

ROLE OF HISTONE VARIANT H3.3 IN TRANSCRIPTION AND MITOTIC PROGRESSION

A DISSERTATION SUBMITTED TO
THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
MOLECULAR BIOLOGY AND GENETICS

By

Ayşegül Örs

April 2017

ROLE OF HISTONE VARIANT H3.3 IN TRANSCRIPTION AND MITOTIC PROGRESSION

By Ayşegül Örs

April 2017

We certify that we have read this dissertation and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy in Molecular Biology and Genetics.

Işık Yuluğ
(Advisor)

Mehmet Öztürk
(Co-advisor)

İhsan Gürsel

Ayşe Elif Erson-Bensan

Uygar Halis Tazebay

Ali Osmay Güre

Approved for the Graduate School of Engineering and Science

Ezhan Karaşan

Director of the Graduate School of Engineering and Science

ABSTRACT

ROLE OF HISTONE VARIANT H3.3 IN TRANSCRIPTION AND MITOTIC PROGRESSION

Ayşegül Örs

Ph.D. in Molecular Biology and Genetics

Advisor: Işık Yuluğ

Co-advisor: Mehmet Öztürk

April 2017

Chromatin structure needs to be dynamic and flexible in order for the eukaryotic cellular processes to function correctly. Incorporation of histone variants into chromatin serves to increase epigenetic plasticity by conferring new structural and functional properties to chromatin. Histone variants are implicated in many cellular processes such as transcription or cell division and their deregulation is involved in tumorigenesis. H3.3 is an evolutionarily well conserved histone variant that differs by only a few amino-acids from its replication-dependent counterparts. With the aim of determining H3.3 function, novel knock-in/conditional knock-out mouse models were established and characterized. In these models, one of the two genes coding for H3.3, *H3f3a* or *H3f3b* has been modified to code for an N-terminal FLAG-FLAG-HA tagged H3.3A or H3.3B which can be depleted upon Cre expression. Nucleosome resolution genome-wide mapping FH-H3.3A and FH-H3.3B determined that H3.3A and H3.3B were similarly enriched at promoter regions and their enrichment levels positively correlated with high expression and gene body enrichment. They were also found enriched in telomeres and some repetitive DNA sequences. In a subset of these repetitive regions H3.3A and H3.3B showed differential enrichment properties. As double H3.3-KO mouse generation resulted lethal, mouse embryonic fibroblasts (MEFs) were isolated from FH-H3.3 mice and transformed. Using a combination of Cre recombinase mediated knock-out and RNA interference technology, a new cellular model was established where H3.3 expression was essentially depleted. Although H3.3 enrichment profiles were indicative of a role in active transcription, whole transcriptome analysis upon single H3.3 depletion in livers and an almost complete H3.3 depletion in MEFs yielded very few differentially regulated genes. Interestingly, H3.3

depleted MEFs showed a high increase in mitotic defects and abnormal nuclear structures. Thus, an important yet often understudied role for H3.3 in genomic maintenance during mitotic progression was highlighted.

Keywords: Histone variants, H3.3, H2A.Z, ChIP-Seq, RNA-Seq, liver, mouse model, transcription, mitotic progression



ÖZET

HİSTON VARYANTI H3.3'ÜN TRANSKRİPSİYONDA VE MİTOZ BÖLÜNME İLERLEMESİNDEKİ ROLÜ

Ayşegül Örs

Moleküler Biyoloji ve Genetik, Doktora

Tez danışmanı: Işık Yuluğ

Eş tez danışmanı: Mehmet Öztürk

Nisan 2017

Ökaryotik hücresel süreçlerin doğru çalışabilmesi için kromatin yapısının dinamik ve esnek olması gerekir. Histon varyantlarının kromatine dahil edilmesi, kromatine yeni yapısal ve işlevsel özellikler kazandırarak, epigenetik esnekliği artırır. Histon varyantları, transkripsiyon veya hücre bölünmesi gibi birçok hücresel fonksiyonların gerçekleştirilmesinde rol alır ve deregülasyonlarının kanserleşmede sürecinde etkisi vardır. H3.3 proteini, replikasyona-bağımlı eşdeğerlerinden sadece birkaç amino asit farklılık gösteren evrimsel olarak iyi korunmuş bir histon varyantıdır. H3.3 fonksiyonunu belirlemek amacıyla, yeni koşullu nakavt (knock-in/ conditional knock-out) fare modelleri oluşturulmuş ve tanımlanmıştır. Bu modellerde, H3.3'ü kodlayan iki genden biri, *H3f3a* veya *H3f3b*, N-terminal etiketli H3.3A veya H3.3B kodlamak üzere modifiye edilmiştir. Ayrıca cre rekombinaz ifadesiyle bu genlerin nakavt edilebilmektedir. Bu çalışmada, karaciğerde, nükleozom çözünürlüğünde, genom çapında FH-H3.3A ve FH-H3.3B yerleşme haritalaması gerçekleştirilmiştir. Bu analiz, H3.3A ve H3.3B'nin promotör bölgelerde benzer şekilde yoğunlaştıklarını ve yoğunluklarının, yüksek gen ekspresyonu ve gen boyunca yoğunlaşma ile pozitif korelasyonda olduğunu göstermiştir. H3.3'ün, ayrıca telomerlerde ve bazı tekrarlayan DNA dizilerinde yoğunlaştığı bulunmuştur. Bazı tekrarlayan bölgelerde H3.3A ve H3.3B, farklı şekilde yoğunlaşmıştır. İki H3.3 geninin de yok edildiği fare elde etme çalışmaları ölümcül sonuçlandığı için, FH-H3.3 farelerinden embriyonik fare fibroblastları (MEF) izole edilmiş ve ölümsüzleştirilmiştir. Cre aracılı nakavt ve RNA enterferans teknolojileri birlikte kullanılarak, H3.3 ifadesinin esas olarak susturulduğu yeni bir hücre modeli geliştirilmiştir. Genom boyunca yoğunlaşma profillerinin H3.3'ün aktif transkripsiyonda bir rolü olduğunu göstermesine rağmen, H3.3

nakavt modelleri üzerine yapılan transkriptom analizleri sonucu çok az sayıda gen etkilenmiştir. İlginç şekilde, H3.3 ifadesi susturulmuş MEF'lerde, mitotik defektler ve anormal nükleer yapılarında yüksek bir artış gözlenmiştir. Böylece, H3.3'ün, mitotik ilerleme sırasındaki genom bütünlüğünün korunmasında önemli ama çoğu zaman az üzerinde durulan bir rol üstlendiği vurgulanmıştır.

Anahtar kelimeler: Histon varyantları, H3.3, H2A.Z, ChIP-Seq, RNA-Seq, karaciğer, fare modeli, MEF, transkripsiyon, mitotik ilerleme.





To my beloved family

Acknowledgements

Firstly, I would like to express my gratitude to my advisor Assoc. Prof. Dr. Işık Yuluğ for taking me on as her PhD student and offering her guidance during the last years of my PhD studies.

I am forever grateful to my co-advisor, Prof Dr. Mehmet Öztürk for his mentorship and for providing me with the opportunity to pursue my research in the Institute of Advanced Biosciences in Grenoble, France. His scientific guidance as well as his immeasurable professional and personal support allowed me to carry out my thesis in the best possible way.

I thank my supervisor Dr. Stefan Dimitrov for welcoming me in his research group at the Institute for Advanced Biosciences in Grenoble, France and providing me with his extensive guidance and the resources to develop my PhD project. He has always been available, supportive, generous and understanding throughout our collaboration and I consider it a privilege to have been part of his research group.

There are no words to describe the extent of my gratitude towards Dr. Kiran Padmanabhan. He has supervised my thesis and has been there with me through the good, the bad and the ugly. His resourceful knowledge and unique enthusiasm along with his everlasting support, guidance and patience allowed me to mature both scientifically and personally. I also thank him for his valuable input in the writing of my dissertation and for taking the time in his busy schedule to attend my defense.

I would like to express my deepest appreciations to the thesis committee members, Prof. Dr. İhsan Gürsel and Assoc. Prof. Dr. Ayşe Elif Erson-Bensan for their availability and valuable suggestions during meetings. Moreover, I would like to thank Prof Dr. Uygur Tazebay and Assoc. Prof. Dr. Ali Osmay Güre for accepting to read and evaluate my dissertation as members of the thesis jury.

A special thanks to our collaborators, Thomas Westerling, Razvan Chereji and especially Christophe Papin in analyzing the ChIP-Seq and RNA-Seq data presented in this study.

I thank Bertrand Favier for his guidance and support with all animal models and experimentation as well as Patrick Vernet for his invaluable help in colony maintenance

and animal experiments. I extend my gratitude to the staff at the animal housing facilities in Plateforme de Haute Technologie Animale (PHTA) and Institute for Advanced Biosciences (IAB). I would also like to spare a line on this page to express my respect to all the mice used in this study and the valuable contribution of laboratory subjects to the advancement of scientific research everywhere.

I am most appreciative of the help and support of Mylène Pezet in the flow cytometry platform and of the staff in the microscopy platform in IAB.

The work presented in this thesis was conducted in part in the Dimitrov - Chromatin and Epigenetics group at the Institute of Advanced Biosciences (IAB) in Grenoble, France. This long collaboration was made possible thanks to funding from Agence National de la Recherche (ANR), the European Molecular Biology Organization (EMBO) short-term fellowship, Scientific and Technological Research Council of Turkey (TUBITAK) 2214/A doctoral research grant and French Ministry of Foreign Affairs scholarship.

During my thesis, I was lucky to be a part of two great laboratory families.

First, my gratitude goes to members of the former Ozturk group at Bilkent-MBG; Haluk Yüzügüllü, Özge Gürsoy Yüzügüllü, Şerif Şentürk, Çiğdem Özen, Dilek Çevik, Gökhan Yıldız, Mustafa Yılmaz, Emre Yurdusev, Hande Topel, Umur Keleş, Engin Demidzen, Derya Soner Cavga, Merve Deniz Abdüsselemoğlu and Yusuf İsmail Ertuna for the collaborative, supportive and positive work and learning environment they helped create during our time in Bilkent-MBG. I would like to sincerely thank Dilek Cevik and Pelin Telkoparan for their priceless friendship and endless support. I especially thank my colleague and house-mate Özlem Tufanlı as well as my other fellow PhD candidates, particularly Verda Bitirim, Gözde Güçlüler and Banu Bayyurt for their moral support and the fun times shared.

Second, a cordial thank you to past and present members of Dimitrov group at IAB; Damien Goutte-Gattat, Véronique Gerson, Geneviève Chevalier, Noémie Mandier, Thierry Gautier, Marc Block, Daniel Bouvard, Anne-Sophie Ribba and Emeline Fontaine with special mention of my fellow PhD candidates, Defne Dalkara, Lorrie Ramos, Yohan Roulland and Hiba Sabra. I particularly thank Sophie Barral for her initial mentorship in teaching me all chromatin based techniques used in the laboratory and for her valuable friendship. Working

in IAB also allowed me to meet amazing people and make great friends and I would like to express my thanks to Ayça Zeybek, Matteo Cattaneo, Hitoshi Shiota, Alexandar Kyumurkov and Mathieu Danguin for their friendship and encouragement over the years.

Whether from a technical, scientific or personal point of view, having such a diverse working environment has trained me for the better and for the worse and was a real pleasure.

I thank my parents, Hülya and Seyhun for all their love, support and patience throughout this thesis and my entire life. They have provided me with the best education possible, encouraged me to always push further and to keep all windows of opportunity open. I appreciate their sacrifices and I would not have been able to get to this stage without them. To my brother Ali Osman, a heartfelt thank you for his emotional, professional and “electronic” support and for being a role model for perfectionism and over-achievement throughout my life. I also wish to thank Onur Dalliag, for his unfailing encouragement throughout these years.

Finally, I extend my gratitude to my other family members and friends in Turkey, in France and all over the world, who have cheered me on along the way and contributed to the successful completion of this thesis. I feel lucky to have a great family, amazing colleagues and lifelong friends.

“Gutta cavat lapidem non bis, sed saepe cadendo; sic homo fit sapiens non bis, sed saepe legendo”

Giordano Bruno, *Candelaio* (1582)

Table of contents

	Page
ABSTRACT	i
ÖZET	iii
Acknowledgements.....	vi
Table of contents	ix
List of Figures.....	xiii
List of Tables	xv
Abbreviations	xvi
Chapter 1. Introduction.....	1
1.1. Structural organization of chromatin	1
1.1.1. Histones	1
1.1.2. Nucleosome Core Particle and the Chromatosome	4
1.1.3. Higher-order organization of chromatin.....	5
1.2. Organization of chromatin function.....	8
1.2.1. Eukaryotic chromosome structure.....	8
1.2.2. The cell cycle and chromatin.....	8
1.2.3. Transcription and chromatin.....	11
1.3. Epigenetic regulation	12
1.3.1. DNA methylation	14
1.3.2. Non-coding RNA.....	14
1.3.3. Post-translational histone modifications.....	15
1.3.4. Remodeling factors.....	18
1.3.5. Histone chaperones.....	19
1.3.6. Histone variants	20
1.4. H3.3 and the histone H3 family	25
1.4.1. Evolution of histone H3 variants.....	25
1.4.2. Genetic structure of H3 and H3.3 in mammals	25
1.4.3. The H3.3 containing nucleosome	29
1.4.4. H3 variant incorporation into chromatin	31
1.5. Genomic localization of H3.3 and function	32

	Page
1.5.1. Early studies on genome-scale H3.3 localization provide elements to speculate on H3.3 function.....	32
1.5.2. Role of H3.3 in active transcription	33
1.5.3. H3.3 interaction with H2A.Z at promoters.....	33
1.5.4. Role of H3.3 in heterochromatin maintenance and genomic stability ...	34
1.5.5. H3.3 in development and sexual reproduction	36
1.5.6. H3.3 in tumorigenesis.....	36
1.6. Aim	39
Chapter 2. Materials and Methods.....	41
2.1. Materials	41
2.1.1. General laboratory chemicals and reagents	41
2.1.2. Cell culture chemicals and reagents	43
2.1.3. Primers.....	43
2.1.4. Lentiviral vectors and shRNA	47
2.1.5. Enzymes	47
2.1.6. Antibodies and beads.....	48
2.1.7. Equipment.....	48
2.1.8. Software.....	50
2.2. Solutions and Media.....	51
2.2.1. Cell culture solutions and media	51
2.2.2. Genomic DNA Extraction and analysis	51
2.2.3. Western blot solutions and buffers	51
2.2.4. Nuclei isolation buffers	52
2.2.5. Chromatin preparation and immunoprecipitation buffers	54
2.3. Mouse models.....	55
2.4. Methods	55
2.4.1. General maintenance and handling of test subjects.....	55
2.4.2. Mouse embryonic fibroblast (MEF) isolation	55
2.4.3. General maintenance and handling of cell lines.....	56
2.4.4. Transformation of bacteria	57
2.4.5. Plasmid DNA isolation.....	57

	Page
2.4.6. Genomic DNA isolation	57
2.4.7. Genotyping by PCR.....	57
2.4.8. Agarose Gel Electrophoresis	59
2.4.9. Lentivirus production for shRNA based knock-down in MEFs.....	59
2.4.10. Adenovirus infection for transient Cre expression in MEFs	59
2.4.11. RNA isolation.....	60
2.4.12. cDNA preparation	61
2.4.13. Quantitative PCR (qPCR) and analysis	61
2.4.14. Nuclei isolation from liver tissue	61
2.4.15. Nuclei isolation from cell lines.....	62
2.4.16. Nuclei quantification and lysis	62
2.4.17. Whole cell protein extraction.	63
2.4.18. Western blot.....	63
2.4.19. Native chromatin immunoprecipitation (N-ChIP).....	64
2.4.20. Cross-linked chromatin immunoprecipitation (X-ChIP).....	66
2.4.21. Immunofluorescent Staining and imaging.....	67
2.4.22. Nuclear and mitotic defect evaluation	67
2.4.23. Flow cytometry cell cycle analysis with propidium iodide DNA staining	67
2.4.24. RNA and ChIP sequencing and bioinformatic analysis	67
2.4.25. Statistical analysis of qPCR data	68
Chapter 3. Results	69
3.1. Novel knock-in/ conditional knock-out mouse lines.....	69
3.1.1. Molecular description and validation of mouse models.....	69
3.1.2. WT, FH-H3.3A, FH-H3.3B, H3.3A-KO and H3.3B-KO mouse lines are phenotypically similar.	73
3.2. Genome-wide distribution of H3.3 at nucleosome resolution in the liver ...	75
3.2.1. Mono-nucleosome preparation from liver tissue ensures high- resolution profiling of H3.3.....	75
3.2.2. H3.3 is highly enriched at TSS and its enrichment positively correlates with gene expression	77

	Page
3.2.3. H3.3A and H3.3B have identical enrichment patterns at most genomic sites except at some retroviral repeat elements	79
3.2.4. H2A.Z is co-localized with H3.3 around the TSS.....	81
3.3. Effect of H3.3 loss on transcriptome.....	83
3.3.1. Knock-out of a single H3.3 coding gene does not affect liver transcriptome	83
3.3.2. H3.3 depleted mouse embryonic fibroblasts	86
3.3.3. H3.3 depletion in MEFs has some, yet minimal effect on the transcriptome	91
3.4. H3.3 involvement in mitotic progression.....	95
3.4.1. H3.3 and H2A.Z are present at TSS of deregulated genes involved in cell cycle.....	95
3.4.2. H3.3 depletion does not affect H2A.Z enrichment at promoter regions	95
3.4.3. H3.3 depletion results in defective mitotic progression	98
Chapter 4. Discussion	100
Chapter 5. Perspectives	110
Complete H3.3 depletion in MEFs.....	110
Implication of the N-terminal tail of H3.3 and its phosphorylation in mitotic regulation.....	110
H3.3 implication in liver regeneration	111
References	112
Appendices	131
Appendix A – Flow cytometry analysis of the cell cycle by DNA content (PI incorporation)	131
Appendix B – Copyright Permissions.....	132
Articles	136

List of Figures

Figure	Page
Figure 1-1. Secondary structure of histones.....	2
Figure 1-2. Nucleosome assembly.	5
Figure 1-3. Levels of chromatin compaction in the eukaryotic nucleus.	6
Figure 1-4. The eukaryotic cell cycle.....	9
Figure 1-5. Main mechanisms of epigenetic regulation.....	13
Figure 1-6. Human core and linker histone variants.	21
Figure 1-7. Genomic organization of H3 coding genes	25
Figure 1-8. Sequence alignment of processed transcripts of H3f3a and H3f3b.....	28
Figure 1-9. Amino-acid sequence alignment of mammalian histone variants H3.3, H3.2 and H3.1.	29
Figure 1-10. Differential amino-acids between H3.3 and H3 variants in reference to the nucleosome structure.....	30
Figure 1-11. H3.3 incorporation into chromatin.	31
Figure 1-12. Contribution of histone mutations and deregulations in their expression to tumorigenesis in humans.....	37
Figure 3-1. Generation of H3.3A and H3.3B mouse models.....	69
Figure 3-2. Genotype validation of FH-H3.3A and FH-H3.3B mice.	70
Figure 3-3. FH-H3.3A and FH-H3.3B expression in mouse livers.	71
Figure 3-4. H3.3 knock-out validation of WT, FH-H3.3A, H3.3A-KO, FH-H3.3B and H3.3B-KO mice by genotyping, RT-qPCR and Western blot.	72
Figure 3-5. Mono-nucleosome preparation for N-ChIP and ChIP validation.....	76
Figure 3-6. ChIP-Seq data reveals enrichment of H3.3.	77
Figure 3-7. H3.3A and H3.3B are enriched around transcription start (TSS) and termination sites (TTS).	78
Figure 3-8. H3.3 is present at TSS and positively correlates with gene expression.	78
Figure 3-9. H3.3A and H3.3B show similar deposition patterns at most genomic regions.	79
Figure 3-10. H3.3A and H3.3B enrichment differs at repetitive sequences.	80

Figure	Page
Figure 3-11. Genome-wide enrichment pattern of FH-H3.3A, FH-H3.3B and H2A.Z at TSS.....	81
Figure 3-12. Normalized densities of H2A.Z, H3.3A and H3.3B within gene bodies expressed at different levels in mouse liver.	82
Figure 3-13. H3.3A, H3.3B and H2A.Z enrichment at TSS correlates with transcription levels.	84
Figure 3-14. Effect of single H3.3 loss on the transcriptome of adult mouse livers.....	85
Figure 3-15. Generation of H3.3B-KO and H3.3B-KO / H3.3A-Kd MEFs.....	86
Figure 3-16. Validation of WT, FH-H3.3A, H3.3A-KO, FH-H3.3B and H3.3B-KO MEFs by genotyping, RT-qPCR and Western blot.....	87
Figure 3-17. H3.3 specific shRNA selection and validation of knock-down efficiency.	89
Figure 3-18. H3.3 expression and doubling times in H3.3B-KO / H3.3A-Kd MEF model.....	90
Figure 3-19. Effect of H3.3 loss on MEF transcriptome.....	91
Figure 3-20. Comparative analysis of the differentially expressed genes in H3.3 deficient embryos and in H3.3 depleted ESCs.....	93
Figure 3-21. H3.3 and H2A.Z enrichment at transcription start sites (TSS) of mitotic genes showing differential mRNA expression.....	97
Figure 3-22. Mitotic defects in H3.3 deficient MEFs.	99
Figure 4-1. Graphical abstract of thesis.	109
Figure A. Cell cycle analysis by flow cytometry of FH-H3.3B MEFs.....	131

List of Tables

Table	Page
Table 1-1. Point mutations in H3.3 and its chaperones observed in human cancer.....	38
Table 2-1. Chemicals, reagents, enzymes and kits used for general laboratory purposes	41
Table 2-2. Chemicals, reagents, kits and media used in cell culture	43
Table 2-3. Primers used in genotyping mouse and cell lines from genomic DNA.....	44
Table 2-4. Primers used in RT-qPCR for gene expression	44
Table 2-5. Primers used in ChIP-qPCR	46
Table 2-6. Plasmids used for lentiviral shRNA transduction.....	47
Table 2-7. Enzymes and kits used in the study	47
Table 2-8. Antibodies used in the study.....	48
Table 2-9. Agarose resins and magnetic beads	48
Table 2-10. Instruments used in the study	48
Table 2-11. General laboratory equipment used in the study	49
Table 2-12. Software used in the study	50
Table 2-13. Cell lines and their relevant culture media	51
Table 2-14. Buffers used in genomic DNA extraction and analysis.....	51
Table 2-15. Buffers and solution used in western blot.....	52
Table 2-16. Polyacrylamide Tris Gel preparation.....	52
Table 2-17. Buffer compositions used in nuclei isolation from liver tissue	53
Table 2-18. Buffer compositions used in nuclei isolation from cells	53
Table 2-19. Buffer compositions used in chromatin preparation and ChIP.....	54
Table 2-20. PCR reaction setup volumes for genotyping.	58
Table 2-21. Primer pairs used in genotyping of FH-H3.3A, FH-H3.3B and Actin-Cre mouse and cell lines	58
Table 2-22. PCR cycling conditions used in genotyping protocols	58
Table 3-1. Average litter size, weaned to born and female (F) to male (M) ratios in mouse models.....	73
Table 3-2. Genotype distribution after heterozygous mating of mouse lines.	73

