From Department of Medical Biochemistry and Biophysics
Karolinska Institutet, Stockholm, Sweden

EXPLORING THE DYNAMIC EPIGENOME IN PLURIPOTENT STEM CELLS USING QUANTITATIVE METHODS

Angelo Salazar Mantero

Karolinska Institutet

Stockholm 2022

Published by Karolinska Institutet.
Printed by Universitetsservice US-AB, 2022

ISBN 978-91-8016-787-1
Cover illustration: Chromatin containing nucleosomes with histone variants and epigenetic regulating enzymes TET1 and PRC2. Artist: Jaime Espinoza Ruiz.

Exploring the Dynamic Epigenome in Pluripotent Stem Cells using Quantitative Methods
THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Angelo Salazar Mantero

The thesis will be defended in public at Eva & Georg Klein seminar hall, Karolinska Institutet, Solnavägen 9, Solna on October 31st, 2022 at 09:30 AM.

Principal Supervisor:
Assoc. Prof. Simon Elsässer
Karolinska Institutet
Dept. of Medical Biochemistry and Biophysics
Division of Genome Biology

Co-supervisors:
Prof. Gonçalo Castelo-Branco
Karolinska Institutet
Dept. of Medical Biochemistry and Biophysics
Division of Molecular Neurobiology

Assoc. Prof. Peter Svensson
Karolinska Institutet
Dept. of Biosciences and Nutrition

Opponent:
Prof. Johan Jakobsson
Lund University
Dept. of Experimental Medical Science
Division of Molecular Neurogenetics

Examination Board:
Prof. Jan Larsson
Umeå University
Dept. of Molecular Biology

Assoc. Prof. Claudio Cantù
Linköping University
Dept. of Bioclinical and Medical Sciences
Division of Molecular Medicine and Virology

Assoc. Prof. Colm Nestor
Linköping University
Dept. of Bioclinical and Medical Sciences
Division of Children's and Women's Health

Dedicated to my family and friends. Thank you for all your support!

Para mi familia que siempre esta a mi lado a pesar de la distancia.

POPULAR SCIENCE SUMMARY OF THE THESIS

Imagine yourself being a child and entering a library containing all the books ever written. Initially you are curious about many subjects, but with time you become inclined towards reading books of a certain topic, like arts or natural sciences. Eventually, you choose an even more dedicated path finally resulting in your current career. Other people entering the very same library chose another set of books to read, leading them to specialize as nurses, constructors or web developers.

Like the child in the library having the potential to pursue any profession, stem cells have the capability to generate highly specialized cells, as neurons, muscle cells and skin cells. The cells' genetic library is its DNA, where the information for generating each cell type resides, but the cell fate is controlled by the combination of genes that are activated and silenced.

The regulation on how genes are turned on or off is of extreme importance to study, to understand the complexity of human development, where one fertilized cell renders a highly complex human being, or to understand erroneous situations where the wrong sets of genes are activated, causing disease progression. In this thesis we have studied the interaction of DNA with different proteins that mediate whether genes are turned on or off. These proteins may be seen as librarians that facilitate the access to certain books or move shelves to restricted areas.

In Paper I, we studied the protein TET1 which normally helps removing DNA methylation, which is a road blocks for gene expression. Surprisingly, we found that it also has a role in maintaining certain repetitive regions silenced. Thus, in the 'library', it acts as a librarian with dual roles, both allowing for certain books to be read, but making sure that potentially harmful books are off limits.

In Paper II, we investigated the initial steps of human development, where we dissected the role of the protein complex PRC2. The findings showed that when this protein complex does not function properly, stem cells specialize toward cells that build up not only the embryo, but also the tissue that forms the placenta. This shows that PRC2 has a pivotal role in stem cell regulation and in mediating its progression throughout development. Hence, in the 'library', PRC2 acts as a librarian that guides the child to the children's section, and in its absence, the child runs toward different juvenile sections.

Lastly, in Papers III and IV, we studied the interaction of DNA with the most abundant protein class it binds to, histones. More specifically, we studied a subset of histones called histone variants. Histone variants through evolution have diverged into having specialized regulations and functions, which are yet to be fully understood. We sought to unravel what proteins the histone variants interact with, where in the DNA they are localized and for how long they stay there. In the 'library' one could see the histones as the shelves that books are put in, where some shelves are more accessible, while others are very narrow with books jammed in tightly.

As you can read, the librarians carry a crucial role in keeping the library in order, choosing what books to put on display at certain times and which ones to hide. In this way the librarian can guide the readers toward their specializations, resulting in the readers acquiring all the knowledge needed on how to perform individual and advanced tasks. The more we understand the complex work that librarians do in the world's most comprehensive library, the better we can understand how they guide the readers towards this end.

Similarly, the more we understand about the molecular mechanisms that shape and control gene regulation, the better we will understand stem cell biology and development, and also when something goes wrong, how diseases initiate and progress. A better understanding of molecular mechanisms will lead to better diagnostics and improved treatments.

ABSTRACT

A totipotent stem cell has the potential to give rise to the trillions of cells in the human body, all carrying the very same genetic information. Through differentiation events, gene expression changes guided by epigenetic mechanisms resulting in specialized phenotypes. The more we understand of the highly dynamic epigenome, the better we will understand key phenomena such as human development, tissue regeneration and disease initiation and progression. In this thesis, focus lays on the role of key players in the field of molecular mechanisms of epigenetics using both mouse and human embryonic stem cells.

In **Paper I** we investigated the non-catalytic function of ten-eleven translocation 1 (TET1) methylcytosine dioxygenase in mouse embryonic stem cells (ESCs). Upon knocking out TET1, endogenous retroviral expression increased, while neither cells expressing wild type TET1 or a non-catalytic mutant showed the same trend. Studying the epigenetic landscape, we found that TET1, independently of its catalytic activity, is important for establishment of the silencing histone marks H3K9me3 and H4K20me3 at endogenous retroviral elements. This suggests that TET1 serves as an interaction hub for chromatin modifying complexes to repress the interstitial heterochromatin that the ERVs reside in.

In **Paper II** we investigated the role of Polycomb Repressive Complex 2 (PRC2) as an epigenetic regulator for cell type specification in human ESCs. By abolishing the function of PRC2 via drugs or knocking out its catalytic subunit we observed spurious differentiation of naïve pluripotent stem cells toward cells belonging to the mesoderm and trophectoderm lineages, indicating that PRC2 has a pivotal role in shielding naïve pluripotent stem cells from differentiating toward trophectoderm differentiation.

In **Paper III** we used genetic code expansion in mouse ESCs to produce acute and defined fractions of labelled histone variant H3.3 to study its chromatin deposition kinetics and turnover rate using quantitative methods for immunocytochemistry, chromatin immunoprecipitation sequencing and protein quantification. We revealed that H3.3 accumulates rapidly in a subnuclear space together with DAXX, ATRX, Smarcad1 and HP1 prior to significant chromatin incorporation both at enhancers and interstitial heterochromatin Moreover, this technique allowed for studying novel interactors of H3.3 in a temporal manner directly after protein synthesis. Furthermore, we found that one of the interactors plays a key role as a chromatin remodeler allowing for H3.3 turnover in enhancers.

Using the same methods, in **Paper IV** we expanded the study from Paper III to also focus on the kinetics of canonical histones H2A and H3 compared to the variants H2A.Z, macroH2A, H3.3 and CENP-A. Our results show that the histones are subjected differently to pre-assembly degradation, have defined individual genomic incorporation rates and distinctive half-life in chromatin. Using quantitative ChIP-seq allowed for studying the incorporation to repetitive elements, which is of essence when studying histone variants. Furthermore, we laid some ground work towards finding the enigmatic histone chaperone of macroH2A.

Taken together, we show that TET1, PRC2 and histone variants play essential and unique roles in the maintenance of homeostasis in ESCs. Continuing to unravel their dynamics and roles will be instrumental for understanding epigenetically regulated diseases and lead to improved diagnostics and treatments.

LIST OF SCIENTIFIC PAPERS

I. *Paul Stolz, **Angelo Salazar Mantero**, Andrey Tvardovskiy, Enes Ugur, Lucas E. Wange, Christopher B. Mulholland, Yuying Cheng, Michael Wierer, Wolfgang Enard, Robert Schneider, Till Bartke, Heinrich Leonhardt, Simon J Elsässer and Sebastian Bultmann*
TET1 regulates gene expression and repression of endogenous retroviruses independent of DNA demethylation
Nucleic Acids Research 2022, vol. 50, p. 8491–8511

II. *Banushree Kumar*, Carmen Navarro*, Nerges Winblad*, John P Schell, Cheng Zhao, Jere Weltner, Laura Baqué-Vidal, **Angelo Salazar Mantero**, Sophie Petropoulos, Fredrik Lanner, Simon J Elsässer*
Polycomb Repressive Complex 2 shields naïve human pluripotent cells from trophectoderm differentiation
Nature Cell Biology 2022, vol. 24, p. 845–857

III. *Anna-Maria Katsori*, **Angelo Salazar Mantero***, Yuying Cheng, Carmen Navarro, Jürgen Eirich, Rozina Caridha, Banushree Kumar, Georgios Mermelekas, Simon J Elsässer*
Quantitative histone H3.3 dynamics resolved by rapid protein pulsing
Manuscript

IV. ***Angelo Salazar Mantero**, Yuying Cheng, Marius Jung, Rozina Caridha, Janice Linne, Jürgen Eirich, Simon J Elsässer*
Quantitative temporal catalogue of histone variant dynamics in mESCs
Manuscript

* Authors contributed equally

Scientific papers not included in the thesis:

I. *Nerges Winblad, **Angelo Salazar Mantero**, Simon J Elsässer, Fredrik Lanner*
SPEN is indispensable for XIST expression in naïve human embryonic stem cells
Manuscript

CONTENTS

LIST OF ABBREVIATIONS

5caC	5-carboxylcytosine
5fC	5-formylcytosine
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
ANP32E	Acidic Nuclear Phosphoprotein 32 Family Member E
ASF1	Anti-Silencing Factor 1
ATRX	Alpha-Thalassemia/mental Retardation X-linked syndrome
CABIN1	Calcineurin Binding Protein 1
CENP-A	Centromeric specific protein A
CHD1	Chromodomain Helicase DNA Binding Protein 1
ChIP-seq	Chromatin immunoprecipitation sequencing
CM	Catalytic mutant
DAXX	Death Domain Associated Protein
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
EED	Embryonic Ectoderm Development
ERV	Endogenous etroviral element
EZH2	Enhancer of Zeste Homolog 2
FACT	Facilitates Chromatin Transcription
GATA3	GATA Binding Protein 3
GCE	Genetic code expansion
H2Aub	H2A ubiquitination
H3K27me3	Histone H3 lysine 27 trimethylation
H3K4me3	Histone H3 lysine 4 trimethylation
H3K9me3	Histone H3 lysine 9 trimethylation
H4K20me3	Histone H4 lysine 20 trimethylation
hESC	Human embryonic stem cell
HIRA	Histone regulator A
HJURP	Holliday Junction Recognition Protein
HP1	Heterochromatin Protein

IAP	intracisternal A-type particles
IPO	Importin
KDM	Lysine-specific demethylase
KO	Kknock out
LINE	Long interspersed nuclear elements
LTR	Long terminal repeat
mESC	Mouse embryonic stem cell
MINUTE-ChIP	Multiplexed indexed unique molecule T7 amplification end-to-end sequencing
mRNA	Messenger RNA
NAP1	Nucleosome Assembly Protein 1
NASP	Nuclear Autoantigenic Sperm Protein
ncAA	Non-canonical amino acid
POI	Protein of interest
PRC2	Polycomb repressive complex 2
PTM	Post-translational modification
PylRS	Pyrrolysyl-tRNA Synthetase
PylT	Pyrrolysyl-tRNA
RNA	Ribonucleic acid
SINE	Short interspersed nuclear elements
Smarcad1	SWI/SNF-Related, Matrix-Associated Actin-Dependent Regulator Of Chromatin, Subfamily A, Containing DEAD/H Box 1
TE	Transposable element
TET1	Ten-Eleven Translocation 1
tRNA	Transfer RNA
WT	Wild type

1 INTRODUCTION

All cells in the human body contain the very same genetic information, yet they differ vastly in their phenotypes. The reason for this is that all cells originate from a single cell, the zygote, which is a totipotent stem cell that during embryonic development progressively divides and generates cells that commit to different lineages (1). Those lineages result in different tissue types ultimately forming the entire human body in its full complexity. For the cells to specialize, i.e. differentiate, different sets of genes need to be activated while other sets are turned off. This process of regulating the activity of the genome is termed epigenetic gene regulation (2). To activate or repress gene activity, there are changes to a gene or its regulating regions in terms of their accessibility, either allowing or impeding access of the transcriptional machinery. This process is reversible as the chromatin structure is highly dynamic, allowing for rapid changes resulting as consequence to different signal inputs.

The dynamic epigenome is kept in balance by different chaperones that deposit histones that form nucleosomes, the basic constituents of chromatin, and by chromatin remodellers that mediate the eviction or sliding of nucleosomes (3). The way for the remodellers to interact with chromatin is through covalent post-translational modifications (PTMs) to histones which are deposited by enzymes termed *writers*, identified by *readers*, and removed by *erasers* (4). This leads to a highly dynamic interplay of histone turnover at sites with *active* histone marks or to a silent state of chromatin in regions with *repressive* marks. Other than histone marks, cytosine methylation at CG-rich promoters results in silencing, but this is also a reversible mark, which is removed by the Ten-Eleven Translocation (TET) family of dioxygenases, the TET enzymes, resulting in reactivation of the once silenced promoter (5). Another level of epigenetic alteration of chromatin is by the replacement of canonical histones with histone variants, which have during the last decades gain progressively more attention in fine-tuning chromatin organization and gene expression (6).

In the studies covered by this thesis, we show the importance of different factors to maintain homeostasis in pluripotent stem cells which is important to further understand in the growing field of stem cells. The more we understand about stem cell biology, the more we can use that knowledge for further understanding development and diseases, leading toward more precise drug development with novel targets and drugs and into the field of tissue engineering.

2 LITERATURE REVIEW

The genome carries the information encoding all proteins that are expressed in all cells. One way that the cell coordinates how the transcriptome and consequently the proteome differ between cell types or through the cell cycle is through regulated chromatin accessibility and compaction (3). DNA is wound onto proteins called histones, that are packaged into structures named nucleosomes. These nucleosomes interact with neighboring nucleosomes and linker histones to form ordered chromatosomes. These structures allow for further compaction of the chromatin fiber, so that all genomic DNA fits into the confined space of the nucleus (7,8). This highly compacted structure needs to allow for opening of certain regions for transcription to occur. Each cell type has a subset of genes that need to be activated in a cell cycle dependent manner, and this is controlled by local decompaction of corresponding regions of the genome, allowing accessibility for the transcription machinery to come in direct contact with the DNA carrying the genetic information. Another biological process that requires unraveling of the entire chromatin structure is DNA replication, where the replication fork mediates melting of the two DNA strands to allow for each molecule of DNA to be copied to two molecules, one for each daughter cell upon division.

Chromatin is not solely composed of histones and DNA, but also other classes of proteins, such as transcription factors and chromatin remodelers, as well as regulatory RNAs such as microRNA and long noncoding RNA that interact with DNA (9,10). ATP-driven chromatin remodelers can alter the structure of chromatin by facilitating nucleosome assembly, increasing DNA accessibility by sliding or eviction of nucleosomes or loosening of the DNA wrapped around the histone octamer (11). The latter results in higher DNA accessibility for proteins required for transcription, replication, repair and recombination.

Altogether, the condensed chromatin structure formed is denoted as *heterochromatin* and the open and active regions are called *euchromatin*. These chromatin states are able to change dynamically, being globally loose in early development, and creating more stable heterochromatic states upon differentiation (12). The plasticity of chromatin is also crucial for the interaction of regions not proximal to each other, either on neighboring chromosomes or nuclear structures as the nuclear envelope or the nucleolus (13). The genome organization changes in the course of differentiation, allowing for altered transcriptional activity as cells generate a specialized proteome and specialized function (14).

2.1 HISTONES

Histone proteins are positively charged proteins containing a poorly structured N-terminal region (‘tail’), a folded core and, in some cases, a C-terminal tail. There are four subtypes of core histones that together form histone octamers, H2A, H2B, H3 and H4. Two dimers of H3 and H4 form a tetramer. The tetramer interacts with two dimers of H2A and H2B to form the histone octamer wrapped by roughly 147bp DNA, with the histone tails protruding from the nucleosome. Additionally, there are linker histones H1 that are placed onto the nucleosomes forming chromatosomes, forming further chromatin compaction (15).

Histone genes are located in clusters in the genome and are expressed exclusively during S phase. Coupled to this, they exhibit replication dependent deposition following the synthesis of DNA. The transcripts of canonical histones also deviate from the majority of transcripts, as they are not polyadenylated, but rather end in a 3′ stem–loop sequence that is crucial in their regulation as it is targeted for degradation by the end of S phase (16).

Histone interacting proteins termed histone chaperones facilitate the dimerization of histones and assembly of the nucleosome. A wide range of histone chaperones with specialized functions have been discovered. For the assembly of H3-H4 dimers, the chaperones Heat shock cognate 71 kDa protein (HSC70), heat shock protein 90 (HSP90) and nuclear autoantigenic sperm protein (NASP) aid in folding and dimerization, Anti-silencing function 1 (ASF1) acts as a shuttle, Importin 4 (IPO4) aids in nuclear import, mini-chromosome maintenance protein 2 (MCM2) shields the DNA-binding interface of histones while Tonsoku-like protein (TONSL) protects the H4 tail of histones before assembly to the chromatin mediated by Chromatin assembly factor-1 (CAF1) (17). For H2A-H2B assembly, Nucleosome Assembly Protein 1 (NAP1) acts as a shuttle, being aided by Importin 4 (IPO9) for the nuclear import and is deposited to replicating chromatin with the help of the complex Facilitates chromatin transcription (FACT) (18,19). Chaperones do not only orchestrate the stepwise assembly of nucleosomes but also their eviction. Nucleosome free regions are often created transiently, followed by deposition of new histone, a process termed histone turnover (20).

2.2 GENOME COMPOSITION AND REPETITIVE ELEMENTS

Roughly 1.5-2.0 % of the human genome consists of protein coding genes, coding for between 20 000 and 25 000 genes (21). The remaining noncoding DNA, previously believed to be ‘junk DNA’, contains important regulatory DNA elements and essential repeat elements (22). The regulatory regions are promoters that provide the binding site for transcriptional machinery, enhancers that provide binding sites for proteins that aid in transcriptional activation, silencers that repress transcription and insulators that can serve as barriers protecting a gene from the activity of neighboring genes (23). But these regulatory regions are only a subset of the noncoding DNA. Roughly half of human and mouse DNA consist of repetitive regions (Figure 1, left) (26). These regions consist of tandemly repetitive and interspersed repetitive DNA. The tandem repeats are positioned in arrays called satellites and occupy different regions of the chromosomes. The bulk of centromere DNA in humans consists of α-satellites and of minor satellites in mouse (Figure 1, middle and right). The pericentromeric DNA also consists predominantly of repetitive arrays, in humans by satellites I, II and II, and in mouse by major satellites (27). Telomeres are also mainly made up of tandem repeats, consisting of hexameric DNA repeats (28). Functional transcripts that are transcribed from non-coding DNA are transfer RNA and (tRNA) and ribosomal RNA (rRNA), both classes having genes in tandem repeats (29,30).

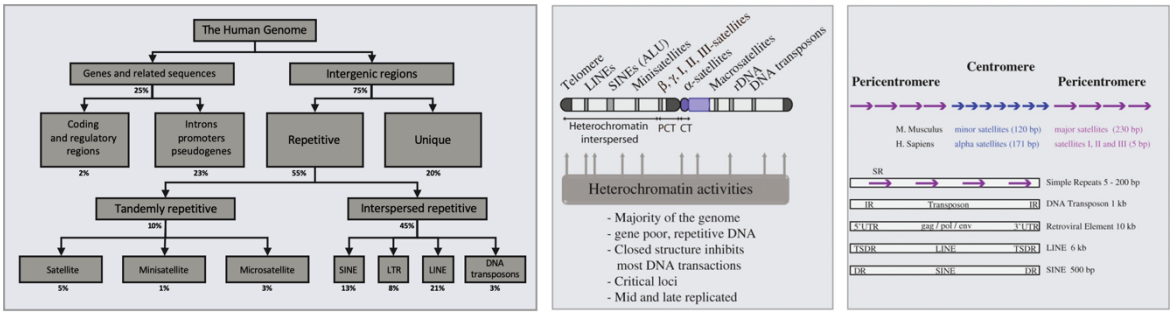

Figure 1. Composition of the human genome. Left) Percentage shares of functional and non-functional genetic sequences. Adapted from Jasinska and Krzyzosiak, 2004 (24). Middle and right) Position of repetitive DNA in the chromosome and representative configuration of selected repetitive sequences, respectively, adapted from Saksouk et al. 2015 (25).

Interspersed repetitive DNA is found throughout the genome and consists mainly of transposable elements (TEs), either of retrotransposons or DNA transposons. The three classes of retrotransposons are interspersed nuclear elements (LINEs), short interspersed nuclear elements (SINEs) and long tandem repeat (LTR) transposons (31). LINEs are autonomous parasitic elements that have two open reading frames, one encoding an RNA binding protein and one with endonuclease and reverse transcriptase activity, allowing for its own amplification and genome insertion. SINEs are smaller than LINEs and do not encode reverse transcriptase, hence they depend on the activity of LINEs for retrotransposition (32). The vast majority of LTR transposons are endogenous retroviral elements (ERVs), which have accumulated in vertebrate genomes due to retroviral infections throughout the course of evolution (33). ERVs encode three genes, group-specific antigen for viral capsid (*gag*), reverse transcriptase (*pol*) and protective lipid envelope (*env*), meaning that they are also autonomous for retrotransposition, and are flanked by long tandem repeats (34). Lastly, DNA transposons encode transposase, but do not generate copies by undergoing reverse transcription, but rather mobilize by excision from their loci and reintegrate elsewhere (35).

TEs are insertional mutagens that due to their mobility can alter gene expression in the locus they have incorporated into or excised from, which has an important role in the plasticity of eukaryotic genome and therefore on eukaryotic evolution (35). Although they do indeed have a positive role in regulatory innovation, they pose a threat to the genome integrity as they can incorporate to genes bodies or regulatory regions and disrupt gene expression, and therefore silencing mechanisms have emerged to keep the TEs repressed (36,37).

