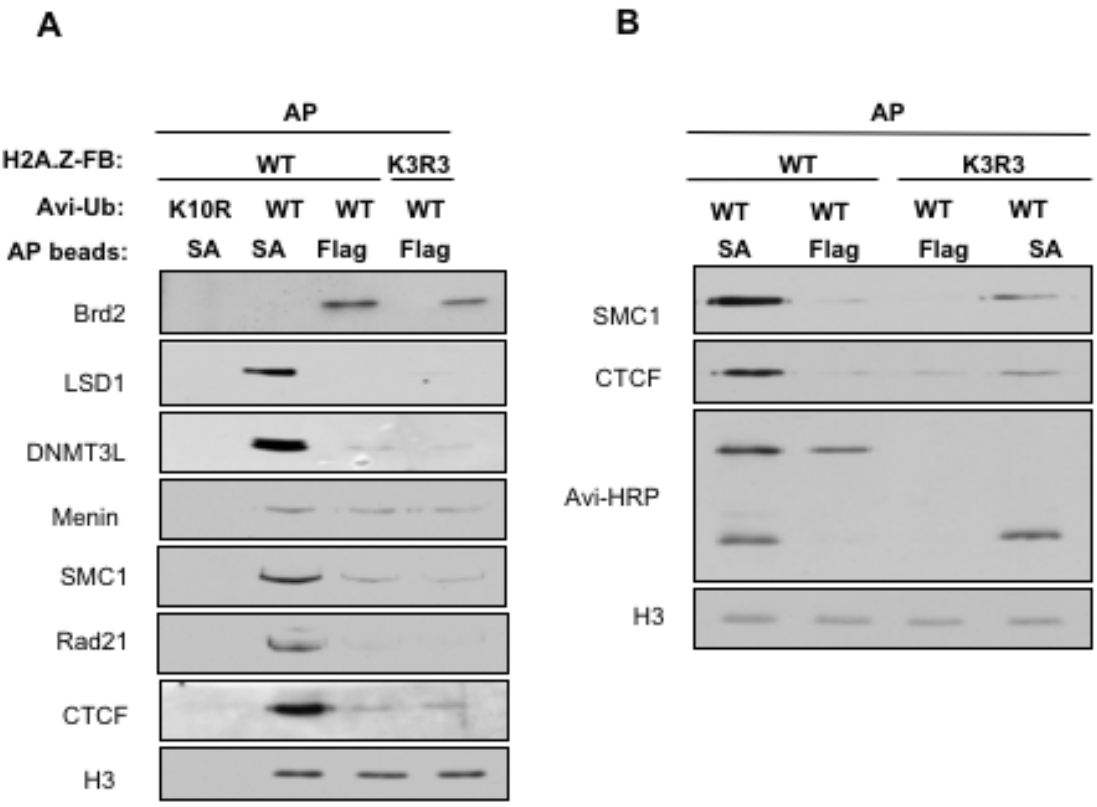


Figure 2-6. H2A.ZUb1 nucleosomes are enriched for specific chromatin-binding proteins. (A) HEK293Ts were co-transfected with H2A.Z-FB (WT or K3R3) and Avi-Ub (WT or K10R). Mononucleosome preparations were immunoprecipitated with either Streptavidin-agarose beads (SA) or Flag-agarose (F) beads. Immunoprecipitates were normalized for nucleosome content using H3 and probed with antibodies indicated. (B) Co-immunoprecipitation of CTCF and cohesin component SMC1 with H2A.ZUb1 were compared with H2AUb1 or H2BUb1 by their immunoprecipitation with Streptavidin-coupled beads in an H2A.Z-K3R3 background.

(H3K4me0) (Argentaro et al., 2007; Ooi et al., 2007; Otani et al., 2009). The maintenance DNA methyltransferase DNMT1 was also recovered in an interaction screen as one the proteins most strongly enriched on H2AK119Ub1-nucleosomes in comparison to non-ubiquitylated H2A (Kalb et al., 2014), which suggests a similar possible connection between a preference of DNA methyltransferases with ubiquitylated H2A-family members.

Finally, given the reported link between H2A.Z and enhancers, and its wider association with differentially packaged chromatin, we surveyed the ability of its modified forms to interact with SMC1 and Rad21, members of the cohesin complex, as well as the chromatin topology and transcription factor CTCF, which are thought to facilitate DNA looping between distal loci (Merkenschlager and Nora, 2016; Ong and Corces, 2014). To our surprise, we observe that both cohesin and CTCF show a marked preference for H2A.ZUb1 (Fig. 2-6). In order to exclude the possibility that this enrichment is due to binding of these proteins to the endogenous, monoubiquitylated lower band, we specifically purified the lower band from cells expressing the non-ubiquitylatable H2A.Z-BirA fusion protein (H2A.Z-K3R3-FB) using streptavidin-coupled beads and compared this to H2A.ZUb1 (Fig. 2-6B). As shown in Figure 2-6B, both CTCF and cohesin preferentially purify with H2A.ZUb1 in comparison to endogenous monoubiquitylated nucleosomes, which may comprise a pool of H2A.ZUb1, H2AUb, and H2BUb1. This result further suggests that the preferential association of cohesin and CTCF with H2A.Z is dependent on the levels of H2A.ZUb1. Altogether, our affinity

purification results suggest that H2A.ZUb1-nucleosomes are enriched for, and thus functionally linked to, effector proteins that are known to function in transcriptional repression.

2.3.4 Genome-wide mapping of H2A.ZUb1 occupancy

The results thus far provide evidence that H2A.ZUb1 is preferentially associated with histone PTMs and interacting factors linked to the control of repressive chromatin states. Next, to determine whether H2A.ZUb1 is more generally associated with repressed genes, we used our purification strategy in combination with high throughput sequencing (which is equivalent to ChIP-Sequencing) to investigate the occupancy of H2A.ZUb1-mononucleosomes on a genome-wide scale. Specifically, we performed ChIP-Seq on H2A.ZUb1- mononucleosomes isolated from HEK293T cells using our streptavidin purification scheme or by Flag-affinity purification (total H2A.Z). The latter was used as a control for the distribution of steady-state-modified ectopic H2A.Z on chromatin.

Overall, the mapping rates for streptavidin-purified H2A.ZUb1 or total H2A.Z were > 97% and >70% unique reads, with 60% of streptavidin-purified H2A.ZUb1 peaks overlapping with Flag-peaks. Most overlap between H2A.ZUb1-mononucleosomes and total H2A.Z was found to occur near genes. The average profile for both streptavidin- and Flag-peaks across > 40,000 annotated genes comprised canonical, well-phased, H2A.Z-containing bimodal nucleosomes flanking promoter NDRs, with signals decreasing towards the gene body (Supplementary Fig. 1). By correlating peaks with

gene expression measurements from RNA-seq data obtained from HEK293Ts, we observe that total levels of H2A.Z (Flag ChIP) in the promoter region correlate positively

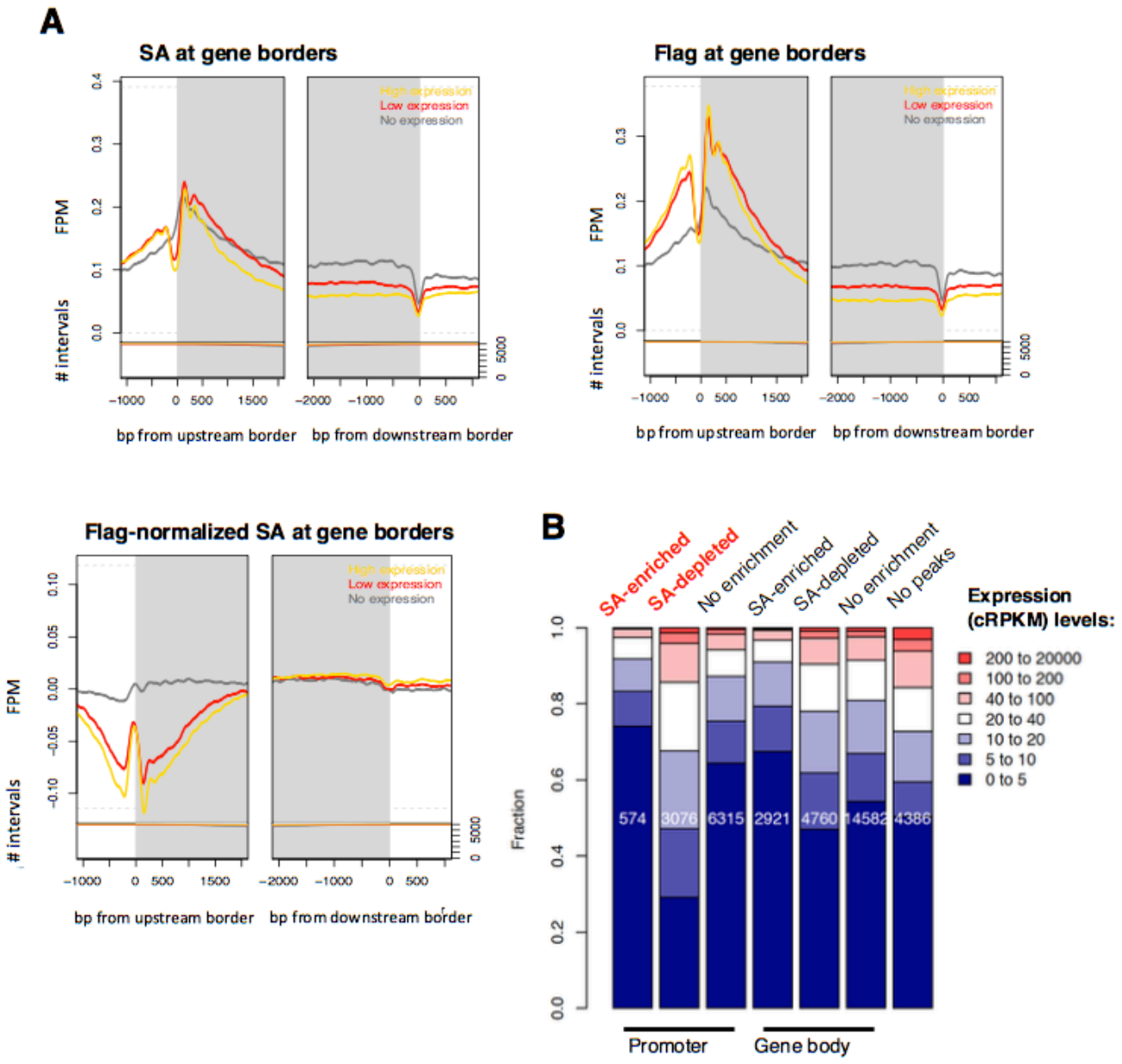


Figure 2-7. H2A.ZUb1 is enriched within the promoter and gene body of non-expressed genes. (A) ChIP signals at genes with different expression levels in HEK293T. Expression groups are defined as follows: high expression, $x > 12$ cRPKM (yellow); low expression, $1 < x < 12$ cRPKM (red); no expression, $x < 1$ cRPKM (grey). (B) Expression of genes with H2A.Z peaks in the indicated position. SA = streptavidin.

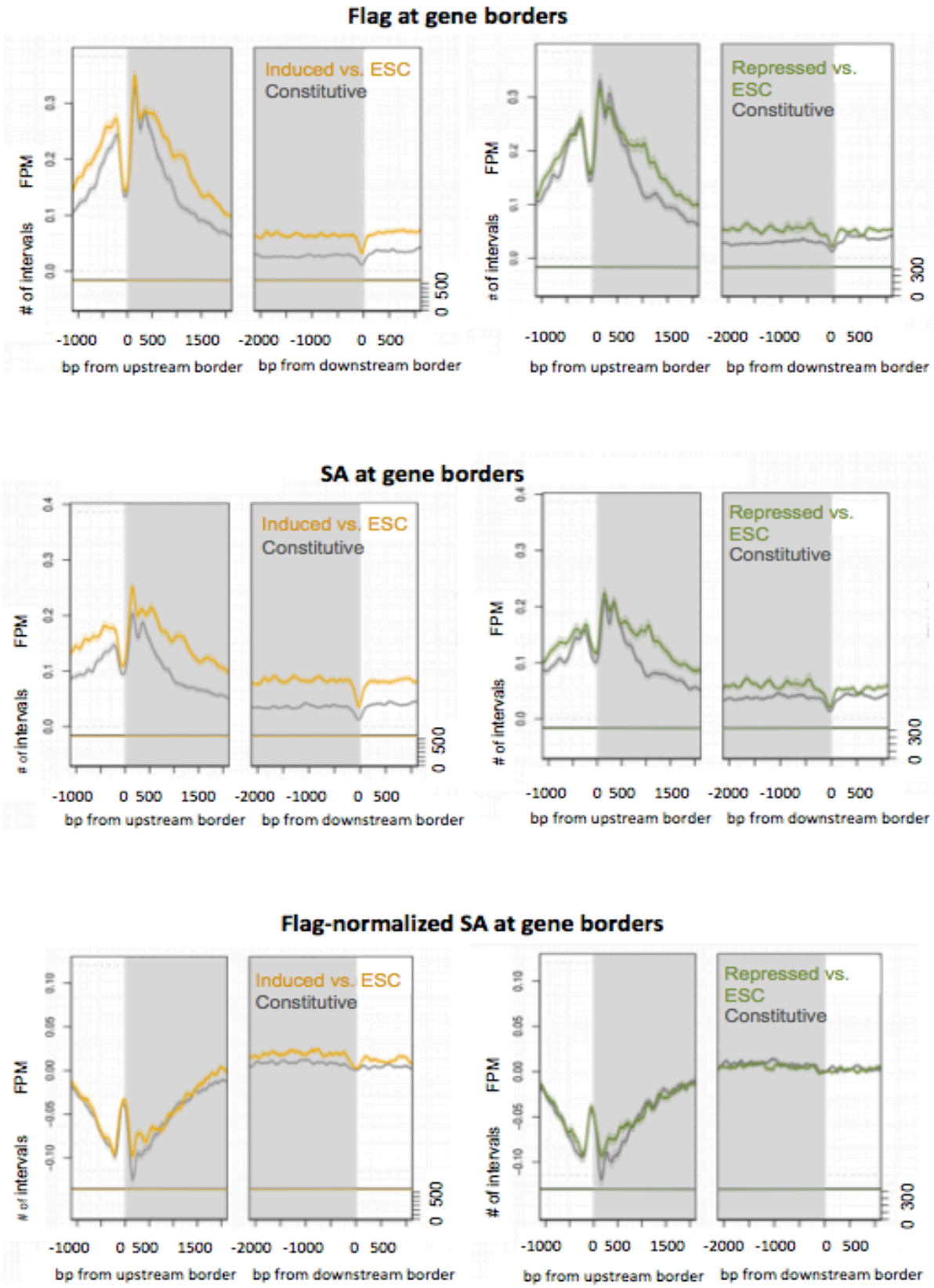
with gene expression, both low and high. In contrast, we observe that levels of H2A.ZUb1-mononucleosomes (streptavidin ChIP) around the TSS do not vary with expression state, and are present at low levels in both expressed and non-expressed genes (Fig. 2-7A).

By merging all peaks detected in either total (Flag)- or streptavidin-purified H2A.ZUb1-mononucleosomes and then comparing the read density of H2A.ZUb1 with total H2A.Z for each peak, we detected two sets of peaks with robust differences in relative signal. A significant enrichment is referred to as an “H2A.ZUb1-enriched” peak, and as an “H2A.ZUb1-depleted” peak if the peak were depleted. If there were no significant difference in read densities, peaks were designated as “no enrichment”. While most peaks have both total H2A.Z- and H2A.ZUb1-nucleosomes at average levels, a noteworthy exception is the presence of high levels of both H2A.ZUb1 and total H2A.Z within the gene bodies of non-expressed genes (Fig. 2-7A; top left and right respectively). Importantly, by comparing the peaks of H2A.ZUb1 with total H2A.Z, we observe that even though non-expressed genes have lower levels of H2A.Z around the TSS, a much higher fraction, of it, is monoubiquitylated and H2A.ZUb1-enriched (Fig. 2-7A; bottom). In contrast, a significantly greater proportion ($p < 0.05$, Fisher’s exact) of H2A.ZUb1-depleted promoters compared to total H2A.ZUb1-bound promoters correspond to highly expressed genes (Fig. 2-7A; bottom)

We next asked whether H2A.Z or H2A.ZUb1 preferentially occupy cell type-dependent or cell type-independent genes. For this, we defined groups of genes based

on their expression in HEK293T versus ESCs. In order to exclude the possibility that differences in occupancy were due to expression differences in our samples, subsets of

A


Flag at gene borders
Induced vs. ESC
Constitutive
Repressed vs.
ESC
Constitutive
FPM
# of intervals
bp from upstream border
bp from downstream border
SA at gene borders
Induced vs. ESC
Constitutive
Repressed vs.
ESC
Constitutive
FPM
# of intervals
bp from upstream border
bp from downstream border
Flag-normalized SA at gene borders
Induced vs. ESC
Constitutive
Repressed vs.
ESC
Constitutive
FPM
# of intervals
bp from upstream border
bp from downstream border

B

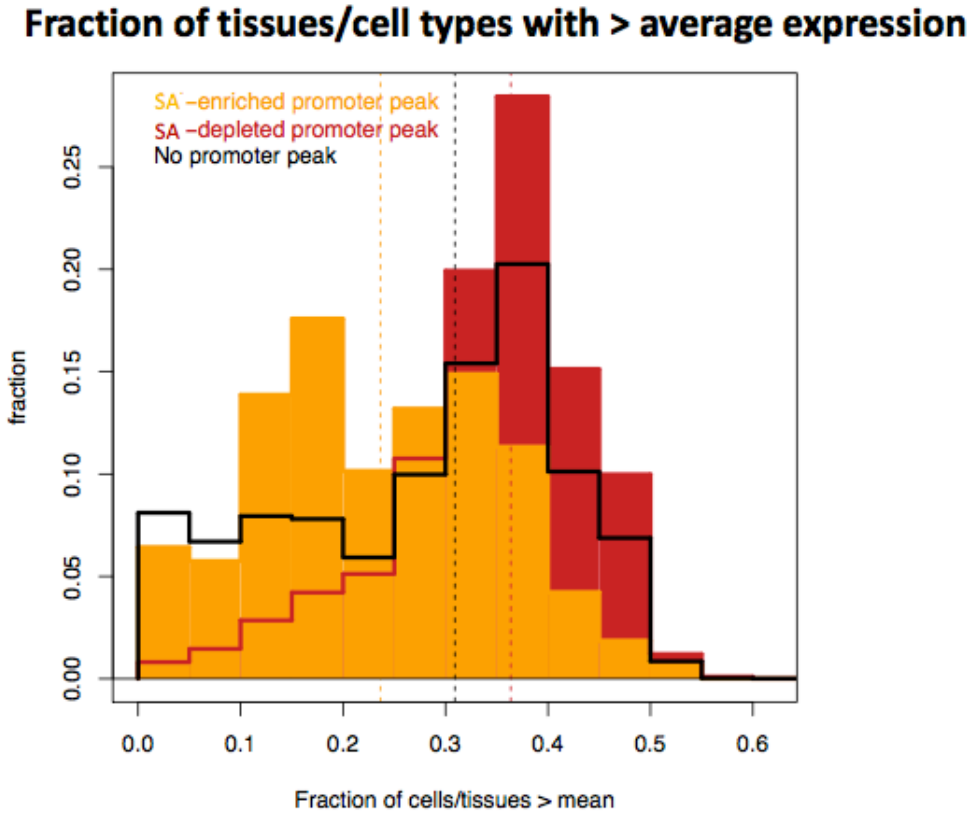


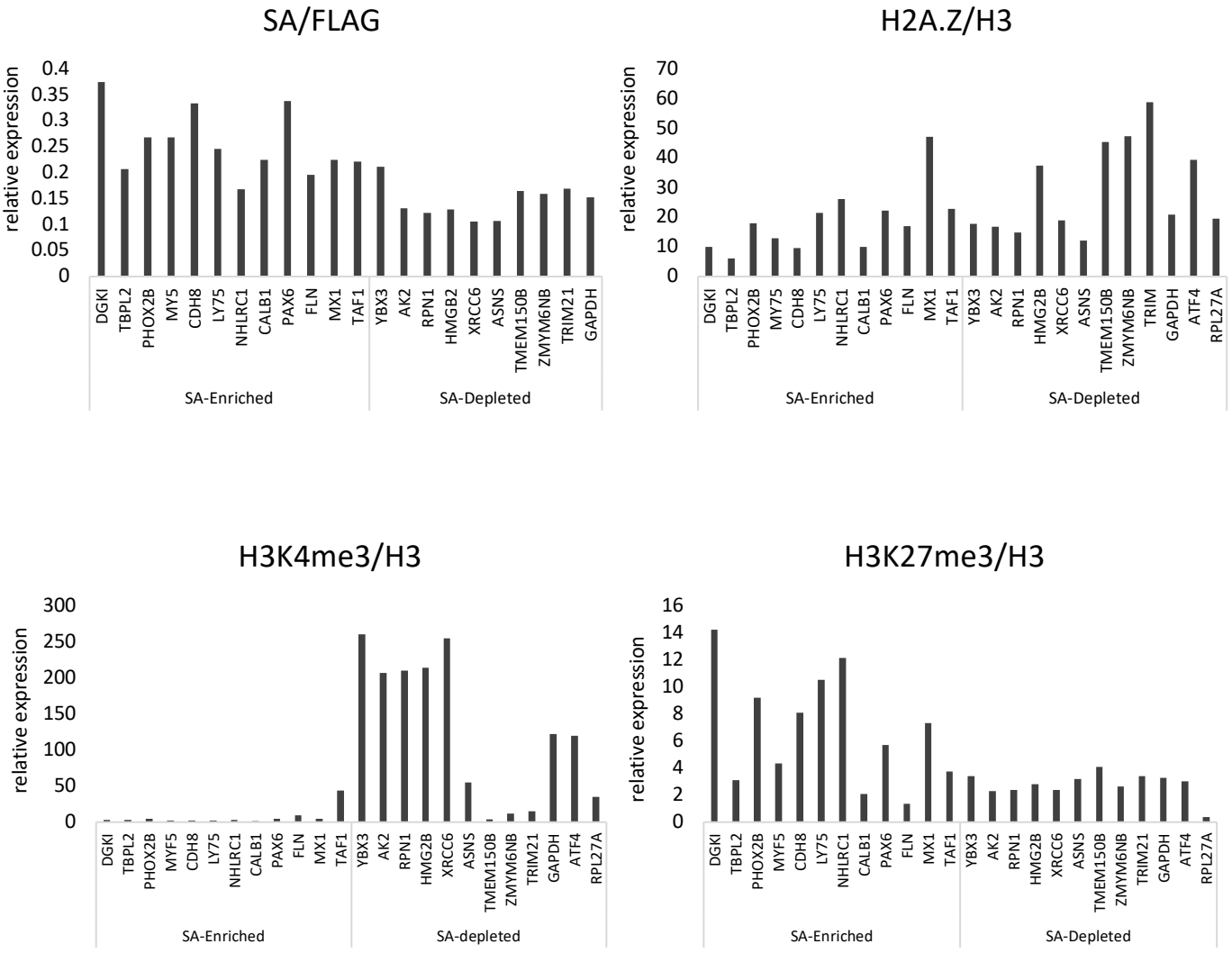
Figure 2-8. ChIP signals at inducible and constitutive genes. (A) Left column, 'induced' are genes with > 2-fold expression relative to ESC and > 5 cRPKM in HEK293Ts; 'constitutive' are genes with SD < 0.5 mean expression. Right column analogously, except 'repressed' are genes with < 2-fold less expression relative to ESC and minimum ESC expression > 5 cRPKM. In each column, only subsets of each group are used that have matching distributions of expression in HEK293Ts. (B) Fraction of 55 tissues and cell lines in which the expression of a gene is greater than the mean. Dashed lines indicate group medians. SA = streptavidin.

the compared groups of genes were matched for expression levels (Fig. 2-8). As shown in Figure 2-7, both total H2A.Z and H2A.ZUb1 occupancy are significantly higher over cell-type dependent genes ($p < 0.05$, Fisher's exact); including genes that are relatively highly expressed in HEK293T versus ESC, as well as those that are comparatively lowly-expressed. We also looked at the "evenness" of expression of H2A.ZUb1-enriched and H2A.ZUb1-depleted genes across a panel of 55 tissues and cell lines in existing RNA-seq data sets (Fig. 2-8). We find that genes that are H2A.ZUb1-enriched are highly expressed in a smaller number of cell types than genes with H2A.ZUb1-depleted peaks, or no peaks at the promoter. These results support the notion that genes with a restricted expression patterns are preferentially regulated by monoubiquitylation of H2A.Z.