Abbreviations

<u>Abbreviation</u>	<u>Explanation</u>	<u>Abbreviation</u>	<u>Explanation</u>
ACF	ATP-utilizing chromatin assembly and remodeling factor	ECM	ExtraCellular Matrix
AEBSF	4-(2-Aminoethyl) benzenesulfonyl fluoride hydrochloride	EDTA	EthyleneDiamineTetra-acetic Acid
ALT	Alternative Lengthening of Telomeres	EGTA	Ethylene Glycol-bis (β -aminoethyl ether)-N,N,N',N'-Tetraacetic Acid
APS	Ammonium PerSulfate	ERV	Endogenous Retroviral element
ARE	AU-Rich Element	ESC	Embryonic Stem Cell
ATRX	Alpha-Thalassemia/Mental Retardation Syndrome	FACS	Fluorescence Assisted Cell Sorting
AUF1	AU-binding Factor 1	FACT	FAcilitates Chromatin Transcription
Bp	Basepairs	FBS	Fetal Bovine Serum
CAF-1	Chromatin Assembly Factor-1	HAT	Histone Acetyl-Transferase
CDK	Cyclin Dependent Kinase	HDAC	Histone DeACetylase
Cds	coding DNA sequence	HDM	Histone DeMethylase
CHD	Chromodomain Helicase DNA binding	HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
CpG	Cytosine-Guanosine motifs	HJURP	Holliday JUncion Recognition Protein
DAXX	Death Associated Protein	HMT	Histone Methyl Transferase
DMSO	Dimethyl sulfoxide	HP1	Heterochromatin Protein 1
DNA	DesoxyriboNucleic Acid	INO80	INOsitol requiring 80
DNMT	DNA Methyl-Transferase	ISWI	Imitation Switch
Dpf	Days post fertilization	KCl	Potassium Chloride
DTT	1,4-dithiothreitol	kD	Kilo Dalton
		Kd	Knock-down

<u>Abbreviation</u>	<u>Explanation</u>	<u>Abbreviation</u>	<u>Explanation</u>
KO	Knock-out	PCR	Polymerase Chain Reaction
LB	Lysogeny/Luria Bertoni Broth	PI	Propidium Iodide
LiCl	Lithium Chloride	PLK	Polo-Like Kinase
lncRNA	Long non-coding RNA	Pol II	RNA-polymerase II
MEF	Mouse Embryonic Fibroblast	PTM	Post-translational Modifications
MNase	Micrococcal Nuclease	qPCR	Quantitative PCR
MOI	Multiplicity of infection	qs	Quantum satis (sufficient quantity for)
mRNA	Messenger RNA	RBP	RNA Binding Protein
MWCO	Molecular Weight Cut-Off	RD	Replication Dependent
NaCl	Sodium Chloride	RNA	RiboNucleic Acid
NaDOC	Sodium Deoxycholate	RNase A	RiboNuclease A
NAP1	Nucleosomes Assembly Protein 1	RT	Room Temperature
NASP	Nuclear Autoantigenic Sperm Protein	RT-qPCR	Reverse transcription-qPCR
N-ChIP	Native Chromatin ImmunoPrecipitation	SAC	Spindle Attachment Checkpoint
NCP	Nucleosome Core Particle	SAM	S-adenosyl methionine
NEAA	Non-Essential Amino-Acids	SDS	Sodium dodecyl Sulfate
NGS	Next Generation Sequencing	Seq	Sequencing
o/n	Overnight	SWI/SNF	Switch/ Non-fermentable Sucrose
OD	Optical Density	tSCE	Telomere sister chromatin exchange
P/S	Penicillin/ Streptomycin	UTR	Untranslated region
PBS	Phosphate Buffered Saline	WT	WildType
PCIA	Phenol Chloroform Isoamyl Alcohol	X-ChIP	Crosslinked Chromatin Immunoprecipitation
PCNA	Proliferating Cell Nuclear Antigen		

Chapter 1. Introduction

In the eukaryotic organism, genomic information is organized into chromosomes. Each chromosome consists of a single, linear, double helix deoxyribonucleic acid (DNA) molecule tightly associated with highly conserved protein complexes. This organizational structure is called chromatin and more specifically corresponds to the complex of nuclear DNA and closely associated RNA and proteins.

The overall composition of chromatin is one third nucleic acids and two thirds proteins. Half of this protein fraction corresponds to very basic proteins called histones, the other half consists of various proteins called "non-histone proteins".

The nucleosome is the fundamental repeating unit of chromatin. It basically consists of DNA that's wrapped in approximately 2 superhelical turns around an octameric core structure created by histone proteins ¹.

Primarily, it was thought that chromatin's only function was to compact DNA so that the large molecule would fit in the small volume of the eukaryotic nucleus. Histones were not subject to much interest since it was believed that the transcription machinery could easily override these small proteins as it was the case in bacteria ². Over the last decades however, the increasing research in the field of epigenetics has seen histones transition from simple static building blocks to important dynamic actors in all physiological processes that involve interaction with DNA such as replication, transcription and maintenance of genomic integrity.

1.1. Structural organization of chromatin

1.1.1. Histones

Histones are the major protein constituents of chromatin. They are present in such large quantities that their mass is almost equal to that of DNA ³. Consequently, DNA quantification is commonly used as a means to determine histone mass when studying histones.

Histones are small and very basic proteins. They have approximately 100 amino-acids in their primary sequences (~15kD) with high lysine and arginine contents. There are five classes of histones: H1, H2A, H2B, H3 and H4. Based on their contribution to chromatin structure they can be divided into two groups: core histones consisting of H2A, H2B, H3, H4 that in pairs, form the core histone octamer and linker histone H1 that binds outwardly to the nucleosome core at entry and exit sites of linker DNA and contribute to the formation of the chromatosome. Histones are incorporated into chromatin through the action of specific histone chaperones ^{4,5}.

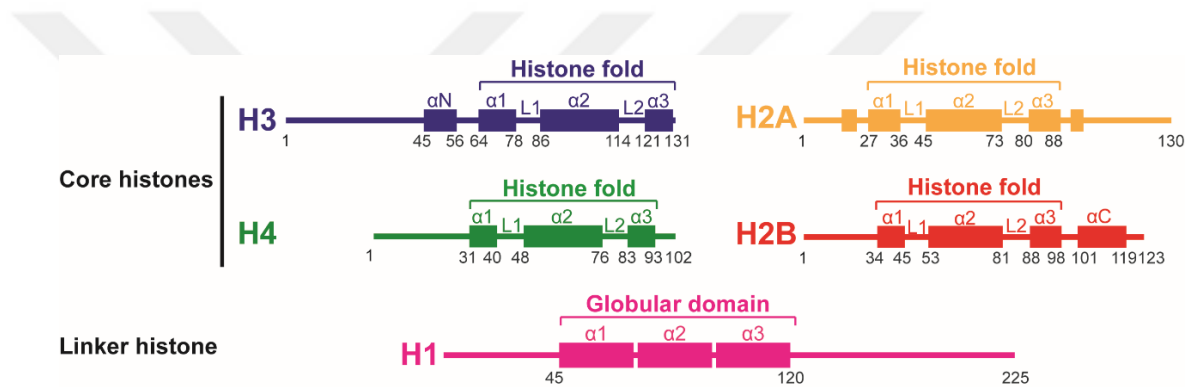


Figure 1-1. Secondary structure of histones. Core histones have a structure comprised of three regions, the characteristic histone fold which consists of a large central alpha helix (α2) connected to two smaller alpha helices (α1, α3) on either side by two loops (L1, L2) and is flanked by the two N-terminal and C-terminal tails which remain mainly unstructured apart from some structures like the αN for histone H3. Linker histone H1 proteins have a central globular domain containing three helical regions flanked by variable N- and C-terminal domains. Representative figures constructed from ^{2,6}

The histone fold motif is highly conserved and consists of a large central alpha helix connected by two loops to two smaller alpha-helices on either side. This central globular domain is flanked by a 15-45 amino-acid long N-terminal region and a much shorter C-terminal region of only a few residues (Figure 1-1) ^{7,8}. While some parts of these extensions can be structured, they remain mostly flexible. Especially the N-terminal tails diverge notably between core histones and are subject to many post-translational modifications that play important roles in epigenetic signaling, as detailed in Section 1.3.3. The presence of the histone-fold motif allows for histones to form stable dimers via the “handshake” interaction which is the basis for the assembly of the histone octamer ⁸ (Figure 1-2).

Linker histones differ greatly from core histones and are far less conserved. They are enriched in lysine residues and lack the histone-fold motif. H1-like linker histones have a three-domain structure: a short unstructured aminoterminal domain of about 45 residues; a highly conserved globular core domain of about 75 residues; and a carboxy terminal domain of about 100 residues (Figure 1-1) ^{6,9}.

Histones can also be classified into two distinct categories based on the time of their incorporation into chromatin during the cell cycle. These categories are replication dependent histones and replication independent histones. Replication-dependent (RD) histones are also known as “canonical”, “conventional”, “bulk” or “major” histones, and as suggested by their name, their expression and association with chromatin are coupled to DNA replication. They are loaded onto chromatin through the action of specialized proteins called histone chaperones. In metazoans, replication dependent histone transcripts present a very specific structure. They are coded by multicopy, intronless gene families organized in clusters with a stem-loop type structure to indicate transcription termination instead of the typical polyadenylation signal. This structure allows fast production of high levels of histone proteins needed for nucleosome assembly in the newly duplicated DNA during cell cycle progression. As an example, in humans, *HIST1*, a large histone cluster that contains 55 genes is located on chromosome 6 and *HIST2* and *HIST3*, two smaller clusters containing 9 genes are located on chromosome 1. Despite some genomic rearrangements, gene number and organization of mouse histone gene clusters are strikingly similar their human counterparts. In mice, the large histone cluster *Hist1* is located on chromosome 13 and contains 51 genes while the two smaller clusters, *Hist2* and *Hist3*, are located on chromosomes 3 and 11 ¹⁰. The expression of replication dependent histones is tightly regulated both at transcriptional and post-transcriptional levels and is restricted to the S-phase of the cell cycle to avoid histone toxicity ¹¹⁻¹³.

In contrast, replication-independent histones, also called histone variants or replacement histones, are expressed at basal levels through-out the cell-cycle independently from DNA replication. Some histone variants have specific chaperones that load them on and evict them from chromatin. The genes that code for each histone variant are located outside the major histone clusters and present a much typical structure. They have at least one intron, a polyadenylation signal and often long 5' and 3' untranslated regions (UTR) ¹⁴. Histone variants can replace their corresponding conventional histones in a nucleosome and confer

new structural and functional properties. Histone variants and their roles in epigenetic regulation are further detailed in sections 1.3.6, 1.4 and 1.5.

1.1.2. Nucleosome Core Particle and the Chromatosome

The nucleosome core particle (NCP) was first discovered during nuclease digestion experiments of purified chromatin. When digested with nucleases for only a short period of time, protein-bound DNA is protected from digestion, whereas “free” DNA such as linker DNA, is accessible to nucleases and is degraded. High-resolution X-ray crystallization studies gave us detailed insight as to the atomic structure of the NCP ^{1,15}.

The histone octamer that constitutes the core of the NCP, consists of a pair of each of the core histones. For nucleosome assembly, first, H2A/H2B and H3/H4 dimers are formed. Then, two H3/H4 dimers associate via an interaction between the $\alpha 2$ and $\alpha 3$ helices of the H3 histones and form a tetramer. It is the H3/H4 tetramer that binds to DNA to form the intermediate core particle and the full nucleosome core particle assembly is completed with the association of two H2A/H2B dimers via the 2α and 3α helices of H2B and H4 (Figure 1-2) ^{1,16}.

The handshake interactions between each pair of histones lead to the formation of β bridges between loops L1 and L2. These bridges form part of the DNA binding sites. The association between the histone octamer and the DNA fragment is mainly due to the insertion of the side chains of the arginine residues in the minor groove of the DNA double helix ^{3,8}. While the structured histone regions are implicated in the majority of DNA-histone interactions, the less structured histone tails protrude from the nucleosome. Histone tails are more accessible to interact with neighboring nucleosomes or other factors. They are also target to many post-translational modifications (PTMs). The implication of PTMs in important cellular processes is detailed in section 1.3.3.

The chromatosome is the structure formed by the association of a H1 linker histone to the nucleosome core particle ¹⁷. Histone H1 binds to the NCP at the entry and exit sites of DNA ¹⁸. It associates over 20 base pairs (bp) of DNA at the entry and exit sites, greatly limiting their movements and thus locking the nucleosome in a closed position ¹⁹. The term nucleosome *per se* refers to the association of the NCP or chromatosome with one of its adjacent linker DNAs (Figure 1-2).

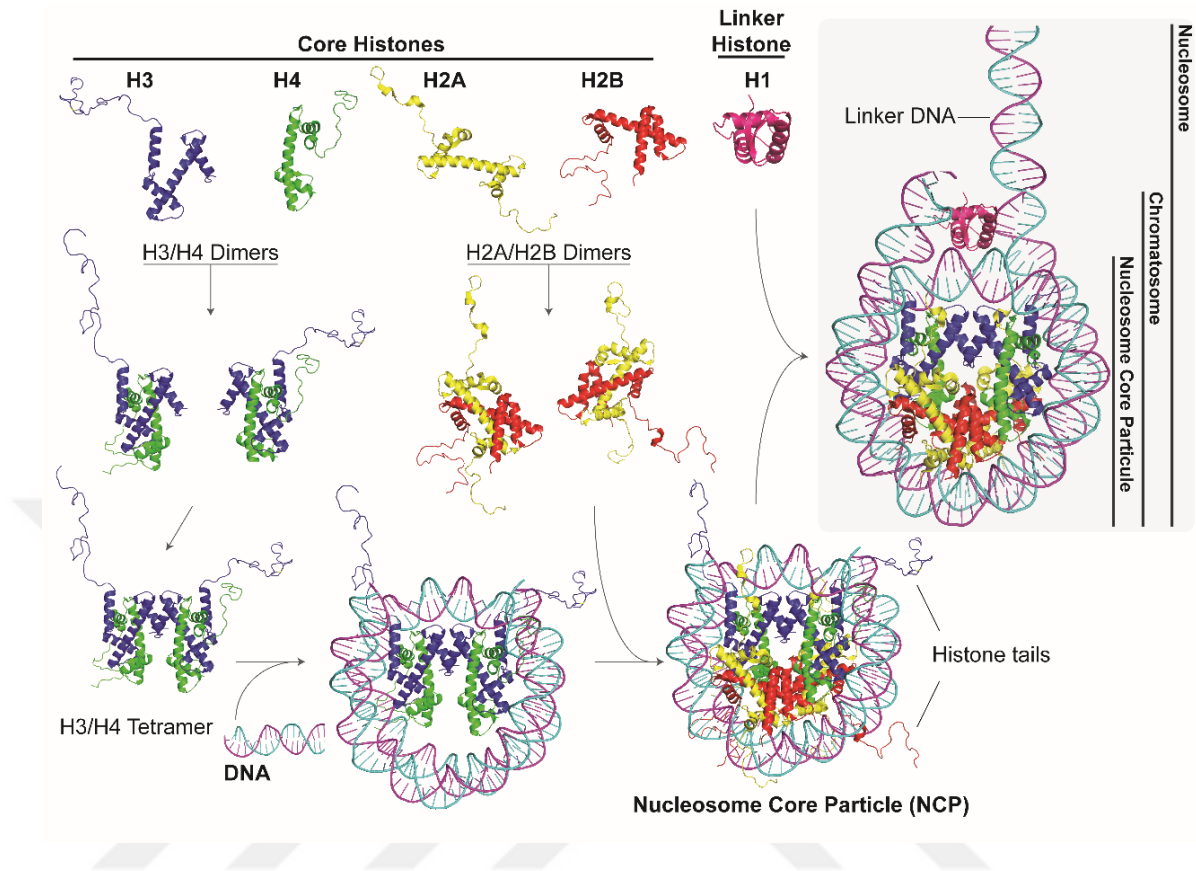


Figure 1-2. Nucleosome assembly. First, H3 (blue), H4 (green), H2A (yellow), H2B (red) dimerize through “hand-shake” interactions to form H2A/H2B and H3/H4 dimers⁸. Then, two H3/H4 dimers associate between each other to form a tetramer. It is the H3/H4 tetramer that binds to DNA (cyan, magenta) and the nucleosome core particle (NCP) assembly is completed with the association of two H2A/H2B dimers. Linker histone H1(hot pink) binds to the NCP forming the chromatosome^{17,18}. The nucleosome refers to the NCP associated to one of its adjacent linker DNA. N-terminal histone tails are usually accessible to various effectors. Image reconstructed from nucleosome core particle high resolution structure (PDB ID: 1KX5¹⁵) and chromatosome structure (PDB ID: 4QLC²⁰) shown in cartoon representation using PyMol software²¹.

1.1.3. Higher-order organization of chromatin

The main function of chromatin is to compact genetic material inside the nucleus all the while allowing access to different regions of DNA for the correct progression of physiological processes. In order to fulfill this function, it needs to remain highly dynamic and flexible. Through rearrangement of nucleosome arrays in various reproducible spatial conformations, chromatin achieves higher orders of organization²².

The first order of compaction of chromatin is defined by the packaging of DNA with histones yielding a chromatin fiber of approximately 10-nm in diameter called the

nucleofilament (Figure 1-3). It is essentially a succession of nucleosomes separated by linker DNA and is familiarly called the “beads on a string” structure where “beads” refer to nucleosomes and “string” refers to DNA. This primary structure reduces DNA length by an approximate 6-fold and is generally permissive to transcription. However, under physiological conditions, nucleosomes are rarely found to stay in their “stretched-out” form. Instead, they are further folded and compacted at different levels to form “higher order structures”^{2,23,24} (Figure 1-3).

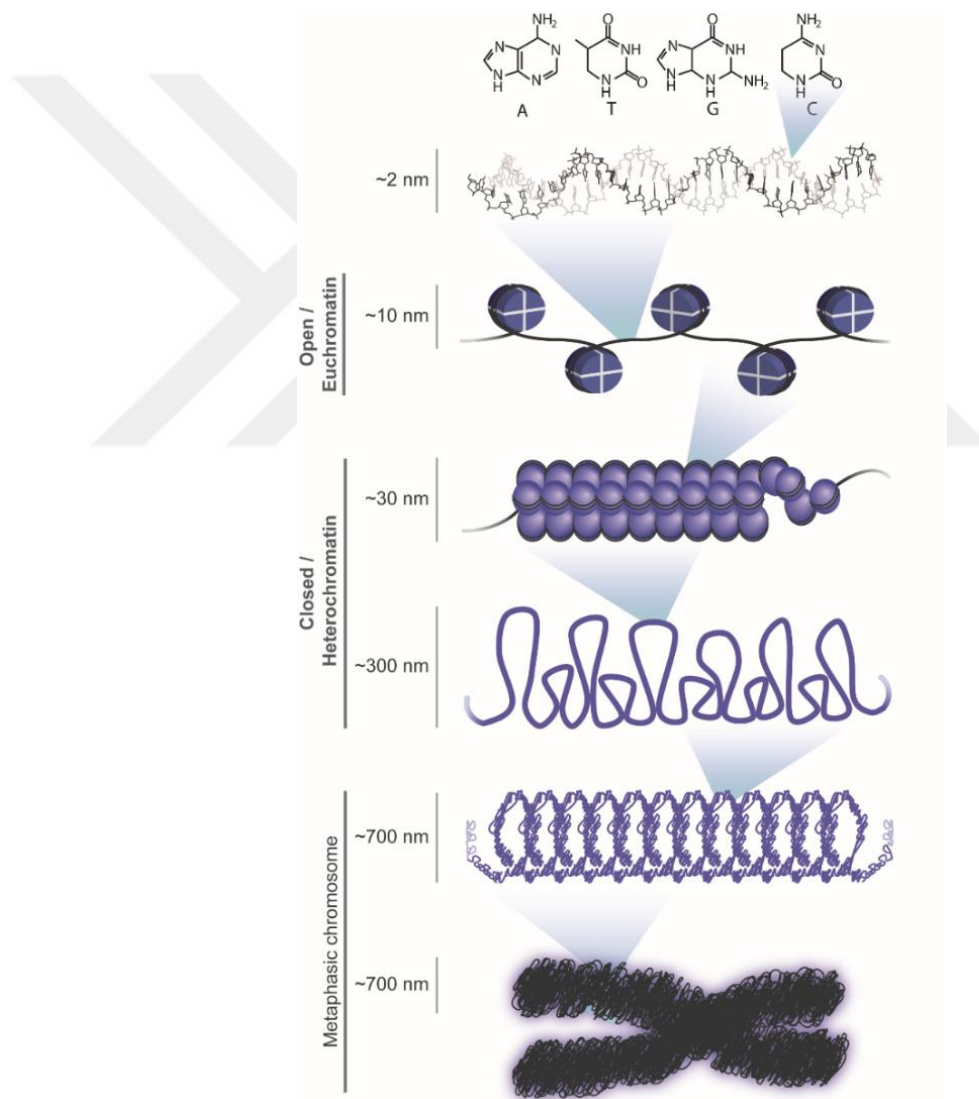


Figure 1-3. Levels of chromatin compaction in the eukaryotic nucleus. DNA is wrapped around a histone octamer to form the first order of compaction called the nucleofilament or the 10-nm fiber. Arrays of nucleofilaments fold further to form the 30-nm fiber. During interphase, this structure goes through different levels of compaction before reaching the highest level of compaction that is observed in the metaphase chromosome².

The secondary order of compaction for chromatin is the 30-nm chromatin fiber which is the result of a helical rearrangement of linear nucleosomes stabilized by linker proteins like H1 and HP1 (Figure 1-3). It allows a DNA compaction of almost 50 fold²⁵. Though still subject to controversy, two models have been put forward to explain the formation of this fiber. The first model, called the solenoid or the one-start helix model suggests that a single array of consecutive nucleosomes connected by linker histones, folds around an axis of symmetry to form the fiber with 6 to 8 nucleosomes per helical turn^{25,26}. The second model called the zig-zag or the two-start helix model, suggests that two nucleosome arrays are assembled in a zig-zag so that consecutive nucleosomes are found alternatively on each side of the fiber²⁷⁻²⁹.

Unfortunately, both models have the disadvantage of being predominantly based on in vitro observations. Recent studies that aimed to elucidate which model is predominant in vivo, have suggested a coexistence of different structures where the formation of higher order structure would actually be determined by the environment of the nucleosomes^{30,31}.

Chromatin architecture is sensitive to a large number of internal and external factors such as length of internucleosomal DNA, presence of histone variants or post-translational modifications, ionic conditions or binding of chromatin architecture proteins^{32,33}.

To date, the in vivo structure of the 30-nm fiber has not been determined and there is even doubt to its actual existence^{34,35}. Although higher levels of organization do exist, with the most evident example being the metaphasic chromosome, the precise structure and the sequential hierarchy of such organization is subject to intensive research.

The ability of chromatin to be arranged and rearranged in different orders of structure following intranuclear signals demonstrates its' dynamic nature. Organisms have evolved different mechanisms to introduce variation into chromatin. Total or partial reorganization of chromatin allows for new functional properties, that allow in particular, transcription, replication or repair of DNA.