2.3 EPIGENETIC LANDSCAPE

The genetic information of DNA is either accessible or repressed, resulting in the regulation of transcriptional activity of genes. What dictates this outcome is a very complex network of epigenetic modifications that lead to allowance or inhibition of transcription factor binding to promoters of genes of interest. The main contributors for epigenetic regulation of chromatin are DNA methylation, the composition of nucleosomes (histone PTMs and histone variants) and regulatory RNA species (**Error! Reference source not found.**).

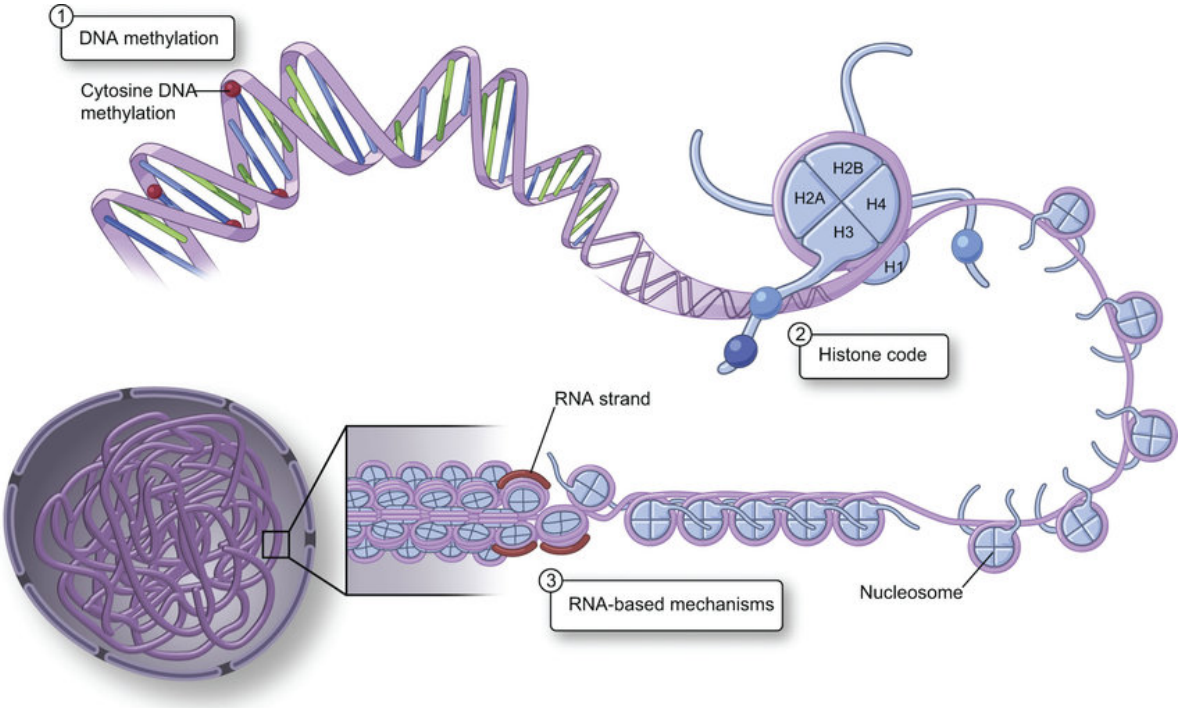

Figure 2. Fundamental mechanisms of epigenetic gene regulation. 1) DNA methylation, 2) histone tail modifications or histone substitution with histone variants and 3) RNA-mediated regulation by noncoding RNA species. Adapted from Yan, et al. 2010 (38).

2.3.1 DNA methylation

Cytosine, one of the four nucleotide bases of DNA, can be covalently modified by DNA methyltransferases (DNMTs) resulting in 5-methylcytosine (5mC) (39). DNMT1 maintains the DNA methylation as DNA is replicated, whereas DNMT3a and DNMT3b conduct *de novo* methylation (40). In GC-rich promoters or enhancers, an accumulation of 5mC results in transcriptional repression of the gene they control. Although this mark is a covalent modification it is also reversible as it can be removed through either passive or active pathways. Passive demethylation occurs through a series of dilution as DNA replicates, but 5mC is not deposited on the newly synthesized DNA strand. Active demethylation requires either deamination of 5mC by the activation-induced cytidine deaminase (AID) and apolipoprotein B mRNA editing enzyme (APOBEC) family members or stepwise oxidation by the Ten-Eleven Translocation family of dioxygenases (TETs). In the first case, 5mC is converted to thymine and due to base excision repair of DNA, the thymine is later converted into cytosine (41). In the second case, the TET enzymes modulate the oxidation of 5mC to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) by the Ten-Eleven Translocation family of dioxygenases (TETs), followed by thymine DNA glycosylase (TDG)-mediated base excision repair where 5caC is substituted with cytosine (42). The catalytic activity of the TET enzymes requires α-ketoglutarate and iron ions that interact with the enzyme's double stranded beta-helix (DSBH) domain (43). For the catalytic reaction, a conserved histidine-x-aspartic acid (HxD) motif has been found to be of essence at the DSBH of the enzymes (44). For binding to DNA, TET1 and TET3 have a cysteine-x-x-cysteine (CXXC) domain on their N-termini, while TET2 is dependent on its interaction with Inhibition Of The Dvl And Axin Complex (IDAX) for DNA binding (45).

DNA methylation has a pivotal role throughout differentiation, as it is depleted from genes that are activated and deposited to silenced regions (46). Deregulated DNA methylation can lead to genomic instability and deregulation of genes essential for cell homeostasis, leading to disease initiation and progression (47,48). To date, several diseases show aberrant DNA methylation, as multiple forms of cancer, immune-mediated diseases, imprinting disorders and atherosclerosis (49–52)

2.3.2 Histone post-translational modifications

Histone modifying enzymes decorate histone tails with covalent post-translational modifications (PTMs) in an often-reversible manner. Histone modifying enzymes can be further classified into 'writers' or 'erasers' according to their activity in establishing or removing PTMS, respectively (53). Histone PTMs include acetylation, methylation, and phosphorylation, amongst others and are recognized by 'reader' proteins, adding another layer of information. Acetylation is carried out by lysine acetyltransferases (a writer), where the acetyl group of acetyl-CoA is transferred to the positively charged lysine, resulting in a residue with neutral charge. This is an event that can be reversed by histone deacetylases (an eraser). Methyltransferases require the donor S-adenosyl-L-methionine for the methyl group transfer to either a lysine or an arginine residue, and it is reversed by demethylases. Two classes of ATP-dependent kinases conduct phosphorylation, one that modifies either serine or threonine, and another that modifies tyrosine residues. The result of phosphorylation is that the modified residue receives a negative charge, and the inverse reaction is conducted by phosphatases. Lysine ubiquitinates deposit ubiquitin onto lysine residues, which can be removed by ubiquitin hydrolases. Often, reader, writer and/or eraser modules are found in the same macromolecular complex or even on the same polypeptide chain, suggesting a dense crosstalk (either synergistic or antagonistic) between different PTMs. A common phenomenon that occurs upon the deposition of a PTM is the recruitment of the same or other modifying enzymes leading to an amplification of the signal or spreading onto neighboring nucleosomes (54,55).

Many of the PTMs have a strong correlation with either active or repressed chromatin, depending on which histone tail residue has received the modification. An example of an active mark is H3K4me3, meaning trimethylation of the Lys4 residue of histone H3. This mark is highly enriched at promoters of active genes, deposited by the MLL family of writer enzymes, read by the chromatin remodeler Chromodomain Helicase DNA Binding Protein 1 (CHD1) and removed by the erasers Lysine Demethylase 5B (KDM5b) and Lysine Demethylase 2B (KDM2B) (56,57). CHD1 chromatin remodeling activity is thought to open promoter regions to facilitate recruitment of transcription factors and RNA Polymerase (57). A contrasting, i.e. repressing, mark is H3K27me3, trimethylation of Lys27 of histone H3, which is deposited by the Polycomb group protein 2 (PRC2), read by PRC1 and erased by Kdm6a/b (58). Interestingly, both active and repressive marks can coincide in the same regions, such as bivalent gene promoters of stem cells, where the abundance of each dictates the expression status of the gene of interest (59–62).

Unique structural features of histone variants, and in some cases their PTMs, confer specificity to interactions between histones and histone chaperones. Chaperones have an essential role in the regulation of chromatin dynamics, as they bind histones, mediate the assembly of nucleosomes and their removal. The specificity of the chaperones, i.e. how many different histones it can interact with, depends on the recognition site of the histones. There are some promiscuous chaperones that bind to the conserved histone backbone (that is also used to contact DNA thus essentially invariant in evolution), allowing for multiple binding. NAP1, for example, binds both to H3-H4 and H2A-H2B dimers (63). Others are highly selective for variable regions of the histones, e.g. CAF1, which exclusively interacts with histone H3 (10,64). The search for histone chaperones has led to the identification of a long list of proteins that interact with histones, but many have yet to be identified.

2.3.3 Histone variants

All eukaryotes have histones, and in all of them, histone variants have been found. Histone variants have been evolutionary diversified from ancient canonical histones (65). While multiple genes encoding canonical histones reside in clusters in the genome, there are single genes of histone variants in distant genomic locations. The genes for the histone variants differ from their canonical counterparts as they have introns and their transcripts can therefore be subjected to splicing and their mRNA is polyadenylated, suggesting a fundamentally different regulation at the level of transcription, translation and RNA degradation (66,67). Histone variants differ from the canonical histones in amino acid sequence, they are replication independent and have different mRNA processing (68,69). Many of the histone variants are conserved, but several, especially isoforms, are specific to certain classes of organisms. Since some histone variants are as conserved as the canonical histones, it is believed that they play an important role for chromatin structure and/or the regulation of gene expression. In embryonic development it is interesting to study when histones and histone variants start being expressed, particularly in the earliest stages after oocyte fertilization where histone take the place of sperm protamines to generate *de novo* chromatin signatures. As the organism develops, these signatures change dynamically as they alter their localization in response to different cues (70). Furthermore, not only one histone variant replaces a canonical histone upon forming a nucleosome. There are both homo- and heterotypic nucleosomes in regard to histone variants, where either one or both canonical histones of a subtype have been exchanged with variants (71–74). Carrying different histone variants might alter the stability interactome of certain nucleosomes.

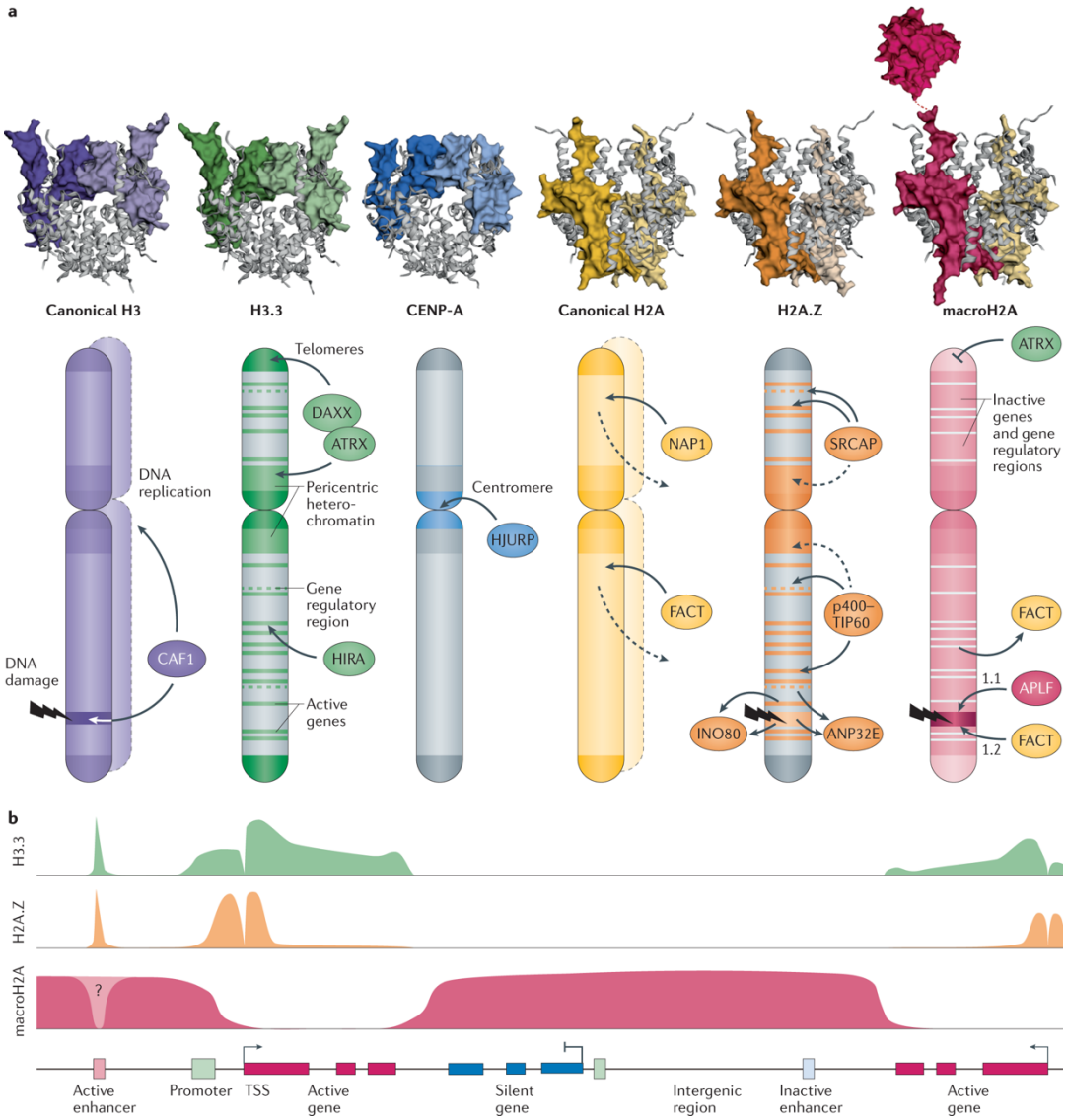

Figure 3. Histone variants have specialized histone chaperones and occupy specific chromatin regions both a) genome wide and b) at single gene resolution. Adapted from Ghiraldini et al. 2021 (75).

Histone variants play an important role in stem cells and development, where they dynamically reside in distinct chromatin regions, regulating chromatin structure for stem cell self-renewal and for progression toward differentiation (76–78). Since the histone variant genomic location changes with the cell cycle, it is interesting to further understand their dynamics, and role in gene expression regulation (79). For example, the H2A variant H2A.X is known be recruited to DNA damage sites, upon being phosphorylated, where it plays a role in DNA damage response (80). Another histone H2A variant that has been linked to a certain biological function is H2A.Z, which is often found adjacent to the transcription start site of genes (81). Two more histone H2A variants, macroH2A and H2A.Bbd have been studied extensively. MacroH2A is the structurally most deviant histone variant in comparison to its canonical counterpart, where the molecular size is three times larger than canonical H2A (82). This protein is highly linked to high chromatin compaction and gene repression and its levels increase upon differentiation (83,84). H2A.Bbd on the other hand acts opposite to macroH2A, as it is smaller than canonical H2A and is linked to gene activation (85–87). Histone H3 has two very well studied variants, H3.3 and CENP-A. H3.3 has been described to have dynamic roles serving as a replication-independent replacement histone (88–90). CENP-A on the other hand is specifically linked to

centromeric chromatin and plays an essential role for correct segregation of sister chromatids during mitosis (91,92).

Deregulation of the histone variants, due to aberrant cell specific expression levels or due to mutations to either them or their chaperones, has been associated with various diseases such as cancer. The more we understand about histone variants and their different mechanisms, the better we will understand diseases that arise due to aberrant function of these proteins, resulting in improved treatment or disease prevention. In this thesis, four histone variants are on focus; the H3 variants H3.3 and CENP-A and two H2A variants H2A.Z and macroH2A.

2.3.3.1 H3.3

Histone variant H3.3 is a highly conserved histone H3 variant in mammals encoded by the two genes *H3f3a* and *H3f3b* (93). Its protein sequence has only four amino acids that differ from canonical H3.2 and five that differ from H3.1 (94,95). These few differences result in H3.3 having a different interactome from its canonical counterparts, resulting in its specialized histone chaperone complexes Histone regulator A (HIRA) together Calcineurin-binding protein 1 (CABIN1) and Ubinuclein 1 (UBN1) and Death Domain Associated Protein (DAXX) which interacts with the chromatin remodeler Alpha Thalassemia Retardation Syndrome X-Linked (ATRX) (96,97). Due to this specific interactome, H3.3 undergoes a nonuniform deposition in the genome (98). Its deposition undergoes incorporation to active regions following a pathway controlled by HIRA, and to repressed regions mediated by DAXX (94,99).

H3.3-H4 dimers bound by UBN1/2 and associated to HIRA, with nuclear import facilitated by ASF1a, are deposited to transcribed genes in a cell cycle independent manner (100,101). For this, HIRA and ASF1a form a complex with UBN1 and CABIN1, resulting in H3.3 incorporation both at active promoters and gene bodies (102,103). Furthermore, HIRA interacts with replication protein A (RPA) for the H3.3 deposition at active enhancers (104).

On the other hand, H3.3 bound by DAXX and ATRX is deposited to repressed chromatin with deposition at pericentromeric chromatin, telomeres and endogenous retroviral elements. At pericentromeres, H3.3 has a pivotal function for heterochromatin establishment, found to be linked to its methylation of lysine 27, as upon mutation of that residue, there is HP1 displacement and pericentromeric transcription upregulation (105). Deposition of H3.3 to telomeres has been proven essential for spreading the repressive histone mark H3K9me3 for conserving a silent chromatin state (106,107). In mouse embryonic stem cells, DAXX and ATRX deposit H3.3 at endogenous retroviral elements (ERVs), which is important for silencing of those regions (88,108).

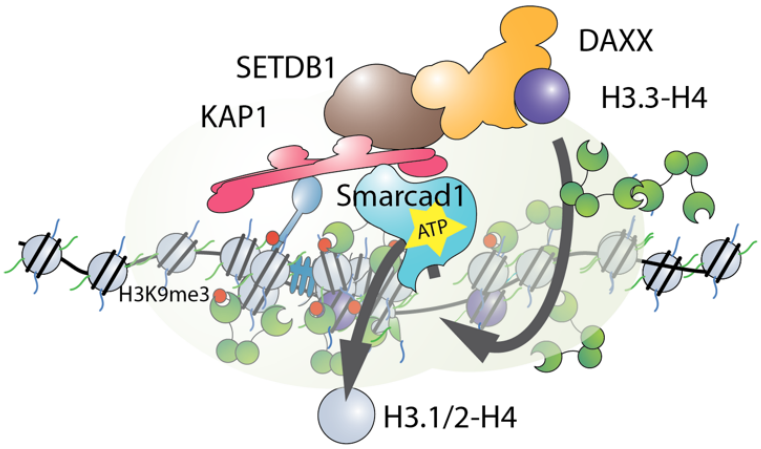

Figure 4. Model for molecular mechanism underlying dynamic heterochromatin at mESC ERVs. KAP1 is recruited by DNA-binding KRAB-ZFPs to interstitial repeats, where it orchestrates heterochromatin formation and maintenance by recruiting the methyltransferase SETDB1, histone H3.3 chaperone DAXX and chromatin remodeler Smarcad1. Through SETDB1-mediated histone H3K9me3 methylation, HP1 is recruited and presumably contributes to heterochromatin formation by compaction. Adapted from Navarro et al. 2020 (109).

At ERVs, Krüppel-associated box (KRAB)-associated protein 1 (KAP1) (also known as Trim28) recruits DAXX, allowing for deposition of H3.3, which in its turn aids in efficient KAP1 association with chromatin, followed by ERG-associated protein with SET domain (ESET) (also called SETDB1) recruitment resulting in H3K9me3 methylation and hence, chromatin compaction (88,110). ERV heterochromatin in mESCs is dynamic due to nucleosome eviction by the chromatin remodeler Smarcad1 followed by redeposition of H3.3 by DAXX with the aid of Morc3 (109,111,112). In H3.3 knock out cells, ERVs have a more open chromatin state, confirming the importance of H3.3 for ERV silencing (109).

Deregulation of proper H3.3 activity either by mutations to the H3.3 genes or to its chaperones has been linked to diseases. Mutation of its residue Lys27 to methionine has been linked to high-grade pediatric gliomas, coinciding with a global reduction of H3K27me3 (113). This H3.3 mutation is also prevalent in midline gliomas, where this mutation not only affects the interaction with PRC2, but also affects the mitotic H3.3 specific Chk1 phosphorylation of Ser31 (114,115). Mutations of Gly34 to arginine or valine are also linked to gliomas, with a poor prognosis both in pediatric and adult patients (116,117). Mutation of the Gly34 to tryptophan or leucin are correlated to giant cell tumors of the bone and Lys36 to methionine is prevalent in chondroblastoma (118). Mutations to DAXX and ATRX have been linked to pancreatic neuroendocrine tumors, neuroblastoma and sarcomas (119–121). Alpha thalassemia X-linked intellectual disability (ATR-X) syndrome is an X-linked condition involving severe intellectual disability in males, caused by ATRX mutations (122).

2.3.3.2 CENP-A

Centromeric specific protein A (CENP-A) is a conserved histone variant in eukaryotes that has an essential role during mitosis and meiosis as it is one of the main factors for centromere-associated network of proteins (CCAN) assembly, together with other centromeric specific proteins (123–125). The centromere has a unique 3D structure, where chromatin containing CENP-A is interspersed with canonical nucleosomes carrying H3K4me2 (126,127). CENP-A

differs significantly from H3 in its N-terminal domain, where lysine residues that can be acetylated in canonical H3 are missing in CENP-A, which could explain why CENP-A is more retained at compact regions (128–130). Aurora proteins and other centromere specific proteins are recruited to where CENP-A has been positioned and form the kinetochore. Kinetochore assembly in M phase is followed by the separation of the sister chromatids, leaving half of the CENP-A proteins in each sister chromatid (91,124,131). For faithful maintenance of CENP-A, new CENP-A need to be placed in the centromeric region, after the DNA has been replicated. Interestingly, the simplest hypothesis that new CENP-A joins existing CENP-A at the centromere at the wake of replication does not occur. Instead, CENP-A is repositioned into the centromere during G1 phase, a process that is essential for maintaining the integrity of the centromere (132,133). During replication, both canonical H3 and histone variant H3.3 are positioned in the centromeric region and there is a model suggesting that H3.3 acts as a placeholder for CENP-A to be deposited in the next G1 phase (134,135). Replenishment of CENP-A also occurs due to centromeric transcription (136,137).