2.3.5 Validation and characterization of H2A.ZUb1-enriched and H2A.ZUb1-depleted promoters by gene-specific ChIP.

Finally, we selected subsets of H2A.ZUb1-enriched or H2A.ZUb1-depleted promoters identified by ChIP-Seq to validate by gene-specific ChIP (12 promoters each). Specifically, we performed ChIPs using streptavidin- or Flag-coupled beads, and used qPCRs with primers corresponding to regions flanking the TSSs of the 24 tested genes for quantitative analyses (see Table S1 for primer sequences). We normalized the streptavidin signals over the Flag signals and plotted the individual gene data as

A


SA/FLAG
relative expression
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
DGKI TBPL2 PHOX2B MY5 CDH8 LY75 NHLRC1 CALB1 PAX6 FLN MX1 TAF1 YBX3 AK2 RPN1 HMGB2 XRCC6 ASNS TMEM150B ZMYM6NB TRIM21 GAPDH
SA-Enriched
SA-Depleted
H2A.Z/H3
relative expression
70
60
50
40
30
20
10
0
DGKI TBPL2 PHOX2B MY75 CDH8 LY75 NHLRC1 CALB1 PAX6 FLN MX1 TAF1 YBX3 AK2 RPN1 HMG2B XRCC6 ASNS TMEM150B ZMYM6NB TRIM GAPDH ATF4 RPL27A
SA-Enriched
SA-Depleted
H3K4me3/H3
relative expression
300
250
200
150
100
50
0
DGKI TBPL2 PHOX2B MYF5 CDH8 LY75 NHLRC1 CALB1 PAX6 FLN MX1 TAF1 YBX3 AK2 RPN1 HMG2B XRCC6 ASNS TMEM150B ZMYM6NB TRIM21 GAPDH ATF4 RPL27A
SA-Enriched
SA-depleted
H3K27me3/H3
relative expression
16
14
12
10
8
6
4
2
0
DGKI TBPL2 PHOX2B MYF5 CDH8 LY75 NHLRC1 CALB1 PAX6 FLN MX1 TAF1 YBX3 AK2 RPN1 HMG2B XRCC6 ASNS TMEM150B ZMYM6NB TRIM21 GAPDH ATF4 RPL27A
SA-Enriched
SA-Depleted

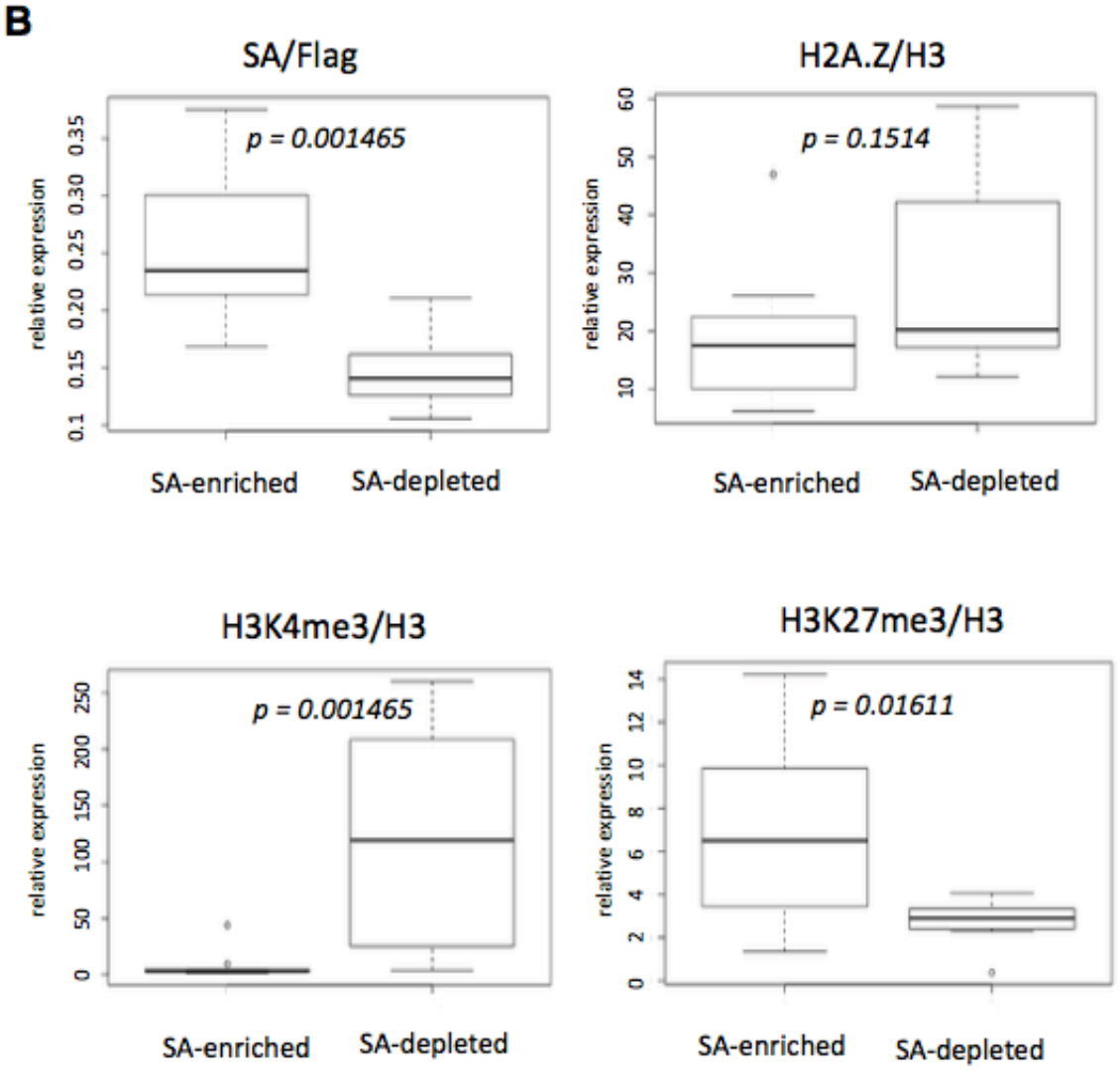


Figure 2-9. Gene-specific ChIP-qPCR of H2A.ZUb1-enriched and H2A.ZUb1-depleted promoters. (A) To calculate relative expression, all ChIP-signals were first normalized using the (1%) input method and then each event was divided by either Flag or H3 as indicated to further normalize each signal. (B) Box plot representations of same ChIP-qPCRs. *P*-values indicate Wilcoxon signed rank test, SA (streptavidin)-enriched or SA-depleted corresponds to H2A.ZUb1-enrichment or H2A.ZUb1-depletion as predicted by ChIP-Seq analysis.

shown in Figure 2-9A, and then combined the ChIP data of the two groups of genes for Box plot presentation and statistical analyses (Fig. 2-9B). In addition, we also examined the H3K4me3, H3K27me3, total H3 and total H2A.Z levels at the same promoters by ChIP analyses as well (Fig. 2-9). Although the ChIP data when viewed as individual genes are somewhat variable, once data from all H2A.ZUb1-enriched and H2A.ZUb1-depleted promoters are grouped together, we were able to see clear and statistically relevant trends. First, as expected, the streptavidin-signal (streptavidin affinity-purification normalized to Flag-purification signal) is significantly higher for the H2A.ZUb1-enriched group compared to the H2A.ZUb1-depleted group, indicating an enrichment of H2A.ZUb1 at those promoters (Fig. 2-9B). In contrast, the H2A.Z ChIP (using an antibody that recognizes both non-ubiquitylated and ubiquitylated H2A.Z) showed that there are statistically similar amounts of total H2A.Z at the promoters of both groups of genes, indicating that the differential streptavidin-signal seen between the two groups is not simply due to different amounts of H2A.Z at those promoters. Also, consistent with the histone PTM trends observed by our mononucleosome-AP-Western blots, the H2A.ZUb1-enriched promoters are hyper-methylated for H3K27 but hypo-methylated for H3K4, and the reverse trend is seen at the H2A.ZUb1-depleted promoters. From these data, the most striking feature of the H2A.ZUb1-enriched promoters is the consistent hypo-methylation of H3K4, which suggests a possible incompatibility or antagonism between H2A.Z ubiquitylation and H3K4 methylation.

2.4 Discussion

2.4.1 Biochemical purification of H2A.ZUb1-mononucleosomes

Our development of this purification system allowed us to circumvent reliance on a conventional antibody system for detection and enrichment of H2A.ZUb1 for which there is no existing antibody. Moreover, while antibodies work well for many histone modifications (e.g. acetylation, methylation), monoubiquitin is a more challenging epitope for antibody development. This is because the C-terminus of ubiquitin must be recognized in addition to the isopeptide bond that links it to the histone, as well as an epitope on the histone surface. Antibodies do exist which specifically detect H2AUb and H2BUb; however, their generation necessitates intensive effort and screening many mono-clonal hybridomas. Our use of the BirA-biotinylation method in contrast allowed us to purify monoubiquitylated H2A.Z with much greater expediency. Additionally, our system is advantageous in that the non-covalent bond between biotin and the AviTag provides the high affinity interaction required to purify sufficient mononucleosomes for biochemical analyses. Our selective enrichment of nucleosome-incorporated monoubiquitylated H2A.Z-Flag-BirA is confirmed by the pull-down of mononucleosomes containing stoichiometric amounts of core histones, which almost exclusively contain Avi-ubiquitylated H2A.Z-FB (Fig. 2-4). In contrast, Flag pull-downs recover mostly non-ubiquitylated H2A.Z-FB. At the same time, a potential disadvantage of our system is that it relies on an intramolecular reaction that is not exclusive to H2A.Z and hence could be susceptible to background biotinylation of Avi-tagged ubiquitin. For example, we found that H2A.Z-FB could potentially reach over and biotinylate Avi-ubiquitylated endogenous histones albeit at much lower efficiency (e.g. the abundance of the top band in comparison to the lower band detected by Avi-conjugates in Fig. 2-2). As a result, minor amounts of endogenous ubiquitylated histones are co-purified with

H2A.ZUb1. Nevertheless, as demonstrated by comparing streptavidin pull-downs of wildtype H2A.Z-FB and mutant H2A.Z-K3R3-FB, we observe increased binding of CTCF and SMC1 in proportion to the level of Avi-ubiquitylated H2A.Z-FB, suggesting that the binding proteins analyzed are selectively co-purified with H2A.ZUb1. Together, these support the efficacy of our system in selectively isolating mononucleosomes containing monoubiquitylated H2A.Z.

2.4.2 Characterization of H2A.Z (Ub versus non-Ub)-mononucleosomes

Using our system to compare H2A.ZUb1-mononucleosomes with equivalent amounts of Flag-IP'd H2A.Z (mostly non-ubiquitylated with minor amounts of ubiquitylated H2A.Z), we find that H2A.ZUb1 is preferentially linked to histone PTM hallmarks for transcriptional silencing whereas non-ubiquitylated H2A.Z is enriched with hallmarks of active transcription (i.e. H3K4 methylation, H3K27Ac, H4 acetylation, Fig. 2-5). The observed co-enrichment of H3K27me3 with H2A.ZUb1 is consistent with the link between H2A.Z ubiquitylation by the PRC1 complex RING1b E3 ligase previously reported by our lab (Sarcinella et al., 2007). Together, these results are in agreement with previous studies linking H2A.ZUb1 to transcriptional silencing (Draker et al., 2011; Sarcinella et al., 2007; Surface et al., 2016) and importantly, they validate that distinct combinations of histone PTMs co-exist with H2A.ZUb1 in the nucleosome context, suggesting the possibility of combinatorial effects and PTM crosstalk. This is further supported by the patterns visualized by Western blot analysis of a subset of co-purified binding proteins with previous links to H2A.Z (Fig. 2-6). For example, we find that transcriptional activators show either no enrichment (i.e. MLL) or are depleted on

H2A.ZUb1-mononucleosomes (i.e. Brd2). Alternatively, factors associated with silencing such as LSD1 and DNMT3L are enriched on H2A.ZUb1-nucleosomes. Interestingly, we observe the co-enrichment of H2A.ZUb1 with CTCF and cohesin components, which play an important role in establishing and maintaining the 3D architecture of the genome (reviewed in Ghirlando and Felsenfeld, 2016). The possible functional connections between the observed histone PTMs and the pattern of these binding factors are discussed in greater detail at the end of this section and in Figure 2-10.

2.4.3 Genome-wide analysis of H2A.ZUb1

To further interrogate the role of H2A.ZUb1, we compared the genome-wide distributions of H2A.Ub1 and total H2A.Z by performing ChIP-Seq using HEK293T cells (Fig. 2-7 and 2-8). For both streptavidin- and Flag-ChIP samples, H2A.Z-nucleosomes appear to flank the TSS, as demonstrated by the characteristic ‘double’ peak. This result also confirms that the BirA fusion to H2A.Z does not interfere with its targeting to the known genome-wide localization of H2A.Z. Additionally, we observe more total H2A.Z at the promoters of active genes, which fits with previous literature linking bulk H2A.Z with transcriptional competency and activity (Giaimo et al., 2019). Unexpectedly, we find roughly equal amounts of streptavidin-ChIP peaks at the promoters of both expressed and silent genes, suggesting that H2A.ZUb1 is not exclusively localized around inactive TSS, but is present around active promoters as well. However, since there are significantly higher levels of total H2A.Z at inactive promoters, this could indicate that it is the ratio of non-ubiquitylated H2A.Z to ubiquitylated H2A.Z at the TSS that regulates transcriptional activity. This possibility is supported by our comparison of

streptavidin-ChIP to Flag-ChIP at each peak and use of statistical analyses to categorize peaks as being H2A.ZUb1-enriched or H2A.ZUb1-depleted. By comparing these two groups, we are able to establish robust differences in their occurrence at the TSS, whereby H2A.ZUb1 is enriched at the promoter region of silent genes and is depleted at active promoters (and to a lesser-extent, is slightly enriched within inactive gene bodies). Notably, by using ChIP-qPCR to compare a subset of H2A.ZUb1-enriched or H2A.ZUb1-depleted promoter regions identified by our ChIP-Seq analysis, we were able to confirm the respective enrichment of H3K27me3 or H3K4me3 at the promoters of these two groups of genes (Fig. 2-8). Moreover, this pattern exactly mirrors our biochemical analyses examining the PTM patterns seen on the bulk mono-nucleosomes purified from SA- vs. Flag-IPs (Fig. 2-5). Cumulatively, these data consistently support the association of H2A.ZUb1 with the inactive state and non-ubiquitylated H2A.Z with gene expression, and moreover, further suggest that it is the balance of ubiquitylated and non-ubiquitylated H2A.Z that links this histone variant to different transcriptional states.

2.4.4 Possible links between H2A.ZUb1 and bivalency

In ES cells, H2A.Z occupies bivalent promoters, which are those characterized by TSS flanking-nucleosomes modified by both H3K27me3 and H3K4me3 (Bernstein et al., 2006; Creyghton et al., 2008). Bivalent promoters are highly enriched in pluripotent cells relative to differentiated cells and maintain developmentally important genes in a transcriptionally silent but poised state. H2A.Z has been shown to play an active role in the establishment of bivalent domains through its mutual interdependent recruitment of

PRC1 and PRC2 (Creyghton et al., 2008; Wang et al., 2018). Both H2AK119Ub1 and H2A.ZUb1 have also been localized to bivalent promoters (Endoh et al., 2012; de Napoles et al., 2004; Surface et al., 2016). H2AK119Ub1 contributes to the maintenance of PRC2 at bivalent domains and prevents the elongation of paused RNA polymerase, creating a transcriptional state primed for rapid activation (Blackledge et al., 2014; Stock et al., 2007; Zhou et al., 2008). More recently, H2A.ZUb1 has also been proposed to function by supporting PRC2 activity, and also found to antagonize the binding of transcriptional co-activator Brd2 at these sites (Surface et al., 2016).

How bivalent domains are resolved into states of transcriptional activity or stable repression marked by the removal of H3K27me3 or H3K4me3, respectively, remains poorly understood (Gao et al., 2018). The majority of bivalent genes are progressively and stably silenced during differentiation by PcG proteins. Interestingly, we have found that in differentiated cells, H2A.ZUb1-nucleosomes are depleted of H3K4 methylation marks and preferentially interact with LSD1, an H3K4-demethylase that functions within the larger NuRD chromatin remodeling and histone deacetylase complex (Whyte et al., 2012). The LSD1-NuRD complex has been implicated in the control of early embryogenesis by maintaining bivalent promoters through control of H3K4me1/2 levels (which can influence the equilibrium of H3K4me3), as well as directly decommissioning of active enhancers (Adamo et al., 2011; Whyte et al., 2012). The NuRD complex also possesses histone deacetylase function from its HDAC1 and HDAC2 (histone deacetylase 1 and histone deacetylase 2) subunits, and has been shown to functionally overlap with Polycomb proteins. Specifically, the HDAC activity of NuRD is required at bivalent genes for the deacetylation of H3K27 and the stable association of PRC2 and

subsequent deposition of H3K27me3 in ES cells (Hu and Wade, 2012; Reynolds et al., 2012). Given that we find H2A.ZUb1-nucleosomes to be hypoacetylated, depleted of H3K4 methylation and enriched for H3K27me3, it is plausible that the interaction we observe between H2A.ZUb1-enriched nucleosomes and LSD1 reflects a functional relationship that not only results in the de-methylation of H3K4, but also in the de-acetylation of H3K27 through the NuRD complex. This interaction could subsequently support H3K27 trimethylation at nucleosomes enriched with H2A.ZUb1. In this manner, H2A.ZUb1 could play an active role in establishing repressive chromatin domains not only by antagonizing the binding of Brd2, but also by promoting the association of LSD1 and possibly other co-repressors.

2.4.5 Possible links between H2A.ZUb1 and DNA methylation

We find further links between H2A.ZUb1 and transcriptional repression through the preferential binding of DNMT3L to H2A.ZUb1-nucleosomes in comparison to those containing non-ubiquitylatable H2A.Z. DNMT3L itself does not possess DNA methyltransferase activity but regulates *de novo* DNA methylation by stimulating the enzymatic activities of DNMT3A and DNMT3B (Chédin et al., 2002; Suetake et al., 2004), and has been shown to target chromatin through its H3K4me0 (i.e. unmethylated H3K4)-binding ADD domain (Argentaro et al., 2007; Ooi et al., 2007; Otani et al., 2009). Cytosine methylation and Polycomb complexes represent two major repressive pathways that primarily act on non-overlapping sets of genes with some notable exceptions such as their colocalization at the inactive X chromosome (Pinheiro and Heard, 2017; Viré et al., 2006). Our finding that H2A.ZUb1 localizes to the inactive X

chromosome (Sarcinella et al., 2007) and promotes the chromatin-association of DNMT3L (current study) could suggest that H2A.ZUb1 is able to regulate *de novo* methylation in certain contexts**.** Although presence of H2A.Z and DNA methylation is generally found to anticorrelate around TSSs and gene bodies (Conerly et al., 2010), H2A.ZUb1 could be linked to *de novo* DNA methylation through the Polycomb pathway at discrete sites in the genome. PRC2 has also been shown to direct cytosine methylation at a subset of genes by directly binding DNMT3A, DNMT3B, as well as the maintenance methyltransferase DNMT1, which restores nascent strand DNA methylation at hemimethylated CpGs following DNA replication (Viré et al., 2006). More recently, DNMT1 was recovered as one of the most enriched proteins pulled-down with H2AK119Ub1 (Kalb et al., 2014). At the same time, FBXL10 (also known as KDM2B), a subunit of the non-canonical PRC1.1 complex, has been shown to directly antagonize *de novo* methylation by binding unmethylated CpGs at loci bound by both PRC1 and PRC2 (Boulard et al., 2015). Moreover, recruitment of FBXL10 to bivalent promoters has been shown to be impaired in ES cells expressing non-ubiquitylatable H2A.Z (Surface et al., 2016). It is therefore possible that Polycomb-directed DNA methylation could outcompete FBXL10 when DNA methyltransferases are stabilized through multiple contacts that include H2A.ZUb1-nucleosomes. One hypothesis is that involvement of H2A.ZUb1 in DNA methylation could function to reinforce the stable repression of PcG domains, whereas de-ubiquitylation of H2A.Z could alternatively contribute to loss of DNA methylation. Interestingly, DNMT1 activity has been shown to be dependent on UHRF1-mediated ubiquitylation of histone H3. Specifically, the direct binding of DNMT1 to monoubiquitylated H3 is essential for the faithful transmission of

cytosine methylation following DNA replication (Nishiyama et al., 2013; Qin et al., 2015). It is tempting to speculate that recognition of monoubiquitylated H2A.Z by DNA methyltransferases could likewise occur to ensure the fidelity of DNA methylation outside of replication.

In differentiated cells, DNA methylation has been shown to antagonize the spread of Polycomb domains and ectopic recruitment of PcG proteins (Lynch et al., 2012; Reddington et al., 2014). Therefore, *de novo* DNA methylation within the Polycomb pathway could also serve a negative-feedback function to restrain the spread of PcG domains. To this end, another possibility is that *de novo* DNA methylation could be promoted through H2A.ZUb1 as part of a pathway to generate cytosine methylation intermediates. For example, the TET family of dioxygenases which catalyze the conversion of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) are recruited to bivalent genes through PRC2 and overlap with Polycomb-repressed domains (Putiri et al., 2014; Williams et al., 2011). 5hmC has been suggested to protect against aberrant DNA hypermethylation while contributing to the repressive state through its interactions with co-repressors such as NuRD and Sin3A (Chandru et al., 2018; Williams et al., 2011; Yildirim et al., 2011). At the same time, 5hmC has also been proposed to promote DNA methylation, presumably through in/direct influences on local chromatin structure (Putiri et al., 2014). It will therefore be of great interest to determine if H2A.ZUb1 is functionally linked to catalysis of 5mC and if so, whether it plays a role in the turnover of methylated cytosines.

2.4.6 Possible links between H2A.ZUb1 and architectural proteins

Polycomb-target genes are also regulated at the level of higher-order chromatin structure. For example, Polycomb genes exist within topologically associating domains (TADs) and these can further interact to form larger domains visualized microscopically as Polycomb bodies or foci (Entrevan et al., 2016; Pirrotta and Li, 2012; Wani et al., 2016). TADs are sub-megabased sized, self-interacting chromatin folds which, in general, confine active versus inactive domains to create insulated neighbourhoods of similar transcriptional activity (Dixon et al., 2012; Nora et al., 2012). Two architectural proteins, CTCF and cohesin, primarily form and maintain the boundaries of TADs, as well as the looping events that occur between promoters and enhancers, and other long-range chromatin interactions. CTCF is an 11-zinc-finger, sequence-specific DNA-binding protein, and cohesin is a ring-shaped complex comprising the core subunits SMC1, SMC3, SCC1 (Rad21), and SA1/SA2 (Parelho et al., 2008; Rao et al., 2014; Rubio et al., 2008; Stedman et al., 2008; Vietri Rudan et al., 2015; Wang et al., 2012). CTCF binding motifs are non-palindromic DNA sequences and CTCF homodimerizes when two specific DNA binding sites are convergently oriented. The bases of these loops are joined and further stabilized by a pair of cohesin molecules (Ghirlando and Felsenfeld, 2016).