1.2. Organization of chromatin function

1.2.1. Eukaryotic chromosome structure

For genetic information to be accurately and effectively passed on during subsequent cell divisions, a chromosome must be able to replicate and the doubled material be equally divided between daughter cells, all the while maintaining a damage free DNA. Three specialized elements in the chromosome control these processes ². First, the presence of many replication origins on eukaryotic chromosomes allows for rapid duplication of the genetic material during S phase of the cell cycle. The duplicated DNA molecules, *i.e.* sister chromatids, are maintained in proximity by cohesins and stay attached at their centromeres. During cell division, centromeres serve as a docking point for kinetochore assembly which will allow the sister chromatids to segregate into two daughter cells ^{36,37}. Finally, telomeres, specialized repetitive sequences, ensure correct replication and protection of the ends of the linear DNA molecule ^{36,38,39}. The architectural structure of the chromosomes can be best observed during metaphase of the cell cycle when DNA has duplicated and chromatin is at its most condensed state (Figure 1-3).

In the eukaryotic nucleus, nucleosomes are not arranged in a regular, homogenous manner. Instead, chromatin shows different levels of compaction dependent on cell-cycle stage and nuclear processes such as replication, transcription and repair. Furthermore, the nucleus is organized into function-specific sub-nuclear domains even though these domains are not architecturally fixed nor delimited by membranes. This compartmentalization links chromatin structure to genomic function under the control of physiologic signaling ⁴⁰.

1.2.2. The cell cycle and chromatin

The cell cycle enables the duplication of genetic information and its division into two identical daughter cells. These two events determine the phases of the cell cycle. DNA replication takes place during the S phase which stands for synthesis and chromosome segregation and cell division takes place during the M phase which stands for mitosis. Between these two phases are the gap or growth phases, G1 that follows the M-phase and G2 that follows the S-phase. These gap phases not only serve to prepare the cell for growth but also allow for checking that conditions are favorable both internally and externally for proper cell cycle progression. As a matter of fact, if conditions are not suitable, cells may extend their stay in G1 or enter a quiescent state called G0 after which they may later reenter

the cell cycle or die. Interphase corresponds to the G1, S and G2 phases. There are three known checkpoints: the G1 checkpoint also called the restriction point because cells pass this point are committed to DNA replication and can no longer extend their stay in G1 or enter G0, the G2/M checkpoint which checks for DNA damage before cell division and the Metaphase checkpoint also called the spindle checkpoint which checks for correct alignment of chromosomes on the metaphase plate (Figure 1-4).

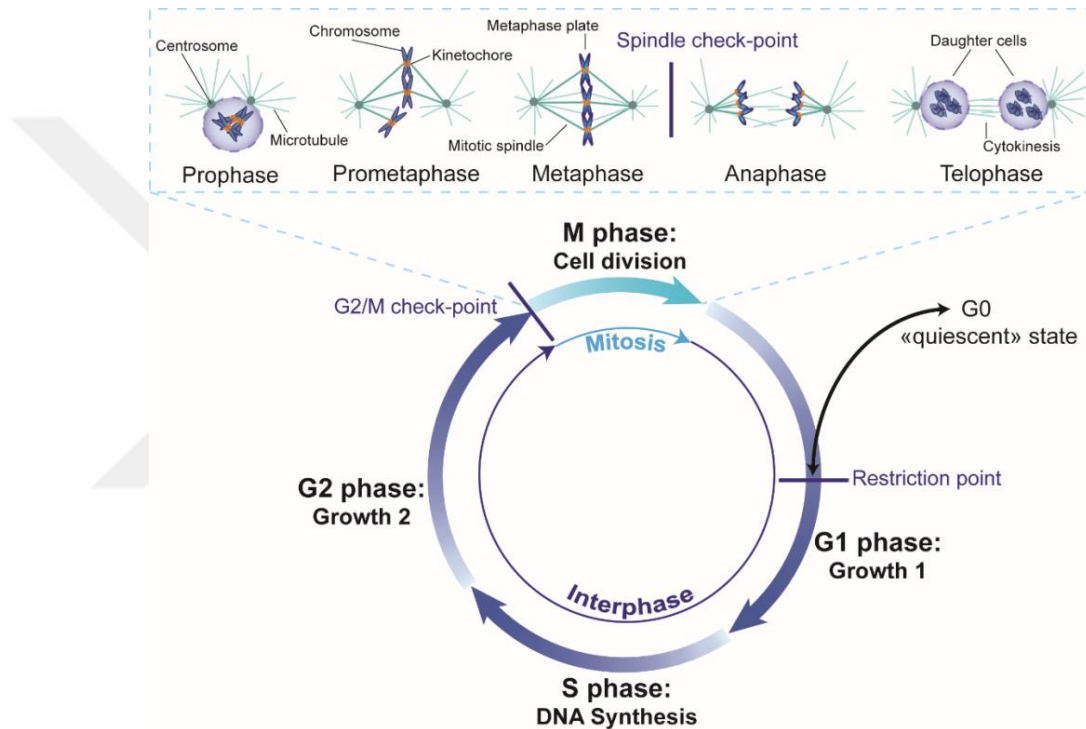


Figure 1-4. The eukaryotic cell cycle. The cell cycle consists of an S phase, when DNA replication takes place, and a mitosis phase, when cell division takes place, separated by the two growth or gap phases G1 and G2 during which the cell prepares for the ensuing phases and checks that conditions are favorable for cell cycle progression. If conditions are not favorable, during G1, the cell may enter a quiescent state called G0 until conditions become favorable or the cell dies. There are 3 known checkpoints: the restriction point at the end of G1 when the cell commits to DNA replication, the G2/M checkpoint which checks for DNA damage before cell division and the spindle checkpoint at the end of metaphase which checks for correct attachment of sister chromatids at their kinetochores and their alignment on the metaphase plate. Interphase corresponds to G1, S and G2 phases. An expanded view of mitosis is presented at the top of the figure. In prophase, chromatin condenses and the centrosomes migrate to opposite poles of the cell emanating microtubules in all directions. In prometaphase, the nuclear membrane is disintegrated and microtubules attach to chromosomes at their kinetochores which migrate toward the center. During metaphase, chromosomes are aligned on the metaphase plate and the cell proceeds onto anaphase only if all chromosomes are correctly aligned and attached to the mitotic spindle. In anaphase, the sister chromatids separate, finally in telophase the DNA decondenses, the nuclear envelope reforms and the two daughter cells form by cytokinesis. Figure adapted from Pines, 2011⁴¹ with permission (Appendix B).

Mitosis is divided into five phases. During the first phase, prophase, chromatin condenses to form individual chromosomes. The two centrosomes each migrate to opposite poles of the cell and emanate microtubules in all directions and the nuclear membrane starts being absorbed by the endoplasmic reticulum. During prometaphase, the nuclear envelope is disintegrated and fully condensed chromosomes migrate towards the center of the cell and align at the spindle equator also called the metaphase plate. The mitotic spindle consists of the set of microtubules linking the chromosomes at their centromere to the centrosomes. In metaphase, chromosomes remain on the metaphase plate and the cells “checks” for correct orientation and attachment of each chromosome to the mitotic spindle. If cells pass the “mitotic checkpoint”, mitosis continues on to anaphase where sister chromatids are separated towards each pole of the spindle which starts to elongate as the daughter cells separate. During telophase, the chromosomes are released from the spindle microtubules as they begin to de-condense and the nuclear envelope begins to form around the daughter cells. Cytokinesis occurs almost simultaneously and leads to the separation of the cytoplasm of the two daughter cells. It consists of the formation of a contractile ring in the middle of the cell that gradually tightens to leave a small very dense structure called the residual body between the cells. The disappearance of this structure marks the separation of the daughter cells and the end of mitosis ² (Figure 1-4).

During mitosis, chromatin is highly condensed with a varying level of compaction depending on the specific phase of mitosis. The highest level of compaction is achieved in metaphase and chromosomes can be easily visualized in their characteristic X-shape. The organization of chromatin in the mitotic chromosomes depends on the arrangements of the fibers into an axial skeleton composed of structural maintenance of chromosome (SMC) proteins ^{42,43}, in particular, condensins I and II ⁴⁴. Mitosis, and therefore the structure of the mitotic chromosome, is associated with an almost complete absence of transcription. Intriguingly, this absence of transcription does not come directly from the high level of compaction of DNA. As a matter of fact, the level of compaction does not necessarily entail a total loss of the accessibility of DNA to various transcription factors ⁴⁵. It would seem that it is rather the fact that the structure of chromatin is relatively "frozen" during mitosis leading to the fact that RNA polymerase can no longer advance along the DNA strand ⁴⁶. As a result of the transcriptional arrest at this stage, all effectors and enzymes needed for mitotic progression should be readily available in the cell.

Mitosis is a tightly regulated process. Mistakes in distribution of sister chromatids could lead to the loss of significant genetic information and cause irreparable damage to the cells. This regulation is carried out by mainly two mechanisms. The first one consists of the degradation of effectors by proteolysis after they have completed their function and the second mechanism uses signaling through phosphorylation by different types of kinases^{47,48}, such as the members of the cyclin dependent kinase (CDK), Aurora⁴⁹ and Polo-like kinase (PLK) families⁵⁰.

At interphase, the mitotic chromosome structure does not persist. Chromosomes adopt a more relaxed conformation and yet remain organized into nuclear domains called chromosome territories⁵¹. A gene's expression can be highly altered depending on its nuclear sub-location. Two main areas can be defined based on chromatin condensation levels. Euchromatin has a loose chromatin structure, is rich in actively transcribed genes and replicates in early S-phase. Heterochromatin on the other hand is highly condensed, gene poor, enriched in repetitive elements and it replicates at the end of S phase. Heterochromatin can be further divided into constitutive and facultative heterochromatin. Constitutive heterochromatin is always compact and is characteristic of centric, pericentric and telomeric regions harboring repetitive DNA elements and imprinted genes which are heritably silenced throughout the whole organism. On the other hand, facultative heterochromatin is formed in gene rich regions but remains transcriptionally inactive. However, it can become de-condensed, hence reactivated, in certain contexts that can be temporal (developmental, cell cycle specific), spatial (changes in nuclear localization due to external signals) or hereditary (mono-allelic gene expression)⁵². Inside the nucleus, heterochromatic regions of a chromosome are often found at the periphery, closely associated to lamina, whereas gene-rich regions are directed towards the center⁵¹. Furthermore, constitutive and facultative heterochromatin are characterized by distinct epigenetic marks such as specific histone post-translational modifications and associated proteins^{33,53}.

1.2.3. Transcription and chromatin

Transcription is an extremely complex process involving many levels of control. Its regulation is essential for development and cellular homeostasis as it allows the differential expression of genes according to cell type and cell cycle stage. Deregulation of transcription

severely impacts cellular functions and is implicated in the development of many pathological conditions. Although transcription is one of the most studied areas in science, many mechanisms still remain to be elucidated.

Like the cell cycle, transcription is a cyclic process. Briefly, during initiation, activators bind specific DNA sequences upstream of the promoter leading to the binding of effectors such as transcription factors which position RNA polymerase II (Pol II) and constitute the pre-initiation complex. RNA synthesis is initiated when a transcription factor melts a short DNA sequence and places it in the Pol II cleft. In higher eukaryotes, Pol II is paused at promoter-proximal regions and the carboxy-terminal domain (CTD) of RNA Pol II needs to be phosphorylated for elongation to proceed. Effective elongation is proceeded by termination which consists of the release of mRNA from Pol II and consequently release of Pol II from DNA.

As stated in the previous section, at interphase, euchromatin and heterochromatin define areas of differential levels of transcription based on condensation levels of chromatin. However, even though euchromatin is less compact than heterochromatin, it still has a dense nucleosome structure. Nucleosomes constitute obstacles to transcription and need to be displaced or replaced by chromatin remodeling to allow access to genetic information⁵⁴. Chromatin structure can block initiation of transcription and delay its elongation^{55,56}. Therefore, uncovering mechanisms that control nucleosome interaction with DNA and other proteins at sites of transcription is an important step in understanding the regulation of transcription.

1.3. Epigenetic regulation

The notion of epigenetics, in its modern sense, is based on the work of biologist Conrad H. Waddington (1942). He initially defined epigenetics as “the branch of biology which studies the causal interactions between genes and their products, which bring the phenotype into being”⁵⁷. Historically, any event that could not be clarified by genetic phenomena would be classified into “epigenetics”. Today, the term “epigenetics” is used to refer to heritable changes in gene expression that are independent from direct modifications to the DNA sequence^{23,58}. It’s a collection of marks affixed to the genome that orchestrate the reorganization of chromatin into the functional domains described above, allowing a selective and directed expression of the genome. Unlike transcription factors, epigenetic

marks persist even after the external stimuli that lead to their deposition are no longer relevant. They are stable and play a role in cellular memory, thus being responsible for long term gene expression profiles. Nonetheless, epigenetic marks can be modified or reversed depending on the environment ^{23,58}.

The main mechanisms of epigenetic regulation of gene expression in mammals are DNA methylation, interactions with non-coding RNAs, remodeling of local chromatin structures by energetically displacing nucleosomes and/or altering histone-DNA interactions, reversible post-translational modifications on histone tails and incorporation of histone variants ^{59–61} (Figure 1-5). These mechanisms are briefly described below with an emphasis on histone variants.

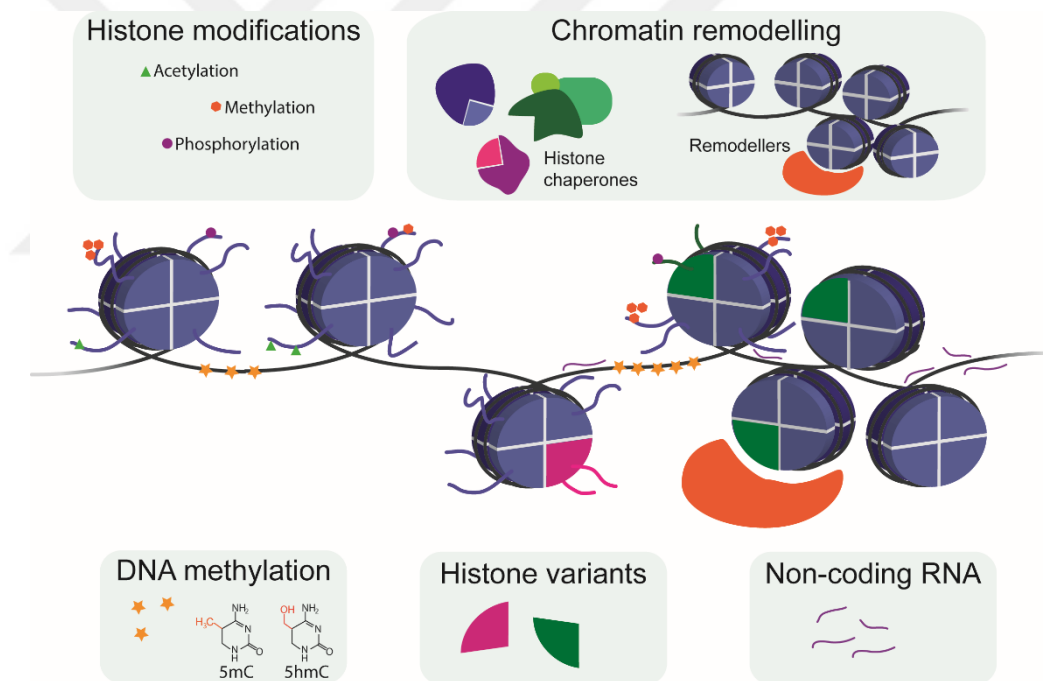


Figure 1-5. Main mechanisms of epigenetic regulation. Several epigenetic mechanisms are employed individually or in combination by eukaryotic cells to introduce variation into chromatin and allow for the proper progression of cellular processes ^{59–61}. The main mechanisms of epigenetic regulation of gene expression in mammals are 1- covalent modifications to DNA such as cytosine methylation (5mC) and hydroxymethylcytosine (5hmC) (depicted as yellow stars), 2- interactions with non-coding RNAs (depicted as purple strands coating chromatin), 3- remodeling of local chromatin structures by energetically displacing nucleosomes and/or altering histone-DNA interactions using chromatin remodellers, 4- reversible post-translational modifications on histone tails such as acetylation (green triangle), phosphorylation (purple circle) and methylation (orange hexagon) and 5- incorporation of histone variants depicted by pink and green quarter circles.

1.3.1. DNA methylation

DNA methylation is a post-replicative modification that, in mammals consists of the binding of a methyl (CH₃) group almost exclusively to the carbon 5 of the pyrimidine ring of Cytosine residues in CpG dinucleotides ⁶². This covalent addition of a methyl group to a cytosine is catalyzed by DNA methyltransferases (DNMTs). Some DNMTs bind new methyl groups (de novo methylation) to DNA while others function to maintain existing methyl marks through mitosis, thus contributing to epigenetic memory (methylation maintenance) ⁶³. Regions with high density of CpG dinucleotides are called CpG islands and are usually located on gene promoters and/or the first exons of mammalian genes ⁶⁴. In combination with other epigenetic marks like histone PTMs and RNA, DNA methylation and its binding proteins are associated with heterochromatin formation thus to silencing of transcription ⁶⁵⁻⁶⁸.

DNA methylation is involved in the establishment and maintenance through cell division of repetitive and centromeric DNA silencing, X chromosome inactivation in females and mammalian imprinting.

1.3.2. Non-coding RNA

Even though only a small fraction of the genome is actually translated, a large number of transcripts are implicated in important cellular functions. In recent years, increasing evidence showed that microRNAs (miRNA), Piwi-interacting RNAs (piRNA), endogenous silencing RNAs (siRNA) and long non-coding RNAs (lncRNA) play important roles in regulatory mechanisms at both transcriptional and post-transcriptional levels ⁶⁹⁻⁷¹.

Non-coding RNAs do not code for any functional proteins by definition and their expression can be strongly tissue-specific ⁷². Through various mechanisms not completely understood for most, non-coding RNAs are able to regulate gene expression. For example, their association with specific proteins involved in gene regulation like transcription or remodeling factors, could serve as a trap that prevents those factors to act on their relevant targets thus impeding their function. On the other hand, association with lncRNAs can also serve to guide proteins to their specific targets.

Finally, lncRNA can also serve as docking platform that would centralize various effector proteins to the same genomic site ⁷³. Non-coding RNAs are known to associate with many

factors involved in chromatin reorganization and methylation such as DNMTs and heterochromatin protein (HP1). One of the most well-known examples of non-coding RNA based epigenetic regulation is the XIST RNA involved in the silencing of the supernumerary female X chromosome in mammals ⁷⁴. Although, the detailed mechanisms by which non-coding RNAs act remain to be elucidated, expanding research has provided insight to their involvement in dynamic regulation of chromatin and epigenetic memory, including genomic imprinting, DNA methylation and transcription regulation ^{68,69,73,75}.

1.3.3. Post-translational histone modifications

Histones are targets to a large number of post-translational modifications (PTMs). These modifications include the addition of small chemical groups such as acetyl, methyl and phosphate groups as well as the addition of bigger globular proteins such as ubiquitin and SUMO ⁷⁶⁻⁷⁸. A majority of these modifications are found on the easily accessible histone termini, but globular domains situated at the core of the nucleosome can also be affected. Some of these modifications are specifically associated with cellular processes such as transcription, DNA repair or chromatin condensation.

Post-translational histone modifications, together with their “writers”, “erasers” and “readers” form the basis of the “histone code” hypothesis. This hypothesis suggests that specific PTMs, alone or in combination with others, serve as signals that regulate cellular processes ⁷⁹.

In the majority of cases, PTMs act indirectly as docking sites for modification recognizing effector proteins that can in turn recruit other proteins or change or maintain chromatin structure ⁷⁶. Some PTMs can also act directly on nucleosome structure by altering the charge of a residue, thereby changing histone-DNA or inter-histone affinities ⁸⁰ or facilitating its remodeling ⁸¹.

The most studied post-translational histone modifying enzyme families are histone acetyltransferases (HAT) / deacetylases (HDAC) for acetylations and histones methyltransferases (HMT) / demethylases (HDM) for methylations ⁸²⁻⁸⁴. Acetylation and methylation along with phosphorylation and ubiquitination, are the most researched histone modifications. The action mechanisms of some principle modifications are briefly detailed in the following paragraphs.

Acetylation of histones

Acetylation occurs with the addition of an acetyl molecule on the amine group of lysine residues and is mainly present on the N-terminus of H3 and H4 histones. It is catalyzed by histone acetyl transferases (HATs) through transfer of an acetyl group from acetyl-CoA molecule to the ϵ -amino terminal group of lysine residues, while the removal of acetyl groups is ensured by histone deacetylases (HDACs) that recognize acetylated lysines via their bromodomains ⁸⁴.

Acetylation serves to neutralize the charge of the residue, thus by loosening the interaction of H3-H4 with DNA, facilitates access to DNA ⁸⁵. Henceforth, acetylation is associated with active transcription and hypoacetylation is a mark of transcriptionally inactive chromatin ⁸⁶. However, studies based on mutation of lysine residues revealed that the observed effect on transcription is not dependent on the specific position of lysine residues but is rather due to the modification of the general charge of the N-terminal tail ^{87,88}.

Active promoters and euchromatic regions are enriched with histone acetylations. One of the best studied PTMs is the acetylation of H4K16 found at promoters and actively transcribed genes. By neutralizing the charge of the lysine residue, H4K16ac alters the interaction of H4 with the neighboring nucleosome and prevents formation of higher order chromatin structure. Notably, this modification needs to be removed from chromatin during the G2/M phases of the cell cycle to allow for proper chromosome condensation ^{89,90}.

Methylation of histones

Histone methylation mainly occurs on lysine and arginine residues prominently on histones H3 and H4. However, recently, other residues such as glutamine, aspartic acid and proline were determined to be targets for methylation. Lysine residues can be mono, di or trimethylated whereas arginine residues can be mono- or dimethylated. Histone methyltransferases (HMT) are responsible for the transfer of a methyl group from the S-adenosyl methionine (SAM). Most HMTs contain a highly-conserved SET domain that catalyzes this transfer. Even though histone methylation was initially thought to be permanent, the identification of histone demethylases indicated that this type of modification could be reversed ^{76,77}.

Contrary to acetylation, methylations do not alter the charge of the residues they target and show no consensus in their function. Each enzyme is generally specific to a residue and the function of a methylation is dependent on the position of the target residue as well as the number of bound methyl groups. In particular, methylations of H3K4, H3K36 and H3K79 are associated with active transcription whereas the di- and trimethylation of H3K9 and H3K27 are rather linked to heterochromatin and repression of transcription^{82,84}.

The presence of these marks allows recruitment of proteins necessary for the formation of heterochromatin. H3K9me2 or H3K9me3 allows the recruitment of HP1^{91,92} whereas H3K27me2 and H3K9me3 are associated with the recruitment of Polycomb proteins⁹³. Interestingly however, the mono-methylation of H3K9 or H3K27 is associated with active transcription. H4K20 is another example for different functions dependent on the number of attached methyl groups. While H4K20me1 presence in promoters is associated with active transcription, H4K20me2 is associated with DNA damage response and H3K20me3 with repression of transcription^{94,95}.

Phosphorylation of histones

Histone phosphorylation is catalyzed by specific kinases that transfer a phosphate group from ATP to the hydroxyl group of mainly serine, tyrosine or threonine residues, thus introducing a negative charge to the histone. This process is reversible through hydrolysis by phosphatases^{96,97}.

In general, phosphorylation is an important element in cellular signaling as it can rapidly regulate protein function and localization. Until recently, phosphorylation of histones seemed to attract lesser attention in comparison to acetylation or methylation. However, recent research has elucidated prominent functions for this modification⁹⁶. One of the most studied examples of histone phosphorylation is the phosphorylation of the serine residue at position 139 of the histone H2A variant H2A.X. This phosphorylation is an important initiator event in DNA damage response and repair⁹⁸.

