

Dissertation

submitted to the
Combined Faculty of Natural Sciences and Mathematics
of the Ruperto Carola University Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Presented by
Dipl.-Biochem. Maja Gehre
born in: Berlin
Oral examination: 30.11.2018

The role of histone H3.3 lysine 4 and lysine 36 residues in mouse embryonic stem cells and neuronal development

Referees:

Dr. Eileen Furlong

Prof. Dr. Frank Lyko

Abstract

Numerous mutations in histone H3 lysine 4 (H3K4) and H3 lysine 36 (H3K36) modifying enzymes have been reported in human disease, yet the role of the H3K4 and H3K36 residues in mammals remain unclear due to the clustered arrays of many histone genes. Replication-dependent canonical H3 (H3.1/H3.2) exists as multiple gene copies and supplies nucleosomes for packaging of newly synthesized DNA during replication. The histone variant H3.3 differs from canonical H3 by only 4 to 5 amino acids, which allow nucleosome assembly independent of DNA replication throughout the cell cycle and in post-mitotic cells. In this study, I set out to investigate the role of the K4 and K36 residues in the histone H3.3 variant, which is enriched at active regions of the mammalian genome and encoded by two isolated genes; therefore amenable to functional analysis. Using CRISPR-Cas9, I mutated the K4 or K36 residue of endogenous H3.3 to unmodifiable alanine (A) in mouse embryonic stem cells (ESCs) and revealed that the K4A mutation, but not K36A, resulted in widespread gene expression changes and impairment of neuronal differentiation into glutamatergic neurons. Furthermore, K4A resulted in significant H3.3 protein depletion at transcription start sites and active enhancers of ESCs - without effects at other sites. Genomic regions depleted of H3.3K4A showed concerted alterations of histone modifications (decreased K27 acetylation and increased K4 methylation) regardless of gene expression changes. In differentiated neurons, the K4A mutation impacted protein stability and resulted in widespread proteasomal degradation of the mutant histone. Thus, H3.3K4 is required for site-specific nucleosome maintenance at regulatory regions, histone stability and cellular differentiation of ESCs. H3.3K36 is not required for H3.3 deposition and turnover inside coding regions, and the K36A mutation affected gene expression at later stages of neurodevelopment. Furthermore, the K36A mutation globally depleted H3K36 di-methylation levels in ESCs, which resulted in a spread of the repressive mark H3K27me₃, suggesting that H3K36 di-methylation is required to restrict the activity of PRC2. This study demonstrates a direct link between a specific histone residue (H3K4) and histone maintenance at promoters and enhancers, and that H3.3 provides a platform for analyzing the role of histone residues in mammals.

Zusammenfassung

Zahlreiche Mutationen in Histon-Methyltransferasen, die H3 Lysin 4 (H3K4) und H3 Lysin 36 (H3K36) modifizieren können, liegen menschlichen Krankheiten zugrunde. Trotzdem bleibt es unklar, welche Relevanz die Lysinreste H3K4 und H3K36 in Säugetieren haben, weil die Histongene in zahlreichen Kopien innerhalb von sich wiederholenden Genclustern vorliegen. Die Replikations-abhängigen H3 Proteine (H3.1/H3.2) werden von mehreren Genen kodiert, die während der Replikation die notwendigen Nukleosomen bereitstellen um die neu synthetisierte DNA im Zellkern zu verpacken. Die Histonvariante H3.3 unterscheidet sich vom kanonischen H3 in nur 4 bis 5 Aminosäuren, welches es ermöglichen, dass Nukleosomen unabhängig von der DNA Replikation während des gesamten Zellzyklus und auch in post-mitotischen Zellen zusammengesetzt werden. In dieser Studie habe ich die Funktion von Lysin 4 und 36 in der Histonvariante H3.3 untersucht, welche vor allem in aktiven Regionen des menschlichen Genoms angereichert ist und von nur zwei unabhängigen Genen kodiert wird, was H3.3 zu einem passenden Kandidaten für eine funktionelle Analyse macht. Mit Hilfe von CRISPR-Cas9 habe ich Lysin 4 oder Lysin 36 im endogenen H3.3 von murinen embryonischen Stammzellen in einen unmodifizierbaren Alaninrest (K4A/K36A) umgewandelt. So konnte ich zeigen, dass die K4A, aber nicht die K36A, Mutation die Genexpression in Stammzellen umfassend beeinflusst und die neuronale Differenzierung verhindert. Desweiteren hat die K4A Mutation eine signifikante Reduktion des H3.3 Proteins an Transkriptionsstartstellen und aktiven Enhancern zur Folge, ohne dass es sein Vorkommen an anderen Stellen beeinflusste. Die Regionen des Genoms, die von der Reduktion betroffen waren, zeigten abgestimmte Änderungen an Histonmodifikationen (verminderte K27 Acetylierung und erhöhte K4 Methylierung), die unabhängig von den Genexpressionsänderungen waren. In differenzierten Neuronen hingegen beeinflusste die H3.3K4A Mutation die Histonstabilität und resultierte in umfangreichen Abbau des mutierten Histons durch das Proteasomensystem. Daraus lässt sich schließen, dass H3.3K4 für die Erhaltung von Nukleosomen an regulatorischen Elementen, Histonstabilität und zelluläre Differenzierung von embryonischen Stammzellen notwendig ist. H3.3K36 hingegen war nicht notwendig fuer die Deposition und den Austausch von H3.3 in Genen und die K36A Mutation beeinflusste die Genexpression erst während der neuronalen Entwicklung. Desweiteren führte die H3.3K36A Mutation zu einer umfassenden Verminderung der H3K36 Dimethylierung und gleichzeitigen Ausbreitung von repressiver H3K27 Trimethylierung. Dies deutet darauf hin, dass H3K36 Dimethylierung die Funktion hat die Aktivität von PRC2 einzuschränken. Diese Arbeit zeigt eine direkte Verbindung zwischen einem spezifischen Histon Lysinrest (H3K4) und der Erhaltung von Histonen an Promotoren und Enhancern auf. Mutationen in H3.3 bieten somit eine geeignete Plattform um die Funktion von Histonresten in Säugetieren zu testen.

Acknowledgement

I would like to express my genuine appreciation for the guidance and support of my supervisor, Kyung-Min Noh. I truly appreciate your constant encouragement and dedication to this project. I have spent four memorable years at EMBL Heidelberg, which have been full of opportunities and joy. I would also like to extend my gratitude to the members of my TAC and PhD defense committees, Eileen Furlong, Frank Lyko, Oliver Stegle, Judith Zaugg, Ingrid Lohmann, and Sandra Hake. Special thanks to the members of the outstanding core facilities at EMBL Heidelberg (Gene Core, FACS, Proteomics), especially to Malte Paulsen, Diana Ordonez and Vladmir Benes for their time and expertise. Special thanks to Nichole Diaz for all her help and guidance in the cell culture, and for being her amazing self. Also, I would like to thank Daria Bunina, Simone Sidoli and Marlena Lübke for their efforts and contributions to this project. And to the current and past members of the Noh group and the Genome Biology Unit for helpful discussions and for providing such a great working environment.

I would also like to thank all my friends who supported and encouraged me and made my time as a PhD student all the more enjoyable. There are so many precious memories I share with you that I will not forget for the rest of my life. I want to thank all my flatmates and friends back home who are simply the best thing to come home to. Also, I want to thank Matthias who has always given me a reason to wake up in the morning. Many thanks to my beloved sisters and grandma.

I dedicate this thesis to my parents who have always been there for me and helped me in my education, I love you.

Contents

Page

List of Figures

List of Tables

1	Introduction	1
1.1	Histones & Chromatin	1
1.2	Histone modifications	2
1.2.1	Histone Acetylation	3
1.2.2	Histone methylation	3
1.2.3	Properties of histone modifications	4
1.3	Histone variants	7
1.3.1	Histone H3 variants	8
1.3.2	H3 in proliferating and quiescent cells	9
1.3.3	Contributions of H3.3 to development and viability of multicellular organisms	10
1.4	Chromatin assembly and remodeling	12
1.4.1	Chromatin assembly by histone chaperones	12
1.4.2	Chromatin remodeling	15
1.4.3	Histone turnover	16
1.5	Systems to probe the contribution of histone residues to nucleosome function	20
1.6	Aims & Motivation	22
2	CRISPR-mediated gene editing of mammalian histone H3.3	23
2.1	Background	23
2.2	Results	26
2.2.1	Delivery of Cas9 and guide into mouse embryonic stem cells	26
2.2.2	Optimization of CRISPR-Cas9 mediated gene editing in ESCs	26
2.2.3	High-throughput screening for CRISPR editing	28
2.2.4	Knockout of H3.3A	32
2.2.5	CRISPR off-target analysis for large chromosomal deletions	33
2.3	Discussion	39
2.4	Materials & Methods	41
3	Functional characterization of histone H3.3 mutants	46
3.1	Background	46
3.2	Results	50
3.2.1	Characterization of H3.3K4A and H3.3K36A mutant ESCs and their developmental potential	50
3.2.2	Impact of H3.3K4A and H3.3K36A mutations on gene expression in ESCs and neurons	53

3.2.3	The epigenetic landscape of H3.3K4A and H3.3K36A mutants	66
3.3	Discussion & Outlook	74
3.4	Materials & Methods	80
4	Supplementary Material	88
	Bibliography	93

List of Figures

1.1	Structure of the nucleosome core particle	1
1.2	Post-translational modifications of the N-terminal histone tails	3
1.3	Schematic genomic region depicting the distribution of histone modifications at promoters, enhancers, gene bodies and boundary elements.	7
1.4	Post-translational modifications of the N-terminal histone tails	8
1.5	Assembly pathways of histone H3 variants throughout the cell cycle.	13
1.6	Schematic representation of histone exchange at regulatory elements and inside coding regions.	19
2.1	CRISPR-Cas9 mediated genome engineering.	24
2.2	Transduction efficiency in ESCs with a plasmid-based CRISPR system using nucleofection.	26
2.3	Comparison of single-stranded and double-stranded DNA repair templates for CRISPR editing in untreated and Scr7-treated ESCs.	28
2.4	Mismatch qPCR screen detects CRISPR-mediated point-mutations in H3.3B.	30
2.5	Sanger sequencing confirming the successful integration of nucleotide changes into H3.3B at lysine 4 and 36.	31
2.6	Successful knock-out of <i>H3f3a</i> in CRISPR clones.	33
2.7	Prediction of chromosomal rearrangements from RNA-Sequencing data.	36
2.8	Comparison between RNA-Seq and DNA-Seq data in predicting chromosomal rearrangements	37
2.9	Overview of CRISPR-editing workflow in mouse ESCs.	38
3.1	Experimental design to study the role of histone H3.3K4 and H3.3K36 residues.	50
3.2	Morphology of control and K4A/K36A mutant ESCs.	51
3.3	Self-renewal of K4A/K36A mutant ESCs is maintained but K4A results in slower cell growth during neuro-differentiation.	51
3.4	Cell cycle analysis	52
3.5	H3.3K4A, but not K36A, results in erratic morphology during neuro-differentiation from ESCs.	53

3.6	H3.3K4A and K36A induce distinct changes in transcription in ESCs and neurons.	54
3.7	Gene expression analysis of H3.3K4/K36A mutant ESCs.	55
3.8	Gene expression analysis of H3.3K4/K36A mutant neurons.	56
3.9	H3.3K4A affects the activation of developmental genes and repression of pluripotency genes.	57
3.10	H3.3K4A, but not K36A, is depleted at TSS and enhancers in ESCs.	58
3.11	H3.3K4A is depleted at TSS, but not TES.	59
3.12	H3.3K36 and H3.3K4A are not required for histone turnover at TES.	60
3.13	H3.3K4A mutation does not affect deposition by chaperones.	61
3.14	Remodelers Chd1, Chd4, Chd8, Ep400, Smarca4 are enriched at regulatory regions and Chd1 displays a similar enrichment profile to H3.3 at promoters in ESCs.	63
3.15	Lysine 4 residue is required for histone H3 protein stability in differentiated neurons.	63
3.16	H3.3K4A instability is not due to lower expression of H3.3-specific chaperones.	64
3.17	Lysine 4 residue is required for histone H3 protein stability in HEK293T cells.	65
3.18	H3.3K36A mutation reduces global levels of H3K36me2/me3 and H3K27ac, and inversely increases H3K27me3 levels, whereas H3.3K4A mutation reduces H3K4me3 levels.	67
3.19	Turnover analysis of histone modifications in H3.3K36A mutants by Mass Spectrometry.	68
3.20	Splicing analysis	70
3.21	H3.3K4A mutation induces histone modification changes at regulatory elements.	73
S1	Gating strategy for Cas9-GFP-positive cells flow cytometry analysis.	88
S2	<i>In vitro</i> differentiation of ESCs into mature glutamatergic neurons.	89
S3	Gating strategy for cell cycle analysis by flow cytometry.	89
S4	Proteasomal degradation is not responsible for H3.3K4A depletion at TSS.	90
S5	Quality control of ChIP-Seq data.	91
S6	Promoters and enhancers of differentially expressed and unchanged genes are similarly affected by H3.3 depletion, but H3K27ac reduction correlates with gene expression in H3.3K4A mutants.	92

List of Tables

2.1	Summary of CRISPR editing screens to introduce point-mutations into histone H3.3.	27
-----	---	----

ATRX α -thalassemia/mental retardation X-linked

ATAC-seq Assay for Transposase-Accessible Chromatin using sequencing

bp base pair

CENP-A Centromere protein A

ChIP-Seq Chromatin immunoprecipitation coupled with DNA sequencing

ChIP Chromatin Immunoprecipitation

Ct Cycle treshold

D. Drosophila

Daxx Death-domain associated protein

DEG differentially expressed genes

DMEM Dulbecco's modified Eagle's medium

DMSO Dimethyl sulfoxide

DNA Deoxyribonucleic Acid

dNTP deoxynucleotidetriphosphate

EB embryoid body

E. Coli Escherichia Coli

e.g. exempli gratia

EDTA Ethylenediaminetetraacetic acid

ESC embryonic stem cell

GFP green fluorescent protein

FCS Fetal Calf Serum

FW Forward

HAT histone acetyltransferase

HDAC histone deacetylase

HDR homology-directed repair

HEK human embryonic kidney cells

HMT Histone methyltransferases

HMTase Histone methyltransferase

hrs hours

HP1 heterochromatin protein 1

i.e. id est

kbp kilo base pair

MEF mouse embryonic fibroblast

min minutes

Mll mixed-lineage leukemia

mRNA messenger-RNA

NGS next generation sequencing

NHEJ Non-homologous end-joining

NPC neural precursor cell

ORF Open Reading Frame

PCR Polymerase Chain Reaction

Pc Polycomb proteins

Pen/Strep Penicillin/Streptomycin

Pol II RNA Polymerase II

qPCR quantitative real-time PCR

RNA Ribonucleic Acid

RPKM reads per kilo base per million

RT Reverse Transcription

rt room temperature

RV Reverse

S. Saccharomyces

SAM S-adenosyl methionine

sec seconds

Seq Sequencing

SWI/SNF SWItch/Sucrose Non-Fermentable

TBE Tris-Borate-EDTA

TES transcription end sites

TSS transcription start sites

UCSC University of California Santa Cruz

UV Ultra Violet

1 | Introduction

1.1 Histones & Chromatin

In the nucleus of eukaryotic cells, histone proteins are tightly associated with DNA molecules and the resulting DNA-protein complex is called chromatin. The fundamental repeating units of chromatin are the nucleosomes (Kornberg and Thomas, 1974). A single nucleosome contains an octamer of the four core histone proteins - H2A, H2B, H3, and H4 - around which a DNA segment of 146 bp is wound (Luger et al., 1997). (Fig. 1.1). The linker histone, H1, interacts with DNA between nucleosomes.

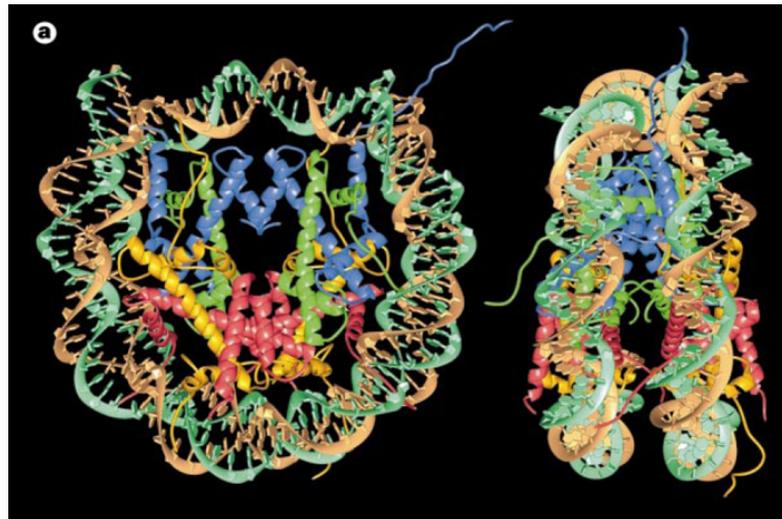


Figure 1.1: Structure of the nucleosome core particle. DNA of 146 bp (brown and green) is wrapped around an histone octamer consisting of H3 (blue), H4 (lightgreen), H2A (yellow), H2B (red). The unstructured histone tails are protruding from the nucleosome core. Nucleosome particle is viewed down the DNA superhelix axis (left) and perpendicular to it (right). Image was taken from Luger et al. (1997).

The association with histones compacts long DNA molecules to fit into the nucleus of cells, which is only 10-20 micrometres in size. Each human cell contains roughly 2 meters of linear DNA, but the packaging into nucleosomes shortens the molecule length about sevenfold and further coiling results in an even shorter chromatin fiber with a diameter of approximately 30 nanometers (Annunziato, 2008). Despite their significant role in packing DNA, histones are much more than just static scaffolding proteins, and many discoveries in recent years have shown that they act as modulators of essen-

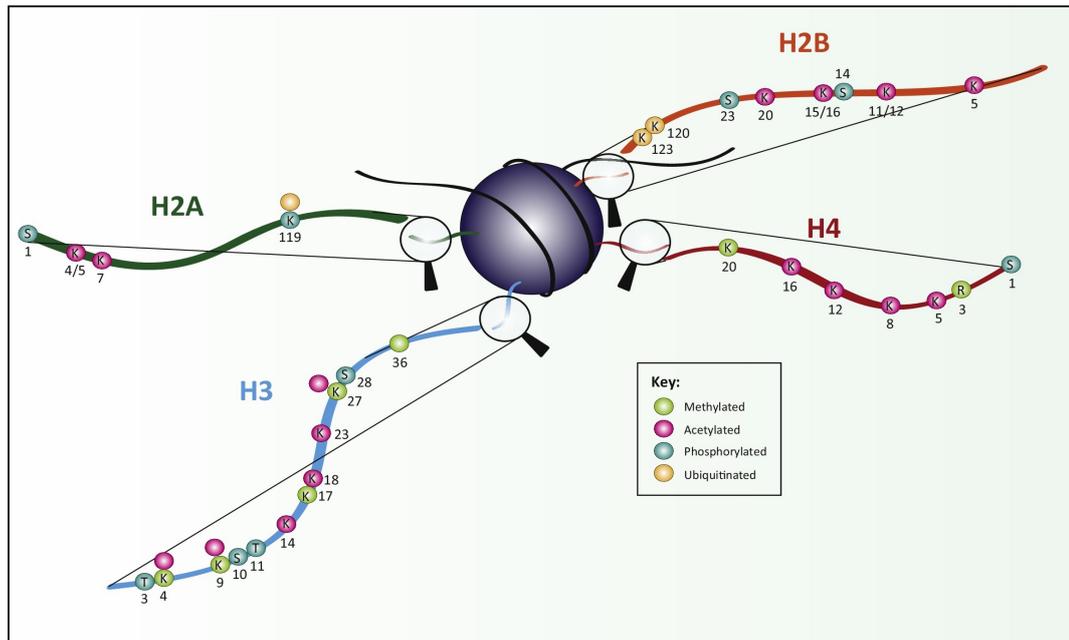
tially all DNA-related processes, including transcription, DNA replication and repair (Kouzarides, 2007). Transcription and replication require that polymerases, enzymes that read and copy DNA, can access the DNA template directly, but chromatin compaction and nucleosomes present barriers to these processes (Li et al., 2007). Therefore, cells require reversible mechanisms for opening up chromatin structure, and these are often regulated by the packaging histones themselves. Histones exhibit at least three dynamic properties that regulate accessibility of the underlying DNA:

1. Histones can be enzymatically modified (e.g. by the addition of acetyl, methyl, phosphate or other groups) (Kouzarides, 2007).
2. Histones are diversified by many histone variants, which can confer specific structural and functional properties to the nucleosome (Talbert and Henikoff, 2010; Henikoff and Smith, 2015).
3. Histones can be displaced or shifted by chromatin remodeling complexes, thereby exposing underlying DNA sequences to polymerases and other enzymes (Narlikar et al., 2013).

These properties are dynamic and reversible because modified or variant nucleosomes can be displaced from chromatin, and higher-order chromatin can be re-established after transcription or replication is completed by protein complexes that facilitate compaction.

1.2 Histone modifications

Histone proteins are basic proteins, which allows them to tightly interact with the negatively charged phosphate backbone of a DNA molecule. Structural analysis of nucleosomes has revealed that the histone octamer forms a core nucleosome particle around which the DNA is wrapped (Fig. 1.1). While the histone core is folded and difficult to access, the histone N- and C-terminal tails are unstructured and protrude from the nucleosome core, making them the most accessible parts of histones (Luger et al., 1997). The N-terminal tails contain many modifiable residues and are the most heavily modified part of histones (Fig. 1.2), but some modifications have also been detected within the globular core (Masumoto et al., 2005; van Leeuwen et al., 2002; Lawrence et al., 2016). The N-terminal tail of histone H3 features the most extended tail sequence with more than 30 known modification events (Bannister and Kouzarides, 2011; Kouzarides, 2007). Often it is the occurrence of multiple modifications that define a distinct chromatin state and they have been proposed to exert their effects via diverse mechanisms. Histone acetylation and methylation are relevant for the content of this thesis and their biological functions are described below in more detail, followed by a general description of the properties of histone modifications.



Trends in Genetics

Figure 1.2: Schematic representation of post-translational modifications of the N-terminal histone tails. Commonly modified histone residues (K = lysine, R = arginine, S = serine, T = threonine) are depicted and the position of the modified amino acid is shown in black. Image taken from Lawrence et al. (2016).

1.2.1 Histone Acetylation

Acetylation of lysines is highly dynamic and regulated by the opposing action of two families of enzymes, histone acetyltransferases (HATs) and histone deacetylases (HDACs) (Bannister and Kouzarides, 2011). HATs catalyse the transfer of an acetyl group from the cofactor acetyl-CoA to the ϵ -amino group of lysine residues. This neutralizes lysine's positive charge and has the potential to reduce the electrostatic interactions between histones and DNA. Generally, HATs are less specific and often modify multiple residues within the histone N-terminal tails. Thanks to their ability to destabilize nucleosomes, HATs are considered to act as transcriptional coactivators. HDAC enzymes oppose the effects of HATs and reverse lysine acetylation, an action that restores the positive charge of the lysine. This potentially stabilizes the nucleosome, which is consistent with HDACs being predominantly transcriptional repressors.

1.2.2 Histone methylation

Histone methylation occurs on the side chains of lysines and arginines and it is diversified by the fact that lysines may be mono-, di- or tri-methylated, whereas arginines may be mono- and di-methylated (symmetrically or asymmetrically) (Bannister and Kouzarides, 2011; Hyun et al., 2017). Unlike acetylation and phosphorylation, histone

methylation does not alter the charge of the histone protein. Histone methyltransferases (HMTases) catalyse the transfer of a methyl group from S-adenosyl methionine (SAM) to a lysine's ϵ -amino group. Most HMTases methylate lysines within the N-terminal tails and are highly specific for their substrate. Strikingly, all of the HMTases that methylate N-terminal lysines contain a so-called SET domain that harbours the enzymatic activity (Dillon et al., 2005) with the only exception of the Dot1 enzyme that methylates H3K79 within the histone globular core and does not contain a SET domain (Feng et al., 2002). Lysine demethylases, which reverse the action of HMTases, also possess a high level of substrate specificity with respect to their target lysine. They are also sensitive to the degree of methylation, and often only demethylate one specific state (e.g. dimethylation), whereas other enzymes are capable of demethylating all states. Remarkably, lysine methylation is recognized by more distinct domain types than any other modification, which could be attributed to the modification's diversity and importance. Domains that promote protein binding to methylated lysines in the tail region of histone H3 include PHD fingers or the so-called Tudor 'royal' family of domains (comprising Tudor, PWWP, chromodomains, and MBT domains) (Lu and Wang, 2013; Kim et al., 2006; Champagne and Kutateladze, 2009; Maurer-Stroh et al., 2003). Proteins associated with the remodeling and manipulation of chromatin often contain chromodomains and famous examples include transcriptional repressors HP1, recruited to heterochromatin by recognition of tri-methylated H3K9 by its chromodomain (Lachner et al., 2001; Bannister et al., 2001). Whereas histone acetylation is mostly associated with open chromatin, histone methylation can recruit both transcriptional activators and repressors.

1.2.3 Properties of histone modifications

Histone modifications can change the overall charge of histone proteins, thus modulating the interaction with DNA.

Acetylation and phosphorylation of residues introduce negative charges to histones, which can disrupt electrostatic interactions between histones and DNA. This presumably destabilizes histones and opens chromatin structure, thereby facilitating DNA access to the transcriptional machinery or other chromatin binding complexes (Bannister and Kouzarides, 2011).

Modification can act indirectly by regulating the recruitment of specialized effector enzymes which change chromatin structure.

The effector proteins contain specialized domains that recognize and bind a distinct modification state of histone N-terminal tail residues, and bring upon a certain downstream effect. In general, activating modifications attract chromatin remodelers and basal transcription factors to promote transcription. For example, the recognition of acetylated histone tails by specialized bromodomains is often the prerequisite for association of chromatin remodeling complexes to nucleosomes, which can then open

chromatin structure (Winston and Allis, 1999). This includes the SWI/SNF-remodeller ATPases that contain a bromodomain that binds acetylated histone tails (Hassan et al., 2002). Repressive modifications generally deter basal transcription factors and attract proteins that promote chromatin packaging and compaction. This is the case for the transcriptional repressor HP1, which is recruited to heterochromatin by recognition of tri-methylated H3K9 by its chromodomain (Bannister et al., 2001; Lachner et al., 2001). Several chromatin-associated factors interact with modified histone tails via multiple domains, whereas others have specific domains within them that allow the simultaneous recognition of several modifications and other nucleosomal features (Lawrence et al., 2016).

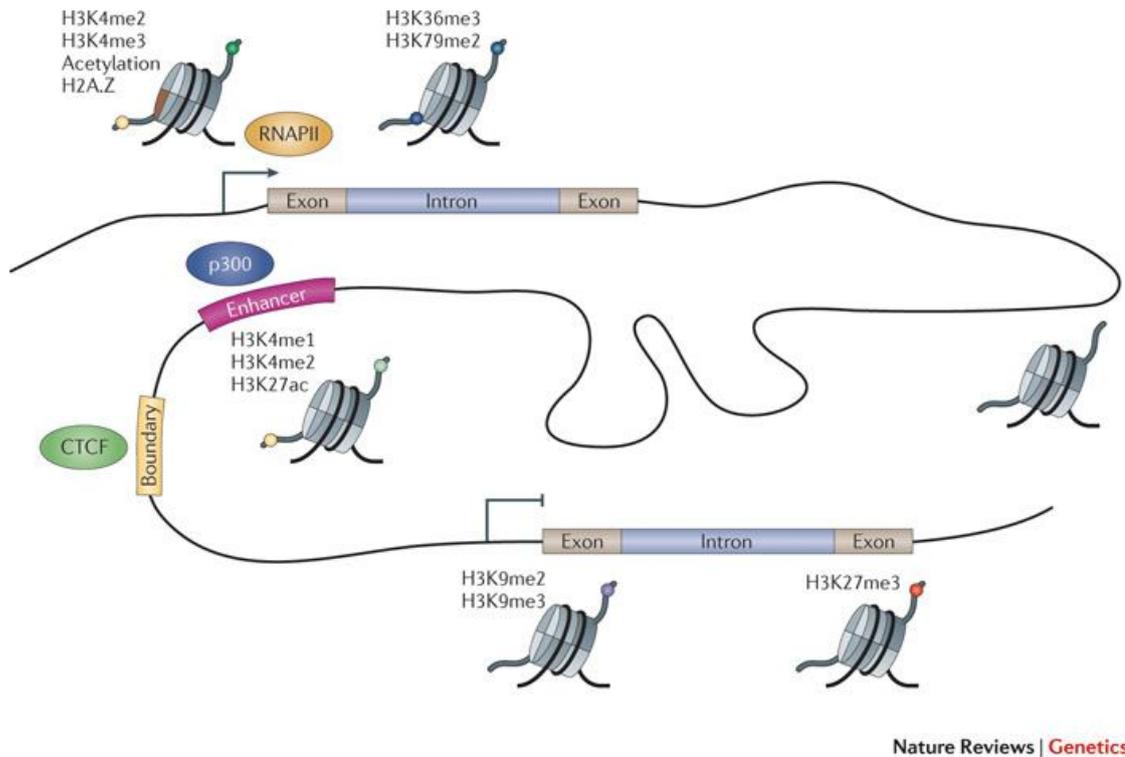
Histone modifications can confer epigenetic memory.

Histone modifications are stable marks that can potentially be inherited and maintained through mitosis (Probst et al., 2009; Almouzni and Cedar, 2016). During replication, DNA molecules are decompacted and pre-established chromatin structure is lost. At replication forks, polymerases require direct access to the DNA template, which requires the transient disruption of histone-DNA interaction of nearby nucleosomes. Parental histones can be recycled and segregated onto nascent DNA behind the replication fork, and gaps are filled by replication-dependent de novo deposition of new histones (Groth et al., 2007). This opens a window of opportunity during S-phase for cells to switch to another lineage as required during differentiation or development, or to stay committed to the same lineage as the mother cell by maintaining chromatin states. For the maintenance of histone modifications, parental histones could be used as a template that promote the modification of new histones in the daughter cells. If parental and new histones are distributed randomly on nascent DNA, the copying of modification from neighboring parental histones to new histones can help in re-establishing chromatin states that maintain cell identity (Probst et al., 2009). Conversely, unequal distribution of parental histones carrying modifications might allow developmental decisions and lineage specifications. These mechanisms are effective in regions where many nucleosomes carry the same histone modification, particularly heterochromatic regions that are marked by repressive H3K9me3 that is recognized by HP1 protein (Bannister et al., 2001; Lachner et al., 2001). However, particular marks are restricted to few nucleosomes, which is often observed for histone marks at regulatory regions. Such modifications might be more difficult to transmit through mitosis, or they require different mechanisms for inheritance, such as splitting of histone tetramers between dividing cells (Xu et al., 2010).

Histone modifications define chromatin regions.

Eukaryotic genomes are roughly divided into two distinct chromatin environments: A relatively relaxed and accessible compartment, containing most regulatory regions and active genes, referred to as euchromatin; and a relatively compact compartment,

containing transcriptionally repressed regions such as centromeres and telomeres, and many of the inactive genes, referred to as heterochromatin (Jenuwein, 2001). However, it should be noted that this view is simplified, and more genomic domains have been defined to describe chromatin architecture *in vivo* (de Wit et al., 2008; Filion et al., 2010). Nevertheless, these two main compartments are enriched with certain histone modifications, whereas they seem relatively devoid of other modifications (Fig. 1.3). Facultative or constitutive heterochromatin is marked by specific histone modifications, such as H3K9me2, H3K9me3, and H3K27me3 (Kharchenko et al., 2011), and it can form attachments between chromosomes and the nuclear envelope (Hochstrasser et al., 1986), which is a repressive environment (Pickersgill et al., 2006). Euchromatin is associated with other modifications than heterochromatin, and regions inside euchromatin that regulate transcription or sites of active transcription are rich in multiple modifications. For instance, active enhancers contain relatively high levels of H3K4me1 and the transcriptional start sites (TSS) of active genes possess a high enrichment of H3K4me3 (Heintzman et al., 2007; Barski et al., 2007; Schneider et al., 2004). Similarly, histone acetylation is associated with active regulatory regions and H3K27ac is detected at both active enhancers and TSS (Wang et al., 2008; Rada-Iglesias et al., 2011; Bonn et al., 2012; Creyghton et al., 2010). In addition, H3K36me3 is highly enriched throughout the entire transcribed region (Bannister et al., 2005). On-going research is trying to unravel how some enzymes are recruited to specific locations and how different modifications are integrated in order to regulate DNA processes such as transcription or gene repression.



Nature Reviews | Genetics

Figure 1.3: Schematic genomic region depicting the distribution of histone modifications at promoters, enhancers, gene bodies and boundary elements. Euchromatin is enriched for histone modifications associated with active regulatory elements (promoters and enhancers) and actively transcribed regions. Active promoters are commonly marked by H3K4 methylation (H3K4me2/me3), acetylation (ac), and the histone variant H2A.Z. Enhancers are relatively enriched for H3K4me1, H3K4me2, H3K27ac and the histone acetyltransferase p300. Transcribed regions are enriched for H3K36me3 and H3K79me2. CTCF is commonly found at boundary elements that act as insulators between different functional elements. Repressed genes inside heterochromatin are often located inside large domains marked by H3K9me2 and/or H3K9me3 or H3K27me3. Image is taken from Zhou et al. (2011).

1.3 Histone variants

Like histone modifications, the use of histone variants contributes to the regulatory repertoire of chromatin. Differences in the amino acid sequence and most importantly different expression and deposition patterns throughout the cell cycle allow for their specialized features. Characterization of histone variants is still ongoing, and the functions of some specialized histone variants have been well characterized. Famous examples include CENP-A, a histone H3 variant that substitutes for histone H3 specifically in nucleosomes of centromeric chromatin and that has unique properties essential for centromere function, kinetochore assembly and chromosome segregation during mitosis (Dunleavy et al., 2005). Or H2A.Z, a H2A variant which is incorporated into and destabilizes nucleosomes at promoter region, thereby promoting gene activation (Zlatanova and Thakar, 2008; Bonisch and Hake, 2012; Zhang et al., 2005; Jin et al., 2009). In the following, I would like to focus on the histone variant H3.3 and its canonical counterparts H3.1/H3.2, because it is relevant for the content of this thesis.

1.3.1 Histone H3 variants

In animals, two main classes of histone H3 genes encode distinct H3 proteins: the ‘canonical’, replication-dependent histone H3 (H3.1 and H3.2) and the replication-independent histone variant H3.3 (Elsaesser et al., 2010). These variants, expressed at distinct times during the cell cycle, do not impart any considerable structural alteration to the nucleosomes, but contribute to the chromatin assembly and the genomic localization of nucleosomes.

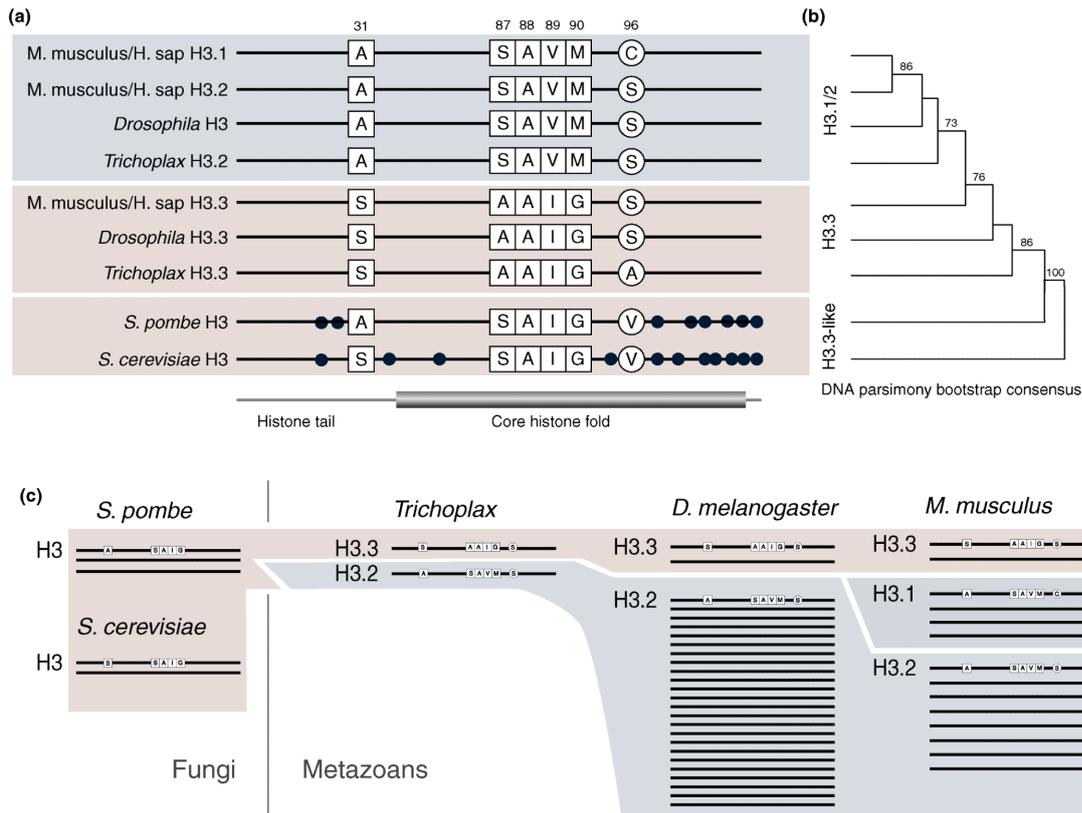


Figure 1.4: Schematic overview of protein sequences and genes of histone H3 variants in animals and yeast. (a) Alignment of the histone H3 protein sequences in mammals, *Drosophila*, *Trichoplax* and two yeast species. The amino acids that distinguish the H3 variants are indicated, and dots indicate additional differences in yeast. H3.3 varies from H3.1/H3.2 at serine 31 in the histone tail and in three amino acids 87-90 in the histone core. Histone H3.1 is specific to mammals and varies from H3.2 in one amino acid at position 96. Yeast species only have one H3 protein. (b) Phylogenetic analysis of the H3 variants suggest that yeast H3 is an H3.3-like protein and displays a higher similarity in its coding sequence to the H3.3 variant than to H3.1/2 in animals. (c) Overview of the genes encoding H3 proteins in the indicated species. The number of H3 genes have expanded greatly from yeast and basal animals (*Trichoplax*) to higher-order species such as *Drosophila* and mammals. Image and analysis was obtained from Elsaesser et al. (2010).

Replication-dependent canonical H3

The majority of histones are synthesized during S phase for rapid deposition behind replication forks in order to package replicating DNA. These histones are generally referred to as canonical or replication-dependent (RD) histones. In higher eukaryotes,

canonical histone H3 genes, together with those of H2A, H2B, and H4, occur as multiple copies inside histone clusters. Increasing copy numbers of canonical histone genes in evolution indicate that histone clusters have expanded to supply sufficient nucleosomes to package increasingly larger genomes. The genes of canonical histone are short, intronless genes that lack a poly(A) tail, but instead have a conserved 3' stem loop (Elsaesser et al., 2010). These common features are likely to be responsible for a regulated expression of canonical histones during S-phase in a replication-dependent manner (Harris et al., 1991). In mammals one can distinguish between two canonical H3 variants: H3.1 and H3.2, which only differ by a single amino acid (Fig. 1.4a). H3.2 is shared by all animals, whereas H3.1 variant additionally occurs exclusively in mammals, thus arguing for a later evolutionary separation (Postberg et al., 2010).

Replication-independent histone variant H3.3

The third H3 protein is histone variant H3.3. Two distinct genes encode mammalian H3.3 (H3.3A and H3.3B) that are located outside of histone cluster, contain introns, and give rise to classical polyadenylated mRNAs. Unlike canonical H3, the expression of H3.3 genes is cell-cycle-independent and H3.3 serves as a substrate for both replication-dependent deposition and histone replacement processes that occur outside of S phase (Elsaesser et al., 2010). Overall, the amino acid sequence of H3 is mostly conserved, with only few amino acids distinguishing H3.3 from its canonical counterparts (5 amino acid positions are different to H3.1) (Fig. 1.4a). The first difference is amino acid variation in residues 87–90 of the histone core region ("AAIG" vs. "SAVM"), which is necessary for assembly of H3.3 into nucleosomes by histone chaperones such as Hira (discussed in section 1.4). The second difference is a substitution in the N-terminal tail at residue 31 ("A" vs "S"). This change is not required for selective deposition, but may be an important site for mitosis-specific phosphorylation (Hake et al., 2005). In yeast species *S. cerevisiae* and *S. pombe*, all H3 genes, apart from centromeric CENP-A, encode for only one H3 variant that shares higher sequence similarity with H3.3 than with canonical H3 (Fig. 1.4b). Therefore, phylogenetic analysis has suggested that H3.3 and yeast H3 share a common ancestor (Fig. 1.4b), which gave rise to all H3 variants in all opisthokonts (animals, fungi, and eukaryotic microorganisms) (Postberg et al., 2010; Elsaesser et al., 2010). With increasing genome complexity during evolution, ancestral H3 genes are thought to have expanded and formed the canonical histone clusters to supply sufficient nucleosomes during replication for packaging of increasingly larger genomes (Fig. 1.4c).

1.3.2 H3 in proliferating and quiescent cells

In proliferating cells, the canonical H3.1 and H3.2 represent the dominant isoforms and H3.3 contributes approximately one fourth of the total pool of histone H3 (McKittrick et al., 2004; Maze et al., 2015). This amount of H3.3 would still be sufficient to package essentially all actively transcribed genes (McKittrick et al., 2004) and based

on the observations that specifically H3.3 contains modifications associated with transcriptionally active chromatin, it has been proposed that H3.3 “barcodes” regions of active chromatin (Hake and Allis, 2006) in proliferating cells. Chromatin immunoprecipitation coupled with DNA sequencing (ChIP-Seq) have allowed a more detailed map of H3.3 deposition, revealing specific H3.3 incorporation throughout the gene body of transcribed genes as well as into promoter regions in dividing *Drosophila* and mammalian cells (Wirbelauer et al., 2005; Mito et al., 2005, 2007; Goldberg et al., 2010). H3.3 enrichment has also been observed at promoters of inactive genes, which could indicate that these promoters are in a poised state (Mito et al., 2007; Tamura et al., 2009). Furthermore, H3.3 enrichment also occurs at genic and intergenic regulatory regions in animals, such as enhancers (Goldberg et al., 2010; Mito et al., 2007). Therefore, H3.3 enriched foci occur at active chromatin regions in dividing cells.

When cells exit the cell cycle and differentiate, they no longer produce or incorporate replication-dependent histones, but require a continuous supply of histones to compensate for nucleosome displacement during transcription and other DNA-related processes. Especially neurons, which are formed early during embryonic development and remain in a non-proliferating state for years, accumulate H3.3-containing nucleosomes (Maze et al., 2015). Indeed, H3.3 constitutes around 90% of all H3 protein in brains of adult mice and rats (Pina and Suau, 1987). Consequently, replication-independent H3.3 is deposited in all chromatin types and becomes the predominant H3 variant in non-proliferating cells (quiescent, G1, and G2 cells) (Maze et al., 2015). Even though a detailed map of H3.3 enriched loci in aged brain by ChIP-Seq is not available yet, it may resemble the broad and even distribution of canonical H3 with no or few enriched foci.

1.3.3 Contributions of H3.3 to development and viability of multicellular organisms

Loss-of-function and replacement studies for genes encoding H3 variants or its chaperones have proven insightful to identify their contribution to animal viability and development in *Drosophila* and mammals.

Drosophila melanogaster

Individual homozygous disruption of either one of the H3.3 genes in *Drosophila* (H3.3A and H3.3B) has little phenotypic effect on the overall organism. In contrast, combined disruption of both genes results sterility in both sexes and slightly reduced viability, but with no obvious morphological defects in adult flies (Hodl and Basler, 2009; Sakai et al., 2009). Viability, but not sterility, can be rescued by replacement of H3.3 with canonical H3, now expressed outside of S-phase. Vice versa, replacement of canonical H3 with H3.3 does not result in any impairment. Similarly, the histone chaperone Hira, which deposits H3.3 into chromatin (discussed in more detail in section 1.4),

is also essential for fertility. Specifically, during fertilization it helps in decondensing paternal chromatin by assembling H3.3-containing nucleosomes in the male pronucleus (Bonnefoy et al., 2007; Loppin et al., 2005). In *Drosophila*, both Hira and H3.3 are required for fertility and for transcriptional regulation of specific genes, but not for viability in development and adult animals (Bonnefoy et al., 2007; Hodl and Basler, 2009; Nakayama et al., 2007; Sakai et al., 2009). Furthermore, histone H3 variants can replace each other in both replication-dependent and replication-independent chromatin assembly (Sakai et al., 2009).