Interestingly, we have found that CTCF and cohesin selectively engage H2A.ZUb1-nucleosomes. It has previously been shown that H2A.Z-nucleosomes flank CTCF binding motifs and is likely important for maintaining a nucleosome-free region permissive to CTCF engagement (Barski et al., 2007; Henikoff, 2009; Jin and Felsenfeld, 2007). At the same time, PRC1 is able to self-polymerize through the sterile alpha motif (SAM) within its Polyhomeotic (Ph) domain (Blackledge et al., 2015; Isono et

al., 2013). It is possible that by associating with CTCF/cohesin as well as PRC1, H2A.ZUb1-nucleosomes are uniquely positioned to regulate long-range chromatin interactions on multiple scales. Certain long-range contacts such as those with barrier function could be strengthened by H2A.ZUb1-nucleosome interactions with CTCF/cohesin with the corollary effect of weakening those mediated through oligomerization of PRC1 or vice versa. If true, H2A.ZUb1-nucleosomes could contribute to plasticity of silent chromatin by ensuring that physical domains do not become fixed.

At the same time, subsets of CTCF binding sites display sensitivity to 5-methylcytosine or 5-hydroxymethylcytosine (Maurano et al., 2015; Wang et al., 2012; Yu et al., 2012). Our finding that H2A.ZUb1-nucleosomes also selectively engage DNMT3L may therefore suggest H2A.ZUb1 could also regulate CTCF binding through methylation turnover in certain contexts and potentially disrupt the CTCF binding by directing *de novo* DNA methylation. At the same time, H2A.Z has been suggested to bookmark CTCF-bound sites during cell division (Oomen et al., 2018) and given that Polycomb patterns could persist through mitosis through self-propagation (Reinberg and Vales, 2018), it is possible that monoubiquitylation of H2A.ZUb1 could also ensure the robust inheritance of organized domains.

In an alternative scenario, monoubiquitylation of H2A.Z could be linked to CTCF and cohesin at promoters that display unidirectional transcription. For example, CTCF and cohesin are found to occupy the upstream TSS nucleosome in unidirectional genes but not in bidirectional ones, and their occupancy has been suggested to function in blocking initiation of antisense transcription (Bagchi and Iyer, 2016; Bornelöv et al., 2015). H2A.Z is present within nucleosomes immediately upstream (the -1 nucleosome)

and downstream (the +1 nucleosome) of the TSS of most promoters where it is presumed to contribute to nucleosome-depletion and hence permissiveness to RNAPII loading (Jin et al., 2009; Weber et al., 2014b). At the +1 nucleosome, H2A.Z-nucleosomes also tend to contain histone variant H3.3 and such dual-histone variant nucleosomes serve to enhance RNAPII processivity through multiple mechanisms including their increased rate of turnover (Weber et al., 2014b). It is therefore tempting hypothesize that monoubiquitylation of H2A.Z at promoters could preferentially occur at the nucleosome immediately upstream of the TSS where it could direct CTCF/cohesin-complexes to supress divergent transcription. Additionally, H2A.ZUb1 at the -1 nucleosome could also serve to increase the activation energy barrier presented to RNAPII through its antagonism of Brd2 and association with other co-repressors (Fig. 2-9). This could also coincide with our observation that low levels of H2A.ZUb1 tend to overlap with H2A.Z-enriched sites globally regardless of transcription-level.

2.4.7 Concluding comments

Together, our finding that both H2A.ZUb1-nuclesomes as well as total-H2A.Z-containing nucleosomes are present at high levels within the promoter region and body of genes that are downregulated or upregulated in HEK293T in comparison to ES cells could suggest that at the latter loci, H2A.ZUb1-nucleosomes could serve to keep transcription levels restrained through its antagonism of BRD2 as well as its association with LSD1. In support of this hypothesis, we find that H2A.Z-nucleosomes that are enriched for monoubiquitylation are associated with genes which show a more constrained pattern of expression across multiple cell types, while in contrast, H2A.Z-

nucleosomes that are depleted of monoubiquitylation are associated with genes that display a broader range of expression levels. Our finding that H2A.ZUb1-mononucleosomes are enriched for histone modifying proteins linked to inactive or repressive histone PTMs suggests that H2A.ZUb1 is a dynamically engaged PTM and future studies to resolve the causal nature of its co-existing modifications await more-targeted interrogation of potential crosstalk.

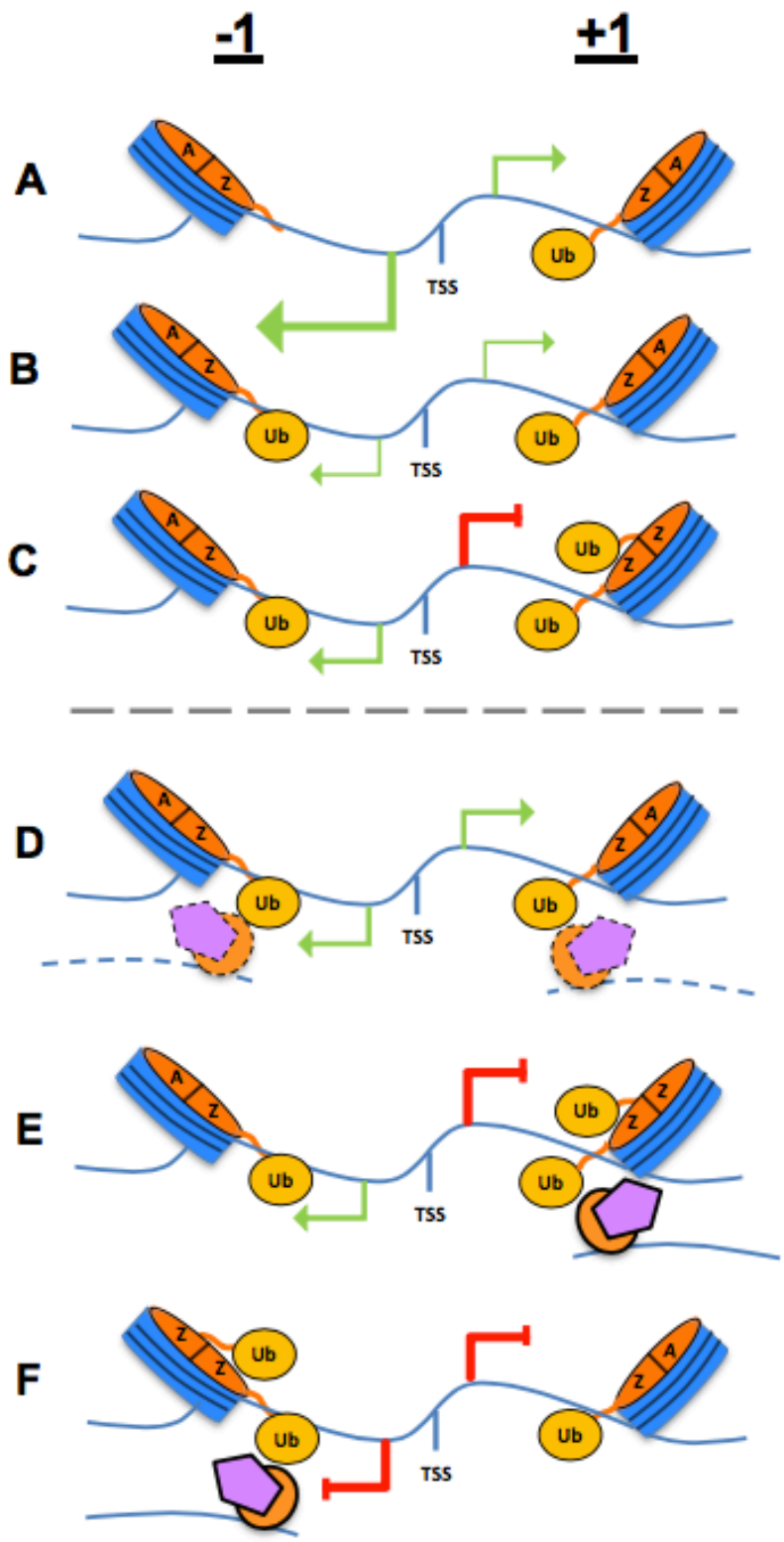


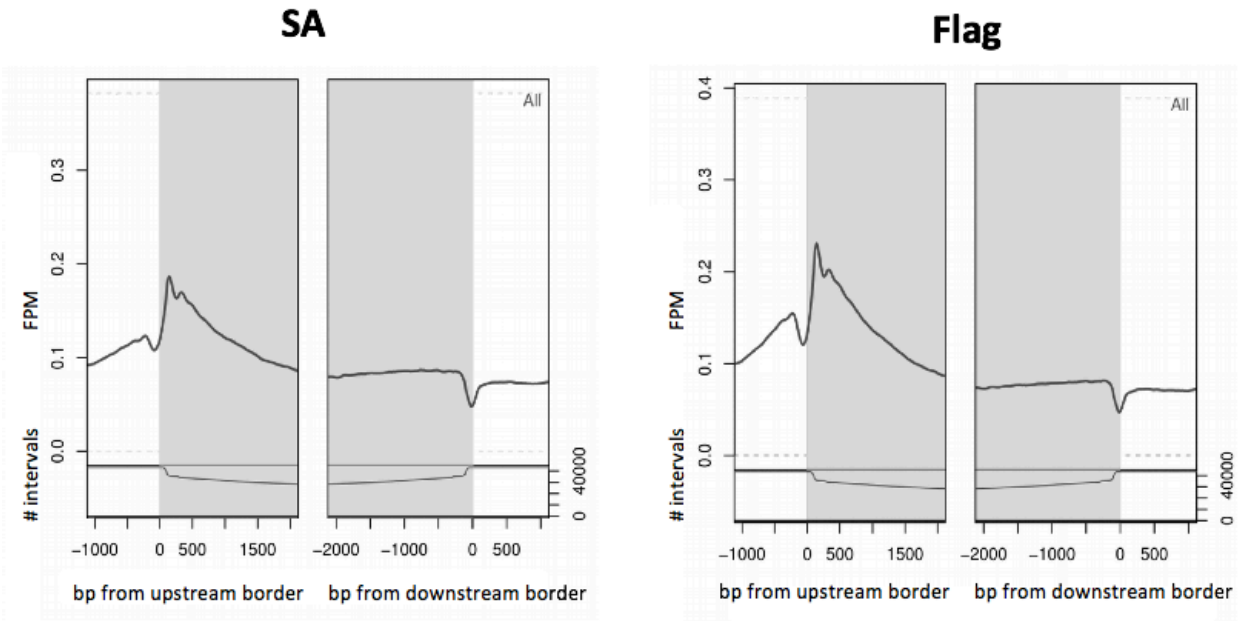
Figure 2-10. Models of H2A.ZUb1 at the promoter. A few hypothetical scenarios in which H2A.ZUb1 could contribute to transcriptional regulation at active promoters (where it is present but not enriched) or to transcriptional silencing at promoters (where it is enriched). TSS are flanked by strongly-positioned H2A.Z-containing nucleosomes. These nucleosomes modulate the permissibility of RNAPII. At many mRNA genes in mammals, RNAPII transcribes in both the sense and antisense directions. Whereas sense transcripts often lead to mature mRNAs, antisense transcripts are typically short-lived and play a role in modulating sense transcript pre-mRNA processing. (A) One copy of H2A.ZUb1 (heterotypic) could be present in the +1 nucleosome (which modulates sense transcription). Here H2A.ZUb1 could act to favour binding of transcriptional repressors and repel binding of transcription factors. In this state, transcription could still occur through the +1 nucleosome since the unmodified H2A or H2A.Z could still recruit positive transcription regulators. (B) The additional presence of H2A.ZUb1 at the -1 nucleosome could also have the same effect, and could negatively impact transcription in the sense direction by lowering the concentration of bound transcription factors or increasing the concentration of repressors. (C) Homotypic nucleosomes comprised of two copies of H2A.ZUb1 could be refractory to transcription due to cumulative effects of the binding events described above. (D) CTCF/cohesin complexes (represented by pentagon and orange donut) could interact with H2A.ZUb1 heterotypic nucleosomes with low residency time and have a negative impact on transcription by interfering with RNAPII. Presence of H2A.ZUb1 at both -1 and +1 promoters could have an additive affect in retaining CTCF/cohesin at the promoter region. (E) Homotypic nucleosomes containing two copies of H2A.ZUb1 could increase residency time of CTCF/cohesin and prevent RNAPII transit. (F) CTCF/cohesin looping interactions at the -1 nucleosome could also result in inhibition of sense transcription through the +1 nucleosome by creating a local topology refractory to RNAPII binding.

2.6 Appendix

Table S1. Primer pairs used for gene-specific ChIP-qPCR

Promoter	Forward	Reverse
DGKI	GCTGCATTTAGGGGAGGGG	CGCCGACGAATTCACTCCAA
TBPL2	GGTAAGAGGGTAAGCGCGG	CCATCGCCCAACGCTTCTC
PHOX2B	GGCTTCCTATATACGGGCGG	GTACGCCGCAGGTAAGGAC
MYF5	CTCTCAGCAGGATGGACGTG	TATGCAGGAGCCGTCGTAGA
CDH8	GGGCCTCTTGCGTACAGAAT	GCGACACACAGCCTCTACAT
LY75	CCTGGTGGGTGGGTTCTATC	GGCTGGAATGGAGAAGTCGT
NHLRC1	GGCTGTCCGGGCATAAAACA	GTGTGTTCGTGTTTTCCGGT
CALB1	TTTGGAAGTGTGAGGACGCA	GGGCCTAGAAAGGCGAACTT
PAX6	AACTAGTCTTGCCGAGTGCG	GAGGAGGGGACAGGGTGATT
MX1	CCTTGAGGACCAAAAGCGAC	CCTCAGGTGATCCCTTGGC
TAF1	AGTGATCGTTCTGGGGGAGA	CTCAGTAGGCGAAACCAGCA
YBX3	TGTCGGTCCTTCCCCTACAT	TGGAAAATGCCTGCGTTTGG
AK2	CTAACTCAGACTGCCCCGAC	GCCGGAGATCTAGAAGCCCT
RPN1	GTCGCCCACACTCACCTG	CGTCCCGAGCTACCTCTTTC
HMG2B	TTGCCCTGCAAAACCGATTG	AGGTTCCCTGCCTTGACTTC
XRCC6	GTAAGCGGGCCGTTATCCAT	CTCCTCGGATTCGCACACTT
ASNS	GGACAGAAAGGTCCTTCCGC	GTGGAGGATGCGGTCTTCAG
TMEM150B	GAGCCTCCATACCCAACTCG	AGGTCACATAAGCACCGTGG
ZMYM6NB	AGTACACGCAGCACCGAA	GGTAGCTCAGGCGAGAGTCTT

TRIM21	AAACTCAGTAGCCCGTGGTC	AGCGTCTAGGTGTGGAGTGA
GAPDH	CCAACTTTCCCGCCTCTCAG	GGACCCTTACACGCTTGGAT
ATF4	GATTTGTGGCCTGCGGAAAC	GCTATGAATGGGGCCTCTGG
RPL27A	CGTGGCCGATACCTCGC	AGGGAGTGGATGACTAGGGG



Supplementary Figure 1. Average profiles of SA- and Flag-ChIP peaks across > 40,000 genes in annotation, including > 20,000 non-coding genes.

Chapter 3

A Role for Variant H2A.Z-1 in Alternative Splicing Regulation

3 A role for histone variant H2A.Z-1 in alternative splicing regulation

3.1 Introduction

Transcripts from at least 95% of mammalian multiexon genes are differentially spliced in one or more cell-types. This alternative splicing (AS) and the resulting isoform diversity generates proteins possessing distinct functions (Pan et al., 2008). Intron removal from pre-mRNA and the joining of exons is catalyzed by the spliceosome, a large complex comprising five small nuclear (sn)RNAs assembled into sub-complexes, known as small nuclear ribonucleoprotein particles (snRNPs; U1, U2, and the U4/U6.U5 tri-snRNP), and over 100 auxiliary proteins. Pre-mRNA splicing largely occurs co-transcriptionally, placing spliceosome machinery in spatiotemporal proximity to chromatin and chromatin binding proteins. Growing evidence suggests that splicing and chromatin regulation are intrinsically coupled (Alexander and Beggs, 2010; Goldstrohm et al., 2001; Luco et al., 2011). For example, at a global level, nucleosomes are non-randomly distributed across transcribed genes, displaying preferential positioning at exon-intron and intron-exon boundaries, and are enriched over (GC-rich) exons, which on average, coincide with the length of a single nucleosome (50 – 250 bp) (Schwartz et al., 2009; Spies et al., 2009; Tilgner et al., 2009). Chromatin can modulate splice-site selection through two non-mutually exclusive mechanisms – indirectly, by influencing the pausing and elongation rate of RNA polymerase II, or more directly, by binding (or repelling) ancillary factors which in turn recruit or stabilize assembly of the spliceosome on nascent transcripts (Herzel et al., 2017). Exonic nucleosomes present a barrier to RNAPII progression, creating an opportunity for the splicing reaction to occur. According to the kinetic model, transcription elongation rate can regulate splicing by modulating accessibility of *cis*-

acting sequences. For example, slow elongation by RNAPII can result in increased exon inclusion by allowing more time for the splicing machinery to recognize weak splice-sites demarcating an exon, whereas faster elongation by RNAPII can lead to skipping of the same exon by favoring usage of stronger, more distal splice-sites (Dujardin et al., 2013; Maslon et al., 2019; de la Mata et al., 2003).

At the same time, histone modifications, including post-translational modification (PTM) and variant histone substitution, occur non-randomly across genes and can further influence the rate of RNAPII progression by impacting chromatin structure (e.g. Weber et al., 2010). Histone modifications have also more recently been implicated in recruiting the spliceosome through adaptor proteins (Luco et al., 2011; Schwartz and Ast, 2010), and in some cases, have been suggested to modulate availability of splicing factors and bind directly to RNA (Soboleva et al., 2017).

Recently, it has been demonstrated that histone variant H2A.Z is required for efficient splicing of weak introns with non-consensus splice-sites in yeast (Neves et al., 2017; Nissen et al., 2017). H2A.Z shares ~ 60% amino acid identity with canonical histone H2A, and is ~ 90% conserved amongst eukaryotes (Thatcher and Gorovsky, 1994). The enrichment of H2A.Z at the promoters of most genes in yeast and higher eukaryotes suggests transcriptional regulation is a key activity of this variant (Subramanian et al., 2015). Vertebrates possess two non-allelic paralogs of H2A.Z, H2A.Z-1 and H2A.Z-2, which differ by three conserved amino acids (Eirín-López et al., 2009). Despite this limited sequence variation, mice that lack *H2AFZ* (H2A.Z-1) but maintain *H2AFV* (H2A.Z-2) fail to develop in utero, indicating that these isoforms possess non-redundant activities (Faast et al., 2001). Consistent with divergent

functions, knockdown of H2A.Z-1 or H2A.Z-2 have been shown to affect the basal expression of mostly non-overlapping sets of genes in neurons and in melanoma cells (Dunn et al., 2017; Vardabasso et al., 2015). Whether H2A.Z isoforms are differentially linked to splicing, however, remains unexplored.

Previously, using an immunoprecipitation-mass spectrometry (IP-MS) approach, we identified chromatin binding proteins that preferentially associate with nucleosomes containing histone H2A.Z-1 over those containing H2A (Draker et al., 2012). Consistent with the transcription regulatory role of H2A.Z, gene ontology (GO) analysis revealed that many of these proteins possess putative transcription-associated functions, including Brd2, which is known to function in transcriptional activation. In addition to Brd2, our H2A.Z-nucleosome IP-MS screen also identified the Ser/Arg (SR) repeat-related protein USP39 (ubiquitin-specific peptidase 39; also known as human Sad1) as another factor that selectively binds H2A.Z-1 over H2A-containing nucleosomes. Proteins with SR repeats have diverse roles in the regulation and assembly of splicing complexes. USP39 is a component of the human U4/U6.U5 tri-snRNP complex, and has been shown to play an important role in the association of this snRNP particle with pre-spliceosomal complexes (Huang et al., 2014; Makarova et al., 2001). In this capacity, USP39 is one several proteins (others include the splicing factor SPF30, other SR proteins, the SR protein kinase SRPK2, and SART1) identified in higher eukaryotes which mediate stable integration of the U4/U6.U5 tri-snRNP into the nascent spliceosome (Makarova et al., 2001), a stage of spliceosome assembly still not well-understood.

Here, we show that USP39 selectively interacts with H2A.Z-1- over H2A.Z-2-mononucleosomes. In support of a direct coupling mechanism between H2A.Z-1 and pre-mRNA splicing, we map the capacity of H2A.Z-1 to interact with USP39 to its C-terminal isoform-specific residue. We further find that H2A.Z-1 and USP39 co-regulate an overlapping subset of alternative splicing events in HEK293T cells. Together, these data demonstrate an H2A.Z-1-specific role in alternative splicing through its interaction with the spliceosomal factor USP39.

3.2 Materials and methods

3.2.1 Cell culture, transfection, plasmids and antibodies

HEK293T cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum (DMEM; Wisent). Expression of H2A.Z-Flag constructs was carried out using polyethylenimine (PEI; Polysciences). Expression constructs were based on the pcDNA 3.1 (+) (Invitrogen) backbone with the Flag-tag cloned in-frame to the C-terminus of H2A.Z. To generate H2A.Z-1 mutants, amino acids T14, S38, or V127 were mutated to A14, T38, or A127, to introduce one H2A.Z-2-specific residue per construct. Conversely, H2A.Z-2 mutants were generated by mutating A14, T38, or A127 to T14, S38, or V127. The following antibodies were used for Western blot analysis: polyclonal Flag (Sigma; F7425), USP39 (Abcam; ab131332), SF3B1 (Abgent; AP13754A), PRP8 (Abcam; ab79237), H2A.Z (Active Motif; AM39113), PHF6 (Bethyl; A301-450A), H3 (Abcam; ab1791).

3.2.2 Mononucleosome co-immunoprecipitation

Mononucleosomes were generated as described previously (Draker et al., 2012). In brief, HEK293T cells were grown in 15 cm-diameter plates and transfected with constructs expressing Flag-tagged wild-type or mutant H2A.Z-1 or H2A.Z-2. Cells were trypsinized, counted, and washed in 1X PBS, 48 hrs following transfection. Cellular pellets were resuspended in buffer A (20mM HEPES, pH 7.5, 10mM KCl, 1.5mM $MgCl_2$, 0.34M sucrose, 10% glycerol, 1mM dithiothreitol, 5mM sodium butyrate, 10mM NEM, and protease inhibitors), pelleted and then resuspended in buffer A containing 0.2% Triton X-100 and incubated on ice for 5 min. The nuclear suspension was centrifuged at 600 x g; nuclei were then washed once in buffer A, then resuspended in cutting buffer (15 mM NaCl, 60 mM KCl, 10 mM Tris pH 7.5, 5mM sodium butyrate, 10mM NEM, and protease inhibitors) plus 2mM $CaCl_2$. Micrococcal nuclease (MNase; Worthington) was added at a concentration of 10 units/1.0 x 10^7 cells then incubated at 37°C for 30 min. The reaction was stopped by the addition of 20mM EGTA. The MNase-digested nuclei were centrifuged at 1300 x g. The resulting supernatant (S1) was saved and kept on ice. The digested nuclear pellet was subjected to hypotonic lysis by resuspension in TE buffer (10mM Tris-HCL, pH 8.0, 1mM EDTA). Samples were incubated on ice for 1 hr, with occasional mixing by pipette. The suspension was then centrifuged at 16 000 x g and the supernatant (S2) was transferred to a new tube. Salt was adjusted in S1 to 150mM NaCl by adding 2X buffer D (30 mM Tris pH 7.5, 225 mM NaCl, 3 mM $MgCl_2$, 20% glycerol, 0.4% Triton-X 100, 5mM sodium butyrate, 10mM NEM, and protease inhibitors) drop-wise, with constant mixing on a vortex set to low speed. S2 was also titrated to 150mM NaCl by the drop-wise addition of buffer E (60 mM HEPES pH 7.5, 450 mM NaCl, 4.5 mM $MgCl_2$, 0.6 mM EGTA, 0.6 % Triton-X 100, 30% glycerol, 5mM

sodium butyrate, 10mM NEM, and protease inhibitors). Insoluble material was pelleted via centrifugation. The clarified supernatants were combined and then used for affinity purification. Flag M2-agarose beads (Sigma) were added and incubated overnight at 4°C on an end- over-end rotator. Beads were washed 4 times in 1X buffer D, followed by 3 washes in 1X buffer D containing 0.5% Triton X-100. Proteins were eluted from the beads by resuspension in 2X SDS sample buffer and boiled for 10min. For Western blot analysis, samples were run on SDS-polyacrylamide electrophoresis gels according to standard practices.