Phosphorylation also plays an important role during mitosis as H3S10ph and H3S28ph are necessary for chromosome condensation at the beginning of mitosis⁹⁹⁻¹⁰¹. In particular, H3S10 phosphorylation by Aurora B kinase allows the recruitment of Hst2p, a histone deacetylase, which removes an acetyl group from the H4 lysine 16. The N-terminal end of H4 can then interact with the same end of the neighboring nucleosome and allow

condensation of chromatin ^{101,102}. The H3S10ph mark is also antagonistic to the H3K9me3 mark which results in the dissociation of HP1 and probably serves in the proper condensation of the mitotic chromosome ¹⁰³. H3S10ph is also linked to regulation of transcription as it is coupled to the acetylation of nearby residues at K14 and K9 ^{104,105}. Likewise, H3S28 phosphorylation interferes with H3K27me3 and interaction with the PRC2 repressive complex, inducing a methyl-acetylation switch ¹⁰⁶. Similarly, the presence of H3T3ph during mitosis weakens the association of TFIID to H3K4me3 that serves to repress transcription during mitosis ¹⁰⁷.

In addition to the three marks briefly described above, histones can be targets to various chemical groups and peptides such as ADP-ribosylation, deamination, propionylation, butyrylation, citrullination and crotonylation and other marks are being discovered constantly ^{77,78}. Notably, with 21 potential modification sites, the N-terminal tail of histone H3 (first 40 residues) is the most heavily modified tail of all four core histones ⁷⁷.

In general, by acting individually or in combination, PTMs offer a means to alter the interaction of nucleosomes with DNA, other nucleosomes and external factors (transcription factors, repair proteins, etc.) with direct implications on chromatin structure and function.

1.3.4. Remodeling factors

Variation in chromatin can also be introduced by non-covalent modification of chromatin thanks to chromatin remodeling factors. Through ATP-dependent mechanisms, chromatin remodeling complexes regulate chromatin accessibility through displacement, removal or reassembly of nucleosomes on DNA. Their effect can be global, throughout the nucleus, or localized to specific genomic sites ^{108–110}. The removal of a nucleosome may be followed by its replacement by another nucleosome containing a histone variant.

There are four families of chromatin remodelers defined by the structure of their ATPase domains. All four families share a high affinity for nucleosomes, histone modifications, DNase dependent ATPase domains involved in DNA/histone interactions and domains or proteins that interact with transcription factors ¹¹¹.

The SWI / SNF family (SWitch / Non-Fermentable Sucrose) is defined by the presence of an N-terminal domain SANT and at least one bromo-domain in the C-terminal part. The

first members of this family were first discovered in yeast ^{112,113}. Later they were described to be involved in all mechanisms requiring nucleosome sliding and removal such as transcription, development, DNA repair, replication. However they haven't been described in chromatin assembly ¹¹⁴.

The INO80 (INOsitol requiring 80) family members have a split ATPase domain with a "spacer" in the middle of its sequence. This structure allows interaction with the Rvb1 or Rvb2 proteins and adds helicase activity to the complex. Members of the INO80 family are involved in diverse functions such as activation of transcription, DNA repair and histone dimer repositioning ¹¹⁵. More specifically, they are responsible for the regulation of the incorporation of H2A.Z into chromatin ¹¹⁶. They are also involved in repair mechanisms, particularly in the repair of double strand breaks ^{117,118}.

The CHD (Chromodomain-Helicase-DNA binding) family is defined by the presence of at least one chromodomain, usually at the N-terminal position. Chromodomains specifically recognize methylated histones. Members of the CHD family are mainly involved in the regulation of gene expression through sliding and removal of nucleosomes ^{119,120} and have important roles in development ¹²¹. NURD (NUcleosome Remodelling and Deacetylase) is a particular member of this family, in addition to its ATP-ase activity, it has a constant HDAC (Histone deacetylase) subunit ¹²² which is associated with repression of transcription.

The ISWI (Imitation SWItch) family members are characterized by the presence of two SANT domains in the C-terminal region, but differ from the SWI / SNF family members by the absence of a bromodomain. This family is also involved in regulation of transcription through nucleosome sliding and reassembly ¹²³. ACF (ATP-utilizing chromatin assembly and remodeling factor) is a member of this family and regulates the nucleosome spacing together with NAP1 chaperone. ACF also enables the association of new histone octamers on DNA after replication ¹¹¹.

1.3.5. Histone chaperones

Histone chaperone proteins ensure the storage, deposition and removal of histones into and from chromatin ¹²⁴⁻¹²⁶. They interact with the hydrophobic surfaces of histones and prevent them from interacting with nucleic acids or any charged cellular constituents other than the

target DNA, during assembly and disassembly of nucleosomes. Therefore, free histones are fundamentally non-existent in the cellular context ¹²⁷.

Some chaperones like the FACT (FACilitates Chromatin Transcription) complex or NAP1 (Nucleosomes Assembly Protein 1) are capable of binding both H2A-H2B and H3-H4 dimers while others like CAF-1 (Chromatin Assembly Factor-1) can only bind H3-H4 dimers ¹²⁸. Also, some chaperones are capable of binding several variants of a given histone while others are specific to only one histone variant. For example, HJURP (Holliday Junction Recognition Protein) is indispensable for the incorporation of the centromere specific histone H3 variant, CENP-A (or CenH3) ¹²⁹.

Additionally, some chaperones are linked to only one function, like NASP (Nuclear Autoantigenic Sperm Protein that serves for stocking histones in the nucleus after their synthesis ¹³⁰. Another example is ANP32E which serves only for the eviction of H2A.Z containing nucleosomes from chromatin but not for their incorporation ¹³¹

1.3.6. Histone variants

Introduction of histone variants into chromatin is another strategy used by eukaryotic cells to engender variation in chromatin structure and is the main focus of this study. As mentioned previously in section 1.1.1, histone variants are non-allelic isoforms of replication-dependent histones, coded by different genes often found as a single copy ¹³. With the exception of histone H4, all replication dependent histones have variants that can replace them throughout different phases of the cell cycle and diversify chromatin structure and function (Figure 1-6). Histone variants also have the particularity of being expressed at different times in development, being associated to specific DNA sequences, and being tissue-specific ^{132,133}. They may differ from their “conventional” counterparts at their N-terminal tail (*i.e.* macroH2A), at their histone-fold domain (*i.e.* H2A.Bbd) or by only a few amino-acids (*i.e.* H3.3) ¹³⁴. By modifying chromatin structure, they play an important role in conferring novel local properties to chromatin and are implicated in many cellular processes such as transcription, repair and cell division ^{135–137}. Notwithstanding, most histone variants keep the same PTM sites as their conventional counterparts, allowing recognition by similar chromatin regulating proteins ^{134,138,139}.

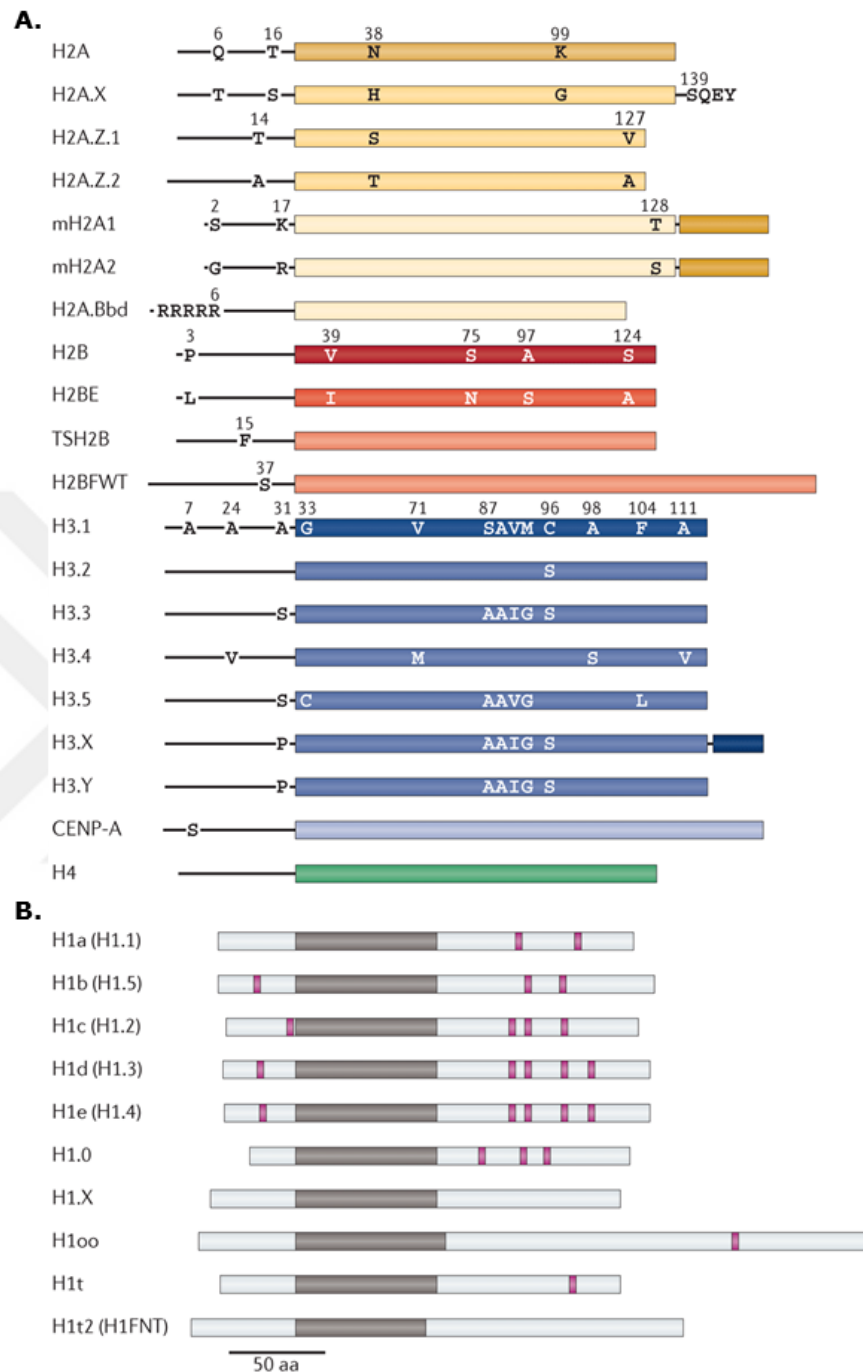


Figure 1-6. Human core and linker histone variants. **A.** Variants of core histones H2A (yellow), H2B (red), H3 (blue) and H4 (green). Unstructured N-terminal tails are depicted in black lines and the major residues that differ are given. Darker shades of color correspond to higher divergence from the replication dependent variants shown on top of the lists. **B.** Variants of human linker histone H1. As H1 variants diverge greatly in sequence, specific residues are not stated. Lighter shade of grey corresponds to unstructured C-terminal and N-terminal regions. Globular domains are depicted in dark grey. Phosphorylation sites are indicated as magenta boxes. Alternative names are given in parentheses. Figure reproduced from Maze et al. 2014¹⁴⁰ with permission (Appendix B).

Histone H1 variants

Linker histone H1 plays a prominent role in higher-order chromatin organization. In mammals, in addition to the five replication dependent somatic H1 histones (H1.1, H1.2, H1.3, H1.4 and H1.5), there are 2 replication independent somatic variants (H1.0 and H1X), three testis-specific variants (H1t, H1T2m and H1LS1) and one oocyte-specific variant (H1oo)^{9,141}. Even though H1 linker histones are strongly implicated in chromatin compaction in general, it remains unclear whether H1 variants have specific functions.

Histone H2A variants

The histone H2A family (H2A.Z, H2A.X, macroH2A.1, macroH2A.2, H2A.Bbd) constitutes one of the largest histone families with known variants in eukaryotes¹⁴². MacroH2A, expressed only in vertebrates, has a distinctly large globular domain (macro) and is by far the largest histone with a molecular weight of 40 kDa. Despite its relatively large size, nucleosomes containing macroH2A have similar structures to nucleosomes containing the conventional H2A histone¹⁴³. However, macroH2A-H2B interaction is stronger than H2A-H2B interaction resulting in a more stable nucleosome¹⁴⁴. This structural property together with its involvement in X chromosome inactivation in female mammals suggests a specific function in gene silencing for this variant¹⁴⁵. Studies have shown that macroH2A is also present at promoters of inducible genes in certain stem cells where it is thought to play a role in preventing reprogramming of differentiated cells by restraining chromatin remodeling^{146–148} and reducing acetylation levels¹⁴⁹.

On the other hand, H2A.Bbd (Barr body deficient), as the name suggests, is excluded from the inactive X chromosome. H2A.Bbd does not have a C-terminal tail. Consequently, H2A.Bbd containing nucleosomes are more relaxed with the core particle binding only 120 base pairs of DNA forming a less stable nucleosome that allows easier access to DNA. It is therefore associated with transcriptionally active chromatin^{134,135,150}.

H2A.X plays a specific role in the repair of double strand DNA breaks. When a break occurs, H2A.X is phosphorylated at serine 139. The phosphorylated form, denoted γ H2A.X, then forms foci of several megabases around the damage¹⁵¹. The presence of this phosphorylation triggers the recruitment of the repair machinery and chromatin remodeling factors¹⁵², always with the aim of facilitating access to DNA at the damage site.

H2A.Z is one of the most evolutionarily preserved histone variants and is known to be essential for the survival of mice ¹⁵³. It only shows 60% homology to the conventional histone H2A. Its genome-wide localization has been extensively studied in yeast and vertebrates. In yeast, it is located both in promoters and in insulator regions between repressed and active chromatin ^{154,155}. In vertebrate cells H2A.Z has been found in active promoters, in a probable acetylated form ¹⁵⁶, in facultative heterochromatin, in a monoubiquitylated form ¹⁵⁷ and at centromeres ¹⁵⁸. These different localizations suggest different roles for H2A.Z, notably roles in both gene activation and inactivation as well as structuring of centromeric regions has been proposed yet has not been shed light on. It is suggested that other factors associated to chromatin at the same site and time determine the role of H2AZ at these sites ¹⁵⁹. H2AZ is also involved in stem cell differentiation ¹⁶⁰, DNA repair ¹⁶¹ and cell senescence ¹⁶².

Histone H2B variants

Unlike the histone H2A and H3 families, the histone H2B family does not have ubiquitously expressed variants. Instead, in mammals, there are two testis-specific variants TH2B and H2B.FWT. TH2B replaces H2B in spermatocytes. Nucleosomes containing TH2B and TH2A are described to be very unstable ¹⁶³ constituting a transient state before histones can be replaced by protamines during spermatogenesis¹⁶⁴. H2B-FWT has only been described in humans and great apes. It would appear to be localized at telomeres during gametogenesis yet its role remains to be elucidated ¹⁶⁵

Histone H3 variants

The histone H3 family consists of at least 5 well-identified mammalian members: the replication dependent H3.1 and H3.2, the centromere specific variant CenH3 (or CENP-A), testis specific H3.4 (also named TH3, H3.1t or H3.t) and replication-independent H3.3¹⁶⁶. CenH3 is common to all eukaryotes, H3.3 and H3.2 are most likely present in all metazoans while H3.1 and H3.1t are only observed in mammals. A part from CenH3 which shows only 50% homology in its histone fold domain to the other H3 variants, all subtypes only differ by a few amino-acids. CenH3 variant marks centromeres and is required for kinetochore assembly during mitosis and meiosis ^{167–169}.

H3.4 variant differs by only four amino-acids from H3.1 and by five amino-acids from H3.2. It is a testis-specific variant which, in humans, is only expressed in spermatocytes

where it represents the majority of H3 histones¹⁷⁰. The H3.4 containing nucleosome is less stable than the conventional nucleosomes as the interaction between the H3.4-H4 tetramer and the H2A-H2B dimers is weaker. H3.4 might therefore have a role in constituting a transient state before histones' replacement by protamines during spermatogenesis¹⁷¹.

Additionally, in Catarrhines, two genes initially identified as pseudogenes were in fact shown to code for the H3.X and H3.Y variants which are very similar in sequence to H3.3¹⁷². They are expressed in both normal and cancerous tissues. Depletion of H3.Y leads to an increase in H3.X expression, indicative of functional links between the two variants. Moreover, genome-wide H3.Y mapping studies showed that H3.Y is enriched around TSS of active genes suggesting a role for this variant in regulation of transcription¹⁷³.

The latest discovered histone H3 variant is H3.5, found in hominids and expressed in seminiferous tubules in humans. This variant is enriched at TSS and is described to be capable of complementing the loss of H3.3^{174,175}. As H3.3 is the focus of this dissertation, its structure and function are detailed in the following chapters.

Chromatin structure needs to be permissible to variation for correct progression of various nuclear functions. Variations are introduced through mechanisms of epigenetic regulation. During replication, transcription and DNA damage repair, chromatin remodelers and histone chaperones serve to regulate positioning and spacing between nucleosomes while histone modifications or replacement by histone variants confer additional properties to the chromatin to ease access or provide binding sites for effectors. Covalent and non-covalent modifications are often linked as some remodeling complexes have subunits that also serve to modify chromatin¹⁷⁶. A very good example to the crosstalk between different actors of epigenetic regulation is the silencing of the inactive X chromosome in female mammals. During this process, the non-coding Xist RNA, DNA methylation, histone variant macroH2A, histone modifications and their effectors all work together to allow for proper restructuration of chromatin in the silenced X chromosome¹⁷⁷.

1.4. H3.3 and the histone H3 family

1.4.1. Evolution of histone H3 variants

Non-centromeric H3 variants, especially H3.3, are extremely well conserved in evolution. In the budding yeast, *Saccharomyces cerevisiae*, there is only one non-centromeric H3 variant which shows more than 91% similarity to the mammalian H3.3.¹⁶⁹ In an evolutionary point of view, it is suggested that histone H3 variants evolved from a replication independent H3.3 like variant rather than the replication-dependent H3.1 or H3.2. Throughout evolutionary time, a number of histone H3 variants have evolved independently within most eukaryotic species. Replication dependent variant genes most likely expanded to keep up with the requirements of larger genomes while replication independent variants undertook more function or tissue specific roles^{167,178,179}.

1.4.2. Genetic structure of H3 and H3.3 in mammals

In mice, H3.1 and H3.2 coding genes are found in two different clusters (Figure 1-7).

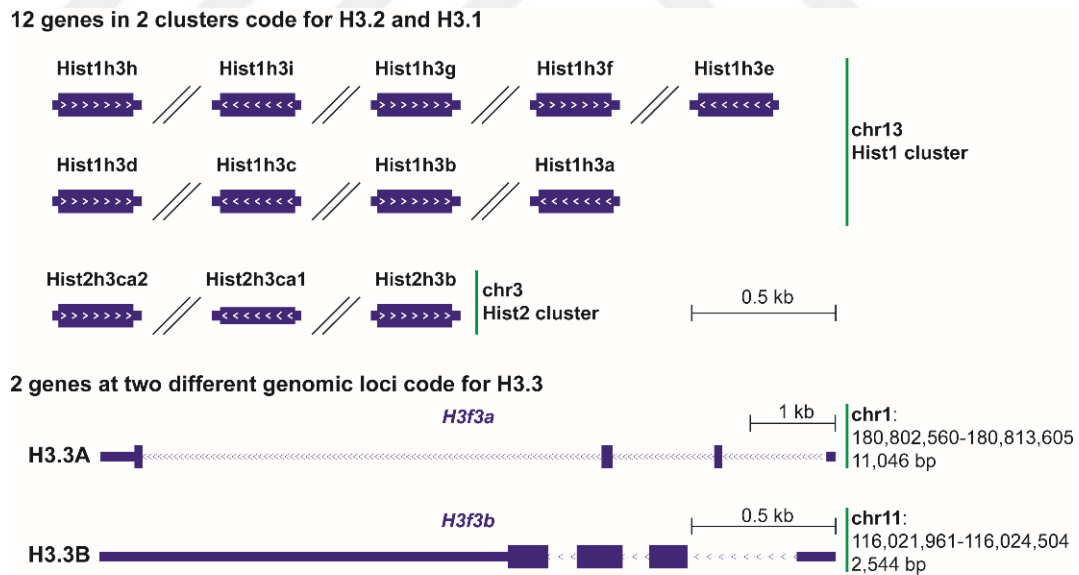


Figure 1-7. Genomic organization of H3 coding genes 12 genes code for the replication dependent H3.1 and H3.2 and are found in the Hist1 (chromosome 13) and Hist2 (chromosome 3) genes clusters along with other genes coding for the other RD histones. These genes do not contain introns. H3.3 is coded by H3f3a (chromosome 1) and H3f3b (chromosome 11). These two genes are single-copy and contain introns. Gene names are indicated above the gene and genomic locations are indicated on the right. Coding exons are depicted by thick rectangles while non-coding exon sequences are depicted by thin rectangles. Arrows indicate the strand from which the gene is transcribed. UCSC Genome Browser view from the mm10/GRCm38 assembly of the mouse genome^{14,180,181}.

Hist1 is located on chromosome 13 and contains 9 genes (Hist1H3a-Hist1H3i). Hist2 is located on chromosome 3 and contains 3 genes (Hist2H3b, Hist2H3ca1, Hist2H3ca2) ¹⁰. As mentioned previously, the mRNA produced by these clusters lack introns and poly-A tails and are processed by a specific machinery that recognizes the stem-loop termination structure. This structure allows for rapid protein production to fulfill the heavy demand for histones at the beginning of replication ¹¹.

H3.3 on the other hand is transcribed from two different single-copy genes, *H3f3a* on chromosome 1 and *H3f3b* on chromosome 11, in a replication-independent manner ¹⁸² (Figure 1-7).

As opposed to H3.1 and H3.2 coding genes, the mRNA produced by the two genes have a more typical structure, with introns and a poly-A tail. Interestingly, even though they code for the exact same amino-acid sequence, they are found in different genomic loci and have different nucleotide sequences. As a matter of fact, even their coding sequences only show 77% similarity (Figure 1-8). However, the main difference between the two mRNA sequences, lies in their unusually long 5' and 3'UTR regions and their developmental expression ^{14,168,182-184}.

Regardless of the ubiquitous expression of the H3.3 protein, the transcription of *H3f3a* and *H3f3b* genes show differential levels of expression according to tissue types¹⁸⁵. A study conducted in *H3f3a* or *H3f3b* deficient mouse fetuses showed a higher expression of *H3f3a* in the brain and similar expression levels of both genes in liver, kidney and lung ¹⁸⁶. Another study also confirmed the differential expression of *H3f3a* and *H3f3b* in different tissues while demonstrating that they are most highly expressed in testes and ovaries with a preference for *H3f3a*, suggesting an important role for H3.3 in germ cell development ^{182,187,188}. Regulation of H3.3 coding gene expression can be at the level of transcription, mRNA processing, or mRNA stability.

Even though the two genes seem to both show a basal expression, they have very different promoter structures. While the *H3f3a* promoter is characterized by a GC-rich sequence upstream of TSS with 4 SP1 binding motifs, the *H3f3b* promoter has a TATA box and several CCAAT boxes as well as an Oct and CRE/TRE element. SP1 is a transcription factor involved in the transcriptional regulation of housekeeping genes ¹⁸⁹. The promoter structure of *H3f3a* together with its expression patterns point to a basally expressed gene ¹⁴.