Mammals

Loss-of-function studies of H3.3 have disrupted one of the two H3.3-encoding genes (*H3f3a* or *H3f3b*) successfully in mice (Bush et al., 2013; Couldrey et al., 1999; Tang et al., 2013; Jang et al., 2015). One study by Jang et al. (2015) found that regardless which H3.3 gene was removed, mice were normal and fertile in both sexes, suggesting that both H3.3 genes are not required for completion of a mouse life cycle. In other studies, disruption of either *H3f3a* or *H3f3b* results in developmental defects, partial neonatal lethality, and reduced fertility (Tang et al., 2013, 2015). The discrepancy between studies in regard to viability and fertility in single-gene knockouts is likely due to genetic background differences. A mixed background (mainly C57BL/6 and 129) and resulting genetic heterogeneity may have contributed to the better viability and fertility of knockouts as a result of hybrid vigor (Jang et al., 2015). In spite of this, all studies conclude a partial redundancy of H3.3 genes and that only one gene is absolutely required for mouse development, viability and fertility. In contrast, complete removal of both H3.3 genes results in severe developmental retardation and early embryonic lethality (Roberts et al., 2002). Similarly, knockdown of both H3.3 genes leads to early embryonic developmental arrest, chromosome missegregation, and chromatin condensation (Lin et al., 2013). At the cellular level, H3.3 loss triggers cell cycle suppression resulting in reduced cell proliferation and increased cell death (Jang et al., 2015). In mouse embryonic stem cells (ESCs) H3.3 has been shown to be required for proper establishment of repressive histone marks at the promoters of developmentally regulated genes, thereby contributing to their silencing *in vitro* (Banaszynski et al., 2013). However, embryonic patterning and marker gene expression patterns remain intact in double-knockout mutant embryos (that rarely survive post-implantation), suggesting that H3.3 is not absolutely required for transcriptional regulation of developmental genes *in vivo*. In line with this, loss of Hira in mice causes defects in early embryogenesis (Roberts et al., 2002). Hira-deficient mouse ESCs or human cells show a dramatic reduction in H3.3 incorporation into genes and regulatory elements, but display only minor transcriptional defects (Goldberg et al., 2010; Ray-Gallet et al., 2002). In mammals, H3.3 and its chaperone Hira are absolutely required for embryonic development and animal's survival. Thus, the presence of different H3 species seems more important in mammals than in *Drosophila* and different chromatin assembly pathways

can only partly compensate for each other. However, more experimental evidence is needed to confidently assess the contributions of each H3 variant.

1.4 Chromatin assembly and remodeling

The bulk of newly synthesized histones are incorporated during DNA replication, a process that requires both the making and breaking of nucleosomes. But also outside of S-phase nucleosomes are continuously displaced to create sites of accessible DNA required for recruitment of transcription factors or to allow passage of transcribing RNA polymerases (Venkatesh and Workman, 2015). In eukaryotic cells, the assembly and disassembly of nucleosomes accompanying genomic processes are complex operations involving many steps and proteins of various functions. Histone chaperones play a crucial role in these processes, but their function is intimately coupled to the making and breaking of histone-DNA interactions by chromatin remodeling complexes. In the following, I want to give a brief summary on how nucleosome turnover, assisted by histone modifications and histone variants, can establish or alter the regional properties of chromatin, restructure higher-order chromatin and provide accessible DNA templates for cellular machinery.

1.4.1 Chromatin assembly by histone chaperones

Histone chaperones are a diverse family of histone-binding proteins that prevent non-nucleosomal interactions between histones and DNA. They are involved at multiple steps after histone synthesis to sequester core histones from DNA until they can be properly assembled into nucleosome, which is the energetically most favorable conformation (Ransom et al., 2010). Histone chaperones must also ensure timely reassembly of chromatin following unwinding and separation of DNA strands during replication and transcription (Winkler and Luger, 2011; Gurard-Levin et al., 2014). Eukaryotes have evolved specific histone chaperones for one or more histone variants and possibly even specific chaperones for the recycling of old histones versus incorporation of newly synthesized histones. Deposition of the canonical histone H3 is coupled to DNA synthesis during replication and possibly DNA repair, whereas histone variant H3.3 serves as the replacement variant for the replication-independent deposition pathway (Fig. 1.5). Several histone chaperone complexes involved in these distinct chromatin assembly pathways have been characterized and they are largely conserved from yeast to mammals. In yeast, the precise role of many chaperones is partially obscured by the functional redundancy between them, whereas in other organisms many of the chaperones are essential, highlighting their critical role in biology.

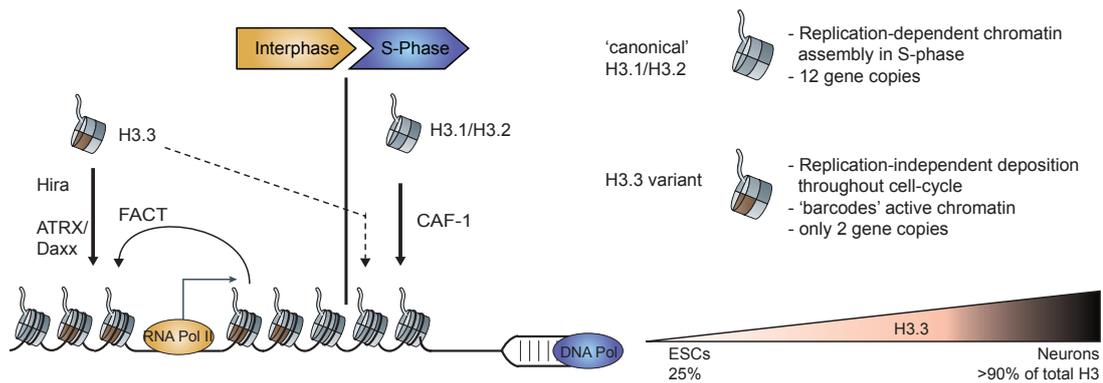


Figure 1.5: Assembly pathways of histone H3 variants throughout the cell cycle. Canonical H3.1/H3.2 are mainly expressed and assembled into nucleosomes in a replication-dependent manner during S-phase by the histone chaperone CAF-1. Variant H3.3 can additionally be deposited onto chromatin outside of S-phase by the chaperone complexes Hira and ATRX/Daxx. Hira mediated H3.3 deposition into genic regions, whereas ATRX/Daxx predominantly assembly H3.3 into nucleosomes at repetitive heterochromatic sites. The replication-independent assembly of H3.3 results in accumulation of H3.3 in nucleosomes of post-mitotic cells such as neurons compared to highly proliferative ESCs. The chaperone FACT destabilizes and reassembles nucleosomes during transcription by Pol II.

1.4.1.1 Replication-dependent assembly of canonical H3

The bulk of newly synthesized canonical histones are incorporated during DNA replication by CAF-1 (chromatin assembly factor 1) (Smith and Stillman, 1989; Burgess and Zhang, 2013). CAF-1 acts as a histone chaperone that facilitates the incorporation of two dimers, or one tetramer, of H3-H4 onto replicating DNA as a first step in nucleosome assembly (Tagami et al., 2004). Histones H2A-H2B can bind to this chromatin precursor subsequent to DNA replication to complete the nucleosome (Krude, 1995; Gurard-Levin et al., 2014). Similarly, nucleosome disassembly, the reversal of the assembly process, is initiated with partial DNA unwrapping and loss of H2A-H2B dimers, followed by the loss of the H3-H4 tetramer from the DNA. In agreement with this model, H3-H4 dimers have been shown to be much more stably associated with DNA than H2A-H2B dimers (Kireeva et al., 2002; Kulaeva et al., 2010). But at the same time, the exchange of any nucleosomal histone contained within the H3-H4 tetramer for a new histone would require disassembly of the entire nucleosome.

1.4.1.2 Replication-independent assembly of H3.3

Histone variant H3.3 is expressed and can be deposited either during DNA replication or in a replication-independent (RI) manner throughout the cell cycle (Ahmad and Henikoff, 2002; Wu et al., 1982). The chaperone systems responsible for replication-independent deposition of newly synthesized H3.3 are the Hira complex and Atrx/Daxx complex. Both complexes recognize the H3.3 variant region "AAIG" in the histone core (Ricketts et al., 2015; Ahmad and Henikoff, 2002), but mediate deposition into different genomic regions.

The Hira complex

The Hira complex deposits H3.3 (Tagami et al., 2004) into euchromatic regions such as promoters, gene bodies and cis-regulatory elements (enhancers), and into transient nucleosome free regions (Goldberg et al., 2010; Ray-Gallet et al., 2011). The loss of Hira results in a dramatic reduction of H3.3 incorporation at promoters and gene bodies, but not at many intergenic transcription factor binding sites (Goldberg et al., 2010). This suggests that H3.3 can be deposited by other (unknown) chaperones or chaperone-independent pathways specifically into nucleosomes at regulatory regions. Hira-mediated deposition throughout the cell cycle is specific to H3.3 and cannot be extended to canonical H3 (Ray-Gallet et al., 2011). On the other hand, in the absence of replication-dependent H3 deposition by CAF-1, Hira can also deposit H3.3 at sites of DNA replication in human cells (Ray-Gallet et al., 2011).

The ATRX/Daxx complex

The ATRX/Daxx complex mediates H3.3 deposition into repetitive heterochromatin, such as retrotransposons, pericentromeric and telomeric regions (Lewis et al., 2010; Delbarre et al., 2013; Drane et al., 2010; Goldberg et al., 2010; Iwase et al., 2011) and into specific genes in response to neuronal signaling (Michod et al., 2012). H3.3 deposition by the ATRX/Daxx chaperone is believed to contribute to epigenetic silencing of a few, but not all repetitive regions, as has been shown for retrotransposons (Elsasser et al., 2015). Loss of the ATRX protein results in aneuploidy and defects in chromosomal segregation (Baumann et al., 2010). Additionally, both ATRX and Daxx mutations are often found in pediatric gliomas and neuroblastomas (Schwartzentruber et al., 2012; Cheung et al., 2012) and in other cancer types (Heaphy et al., 2011; Jiao et al., 2011), suggesting that these mutations drive disease in a particular developmental context (Dyer 2017). Thus, the interplay between ATRX/Daxx and H3.3 histones at telomeres and other repetitive sequences could contribute to silencing of these elements, which may be critical to maintain genome integrity and prevent formation of cancer.

FACT complex

FACT (facilitates chromatin transcription), is a histone chaperone that can deposit histones onto DNA (Belotserkovskaya et al., 2003) and is also involved in nucleosome reorganization during transcription, replication and DNA repair (Winkler and Luger, 2011). During transcriptional elongation, FACT destabilizes the nucleosome by displacing the H2A/H2B dimers to allow passage of the transcribing RNA polymerase through a nucleosomal DNA template (Belotserkovskaya et al., 2003). It also facilitates the subsequent recovery of the nucleosome after transcription by holding the H3/H4 tetramer on DNA and promoting the deposition of the H2A/H2B dimer to reform the nucleosome (Hsieh et al., 2013; Chen et al., 2018). Therefore, FACT has a dual function in destabilizing the nucleosome to facilitate transcription and in maintaining the original nucleosome composition. In this way, FACT could play a role in preserving

the original histones with pre-established histone modifications during transcription or DNA replication (Chen et al., 2018).

1.4.2 Chromatin remodeling

Chromatin remodelers are large, multiprotein complexes that can mobilize nucleosomes to provide regulated access to DNA sequences. They play important roles in many chromatin functions: to space nucleosomes properly during chromatin assembly, to enable the controlled access of transcription factors to specific genes, to regulate the assembly of RNA-polymerases at promoters and allow the access of DNA-repair factors to sites of DNA lesions (Saha et al., 2006). Furthermore, remodelers may help the replication-independent nucleosome assembly machinery to fill in gaps where nucleosomes have been ejected (Lewis et al., 2010; Konev et al., 2007). Often, chromatin remodelers are driven energetically by ATP hydrolysis inside their ATPase domain and they are subgrouped into the four largest families: SWI/SNF, ISWI, CHD (NuRD), and INO80 (Clapier and Cairns, 2009). Members of these families function in different biological processes and prefer different nucleosomal substrates. Most remodelers can move nucleosomes to occupy alternative positions by sliding the histone octamer along the DNA. This activity depends on ATP-dependent DNA translocation by the ATPase domain (Saha et al., 2002, 2005; Zofall et al., 2006; Clapier and Cairns, 2009). The sliding activity can create an evenly spaced arrangement of nucleosomes on DNA, or distribute nucleosomes unevenly, potentially to open up binding sites on chromatin. In contrast to these sliders, other remodelers have evolved a mechanism for ejection of nucleosomes or the exchange of histone dimers (Saha et al., 2006). Histone exchange mediated by chromatin remodeling complexes can contribute to the deposition of histone variants: A member of the INO80 family deposits histone variant H2A.Z by ejecting a H2A/H2B dimer from the nucleosome and inserting a less stable H2A.Z/H2B dimer in its place (Luk et al., 2010; Mizuguchi et al., 2004; Krogan et al., 2003), which can promote gene activation (Bonisch and Hake, 2012; Zhang et al., 2005; Jin et al., 2009). Furthermore, ATRX, which deposits histone variant H3.3 into heterochromatic regions as part of the Daxx/Atrx complex, also contains an ATPase domain and belongs to the SWI/SNF family of chromatin remodeling proteins (Lewis et al., 2010).

1.4.2.1 Recognition of histone substrates by chromatin remodeling complexes

The binding of chromatin remodelers to histones and nucleosomal DNA can be regulated by the histone modification state. The modification state helps determine whether the nucleosome is an appropriate substrate for a remodeler complex (Saha et al., 2006). This recognition is often mediated by protein domains that chromatin remodelers share with histone modifying enzymes, such as bromodomains and chromodomains. Bromodomains allow remodeling complexes to selectively interact with

acetylated histone tails (Winston and Allis, 1999; Dhalluin et al., 1999; Jacobson et al., 2000) as seen for SWI/SNF-remodeling complexes (Hassan et al., 2002; Awad and Hassan, 2008). Some remodelers contain bromodomains that selectively binds one particular acetylated lysine residue in the histone tail, as seen for the yeast RSC complex that binds acetylated H3K14 (Kasten et al., 2004), which might help in targeting remodelers to particular genes. A subset of chromodomains might carry out a corresponding function to mediate the recognition of methylated lysine residues (Fischle et al., 2003; Min et al., 2003). For example, the mammalian remodeler Chd1 localizes to actively transcribed genes through recognition of H3K4 methylation by its double-chromodomain (Flanagan et al., 2005; Sims et al., 2005), a hallmark of active chromatin (Schneider et al., 2004). In yeast and *Drosophila*, the chromodomain of Chd1 does not selectively bind H3K4-methylated histone tails and is not required for the recruitment of Chd1 to chromatin (Morettini et al., 2011; Sims et al., 2005), but might stimulate the remodeling activity (Morettini et al., 2011; Hauk et al., 2010; Radman-Livaja et al., 2012). Thereby, Chd1 can accelerate the H3 exchange close to promoters, and reduce exchange towards transcriptional end sites (Radman-Livaja et al., 2012). This example shows that deposition of histone modifications and nucleosome exchange are tightly connected processes.

1.4.3 Histone turnover

The process of swapping histones incorporated into chromatin for free histones is called histone exchange and the rate at which this exchange occurs is referred to as histone turnover (Venkatesh and Workman, 2015). Accordingly, regions with high turnover have a constant influx of newly synthesized histones, whereas low turnover regions contain longer lived nucleosomes, as seen inside compacted heterochromatin (Aygun et al., 2013) or at sites where old histones are actively recycled. The availability of histone variants outside of S-phase results in the substitution of canonical histones in nucleosomes with variant histones during exchange processes (Wirbelauer et al., 2005). Specifically, Hira promotes the assembly of H3.3–H4 to substitute for displaced nucleosomes (Tagami et al., 2004; Ahmad and Henikoff, 2002; Ray-Gallet et al., 2011), which results in the accumulation of H3.3 at sites of high histone turnover. The rate at which nucleosomal H3.3 incorporation (or displacement) occurs, has been established as a measure for histone turnover (Deaton et al., 2016; Kraushaar et al., 2013).

1.4.3.1 Site-specific histone turnover

Sequence-specific transcription factors have to compete with nucleosomes for binding sites inside regulatory regions of the genome (Henikoff, 2008). Consequently, the binding of transcription factors, with assistance from co-activators like chromatin remodelers, may lead to nucleosome displacement. In support of this, promoter regions often show low nucleosome densities (Mito et al., 2005; Sekinger et al., 2005) and regula-

tory regions are bound by multiple chromatin remodelers (de Dieuleveult et al., 2016). Furthermore, promoter and enhancer regions are enriched with H3.3-containing nucleosomes, which have been shown to undergo rapid turnover (Kraushaar et al., 2013; Deaton et al., 2016; Huang and Zhu, 2014). Inside coding regions, histone exchange occurs at lower turnover rates is coupled to transcription by Pol II (An overview is depicted in Fig. 1.6b).

Promoter regions

Promoters display high turnover rates compared to other genomic regions and the highest turnover generally occurs at highly transcribed genes (Kraushaar et al., 2013; Dion et al., 2007; Deaton et al., 2016). This implies that transcription initiation by Pol II could lead to substantial H3.3 turnover. However, H3.3 turnover at promoter regions does not linearly correlate with transcriptional activity (Kraushaar et al., 2013; Huang and Zhu, 2014). Interestingly, histone exchange also occurs at inactive promoters (Mito et al., 2007; Kraushaar et al., 2013; Dion et al., 2007; Rufiange et al., 2007), although with a lower turnover rate, possibly to keep promoters in a poised state and to facilitate their activation upon induction. Thus, promoters are either active and display high H3.3 turnover or are inactive and display low H3.3 turnover. Similar observations have been made for replication-independent H3 turnover in *S. cerevisiae* (Dion et al., 2007). Therefore, the engagement of RNA Pol II and likely the binding of sequence-specific transcription factors, but not necessarily the transcription frequency of Pol II, correlate with H3.3 turnover at promoters.

Enhancer regions

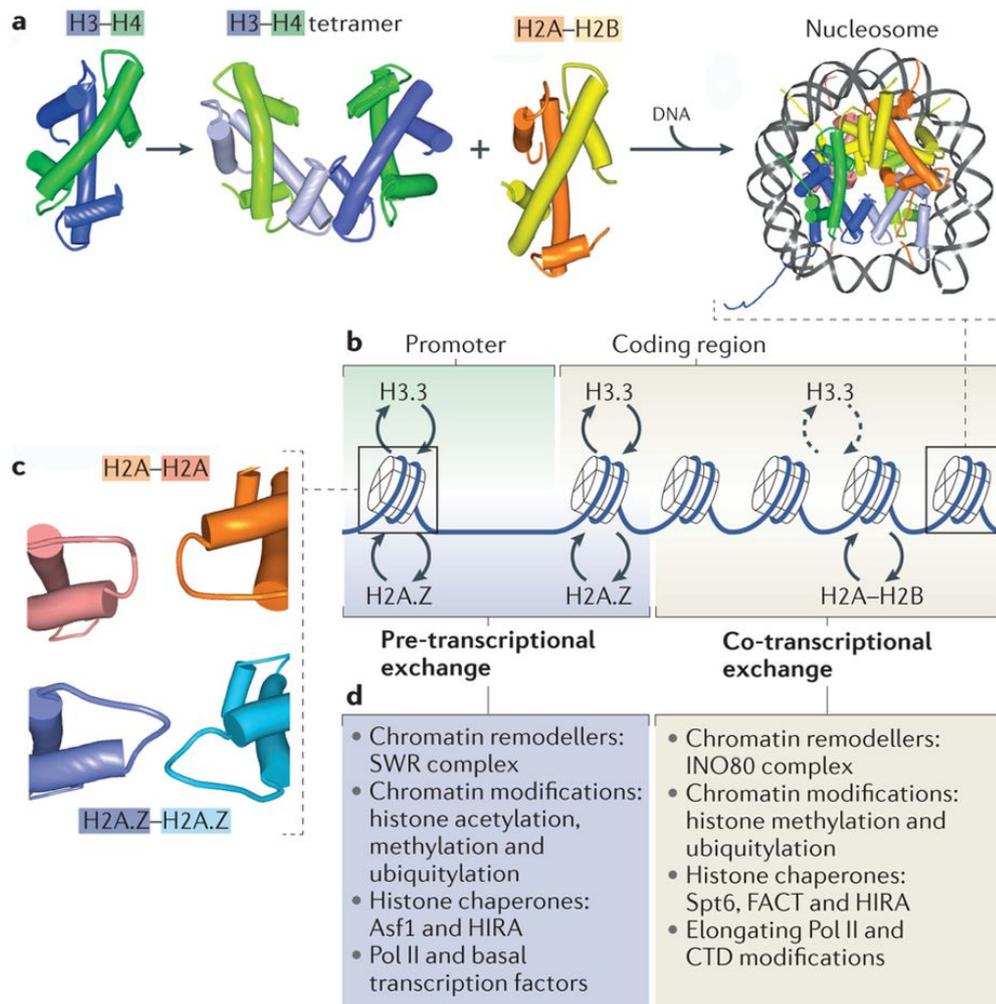
Enhancers also display relatively high H3.3 turnover (Kraushaar et al., 2013; Deaton et al., 2016). Given that enhancers are enriched for sequence-specific transcription factor binding sites and that enhancers are often occupied by Pol II (Kim et al., 2010), high H3.3 turnover at enhancers may be the result of very similar mechanisms that are in action at promoter regions. The highest turnover rates have been observed at regulatory regions occupied by multiple transcription factors, suggesting that transcription factor binding events indeed contribute to turnover rates (Deaton et al., 2016). Furthermore, cis-regulatory elements that contribute to transcriptional repression instead of activation, such as Polycomb response elements, are enriched with H3.3 and display high turnover in *Drosophila* (Mito et al., 2007). This suggests that higher accessibility of regulatory regions does not necessarily result in gene activation, and H3.3 turnover may have multiple functions in distinct genomic locations that are not always correlated with activity of chromatin (Banaszynski et al., 2013).

Gene bodies

Given the physical disruption of nucleosomes that must accompany RNA polymerase passage, it follows that histone exchange occurs frequently inside gene bodies of

highly transcribed genes (Schwartz and Ahmad, 2005; Schwabish and Struhl, 2004; Wirbelauer et al., 2005). Indeed, H3.3 is enriched in all transcribed genes (Wirbelauer et al., 2005) and H3.3 exchange at gene bodies closely correlates with transcription rates (Deal et al., 2010), which suggests that the passage of Pol II and its associated factors may regulate histone turnover in gene bodies directly (Venkatesh and Workman, 2015). However, compared to promoter and enhancer regions, gene bodies show relatively low levels of nucleosome turnover and H3.3 exchange rates (Dion et al., 2007; Kraushaar et al., 2013; Rufiange et al., 2007; Deal et al., 2010). One explanation for the observed lower turnover is that during transcription, nucleosomes are temporarily disassembled ahead of, and subsequently assembled behind, the transcription machinery, preferentially by reassembly of parental histones and, to a lesser extent, by the incorporation of newly synthesized histones (Svensson et al., 2015; Hsieh et al., 2013). Histone recycling in gene bodies depends on the protein complex FACT (Belotserkovskaya et al., 2003; Chen et al., 2018; Jamai et al., 2009). Alternatively, engagement and initiation of Pol II may be a more dynamic process than Pol II elongation, and possibly the transcription initiation complex is assembled multiple times before progressing into the gene body, which would also result in comparably higher turnover rates at promoters.

In conclusion, the highest dynamic exchange occurs at promoter and enhancer regions, which display continuous binding and removal, respectively, of the transcription machinery and sequence-specific transcription factors. The continuous rapid exchange of H3.3 nucleosomes at promoters and enhancers could transiently expose transcription factor binding sites, thus keeping these sites accessible (Henikoff, 2008). The binding of transcription factors with chromatin in turn is also dynamic (Lickwar et al., 2012), allowing reassembly of nucleosomes after their dissociation. Collectively, these highly dynamic regions with fast histone turnover are most likely the result of the combined action of chromatin remodelers, the binding and dissociation of transcription factors and the engagement of the transcriptional machinery (Huang and Zhu, 2014). Histone turnover at gene bodies is slower, coupled to transcription by Pol II, and appears to be regulated by mechanisms that are distinct from the ones that function at cis-regulatory elements.



Nature Reviews | Molecular Cell Biology

Figure 1.6: Schematic representation of histone exchange at regulatory elements and inside coding regions. (a) During nucleosomal assembly, two H3-H4 heterodimers form a tetramer and are joined by two dimers of H2A-H2B to form a full nucleosomes around which DNA is wrapped. The replacement of H3 and H4 requires the preceding removal of the H2A-H2B dimers from the nucleosome, whereas H2A-H2B dimers can be exchanged without disrupting the H3-H4 tetramer. (b) Histone exchange processes at promoter and coding regions result in the replacement of canonical H3 with the H3.3 variant, resulting in the enrichment of H3.3 at these sites. At promoters and the +1 nucleosome, canonical H2A is additionally exchanged for the H2A.Z variant. The rates of histone exchange differ between the promoter and coding regions. Pre-transcriptional exchange at promoters occurs more rapidly, whereas co-transcriptional exchange occurs more slowly. (c) Specific histone variants such as H2A.Z can destabilize nucleosomes. Schematic representation depicts the loop (L1) region mediating the interaction between two H2A or two H2A.Z proteins. H2A.Z destabilizes heterotypic nucleosomes (H2A.Z-H2A) by weakening the interaction between these loop structures. (d) Factors that contribute to histone exchange over promoters and coding regions. Proteins that regulate the exchange include histone chaperones, chromatin remodeling complexes and the Pol II transcriptional machinery itself. Modifications of histones and of CTD (carboxy-terminal domain) of Pol II are also associated with histone exchange processes and these can vary between promoters and coding regions. Image was taken from Venkatesh and Workman (2015).

1.4.3.2 Histone modifications and turnover

Histone modifications are marking different regions of genes, and have been implicated to contribute to local turnover properties. H3.3-containing nucleosomes at enhancers and promoters display rapid turnover and these regions are associated with the histone modifications H3K4me1, H3K4me3, H3K9ac, H3K27ac and co-occupancy with the histone variant H2A.Z (Kraushaar et al., 2013). Furthermore, H3.3 itself is often marked by active histone modifications (McKittrick et al., 2004; Hake and Allis, 2006). In contrast, turnover is anti-correlated with H3K27me3 (Kraushaar et al., 2013), a repressive mark often marking poised promoters and silenced chromatin. H3K27me3 can recruit reader proteins to promote the formation of higher order chromatin structure (Li and Reinberg, 2011; Margueron et al., 2008), which may reduce histone turnover. Interestingly, in mammalian cells histone turnover is anti-correlated with H3K36me3 (Kraushaar et al., 2013), an active modification marking transcribed gene bodies. This negative correlation has also been observed in *S. cerevisiae*, where H3K36 methylation is thought to suppress histone exchange at active genes (Smolle et al., 2012). Gene bodies themselves have lower histone turnover than regulatory regions, suggesting that there may indeed be a functional link with H3K36me3. Similarly, nucleosomes inside heterochromatic regions have low turnover rates and are marked by H3K9me3 in yeast (Aygun et al., 2013). Thus, the association of distinct histone marks with regions of high and low turnover, respectively, suggests that there may be a causal relationship between histone modifications and the rate of nucleosome exchange. However, experimental evidence is controversial and other *in vitro* experiments in yeast suggest that turnover and modifications are largely independent events (Ferrari and Strubin, 2015). Especially in mammalian cells, it remains to be tested whether H3K36 methylation does indeed contribute to the comparably low histone turnover rates observed at gene bodies, and whether H3K4 methylation, H3K9 and H3K27 acetylation instead accelerate turnover at cis-regulatory elements.

1.5 Systems to probe the contribution of histone residues to nucleosome function

A particularly powerful approach for studying the biological function of specific histone residues is to change the residue to a different amino acid that cannot be appropriately modified or that mimics a certain modification state. In this way, one can address the importance of the residue itself and gain information about histone modifications at this site. Systematic histone H3 mutation studies in *Saccharomyces cerevisiae* and *Drosophila melanogaster* have delivered direct clues for the functional significance of histone residues in individual organisms (Dai et al., 2008; Hodl and Basler, 2012; Pengelly et al., 2013). However, similar studies in mammals are challenged due to the presence of large multi-loci gene clusters. In yeast, genome engineering based on

homologous recombination is highly efficient and has been successfully applied to histones to generate systematic histone H3 and H4 substitutions and deletion mutants to probe the contribution of each residue to nucleosome function (Dai et al., 2008). *S. cerevisiae* carries only two gene copies for histone H3 and Boeke and colleagues have removed the first of these gene copies and exchanged the second gene (Dai et al., 2008). In this way, they could systematically substitute each residue of H3 with alanine, and additionally substituted all modifiable residues with amino acids mimicking modified and unmodified states. Furthermore, they introduced deletion mutants that are lacking parts of or the whole N-terminal tail of H3. In *D. melanogaster*, the replication-dependent histone genes are found at a single locus (Lifton et al., 1978). Herzig and colleagues successfully combined the deletion of this endogenous histone gene cluster with its substitution by transgenes (Gunesdogan et al., 2010), thereby allowing the first characterization of histone residue and histone variant function in animal development (Pengelly et al., 2013; Hodl and Basler, 2009, 2012; McKay et al., 2015). This system has been further developed to simplify the exchange using BACs to deliver the transgenes (McKay 2015). These studies have revealed an increasing dependency on histone tail residues in multicellular organisms, potentially because they are required during animal development. Viable deletions in yeast, including deletions of the entire N-terminal tail or mutation of individual residues Dai et al. (2008), can have dramatic consequences for *Drosophila* development and are often lethal at a certain developmental stage (McKay et al., 2015; Meers et al., 2017). Thus, probing the importance of histone tail residues in higher organisms is essential to determine their function in humans. However, implementing systematic histone exchange tools in mammals is technically challenging because canonical histones are encoded inside gene clusters located at multiple chromosomal locations (Marzluff et al., 2002), e.g. the human genome contains more than 60 canonical histone genes that are located at three different loci (Marzluff et al., 2002).

These complications prevented systematic testing of replication-dependent canonical histones so far, and the complete exchange of all genes may remain impossible. The genes of replication-independent histone variants however, such as H3.3, exist in fewer copies outside the histone clusters (Elsaesser et al., 2010). Recent advances in genome engineering techniques, namely the CRISPR-Cas9-System, made targeted mutations inside these genes feasible, opening the possibility for systematic histone mutations in the mammalian system.

1.6 Aims & Motivation

Numerous mutations in histone H3 modifying enzymes have been reported in human disease, yet the role of the histone residues themselves remains unclear in mammals due to the clustered arrays of many histone genes. In this study, I aim to investigate the role of histone residues (e.g. K4 and K36) in the histone H3.3 variant, which is enriched at active genes and regulatory regions of genome. H3.3 is encoded by two isolated genes outside of histone clusters; therefore amenable to functional analysis. I set out to mutate the H3.3 encoding genes as a platform to analyze the role of histone residues in mammals. In this way, I want to address the long-standing questions of the functional requirement of histone residues and their modifications for transcription, histone turnover and the epigenetic landscape in neurodevelopment. In the second chapter of this thesis, I describe the development of an experimental strategy that allows the investigation of histone residues in histone variant H3.3 in mammalian cells. Using CRISPR-Cas9, I aim to mutate specific lysine residues (e.g. K4 and K36) of endogenous H3.3 to alanine in mouse embryonic stem cells (ESCs). The mutation will change the property of the residue itself and prevent modifications at this site. In Chapter 3, I describe the experimental results obtained from H3.3K4A and H3.3K36A mutant cells. First, I assess the contributions of H3.3K4 and H3.3K36 residues to transcriptional regulation during neuronal development. Using an *in vitro* differentiation system, I differentiate mutant ESCs into neural precursor cells (NPCs) and post-mitotic glutamatergic neurons to investigate whether specific H3.3 residues are required to give rise to specialized cell types. Gene expression profiles are analyzed by means of RNA sequencing at different developmental stages to assess if gene regulation is perturbed. Second, I profile how lysine residues in H3.3 contribute to H3.3 deposition into distinct chromatin types, such as active genes and regulatory regions. Last, I address how the epigenetic landscape of mouse ESCs is perturbed by H3.3 mutations using antibody-based approaches such as immunoblot and ChIP-Seq and antibody-independent approaches such as mass spectrometry.

2 | CRISPR-mediated gene editing of mammalian histone H3.3

2.1 Background

CRISPR-mediated genome engineering provides a powerful tool to study the role and function of genes and proteins. In the past decades, the advances in genome and transcriptome sequencing techniques have shed light on the genetic causes underlying many human diseases, such as neurodevelopmental disabilities or cancer. Sometimes, a single point-mutation in a protein coding gene has been identified as the primary cause of the disease. CRISPR-Cas offers the possibility to introduce or remove such a mutation of interest to understand disease mechanisms and even bears therapeutic potential. Precise genome editing can also help in understanding the role of histone residues and their post-translational modifications in mammalian gene regulation.

Generally, CRISPR-Cas is an adaptive immunity system that protects bacteria and archae against foreign DNA (Makarova et al., 2011; Jinek et al., 2012). In recent years, components of this system have been modified and made applicable for genome engineering in mammalian cells (Charpentier and Doudna, 2013; Ran et al., 2013). The main components are the endonuclease Cas9 that can cleave double-stranded DNA molecules, and a single-guide RNA (sgRNA). The sgRNA acts as a scaffold and directs Cas9 to a genomic site of interest by a short 20 nucleotide complementary guide sequence. The requirement for Cas9 to bind and cleave the targeted genomic sequence is a protospacer adjacent motif (PAM) in the DNA, most commonly a „5'-NGG“ motif where N is any nucleotide followed by two guanine nucleotides.

CRISPR-Cas mediated double-strand cleavage can be repaired in cells by two different repair mechanisms: Non-homologous end-joining (NHEJ) and homology-directed repair (HDR) (Fig. 2.1). Both pathways repair the double-strand break, but the resulting genotype is different. During NHEJ, the two DNA strands are ligated without a homologous DNA sequence as a template for repair. This often results in deletions, insertions or indels at the repair site, which is useful to generate a knock-out of a particular gene. Meanwhile, HDR is a more precise repair mechanism that requires a DNA template with homology to the repair site such as a homologous chromosome or an exogenously supplied repair template. The HDR pathway can be exploited to

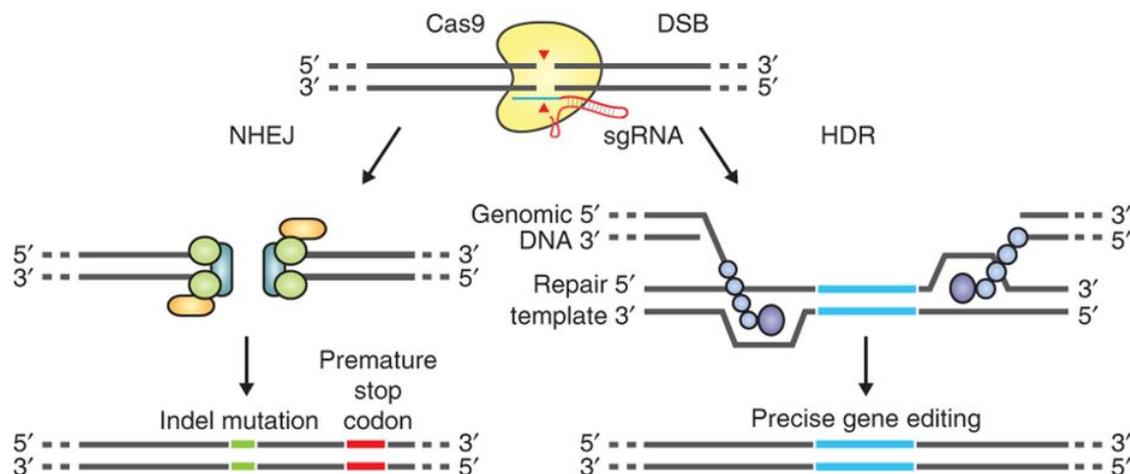


Figure 2.1: CRISPR-Cas9 mediated genome engineering. Cas9 introduces double-strand breaks (DSB) into the DNA, which can be repaired by the Non-homologous end-joining (NHEJ) or the homology-directed repair (HDR) pathway. NHEJ ligates the DNA strands in an error-prone way that results in insertion or deletion (Indel) mutations at the repair site. Indels can cause frameshift and the formation of premature stop codons, resulting in gene knock-out. The more precise HDR pathway repairs the DNA according to a repair template, which can be a plasmid or single-stranded DNA with homology to the DSB site. This pathway allows precise gene editing by introduction of nucleotide changes of interest. Image was obtained from Ran et al. (2013).

generate small, precise modifications in the genome in the presence of an exogenously introduced repair template that carries the desired nucleotide changes (editing) (Ran et al., 2013). Studies showed that the frequency of CRISPR-editing via the HDR pathway is low (Ran et al., 2013). Thus, it is required to establish a robust way for high-throughput screening of many clonal cell lines. So far, screening for genome editing is mostly done by restriction digest (Ran et al., 2013) or Sanger sequencing. Restriction digest requires the introduction of nucleotide changes that give rise to a restriction site to help identifying edited clones. However, this procedure is tedious and it is not always feasible to change nucleotides in a way to create a restriction site, while maintaining the amino acid sequence of the encoded protein. Alternatively, Sanger sequencing is the most reliable technique to identify editing events in cells, but it is expensive if applied to large numbers of clonal cell lines.

The first chapter of my thesis describes how I tested and adapted the CRISPR-Cas9 system developed by Ran et al. (2013) to be applicable for mutation of histone variant H3.3 in ESCs, specifically to exchange lysine residues 4 and 36 (H3.3K4 and H3.3K36) with alanine. As only one of the two H3.3-encoding genes is required for animal survival and development (Tang et al., 2013, 2015; Jang et al., 2015), I knock-out *H3f3a* and introduce point-mutations in the second gene, *H3f3b*. Furthermore, I develop a new screening method to reliably detect editing events in clonal cell lines as an alternative to restriction digest or Sanger sequencing. The new strategy should 1) be more time-effective than screening by restriction digest, 2) reduce the costs of screening many cell lines compared to Sanger sequencing and 3) abolish the need to insert a restriction site into the genome, which is not always possible depending on the targeted genomic

sequence. Finally, I address the issue of large chromosomal deletions and rearrangements that have been reported to occur during CRISPR editing (Lee and Kim, 2018; Zhang et al., 2015). For this, I tested whether RNA-Sequencing data generated from clonal cell lines can be used for a large-scale CRISPR off-target analysis.

2.2 Results

I conducted the experiments, collected and analyzed the data in this chapter, unless indicated otherwise. Nichole Diaz assisted me with cell culture work. Christopher Buccitelli (in the group of Jan Korbel) analyzed DNA-Sequencing data in Fig. 2.8 and advised in the off-target analysis from RNA-Seq data.

2.2.1 Delivery of Cas9 and guide into mouse embryonic stem cells

To target the H3.3 encoding genes *H3f3a* and *H3f3b* for gene editing, I used the plasmid-based CRISPR system developed by Ran et al. (2013). The plasmid encodes a fusion protein of Cas9 and green fluorescent protein (GFP), which allows cell selection by flow cytometry, and a single-guide RNA for targeting of Cas9 to a genomic site. The delivery of this plasmid and a repair template is crucial for successful editing, but primary cells, including ESCs, are difficult to transfect using traditional transduction methods such as liposomal reagents. I tested if electroporation with a nucleofector system (nucleofection) is suitable for the delivery of the CRISPR plasmid into ESCs and analyzed the efficiency by flow cytometry. Overall, around 5.4% of all sorted cells were GFP-positive and therefore successfully transduced (Fig. 2.2). This proportion is sufficiently high to obtain the required cell numbers for gene editing. Next, I continued with optimizing the conditions for gene editing of endogenous H3.3 using this CRISPR-Cas9 system.

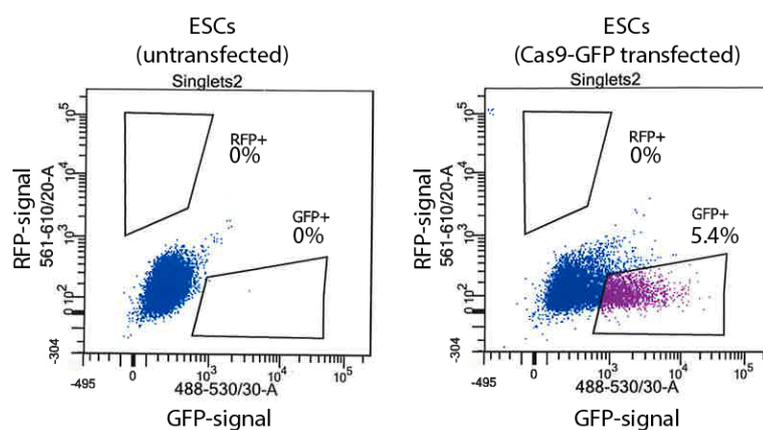


Figure 2.2: Transduction efficiency in ESCs with a plasmid-based CRISPR system using nucleofection. Nucleofected ESCs were analyzed by flow cytometry for GFP expression, which represents successful delivery of the Cas9-encoding plasmid. Displayed are GFP-signal (x-axis) against RFP-signal (y-axis) and rectangles indicate areas of positive cells expressing the analyzed Cas9-GFP fusion protein. Gating strategy for this experiment is depicted in Supplementary Fig. S1.

2.2.2 Optimization of CRISPR-Cas9 mediated gene editing in ESCs

Gene editing through the HDR pathway occurs at lower frequencies than gene knock-out via the NHEJ pathway, but treatment with small molecules has been proposed to

promote the frequency of HDR in cells. I tested the efficiency of two small molecules, Scr7 and L755,507, to promote gene editing of the H3.3 genes (*H3f3a* or *H3f3b*). Scr7 has been reported to promote editing via HDR by inhibiting the activity of DNA ligase IV, an important enzyme in the competing NHEJ pathway (Maruyama et al., 2015; Chu et al., 2015). L755,507 is a β 3-adrenergic receptor partial agonist reported to enhance gene editing, but the mode of action is unknown (Yu et al., 2015). The treatment with individual small molecules at concentrations between 1-10 μ M did not visibly reduce cell viability. To minimize potential toxicity, I treated the cells with the small molecules only for a short period of the culture (12 hours before and 24 hours after delivery of the Cas9-plasmid and single-stranded repair template by nucleofection). In untreated ESCs, I did not obtain edited cell lines carrying the mutation of interest, neither with L755,507 treatment. Using the Scr7 inhibitor, I obtained edited cell lines that had incorporated nucleotide changes according to a supplied repair template. Thus, treatment with Scr7 inhibitor resulted in a higher frequency of editing events than without treatment or by treatment with L755,507 and in this experimental set-up proved to be essential for obtaining edited cell lines (Table 2.1). Overall, the editing frequency of either *H3f3a* or *H3f3b* is around 0-2% of transduced cells for homozygous editing and 1-10% for heterozygous editing.

Table 2.1: Summary of CRISPR editing screens to introduce point-mutations into histone H3.3. Non-treated or treated ESCs were transduced with two CRISPR-Cas9 plasmids targeting both the *H3f3a* and *H3f3b* gene. Total number of clonal cell lines that were screened and the detected edited clonal cell lines (heterozygous or homozygous) are indicated.

Gene	Treatment	Clones total	Edited	
			Heterozygous	Homozygous
H3f3a	no drug	56	0	0
H3f3b	no drug	56	0	0
H3f3a	L755,507	49	0	0
H3f3b	L755,507	49	0	0
H3f3a	Scr7	43	2 (4.7%)	0
H3f3b	Scr7	43	0	1 (2.3%)
H3f3a	Scr7	57	10 15.8%	1 (1.8%)
H3f3b	Scr7	57	1 (1.8%)	1 (1.8%)

To further maximize the editing frequency in ESCs, I tested the efficiency of short templates (200 bp) that are either single- or double-stranded in the presence or absence of Scr7 inhibitor treatment. The repair templates carried a restriction site that gets inserted at the repair site in case of successful editing and the efficiency of editing can be estimated by restriction digest in a bulk of cells sorted by flow cytometry. The combination of Scr7 treatment with a single-stranded repair donor (ssDNA) yielded the highest frequency of editing (Fig. 2.3). This is in agreement with studies that have

reported the advantages of using a short single-stranded DNA sequence (ssDNA) over a long double-stranded plasmid DNA as the repair template (Ran et al., 2013; Chen et al., 2011). For editing of H3.3, the combination of ssDNA with a Scr7 treatment proved to be the most suitable conditions.

Despite the efforts to optimize the conditions for CRISPR-targeting, the observed frequencies (Table 2.1) for precise genome editing were too low to mutate both H3.3 encoding genes in one round of CRISPR targeting. Thus, I decided to mutate only the *H3f3b* gene (encoding H3.3B protein) and remove the *H3f3a* gene by CRISPR-mediated knock-out. The editing frequencies I observed in the initial experiments for homozygous editing of *H3f3b* were between 0-2.3% per screen and I estimated that it would require screening of 100-200 clonal cell lines to obtain a successfully edited clone per mutation.

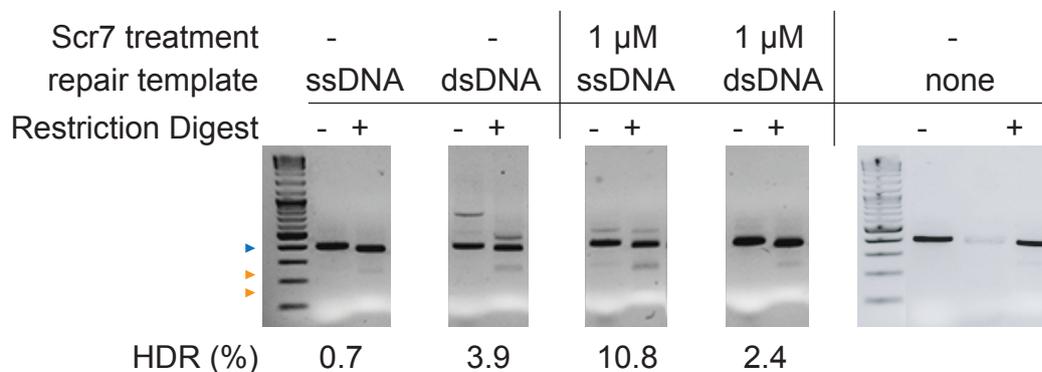


Figure 2.3: Comparison of single-stranded and double-stranded DNA repair templates for CRISPR editing in untreated and Scr7-treated ESCs. Cells were transduced with CRISPR-Cas9 plasmids targeting the *H3f3b* gene and indicated repair templates carrying a restriction site for genomic insertion. A bulk of transduced cells expressing Cas9-GFP was selected by flow cytometry. The targeted *H3f3b* locus was amplified by PCR and subjected to restriction digest. Digestion pattern was analyzed by agarose gel electrophoresis. Successful integration of the restriction site by HDR results in cleavage of the PCR product (blue arrowhead) and the occurrence of smaller digestion products (orange arrowheads). HDR frequency was calculated as the ratio of band intensities.