The centromeric region contains repetitive sequences; alpha satellites in human cells and minor satellites in murine cells (126,138,139). CENP-A is deposited into the repetitive sequences, but it is not DNA sequence identification by CENP-A that dictates the centromeric deposition, but rather the interactors that lead them there (140,141). An essential interactor for the centromeric deposition of CENP-A is the histone chaperone HJURP (Holliday junction recognition protein) (142). HJURP forms a complex with CENP-A and interacts with the Mis18 complex, which associates to the centromere in late anaphase or telophase to early G1 (143–146). The timing on when HJURP positions CENP-A in the centromere is regulated by cell cycle dynamic PTMs on the histone tail of CENP-A. For example, Cdk1 phosphorylates Ser68 of CENP-A, which stalls its interaction with HJURP in early mitosis until Cdk1 levels decrease and Ser68 dephosphorylation occur toward late mitosis (147). Another example is that Lys79 and Lys124 are acetylated in G1 to S, but the same residue is instead altered to monomethylation during mid to late S phase (148,149). Many mutations of the amino acid sequence of CENP-A have been investigated resulting in a further understanding of each domain (150,151). This strategy has been crucial for the identification of essential sites required for interaction with chromatin remodelers and the characterization of what residues that are modified. One PTM that occurs to all CENP-A proteins is trimethylation of Gly1 by N-terminal methyltransferase 1 (NRMT1), which is essential for the correct deposition of CENP-A into centromeric chromatin (152). Several cell cycle dependent PTMs have been identified, but for some their role is still unclear. For example, it is debated whether phosphorylation of S7 by Aurora B is necessary for the centromeric incorporation of CENP-A (153–155). The same uncertainty exists for phosphorylation of Ser68p and ubiquitinylation of K124 (147,156,157). Upon overexpression, CENP-A can be erroneously incorporated into chromosome arms and this has been observed in several cancers, such as solid tumors in the breast, central nervous system, lung and endometrium (158,159). (160–162). In all of those, CENP-A overexpression is a marker for lower time of tumor recurrence and lower survival rates for patients

The chaperone misplacing the excess CENP-A to the chromosome arms are DAXX and ATRX, two chaperones of the H3 variant H3.3 (160,163,164). If CENP-A is not evicted from the chromosome arms it can result in chromosome instability, where the excess CENP-A accumulates in hotspots and fragility is increased (134,160). Neocentromeres can follow as a consequence in the these sites and can cause detrimental effects (162,165). Other issues related to this CENP-A generated chromosome instability are lagging chromosomes and DNA double strand breaks (125,166–168). Inversely, CENP-A has also been linked to be present in DNA double strand breaks (169). Even in normal cells, there is a low fraction of CENP-A that is deposited to the chromosome arms, but there is a mechanism to prune out these ectopically deposited histones (170). During DNA replication the non-centromeric CENP-A is evicted and exchanged with canonical H3 or its variant H3.3 (137,171,172). HJURP overexpression has been documented in breast cancer, where it is an indicator of a poor prognosis for disease-free survival (173). In gliomas it has a similar effect, where overexpression of HJURP is also inversely correlated with patient survival and HJURP knockdown experiments of glioblastoma cell lines result in higher sensitivity to radiotherapy (174,175). Furthermore, studies in hepatocellular carcinomas have shown that the increased levels of HJURP drive cell cycle progression by destabilizing the tumor suppressor p21 (176).

2.3.3.3 H2A.Z

H2A.Z is often found in the mammalian genome flanking the nucleosome-depleted region of active genes (177). Due to this, H2A.Z has been used for decades as a marker for active gene expression (71,177,178). The role it plays for gene activation is still not fully understood, but one of the reasons might be that H2A.Z/H2B dimer is less stable in the nucleosome than its canonical counterpart (179). In active regions, it is often found together with the histone variant H3.3, and coinciding with histone markers of active gene expression, suggesting that H2A.Z might be involved in sequential recruitment of active markers (180–182). Another role of H2A.Z appears to be the recruitment of SUV420H1 methyltransferase to H2A.Z nucleosomes, which promotes dimethylation of H4K20, enabling the recruitment of the origin-recognition complex ORC1 (183). Furthermore, H2A.Z is also found at enhancer regions, together with hypomethylated DNA, RNA polymerase II recruitment and eRNA production (177,184–186). Another region where H2A.Z is found is at the pericentromeric regions, where it colocalizes with H3K9 methylation, and whereto the chromatin architectural and phase separation related protein HP1 is recruited (187). The histone chaperones that position H2A.Z into chromatin are PWWP Domain Containing 2A (PWWP2A), Snf2 Related CREBBP Activator Protein (SRCAP) and NuA4/Tip60/p400 (188–190). PWWP2A is one of the most recently discovered histone chaperones of H2A.Z, which has one domain that binds to the C-terminal region of H2A.Z and another DNA binding domain. PWWP2A binds predominantly to H2A.Z nucleosomes placed in promoter regions or highly transcribed genes (188). Absence of PWWP2A results in increased levels of acetylated H2A.Z and H3K27 (191). Both SRCAP and NuA4/Tip60/p400 contain the subunit YL1 which has been identified to bind specifically to the alpha helix C domain of H2A.Z (189,190). Acidic Nuclear Phosphoprotein 32 Family Member E (ANP32E) is a chaperone linked to H2A.Z chromatin eviction (178,184,190).

H2A.Z at the first nucleosome downstream of the transcription start site has been found to be globally lost in overexpression of Anp32e, and the opposite effect has been seen in knock down experiments of the chaperone (192). A second chaperone known to evict H2A.Z from chromatin is INO80, which has also been linked to homologous recombination (193,194).

Many PTMs of H2A.Z have been characterized and research is ongoing to further understand what roles each play in different biological processes and through the cell cycle (195,196). They are dynamic throughout the cell cycle in regard to H2A.Z, and ongoing studies are trying to map the cell cycle resolved dependency and effect of each. Seven residues on the N-terminal tail of H2A.Z have been known to be acetylated; Lys4, Lys7, Lys11, Lys13, Lys15, Lys37 and Lys 101 (197,198). The first five of them are linked to active gene expression. The proteins Tip60 and p300 acetylate the first three residues, while the histone acetyl transferase for the remaining is still not characterized (197). Lys37 has not only the possibility to be acetylated, but it can also be ubiquitinylated, in which case it is linked to gene repression. Ubiquitinylation has been discovered for eight residues; Lys13, lys15, Lys74, Lys77, Lys115, Lys120, Lys121 and Lys125 (199). Of these, the last three are ubiquitinylated by Ring1b, and have been linked to gene repression. Deubiquitylation is carried out by Usp10, where removal of this repressive marker aligns with transcriptional activation (200). Setd6 or Set7 can methylate Lys4, Lys7 and Lys13, where methylation of Lys7 has been linked to repression(201). Oppositely, C-terminal methylation of Lys101 by Smyd3 has been linked to gene activation, as Anp32e interaction is prevented (202). Phosphorylation of four residues has been identified; Ser9, Tyr60, Ser98 and Ser116, but their modes of action remain to be characterized (203). Lys120 can be subjected to acetylation, ubiquitinylation and SUMOylation, where only the ubiquitinylated residue has been linked to a (repressive) role (196,204).

H2A.Z exists as three isoforms in human and two in mouse, where they share H2A.Z.1 and H2A.Z.2, while H2A.Z.2.2 is a human specific isoform (190,205). Although the isoforms result in globally similar distribution, H2A.Z.2 associates more with H3K4me3, and the two major isoforms regulate different neuronal genes (179,205). H2A.Z has higher expression levels in the brain, where it is believed to have specific modes of action to regulate synaptic genes (206–209). In brain tissue, H2A.Z interacts with Setd2, promoting H3K36 trimethylation of promoters for activation of Nkx2 homeobox 4, an important factor of neurogenesis (206). Upon depleting H2A.Z in the brain, neuronal differentiation and correct dendrite formation are disrupted (206). In aging neurons of the hippocampus, H2A.Z is observed to accumulate, especially on exons and 5'UTR regions (208). H2A.Z deregulation has been identified in different disease models, such as cancers and mental retardation (81,190,210). Overexpression of H2A.Z has been linked to melanoma, and solid tumors in the bladder, prostate and breast. In prostate cancer, acetylated H2A.Z has an increased deposition to TSS of oncogenes and reduced levels at tumor suppressor genes (211). Furthermore, floating harbor syndrome is a mental retardation syndrome that is directly linked to a C-terminal truncation of SRCAP, altering its localization from the nucleus to the cytoplasm, resulting in a reduced placement of H2A.Z.2 into the AT-rich enhancers, and the disease phenotype can be rescued upon overexpression of the H2A.Z isoform (212).

2.3.3.4 MacroH2A

MacroH2A is the most deviant variant of histone H2A, having a molecular weight of three times its canonical counterpart (213). This protein contains three major compartments; a H2A histone fold, a linker domain and a globular macrodomain (82,214–217). The histone fold only shares 64% similarity with canonical H2A (73,218). The linker region has a basic nature as it contains a series of lysine residues, and it is believed to help in maintaining a compact chromatin formation, often found to be similar to the function of the linker histone H1 (215,216,219,220). The macrodomain has the ability to bind ADP-ribose and histone deacetylases (82,214,221,222).

This protein is most often detected in heterochromatin, coinciding with repressive histone markers H3K27me3 and H3K9me3 (85,223). The expression levels of macroH2A vary significantly in different tissues, with very low levels in stem cells and increased in differentiated cells (84,214,224). There are two genes encoding macroH2A, where one has two splice variants, resulting in a total of three isoforms; macroH2A1.1, macroH2A1.2 and macroH2A2. The three have been described to play different roles, where macroH2A1.1 has been found to interact though its macrodomain with the poly-ADP-ribose polymerase 1 (PARP1) enzyme, which is a cellular stress sensor (214,224–228). To study the essentiality of macroH2A, transgenic macroH2A null mice were created, where developmental and postnatal growth was reduced and reproductive efficiency decreased (229,230). In knock down experiments the heterochromatin architecture and nuclear organization were disrupted (223). Several posttranslational modifications have been discovered for macroH2A, which are dynamic throughout the cell cycle. On the N-terminal tail, Lys17 acquires a methylation mark, while several modifications occur on the C-terminal tail. Lys115, Lys116 and Lys122 are acetylated, Lys122 can also be dimethylated, Thr128 and Ser167 are phosphorylated, and the residue Lys237 in the macrodomain is trimethylated (231–233). Studies have linked transcription to be an active way to prune out macroH2A from active chromatin sites, leaving higher accumulation in repressed chromatin (234). MacroH2A plays a role in stem cells, although its expression levels are low (83,84,235,236). Upon differentiation of the naïve to primed stem cells, global chromatin changes occur, where chromosomal domains are generated and a 3D genome structure starts to occur and macroH2A is observed to be linked to the compact heterochromatic structures (224,237,238). Deregulation of macroH2A has been linked to several diseases, for example downregulation of this variant has been documented in melanoma, lung cancer and colon cancer while its upregulation is evident in hepatocellular carcinoma (239–242). In gastrointestinal cancers, the isoforms macroH2A1.1 and macroH2A1.2 have opposite roles as oncosuppressors or oncogenes, in an organ specific manner (243). Another example is the overexpressed levels of macroH2A observed in Huntington disease. This disorder is characterized by the loss of function of huntingtin protein, resulting in DNA damage and where macroH2A is suggested to prevent the recruitment of DNA repair proteins (244).

2.4 EPIGENETIC MECHANISMS IN PLURIPOTENT STEM CELLS

Early embryogenesis follows a conserved trajectory in mammalians species (245). The fertilized egg, i.e. the zygote, rapidly undergoes multiple rounds of cell division, increasing the cell mass with progressively more specialized cells. The zygote is a totipotent stem cell that gives rise both to the inner cell mass in the blastocyst and to the extraembryonic trophectoderm, which envelopes the embryo and later in development forms the placenta (**Error! Reference s ource not found.**). The cells in the inner cell mass of the blastocyst are pluripotent stem cells that have the potential to specialize into the three germ layers (ectoderm, mesoderm and endoderm). An important event of embryogenesis is implantation of the embryo, where it attaches to the uterus, separating the early stages of development into pre- and post-implantation (246,247).

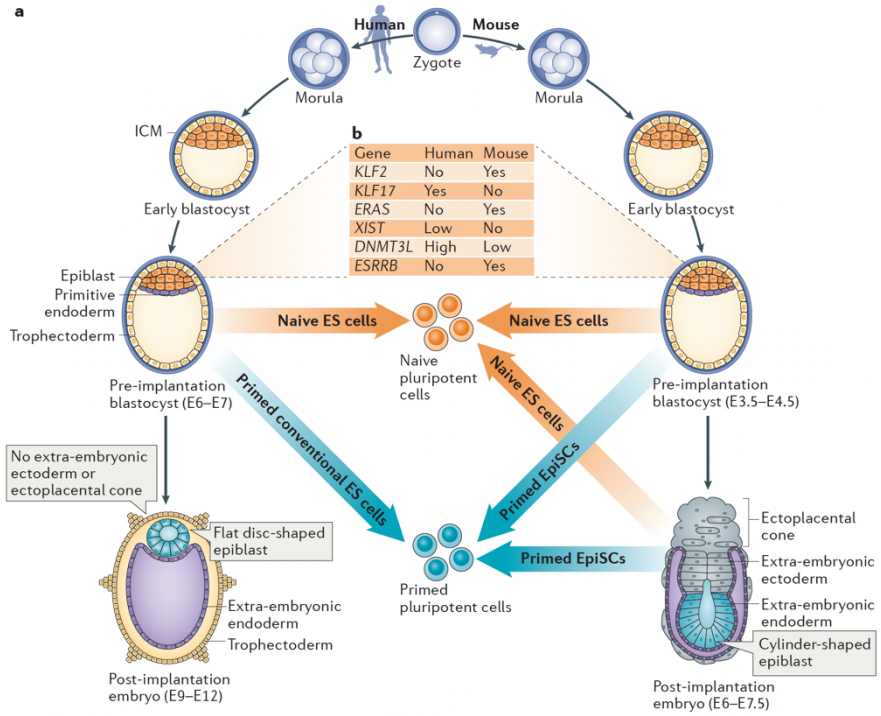

Figure 5. A model that classifies 'relative naïvety' within the spectrum of naïve to primed stem cells. a) similarities and differences in early pre- and post-implantation in vivo development in mice and humans. b) Important differences in gene expression in both models. Adapted from Weinberger et al. 2016 (245).

Both human and mouse pluripotent stem cells can be extracted from embryos and cultured in *ex vivo* where they self-renew, which is of essence for studying the molecular mechanisms in early development. The pluripotent cells extracted from the inner cell mass are called naïve (or ground state) stem cells, whereas cells extracted from the post implantation stages are termed primed pluripotent stem cells. In culture, depending on the growth conditions, pluripotent cells can be controlled to switch between those two pluripotent stages (245). Although human and mouse early development is similar, it should be noted that there are important differences as in the timing of when the zygotic genome is activated, when X chromosome inactivation occurs and there are important transcriptomic differences (248–250).

The progression through development is mediated by epigenetic mechanisms. Upon fertilization, the paternal genome DNA methylome is drastically abolished, while the maternal genome undergoes gradual demethylation leaving the genome in a globally hypomethylated state at the morula stage. This active demethylation is controlled by the TET enzymes. At the blastocyst state, *de novo* methylation is initiated and carried out by DNMT3a and DNMT3b, where new methylation patterns are re-established in the growing embryo. This methylation is later maintained by DNMT1 (41,251). The active demethylation in early development is attributed to the TET1 paralogs TET1 and TET2. As DNA methylation activity increases in primed pluripotency, both 5mC and its oxidation derivatives increase in genomic abundance (Figure 6) (252).

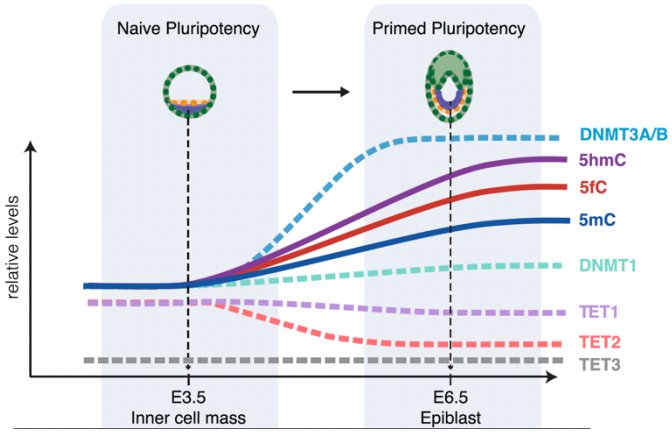

Figure 6. Distinct contributions of different DNA modifying enzymes during the transition from naïve to primed pluripotency in relative levels. Adapted from Mulholland et al. 2020 (252).

As the repressive DNA methylation mark is lost in naïve pluripotency, the repressive chromatin marker H3K27me3 is globally at high levels and the active mark H3K4me3 has low global levels. Upon progressing to primed pluripotency, there is a slight global reduction of H3K27me3 and increase of H3K4me3 (253). This is important as these two marks coincide in bivalent promoters of developmental genes that are poised for rapid activation as results of differentiation cues. In naïve cells, the genes under the control of bivalent promoters are transcribed at very low levels, but depending on the developmental stimuli they either have an increase of H3K27me3 and are silenced, or gain H3K4me3 facilitating transcription (61).

Histone variants also play important roles during development. H2A.Z has a role in mediating the general gene expression for self-renewal maintenance and differentiation potential (254–257). H2A.Z has been found to be enriched in the majority of homeodomain genes, which regulate embryonic development (258). H2A.Z also co-localizes with Polycomb group proteins at bivalent gene promoters of stem cells and it has been shown that loss of one leads to loss of the other, suggesting interdependency and non-redundant function (259–261). H2A.Z nucleosomes have higher H3K27me3 levels than the canonical counterpart, and H2A.Z depletion leads to global decrease of that repressive marker (260). H3.3 has also been linked to interact with PRC2 in embryonic stem cells, aiding in the correct deposition of H3K27me3 at developmental genes (262). As mouse cells differentiate toward primed, X chromosome inactivation occurs and the main genomic territory acquiring an accumulation of macroH2A is one of the two female X chromosomes, which gets silenced during differentiation. MacroH2A

is detained in that inactive chromatin and it is believed to have an important role in maintaining that specific chromosome repressed (263–265). Interestingly, the phosphorylated subset of the pool of macroH2A is excluded from the inactive X chromosome (266–268).

During early embryogenesis, transcriptional silencing of transposable elements occurs, mediated by KAP1 and DNA methylation (110). But there is a subset of endogenous retroviral elements that avoid silencing and are active during development and might aid in mediating the embryonic genome activation. Some of the ones found in humans to date are MLT2A1, LTR7B, LTR7Y and LTR5_Hs (269,270). In mouse embryonic stem cells, the ERV subfamily of MuERV-L are very rapidly transcribed and essential for progression during development (271)

3 RESEARCH AIMS

The overall aim of this thesis was to study the dynamic epigenome in stem cells in regard to key epigenetic modulators involved in DNA demethylation, heterochromatin establishment and chromatin organization fine-tuning histone variants.

The specific aims of the individual projects were:

I. To unravel the non-catalytic functions of TET1 in mESCs by studying the epigenome of cells with TET1 wild type, catalytic mutant and knock out.

II. To functionally characterize the epigenetic role of PRC2 in naïve and primed hESCs.

III. To quantitatively study the dynamics of histone variant H3.3 in mESCs in regards to its chromatin deposition and turnover and to dissect its temporal interactome.

IV. To quantitatively catalogue the dynamics of histone variants H3.3, CENP-A, H2A.Z and macroH2A in mESCs, focusing on their deposition to repetitive DNA regions.

4 MATERIALS AND METHODS

Multiple methods were used in the collaborative studies. A thorough description of those are found on the individual papers attached to this thesis. Selected methods that were highly used are covered in this chaper.

4.1 CELL CULTURE

Mouse embryonic stem cells (mESCs) were cultured feeder-free in 0.1% gelatin-coated flasks under standard conditions (5% CO_2, 5% O_2, 37 °C) in knockout DMEM medium, 2 mM Alanyl-glutamine, 0.1 mM non-essential amino acids, 15% fetal bovine serum FBS, 0.1 mM β-mercaptoethanol, ESGRO Leukemia Inhibitory Factor (LIF) and 2i;1 μM MEK inhibitor PD032590 and 3 μM GSK3 inhibitor CHIR99021. Cells were passaged every other day using TrypLE to avoid differentiation. Prior to microscopy experiments, to obtain naïve-like colonies, cells were grown in serum-free conditions with ESGRO complete basal medium supplemented with 2i and LIF.

4.2 AMBER SUPPRESSION

Genetic code expansion (GCE) is a synthetic biology approach that allows for site specific incorporation of non-canonical amino acids (ncAA) with defined chemical properties into proteins of interest (POIs). In Paper III and IV we used amber suppression, as the amber codon is the least used stop codon in mammals (272). In this approach, the amber stop codon is reassigned into an elongation codon by expanding the translation machinery of the cell by introducing a tRNA and tRNA synthetase pair that is orthogonal to the host. This pair, being the amber suppression machinery, has its origin in the archaeon *Methanosarcina mazei*, and allows for incorporation of pyrrolysine. In Elässer Lab, efforts have been made for allowing the amber suppression machinery to also allow for incorporation of other ncAAs.

4.2.1 Generation of amber suppression cell lines

To generate the stable amber suppression cell lines, we used three plasmids. One carrying the genes for expression of the amber suppression machinery, a second containing the gene of interest with an early amber stop codon and a third encoding PiggyBac transposase for transposase-mediated integration of the first two constructs following a protocol established in Elsässer Lab (273). The constructs also carry independent drug resistance markes, enabling the selection of stably transfected cells through double drug selection. We selected cells in with progressively higher drug concentrations to ensure high amber suppression and abundance of the transcripts of interest. We used polyclonal cells in bulk for validation of cell lines, but monoclonal cell lines were selected and used for imaging experiments to ensure that cells were genetically identical, minimizing the population heterogeneity.