The CRE/TRE element is indicative of a responsive promoter and may explain the differences in *H3f3b* transcription levels during differentiation and development^{190,191}. Taken together, mRNA expression and promoter analyses suggest that although *H3f3b* also contributes to basal H3.3 expression, *H3f3a* is the basally expressed H3.3 gene¹⁴.

The length of 3'UTRs is described to have an important role in translational regulation of protein expression¹⁹². Both H3.3 coding genes have considerably long 3'UTRs. Furthermore, *H3f3b* 3'UTR is twice as long and possesses several polyadenylation sites (Figure 1-8). Through alternative use of these sites, *H3f3b* mRNA can produce at least 3 transcripts of different lengths. Moreover, it has been reported that the polyadenylation site used is indicative of tissue-specific expression in *Drosophila melanogaster* and in *Drosophila hydei*¹⁹³. UTRs contain many regulatory elements that interact with RNA binding proteins (RBP) that can either be permissive or repressive to gene expression. An RNA-binding protein, PLAUF, similar to the human or murine AU-binding factor 1 (AUF1) family proteins was identified to bind to 3'UTR of the H3.3 coding gene in the sea urchin (*P. lividus*)¹⁹⁴. The single H3.3 coding gene in the sea urchin is most similar to the human *H3f3b* gene with a 3'UTR that contains several AU-rich element (ARE) motifs¹⁹⁵. The presence of ARE motifs has been associated with faster mRNA degradation, however, binding of AUF1 to these elements could have either positive or negative effects on mRNA translation or stabilization¹⁹⁶. Additionally, AUF1 proteins, like H3.3, show differential distribution in mouse tissues and the two proteins colocalize at specific tissues or stages of development^{194,197}. Recent advances in the field of non-coding RNA mediated regulation have shown that direct interaction between AUF1 proteins non-coding RNA such as miRNA and lncRNAs likely affects post-transcriptional regulation of gene expression¹⁹⁸.

Lastly, *H3f3a* and *H3f3b* genes also differ greatly in secondary structure and sequence in their 5' UTR as well as in their introns. It is therefore likely for regulatory elements found in these sites to contribute to the differential regulation of these genes.

1.4.3. The H3.3 containing nucleosome

H3.3 is a 136 amino-acids long protein with an approximate 15kD molecular weight and it differs by only 4 or 5 amino-acids from the replication dependent H3.2 and H3.1 respectively (Figure 1-9). Structurally, these differences are mostly located towards the core of the nucleosome (Figure 1-10). In fact, only the H3.3S31 residue, found on the H3 N-terminal tail, is accessible. The remaining different residues, A87, I89 and G90, being in the H3 α -helix 2 covered by DNA, are inaccessible in the context of the nucleosome^{1,15,59}. The C-terminal tail of H3 is only a few amino-acids long and is located towards the center of the nucleosome.

Studies on the crystal structure of H3 variant containing nucleosomes showed no considerable differences between H3.1, H3.2 and H3.3 containing nucleosome structures. Instead, the differing amino-acids in the histone-fold play a role in histone recognition by H3.3 specific chaperones^{141,171,200}. In fact, a study where the differing residues of replication dependent H3 histones were substituted by the corresponding residues of H3.3 in *Drosophila melanogaster*, resulted in a replication independent incorporation of mutant histones²⁰¹.

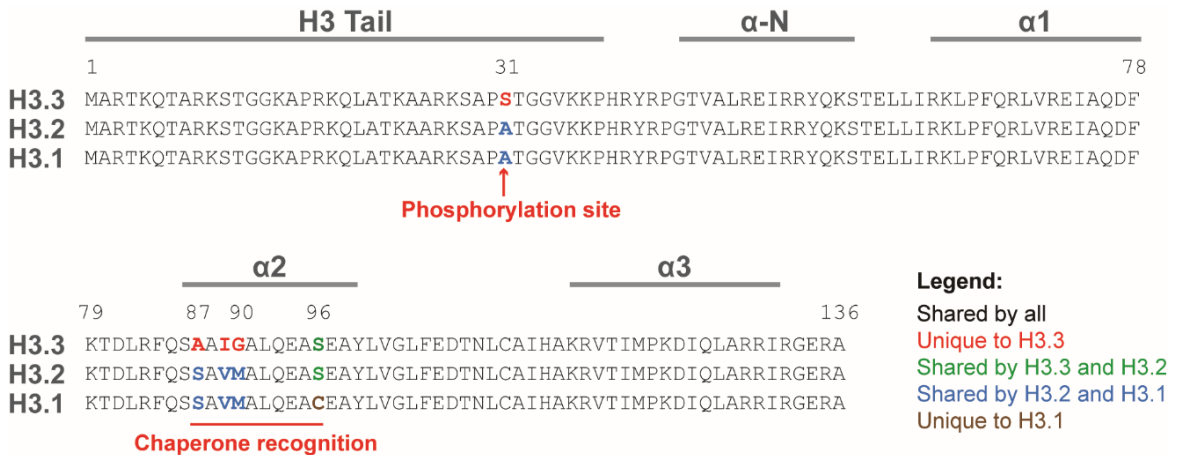


Figure 1-9. Amino-acid sequence alignment of mammalian histone variants H3.3, H3.2 and H3.1. ... H3.3 differs from its replication-dependent counterparts at two regions. The H3.3 specific serine residue at position 31 is a target for phosphorylation whereas residues at positions 87, 89, 90 and 96 are implicated in H3.3-specific chaperone recognition and replication independent incorporation^{141,200,202}. The replication dependent variants H3.2 and H3.1 differ only at position 96. Secondary structure elements are indicated¹

While the H3.3 specific S31 residue is not implicated in the recognition of H3.3 by specific chaperones, it constitutes an additional phosphorylation site on the N-terminal tail on H3.3. A hydroxyl group at position 31, so close to the important K27 modification site, could create additional repulsive electrostatic forces that could alter interactions between the nucleosome and chromatin modifying enzymes or other effectors^{202,203}.

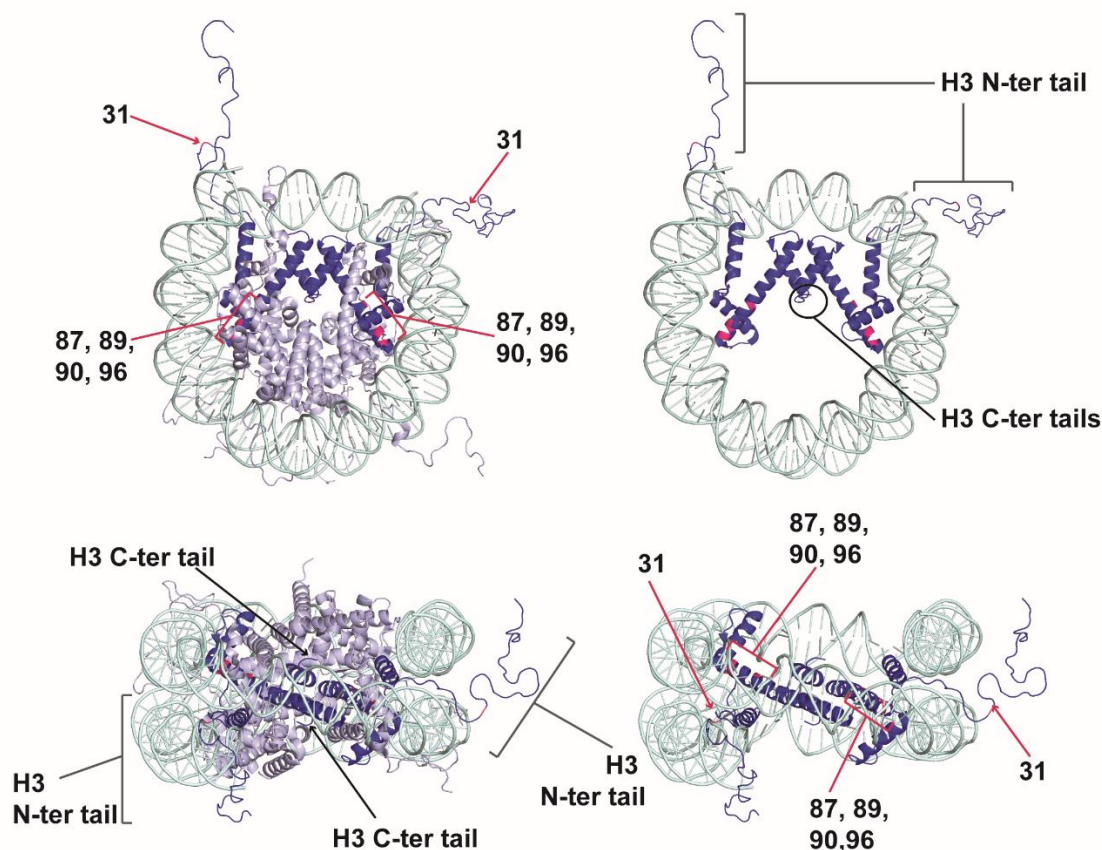


Figure 1-10. Differential amino-acids between H3.3 and H3 variants in reference to the nucleosome structure.. In the context of the nucleosome, the residue at position 31 of H3.1, H3.2 and H3.3 variants, is on the N-terminal tail and is accessible. The remaining differing residues, at positions 87, 89, 96 and G96, are in the $\alpha 2$ helix of H3 and are covered by DNA, therefore inaccessible^{1,15,59}. The short C-terminal tail of H3 is located in the center of the nucleosome. Crystal structure of the nucleosome is viewed from two angles. On the left is the whole nucleosome while on the right all histones but the two H3 have been removed for clear visualization. Histone H3 is in blue with varying residues highlighted in red, the other histones are in light blue, DNA double helix is in pale cyan. Red arrows point to the location of residues at position 31 of H3, red brackets indicate positions 87, 89, 90, 96 of H3, black arrows point to C-terminal terminal of H3 and black brackets indicate N-terminal tails of H3. Image reconstructed from nucleosome core particle high resolution structure (PDB ID: 1KX5¹⁵) shown in cartoon representation using PyMol software²¹

1.4.4. H3 variant incorporation into chromatin

Replication dependent H3 histones are homogenously distributed through chromatin and are incorporated by chromatin assembly factor-1 subunit A (CAF-1) which interacts directly with the DNA polymerase clamp, PCNA (Proliferating Cell Nuclear Antigen) at replication forks¹⁶. On the other-hand, H3.3 incorporation is independent of cell cycle stage and is carried out by two distinct chaperones depending on the genomic site of incorporation since H3.3 distribution is not homogenous as detailed in the following sections^{16,204–207}.

The histone chaperone complex HIRA/UBN1/CABN1 is accountable for H3.3 localization at promoters and gene bodies through the recognition of the H3.3-specific G90 residue by UBN1^{208,209}. With a distinct mechanism, yet still through recognition of the H3.3G90 residue, DAXX (Death domain Associated protein 6) and its chromatin remodeling partner ATRX (Alpha-Thalassemia Mental Retardation Syndrome protein) are responsible for H3.3 deposition at repetitive sequences like pericentromeric heterochromatin, telomeres and ERV elements^{205,210–212} (Figure 1-11).

H3.3 also shows distinct turnover rates depending on the genomic site of incorporation. H3.3 containing nucleosomes found in enhancers and promoters and associated with active chromatin marks are subject to a rapid turnover while silenced regions containing repressive marks show slow turnover^{213–215}.

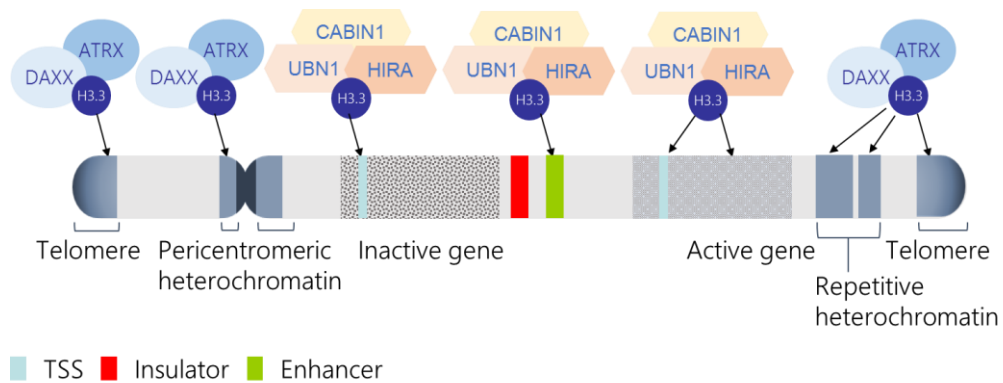


Figure 1-11. H3.3 incorporation into chromatin. H3.3 incorporation into chromatin is heterogenous and is carried out mainly by two distinct chaperone complexes depending on the genomic site of incorporation. HIRA complex targets H3.3 to promoter regions, regulatory sequences and gene bodies via interaction of the UBN1 subunit with H3.3 specific G90 residue. DAXX together with its partner ATRX, targets H3.3 to telomeres, pericentromeric regions and other heterochromatic sites containing repetitive DNA sequences^{205,210–212}.

1.5. Genomic localization of H3.3 and function

1.5.1. Early studies on genome-scale H3.3 localization provide elements to speculate on H3.3 function

Genome-wide mapping studies in drosophila produced consistent results on the role of H3.3 in regulation of gene expression. Ectopically expressed tagged-H3.3 was strongly enriched at transcriptionally active genes bodies ^{216,217} but also at cis-regulatory elements found near or outside of both active and repressed genes ²¹⁸.

However, the first genome-wide mapping studies in vertebrates were less clear and even conflicting at times. In one study conducted in mouse hematopoietic cell lines, H3.3 was found enriched only in promoter regions in ²¹⁹ while in another study in chicken erythrocytes, H3.3 was found enriched at either upstream regulator regions or gene bodies, or neither, or both of active genes. Furthermore, the latter reported that there was no link between H3.3, transcription levels and active transcription marks as H3.3 was enriched in both active and inactive genes ²²⁰. Though originally not present at TSS of inducible genes in human T-cells, upon induction, H3.3 was found to localize at gene-bodies ²²¹. A study in HeLa cells showed a positive correlation between enrichment of H3.3 in TSS and gene bodies and transcriptional activity combined with a lack of enrichment of H3.3 in inactive genes ¹⁵⁹. However in another study, H3.3 was reported to be constitutively present in the promoters of inactive inducible genes in mouse embryonic fibroblasts (MEFs) and extended to their gene bodies upon induction ²²².

While most of these studies used transfected cell lines that ectopically expressed tagged-H3.3, in 2010 a group reported the use of genetically modified mouse ESCs that produce tagged-H3.3 proteins endogenously ²⁰⁵. This study showed that H3.3 was enriched at the TSS of active and repressed genes and extended to gene bodies of activating genes upon induction of differentiation. H3.3 at these regions also correlated with markers of active transcription such as phosphorylated RNA Pol II and active histone marks. Surprisingly, this same study found that H3.3 was not only confined to genes and their promoters but was also enriched at telomeres. It also showed that the replication-independent incorporation of H3.3 at telomeres depended on the DAXX/ATRX chaperone in contrast to the HIRA complex used for the other sites. This result was in line with other studies that involved H3.3 in the maintenance of telomeres in association with ATRX ^{204,223}. This wasn't the first

account of H3.3 positioning at repetitive sequences as H3.3 was found enriched in its phosphorylated form at pericentromeric regions during mitosis ²⁰². Additionally, in more recent years, H3.3 has also been reported enriched in endogenous retroviral repeat elements and involved in their silencing ²²⁴.

The heterogenous enrichment of H3.3 in chromatin suggests roles in different cellular processes and advances made in the last years have provided crucial elements in understanding the mechanisms behind H3.3 involvement in these processes.

1.5.2. Role of H3.3 in active transcription

Unlike the replication dependent H3 histones, H3.3 distribution on the genome is heterogenous and provides clues to the protein's function. Early genome-wide mapping studies position H3.3 in sites of active transcription ^{201,219}. More precisely its enriched at promoters, in regions around transcription start sites (TSS) and in gene bodies of highly expressed genes ²⁰⁵. Furthermore, H3.3 is enriched by more than 2-fold in modifications representative of active chromatin such as K9, K14, K37 and K18 acetylations and K36me1, K36me2, K79me2 methylations. Similarly, H3.3 enriched promoters show higher H3K4me3 enrichment than promoters containing H3.2 promoters ²²⁵. Moreover, H3.3 is deprived of repressive methylation marks on K27 ^{166,226}. These modifications have also been observed in *Drosophila* and *Arabidopsis* ^{138,227,228}. Also, H3.3 has been found to antagonize H1 incorporation, favoring a decondensed chromatin state ^{229,230}. The activating effect of H3.3 was observed in chicken erythroid cells when expression of specific genes varied in response to the overexpression of ectopic H3.2 and H3.3 ²²⁰. Another example was seen when the expression of interferon stimulated genes decreased upon knock-down of H3.3 in mouse fibroblasts treated with interferon ²²². Surprisingly however, recent transcriptomic studies in *H3f3b* depleted MEFs and in developing H3.3-null mouse embryos showed that in the complete absence of H3.3, gene regulation was not seriously affected ¹⁸⁷.

1.5.3. H3.3 interaction with H2A.Z at promoters

The histone variants H2A.Z and H3.3 appear to play an important role in promoter structure ^{60,159,231}. Colocalization experiments showed that there is an enrichment of nucleosomes containing both H3.3 and H2AZ at promoters and enhancers of transcriptionally active genes as well as in the gene-body of some very active genes ¹⁵⁹. Furthermore, nucleosomes

containing both H3.3 and H2AZ were described to be less stable than nucleosomes containing H3.3 and H2A which in turn are less stable than H3-H2A containing nucleosomes ²³¹. This instability could have an important role in conferring ease of access to transcription factors at promoters and other regulatory sites. The occupancy of sites related to active transcription by these unstable nucleosomes could also play a role as “place holder” to prevent the remodeling of these regions with stable nucleosomes containing canonical histones ²³². It is thus suggested that the chromatin structure at vertebrate promoters and other regulatory sites is a very dynamic one, characterized by a rapid turnover where stable nucleosomes are replaced by less stable ones containing histone variants which can then be replaced by transcription factors ^{159,233}.

However, this notion is subject to controversy and in vitro studies showed that nucleosomes containing either or both variants have identical stability ^{141,234,235}. Nonetheless, the instability of H3.3 and H2A.Z containing nucleosomes observed in vivo can be accounted by the presence of PTMs or the existence of recently described subnucleosomal structures as a result of the action of chromatin remodeling complexes not yet identified ^{166,236}. Furthermore, the replacement of replication-dependent H3 histones with H3.3 with time in non-dividing cells also casts doubt on the lack of stability in H3.3 containing nucleosomes.

1.5.4. Role of H3.3 in heterochromatin maintenance and genomic stability

Although initially described to be a marker of active chromatin, growing evidence based on the development of ChIP-Seq studies, showed that H3.3 is also enriched at heterochromatic repeat regions such as pericentromeric heterochromatin, telomeres and more recently, endogenous retroviral repeat elements (ERVs) and silenced imprinted differentially methylated regions ^{202,205,224,237}. These repetitive sequences are usually found in the form of constitutive heterochromatin to prevent aberrant transcription which can disturb genome stability.

Telomeres consist of tandem DNA repeats that protect the ends of chromosomes from DNA degradation and inappropriate DNA repair. The maintenance of these structures is crucial for genomic integrity and cell survival and is effected by two mechanisms: the reverse transcriptase enzyme telomerase activity and the alternative lengthening of telomeres pathway (ALT) ²³⁸. Telomeres are enriched for the heterochromatin specific H3K9me3, H4K20me3 and DNA methylation and are marked by hypoacetylation of H3 and H4 ^{237,239}.

Furthermore, the absence of H3.3 leads to an increase in telomere sister chromatid exchange events (tSCE) indicating dysfunction in telomere maintenance and an important role for H3.3 at this site ¹⁸⁷.

Endogenous retroviruses (ERVs) are a subclass of retrotransposons which can be packaged and moved within the genome. Approximately 10% of the mouse genome is constituted of ERV repeats. Contrary to humans, ERVs can be transcriptionally active in the mouse genome ²⁴⁰. As these elements may contain regulatory sequences such as promoters, enhancers, polyadenylation signals and factor binding sites, they can disrupt or influence expression of nearby genes. Owing to the deleterious effects of a de novo integration, cells have evolved mechanisms to restrict retroviral activity. ERVs are silenced mainly by DNA methylation but also by H3K9 and H3K27 trimethylation ²⁴¹.

H3.3 may further be implicated in the maintenance of heterochromatic state in repetitive regions through its enrichment with the repressive K9me3 modification at these sites ^{226,242}.

During mitosis, the serine residues H3S10 and H3S28 found in the amino-terminal tail of canonical histone H3 are phosphorylated ¹⁰¹. The implication of these modifications on chromosome condensation during mitosis and meiosis in eukaryotes is well documented ^{99,243–247}. In addition to H3S10 and H3S28 residues, H3.3 is phosphorylated on its specific S31 residue at pericentromeric regions during mitosis ^{202,203,246}. H3.3S31 phosphorylation has also been observed in *Oikopleura dioica*, during mitosis and meiosis suggesting evolutionary conservation ²⁴⁸.

These phosphorylations are thought to interact with trans-acting proteins that regulate chromatin structure. However, the significance and action mechanism of H3.3S31 remains to be elucidated. A recent study showed that in the event of defects in chromosome alignment or segregation, the H3.3S31 mark extends over both chromatids amplifying the stress signal and resulting in p53 dependent cell-cycle arrest ²⁴⁹. Similarly, in ALT deficient cells, H3.3S31 phosphorylation was shown to extend to the entire chromosome following the increase of the CHK-1 kinase activity suggesting a role in maintenance of genomic integrity ²⁵⁰.

1.5.5. H3.3 in development and sexual reproduction

Over the past decade, several knock-out or knock-down studies in mice with the aim to elucidate H3.3 function were conducted. Inactivation of *H3f3a* by gene-trap²⁵¹ and production of a null-allele by targeted deletion^{186,252} resulted in reduced viability, growth deficiency and subfertility. *H3f3b* knock-out mice also showed reduced viability and infertility^{188,191}. Knock-down of H3.3 using specific siRNA yielded in defective nuclear envelope formation while another study using morpholinos showed variation in chromatin condensation^{229,253}. Interestingly however, a recent study showed that single knock-out mice did not show any observable defects in growth or fertility¹⁸⁷. Knock-out of both H3.3 coding genes resulted in inviable oocytes¹⁸⁶ and embryonic lethality as early as E6.5¹⁸⁷.

In vertebrates, similar to other variants, H3.3 replaces H3.1 and H3.2 in slowly- or non-dividing cells as replication-dependent H3 deposition is no longer in effect. This is particularly the case in neural tissues where H3.3 replaces H3.1 and H3.2 almost entirely. As a result, the over-all ratios of H3.3/ H3.1 or H3.2 vary with age. Furthermore, during senescence, H3.3 is replaced by its cleaved form which lacks the first 21 amino-acids. This replacement is followed by a decrease in cell cycle gene expression probably due to the loss of H3K4 related signaling²⁵⁴. Although not completely clarified, these data suggest a role for H3.3 in senescence and neuronal function^{255–258}.

Consistent with its conflicting genomic distribution as well as tissue and developmental stage-specific expression, H3.3 appears to be essential for development and is implicated in multiple cellular processes including transcription, genomic maintenance, cell division and senescence.