2.2.3 High-throughput screening for CRISPR editing

The systematic exchange of multiple residues of H3.3 in mammalian cells can only be achieved if a reliable method allows high-throughput screening of many clonal cell lines at reduced costs. Whereas Sanger sequencing is a fast and precise method for screening, it is expensive if applied to many cell lines. Instead, screening by restriction digest is inexpensive, but tedious and requires the insertion of a restriction site into the targeted genomic locus.

Studies used real-time quantitative PCR (qPCR) to measure the methylation status of a single cytosine at a specific positions in the DNA. For this, specific primers are designed to recognize a single nucleotide polymorphism that arises after bisulfite con-

version of methylated DNA at the most 3' end of the primer (Dugast-Darzacq and Grange, 2009). I tested if a similar strategy can be employed to detect CRISPR-editing by designing mutation-specific primers that recognize the inserted nucleotide changes of interest (Fig. 2.4a,b). I will refer to this screening method as Mismatch-qPCR. I sorted transduced cells by single cell sorting and performed Mismatch qPCR after CRISPR targeting. Using this method, I was able to separate edited clones from wild type clones by shifts to lower cycle numbers (rounds of amplification) (Fig. 2.4c). In combination with a primer that recognizes the unchanged wild type allele, it was possible to distinguish heterozygous from homozygous clones. Heterozygous clones with one mutant and one wild type allele amplify in a qPCR reaction with both primer sets, while homozygous clones only amplify using the mutation-specific primer. Using the already established restriction digest method, I confirmed homozygosity and heterozygosity of the clonal lines, which can be identified by the complete or incomplete digestion of a PCR product (Fig. 2.4c). The results from restriction digest were in agreement with the results from the qPCR screen.

Using the optimized conditions for CRISPR-editing in ESCs and the developed Mismatch qPCR screen, I generated cell lines edited homozygously at either lysine 4 or lysine K36 of H3.3B. By means of Sanger sequencing, I determined the exact genotype of the edited clones and confirmed that the lysine to alanine mutation was introduced in addition to the synonymous mutations inside the guide binding site in both alleles of *H3f3b* (Fig. 2.5a,b). For some other clones that were detected by screening, I observed incomplete repair resulting in additional small deletions around the guide binding site. These clones were discarded and only clones that have incorporated all nucleotide changes correctly from the repair template were used for downstream analysis. Next, I went on to remove the *H3f3a* gene by CRISPR-mediated knock-out.

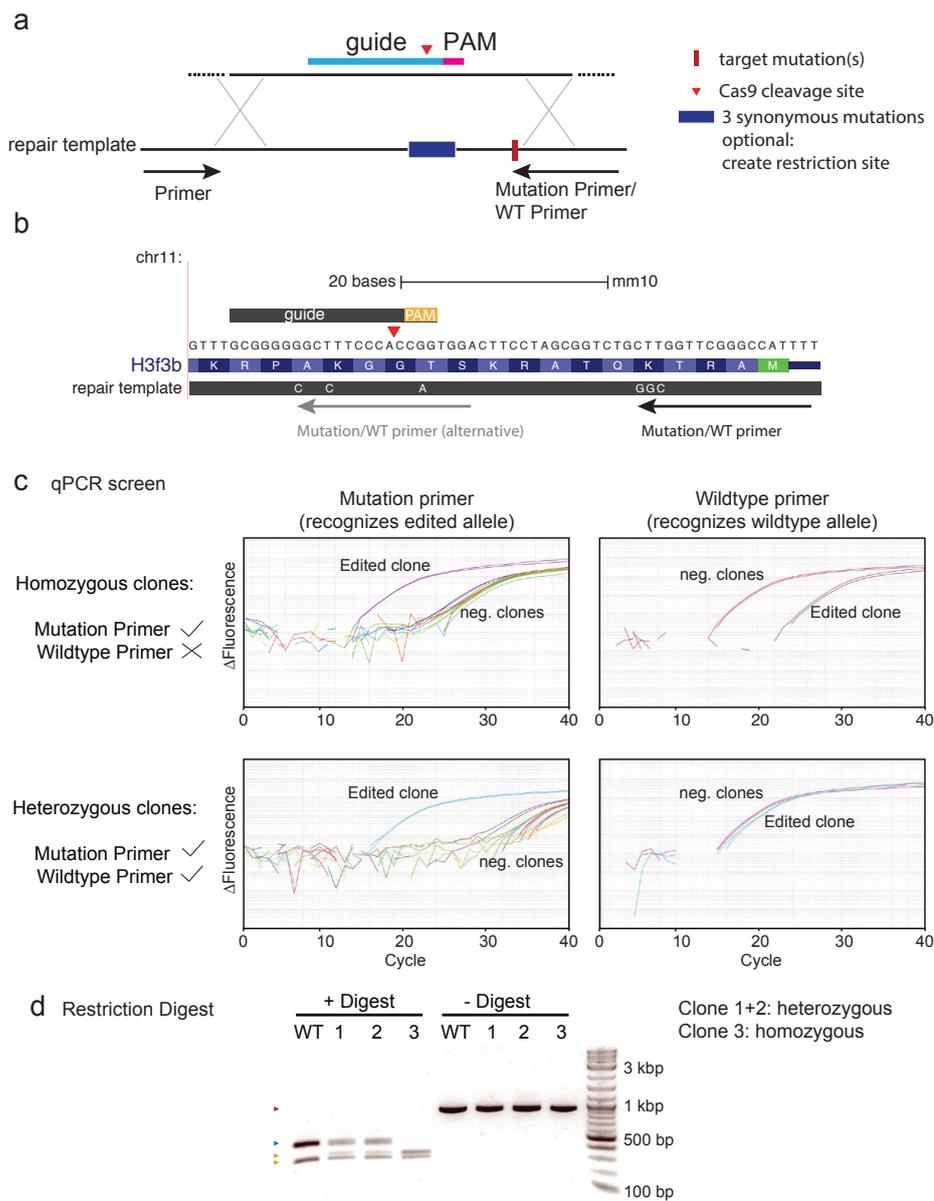
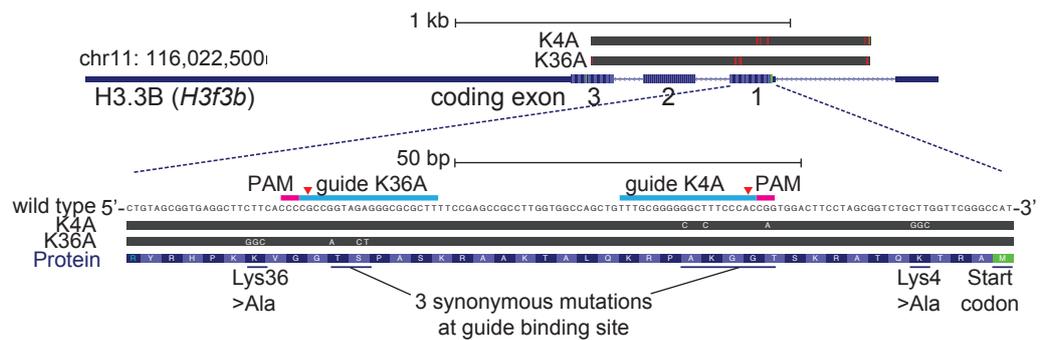


Figure 2.4: Mismatch qPCR screen detects CRISPR-mediated point-mutations in H3.3B. (a) Guide sequences were designed to target Cas9 close to the mutation site of interest. The single-stranded DNA repair template contains nucleotide changes to introduce a target mutation (lysine-to-alanine), and 3 additional synonymous mutations into the guide binding site or into the PAM to prevent re-cleavage after repair. Optionally, synonymous mutations can give rise to a new restriction site used to validate clones. The mutation-specific primer includes nucleotide changes that arise after CRISPR-editing at the most 3' end. The wild-type (WT) primer recognizes the same, but unmodified genomic site. (b) Examples of two mutation-specific primers for Mismatch qPCR that detect editing of lysine 4 to alanine in H3.3B by recognizing either the K4A mutation or the synonymous mutations inside the guide. (c) Mismatch qPCR screen of CRISPR cell lines using mutation-specific and wild-type primers. Successful amplification results in an increase of the fluorescent signal (y-axis) at lower cycle numbers (x-axis). DNA of homozygously edited clones is amplified only with mutation-specific primers, whereas heterozygous clones are also amplified using the wild-type primer. (d) Confirmation of editing events by restriction digest using a newly introduced restriction site after CRISPR targeting. DNA of wildtype cells (WT) and positive clones predicted by Mismatch qPCR screening were used for PCR amplification followed by restriction digest with BanI. Digestion pattern was analyzed by agarose gel electrophoresis. Digestion of the PCR product (red arrowhead) of wild type DNA results in a larger product (blue arrowhead) than from edited DNA (orange arrowheads) with an additional integrated restriction site. Restriction digestion confirms the detected editing events by qPCR.

a Sanger sequencing



b Chromatograms

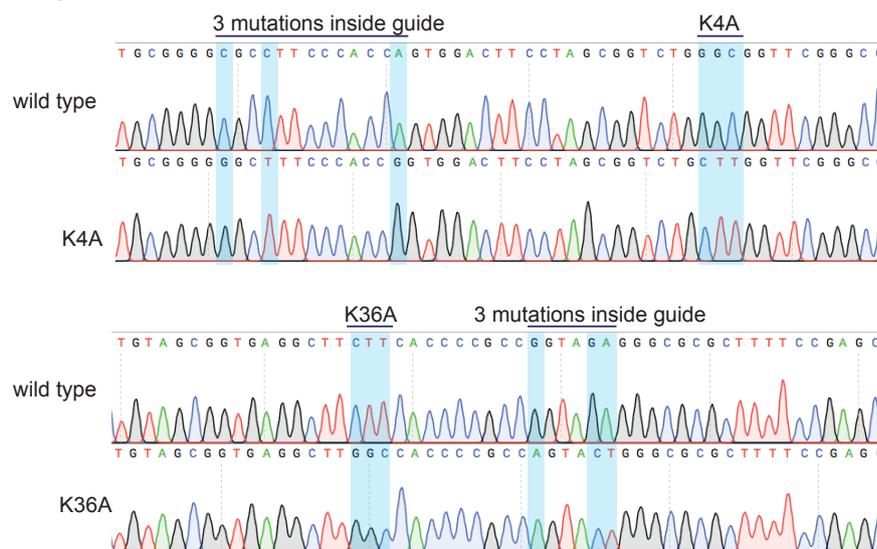


Figure 2.5: Sanger sequencing confirming the successful integration of nucleotide changes into H3.3B at lysine 4 and 36. (a) Sanger-sequencing results of the H3.3B locus for H3.3K4A and H3.3K36A mutant cells. Blat tool is used to compare the results with the wildtype genome (mm10) and the search is visualized using UCSC. Zoom into the region of the first coding exon reveals successful introduction of lysine-to-alanine mutation at either K4 or K36, respectively, and the introduction of 3 additional synonymous mutations inside guide recognition site or PAM. (b) Analysis of chromatograms from Sanger sequencing confirms the homozygous exchange of targeted nucleotides in H3.3K4A/K36A mutant cell lines.

2.2.4 Knockout of H3.3A

Individual knock-out experiments of one of the two H3.3 encoding genes have shown that only one copy of H3.3 is absolutely required for viability and reproduction in mice (Tang et al., 2013, 2015; Jang et al., 2015). Thus, I deleted H3.3A, after having introduced point-mutations into both alleles of H3.3B. This step was taken because removal of one gene copy is not expected to compromise differentiation potential and pluripotency state of ESCs, but will facilitate the generation of mutant cell lines. To remove H3.3A from ESCs, I used Cas9 and a guide targeting the first coding exon of the *H3f3a* gene. RNA-Seq of knock-out clones revealed significantly reduced levels of the entire H3.3A transcript compared to wild type cells (Fig. 2.6) suggesting that deletions inside the first coding exon result in non-sense mediated decay of the remaining mRNA. The few remaining reads detected for H3.3A are mapping to other non-targeted exons, which confirms the successful targeting of the first coding exon by CRISPR-Cas9 and thereby knock-out of H3.3A (Fig. 2.6).

As a next step, I want to confirm that during clonal selection and CRISPR targeting the integrity of the genome was not affected, e.g. by chromosomal rearrangements.

In agreement with knock-out of one gene copy, I consistently observed reduction of H3.3 protein levels for control (H3.3A knock-out) and H3.3K4A and H3.3K36A mutants (H3.3A knock-out + H3.3B editing) compared to wild-type cells by approximately 50% as analyzed by immunoblot (Fig. 2.6). This confirmed the successful knock-out of *H3f3a* in CRISPR clones on the protein levels, and furthermore indicated that gene editing of *H3f3b* did not significantly reduce protein expression of H3.3B in ESCs.

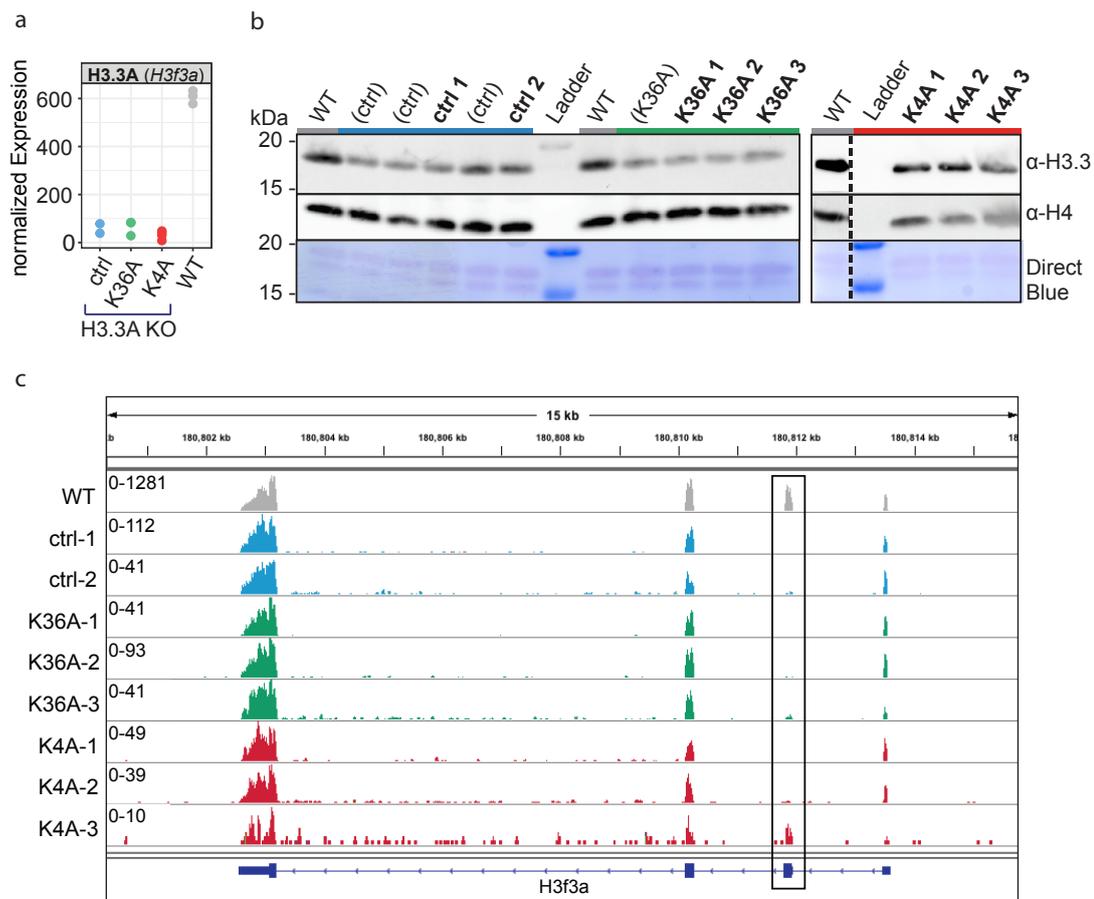


Figure 2.6: Successful knock-out of *H3f3a* in CRISPR clones. (a) Normalized mRNA expression of the *H3f3a* gene measured by RNA-Sequencing shows strongly reduced gene expression levels and successful gene knock-out in control, K4A and K36A CRISPR cell lines. Each dot represents one biological replicate. (b) Immunoblot analysis of H3.3 protein levels in CRISPR cell lines. Control, K4A and K36A mutant cells have reduced H3.3 protein expression compared to wild type cells (WT) after deletion of H3.3A. (c) Visualization of RNA-Sequencing reads in the *H3f3a* gene for CRISPR cell lines and wild type cells. Black box indicates the first coding exon that was targeted for deletion with CRISPR. Only few reads are detected for the entire *H3f3a* gene. Track height is indicated on the left and shows strongly reduced expression for the CRISPR clones compared to wild type, suggesting non-sense mediated decay of the mRNA.

2.2.5 CRISPR off-target analysis for large chromosomal deletions

Double-strand cleavage by Cas9 can cause unintended off-target effects that affect genome integrity. (Lee and Kim, 2018; Zhang et al., 2015). As a part of this study, I generated RNA-Seq data from the CRISPR-Cas9 targeted cell lines, therefore I wanted to test if RNA-Seq data can be used to detect whether genomic integrity is affected in these cells. The advantage would be to exclude clonal cell lines with chromosomal deletions or duplications, which could otherwise complicate downstream analysis. Using mRNA-Seq, I determined gene expression changes in CRISPR cell lines relative to their wild type ESC line of the same genetic background. The gene expression changes were compared to the genomic coordinates of the respective gene. Systematically down- or up-regulated genes that are located in proximity to each other indicate

large-scale chromosomal abnormalities (Fig. 2.7a).

Using this strategy, I found that incomplete repair of a chromosome by NHEJ or HDR can result in deletions of chromosome pieces, ranging from the CRISPR target site to the end of a chromosome (Fig. 2.7b,c,d). Such deletions result in one-allelic loss of hundreds of genes, and I will refer to these unintended on-target effects as chromosome arm-loss. I observed that chromosome arm-losses can occur independent of the exact CRISPR guide sequence and on different chromosomes, because they were detected during the targeting of *H3f3a* on chromosome 1 or *H3f3b* on chromosome 11 (Fig. 2.7b,c). Since both H3.3-encoding genes are located at the periphery of chromosome 1 and 11, respectively, it is possible that CRISPR-targeting of genes at the ends of chromosomes are more likely to result in one-allelic chromosome-arm loss. It should be noted that such deletions are not detectable by traditional Sanger Sequencing, because only the intact allele is amplified in a PCR reaction. Additionally, I also observed rearrangements of chromosomes that were not targeted by CRISPR, e.g. of chromosome 6 (Fig. 2.7d). These rearrangements can potentially be CRISPR off-target effect, but may also have occurred spontaneously during clonal selection.

To increase the confidence in genome integrity predictions from RNA-Seq data and to validate the method, I tested whether the predicted chromosomal deletions/duplications can be confirmed on the genomic level by means of DNA sequencing.

From RNA-Seq data of a chosen cell line, a cluster of up-regulated genes and a cluster of down-regulated genes between the CRISPR target site and the chromosome end were detected (Fig. 2.8a). The same chromosomal abnormality was also detectable using DNA-Seq data (Fig. 2.8b) and this confirmed that the gene expression changes were the result of a duplication-deletion rearrangement on the genomic level. Compared to the DNA-Seq analysis, RNA-Seq data yields a lower resolution because the predictions are dependent on the gene-density per chromosome, which is rather sparse considering that only 62% of the genome is transcribed, and an even smaller fraction of this corresponds to coding exons (5.5%) (ENCODE, 2012). Thus, DNA-Seq data allows a more detailed analysis of chromosomal abnormalities with higher precision and confidence, but with RNA-Seq data it is possible to make similar predictions especially in the case of large chromosomal abnormalities.

The occurrence of CRISPR-dependent and -independent effects on genome integrity suggests that an extensive on- and off-targeted analysis for generated clonal cell lines is recommended and should be integrated into the standard workflow for validation of CRISPR cell lines. In addition, these results suggest that for generating knock-out cell lines, it would be best to choose cell lines that show deletions of different lengths in the two targeted alleles, which can help in avoiding chromosome-arm loss and ensure that two intact chromosomes remain. The two alleles need to be amplified and appear as amplicons of different sizes on an agarose gel after PCR and can be extracted and sequenced individually. However, large-scale deletions can also occur during gene editing and it is necessary to test genome integrity either by sequencing or by qPCR

for systematic down- or up-regulation of proximal genes.

In conclusion, the summarized strategy for gene editing in ESCs is displayed in Fig. 2.9. It was developed for editing of H3.3 encoding genes in mouse ESCs, but should be applicable to any gene of interest and cell line.

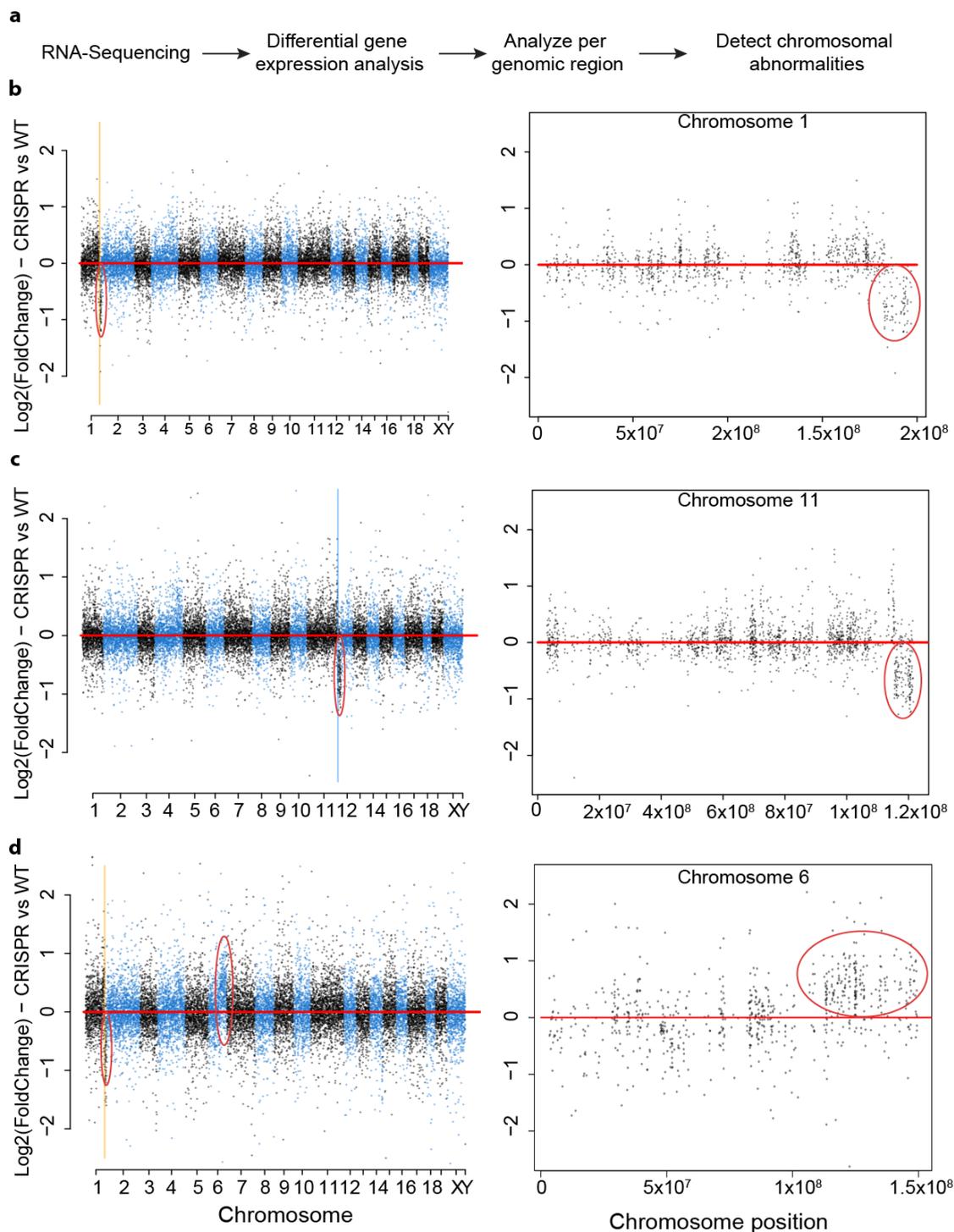


Figure 2.7: Prediction of chromosomal rearrangements from RNA-Sequencing data. (a) Strategy for analyzing genome integrity from RNA-Seq data. (b,c,d) CRISPR off-target analysis from differentially expressed genes. $\log_2(\text{FoldChanges})$ of gene expression in three different CRISPR clones compared to unmodified wild type ESCs were determined by DESeq2 and plotted over chromosome position for all (left) or a specific chromosome (right). Lines indicate CRISPR cleavage site inside *H3f3a* gene (yellow) and *H3f3b* gene (blue). Loss or duplication of a chromosome part can be detected by coordinated up- or down-regulation of proximal genes. (b) CRISPR clone showing a cluster of systematically down-regulated genes on chromosome 1 close to the CRISPR targeting site in *H3f3a* gene. (c) CRISPR clone showing a cluster of systematically down-regulated genes on chromosome 11 close to the CRISPR targeting site in *H3f3b* gene. (d) CRISPR clone showing a cluster of systematically down-regulated genes on chromosome 1 close to the CRISPR targeting site in *H3f3a* gene and a large cluster of up-regulated genes on chromosome 6.

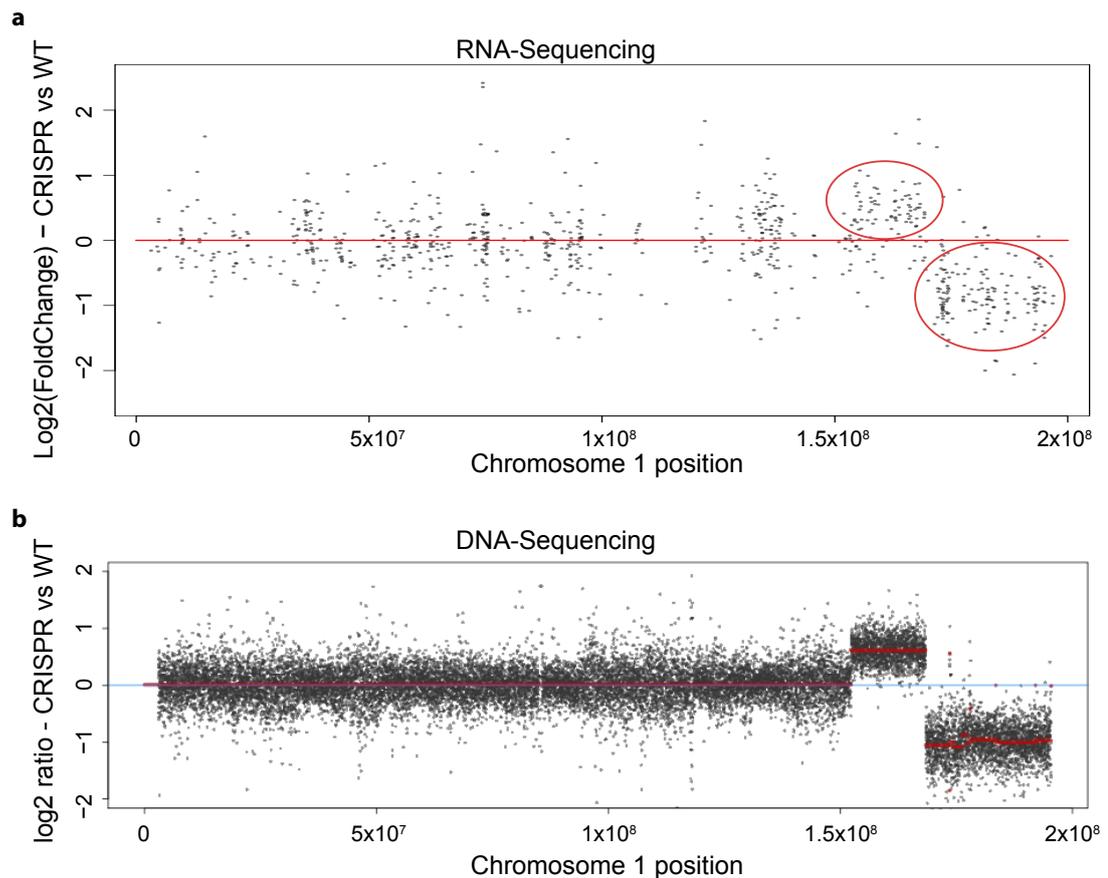


Figure 2.8: Comparison between RNA-Seq and DNA-Seq data in predicting chromosomal rearrangements.(a) Differential gene expression analysis with regard to the genomic coordinates from RNA-Seq data. CRISPR clone shows two clusters of systematically up- and down-regulated genes on chromosome 1 close to the CRISPR targeting site, suggesting a partial chromosomal duplication and deletion. (b) Analysis of DNA-Seq reads with respect to genomic coordinates confirms the duplication-deletion rearrangement at the CRISPR targeting site on chromosome 1 on the genomic level. *Analysis of genomic DNA-Seq data in (b) was performed by Christopher Buccitelli from Jan Korbel's group.*

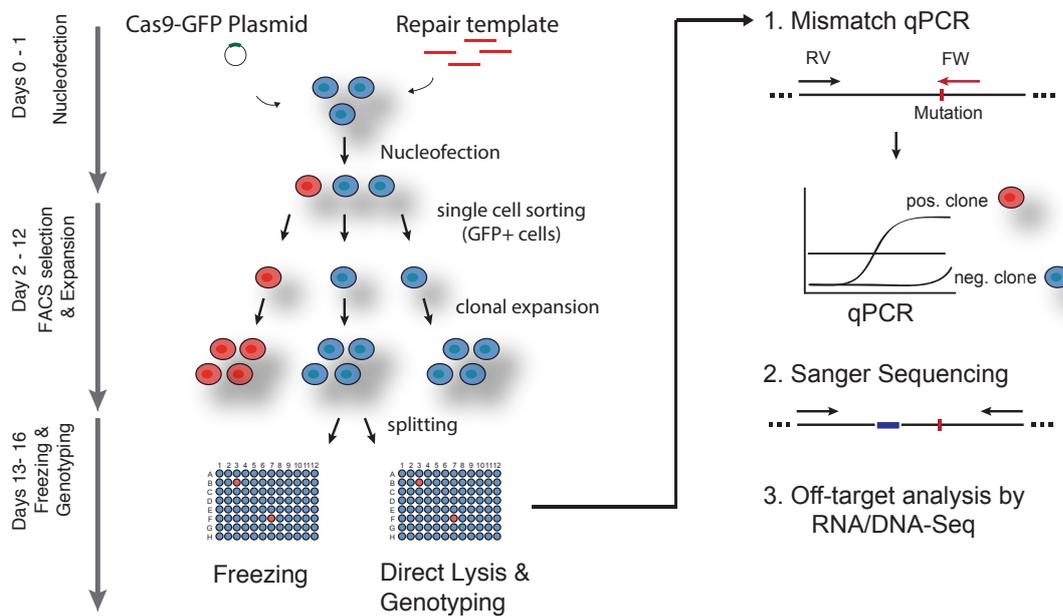


Figure 2.9: Overview of CRISPR-editing workflow in mouse ESCs. General scheme used to introduce point-mutations into H3.3B. Cas-9 plasmid with guide and single-stranded repair templates are delivered into ESCs by nucleofection. Transduced GFP-positive cells are selected by flow cytometry and single cells are sorted into 96-well plates. Clonal cell lines are expanded and split into two identical plates for freezing and screening. For Mismatch-qPCR, cells are lysed directly in a 96-well plate and successful editing events are detected in a qPCR reaction. Editing events can be confirmed by Sanger sequencing and optionally by restriction digest. Large-scale on- and off-target analysis is performed in selected clonal lines from RNA- or DNA-Seq data.

2.3 Discussion

Genome engineering by CRISPR-Cas9 provides a powerful tool to exchange endogenous histone residues and to probe their function in mammalian development. Targeting of the H3.3 genes (*H3f3a* and *H3f3b*) showed that gene editing occurred at low frequencies in mouse ESCs and it required extensive screening of many clonal cell lines to obtain successfully edited clones. The observed frequencies were not sufficiently high to achieve editing of both H3.3 genes in a single targeting step and two consecutive rounds of CRISPR targeting would be required to exchange both genes. However, knock-out studies in mice suggest a partial redundancy of the two H3.3 genes, with only one of them being absolutely required for the completion of an entire mouse life cycle (Jang et al., 2015; Tang et al., 2015). Therefore, I introduced mutations into one of the two H3.3 encoding genes (*H3f3b*/H3.3B) in combination with the knock-out of the second gene (*H3f3a*/H3.3A), thereby generating clonal cell lines that have the entire protein pool of H3.3 mutated. Specifically, I introduced homozygous point-mutations into H3.3B at lysine 4 and 36, respectively. The resulting half-dosage of H3.3 could potentially become a problem in post-mitotic cells, which dependent on H3.3 as the predominant H3 variant (Maze et al., 2015). Furthermore, H3.3 has been shown to be critical to maintain genome integrity and complete loss of H3.3 in mice results in embryonic lethality due to DNA segregation defects during mitosis (Jang et al., 2015). Therefore, it is crucial to include H3.3A knock-out cells as controls for downstream analysis rather than wild type cells that have both H3.3 genes intact. This will allow to discriminate the effect of the introduced H3.3 mutation from a dosage-dependent effects of H3.3 itself.

To this end, I generated numerous clonal cell lines to have a complete set of controls and H3.3 mutants for downstream characterization. To simplify this effort and to screen clonal cell lines with high through-put, I developed a qPCR-based method (Mismatch qPCR) that could reliably identify CRISPR-edited clones. By direct comparison with the restriction digest method (Ran et al., 2013), Mismatch qPCR proved to be a faster screening method that did not require the insertion of a restriction site into the genome. The read-out can be observed during the qPCR reaction without requiring subsequent analysis steps. Sanger sequencing was required to exclude false positive clones and to confirm the precise genotype of the clonal cell lines. Nevertheless, sequencing of few clones after screening by Mismatch qPCR was more economical than to sequence all generated clonal cell lines. The limitation of Mismatch qPCR is certainly the requirement of suitable primer pairs for screening. Dependent on the DNA sequence and GC-content of the targeted locus, it is not always possible to design primers that fall into the recommended property range (e.g. melting temperature), which is predicted to result in less efficient PCR amplification. However, the read-out of Mismatch qPCR is qualitative and not quantitative and should not be compromised by less efficient primers.

Following the generation on CRISPR-edited clones, I wanted to confirm that their genomic integrity had not been compromised by the targeting. Chromosomal rearrangements can occur simply during clonal selection of ESCs, and recently, it has been reported that large chromosomal deletions and complex genomic rearrangements can occur at the CRISPR-targeted site in mouse ESCs and other cell types (Kosicki et al., 2018). These results caution against neglecting the risk of on- and off-target effects introduced by CRISPR-Cas9. Testing the genomic integrity of generated clonal lines from commonly available genomic datasets such as RNA-Seq and DNA-Seq data can help in excluding affected cell lines. Using RNA-Seq data, I frequently observed deletions at the site of CRISPR-targeting, and less commonly rearrangements on other chromosomes. At the CRISPR-target site, double-strand DNA cleavage resulted in the one-allelic loss of a chromosome arm and thus down-regulation of hundreds of genes. Large-scale deletions and rearrangements severely affected genome integrity, and suggest that an extensive on- and off-target analysis for generated clonal cell lines is indeed necessary and should be integrated into the standard workflow for CRISPR editing. Depending on the availability, both RNA- or DNA-Sequencing data are suitable for this analysis and these datasets are often publicly available for already published studies, e.g. input DNA sequencing data from Chromatin Immunoprecipitation (ChIP) experiments. A sequencing based approach for identifying off-target clones will be easily implemented and might be comparable with the traditional karyotype analysis.

With the growing list of mutations associated with human diseases, CRISPR-Cas9 mediated editing is becoming increasingly important to study disease mechanisms. Economical screening methods with high-throughput such as Mismatch qPCR in combination with a large-scale off-target analysis can facilitate the generation of multiple biological replicates for a mutation, which is essential for data interpretation and reproducibility. Furthermore, several publications have proposed alternative CRISPR based platforms for genome editing, such as Cas9 proteins from various bacterial species or different endonucleases like Cpf1 (Ran et al., 2015; Zetsche et al., 2015). Expanding Mismatch qPCR screening to other CRISPR systems would allow the direct comparison of their efficiency in a quantitative manner.

2.4 Materials & Methods

Cell culture

Murine ESCs (129XC57BL/6J) were cultured in ESC media containing Knockout-DMEM (Thermo Fisher) with 15% EmbryoMax FBS (Millipore) and 20 ng/ml leukemia inhibitory factor (LIF, produced by Protein Expression Facility at EMBL Heidelberg), 1% non-essential amino acids, 1% Glutamax, 1% Pen/Strep, 1% of 55mM beta-Mercaptoethanol solution. Cells were maintained at 37°C with 5% CO₂. ESCs were routinely tested for mycoplasma.

CRISPR editing

This paragraph summarizes the final CRISPR editing strategy, individual steps are described in more details underneath this paragraph. Guide sequences were cloned into pSpCas9(BB)-2A-GFP (PX458, Addgene) according to instructions by Ran et al. (2013). For gene editing, 2×10^6 ESCs were transfected with 2 µg Cas9-GFP plasmid and 100 µM ssODN repair templates (180 bp, IDT ultrameres) using electroporation (Nucleofector, Lonza). Prior to nucleofection, ESCs were treated with 10 µM Scr7 inhibitor (Xcessbio, M60082-2s) for 12 hours to promote gene editing (Maruyama et al., 2015). Single-cell sorting was performed 48 hours post-electroporation for GFP positive cells. Colonies were expanded for genotyping and freezing. To detect gene editing events, clonal cell lines were lysed in a 96-well plate and screened for successful mutation. Editing events were confirmed and checked for homozygosity by Sanger Sequencing.

We used sgRNA guides that recognized the following genomic sequences:

Guide	target	5' → 3' Sequence
sgRNA_H3f3a_g3	Knock-out H3.3A	CCTGGGTGCTTTACCACCGG
sgRNA_H3f3b_g7	Edit H3.3B K36A	AAGCGCGCCCTCTACCGGCG
sgRNA_H3f3b_g3	Edit H3.3BK4A	TTTGC GGGGGGCTTTCCCAC

Repair template	5' → 3' Sequence
ssODN_H3f3b_g3_H3.3K4A	GGAGCCCGGTGACCTGGCCTTGAACGTCGCTTGTCTCGCAGG TGAAAAAAAAATGGCCCGAACCGCCCAGACCGCTAGGAAGTCC ACTGGTGGGAAGGCGCCCCGCAAACAGCTGGCCACCAAGGCG GCTCGGAAAAGCGCGCCCTCTACCGGCGGGGTGAAGAAGCCT CACCGCTACAGGT
ssODN_H3f3b_g7_H3.3K36A	CCAAGCAGACCGCTAGGAAGTCCACCGGTGGGAAAGCCCCC GCAAACAGCTGGCCACCAAGGCGGCTCGGAAAAGCGCGCCCA GTACTGGCGGGGTGGCCAAGCCTCACCGCTACAGGTAGGCAG AGGGCTGGGAACAATGACTTGGCCCGCGGCTTGC GGGCGGGC GCTCTCTCCCTTC

Guide design and cloning into CRISPR plasmid

Guides were designed with homology to a sequence close to the mutation site of interest using MIT's Optimized CRISPR design tool. As a general guideline, the guide binding site should ideally be less than 30 nucleotides away from the mutation site of interest, and can also overlap the mutation site. If the mutation site is close to an intron, it is recommended to use an intronic guide sequence in case additional indels occur at the CRISPR cutting site, but this is optional. Guide sequences with an aggregate score of greater than 50% were selected and cloned into pSpCas9(BB)-2A-GFP (PX458, Addgene) or pSpCas9(BB)-2A-RFP (modified from PX458) according to instructions by Ran et al. (2013). For this purpose, phosphorylated DNA oligos (5'-Phos) were ordered from Eurofins according to this scheme:

CACC + G + guide sequence forward
AAAC + guide sequence reverse + C

pSpCas9(BB)-2A-GFP was digested with BbsI, followed by dephosphorylation using Antarctic Phosphatase (NEB) and separated from undigested plasmid by 1% agarose gel electrophoresis. Digested plasmid was extracted from the gel (Gel extraction Kit, Qiagen). Complementary guide oligos were annealed and cloned into pSpCas9(BB)-2A-GFP/RFP plasmid, from here on referred to as Cas9-GFP-guide or Cas9-RFP-guide plasmid.

Design of repair template

The single-stranded repair template was designed to encompass the DNA 90 bp upstream and downstream of the Cas9 cutting site as suggested by Ran et al. (2013). I used IDT's DNA ultrameres (up to 200 bp) as DNA template. The repair template contains the mutation of interest (e.g. a lysine to alanine exchange in H3.3 at lysine 4 or lysine 36) and 3 synonymous mutations inside the guide binding site to prevent repeated cleavage by Cas9. Editing efficiency can be improved if one of these additional mutations changes the PAM sequence into a non-PAM sequence. Synonymous mutations do not change the resulting protein sequence, and should be chosen in a way that codon usage frequency is considered and codons with very low frequencies should not be used as they can alter expression levels of the encoded protein. To generate double-stranded repair templates (dsDNA) the single-stranded repair template was amplified in a PCR reaction using one primer annealing to the 5' end of the template. This results in linear amplification of the complementary strand.

Nucleofection

For gene editing 2×10^6 ESCs were transfected with 2 μ g Cas9-GFP-guide plasmid and 5 μ l of 100 μ M ssODN repair templates (180 bp, IDT ultrameres) using electroporation (Nucleofector, Lonza). Cells were resuspended in 100 μ l of P3 solution (Lonza) and

2 µg plasmid DNA and 5 µl donor template were added. Cells were transferred into a cuvette and electroporated with pulse code CG 104 for mouse ESCs (“ES, mouse”). After electroporation cells were plated into a T-25s flask with pre-plated MEFs containing ESC media. For drug treatment, ESCs were treated with 5 µM L755,505 (Xcessbio, M60237-2s) or 10 µM Scr7 inhibitor (Xcessbio, M60082-2s) for 12 hours prior to nucleofection and for additional 24 hours after nucleofection to promote gene editing (Maruyama et al., 2015; Chu et al., 2015; Yu et al., 2015).

Single-cell sorting by flow cytometry

Single-cell sorting by flow cytometry was performed 48 hours post-electroporation for GFP positive cells. Single ESCs were sorted into the wells of a 96-well plate containing pre-plated MEFs and 150 µl of ESC media. Cells were sorted on an FACS Aria Fusion sorter (BD Biosciences). Gating was performed with BD’s FACSDiva 8.0.1 software and single cells were chosen for analysis after doublet discrimination by detection of disproportions between cell size (FSC-A) vs. cell signal (FSC-H). Correlation of FSC A vs. H, the same cell will have the same (or very similar) values in both axis. Therefore, all singlet events will fall more on a diagonal than doublets. The gating strategy for GFP-positive transfected cells is depicted in Supp. Fig. S1.

Cell expansion and lysis

Growing cell colonies were dispersed 6 days after sorting and 9 days after sorting split into 2 identical 96-well plates. One plate was used for freezing in DMSO-containing medium and the second was used for screening of clones. To detect gene editing events, clonal cell lines were lysed directly in a 96-well plate. For cell lysis medium was removed from wells and 70 µl of direct-lysis solution (301-C, Viagen Biotech) supplemented with 0.5 mg/ml proteinase K (Roche 03115887001) were added to each well. Cells were lysed at 55°C for 2 hours, shaking at 350 rpm and afterwards proteinase K was heat deactivated for 45 min at 85°C. Cell lysates can be stored at 4°C for up to one week, for long-time storage (1-2 months) lysates were frozen at -20°C.

Mismatch qPCR

Screening primers to detect genome editing events by qPCR should be suitable for standard qPCR reaction and the total amplicon size should be under 150 bp to guarantee successful amplification during elongation step. One of the primers encompasses the editing site, ideally directly ending with a point mutation on the 3’ end. Thus, amplification with this primer should not work on the wild type sequence. Instead, the wild-type primer is designed to recognize the unmodified genomic sequence. Reverse primer recognizes the wild-type sequence away from the CRISPR editing site. For editing screen, 0.5 µl of the crude lysate are sufficient for a 20 µl qPCR reaction (96-well plate) or 0.25 µl for a 10 µl qPCR reaction (384-well plate), higher amounts of crude lysate can interfere with PCR reaction. DNA from a bulk sorting should be

included as a positive control and wild type DNA from unmodified cells as a negative control. qPCR is run with a standard cycling program to detect cycle threshold (Ct) values. All clones with significantly lower Ct numbers than the negative control (wild type), and similar or lower Ct values than the positive control were used for downstream validation. For sanger sequencing, a region of 1-1.5 kb around the mutation site was amplified by PCR and sequenced from both ends to confirm editing. The crude lysates can also be used for genotyping, but may result in low-quality Sanger sequencing. In this case it is recommended to thaw all positive clones, expand them and extract genomic DNA (Puregene Core Kit B, Qiagen) to improve the quality of the sequencing reaction. Editing events were confirmed and checked for homozygosity by analyzing the chromatogram (SnapGene Viewer) of the Sanger Sequencing reaction.

Restriction Digest

A DNA fragment of 1kb around the CRISPR cutting site was amplified using PCR. PCR product was digested with a suitable restriction enzyme (for which restriction site was inserted into the genome) for 30 minutes at 37°C. Digestion products were analyzed by agarose gel electrophoresis.