To produce the proteins of interest, ncAA was supplemented to the growth media of the cells, which is labelled the pulse. To stop translation of the protein of interest, the growth medium was aspirated, cells were washed once with phosphate buffer saline and new medium was **replenished (**Error! Reference source not found.**). In the** absence of ncAA no more POI was g enerated, allowing for studies of produced protein during a limited window of time.

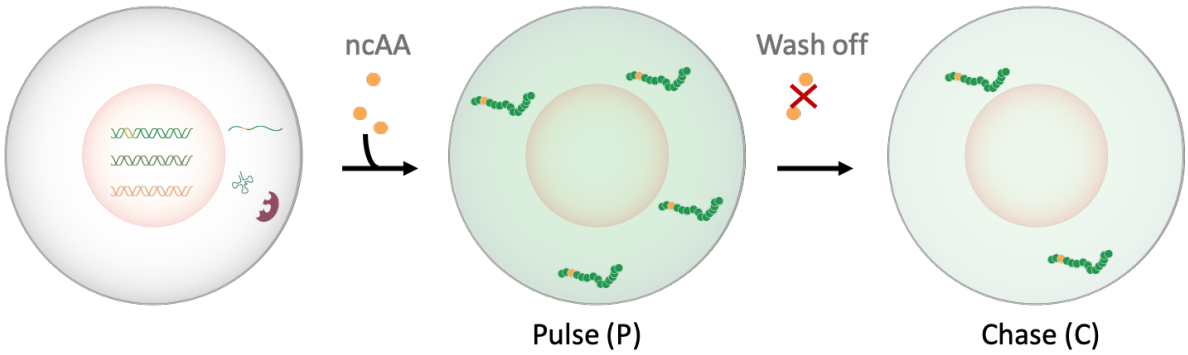

Figure 7. Pulse and chase experiments conducted on engineered mESCs. Left) stably integrated genes for pylT and PylRS are constitutively expressed, generating an always present but inactive amber suppression macihnery. A gene of interest, in this case GFP (green DNA helix), carries an early amber stop codon (yellow). The gene is transcribed, but as the amber suppression machinery is inactive, translation only occurs until it reaches the amber stop codon and there is no full-length protein production. Middle) upon pulsing, i.e. supplementing the growth medium with non-canonical amino acid (ncAA), the amber suppression machinery is activated, allowing for the amber stop codon to be used as an elongation codon, generating full length protein of interest carrying the ncAA. Right) amber suppression is simply stopped by exchanging the media. As no more ncAA is present, the amber suppression machinery activity stops and with it, also the translation of the protein of interest. This allows for a controlled window of time for protein pulsing, generating controlled protein amounts that can be studied over time.

In Paper III, the cells expressed a construct for H3.3 with a C-terminal triple HA-tag. The amber stop codon was positioned on position K27, where the ncAA acetyl-lysine was placed upon pulsing the cells. It should be noted that it has been shown that the pulsed H3.3 protein carrying an acetylation mark rapidly was deacetylated, and shoud therefore not affect the native dynamics of the protein (274). Dendra2, a fluorescent reporter, was pulsed at the same time as the labelled H3.3. Dendra2 was placed on the amber suppression construct, downstream of the synthetase gene (PylRS-TAG-P2A-Dendra2, where 'TAG' is the stop codon) and therefore was synthetised in parallel to labeled H3.3, only upon pulsing the cells.

In paper IV, a fluorescent reporter and the histone variants were encoded on the same construct, consisting of GFP (with early TAG)-TEVp-TEVp_cutting_site-HV-3xHA. Upon pulsing a fusion protein was generated consisting of GFP, Tobacco Etch Virus (TEVp), TEVp cutting site, histone variant of interest and C-terminal triple HA-tag. TEVp was shown to have a high protease activity, separating the GFP-TEVp from HV-3xHA in a rapid (almost immediate) manner. All cell lines in this paper were pulsed with the ncAA cyclopropene-L-lysine. Single clones were selected for each cell line, where the clones selected for downstream experiments were the ones exhibiting naïve morphology and high amber suppression efficiency with GFP intensity as read out.

4.3 MINUTE-CHIP

For quantitative multiplexed ChIP-seq, MINUTE-ChIP was used in all the papers, following a protocol developed in Elsässer Lab (253). MINUTE stands for multiplexed indexed unique molecule T7 amplification end-to-end sequencing and is a protocol based on Mint-ChIP, developed in the Bernstein (275). In short, cells were lysed and chromatin was digested with micrococcal nuclease, producing mono- or di-nucleosome fragments. Each fragment was barcoded by ligating on a dsDNA adapter comprising T7 RNA polymerase promoter (T7), random 6-bp sequence (UMI), and 8-bp barcode (BC). The BC was unique to each sample.

This allowed for pooling up to 16 barcoded samples for subsequent one-pot reaction, which is ideal for experiments with multiple conditions or replicates. From this point on, all reactions occured from a mixture of all samples. Fractions of the barcoded pooled lysate were then used for immunoprecipitation (with a small volume reserved as input material for later normalization as total global chromatin representation). Immunoprecipitation was performed with desired antibodies bound to Protein A beads (when using rabbit derived antibodies) or Protein G beads (for mouse derived antibodies). Immunoprecipitation was conducted for 4 h at 4°C, followed by magnetic pull down, and pull-down material was cleaned with short salt buffer washes. Proteinase K was added for 1 h at 63°C to denature the proteins and DNA was purified. To construct final libraries, DNA from each immunoprecipitation was in vitro transcribed using the T7 promoter. A RNA 3' adapter (RA3) was thereafter appended to the produced RNA to allow for paired end sequencing. The RA3 also allowed for a reverse transcription reaction, for generation of cDNA. As a final step, library PCR was conducted using PCR primers comprising both Illumina-compatible sequences and a second barcode, so multiple libraries can be pooled into a single sequencing run. The resulting libraries were diluted to 4nM, pooled and sequenced on the Illumina platform.

4.4 IMMUNOCYTOCHEMISTRY AND HIGH-CONTENT MICROSCOPY

Cells were grown on 0.1% gelatin coated 96-well imaging plates and at the end of an experiment the cells were fixed for 10 min in the dark using 4% paraformaldehyde. After three phosphate buffer saline (PBS) washes, the plates were stored at 4°C. For immunocytochemistry, cells were permeabilized for 10 min using 0.1% Triton X-100. Thereafter, blocking solution caontaining 1% bovine serum albumin in PBS supplemeted with 0.05% Triton X-100 was conducted at room temperature (RT) for 1 h. Primary antibodies were diluted in blocking solution and samples were incubated with the primary antibody for 1 h at RT. Three PBS washes were carried out prior to 1 h of incubation in the dark with secondary antibodies conjugated to flurophores diluted in blocking solution. Three PBS washes were used to wash off unbound antibodies, and DNA was stained with 4′,6-diamidino-2-phenylindole (DAPI) for 10 min in the dark, followed by a final PBS wash and supplement of PBS.

ImageXpress Micro confocal high-content imaging system was used to image the immunostained plates. This allowed for running up to four filters at the same time. Routinely, four fields were chosen per well, and samples were placed in the plates at least as duplicates. 20x magnification was used to have large amounts of cells to analyse. The obtained micrographs were analysed on CellProfiler v4.0.6. Nuclei were segmented using the DAPI channel and using those masks, integrated and mean nuclear intensity of all wavelengths were calculated. For speckle identification, nuclear objects of a small particle size were enhanced using a module for speckle enhancement. Those foci were selected by primary object identification using a global strategy and Otsu thresholding method. The resulting data was quantified using RStudio v1.3.959. Thresholding strategies based on the mean intensity of the small objects were used compared to the nuclear mean intensity to discriminate between defined or diffused particles.

5 SUMMARY OF RESEARCH PAPERS

TET1 REGULATES GENE EXPRESSION AND REPRESSION OF ENDOGENOUS RETROVIRUSES INDEPENDENT OF DNA METHYLATION (PAPER I)

The TET enzymes have different expression levels during the course of development and might therefore also have different functions (252). Out of the three TET enzymes, TET1 and TET2 are the most abundant paralogs throughout development, with TET1 being more active in the early stages (276). Although TET1 has a well-established function for oxidizing 5mC (5 methylcytosine), studies also show that it also has non-catalytic functions (276,277). The TET1 non-catalytic activity was therefore investigated further in this study to understand its role in early pluripotency, where a TET1 WT (WT) cell was compared to a TET1 KO (Tet1 KO) and an engineered TET1 catalytic mutant (Tet1 CM) in naïve mouse embryonic stem cells (mESCs).

The first noticeable result was that the WT cell line grew more rapidly than the Tet1 KO and tet1 CM cells, but only Tet1 KO exhibited a difference in cell morphology, resembling reduced pluripotency. In a transcript level, the absence of TET1 lead to a prominent dysregulation of gene expression, where markers for cell differentiation where upregulated, in line with the phenotype observed in the KO cells. In the Tet1 CM on the other hand, the transcriptional deregulation was less pronounced, indicating that TET1 regulates gene expression independent of its catalytic activity. To further dissect whether the transcriptional deregulation was due to the catalytic activity of TET1, enzymatic methylome sequencing was conducting, showing that DNA methylation indeed was drastically higher in Tet1 KO and somewhat increased in Tet1 CM cells. Interestingly, although there were alterations to promoter methylation of genes, this did not always correlate with higher gene expression.

Since the gene expression deregulation cannot be fully explained by the DNA methylome differences, there was an interest in studying the epigenetic landscape in the context of histone tail post translational modifications using mass spectrometry. This revealed that the repressive histone marks H3K27me3, H4K20me3 and the active mark H4 lysine acetylation (pH4Kac) decreased significantly in Tet1 KO mESCs, but were only mildly affected in the Tet1 CM cells. To dissect this further, we used quantitative ChIP-seq which confirmed the proteomics results.

Focusing on the genes that were deregulated in the transcriptomic data, there was no alteration in the H3K4me3 profiles, while H3K27me3 signal decreased in both the upregulated and downregulated genes of the Tet1 KO cells, while the opposite occurred in the Tet1 CM cells. The most dramatic effects were observed for the repressive histone mark H4K20me3 at the endogenous retroelements ERV1, ERVK and ERVL, where similar profiles were observed for Tet1 WT and Tet1 CM, but highly decreased signal in Tet1 KO, indicating that TET1 aids in heterochromatin maintenance at ERVs independently of its catalytic activity. Transcriptomics analysis confirmed that ERV expression increased in the absence of the repressive marks.

Lastly, interactomics data showed that TET1 interacts with the SIN3A/HDAC complex, PRC2 complex and HP1, which all have key roles in maintaining the endogenous retroviral elements in a silent state by inducing and maintaining heterochromatin formation. To validate the importance of the interaction of TET1 with SIN3A, constructs of Tet1 WT, Tet1 CM and Tet1 Sin3a mutant were transfected into Tet1 KO cells, where HP1 staining patterns were used as read out. In the case of Tet1 WT and Tet1 CM, defined foci were observed for HP1, while Tet1 KO and Tet1 Sin3a did not manage to rescue the staining pattern, but showed a more diffuse pattern for HP1. All together, we showed that TET1 has a non-catalytic function in interstitial heterochromatin silencing in mESCs, potentially by acting as an interaction hub for chromatin modifying complexes that deposit repressive markers at ERVs (Figure 8).

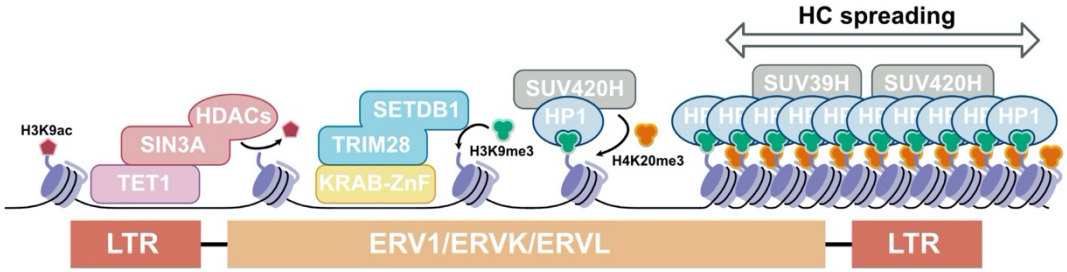

Figure 8. Proposed model of the non-catalytic role of TET1 in ERV silencing. TET1 recruits the SIN3A/ HDAC complex to the ERV elements ERV1, ERVK and ERVL. This leads to local deacetylation of H3K9ac, facilitating the recruitment of the KRAB-ZnF/TRIM28/SETDB1 silencing complex. As a result H3K9me3 is deposited, which HP1 binds to, leading to the recruitment of SUV39H and SUV4-20H that establish H3K9me3 and H4K20me3 domains. This leads ultimately to heterochromatin (HC) spreading. Adapted from Stolz et al. 2022 (278).

POLYCOMB REPRESSIVE COMPLEX 2 SHIELDS NAÏVE HUMAN PLURIPOTENT CELLS FROM TROPHECTODERM DIFFERENTIATION (PAPER II)

Polycomb repressive complex 2 (PRC2) is a *writer* that deposits the repressive H3K27me3 mark on silent and bivalent chromatin (279). The focus of this paper was to study the role of PRC2 in maintenance of human embryonic pluripotency. Three histone marks were studied; active H3K4me3, and the two repressive marks H3K27me3 and H2AK119ub (hereafter denoted as H2Aub), which are deposited by MLL2, PRC2 and PRC1, respectively (280–282).

Utilizing quantitative ChiP-seq, immunoblotting and immunocytochemistry, it was found that naïve cells had a diffuse and high global H3K27me3 profile, over 3-fold higher than in primed cells. This is in line with previous observations made in mouse embryonic stem cells (253). H2Aub levels were also higher in naïve cells, indicating that there is a crosstalk between PRC2 and PRC1 in maintaining both repressive marks. H3K4me3 on the other hand was negligibly altered in the two pluripotency states.

To unravel the function of PRC2, its catalytic function was inhibited with the drug Tazemetostat which targets the Enhancer of zeste homolog 2 (EZH2) subunit. Global quantification pst 7 days of this treatment showed almost complete abolishment of H3K27me3 on both naïve and primed cells, but no effect on H2Aub. H3K4me3 on the other hand decreased to roughly half its original levels upon the PRC2 catalytic inhibition.

Continuing focusing on H3K27me3, it was observed that the X chromosome had a 4-fold difference of H3K27me3 content in naïve cells in comparison to primed cells, while autosomes showed no significant differences. H2Aub was also higher in the X chromosome of naïve cells, but it had a less pronounced enrichment than H3K27me3. Interestingly, the higher abundance of the repressive marks had no chromosome-wide transcriptional consequence.

Next, promoter bivalency, i.e. co-occurrence of H3K4me3 and H3K27me3 at promoters, was studied and new unique as well as shared bivalent promoters in naïve vs primed cells were defined. These categories together with transcriptomics data revealed that naïve and primed cells had very different transcriptional responses to EZH2 inhibition, despite sharing a large number of bivalent genes. Predominantly naïve cells showed a transcriptional disbalance due to the treatment, where genes important for development of intra cellular mass and epiblast were downregulated, while marker genes for mesoderm and extraembryonic trophectoderm and amnion were upregulated. This effect was not recapitulated in the treated primed hESCs.

Of particular interest, GATA Binding Protein 3 (GATA3) was one of the upregulated genes, and as it is a transcription factor that drives trophectoderm differentiation it was further investigated. To validate these results, PRC2 function was disrupted in naïve cells with a compound that inhibits its subunit Embryonic Ectoderm Development (EED) and by clustered regularly interspaced short palindromic repeats (CRISPR) knock out of EED. Immunocytochemistry revealed that in all cases H3K27me3 signals decreased and GATA3

signal increased in a subpopulation, ensuring that PRC2 shields the naïve hESCs from expressing a key regulator for trophectoderm differentiation.

Lastly, following a time course of drug treatment, extensive analysis of the single cell transcriptomic data showed that the PRC2 disruption had a very mild effect on primed cells, while naïve cells expressed markers of multiple lineage progenitors which could be used for trajectory-based differential expression analysis. From this, a pseudotime axis was created showing how the naïve cells with PRC2 disruption progressively differentiate from a ground state, through an activated state, to trophectoderm and mesoderm cells.

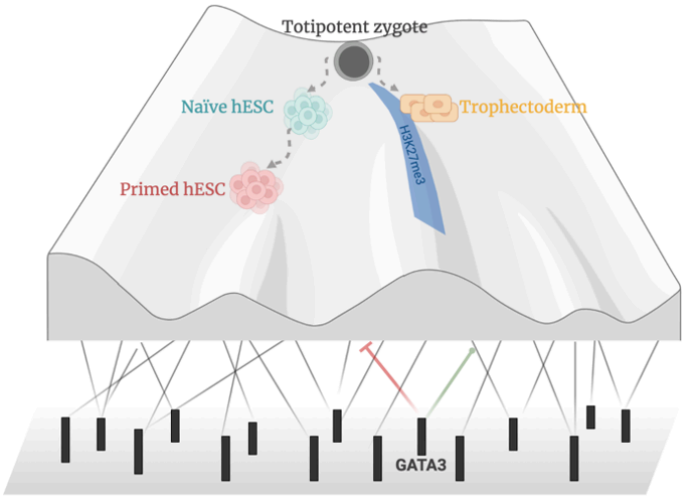

Figure 9. The zygote is a totipotent stem cell that gives rise to the embrypnic and extraembryonic tissue. In the initial steps of embryogenesis, cells divide and commit to either forming the inner cell mass or the extraembryonic trophoblast of the embryo. Naïve pluripotent stem cells of the epiblast progress toward primed pluripotent and subsequently form the three germ layers. In this study, an epigenetic barrier has been discovered, where PRC2, through deposition of H3K27me3, blocks the potential of progression from naïve to trophectoderm differentiation, with GATA3 being its master regulator. Artist: Banushree Kumar.

QUANTITATIVE HISTONE H3.3 DYNAMICS RESOLVED BY RAPID PROTEIN PULSING (PAPER III)

H3.3 is a histone variant with important clinical significance, as it is linked to several diseases when the genes encoding this protein are mutated. Similarly, mutations or deregulation of its histone chaperones are also linked to failure in protection of genome integrity. It is therefore important to dissect the molecular function of H3.3 in cell homeostasis in the context of its interactome and its chromatin dynamics. In this study, we used genetic code expansion in mouse embryonic stem cells, where we produced labelled H3.3 in mouse embryonic stem cells for dedicated windows of time and used high content microscopy, mass spectrometry and quantitative ChIP-seq as readouts.

We showed that amber suppression is an approach for generating rapid protein pulses, where nuclear accumulation of the pulsed histone was observed after only 30 min. The system is also switchable, allowing for time windows of protein production to be controlled simply by changing the growth media of the cells. Importantly, it is also tunable, as the added concentration of non-canonical amino acid regulates the amount of protein that can be produced, which is of essence to obtain trace levels of the pulsed protein, to not compete with the endogenous pool of H3.3.

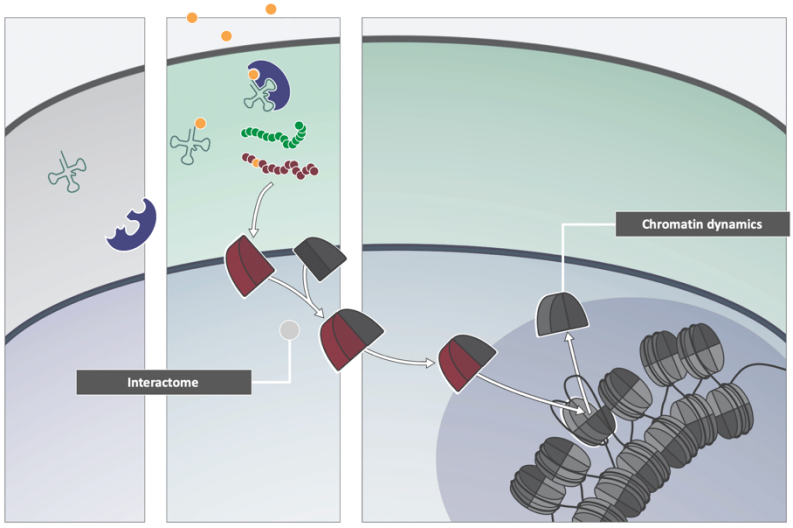

Figure 10. Amber suppression mediated labelled histone H3.3-3xHA expression. H3.3-3xHA is constitutively transcribed, but translated to its full length only in the presence of ncAA (yellow). For this, the amber suppression machinery is also constitutively expressed. This allows for a fast, tunable and switchable generation of histone pool that is followed through a time course, where focus lays on its temporal interactome and its chromatin dynamics.

Studying the accumulation of pulsed H3.3 in a subnuclear compartment we identified that it overlapped with its histone chaperones DAXX and ATRX, but also with the chromatin remodeler Smarcad1 and HP1 paralogs. This indicates that phase separation could potentially play a role in the incorporation of H3.3 to regions of the genome. In line with this, those interactions were also observed as result of an interactome experiment of a pulse chase time course of H3.3. Its results revealed known interactors, which confirms that the pulsed proteins behave identically to the endogenous proteins, but also novel interactors were found. The interactors could due to the temporal resolution be subcategorized into transient, early, mid and late interactors. Importins and chaperones were found amongst the early interactors, followed

by important chromatin factors for both active and repressive chromatin, where the latter was progressively more prevalent in the later time points.

Time-lapse multiplexed quantitative ChIP-seq was used to study locus-specific incorporation, enabling future identification of kinetic parameters. Early time points revealed a similar incorporation rate to both active and repressed chromatin states, mediated by HIRA and DAXX pathways, respectively. Of particular interest, the latter showed a high nucleosome turnover rate at retroviral elements, which typically is absent of DNA accessibility. This sheds light into the rate of incorporation of H3.3 into interstitial heterochromatin, where H3.3 was found to have a fundamental role for local compaction of chromatin (109). Lastly, H3.3 incorporation to enhancers was further dissected by pertrubing chromatin remodeling, revealing the source of nucleosome turnover and importance in the assembly of H3.3 at functionally active chromatin states in mouse embryonic stem cells. In summary, amber suppression assisted protein pulsing is a rapid, tunable and switchable technique that allows for studies of histone variants in terms of their temporal interactome and chromatin assembly and turnover kinetics.