1.5.6. H3.3 in tumorigenesis

Cancer cells present a disturbed epigenetic landscape in addition to the classic genetic aberrations²⁵⁹. Histone variants, as an important element in epigenetic regulation, have been implicated in many different cancer types summarized in Figure 1-12²⁶⁰.

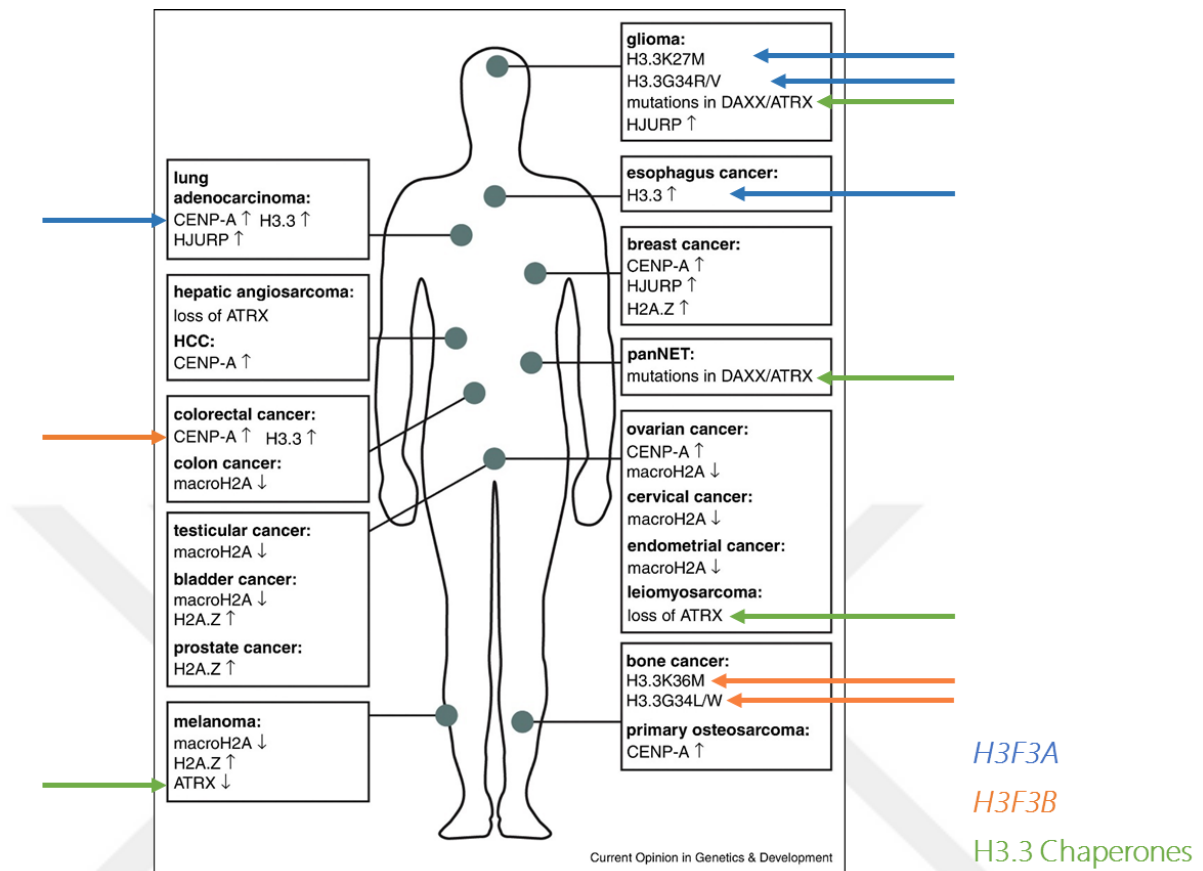


Figure 1-12. Contribution of histone mutations and deregulations in their expression to tumorigenesis in humans. Representative image of human tumors with histone mutations or due to deregulation of histone variant or chaperone expression. ↑: upregulation, ↓: downregulation of indicated variant. Blue arrows point to *H3F3A* mutations, orange arrows point to *H3F3B* mutations, and green arrows point to H3.3 chaperones. Figure adapted from Zink and Hake, 2016²⁶⁰ with permission (Appendix B).

More particularly regarding H3.3, missense somatic mutations in the human *H3F3A* gene identified in childhood brain tumors such as diffuse intrinsic pontine glioma (DIPG) and glioblastoma multiform were the first direct alterations in histones reported in cancer^{261,262}. These mutations lead to the substitution of the Lysine 27 residue to methionine (K27M) or isoleucine (K27I) and the Glycine 34 residue to arginine (G34R) or valine (G34V)^{261–263}. These mutations are located in sites of extensive post-translational modification and alter global methylation levels of H3K27 and H3K36. Even though the precise mechanisms of action are not known, a transcriptional effect is suggested given the function of the affected modifications^{264–266}. In addition to direct mutations, overexpression of *H3F3A* has been observed in esophageal and lung cancers^{267,268}.

Similarly, mutations in *H3F3B* resulting in H3.3K36M substitution have been identified in chondroblastoma and G34W or G34L substitutions resulting in giant cell tumors of bone ²⁶⁹. H3.3 mutations are also observed in osteosarcoma ²⁶⁹ and epithelial ovarian cancer ²⁷⁰. A recent study demonstrated that *H3F3B* expression increases significantly in colorectal tumor tissues and correlates with some clinicopathological properties ²⁷¹.

Mutations in the H3.3 chaperone, ATRX/DAXX have been found to co-exist with some K27M and all G34V substitutions in brain tumors ²⁶¹. ATRX/DAXX is also mutated in pancreatic neuroendocrine tumors ^{272,273}. As this chaperone is responsible for H3.3 deposition at repetitive sequences, it is suggested that tumorigenesis occurs as a result of disrupted genomic stability underlining the importance of H3.3 in this process ²⁶⁴. Table 1-1 recapitulates point mutations found in H3.3 and H3.3 chaperones observed in human cancer.

Table 1-1. Point mutations in H3.3 and its chaperones observed in human cancer Adapted from Buschbeck and Hake, 2017 ²⁷⁴ with permission (Appendix B).

Mutated Protein	Mutation	Cancer Tissue	Function
H3.3	K27M	Pediatric brain tumors and juvenile bone tumors	Inhibits the K27 methyltransferase activity of Polycomb repressive complex 2
	G34R, G34V and G34L	Pediatric brain tumors and juvenile bone tumors	Unknown function, but tight association with ATRX and DAXX mutations
	K36M	Chondroblastoma	Inhibits K36 methyltransferases
ATRX	Various in coding sequence	Cancers of the peripheral and central nervous system; less frequent in various other types of cancer	Mostly inactivating
DAXX	Various	Various, most frequent in the nervous system	Considered to be inactivating, less frequent than ATRX mutations

Although mutations in histone variants or changes in their expression have been clearly linked to tumorigenesis, the precise mechanisms of effect remain to be elucidated. It is therefore crucial to uncover the specific functions of histone variants such as H3.3 to give insight to disease onset and development while providing elements for developing new specific therapeutic strategies.

1.6. Aim

In order for the eukaryotic cellular processes to function correctly, chromatin structure needs to be dynamic and flexible. The replacement of core histone proteins by their variants confer new structural and functional properties to chromatin and serve to increase epigenetic plasticity. Two genes, *H3f3a* and *H3f3b*, with different nucleotide sequences and genomic loci code for the exact same amino-acid sequence of the histone variant H3.3 in mice. They vary in nucleotide sequence in their untranslated regions as well as in expression depending on developmental stage and tissue-specificity. In this dissertation, the products of these genes will respectively be called H3.3A and H3.3B for clarity purposes.

Contrary to the replication dependent H3.1 or H3.2, H3.3 is incorporated in a replication independent manner by at least two different chaperone complexes at distinct genomic sites²⁰⁵. Genome-wide H3.3B mapping studies point to a role in both active transcription and maintenance of heterochromatin.

Due to the minimal difference in amino-acid sequence between the replication dependent H3.1 (or H3.2) and H3.3, it has proven very difficult to produce specific antibodies that differentiate between these variants let alone ChIP-grade ones. Additionally, since the two H3.3 genes code for the exact same protein, it has been considered that the products of these two genes are complementary and interchangeable. As a result, the initial genome-wide mapping studies relied on epitope-tagging of the H3.3B coding *H3f3b* gene. The first genome-scale enrichment studies used ectopically expressed tagged H3.3^{159,219,220,222,223}. It wasn't until 2010 that a group reported results where the endogenous H3 coding genes were replaced with epitope tagged variants in mouse ESCs²⁰⁵.

A common point in all of these studies was that the tag was placed at the C-terminal end of H3.3B. In the context of the nucleosome, the C-terminal region of H3 places towards the center¹. Placing a tag at this region might therefore constitute a structural disturbance to the physiological reconstitution of nucleosomes hence altering H3.3 positioning and function. It can also produce a barrier for epitope recognition for antibodies and result in a preference for binding to non-nucleosomal targets. As a result, these studies are marked by genome-wide incorporation studies of poor resolution.

We generated novel knock-in/conditional knock-out (KI/cKO) mouse models targeting either one of the two H3.3 coding genes to produce a FLAG-FLAG-HA tagged H3.3 protein. In our models, the tag is placed at the N-terminal region for minimal structural disturbance to the nucleosome. Furthermore, the regions containing the translation start site of these genes are flanked with LoxP sites which serve to obtain a functional knock-out of H3.3 upon Cre mediated recombination. These mouse models constitute valuable tools and put us in a unique position to study the separate functions of H3.3A and H3.3B in vivo.

The purpose of this study was to provide elements to elucidate the precise function of H3.3 in transcription and mitotic progression.

More particularly, this study aimed to:

- 1- Determine the genome-wide distribution of H3.3A and H3.3B separately at nucleosomal resolution in the adult liver of novel mouse models using a high-resolution native ChIP technique coupled to next-generation (NGS) paired-end sequencing and qPCR.
- 2- Determine the effect of loss of H3.3 on the transcriptome by RNA-Seq and RT-qPCR.
- 3- Study the role of H3.3 in mitosis and the effect of loss of H3.3 on mitotic progression based on cell cycle analysis and mitotic and nuclear defect quantification by immunofluorescence.

Chapter 2. Materials and Methods

2.1. Materials

2.1.1. General laboratory chemicals and reagents

Chemicals and reagents used in common laboratory techniques are listed in alphabetical order in Table 2-1 with the providing company and catalog numbers.

Table 2-1. Chemicals, reagents, enzymes and kits used for general laboratory purpose:

Product	Provider	Catalog No.
β -mercaptoethanol	Sigma-Aldrich	AC125470100
Acryl:Bis-acryl 30% solution	Sigma-Aldrich	A3574
AEBSF (4-(2-Aminoethyl) benzene sulfonyl fluoride hydrochloride)	Euromedex	50985
Agarose	Euromedex	D5
Amersham ECL Prime Western Blotting Detection Reagent	GE Healthcare	RPN2232
Aprotinin	Sigma-Aldrich	A4529
Calcium chloride (CaCl ₂)	Sigma-Aldrich	C5080
Chloroform	Carlo Erba	438601
Chymostatin	Sigma-Aldrich	C7268
DNA ladder Mix, GeneRuler	Thermo Fisher Scientific	SM0332
DTT (1,4-dithiothreitol)	Sigma-Aldrich	GE17-1318-01
Ethanol	Acros Orgaics	4146072
EGTA (Ethylene glycol-bis (β -aminoethyl ether)-N,N,N',N'-tetraacetic acid)	Sigma-Aldrich	E4378
EDTA (EthyleneDiamineTetracetic Acid)	Sigma-Aldrich	E5134
Fluorescent mounting medium	Dako	S3023
Formaldehyde 37% (HCHO)	Sigma-Aldrich	F8775
Formalin 4%	Sigma-Aldrich	HT5014
Glycerol	Euromedex	50405
Glycine	Euromedex	26-128-6405-C
HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)	Euromedex	10-110-C

Table 2-1. Chemicals, reagents, enzymes and kits used for general laboratory purposes (continued)

Product	Provider	Catalog No.
Hoechst 33342	Invitrogen	H3570
Isopropanol	Carlo Erba	415154
Leupeptin	Sigma-Aldrich	L2884
Lithium chloride (LiCl)	Sigma-Aldrich	L-9650
Nicotinamide	Sigma-Aldrich	N3376
NP-40	Fluka	74385
NucleoSpin Plasmid Maxi Prep kit	Macherey Nagel	740588
PageRuler Prestained Protein Ladder, 10 to 180 kDa	Thermo Fisher Scientific	26617
PBS 10X	Euromedex	ET330A
PCR Purification Kit (QIAquick)	Qiagen	28106
Pepstatin	Sigma-Aldrich	P5318
Phenol chloroform:isoamyl alcohol	Invitrogen	15593-031
Potassium chloride (KCl)	Carlo Erba	471177
Propidium iodide	Sigma-Aldrich	P4170
Protease inhibitor cocktail EDTA free	Roche	5056489001
Quant-iT PicoGreen dsDNA Assay Kit	Thermo Fisher Scientific	P11496
RNA Mini Kit (PureLink)	Thermo Fisher Scientific	12183018A
SDS 10%	Euromedex	EU0760
Sodium bicarbonate (NaHCO ₃)	Sigma-Aldrich	S5761
Sodium butyrate	Sigma-Aldrich	B5887
Sodium chloride (NaCl)	Carlo Erba	479687
Sodium deoxycholate (NaDOC)	Sigma-Aldrich	D6750
Spermidine	Sigma-Aldrich	S2626
Spermine	Sigma-Aldrich	S3256
Sucrose	Euromedex	200-301-B
TAE 50X	Euromedex	EU0201
Tris Base	Euromedex	200923-A
Triton X-100	Sigma-Aldrich	T9284
Trizol reagent	Life technologies	15596026
Water, DNase/RNase-Free, Distilled	Thermo Fisher Scientific	10977049

2.1.2. Cell culture chemicals and reagents

Table 2-2 lists chemicals, reagents, kits and media used in mammalian and bacterial cell culture experiments conducted in the course of the study.

Table 2-2. Chemicals, reagents, kits and media used in cell culture

Product	Provider	Catalog number
Dimethyl sulfoxide (DMSO)	Euromedex	UD8050-05-A
DMEM, High Glucose, Glutamax, pyruvate	Gibco	31966021
Fetal bovine serum (FBS)	Gibco	10270
Non-essential amino-acids (NEAA)	PAA	M11-003
Normal goat serum	Abcam	ab7481
OptiMEM	Gibco	31985-062
Penicillin/Streptomycin (P/S)	Euromedex	E4413-B
Phosphate buffered saline (PBS)	Gibco	14190-094
Puromycin	PAA	P11-019
Trypsin/EDTA 0.25%	Dominique Dutscher	L0930-100
Lipofectamine 2000 transfection reagent	Invitrogen	11668
Ad-CMV-iCre,	Vector Biolabs	1045
LB Broth (LB)	AthenaES	0103
Ampicillin	Euromedex	EU0400

2.1.3. Primers

Primers used in genotyping, RT-qPCR and ChIP qPCR are presented in Table 2-3, Table 2-4 and Table 2-5. All primers were synthesized by Eurogentec (Belgium) with the Sepop desalting purification method. Primers were received at a concentration of 100µM in dH₂O and were diluted 1:10 in nuclease free water before reaction setup.

Genotyping Primers

Primers used in PCR experiments to verify mouse and cell line genotypes are listed in Table 2-3.

Table 2-3. Primers used in genotyping mouse and cell lines from genomic DNA

Name	Description	Sequence
Lf 3976	Presence of the distal LoxP around <i>H3f3a</i> (FH-H3.3A)	TTTGCAGACGTTTCTAATTCTACT
Lr 3977	Presence of the distal LoxP around <i>H3f3a</i> (FH-H3.3A)	ATATCGGATTCAACTAAAACATAAC
Ed_AK_F	Excision of the floxed exons (H3.3A-KO)	TTCTGTGTTTGTGGCTTCGTT
Ed_AK_R	Excision of the floxed exons (H3.3A-KO)	ATTTAAATGCCCCACCACTGC
Er 3956	Excision of the selection marker / H3.3B Knock-in (H3.3B-KO / FH-H3.3B)	TCAATCTAGGCCTAAGACCAAA
Ef 3955	Presence of the distal LoxP around <i>H3f3b</i> (FH-H3.3B)	TCCTCATTCTACCACATGTTCA
Lf 3953	Excision of the floxed exons (H3.3B-KO)	CTGCCCCGTTCTGCTCGCCGATT
TK139	Presence of the Cre transgene	ATTTGCCTGCATTACCGGTC
TK141	Presence of the Cre transgene	ATCAACGTTTTCTTTTCGGA

RT-qPCR Primers

Primers used in RT-qPCR experiments to verify gene expression levels are listed in Table 2-4.

Table 2-4. Primers used in RT-qPCR for gene expression

Name	Gene	Sequence	Accession No./Location
AO257	Endothelin 1 (<i>Edn1</i>)	CCCACTCTTCTGACCCCTTT	NM_010104.4/ Ex2
AO258	Endothelin 1 (<i>Edn1</i>)	GGCTCTGCACTCCATTCTCA	NM_010104.4/ Ex3
AO239	Growth arrest and DNA-damage-inducible 45 alpha (<i>Gadd45a</i>)	TGCTGCTACTGGAGAACGAC	NM_007836.1/ Ex3
AO240	Growth arrest and DNA-damage-inducible 45 alpha (<i>Gadd45a</i>)	TCCATGTAGCGACTTTCCCG	NM_007836.1/ Ex4
AO301	Growth arrest specific 2 (<i>Gas2</i>)	GCCGAGATTTGGGAGTTGAT	NM_008087/ Ex4

Table 2-4. Primers used in RT-qPCR for gene expression (continued)

Name	Gene	Sequence	Accession No./Location
AO302	Growth arrest specific 2 (<i>Gas2</i>)	GCTTTATCAGACCAGGAGGC	NM_008087/ Ex6
AO11	H3 histone, family 3A (<i>H3f3a</i>)	ACAAAAGCCGCTCGCAAGAG	NM_008210.5/ Ex2
AO12	H3 histone, family 3A (<i>H3f3a</i>)	ATTTCTCGCACCAGACGCTG	NM_008210.5/ Ex3
AO13	H3 histone, family 3B (<i>H3f3b</i>)	TGGCTCTGAGAGAGATCCGTC GTT	NM_008211.3 Ex3
AO14	H3 histone, family 3B (<i>H3f3b</i>)	GGATGTCTTTGGGCATGATGG TGAC	NM_008211.3 Ex4
AO265	Platelet derived growth factor, B (<i>Pdgfb</i>)	GAGTCGGCATGAATCGCTG	NM_011057.3 Ex1
AO266	Platelet derived growth factor, B (<i>Pdgfb</i>)	GCCCCATCTTCATCTACGGA	NM_011057.3 Ex2
MG7	Ribosomal protein S9 (<i>Rps9</i>)	TTGTCGCAAAACCTATGTGAC C	NM_029767.2 Ex2
MG8	Ribosomal protein S9 (<i>Rps9</i>)	GCCGCCTTACGGATCTTGG	NM_029767.2 Ex3
AO253	Seh1-like (<i>Seh1l</i>)	ATGACGGCTGTGTTAGGTTGT	NM_001039088.1 Ex7
AO254	Seh1-like (<i>Seh1l</i>)	TACTCAGCTGTGCTTTCTGCT	NM_001039088.1 Ex8
AO247	SMAD family member 6 (<i>Smad6</i>)	GCCACTGGATCTGTCCGATT	NM_008542.3 Ex3
AO248	SMAD family member 6 (<i>Smad6</i>)	GGTCGTACACCGCATAGAGG	NM_008542.3 Ex4

ChIP-qPCR Primers

Primers used in ChIP-qPCR experiments to verify genomic localization of target proteins are listed in **Table 2-5**.

Table 2-5. Primers used in ChIP-qPCR

Name	Gene	Sequence	Genomic coordinates
AO7	β -actin (<i>Actb</i>)	GCCCCATTCAATGTCTCG GT	Chr5: 143668708-728
AO8	β -actin (<i>Actb</i>)	CCACACAAATAGGGTCC GGG	Chr5: 143668832-851
AO259	Endothelin 1 (<i>Edn1</i>)	AACTAATCTGGTTCCCCG CC	Chr13:42301295-393
AO260	Endothelin 1 (<i>Edn1</i>)	GAGGTGGGGCTGATCATT GT	Chr13:42301399-418
AO287	Growth arrest and DNA- damage-inducible 45 alpha (<i>Gadd45a</i>)	TTTCCGCTCAACTCTGCC TT	Chr6:67037562-543
AO288	Growth arrest and DNA- damage-inducible 45 alpha (<i>Gadd45a</i>)	ACTCTGCACTGCTGCCTC	Chr6:67037379-396
AO300	Growth arrest specific 2 (<i>Gas2</i>)	CCCAAACACTAAGCTAA GACAGA	Chr7:51878844-822
AO3	Interferon α receptor 2 (<i>Ifnar2</i>)	CCCCGATCCGTTAACTCT GG	Chr16:91372937-956
AO4	Interferon α receptor 2 (<i>Ifnar2</i>)	GACAAATGGGCACCTTCG CA	Chr16: 91372997-3016
AO267	Platelet derived growth factor, B (<i>Pdgfb</i>)	AGCTCTGCGCTTTCTGAT CT	Chr15:80014463-482
AO268	Platelet derived growth factor, B (<i>Pdgfb</i>)	GATGGTTCGTCTTCACTC GC	Chr15:80014376-395
AO297	Seh1-like (<i>Seh1l</i>)	TCATCACTGACTGCTGCT TC	Chr18:67774047-066
AO298	Seh1-like (<i>Seh1l</i>)	CTTAGGAATGATGGGGA CGC	Chr18:67774148-167
AO295	SMAD family member 6 (<i>Smad6</i>)	ATATCCTTCTGGGTCTTG CCA	Chr9:64022126-146
AO296	SMAD family member 6 (<i>Smad6</i>)	GCTCAAGGGTGTCAGCA AAA	Chr9:64022208-189

2.1.4. Lentiviral vectors and shRNA

Plasmids used to produce shRNA containing lentiviral particles are detailed in Table 2-6.

Table 2-6. Plasmids used for lentiviral shRNA transduction

Plasmid	Provider	Description
pLP1	Addgene	Lentiviral packaging plasmid containing the HIV-1 gag and pol genes.
pLP2	Addgene	Lentiviral packaging plasmid containing the HIV-1 rev gene.
pCMV-VSV-G	Addgene	Envelope protein for producing lentiviral particles.
pLKO.1 empty	GE Dharmacon	Control shRNA
pLKO.1- TRCN0000012026	GE Dharmacon	shRNA against <i>H3f3a</i> transcript
pLKO.1- TRCN0000092918	GE Dharmacon	shRNA 1 against <i>H3f3b</i> transcript
pLKO.1- TRCN0000092919	GE Dharmacon	shRNA 2 against <i>H3f3b</i> transcript
pLKO.1- TRCN0000092920	GE Dharmacon	shRNA 3 against <i>H3f3b</i> transcript
pLKO.1- TRCN0000092922	GE Dharmacon	shRNA 4 against <i>H3f3b</i> transcript

2.1.5. Enzymes

Enzymes and kits used in the study are detailed in Table 2-7

Table 2-7. Enzymes and kits used in the study

Product	Provider	Catalog number
My Taq Red DNA polymerase	Bioline	BIO-21110
Nuclease S7 Micrococcal nuclease	Roche	10107921001
Proteinase K	Euromedex	09-0912
PureLink DNase Set	ThermoScientific	12185010
Ribonuclease A (RNase A)	Sigma-Aldrich	000000011579681001
RNase OUT	ThermoScientific	10777019
SuperScript IV First-Strand Synthesis System	ThermoScientific	18091050
SYBR qPCR Premix Ex Taq (Tli RNaseH Plus)	Ozyme	TAKRR420W

2.1.6. Antibodies and beads

Antibodies used in Western blot, chromatin immunoprecipitation and immunofluorescence experiments are detailed in Table 2-8 and Table 2-9 lists resins and magnetic beads used in protein precipitation experiments.