Knock-out of H3.3A

For knockout of H3.3A, ESCs were nucleofected with 2 µg Cas9-GFP plasmid as described above for gene editing, but without Scr7 treatment or a repair template. For screening of knock-out, RNA was extracted from cells and 1 µg of total RNA was reverse transcribed using random primers (High Capacity Reverse Transcriptase, Applied Biosystems). Resulting cDNA was used as template for PCR reaction. Primers were complementary for the H3.3A coding sequence. Length of the amplicon was visualized by agarose gel electrophoresis and shorter amplicons were indicative for successful removal of the first exon and therefore knock-out of H3.3A. Alternatively, deletions were detected by PCR on genomic DNA of clonal cell lines followed by agarose gel electrophoresis. Deletions were visible by shifts to lower amplicon sizes. Finally, deletions were confirmed by Sanger sequencing and RNA-Seq analysis.

CRISPR off-target analysis by RNA-Seq analysis

CRISPR off-target effects in the form of chromosomal duplications/deletions were ruled out by RNA-sequencing. RNA-Seq analysis. RNAs were extracted from approx. 1×10^6 cells using RNeasy Kit (Qiagen), followed by DNase digestion using TURBO DNase (Thermo Fisher). mRNAs were isolated from 1 µg of total RNA using a PolyA selection kit (NEB) and sequencing libraries were prepared following instructions from NEBs Ultra Library Preparation Kit for Illumina. All samples were barcoded, pooled and sequenced on a HiSeq2000 Sequencer (Illumina) using a 50 bp single-end run. Sequencing reads were mapped to mouse reference genome (mm10 assembly) using Tophat2 aligner with default settings for single-end reads. Reads per gene

were counted using HTSeqCount union or intersection_nonempty mode. We used Ensembl gene annotation Mus_musculus.GRCm38.83. Differential RNA-Seq analysis between clone and wild type cells was performed and $\log_2(\text{FoldChanges})$ were plotted over chromosome position to obtain distribution profiles of overall gene expression changes. Cell lines that displayed deletions or duplications of chromosome regions, as seen by concomitant up- or down-regulation of close-by genes, were discarded and not used for analysis.

Protein extraction & Immunoblotting

For crude nuclear extraction, cells were resuspended in PBS+0.05% Triton-X and rotated for 30 min at 4°C to lyse cells. Nuclei were pelleted by centrifugation at 1000xg for 5 min at 4°C, resuspended in SDS loading dye and sonicated with EpiShear Sonicator (Active Motif). Prepared lysates were separated in MES buffer on precast 4–12% Bis-Tris SDS-PAGE gels (Thermo) and transferred onto nitrocellulose membranes. Blots were incubated overnight with primary antibodies, for 1 hour with HRP-conjugated secondary antibody and developed using enhanced chemiluminescent substrate (Thermo Fisher). I we used the following antibodies: H3.3 (09-838, Millipore), H3 (ab18521, abcam), H4 (ab10158, abcam), and secondary antibodies goat anti-mouse IgG-HRP (1721011, BioRad), goat anti-rabbit IgG-HRP (1706515, BioRad).

3 | Functional characterization of histone H3.3 mutants

3.1 Background

Histone H3 plays a central role in the chromatin-based epigenetic regulation of gene expression. The dependence on histone H3 tail residues can vary greatly between different organisms, and single mutations that are viable in yeast can have severe phenotypic outcomes in animals. In yeast, remarkably few residues within the N-terminal tail are essential for viability, and even long deletions of the N-terminal region of H3 are tolerable to a great extent (Dai et al., 2008; Nakanishi et al., 2008). In animals, mutations of H3 tail residues can induce morphological defects (Pengelly et al., 2013; Herz et al., 2014), result in lethality (McKay et al., 2015) or can contribute to the development of cancer (Schwartzentruber et al., 2012; Wu et al., 2012; Lu et al., 2016). Thus, histone tail residues appear to be increasingly important in organisms that exhibit diverse regulation of genome activity across different cell types and developmental stages. The expansion of histone mutation studies to mammalian development will greatly increase our understanding of the functional role of the H3K4 and H3K36, and other histone residues. In the following, I will summarize the roles that have been attributed to the methylation of H3K4 and H3K36, respectively.

H3K4 methylation

H3K4 methylation is an evolutionarily conserved histone modification that marks transcription start sites (TSS) and promoters (Hyun et al., 2017; Shilatifard, 2012). In yeast, the only H3K4 Histone methyltransferase (HMTase) is Set1, which acts as part of the multi-subunit COMPASS complex (Briggs et al., 2001; Miller et al., 2001; Roguev et al., 2001). Set1 and its H3K4 methyltransferase activity are conserved from yeast to humans, but animals carry multiple homologs of this protein. *D. melanogaster* has three Set1 homologs (Set1, Trx and Trr), and mammals harbor six homologs (Set1a, Set1b, Mll1, Mll2, Mll3, and Mll4), which form different complexes with non-redundant functions (Hyun et al., 2017; Shilatifard, 2012). Among these, mammalian Set1a and Set1b are the orthologs of yeast Set1 and can produce all three H3K4 methylation states,

Mll1/2 can catalyze H3K4 mono- and di-methylation, whereas the activity of Mll3/4 is restricted to H3K4 mono-methylation (Shinsky et al., 2015). Mutations in H3K4 HMTases are highly associated with the development of cancer (Rao and Dou, 2015) and various translocations of the Mll gene underlie a variety of lymphoid tumors, as well different types of leukemia in children and adults (Shilatifard, 2012; Tenney and Shilatifard, 2005).

H3K4 methylation is associated with regulatory regions of the genome and the degree of methylation differs depending on the genomic site. H3K4me1 is highly enriched at enhancers, H3K4me2 is highest toward the 5' end of transcribed genes (Kim and Buratowski, 2009), and H3K4me3 is a hallmark of promoter (Heintzman et al., 2007). In yeast, H3K4 methylation is deposited downstream of transcription and is mediated by the recruitment of the HMTase Set1 to the 5' end of the coding region by the transcribing Pol II (Ng et al., 2003; Haberle and Stark, 2018). Thus, H3K4me3 only marks promoters of actively expressed genes in yeast (Santos-Rosa et al., 2002), and it has been suggested to provide a memory mark of recent transcriptional activity, and to facilitate new rounds of transcription (Ng et al., 2003). Instead, the mammalian Set1 complex is targeted to many promoters by binding to unmethylated CpG islands through its Cfp1 subunit (Clouaire et al., 2014, 2012; Brown et al., 2017). In animals, H3K4me3 marks both inactive and active promoters (Barski et al., 2007; Guenther et al., 2007; Mikkelsen et al., 2007), suggesting that H3K4 methylation alone is not sufficient to activate transcription. The co-occurrence of H3K4 methylation with H3K27 acetylation distinguishes active promoters (Wang et al., 2008) and enhancers (Rada-Iglesias et al., 2011; Bonn et al., 2012; Creighton et al., 2010) from their disengaged state in mammals and *Drosophila*. Conversely, at poised promoters H3K4me3 is found together with PRC2-mediated H3K27me3, which keeps developmental genes in a repressed state primed for fast activation during development (Mikkelsen et al., 2007). Although H3K4me3 and H3K27ac correlate with transcriptional activity (Wang et al., 2013; Barski et al., 2007; Guenther et al., 2007; Heintzman et al., 2007), whether they are causally involved in mammalian transcription is not clear.

H3K36

Similar to H3K4, methylation of H3K36 is highly conserved from yeast to humans (Wagner and Carpenter, 2012). Set2 is the sole H3K36 methyltransferase in yeast, which catalyzes all three methylation states (Kizer et al., 2005). Mammalian cells contain at least eight different H3K36 HMTases and among these Setd2, Nsd1, Nsd2 and Nsd3 are considered to catalyze the majority of H3K36 methylation (Hyun et al., 2017). Only Setd2 can produce H3K36 tri-methylation, whereas the other enzymes including Nsd1/2/3 are restricted to mono- and/or di-methylation (Edmunds et al., 2008). Mutations of H3K36 HMTases have been reported in numerous human diseases, such as Sotos syndrome (Kurotaki et al., 2002; Douglas et al., 2003), Wolf-Hirschhorn

syndrome (Stec et al., 1998; Nimura et al., 2009) and blood cancer (acute myeloid leukaemia) (Jaju et al., 2001).

Genes display a progressive shift from di- to tri-methylation of H3K36 between their promoters and 3' end (Wagner and Carpenter, 2012; Bannister et al., 2005) and various functions have been attributed to H3K36 methylation. In yeast, Set2 is recruited co-transcriptionally by Pol II (Li et al., 2003; Krogan et al., 2003) and deposits H3K36 methylation during transcriptional elongation, resulting in high levels of H3K36me3 in gene bodies of actively transcribed genes (Kizer et al., 2005; Li et al., 2003, 2002). Because of its association with active transcription, H3K36me3 is generally viewed as an activating mark. However, in other cases it has been reported to recruit repressors of transcriptional elongation to specific target genes (Wen et al., 2014).

Another reported function of H3K36 methylation is the suppression of cryptic transcription initiation. In yeast, H3K36 methylation (H3K36me1/me2) by Set2 recruits histone deacetylases to coding regions to suppress transcription initiation inside genes (Carrozza et al., 2005). Similarly, associations of mammalian Set2 homologs with elongating Pol II have also been observed (Yuan et al., 2009; Sun et al., 2005). In mammalian cells, H3K36 methylation has been reported to prevent cryptic transcription initiation either by promoting the demethylation of H3K4 (Fang et al., 2010) or by recruiting DNA methyltransferase Dnmt3a to gene bodies (Dhayalan et al., 2010).

Furthermore, low rates of histone exchange have been observed inside coding regions (Dion et al., 2007; Kraushaar et al., 2013; Rufiange et al., 2007; Deal et al., 2010) and methylation of H3K36 in newly deposited histones has been reported to suppress histone exchange inside genes of the yeast genome (Smolle et al., 2012). This suggests that H3K36 modifications could be important for recycling of old histones during transcriptional elongation.

Furthermore, H3K36 methylation has an antagonistic relationship with H3K27 methylation (Young et al., 2011), and methylation of H3K36 in a nucleosome usually precludes deposition of H3K27 methylation (Yuan et al., 2011). This inhibition of PRC2 by H3K36me2/me3 is conserved in mammals and flies, and could restrict the expansion of PRC2-mediated repressive H3K27me3 to transcriptionally active chromatin and other regions (Schmitges et al., 2011).

In animals, exons are generally marked by higher H3K36me3 levels than introns (Kolasinska-Zwierz et al., 2009; Wilhelm et al., 2011) and it has been repeatedly suggested that H3K36 methylation by Setd2 regulates alternative splicing (Luco et al., 2010; Kim et al., 2011). Evidence for this comes from observations that Setd2 depletion causes deregulation of exon exclusion (Luco et al., 2010). Furthermore, splicing events result in enhanced recruitment of Setd2 to genes, which is in line with observations that intron-less genes have generally lower H3K36me3 levels than intron-containing genes (de Almeida et al., 2011).

Together, this demonstrates the need of directly testing the function attributed to hi-

stone residues, rather than relying purely on correlation or indirect interference from phenotypes of mutant enzymes that catalyze modifications. In this chapter, I will describe the characterization of mouse ESCs that carry homozygous point mutations in histone H3.3 at lysine 4 and 36 (H3.3K4A/K36A). I aim to test the impact of the mutation on transcriptional regulation and splicing in pluripotent ESCs and neuronal development by RNA-Seq. Second, I profile how lysine residues in H3.3 contribute to H3.3 deposition and histone turnover at distinct chromatin types, such as active genes and regulatory regions. Last, I address how the epigenetic landscape of mouse ESCs is perturbed by H3.3 mutations using antibody-based approaches such as immunoblot and CHIP-Seq and antibody-independent approaches such as mass spectrometry.

3.2 Results

I conducted the experiments and analyzed the data described in this chapter, unless indicated otherwise. Nichole Diaz repeatedly assisted me with cell culture work and she performed the growth assay of mutant and wild type ESCs in Fig. 3.3a. Marlena Lübke, a master student, who conducted her master thesis under the supervision of Kyung-Min Noh and myself, performed immunofluorescence in differentiated neuron cultures and the images are depicted in Fig. 3.5b. She also collected the data for the cell cycle analysis in Fig. 3.4. Simone Sidoli performed mass spectrometry measurements of histone modifications and analyzed the data in Fig. 3.18. Daria Bunina performed the clustering analysis of combined ChIP-Seq and RNA-Seq data and generated the heatmap displayed in Fig. 3.21a).

3.2.1 Characterization of H3.3K4A and H3.3K36A mutant ESCs and their developmental potential

In ESCs, H3.3 provides approximately 25% of the entire H3 pool, but it accumulates in cells of the brain throughout development and into adulthood, replacing canonical H3 in neuronal nucleosomes (Pina and Suau, 1987; Maze et al., 2015). Thus, I expected that H3.3 mutations would have a severe impact in post-mitotic neurons but not in rapidly-dividing, pluripotent ESCs. I characterized the control and K4A/K36A mutants at the ESC stage and I differentiated them into neural precursor cells (NPCs) and post-mitotic glutamatergic neurons to investigate whether H3.3 mutants can give rise to specialized cell types (Fig. 3.1, overview of differentiation protocol in Fig. S2).

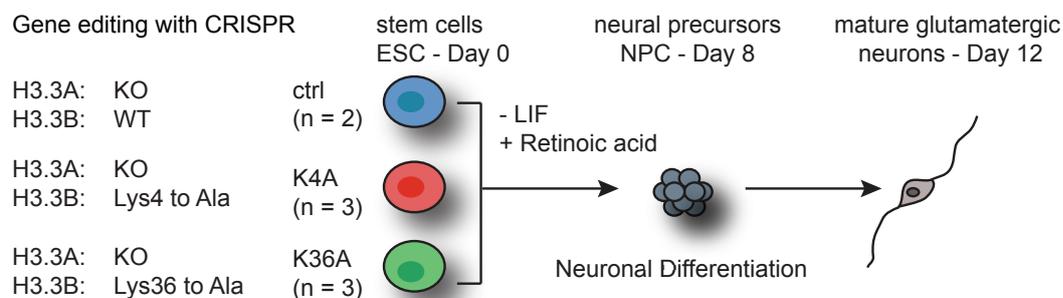


Figure 3.1: Experimental design to study the role of histone H3.3K4 and H3.3K36 residues. Three independent clonal lines with a homozygous *H3f3a* gene knock-out and homozygous *H3f3b* mutations of lysine 4 or 36 to alanine, along with two independent clonal lines with a homozygous *H3f3a* knock-out as controls are established using CRISPR-Cas9. Control and mutant ESCs are differentiated by formation of embryoid bodies (EBs) into neural precursors (Day 8), then further into glutamatergic neurons (Day 12) according to a modified protocol by Bibel et al. (2007).

At the ESC stage, K4A and K36A mutants grown on feeder layer of mouse embryonic fibroblasts (MEFs) displayed normal morphology and grew in round, dense colonies comparable to controls (Fig. 3.2). When ESCs were removed from the feeder layer, K4A mutants displayed a minor morphology change and cells were slightly flatter and larger than those of controls and K36A cells, which grew more densely packed.

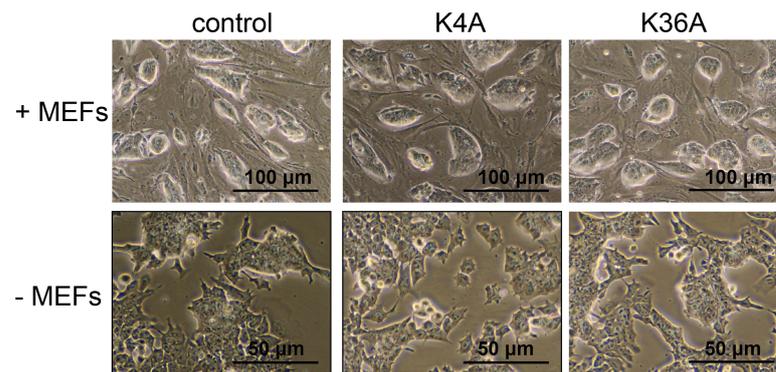


Figure 3.2: Morphology of control and K4A/K36A mutant ESCs. ESCs of control and K4A/K36A mutants were grown either on feeder layer of MEFs (+ MEFs) or as feeder-free culture on gelatine (- MEFs).

K4A and K36A clones showed comparable growth rates with wild type and control clones at the ESC stage (Fig. 3.3a). Cell cycle analysis revealed that ESCs are highly proliferative with approximately 70% of all cells residing in S-Phase (Fig. 3.3b). K4A and K36A mutants had comparable, but slightly elevated proliferation rates compared to controls. This suggests that mutants displayed no cell cycle arrest and that self-renewal of ESCs was maintained.

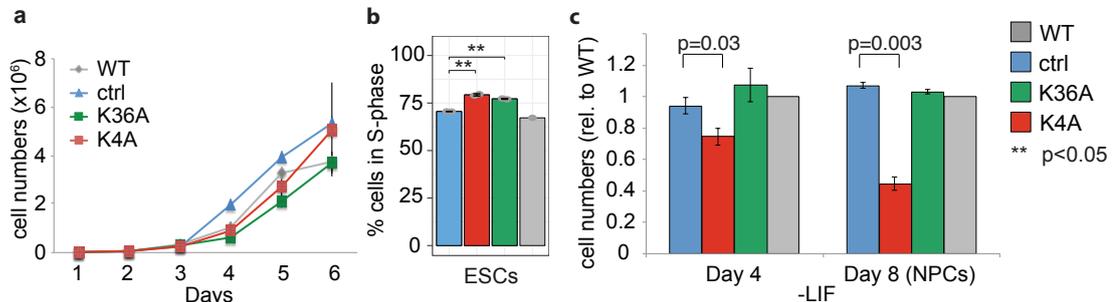


Figure 3.3: Self-renewal of K4A/K36A mutant ESCs is maintained but K4A results in slower cell growth during neuro-differentiation. (a) Cell growth assay of wild type, control and mutant ESCs. Error bars indicate standard deviation of biological replicates; n=3. Wild type are ESCs without CRISPR targeting. (b) Cell proliferation analysis of control and mutant ESCs. The percentage of proliferating cells residing in S-Phase were measured in ESCs by detecting the incorporation of labeled nucleotides (EdU) during DNA synthesis followed by flow cytometry analysis. Displayed are the percentages of cells detected in S-Phase (EdU+). Flow cytometry analysis was performed according to gating strategy in Fig. S3. Biological replicates were n=2, except for wild type (n=1). Unpaired t-test was used to calculate significance. (c) Quantification of cell numbers obtained from EBs on differentiation day 4 and 8 of wild type, control, K36A and K4A cells. Error bars indicate standard deviation of biological replicates. Unpaired t-test was used to calculate significance. Biological replicates were n=2 (control), n=3 (K4A or K36A).

Once neuronal differentiation into NPCs was induced, K4A mutants displayed significantly reduced cell numbers by 25% on day 4 and by 50% on day 8 compared to controls and K36A mutants (Fig. 3.3c). Cell cycle analysis revealed that NPCs are less proliferative than ESCs and only 20% of cells reside in S-phase (Fig. 3.4). Whereas

K36A mutant NPCs displayed comparable proliferation rates to controls, the K4A mutant NPCs had a higher frequency of proliferating cells in the population, which could indicate that K4A mutants were slower at exiting the cell cycle during neuronal differentiation compared to controls. Alternatively, this data can also suggest that K4A mutants have a prolonged S-phase and require longer time to replicate their DNA. It will require further investigation to test if a prolonged S-phase caused the strongly reduced cell numbers of the K4A mutants at the NPC stage, or whether other factors, such as increased apoptosis, contributed to this.

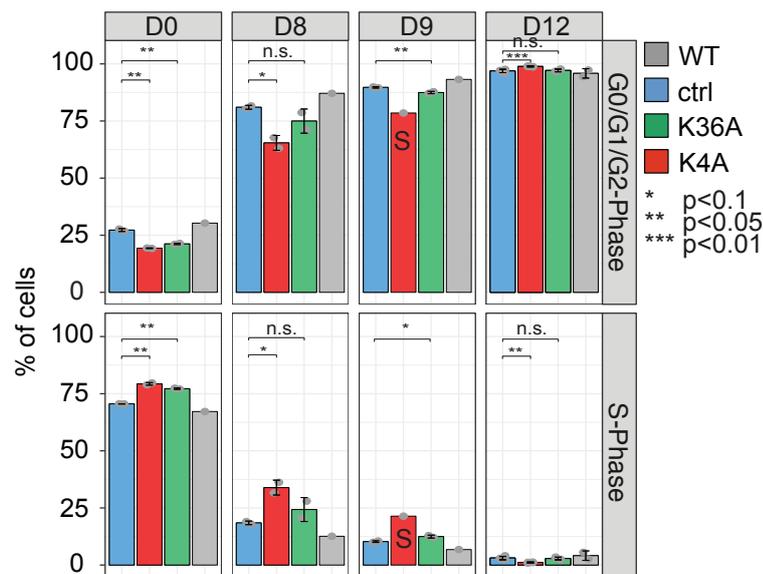


Figure 3.4: Cell cycle analysis of H3.3K4A/K36A mutants in neuronal development. The percentage of proliferating cells residing in S-Phase were measured in ESCs and during neuronal differentiation by detecting the incorporation of labeled nucleotides (EdU) during DNA synthesis followed by flow cytometry analysis. Analyzed time points include pluripotent ESCs, NPCs (D8), immature neurons (D9) and mature glutamatergic neurons (D12) and displayed are the percentages of cells detected either in S-Phase (EdU+), G0/G1/G2/M Phase (EdU-). Flow cytometry analysis was done according to gating strategy in Fig. S3. Biological replicates were n=2, except for wild type (WT) cells and for K4A cells on day 9 (n=1, indicated by "S" in plot).

Upon further neuro-differentiation into glutamatergic neurons, K4A cells formed a heterogeneous population consisting of non-neural cells and neuron-like cells that differ from control neurons both in their reduced interconnectivity, and viability (Fig. 3.5a). K4A neurons showed less expression of mature neuronal marker (Map2) at dendritic and axonal structures, and lacked structural polarity (Fig. 3.5b). We also noted densely packed areas of the cells expressing neuronal stem cell marker (Sox2), but not Map2, suggesting that these cells were immature progenitor cells (Fig. 3.5b). In contrast, K36A cells differentiated into neurons with normal morphology, forming Map2 positive axonal networks, but at a higher cell density than controls (Fig. 3.5a,b). Cell cycle analysis revealed that the fraction of proliferating cells has further decreased to 10% in immature neurons (D9) and to less than 5% in mature glutamatergic neurons (D12), supporting that 12 days of neuronal differentiation were sufficient to obtain

post-mitotic neurons that exit the cell cycle and enter quiescent G0 phase (Fig. 3.4). In immature neurons (D9), K36A mutants displayed slightly elevated levels of proliferating cells, which could contribute to the observed higher cell density of the neuron culture. At the final neuron stage of differentiation (D12), the majority of K4A and K36A mutants have exited the cell cycle comparable to controls and reside in quiescent G0 phase (Fig. 3.4). Overall, morphological differences of derived neurons and reduced cell numbers suggest that neuro-development of K4A mutants is impaired, and that defects already occurred before commitment to the neuronal lineage. Thus, the majority of H3.3K4A cells, but not K36A cells, failed to undergo cell lineage specification required for the formation of glutamatergic neurons.

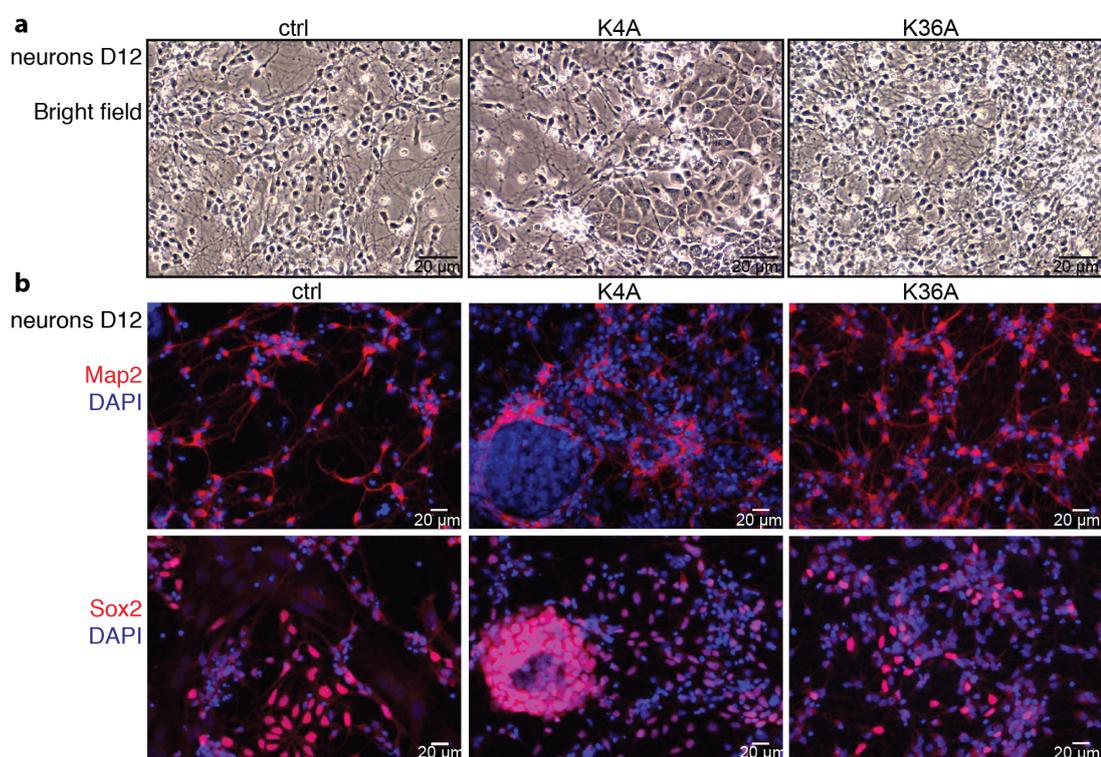


Figure 3.5: H3.3K4A, but not K36A, results in erratic morphology during neuro-differentiation from ESCs. (a) Representative bright field images of neuron populations on day 12 of differentiation. (b) Merged immunofluorescence images of neuron populations on day 12 stained with antibodies against Map2 or Sox2 and DAPI to detect nuclei. Map2 stains axonal and dendritic structures of mature neurons and Sox2 stains neural stem cells

3.2.2 Impact of H3.3K4A and H3.3K36A mutations on gene expression in ESCs and neurons

To investigate the effects of the H3.3 mutants on transcription during differentiation, I used mRNA-sequencing at the ESC and neuron stage. Hierarchical clustering revealed altered gene expression profiles for H3.3K36A mutants only after differentiation into neurons, but for H3.3K4A mutants already in ESCs (Fig. 3.6a). The H3.3K36A mutation caused relatively small gene expression changes in ESCs (409 differentially

expressed genes (DEGs) in ESCs, $FDR < 0.05$, Fig. 3.6b), which indicates that the H3.3K36A has only a minor impact in ESCs. Notably, H3.3K4A mutants exhibit widespread gene expression changes in both ESCs and neurons (2766 DEG in ESCs, 4969 DEG in neurons, $FDR < 0.05$, Fig. 3.6b).

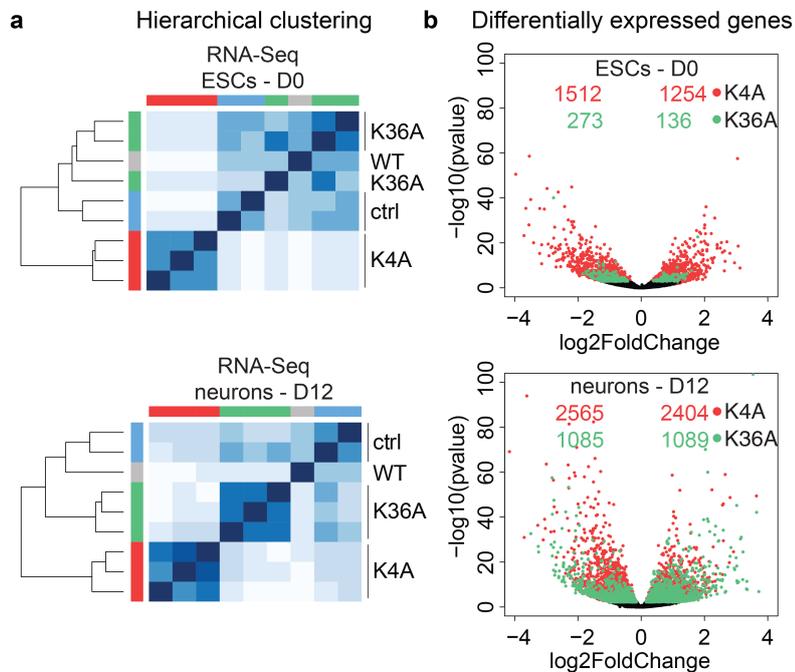


Figure 3.6: H3.3K4A and K36A induce distinct changes in transcription in ESCs and neurons. (a) Heat map representation of sample relationships of \log_2 -transformed gene counts obtained from RNA-Seq data obtained from ESCs (top) and differentiated neurons (bottom). Dendrograms are derived from unsupervised hierarchical clustering and are indicative of the similarities between samples. (b) Volcano plot of differentially expressed genes (DEG) for H3.3K4A and H3.3K36A mutants compared to controls at the ESCs (top) and neuron stage (bottom). Significant DEG with $FDR < 0.05$ are marked by color and the total number of significant up- and down-regulated genes are displayed for each condition. Biological replicates were $n=2$ (controls) or $n=3$ (K4A and K36A).

Gene ontology (GO) analysis of DEGs in ESCs showed that down-regulated genes in K4A mutants were enriched for genes associated with several developmental pathways, including brain development, angiogenesis and heart morphogenesis (Fig. 3.7a). Up-regulated genes were associated with stem cell maintenance and response to leukemia inhibitory factor LIF, but also with meiosis-related functions (spermatogenesis, oogenesis) and developmental processes. Furthermore, analysis of the expression levels of pluripotent factors revealed that many pluripotency factors including *Klf4*, *Esrrb*, *Oct4* and *Sox2*, but not *Nanog*, were significantly higher expressed in K4A mutants, but their expression was unaffected in K36A mutants (Fig. 3.7b). RNA-Seq analysis suggests that the developmental potential of K4A ESCs is affected, possibly due to elevated expression levels of pluripotency factors and widespread gene expression changes, including the misregulation of genes linked to neuronal development.

In agreement with the expectations of an increasing importance of H3.3 in neurons, the H3.3K36A mutation caused greater gene expression changes in neurons than in

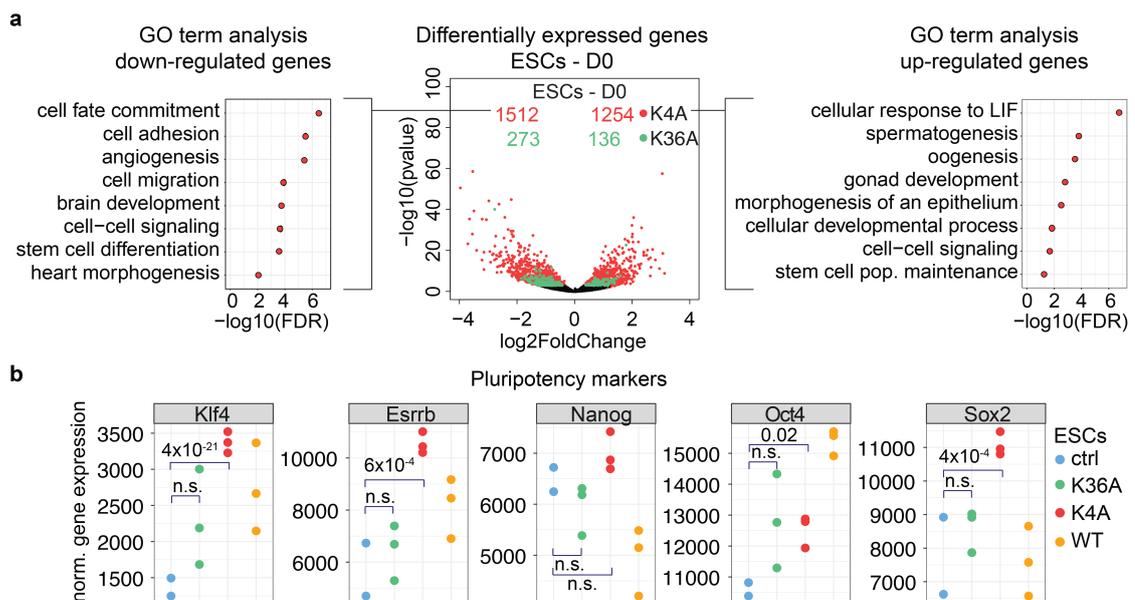


Figure 3.7: Gene expression analysis of H3.3K4/K36A mutant ESCs. (a) Some of the top most enriched GO terms within biological process categories for down-regulated genes (left) or up-regulated genes (right) in H3.3K4A ESCs ($\log_2\text{Fold change} < -0.58$, $\text{FDR} < 0.05$). Analysis was performed using the topGO package (Bioconductor). Significant GO terms were selected using Fisher classic and Fisher elim algorithm and depicted FDR values are derived from Fisher elim test. (b) Normalized RNA-Seq counts representing gene expression for pluripotency stem cell markers (Klf4, Esrrb, Nanog, Oct4 and Sox2) in controls, H3.3K36A, H3.3K4A and wild type (WT) ESCs. Depicted p values for differentially expressed genes were calculated with DESeq2 and adjusted using the Benjamini Hochberg's method (padj).

ESCs (409 DEGs in ESCs, 2174 DEG in neurons, $\text{FDR} < 0.05$, Fig. 3.6b). In neurons, up-regulated genes were enriched with GO terms related to neuron maturation such as axonogenesis, whereas down-regulated genes were related to proliferation, cell death, and mature neuronal signaling pathways (Fig. 3.8b). This indicates that K36A mutant form more immature neurons than control cells, resulting in higher cell density and significantly lower expression of mature glutamatergic markers (Fig. 3.8c). These results also indicate that the H3.3K36A mutation induces aberrant expression of genes that are necessary for development into post-mitotic glutamatergic neurons, but has only a minor impact in ESCs.

H3.3K4A cells displayed gene expression changes superficially similar to controls during development, overall expressing neurodevelopmental genes and repressing pluripotency genes, but with detailed transcriptional regulation clearly perturbed. Specifically, H3.3K4A neurons showed higher expression levels of genes that should have been repressed during differentiation relative to controls (e.g., genes linked with endoderm/mesoderm development and mitosis/meiosis, Fig. 3.8a and Fig. 3.9a,c), and higher levels of neural stem cell markers comparable to those seen in the neural precursor state of wild type cells (Fig. 3.8c). In addition, we observed atypically low expression of genes involved in cell signaling and ectoderm development (neuronal lineage) (Fig. 3.8 and Fig. 3.9a,c) and mature glutamatergic markers (Fig. 3.8c). Thus,

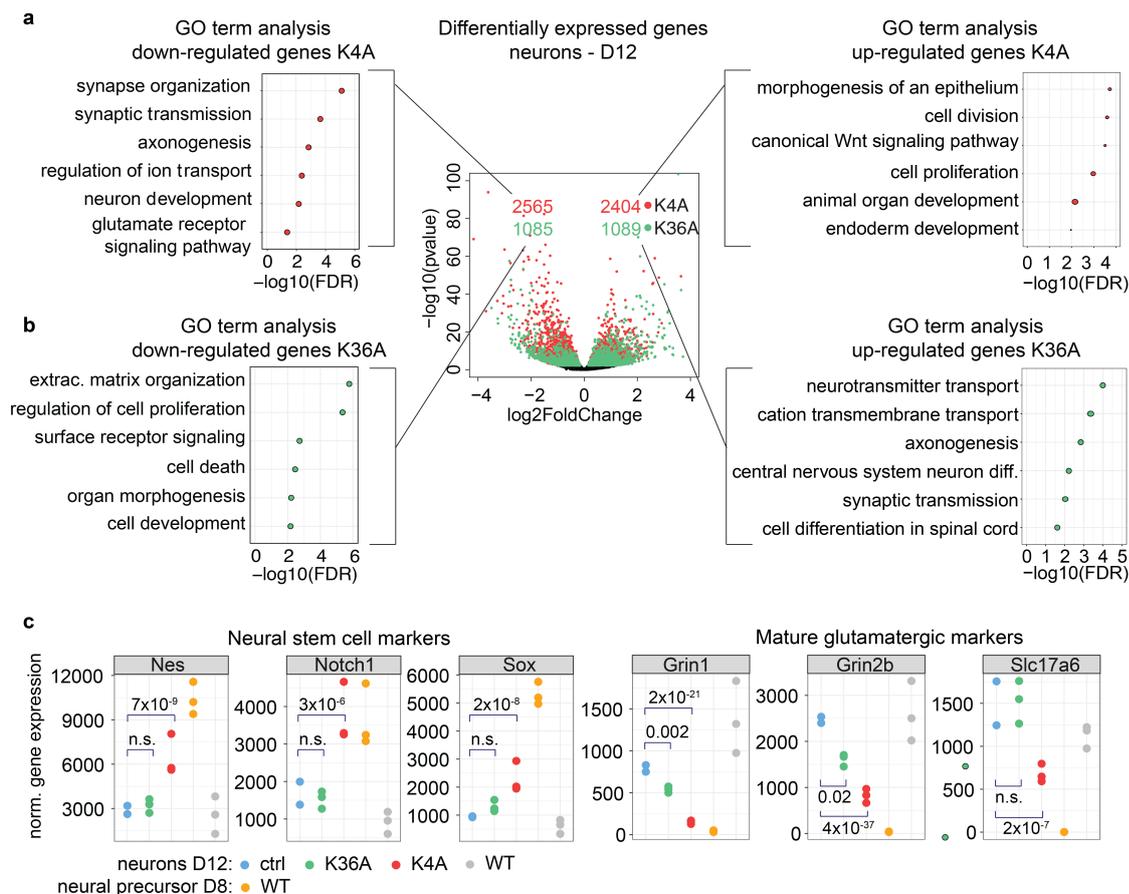


Figure 3.8: Gene expression analysis of H3.3K4/K36A mutant neurons. (a) Some of the top most enriched GO terms within biological process categories for down-regulated (left) or up-regulated genes (right) in H3.3K4A neurons on day 12 (\log_2 Fold change < -0.58 , FDR < 0.05). Analysis using the topGO package (Bioconductor). Significant GO terms were selected using Fisher classic and Fisher elim algorithm and depicted FDR values are derived from Fisher elim test. (b) GO term analysis as in (a) but for H3.3K36A mutant neurons. (c) Normalized RNA-Seq counts representing gene expression for immature neural stem cell markers (Nestin, Notch1, Sox2) and mature glutamatergic neuron markers (Grin2b, Slc17a6, Grin1) of controls, H3.3K36A, H3.3K4A and wild type (WT) neurons in comparison to WT neural precursors derived on day 8. Depicted p values for differentially expressed genes were calculated with DESeq2 and adjusted using Benjamini Hochberg's method (padj).

H3.3K4A mutants seem to suffer a developmental arrest, suggesting that the H3.3K4 residue is required to maintain gene expression levels in stem cells that in turn coordinate appropriate lineage specifications, such as development into glutamatergic neurons. A study using H3.3 knockdown (Banaszynski et al., 2013) reported less pronounced transcriptome changes in ESCs than observed in K4A cells, implying that a H3.3K4A mutation may induce more extensive changes in the chromatin environment than expected for a loss-of-function mutation.

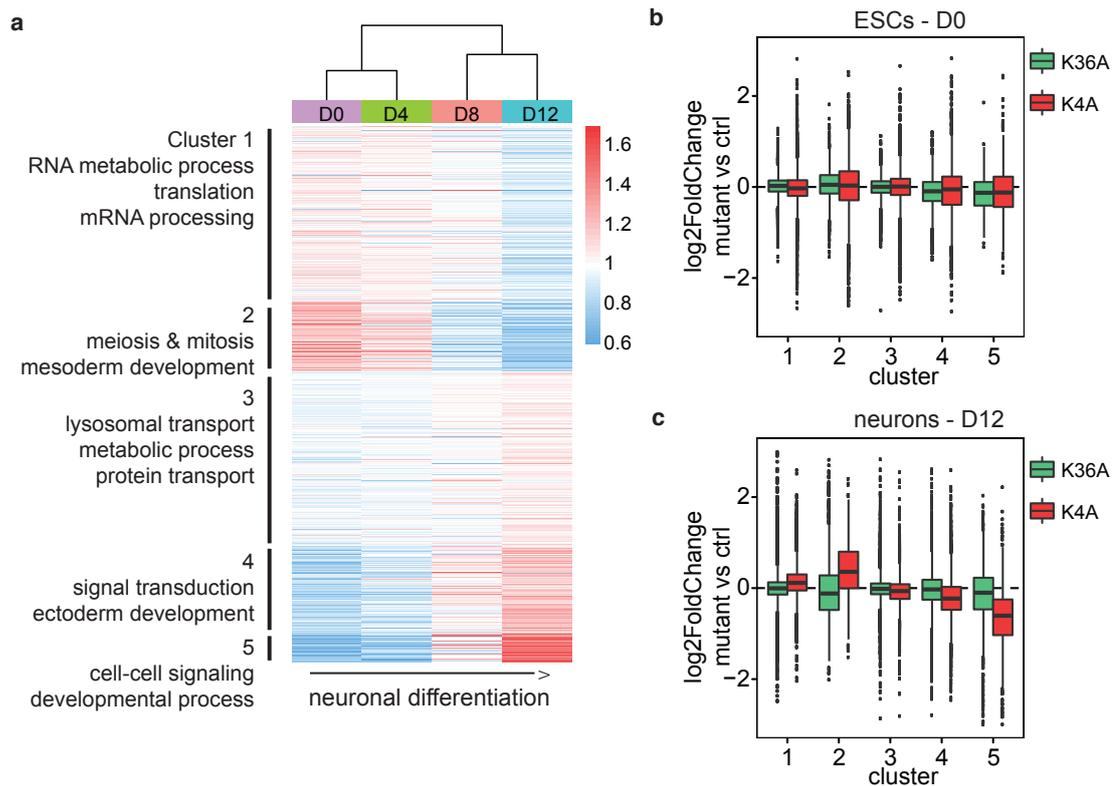


Figure 3.9: H3.3K4A affects the activation of developmental genes and repression of pluripotency genes. (a) Heatmap of RNA-Seq data displaying gene expression changes during neuronal development of wild type ESCs into glutamatergic neurons. All genes were sorted into five clusters (k-means clustering) according to their expression during neurodevelopment (D0, D4, D8, D12). Most enriched biological GO terms for each cluster are indicated and were identified using DAVID database. (b,c) Gene expression changes of H3.3K4A/K36A ESCs (b) and neurons (c) compared to controls were plotted per gene cluster defined in (a). Positive values represent overall higher expression of genes inside the cluster compared to control cells, negative values represent lower expression of genes inside the cluster compared to control cells.

3.2.2.1 Genomic enrichment of wild type and mutant H3.3 determined by H3.3-ChIP-Seq

H3.3 enrichment

H3.3 is expressed and assembled into nucleosomes throughout the cell cycle and replaces replication-dependent canonical H3 in nucleosomes at regions with high histone exchange, such as regulatory regions of the genome and actively transcribed genes. As a consequence, H3.3 becomes relatively enriched at these regions compared to intergenic sites in ESCs (Goldberg et al., 2010). I mapped H3.3 enriched loci in the ESC genome using chromatin immunoprecipitation coupled with DNA sequencing (ChIP-Seq) to address if histone mutation influence the localization of H3.3 at distinct genomic regions. Interestingly, average ChIP-Seq profiles for control and mutant H3.3 at annotated genes revealed a significant depletion of the H3.3K4A protein from the TSS, but not from gene bodies and transcriptional end sites (TES) (Fig. 3.10a,c). The most pronounced depletion of H3.3K4A was observed at active TSS (Pol II occupied),

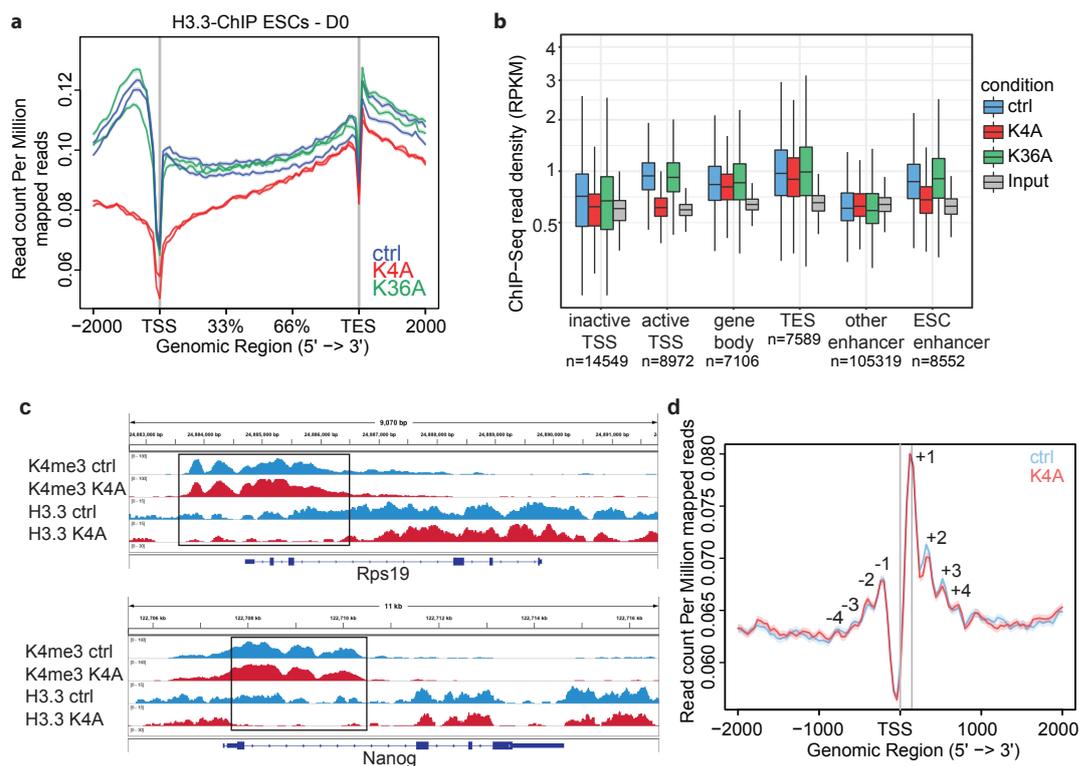


Figure 3.10: H3.3K4A, but not K36A, is depleted at TSS and enhancers in ESCs. (a) Density profile of H3.3-ChIP-Seq reads over annotated genes in H3.3 control and K4A/K36A ESCs. Reads per million mapped reads were plotted over TSS, scaled gene bodies and TES using NGSplot (Shen et al., 2014). $n=2$ biological replicates. (b) Analysis of H3.3 enrichment in control, H3.3K4A and H3.3K36A ESCs at indicated genomic regions. H3.3-ChIP-Seq reads per kilobase per million mapped reads (RPKM) were calculated for the following features: inactive TSS (± 1.5 kb), active TSS (± 1.5 kb); gene body and TES (± 1.5 kb) of active genes; ESC-specific enhancers and enhancers of other tissue-types. Enhancers and activity status of TSS based on Pol II occupancy were predicted by Shen et al. (2012) Number of analyzed genomic features (n) are depicted underneath boxplot. RPKM levels higher than those of input indicate enrichment of H3.3. ChIP $n=2$, Input $n=3$ biological replicates. (c) ChIP-Seq tracks of representative genes that display depletion of H3.3K4A mutant around the TSS, marked by H3K4me3. Black boxes indicate the depleted region. (d) Nucleosome positioning profile around the TSS in H3.3 control and H3.3K4A ESCs. MNase-digested DNA was prepared for MNase-Seq and reads per million mapped reads were plotted over TSS of all mm10 genes using NGSplot (Shen et al., 2014). Typical nucleosome pattern (150 bp) is visible.

inactive TSS (Pol II unoccupied) and predicted ESC-specific enhancers (Shen et al., 2012), compared to wild type H3.3 (Fig. 3.10b). Inside gene bodies and around TES of active genes, H3.3 was enriched independent of the mutation (Fig. 3.10b). Mapping the positioning of nucleosomes using MNase-seq showed only a slight reduction in total nucleosome density around TSS in K4A cells (Fig. 3.10d), suggesting that H3.3K4A depletion was mostly compensated by nucleosomes containing canonical H3.1/H3.2. To confirm the TSS-specific depletion of H3.3K4A observed by ChIP-Seq, enrichment of wild type and mutant H3.3 was measured at the TSS and TES in ESC lines expressing a SNAP-tagged version of H3.3 (Deaton et al., 2016). For this, H3.3-SNAP expression was induced with doxycycline from an inducible promoter followed by pulse-labeling of the available protein with SNAP-biotin (Fig. 3.11a). ChIP-qPCR

for labeled H3.3 showed again that H3.3K4A was depleted from TSS, but not TES, compared to wild type and H3.3K36A (Fig. 3.11b). A lysine-to-arginine mutation (H3.3K4R) that preserves the positive charge on the mutated residue also showed histone depletion at TSS (Fig. 3.11b), suggesting that the H3.3K4 residue itself, and not its positive charge, is essential to maintain histones at the TSS.