3.2.3 siRNA knockdown and RNA analysis

HEK29T cells were transfected with 10nM of ON-TARGETplus siRNA SMART-pools (Dharmacon) targeting either *H2AFZ* or *Usp39* using Lipofectamine RNAiMAX (Invitrogen) and collected 72hr post-transfection. RNA was DNase-treated and isolated using an RNeasy kit (Qiagen). Semi-quantitative RT-PCR assays were performed using a OneStep RT-PCR kit (Qiagen) as per the manufacturer's recommendations, with exception of using 40ng of total RNA as input in 20uL reactions. The number of amplification cycles was 22 for *Gapdh* and 32 for all other transcripts analyzed. Cassette exon events were analyzed using sense and antisense primers designed to hybridize constitutive exons flanking the alternative exon. Reaction products were resolved in 3% agarose gels and relative isoform abundance was quantified by densitometry using Image Studio Lite.

3.2.4 RNA-Seq analysis (performed by Ulrich Braunschweig through collaboration with Dr. Benjamin Blencowe at U of Toronto)

RNA libraries were made by the Donnelly Sequencing Centre using Illumina TruSeq stranded mRNA sample preparation. Sequencing was also performed by the Donnelly Sequencing Centre. For analysis, we utilized a multi-modular pipeline called VAST-TOOLS to detect and quantify all major AS events in RNA-Seq reads, using the hg19 genome release (https://github.com/vastgroup/vast-tools). Details of this pipeline are published (Irimia et al., 2014; Tapial et al., 2017). VAST-TOOLS is able to detect and quantify AS events by assembling libraries of exon-exon junctions (EEJs) for subsequent alignment of RNA-Seq reads. To detect cassette exon events, three complementary modules are used to assemble EEJs: first, a “transcription-based module” that employs cufflink and aligns expressed sequence tags (ESTs) and cDNAs with genomic sequence; second, a “splice site-based module” that employs the joining of all hypothetically possible EEJ combinations from annotated splice sites; and third, a “microexon module” which searches for pairs of donor and acceptor splice sites in intronic sequences to detect very short (i.e. 3-15nt) microexons. Alternative 5’- and 3’- events (Alt3 and Alt5) are quantified based on the fraction of reads supporting the usage of alternative 5’- or 3’- splice sites. Intron retention events are detected using the pipeline recently described (Braunschweig et al., 2014), which uses a comprehensive set of reference exon-intron junctions (EIJs), intron midpoint sequences, and EEJs formed upon intron removal. Introns are classified as retained when there is a balanced accumulation of reads that map to the 5’ and 3’ EEJs as well as the midpoint sequence of the intron, relative to the EEJ sequence (Braunschweig et al., 2014; Irimia et al., 2014).

3.3 Results

3.3.1 Spliceosome components preferentially interact with H2A.Z-1

Previously, we used an IP-MS approach to identify proteins that preferentially bind to H2A.Z-1-containing nucleosomes over H2A-containing nucleosomes. Specifically, we had compared Flag-immunoprecipitated mononucleosomes from HEK293T cells expressing either Flag-tagged H2A.Z-1 or Flag-tagged H2A (Draker et al., 2012). Of the top hits identified (Fig. 3-1A), we chose to further investigate the interaction between USP39 and H2A.Z-1-nucleosomes because of the role USP39 plays in spliceosome assembly, and the links between H2A.Z and splicing suggested by its association with other spliceosome components (Draker et al., 2012). Due to the lack of information regarding the roles of H2A.Z-1 and H2A.Z-2 and how these paralogs may differ in activity, we first extended our comparison to include Flag-tagged H2A.Z-2. These experiments confirmed the preferential binding of USP39 to Flag-tagged H2A.Z-1-nucleosomes over Flag-tagged H2A-nucleosomes, using our previously described mononucleosome IP method followed by Western blot (Fig. 3-1B). Interestingly, while more USP39 binds to nucleosomes containing either H2A.Z-1 or H2A.Z-2 over those containing H2A, we also found a clear preference of USP39 for H2A.Z-1 over H2A.Z-2, suggesting that these almost identical isoforms are differentially associated with different amounts of USP39 (long exposure blot for USP39 on Fig. 3-1B). Our initial IP-MS screen also identified SF3B1, a component of the U2 snRNP, as a spliceosomal factor with a preference for interacting with H2A.Z-1-nucleosomes over H2A-nucleosomes, and this preferential interaction was also confirmed by IP-Western blot (Fig. 3-1B). However, while SF3B1 also clearly prefers H2A.Z-1 nucleosomes over

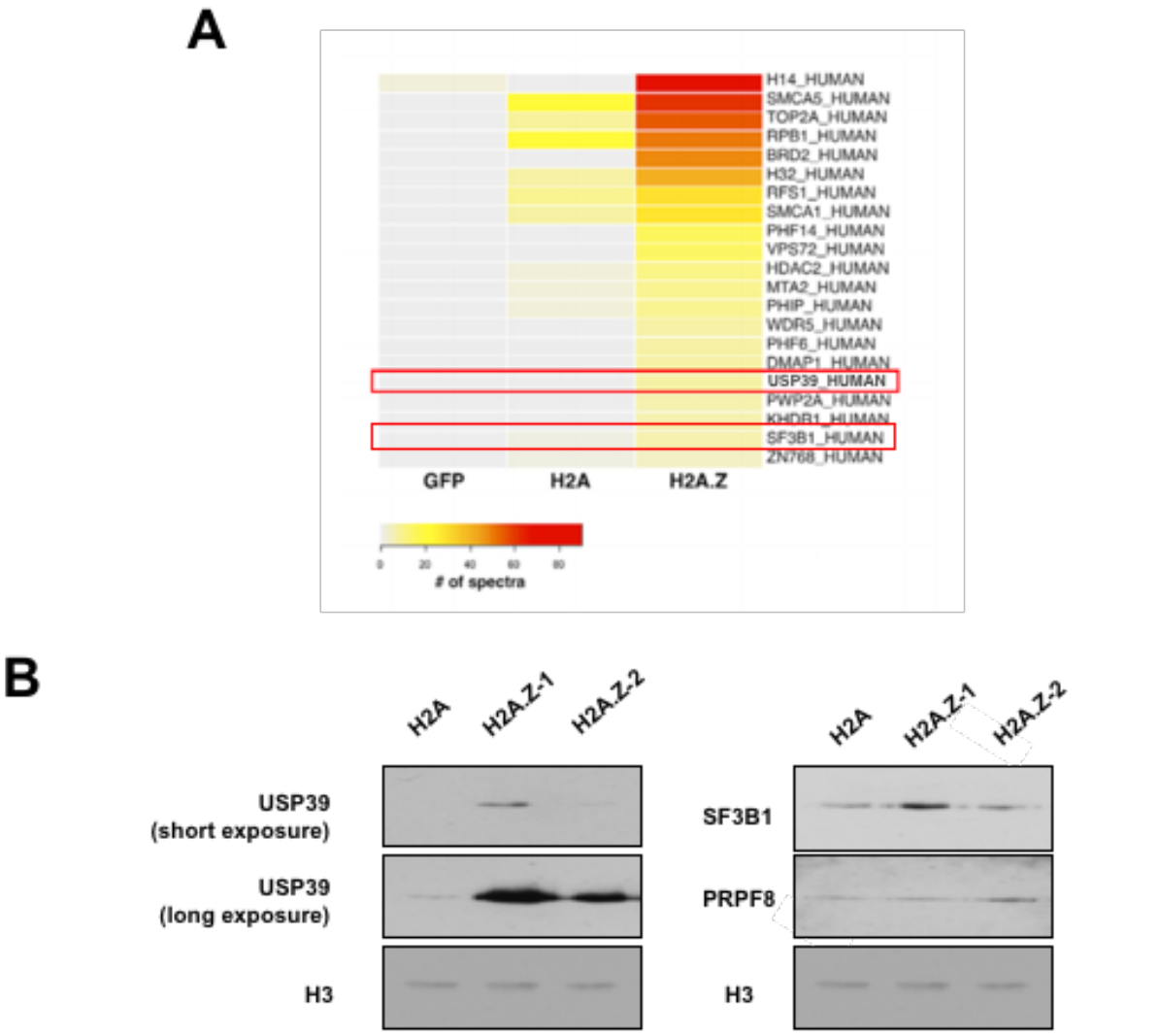


Figure 3-1. H2A.Z is enriched for spliceosomal factors. (A) Top hits of proteins identified by IP-MS to preferentially interact with H2A.Z- over H2A-mononucleosomes. (C) IP-Western blot showing the interaction of USP39 and (D) SF3B1 and PRPF8 amongst H2A.Z isoforms. Heatmap of IP-MS hits reprinted from Draker et. al 2012.

H2A.Z-2 nucleosomes, there was not difference in the amounts of SF3B1 binding to H2A.Z-2 or H2A nucleosomes. Lastly, we also tested the binding of PRPF8, a core component of U5 snRNPs, to H2A-, H2A.Z-1 and H2A.Z-2 nucleosomes. Unlike the other splicing factors, PRPF8 did not show any preference for any of these H2A/H2A.Z-containing nucleosomes. (Fig. 3-1B). Altogether, these results provide evidence that USP39 has a specific preference for mononucleosomes containing H2A.Z-1 over H2A.Z-2, but the pattern of binding and preference of other splicing factors to the different types of nucleosomes are variable and possibly mediated by additional factors.

3.3.2 A single amino acid difference between H2A.Z-1 over H2A.Z-2 confers a binding preference for USP39.

We next examined the cause of the differential binding of USP39 to H2A.Z-1 in comparison to H2A.Z-2 and asked whether any of the H2A.Z-1-specific residues is specifically important or responsible for mediating the preferential interaction of H2A.Z-1 for USP39. To this end, point mutants of H2A.Z-1 (T14A, S38T, V127A) and H2A.Z-2 (A14T, T38S, A127V) were generated which express one swapped isoform-specific residue at a time. Using our mononucleosome IP approach, we found that the preference of USP39 for H2A.Z-1- over H2A.Z-2-nucleosomes is abolished when the C-terminal-most amino acid of H2A.Z-1 is converted to that of H2A.Z-2 (i.e. V127A; Fig. 3-2A), and is conversely gained when the C-terminal residue of H2A.Z-2 is converted to that of H2A.Z-1 (i.e. A127V; Fig. 3-2B). Sequence diversity amongst the H2A family is most pronounced in their C-terminal tails suggesting that this region is an important determinant of functional distinction amongst the H2A variants (Ausió and Abbott, 2002), and this is consistent with a role for the C-terminus in mediating the differential

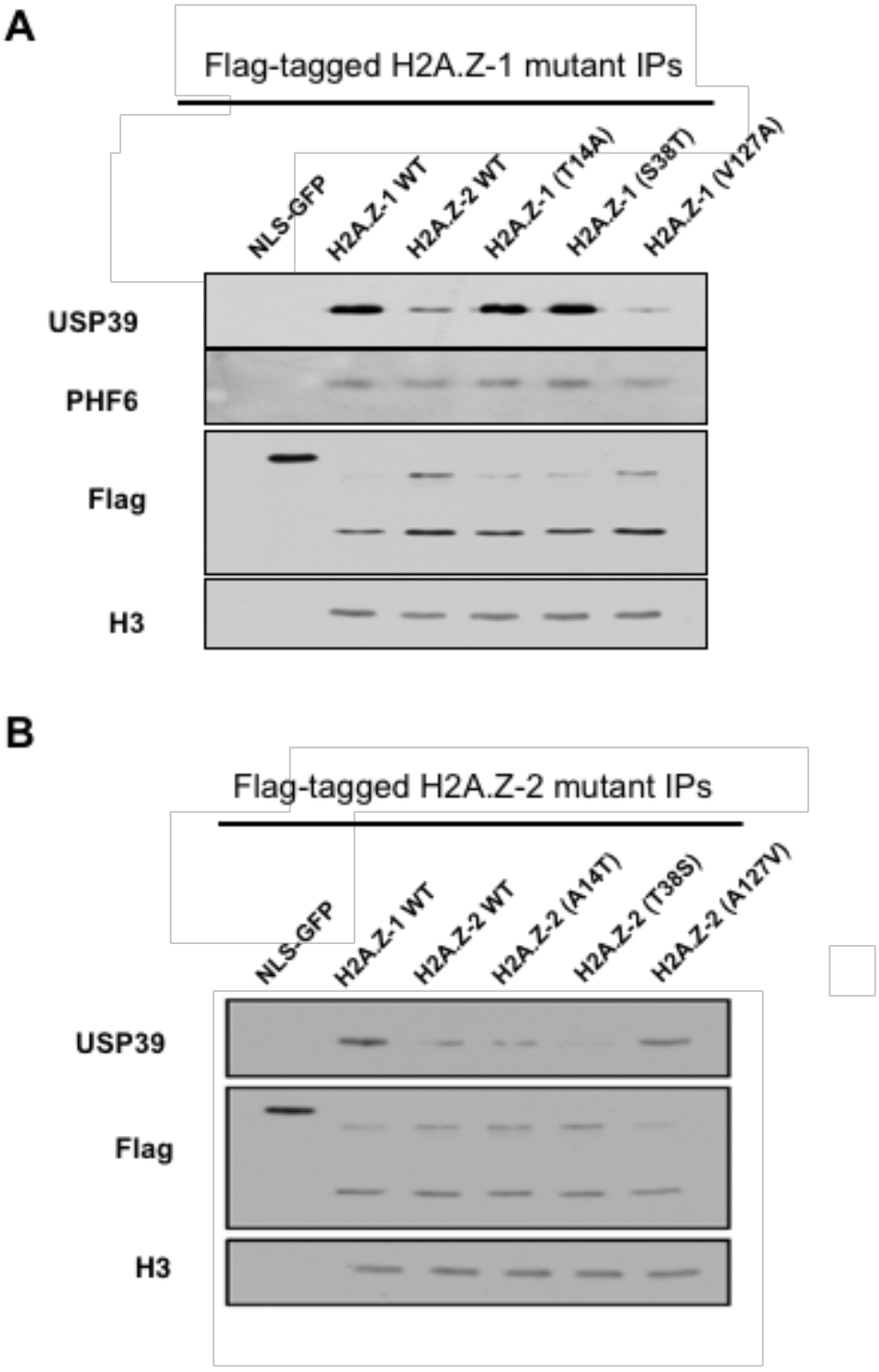


Figure 3-2. USP39 engages H2A.Z through its C-terminal isoform-specific residue. (A) mononucleosome IPs of H2A.Z-1 and H2A.Z-2, and H2A.Z-1 mutants. IPs were blotted for PHF6 to show specificity. (B) mononucleosome IPs of H2A.Z-1 and H2A.Z-2, and H2A.Z-2 mutants.

binding of USP39 between H2A.Z isoforms. To further validate the specificity of our H2A.Z mutants, we used another H2A.Z-specific binding protein, PHF6 to test for differential binding to the different point mutants. In contrast to USP39, PHF6 interacts similarly with both isoforms and accordingly it does not display any altered binding among the H2A.Z-1 mutants (Fig. 3-2A). Together, these results demonstrate that the preferential interaction of USP39 with H2A.Z-1 compared H2A.Z-2 is conferred by a single amino acid at the C-terminus of H2A.Z-1.

3.3.3 H2A.Z-1 and USP39 co-regulate a subset of alternative splicing events in human cells.

To address whether the physical interaction between H2A.Z-1 and USP39 may be functionally linked to pre-mRNA splicing, we performed RNA-Seq on HEK293T cells in which these factors were depleted by transfecting cells with siRNA pools targeting the corresponding transcripts (expressed from *H2AFZ* or *Usp39*; Fig. 3A). The RNA-Seq data were analyzed to identify changes in alternative splicing events on a genome-wide scale (by Dr. Ulrich Braunschweig at U of T). The RNA-Seq analysis pipeline we utilized (VAST-TOOLS) can detect and quantify all main classes of alternative splicing (AS) events (Irimia et al., 2014; Tapial et al., 2017). Using this pipeline, we generated quantitative estimates of “percent spliced in” (PSI) for alternative cassette exons (CE), microexons (MIC), alternative 5´ and 3´ splice site selection, and intron retention (IR) events and focused on significant (i.e. >10 PSI, $p < 0.05$; Fisher’s exact test) changes (Fig. 3b). In yeast, IR is the primary mechanism of AS and is disrupted in strains depleted of *HTZ1* (yeast H2A.Z) (Neves et al., 2017; Nissen et al., 2017). In contrast, IR relatively is a rare AS event in invertebrates and vertebrates, which instead more

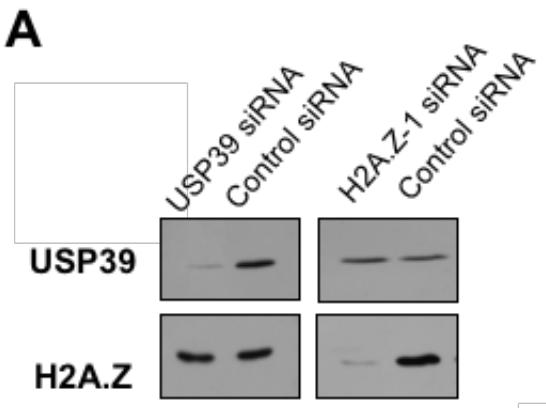
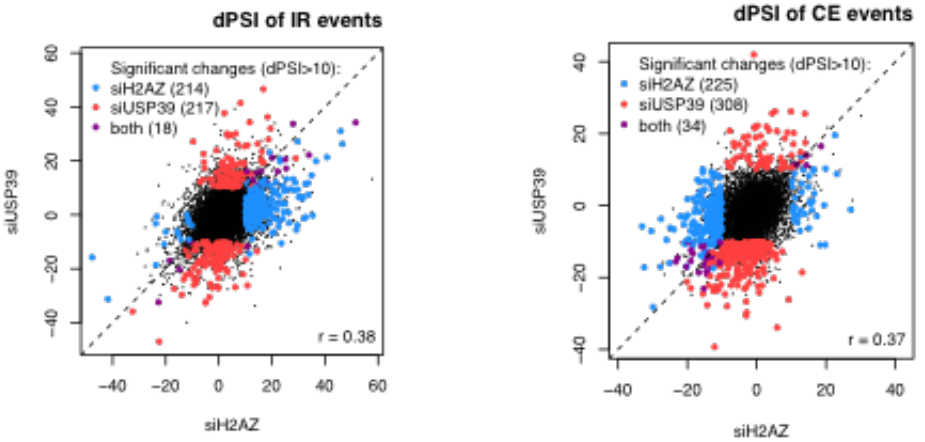
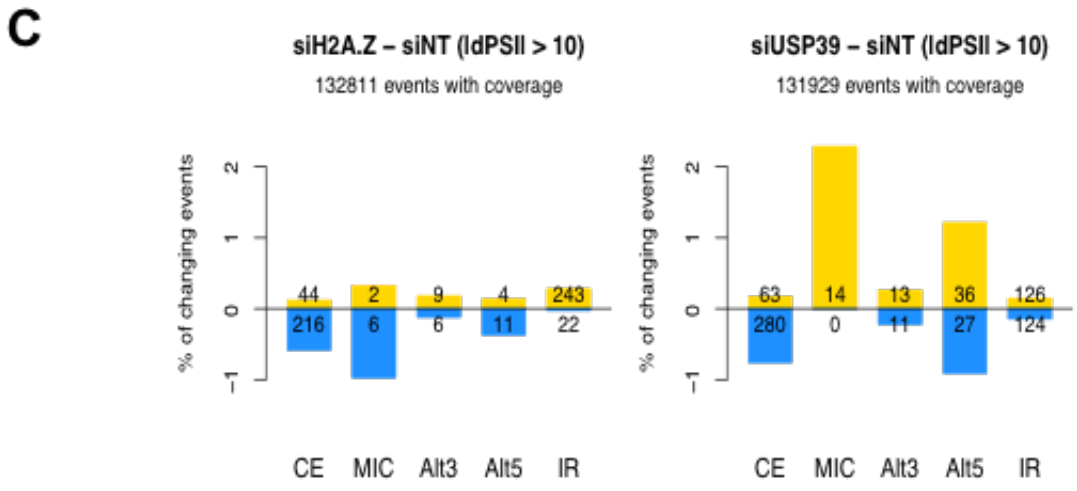


Figure 3-3. H2A.Z-1 and USP39 are functionally linked at a subset of alternative splicing events. (A) Western blot of whole-cell extracts showing the specificity of transiently transfected siRNA pools against *USP39* or *H2AFZ*. (B) Ccorrelation of alternative splicing changes (expressed as dPSI; change in percent-spliced-in [vs control]) between knockdowns categorized by alternative splicing type. CE, cassette exon; IR, intron retention. Differential events are highlightA group of 34 CEs display consistent and significant (>10 PSI, p<0.05) changes in both USP39- and H2A.Z-1-depleted cells (left plot) and 18 overlapping introns are affected (right plot). (C) Number of AS events by type, in cells expressing either *H2AFZ* or *USP39* siRNA pools, with increasing (yellow) and decreasing (blue) inclusion, as a percentage of all events that could be measured.





commonly utilize alternative CEs (Kim et al., 2008). Upon knockdown of either H2A.Z-1 or USP39, we observe modest but significant changes in both CE and IR which comprise both skipping and inclusion events. Although correlations between CE and IR events shared by H2A.Z-1 and USP39 are weak (r = 0.37 or 0.38, respectively), these overlaps were found to be significant. We also observed an enrichment of IR events for intron skipping in *H2AFZ* depleted cells, which was not seen upon knockdown of *USP39* (where intron retention and skipping events where found to be roughly equal) (Fig. 3C).

Finally, in order to confirm these patterns seen in genome-wide analyses, I performed RT-PCR analyses of a representative set of 20 CE events using RNA harvested from the siRNA treated cells analyzed by RNA-Seq (Fig. 3-4). For this, we chose events that were found to be affected by either H2A.Z depletion, USP39 depletion, or both. Significantly, we find a modest correlation between the data analyzed by RNA-Seq and by RT-PCR by performing linear regression analysis ($r^2 = 0.6124$).