Table 2-8. Antibodies used in the study

Antibody	Provider	Catalog Number	Conditions
FLAG M2	Sigma-Aldrich	F3165	WB 1:2000
H2A.Z (rabbit) Serum 98	Home-made		WB 1:1000 ChIP 5µg/ul
H3	Millipore	05-928	WB 1:5 000
H4	Abcam	ab10158	WB 1:5 000
IgG (rabbit) Isotype Control	Abcam	ab171870	ChIP 5µg/ul
Lamin B	Santa-cruz	sc-6217	IF: 1:300
Anti-goat Cy3 conjugated	Jackson	705-165-147	IF: 1:300
Anti-mouse HRP conjugated	Amersham	NA9340V	WB: 1:5000
Anti-rabbit HRP conjugated	Amersham	NA9310V	WB: 1:5000

Table 2-9. Agarose resins and magnetic beads

Resins/magnetic beads	Provider	Catalog Number
Anti-HA affinity matrix	Roche	118150160001
Dynabeads Protein A	Novex (Thermo Fisher Scientific)	10002D
Dynabeads Protein G	Novex (Thermo Fisher Scientific)	10004D

2.1.7. Equipment

Instruments

Table 2-10 lists instruments used in the study as well as their models and providers.

Table 2-10. Instruments used in the study

Instrument	Provider
Accuri C6 flow cytometer	Becton Dickinson
Agarose gel imaging system - Gene flash	Labgene
Auto-Densi Flow IIC Gradient Fractionator	Buchler Instruments

Table 2-10. Instruments used in the study (continued)

Instrument	Provider
Axio Imager Z1 (ApoTome)	Zeiss
Axiovert 135 with Infinity 2 camera	Axiovert and Lumenera
Bioruptor waterbath sonicator	Diagenode
Cell culture hood - HS12	Thermo Fisher Scientific
Cell culture incubators - MCO-20AI	Sanyo
Centrifuges 5810R, 5702, 5415R, 5424	Eppendorf
Chemical fume hood Type 1200	Hotter mann
Chemidoc	Bio-rad
Film developer - Hyper processor	Amersham
Mastercycler – Nexus gradient	Eppendorf
Mini PROTEAN Tetra Cell and Mini Transblot Cell	Bio-Rad
Nanodrop ND-2000	Nanodrop
Olympus BX41	Olympus
Olympus CXX31	Olympus
PowerPac Basic Power Supply	Bio-rad
Roche LightCycler 480 - Real-Time thermal cycler	Roche Life Science
Roto-shaker (MHR23)	HLC Biotech
Spectrophotometer	Eppendorf
Sub-Cell GT Cell for agarose gel electrophoresis	Bio-rad
Ultracentrifuge (CO-LE80K) with Swing-bucket rotors SW28 and SW41Ti and their adapters	Beckman Coulter
VICTOR Multilabel Plate Reader	Perkin-Elmer
Water bath	

General laboratory equipment

Table 2-11 lists general laboratory equipment and their providers used in the study.

Table 2-11. General laboratory equipment used in the study

Equipment	Provider
Dissection material: curved scissors, curved and straight forceps, razor blades	Dominique Dutscher
Needles, Terumo 21G, 23G, and 26G	Dominique Dutscher
Cell culture plates	Falcon

Table 2-11. General laboratory equipment used in the study (continued)

Equipment	Provider
Cryovials	Corning
Graduated pipettes (5ml: 357443, 10ml: 357451 and 25ml: 357535)	Falcon
0.2, 1.5, 2.0ml micro centrifuge tubes	Eppendorf
15 and 50ml conical tubes	Falcon
Ultracentrifuge tubes (11ml: 347357 and 33ml: 331372)	Beckman Coulter
Neubauer improved, cell counting chamber	Neubauer
Glass tissue grinder sets 2ml and 15ml	Dounce
Gauze	Dominique Dutscher
Nitrocellulose membrane, HyBond, 0.2µm	Amersham
X-ray films Blue Devil (30-101)	Genesee scientific
Centrifugal filter, Amicon Ultra, 0.5ml and 15ml	Millipore
Dialysis tubing, Regenerated cellulose, 6k MWCO	Spectrapor
Cellulose acetate filters for syringes (0.45µm)	Dominique Dutscher
Coverslips 14mm diameter (002130)	Waldemar Knittel
Glass slides (8037/1)	Waldemar Knittel
FluoroNunc, flat bottom black 96-well plates for fluorescence	Dominique Dutscher
LightCycler 480, 96-well plates, white, with sealing foils	Roche Life Sciences

2.1.8. Software

The software used in this study are listed in Table 2-12.

Table 2-12. Software used in the study

Software	Developer
BLAST	NCBI, Bethesda, Maryland, USA ²⁷⁵
Primer BLAST	NCBI, Bethesda, Maryland, USA ²⁷⁶
Genome Browser	UCSC, Santa Cruz, California, USA ¹⁸⁰
GraphPad Prism 6	GraphPad Software, La Jolla, California, USA
Modfit LT	Verity Software House, Topsham, Maine, USA
FCS Express 6	De Novo Software, Glendale, California, USA
Pymol Version 1.8	The PyMOL Molecular Graphics System, Schrödinger, LLC ²¹

2.2. Solutions and Media

2.2.1. Cell culture solutions and media

Mammalian and bacterial cell lines used in the study and their relevant growth media are listed in Table 2-13.

Table 2-13. Cell lines and their relevant culture media

Cell line	Medium
Primary MEFs	DMEM, high glucose, GlutaMAX Supplement, pyruvate, 10% FBS, 1% penicillin-streptomycin, 1%NEAA.
Immortalized MEFs	DMEM, high glucose, GlutaMAX Supplement, pyruvate, 10% FBS, 1% penicillin-streptomycin.
293 T	
293 Phoenix Eco	
Transformed DH5- α Escherichia coli	LB Broth supplemented with 100 μ g/ml ampicillin or LB-Agar supplemented with 100 μ g/ml ampicillin

2.2.2. Genomic DNA Extraction and analysis

Buffer and solution compositions used in genomic DNA extraction and analysis are listed in Table 2-14

Table 2-14. Buffers used in genomic DNA extraction and analysis

Buffer	Composition
Tail Buffer	50mM Tris-HCl pH 8, 100mM EDTA pH 8, 100mM NaCl, 1% SDS
6X DNA loading dye	40% Sucrose, 0.25% Xylene cyanol (or bromophenol blue)

2.2.3. Western blot solutions and buffers

Buffer and solution compositions used in Western blot and protein and analysis are listed in Table 2-15 and Table 2-16

Table 2-15. Buffers and solution used in western blot

Buffer	Composition
2X Laemmli lysis buffer	4% SDS, 10% 2- β -mercaptoethanol, 20% Glycerol, 0.004% Bromophenol blue, 0.125 M Tris-HCl pH 6.8
10X SDS Running Buffer	30g Tris-Base, 144g Glycine, 100ml SDS, qs 1L dH ₂ O Dilute 10x before use
1X SDS Transfer Buffer	1X SDS Running Buffer, 10% Ethanol
Blocking solution	5% (w/v) milk in PBS
Ponceau S	0.1% (w/v) Ponceau S and 5% (v/v) acetic acid in ddH ₂ O
WB Wash Buffer	350mM NaCl in PBS

Table 2-16. Polyacrylamide Tris Gel preparation

For 5ml of:	5% Stacking	10% Resolving	12% Resolving	15% Resolving	18% Resolving
Acrylamide 30 %	830 μ l	1,7ml	2ml	2,5ml	3ml
Tris-HCl 1,5M pH 8.8		1,3ml	1,3ml	1,3ml	1,3ml
Tris-HCl 1M pH 6.8	630 μ l				
SDS 10%	50 μ l	50 μ l	50 μ l	50 μ l	50 μ l
APS	50 μ l	50 μ l	50 μ l	50 μ l	50 μ l
TEMED	5 μ l	5 μ l	5 μ l	5 μ l	5 μ l
H ₂ O	3,4ml	1,9ml	1,6ml	1,1ml	0,6ml

2.2.4. Nuclei isolation buffers

Nuclei isolation from livers

Buffer and solution compositions used in isolation of nuclei from liver tissues are detailed in **Table 2-17**.

Table 2-17. Buffer compositions used in nuclei isolation from liver tissue

Buffer	Composition
2.2M Sucrose Buffer	2.2M Sucrose, 10mM HEPES KOH pH 7.6, 15mM KCl, 2mM EDTA, 125mM Glycine*, 0.15mM Spermine [†] , 0.5mM Spermidine [†] , 0.5M DTT [†] , 2mM NaBu [†] , 5mM Nicotinamide [†] , 10µg/ml of AEBSF [†] , Pepstatin [†] , Aprotinin [†] , Leupeptin [†] , Chymostatin [†]
2.05M Sucrose Buffer (Cushion)	2.05M Sucrose, 10% glycerol, 10mM HEPES KOH pH 7.6, 15mM KCl, 2mM EDTA, 150mM Glycine [†] , 0.15mM Spermine [†] , 0.5mM Spermidine [†] , 0.5M DTT [†] , 2mM NaBu [†] , 5mM Nicotinamide [†] , 10µg/ml of AEBSF [†] , Pepstatin [†] , Aprotinin [†] , Leupeptin [†] , Chymostatin [†]
Nuclei Buffer	10mM HEPES KOH pH 7.6, 100mM KCl, 0.1mM EDTA, 10% Glycerol, 0.15mM Spermine [†] , 0.5mM Spermidine [†] , 0.5mM DTT, 2mM NaBu [†] , 5mM Nicotinamide [†] , 10µg/ml of AEBSF [†] , Pepstatin [†] , Aprotinin [†] , Leupeptin [†] , Chymostatin [†]

Nuclei isolation from cells

Buffer and solution compositions used in isolation of nuclei from adherent cell cultures are detailed in Table 2-18.

Table 2-18. Buffer compositions used in nuclei isolation from cells

Buffer	Composition
Buffer A	150mM Tris pH 7.5 pH, 150mM NaCl, 600M KCl, adjust pH to 7.6
Buffer B	10% Buffer A, 340mM Sucrose, 2mM EDTA, 0.5mM EGTA, 1mM DTT, 0.2mM Spermine [†] , 0.65mM Spermidine [†] , 0.5mM PMSF [†] , 2mM NaBu [†] , 5mM Nicotinamide [†] , 1X Roche Protease Inhibitor Cocktail [†] , adjust pH to 7.4
Lysis Buffer	10% Buffer A, 340mM Sucrose, 2mM EDTA, 0.5mM EGTA, 0.4% NP40, 1mM DTT, 0.2mM Spermine [†] , 0.65mM Spermidine [†] , 0.5mM PMSF [†] , 2mM NaBu [†] , 5mM Nicotinamide [†] , 1X Roche Protease Inhibitor Cocktail [†] , adjust pH to 7.4
Buffer D	10% Buffer A, 340mM Sucrose, 1mM DTT, 0.2mM Spermine [†] , 0.65mM Spermidine [†] , 0.5mM PMSF [†] , 2mM NaBu [†] , 5mM Nicotinamide [†] , 1X Roche Protease Inhibitor Cocktail [†] , adjust pH to 7.4

* Only add if using for cross-linked ChIP

[†] Add extemporaneously

2.2.5. Chromatin preparation and immunoprecipitation buffers

Buffer and solution compositions used in native and crosslinked ChIP are detailed in Table 2-19

Table 2-19. Buffer compositions used in chromatin preparation and ChIP

Buffer	Composition
Digestion Buffer	5mM CaCl ₂ , 10mM Tris pH 7.5
Stop Buffer	1mg/ml proteinase K, 10mM EDTA pH 8.0
TNE Buffer	10mM Tris pH 7.5, 600mM NaCl, 5mM EDTA
Sucrose Gradient	10% Sucrose, 600mM NaCl, 5mM EDTA, 10mM Tris pH 7.5
Dialysis Buffer	10mM Tris pH 7.5, 50mM NaCl, 1mM EDTA
N-ChIP Buffer	10mM Tris, pH 7.5, 80mM NaCl, 1mM EDTA, 0.5% Triton X-100
N-Wash Buffer I	10mM Tris, pH 7.5, 80mM NaCl, 1mM EDTA, 0.5% Triton X-100
N-Wash Buffer II	10mM Tris, pH 7.5, 100mM NaCl, 1mM EDTA, 0.5% Triton X-100
N-Wash Buffer III	10mM Tris, pH 7.5, 150mM NaCl, 1mM EDTA, 0.5% Triton X-100
Nuclei Wash Buffer	20mM Tris, pH 7.5, 150mM NaCl, 2mM EDTA
Nuclei Lysis Buffer (2X)	20mM Tris, pH 7.5, 150mM NaCl, 2mM EDTA, 2% SDS
X-ChIP Buffer	20mM Tris, pH 7.5, 150mM NaCl, 2mM EDTA, 1% Triton X-100
X-Wash Buffer I	20mM Tris, pH 7.5, 150mM NaCl, 2mM EDTA, 1% Triton X-100, 0.1% SDS.
X-Wash Buffer II	20mM Tris-HCl, pH 7.5, 2mM EDTA, 500mM NaCl, 1% Triton X-100, 0.1% SDS.
X-Wash Buffer III	10mM Tris-HCl, pH 7.5, 1mM EDTA, 0.25M LiCl, 1% NP-40, 1% deoxycholate (Na-DOC).
ChIP Elution Buffer	100mM NaHCO ₃ , 1% SDS.

2.3. Mouse models

The FH-H3.3A (K481) and FH-H3.3B (K480) mutant mouse lines were established at the Phenomin-iCS (Phenomin-Institut Clinique de la Souris, Illkirch, France; <http://www.ics-mci.fr/en/>). For the construction of the targeting vector a 0.5 kb fragment encompassing exon 2 of either *H3f3a* or *H3f3b* was amplified by PCR (from 129S2/SvPas ES cells genomic DNA) and subcloned in an iCS proprietary vector. This iCS vector contains a LoxP site as well as a floxed and flipped Neomycin resistance cassette. A DNA element encoding the FLAG-FLAG-HA amino acids was inserted in frame with the N-terminus of H3.3B. A 4.5 kb fragment (corresponding to the 5' homology arm) and 3.5 kb fragment (corresponding to the 3' homology arms) were amplified by PCR and subcloned in step1 plasmid to generate the final targeting construct. The linearized construct was electroporated in 129S2/SvPas mouse embryonic stem (ES) cells. After selection, targeted clones were identified by PCR using external primers and further confirmed by Southern blot with 5' and 3' external probes. 3 positive ES clones for *H3f3a* and 2 positive ES clones for *H3f3b* were injected into C57BL/6N blastocysts, and male chimaeras derived gave germline transmission.

Complete H3.3A Knock-out (H3.3A-KO) and H3.3B Knock-out (H3.3B-KO) mice were generated in the Plateforme de Haute Technologie Animale (PHTA, Grenoble, France) mouse facility by breeding FH-H3.3A and FH-H3.3B mice with hActin-Cre expressing mice.

2.4. Methods

2.4.1. General maintenance and handling of test subjects

Mice were housed in the Plateforme de Haute Technologie Animale (PHTA, Grenoble, France) mouse facility (agreement number C 38 516 10001, registered protocol n° 321 at ethical committee C2EA-12).

2.4.2. Mouse embryonic fibroblast (MEF) isolation

Timed breeding was setup, vaginal plug date was recorded and the pregnant female was moved to a separate cage to ensure proper age of embryos at harvest. Female was euthanized at 13.5 days post fertilization (dpf). Both uterine horns containing embryos were removed and placed in a 10cm tissue culture dish containing 20ml of sterile PBS. The dish containing

the uterine horns with the embryos was transferred to the cell culture hood and the uterine horn were re-transferred to a new dish containing sterile PBS. This was repeated until most blood was removed (2-3 times). Using sharp forceps, the uterine wall was torn to release the first embryo. The embryo was moved to a new dish containing sterile PBS. With clean forceps, the head of the embryo was removed and transferred into a new 1.5ml microcentrifuge tube for genotyping. Using clean curved forceps, all internal organs were removed and discarded. The remainder of the torso was transferred into a 6-well plate containing complete MEF culture medium. With sterile 23G needles and/or razor blades, tissues were minced and roughly homogenized by pipetting 10-15x using a 5ml graduated pipette. After 24h of culture, medium containing tissue pieces was transferred to a new 6-well plate to obtain more MEFs. Cells were grown till pre-confluence and passaged for immortalization with the serial passage (3T3) method ²⁷⁷ or cryopreserved.

2.4.3. General maintenance and handling of cell lines

All cell lines were handled using aseptic technique in a dedicated cell culture room under a vertical laminar flow hood.

MEFs and 293T packaging cell lines are adherent cell lines. For subculturing from a 10cm tissue culture dish; cell media was removed, cells were washed with 10ml PBS, 1ml of 0.25% Trypsin-EDTA solution was added and cells were placed in a cell culture incubator at 37°C, 5% CO₂. After 2-3 minutes of incubation to allow cell detachment, trypsin was quenched by adding 9ml of complete cell culture medium to the cells. Cells were homogenized by pipetting up and down and either split 1:5 or 1:10 in to a new 10cm tissue culture dish for culture maintenance or counted with a Neubauer Improved cell counting chamber for infection experiments.

For cryopreservation, 70-80% confluent cells in growth phase were detached by trypsinization as described above then transferred into a 50ml conical tube and centrifuged at 500 x g for 3-5 minutes. Cells were slowly resuspended in 1ml of freshly prepared freezing medium consisting of 10% dimethyl sulfoxide (DMSO) in complete cell culture medium and transferred into sterile cryovials. Cryovials were kept in cryo-boxes containing isopropanol at -80°C for minimum 2h before being transferred into liquid nitrogen tanks.

For thawing the cells, cryovials were rapidly placed in a water bath heated at 37°C. As soon as the cells were thawed they were transferred into a new 10cm plate containing at least

10ml of pre-warmed culture medium and placed in the incubator. Cell medium was changed after cells had adhered to remove traces of DMSO.

2.4.4. Transformation of bacteria

Chemically competent DH5- α Escherichia coli were thawed on ice and 50 μ l mixed with 2 μ l of plasmid. The mixture was incubated on ice for 20 minutes, then placed at 42°C for 1 minute and immediately transferred on ice for 1-2 minutes. 450 μ l of LB was added and mixed by gentle tapping. Cells were incubated at 37°C for 1 hour, plated on LB agar plates containing 100 μ g/ml ampicillin and incubated at 37°C overnight. Single colonies were picked and allowed to grow in 200ml of LB at 37°C o/n.

2.4.5. Plasmid DNA isolation

Plasmid DNA was isolated from bacterial culture using NucleoSpin Plasmid kit (Macherey-Nagel) per manufacturer's instructions.

2.4.6. Genomic DNA isolation

Samples (tissues, cells) were placed in 1.5ml microcentrifuge tubes and digested in 750 μ l Tail buffer supplemented with 40 μ l of Proteinase K (10mg/ml) at +55°C overnight (o/n) for tissues or at +37°C for several hours for mammalian cells. Dissociation of larger pieces was facilitated by placing the tubes on a roto-shaker. After digestion, 250 μ l (~1/3 of total volume) of saturated NaCl (>6M) was added, and samples were placed in a roto-shaker for 5 minutes. Tubes were spun at maximum speed (16 873 $\times$ g) for 10 minutes. 750 μ l of supernatant was transferred to a new tube and 500 μ l of isopropanol (0.6 x Volume) was added and mixed by inverting the tube 5-10 times. DNA was precipitated by centrifugation at maximum speed for 10 minutes. The supernatant was discarded and the pellet was washed once in 750 μ l of 75% ethanol. Pellets were left to air dry at room temperature for approximately 30 minutes and DNA was resuspended in an adequate volume of DNase/RNase free water. DNA concentrations were quantified using the Nanodrop ND 2000 microvolume spectrophotometer. Concentrations were adjusted to 100ng/ μ l and DNA samples were stored at -20°C for long term or at +4°C for short term storage.

2.4.7. Genotyping by PCR

Genotyping was carried out on genomic DNA. PCR reaction mix was prepared according to Table 2-20 per manufacturer's instructions.

Table 2-20. PCR reaction setup volumes for genotyping.

Reagent	Volume
5X My Taq Red Buffer	5 µl
Primer 1 10mM	0.4 µl
Primer 2 10mM	0.4 µl
MyTaq	0.4 µl
H2O qs 23µl	16.8 µl
DNA 100-200 ng/µl	2 µl
Total	25 µl

Primer pairs and cycling conditions are detailed in Table 2-3, Table 2-21 and Table 2-22.

Table 2-21. Primer pairs used in genotyping of FH-H3.3A, FH-H3.3B and Actin-Cre mouse and cell lines. FL refers to the floxed allele, WT refers to the wild-type allele, L- refers to the deleted allele and Tg refers to the Cre recombinase transgene.

Mouse Line	Primers	Description	PCR program	Expected band size
FH-H3.3A	Lf 3976 Lr 3977	Excision of the selection marker: Knock-in	Cre40	FL: 361 kb WT: 291 kb
	Ed_AK_F Ed_AK_R	Excision of the floxed exon: Knock-out	Long	FL: 1655 kb WT: 1300 kb L- : 489 kb
FH-H3.3B	Ef 3955 Er 3956	Excision of the selection marker: Knock-in	Cre40	FL: 303 kb WT: 177 kb
	Lf 3953 Er 3956	Excision of the floxed exon: Knock-out	Long	FL: 2270 kb WT: 1908 kb L- : 493 kb
Cre	TK139 TK141	Cre Recombinase presence	Cre40	Tg: 349 kb

Table 2-22. PCR cycling conditions used in genotyping protocols

Cre40		
Temp	Time	#Cycles
94°C	3min	1
94°C	1min	2
58°C	1min	
72°C	1min	38
94°C	30s	
58°C	30s	
72°C	30s	
72°C	3min	1
4°C	∞	

Long		
Temp	Time	#Cycles
95°C	3min	1
95°C	30sec	32
60°C	1min	
72°C	2min	
72°C	10min	1
4°C	∞	

2.4.8. Agarose Gel Electrophoresis

DNA fragments were separated by agarose gel electrophoresis. The Biorad Sub-cell GT system was used for electrophoresis. Agarose gels at 1.5% (w/v) were prepared by dissolving 1.5g of agarose in 100ml TAE buffer in a microwave-oven. After brief cooling, ethidium bromide was added at a final concentration of 10µg/mL and the agarose gel was cast into gel trays with the appropriate comb and kept under a chemical fume hood until cooled and solidified completely. DNA samples mixed 1:5 with 6X DNA loading dye were loaded into the gel. The electrophoresis was done at 120 V until the bands had separated properly (approximately 40 minutes). At the end of the migration, DNA was visualized under UV light and gel pictures were taken with gel imaging system. DNA size was determined with a 1kb mix DNA ladder.