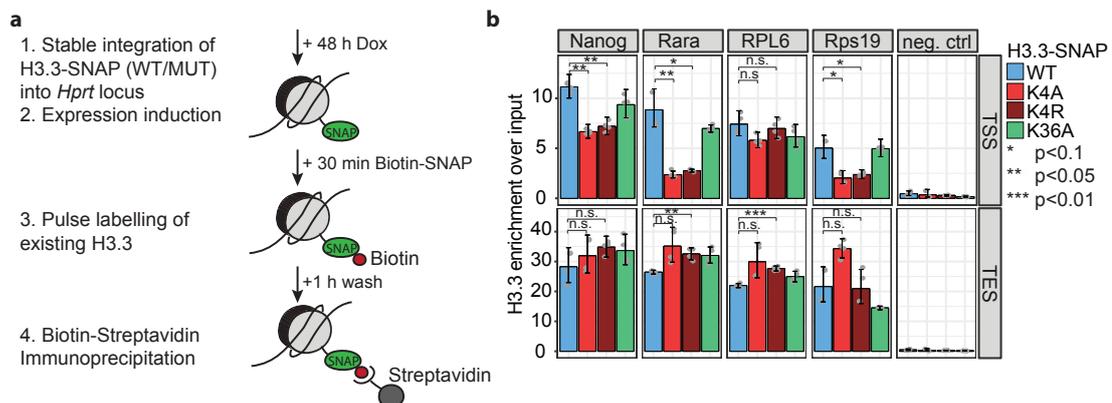


Figure 3.11: H3.3K4A is depleted at TSS, but not TES. (a) Experimental overview for a pulse labeling experiment of H3.3 protein. ESC lines were generated by stably integrating wild type or mutant H3.3 with a C-terminal SNAP-tag into the *Hprt* locus (Deaton et al., 2016; Iacovino et al., 2011). Expression of H3.3-SNAP was induced with Doxycycline (Dox) for 48 hours. Cells were incubated with biotin-SNAP for 30 min to label all available H3.3-SNAP. Cells were washed for 1 hour to remove free biotin. Labeled H3.3 was precipitated with Streptavidin beads. (b) ChIP-qPCR results for H3.3-SNAP Biotin labeling experiment. Relative enrichment over input of H3.3 wild type, K4A, K36A and K4R (lysine to arginine) mutants was measured at TSS and TES of indicated genes, or a gene-free region on chromosome 6 (neg. ctrl). n=2 biological replicates measured in 2 technical replicates. P-values were measured by Student's t-test.

Co-transcriptional H3.3 exchange at coding regions

Histone exchange of H3.3 occurs at both the promoter and the coding regions, but the turnover rates can vary and coding regions undergo less H3.3 exchange than the promoters (Venkatesh and Workman, 2015; Deaton et al., 2016; Kraushaar et al., 2013). The exchange in coding regions of genes occurs in a co-transcriptional manner (Venkatesh and Workman, 2015), and transcriptional elongation by Pol II recruits Set2 and results in higher H3K36me3 levels at these sites (Kizer et al., 2005). Thus, it has been speculated that H3K36me3 in coding regions prevents histone exchange behind transcribing Pol II (e.g. recycling through reassembly or preventing disassembly) to facilitate transcriptional elongation, thereby resulting in lower turnover rates. This prediction was unfulfilled in mutant cells as the H3.3K4A and H3.3K36A mutations did not impact enrichment of H3.3 inside coding regions or at TES as measured by ChIP-Seq and ChIP-qPCR (Fig. 3.10b and Fig. 3.11b). Additionally, I tested whether mutant H3.3 shows more rapid exchange at TES than wild type H3.3 using the previously described SNAP-system (Fig. 3.11a). Wild type and mutant H3.3-SNAP were pulse-labeled in ESCs and enrichment of labeled protein at the TES was quantified after 3 and 6 hours.

Exchange with newly synthesized, unlabelled protein is expected to result in lower H3.3 enrichment as measured by ChIP-qPCR. After 3 hours, H3.3 was still readily detectable at the TES of all tested genes independent of the mutation (Fig. 3.12). After 6 hours, H3.3 was lost from the TES of highly expressed genes Rps19, Slc2a3 and Nanog, whereas TES of developmental genes Rara and Phc1 still maintained enrichment of labeled H3.3. However, I did not observe a significant difference between wild type and mutant H3.3 at both time points. Thus, histone exchange of H3.3 at coding regions can vary from gene to gene, but remains unaffected by a K4A or K36A mutation.

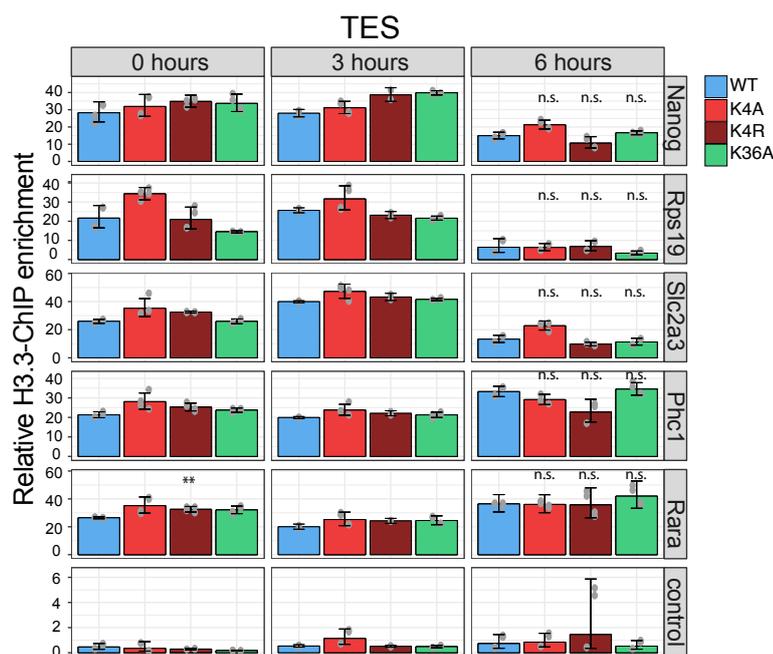


Figure 3.12: H3.3K36 and H3.3K4A are not required for histone turnover at TES. ChIP-qPCR results for H3.3-SNAP Biotin labeling experiment in cells harvested 0, 3, and 6 hours after H3.3 labeling. Relative enrichment over input of H3.3 wild type, K4A, K36A and K4R (lysine to arginine) mutants was measured at TES of indicated genes, or a gene-free region on chromosome 6 (control). Decrease in H3.3 enrichment over time indicates exchange for newly synthesized (unlabeled) histones. Two biological replicates were measured in 2 technical replicates (n=4), except for wild type H3.3 at timepoint 3 hours (n=2). P-values were measured by Student's t-test.

Pre-transcriptional H3.3 exchange at regulatory regions

The results from H3.3-ChIP-Seq analysis suggested that the depletion of H3.3K4A occurs at regulatory regions, where histone exchange occurs pre-transcriptionally, but not inside coding regions. Furthermore, the depletion occurred at both active and inactive promoters and therefore independent of Pol II transcription or occupancy. Factors that contribute to the depletion should therefore be present at regulatory regions, but not gene bodies. Possible explanations for the depletion could be: 1) The H3.3K4A mutation interferes with initial chromatin deposition at regulatory regions by H3.3-specific chaperones, or 2) the H3.3K4A mutation interferes with chromatin remodeling and pre-transcriptional histone exchange at regulatory regions, but not with

co-transcriptional histone exchange inside coding regions. To address the first hypothesis, I looked at the chaperone complexes which are responsible for H3.3 chromatin assembly. The histone chaperone complex Hira is specific for assembly of newly synthesized H3.3 into nucleosomes at both gene bodies and regulatory elements (Goldberg et al., 2010). Structural and biochemical data suggests that a H3.3K4A mutation is unlikely to interfere with Hira function, because it recognizes the four variant amino acids within the H3.3 core (Ahmad and Henikoff, 2002; Ricketts et al., 2015). Furthermore, the enrichment of H3.3K4A at gene bodies and TES remained unaffected (Fig. 3.10b), as opposed to H3.3 depletion at entire genic regions in Hira knockout cells (Goldberg et al., 2010) (Fig. 3.13a).

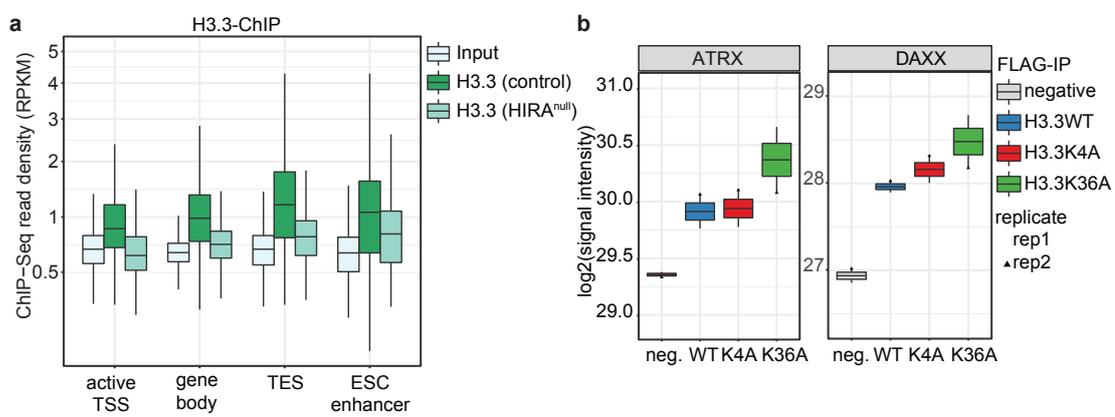


Figure 3.13: H3.3K4A mutation does not affect deposition by chaperones. (a) Analysis of H3.3 enrichment at genic regions and enhancers in control and HIRA knockout (HIRA^{null}) ESCs. ChIP-Seq data was obtained from Goldberg et al. (2010) and sequencing reads per kilobase per million mapped reads (RPKM) were calculated for active TSS (± 1.5 kb); gene bodies and TES (± 1.5 kb) of active genes; and ESC-specific enhancers (Shen et al., 2012). RPKM levels higher than those of input indicate enrichment of H3.3. (b) Interaction analysis between wild type and mutant H3.3 with histone chaperones Daxx and Atrx. ESC lines were generated by stably integrating wild type or mutant H3.3 (K4A/K36A) with a C-terminal HA-FLAG-tag into the Rosa26 locus (Perez-Pinera et al., 2012; Chu et al., 2016). MNase-digested mono-nucleosomes were immunoprecipitated using anti-FLAG-beads for cell lines expressing FLAG-tagged H3.3WT, H3.3K4A, H3.3K36A or in unmodified wild type cells (negative control) and measured by mass spectrometry analysis. Histone chaperones Daxx and Atrx were identified to bind mutant and wildtype nucleosomes and log₂(signal intensities) are displayed. n=2 biological replicates.

The other known chaperone of H3.3 is Daxx, which acts together in a complex with the SWI/SNF-like chromatin remodeling protein ATRX (Drane et al., 2010; Lewis et al., 2010; Xue et al., 2003). The Daxx/ATRX complex mediates H3.3 deposition into repetitive genomic regions such as pericentric DNA and telomeres (Drane et al., 2010; Wong et al., 2010; Lewis et al., 2010), but not regulatory elements or genes (Goldberg et al., 2010), suggesting that Daxx/ATRX complex is not responsible for the observed H3.3K4A depletion. Furthermore, immunoprecipitation coupled with mass spectrometry (IP-MS) analysis showed that both H3.3 wild type and H3.3K4A/K36A mutant nucleosomes were still bound by both Daxx and ATRX (Fig. 3.13b). The chaperone Hira was not detected in this experiment, likely because it binds free histone H3.3 rather than already assembled nucleosomes, which were mostly immunoprecipitated

in this analysis. Together, these results and evidence from previous studies indicates that the H3.3K4A mutation does not impair binding to known chaperons, nor initial chromatin assembly in general.

Addressing the second hypothesis is more difficult, because chromatin remodeling and histone exchange are mediated by many different remodeling complexes. To this end, I identified remodelers that mediate chromatin remodeling pre-transcriptionally (promoters/enhancers), but not co-transcriptionally (gene bodies) in mouse ESCs. Analysis of publicly available ChIP-Seq data for chromatin remodelers in ESCs (de Dieuleveult et al., 2016) showed that specific remodelers are enriched mostly inside coding regions (Chd2, Chd6 and Chd9), whereas others are specifically enriched at promoters and enhancers (Chd1, Chd4, Chd8, Brg1 and Ep400) (Fig. 3.14a). Of the latter category, Ep400, Brg1 and Chd4 displayed a relatively narrow peak at the TSS and $-/+1$ nucleosome (de Dieuleveult et al., 2016), and only the remodeler Chd1 showed a broader enrichment and occupied multiple nucleosomes around the TSS. This enrichment pattern strongly resembled H3.3 WT enrichment around promoters of control cells (Fig. 3.14b). Comparison with the region of the H3.3K4A depletion showed that Chd1 was strongly enriched throughout the entire region of the depletion, but not inside gene bodies that display maintained occupancy by H3.3K4A (Fig. 3.14b). Interestingly, Chd1 has been previously reported as a potential H3.3 chaperone in *Drosophila* (Konev et al., 2007), but no evidence for a similar role in mammals has been reported at this point. Human Chd1 has also been reported to be recruited by binding to methylated H3K4 (Sims et al., 2005). In summary, Ep400, Brg1, Chd4, Chd8 and especially Chd1 are potential candidates that might be responsible for the depletion of H3.3K4A from regulatory regions. Further experiments will be required to address if the mode of action for these remodeling complexes depends on the lysine 4 residue of histone H3 for nucleosome binding.

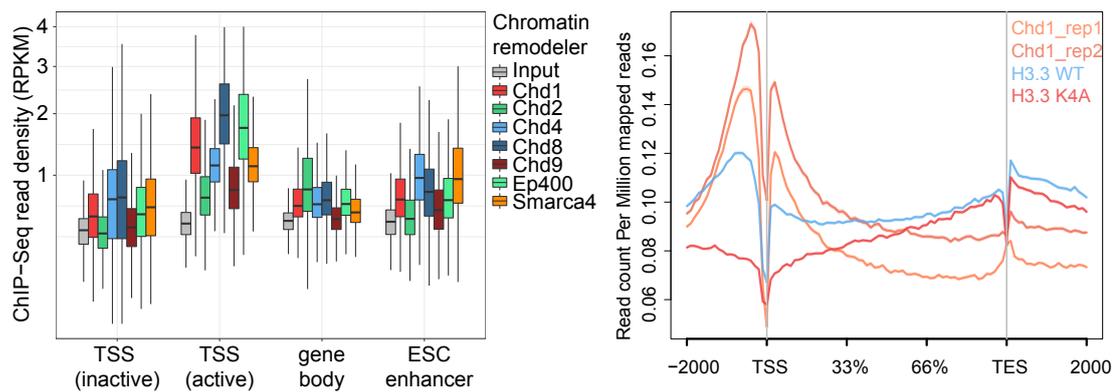


Figure 3.14: Remodelers Chd1, Chd4, Ep400, Smarca4 are enriched at regulatory regions and Chd1 displays a similar enrichment profile to H3.3 at promoters in ESCs. (a) Analysis of chromatin remodeler enrichment in wild type ESCs at indicated genomic regions for Chd1, Chd2, Chd4, Chd8, Chd9, Ep400 and Smarca4. ChIP-Seq data was obtained from de Dieuleveult et al. (2016) and reads per kilobase per million mapped reads (RPKM) were calculated for the following features: inactive TSS (± 1.5 kb), active TSS (± 1.5 kb); gene body of active genes; ESC-specific enhancers. ESC-specific enhancers and activity status of TSS based on Pol II occupancy were predicted by Shen et al. (2012). RPKM levels higher than those of input indicate enrichment of the remodeling complex. Chd1, Chd4, Chd8, Ep400 and Smarca4 are enriched at regulatory regions compared to gene bodies. ChIP n=1-2, Input n=1 biological replicates. (b) Density profile of Chd1 and H3.3-ChIP-Seq reads over annotated genes. Chd1 ChIP-Seq was performed in wild type ESCs and data was obtained from de Dieuleveult et al. (2016) (n=2). H3.3 ChIP-Seq was performed in H3.3 control and K4A ESCs as indicated. Reads per million mapped reads were plotted over TSS, scaled gene bodies and TES using NGSplot (Shen et al., 2014).

Protein stability

The depletion of H3.3 at regulatory regions could suggest that K4A mutant histone is less stable, therefore I investigated whether the H3K4 residue is required for histone stability. Using specific antibodies recognizing either the variant H3.3 (wild type and mutant) or canonical H3.1/H3.2, I tested the abundance of H3 histones during neuronal differentiation (Fig. 3.15). In ESCs, protein levels of H3.3K4A were comparable to the controls implying that the observed depletion at active regulatory regions was not coupled with changes in total protein abundance. Surprisingly, H3.3K4A was less abundant in NPCs and strongly reduced in neurons. Protein abundance of canonical H3.1/H3.2 was not affected by this depletion and levels remained unchanged. H3.3K36A mutant histone was stably expressed throughout differentiation comparable to levels of wild type H3.3 in control cells.

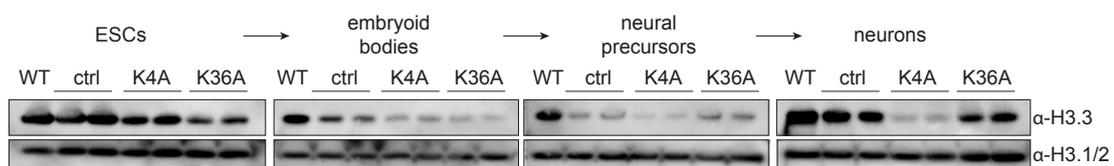


Figure 3.15: Lysine 4 residue is required for histone H3 protein stability in differentiated neurons. Immunoblot of nuclear fraction of wild type, control and H3.3K4A/K36A mutants in ESCs and during neuronal differentiation to detect levels of endogenous H3.3 protein. Embryoid bodies were obtained on day 4, neural precursors on day 8 and neurons on day 12 of differentiation. Antibody is recognizing endogenous H3.3 variant or canonical H3.1/H3.2. WT refers to unmodified ESCs with two H3.3 genes intact; and control, K4A, K36A cells have H3.3A gene removed.

Reduced protein abundance of H3.3K4A in neurons was not due to changes in the RNA expression levels of this gene, as shown from RNA-Seq analysis (Fig. 3.16a). In ESCs, RNA expression of H3.3K4A was significantly up-regulated, suggesting that lower stability of the protein was compensated by higher gene expression and is therefore less obvious as measured by immunoblot (Fig. 3.15). Furthermore, gene expression of the chaperone complexes was maintained in H3.3K4A neurons, suggesting that histone instability was not due to misregulation of chaperone expression (Fig. 3.15b).

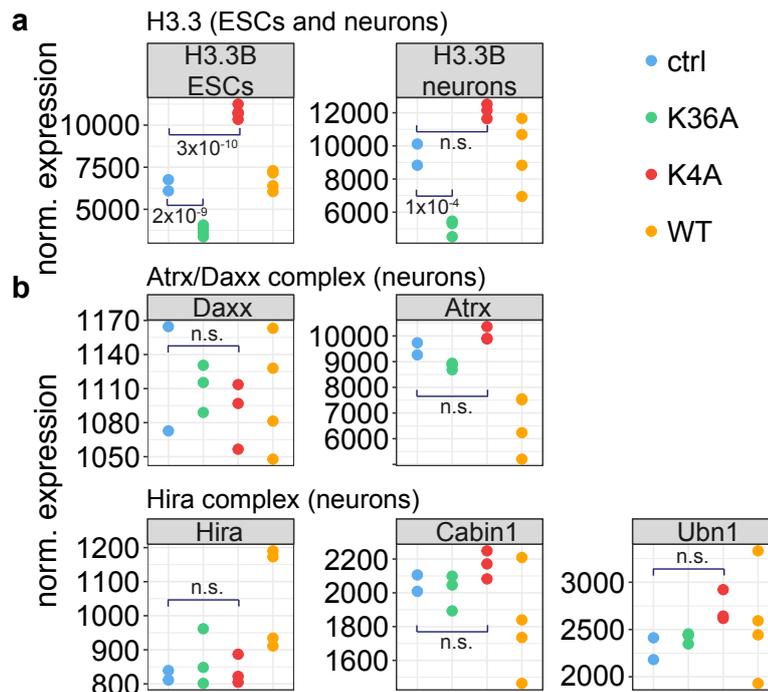


Figure 3.16: H3.3K4A instability is not due to lower expression of H3.3-specific chaperones. (a) Normalized RNA-Seq counts representing gene expression for H3.3B of controls, H3.3K36A, H3.3K4A and wild type (WT) ESCs (left) and neurons (right). Depicted p values for differentially expressed genes were calculated with DESeq2 and adjusted using Benjamini Hochberg's method (padj). Significantly lower gene expression of H3.3 in K36A ESCs and neurons is due to difficulties in mapping reads inside the mutated exon to the wild type reference genome (5 nucleotide changes introduced). (b) Normalized RNA-Seq counts representing gene expression for components of the histone chaperone complexes Atrx/Daxx or Hira (Ubinuclein-1, Hira, Cabin-1) of controls, H3.3K36A, H3.3K4A and wild type (WT) neurons. p values for differentially expressed genes were calculated with DESeq2 and adjusted using Benjamini Hochberg's method (padj) and padj values > 0.05 were considered not significant (n.s.).

Ectopic expression of wild type and mutant H3.3 in human embryonic kidney cells (HEK) 293T cells confirmed lower abundance of H3.3K4A protein, supporting that lower expression was not due to differentiation defects or a developmental delay (Fig. 3.17). Furthermore, cellular fractionation followed by immunoblotting revealed that H3.3K4A, but not H3.3K36A, was degraded selectively in the nuclear chromatin fraction, but did not affect protein stability in the cytoplasm (Fig. 3.17b). Mutants mimicking acetylated lysine (H3.3K4Q) and non-acetylated lysine (H3.3K4R) were also depleted in nuclear but not cytosolic fractions (Fig. 3.17c). Moreover, K4A substitution in the canonical H3 protein (H3.1) also shows degradation comparable to H3.3K4A (Fig.

3.17d). Inhibition of the proteasome by MG-132 treatment rescues protein levels of H3.3K4A to those of wild type H3.3 in whole cell lysates (Fig. 3.17e). However, inhibition of the proteasome in ESCs could not rescue the observed site-specific depletion of H3.3K4A from promoters, as measured by ChIP-qPCR (Supplementary Fig. S4), suggesting that degradation occurs downstream of chromatin eviction. Taken together, these results are consistent with the model that the H3K4 residue itself, and not its charge state, maintains histone H3 protein stability of all H3 variants on chromatin.

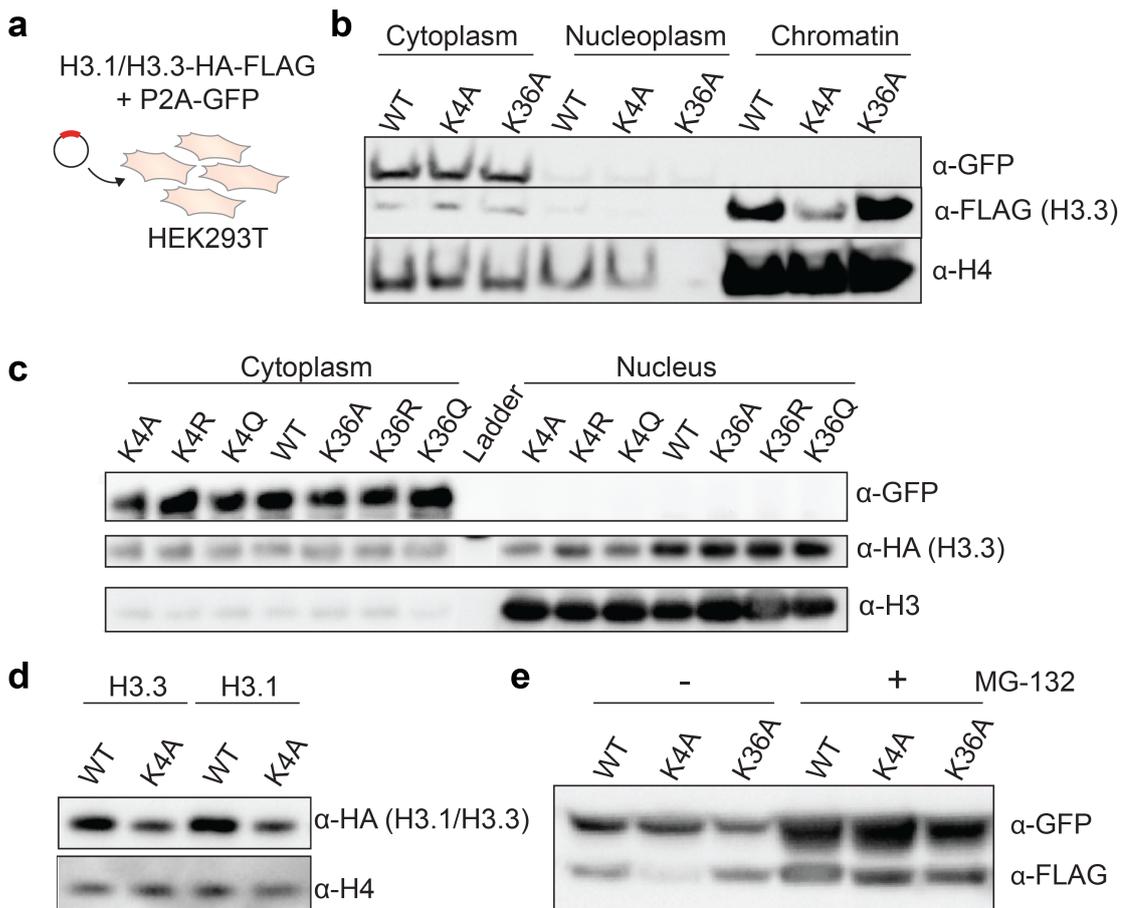


Figure 3.17: Lysine 4 residue is required for histone H3 protein stability in HEK293T cells. (a) Scheme for stability analysis of H3.3 mutants in HEK293T cells. Cells were transfected with H3.3 or H3.1 encoding constructs. H3.3/H3.1 was fused to HA and FLAG-tag for antibody detection and is transcribed together with P2A-GFP. Construct expresses two proteins, mostly nuclear H3.3-HA-FLAG and mostly cytosolic P2A-GFP (expression control). (b) Immunoblots after cellular fractionation of HEK293T cells. HEK293T cells were transfected with either wild type or K4A/K36A mutant H3.3. FLAG antibody was used to detect H3.3 and GFP antibody is used as transfection control. (c) Immunoblots of cytosolic and nuclear fractions of HEK293T cells. HEK293T cells were transfected with either wild type or K4A/R/Q or K36A/R/Q mutant H3.3 (Q for mutation to glutamine and R for mutation to arginine). HA antibody was used to detect H3.3 and GFP antibody is used as expression control. (d) Immunoblots of whole cell lysates of HEK293T cells to measure expression of wild type compared to K4A mutant H3.1/H3.3. (e) Immunoblots of whole cell lysates of HEK293T cells that were either untreated or treated with the proteasome inhibitor MG-132 (5 μ M for 5 hours) prior to harvesting. The expression of wild type compared to K4A mutant H3.3-FLAG was detected with FLAG antibody and antibody against GFP was used as expression control.

3.2.3 The epigenetic landscape of H3.3K4A and H3.3K36A mutants

3.2.3.1 Assessment of global differences in histone modifications by immunoblot and mass spectrometry

Presence of mutant histone H3.3 can alter global levels of histone methylation levels. Especially lysine-to-methionine mutations in a fraction of H3 can act in a dominant manner to deplete global methylation levels on the remaining wild type histones (Herz et al., 2014). Immunoblotting of H3.3K36A mutant ESCs revealed a marked reduction of H3K36me₂ (10-fold), a moderate reduction of H3K36me₃ (1.7-fold), a moderate reduction of H3K27ac (2.6-fold), and an increase of H3K27me₃ (4-fold) (Fig. 3.18a,b). In parallel, we used mass spectrometry to identify changes in the global abundance of other H3 modifications (H3.1/H3.2) and confirmed the concomitant change of H3K36me₂, H3K36me₃ and H3K27me₃ (Fig. 3.18c). This crosstalk between H3K27me₃ and H3K36me₂ is consistent with observed results in mesenchymal tumor cells carrying a H3.3K36M mutation (Lu et al., 2016), which has been shown to result in loss of H3K36 methylation and spread of repressive H3K27me₃ to intergenic regions (Streubel et al., 2018).

In H3.3K4A mutant ESCs, we only detected globally reduced levels of H3K4me₃ (5-fold) by immunoblotting, whereas other tested modifications remain unchanged (Fig. 3.18a,b). Also mass spectrometry analysis detected no significant effect of the H3.3K4A mutation on other histone modifications present on wildtype H3.1/H3.2 (Fig. 3.18c).

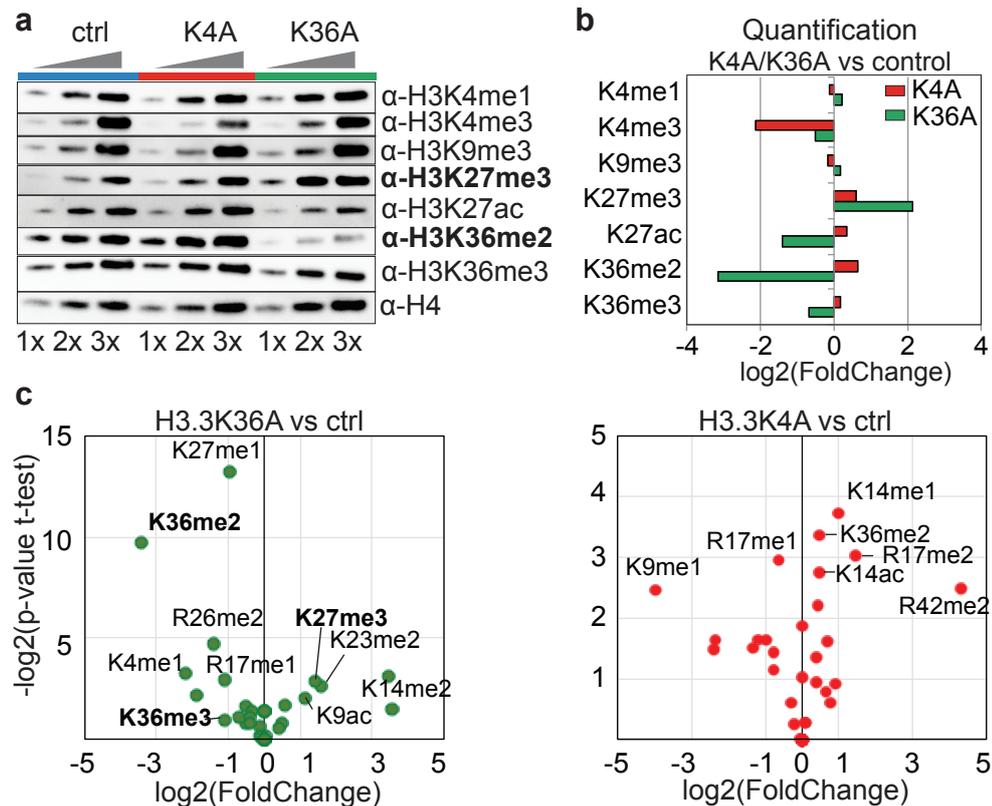


Figure 3.18: H3.3K36A mutation reduces global levels of H3K36me2/me3 and H3K27ac, and inversely increases H3K27me3 levels, whereas H3.3K4A mutation reduces H3K4me3 levels. (a) Immunoblot analysis of H3.3 control and K4A/K36A mutants using antibodies recognizing different modification states of the N-terminal H3 tail (me - methylation, ac - acetylation). Antibodies are not specific for canonical H3.1/2 or H3.3 variant, but recognize entirety of H3 histone. Protein samples were loaded 3x in increasing amounts. H4 serves as loading control and one representative immunoblot is depicted. (b) Western blot quantification of H3 modifications relative to H4. Log₂(FoldChanges) of histone modification were calculated normalized to H4 loading controls. (c) Volcano plot displaying global changes in histone modifications in H3.3K36A (left) or H3.3K4A (right) mutants compared to controls. The abundance of histone modifications was measured by middle-down mass spectrometry analysis of N-terminal histone tails (Sidoli et al., 2014). Log₂FoldChanges were calculated for canonical H3.1/H3.2 in mutant ESC compared to control ESCs and plotted against -log₂(p-values) derived from Student's t-test. The 10 (K36A) or 7 (K4A) most significant modifications are depicted in black. H3K4me3 was not detected in this analysis. n=3 biological replicates. *Mass spectrometry analysis was done by Simone Sidoli.*

Turnover of histone modifications

In collaboration with Simone Sidoli in the laboratory of Benjamin A. Garcia, I tested if the H3.3K36A mutation impacts the turnover of histone modifications. The turnover is representative of how compacted the genomic regions are that are marked by a specific histone modification. For this, we grew cells in heavy SILAC medium for 24 hours and measured the ratios of histone modifications on old (light) and newly synthesized (heavy) histones. Surprisingly, we did not detect changes in turnover of H3K27me3 or H3K36me2 (Fig. 3.19), despite the observed global changes in abundance observed previously, and the predicted redistribution of H3K27me3 to intergenic regions. However, we observed significantly higher turnover rates of H3K27 acetylation (on canonical H3.1/H3.2) in H3.3K36A mutants (4-Fold change), which was also globally reduced as measured by immunoblot (Fig. 3.18a,b) and mass spectrometry analysis (Fig. 3.19). The higher turnover rate suggests that chromatin regions marked by H3K27ac are more open, resulting in the more rapid deposition of H3K27ac onto newly synthesized histones. Possibly, the turnover of H3K27ac could be directly or indirectly affected by the global decrease in H3K36me2 and spread of H3K27me3 in H3.3K36A mutants. These results are preliminary and further experiments will be required to define this interactions. To this end, I propose to perform ATAC-Sequencing, a technique that measures the accessibility of chromatin. In this way, I would test if regions marked by H3K27ac are actually more open (higher turnover) in H3.3K36A mutants.

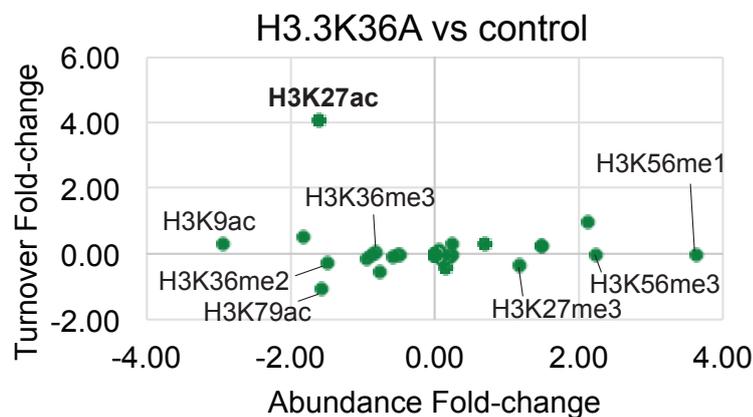


Figure 3.19: Analysis of turnover of histone modifications in H3.3K36A mutants by Mass Spectrometry. Scatter plot displaying changes in the abundance (x-axis) and the turnover (y-axis) of histone modifications in H3.3K36A mutants compared to controls. The turnover is a measure for the chromatin compaction state of genomic regions marked by a modification. ESCs were grown for 24 hours in heavy SILAC medium and the abundance of histone modifications was measured by bottom-up mass spectrometry analysis (Sidoli et al., 2014). The turnover was calculated from the ratio of heavy to light peptides detected for each modification. FoldChanges were calculated in H3.3K36A mutant ESCs compared to control ESCs. Some of the most significant modifications are depicted in black. n=3 biological replicates. *Mass spectrometry analysis was done by Simone Sidoli.*

3.2.3.2 ChIP-Seq analysis of region-specific histone modifications

H3.3K36A

H3K36me3 is enriched inside coding regions and has been proposed to regulate transcriptional outcome and alternative splicing of the underlying gene (Wagner and Carpenter, 2012). This association is based on observations that the H3K36 HMTase Set2 is recruited co-transcriptionally by Pol II (Li et al., 2003; Krogan et al., 2003) and deposit H3K36 methylation during transcriptional elongation, resulting in high levels of H3K36me3 in gene bodies of actively transcribed genes (Kizer et al., 2005; Li et al., 2003, 2002). However, it remains unclear whether the deposition of H3K36me3 in coding regions facilitates transcription and splicing, or whether it is just a by-product of active transcription. To address these questions, I used ChIP-Seq and RNA-Seq in wild type and H3.3K36A mutants and correlated changes of H3K36me3 inside genes with gene expression in ESCs, during neuronal development (NPCs) and in terminally differentiated neurons. I observed a positive correlation in ESCs (Pearson=0.3), neural precursor cells (Pearson=0.6) and neurons (Pearson=0.4), but H3K36me3 levels do not necessarily alter gene expression of the majority of genes (Fig. 3.20a). These results suggest that H3K36me3 is deposited dependent on transcription, but that a reduction in H3K36me3 levels, as induced by a H3.3K36A mutation, has a minor impact on gene expression especially in undifferentiated ESCs. The increasing correlation in developing NPCs could indicate that Setd2 is recruited to developmental genes to mediate its function there.

To test whether splicing is affected, I used DEX-Seq to identify differentially spliced exons in K36A mutants compared to controls. In ESCs, I observed small numbers of alternative splicing events (Fig. 3.20b), and the inclusion or exclusion of an exon did not correlate with H3K36me3 levels inside the affected exon (Fig. 3.20c). K36A neurons displayed more alternative splicing events than observed in ESCs (Fig. 3.20b), but these events again did not correlate with H3K36me3 levels inside the exon (Fig. 3.20c). This suggests that in our system, alternative splicing of an exon is not the consequence of reduced or increased H3K36me3 levels.

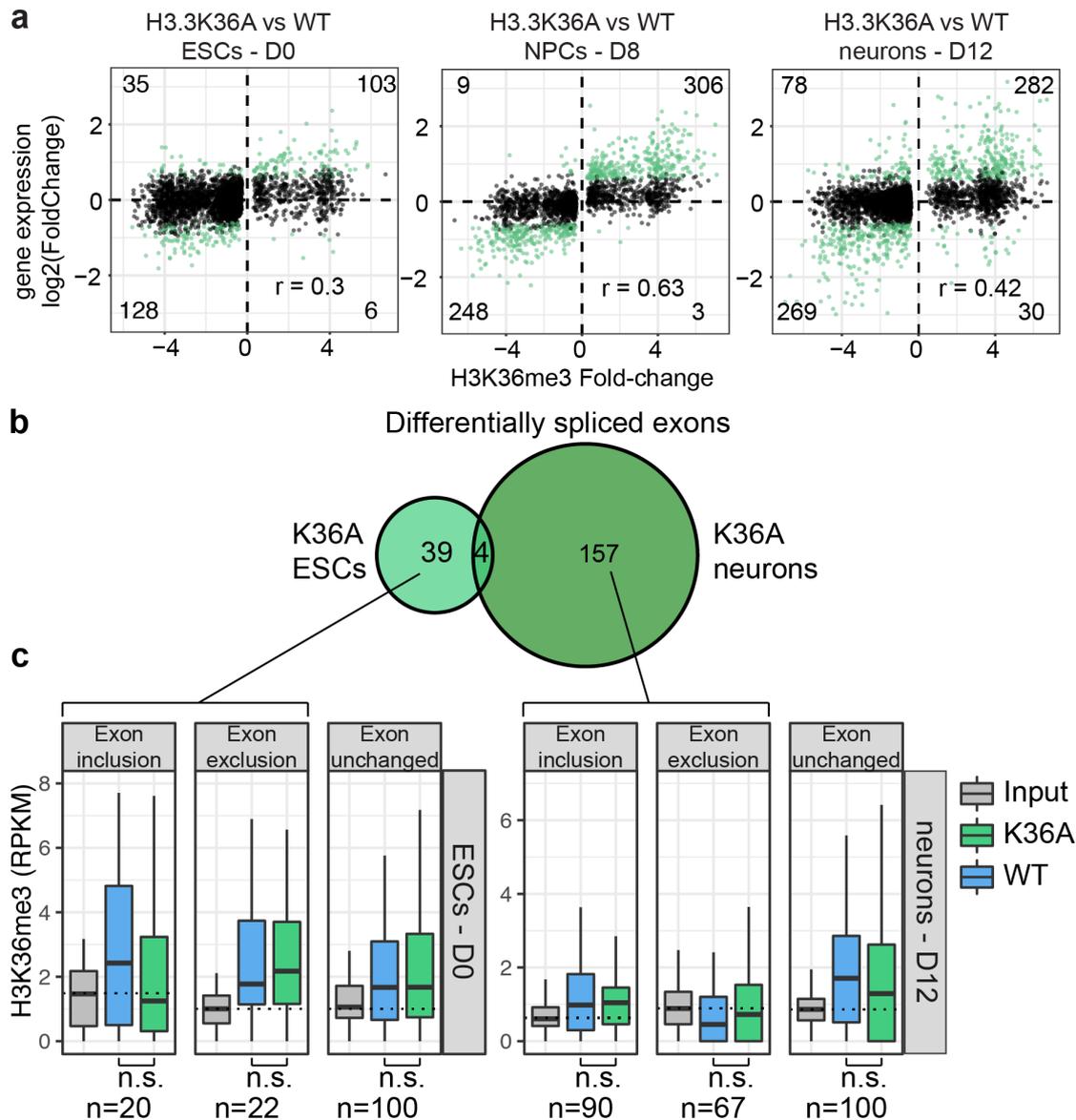


Figure 3.20: H3K36me3 levels do not directly regulate gene expression or differential splicing. (a) Scatter plot correlating genic H3K36me3 levels (ChIP-Seq) with gene expression changes (RNA-Seq) in ESCs (D0), NPCs (D8) and neurons (D12). Significant Fold-changes (FDR < 0.05) of genic H3K36me3 levels (x-axis) are plotted against log₂(FoldChanges) in gene expression (y-axis) measured in H3.3K36A mutant cells compared to wild type cells. The colour of dots indicates whether gene expression changes are significant (FDR < 0.05, absolute log₂(FoldChange)>0.58) and numbers of significant changes are indicated per quadrant. Pearson correlation coefficients (r) are depicted in plots. (b) Venn Diagram depicting total numbers of significant alternative splicing events in H3.3K36A ESCs and neurons compared to controls. Analysis was done using DEX-Seq (Anders et al., 2012) with a significance threshold of FDR<0.1. Biological replicates were n=2. (c) Boxplot depicting H3K36me3 levels in alternatively spliced exons in H3.3K36A and control ESCs (left) and neurons (right). H3K36me3-ChIP-Seq reads per kilobase per million mapped reads (RPKM) were calculated for exons that were significantly more included or excluded from the mRNA, or for randomly selected exons that were not differentially spliced. Number of analyzed genomic features (n) are depicted underneath boxplot. Significance was calculated using an unpaired Wilcoxon test. All comparisons were not significant (n.s.).