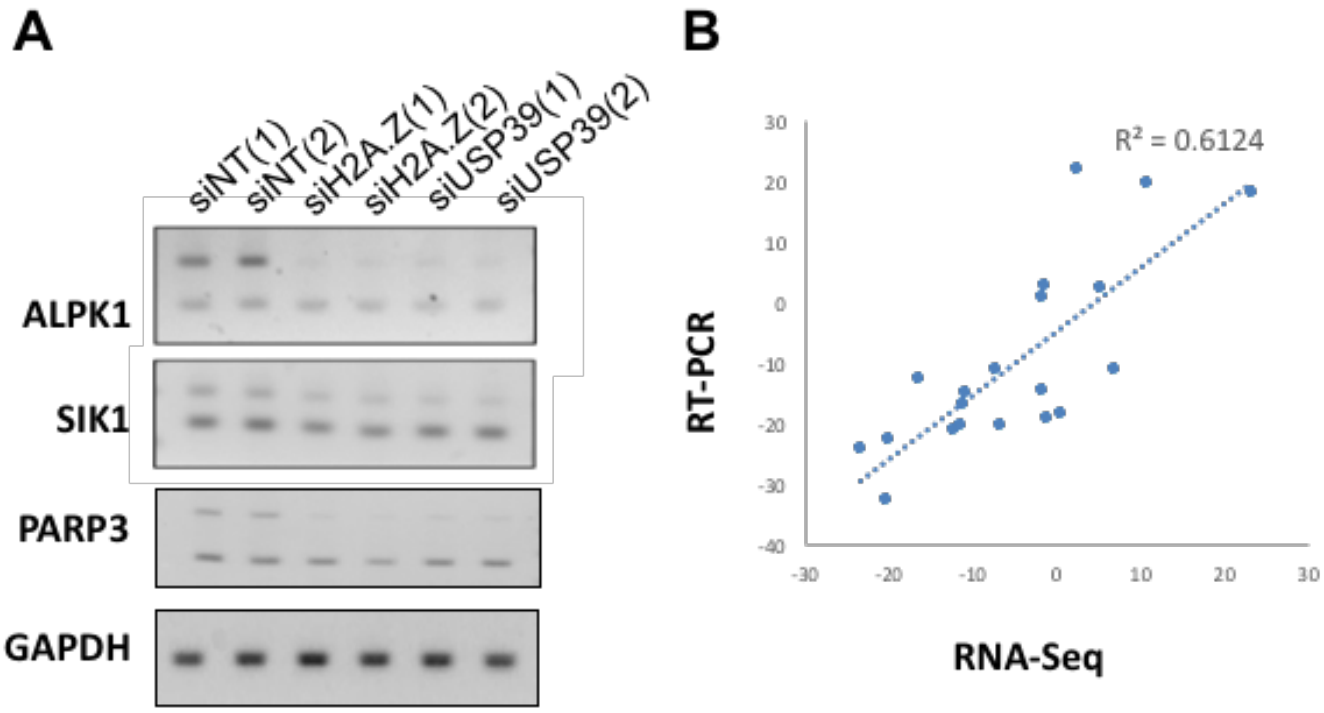


Figure 3-4. Correlation between splicing changes in USP39 and H2A.Z-1 knockdown cells assessed by RNA-Sequencing and RT- PCR. (A) A sample from the 20 candidate events, identified by RNA- Sequencing and validated by RT-PCR performed in replicate, which are regulated by both USP39 and H2A.Z-1. (B) Scatter plot of RT-PCR dPSIs (change in percent spliced-in) versus dPSIs obtained by densiometric analysis of RT-PCR samples.

3.4 **Discussion**

Together, the results from co-IP and RNA-Seq analysis indicate that the U4/U6.U5 tri-snRNP component USP39 specifically interacts with H2A.Z-1 in a manner dependent on its C-terminal isoform-specific residue and further suggest that this interaction reflects a coupling mechanism between USP39 and H2A.Z-1 that is responsible for the co-regulation of a subset of alternative splicing events in human cells.

The spliceosome catalyzes the excision of non-coding introns and ligates exons through a two-step transesterification reaction (Fig. 3-5). For each intron, this multi-megadalton ribonucleoprotein complex is assembled *de novo* from its five canonical snRNPs and various other non-snRNP factors (Will and Luhrmann, 2011). This process begins with recognition of the 5' splice site (5' SS) and the branch point sequence (BPS) by the U1 snRNP and U2 snRNPs, respectively, followed by recruitment of the U4/U6.5 snRNP, which leads to assembly of pre-catalytic complex B (Will and Luhrmann, 2011). Major structural rearrangements within the spliceosome catalyzed by two helicases, Prp28 and Brr2, lead to subsequent formation of the catalytically competent complex B*, which carries out the first transesterification reaction. Prp28 catalyzes the exchange of U1 for U6 at the 5' SS, whereas Brr2 catalyzes unwinding of the U4 and U6 snRNA duplex. The latter event allows liberated U6 snRNA to fold and pair with U2 snRNA, forming the catalytic centre of the spliceosome (Charenton et al., 2019; Mathew et al., 2008; Staley and Guthrie, 1999; Will and Luhrmann, 2011). In the pre-catalytic complex, Brr2 is kept away from its U4 substrate by multiple interactions including its direct binding to USP39 (Agafonov et al., 2016; Charenton et al., 2019). Cryo-electron

microscopy has mapped USP39 to the interface between U5 and U4/U6 snRNPs, where it has also been proposed to stabilize the U4/U6.U5 tri-snRNP (Agafonov et al., 2016). Our finding that USP39 is able to selectively bind H2A.Z-1-containing mononucleosomes suggests that USP39 is a splicing factor which can also engage chromatin outside the context of the large spliceosome complex. In this manner, it is possible that H2A.Z-1 is able to recruit USP39 to introns and in this way contribute to the regulation of splicing.

In addition to USP39, we find that the U2 snRNP complex component SF3B1 selectively associates with H2A.Z-1-mononucleosomes, further suggesting a nucleosome isoform-specific role in splicing. SF3B1 binds to branchpoint adenosines in the early stages of spliceosome assembly, and has been previously shown to bind chromatin and localize with exonic nucleosomes (Kfir et al., 2015). It is presumed that upon emergence of the 3' ends of certain introns from RNAPII, SF3B1 and the U2 snRNP translocate from exonic nucleosomes to the nascent mRNA (Kfir et al., 2015). It is also suggested that this transition is triggered by phosphorylation of SF3B1, as this modified form of SF3B1 is present within active spliceosome complexes (Kfir et al., 2015). Our finding that SF3B1 selectively associates with H2A.Z-1-nucleosomes could therefore also suggest that functional links between H2A.Z-1 and SF3B1 in active splicing could involve H2A.Z-1-nucleosome mediated recruitment of an SF3B1 kinase. Incidentally, it has been reported that USP39 can be modified by SUMOylation (Wen et al., 2014a), although a role for USP39 SUMOylation in splicing has not been reported.

Significantly, our finding that USP39 and SF3B1 are differentially enriched on H2A.Z-1- versus H2A.Z-2 nucleosomes implies that a single, conserved amino acid

substitution can be directly read by chromatin binding proteins. In the case of USP39, we demonstrate that the terminal isoform-specific residue of H2A.Z-1 (V127) enhances binding of USP39 to mononucleosomes. Interestingly, the terminal isoform-specific amino acid of H2A.Z is the only isoform-specific residue of the three that resides within the C-terminal domain. The C-terminal domain displays the greatest divergence amongst the histone H2A variants (Ausió and Abbott, 2002) and is consistent with this region mediating isoform-specific roles for H2A.Z-1 and H2A.Z-2.

Recently, yeast H2A.Z (HTZ1) has also been linked to splicing (Neves et al., 2017; Nissen et al., 2017). Specifically, it has been demonstrated that yeast H2A.Z can recruit the U2 snRNP to chromatin, and that depletion of H2A.Z primarily affects introns containing non-consensus splice sites or branchpoint sequences (Neves et al., 2017; Nissen et al., 2017). Our findings in human cells are therefore consistent with a conserved role for H2A.Z in splicing through its interaction with components of the U2 snRNP. However, it should be noted that the amino acid sequence of the C-terminus of yeast H2A.Z is completely different from the human H2A.Z-1/H2A.Z-2 sequence and; therefore, the C-terminal residue mediating H2A.Z-1-specific engagement of USP39 is unlikely to be conserved in yeast H2A.Z. In addition, although human and yeast USP39 share 65% identity in amino acid sequence, human USP39 is unique in that it possess an N-terminal RS-domain (Makarova et al., 2001), though it is not yet known if USP39 is regulated by phosphorylation. RS domains mediate protein-protein and protein-RNA interactions that are important for splicing and its regulation (Long and Caceres, 2009). As such, the RS domain of USP39 could participate in mediating interactions with H2A.Z-1 or other factors that control the alternative splicing events affected by the

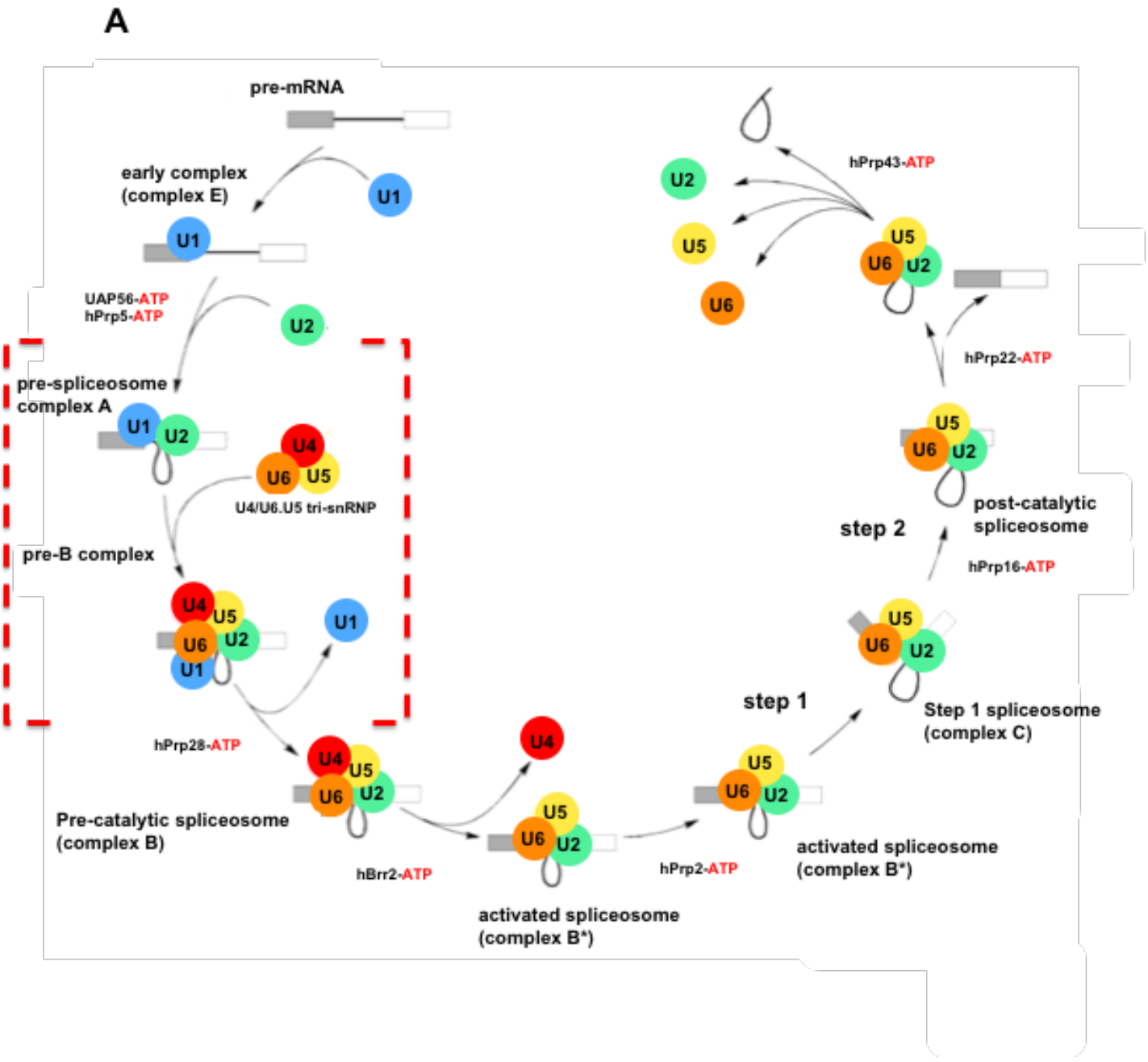
depletion of these factors. Consequently, while yeast and human H2A.Z and USP39 are both implicated in splicing regulation, and their functional roles may be conserved, it is possible the interactions underlying their mechanism of regulation are divergent. However, we have not mapped the interaction domain of USP39 that mediates its interaction with H2A.Z-1, and therefore do not know whether that domain is conserved in yeast USP39. In the broader context, the question of whether the physical and functional links between human USP39 and H2A.Z-1 is evolutionary conserved could be further explored in future studies.

In addition to our interaction data and consistent with a proposed functional coupling between H2A.Z-1 and USP39, we identify shared alternative splicing events in human cells that are affected by depletion of H2A.Z-1 or USP39. Specifically, we find a significant enrichment in cassette exon (CE) skipping events upon depletion of either USP39 or H2A.Z-1 and a similar bias towards CE skipping of the shared events. CE skipping is the predominant form of alternative splicing in animals (Barbosa-Morais et al., 2012; Kornblihtt et al., 2013). In general, the strength of a splice site is inversely proportional to its divergence from the consensus sequence, but its use is also context-dependent, and influenced through the interplay with chromatin and transcriptional machineries (Kornblihtt et al., 2013). However, we have yet to analyze the splice sites affected by H2A.Z-1 or USP39 depletion.

Interestingly, in addition to disrupting splicing of a subset of cassette exons, depletion of H2A.Z-1 also results in the inclusion of a large number of introns, whereas introns affected by USP39 depletion are equally included or skipped. This could indicate that H2A.Z-1 is involved in supporting constitutive intron splicing and regulates a subset

of splicing events independently of USP39. Previously, it has been reported that IR correlates with the accumulation of paused elongating RNAPII (i.e. Ser2 hyperphosphorylated CTD RNAPII), and that it often elicits nonsense-mediated decay (NMD) to negatively regulate cytoplasmic transcript levels (Braunschweig et al., 2014). It is therefore possible that depletion of H2A.Z-1 results in additional IR events which were not detected by RNA-Seq due to mRNA instability. At the same time, IR in cells depleted of H2A.Z-1 could reflect a general role for H2A.Z in maintaining RNAPII elongation kinetics within the gene body as H2A.Z nucleosomes are reportedly favourable to the passage of RNAPII (Weber et al., 2010, 2014a).

Altogether, our data suggest that H2A.Z-1 can modulate splice-site selection in human cells and that H2A.Z-1-mononucleosomes selectively interact with splicing factors involved in spliceosome assembly. Significantly, a direct-coupling mechanism may exist whereby H2A.Z-1-nucleosomes stabilize or recruit USP39 or the U2 snRNP to mediate a subset of splicing outcomes. As we observe USP39 to also preferentially interact with H2A.Z-2 nucleosomes in comparison to those containing H2A, it is possible that H2A.Z-2 can also regulate splicing events. In this regard, the ability of H2A.Z to form homotypic nucleosomes (i.e. those containing H2A.Z-1/H2A.Z-1, H2A.Z-1/H2A.Z-2, or H2A.Z-2/H2A.Z-2) or heterotypic nucleosomes with one copy of H2A.Z and one copy of H2A might differentially modulate splicing outcomes by fine-tuning the levels of chromatin-bound splicing factors available at splice sites.


A
pre-mRNA
early complex
(complex E)
U1
U1
UAP56-ATP
hPrp5-ATP
U2
pre-spliceosome
complex A
U1
U2
U4
U6
U5
U4/U6.U5 tri-snRNP
pre-B complex
U4
U5
U6
U2
U1
U1
hPrp28-ATP
U4
U5
U6
U2
Pre-catalytic spliceosome
(complex B)
hBrr2-ATP
U4
U5
U6
U2
activated spliceosome
(complex B*)
hPrp2-ATP
U5
U6
U2
activated spliceosome
(complex B*)
step 1
U5
U6
U2
Step 1 spliceosome
(complex C)
hPrp16-ATP
step 2
U5
U6
U2
post-catalytic
spliceosome
hPrp22-ATP
U5
U6
U2
hPrp43-ATP
U2
U5
U6

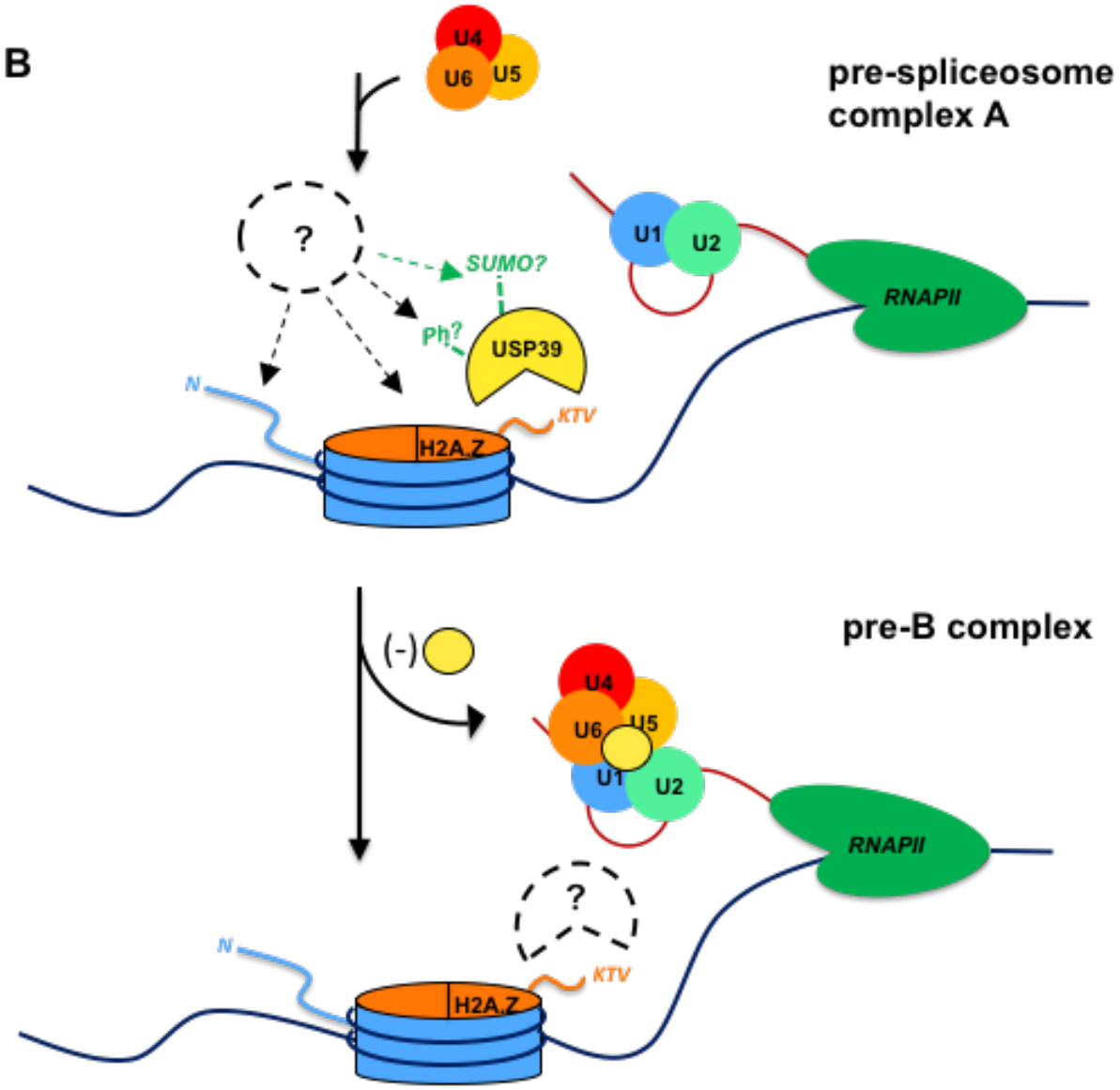


Figure 3-5. Potential coupling of H2A.Z-1 and USP39. (A) Schematic of the splicing cycle. The eight DExD/H-box ATPases/helicases that are conserved amongst yeast and human and are important for spliceosome remodelling events are indicated. For pre-mRNA, rectangles represent exons while the black like denotes an intron. The step during spliceosome assembly where H2A.Z may play a USP39-linked role is bracketed by red dashed lines (adapted from Will and Lührmann 2011). (B) H2A.Z could recruit USP39 to intronic regions. In order for USP39 to associate with the U4/U6.U5 tri-snRNP and the spliceosome, its dissociation from H2A.Z could be triggered by its PTM (e.g. SUMOylation or phosphorylation) by additional chromatin binding effectors (represented by black, dashed-line circles) that are recruited or bound to H2A.Z-nucleosomes. Alternatively, USP39 could be out-competed for binding to H2A.Z by proteins that are recruited to H2A.Z-nucleosomes. Pre-mRNA is depicted as red line.

Chapter 4

Closing Remarks and Future Directions

4 Closing remarks and future directions

The term ‘epigenetics’ was coined by Conrad Waddington in 1940 to reference mechanisms through which an organism’s environment could affect its acquired phenotype without parallel mutations to its genotype (Waddington, 1953). Waddington had observed, for example, that *Drosophila* pupae exposed to heat shock or ether vapors developed phenotypes that could be passed to successive generations at a frequency exceeding that of genetic mutation (Skvortsova et al., 2018; Waddington, 1953). We now know that heritable (or potentially heritable) changes to the genome which act “above” the linear sequence of DNA are executed by the information possessed within covalent modifications to DNA such as methylation, histone modifications, as well as by specialized non-coding RNAs. These factors establish cell type-specific regulatory programs and have the potential to be mitotically or meiotically inherited, and can thus be transmitted from parental gametes to the zygote, both directly or by proxy, through the ensuing programs they orchestrate. Epigenetic mechanisms provide an interface between the cell’s micro/environment and the regulation of cellular pathways in real-time (vs. evolutionarily) throughout its lifespan, and collectively these relationships refer to the cell’s ‘epigenome’ (Laubach et al., 2018; Roadmap Epigenomics Consortium et al., 2015). Importantly, while an organism possesses a single genome, in principle, it is characterized by at least as many epigenomes as it has cell types, and hence epigenetic networks can be inherently nuanced and contextual. This context-dependency underscores the importance of deciphering cross-talk amongst epigenetic mechanisms at the nucleosome-level on which multiple cellular pathways converge. Evolutionarily conserved and essential histone variant H2A.Z is

involved in many cellular processes; however, our understanding of its multi-faceted biological roles is still lacking due to our incomplete characterization of nucleosomes containing its modified forms. Additionally, true epigenome profiling studies necessitate cellular uniformity (Ganesan A., 2018). With respect to this, recent advances in low-cell and single-cell multi-omics techniques (including RNA-Seq, ChIP-Seq, chromatin accessibility and conformation, bisulfite sequencing of DNA methylomes, as well as genomic and proteomic techniques) are revolutionizing our ability to unravel the complexity of multicellular biological systems at unprecedented resolution (Kelsey et al., 2017; Mincarelli et al., 2018). These approaches will become increasingly important in future studies to best infer the cause and effect relationships of complex processes.

4.1 Interrogating the activities of H2A.ZUb1

Monoubiquitylation of histone H2A at K119 is catalyzed by RING1A/B of the PRC1 complex. Work in our lab previously established that H2A.ZUb1 is also mediated by RING1B, and thus H2A.ZUb1 and H2AUb1 may share some overlapping functions (Sarcinella et al., 2007). Recently, the complexity of PRC1 complexes is becoming more appreciated as they have been found extant in a large number of different forms that have the same general structure but with different proteins substituting each component. The variation of PRC1 complexes plays an important role in their targeting and function, though the extent of this diversity *in vivo* or their cell-type specificity is unknown. While both cPRC1 and ncPRC1 catalyze H2AK119Ub, less is known about the mechanisms of ncPRC1 complexes, which are not targeted to H3K27me3 through a CBX domain. (Bajusz et al., 2018). In the future, (conditional) knockout studies

systematically targeting PRC1 components will be useful in determining which complexes spatiotemporally mediate H2A.Z monoubiquitylation, and how or if they functionally overlap with PRC2.