QUANTITATIVE TEMPORAL CATALOGUE OF HISTONE VARIANT DYNAMICS IN mESCs (PAPER IV)

In Paper III we showed how amber suppression can be used to dissect histone dynamics of H3.3 in mouse embryonic stem cells. Following that up, we aimed to further the scope to other study other histone variants, that also have regulatory functions in the cell and are linked to diseases in the case of their deregulation. For this, we generated cell lines expressing a construct for equimolar expression of labelled histone variants and a green fluorescent protein reporter. This allows not only for studying the temporal localization of the labelled protein, but also to study its abundance in comparison to how much that was initially produced, aiding in studies of protein decay.

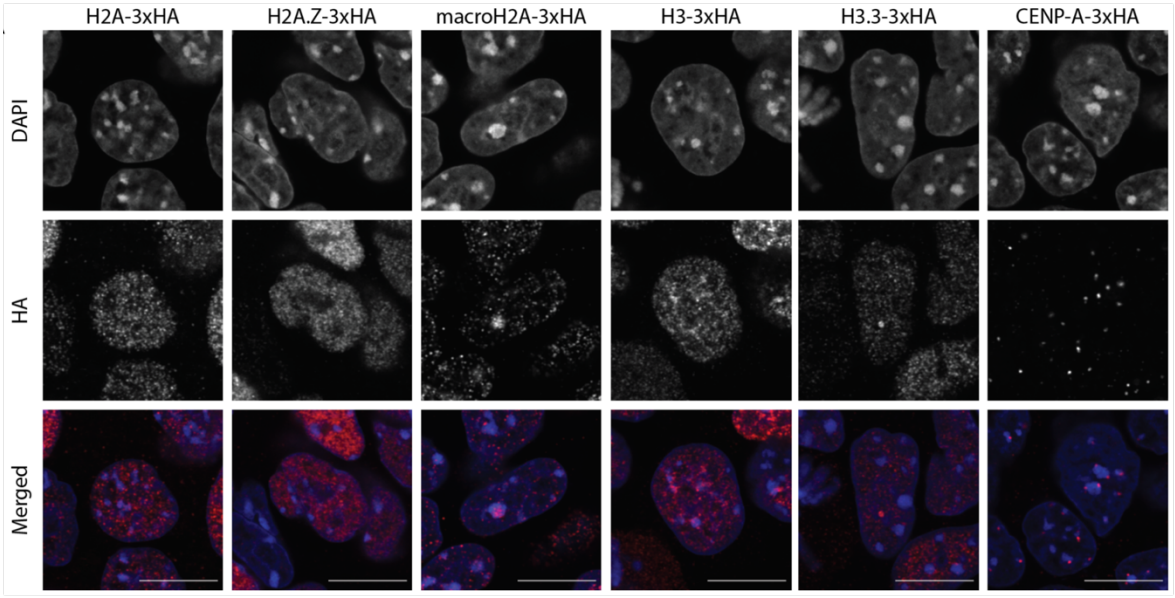

Figure 11. Amber suppression histone variants show distinct staining patterns in mESCs, showing that they have highly specific chromatin deposition. DNA (DAPI) in blue and HA in red. Scale bar: 10 μm.

Six cell lines were used, each expressing a construct encoding a histone under amber suppression control. The six histones of interest were the canonical histone H3.2 and H2A, and the histone variants H3.3, CENP-A, H2A.Z and macroH2A. Individual subnuclear localization patterns were confirmed with immunocytochemistry, indicating that the labeled proteins indeed behaved as the endogenous proteins. Of importance, CENP-A was pulsed differently than the other proteins to minimize any risk of overexpression, since this leads to mitotic defects and aneuploidy (283).

Staining patterns of pulsed CENP-A were highly heterogeneous, although all cells evidently were pulsed, seen by their positive signal of the fluorescent reporter, indicating that a larger fraction of the pulsed cells degraded the pulsed CENP-A in a rapid manner. Upon treating cells with proteasomal or lysosomal inhibitors, it was observed that CENP-A was indeed subjected to proteasomal degradation, with a much more pronounced effect than the other pulsed histones. This indicates that the pre-assembly histone life time is highly individual. An explanation for this could be that histones are highly dependent on binding to their chaperones in a cell cycle dependent manner to avoid proteasomal decay. To study the post-assembly stability, total protein quantification of chromatinized protein was conducted during a time

course of three days. Overall, H2A and its histone variants showed a much higher decay rate than H3 and its variants.

Lastly, we examined chromatin assembly kinetics by conducting amber suppression pulse labelling time course followed by quantitative multiplexed ChIP-seq. For this, we also added conditions perturbing transcription, replication or activity of a chromatin remodeler to further dissect the effects on each histone protein. H2A and H3 showed rapid incorporation to all chromatin states. H3.3 incorporated to active regions and interstitial heterochromatin (as in Paper III). H2A.Z showed specific incorporation to active regions, poised promoters and weekly at insulators, macroH2A had a slow incorporation rate to heterochromatin. Using the potential of MINUTE-ChIP to also quantitatively analyze repetitive regions, we could follow the incorporation of CENP-A to centromeric minor satellites, H2A.Z incorporation to constitutive inter-lamin associated domains and macroH2A deposition to pericentric major satellites.

To summarize, using amber suppression allows for studying histone pre- and post- assembly decay rates, nuclear shuffling. Adding the quantitative genomics approach of MINUTE-ChIP permits quantitative analysis of chromatin assembly and turnover rates with chromatin state specific dynamics, even at repetitive regions. Pooling together all this information will be essential for determining histone specific kinetic constants leading to kinetic modeling approaches. Further, time-resolved interactomics will show the key players that modulate dynamics of each histone.

6 CONCLUDING REMARKS AND PERSPECTIVE

For this thesis, multiple omic methods have been used to study epigenetic mechanisms in stem cells. In **Paper I**, the focus was on the role of Polycomb repressive complex 2 (PRC2) in conserving the integrity of naïve human embryonic stem cells, showing that in its absence there was differentiation toward mesoderm and trophectoderm lineages. This shows that PRC2 and H3K27me3 generate an essential barrier that limits the pluripotent cells that for the inner cell mass of the embryo to erroneously generate extraembryonic tissue. In **Paper II**, a non-catalytic function for TET1 in mouse embryonic stem cells (mESCs) was discovered, showing that TET1 potentially acts as an interaction hub for chromatin factors that mediate silencing of endogenous retroviral elements. In **Paper III**, it was shown that H3.3 has a very rapid chromatin association and incorporation in mESCs, with accumulation at active genes, active enhancers and interstitial heterochromatin. As time course experiments were conducted, the temporal interactome of H3.3 could also be defined. Lastly, **Paper IV** we broaden the scope of chromatin dynamics of histones in mESCs by studying the histone variants H3.3, CENP-A, H2A.Z and macroH2A together with their respective canonical histone counterparts. In this study we generated pools of these proteins, and studied their unique pre-assembly and post assembly decay rates, dynamic incorporation to dedicated chromatin states, including highly repetitive regions. Altogether we show that the multiplexed quantitative ChIP-seq (MINUTE-ChIP), together with transcriptomics, immunocytochemistry and extensive image analysis, quantitative protein quantification and proteomics yield numerous novel findings in the field of epigenetic regulation of stem cells.

Quantitative Methods

During the last decades there has been incredible advances in the field of fundamental research as a plethora of methods for biomolecular studies have been developed. With it, there is a need for conducting quantitative studies, to properly evaluate and understand the results. In this thesis, MINUTE-ChIP, a quantitative multiplexed ChIP-seq protocol, has been used in each of the studies, where multiple samples have been barcoded, pooled together, and from that pool, various immunoprecipitations have been conducted. This ensures that there is no technical variation between the samples throughout the process. Furthermore, a unique molecular identifier is introduced to all DNA fragments in an early step, which allows for distinguishing between technical amplicons due to PCR amplification versus unique fragments, which is the key for studying repetitive elements. Further, we also always run a sample of the pooled lysate, without any selection. This sample is used for normalization and represents the baseline of genome coverage in the various samples which allows us to account for biases due to over- or underrepresentation of chromatin in certain chromatin states. This is important, as heterochromatin routinely is challenging to study due to potential under digestion. Altogether, MINUTE-ChIP is a method for accurately quantifying relative differences in genome-wide histone modification patterns and histone variant genome code across multiple pooled samples.

Other quantitative methods used have been protein quantification using near infrared and infrared secondary antibodies instead of the conventional horseradish perish reaction for western blots. This allows for a low background and high sensitivity when measuring the intensity of fluorescence instead of chemical reaction, which is highly dependent on the duration of substrate incubation and is susceptible to signal saturation. Another quantitative method used was high content microscopy, which allows for ensuring that imaging settings are conducted within a linear range of the instrument, and multiple conditions are detected in parallel, limiting the batch-to-batch differences between experiments. To include even more quantitative aspects to the techniques, it would be favorable to do use tandem mass tags for labelling of protein samples, allowing for multiplexing samples, for mass spectrometry. With this approach it will be easier to quantify differences between proteomics samples. Altogether, the more quantitative experiments are, the more reliable the results, leading to better interpretation of the data.

Genetic code expansion in mammalians cells

Genetic code expansion (GCE) has been used for residue- and site-specific incorporation of non-canonical amino acids (ncAAs) to proteins of interest. There are many uses of this technique, such as biorthogonal ncAA-tagging labeling for protein purification, protein visualization using fluorescent molecules and adding ncAA that can crosslink to specific residues to study protein-protein interactions (284). In Papers III and IV, we used GCE to generate a switchable system for protein expression to avoid overexpression of the labelled proteins. Since the amber stop codon is the least abundant mammalian stop codon, we used amber suppression in all our mESC cell lines (272). Although there is a short lag phase for generating the histones of interest as the ncAAs need to be loaded to the amber suppression machinery, this lag phase in protein synthesis is in the range of minutes and is therefore much shorter than in inducible approaches where a compound is added to initiate gene expression. Importantly, this system is rapidly turned off simply by washing off the medium containing ncAA. There is indeed a delay in switching off the system, as amber suppression occurs until the fraction of intracellular ncAA is depleted. To overcome this obstacle, we have tried challenging the system using compounds that compete with the ncAA to obtain an even more controlled amber suppression turn off rate.

Another benefit of using this system has been the potential to use a small protein tag, in contrast to many studies where the histone variant has been fused to a bulky fluorescent protein. By adding a small tag at the C-terminus of the protein we believe that the risk of affecting proper interaction and deposition of the histones are minimal. Further, adding the very same tag to multiple histone variants has major technical advantages, as it decreases the number of variables when using different antibodies for immunostaining or immunoprecipitation, resulting in the possibility for highly comparative studies of the panel of proteins of interest. This is of extra importance as several commercial antibodies for histone variants are not specific enough.

There are endless possibilities of implementation for amber suppression for generation of protein pools for temporal studies. This system is perfectly suited for studying the effect of gene mutations with clinical relevance, as in the case of oncohistones, to dissect what differences there are in comparison to the wild type case. Another possible and very simple approach is to generate overexpression of the protein of interest on demand to see what the effects are with a tunable and temporal resolution. Yet another possibility is to use non-canonical amino acids that are not possible to be recognized by readers, inhibiting site specific post translational modification to study the role of different histone marks.

Maintaining stem cell integrity

In the early steps of embryogenesis there is a global DNA demethylation event, zygotic genome activation, and *de novo* DNA methylation. TET1 and TET2 are the main regulators of the first event, where they catalyze the oxidation for active DNA demethylation. A question that arises as a consequence of this is how genes are not globally upregulated in the DNA hypomethylated state. In Paper 1 we observed that TET1 interacts with the repressive chromatin methyltransferase PRC2. In a study conducted in parallel to ours, this interaction was further dissected, where they showed that TET regulates developmental genes together with PRC2 and SIN3A independent of its DNA demethylation activity, which is in line with our observations (285). In Paper 2 we showed that H3K27me3, the repressive histone mark deposited by PRC2, was at very high levels in the early stages of pluripotency, indicating that the DNA hypomethylation is balanced by a high global H3K27me3 abundance. Disruption of TET1 or PRC2 function in each of those studies showed rapid differentiation of the cells. In the first case, genes active in gastrulation were upregulated and cells gained a more differentiated morphology in the absence of TET1. In the second case, cells differentiated in a heterogeneous manner, but surprisingly also into trophectoderm lineage. Taken together, TET1 and PRC2 are essential for maintaining the stem cell integrity.

For all cells it is also of essence to regulate the histone variants for maintaining homeostasis. In the first steps of embryogenesis, the protamines from the gametes are degraded and replaced with histones. At this step it is important that the correct histone variants end up in the correct place. Histone H3.3 is abundant from a very early stage, whereas histones H2A.Z and macroH2A have increased expression throughout development. Even though the latter two are less expressed, they carry important functions. For example, H2A.Z aids in gene regulation by mediating interaction with RNA Pol II and macroH2A highly accumulates in the inactive X chromosome in the post-implantation stages of development.

There is growing evidence for the importance of the three-dimensional structure of chromatin for gene expression, where the differentiating cell acquires a more structured genome order with repressive regions clustering at the nuclear periphery through interactions with the nuclear lamina and topologically associated domains (TADs) being formed, where chromatin is clustered with regions sharing similar epigenetic profiles (13). It would be extremely interesting to study the role played by histone variants in the 3D genome regulation. In paper IV, we showed that some histones (H2A and H2A.Z) have a diffused distribution in the cell,

while others have dedicated regions of incorporation. H3.2 overlaps with regions of ongoing replication, H3.3 goes through a subnuclear (possibly phase separated) compartment, macroH2A accumulates at DNA-rich regions and CENP-A is deposited to centromeres. The genomics information of the pulsed histone panel also showed that H2A.Z accumulates very rapidly to active regions and slower at heterochromatin and constitutive inter-lamin associated domains, H3.3 enters active and repressed chromatin and macroH2A is predominantly deposited to heterochromatin and pericentromere. It would be interesting to study if the histones also play a role in establishing TADs, which would be expected, as the histone variants are associated with specific histone marks.

Lastly, it will be of interest to fully understand the role of cell cycle on histone deposition and eviction. Many studies assessing cell cycle dependency of histone variant deposition have been carried out with chemically induced cell cycle arrest. The conducted drug treatment might cause DNA damage, and therefore the literature should be read with caution, as deposition of the histone variants might be altered due to cellular stress. The same is the case for studies where the histone variants have been fused to bulky fluorescent proteins in one of the termini, where thorough studies of the effects of alteration of interaction partners and PTMs have not been carried out. Therefore, amber suppression with genes fused to small molecular tags serve as excellent alternatives and for a higher cell cycle resolution it would be interesting to use a cell with fluorescent reporters for the cell cycle, as the highly used Fluorescence Ubiquitin Cell Cycle Indicator (FUCCI) system, for determination of cell cycle dynamics of pulsed histones (286,287).

Regulation of endogenous retroviral elements in mESCs

Endogenous retroviral elements could pose a threat to the genome if they are activated and initiate retrotransposition events. The cell has therefore developed a mechanism for silencing the ERV regions by inducing heterochromatin compaction. This is established by Krüppel-associated box zinc finger proteins (KRAB-ZFPs) that recognize the DNA sequences, followed by TRIM28 (KAP1) binding, histone deacetylase and SETDB1 (ESET) recruitment, H3K9me3 deposition and HP1 binding, resulting in chromatin compaction. In Paper I and Paper III we studied TET1 and H3.3, and what role they might play in the maintenance of a silent state of ERVs in mESCs.

In Paper I, there was upregulation of ERVs in TET1 KO cells, but this was much subtler in engineered cells expressing a catalytic dead version of TET1, attributing its ERV silencing role to a non-catalytic function. Interestingly, SIN3A was found to be an interactor of TET1 that also localizes to ERVs and potentially aids in the silencing through its interaction with HDACs. Further, displacement of HP1 was observed in TET1 KO cells, but partially re-established upon rescuing TET1 expression, adding to the functional role of TET1 for HP1 establishment.

In Paper III, we showed that H3.3 rapidly accumulates at interstitial heterochromatin, indicating active histone turnover, showing that although the ERV chromatin in mESCs is transcriptionally silent it is still highly dynamic. An explanation for this highly dynamic nature

is proposed in a recent study, where it is shown that the chromatin remodeler Smarcad1 evicts nucleosomes at interstitial heterochromatin, where H3.3 is selectively and continuously deposited (109).

Outlook

Deregulation of PRC2, TET1 and histone variants and their chaperones have all been linked to diseases. It is therefore of utmost importance to continue unraveling the regulation of all these epigenetic factors to understand how they mediate gene expression. Several studies have shown that mutations to histones ('oncohistones') can be linked to prognosis of patient survival. This shows that there is a clinical need for further understanding the healthy epigenome to better understand what occurs when it is deregulated. With this knowledge we could derive more epigenetic drugs or approaches. As an example, we could use the deregulation of retrotransposons to our advantage for forcing cancer stem cells to activate expression of interferon genes, making them a target from the once evaded innate immune system (288). Knowing exactly the clinical effect of mutations, loss- or gain- of function of these epigenetic regulators will lead to improved diagnostics, better treatment and therefore better patient prognostics.

7 ACKNOWLEDGEMENTS

Pluripotent stem cells convert from naïve to primed state when being supplemented with certain growth factors. I would like to believe that I've made a similar transition, with my PhD journey and everyone involved in it as growth factors. Now I'm primed for the next chapter in life, looking forward to the upcoming steps of differentiation.

First, I would like to thank my supervisor **Simon Elsässer** for inviting me to his group, for encouraging me to learn multiple methods, for giving me a truly interesting research project and for trusting me to work on collaborative projects with people both in and outside his research group. I've been very lucky to have had two caring co-supervisors as well, **Gonçalo Castelo-Branco** and **Peter Svensson**, who whenever we met showed high interest in being updated on the progress of my doctoral studies and gave great feedback on the occasions where we met.

I want to express my utmost gratitude to my opponent **Prof. Johan Jakobsson** and the examination board members **Prof. Jan Larsson**, **Assoc. Prof. Claudio Cantù** and **Assoc. Prof. Colm Nestor** for accepting the invitation to read this thesis, examine my work and being essential members of my public defence. I highly appreciate all the time you have put on this task. Thank you as well to **Dr. Per Moberg** who kindly accepted my request on acting as a chairperson for the defence in less than a blink of an eye.

To the current and former members of the **Elsässer Lab**: thank you for all the questions, comments, feedback, input, reagents, laughter, troubleshooting, fikas and every late-night lab chat. I've learnt lots from you all through the years, both experimentally and regarding your individual (and diverse) research topics.

Banu, thank you for teaching me so much in the lab (WB, Wes and ChIP-seq). Thanks for always being available and answering all thousands of questions I've had through the years. Also thank you for giving feedback on my work, presentations, micrographs and now also on my thesis (twice!). **Nerges**, vad hade jag gjort utan ditt dagliga stöd och peppande? Vilken resa vi haft hittills, och det är helt fantastiskt att den inte tar slut, utan snarare leder oss ännu närmare. Jag kunde inte be om en bättre kosmisk tvilling! **Dörte**, thank you for cheering me up with pictures, dancing, singing and the cutest post-its. I will miss our conversations, especially the 'super coherent' ones. Let's go dancing more often! **Yuying**, I'm so impressed by how fast you became a wet lab person and thanks for all your help with bioinformatics. All four of you, thank you for providing a steady ground I could always rely on.

Rozina, thank you for being super helpful and for sharing my enthusiasm for colorful micrographs. **Lorenzo**, thanks for all the chats inside and outside the cell culture room. **Birthe**, thanks for always answering my questions regarding cloning. **Jing**, I'll miss our nightly discussions and thanks for feeding me sweets! Thanks **Lea** for sharing borrowed ice cream and microscopy interest, and for your amazing attitude in the lab, especially when experiments failed. **Tomek**, thanks for being a good friend and enjoying my pun (yes, singular). **Reza** and

Eleni, thank you for your genuine kindness and thoughtfulness! **Jaime**, thank you for all discussions and for the amazing cover picture! **Alba**, thanks for being a great *vecinaaa*!! I'm very grateful to my students **Marius**, **Christian**, **Janice**, **Maximillian** and **Javier**, it was an honor working together with you! **Carmen**, thank you for being an amazingly kind bioinformagician! **Anna-Maria**, thank you for sharing the experience with Paul S. Chase. **Philip**, thanks for the input on presentations. **Rahul**, thank you for your supervision when I joined the lab. **Laura**, I truly appreciate that you and I have such similar humor, and thanks for helping me out when I had covid! **Rui**, **Hannes**, **Kyle**, **Chris**, **Anna**, **Ruth**, **Bruno**, **Xinhe**, **Eva**, **Christina**, **Theresa**, **Helena**, **Turid**, **Marianne**, and everyone listed above, and many more that I'm forgetting; I truly enjoyed sharing the lab space with you and I hope to hang out with you also in the future. Also, an enormous thanks to all amazing people in neighbouring groups who I had the pleasure to work alongside; **Michele**, **Ana**, **Constantin**, **Gustaw**, **Gabrielle**, **Su**, **Marcin**, **Joaquin**, **Masahiro**, **Maud**, **Luuk**, **Bennie**, **Johana**, **Asimina**, **Myriam**, **Daniela**, **Maria**, **Katta**, **Louise**, **Dimitris**, **Ann-Sofie**, **Sidorela**, **Vamakshi**, **Sharath**, **Tharagan**, **Bishoy** and **Marianna**.

I'm truly grateful for being invited to very interesting collaborations. **Paul**, thank you for your openness and your enthusiasm! It was a truly fantastic experience working with you, **Sebastian** and your colleagues. Thank you to the authors of the PRC2 story for inviting me to collaborate and thanks **JP** for kindly driving over with cells back when we generated the first cell lines.

Duong and **Anna D**, thank you for being my mentors and sharing your experiences with me! **Anders** and **Marie**, I'm truly grateful for your supervision during the internship at Pfizer, you both really made me feel welcomed and I feel like I learnt plenty. **Claudia**, I'm thankful for your input at multiple presentations, your feedback at my half-time seminar and your amazing attitude for helping with the last step of the jorney! **Erika S**, tack för att du hjälpte mig under en väldigt viktig period.

The PhD years also involved some time outside the lab, which was shared with fantastic people. A big hug to my friends that have supported in various ways, especially **Yuqi**, **Matilde**, **Emelie**, **Christina**, **Bernie**, **Johanna**, **Beatrice**, **Sara, Victor**, **Anna W**, **Sandra**, **Franziska**, **Miriam**, **Georgia**, **Christoffer**, **Olov**, **Alvaro**, **Linda** and **Jorge**.