H3.3K4A

In mammalian cells, H3K4me1 is found at enhancer regions and H3K4me3 is a hallmark of promoters (Barski et al., 2007; Guenther et al., 2007; Mikkelsen et al., 2007). Additionally, regulatory elements can be marked by H3K27ac, which is predictive of their activity state (Rada-Iglesias et al., 2011; Bonn et al., 2012; Creyghton et al., 2010; Wang et al., 2008).

H3.3K4A ESCs displayed widespread gene expression changes, and I investigated whether changes in H3K4 methylation or H3K27 acetylation were causally involved in these changes. I used ChIP-Seq to examine the genomic distributions of H3K27ac, H3K4me1, H3K4me3, and H3.3 in control and H3.3K4A ESCs (Supplementary Fig. S5), and correlated them with transcriptional changes of genes (Fig. 3.21). Clustering analysis of all datasets revealed that distinct biological GO term categories were associated with groups of up- and down-regulated genes and histone modification profiles (Fig. 3.21a and Supplementary Fig. S6a). Notably, H3.3K4A ESCs showed a substantial decrease in H3K27ac that was highly correlated with H3.3 depletion at TSS and promoter-proximal active enhancers (Fig. 3.21a). Moreover, the H3.3 depletion at TSS and enhancers occurred regardless of differential gene expression changes (Fig. 3.21b), but the extent of the H3K27ac reduction correlated with gene expression, and down-regulated genes displayed no or very low H3K27ac levels at their TSS compared to other genes (Fig. 3.21b). The same trend was observed for promoter-proximal enhancers of up- and down-regulated genes, which both showed reduced H3K27ac, but the extent of the reduction correlated with the expression of the proximal gene (Supplementary Fig. S6b).

In contrast, H3.3K4A ESCs showed a significant increase of H3K4me1 signal at H3.3 depletion sites around TSS independent of gene expression changes (Fig. 3.21a). ESC-specific enhancers displayed overall reduced levels of H3K4me1 (Fig. 3.21b), and at promoter-proximal enhancers H3K4me1 levels correlated with gene expression of the nearby gene (Fig. S6b). I found that H3K4me3 was mildly increased directly at the TSS (Fig. 3.21a,b). Instead, H3K4me3 peaks that reside within or upstream/downstream of genes, including enhancers, were significantly reduced (Supplementary Fig. S6b,c), consistent with the observed global decrease of this mark detected by immunoblot (Fig. 3.18a).

Collectively, the depletion of H3.3K4A occurred at all genes independent of transcriptional alterations, and induced correlated changes in H3K27ac and anti-correlated changes in H3K4me1 at TSS. The extent of H3K27ac reduction at promoters and enhancers was predictive of the differential gene expression status. The reduction in H3K27ac levels could be the result of a destabilization of HATs, such as p300, by the H3.3K4A mutation. The severity of this destabilization at genes might in turn be associated with the transcriptional outcome in the K4A mutant. Alternatively, if H3.3K4A is the target of acetylation, the removal of mutant H3.3 at depletion sites could also be responsible for reduced H3K27ac levels. The increase in H3K4me1 could be the result

of the replacement of H3.3 with canonical H3.1/H3.2, which might be the preferred target for HMTases that catalyze H3K4me1.

My study revealed that a H3.3 lysine 36 mutation affected differentiation and transcription at later stages of neurodevelopment, whereas H3.3 lysine 4 mutation lead to widespread changes in both the histone modification landscape and transcription program in ESCs, which resulted in impaired neuronal differentiation. In particular, the H3.3K4A mutation caused a significant decrease in H3K27ac at active regulatory elements, which is not observed in H3.3 knockdown cells (Banaszynski et al., 2013), and clearly differentiates the effect of the mutation from a loss-of-function phenotype. Despite the overall decrease of H3K27ac enrichment at active regulatory regions, the extent of H3K27ac reduction at promoters correlated with changes in gene expression, suggesting that subsets of genes were more affected by the H3.3K4A mutation. Furthermore, I showed that lysine 4 itself is crucial for enrichment of H3.3-containing nucleosomes at TSS and active enhancers in ESCs. This result suggests that the cellular machinery mediating nucleosome exchange/remodeling at regulatory elements depends on H3 lysine 4 itself. I also found that in differentiated cells, the depletion of H3.3K4A was not restricted to regulatory regions, but resulted in widespread proteasomal degradation of the mutant histone from the chromatin fraction, strongly arguing that lysine 4 is required for stability.

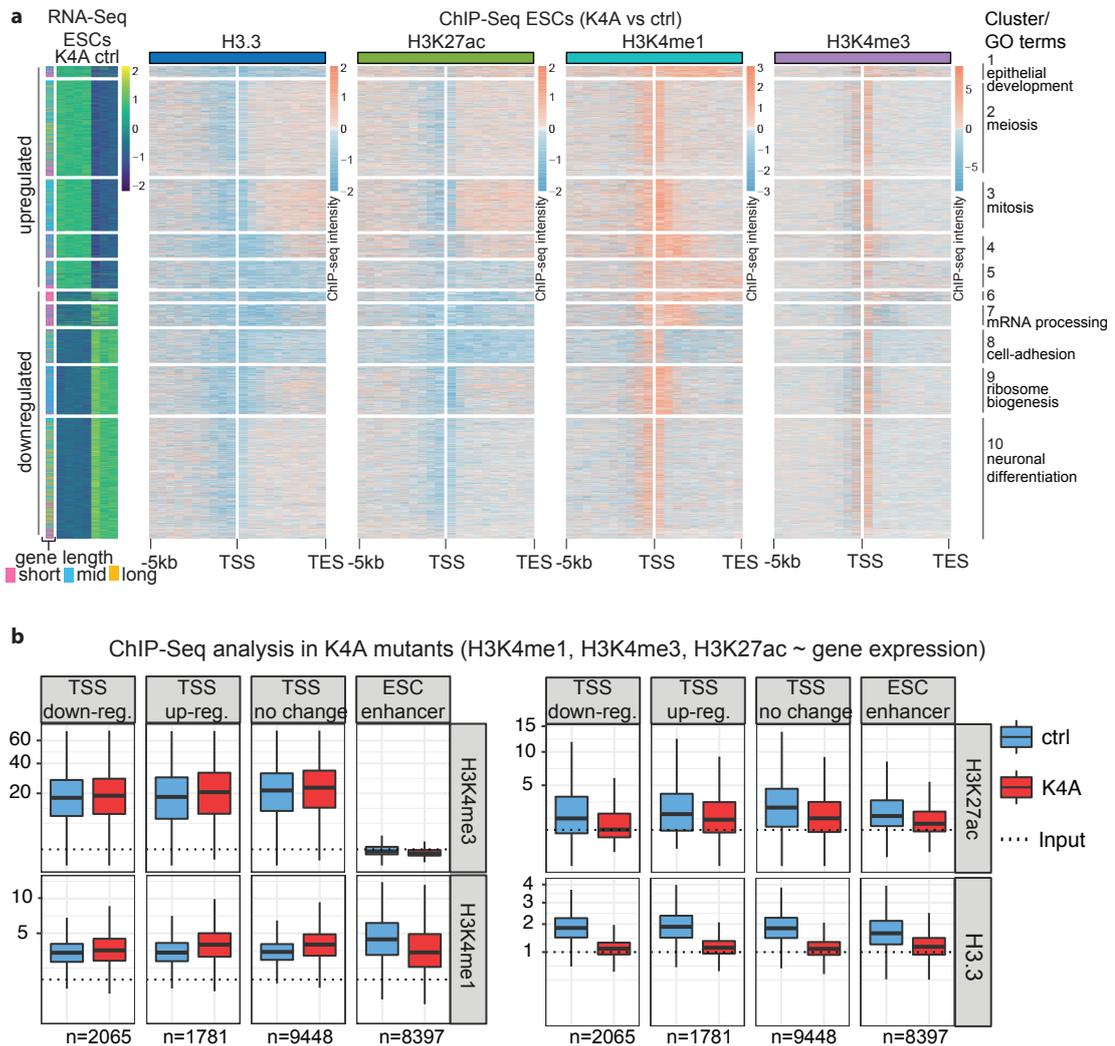


Figure 3.21: H3.3K4A mutation induces histone modification changes at regulatory elements. (a) Heatmap clustering of differentially expressed genes in H3.3K4A mutants (FDR<0.05). Genes were grouped by RNA expression changes and signal intensity changes of H3.3, H3K27ac, H3K4me1, H3K4me3 ChIPSeq data using k-means clustering. Clusters 1-5 contain up-regulated genes (n=2859) and 6-10 downregulated genes (n=3196). Representative GO-terms are indicated per cluster (full list in Supplementary Fig. 1.2). Gene length is indicated by colour (left). RNA-expression values are z-scored across samples and based on normalized read counts per gene. Biological replicates for RNA-Seq n=3 (ctrl), n=4 (K4A). For ChIP-Seq, changes in signal intensity (K4A vs ctrl) were plotted from 5 kb upstream of the TSS (500bp per window) and for scaled gene bodies (10% of gene per window). Biological replicates for ChIPSeq n=2 (ctrl), n=3 (K4A). (b) Analysis of H3K4me1, H3K4me3, H3K27ac and H3.3 enrichment in control and H3.3K4A ESCs at TSS and enhancers. ChIP-Seq reads per kilobase per million mapped reads (RPKM) were calculated for TSS (± 1.5 kb) of differentially expressed genes and unchanged genes, and for all ESC-specific enhancers (Shen et al., 2012). Differential gene expression status for up- and down-regulated genes was determined by RNA-Seq analysis in H3.3K4A ESCs compared to control. Number of analyzed genomic features (n) are depicted underneath boxplot. RPKM levels higher than those of input indicate enrichment of the histone modification. ChIP n=3, Input n=2 biological replicates. Boxes display median RPKM values, third and first quartile, whiskers show $1.5 \times$ the interquartile range above and below the box. *Daria Bunina performed clustering analysis of combined RNA-Seq and ChIP-Seq data and generated heatmap displayed in (a).*

3.3 Discussion & Outlook

In yeast and *Drosophila*, systematic mutation of histone H3 has proven to be an invaluable tool to discern the functional role of a histone residue. In humans, the study of oncogenic histone mutations has recently emerged and helped in understanding the function of histone residues in healthy and diseased cells (Lu et al., 2016; Fang et al., 2016; Chan et al., 2013; Lewis et al., 2013). These oncogenic mutations often occur in only one allele of the histone variant H3.3, but nevertheless can have severe consequences on the epigenetic landscape of cells and can contribute to the development of cancer (Schwartzentruber et al., 2012; Wu et al., 2012; Behjati et al., 2013). Thus, mutations of the histone variant H3.3 could provide a suitable system to test the functional role of histone residues in mammalian development, without the need to exchange the canonical H3 genes. Using this system, we aimed to test whether histone modifications, such as H3K36me3, H3K4me3, H3K27me3 and H3K27ac, which strongly correlate with transcriptional activity (Barski et al., 2007; Guenther et al., 2007; Heintzman et al., 2007), are causally involved in mammalian transcription and to identify uncharacterized functions of histone residues.

In this study, I used the mutation of histone residues K4 and K36 to complement the studies that have perturbed the respective histone modifying enzymes, thereby assessing their contributions to chromatin regulation. This strategy has been successful in the past to unravel the functions of a HMTase from that of the deposited histone modification, as seen for pioneering histone mutation studies in *Drosophila melanogaster*, which have delivered evidence for the role of H3K27me3 and H3K27ac for transcriptional regulation. Mutations of H3K27 to arginine (H3K27R) in wing imaginal disc cells of *Drosophila melanogaster* fail to repress transcription of PRC2-target genes and these flies display morphological differences similar to PRC2 mutants (Pengelly et al., 2013). At the same time, H3K27R mutant cells do not show widespread transcriptional changes of other genes or impaired transcriptional activation (Pengelly et al., 2013; McKay et al., 2015). While these results emphasize the importance of H3K27me3 for transcriptional repression by PRC2 (Young et al., 2011), it also suggests that transcriptional activation might not require H3K27 acetylation on canonical H3 (Pengelly et al., 2013; McKay et al., 2015). It is a compelling idea that H3K27ac is only a byproduct of histone acetyl-transferase p300, whose relevant targets include transcription factors and the Pol II complex itself (Imhof et al., 1997; Weinert et al., 2018; Schroder et al., 2013; Haberle and Stark, 2018).

Alternatively, it would be possible that acetylation of histone variant H3.3, which remained intact in this *Drosophila* study, might be sufficient to promote transcriptional activation. Nonetheless, a H3.3 knockdown study in mouse ESCs did not result in reduced H3K27ac enrichment at promoters (Banaszynski et al., 2013), arguing that H3K27ac can be sufficiently maintained by canonical H3 in ESCs and that H3.3 depletion itself does not result in a decrease of H3K27ac.

In H3.3K4A mutant ESCs, I observed a reduction of H3K27ac enrichment, which was highly correlated with depletion of mutant H3.3 from regulatory regions. This poses the question how a H3.3K4A mutation can have such a severe impact on H3K27ac at active promoters and enhancers. H3K4 methylation is highly correlated with H3K27ac at active regulatory regions, suggesting that H3K4 HMTases and HATs function in proximity to each other. Thus, the H3.3K4A mutation could directly or indirectly destabilize the binding of HATS, such as p300, to regulatory regions, thereby perturbing gene expression and resulting in the reduction of H3K27ac levels at sites of H3.3K4A depletion.

Transcriptional regulation by H3K4 methylation

H3K4 methylation is associated with regulatory regions of the genome and the degree of methylation differs depending on the genomic site. H3K4me1 is highly enriched at enhancers and H3K4me3 is a hallmark of promoters (Heintzman et al., 2007). In yeast, H3K4me3 only marks promoters of actively expressed genes (Santos-Rosa et al., 2002; Ng et al., 2003), whereas in mammalian cells H3K4me3 marks both inactive and active promoters (Barski et al., 2007; Guenther et al., 2007; Mikkelsen et al., 2007), and its implication in transcription is less clear. Recently, the pervasive enhancer mark H3K4me1 has been shown to be dispensable for enhancer activity and gene regulation by HMTases in pluripotent cells and *Drosophila* development (Dorighi et al., 2017; Rickels et al., 2017). In these two studies, the authors used catalytically dead enzymes to analyze if the H3K4 HMTases Mll 3 and 4 can mediate their regulatory function at enhancers independent of their HMTase activity. Indeed, they found that enhancer activity, as measured by enhancer RNA synthesis by Pol II, and regulation of associated genes can occur in the absence of H3K4me1 deposition. Similarly, H3K4me3 at promoters of constitutively active genes was shown to be dispensable for gene transcription in mouse ESCs (Clouaire et al., 2012, 2014). Such data demonstrate the need of directly testing the function attributed to histone residues, rather than relying purely on correlation or indirect inference from phenotypes of mutant enzymes that catalyze modifications (Pollex and Furlong, 2017).

Efforts to understand the role of H3K4 methylation for transcription have been made in *Drosophila* by Hödl et al. The authors have replaced H3K4 with alanine in canonical and variant H3 of *D. melanogaster* and the results suggest that gene regulation and transcription can occur in the absence of H3K4 methylation (Hödl and Basler, 2009, 2012). H3K4A wing imaginal disc cells can respond to developmental cues by activating key transcription factors, but the mutant cell patches remain small and display a strongly diminished proliferative capacity. Similarly, mutations of H3K4 to alanine or arginine in *Saccharomyces cerevisiae* result in slow-growth defects, which are also observed for deletions of the sole H3K4 HMTase, Set1 (Briggs et al., 2001). This data suggests that H3K4 HMTases regulate transcription often independent of histone methylation, but H3K4 might have other uncharacterized functions that are relevant for normal cell sur-

vival and growth. In agreement with this, H3.3K4A ESCs in this study maintained the ability to activate developmental genes and repress pluripotency genes during neuronal differentiation, but displayed reduced fitness and a developmental delay. As a result, H3.3K4A mutants yielded lower cell numbers in differentiation and were not able to give rise to a homogenous population of glutamatergic neurons. Furthermore, the transcriptional program of H3.3K4A ESCs and neurons was severely perturbed, and this was not due to diminished H3K4me3 or H3K4me1 levels at promoters and enhancers. Instead, the H3.3K4A mutation resulted in depletion of H3.3 from regulatory regions, which may explain the minor impact of the mutation on H3K4 methylation levels.

Histone exchange

The reasons for the depletion of H3.3K4A from regulatory regions is less obvious. The mutation is unlikely to affect initial chromatin assembly of newly synthesized H3.3 by its chaperone Hira, which does not bind the histone tail, but instead recognizes four variant amino acids in the folded histone core (Ahmad and Henikoff, 2002; Ricketts et al., 2015). Several chromatin remodeling complexes have already been identified to also act as histone chaperones that can assemble H3 into nucleosomes (Konev et al., 2007; Pradhan et al., 2016; Lewis et al., 2010). In proliferating cells, H3.3 is enriched at genomic sites with active histone exchange (Kraushaar et al., 2013; Deaton et al., 2016) and H3.3K4A mutant protein might interfere with the action of some chromatin remodeling complexes.

Co-transcriptional histone exchange inside coding regions occurs more slowly and is mediated by different proteins, which catalyze restructuring of the nucleosome before and behind transcribing Pol II. Histone modifications in newly deposited histones, such as methylation of H3K36, have been reported to suppress histone exchange in coding regions of the yeast genome (Smolle et al., 2012; Venkatesh and Workman, 2015), suggesting that H3K36 methylation could be important for recycling of old histones during transcriptional elongation. Unlike the yeast study, I did not observe that the deposition, or the enrichment of H3.3K36A or H3.3K4A inside coding regions of ESCs was impaired compared to wild type H3.3. Furthermore, mutant H3.3K4A/K36A was not exchanged more rapidly at TES than wild type H3.3. Together, these results indicate that H3K4 and H3K36 methylation of histone H3.3 is not required for histone recycling behind transcribing Pol II in mammals.

Compared to coding regions, histone exchange at regulatory regions occurs in the absence of transcription (Venkatesh and Workman, 2015) and higher turnover rates facilitate transcription factor binding and assembly of the transcriptional machinery on DNA (Henikoff, 2008). In ESCs, specific chromatin remodeling complexes have been shown to preferentially bind at regulatory regions (de Dieuleveult et al., 2016) and it is possible that one or multiple of these remodelers depend on the H3K4 residue for histone exchange and nucleosome re-assembly at regulatory regions. Comparison of

the enrichment profiles of these remodelers revealed that Chd1 occupies similar nucleosomes as H3.3 at promoters. Chd1 is evolutionary conserved from yeast to humans, and human Chd1 has developed the ability to distinguish and bind different H3K4 methylation states (Sims et al., 2005) compared to its yeast counterpart. The H3K4 residue might be required for viability in animals, because it is recognized by one or multiple chromatin remodelers that mediate histone exchange processes at regulatory regions of higher organisms, e.g. Chd1.

Interestingly, in differentiated cells depletion of H3.3K4A was not restricted to regulatory elements and I observed widespread degradation of H3.3K4A in the chromatin fraction. Chromatin of differentiated cells is more compacted than ESCs and the action of chromatin remodeling complexes might be increasingly important in differentiation at all chromatin types to enable cellular plasticity and access to the DNA template. Furthermore, differentiating cells are proliferating more slowly than ESCs and increasingly depend on the H3.3 variant for replication-independent nucleosome assembly (Maze et al., 2015). Functional analysis of the association of chromatin remodelers and mutant H3.3K4A will be required to address the question how and if histone exchange is involved in the observed depletion of H3.3K4A.

Transcriptional regulation by H3K36 methylation

The co-transcriptional recruitment of H3K36 HMTases during transcriptional elongation results in high levels of H3K36 methylation at gene bodies. Because of the association with active transcription, H3K36me3 is generally viewed as an activating modification that facilitates alternative splicing and transcriptional elongation, e.g. by preventing cryptic transcription initiation from inside coding regions (Carrozza et al., 2005; Fang et al., 2010; Dhayalan et al., 2010). In H3.3K36A mutant ESCs, transcription remained largely unaffected and resulted in only few differentially expressed genes and alternative splicing events, despite a significant reduction in H3K36me3 levels. While I observed increasing numbers of differentially expressed genes and alternative splicing events in H3.3K36A neurons compared to ESCs, these events were not directly regulated by H3K36me3 levels. These results suggest that transcriptional elongation and alternative splicing are not directly affected by changes in H3K36me3 levels inside coding regions. However, changes in H3K36me3 levels moderately correlate with changes in transcription (H3.3K36A mutant vs. control), supporting the co-transcriptional deposition of H3K36me3 by HMTases. Furthermore, I observed the highest correlation between changes in H3K36me3 levels and changes in transcription in neural precursor cells, suggesting that Setd2, which is the only HMTase depositing H3K36me3 (Yuan et al., 2009; Edmunds et al., 2008), might be recruited during the activation of developmental genes to mediate its function.

Our findings are in agreement with histone mutation studies in *Drosophila*, in which H3K36R mutant embryos did not show perturbed transcriptional elongation by cryptic transcription initiation or alternative splicing (McKay et al., 2015). Thus, the role of

H3K36 methylation in transcriptional elongation might not be conserved from yeast to flies and mammals, or the HMTases might have evolved to mediate these function independent of H3K36 methylation. However, H3K36R mutant larvae die before the completion of development (Meers et al., 2017), which is also observed for the depletion of dSet2, the sole H3K36 HMTase in *Drosophila* (Stabell et al., 2007). Similarly, the loss of Setd2 in mice results in early embryonic lethality by severe vascular defects (Hu et al., 2010). This suggests that H3K36 HMTases have a developmentally essential function that directly depends on H3K36 methylation.

PRC2 activity

Furthermore, H3K36 methylation has been suggested to restrict the activity of PRC2 and prevent expansion of repressive H3K27me3 to active chromatin, because H3K36 and H3K27 methylation are mutually exclusive for the same nucleosome (Young et al., 2011; Schmitges et al., 2011; Yuan et al., 2011). H3K36 mutations have been shown to inhibit the activity of H3K36 HMTases *in vitro* (Lu et al., 2016). Specifically, the oncogenic H3.3K36M (lysine-to-methionine) mutation dominantly inhibited the major H3K36 HMTases Setd2 and Nsd1/2, and thereby reduced all three methylation states of H3K36, which resulted in spread of repressive H3K27me3 to intergenic regions (Lu et al., 2016). The H3.3K36A (alanine) mutation showed a more pronounced inhibitory effect on Nsd1/Nsd2 *in vitro*, which deposit only H3K36me1/me2 (Yuan et al., 2009; Edmunds et al., 2008). In agreement with this, we observed that H3K36me2 on total H3 is the most reduced histone mark in H3.3K36A ESCs, suggesting that H3.3K36A does indeed have a dominant inhibitory effect on Nsd1 and/or Nsd2 *in vivo*.

Additionally, H3.3K36A mutant ESCs displayed a global gain of H3K27me3, which supports the role of H3K36 methylation to restrict PRC2 activity. In agreement with this, inhibition of the H3K36 HMTase Nsd2 has been shown to cause a switch in H3K36me2-occupied chromatin domains to H3K27me3-occupied domains, strongly arguing that H3K36me2 specifically restricts H3K27me3 (Streubel et al., 2018). This role of H3K36me2 in restricting PRC2 might be limited to mammals, because a H3K36 to arginine (H3K36R) mutation in *Drosophila* does not result in a global increase of PRC2-mediated H3K27me3 (McKay et al., 2015; Meers et al., 2017), or the H3K36R mutation might sufficiently mimic methylated H3K36. Surprisingly, these severe changes in the epigenetic landscape had no major impact on the transcriptional program of ESCs, whereas gene expression was perturbed only at later stages of neuronal development. The restriction of PRC2-mediated H3K27me3 might not be as important in pluripotent ESCs, which have been shown to have a more open and permissive chromatin state compared to differentiated cell types (Bhattacharya et al., 2009; Bickmore and Zaret, 2010; Mattout and Meshorer, 2010; Meshorer and Misteli, 2006). During differentiation, chromatin gets more compacted, and the distribution of H3K36 and H3K27 methylation might contribute in setting a balance of active and repressed chromatin states.

Cell proliferation and growth

The oncogenic histone mutation H3.3K36M is detected in over 95% of chondroblastomas (Behjati et al., 2013), which are characterized by the accumulation of immature chondroblasts and the growth of this cancer has been suggested to arise from uncontrolled proliferation and incomplete differentiation of mesenchymal progenitor cells (Lu et al., 2016). Similarly, I observed that H3.3K36A neurons in this study were less committed to the glutamatergic lineage and grew to a much higher density compared to wild type neurons. Interestingly, similar tissue overgrowth phenotypes have also been observed in human diseases with underlying mutations in a H3K36 HMTase (Kurotaki et al., 2002). Specifically, loss-of-function mutations of Nsd1 are the major cause of Sotos syndrome (Tatton-Brown et al., 2005; Douglas et al., 2003; Cecconi et al., 2005), which is characterized by pre-natal and post-natal overgrowth, distinctive facial appearance, advanced bone age and developmental delay of mental and movement abilities. Strikingly, the affected children tend to grow quickly and have unusually large heads, which could indicate neuron over-growth (Tatton-Brown et al., 2005; Cecconi et al., 2005). It would be interesting to test if deletions of Nsd1 result in an increased density of the neuron population as observed for the H3.3K36A mutation, which would suggest that the H3K36 HMTase activity of Nsd1 and thus H3K36 methylation is required to restrict organism growth and uncontrolled cell proliferation. Together, this suggests that the H3K36 residue is important to restrict cell growth by facilitating lineage commitment of cells into the terminally differentiated cell type.

While the exact mechanism of H3.3 turnover has yet to be elucidated, this study identifies a direct link between a specific histone residue (H3K4) and histone enrichment at promoters and enhancers. In the light of the growing list of H3 mutations associated with human diseases, this study shows that endogenous H3.3 mutations can serve as an experimental platform to study the functional role of histone residues in mammalian development.

3.4 Materials & Methods

Cell culture. HEK293T cells were cultured in Dulbecco's modified Eagle's medium (DMEM, Thermo Fisher) containing 10% fetal bovine serum. Murine ESCs (129XC57BL/6J) were cultured in ESC media containing Knockout-DMEM (Thermo Fisher) with 15% EmbryoMax FBS (Millipore) and 20 ng/ml leukemia inhibitory factor (LIF, produced by Protein Expression Facility at EMBL Heidelberg), 1% non-essential amino acids, 1% Glutamax, 1% Pen/Strep, 1% of 55 mM beta-Mercaptoethanol solution. Cells were maintained at 37°C with 5% CO₂. ESCs were routinely tested for mycoplasma. For ESC growth assay, cells were seeded at a density of 10⁴ cells per well. Per cell line 12 wells were seeded on day 0. Two wells per cell line were trypsinized and counted daily for a total of 6 days, while the remaining wells received fresh media. Growth curves were generated using the averaged duplicate cell counts.

Neuronal differentiation. ESCs were differentiated into glutamatergic neurons according to Bibel:2007 with small modifications. 4x10⁶ ESCs were resuspended in 15 ml differentiation media without LIF (regular DMEM supplemented with 10% FBS (Gibco), 1% non-essential amino acids, 1% Glutamax, 1% Pen/Strep, 1% Sodium Pyruvate) and grown in suspension using non-adherent bacteriological petri dishes (Greiner, 633102, 10 cm with vents) to promote the formation of embryoid bodies. Differentiation media was changed every two days. For this embryoid bodies were transferred to 50 ml Falcon tubes, collected at the bottom of the tube by allowing them to settle within 5 minutes. Supernatant medium was removed and EBs were resuspended in 15 ml fresh differentiation media and transferred to a new bacteriological petri dish. On day 4, 5 µM retinoic acid was added to the media. Embryoid bodies were cultured for 4 additional days in the presence of retinoic acid to obtain neural precursor cells on differentiation day 8. Embryoid bodies were dissociated with trypsin and cells were plated in N2 media composed of regular DMEM supplemented with 1xN2 and 1xB27-VitaminA (Thermo Fisher) at a density of 2x10⁵ cells/cm². For neural culture, plates were pre-coated with Poly-D-Lysine (Sigma) and Laminin (Roche). N2 media was changed every two days without exposing neuron culture to air. Four days after plating, neurons were harvested after a total of 12 differentiation days.

RNA-Seq analysis. RNAs were extracted from approx. 1x10⁶ cells using RNeasy Kit (Qiagen), followed by DNase digestion using TURBO DNase (Thermo Fisher). mRNAs were isolated from 1 µg of total RNA using a PolyA selection kit (NEB) and sequencing libraries were prepared following instructions from NEBs Ultra Library Preparation Kit for Illumina. All samples were barcoded, pooled and sequenced on a HiSeq2000 Sequencer (Illumina) using a 50 bp single-end run. Sequencing reads were mapped to mouse reference genome (mm10 assembly) using Tophat2 aligner with default settings for single-end reads. Reads per gene were counted using HT-

SeqCount union or intersection_nonempty mode. We used Ensembl gene annotation Mus_musculus.GRCm38.83. Differential gene expression analysis was carried out using DESeq2 (Love et al., 2014). Genes with a FDR of $FDR < 0.05$ were considered differentially expressed. For visualization of gene expression changes, gene counts were normalized using DESeq2 and plotted per gene. Enrichment analysis for GO terms associated with biological processes was performed using topGO package using differentially expressed genes with $FDR < 0.05$ and absolute $\log_2(\text{FoldChange}) > 0.58$. Significantly enriched GO terms were identified using Fisher classic and Fisher elim algorithm. Numbers of biological replicates used are indicated underneath each figure. For Fig. 3.6 two biological replicates were used to detect differentially expressed genes to compare between timepoints and conditions. For heatmap in Fig. 3.21, three biological replicates for control and 4 biological replicates for K4A mutants were used to calculate differentially expressed genes.

SNAP experiment. ESC lines stably expression H3.3-SNAP were generated using a p2lox plasmid containing cDNA of H3.3 (wildtype or mutant) with a C-terminal HA- and SNAP-tag. H3.3-HA-SNAP was integrated into the *hprt* locus of A2lox.Cre ESCs (Iacovino et al., 2011) according to instructions in Deaton et al. (2016). A2lox.Cre ESCs were treated with 500 ng/ml doxycycline for 24 hours, nucleofected (Lonza) using 2 μg of p2lox-H3.3-HA-SNAP plasmid and plated on drug-resistant MEFs. 24 hours after nucleofection, cells were treated with 300 $\mu\text{g}/\text{ml}$ G418 for 7 days. Visible colonies were picked and screened for integration of H3.3-HA-SNAP using immunoblotting. For biotin pulse-labelling experiment, cells were expanded and H3.3-SNAP expression was induced by the addition of doxycycline for 48 hours. ESC were harvested and 30×10^6 cells were incubated in 150 μl of media supplemented with 10 μM SNAP-Biotin (NEB) for 30 min at 37°C. Cells were washed twice by resuspension in 10 ml of fresh media followed by centrifugation at 1100 rpm, 4 min. Cells were resuspended in 10 ml media and incubated for 30 min at 37°C to allow unbound biotin to soak out. Cells were washed one last time. For time point 0 hrs, 30×10^6 cells were harvested, for later time points 30×10^6 cells were re-plated in a 15 cm plate and harvested after 3 or 6 hours, respectively. Cell lysates were prepared from 30×10^6 cells for native ChIP.

Native ChIP. DNA for native ChIP was digested by MNase treatment to obtain mainly mononucleosomes and a fraction of dinucleosomes using a modified protocol from Barski et al. (2007). In each sample, 10^7 cells were resuspended in digestion buffer, treated with 100 U MNase (Worthington) and incubated at 37°C for 5 minutes while shaking at 500 rpm. Samples were immediately moved to ice and MNase was quenched by addition of EDTA. Lysates were sonicated using Bioruptor Pico (Diagenode) and dialyzed against RIPA buffer for 3 hours at 4°C. Insoluble materials was pelleted at 10,000rpm for 10 min, 4°C and supernatant was used as input for ChIP. To check that mostly mono-nucleosomes and di-nucleosomes were obtained, DNA fragment

sizes of inputs were analyzed by agarose gel electrophoresis. 5% of inputs were set aside for sequencing. For native ChIP of histone modifications, Protein-G-Dynabeads (Invitrogen) were pre-coated with antibodies for 4 hours at 4°C. For ChIP we used the following antibodies: H3K4me1 (39297, active motif), H3K4me3 (39915, Active Motif), H3K27ac (39685, Active Motif). For SNAP-Biotin ChIP we used MyOne-StreptavidinT1-Dynabeads (Thermo Fisher). Beads were added to lysates and rotated over night at 4°C. On the next day, beads were washed and finally rinsed with TE+50mM NaCl Buffer. Samples were eluted from ProteinG Dynabeads using SDS-elution buffer (50 mM Tris-HCl pH8.0, 10 mM EDTA, 1% SDS) at 65°C for 30 minutes, shaking at 1500 rpm. For streptavidin dynabeads, samples were eluted at 65°C for 10 min followed by incubation 85°C for 10 min while shaking at 1500rpm. Proteinase K was added (0.2 mg/ml final) to eluted samples and digestion was carried out for 2 hours at 55°C followed by PCR purification (Qiagen).

Crosslink-ChIP. For H3.3-ChIP, 10^6 cells were harvested and resuspended in 20 ml of 1% fresh paraformaldehyde (PFA) diluted in PBS and rotated for 10 min at room temperature. Unreacted PFA was quenched with glycine. Cells were washed twice to remove PFA. Lysates were prepared by resuspending crosslinked cells in hypotonic buffer (15 mM HEPES pH8, 5 mM MgCl₂, 10 mM KCl, 0.5mM EDTA, 2x Protease Inhibitors) followed by douncing to release intact nuclei. Nuclei were pelleted and resuspended in digeston buffer (15 mM HEPES pH8, 5 mM MgCl₂, 30 mM KCl, 3mM CaCl₂, 2x Protease Inhibitors) by douncing. In-nuclei MNase digestion was started by addition of 150U MNase (Worthington). Samples were then quickly vortexed and incubated at 37°C for 5 min while shaking at 500 rpm. Reaction was quenched by addition of 10x Quenching buffer (150mM HEPES pH8, 1500mM NaCl, 50 mM EDTA, 30 mM EGTA, 0.2% TritonX-100) and by placing samples on ice. 0.1% sodium deoxycholate, 0.5% sarcosine and 1% Triton-X were added to lysates and samples were sonicated in Bioruptor Pico (Diagenode) to further promote DNA fragmentation. Insoluble material was removed by centrifugation at 20.000rpm for 10 min, 4°C and the soluble supernatant was used as ChIP Input. DNA fragment sizes of lysates were checked by agarose gel electrophoresis. Mainly mono-, di-, tri-nucleosomes were obtained. Protein-G Dynabeads were precoated with H3.3 antibody (Millipore) for 4 hours at 4°C. Beads were washed, added to lysates and incubated over night at 4°C while rotating. On the next day, beads were washed 8x with RIPA buffer (50 mM HEPES pH7.6, 100 mM LiCl, 1 mM EDTA, 1% NP-40, 0.7% Sodium deoxycholate) and gently rinsed once with TE buffer+50 mM NaCl. Elution was carried out using SDS elution buffer (50 mM Tris-HCl pH8.0, 10 mM EDTA, 1% SDS) at 65°C for 30 minutes while shaking samples at 1500rpm. Samples were Proteinase K digested (0.2 mg/ml final) for 2 hours at 55°C and crosslink was reversed over night at 65°C. Samples were purified using Qiagen's PCR purification kit and used for ChIP-library preparation.

ChIP-Seq Data Analysis. Sequencing libraries were prepared using DNA Ultra II library preparation kit (NEB) according to the manufacturer's instructions and sequenced either on Illumina's HiSeq2000 Sequencer (50 bp single-end mode) or NextSeq 500 Sequencer (75 bp single-end mode). Sequencing reads were aligned to mouse reference genome (mm10 assembly) using Bowtie2 (Langmead and Salzberg, 2012). Only non-duplicated, uniquely mapped reads were retained for further analysis. ChIP-Seq peak calling was done using MACS2 (Zhang et al., 2008) (for H3K4me1, H3K4me3, K3K27ac) or SICER (Xu et al., 2014) (for H3.3). Differential peak calling of ChIP-datasets for H3.3K4A mutants and controls was done using R and DiffBind package (Bioconductor). Reproducibility between ChIP-Seq replicates was assessed by PCA analysis on binned and library size normalized bigwig files using deeptools (Ramirez et al., 2016). PCA plots and Differential Peak analysis for ChIP-Seq experiments are found in Supplementary Fig. S5. Metagene plot to visualize distribution of ChIP-Seq density over scaled gene bodies was generated using NGSplot (Shen et al., 2014). For Boxplot in Fig. 3.10, genomic coordinates of predicted ESC-specific enhancers and enhancers of other tissue-types were obtained from Shen et al. (2012). To obtain genomic windows, enhancer coordinates were extended by ± 1 kb. Active TSS were defined as annotated RefSeq TSS ± 1.5 kb that display Pol II occupancy in ESCs and inactive TSS ± 1.5 kb as those without Pol II occupancy in ESCs (Shen et al., 2012). Genome coordinates were lifted to mm10 using UCSC's LiftOver Tool. TES of active genes (at least one active TSS) were defined as ± 1.5 kb around gene end coordinates obtained from Ensembl biomart (version: "Ensembl Genes 91"). Gene bodies were defined from gene start coordinates+1.5kb until gene end coordinates-1.5kb as obtained from Ensembl biomart. ChIP-Seq reads or Input-reads in defined regions were counted using multiBamSummary (deepTools) and normalized to sequencing depth and gene length to obtain RPKM values in R.

For ChIP-Seq heatmaps in Fig. 3.21, read coverage for each sample was normalized by the total number of reads, then average coverage for each feature was calculated for all mutant (or control) samples and subtracted from each other to get the final score. The locations of TSS and TES of differentially expressed genes were taken from Ensembl biomart. The heatmap was produced in R using a pheatmap package. For the boxplots in the Supplementary Fig. S6 we calculated the fold change of similarly obtained average coverage scores for mutant vs. control samples within each region of interest. TSS regions were defined as ± 1.5 kb from the gene start position, TES regions - as ± 1.5 kb from the gene end position. Gene bodies were defined from TSS+1.5kb until TES-1.5kb (for genes shorter than 3 kb the entire TSS to TES region was taken). Enhancers were selected from the list of predicted ESCs-specific enhancers¹⁴ overlapping the region 0-50kb upstream of the gene's TSS.

ChIP-qPCR. Prepared libraries from ChIP experiment were diluted 1:20 and used as input for qPCR reaction with SYBR Green PCR Master Mix (Applied Biosystems).

qPCR was performed on a StepOnePlus Real-Time PCR machine. For each condition we used two biological replicates (ChIP material from two different Knock-in Cell lines) which were measured in technical duplicates. The following qPCR primers were used:

Nanog_TSS: F 5-CATGAGTGTGGGTCTTCCTG, R 5-CAGAATTCGATGCTTCCTCA;
Nanog_TES: F 5-AGTTGGAATGGCCTGGTAAC, R 5-TCCAGTTCTTGACTGGATGC;
Rara_TSS: F 5-GGAATCTAAGCCACGGAGAA, R 5-GTCACCGAGCACTTCCTTTT;
Rara_TES: F 5-ATTCTGGGAAAGGAGAGCTG, R 5-CAAGGACTGGCTTCCTCTCT;
Rps19_TSS: F 5- GCATGGGTTTGGATTTATGG, R 5-AGGCCTCAGTTCCAACAAAC;
Rps19_TES: F 5-GCTGCCAACAAGAAGCATT, R 5-ACCAAAACCAGATCCAGACC;
Rpl6_TSS: F 5-TAACCAGCAGGAGTTCGTCA, R 5-TGCTCTAGCACGGGACACTA;
Rpl6_TES: F 5-GGCTTTGGATTAGAGGCACA, R 5-GCAAAACCACAGCCACAGTA;
Phc1_TSS: F 5-GCAAAAGATGTGGGGAGAAA, R 5-ACTCTGGGACCGAACAAATG;
Phc1_TES: F 5-AGGGTCACCAGCCATCATAG, R 5-TGCAGAAAGAATGGCAACAG;
Slc2a3_TSS: F 5-TGCTCCGGTTTCTCCTAAGT, R 5-TATCGAATTGCTTGCTGGAG;
Slc2a3_TES: F 5-TCAGACCCTTTAGGCAAACC, R 5-CACAGGGAGTGGAGAGTTGA;
GD_Chr6: F 5-ACGCCATATACAGCAAACCA, R 5-GCCTTGACTTGTCCTGATT

Immunofluorescent staining. *Immunofluorescence staining was performed by Marlena Lübbe.* A modified protocol from Hycult biotech (Version: 04-2010) was used for IF staining. Neurons were grown on coverslips and on neuronal differentiation day 12, the cells were fixed in PBS with 2% paraformaldehyde (Electron Microscopy Sciences) for 20 min at RT, rinsed three times with PBS, and permeabilized with ice-cold methanol for 10 min at -20 °C. Cells were incubated with primary antibody in PBS+2% BSA for 60 min and then incubated with 1:200 diluted Alexa Fluor 594-conjugated secondary antibody for 30 min at room temperature. Cells were then counterstained with 100 ng/ml DAPI for 5 min and coverslips were mounted on slides with Prolong Gold antifade reagent (Life Technologies). Primary Antibodies in this study include MAP2 (M9942, Sigma), Sox2 (AB5603, UBI) and secondary antibodies Alexa Fluor 594 goat α -rabbit or α -mouse (A11012 and A11005, Life Technologies). Images were taken with an Olympus ScanR (IX81, cube4, 10x UplanSApo NA 0.4 Air, Dapi: Ex: 347/50 EM: 460/50 (AHF F31-000 Set Dapi), Cy3: Ex: 545/30 EM: 610/72 (AHF F46-04 Cy3 ET Set)). Image analysis was done in Fiji. For the depicted images the contrast was enhanced as follows: MAP2 (0.7%)/DAPI (2.5%), Sox2 (1%)/DAPI (3 %).

Cell cycle analysis. *Cell cycle analysis was performed by Marlena Lübbe.* The percentage of proliferating cells in S-Phase was measured using the The Click-iT Plus EdU Alexa Fluor 488 Flow Cytometry Assay Kit (Invitrogen, C10632) in ESCs (D0), NPCs (D8), immature neurons (D9) and mature neurons (D12). Cells in 6-well plate were incubated in media supplemented with 10 μ M EdU for 1 hour at 37°C. As a negative staining control, DMSO was used instead of EdU. ESCs were harvested by trypsiniza-

tion. Neurons were harvested by removing half the media, adding cOmplete protease inhibitor (P.I., Roche, 11873580001) to a final concentration of 1x, and using a cell lifter to detach the neurons. The cells were washed once with PBS/1% (w/v) BSA. For fixation, 100 μ l Click-iT fixative was added to the cell pellets for 15 min at room temperature in the dark. Cells were stored temporarily in PBS/1% (v/v) formaldehyde (Merck, 1.04003.1000) at 4°C until analysis. For permeabilization cell pellets were re-suspended in 1 x Click-iT saponin-based permeabilization and wash reagent for up to 15 min. The Click-iT Plus reaction cocktail (PBS, copper protectant, fluorescent dye picolyl azide, reaction buffer additive) was added for 30 min at room temperature in the dark, followed by a wash with 1 x Click-iT saponin-based permeabilization and wash reagent. Cells were then stained for DNA content by adding a final concentration of 1 μ g/ μ l DAPI in 500 μ l 1 x Click-iT saponin-based permeabilization and wash buffer. On D0 and D12 the cells were incubated at 4 °C for 2 h before FACS analysis and an additional measurement was performed after 3 days. On D8 and D9, the DAPI staining was done solely overnight. A BD LSRFortessa was used for FACS and FlowJo for subsequent data analysis.

Histone extraction and mass spectrometry analysis of histone modifications. *I collected and purified histone samples and mass spectrometry data collection and analysis was performed by Simone Sidoli.* Histones were purified from 10^7 ESCs by acidic extraction as described by Sidoli and Garcia (2017); Sidoli et al. (2014). Briefly, cell nuclei were isolated by using a nuclear isolation buffer enriched with protease/deacetylase/phosphatase inhibitors and 0.2% NP-40. Histones were extracted using 0.2M H₂SO₄, and they were precipitated with 33% trichloroacetic acid. Histones were digested with the protease GluC at an enzyme:sample ratio of 1:20 at room temperature overnight in 5 mM ammonium acetate (pH:4.0). Samples were loaded onto a 150x0.075 mm nanocolumn in-house packed using Hypercarb™ Porous Graphitic Carbon (Thermo Scientific) using an EasyLC-1000 nanoHPLC (Thermo Scientific, San Jose, CA, USA) coupled online with an Orbitrap Fusion mass spectrometer (Thermo). Histone tails were separated using a 90 min gradient from 10% to 15% buffer B (buffer A: 0.1% formic acid, buffer B: 80% acetonitrile, 0.1% formic acid). The flowrate for the analysis was set to 300 nL/min. Data acquisition was performed in the Orbitrap for both precursor and product ions, with a mass resolution of 60,000 for MS and 30,000 for MS/MS. MS acquisition window was set at 665-705 m/z. Only charge state 8+ was accepted for MS/MS fragmentation. The isolation width was set at 2 m/z. The most intense ions with MS signal higher than 5,000 counts were isolated for fragmentation using ETD with an activation time of 20 msec. Data processing was performed as described in Sidoli et al. (2014). Briefly, spectra were deconvoluted with Xtract (Thermo) and searched with Mascot (v2.5, Matrix Science, London, UK), including mono- and dimethylation (KR), trimethylation (K) and acetylation (K) as dynamic modifications. The histone sequence database was customized to include mutated forms. Mass tolerance was set to

2.1 Da for precursor mass and 0.01 Da for product mass. Mascot results were filtered for unambiguous identifications and peptides were quantified using the latest version of isoScale (Sidoli et al., 2017). For turnover experiment in Fig. 3.19, ESCs were grown for 24 hours in ESC medium supplemented with heavy L-Arginine-HCl (13C- and 15N-labeled amino acid, Thermo Fisher 89989) prior to harvest. Histones were cleaved by tryptic digestion and analyzed by mass spectrometry as described previously. The ratio of heavy (newly synthesized histones) and light (old histones) was calculated to determine turnover.