The Polycomb pathway is essential for maintaining the balance between pluripotency and differentiation, and preserves cell-type specific silencing programs. At the same time, a detailed mechanistic understanding of Polycomb silencing, including the direct role of H2AUb1, is lacking. H2AUb1 deubiquitination has been causally linked to H3K4 methylation through the activity of a growing number of H2AK119-specific deubiquitinases (DUBs), and our lab previously established that USP16, USP10, and 2A-DUB are active towards H2A.ZUb1 (Draker et al., 2011; Sarcinella et al., 2007). To date, more than 100 putative DUBs have been annotated in the human genome. In the future, unbiased screening approaches will be useful to identify additional H2A.ZUb1 DUBs, which may have further utility in targeted mechanistic studies. Interestingly, work in this thesis demonstrates that H2A.ZUb1-enriched nucleosomes preferentially associate with the H3K4 demethylase LSD1, raising the possibility H2A.ZUb1 is also linked to removal of H3K4 methylation and hence the active establishment of PcG domains. H2A.ZUb1 is reportedly present on bivalent nucleosomes in ESCs, and LSD1 plays a role in decommissioning bivalent promoters and enhancers during lineage-specific programming (Ku et al., 2012; Whyte et al., 2012). Studies interrogating a causal role for H2A.ZUb1 in demethylating H3K4 could be facilitated by assessing co-localization of PTMs at regions of putative crosstalk in cell-lines exclusively expressing non-ubiquitylatable H2A.Z, which could be generated through CRISPR/Cas9 mutagenesis.

At the same time, correlation of ChIP- and RNA-Seq data presented in this thesis reveals the enrichment of H2A.ZUb1 within genes that are silent, though not actively repressed under basal conditions by H2A.Z. This could reflect our use of a committed cell-line (HEK293T), in which case silencing by H2A.ZUb1 could be maintained through persistent crosstalk, or that H2A.ZUb1-targeted genes might require additional signals for activation or de-repression. Further, H2A.ZUb1 could play a delocalized regulatory role, affecting the transcription of genes irrespective of its location in relation to the TSS. To this end, characterizing exonic sites of H2A.ZUb1-nucleosome enrichment could provide functional insight. For example, I have shown that H2A.ZUb1-enriched nucleosomes promote chromatin binding by CTCF and cohesin. It will be interesting to determine where in the genome H2A.ZUb1 colocalizes with these architectural proteins, and whether their overlapping sites correspond to higher-order physical contacts identified in Hi-C data.

Finally, one can envision other, non-exclusive functional consequences of an interaction between H2A.ZUb1, CTCF and cohesin. For instance, CTCF and cohesin localize to the -1 nucleosome at the TSS of divergent promoters, which transcribe RNA in both directions. At these sites, CTCF and cohesin have been proposed to supress spurious antisense transcription, the RNA products of which can otherwise in turn influence sense transcription (Bornelöv et al., 2015). It is tempting to speculate that H2A.ZUb1 plays a role in defining the -1 nucleosome of divergent promoters and is involved in targeting or stabilizing CTCF/cohesin, perhaps through a functional interaction with DNMTs. Testing this hypothesis directly could exploit the use of deactivated Cas9 (dCas9) to induce sequence-guided deubiquitination of H2A.ZUb1 on

the -1 nucleosome at candidate promoters, followed by targeted evaluation of sense and antisense transcripts.

4.2 Studying the interaction of H2A.Z with splicing

In the second part of this thesis I propose an H2A.Z-1-specific role for H2A.Z in alternative splicing (AS) regulation through its interaction with USP39. This is based on identification of an H2A.Z-1 residue that mediates selective binding of USP39 to H2A.Z-nucleosomes, and our detection of significantly overlapping changes in AS upon knockdown of H2A.Z-1 and USP39. It is important to note that work in this thesis has not explored the possibility of an H2A.Z-2-dependent role in AS which, if extant, could be dependent (antagonistically or cooperatively) or independent of H2A.Z-1/USP39. Nevertheless, several questions should be addressed regarding the interaction of H2A.Z-1 and USP39 in future studies.

First, since the AS changes observed upon siRNA knockdown were modest , future RNA-Seq experiments may reveal greater disruption of events by using inducible knockout cell-lines of H2A.Z-1 or USP39, as both proteins are present in high copy number. Moreover, to validate and further characterize the functional specificity of H2A.Z-1 and H2A.Z-2, cell-lines could be generated in which the C-terminal, isoform-specific residue of endogenous H2A.Z-1 has been mutated to the terminal amino acid of H2A.Z-2 (i.e. V127A) using CRISPR-based genome editing methods, and then profiled for differential RNA-Seq patterns. These cell-lines could also serve as the basis of subsequent rescue experiments, where the wildtype H2A.Z-1 sequence is re-introduced to rule out off-target effects. As additional controls, and because we have found that

H2A.Z-2 is expressed in much lower amounts than H2A.Z-1 in HEK293T, the three isoform-specific residues of H2A.Z-1 can be mutated to those of H2A.Z-2 in order to compare the effects of H2A.Z-2 on splicing efficiency, as well as an additional cell-line derived from these, where the H2A.Z-2 terminal residue has been re-introduced.

Secondly, future studies are needed to resolve how and where H2A.Z-1 and USP39 interact to impart regulation of AS. As a starting point, this could be investigated by ChIP-Seq analysis of USP39 binding sites in WT or H2A.Z-1 knockdown cells to test whether H2A.Z-1 is required or facilitate recruitment of USP39 to their shared target loci. While it is possible H2A.Z-1 and USP39 are functionally coupled through direct recruitment at shared loci, H2A.Z is also known to influence RNAPII promoter-proximal pausing and elongation rate (Weber and Henikoff, 2014; Weber et al., 2010), and hence it may affect target AS events by modulating RNAPII elongation as well. In this manner, the association of USP39 with H2A.Z-1 could reflect a window-of-opportunity favorable to local interactions . Mammalian native elongation transcript-sequencing (mNET-Seq) and similar techniques could be used to compare the extent of RNAPII pausing at promoters and other *cis* elements. mNET-Seq uses IP to capture nascent RNA bound to different C-terminal domain (CTD) phosphorylated forms of RNAPII (Nojima et al., 2015). In particular, sequencing from the 3’ ends of nascent transcripts bound to Ser5 hyperphosphorylated RNAPII could be used to detect the locations of paused or slowed RNAPII in cells expressing wild-type H2A.Z-1 or H2A.Z-1(V127A). In addition, directly assessing the impact of H2A.Z-1, H2A.Z-2, and C-terminal mutants of H2A.Z-1/2 on co-transcriptional splicing could be achieved using *in vitro* reporter assays. To this end, splicing efficiencies on pre-mRNA transcribed from a chromatinized template comprised

of recombinant nucleosomes from different H2A.Z constructs could be compared in the presence or absence of USP39.

Finally, it would be important to analyse chromatin-bound USP39 for its association with different spliceosomal snRNPs (e.g. U1, U2, U4/U6 snRNPs, and U4/U6.U5 tri-snRNPs), as well as additional spliceosome factors, to infer the functional state of USP39-H2A.Z-1-nucleosome complexes with respect to stage of assembly of associated splicing complexes. To this end, techniques such as co-IP or sucrose gradient ultracentrifugation combined with tandem mass spectrometry, or analysis of putative factors *a posteriori*, could yield greater insight.

4.3 Closing remark

As exemplified by its indispensability during lineage commitment, variant histone H2A.Z is closely integrated with a number of fundamental processes, and these specific activities appear to correlate with distinct patterns of histone PTMs. The functionalization of individual nucleosomes into chromatin domains and subsequent higher-order organizations - which then feedback to nucleosome regulation – could in theory precipitate from individually-targeted histone modifications. An extreme example is provided by H2A.Z-1 and H2A.Z-2 which, notwithstanding possible differences in their temporal regulation, can have varied effects on chromatin, implying that these isoforms are distinguishable by a single, conserved residue having no obvious impact on nucleosome structure. In contrast to this high degree of selectivity, histone modifications converge on the same nucleosome and engage in crosstalk. Therefore, a highly

selective signal within a regulatory program can in principle be modulated through multiple different histone modifications. Redundancy presumably acts to maintain the responsiveness of a system to a variety of signals, and deciphering how selectivity is maintained by the cell in light of this necessitates mechanistic studies which are able to assign some degree of causality. Hopefully the data accumulated in this thesis will help us develop testable hypotheses to address this aim in the future.

4.4 References

Abbott, D.W., Ivanova, V.S., Wang, X., Bonner, W.M., and Ausió, J. (2001). Characterization of the Stability and Folding of H2A.Z Chromatin Particles IMPLICATIONS FOR TRANSCRIPTIONAL ACTIVATION. J. Biol. Chem. *276*, 41945–41949.

Adamo, A., Sesé, B., Boue, S., Castaño, J., Paramonov, I., Barrero, M.J., and Belmonte, J.C.I. (2011). LSD1 regulates the balance between self-renewal and differentiation in human embryonic stem cells. Nature Cell Biology *13*, 652–659.

Agafonov, D.E., Kastner, B., Dybkov, O., Hofele, R.V., Liu, W.-T., Urlaub, H., Lührmann, R., and Stark, H. (2016). Molecular architecture of the human U4/U6.U5 tri-snRNP. Science *351*, 1416–1420.

Ahmed, S., Dul, B., Qiu, X., and Walworth, N.C. (2007). Msc1 acts through histone H2A.Z to promote chromosome stability in Schizosaccharomyces pombe. Genetics *177*, 1487–1497.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002). Chromosomal DNA and Its Packaging in the Chromatin Fiber. Molecular Biology of the Cell. 4th Edition.

Alexander, R., and Beggs, J.D. (2010). Cross-talk in transcription, splicing and chromatin: who makes the first call? Biochemical Society Transactions *38*, 1251–1256.

Ali, T., Renkawitz, R., and Bartkuhn, M. (2016). Insulators and domains of gene expression. Current Opinion in Genetics & Development *37*, 17–26.

Allfrey, V.G., Faulkner, R., and Mirsky, A.E. (1964). Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. PNAS *51*, 786–794.

Allshire, R.C., and Madhani, H.D. (2018). Ten principles of heterochromatin formation and function. Nature Reviews Molecular Cell Biology *19*, 229–244.

Aranda, S., Mas, G., and Croce, L.D. (2015). Regulation of gene transcription by Polycomb proteins. Science Advances *1*, e1500737.

Argentaro, A., Yang, J.-C., Chapman, L., Kowalczyk, M.S., Gibbons, R.J., Higgs, D.R., Neuhaus, D., and Rhodes, D. (2007). Structural consequences of disease-causing mutations in the ATRX-DNMT3-DNMT3L (ADD) domain of the chromatin-associated protein ATRX. PNAS *104*, 11939–11944.

Ausió, 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.

Babiarz, J.E., Halley, J.E., and Rine, J. (2006). Telomeric heterochromatin boundaries require NuA4-dependent acetylation of histone variant H2A.Z in Saccharomycescerevisiae. Genes Dev. *20*, 700–710.

Bagchi, D.N., and Iyer, V.R. (2016). The Determinants of Directionality in Transcriptional Initiation. Trends in Genetics *32*, 322–333.

Bajusz, I., Kovács, G., and Pirity, M.K. (2018). From Flies to Mice: The Emerging Role of Non-Canonical PRC1 Members in Mammalian Development. Epigenomes *2*, 4.

Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature *410*, 120–124.

Barbosa-Morais, N.L., Irimia, M., Pan, Q., Xiong, H.Y., Gueroussov, S., Lee, L.J., Slobodeniuc, V., Kutter, C., Watt, S., Çolak, R., et al. (2012). The Evolutionary Landscape of Alternative Splicing in Vertebrate Species. Science *338*, 1587–1593.

Barker, D.F., and Campbell, A.M. (1981). The birA gene of Escherichia coli encodes a biotin holoenzyme synthetase. J. Mol. Biol. *146*, 451–467.

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.

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.

Beckett, D., Kovaleva, E., and Schatz, P.J. (1999). A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. Protein Sci. *8*, 921–929.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006). A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. Cell *125*, 315–326.

Bickmore, W.A., and Sumner, A.T. (1989). Mammalian chromosome banding — an expression of genome organization. Trends in Genetics *5*, 144–148.

Blackledge, N.P., Farcas, A.M., Kondo, T., King, H.W., McGouran, J.F., Hanssen, L.L.P., Ito, S., Cooper, S., Kondo, K., Koseki, Y., et al. (2014). Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. Cell *157*, 1445–1459.

Blackledge, N.P., Rose, N.R., and Klose, R.J. (2015). Targeting Polycomb systems to regulate gene expression: modifications to a complex story. Nature Reviews Molecular Cell Biology *16*, 643–649.

Blackledge, N.P., Fursova, N.A., Kelley, J.R., Huseyin, M.K., Feldmann, A., and Klose, R.J. (2019). PRC1 catalytic activity is central to Polycomb system function. BioRxiv 667667.

Bonenfant, D., Towbin, H., Coulot, M., Schindler, P., Mueller, D.R., and Oostrum, J. van (2007). Analysis of Dynamic Changes in Post-translational Modifications of Human Histones during Cell Cycle by Mass Spectrometry. Molecular & Cellular Proteomics *6*, 1917–1932.

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

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

Bornelöv, S., Komorowski, J., and Wadelius, C. (2015). Different distribution of histone modifications in genes with unidirectional and bidirectional transcription and a role of CTCF and cohesin in directing transcription. BMC Genomics *16*, 300.

Boulard, M., Edwards, J.R., and Bestor, T.H. (2015). FBXL10 protects Polycomb-bound genes from hypermethylation. Nature Genetics *47*, 479–485.

Branco, M.R., and Pombo, A. (2006). Intermingling of Chromosome Territories in Interphase Suggests Role in Translocations and Transcription-Dependent Associations. PLOS Biology *4*, e138.

Braunschweig, U., Barbosa-Morais, N.L., Pan, Q., Nachman, E.N., Alipanahi, B., Gonatopoulos-Pournatzis, T., Frey, B., Irimia, M., and Blencowe, B.J. (2014). Widespread intron retention in mammals functionally tunes transcriptomes. Genome Res. *24*, 1774–1786.

Bruce, K., Myers, F.A., Mantouvalou, E., Lefevre, P., Greaves, I., Bonifer, C., Tremethick, D.J., Thorne, A.W., and Crane-Robinson, C. (2005). The replacement histone H2A.Z in a hyperacetylated form is a feature of active genes in the chicken. Nucleic Acids Res *33*, 5633–5639.

Brunelle, M., Nordell Markovits, A., Rodrigue, S., Lupien, M., Jacques, P.-É., and Gévry, N. (2015). The histone variant H2A.Z is an important regulator of enhancer activity. Nucleic Acids Res *43*, 9742–9756.

Buschbeck, M., and Hake, S.B. (2017). Variants of core histones and their roles in cell fate decisions, development and cancer. Nature Reviews Molecular Cell Biology *18*, 299–314.

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.

Cain, C.E., Blekhman, R., Marioni, J.C., and Gilad, Y. (2011). Gene Expression Differences Among Primates Are Associated With Changes in a Histone Epigenetic Modification. Genetics *187*, 1225–1234.

Cavalli, G., and Misteli, T. (2013). Functional implications of genome topology. Nature Structural & Molecular Biology *20*, 290–299.

Chakravarthy, S., Gundimella, S.K.Y., 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.

Chandru, A., Bate, N., Vuister, G.W., and Cowley, S.M. (2018). Sin3A recruits Tet1 to the PAH1 domain via a highly conserved Sin3-Interaction Domain. Scientific Reports *8*, 14689.

Charenton, C., Wilkinson, M.E., and Nagai, K. (2019). Mechanism of 5' splice site transfer for human spliceosome activation. Science *364*, 362–367.

Chédin, F., Lieber, M.R., and Hsieh, C.-L. (2002). The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. PNAS *99*, 16916–16921.

Chen, P., Zhao, J., Wang, Y., Wang, M., Long, H., Liang, D., Huang, L., Wen, Z., Li, W., Li, X., et al. (2013). H3.3 actively marks enhancers and primes gene transcription via opening higher-ordered chromatin. Genes & Development *27*, 2109–2124.

Cheung, P., Tanner, K.G., Cheung, W.L., Sassone-Corsi, P., Denu, J.M., and Allis, C.D. (2000). Synergistic Coupling of Histone H3 Phosphorylation and Acetylation in Response to Epidermal Growth Factor Stimulation. Molecular Cell *5*, 905–915.

Choi, J., Heo, K., and An, W. (2009). Cooperative action of TIP48 and TIP49 in H2A.Z exchange catalyzed by acetylation of nucleosomal H2A. Nucleic Acids Res *37*, 5993–6007.

Ciabrelli, F., and Cavalli, G. (2015). Chromatin-Driven Behavior of Topologically Associating Domains. Journal of Molecular Biology *427*, 608–625.

Clayton, A.L., Rose, S., Barratt, M.J., and Mahadevan, L.C. (2000). Phosphoacetylation of histone H3 on c-fos- and c-jun-associated nucleosomes upon gene activation. The EMBO Journal *19*, 3714–3726.

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.

Cooper, S., Dienstbier, M., Hassan, R., Schermelleh, L., Sharif, J., Blackledge, N.P., De Marco, V., Elderkin, S., Koseki, H., Klose, R., et al. (2014). Targeting Polycomb to Pericentric Heterochromatin in Embryonic Stem Cells Reveals a Role for H2AK119u1 in PRC2 Recruitment. Cell Rep *7*, 1456–1470.

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

Craig, J.M., and Bickmore, W.A. (1993). Genes and genomes: Chromosome bands – flavours to savour. BioEssays *15*, 349–354.

Creyghton, M.P., Markoulaki, S., Levine, S.S., Hanna, J., Lodato, M.A., Sha, K., Young, R.A., Jaenisch, R., and Boyer, L.A. (2008). H2AZ is enriched at polycomb complex target genes in ES cells and is necessary for lineage commitment. Cell *135*, 649–661.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. PNAS *107*, 21931–21936.

Croft, J.A., Bridger, J.M., Boyle, S., Perry, P., Teague, P., and Bickmore, W.A. (1999). Differences in the Localization and Morphology of Chromosomes in the Human Nucleus. The Journal of Cell Biology *145*, 1119–1131.

Cui, Y., Cai, M., and Stanley, H.E. (2017). Comparative Analysis and Classification of Cassette Exons and Constitutive Exons.

Cull, M.G., and Schatz, P.J. (2000). Biotinylation of proteins in vivo and in vitro using small peptide tags. Meth. Enzymol. *326*, 430–440.

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

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing Chromosome Conformation. Science *295*, 1306–1311.

Di Croce, L., and Helin, K. (2013). Transcriptional regulation by Polycomb group proteins. Nature Structural & Molecular Biology *20*, 1147–1155.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380.

Dowen, J.M., Fan, Z.P., Hnisz, D., Ren, G., Abraham, B.J., Zhang, L.N., Weintraub, A.S., Schuijers, J., Lee, T.I., Zhao, K., et al. (2014). Control of Cell Identity Genes Occurs in Insulated Neighborhoods in Mammalian Chromosomes. Cell *159*, 374–387.

Draker, R., Sarcinella, E., and Cheung, P. (2011). USP10 deubiquitylates the histone variant H2A.Z and both are required for androgen receptor-mediated gene activation. Nucleic Acids Res *39*, 3529–3542.

Draker, R., Ng, M.K., Sarcinella, E., Ignatchenko, V., Kislinger, T., and Cheung, P. (2012). A Combination of H2A.Z and H4 Acetylation Recruits Brd2 to Chromatin during Transcriptional Activation. PLOS Genetics *8*, e1003047.

Dryhurst, D., McMullen, B., Fazli, L., Rennie, P.S., and Ausió, 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 Letters *315*, 38–47.

Dujardin, G., Lafaille, C., Petrillo, E., Buggiano, V., Gómez Acuña, L.I., Fiszbein, A., Godoy Herz, M.A., Nieto Moreno, N., Muñoz, M.J., Alló, M., et al. (2013). Transcriptional elongation and alternative splicing. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms *1829*, 134–140.

Dunn, C.J., Sarkar, P., Bailey, E.R., Farris, S., Zhao, M., Ward, J.M., Dudek, S.M., and Saha, R.N. (2017). Histone Hypervariants H2A.Z.1 and H2A.Z.2 Play Independent and Context-Specific Roles in Neuronal Activity-Induced Transcription of Arc/Arg3.1 and Other Immediate Early Genes. ENeuro *4*.

Eirín-López, J.M., González-Romero, R., Dryhurst, D., Ishibashi, T., and Ausió, 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.

Endoh, M., Endo, T.A., Endoh, T., Isono, K., Sharif, J., Ohara, O., Toyoda, T., Ito, T., Eskeland, R., Bickmore, W.A., et al. (2012). Histone H2A Mono-Ubiquitination Is a Crucial Step to Mediate PRC1-Dependent Repression of Developmental Genes to Maintain ES Cell Identity. PLOS Genetics *8*, e1002774.

Entrevan, M., Schuettengruber, B., and Cavalli, G. (2016). Regulation of Genome Architecture and Function by Polycomb Proteins. Trends in Cell Biology *26*, 511–525.

Erdel, F., and Rippe, K. (2018). Formation of Chromatin Subcompartments by Phase Separation. Biophysical Journal *114*, 2262–2270.

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

Fan, 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 & Molecular Biology *9*, 172–176.

Fan, J.Y., Rangasamy, D., Luger, K., and Tremethick, D.J. (2004). H2A.Z Alters the Nucleosome Surface to Promote HP1α-Mediated Chromatin Fiber Folding. Molecular Cell *16*, 655–661.

Farrelly, L.A., Thompson, R.E., Zhao, S., Lepack, A.E., Lyu, Y., Bhanu, N.V., Zhang, B., Loh, Y.-H.E., Ramakrishnan, A., Vadodaria, K.C., et al. (2019). Histone serotonylation is a permissive modification that enhances TFIID binding to H3K4me3. Nature *567*, 535.

Filippova, D., Patro, R., Duggal, G., and Kingsford, C. (2014). Identification of alternative topological domains in chromatin. Algorithms for Molecular Biology *9*, 14.

Fischle, W., Wang, Y., Jacobs, S.A., Kim, Y., Allis, C.D., and Khorasanizadeh, S. (2003). Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. Genes Dev. *17*, 1870–1881.

Fischle, W., Tseng, B.S., Dormann, H.L., Ueberheide, B.M., Garcia, B.A., Shabanowitz, J., Hunt, D.F., Funabiki, H., and Allis, C.D. (2005). Regulation of HP1–chromatin binding by histone H3 methylation and phosphorylation. Nature *438*, 1116–1122.

Flanagan, J.F., Mi, L.-Z., Chruszcz, M., Cymborowski, M., Clines, K.L., Kim, Y., Minor, W., Rastinejad, F., and Khorasanizadeh, S. (2005). Double chromodomains cooperate to recognize the methylated histone H3 tail. Nature *438*, 1181–1185.

Fritz, A.J., Barutcu, A.R., Martin-Buley, L., Wijnen, A.J. van, Zaidi, S.K., Imbalzano, A.N., Lian, J.B., Stein, J.L., and Stein, G.S. (2016). Chromosomes at Work: Organization of Chromosome Territories in the Interphase Nucleus. Journal of Cellular Biochemistry *117*, 9–19.

Ganesan A. (2018). Epigenetics: the first 25 centuries. Philosophical Transactions of the Royal Society B: Biological Sciences *373*, 20170067.

Gao, Y., Gan, H., Lou, Z., and Zhang, Z. (2018). Asf1a resolves bivalent chromatin domains for the induction of lineage-specific genes during mouse embryonic stem cell differentiation. Proceedings of the National Academy of Sciences *115*, E6162–E6171.