Tusen tack till hela min familj som varit med på varje steg av resan och varit en fantastisk hejarklack! **Mama**, **Papa**, **Sebastian**, **Cristian**, **Anna**, **Anna**, **Lea**, **Claudia**, **Nicolas**, **Elizabeth**, och **Andréas**, jag är otroligt lyckligt lottad att ha er i min familjekonstellation. Jag ser fram emot att umgås med er mycket mer nu när nästa kapitel i livet påbörjas. **Erik**, tack för att du introducerat mig till nya hobbies, varit andförälder med mig och hjälpt mig att få perspektiv på en hel del i livet. Tack också till din familj som öppnat upp sina dörrar för mig. **Familia**, muchisimas gracias por todo el apoyo y por acompañarme en cada paso del camino. Se que siempre puedo contar con cada uno de ustedes y los quiero muchísimo. Un abrazo enorme también para los familiares en Chile!

There are so many more people to thank, but I'll make sure to do it next time we meet in person.

And a big thanks to you, the reader, for getting to this point of the text. Especially if you've read more than just the acknowledgements section (but no hard feelings if you didn't).

Lastly, I hope that I can continue with a career involving FISH and ChIPs!

8 REFERENCES

1. Molè MA, Weberling A, Zernicka-Goetz M. Comparative analysis of human and mouse development: From zygote to pre-gastrulation. Curr Top Dev Biol. 2020;136:113–38.

2. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet. 2003 Mar;33 Suppl:245–54.

3. Cosgrove MS, Boeke JD, Wolberger C. Regulated nucleosome mobility and the histone code. Nat Struct Mol Biol. 2004 Nov;11(11):1037–43.

4. Gillette TG, Hill JA. Readers, writers, and erasers: chromatin as the whiteboard of heart disease. Circ Res. 2015 Mar 27;116(7):1245–53.

5. Jin B, Li Y, Robertson KD. DNA methylation: superior or subordinate in the epigenetic hierarchy? Genes Cancer. 2011 Jun;2(6):607–17.

6. Martire S, Banaszynski LA. The roles of histone variants in fine-tuning chromatin organization and function. Nat Rev Mol Cell Biol. 2020 Sep;21(9):522–41.

7. Eykelenboom JK, Gierliński M, Yue Z, Hegarat N, Pollard H, Fukagawa T, et al. Live imaging of marked chromosome regions reveals their dynamic resolution and compaction in mitosis. J Cell Biol. 2019 May 6;218(5):1531–52.

8. Warren C, Shechter D. Fly fishing for histones: catch and release by histone chaperone intrinsically disordered regions and acidic stretches. J Mol Biol. 2017 Aug 4;429(16):2401–26.

9. Franco AA, Lam WM, Burgers PM, Kaufman PD. Histone deposition protein Asf1 maintains DNA replisome integrity and interacts with replication factor C. Genes Dev. 2005 Jun 1;19(11):1365–75.

10. Liu WH, Roemer SC, Port AM, Churchill MEA. CAF-1-induced oligomerization of histones H3/H4 and mutually exclusive interactions with Asf1 guide H3/H4 transitions among histone chaperones and DNA. Nucleic Acids Res. 2012 Dec;40(22):11229–39.

11. Giles KA, Gould CM, Du Q, Skvortsova K, Song JZ, Maddugoda MP, et al. Integrated epigenomic analysis stratifies chromatin remodellers into distinct functional groups. Epigenetics Chromatin. 2019 Feb 12;12(1):12.

12. Ho L, Crabtree GR. Chromatin remodelling during development. Nature. 2010 Jan 28;463(7280):474–84.

13. Zheng H, Xie W. The role of 3D genome organization in development and cell differentiation. Nat Rev Mol Cell Biol. 2019 Sep;20(9):535–50.

14. Ma Y, Buttitta L. Chromatin organization changes during the establishment and maintenance of the postmitotic state. Epigenetics Chromatin. 2017 Nov 10;10(1):53.

15. Rudnizky S, Khamis H, Ginosar Y, Goren E, Melamed P, Kaplan A. Extended and dynamic linker histone-DNA Interactions control chromatosome compaction. Mol Cell. 2021 Aug 19;81(16):3410-3421.e4.

16. Marzluff WF, Wagner EJ, Duronio RJ. Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. Nat Rev Genet. 2008 Nov;9(11):843–54.

17. Hammond CM, Strømme CB, Huang H, Patel DJ, Groth A. Histone chaperone networks shaping chromatin function. Nat Rev Mol Cell Biol. 2017 Mar;18(3):141–58.

18. Padavannil A, Sarkar P, Kim SJ, Cagatay T, Jiou J, Brautigam CA, et al. Importin-9 wraps around the H2A-H2B core to act as nuclear importer and histone chaperone. eLife. 2019 Mar 11;8.

19. Tsunaka Y, Ohtomo H, Nishimura Y. FACT modulates the conformations of histone H2A and H2B N-terminal tails within nucleosomes. Commun Biol. 2022 Aug 13;5(1):814.

20. Park Y-J, Luger K. Histone chaperones in nucleosome eviction and histone exchange. Curr Opin Struct Biol. 2008 Jun 3;18(3):282–9.

21. Clamp M, Fry B, Kamal M, Xie X, Cuff J, Lin MF, et al. Distinguishing protein-coding and noncoding genes in the human genome. Proc Natl Acad Sci USA. 2007 Dec 4;104(49):19428–33.

22. Ge SX. Exploratory bioinformatics investigation reveals importance of "junk" DNA in early embryo development. BMC Genomics. 2017 Feb 23;18(1):200.

23. Raab JR, Kamakaka RT. Insulators and promoters: closer than we think. Nat Rev Genet. 2010 Jun;11(6):439–46.

24. Jasinska A, Krzyzosiak WJ. Repetitive sequences that shape the human transcriptome. FEBS Lett. 2004 Jun 1;567(1):136–41.

25. Saksouk N, Simboeck E, Déjardin J. Constitutive heterochromatin formation and transcription in mammals. Epigenetics Chromatin. 2015 Jan 15;8:3.

26. International Human Genome Sequencing Consortium, Whitehead Institute for Biomedical Research, Center for Genome Research:, Lander ES, Linton LM, Birren B, Nusbaum C, et al. Initial sequencing and analysis of the human genome. Nature. 2001 Feb 15;409(6822):860–921.

27. Thakur J, Packiaraj J, Henikoff S. Sequence, chromatin and evolution of satellite DNA. Int J Mol Sci. 2021 Apr 21;22(9).

28. Calado RT, Dumitriu B. Telomere dynamics in mice and humans. Semin Hematol. 2013 Apr;50(2):165–74.

29. Stults DM, Killen MW, Williamson EP, Hourigan JS, Vargas HD, Arnold SM, et al. Human rRNA gene clusters are recombinational hotspots in cancer. Cancer Res. 2009 Dec 1;69(23):9096–104.

30. Darrow EM, Chadwick BP. A novel tRNA variable number tandem repeat at human chromosome 1q23.3 is implicated as a boundary element based on conservation of a CTCF motif in mouse. Nucleic Acids Res. 2014 Jun;42(10):6421–35.

31. Martens JHA, O'Sullivan RJ, Braunschweig U, Opravil S, Radolf M, Steinlein P, et al. The profile of repeat-associated histone lysine methylation states in the mouse epigenome. EMBO J. 2005 Feb 23;24(4):800–12.

32. Göke J, Ng HH. CTRL+INSERT: retrotransposons and their contribution to regulation and innovation of the transcriptome. EMBO Rep. 2016 Aug;17(8):1131–44.

33. Gifford RJ, Blomberg J, Coffin JM, Fan H, Heidmann T, Mayer J, et al. Nomenclature for endogenous retrovirus (ERV) loci. Retrovirology. 2018 Aug 28;15(1):59.

34. Löwer R, Löwer J, Kurth R. The viruses in all of us: characteristics and biological significance of human endogenous retrovirus sequences. Proc Natl Acad Sci USA. 1996 May 28;93(11):5177–84.

35. Lanciano S, Cristofari G. Measuring and interpreting transposable element expression. Nat Rev Genet. 2020 Dec;21(12):721–36.

36. Ward MC, Zhao S, Luo K, Pavlovic BJ, Karimi MM, Stephens M, et al. Silencing of transposable elements may not be a major driver of regulatory evolution in primate iPSCs. eLife. 2018 Apr 12;7.

37. Hsu P-S, Yu S-H, Tsai Y-T, Chang J-Y, Tsai L-K, Ye C-H, et al. More than causing (epi)genomic instability: emerging physiological implications of transposable element modulation. J Biomed Sci. 2021 Aug 7;28(1):58.

38. Yan MSC, Matouk CC, Marsden PA. Epigenetics of the vascular endothelium. J Appl Physiol. 2010 Sep 1;109(3):916–26.

39. Breiling A, Lyko F. Epigenetic regulatory functions of DNA modifications: 5-methylcytosine and beyond. Epigenetics Chromatin. 2015 Jul 21;8:24.

40. Ambrosi C, Manzo M, Baubec T. Dynamics and Context-Dependent Roles of DNA Methylation. J Mol Biol. 2017 May 19;429(10):1459–75.

41. Bhutani N, Burns DM, Blau HM. DNA demethylation dynamics. Cell. 2011 Sep 16;146(6):866–72.

42. Kohli RM, Zhang Y. TET enzymes, TDG and the dynamics of DNA demethylation. Nature. 2013 Oct 24;502(7472):472–9.

43. Mohr F, Döhner K, Buske C, Rawat VPS. TET genes: new players in DNA demethylation and important determinants for stemness. Exp Hematol. 2011 Mar;39(3):272–81.

44. Zhang L, Wang J-C, Hou L, Cao P-R, Wu L, Zhang Q-S, et al. Functional Role of Histidine in the Conserved His-x-Asp Motif in the Catalytic Core of Protein Kinases. Sci Rep. 2015 May 11;5:10115.

45. Ko M, An J, Bandukwala HS, Chavez L, Aijö T, Pastor WA, et al. Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX. Nature. 2013 May 2;497(7447):122–6.

46. Suelves M, Carrió E, Núñez-Álvarez Y, Peinado MA. DNA methylation dynamics in cellular commitment and differentiation. Brief Funct Genomics. 2016 Nov;15(6):443–53.

47. Baylin SB, Esteller M, Rountree MR, Bachman KE, Schuebel K, Herman JG. Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer. Hum Mol Genet. 2001 Apr;10(7):687–92.

48. Jelinek J, Gharibyan V, Estecio MRH, Kondo K, He R, Chung W, et al. Aberrant DNA methylation is associated with disease progression, resistance to imatinib and shortened survival in chronic myelogenous leukemia. PLoS ONE. 2011 Jul 8;6(7):e22110.

49. Skvortsova K, Stirzaker C, Taberlay P. The DNA methylation landscape in cancer. Essays Biochem. 2019 Dec 20;63(6):797–811.

50. Calle-Fabregat C de la, Morante-Palacios O, Ballestar E. Understanding the relevance of DNA methylation changes in immune differentiation and disease. Genes (Basel). 2020 Jan 18;11(1).

51. Elhamamsy AR. Role of DNA methylation in imprinting disorders: an updated review. J Assist Reprod Genet. 2017 May;34(5):549–62.

52. Tabaei S, Tabaee SS. DNA methylation abnormalities in atherosclerosis. Artif Cells Nanomed Biotechnol. 2019 Dec;47(1):2031–41.

53. Huang H, Lin S, Garcia BA, Zhao Y. Quantitative proteomic analysis of histone modifications. Chem Rev. 2015 Mar 25;115(6):2376–418.

54. Loyola A, Bonaldi T, Roche D, Imhof A, Almouzni G. PTMs on H3 variants before chromatin assembly potentiate their final epigenetic state. Mol Cell. 2006 Oct 20;24(2):309–16.

55. Sidoli S, Lopes M, Lund PJ, Goldman N, Fasolino M, Coradin M, et al. A mass spectrometry-based assay using metabolic labeling to rapidly monitor chromatin accessibility of modified histone proteins. Sci Rep. 2019 Sep 20;9(1):13613.

56. Zhang H, Gayen S, Xiong J, Zhou B, Shanmugam AK, Sun Y, et al. MLL1 inhibition reprograms epiblast stem cells to naive pluripotency. Cell Stem Cell. 2016 Apr 7;18(4):481–94.

57. Skene PJ, Hernandez AE, Groudine M, Henikoff S. The nucleosomal barrier to promoter escape by RNA polymerase II is overcome by the chromatin remodeler Chd1. eLife. 2014 Apr 15;3:e02042.

58. Tavares L, Dimitrova E, Oxley D, Webster J, Poot R, Demmers J, et al. RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3. Cell. 2012 Feb 17;148(4):664–78.

59. Sen S, Block KF, Pasini A, Baylin SB, Easwaran H. Genome-wide positioning of bivalent mononucleosomes. BMC Med Genomics. 2016 Sep 15;9(1):60.

60. Harikumar A, Meshorer E. Chromatin remodeling and bivalent histone modifications in embryonic stem cells. EMBO Rep. 2015 Dec;16(12):1609–19.

61. Mas G, Blanco E, Ballaré C, Sansó M, Spill YG, Hu D, et al. Promoter bivalency favors an open chromatin architecture in embryonic stem cells. Nat Genet. 2018 Oct;50(10):1452–62.

62. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell. 2006 Apr 21;125(2):315–26.

63. D'Arcy S, Martin KW, Panchenko T, Chen X, Bergeron S, Stargell LA, et al. Chaperone Nap1 shields histone surfaces used in a nucleosome and can put H2A-H2B in an unconventional tetrameric form. Mol Cell. 2013 Sep 12;51(5):662–77.

64. Tyler JK, Adams CR, Chen SR, Kobayashi R, Kamakaka RT, Kadonaga JT. The RCAF complex mediates chromatin assembly during DNA replication and repair. Nature. 1999 Dec 2;402(6761):555–60.

65. Cheema MS, Ausió J. The Structural Determinants behind the Epigenetic Role of Histone Variants. Genes (Basel). 2015 Jul 23;6(3):685–713.

66. Kari V, Karpiuk O, Tieg B, Kriegs M, Dikomey E, Krebber H, et al. A subset of histone H2B genes produces polyadenylated mRNAs under a variety of cellular conditions. PLoS ONE. 2013 May 22;8(5):e63745.

67. Dávila López M, Samuelsson T. Early evolution of histone mRNA 3' end processing. RNA. 2008 Jan;14(1):1–10.

68. Wagner EJ, Burch BD, Godfrey AC, Salzler HR, Duronio RJ, Marzluff WF. A genome-wide RNA interference screen reveals that variant histones are necessary for replication-dependent histone pre-mRNA processing. Mol Cell. 2007 Nov 30;28(4):692–9.

69. Biterge B, Schneider R. Histone variants: key players of chromatin. Cell Tissue Res. 2014 Jun;356(3):457–66.

70. Ray-Gallet D, Almouzni G. Nucleosome dynamics and histone variants. Essays Biochem. 2010 Sep 20;48(1):75–87.

71. Weber CM, Henikoff JG, Henikoff S. H2A.Z nucleosomes enriched over active genes are homotypic. Nat Struct Mol Biol. 2010 Dec;17(12):1500–7.

72. Dimitriadis EK, Weber C, Gill RK, Diekmann S, Dalal Y. Tetrameric organization of vertebrate centromeric nucleosomes. Proc Natl Acad Sci USA. 2010 Nov 23;107(47):20317–22.

73. Chakravarthy S, Luger K. The histone variant macro-H2A preferentially forms "hybrid nucleosomes". J Biol Chem. 2006 Sep 1;281(35):25522–31.

74. Horikoshi N, Arimura Y, Taguchi H, Kurumizaka H. Crystal structures of heterotypic nucleosomes containing histones H2A.Z and H2A. Open Biol. 2016 Jun;6(6).

75. Ghiraldini FG, Filipescu D, Bernstein E. Solid tumours hijack the histone variant network. Nat Rev Cancer. 2021 Apr;21(4):257–75.

76. Yildirim O, Hung J-H, Cedeno RJ, Weng Z, Lengner CJ, Rando OJ. A system for genome-wide histone variant dynamics in ES cells reveals dynamic MacroH2A2 replacement at promoters. PLoS Genet. 2014 Aug 7;10(8):e1004515.

77. Schlesinger S, Kaffe B, Melcer S, Aguilera JD, Sivaraman DM, Kaplan T, et al. A hyperdynamic H3.3 nucleosome marks promoter regions in pluripotent embryonic stem cells. Nucleic Acids Res. 2017 Dec 1;45(21):12181–94.

78. Turinetto V, Giachino C. Histone variants as emerging regulators of embryonic stem cell identity. Epigenetics. 2015;10(7):563–73.

79. Mendiratta S, Gatto A, Almouzni G. Histone supply: Multitiered regulation ensures chromatin dynamics throughout the cell cycle. J Cell Biol. 2019 Jan 7;218(1):39–54.

80. McManus KJ, Hendzel MJ. ATM-dependent DNA damage-independent mitotic phosphorylation of H2AX in normally growing mammalian cells. Mol Biol Cell. 2005 Oct;16(10):5013–25.

81. Nekrasov M, Soboleva TA, Jack C, Tremethick DJ. Histone variant selectivity at the transcription start site: H2A.Z or H2A.Lap1. Nucleus. 2013 Dec;4(6):431–8.

82. Chakravarthy S, Gundimella SKY, Caron C, Perche P-Y, Pehrson JR, Khochbin S, et al. Structural characterization of the histone variant macroH2A. Mol Cell Biol. 2005 Sep;25(17):7616–24.

83. Creppe C, Posavec M, Douet J, Buschbeck M. MacroH2A in stem cells: a story beyond gene repression. Epigenomics. 2012 Apr;4(2):221–7.

84. Creppe C, Janich P, Cantariño N, Noguera M, Valero V, Musulén E, et al. MacroH2A1 regulates the balance between self-renewal and differentiation commitment in embryonic and adult stem cells. Mol Cell Biol. 2012 Apr;32(8):1442–52.

85. Ioudinkova ES, Barat A, Pichugin A, Markova E, Sklyar I, Pirozhkova I, et al. Distinct distribution of ectopically expressed histone variants H2A.Bbd and

MacroH2A in open and closed chromatin domains. PLoS ONE. 2012 Oct 30;7(10):e47157.

86. Sansoni V, Casas-Delucchi CS, Rajan M, Schmidt A, Bönisch C, Thomae AW, et al. The histone variant H2A.Bbd is enriched at sites of DNA synthesis. Nucleic Acids Res. 2014 Jun;42(10):6405–20.

87. González-Romero R, Méndez J, Ausió J, Eirín-López JM. Quickly evolving histones, nucleosome stability and chromatin folding: all about histone H2A.Bbd. Gene. 2008 Apr 30;413(1–2):1–7.

88. Elsässer SJ, Noh K-M, Diaz N, Allis CD, Banaszynski LA. Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells. Nature. 2015 Jun 11;522(7555):240–4.

89. Akiyama T, Suzuki O, Matsuda J, Aoki F. Dynamic replacement of histone H3 variants reprograms epigenetic marks in early mouse embryos. PLoS Genet. 2011 Oct 6;7(10):e1002279.

90. Corpet A, Olbrich T, Gwerder M, Fink D, Stucki M. Dynamics of histone H3.3 deposition in proliferating and senescent cells reveals a DAXX-dependent targeting to PML-NBs important for pericentromeric heterochromatin organization. Cell Cycle. 2014;13(2):249–67.

91. Régnier V, Vagnarelli P, Fukagawa T, Zerjal T, Burns E, Trouche D, et al. CENP-A is required for accurate chromosome segregation and sustained kinetochore association of BubR1. Mol Cell Biol. 2005 May;25(10):3967–81.

92. Ross JE, Woodlief KS, Sullivan BA. Inheritance of the CENP-A chromatin domain is spatially and temporally constrained at human centromeres. Epigenetics Chromatin. 2016 May 31;9:20.

93. Frank D, Doenecke D, Albig W. Differential expression of human replacement and cell cycle dependent H3 histone genes. Gene. 2003 Jul;312:135–43.

94. Szenker E, Ray-Gallet D, Almouzni G. The double face of the histone variant H3.3. Cell Res. 2011 Mar;21(3):421–34.

95. Ray-Gallet D, Ricketts MD, Sato Y, Gupta K, Boyarchuk E, Senda T, et al. Functional activity of the H3.3 histone chaperone complex HIRA requires trimerization of the HIRA subunit. Nat Commun. 2018 Aug 6;9(1):3103.

96. Elsässer SJ, Huang H, Lewis PW, Chin JW, Allis CD, Patel DJ. DAXX envelops a histone H3.3-H4 dimer for H3.3-specific recognition. Nature. 2012 Nov 22;491(7425):560–5.

97. Elsaesser SJ, Allis CD. HIRA and Daxx constitute two independent histone H3.3-containing predeposition complexes. Cold Spring Harb Symp Quant Biol. 2010 Nov 3;75:27–34.

98. Goldberg AD, Banaszynski LA, Noh K-M, Lewis PW, Elsaesser SJ, Stadler S, et al. Distinct factors control histone variant H3.3 localization at specific genomic regions. Cell. 2010 Mar 5;140(5):678–91.

99. Ahmad K, Henikoff S. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. Mol Cell. 2002 Jun;9(6):1191–200.

100. Song Y, Seol J-H, Yang J-H, Kim H-J, Han J-W, Youn H-D, et al. Dissecting the roles of the histone chaperones reveals the evolutionary conserved mechanism of transcription-coupled deposition of H3.3. Nucleic Acids Res. 2013 May 1;41(10):5199–209.

101. Rai TS, Cole JJ, Nelson DM, Dikovskaya D, Faller WJ, Vizioli MG, et al. HIRA orchestrates a dynamic chromatin landscape in senescence and is required for suppression of neoplasia. Genes Dev. 2014 Dec 15;28(24):2712–25.

102. Tang Y, Puri A, Ricketts MD, Rai TS, Hoffmann J, Hoi E, et al. Identification of an ubinuclein 1 region required for stability and function of the human HIRA/UBN1/CABIN1/ASF1a histone H3.3 chaperone complex. Biochemistry. 2012 Mar 27;51(12):2366–77.

103. Rai TS, Puri A, McBryan T, Hoffman J, Tang Y, Pchelintsev NA, et al. Human CABIN1 is a functional member of the human HIRA/UBN1/ASF1a histone H3.3 chaperone complex. Mol Cell Biol. 2011 Oct;31(19):4107–18.