H3.3-FLAG-Immunoprecipitation. ESC lines stably expression tagged H3.3 were generated using CRISPR-mediated integration into the Rosa26 gene locus (Perez-Pinera et al., 2012; Chu et al., 2016). We generated a MCS-Rosa26 plasmid (37200, Addgene) containing cDNA of H3.3 (wildtype or mutant) with a C-terminal HA-FLAG-tag and a self-cleaved P2A-GFP. A Guide sequence targeting Rosa26 locus19 was cloned into pSpCas9(BB)-2A-GFP (PX458, Addgene). For integration 2×10^6 ESCs were transfected with 2 μ g Cas9-GFP plasmid and 2 μ g MCS-Rosa26 plasmid using electroporation (Nucleofector, Lonza). Bulk cell sorting was performed 48 hours and 4 days post-electroporation for GFP positive cells. Single-cell sorting was performed after 8 days post-electroporation. Clonal cell lines were screened for integration of H3.3-HA-FLAG using immunoblotting. ESC were harvested and 10^8 cells were lysed using hypotonic lysis buffer and douncing to release nuclei. DNA was fragmented using MNase digestion in nuclei to obtain mostly mononucleosomes. Soluble fraction was used as input for immunoprecipitation. Magnetic M2 Flag-beads (Thermo Fisher) were added to lysates and incubated over night while rotating. On the next day, Flag-beads were washed and samples were eluted for 30 min at 4°C using Flag-peptide. *Eluates were processed for mass spectrometry analysis by the proteomics facility of EMBL Heidelberg. Specifically, Frank Stein performed the data analysis.* In brief, for quantitative comparison between wildtype and mutant H3.3, samples were labeled using TMT and labelled peptides were measured. Measured peptides were mapped and quantified by MaxQuant software (Tyanova et al., 2016). Data was normalized by applying a variance stabilization normalization method with the R package vsn (v 3.44) and analyzed using R. As negative control, unmodified wild type cells without tagged H3.3 were used (background for knock-in cell lines).

Plasmid construction and HEK293T cell ectopic expression. Wildtype H3.3 or H3.1 cDNA sequences were cloned into pBluescript(+) to introduce point-mutations at H3K4 and H3K36 using standard site-directed mutagenesis techniques. For expression in HEK293T cells, cDNA sequences were subcloned into pcDNA3.1-CMV or Rosa26-EF1 α -MCS. HEK293T cells were transfected with H3.1/H3.3-encoding plasmids using Fugene (Promega) or Lipofectamine3000 (Thermo Fisher) reagent. Cells were harvested after 48 hr post-transfection for protein expression analysis.

Protein extraction. Whole cell lysates were prepared by resuspending cell pellets in SDS loading dye and DNA was fragmented by sonication with EpiShear Sonicator (Active Motif). For subcellular fractionation into cytoplasm, nucleoplasm and chromatin, one 10 cm plate of HEK293T cells or 10×10^6 ESCs were fractionated according to Conrad et al. (2014) with small adjustments. To pellet nuclei or chromatin, centrifugation speed was reduced to 1000xg. Chromatin pellet was resuspended directly in SDS loading dye to extract proteins and samples were sonicated until viscosity was reduced. Nucleoplasm and cytoplasm fraction were diluted in SDS-loading dye and sonicated if required. For crude nuclear extraction, cells were resuspended in PBS+0.05% Triton-X and rotated for 30 min at 4°C to lyse cells. Nuclei were pelleted by centrifugation at 1000xg for 5 min at 4°C, resuspended in SDS loading dye and sonicated. For inhibition of the proteasome, HEK293T cells or mESCs were treated at a final concentration of 5 μ M MG-132 for 4 hours prior to harvesting.

Immunoblotting. Prepared lysates were separated in MES buffer on precast 4–12% Bis-Tris SDS-PAGE gels (Thermo) and transferred onto nitrocellulose membranes. Blots were incubated overnight with primary antibodies, for 1 hour with HRP-conjugated secondary antibody and developed using enhanced chemiluminescent substrate (Thermo Fisher). In this study we used the following antibodies: H3K4me1 (39297, active motif), H3K4me3 (39915, active motif), H3K27ac (39685, Active Motif), H3K27me3 (9733, Cell Signaling), H3K36me3 (ab9050, abcam), H3K36me2 (07-369, Millipore), H3.3 (09-838, Millipore), H3 (ab18521, abcam), H4 (ab10158, abcam), FLAG-M2 (F3165, Sigma), HA (ab9110, abcam) and secondary antibodies goat anti-mouse IgG-HRP (1721011, BioRad), goat anti-rabbit IgG-HRP (1706515, BioRad).

4 | Supplementary Material

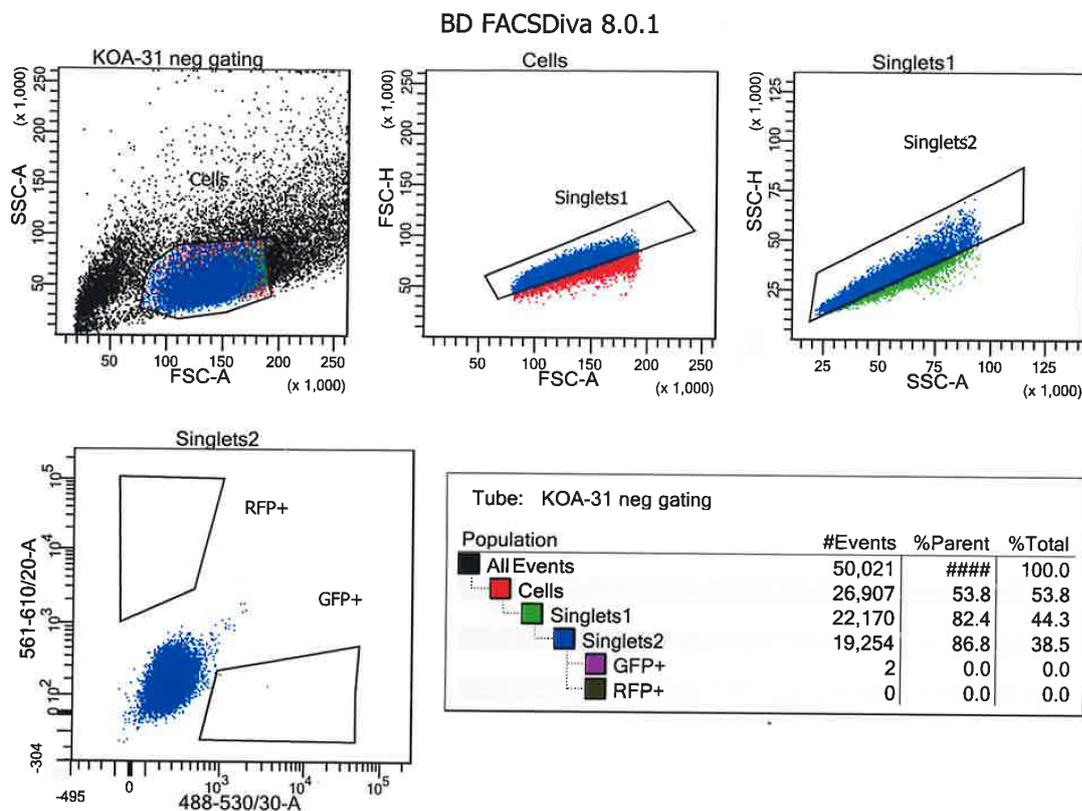


Figure S1: Gating strategy for Cas9-GFP-positive cells flow cytometry analysis. Representative gating strategy for an untransfected wild type cells is displayed. Single cells were chosen for analysis after doublet discrimination by detection of disproportions between cell size (FSC-A) vs. cell signal (FSC-H). The same cell displays higher correlation on the two axis (FSC-A/SSC-A and FSC-H/SSC-H) and all singlet events will fall more on a diagonal than doublets. Transduced GFP-positive cells can be detected outside of the negative population of cells measured with a 488-530 nm laser.

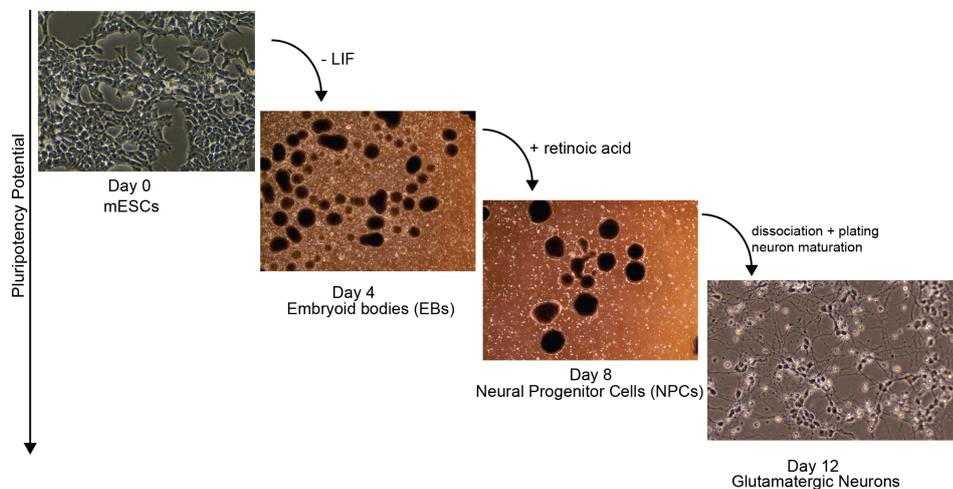


Figure S2: *In vitro* differentiation of ESCs into mature glutamatergic neurons. Bright field images cells during neuronal differentiation into mature glutamatergic neurons according to a modified protocol developed by (Bibel et al., 2007). Cells are grown in suspension for 4 days without LIF, resulting in cell accumulation into embryoid bodies, which grow larger in size over a course of 8 days. The addition of retinoic acid to the medium directs the cells towards the neuronal lineage, yielding neuronal precursor cells on day 8. Neural precursors are dissociated and plated to give rise to terminally differentiated glutamatergic neurons after 12 days of total differentiation.

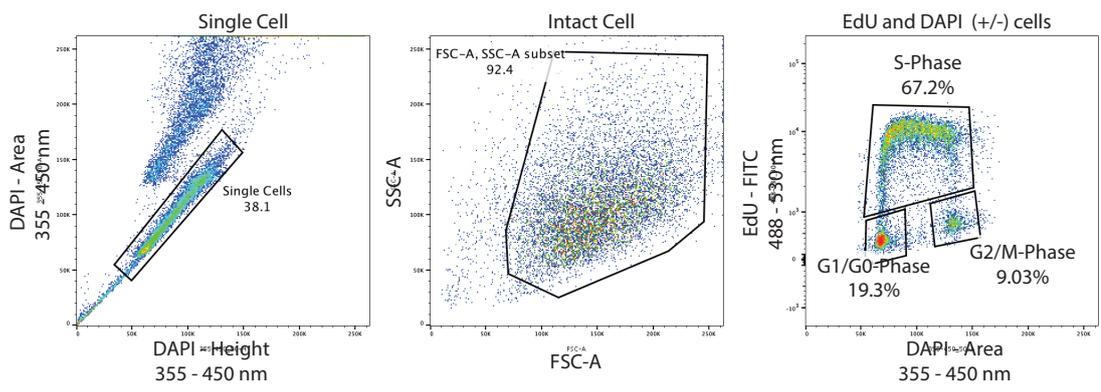


Figure S3: Gating strategy for cell cycle analysis by flow cytometry. Wild type ESCs were labeled with EdU to detect proliferating cells and with DAPI to measure DNA content. Single cells were detected by analysing the DAPI cell area versus the DAPI height, singlets fall on the diagonal between the two axis. Intact cells were selected by comparison of cell size (FSC-A) and cell granularity (SSC-A). For detection of cell cycle phase, EdU signal (incorporation of EdU into DNA during S-Phase) was compared to DAPI (DNA content). *Experiment was performed by Marlena Lübbe.*

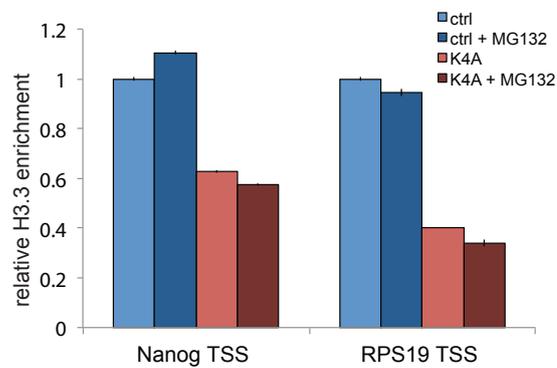


Figure S4: Proteasomal degradation is not responsible for H3.3K4A depletion at TSS. H3.3-ChIP-qPCR results control and K4A mutant ESCs that were either untreated or treated with the proteasome inhibitor MG-132. Relative enrichment of H3.3 was measured at the promoter (TSS) of the highly expressed genes Nanog and Rps19. H3.3K4A is depleted compared to control and depletion is not rescued by MG-132 treatment. Enrichment of H3.3 is depicted relative to the untreated control. Ct values were normalized to a intergenic region. *Experiment was performed by Marlena Lübkke.*

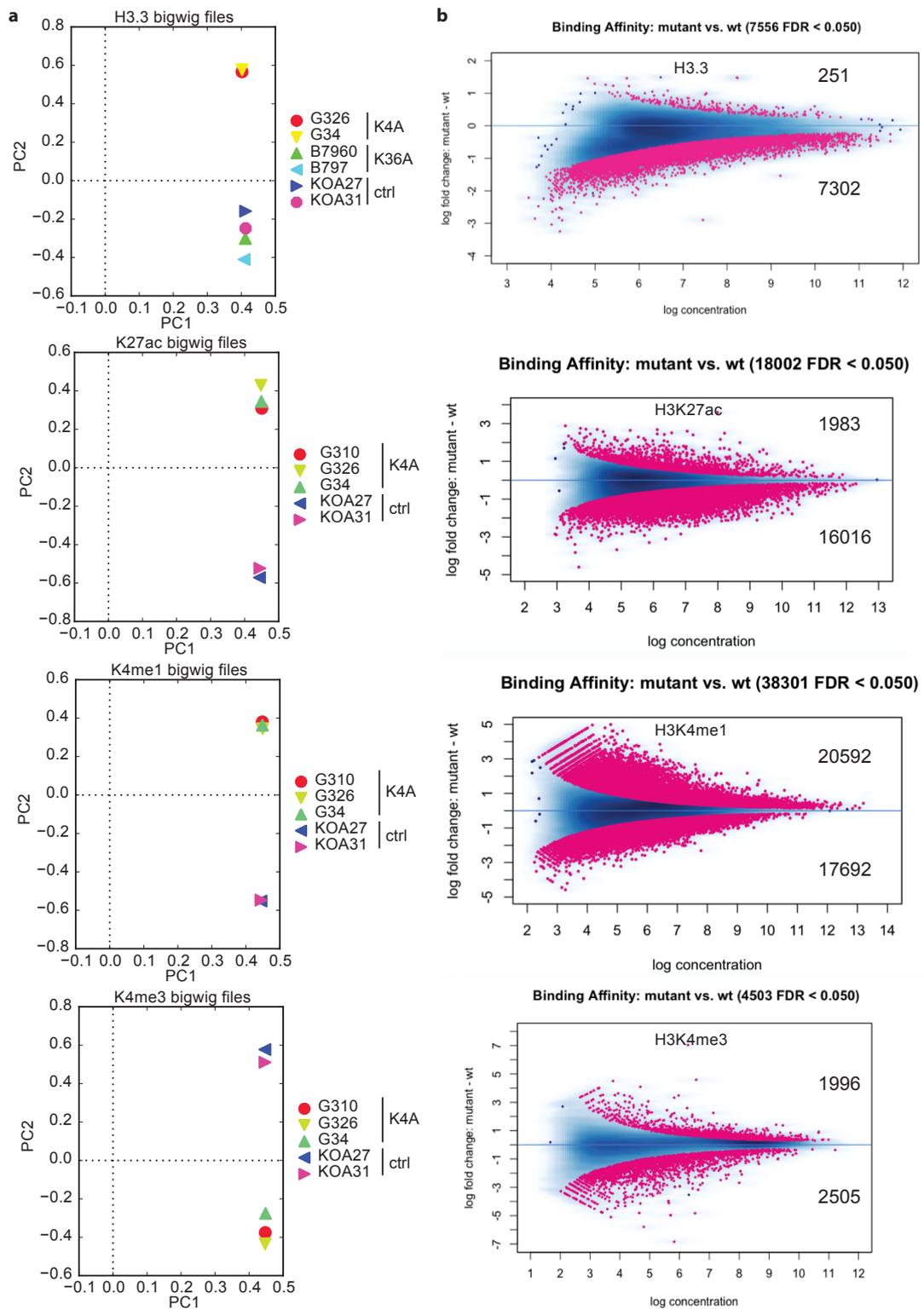


Figure S5: Quality control of ChIP-Seq data. (a) PCA analysis plots displaying differences/similarities between samples for H3.3-, H3K27ac-, H3K4me1- and H3K4me3-ChIP-Seq experiments. PCA analysis was done on binned and library size normalized bigWig files. (b) MA-plot displaying logFoldChanges in signal intensities in ChIP-Seq data of H3.3K4A mutants compared to controls. Sites that are identified to be significantly differentially (FDR < 0.05) bound are shown in red and were identified using DiffBind package. Total number of significant changes (more or less bound than control) are depicted each plot. *Plots in this figure were generated by Daria Bunina.*

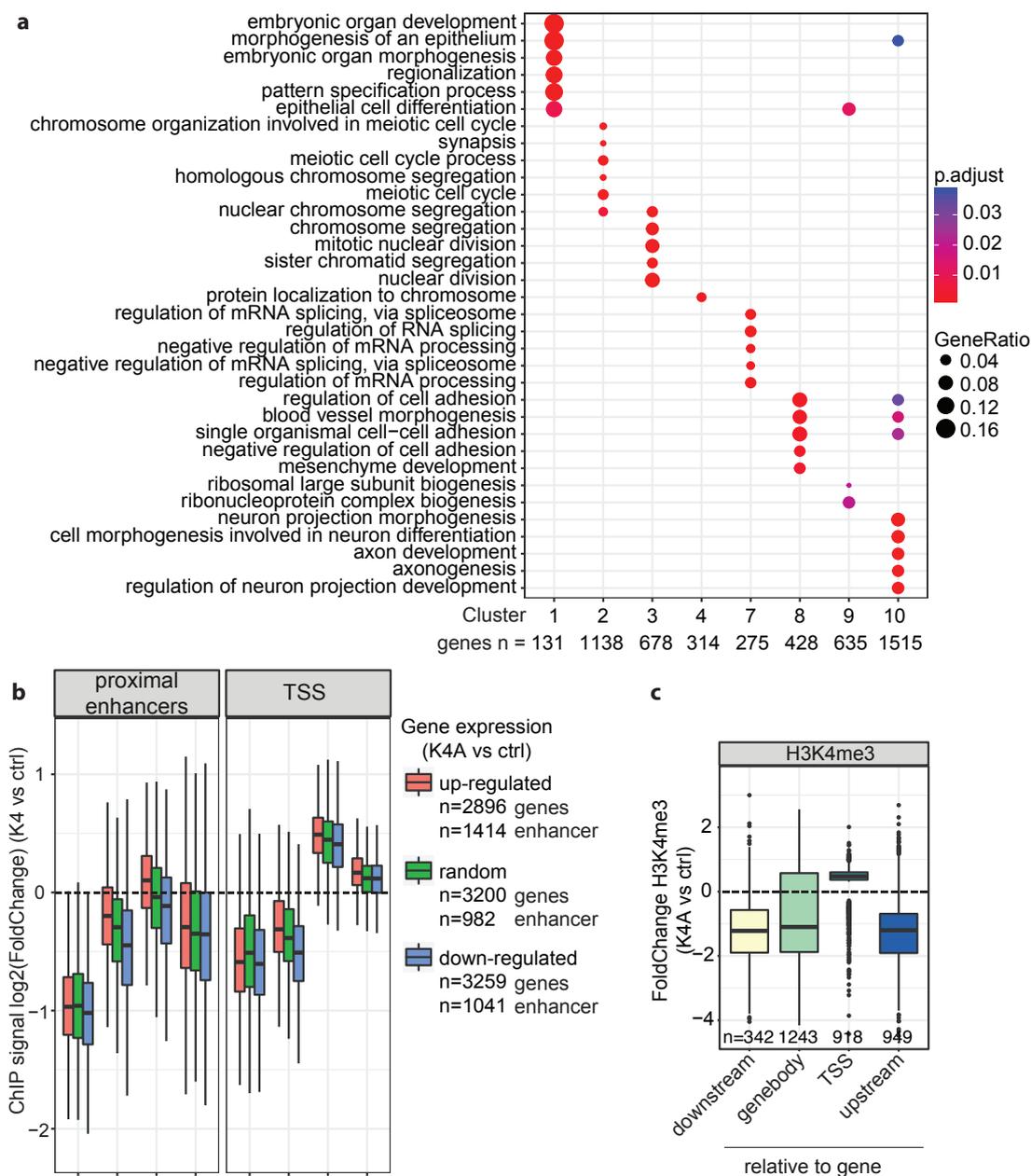


Figure S6: Promoters and enhancers of differentially expressed and unchanged genes are similarly affected by H3.3 depletion, but H3K27ac reduction correlates with gene expression in H3.3K4A mutants. (a) Gene ontology terms (biological processes) significantly enriched in the indicated clusters of differentially expressed genes from the heatmap in Fig. 3.21. (b) Distribution of histone marks at enhancers and TSS in the vicinity of either differentially expressed genes in K4A mutants (up- or down-regulated) or unchanged genes (random). TSS represent ± 1.5 kb around gene start site and promoter-proximal enhancers are ESC-specific enhancers overlapping region 0 to 50kb upstream of the gene start site. Boxplots show the distribution of the log2 fold change of the average ChIP-Seq signal between K4A mutant and control cell lines. Boxes display median log2 fold change, third and first quartile, whiskers show $1.5 \times$ the interquartile range above and below the box. (c) Boxplot displaying significant FoldChanges of H3K4me3 signal in H3.3K4A mutant relative to controls. Significant differentially methylated regions of H3K4me3 were detected using the DiffBind package (FDR < 0.05) and grouped based on their location relative to the nearest gene. Numbers of significant differential peaks are depicted each box. Boxes display median fold change, third and first quartile, whiskers show $1.5 \times$ the interquartile range above and below the box, dots represent outliers. *Daria Bunina performed clustering analysis of combined RNA-Seq and ChIP-Seq data and generated plots seen in (a) and (b).*

Bibliography

Ahmad, K. and S. Henikoff

2002. The histone variant h3.3 marks active chromatin by replication-independent nucleosome assembly. *Molecular Cell*, 9(6):1191–1200.

Almouzni, G. and H. Cedar

2016. Maintenance of epigenetic information. *Cold Spring Harb Perspect Biol*, 8(5).

Anders, S., A. Reyes, and W. Huber

2012. Detecting differential usage of exons from rna-seq data. *Genome Res*, 22(10):2008–17.

Annunziato, A.

2008. Dna packaging: Nucleosomes and chromatin.

Awad, S. and A. H. Hassan

2008. The swi2/snf2 bromodomain is important for the full binding and remodeling activity of the swi/snf complex on h3- and h4-acetylated nucleosomes. *Ann N Y Acad Sci*, 1138:366–75.

Aygun, O., S. Mehta, and S. I. Grewal

2013. Hdac-mediated suppression of histone turnover promotes epigenetic stability of heterochromatin. *Nat Struct Mol Biol*, 20(5):547–54.

Banaszynski, L., D. Wen, S. Dewell, S. Whitcomb, M. Lin, N. Diaz, S. Elsässer, A. Chapgier, A. Goldberg, E. Canaani, S. Rafii, D. Zheng, and C. Allis

2013. Hira-dependent histone h3.3 deposition facilitates prc2 recruitment at developmental loci in es cells. *Cell*, 155(1):107–120.

Bannister, A. J. and T. Kouzarides

2011. Regulation of chromatin by histone modifications. *Cell Res*, 21(3):381–95.

Bannister, A. J., R. Schneider, F. A. Myers, A. W. Thorne, C. Crane-Robinson, and T. Kouzarides

2005. Spatial distribution of di- and tri-methyl lysine 36 of histone h3 at active genes. *J Biol Chem*, 280(18):17732–6.

Bannister, A. J., P. Zegerman, J. F. Partridge, E. A. Miska, J. O. Thomas, R. C. Allshire, and T. Kouzarides

2001. Selective recognition of methylated lysine 9 on histone h3 by the hp1 chromo domain. *Nature*, 410(6824):120–4.

Bibliography

- Barski, A., S. Cuddapah, K. Cui, T. Y. Roh, D. E. Schones, Z. Wang, G. Wei, I. Chepelev, and K. Zhao
2007. High-resolution profiling of histone methylations in the human genome. *Cell*, 129(4):823–37.
- Baumann, C., M. M. Viveiros, and R. De La Fuente
2010. Loss of maternal atrx results in centromere instability and aneuploidy in the mammalian oocyte and pre-implantation embryo. *PLoS Genet*, 6(9):e1001137.
- Behjati, S., P. S. Tarpey, N. Presneau, S. Scheipl, N. Pillay, P. Van Loo, D. C. Wedge, S. L. Cooke, G. Gundem, H. Davies, S. Nik-Zainal, S. Martin, S. McLaren, V. Goodie, B. Robinson, A. Butler, J. W. Teague, D. Halai, B. Khatri, O. Myklebost, D. Baumhoer, G. Jundt, R. Hamoudi, R. Tirabosco, M. F. Amary, P. A. Futreal, M. R. Stratton, P. J. Campbell, and A. M. Flanagan
2013. Distinct h3f3a and h3f3b driver mutations define chondroblastoma and giant cell tumor of bone. *Nat Genet*, 45(12):1479–82.
- Belotserkovskaya, R., S. Oh, V. A. Bondarenko, G. Orphanides, V. M. Studitsky, and D. Reinberg
2003. Fact facilitates transcription-dependent nucleosome alteration. *Science*, 301(5636):1090–3.
- Bhattacharya, D., S. Talwar, A. Mazumder, and G. V. Shivashankar
2009. Spatio-temporal plasticity in chromatin organization in mouse cell differentiation and during drosophila embryogenesis. *Biophys J*, 96(9):3832–9.
- Bibel, M., J. Richter, E. Lacroix, and Y. A. Barde
2007. Generation of a defined and uniform population of cns progenitors and neurons from mouse embryonic stem cells. *Nat Protoc*, 2(5):1034–43.
- Bickmore, W. A. and K. S. Zaret
2010. Altered states: how gene expression is changed during differentiation. *Curr Opin Genet Dev*, 20(5):467–9.
- Bonisch, C. and S. B. Hake
2012. Histone h2a variants in nucleosomes and chromatin: more or less stable? *Nucleic Acids Res*, 40(21):10719–41.
- Bonn, S., R. P. Zinzen, C. Girardot, E. H. Gustafson, A. Perez-Gonzalez, N. Delhomme, Y. Ghavi-Helm, B. Wilczynski, A. Riddell, and E. E. Furlong
2012. Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. *Nat Genet*, 44(2):148–56.
- Bonnefoy, E., G. A. Orsi, P. Couble, and B. Loppin
2007. The essential role of drosophila hira for de novo assembly of paternal chromatin at fertilization. *PLoS Genet*, 3(10):1991–2006.
- Briggs, S. D., M. Bryk, B. D. Strahl, W. L. Cheung, J. K. Davie, S. Y. Dent, F. Winston, and C. D. Allis
2001. Histone h3 lysine 4 methylation is mediated by set1 and required for cell growth and rdna silencing in saccharomyces cerevisiae. *Genes Dev*, 15(24):3286–95.

- Brown, D. A., V. Di Cerbo, A. Feldmann, J. Ahn, S. Ito, N. P. Blackledge, M. Nakayama, M. McClellan, E. Dimitrova, A. H. Turberfield, H. K. Long, H. W. King, S. Kriaucionis, L. Schermelleh, T. G. Kutateladze, H. Koseki, and R. J. Klose
2017. The set1 complex selects actively transcribed target genes via multivalent interaction with cpg island chromatin. *Cell Rep*, 20(10):2313–2327.
- Burgess, R. J. and Z. Zhang
2013. Histone chaperones in nucleosome assembly and human disease. *Nat Struct Mol Biol*, 20(1):14–22.
- Bush, K. M., B. T. Yuen, B. L. Barrilleaux, J. W. Riggs, H. O’Geen, R. F. Cotterman, and P. S. Knoepfler
2013. Endogenous mammalian histone h3.3 exhibits chromatin-related functions during development. *Epigenetics Chromatin*, 6(1):7.
- Carrozza, M. J., B. Li, L. Florens, T. Suganuma, S. K. Swanson, K. K. Lee, W. J. Shia, S. Anderson, J. Yates, M. P. Washburn, and J. L. Workman
2005. Histone h3 methylation by set2 directs deacetylation of coding regions by rpd3s to suppress spurious intragenic transcription. *Cell*, 123(4):581–92.
- Cecconi, M., F. Forzano, D. Milani, S. Cavani, C. Baldo, A. Selicorni, C. Pantaleoni, M. Silengo, G. B. Ferrero, G. Scarano, M. Della Monica, R. Fischetto, P. Grammatico, S. Majore, G. Zampino, L. Memo, E. L. Cordisco, G. Neri, M. Pierluigi, F. D. Bricarelli, M. Grasso, and F. Faravelli
2005. Mutation analysis of the nsd1 gene in a group of 59 patients with congenital overgrowth. *Am J Med Genet A*, 134(3):247–53.
- Champagne, K. S. and T. G. Kutateladze
2009. Structural insight into histone recognition by the ing phd fingers. *Curr Drug Targets*, 10(5):432–41.
- Chan, K. M., D. Fang, H. Gan, R. Hashizume, C. Yu, M. Schroeder, N. Gupta, S. Mueller, C. D. James, R. Jenkins, J. Sarkaria, and Z. Zhang
2013. The histone h3.3k27m mutation in pediatric glioma reprograms h3k27 methylation and gene expression. *Genes Dev*, 27(9):985–90.
- Charpentier, E. and J. A. Doudna
2013. Biotechnology: Rewriting a genome. *Nature*, 495(7439):50–1.
- Chen, F., S. M. Pruetz-Miller, Y. Huang, M. Gjoka, K. Duda, J. Taunton, T. N. Collingwood, M. Frodin, and G. D. Davis
2011. High-frequency genome editing using ssdna oligonucleotides with zinc-finger nucleases. *Nat Methods*, 8(9):753–5.
- Chen, P., L. Dong, M. Hu, Y. Z. Wang, X. Xiao, Z. Zhao, J. Yan, P. Y. Wang, D. Reinberg, M. Li, W. Li, and G. Li
2018. Functions of fact in breaking the nucleosome and maintaining its integrity at the single-nucleosome level. *Mol Cell*, 71(2):284–293.e4.

Bibliography

- Cheung, N. K., J. Zhang, C. Lu, M. Parker, A. Bahrami, S. K. Tickoo, A. Heguy, A. S. Pappo, S. Federico, J. Dalton, I. Y. Cheung, L. Ding, R. Fulton, J. Wang, X. Chen, J. Becksfort, J. Wu, C. A. Billups, D. Ellison, E. R. Mardis, R. K. Wilson, J. R. Downing, and M. A. Dyer
2012. Association of age at diagnosis and genetic mutations in patients with neuroblastoma. *Jama*, 307(10):1062–71.
- Chu, V. T., T. Weber, R. Graf, T. Sommermann, K. Petsch, U. Sack, P. Volchkov, K. Rajewsky, and R. Kuhn
2016. Efficient generation of rosa26 knock-in mice using crispr/cas9 in c57bl/6 zygotes. *BMC Biotechnol*, 16:4.
- Chu, V. T., T. Weber, B. Wefers, W. Wurst, S. Sander, K. Rajewsky, and R. Kuhn
2015. Increasing the efficiency of homology-directed repair for crispr-cas9-induced precise gene editing in mammalian cells. *Nat Biotechnol*, 33(5):543–8.
- Clapier, C. R. and B. R. Cairns
2009. The biology of chromatin remodeling complexes. *Annu Rev Biochem*, 78:273–304.
- Clouaire, T., S. Webb, and A. Bird
2014. Cfp1 is required for gene expression-dependent h3k4 trimethylation and h3k9 acetylation in embryonic stem cells. *Genome Biol*, 15(9):451.
- Clouaire, T., S. Webb, P. Skene, R. Illingworth, A. Kerr, R. Andrews, J. H. Lee, D. Skalnik, and A. Bird
2012. Cfp1 integrates both cpg content and gene activity for accurate h3k4me3 deposition in embryonic stem cells. *Genes Dev*, 26(15):1714–28.
- Conrad, T., A. Marsico, M. Gehre, and U. A. Orom
2014. Microprocessor activity controls differential mirna biogenesis in vivo. *Cell Rep*, 9(2):542–54.
- Couldrey, C., M. B. Carlton, P. M. Nolan, W. H. Colledge, and M. J. Evans
1999. A retroviral gene trap insertion into the histone 3.3a gene causes partial neonatal lethality, stunted growth, neuromuscular deficits and male sub-fertility in transgenic mice. *Hum Mol Genet*, 8(13):2489–95.
- Creyghton, M. P., A. W. Cheng, G. G. Welstead, T. Kooistra, B. W. Carey, E. J. Steine, J. Hanna, M. A. Lodato, G. M. Frampton, P. A. Sharp, L. A. Boyer, R. A. Young, and R. Jaenisch
2010. Histone h3k27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A*, 107(50):21931–6.
- Dai, J., E. M. Hyland, D. S. Yuan, H. Huang, J. S. Bader, and J. D. Boeke
2008. Probing nucleosome function: a highly versatile library of synthetic histone h3 and h4 mutants. *Cell*, 134(6):1066–78.
- de Almeida, S. F., A. R. Grosso, F. Koch, R. Fenouil, S. Carvalho, J. Andrade, H. Levezinho, M. Gut, D. Eick, I. Gut, J. C. Andrau, P. Ferrier, and M. Carmo-Fonseca
2011. Splicing enhances recruitment of methyltransferase hypb/setd2 and methylation of histone h3 lys36. *Nat Struct Mol Biol*, 18(9):977–83.

- de Dieuleveult, M., K. Yen, I. Hmitou, A. Depaux, F. Boussouar, D. Bou Dargham, S. Jounier, H. Humbertclaude, F. Ribierre, C. Baulard, N. P. Farrell, B. Park, C. Keime, L. Carriere, S. Berlivet, M. Gut, I. Gut, M. Werner, J. F. Deleuze, R. Olaso, J. C. Aude, S. Chantalat, B. F. Pugh, and M. Gerard
2016. Genome-wide nucleosome specificity and function of chromatin remodellers in es cells. *Nature*, 530(7588):113–6.
- de Wit, E., U. Braunschweig, F. Greil, H. J. Bussemaker, and B. van Steensel
2008. Global chromatin domain organization of the drosophila genome. *PLoS Genet*, 4(3):e1000045.
- Deal, R. B., J. G. Henikoff, and S. Henikoff
2010. Genome-wide kinetics of nucleosome turnover determined by metabolic labeling of histones. *Science*, 328(5982):1161–4.
- Deaton, A. M., M. Gomez-Rodriguez, J. Mieczkowski, M. Y. Tolstorukov, S. Kundu, R. I. Sadreyev, L. E. Jansen, and R. E. Kingston
2016. Enhancer regions show high histone h3.3 turnover that changes during differentiation. *Elife*, 5.
- Delbarre, E., K. Ivanauskiene, T. Kuntziger, and P. Collas
2013. Daxx-dependent supply of soluble (h3.3-h4) dimers to pml bodies pending deposition into chromatin. *Genome Res*, 23(3):440–51.
- Dhalluin, C., J. E. Carlson, L. Zeng, C. He, A. K. Aggarwal, and M. M. Zhou
1999. Structure and ligand of a histone acetyltransferase bromodomain. *Nature*, 399(6735):491–6.
- Dhayalan, A., A. Rajavelu, P. Rathert, R. Tamas, R. Z. Jurkowska, S. Ragozin, and A. Jeltsch
2010. The dnmt3a pwwp domain reads histone 3 lysine 36 trimethylation and guides dna methylation. *J Biol Chem*, 285(34):26114–20.
- Dillon, S. C., X. Zhang, R. C. Trievel, and X. Cheng
2005. The set-domain protein superfamily: protein lysine methyltransferases. *Genome Biol*, 6(8):227.
- Dion, M. F., T. Kaplan, M. Kim, S. Buratowski, N. Friedman, and O. J. Rando
2007. Dynamics of replication-independent histone turnover in budding yeast. *Science*, 315(5817):1405–8.
- Dorigi, K. M., T. Swigut, T. Henriques, N. V. Bhanu, B. S. Scruggs, N. Nady, n. Still, C. D., B. A. Garcia, K. Adelman, and J. Wysocka
2017. Mll3 and mll4 facilitate enhancer rna synthesis and transcription from promoters independently of h3k4 monomethylation. *Mol Cell*, 66(4):568–576.e4.
- Douglas, J., S. Hanks, I. K. Temple, S. Davies, A. Murray, M. Upadhyaya, S. Tomkins, H. E. Hughes, T. R. Cole, and N. Rahman
2003. Nsd1 mutations are the major cause of sotos syndrome and occur in some cases of weaver syndrome but are rare in other overgrowth phenotypes. *Am J Hum Genet*, 72(1):132–43.

Bibliography

- Drane, P., K. Ouararhni, A. Depaux, M. Shuaib, and A. Hamiche
2010. The death-associated protein daxx is a novel histone chaperone involved in the replication-independent deposition of h3.3. *Genes Dev*, 24(12):1253–65.
- Dugast-Darzacq, C. and T. Grange
2009. Methylquant: a real-time pcr-based method to quantify dna methylation at single specific cytosines. *Methods Mol Biol*, 507:281–303.
- Dunleavy, E., A. Pidoux, and R. Allshire
2005. Centromeric chromatin makes its mark. *Trends Biochem Sci*, 30(4):172–5.
- Edmunds, J. W., L. C. Mahadevan, and A. L. Clayton
2008. Dynamic histone h3 methylation during gene induction: Hypb/setd2 mediates all h3k36 trimethylation. *Embo j*, 27(2):406–20.
- Elsaesser, S. J., A. D. Goldberg, and C. D. Allis
2010. New functions for an old variant: no substitute for histone h3.3. *Curr Opin Genet Dev*, 20(2):110–7.
- Elsasser, S. J., K. M. Noh, N. Diaz, C. D. Allis, and L. A. Banaszynski
2015. Histone h3.3 is required for endogenous retroviral element silencing in embryonic stem cells. *Nature*, 522(7555):240–4.
- ENCODE
2012. An integrated encyclopedia of dna elements in the human genome. *Nature*, 489(7414):57–74.
- Fang, D., H. Gan, J. H. Lee, J. Han, Z. Wang, S. M. Riester, L. Jin, J. Chen, H. Zhou, J. Wang, H. Zhang, N. Yang, E. W. Bradley, T. H. Ho, B. P. Rubin, J. A. Bridge, S. N. Thibodeau, T. Ordog, Y. Chen, A. J. van Wijnen, A. M. Oliveira, R. M. Xu, J. J. Westendorf, and Z. Zhang
2016. The histone h3.3k36m mutation reprograms the epigenome of chondroblastomas. *Science*, 352(6291):1344–8.
- Fang, R., A. J. Barbera, Y. Xu, M. Rutenberg, T. Leonor, Q. Bi, F. Lan, P. Mei, G. C. Yuan, C. Lian, J. Peng, D. Cheng, G. Sui, U. B. Kaiser, Y. Shi, and Y. G. Shi
2010. Human lsd2/kdm1b/aof1 regulates gene transcription by modulating intragenic h3k4me2 methylation. *Mol Cell*, 39(2):222–33.
- Feng, Q., H. Wang, H. H. Ng, H. Erdjument-Bromage, P. Tempst, K. Struhl, and Y. Zhang
2002. Methylation of h3-lysine 79 is mediated by a new family of hmtases without a set domain. *Curr Biol*, 12(12):1052–8.
- Ferrari, P. and M. Strubin
2015. Uncoupling histone turnover from transcription-associated histone h3 modifications. *Nucleic Acids Res*, 43(8):3972–85.
- Filion, G. J., J. G. van Bommel, U. Braunschweig, W. Talhout, J. Kind, L. D. Ward, W. Brugman, I. J. de Castro, R. M. Kerkhoven, H. J. Bussemaker, and B. van Steensel
2010. Systematic protein location mapping reveals five principal chromatin types in drosophila cells. *Cell*, 143(2):212–24.

- Fischle, W., Y. Wang, S. A. Jacobs, Y. Kim, C. D. Allis, and S. Khorasanizadeh
2003. Molecular basis for the discrimination of repressive methyl-lysine marks in histone h3 by polycomb and hp1 chromodomains. *Genes Dev*, 17(15):1870–81.
- Flanagan, J. F., L. Z. Mi, M. Chruszcz, M. Cymborowski, K. L. Clines, Y. Kim, W. Minor, F. Rastinejad, and S. Khorasanizadeh
2005. Double chromodomains cooperate to recognize the methylated histone h3 tail. *Nature*, 438(7071):1181–5.
- Goldberg, A. D., L. A. Banaszynski, K. M. Noh, P. W. Lewis, S. J. Elsaesser, S. Stadler, S. Dewell, M. Law, X. Guo, X. Li, D. Wen, A. Chapgier, R. C. DeKever, J. C. Miller, Y. L. Lee, E. A. Boydston, M. C. Holmes, P. D. Gregory, J. M. Greally, S. Rafii, C. Yang, P. J. Scambler, D. Garrick, R. J. Gibbons, D. R. Higgs, I. M. Cristea, F. D. Urnov, D. Zheng, and C. D. Allis
2010. Distinct factors control histone variant h3.3 localization at specific genomic regions. *Cell*, 140(5):678–91.
- Groth, A., W. Rocha, A. Verreault, and G. Almouzni
2007. Chromatin challenges during dna replication and repair. *Cell*, 128(4):721–33.
- Guenther, M. G., S. S. Levine, L. A. Boyer, R. Jaenisch, and R. A. Young
2007. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell*, 130(1):77–88.
- Gunesdogan, U., H. Jackle, and A. Herzig
2010. A genetic system to assess in vivo the functions of histones and histone modifications in higher eukaryotes. *EMBO Rep*, 11(10):772–6.
- Gurard-Levin, Z. A., J. P. Quivy, and G. Almouzni
2014. Histone chaperones: assisting histone traffic and nucleosome dynamics. *Annu Rev Biochem*, 83:487–517.
- Haberle, V. and A. Stark
2018. Eukaryotic core promoters and the functional basis of transcription initiation. *Nat Rev Mol Cell Biol*, 19(10):621–637.
- Hake, S. B. and C. D. Allis
2006. Histone h3 variants and their potential role in indexing mammalian genomes: the "h3 barcode hypothesis". *Proc Natl Acad Sci U S A*, 103(17):6428–35.
- Hake, S. B., B. A. Garcia, M. Kauer, S. P. Baker, J. Shabanowitz, D. F. Hunt, and C. D. Allis
2005. Serine 31 phosphorylation of histone variant h3.3 is specific to regions bordering centromeres in metaphase chromosomes. *Proc Natl Acad Sci U S A*, 102(18):6344–9.
- Harris, M. E., R. Bohni, M. H. Schneiderman, L. Ramamurthy, D. Schumperli, and W. F. Marzluff
1991. Regulation of histone mrna in the unperturbed cell cycle: evidence suggesting control at two posttranscriptional steps. *Mol Cell Biol*, 11(5):2416–24.
- Hassan, A. H., P. Prochasson, K. E. Neely, S. C. Galasinski, M. Chandy, M. J. Carrozza, and J. L. Workman
2002. Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. *Cell*, 111(3):369–79.