Gao, Z., Zhang, J., Bonasio, R., Strino, F., Sawai, A., Parisi, F., Kluger, Y., and Reinberg, D. (2012). PCGF Homologs, CBX Proteins, and RYBP Define Functionally Distinct PRC1 Family Complexes. Mol Cell *45*, 344–356.

Gévry, N., Hardy, S., Jacques, P.-É., Laflamme, L., Svotelis, A., Robert, F., and Gaudreau, L. (2009). Histone H2A.Z is essential for estrogen receptor signaling. Genes Dev.

Ghirlando, R., and Felsenfeld, G. (2016). CTCF: making the right connections. Genes Dev. *30*, 881–891.

Giaimo, B.D., Ferrante, F., Herchenröther, A., Hake, S.B., and Borggrefe, T. (2019). The histone variant H2A.Z in gene regulation. Epigenetics & Chromatin *12*, 37.

Gibcus, J.H., and Dekker, J. (2013). The Hierarchy of the 3D Genome. Molecular Cell *49*, 773–782.

Gilbert, N., Boyle, S., Fiegler, H., Woodfine, K., Carter, N.P., and Bickmore, W.A. (2004). Chromatin Architecture of the Human Genome: Gene-Rich Domains Are Enriched in Open Chromatin Fibers. Cell *118*, 555–566.

Goldstrohm, A.C., Greenleaf, A.L., and Garcia-Blanco, M.A. (2001). Co-transcriptional splicing of pre-messenger RNAs: considerations for the mechanism of alternative splicing. Gene *277*, 31–47.

Gonzalez-Sandoval, A., and Gasser, S.M. (2016). On TADs and LADs: Spatial Control Over Gene Expression. Trends in Genetics *32*, 485–495.

Grigoryev, S.A., Nikitina, T., Pehrson, J.R., Singh, P.B., and Woodcock, C.L. (2004). Dynamic relocation of epigenetic chromatin markers reveals an active role of constitutive heterochromatin in the transition from proliferation to quiescence. Journal of Cell Science *117*, 6153–6162.

Guillemette, B., and Gaudreau, L. (2006). Reuniting the contrasting functions of H2A.ZThis paper is one of a selection of papers published in this Special Issue, entitled 27th International West Coast Chromatin and Chromosome Conference, and has undergone the Journal's usual peer review process. Biochem. Cell Biol. *84*, 528–535.

Guo, R., Zheng, L., Park, J.W., Lv, R., Chen, H., Jiao, F., Xu, W., Mu, S., Wen, H., Qiu, J., et al. (2014). BS69/ZMYND11 Reads and Connects Histone H3.3 Lysine 36 Trimethylation-Decorated Chromatin to Regulated Pre-mRNA Processing. Molecular Cell *56*, 298–310.

Halley, J.E., Kaplan, T., Wang, A.Y., Kobor, M.S., and Rine, J. (2010). Roles for H2A.Z and Its Acetylation in GAL1 Transcription and Gene Induction, but Not GAL1-Transcriptional Memory. PLOS Biology *8*, e1000401.

Hamiche, A., and Shuaib, M. (2012). Chaperoning the histone H3 family. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms *1819*, 230–237.

Henikoff, S. (2009). Labile H3.3+H2A.Z nucleosomes mark "nucleosome-free regions." Nature Genetics *41*, 865–866.

Henikoff, S., and Greally, J.M. (2016). Epigenetics, cellular memory and gene regulation. Current Biology *26*, R644–R648.

Herzel, L., Ottoz, D.S.M., Alpert, T., and Neugebauer, K.M. (2017). Splicing and transcription touch base: co-transcriptional spliceosome assembly and function. Nature Reviews Molecular Cell Biology *18*, 637–650.

Holmquist, G.P. (1992). Chromosome bands, their chromatin flavors, and their functional features. Am J Hum Genet *51*, 17–37.

Holmquist, G., Gray, M., Porter, T., and Jordan, J. (1982). Characterization of Giemsa dark- and light-band DNA. Cell *31*, 121–129.

Holoch, D., and Margueron, R. (2017). Mechanisms Regulating PRC2 Recruitment and Enzymatic Activity. Trends Biochem. Sci. *42*, 531–542.

Horikoshi, N., Arimura, Y., Taguchi, H., and Kurumizaka, H. (2016). Crystal structures of heterotypic nucleosomes containing histones H2A.Z and H2A. Open Biology *6*, 160127.

Hou, H., Wang, Y., Kallgren, S.P., Thompson, J., Yates, J.R., and Jia, S. (2010). Histone variant H2A.Z regulates centromere silencing and chromosome segregation in fission yeast. J. Biol. Chem. *285*, 1909–1918.

Hu, G., and Wade, P.A. (2012). NuRD and Pluripotency: A Complex Balancing Act. Cell Stem Cell *10*, 497–503.

Huang, J., Marco, E., Pinello, L., and Yuan, G.-C. (2015). Predicting chromatin organization using histone marks. Genome Biology *16*, 162.

Huang, J., Li, K., Cai, W., Liu, X., Zhang, Y., Orkin, S.H., Xu, J., and Yuan, G.-C. (2018). Dissecting super-enhancer hierarchy based on chromatin interactions. Nature Communications *9*, 943.

Huang, Y.-H., Chung, C.-S., Kao, D.-I., Kao, T.-C., and Cheng, S.-C. (2014). Sad1 Counteracts Brr2-Mediated Dissociation of U4/U6.U5 in Tri-snRNP Homeostasis. Molecular and Cellular Biology *34*, 210–220.

Hughes, A.L., and Rando, O.J. (2014). Mechanisms Underlying Nucleosome Positioning In Vivo. Annual Review of Biophysics *43*, 41–63.

Hunter, T. (2007). The Age of Crosstalk: Phosphorylation, Ubiquitination, and Beyond. Molecular Cell *28*, 730–738.

Ikura, T., Ogryzko, V.V., Grigoriev, M., Groisman, R., Wang, J., Horikoshi, M., Scully, R., Qin, J., and Nakatani, Y. (2000). Involvement of the TIP60 Histone Acetylase Complex in DNA Repair and Apoptosis. Cell *102*, 463–473.

Imakaev, M., Fudenberg, G., McCord, R.P., Naumova, N., Goloborodko, A., Lajoie, B.R., Dekker, J., and Mirny, L.A. (2012). Iterative correction of Hi-C data reveals hallmarks of chromosome organization. Nature Methods *9*, 999–1003.

Iñiguez, L.P., and Hernández, G. (2017). The Evolutionary Relationship between Alternative Splicing and Gene Duplication. Front. Genet. *8*.

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

Irimia, M., Weatheritt, R.J., Ellis, J.D., Parikshak, N.N., Gonatopoulos-Pournatzis, T., Babor, M., Quesnel-Vallières, M., Tapial, J., Raj, B., O'Hanlon, D., et al. (2014). A Highly Conserved Program of Neuronal Microexons Is Misregulated in Autistic Brains. Cell *159*, 1511–1523.

Isono, K., Endo, T.A., Ku, M., Yamada, D., Suzuki, R., Sharif, J., Ishikura, T., Toyoda, T., Bernstein, B.E., and Koseki, H. (2013). SAM Domain Polymerization Links Subnuclear Clustering of PRC1 to Gene Silencing. Developmental Cell *26*, 565–577.

Jenuwein, T., and Allis, C.D. (2001). Translating the Histone Code. Science *293*, 1074–1080.

Jia, D., Jurkowska, R.Z., Zhang, X., Jeltsch, A., and Cheng, X. (2007). Structure of Dnmt3a bound to Dnmt3L suggests a model for *de novo* DNA methylation. Nature *449*, 248–251.

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

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 in the human genome. Nat Genet *41*, 941–945.

Junco, S.E., Wang, R., Gaipa, J.C., Taylor, A.B., Schirf, V., Gearhart, M.D., Bardwell, V.J., Demeler, B., Hart, P.J., and Kim, C.A. (2013). Structure of the polycomb group protein PCGF1 in complex with BCOR reveals basis for binding selectivity of PCGF homologs. Structure *21*, 665–671.

Jürgens, G. (1985). A group of genes controlling the spatial expression of the bithorax complex in Drosophila. Nature *316*, 153.

Kalashnikova Anna A., Porter-Goff Mary E., Muthurajan Uma M., Luger Karolin, and Hansen Jeffrey C. (2013). The role of the nucleosome acidic patch in modulating higher order chromatin structure. Journal of The Royal Society Interface *10*, 20121022.

Kalb, R., Latwiel, S., Baymaz, H.I., Jansen, P.W.T.C., Müller, C.W., Vermeulen, M., and Müller, J. (2014). Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression. Nature Structural & Molecular Biology *21*, 569–571.

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. Mol. Cell *33*, 335–343.

Karras, G.I., Kustatscher, G., Buhecha, H.R., Allen, M.D., Pugieux, C., Sait, F., Bycroft, M., and Ladurner, A.G. (2005). The macro domain is an ADP-ribose binding module. The EMBO Journal *24*, 1911–1920.

Kelsey, G., Stegle, O., and Reik, W. (2017). Single-cell epigenomics: Recording the past and predicting the future. Science *358*, 69–75.

Keogh, M.-C., Mennella, T.A., Sawa, C., Berthelet, S., Krogan, N.J., Wolek, A., Podolny, V., Carpenter, L.R., Greenblatt, J.F., Baetz, K., et al. (2006). The Saccharomyces cerevisiae histone H2A variant Htz1 is acetylated by NuA4. Genes Dev. *20*, 660–665.

Kfir, N., Lev-Maor, G., Glaich, O., Alajem, A., Datta, A., Sze, S.K., Meshorer, E., and Ast, G. (2015). SF3B1 Association with Chromatin Determines Splicing Outcomes. Cell Reports *11*, 618–629.

Kikuchi, M., Okumura, F., Tsukiyama, T., Watanabe, M., Miyajima, N., Tanaka, J., Imamura, M., and Hatakeyama, S. (2009). TRIM24 mediates ligand-dependent activation of androgen receptor and is repressed by a bromodomain-containing protein, BRD7, in prostate cancer cells. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research *1793*, 1828–1836.

Kim, E., Goren, A., and Ast, G. (2008). Alternative splicing: current perspectives. BioEssays *30*, 38–47.

Kim, K., Punj, V., Choi, J., Heo, K., Kim, J.-M., Laird, P.W., and An, W. (2013). Gene dysregulation by histone variant H2A.Z in bladder cancer. Epigenetics & Chromatin *6*, 34.

Klymenko, T., and Müller, J. (2004). The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. EMBO Reports *5*, 373–377.

Kolasinska-Zwierz, P., Down, T., Latorre, I., Liu, T., Liu, X.S., and Ahringer, J. (2009). Differential chromatin marking of introns and expressed exons by H3K36me3. Nature Genetics *41*, 376–381.

Kornberg, R.D. (1974). Chromatin Structure: A Repeating Unit of Histones and DNA. Science *184*, 868–871.

Kornberg, R.D. (1977). Structure of Chromatin. Annual Review of Biochemistry *46*, 931–954.

Kornblihtt, A.R., Schor, I.E., Alló, M., Dujardin, G., Petrillo, E., and Muñoz, M.J. (2013). Alternative splicing: a pivotal step between eukaryotic transcription and translation. Nature Reviews Molecular Cell Biology *14*, 153–165.

Kossel, A. (1883). Ueber einen peptonartigen Bestandtheil des Zellkerns. Zeitschrift für Physiologische Chemie *8*, 511–515.

Kozlowski, M., Corujo, D., Hothorn, M., Guberovic, I., Mandemaker, I.K., Blessing, C., Sporn, J., Gutierrez-Triana, A., Smith, R., Portmann, T., et al. (2018). MacroH2A

histone variants limit chromatin plasticity through two distinct mechanisms. EMBO Reports *19*, e44445.

Kraushaar, D.C., Jin, W., Maunakea, A., Abraham, B., Ha, M., and Zhao, K. (2013). Genome-wide incorporation dynamics reveal distinct categories of turnover for the histone variant H3.3. Genome Biology *14*, R121.

Krogan, N.J., Baetz, K., Keogh, M.-C., Datta, N., Sawa, C., Kwok, T.C.Y., Thompson, N.J., Davey, M.G., Pootoolal, J., Hughes, T.R., et al. (2004). Regulation of chromosome stability by the histone H2A variant Htz1, the Swr1 chromatin remodeling complex, and the histone acetyltransferase NuA4. PNAS *101*, 13513–13518.

Ku, M., Jaffe, J.D., Koche, R.P., Rheinbay, E., Endoh, M., Koseki, H., Carr, S.A., and Bernstein, B.E. (2012). H2A.Z landscapes and dual modifications in pluripotent and multipotent stem cells underlie complex genome regulatory functions. Genome Biol *13*, R85.

Kulaeva, O.I., Hsieh, F.-K., and Studitsky, V.M. (2010). RNA polymerase complexes cooperate to relieve the nucleosomal barrier and evict histones. Proc. Natl. Acad. Sci. U.S.A. *107*, 11325–11330.

Kusch, T., Florens, L., MacDonald, W.H., Swanson, S.K., Glaser, R.L., Yates, J.R., Abmayr, S.M., Washburn, M.P., and Workman, J.L. (2004). Acetylation by Tip60 Is Required for Selective Histone Variant Exchange at DNA Lesions. Science *306*, 2084–2087.

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.

de Laat, W., and Duboule, D. (2013). Topology of mammalian developmental enhancers and their regulatory landscapes. Nature *502*, 499–506.

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature *410*, 116–120.

Lamond, A.I., and Spector, D.L. (2003). Nuclear speckles: a model for nuclear organelles. Nature Reviews Molecular Cell Biology *4*, 605–612.

Lantermann, A.B., Straub, T., Strålfors, A., Yuan, G.-C., Ekwall, K., and Korber, P. (2010). *Schizosaccharomyces pombe* genome-wide nucleosome mapping reveals positioning mechanisms distinct from those of *Saccharomyces cerevisiae*. Nature Structural & Molecular Biology *17*, 251–257.

Laubach, Z.M., Perng, W., Dolinoy, D.C., Faulk, C.D., Holekamp, K.E., and Getty, T. (2018). Epigenetics and the maintenance of developmental plasticity: extending the signalling theory framework. Biological Reviews *93*, 1323–1338.

Law, C., and Cheung, P. (2015). Expression of Non-acetylatable H2A.Z in Myoblast Cells Blocks Myoblast Differentiation through Disruption of MyoD Expression. J. Biol. Chem. *290*, 13234–13249.

Lawrence, M., Daujat, S., and Schneider, R. (2016). Lateral Thinking: How Histone Modifications Regulate Gene Expression. Trends in Genetics *32*, 42–56.

Lee, C.-K., Shibata, Y., Rao, B., Strahl, B.D., and Lieb, J.D. (2004). Evidence for nucleosome depletion at active regulatory regions genome-wide. Nature Genetics *36*, 900–905.

Lee, J.-S., Smith, E., and Shilatifard, A. (2010). The Language of Histone Crosstalk. Cell *142*, 682–685.

Lewis, E.B. (2004). A Gene Complex Controlling Segmentation in Drosophila. In Genes, Development and Cancer: The Life and Work of Edward B. Lewis, H.D. Lipshitz, ed. (Boston, MA: Springer US), pp. 205–217.

Li, Q., Brown, J.B., Huang, H., and Bickel, P.J. (2011). Measuring reproducibility of high-throughput experiments. Ann. Appl. Stat. *5*, 1752–1779.

Lieberman-Aiden, E., Berkum, N.L. van, Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. Science *326*, 289–293.

Lindblad-Toh, K., Garber, M., Zuk, O., Lin, M.F., Parker, B.J., Washietl, S., Kheradpour, P., Ernst, J., Jordan, G., Mauceli, E., et al. (2011). A high-resolution map of human evolutionary constraint using 29 mammals. Nature *478*, 476–482.

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

Long, J.C., and Caceres, J.F. (2009). The SR protein family of splicing factors: master regulators of gene expression. Biochemical Journal *417*, 15–27.

Long, M., Sun, X., Shi, W., Yanru, A., Leung, S.T.C., Ding, D., Cheema, M.S., MacPherson, N., Nelson, C.J., Ausio, J., et al. A novel histone H4 variant H4G regulates rDNA transcription in breast cancer. Nucleic Acids Res.

Loomis, R.J., Naoe, Y., Parker, J.B., Savic, V., Bozovsky, M.R., Macfarlan, T., Manley, J.L., and Chakravarti, D. (2009). Chromatin Binding of SRp20 and ASF/SF2 and Dissociation from Mitotic Chromosomes Is Modulated by Histone H3 Serine 10 Phosphorylation. Molecular Cell *33*, 450–461.

Luco, R.F., Pan, Q., Tominaga, K., Blencowe, B.J., Pereira-Smith, O.M., and Misteli, T. (2010). Regulation of Alternative Splicing by Histone Modifications. Science *327*, 996–1000.

Luco, R.F., Allo, M., Schor, I.E., Kornblihtt, A.R., and Misteli, T. (2011). Epigenetics in Alternative Pre-mRNA Splicing. Cell *144*, 16–26.

Luger, K., and Richmond, T.J. (1998). The histone tails of the nucleosome. Current Opinion in Genetics & Development *8*, 140–146.

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

Lynch, M.D., Smith, A.J.H., De Gobbi, M., Flenley, M., Hughes, J.R., Vernimmen, D., Ayyub, H., Sharpe, J.A., Sloane-Stanley, J.A., Sutherland, L., et al. (2012). An interspecies analysis reveals a key role for unmethylated CpG dinucleotides in vertebrate Polycomb complex recruitment. EMBO J *31*, 317–329.

Makarova, O.V., Makarov, E.M., and Lührmann, R. (2001). The 65 and 110 kDa SR-related proteins of the U4/U6·U5 tri-snRNP are essential for the assembly of mature spliceosomes. The EMBO Journal *20*, 2553–2563.

Maslon, M.M., Braunschweig, U., Aitken, S., Mann, A.R., Kilanowski, F., Hunter, C.J., Blencowe, B.J., Kornblihtt, A.R., Adams, I.R., and Cáceres, J.F. (2019). A slow transcription rate causes embryonic lethality and perturbs kinetic coupling of neuronal genes. The EMBO Journal *38*, e101244.

de la Mata, M., Alonso, C.R., Kadener, S., Fededa, J.P., Blaustein, M., Pelisch, F., Cramer, P., Bentley, D., and Kornblihtt, A.R. (2003). A Slow RNA Polymerase II Affects Alternative Splicing In Vivo. Molecular Cell *12*, 525–532.

Mathew, R., Hartmuth, K., Möhlmann, S., Urlaub, H., Ficner, R., and Lührmann, R. (2008). Phosphorylation of human PRP28 by SRPK2 is required for integration of the U4/U6-U5 tri-snRNP into the spliceosome. Nature Structural & Molecular Biology *15*, 435–443.

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

Maurano, M.T., Wang, H., John, S., Shafer, A., Canfield, T., Lee, K., and Stamatoyannopoulos, J.A. (2015). Role of DNA Methylation in Modulating Transcription Factor Occupancy. Cell Reports *12*, 1184–1195.

Mavrich, T.N., Jiang, C., Ioshikhes, I.P., Li, X., Venters, B.J., Zanton, S.J., Tomsho, L.P., Qi, J., Glaser, R.L., Schuster, S.C., et al. (2008). Nucleosome organization in the *Drosophila* genome. Nature *453*, 358–362.

Meneghini, M.D., Wu, M., and Madhani, H.D. (2003). Conserved histone variant H2A.Z protects euchromatin from the ectopic spread of silent heterochromatin. Cell *112*, 725–736.

Merkenschlager, M., and Nora, E.P. (2016). CTCF and Cohesin in Genome Folding and Transcriptional Gene Regulation. Annual Review of Genomics and Human Genetics *17*, 17–43.

Merkenschlager, M., and Odom, D.T. (2013). CTCF and Cohesin: Linking Gene Regulatory Elements with Their Targets. Cell *152*, 1285–1297.

Millar, C.B., Xu, F., Zhang, K., and Grunstein, M. (2006). Acetylation of H2AZ Lys 14 is associated with genome-wide gene activity in yeast. Genes Dev *20*, 711–722.

Mills, A.A. (2010). Throwing the cancer switch: reciprocal roles of polycomb and trithorax proteins. Nature Reviews Cancer *10*, 669–682.

Min, J., Zhang, Y., and Xu, R.-M. (2003). Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. Genes Dev. *17*, 1823–1828.

Mincarelli, L., Lister, A., Lipscombe, J., and Macaulay, I.C. (2018). Defining Cell Identity with Single-Cell Omics. PROTEOMICS *18*, 1700312.

Misteli, T. (2005). Concepts in nuclear architecture. BioEssays *27*, 477–487.

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.

Montavon, T., Soshnikova, N., Mascrez, B., Joye, E., Thevenet, L., Splinter, E., de Laat, W., Spitz, F., and Duboule, D. (2011). A Regulatory Archipelago Controls Hox Genes Transcription in Digits. Cell *147*, 1132–1145.

Musselman, C.A., Ramírez, J., Sims, J.K., Mansfield, R.E., Oliver, S.S., Denu, J.M., Mackay, J.P., Wade, P.A., Hagman, J., and Kutateladze, T.G. (2012). Bivalent recognition of nucleosomes by the tandem PHD fingers of the CHD4 ATPase is required for CHD4-mediated repression. PNAS *109*, 787–792.

Mylonas, C., and Tessarz, P. (2019). NET-prism enables RNA polymerase-dedicated transcriptional interrogation at nucleotide resolution. RNA Biology *0*, 1–10.

Nakayama, J., Rice, J.C., Strahl, B.D., Allis, C.D., and Grewal, S.I. (2001). Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. Science *292*, 110–113.

de Napoles, M., Mermoud, J.E., Wakao, R., Tang, Y.A., Endoh, M., Appanah, R., Nesterova, T.B., Silva, J., Otte, A.P., Vidal, M., et al. (2004). Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. Dev. Cell *7*, 663–676.

Neves, L.T., Douglass, S., Spreafico, R., Venkataramanan, S., Kress, T.L., and Johnson, T.L. (2017). The histone variant H2A.Z promotes efficient cotranscriptional splicing in S. cerevisiae. Genes Dev. *31*, 702–717.

Ng, M.K., and Cheung, P. (2015). A brief histone in time: understanding the combinatorial functions of histone PTMs in the nucleosome context. Biochem. Cell Biol. *94*, 33–42.

Nguyen, U.T.T., Bittova, L., Müller, M.M., Fierz, B., David, Y., Houck-Loomis, B., Feng, V., Dann, G.P., and Muir, T.W. (2014). Accelerated chromatin biochemistry using DNA-barcoded nucleosome libraries. Nature Methods *11*, 834–840.

Nishiyama, A., Yamaguchi, L., Sharif, J., Johmura, Y., Kawamura, T., Nakanishi, K., Shimamura, S., Arita, K., Kodama, T., Ishikawa, F., et al. (2013). Uhrf1-dependent H3K23 ubiquitylation couples maintenance DNA methylation and replication. Nature *502*, 249–253.

Nissen, K.E., Homer, C.M., Ryan, C.J., Shales, M., Krogan, N.J., Patrick, K.L., and Guthrie, C. (2017). The histone variant H2A.Z promotes splicing of weak introns. Genes Dev. *31*, 688–701.

Nojima, T., Gomes, T., Grosso, A.R.F., Kimura, H., Dye, M.J., Dhir, S., Carmo-Fonseca, M., and Proudfoot, N.J. (2015). Mammalian NET-Seq Reveals Genome-wide Nascent Transcription Coupled to RNA Processing. Cell *161*, 526–540.

Noordermeer, D., Leleu, M., Splinter, E., Rougemont, J., Laat, W.D., and Duboule, D. (2011). The Dynamic Architecture of Hox Gene Clusters. Science *334*, 222–225.