104. Zhang H, Gan H, Wang Z, Lee J-H, Zhou H, Ordog T, et al. RPA Interacts with HIRA and Regulates H3.3 Deposition at Gene Regulatory Elements in Mammalian Cells. Mol Cell. 2017 Jan 19;65(2):272–84.

105. Santenard A, Ziegler-Birling C, Koch M, Tora L, Bannister AJ, Torres-Padilla M-E. Heterochromatin formation in the mouse embryo requires critical residues of the histone variant H3.3. Nat Cell Biol. 2010 Sep;12(9):853–62.

106. Lewis PW, Elsaesser SJ, Noh K-M, Stadler SC, Allis CD. Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. Proc Natl Acad Sci USA. 2010 Aug 10;107(32):14075–80.

107. Udugama M, M Chang FT, Chan FL, Tang MC, Pickett HA, R McGhie JD, et al. Histone variant H3.3 provides the heterochromatic H3 lysine 9 tri-methylation mark at telomeres. Nucleic Acids Res. 2015 Dec 2;43(21):10227–37.

108. Sadic D, Schmidt K, Groh S, Kondofersky I, Ellwart J, Fuchs C, et al. Atrx promotes heterochromatin formation at retrotransposons. EMBO Rep. 2015 Jul;16(7):836–50.

109. Navarro C, Lyu J, Katsori A-M, Caridha R, Elsässer SJ. An embryonic stem cell-specific heterochromatin state promotes core histone exchange in the absence of DNA accessibility. Nat Commun. 2020 Oct 9;11(1):5095.

110. Rowe HM, Jakobsson J, Mesnard D, Rougemont J, Reynard S, Aktas T, et al. KAP1 controls endogenous retroviruses in embryonic stem cells. Nature. 2010 Jan 14;463(7278):237–40.

111. Sachs P, Ding D, Bergmaier P, Lamp B, Schlagheck C, Finkernagel F, et al. SMARCAD1 ATPase activity is required to silence endogenous retroviruses in embryonic stem cells. Nat Commun. 2019 Mar 22;10(1):1335.

112. Groh S, Milton AV, Marinelli LK, Sickinger CV, Russo A, Bollig H, et al. Morc3 silences endogenous retroviruses by enabling Daxx-mediated histone H3.3 incorporation. Nat Commun. 2021 Oct 14;12(1):5996.

113. Chan K-M, Fang D, Gan H, Hashizume R, Yu C, Schroeder M, et al. The histone H3.3K27M mutation in pediatric glioma reprograms H3K27 methylation and gene expression. Genes Dev. 2013 May 1;27(9):985–90.

114. Schulte JD, Buerki RA, Lapointe S, Molinaro AM, Zhang Y, Villanueva-Meyer JE, et al. Clinical, radiologic, and genetic characteristics of histone H3 K27M-mutant diffuse midline gliomas in adults. Neurooncol Adv. 2020 Dec;2(1):vdaa142.

115. Day CA, Grigore F, Hakkim FL, Langfald A, Fadness S, Schwab P, et al. The histone H3.3 K27M mutation found in diffuse midline gliomas coordinately disrupts adjacent H3.3 Ser31 phosphorylation and the fidelity of chromosome segregation. BioRxiv. 2022 May 28;

116. Wang L, Shao L, Li H, Yao K, Duan Z, Zhi C, et al. Histone H3.3 G34-mutant Diffuse Gliomas in Adults. Am J Surg Pathol. 2022 Feb 1;46(2):249–57.

117. Lim KY, Won JK, Park C-K, Kim S-K, Choi SH, Kim T, et al. H3 G34-mutant high-grade glioma. Brain Tumor Pathol. 2021 Jan;38(1):4–13.

118. Behjati S, Tarpey PS, Presneau N, Scheipl S, Pillay N, Van Loo P, et al. Distinct H3F3A and H3F3B driver mutations define chondroblastoma and giant cell tumor of bone. Nat Genet. 2013 Dec;45(12):1479–82.

119. Chan CS, Laddha SV, Lewis PW, Koletsky MS, Robzyk K, Da Silva E, et al. ATRX, DAXX or MEN1 mutant pancreatic neuroendocrine tumors are a distinct alpha-cell signature subgroup. Nat Commun. 2018 Oct 12;9(1):4158.

120. Kurihara S, Hiyama E, Onitake Y, Yamaoka E, Hiyama K. Clinical features of ATRX or DAXX mutated neuroblastoma. J Pediatr Surg. 2014 Dec;49(12):1835–8.

121. Ren X, Tu C, Tang Z, Ma R, Li Z. Alternative lengthening of telomeres phenotype and loss of ATRX expression in sarcomas. Oncol Lett. 2018 May;15(5):7489–96.

122. Badens C, Lacoste C, Philip N, Martini N, Courrier S, Giuliano F, et al. Mutations in PHD-like domain of the ATRX gene correlate with severe psychomotor impairment and severe urogenital abnormalities in patients with ATRX syndrome. Clin Genet. 2006 Jul;70(1):57–62.

123. Warburton PE, Cooke CA, Bourassa S, Vafa O, Sullivan BA, Stetten G, et al. Immunolocalization of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres. Curr Biol. 1997 Nov 1;7(11):901–4.

124. Guse A, Carroll CW, Moree B, Fuller CJ, Straight AF. In vitro centromere and kinetochore assembly on defined chromatin templates. Nature. 2011 Aug 28;477(7364):354–8.

125. Van Hooser AA, Ouspenski II, Gregson HC, Starr DA, Yen TJ, Goldberg ML, et al. Specification of kinetochore-forming chromatin by the histone H3 variant CENP-A. J Cell Sci. 2001 Oct 1;114(Pt 19):3529–42.

126. Schueler MG, Sullivan BA. Structural and functional dynamics of human centromeric chromatin. Annu Rev Genomics Hum Genet. 2006;7:301–13.

127. Greaves IK, Rangasamy D, Ridgway P, Tremethick DJ. H2A.Z contributes to the unique 3D structure of the centromere. Proc Natl Acad Sci USA. 2007 Jan 9;104(2):525–30.

128. Geiss CP, Keramisanou D, Sekulic N, Scheffer MP, Black BE, Frangakis AS. CENP-A arrays are more condensed than canonical arrays at low ionic strength. Biophys J. 2014 Feb 18;106(4):875–82.

129. Guo LY, Allu PK, Zandarashvili L, McKinley KL, Sekulic N, Dawicki-McKenna JM, et al. Centromeres are maintained by fastening CENP-A to DNA and directing an arginine anchor-dependent nucleosome transition. Nat Commun. 2017 Jun 9;8:15775.

130. Arimura Y, Tachiwana H, Takagi H, Hori T, Kimura H, Fukagawa T, et al. The CENP-A centromere targeting domain facilitates H4K20 monomethylation in the nucleosome by structural polymorphism. Nat Commun. 2019 Feb 4;10(1):576.

131. Goutte-Gattat D, Shuaib M, Ouararhni K, Gautier T, Skoufias DA, Hamiche A, et al. Phosphorylation of the CENP-A amino-terminus in mitotic centromeric chromatin is required for kinetochore function. Proc Natl Acad Sci USA. 2013 May 21;110(21):8579–84.

132. Bodor DL, Valente LP, Mata JF, Black BE, Jansen LET. Assembly in G1 phase and long-term stability are unique intrinsic features of CENP-A nucleosomes. Mol Biol Cell. 2013 Apr;24(7):923–32.

133. Silva MCC, Bodor DL, Stellfox ME, Martins NMC, Hochegger H, Foltz DR, et al. Cdk activity couples epigenetic centromere inheritance to cell cycle progression. Dev Cell. 2012 Jan 17;22(1):52–63.

134. Dunleavy EM, Almouzni G, Karpen GH. H3.3 is deposited at centromeres in S phase as a placeholder for newly assembled CENP-A in G_1 phase. Nucleus. 2011 Apr;2(2):146–57.

135. Shukla M, Tong P, White SA, Singh PP, Reid AM, Catania S, et al. Centromere DNA Destabilizes H3 Nucleosomes to Promote CENP-A Deposition during the Cell Cycle. Curr Biol. 2018 Dec 17;28(24):3924-3936.e4.

136. Swartz SZ, McKay LS, Su K-C, Bury L, Padeganeh A, Maddox PS, et al. Quiescent Cells Actively Replenish CENP-A Nucleosomes to Maintain Centromere Identity and Proliferative Potential. Dev Cell. 2019 Oct 7;51(1):35-48.e7.

137. Liu H. Insights into centromeric transcription in mitosis. Transcription. 2016;7(1):21–5.

138. Lee H-R, Hayden KE, Willard HF. Organization and molecular evolution of CENP-A--associated satellite DNA families in a basal primate genome. Genome Biol Evol. 2011 Aug 9;3:1136–49.

139. Zeng K, de las Heras JI, Ross A, Yang J, Cooke H, Shen MH. Localisation of centromeric proteins to a fraction of mouse minor satellite DNA on a mini-chromosome in human, mouse and chicken cells. Chromosoma. 2004 Sep;113(2):84–91.

140. Fujita R, Otake K, Arimura Y, Horikoshi N, Miya Y, Shiga T, et al. Stable complex formation of CENP-B with the CENP-A nucleosome. Nucleic Acids Res. 2015 May 26;43(10):4909–22.

141. Moree B, Meyer CB, Fuller CJ, Straight AF. CENP-C recruits M18BP1 to centromeres to promote CENP-A chromatin assembly. J Cell Biol. 2011 Sep 19;194(6):855–71.

142. Zasadzińska E, Barnhart-Dailey MC, Kuich PHJL, Foltz DR. Dimerization of the CENP-A assembly factor HJURP is required for centromeric nucleosome deposition. EMBO J. 2013 Jul 31;32(15):2113–24.

143. Dunleavy EM, Roche D, Tagami H, Lacoste N, Ray-Gallet D, Nakamura Y, et al. HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. Cell. 2009 May 1;137(3):485–97.

144. Zasadzińska E, Huang J, Bailey AO, Guo LY, Lee NS, Srivastava S, et al. Inheritance of CENP-A Nucleosomes during DNA Replication Requires HJURP. Dev Cell. 2018 Nov 5;47(3):348-362.e7.

145. Barnhart-Dailey MC, Trivedi P, Stukenberg PT, Foltz DR. HJURP interaction with the condensin II complex during G1 promotes CENP-A deposition. Mol Biol Cell. 2017 Jan 1;28(1):54–64.

146. Foltz DR, Jansen LET, Bailey AO, Yates JR, Bassett EA, Wood S, et al. Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. Cell. 2009 May 1;137(3):472–84.

147. Yu Z, Zhou X, Wang W, Deng W, Fang J, Hu H, et al. Dynamic phosphorylation of CENP-A at Ser68 orchestrates its cell-cycle-dependent deposition at centromeres. Dev Cell. 2015 Jan 12;32(1):68–81.

148. Bui M, Pitman M, Nuccio A, Roque S, Donlin-Asp PG, Nita-Lazar A, et al. Internal modifications in the CENP-A nucleosome modulate centromeric dynamics. Epigenetics Chromatin. 2017 Apr 4;10:17.

149. Bui M, Dimitriadis EK, Hoischen C, An E, Quénet D, Giebe S, et al. Cell-cycle-dependent structural transitions in the human CENP-A nucleosome in vivo. Cell. 2012 Jul 20;150(2):317–26.

150. Logsdon GA, Barrey EJ, Bassett EA, DeNizio JE, Guo LY, Panchenko T, et al. Both tails and the centromere targeting domain of CENP-A are required for centromere establishment. J Cell Biol. 2015 Mar 2;208(5):521–31.

151. Maddox PS, Hyndman F, Monen J, Oegema K, Desai A. Functional genomics identifies a Myb domain-containing protein family required for assembly of CENP-A chromatin. J Cell Biol. 2007 Mar 12;176(6):757–63.

152. Sathyan KM, Fachinetti D, Foltz DR. α-amino trimethylation of CENP-A by NRMT is required for full recruitment of the centromere. Nat Commun. 2017 Mar 7;8:14678.

153. Kunitoku N, Sasayama T, Marumoto T, Zhang D, Honda S, Kobayashi O, et al. CENP-A phosphorylation by Aurora-A in prophase is required for enrichment of Aurora-B at inner centromeres and for kinetochore function. Dev Cell. 2003 Dec;5(6):853–64.

154. Zeitlin SG, Shelby RD, Sullivan KF. CENP-A is phosphorylated by Aurora B kinase and plays an unexpected role in completion of cytokinesis. J Cell Biol. 2001 Dec 24;155(7):1147–57.

155. Barra V, Logsdon GA, Scelfo A, Hoffmann S, Hervé S, Aslanian A, et al. Phosphorylation of CENP-A on serine 7 does not control centromere function. Nat Commun. 2019 Jan 11;10(1):175.

156. Fachinetti D, Logsdon GA, Abdullah A, Selzer EB, Cleveland DW, Black BE. CENP-A Modifications on Ser68 and Lys124 Are Dispensable for Establishment, Maintenance, and Long-Term Function of Human Centromeres. Dev Cell. 2017 Jan 9;40(1):104–13.

157. Wang K, Yu Z, Liu Y, Li G. Ser68 Phosphorylation Ensures Accurate Cell-Cycle-Dependent CENP-A Deposition at Centromeres. Dev Cell. 2017 Jan 9;40(1):5–6.

158. Sharma AB, Dimitrov S, Hamiche A, Van Dyck E. Centromeric and ectopic assembly of CENP-A chromatin in health and cancer: old marks and new tracks. Nucleic Acids Res. 2019 Feb 20;47(3):1051–69.

159. Sun X, Clermont P-L, Jiao W, Helgason CD, Gout PW, Wang Y, et al. Elevated expression of the centromere protein-A(CENP-A)-encoding gene as a prognostic and predictive biomarker in human cancers. Int J Cancer. 2016 Aug 15;139(4):899–907.

160. Shrestha RL, Ahn GS, Staples MI, Sathyan KM, Karpova TS, Foltz DR, et al. Mislocalization of centromeric histone H3 variant CENP-A contributes to chromosomal instability (CIN) in human cells. Oncotarget. 2017 Jul 18;8(29):46781–800.

161. Tomonaga T, Matsushita K, Yamaguchi S, Oohashi T, Shimada H, Ochiai T, et al. Overexpression and mistargeting of centromere protein-A in human primary colorectal cancer. Cancer Res. 2003 Jul 1;63(13):3511–6.

162. Athwal RK, Walkiewicz MP, Baek S, Fu S, Bui M, Camps J, et al. CENP-A nucleosomes localize to transcription factor hotspots and subtelomeric sites in human cancer cells. Epigenetics Chromatin. 2015 Jan 13;8:2.

163. Nye J, Sturgill D, Athwal R, Dalal Y. HJURP antagonizes CENP-A mislocalization driven by the H3.3 chaperones HIRA and DAXX. PLoS ONE. 2018 Oct 26;13(10):e0205948.

164. Lacoste N, Woolfe A, Tachiwana H, Garea AV, Barth T, Cantaloube S, et al. Mislocalization of the centromeric histone variant CenH3/CENP-A in human cells depends on the chaperone DAXX. Mol Cell. 2014 Feb 20;53(4):631–44.

165. Chan KL, Palmai-Pallag T, Ying S, Hickson ID. Replication stress induces sister-chromatid bridging at fragile site loci in mitosis. Nat Cell Biol. 2009 Jun;11(6):753–60.

166. Burrack LS, Berman J. Neocentromeres and epigenetically inherited features of centromeres. Chromosome Res. 2012 Jul;20(5):607–19.

167. Shang W-H, Hori T, Martins NMC, Toyoda A, Misu S, Monma N, et al. Chromosome engineering allows the efficient isolation of vertebrate neocentromeres. Dev Cell. 2013 Mar 25;24(6):635–48.

168. Hasson D, Alonso A, Cheung F, Tepperberg JH, Papenhausen PR, Engelen JJM, et al. Formation of novel CENP-A domains on tandem repetitive DNA and across chromosome breakpoints on human chromosome 8q21 neocentromeres. Chromosoma. 2011 Dec;120(6):621–32.

169. Zeitlin SG, Baker NM, Chapados BR, Soutoglou E, Wang JYJ, Berns MW, et al. Double-strand DNA breaks recruit the centromeric histone CENP-A. Proc Natl Acad Sci USA. 2009 Sep 15;106(37):15762–7.

170. Nechemia-Arbely Y, H. Miga K, Shoshani O, Aslanian A, A. McMahon M, Lee AY, et al. DNA replication-mediated error correction of ectopic CENP-A deposition maintains centromere identity. BioRxiv. 2018 Sep 27;

171. Foltz DR, Jansen LET, Black BE, Bailey AO, Yates JR, Cleveland DW. The human CENP-A centromeric nucleosome-associated complex. Nat Cell Biol. 2006 May;8(5):458–69.

172. Bouzinba-Segard H, Guais A, Francastel C. Accumulation of small murine minor satellite transcripts leads to impaired centromeric architecture and function. Proc Natl Acad Sci USA. 2006 Jun 6;103(23):8709–14.

173. Hu Z, Huang G, Sadanandam A, Gu S, Lenburg ME, Pai M, et al. The expression level of HJURP has an independent prognostic impact and predicts the sensitivity to radiotherapy in breast cancer. Breast Cancer Res. 2010 Mar 8;12(2):R18.

174. Valente V, Serafim RB, de Oliveira LC, Adorni FS, Torrieri R, Tirapelli DP da C, et al. Modulation of HJURP (Holliday Junction-Recognizing Protein) levels is correlated with glioblastoma cells survival. PLoS ONE. 2013 Apr 25;8(4):e62200.

175. Serafim RB, Cardoso C, Di Cristofaro LFM, Pienna Soares C, Araújo Silva W, Espreafico EM, et al. HJURP knockdown disrupts clonogenic capacity and increases radiation-induced cell death of glioblastoma cells. Cancer Gene Ther. 2020 May;27(5):319–29.

176. Chen T, Huang H, Zhou Y, Geng L, Shen T, Yin S, et al. HJURP promotes hepatocellular carcinoma proliferation by destabilizing p21 via the MAPK/ERK1/2 and AKT/GSK3β signaling pathways. J Exp Clin Cancer Res. 2018 Aug 15;37(1):193.

177. Brunelle M, Nordell Markovits A, Rodrigue S, Lupien M, Jacques P-É, Gévry N. The histone variant H2A.Z is an important regulator of enhancer activity. Nucleic Acids Res. 2015 Nov 16;43(20):9742–56.

178. Luk E, Ranjan A, Fitzgerald PC, Mizuguchi G, Huang Y, Wei D, et al. Stepwise histone replacement by SWR1 requires dual activation with histone H2A.Z and canonical nucleosome. Cell. 2010 Nov 24;143(5):725–36.

179. Placek BJ, Harrison LN, Villers BM, Gloss LM. The H2A.Z/H2B dimer is unstable compared to the dimer containing the major H2A isoform. Protein Sci. 2005 Feb;14(2):514–22.

180. Chen P, Wang Y, Li G. Dynamics of histone variant H3.3 and its coregulation with H2A.Z at enhancers and promoters. Nucleus. 2014 Feb 3;5(1):21–7.

181. Jin C, Felsenfeld G. Nucleosome stability mediated by histone variants H3.3 and H2A.Z. Genes Dev. 2007 Jun 15;21(12):1519–29.

182. Mchaourab ZF, Perreault AA, Venters BJ. ChIP-seq and ChIP-exo profiling of Pol II, H2A.Z, and H3K4me3 in human K562 cells. Sci Data. 2018 Mar 6;5:180030.

183. Long H, Zhang L, Lv M, Wen Z, Zhang W, Chen X, et al. H2A.Z facilitates licensing and activation of early replication origins. Nature. 2020 Jan;577(7791):576–81.

184. Segala G, Bennesch MA, Pandey DP, Hulo N, Picard D. Monoubiquitination of Histone H2B Blocks Eviction of Histone Variant H2A.Z from Inducible Enhancers. Mol Cell. 2016 Oct 20;64(2):334–46.

185. Wang Y, Liu S, Sun L, Xu N, Shan S, Wu F, et al. Structural insights into histone chaperone Chz1-mediated H2A.Z recognition and histone replacement. PLoS Biol. 2019 May 20;17(5):e3000277.

186. Cauchy P, Koch F, Andrau J-C. Two possible modes of pioneering associated with combinations of H2A.Z and p300/CBP at nucleosome-occupied enhancers. Transcription. 2017 May 27;8(3):179–84.

187. Ryan DP, Tremethick DJ. The interplay between H2A.Z and H3K9 methylation in regulating HP1α binding to linker histone-containing chromatin. Nucleic Acids Res. 2018 Oct 12;46(18):9353–66.

188. Pünzeler S, Link S, Wagner G, Keilhauer EC, Kronbeck N, Spitzer RM, et al. Multivalent binding of PWWP2A to H2A.Z regulates mitosis and neural crest differentiation. EMBO J. 2017 Aug 1;36(15):2263–79.

189. Liang X, Shan S, Pan L, Zhao J, Ranjan A, Wang F, et al. Structural basis of H2A.Z recognition by SRCAP chromatin-remodeling subunit YL1. Nat Struct Mol Biol. 2016 Apr;23(4):317–23.

190. Latrick CM, Marek M, Ouararhni K, Papin C, Stoll I, Ignatyeva M, et al. Molecular basis and specificity of H2A.Z-H2B recognition and deposition by the histone chaperone YL1. Nat Struct Mol Biol. 2016 Apr;23(4):309–16.

191. Link S, Spitzer RMM, Sana M, Torrado M, Völker-Albert MC, Keilhauer EC, et al. PWWP2A binds distinct chromatin moieties and interacts with an MTA1-specific core NuRD complex. Nat Commun. 2018 Oct 16;9(1):4300.

192. Mao Z, Pan L, Wang W, Sun J, Shan S, Dong Q, et al. Anp32e, a higher eukaryotic histone chaperone directs preferential recognition for H2A.Z. Cell Res. 2014 Apr;24(4):389–99.

193. Ding J, Yu C, Sui Y, Wang L, Yang Y, Wang F, et al. The chromatin remodeling protein INO80 contributes to the removal of H2A.Z at the p53-binding site of the p21 gene in response to doxorubicin. FEBS J. 2018 Sep;285(17):3270–85.

194. Alatwi HE, Downs JA. Removal of H2A.Z by INO80 promotes homologous recombination. EMBO Rep. 2015 Aug;16(8):986–94.

195. Nekrasov M, Amrichova J, Parker BJ, Soboleva TA, Jack C, Williams R, et al. Histone H2A.Z inheritance during the cell cycle and its impact on promoter organization and dynamics. Nat Struct Mol Biol. 2012 Nov;19(11):1076–83.