- Hauk, G., J. N. McKnight, I. M. Nodelman, and G. D. Bowman
2010. The chromodomains of the chd1 chromatin remodeler regulate dna access to the atpase motor. *Mol Cell*, 39(5):711–23.
- Heaphy, C. M., R. F. de Wilde, Y. Jiao, A. P. Klein, B. H. Edil, C. Shi, C. Bettegowda, F. J. Rodriguez, C. G. Eberhart, S. Hebbar, G. J. Offerhaus, R. McLendon, B. A. Rasheed, Y. He, H. Yan, D. D. Bigner, S. M. Oba-Shinjo, S. K. Marie, G. J. Riggins, K. W. Kinzler, B. Vogelstein, R. H. Hruban, A. Maitra, N. Papadopoulos, and A. K. Meeker
2011. Altered telomeres in tumors with atrx and daxx mutations. *Science*, 333(6041):425.
- Heintzman, N. D., R. K. Stuart, G. Hon, Y. Fu, C. W. Ching, R. D. Hawkins, L. O. Barrera, S. Van Calcar, C. Qu, K. A. Ching, W. Wang, Z. Weng, R. D. Green, G. E. Crawford, and B. Ren
2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet*, 39(3):311–8.
- Henikoff, S.
2008. Nucleosome destabilization in the epigenetic regulation of gene expression. *Nat Rev Genet*, 9(1):15–26.
- Henikoff, S. and M. M. Smith
2015. Histone variants and epigenetics. *Cold Spring Harb Perspect Biol*, 7(1):a019364.
- Herz, H.-M., M. Morgan, X. Gao, J. Jackson, R. Rickels, S. K. Swanson, L. Florens, M. P. Washburn, J. C. Eissenberg, and A. Shilatifard
2014. Histone h3 lysine-to-methionine mutants as a paradigm to study chromatin signaling. *Science*, 345(6200):1065–1070.
- Hochstrasser, M., D. Mathog, Y. Gruenbaum, H. Saumweber, and J. W. Sedat
1986. Spatial organization of chromosomes in the salivary gland nuclei of drosophila melanogaster. *J Cell Biol*, 102(1):112–23.
- Hodl, M. and K. Basler
2009. Transcription in the absence of histone h3.3. *Curr Biol*, 19(14):1221–6.
- Hodl, M. and K. Basler
2012. Transcription in the absence of histone h3.2 and h3k4 methylation. *Curr Biol*, 22(23):2253–7.
- Hsieh, F. K., O. I. Kulaeva, S. S. Patel, P. N. Dyer, K. Luger, D. Reinberg, and V. M. Studitsky
2013. Histone chaperone fact action during transcription through chromatin by rna polymerase ii. *Proc Natl Acad Sci U S A*, 110(19):7654–9.
- Hu, M., X. J. Sun, Y. L. Zhang, Y. Kuang, C. Q. Hu, W. L. Wu, S. H. Shen, T. T. Du, H. Li, F. He, H. S. Xiao, Z. G. Wang, T. X. Liu, H. Lu, Q. H. Huang, S. J. Chen, and Z. Chen
2010. Histone h3 lysine 36 methyltransferase hypb/setd2 is required for embryonic vascular remodeling. *Proc Natl Acad Sci U S A*, 107(7):2956–61.
- Huang, C. and B. Zhu
2014. H3.3 turnover: a mechanism to poise chromatin for transcription, or a response to open chromatin? *Bioessays*, 36(6):579–84.

- Hyun, K., J. Jeon, K. Park, and J. Kim
2017. Writing, erasing and reading histone lysine methylations. *Exp Mol Med*, 49(4):e324.
- Iacovino, M., D. Bosnakovski, H. Fey, D. Rux, G. Bajwa, E. Mahen, A. Mitanoska, Z. Xu, and M. Kyba
2011. Inducible cassette exchange: a rapid and efficient system enabling conditional gene expression in embryonic stem and primary cells. *Stem Cells*, 29(10):1580–8.
- Imhof, A., X. J. Yang, V. V. Ogryzko, Y. Nakatani, A. P. Wolffe, and H. Ge
1997. Acetylation of general transcription factors by histone acetyltransferases. *Curr Biol*, 7(9):689–92.
- Iwase, S., B. Xiang, S. Ghosh, T. Ren, P. W. Lewis, J. C. Cochrane, C. D. Allis, D. J. Picketts, D. J. Patel, H. Li, and Y. Shi
2011. Atrx add domain links an atypical histone methylation recognition mechanism to human mental-retardation syndrome. *Nat Struct Mol Biol*, 18(7):769–76.
- Jacobson, R. H., A. G. Ladurner, D. S. King, and R. Tjian
2000. Structure and function of a human tafii250 double bromodomain module. *Science*, 288(5470):1422–5.
- Jaju, R. J., C. Fidler, O. A. Haas, A. J. Strickson, F. Watkins, K. Clark, N. C. Cross, J. F. Cheng, P. D. Aplan, L. Kearney, J. Boulwood, and J. S. Wainscoat
2001. A novel gene, nsd1, is fused to nup98 in the t(5;11)(q35;p15.5) in de novo childhood acute myeloid leukemia. *Blood*, 98(4):1264–7.
- Jamai, A., A. Puglisi, and M. Strubin
2009. Histone chaperone spt16 promotes redeposition of the original h3-h4 histones evicted by elongating rna polymerase. *Mol Cell*, 35(3):377–83.
- Jang, C. W., Y. Shibata, J. Starmer, D. Yee, and T. Magnuson
2015. Histone h3.3 maintains genome integrity during mammalian development. *Genes Dev*, 29(13):1377–92.
- Jenuwein, T.
2001. Re-set-ting heterochromatin by histone methyltransferases. *Trends Cell Biol*, 11(6):266–73.
- Jiao, Y., C. Shi, B. H. Edil, R. F. de Wilde, D. S. Klimstra, A. Maitra, R. D. Schulick, L. H. Tang, C. L. Wolfgang, M. A. Choti, V. E. Velculescu, J. Diaz, L. A., B. Vogelstein, K. W. Kinzler, R. H. Hruban, and N. Papadopoulos
2011. Daxx/atrx, men1, and mtor pathway genes are frequently altered in pancreatic neuroendocrine tumors. *Science*, 331(6021):1199–203.
- Jin, C., C. Zang, G. Wei, K. Cui, W. Peng, K. Zhao, and G. Felsenfeld
2009. H3.3/h2a.z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nat Genet*, 41(8):941–5.
- Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, and E. Charpentier
2012. A programmable dual-rna-guided dna endonuclease in adaptive bacterial immunity. *Science*, 337(6096):816–21.

Bibliography

- Kasten, M., H. Szerlong, H. Erdjument-Bromage, P. Tempst, M. Werner, and B. R. Cairns
2004. Tandem bromodomains in the chromatin remodeler rsc recognize acetylated histone h3 lys14. *Embo j*, 23(6):1348–59.
- Kharchenko, P. V., A. A. Alekseyenko, Y. B. Schwartz, A. Minoda, N. C. Riddle, J. Ernst, P. J. Sabo, E. Larschan, A. A. Gorchakov, T. Gu, D. Linder-Basso, A. Plachetka, G. Shanower, M. Y. Tolstorukov, L. J. Luquette, R. Xi, Y. L. Jung, R. W. Park, E. P. Bishop, T. K. Canfield, R. Sandstrom, R. E. Thurman, D. M. MacAlpine, J. A. Stamatoyannopoulos, M. Kellis, S. C. Elgin, M. I. Kuroda, V. Pirrotta, G. H. Karpen, and P. J. Park
2011. Comprehensive analysis of the chromatin landscape in drosophila melanogaster. *Nature*, 471(7339):480–5.
- Kim, J., J. Daniel, A. Espejo, A. Lake, M. Krishna, L. Xia, Y. Zhang, and M. T. Bedford
2006. Tudor, mbt and chromo domains gauge the degree of lysine methylation. *EMBO Rep*, 7(4):397–403.
- Kim, S., H. Kim, N. Fong, B. Erickson, and D. L. Bentley
2011. Pre-mrna splicing is a determinant of histone h3k36 methylation. *Proc Natl Acad Sci U S A*, 108(33):13564–9.
- Kim, T. and S. Buratowski
2009. Dimethylation of h3k4 by set1 recruits the set3 histone deacetylase complex to 5' transcribed regions. *Cell*, 137(2):259–72.
- Kim, T. K., M. Hemberg, J. M. Gray, A. M. Costa, D. M. Bear, J. Wu, D. A. Harmin, M. Laptewicz, K. Barbara-Haley, S. Kuersten, E. Markenscoff-Papadimitriou, D. Kuhl, H. Bito, P. F. Worley, G. Kreiman, and M. E. Greenberg
2010. Widespread transcription at neuronal activity-regulated enhancers. *Nature*, 465(7295):182–7.
- Kireeva, M. L., W. Walter, V. Tchernajenko, V. Bondarenko, M. Kashlev, and V. M. Studitsky
2002. Nucleosome remodeling induced by rna polymerase ii: loss of the h2a/h2b dimer during transcription. *Mol Cell*, 9(3):541–52.
- Kizer, K. O., H. P. Phatnani, Y. Shibata, H. Hall, A. L. Greenleaf, and B. D. Strahl
2005. A novel domain in set2 mediates rna polymerase ii interaction and couples histone h3 k36 methylation with transcript elongation. *Mol Cell Biol*, 25(8):3305–16.
- Kolasinska-Zwierz, P., T. Down, I. Latorre, T. Liu, X. S. Liu, and J. Ahringer
2009. Differential chromatin marking of introns and expressed exons by h3k36me3. *Nat Genet*, 41(3):376–81.
- Konev, A. Y., M. Tribus, S. Y. Park, V. Podhraski, C. Y. Lim, A. V. Emelyanov, E. Vershilova, V. Pirrotta, J. T. Kadonaga, A. Lusser, and D. V. Fyodorov
2007. Chd1 motor protein is required for deposition of histone variant h3.3 into chromatin in vivo. *Science*, 317(5841):1087–90.
- Kornberg, R. D. and J. O. Thomas
1974. Chromatin structure; oligomers of the histones. *Science*, 184(4139):865–8.

- Kosicki, M., K. Tomberg, and A. Bradley
2018. Repair of double-strand breaks induced by crispr-cas9 leads to large deletions and complex rearrangements. *Nat Biotechnol*, 36(8):765–771.
- Kouzarides, T.
2007. Chromatin modifications and their function. *Cell*, 128(4):693–705.
- Kraushaar, D. C., W. Jin, A. Maunakea, B. Abraham, M. Ha, and K. Zhao
2013. Genome-wide incorporation dynamics reveal distinct categories of turnover for the histone variant h3.3. *Genome Biol*, 14(10):R121.
- Krogan, N. J., M. C. Keogh, N. Datta, C. Sawa, O. W. Ryan, H. Ding, R. A. Haw, J. Pootoolal, A. Tong, V. Canadien, D. P. Richards, X. Wu, A. Emili, T. R. Hughes, S. Buratowski, and J. F. Greenblatt
2003. A snf2 family atpase complex required for recruitment of the histone h2a variant htz1. *Mol Cell*, 12(6):1565–76.
- Krude, T.
1995. Chromatin. nucleosome assembly during dna replication. *Curr Biol*, 5(11):1232–4.
- Kulaeva, O. I., F. K. Hsieh, and V. M. Studitsky
2010. Rna polymerase complexes cooperate to relieve the nucleosomal barrier and evict histones. *Proc Natl Acad Sci U S A*, 107(25):11325–30.
- Kurotaki, N., K. Imaizumi, N. Harada, M. Masuno, T. Kondoh, T. Nagai, H. Ohashi, K. Naritomi, M. Tsukahara, Y. Makita, T. Sugimoto, T. Sonoda, T. Hasegawa, Y. Chinen, H. A. Tomita Ha, A. Kinoshita, T. Mizuguchi, K. Yoshiura Ki, T. Ohta, T. Kishino, Y. Fukushima, N. Niikawa, and N. Matsumoto
2002. Haploinsufficiency of nsd1 causes sotos syndrome. *Nat Genet*, 30(4):365–6.
- Lachner, M., D. O'Carroll, S. Rea, K. Mechtler, and T. Jenuwein
2001. Methylation of histone h3 lysine 9 creates a binding site for hp1 proteins. *Nature*, 410(6824):116–20.
- Langmead, B. and S. L. Salzberg
2012. Fast gapped-read alignment with bowtie 2. *Nat Methods*, 9(4):357–9.
- Lawrence, M., S. Daujat, and R. Schneider
2016. Lateral thinking: How histone modifications regulate gene expression. *Trends Genet*, 32(1):42–56.
- Lee, H. and J. S. Kim
2018. Unexpected crispr on-target effects. *Nat Biotechnol*, 36(8):703–704.
- Lewis, P. W., S. J. Elsaesser, K. M. Noh, S. C. Stadler, and C. D. Allis
2010. Daxx is an h3.3-specific histone chaperone and cooperates with atrx in replication-independent chromatin assembly at telomeres. *Proc Natl Acad Sci U S A*, 107(32):14075–80.
- Lewis, P. W., M. M. Muller, M. S. Koletsky, F. Cordero, S. Lin, L. A. Banaszynski, B. A. Garcia, T. W. Muir, O. J. Becher, and C. D. Allis
2013. Inhibition of prc2 activity by a gain-of-function h3 mutation found in pediatric glioblastoma. *Science*, 340(6134):857–61.

Bibliography

- Li, B., M. Carey, and J. L. Workman
2007. The role of chromatin during transcription. *Cell*, 128(4):707–19.
- Li, B., L. Howe, S. Anderson, r. Yates, J. R., and J. L. Workman
2003. The set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of rna polymerase ii. *J Biol Chem*, 278(11):8897–903.
- Li, G. and D. Reinberg
2011. Chromatin higher-order structures and gene regulation. *Curr Opin Genet Dev*, 21(2):175–86.
- Li, J., D. Moazed, and S. P. Gygi
2002. Association of the histone methyltransferase set2 with rna polymerase ii plays a role in transcription elongation. *J Biol Chem*, 277(51):49383–8.
- Lickwar, C. R., F. Mueller, S. E. Hanlon, J. G. McNally, and J. D. Lieb
2012. Genome-wide protein-dna binding dynamics suggest a molecular clutch for transcription factor function. *Nature*, 484(7393):251–5.
- Lifton, R. P., M. L. Goldberg, R. W. Karp, and D. S. Hogness
1978. The organization of the histone genes in drosophila melanogaster: functional and evolutionary implications. *Cold Spring Harb Symp Quant Biol*, 42 Pt 2:1047–51.
- Lin, C. J., M. Conti, and M. Ramalho-Santos
2013. Histone variant h3.3 maintains a decondensed chromatin state essential for mouse preimplantation development. *Development*, 140(17):3624–34.
- Loppin, B., E. Bonnefoy, C. Anselme, A. Laurencon, T. L. Karr, and P. Couble
2005. The histone h3.3 chaperone hira is essential for chromatin assembly in the male pronucleus. *Nature*, 437(7063):1386–90.
- Love, M. I., W. Huber, and S. Anders
2014. Moderated estimation of fold change and dispersion for rna-seq data with deseq2. *Genome Biol*, 15(12):550.
- Lu, C., S. U. Jain, D. Hoelper, D. Bechet, R. C. Molden, L. Ran, D. Murphy, S. Venneti, M. Hameed, B. R. Pawel, J. S. Wunder, B. C. Dickson, S. M. Lundgren, K. S. Jani, N. De Jay, S. Papillon-Cavanagh, I. L. Andrulis, S. L. Sawyer, D. Grynspan, R. E. Turcotte, J. Nadaf, S. Fahiminiyah, T. W. Muir, J. Majewski, C. B. Thompson, P. Chi, B. A. Garcia, C. D. Allis, N. Jabado, and P. W. Lewis
2016. Histone h3k36 mutations promote sarcomagenesis through altered histone methylation landscape. *Science*, 352(6287):844–9.
- Lu, R. and G. G. Wang
2013. Tudor: a versatile family of histone methylation ‘readers’. *Trends Biochem Sci*, 38(11):546–55.
- Luco, R. F., Q. Pan, K. Tominaga, B. J. Blencowe, O. M. Pereira-Smith, and T. Misteli
2010. Regulation of alternative splicing by histone modifications. *Science*, 327(5968):996–1000.

- Luger, K., A. W. Mader, R. K. Richmond, D. F. Sargent, and T. J. Richmond
1997. Crystal structure of the nucleosome core particle at 2.8 a resolution. *Nature*, 389(6648):251–60.
- Luk, E., A. Ranjan, P. C. Fitzgerald, G. Mizuguchi, Y. Huang, D. Wei, and C. Wu
2010. Stepwise histone replacement by swr1 requires dual activation with histone h2a.z and canonical nucleosome. *Cell*, 143(5):725–36.
- Makarova, K. S., D. H. Haft, R. Barrangou, S. J. Brouns, E. Charpentier, P. Horvath, S. Moineau, F. J. Mojica, Y. I. Wolf, A. F. Yakunin, J. van der Oost, and E. V. Koonin
2011. Evolution and classification of the crispr-cas systems. *Nat Rev Microbiol*, 9(6):467–77.
- Margueron, R., G. Li, K. Sarma, A. Blais, J. Zavadil, C. L. Woodcock, B. D. Dynlacht, and D. Reinberg
2008. Ezh1 and ezh2 maintain repressive chromatin through different mechanisms. *Mol Cell*, 32(4):503–18.
- Maruyama, T., S. K. Dougan, M. C. Truttmann, A. M. Bilate, J. R. Ingram, and H. L. Ploegh
2015. Increasing the efficiency of precise genome editing with crispr-cas9 by inhibition of nonhomologous end joining. *Nat Biotechnol*, 33(5):538–42.
- Marzluff, W. F., P. Gongidi, K. R. Woods, J. Jin, and L. J. Maltais
2002. The human and mouse replication-dependent histone genes. *Genomics*, 80(5):487–498.
- Masumoto, H., D. Hawke, R. Kobayashi, and A. Verreault
2005. A role for cell-cycle-regulated histone h3 lysine 56 acetylation in the dna damage response. *Nature*, 436(7048):294–8.
- Mattout, A. and E. Meshorer
2010. Chromatin plasticity and genome organization in pluripotent embryonic stem cells. *Curr Opin Cell Biol*, 22(3):334–41.
- Maurer-Stroh, S., N. J. Dickens, L. Hughes-Davies, T. Kouzarides, F. Eisenhaber, and C. P. Ponting
2003. The tudor domain 'royal family': Tudor, plant agenet, chromo, pwwp and mbt domains. *Trends Biochem Sci*, 28(2):69–74.
- Maze, I., W. Wenderski, K. M. Noh, R. C. Bagot, N. Tzavaras, I. Purushothaman, S. J. Elsassner, Y. Guo, C. Ionete, Y. L. Hurd, C. A. Tamminga, T. Halene, L. Farrelly, A. A. Soshnev, D. Wen, S. Rafii, M. R. Birtwistle, S. Akbarian, B. A. Buchholz, R. D. Blitzler, E. J. Nestler, Z. F. Yuan, B. A. Garcia, L. Shen, H. Molina, and C. D. Allis
2015. Critical role of histone turnover in neuronal transcription and plasticity. *Neuron*, 87(1):77–94.
- McKay, D. J., S. Klusza, T. J. Penke, M. P. Meers, K. P. Curry, S. L. McDaniel, P. Y. Malek, S. W. Cooper, D. C. Tatomer, J. D. Lieb, B. D. Strahl, R. J. Duronio, and A. G. Matera
2015. Interrogating the function of metazoan histones using engineered gene clusters. *Dev Cell*, 32(3):373–86.
- McKittrick, E., P. R. Gafken, K. Ahmad, and S. Henikoff
2004. Histone h3.3 is enriched in covalent modifications associated with active chromatin. *Proc Natl Acad Sci U S A*, 101(6):1525–30.

- Meers, M. P., T. Henriques, C. A. Lavender, D. J. McKay, B. D. Strahl, R. J. Duronio, and K. Adelman
2017. Histone gene replacement reveals a post-transcriptional role for h3k36 in maintaining metazoan transcriptome fidelity. *6*.
- Meshorer, E. and T. Misteli
2006. Chromatin in pluripotent embryonic stem cells and differentiation. *Nat Rev Mol Cell Biol*, 7(7):540–6.
- Michod, D., S. Bartesaghi, A. Khelifi, C. Bellodi, L. Berliocchi, P. Nicotera, and P. Salomoni
2012. Calcium-dependent dephosphorylation of the histone chaperone daxx regulates h3.3 loading and transcription upon neuronal activation. *Neuron*, 74(1):122–35.
- Mikkelsen, T. S., M. Ku, D. B. Jaffe, B. Issac, E. Lieberman, G. Giannoukos, P. Alvarez, W. Brockman, T. K. Kim, R. P. Koche, W. Lee, E. Mendenhall, A. O'Donovan, A. Presser, C. Russ, X. Xie, A. Meissner, M. Wernig, R. Jaenisch, C. Nusbaum, E. S. Lander, and B. E. Bernstein
2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*, 448(7153):553–60.
- Miller, T., N. J. Krogan, J. Dover, H. Erdjument-Bromage, P. Tempst, M. Johnston, J. F. Greenblatt, and A. Shilatifard
2001. Compass: a complex of proteins associated with a trithorax-related set domain protein. *Proc Natl Acad Sci U S A*, 98(23):12902–7.
- Min, J., Y. Zhang, and R. M. Xu
2003. Structural basis for specific binding of polycomb chromodomain to histone h3 methylated at lys 27. *Genes Dev*, 17(15):1823–8.
- Mito, Y., J. G. Henikoff, and S. Henikoff
2005. Genome-scale profiling of histone h3.3 replacement patterns. *Nat Genet*, 37(10):1090–7.
- Mito, Y., J. G. Henikoff, and S. Henikoff
2007. Histone replacement marks the boundaries of cis-regulatory domains. *Science*, 315(5817):1408–11.
- Mizuguchi, G., X. Shen, J. Landry, W. H. Wu, S. Sen, and C. Wu
2004. Atp-driven exchange of histone h2az variant catalyzed by swr1 chromatin remodeling complex. *Science*, 303(5656):343–8.
- Morettini, S., M. Tribus, A. Zeilner, J. Sebald, B. Campo-Fernandez, G. Scheran, H. Worle, V. Podhraski, D. V. Fyodorov, and A. Lusser
2011. The chromodomains of chd1 are critical for enzymatic activity but less important for chromatin localization. *Nucleic Acids Res*, 39(8):3103–15.
- Nakanishi, S., B. W. Sanderson, K. M. Delventhal, W. D. Bradford, K. Staehling-Hampton, and A. Shilatifard
2008. A comprehensive library of histone mutants identifies nucleosomal residues required for h3k4 methylation. *Nat Struct Mol Biol*, 15(8):881–8.
- Nakayama, T., K. Nishioka, Y. X. Dong, T. Shimojima, and S. Hirose
2007. Drosophila gaga factor directs histone h3.3 replacement that prevents the heterochromatin spreading. *Genes Dev*, 21(5):552–61.

- Narlikar, G. J., R. Sundaramoorthy, and T. Owen-Hughes
2013. Mechanisms and functions of atp-dependent chromatin-remodeling enzymes. *Cell*, 154(3):490–503.
- Ng, H. H., F. Robert, R. A. Young, and K. Struhl
2003. Targeted recruitment of set1 histone methylase by elongating pol ii provides a localized mark and memory of recent transcriptional activity. *Mol Cell*, 11(3):709–19.
- Nimura, K., K. Ura, H. Shiratori, M. Ikawa, M. Okabe, R. J. Schwartz, and Y. Kaneda
2009. A histone h3 lysine 36 trimethyltransferase links nkx2-5 to wolf-hirschhorn syndrome. *Nature*, 460(7252):287–91.
- Pengelly, A. R., O. Copur, H. Jackle, A. Herzig, and J. Muller
2013. A histone mutant reproduces the phenotype caused by loss of histone-modifying factor polycomb. *Science*, 339(6120):698–9.
- Perez-Pinera, P., D. G. Ousterout, M. T. Brown, and C. A. Gersbach
2012. Gene targeting to the rosa26 locus directed by engineered zinc finger nucleases. *Nucleic Acids Res*, 40(8):3741–52.
- Pickersgill, H., B. Kalverda, E. de Wit, W. Talhout, M. Fornerod, and B. van Steensel
2006. Characterization of the drosophila melanogaster genome at the nuclear lamina. *Nat Genet*, 38(9):1005–14.
- Pina, B. and P. Suau
1987. Changes in histones h2a and h3 variant composition in differentiating and mature rat brain cortical neurons. *Dev Biol*, 123(1):51–8.
- Pollex, T. and E. E. M. Furlong
2017. Correlation does not imply causation: Histone methyltransferases, but not histone methylation, set the stage for enhancer activation. *Mol Cell*, 66(4):439–441.
- Postberg, J., S. Forcob, W. J. Chang, and H. J. Lipps
2010. The evolutionary history of histone h3 suggests a deep eukaryotic root of chromatin modifying mechanisms. *BMC Evol Biol*, 10:259.
- Pradhan, S. K., T. Su, L. Yen, K. Jacquet, C. Huang, J. Cote, S. K. Kurdistani, and M. F. Carey
2016. Ep400 deposits h3.3 into promoters and enhancers during gene activation. *Mol Cell*, 61(1):27–38.
- Probst, A. V., E. Dunleavy, and G. Almouzni
2009. Epigenetic inheritance during the cell cycle. *Nat Rev Mol Cell Biol*, 10(3):192–206.
- Rada-Iglesias, A., R. Bajpai, T. Swigut, S. A. Brugmann, R. A. Flynn, and J. Wysocka
2011. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature*, 470(7333):279–83.
- Radman-Livaja, M., T. K. Quan, L. Valenzuela, J. A. Armstrong, T. van Welsem, T. Kim, L. J. Lee, S. Buratowski, F. van Leeuwen, O. J. Rando, and G. A. Hartzog
2012. A key role for chd1 in histone h3 dynamics at the 3' ends of long genes in yeast. *PLoS Genet*, 8(7):e1002811.

Bibliography

- Ramirez, F., D. P. Ryan, B. Gruning, V. Bhardwaj, F. Kilpert, A. S. Richter, S. Heyne, F. Dunder, and T. Manke
2016. deeptools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res*, 44(W1):W160–5.
- Ran, F. A., L. Cong, W. X. Yan, D. A. Scott, J. S. Gootenberg, A. J. Kriz, B. Zetsche, O. Shalem, X. Wu, K. S. Makarova, E. V. Koonin, P. A. Sharp, and F. Zhang
2015. In vivo genome editing using staphylococcus aureus cas9. *Nature*, 520(7546):186–91.
- Ran, F. A., P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott, and F. Zhang
2013. Genome engineering using the crispr-cas9 system. *Nat Protoc*, 8(11):2281–308.
- Ransom, M., B. K. Dennehey, and J. K. Tyler
2010. Chaperoning histones during dna replication and repair. *Cell*, 140(2):183–95.
- Rao, R. C. and Y. Dou
2015. Hijacked in cancer: the kmt2 (mll) family of methyltransferases. *Nat Rev Cancer*, 15(6):334–46.
- Ray-Gallet, D., J. P. Quivy, C. Scamps, E. M. Martini, M. Lipinski, and G. Almouzni
2002. Hira is critical for a nucleosome assembly pathway independent of dna synthesis. *Mol Cell*, 9(5):1091–100.
- Ray-Gallet, D., A. Woolfe, I. Vassias, C. Pellentz, N. Lacoste, A. Puri, D. C. Schultz, N. A. Pchelintsev, P. D. Adams, L. E. Jansen, and G. Almouzni
2011. Dynamics of histone h3 deposition in vivo reveal a nucleosome gap-filling mechanism for h3.3 to maintain chromatin integrity. *Mol Cell*, 44(6):928–41.
- Rickels, R., H. M. Herz, C. C. Sze, K. Cao, M. A. Morgan, C. K. Collings, M. Gause, Y. H. Takahashi, L. Wang, E. J. Rendleman, S. A. Marshall, A. Krueger, E. T. Bartom, A. Piunti, E. R. Smith, N. A. Abshiru, N. L. Kelleher, D. Dorsett, and A. Shilatifard
2017. Histone h3k4 monomethylation catalyzed by trr and mammalian compass-like proteins at enhancers is dispensable for development and viability. *Nat Genet*, 49(11):1647–1653.
- Ricketts, M. D., B. Frederick, H. Hoff, Y. Tang, D. C. Schultz, T. Singh Rai, M. Grazia Vizioli, P. D. Adams, and R. Marmorstein
2015. Ubinuclein-1 confers histone h3.3-specific-binding by the hira histone chaperone complex. *Nat Commun*, 6:7711.
- Roberts, C., H. F. Sutherland, H. Farmer, W. Kimber, S. Halford, A. Carey, J. M. Brickman, A. Wynshaw-Boris, and P. J. Scambler
2002. Targeted mutagenesis of the hira gene results in gastrulation defects and patterning abnormalities of mesoendodermal derivatives prior to early embryonic lethality. *Mol Cell Biol*, 22(7):2318–28.
- Roguev, A., D. Schaft, A. Shevchenko, W. W. Pijnappel, M. Wilm, R. Aasland, and A. F. Stewart
2001. The saccharomyces cerevisiae set1 complex includes an ash2 homologue and methylates histone 3 lysine 4. *Embo j*, 20(24):7137–48.
- Rufiange, A., P. E. Jacques, W. Bhat, F. Robert, and A. Nourani
2007. Genome-wide replication-independent histone h3 exchange occurs predominantly at promoters and implicates h3 k56 acetylation and asf1. *Mol Cell*, 27(3):393–405.

- Saha, A., J. Wittmeyer, and B. R. Cairns
2002. Chromatin remodeling by rsc involves atp-dependent dna translocation. *Genes Dev*, 16(16):2120–34.
- Saha, A., J. Wittmeyer, and B. R. Cairns
2005. Chromatin remodeling through directional dna translocation from an internal nucleosomal site. *Nat Struct Mol Biol*, 12(9):747–55.
- Saha, A., J. Wittmeyer, and B. R. Cairns
2006. Chromatin remodelling: the industrial revolution of dna around histones. *Nat Rev Mol Cell Biol*, 7(6):437–47.
- Sakai, A., B. E. Schwartz, S. Goldstein, and K. Ahmad
2009. Transcriptional and developmental functions of the h3.3 histone variant in drosophila. *Curr Biol*, 19(21):1816–20.
- Santos-Rosa, H., R. Schneider, A. J. Bannister, J. Sherriff, B. E. Bernstein, N. C. Emre, S. L. Schreiber, J. Mellor, and T. Kouzarides
2002. Active genes are tri-methylated at k4 of histone h3. *Nature*, 419(6905):407–11.
- Schmitges, F. W., A. B. Prusty, M. Faty, A. Stutzer, G. M. Lingaraju, J. Aiwazian, R. Sack, D. Hess, L. Li, S. Zhou, R. D. Bunker, U. Wirth, T. Bouwmeester, A. Bauer, N. Ly-Hartig, K. Zhao, H. Chan, J. Gu, H. Gut, W. Fischle, J. Muller, and N. H. Thoma
2011. Histone methylation by prc2 is inhibited by active chromatin marks. *Mol Cell*, 42(3):330–41.
- Schneider, R., A. J. Bannister, F. A. Myers, A. W. Thorne, C. Crane-Robinson, and T. Kouzarides
2004. Histone h3 lysine 4 methylation patterns in higher eukaryotic genes. *Nat Cell Biol*, 6(1):73–7.
- Schroder, S., E. Herker, F. Itzen, D. He, S. Thomas, D. A. Gilchrist, K. Kaehlcke, S. Cho, K. S. Pollard, J. A. Capra, M. Schnolzer, P. A. Cole, M. Geyer, B. G. Bruneau, K. Adelman, and M. Ott
2013. Acetylation of rna polymerase ii regulates growth-factor-induced gene transcription in mammalian cells. *Mol Cell*, 52(3):314–24.
- Schwabish, M. A. and K. Struhl
2004. Evidence for eviction and rapid deposition of histones upon transcriptional elongation by rna polymerase ii. *Mol Cell Biol*, 24(23):10111–7.
- Schwartz, B. E. and K. Ahmad
2005. Transcriptional activation triggers deposition and removal of the histone variant h3.3. *Genes Dev*, 19(7):804–14.
- Schwartzentruber, J., A. Korshunov, X. Y. Liu, D. T. Jones, E. Pfaff, K. Jacob, D. Sturm, A. M. Fontebasso, D. A. Quang, M. Tonjes, V. Hovestadt, S. Albrecht, M. Kool, A. Nantel, C. Konermann, A. Lindroth, N. Jager, T. Rausch, M. Ryzhova, J. O. Korbel, T. Hielscher, P. Hauser, M. Garami, A. Klekner, L. Bognar, M. Ebinger, M. U. Schuhmann, W. Scheurlen, A. Pekrun, M. C. Fruhwald, W. Roggendorf, C. Kramm, M. Durken, J. Atkinson, P. Lepage, A. Montpetit, M. Zakrzewska, K. Zakrzewski, P. P. Liberski, Z. Dong, P. Siegel, A. E. Kulozik, M. Zapatka, A. Guha, D. Malkin, J. Felsberg, G. Reifenberger, A. von Deimling, K. Ichimura, V. P.

- Collins, H. Witt, T. Milde, O. Witt, C. Zhang, P. Castelo-Branco, P. Lichter, D. Faury, U. Tabori, C. Plass, J. Majewski, S. M. Pfister, and N. Jabado
2012. Driver mutations in histone h3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature*, 482(7384):226–31.
- Sekinger, E. A., Z. Moqtaderi, and K. Struhl
2005. Intrinsic histone-dna interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast. *Mol Cell*, 18(6):735–48.
- Shen, L., N. Shao, X. Liu, and E. Nestler
2014. ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. *BMC Genomics*, 15:284.
- Shen, Y., F. Yue, D. F. McCleary, Z. Ye, L. Edsall, S. Kuan, U. Wagner, J. Dixon, L. Lee, V. V. Lobanenkov, and B. Ren
2012. A map of the cis-regulatory sequences in the mouse genome. *Nature*, 488(7409):116–20.
- Shilatifard, A.
2012. The compass family of histone h3k4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annu Rev Biochem*, 81:65–95.
- Shinsky, S. A., K. E. Monteith, S. Viggiano, and M. S. Cosgrove
2015. Biochemical reconstitution and phylogenetic comparison of human set1 family core complexes involved in histone methylation. *J Biol Chem*, 290(10):6361–75.
- Sidoli, S. and B. A. Garcia
2017. Characterization of individual histone posttranslational modifications and their combinatorial patterns by mass spectrometry-based proteomics strategies. *Methods Mol Biol*, 1528:121–148.
- Sidoli, S., C. Lu, M. Coradin, X. Wang, K. R. Karch, C. Ruminowicz, and B. A. Garcia
2017. Metabolic labeling in middle-down proteomics allows for investigation of the dynamics of the histone code. *Epigenetics Chromatin*, 10(1):34.
- Sidoli, S., V. Schwammle, C. Ruminowicz, T. A. Hansen, X. Wu, K. Helin, and O. N. Jensen
2014. Middle-down hybrid chromatography/tandem mass spectrometry workflow for characterization of combinatorial post-translational modifications in histones. *Proteomics*, 14(19):2200–11.
- Sims, R. J., r., C. F. Chen, H. Santos-Rosa, T. Kouzarides, S. S. Patel, and D. Reinberg
2005. Human but not yeast chd1 binds directly and selectively to histone h3 methylated at lysine 4 via its tandem chromodomains. *J Biol Chem*, 280(51):41789–92.
- Smith, S. and B. Stillman
1989. Purification and characterization of caf-i, a human cell factor required for chromatin assembly during dna replication in vitro. *Cell*, 58(1):15–25.
- Smolle, M., S. Venkatesh, M. M. Gogol, H. Li, Y. Zhang, L. Florens, M. P. Washburn, and J. L. Workman
2012. Chromatin remodelers isw1 and chd1 maintain chromatin structure during transcription by preventing histone exchange. *Nat Struct Mol Biol*, 19(9):884–92.

- Stabell, M., J. Larsson, R. B. Aalen, and A. Lambertsson
2007. *Drosophila* dset2 functions in h3-k36 methylation and is required for development. *Biochem Biophys Res Commun*, 359(3):784–9.
- Stec, I., T. J. Wright, G. J. van Ommen, P. A. de Boer, A. van Haeringen, A. F. Moorman, M. R. Altherr, and J. T. den Dunnen
1998. Whsc1, a 90 kb set domain-containing gene, expressed in early development and homologous to a *drosophila* dysmorphia gene maps in the wolf-hirschhorn syndrome critical region and is fused to igh in t(4;14) multiple myeloma. *Hum Mol Genet*, 7(7):1071–82.
- Streubel, G., A. Watson, S. G. Jammula, A. Scelfo, D. J. Fitzpatrick, G. Oliviero, R. McCole, E. Conway, E. Glancy, G. L. Negri, E. Dillon, K. Wynne, D. Pasini, N. J. Krogan, A. P. Bracken, and G. Cagney
2018. The h3k36me2 methyltransferase nsd1 demarcates prc2-mediated h3k27me2 and h3k27me3 domains in embryonic stem cells. *Mol Cell*, 70(2):371–379.e5.
- Sun, X. J., J. Wei, X. Y. Wu, M. Hu, L. Wang, H. H. Wang, Q. H. Zhang, S. J. Chen, Q. H. Huang, and Z. Chen
2005. Identification and characterization of a novel human histone h3 lysine 36-specific methyltransferase. *J Biol Chem*, 280(42):35261–71.
- Svensson, J. P., M. Shukla, V. Menendez-Benito, U. Norman-Axelsson, P. Audergon, I. Sinha, J. C. Tanny, R. C. Allshire, and K. Ekwall
2015. A nucleosome turnover map reveals that the stability of histone h4 lys20 methylation depends on histone recycling in transcribed chromatin. *Genome Res*, 25(6):872–83.
- Tagami, H., D. Ray-Gallet, G. Almouzni, and Y. Nakatani
2004. Histone h3.1 and h3.3 complexes mediate nucleosome assembly pathways dependent or independent of dna synthesis. *Cell*, 116(1):51–61.
- Talbert, P. B. and S. Henikoff
2010. Histone variants—ancient wrap artists of the epigenome. *Nat Rev Mol Cell Biol*, 11(4):264–75.
- Tamura, T., M. Smith, T. Kanno, H. Dasenbrock, A. Nishiyama, and K. Ozato
2009. Inducible deposition of the histone variant h3.3 in interferon-stimulated genes. *J Biol Chem*, 284(18):12217–25.
- Tang, M. C., S. A. Jacobs, D. M. Mattiske, Y. M. Soh, A. N. Graham, A. Tran, S. L. Lim, D. F. Hudson, P. Kalitsis, M. K. O'Bryan, L. H. Wong, and J. R. Mann
2015. Contribution of the two genes encoding histone variant h3.3 to viability and fertility in mice. *PLoS Genet*, 11(2):e1004964.
- Tang, M. C., S. A. Jacobs, L. H. Wong, and J. R. Mann
2013. Conditional allelic replacement applied to genes encoding the histone variant h3.3 in the mouse. *Genesis*, 51(2):142–6.
- Tatton-Brown, K., J. Douglas, K. Coleman, G. Baujat, T. R. Cole, S. Das, D. Horn, H. E. Hughes, I. K. Temple, F. Faravelli, D. Waggoner, S. Turkmen, V. Cormier-Daire, A. Irrthum, and

Bibliography

- N. Rahman
2005. Genotype-phenotype associations in sotos syndrome: an analysis of 266 individuals with *nsd1* aberrations. *Am J Hum Genet*, 77(2):193–204.
- Tenney, K. and A. Shilatifard
2005. A compass in the voyage of defining the role of trithorax/*mll*-containing complexes: linking leukemogenesis to covalent modifications of chromatin. *J Cell Biochem*, 95(3):429–36.
- Tyanova, S., T. Temu, and J. Cox
2016. The maxquant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc*, 11(12):2301–2319.
- van Leeuwen, F., P. R. Gafken, and D. E. Gottschling
2002. Dot1p modulates silencing in yeast by methylation of the nucleosome core. *Cell*, 109(6):745–56.
- Venkatesh, S. and J. L. Workman
2015. Histone exchange, chromatin structure and the regulation of transcription. *Nat Rev Mol Cell Biol*, 16(3):178–89.
- Wagner, E. J. and P. B. Carpenter
2012. Understanding the language of lys36 methylation at histone h3. *Nat Rev Mol Cell Biol*, 13(2):115–26.
- Wang, H., H. Yang, C. S. Shivalila, M. M. Dawlaty, A. W. Cheng, F. Zhang, and R. Jaenisch
2013. One-step generation of mice carrying mutations in multiple genes by crispr/cas-mediated genome engineering. *Cell*, 153(4):910–8.
- Wang, Z., C. Zang, J. A. Rosenfeld, D. E. Schones, A. Barski, S. Cuddapah, K. Cui, T. Y. Roh, W. Peng, M. Q. Zhang, and K. Zhao
2008. Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet*, 40(7):897–903.
- Weinert, B. T., T. Narita, S. Satpathy, B. Srinivasan, B. K. Hansen, C. Scholz, W. B. Hamilton, B. E. Zucconi, W. W. Wang, W. R. Liu, J. M. Brickman, E. A. Kesicki, A. Lai, K. D. Bromberg, P. A. Cole, and C. Choudhary
2018. Time-resolved analysis reveals rapid dynamics and broad scope of the cbp/p300 acetylome. *Cell*, 174(1):231–244.e12.
- Wen, H., Y. Li, Y. Xi, S. Jiang, S. Stratton, D. Peng, K. Tanaka, Y. Ren, Z. Xia, J. Wu, B. Li, M. C. Barton, W. Li, H. Li, and X. Shi
2014. Zmynd11 links histone h3.3k36me3 to transcription elongation and tumour suppression. *Nature*, 508(7495):263–268.
- Wilhelm, B. T., S. Marguerat, S. Aligianni, S. Codlin, S. Watt, and J. Bahler
2011. Differential patterns of intronic and exonic dna regions with respect to rna polymerase ii occupancy, nucleosome density and h3k36me3 marking in fission yeast. *Genome Biol*, 12(8):R82.
- Winkler, D. D. and K. Luger
2011. The histone chaperone fact: structural insights and mechanisms for nucleosome reorganization. *J Biol Chem*, 286(21):18369–74.

- Winston, F. and C. D. Allis
1999. The bromodomain: a chromatin-targeting module? *Nat Struct Biol*, 6(7):601–4.
- Wirbelauer, C., O. Bell, and D. Schubeler
2005. Variant histone h3.3 is deposited at sites of nucleosomal displacement throughout transcribed genes while active histone modifications show a promoter-proximal bias. *Genes Dev*, 19(15):1761–6.
- Wong, L. H., J. D. McGhie, M. Sim, M. A. Anderson, S. Ahn, R. D. Hannan, A. J. George, K. A. Morgan, J. R. Mann, and K. H. Choo
2010. Atrx interacts with h3.3 in maintaining telomere structural integrity in pluripotent embryonic stem cells. *Genome Res*, 20(3):351–60.
- Wu, G., A. Broniscer, T. A. McEachron, C. Lu, B. S. Paugh, J. Becksfort, C. Qu, L. Ding, R. Huether, M. Parker, J. Zhang, A. Gajjar, M. A. Dyer, C. G. Mullighan, R. J. Gilbertson, E. R. Mardis, R. K. Wilson, J. R. Downing, D. W. Ellison, J. Zhang, and S. J. Baker
2012. Somatic histone h3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. *Nat Genet*, 44(3):251–3.
- Wu, R. S., S. Tsai, and W. M. Bonner
1982. Patterns of histone variant synthesis can distinguish g0 from g1 cells. *Cell*, 31(2 Pt 1):367–74.
- Xu, M., C. Long, X. Chen, C. Huang, S. Chen, and B. Zhu
2010. Partitioning of histone h3-h4 tetramers during dna replication-dependent chromatin assembly. *Science*, 328(5974):94–8.
- Xu, S., S. Grullon, K. Ge, and W. Peng
2014. Spatial clustering for identification of chip-enriched regions (sicer) to map regions of histone methylation patterns in embryonic stem cells. *Methods Mol Biol*, 1150:97–111.
- Xue, Y., R. Gibbons, Z. Yan, D. Yang, T. L. McDowell, S. Sechi, J. Qin, S. Zhou, D. Higgs, and W. Wang
2003. The atrx syndrome protein forms a chromatin-remodeling complex with daxx and localizes in promyelocytic leukemia nuclear bodies. *Proc Natl Acad Sci U S A*, 100(19):10635–40.
- Young, M. D., T. A. Willson, M. J. Wakefield, E. Trounson, D. J. Hilton, M. E. Blewitt, A. Oshlack, and I. J. Majewski
2011. Chip-seq analysis reveals distinct h3k27me3 profiles that correlate with transcriptional activity. *Nucleic Acids Res*, 39(17):7415–27.
- Yu, C., Y. Liu, T. Ma, K. Liu, S. Xu, Y. Zhang, H. Liu, M. La Russa, M. Xie, S. Ding, and L. S. Qi
2015. Small molecules enhance crispr genome editing in pluripotent stem cells. *Cell Stem Cell*, 16(2):142–7.
- Yuan, W., J. Xie, C. Long, H. Erdjument-Bromage, X. Ding, Y. Zheng, P. Tempst, S. Chen, B. Zhu, and D. Reinberg
2009. Heterogeneous nuclear ribonucleoprotein l is a subunit of human kmt3a/set2 complex required for h3 lys-36 trimethylation activity in vivo. *J Biol Chem*, 284(23):15701–7.

- Yuan, W., M. Xu, C. Huang, N. Liu, S. Chen, and B. Zhu
2011. H3k36 methylation antagonizes prc2-mediated h3k27 methylation. *J Biol Chem*, 286(10):7983–9.
- Zetsche, B., J. S. Gootenberg, O. O. Abudayyeh, I. M. Slaymaker, K. S. Makarova, P. Essletzbichler, S. E. Volz, J. Joung, J. van der Oost, A. Regev, E. V. Koonin, and F. Zhang
2015. Cpf1 is a single rna-guided endonuclease of a class 2 crispr-cas system. *Cell*, 163(3):759–71.
- Zhang, H., D. N. Roberts, and B. R. Cairns
2005. Genome-wide dynamics of htz1, a histone h2a variant that poises repressed/basal promoters for activation through histone loss. *Cell*, 123(2):219–31.
- Zhang, X. H., L. Y. Tee, X. G. Wang, Q. S. Huang, and S. H. Yang
2015. Off-target effects in crispr/cas9-mediated genome engineering. *Mol Ther Nucleic Acids*, 4:e264.
- Zhang, Y., T. Liu, C. A. Meyer, J. Eeckhoutte, D. S. Johnson, B. E. Bernstein, C. Nusbaum, R. M. Myers, M. Brown, W. Li, and X. S. Liu
2008. Model-based analysis of chip-seq (macs). *Genome Biol*, 9(9):R137.
- Zhou, V. W., A. Goren, and B. E. Bernstein
2011. Charting histone modifications and the functional organization of mammalian genomes. *Nat Rev Genet*, 12(1):7–18.
- Zlatanova, J. and A. Thakar
2008. H2a.z: view from the top. *Structure*, 16(2):166–79.
- Zofall, M., J. Persinger, S. R. Kassabov, and B. Bartholomew
2006. Chromatin remodeling by isw2 and swi/snf requires dna translocation inside the nucleosome. *Nat Struct Mol Biol*, 13(4):339–46.