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., Berkum, N.L. van, Meisig, J., Sedat, J., et al. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature *485*, 381–385.

Obri, A., Ouararhni, K., Papin, C., Diebold, M.-L., Padmanabhan, K., Marek, M., Stoll, I., Roy, L., Reilly, P.T., Mak, T.W., et al. (2014). ANP32E is a histone chaperone that removes H2A.Z from chromatin. Nature *505*, 648–653.

Olins, D.E., and Olins, A.L. (2003). Chromatin history: our view from the bridge. Nature Reviews Molecular Cell Biology *4*, 809–814.

Oliver, S.S., Musselman, C.A., Srinivasan, R., Svaren, J.P., Kutateladze, T.G., and Denu, J.M. (2012). Multivalent Recognition of Histone Tails by the PHD Fingers of CHD5. Biochemistry *51*, 6534–6544.

Ong, C.-T., and Corces, V.G. (2014). CTCF: an architectural protein bridging genome topology and function. Nature Reviews Genetics *15*, 234–246.

Ooi, S.K.T., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S.-P., Allis, C.D., et al. (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to *de novo* methylation of DNA. Nature *448*, 714–717.

Oomen, M.E., Hansen, A., Liu, Y., Darzacq, X., and Dekker, J. (2018). CTCF sites display cell cycle dependent dynamics in factor binding and nucleosome positioning. BioRxiv.

Osley, M.A. (2006). Regulation of histone H2A and H2B ubiquitylation. Brief Funct Genomics *5*, 179–189.

Otani, J., Nankumo, T., Arita, K., Inamoto, S., Ariyoshi, M., and Shirakawa, M. (2009). Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX–DNMT3–DNMT3L domain. EMBO Reports *10*, 1235–1241.

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

Pan, Q., Shai, O., Lee, L.J., Frey, B.J., and Blencowe, B.J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nature Genetics *40*, 1413–1415.

Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H.C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T., et al. (2008). Cohesins functionally associate with CTCF on mammalian chromosome arms. Cell *132*, 422–433.

Passarge, E. (1979). Emil Heitz and the concept of heterochromatin: longitudinal chromosome differentiation was recognized fifty years ago. Am J Hum Genet *31*, 106–115.

Peng, Y., and Zhang, Y. (2018). Enhancer and super-enhancer: Positive regulators in gene transcription. Animal Models and Experimental Medicine *1*, 169–179.

Pinheiro, I., and Heard, E. (2017). X chromosome inactivation: new players in the initiation of gene silencing. F1000Res *6*.

Pirrotta, V., and Li, H.-B. (2012). A view of nuclear Polycomb bodies. Current Opinion in Genetics & Development *22*, 101–109.

Plank, J.L., and Dean, A. (2014). Enhancer Function: Mechanistic and Genome-Wide Insights Come Together. Molecular Cell *55*, 5–14.

Poux, S., Horard, B., Sigrist, C.J.A., and Pirrotta, V. (2002). The Drosophila Trithorax protein is a coactivator required to prevent re-establishment of Polycomb silencing. Development *129*, 2483–2493.

Pradeepa, M.M., Sutherland, H.G., Ule, J., Grimes, G.R., and Bickmore, W.A. (2012). Psip1/Ledgf p52 Binds Methylated Histone H3K36 and Splicing Factors and Contributes to the Regulation of Alternative Splicing. PLOS Genetics *8*, e1002717.

Putiri, E.L., Tiedemann, R.L., Thompson, J.J., Liu, C., Ho, T., Choi, J.-H., and Robertson, K.D. (2014). Distinct and overlapping control of 5-methylcytosine and 5-hydroxymethylcytosine by the TET proteins in human cancer cells. Genome Biology *15*, R81.

Qin, W., Wolf, P., Liu, N., Link, S., Smets, M., Mastra, F.L., Forné, I., Pichler, G., Hörl, D., Fellinger, K., et al. (2015). DNA methylation requires a DNMT1 ubiquitin interacting motif (UIM) and histone ubiquitination. Cell Research *25*, 911–929.

Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S.A., Flynn, R.A., and Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. Nature *470*, 279–283.

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.

Rangasamy, D., Berven, L., Ridgway, P., and Tremethick, D.J. (2003). Pericentric heterochromatin becomes enriched with H2A.Z during early mammalian development. The EMBO Journal *22*, 1599–1607.

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.

Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., et al. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665–1680.

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., et al. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature *406*, 593–599.

Reddington, J.P., Sproul, D., and Meehan, R.R. (2014). DNA methylation reprogramming in cancer: Does it act by re-configuring the binding landscape of Polycomb repressive complexes? Bioessays *36*, 134–140.

Reinberg, D., and Vales, L.D. (2018). Chromatin domains rich in inheritance. Science *361*, 33–34.

Reynolds, N., Salmon-Divon, M., Dvinge, H., Hynes-Allen, A., Balasooriya, G., Leaford, D., Behrens, A., Bertone, P., and Hendrich, B. (2012). NuRD-mediated deacetylation of

H3K27 facilitates recruitment of Polycomb Repressive Complex 2 to direct gene repression. The EMBO Journal *31*, 593–605.

Ricketts, M.D., and Marmorstein, R. (2017). A Molecular Prospective for HIRA Complex Assembly and H3.3-Specific Histone Chaperone Function. Journal of Molecular Biology *429*, 1924–1933.

Ringrose, L., and Paro, R. (2004). Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. Annu. Rev. Genet. *38*, 413–443.

Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., et al. (2015). Integrative analysis of 111 reference human epigenomes. Nature *518*, 317–330.

Rose, N.R., King, H.W., Blackledge, N.P., Fursova, N.A., Ember, K.J., Fischer, R., Kessler, B.M., and Klose, R.J. (2016). RYBP stimulates PRC1 to shape chromatin-based communication between Polycomb repressive complexes. ELife *5*.

Rothbart, S.B., and Strahl, B.D. (2014). Interpreting the language of histone and DNA modifications. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms *1839*, 627–643.

Rubio, E.D., Reiss, D.J., Welcsh, P.L., Disteche, C.M., Filippova, G.N., Baliga, N.S., Aebersold, R., Ranish, J.A., and Krumm, A. (2008). CTCF physically links cohesin to chromatin. PNAS *105*, 8309–8314.

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.

Ruthenburg, A.J., Li, H., Milne, T.A., Dewell, S., McGinty, R.K., Yuen, M., Ueberheide, B., Dou, Y., Muir, T.W., Patel, D.J., et al. (2011). Recognition of a Mononucleosomal Histone Modification Pattern by BPTF via Multivalent Interactions. Cell *145*, 692–706.

Sabo, P.J., Humbert, R., Hawrylycz, M., Wallace, J.C., Dorschner, M.O., McArthur, M., and Stamatoyannopoulos, J.A. (2004). Genome-wide identification of DNaseI hypersensitive sites using active chromatin sequence libraries. PNAS *101*, 4537–4542.

Saldi, T., Cortazar, M.A., Sheridan, R.M., and Bentley, D.L. (2016). Coupling of RNA Polymerase II Transcription Elongation with Pre-mRNA Splicing. Journal of Molecular Biology *428*, 2623–2635.

Santos-Rosa, H., Schneider, R., Bannister, A.J., Sherriff, J., Bernstein, B.E., Emre, N.C.T., Schreiber, S.L., Mellor, J., and Kouzarides, T. (2002). Active genes are tri-methylated at K4 of histone H3. Nature *419*, 407–411.

Sarcinella, E., Zuzarte, P.C., Lau, P.N.I., Draker, R., and Cheung, P. (2007). Monoubiquitylation of H2A.Z Distinguishes Its Association with Euchromatin or Facultative Heterochromatin. Mol Cell Biol *27*, 6457–6468.

Sashida, G., and Iwama, A. (2017). Multifaceted role of the polycomb-group gene EZH2 in hematological malignancies. Int J Hematol *105*, 23–30.

Sawicka, A., and Seiser, C. (2014). Sensing core histone phosphorylation — A matter of perfect timing. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms *1839*, 711–718.

Schatz, P.J. (1993). Use of Peptide Libraries to Map the Substrate Specificity of a Peptide-Modifying Enzyme: A 13 Residue Consensus Peptide Specifies Biotinylation in Escherichia coli. Nature Biotechnology *11*, 1138–1143.

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.

Schwartz, S., and Ast, G. (2010). Chromatin density and splicing destiny: on the cross-talk between chromatin structure and splicing. The EMBO Journal *29*, 1629–1636.

Schwartz, S., Meshorer, E., and Ast, G. (2009). Chromatin organization marks exon-intron structure. Nature Structural & Molecular Biology *16*, 990–995.

Sevilla, A., and Binda, O. (2014). Post-translational modifications of the histone variant h2az. Stem Cell Res *12*, 289–295.

Shi, X., Kachirskaia, I., Walter, K.L., Kuo, J.-H.A., Lake, A., Davrazou, F., Chan, S.M., Martin, D.G.E., Fingerman, I.M., Briggs, S.D., et al. (2007). Proteome-wide analysis in Saccharomyces cerevisiae identifies several PHD fingers as novel direct and selective binding modules of histone H3 methylated at either lysine 4 or lysine 36. J. Biol. Chem. *282*, 2450–2455.

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.

Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., and de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture–on-chip (4C). Nature Genetics *38*, 1348–1354.

Sims, R.J., Millhouse, S., Chen, C.-F., Lewis, B.A., Erdjument-Bromage, H., Tempst, P., Manley, J.L., and Reinberg, D. (2007). Recognition of Trimethylated Histone H3 Lysine 4 Facilitates the Recruitment of Transcription Postinitiation Factors and Pre-mRNA Splicing. Molecular Cell *28*, 665–676.

Skvortsova, K., Iovino, N., and Bogdanović, O. (2018). Functions and mechanisms of epigenetic inheritance in animals. Nature Reviews Molecular Cell Biology *19*, 774.

Soboleva, T.A., Parker, B.J., Nekrasov, M., Hart-Smith, G., Tay, Y.J., Tng, W.-Q., Wilkins, M., Ryan, D., and Tremethick, D.J. (2017). A new link between transcriptional initiation and pre-mRNA splicing: The RNA binding histone variant H2A.B. PLOS Genetics *13*, e1006633.

Song, F., Chen, P., Sun, D., Wang, M., Dong, L., Liang, D., Xu, R.-M., Zhu, P., and Li, G. (2014). Cryo-EM Study of the Chromatin Fiber Reveals a Double Helix Twisted by Tetranucleosomal Units. Science *344*, 376–380.

Spies, N., Nielsen, C.B., Padgett, R.A., and Burge, C.B. (2009). Biased Chromatin Signatures around Polyadenylation Sites and Exons. Molecular Cell *36*, 245–254.

Staley, J.P., and Guthrie, C. (1999). An RNA Switch at the 5′ Splice Site Requires ATP and the DEAD Box Protein Prp28p. Molecular Cell *3*, 55–64.

Stedman, W., Kang, H., Lin, S., Kissil, J.L., Bartolomei, M.S., and Lieberman, P.M. (2008). Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators. EMBO J. *27*, 654–666.

Stock, J.K., Giadrossi, S., Casanova, M., Brookes, E., Vidal, M., Koseki, H., Brockdorff, N., Fisher, A.G., and Pombo, A. (2007). Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. Nat. Cell Biol. *9*, 1428–1435.

Stoop, P. van der, Boutsma, E.A., Hulsman, D., Noback, S., Heimerikx, M., Kerkhoven, R.M., Voncken, J.W., Wessels, L.F.A., and Lohuizen, M. van (2008). Ubiquitin E3 Ligase Ring1b/Rnf2 of Polycomb Repressive Complex 1 Contributes to Stable Maintenance of Mouse Embryonic Stem Cells. PLOS ONE *3*, e2235.

Strahl, B.D., Ohba, R., Cook, R.G., and Allis, C.D. (1999). Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in Tetrahymena. Proc. Natl. Acad. Sci. U.S.A. *96*, 14967–14972.

Su, Z., and Denu, J.M. (2016). Reading the Combinatorial Histone Language. ACS Chem. Biol. *11*, 564–574.

Subramanian*, V., Fields*, P.A., and Boyer, L.A. (2015). H2A.Z: a molecular rheostat for transcriptional control. F1000Prime Rep *7*.

Suetake, I., Shinozaki, F., Miyagawa, J., Takeshima, H., and Tajima, S. (2004). DNMT3L Stimulates the DNA Methylation Activity of Dnmt3a and Dnmt3b through a Direct Interaction. J. Biol. Chem. *279*, 27816–27823.

Surface, L.E., Fields, P.A., Subramanian, V., Behmer, R., Udeshi, N., Peach, S.E., Jaffe, J.D., and Boyer, L.A. (2016). H2A.Z.1 mono-ubiquitylation antagonizes BRD2 to maintain poised chromatin in ESCs. Cell Rep *14*, 1142–1155.

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

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

Tanabe, H., Müller, S., Neusser, M., Hase, J. von, Calcagno, E., Cremer, M., Solovei, I., Cremer, C., and Cremer, T. (2002). Evolutionary conservation of chromosome territory arrangements in cell nuclei from higher primates. PNAS *99*, 4424–4429.

Tapial, J., Ha, K.C.H., Sterne-Weiler, T., Gohr, A., Braunschweig, U., Hermoso-Pulido, A., Quesnel-Vallières, M., Permanyer, J., Sodaei, R., Marquez, Y., et al. (2017). An atlas of alternative splicing profiles and functional associations reveals new regulatory programs and genes that simultaneously express multiple major isoforms. Genome Res. *27*, 1759–1768.

Tavares, L., Dimitrova, E., Oxley, D., Webster, J., Poot, R., Demmers, J., Bezstarosti, K., Taylor, S., Ura, H., Koide, H., et al. (2012). RYBP-PRC1 Complexes Mediate H2A Ubiquitylation at Polycomb Target Sites Independently of PRC2 and H3K27me3. Cell *148*, 664–678.

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

Thatcher, T.H., and Gorovsky, M.A. (1994). Phylogenetic analysis of the core histones H2A, H2B, H3, and H4. Nucleic Acids Res *22*, 174–179.

Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., Haugen, E., Sheffield, N.C., Stergachis, A.B., Wang, H., Vernot, B., et al. (2012). The accessible chromatin landscape of the human genome. Nature *489*, 75–82.

Tilgner, H., Nikolaou, C., Althammer, S., Sammeth, M., Beato, M., Valcárcel, J., and Guigó, R. (2009). Nucleosome positioning as a determinant of exon recognition. Nature Structural & Molecular Biology *16*, 996–1001.

Tollervey, J.R., and Lunyak, V.V. (2012). Epigenetics. Epigenetics *7*, 823–840.

Torres, I.O., and Fujimori, D.G. (2015). Functional coupling between writers, erasers and readers of histone and DNA methylation. Current Opinion in Structural Biology *35*, 68–75.

Trojer, P., and Reinberg, D. (2007). Facultative Heterochromatin: Is There a Distinctive Molecular Signature? Molecular Cell *28*, 1–13.

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 Research *22*, 307–321.

Vardabasso, C., Gaspar-Maia, A., Hasson, D., Pünzeler, S., Valle-Garcia, D., Straub, T., Keilhauer, E.C., Strub, T., Dong, J., Panda, T., et al. (2015). Histone Variant H2A.Z.2 Mediates Proliferation and Drug Sensitivity of Malignant Melanoma. Molecular Cell *59*, 75–88.

Venkatesh, S., and Workman, J.L. (2015). Histone exchange, chromatin structure and the regulation of transcription. Nature Reviews Molecular Cell Biology *16*, 178–189.

Venne, A.S., Kollipara, L., and Zahedi, R.P. (2014). The next level of complexity: Crosstalk of posttranslational modifications. PROTEOMICS *14*, 513–524.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., et al. (2001). The Sequence of the Human Genome. Science *291*, 1304–1351.

Vietri Rudan, M., Barrington, C., Henderson, S., Ernst, C., Odom, D.T., Tanay, A., and Hadjur, S. (2015). Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. Cell Rep *10*, 1297–1309.

Viré, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Eynde, A.V., Bernard, D., Vanderwinden, J.-M., et al. (2006). The Polycomb group protein EZH2 directly controls DNA methylation. Nature *439*, 871.

Waddington, C.H. (1953). Genetic Assimilation of an Acquired Character. Evolution *7*, 118–126.

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.

Wang, H., Maurano, M.T., Qu, H., Varley, K.E., Gertz, J., Pauli, F., Lee, K., Canfield, T., Weaver, M., Sandstrom, R., et al. (2012). Widespread plasticity in CTCF occupancy linked to DNA methylation. Genome Res. *22*, 1680–1688.

Wang, Y., Zhang, H., Chen, Y., Sun, Y., Yang, F., Yu, W., Liang, J., Sun, L., Yang, X., Shi, L., et al. (2009). LSD1 Is a Subunit of the NuRD Complex and Targets the Metastasis Programs in Breast Cancer. Cell *138*, 660–672.

Wang, Y., Long, H., Yu, J., Dong, L., Wassef, M., Zhuo, B., Li, X., Zhao, J., Wang, M., Liu, C., et al. (2018). Histone variants H2A.Z and H3.3 coordinately regulate PRC2-

dependent H3K27me3 deposition and gene expression regulation in mES cells. BMC Biology *16*, 107.

Wang, Z., Song, J., Milne, T.A., Wang, G.G., Li, H., Allis, C.D., and Patel, D.J. (2010). Pro isomerization in MLL1 PHD3-bromo cassette connects H3K4me readout to CyP33 and HDAC-mediated repression. Cell *141*, 1183–1194.

Wani, A.H., Boettiger, A.N., Schorderet, P., Ergun, A., Münger, C., Sadreyev, R.I., Zhuang, X., Kingston, R.E., and Francis, N.J. (2016). Chromatin topology is coupled to Polycomb group protein subnuclear organization. Nature Communications *7*, 10291.

Weber, C.M., and Henikoff, S. (2014). Histone variants: dynamic punctuation in transcription. Genes & Development *28*, 672–682.

Weber, C.M., Henikoff, J.G., and Henikoff, S. (2010). H2A.Z nucleosomes enriched over active genes are homotypic. Nature Structural & Molecular Biology *17*, 1500–1507.

Weber, C.M., Ramachandran, S., and Henikoff, S. (2014a). Nucleosomes Are Context-Specific, H2A.Z-Modulated Barriers to RNA Polymerase. Molecular Cell *53*, 819–830.

Weber, C.M., Ramachandran, S., and Henikoff, S. (2014b). Nucleosomes Are Context-Specific, H2A.Z-Modulated Barriers to RNA Polymerase. Molecular Cell *53*, 819–830.

Weil, M.R., Widlak, P., Minna, J.D., and Garner, H.R. (2004). Global Survey of Chromatin Accessibility Using DNA Microarrays. Genome Res. *14*, 1374–1381.

Wen, D., Xu, Z., Xia, L., Liu, X., Tu, Y., Lei, H., Wang, W., Wang, T., Song, L., Ma, C., et al. (2014a). Important Role of SUMOylation of Spliceosome Factors in Prostate Cancer Cells. J. Proteome Res. *13*, 3571–3582.

Wen, H., Li, Y., Xi, Y., Jiang, S., Stratton, S., Peng, D., Tanaka, K., Ren, Y., Xia, Z., Wu, J., et al. (2014b). ZMYND11 links histone H3.3K36me3 to transcription elongation and tumour suppression. Nature *508*, 263–268.

Whyte, W.A., Bilodeau, S., Orlando, D.A., Hoke, H.A., Frampton, G.M., Foster, C.T., Cowley, S.M., and Young, R.A. (2012). Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. Nature *482*, 221–225.

Will, C.L., and Luhrmann, R. (2011). Spliceosome Structure and Function. Cold Spring Harbor Perspectives in Biology *3*, a003707–a003707.

Williams, K., Christensen, J., Pedersen, M.T., Johansen, J.V., Cloos, P.A.C., Rappsilber, J., and Helin, K. (2011). TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature *473*, 343–348.

Wolffe, A. (1998). Chromatin: Structure and Function (Academic Press).

Wu, L., Lee, S.Y., Zhou, B., Nguyen, U.T.T., Muir, T.W., Tan, S., and Dou, Y. (2013). ASH2L Regulates Ubiquitylation Signaling to MLL: trans-Regulation of H3 K4 Methylation in Higher Eukaryotes. Molecular Cell *49*, 1108–1120.

Würtele, H., and Chartrand, P. (2006). Genome-wide scanning of HoxB1-associated loci in mouse ES cells using an open-ended Chromosome Conformation Capture methodology. Chromosome Res *14*, 477–495.

Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T., Badenhorst, P., et al. (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature *442*, 86–90.

Yildirim, O., Li, R., Hung, J.-H., Chen, P.B., Dong, X., Ee, L.-S., Weng, Z., Rando, O.J., and Fazzio, T.G. (2011). Mbd3/NURD Complex Regulates Expression of 5-Hydroxymethylcytosine Marked Genes in Embryonic Stem Cells. Cell *147*, 1498–1510.

Yu, M., Hon, G.C., Szulwach, K.E., Song, C.-X., Zhang, L., Kim, A., Li, X., Dai, Q., Shen, Y., Park, B., et al. (2012). Base-Resolution Analysis of 5-Hydroxymethylcytosine in the Mammalian Genome. Cell *149*, 1368–1380.

Yuan, G., Ma, B., Yuan, W., Zhang, Z., Chen, P., Ding, X., Feng, L., Shen, X., Chen, S., Li, G., et al. (2013). Histone H2A ubiquitination inhibits the enzymatic activity of H3 Lysine 36 methyltransferases. J. Biol. Chem. jbc.M113.475996.

Yun, M., Wu, J., Workman, J.L., and Li, B. (2011). Readers of histone modifications. Cell Research *21*, 564–578.

Zentner, G.E., Tesar, P.J., and Scacheri, P.C. (2011). Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. Genome Res.

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.

Zhang, T., Cooper, S., and Brockdorff, N. (2015). The interplay of histone modifications - writers that read. EMBO Reports *16*, 1467–1481.

Zhang, Y., McCord, R.P., Ho, Y.-J., Lajoie, B.R., Hildebrand, D.G., Simon, A.C., Becker, M.S., Alt, F.W., and Dekker, J. (2012). Spatial Organization of the Mouse Genome and Its Role in Recurrent Chromosomal Translocations. Cell *148*, 908–921.

Zhao, Y., and Garcia, B.A. (2015). Comprehensive Catalog of Currently Documented Histone Modifications. Cold Spring Harb Perspect Biol *7*, a025064.

Zhao, Z., Tavoosidana, G., Sjölinder, M., Göndör, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Singh Sandhu, K., Singh, U., et al. (2006). Circular chromosome

conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nature Genetics *38*, 1341–1347.

Zhou, K., Gaullier, G., and Luger, K. (2019). Nucleosome structure and dynamics are coming of age. Nature Structural & Molecular Biology *26*, 3.

Zhou, W., Zhu, P., Wang, J., Pascual, G., Ohgi, K.A., Lozach, J., Glass, C.K., and Rosenfeld, M.G. (2008). Histone H2A Monoubiquitination Represses Transcription by Inhibiting RNA Polymerase II Transcriptional Elongation. Mol Cell *29*, 69–80.

Zilberman, D., Coleman-Derr, D., Ballinger, T., and Henikoff, S. (2008). Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. Nature *456*, 125–129.

Zippo, A., Serafini, R., Rocchigiani, M., Pennacchini, S., Krepelova, A., and Oliviero, S. (2009). Histone Crosstalk between H3S10ph and H4K16ac Generates a Histone Code that Mediates Transcription Elongation. Cell *138*, 1122–1136.

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

Zofall, M., Fischer, T., Zhang, K., Zhou, M., Cui, B., Veenstra, T.D., and Grewal, S.I.S. (2009). Histone H2A.Z cooperates with RNAi and heterochromatin factors to suppress antisense RNAs. Nature *461*, 419–422.