196. Fukuto A, Ikura M, Ikura T, Sun J, Horikoshi Y, Shima H, et al. SUMO modification system facilitates the exchange of histone variant H2A.Z-2 at DNA damage sites. Nucleus. 2018 Jan 1;9(1):87–94.

197. Colino-Sanguino Y, Cornett EM, Moulder D, Smith GC, Hrit J, Cordeiro-Spinetti E, et al. A Read/Write Mechanism Connects p300 Bromodomain Function to H2A.Z Acetylation. iScience. 2019 Nov 22;21:773–88.

198. Ishibashi T, Dryhurst D, Rose KL, Shabanowitz J, Hunt DF, Ausió J. Acetylation of vertebrate H2A.Z and its effect on the structure of the nucleosome. Biochemistry. 2009 Jun 9;48(22):5007–17.

199. Ng MK, Braunschweig U, Blencowe B, Cheung P. Ubiquitylated H2A.Z nucleosomes are associated with nuclear architectural proteins and global transcriptional silencing. BioRxiv. 2019 Sep 5;

200. Draker R, Sarcinella E, Cheung P. USP10 deubiquitylates the histone variant H2A.Z and both are required for androgen receptor-mediated gene activation. Nucleic Acids Res. 2011 May;39(9):3529–42.

201. Feldman M, Vershinin Z, Goliand I, Elia N, Levy D. The methyltransferase SETD6 regulates Mitotic progression through PLK1 methylation. Proc Natl Acad Sci USA. 2019 Jan 22;116(4):1235–40.

202. Tsai C-H, Chen Y-J, Yu C-J, Tzeng S-R, Wu I-C, Kuo W-H, et al. SMYD3-Mediated H2A.Z.1 Methylation Promotes Cell Cycle and Cancer Proliferation. Cancer Res. 2016 Oct 15;76(20):6043–53.

203. Wang L-Y, He Y-X, Li M, Ding J, Sui Y, Conaway JW, et al. DNA-dependent protein kinase catalytic subunit (DNA-PKcs) contributes to incorporation of histone variant H2A.Z into nucleosomes. Protein Cell. 2019 Sep;10(9):694–9.

204. Takahashi D, Orihara Y, Kitagawa S, Kusakabe M, Shintani T, Oma Y, et al. Quantitative regulation of histone variant H2A.Z during cell cycle by ubiquitin proteasome system and SUMO-targeted ubiquitin ligases. Biosci Biotechnol Biochem. 2017 Aug;81(8):1557–60.

205. Dunn CJ, Sarkar P, Bailey ER, Farris S, Zhao M, Ward JM, et al. 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. 2017 Aug 24;4(4).

206. Shen T, Ji F, Wang Y, Lei X, Zhang D, Jiao J. Brain-specific deletion of histone variant H2A.z results in cortical neurogenesis defects and neurodevelopmental disorder. Nucleic Acids Res. 2018 Mar 16;46(5):2290–307.

207. Zovkic IB, Paulukaitis BS, Day JJ, Etikala DM, Sweatt JD. Histone H2A.Z subunit exchange controls consolidation of recent and remote memory. Nature. 2014 Nov 27;515(7528):582–6.

208. Stefanelli G, Azam AB, Walters BJ, Brimble MA, Gettens CP, Bouchard-Cannon P, et al. Learning and Age-Related Changes in Genome-wide H2A.Z Binding in the Mouse Hippocampus. Cell Rep. 2018 Jan 30;22(5):1124–31.

209. Narkaj K, Stefanelli G, Wahdan M, Azam AB, Ramzan F, Steininger CFD, et al. Blocking H2A.Z incorporation via tip60 inhibition promotes systems consolidation of fear memory in mice. eNeuro. 2018 Oct;5(5).

210. Kusakabe M, Oku H, Matsuda R, Hori T, Muto A, Igarashi K, et al. Genetic complementation analysis showed distinct contributions of the N-terminal tail of H2A.Z to epigenetic regulations. Genes Cells. 2016 Feb;21(2):122–35.

211. Dryhurst D, Ausió J. Histone H2A.Z deregulation in prostate cancer. Cause or effect? Cancer Metastasis Rev. 2014 Sep;33(2–3):429–39.

212. Greenberg RS, Long HK, Swigut T, Wysocka J. Single amino acid change underlies distinct roles of H2A.Z subtypes in human syndrome. Cell. 2019 Sep 5;178(6):1421-1436.e24.

213. Gamble MJ, Kraus WL. Multiple facets of the unique histone variant macroH2A: from genomics to cell biology. Cell Cycle. 2010 Jul 1;9(13):2568–74.

214. Kozlowski M, Corujo D, Hothorn M, Guberovic I, Mandemaker IK, Blessing C, et al. MacroH2A histone variants limit chromatin plasticity through two distinct mechanisms. EMBO Rep. 2018 Oct;19(10).

215. Chakravarthy S, Patel A, Bowman GD. The basic linker of macroH2A stabilizes DNA at the entry/exit site of the nucleosome. Nucleic Acids Res. 2012 Sep 1;40(17):8285–95.

216. Muthurajan UM, McBryant SJ, Lu X, Hansen JC, Luger K. The linker region of macroH2A promotes self-association of nucleosomal arrays. J Biol Chem. 2011 Jul 8;286(27):23852–64.

217. Ausió J. Histone variants--the structure behind the function. Brief Funct Genomic Proteomic. 2006 Sep;5(3):228–43.

218. Nashun B, Yukawa M, Liu H, Akiyama T, Aoki F. Changes in the nuclear deposition of histone H2A variants during pre-implantation development in mice. Development. 2010 Nov;137(22):3785–94.

219. Zhang R, Poustovoitov MV, Ye X, Santos HA, Chen W, Daganzo SM, et al. Formation of MacroH2A-containing senescence-associated heterochromatin foci and senescence driven by ASF1a and HIRA. Dev Cell. 2005 Jan;8(1):19–30.

220. Lavigne MD, Vatsellas G, Polyzos A, Mantouvalou E, Sianidis G, Maraziotis I, et al. Composite macroH2A/NRF-1 Nucleosomes Suppress Noise and Generate Robustness in Gene Expression. Cell Rep. 2015 May 19;11(7):1090–101.

221. Chen D, Vollmar M, Rossi MN, Phillips C, Kraehenbuehl R, Slade D, et al. Identification of macrodomain proteins as novel O-acetyl-ADP-ribose deacetylases. J Biol Chem. 2011 Apr 15;286(15):13261–71.

222. Abbott DW, Laszczak M, Lewis JD, Su H, Moore SC, Hills M, et al. Structural characterization of macroH2A containing chromatin. Biochemistry. 2004 Feb 10;43(5):1352–9.

223. Douet J, Corujo D, Malinverni R, Renauld J, Sansoni V, Posavec Marjanović M, et al. MacroH2A histone variants maintain nuclear organization and heterochromatin architecture. J Cell Sci. 2017 May 1;130(9):1570–82.

224. Pasque V, Radzisheuskaya A, Gillich A, Halley-Stott RP, Panamarova M, Zernicka-Goetz M, et al. Histone variant macroH2A marks embryonic differentiation in vivo and acts as an epigenetic barrier to induced pluripotency. J Cell Sci. 2012 Dec 15;125(Pt 24):6094–104.

225. Gaspar-Maia A, Qadeer ZA, Hasson D, Ratnakumar K, Leu NA, Leroy G, et al. MacroH2A histone variants act as a barrier upon reprogramming towards pluripotency. Nat Commun. 2013;4:1565.

226. Hodge DQ, Cui J, Gamble MJ, Guo W. Histone Variant MacroH2A1 Plays an Isoform-Specific Role in Suppressing Epithelial-Mesenchymal Transition. Sci Rep. 2018 Jan 16;8(1):841.

227. Rasmussen TP, Huang T, Mastrangelo MA, Loring J, Panning B, Jaenisch R. Messenger RNAs encoding mouse histone macroH2A1 isoforms are expressed at similar levels in male and female cells and result from alternative splicing. Nucleic Acids Res. 1999 Sep 15;27(18):3685–9.

228. Hurtado-Bagès S, Guberovic I, Buschbeck M. The MacroH2A1.1 - PARP1 Axis at the Intersection Between Stress Response and Metabolism. Front Genet. 2018 Oct 9;9:417.

229. Pehrson JR, Changolkar LN, Costanzi C, Leu NA. Mice without macroH2A histone variants. Mol Cell Biol. 2014 Dec;34(24):4523–33.

230. Changolkar LN, Costanzi C, Leu NA, Chen D, McLaughlin KJ, Pehrson JR. Developmental changes in histone macroH2A1-mediated gene regulation. Mol Cell Biol. 2007 Apr;27(7):2758–64.

231. Chadwick BP, Willard HF. Cell cycle-dependent localization of macroH2A in chromatin of the inactive X chromosome. J Cell Biol. 2002 Jun 24;157(7):1113–23.

232. Ishov AM, Vladimirova OV, Maul GG. Heterochromatin and ND10 are cell-cycle regulated and phosphorylation-dependent alternate nuclear sites of the transcription repressor Daxx and SWI/SNF protein ATRX. J Cell Sci. 2004 Aug 1;117(Pt 17):3807–20.

233. Chu F, Nusinow DA, Chalkley RJ, Plath K, Panning B, Burlingame AL. Mapping post-translational modifications of the histone variant MacroH2A1 using tandem mass spectrometry. Mol Cell Proteomics. 2006 Jan;5(1):194–203.

234. Sun Z, Filipescu D, Andrade J, Gaspar-Maia A, Ueberheide B, Bernstein E. Transcription-associated histone pruning demarcates macroH2A chromatin domains. Nat Struct Mol Biol. 2018 Oct 5;25(10):958–70.

235. Cedeno RJ, Nakauka-Ddamba A, Yousefi M, Sterling S, Leu NA, Li N, et al. The histone variant macroH2A confers functional robustness to the intestinal stem cell compartment. PLoS ONE. 2017 Sep 21;12(9):e0185196.

236. Park SJ, Shim JW, Park HS, Eum DY, Park MT, Mi Yi J, et al. MacroH2A1 downregulation enhances the stem-like properties of bladder cancer cells by transactivation of Lin28B. Oncogene. 2016 Mar 10;35(10):1292–301.

237. Soma A, Sato K, Nakanishi T. Visualization of inactive X chromosome in preimplantation embryos utilizing MacroH2A-EGFP transgenic mouse. Genesis. 2013 Apr;51(4):259–67.

238. Ma Y, Jacobs SB, Jackson-Grusby L, Mastrangelo M-A, Torres-Betancourt JA, Jaenisch R, et al. DNA CpG hypomethylation induces heterochromatin reorganization involving the histone variant macroH2A. J Cell Sci. 2005 Apr 15;118(Pt 8):1607–16.

239. Sporn JC, Jung B. Differential regulation and predictive potential of MacroH2A1 isoforms in colon cancer. Am J Pathol. 2012 Jun;180(6):2516–26.

240. Lo Re O, Fusilli C, Rappa F, Van Haele M, Douet J, Pindjakova J, et al. Induction of cancer cell stemness by depletion of macrohistone H2A1 in hepatocellular carcinoma. Hepatology. 2018 Feb;67(2):636–50.

241. Novikov L, Park JW, Chen H, Klerman H, Jalloh AS, Gamble MJ. QKI-mediated alternative splicing of the histone variant MacroH2A1 regulates cancer cell proliferation. Mol Cell Biol. 2011 Oct;31(20):4244–55.

242. Kapoor A, Goldberg MS, Cumberland LK, Ratnakumar K, Segura MF, Emanuel PO, et al. The histone variant macroH2A suppresses melanoma progression through regulation of CDK8. Nature. 2010 Dec 23;468(7327):1105–9.

243. Giallongo S, Lo Re O, Vinciguerra M. Macro histone variants: emerging rheostats of gastrointestinal cancers. Cancers (Basel). 2019 May 15;11(5).

244. Ehrlich ME, Gandy S. Chromatin plasticity and the pathogenesis of Huntington disease. Proc Natl Acad Sci USA. 2011 Oct 11;108(41):16867–8.

245. Weinberger L, Ayyash M, Novershtern N, Hanna JH. Dynamic stem cell states: naive to primed pluripotency in rodents and humans. Nat Rev Mol Cell Biol. 2016 Mar;17(3):155–69.

246. Wang H, Dey SK. Roadmap to embryo implantation: clues from mouse models. Nat Rev Genet. 2006 Mar;7(3):185–99.

247. Rossant J, Tam PPL. Early human embryonic development: Blastocyst formation to gastrulation. Dev Cell. 2022 Jan 24;57(2):152–65.

248. Taubenschmid-Stowers J, Rostovskaya M, Santos F, Ljung S, Argelaguet R, Krueger F, et al. 8C-like cells capture the human zygotic genome activation program in vitro. Cell Stem Cell. 2022 Mar 3;29(3):449-459.e6.

249. Fan G, Tran J. X chromosome inactivation in human and mouse pluripotent stem cells. Hum Genet. 2011 Aug;130(2):217–22.

250. Blakeley P, Fogarty NME, Del Valle I, Wamaitha SE, Hu TX, Elder K, et al. Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. Development. 2015 Oct 15;142(20):3613.

251. Seisenberger S, Peat JR, Hore TA, Santos F, Dean W, Reik W. Reprogramming DNA methylation in the mammalian life cycle: building and breaking epigenetic barriers. Philos Trans R Soc Lond B Biol Sci. 2013 Jan 5;368(1609):20110330.

252. Mulholland CB, Traube FR, Ugur E, Parsa E, Eckl E-M, Schönung M, et al. Distinct and stage-specific contributions of TET1 and TET2 to stepwise cytosine oxidation in the transition from naive to primed pluripotency. Sci Rep. 2020 Jul 21;10(1):12066.

253. Kumar B, Elsässer SJ. Quantitative Multiplexed ChIP Reveals Global Alterations that Shape Promoter Bivalency in Ground State Embryonic Stem Cells. Cell Rep. 2019 Sep 17;28(12):3274-3284.e5.

254. Subramanian V, Mazumder A, Surface LE, Butty VL, Fields PA, Alwan A, et al. H2A.Z acidic patch couples chromatin dynamics to regulation of gene expression programs during ESC differentiation. PLoS Genet. 2013 Aug 22;9(8):e1003725.

255. Hsu C-C, Zhao D, Shi J, Peng D, Guan H, Li Y, et al. Gas41 links histone acetylation to H2A.Z deposition and maintenance of embryonic stem cell identity. Cell Discov. 2018 Jun 12;4:28.

256. Zhao B, Chen Y, Jiang N, Yang L, Sun S, Zhang Y, et al. Znhit1 controls intestinal stem cell maintenance by regulating H2A.Z incorporation. Nat Commun. 2019 Mar 6;10(1):1071.

257. Li Z, Gadue P, Chen K, Jiao Y, Tuteja G, Schug J, et al. Foxa2 and H2A.Z mediate nucleosome depletion during embryonic stem cell differentiation. Cell. 2012 Dec 21;151(7):1608–16.

258. Creyghton MP, Markoulaki S, Levine SS, Hanna J, Lodato MA, Sha K, et al. H2AZ is enriched at polycomb complex target genes in ES cells and is necessary for lineage commitment. Cell. 2008 Nov 14;135(4):649–61.

259. Sarcinella E, Zuzarte PC, Lau PNI, Draker R, Cheung P. Monoubiquitylation of H2A.Z distinguishes its association with euchromatin or facultative heterochromatin. Mol Cell Biol. 2007 Sep;27(18):6457–68.

260. Wang Y, Long H, Yu J, Dong L, Wassef M, Zhuo B, et al. Histone variants H2A.Z and H3.3 coordinately regulate PRC2-dependent H3K27me3 deposition and gene expression regulation in mES cells. BMC Biol. 2018 Sep 24;16(1):107.

261. Surface LE, Fields PA, Subramanian V, Behmer R, Udeshi N, Peach SE, et al. H2A.Z.1 monoubiquitylation antagonizes BRD2 to maintain poised chromatin in escs. Cell Rep. 2016 Feb 9;14(5):1142–55.

262. Banaszynski LA, Wen D, Dewell S, Whitcomb SJ, Lin M, Diaz N, et al. Hira-dependent histone H3.3 deposition facilitates PRC2 recruitment at developmental loci in ES cells. Cell. 2013 Sep 26;155(1):107–20.

263. Vallot C, Ouimette J-F, Makhlouf M, Féraud O, Pontis J, Côme J, et al. Erosion of X Chromosome Inactivation in Human Pluripotent Cells Initiates with XACT Coating and Depends on a Specific Heterochromatin Landscape. Cell Stem Cell. 2015 May 7;16(5):533–46.

264. Keohane AM, O'neill LP, Belyaev ND, Lavender JS, Turner BM. X-Inactivation and histone H4 acetylation in embryonic stem cells. Dev Biol. 1996 Dec 15;180(2):618–30.

265. Heard E, Disteche CM. Dosage compensation in mammals: fine-tuning the expression of the X chromosome. Genes Dev. 2006 Jul 15;20(14):1848–67.

266. Bernstein E, Muratore-Schroeder TL, Diaz RL, Chow JC, Changolkar LN, Shabanowitz J, et al. A phosphorylated subpopulation of the histone variant macroH2A1 is excluded from the inactive X chromosome and enriched during mitosis. Proc Natl Acad Sci USA. 2008 Feb 5;105(5):1533–8.

267. Chadwick BP, Willard HF. Histone H2A variants and the inactive X chromosome: identification of a second macroH2A variant. Hum Mol Genet. 2001 May 1;10(10):1101–13.

268. Rasmussen TP, Mastrangelo MA, Eden A, Pehrson JR, Jaenisch R. Dynamic relocalization of histone MacroH2A1 from centrosomes to inactive X chromosomes during X inactivation. J Cell Biol. 2000 Sep 4;150(5):1189–98.

269. Liu L, Leng L, Liu C, Lu C, Yuan Y, Wu L, et al. An integrated chromatin accessibility and transcriptome landscape of human pre-implantation embryos. Nat Commun. 2019 Jan 21;10(1):364.

270. Göke J, Lu X, Chan Y-S, Ng H-H, Ly L-H, Sachs F, et al. Dynamic transcription of distinct classes of endogenous retroviral elements marks specific populations of early human embryonic cells. Cell Stem Cell. 2015 Feb 5;16(2):135–41.

271. Kigami D, Minami N, Takayama H, Imai H. MuERV-L is one of the earliest transcribed genes in mouse one-cell embryos. Biol Reprod. 2003 Feb;68(2):651–4.

272. Sun J, Chen M, Xu J, Luo J. Relationships among stop codon usage bias, its context, isochores, and gene expression level in various eukaryotes. J Mol Evol. 2005 Oct;61(4):437–44.

273. Elsässer SJ. Generation of stable amber suppression cell lines. Methods Mol Biol. 2018;1728:237–45.

274. Elsässer SJ, Ernst RJ, Walker OS, Chin JW. Genetic code expansion in stable cell lines enables encoded chromatin modification. Nat Methods. 2016 Feb;13(2):158–64.

275. van Galen P, Viny AD, Ram O, Ryan RJH, Cotton MJ, Donohue L, et al. A multiplexed system for quantitative comparisons of chromatin landscapes. Mol Cell. 2016 Jan 7;61(1):170–80.

276. Khoueiry R, Sohni A, Thienpont B, Luo X, Velde JV, Bartoccetti M, et al. Lineage-specific functions of TET1 in the postimplantation mouse embryo. Nat Genet. 2017 Jul;49(7):1061–72.

277. Gao J, Ma Y, Fu H-L, Luo Q, Wang Z, Xiao Y-H, et al. Non-catalytic roles for TET1 protein negatively regulating neuronal differentiation through srGAP3 in neuroblastoma cells. Protein Cell. 2016 May;7(5):351–61.

278. Stolz P, Mantero AS, Tvardovskiy A, Ugur E, Wange LE, Mulholland CB, et al. TET1 regulates gene expression and repression of endogenous retroviruses independent of DNA demethylation. Nucleic Acids Res. 2022 Aug 26;50(15):8491–511.

279. Uckelmann M, Davidovich C. Not just a writer: PRC2 as a chromatin reader. Biochem Soc Trans. 2021 Jun 30;49(3):1159–70.

280. Denissov S, Hofemeister H, Marks H, Kranz A, Ciotta G, Singh S, et al. Mll2 is required for H3K4 trimethylation on bivalent promoters in embryonic stem cells, whereas Mll1 is redundant. Development. 2014 Feb;141(3):526–37.

281. Perino M, van Mierlo G, Loh C, Wardle SMT, Zijlmans DW, Marks H, et al. Two Functional Axes of Feedback-Enforced PRC2 Recruitment in Mouse Embryonic Stem Cells. Stem Cell Reports. 2020 Dec 8;15(6):1287–300.

282. Cohen I, Bar C, Ezhkova E. Activity of PRC1 and histone H2AK119 monoubiquitination: revising popular misconceptions. Bioessays. 2020 May;42(5):e1900192.

283. Shrestha RL, Rossi A, Wangsa D, Hogan AK, Zaldana KS, Suva E, et al. CENP-A overexpression promotes aneuploidy with karyotypic heterogeneity. J Cell Biol. 2021 Apr 5;220(4).

284. Chung CZ, Amikura K, Söll D. Using genetic code expansion for protein biochemical studies. Front Bioeng Biotechnol. 2020 Oct 19;8:598577.

285. Chrysanthou S, Tang Q, Lee J, Taylor SJ, Zhao Y, Steidl U, et al. The DNA dioxygenase Tet1 regulates H3K27 modification and embryonic stem cell biology independent of its catalytic activity. Nucleic Acids Res. 2022 Apr 8;50(6):3169–89.

286. Singh AM, Trost R, Boward B, Dalton S. Utilizing FUCCI reporters to understand pluripotent stem cell biology. Methods. 2016 May 15;101:4–10.

287. Zielke N, Edgar BA. FUCCI sensors: powerful new tools for analysis of cell proliferation. Wiley Interdiscip Rev Dev Biol. 2015 Oct;4(5):469–87.

288. Kermi C, Lau L, Asadi Shahmirzadi A, Classon M. Disrupting Mechanisms that Regulate Genomic Repeat Elements to Combat Cancer and Drug Resistance. Front Cell Dev Biol. 2022 May 4;10:826461.