2.4.9. Lentivirus production for shRNA based knock-down in MEFs

All virus production was confined to a virology specific cell culture room. 293T cells were grown to 80-90% confluency in 10cm plates. Cells were washed once with PBS and cell culture medium was replaced with 12ml OPTIMEM. Lipofectamine 2000 (Invitrogen) transfection reagent was used following manufacturer's protocol. 5µg of pLP1, 2.5µg of pLP2, 3µg of pCMV-VSV-G and 10µg pLKO.1-shRNA vectors were mixed and added to 1.5ml OPTIMEM. 50µl of lipofectamine 2000 transfection reagent was diluted in 1.5ml OPTIMEM. Lipofectamine and DNA mixtures were mixed and complexes were allowed to form for 20 minutes at room temperature. Complexes were added dropwise on cells and plates were placed in the incubator. After 8-16 hours, transfection medium was replaced with 10ml fresh culture medium. Viral supernatant (10ml) was collected at 48 hours post transfection and filtered through 0.45µm cellulose acetate filters. Target MEFs, at ~50% confluence, were infected with the viral supernatant for 24 hours and cells were selected with 3µg/ml puromycin. Cells were analyzed by RT-qPCR for knock-down efficiency after 3 days.

2.4.10. Adenovirus infection for transient Cre expression in MEFs

MEFs were infected with adenovirus expressing Cre recombinase (Ad-CMV-iCre, Vector Biolabs, Philadelphia, PA) to disrupt endogenous H3.3 coding genes. Target cells were plated at 30 000 cells/well in a 6-well plate. The next day, cell media was removed and cells were washed once with PBS. Virus was diluted in serum-free DMEM at a multiplicity of

infection (MOI) of 500 and added on the cells. The next day, infection medium was replaced with fresh complete medium and cells were analyzed by RT-qPCR and Western blot for knock-out efficiency after 3 days.

2.4.11. RNA isolation

All RNA was isolated in an RNase free environment, using DNase/RNase free labware and solutions.

RNA was isolated using Trizol reagent (Invitrogen) coupled to Purelink RNA Mini kit (Ambion) with on-column DNase digestion following manufacturer's instructions.

For sample homogenization, mouse tissues (50-100mg) were collected in 1ml Trizol Reagent at room temperature (RT) and homogenized using a 2ml glass Dounce tissue grinder set using both pestles for at least 20 strokes each. Following homogenization, samples were incubated at RT for 5 minutes to allow for complete dissociation of nucleoprotein complexes, then they were centrifuged at 12 000 x g for 10 minutes at 4°C. The resulting pellet contains ECM, polysaccharides, and high molecular weight DNA, while the supernatant contains the RNA. In high fat content samples like liver tissue, a layer of fat collects above the supernatant. The fatty layer was discarded and all of the cleared supernatant was transferred into a new tube.

For adherent cells grown in a 6-well plate, the culture medium was discarded and 1ml Trizol Reagent was added directly on the cells. Following 5 minutes of incubation at RT, samples were homogenized by pipetting up and down and transferred to new microcentrifuge tubes.

After column purification, RNA was resuspended in an adequate amount of RNase-free water (20–50µl for cells, 100-200µl for tissues) by passing the solution up and down several times through a pipette tip. Samples were incubated in a heat block at 60°C for 10 minutes. The concentration, 260/280 and 230/260 ratios were measured using the NanoDrop ND 2000. 260/280 ratios of all RNA used were 1.9-2.2 and 230/260 ratios were above 1.8. All RNAs were adjusted to the same concentration and aliquoted before storage at -80°C or downstream applications. RNA integrity was verified by running samples in a 1% agarose gel at 50V or with the 2100 Bioanalyzer (Agilent).

2.4.12. cDNA preparation

SuperScript IV First-Strand Synthesis System (Invitrogen) was used for cDNA synthesis. Reverse transcription reaction was setup in 20µl following manufacturer's instructions using 0.5-1µg of RNA and random hexamers. The cDNA samples were diluted 4X with water prior to downstream applications.

2.4.13. Quantitative PCR (qPCR) and analysis

Takara SYBR qPCR Premix Ex Taq (Tli RNaseH Plus) and LightCycler 480 (Roche) real-time system were used for qPCR. Reactions were set up in duplicates in 20µl using 2µl of DNA, 2.5µl of primer mix at 10µmol/l, 10µl of premix and 5.5µl of RNase/ DNase free water. Cycling conditions were as follows: 3min at 95°C and 40 times 10s 95°C, 30s 60°C.

Primers were designed using the NCBI Primer Blast online primer design tool ²⁷⁶. The mRNA expression primers (Table 2-4) were designed so that the amplicon included an exon-exon junction. The primers used in ChIP-qPCR were designed against specific genomic regions identified in the mm9 assembly of the mouse genome ²⁷⁸ (Table 2-5).

2.4.14. Nuclei isolation from liver tissue

Livers were isolated from mice of relevant genotypes and either used immediately or snap-frozen in liquid nitrogen for storage at -80°C. Livers were allowed to thaw on ice and minced into small pieces with sterile blades in a Petri dish. Volume was completed to 5ml with PBS and samples were homogenized roughly in a 15ml glass Dounce tissue grinder.

For use in cross-linked ChIP, livers were crosslinked by adding 135µl of 37% formaldehyde to obtain a final concentration of 1% and incubated at RT for 7 minutes with rotation. Formaldehyde was quenched by adding 556µl of 1.25M glycine in PBS.

Livers were further homogenized with pestle A after adding 7-10ml of 2.2M sucrose buffer. Homogenate was filtered through 2 layers of gauze into a 50ml conical to remove hard to break membranes. The glass recipient was washed two times with 5ml 2.2M sucrose buffer. 9ml of 2.05 sucrose cushion buffer was added into a 38.5ml thin wall polypropylene ultracentrifugation tube for use in a Beckman Coulter SW28 swing-bucket rotor. Homogenate was layered on top of the cushion buffer by pouring it slowly in to the tube. Samples were centrifuge at 100 000 x g at +4°C during 1h. After centrifugation, the crust formed on the top of the tube was removed with the help of a pipette tip and the supernatant

was discarded by inverting the tube. The pellet containing nuclei was resuspended in 1-2ml Nuclei Buffer using a cut 1ml tip and transferred into 15ml conical tubes. Nuclei were centrifuged at 1500 x g for 2min at +4°C and the nuclear pellet was resuspended in Nuclei Buffer containing 10% glycerol at 1.2ml/liver. Liver nuclei were aliquoted in 150µl volumes and snap-frozen in liquid nitrogen for storage at -80°C.

2.4.15. Nuclei isolation from cell lines

Cells were grown to preconfluence (~80-90%) in a 15cm dish. For ChIP experiments, 4X15 cm plates were used and cells were collected by scraping in PBS-10µg/ml AEBSF. For other uses cells were detached with Trypsin-EDTA 0.25%.

For use in cross-linking ChIP, cells were washed once with PBS at room temperature and crosslinked by incubating with 10ml of PBS-1% formaldehyde solution (HCHO) for 7 minutes at room temperature. After the incubation period, formaldehyde was quenched by adding 1.25M glycine at a final concentration of 125mM. Cells were washed 2X with cold PBS-125mM glycine. All PBS was removed, cells were scraped in 10ml cold PBS-125mM glycine containing 10µg/ml AEBSF, 5mM nicotinamide and 2mM NaBu and transferred into a 50ml conical tube. Plates were washed with the same buffer to collect remaining cells.

Cells were centrifuged for 5 minutes at 500 x g. Supernatant was discarded and samples were snap-frozen in liquid nitrogen and stored at -80°C or used immediately for nuclei isolation.

Cell pellets were resuspended in 5ml of PBS-10µg/ml AEBSF and centrifuged 5min at 500 x g at 4°C. The supernatant was discarded and the pellet was resuspended in 1ml Buffer B. An equal volume of Lysis Buffer was added and mixed gently by inverting the tube. Tubes were incubated on ice for no longer than 5minutes and were inverted every minute to allow for homogenization. Samples were centrifuged 5min at 500 x g at 4°C. Supernatant containing cytosolic extracts were discarded using a pipette. The nuclear pellet was washed with 2ml Buffer D and resuspended in buffer D with 10% glycerol for storage at -80°C.

2.4.16. Nuclei quantification and lysis

Two aliquots of 10µl were drawn from isolated native nuclei and diluted in 85µl of digestion buffer. 25U of micrococcal nuclease S7 (MNase) was added and nuclei were

digested for 5-10 minutes at 37°C. The digestion was diluted 10X in 2M NaCl and the optical density at 260nm was read with a spectrophotometer. The double-strand DNA (ds DNA) quantity was calculated using the formula: $1 \text{ OD}_{260} = 50 \mu\text{g/ml}$ of ds DNA.

For nuclear lysis, equal quantities of nuclei were drawn and pelleted 5 minutes at 500 x g at 4°C. 2X Laemmli lysis buffer was added to nuclear pellets and samples were sonicated for 3 cycles 30s ON/ 30s OFF at high amplitude in the Diagenode Bioruptor waterbath sonicator.

2.4.17. Whole cell protein extraction.

Equal number of cells were plated in 6-well plates one day prior to harvest. Cells were washed once with PBS and 200 μ l of 2X Laemmli lysis buffer was added for whole cell lysis. Lysates were collected in microcentrifuge tubes by scraping. To shear DNA, lysates were sonicated for 3 cycles 30s ON/ 30s OFF at high amplitude in the Diagenode Bioruptor waterbath sonicator.

2.4.18. Western blot

Denaturing polyacrylamide gels were prepared according to Table 2-16 and were cast in the Bio-Rad Mini PROTEAN Tetra Cell system as instructed by the manufacturer. For analysis of small proteins such as histones 15 or 18% gels were used while bigger proteins were separated in 12% gels. Gels were placed in Western blot tanks filled with 1X SDS Running buffer. Samples were loaded in the gel and electrophoresed at 20mA. Once protein separation completed, proteins were transferred to a 0.2 μ m nitrocellulose membrane by wet-transfer in 1X SDS Transfer buffer using the Bio-Rad Mini Trans-blot electrophoretic transfer cell following manufacturer's instructions. Transfer was performed for 75 minutes at 100V. Transfer efficiency was assessed by Ponceau red staining. Membrane was incubated in 5% milk-PBS blocking solution for 1 hour at RT. Primary antibodies were diluted in 5%BSA-PBS according to empirically determined concentrations (Table 2-8) and incubated with the membrane for 1 hour at RT or overnight at 4°C. For the antibodies used in this study, incubation times did not significantly affect the strength of signal. Membranes were washed 2 x 5 minutes with 350mM NaCl-PBS wash buffer and 1 x 5 minutes with PBS to remove residual salt. Membrane was then incubated with HRP conjugated secondary antibody diluted 1:5000 in blocking solution for 1h at RT. Membranes were washed 2 x 5 minutes with 350mM NaCl-PBS wash buffer and 1 x 5 minutes with PBS to

remove residual salt. Chemiluminescent signal detection was performed with ECL Prime chemiluminescence detection kit (Amersham) and Biorad ChemiDoc Imaging system for semi-quantitative analysis. Alternatively, X-ray films were exposed to the membrane and developed in an X-ray film developer.

2.4.19. Native chromatin immunoprecipitation (N-ChIP)

Chromatin preparation – Enzymatic digestion

Nuclei isolated using the native nuclei isolation protocols detailed in 2.4.14 for livers and 2.4.15 for cells. Nuclei were quantified based on protocol in 2.4.16. Digestion efficiency is dependent on type of sample, nuclease quality, temperature and time. Therefore, a digestion assay prior to chromatin preparation is necessary to determine optimal digestion conditions. For this purpose, micrococcal nuclease S7 (MNase, 5 U/ μ l) was diluted 20X in digestion buffer supplemented with protease inhibitors (complete digestion buffer). 100 μ g of nuclei were transferred into a 1.5ml microcentrifuge tube and pelleted for 5 minutes at 500 x g at +4°C. Supernatant was discarded and nuclei were resuspended in 100 μ l of complete digestion buffer. 10 μ l of diluted MNase (2.5 U) was added and digestion started by placing the tube in a waterbath heated at 37°C. 6 tubes, each containing 100 μ l of stop solution (1mg/ml Proteinase K, 10mM EDTA pH 8.0) were prepared. 15 μ l of digested sample was added to the tubes with stop solution at 5', 10', 15', 20', 30', 40' and placed on ice immediately. Once all time points collected, DNA was purified with phenol chloroform:isoamyl alcohol (PCIA) phase separation. Briefly, an equal volume of PCIA was added on digested samples (115 μ l) quick vortexed and centrifuged for 15 minutes at maximum speed at +4°C. 20 μ l of the upper aqueous phase containing DNA was transferred to a new tube with 5 μ l of 75% glycerol and analyzed on a 1.5% agarose gel. An adequate digestion time was chosen so that the chromatin would contain a majority of mono-nucleosomes without too many over-digested fragments.

A quantity of 1mg of nuclei was pelleted and resuspended in 900 μ l digestion buffer, 100 μ l of 0.25 U/ μ l MNase was added, reaction was split into 10 tubes of equal volume and placed in a waterbath heated at 37°C for the empirically determined optimal digestion time (13 minutes in this example). Samples were mixed by gentle tapping every 2 minutes to homogenize nuclei that might have pelleted. After digestion, samples were centrifuged for 3 minutes at 800 x g and supernatant containing MNase was discarded. The digested nuclei

pellets were pooled in 100µl of TNE buffer. Nuclear extracts were centrifuged for 2 minutes at 10 000 x g at +4°C. Supernatant containing solubilized nucleosomes were collected and pellet resuspended in 100µl of TNE buffer. These last 2 steps were repeated 2 times and all three supernatants were pooled. DNA concentration was calculated by diluting an aliquot of sample in 2M NaCl and measuring the OD at 260nm on a spectrophotometer and digestion profile was controlled by running the sample on a 1.5% agarose gel.

Mono-nucleosome preparation

For high resolution genome-wide mapping of H3.3, mono-nucleosomes were isolated through separation of the digestion product on a 5-10% sucrose gradient. 10% sucrose gradient solution supplemented with protease inhibitors was prepared in SW41 Beckman ultracentrifuge tubes and frozen o/n at -80°C. Tubes were thawed at 4°C to allow formation of a 5-20% linear density sucrose gradient. Samples were carefully loaded on gradients drop by drop and centrifuged using a SW41 swing-bucket rotor (Beckman Coulter) for 20h at 100 000 x g.

Fractions of 500µl were collected using a gradient fractionator pump. 10µl aliquots of every other fraction were transferred to a new tube containing 10µl of dH₂O and 10µl of Proteinase K (10mg/ml) and digested for 15 minutes at 37°C. A mix of 10µl of 75% glycerol and 5µl of 1% SDS was added and samples were loaded in a 1.5% agarose gel. Fractions containing a majority of mono-nucleosomes were determined and pooled.

Pooled fractions were transferred into 6k MWCO regenerated cellulose dialysis membranes (Spectrapor) and dialyzed o/n at +4°C in 4.5 L of dialysis buffer. Nucleosomes were quantified by reading OD at 260 nm and concentrated using centrifugal filters (Millipore) if needed.

Native chromatin immunoprecipitation

For chromatin immunoprecipitation, 100µg of input chromatin was diluted in native ChIP buffer containing 1X protease inhibitor cocktail (Roche) and mixed with 20µl of anti-HA affinity matrix (Roche) previously washed once with native ChIP Buffer. The mixture was incubated o/n at +4°C on a rotator to allow for antigen binding. Chromatin isolated from wild-type MEFs with no HA epitope was used as negative control. Beads bound to immunoprecipitated material were washed consecutively with N-ChIP Wash Buffers I-III

for 5 minutes each at RT with rotation. DNA was eluted from beads by adding 2x100µl of freshly prepared Elution Buffer and incubating 15min at RT on a roto-shaker. DNA was purified with Qiagen Qiaquick PCR columns and eluted in 80µl of dH₂O. ChIP DNA was quantified using the Quant-iT PicoGreen dsDNA assay per manufacturer's instructions.

2.4.20. Cross-linked chromatin immunoprecipitation (X-ChIP)

Chromatin preparation - Sonication

Cross-linked nuclei were isolated using the protocols detailed in 2.4.14 for livers and 2.4.15 for cells. Nuclei were resuspended in 125µl of Nuclei Wash Buffer. An equal volume of Nuclei Lysis Buffer was added and samples were incubated on ice for 5 minutes. Lysates were sonicated for 15 cycles of 30 sec ON/ 30sec OFF at high amplitude to shear DNA to an average fragment size of 400-700bp. Lysates were centrifuged for 5min at 4°C at 12 000 x g and the supernatant containing soluble chromatin was transferred into a new tube. An aliquot of chromatin (10-20µl) was reverse crosslinked in 200µl elution buffer for 6-16h at 65°C. The aliquot was treated with RNase A and Proteinase K. DNA was purified using Qiagen PCR columns, quantified with the Nanodrop 2000 and electrophoresed in a 1% agarose gel to verify average fragment size.

Crosslinked chromatin immunoprecipitation

Approximately 25µg of DNA per ChIP was used. Samples were diluted 10X in ChIP Buffer supplemented with 1X protease inhibitor cocktail (Roche). 1-5% input chromatin was kept. 2.5µg of primary antibody was added to samples. Non-specific (mock) IgG from the same species as the primary antibody was used as negative control. Chromatin was incubated with antibody overnight at 4°C with rotation. For H2A.Z ChIP, 7.5µl of protein A-coupled and 7.5µl of protein G-coupled Dynabeads were added to samples and incubated for 2 hours at +4°C with rotation. Tubes were placed on a magnet and the supernatant was kept as the unbound fraction. The beads were washed consecutively with X-Wash Buffer I-III for 5 minutes with rotation.

Elution and reverse cross-linking

Immunoprecipitates were eluted from beads by adding 2x100µl of freshly prepared Elution buffer and incubating 15min at RT on a shaker. Inputs were also completed to a final volume of 200µl with elution buffer. The eluates were reverse crosslinked by incubation at 65°C for 16hours. DNA was purified with Qiagen Qiaquick PCR columns and eluted in

80µl of dH₂O. ChIP DNA was quantified using the Quant-iT PicoGreen dsDNA assay per manufacturer's instructions.

2.4.21. Immunofluorescent Staining and imaging

Cells were fixed in formalin solution for 15min at 37°C, permeabilized with 0.2% Triton-X and incubated with lamin-B antibody (Santa-Cruz sc-6217) diluted 1/300 in 10% goat serum-PBS. Anti-goat IgG coupled with Cyanine 3 (Jackson 705-165-147) was used as secondary antibody. DNA was stained with Hoechst 33342 intercalating dye. All fluorescent microscopy imaging was performed on fixed cells with a Zeiss Axio Imager Z1 microscope with a Plan-Apochromat x63 objective. Z-stack images were acquired with a Zeiss AxioCam camera piloted with the Zeiss Axiovision 4.8.10 software. All image treatment was performed using Fiji (ImageJ2-rc14).

2.4.22. Nuclear and mitotic defect evaluation

The three main mitotic defects observed in this study were: misaligned chromosomes during metaphase, lagging chromosomes during early anaphase and chromatin bridges during late anaphase or beginning of telophase. Cells with nuclear defects had polylobed nuclei or micronuclei. Each experiment was repeated 3 times and 125 mitotic events and 400 nuclei were counted per cell type in each experiment.

2.4.23. Flow cytometry cell cycle analysis with propidium iodide DNA staining

Cells were detached with trypsin/EDTA solution, washed once with PBS and pelleted. After resuspension in a small volume of PBS, ice-cold 70% ethanol was added and cells were fixed overnight at 4°C on a rotator. Cells were washed once in PBS and stained with propidium iodide (PI) solution containing 5µg/µl PI and 200µg/ml RNaseA for 30 minutes at 37°C. Approximately 100 000 cells per condition were analyzed using the Accuri C6 flow cytometer (Becton Dickinson). The percentage of cells in each phase of the cell cycle was determined using MultiCycle AV module embedded in the FCS Express 6 software (De Novo Software, California, USA) and Modfit LT (Verity Software House, Maine, USA). An example of analysis performed is shown in Appendix A, Figure A.

2.4.24. RNA and ChIP sequencing and bioinformatic analysis

H3.3 immunoprecipitated DNA library preparation and paired-end sequencing was conducted in collaboration with Dr. Thomas Westerling in Dana Farber Cancer Institute,

Boston, MA, USA. H3.3 genome-wide mapping analysis was conducted in collaboration with Dr. Razvan Chereji in NIH, Bethesda, MD, USA. RNA library preparation, sequencing and bioinformatic analysis were undertaken in collaboration with Dr. Christophe Papin in IGBMC, Strasbourg, France ²⁷⁹.

For ChIP-Seq, image analysis and base calling were performed using RTA 1.17.20 and CASAVA 1.8.2. (Illumina). Reads were mapped to the mouse genome (mm9) using Bowtie ²⁸⁰ using the following arguments “-m 1 --strata --best -y -S -l 40 -p 2”. Heatmaps and quantitative comparisons of the ChIP-Seq data were performed running seqMINER ²⁸¹ (<http://bips.u-strasbg.fr/seqminer/>), using datasets normalized to 10 million uniquely mapped reads (rp10m, reads per 10 million). As reference coordinates for genes, the Ensembl 67 database of the mouse genome (mm9) was used. Tag densities were collected in 100 bp sliding windows spanning 2 kb (divided in 10 bins) of the length-normalized gene bodies (divided in 50 bins) ²⁸².

RNA-Seq analysis was performed as described in Ors *et al.* ²⁷⁹

The RNA-Seq datasets (raw data as well as processed expression datasets) obtained in MEFs have been deposited in the Gene Expression Omnibus (GEO) under the accession number GSE84308.

2.4.25. Statistical analysis of qPCR data

To determine statistical significance of differences observed between groups in qPCR or FACS analyses, one-way or two-way ANOVA and Student's t-tests were carried out where indicated using GraphPad Prism software (version 6, San Diego, CA, USA). Differences with a $P < 0.05$ were considered significant. The degree of significance was depicted using * for $P < 0.05$, ** for $P < 0.01$, *** for $P < 0.001$ and **** for $P < 0.0001$.

Chapter 3. Results

3.1. Novel knock-in/ conditional knock-out mouse lines

One of the major problems faced with studying H3.3, is the lack of specific antibodies against H3.3 due to its similarity to the more abundant replication dependent histone H3 variants. To counter this problem, researchers developed cell models where a tagged H3.3 was either expressed exogenously^{159,213,220} or replaced endogenous *H3f3b* gene to express a tagged H3.3B²⁰⁵. An important point to take into consideration in these models is that the tag is placed at the C-terminal end of H3.3 which is localized towards the core of the nucleosome. This renders the tag harder to access and more likely to disrupt the core structure, therefore the function, of the nucleosome. The resulting genome-wide mapping of H3.3 based on ChIP-Seq experiments from these studies are of poor resolution^{205,213}.

3.1.1. Molecular description and validation of mouse models

To address the aforementioned issue, we have generated transgenic knock-in mice that express a FLAG-FLAG-HA tagged histone H3.3 with the tag placed at N-terminal of the protein. These mice also contain a LoxP recombination site at either end of exon 2 of one of the two genes coding for H3.3: *H3f3a* or *H3f3b*. The start codon being localized in the second exon, upon its deletion following expression of the Cre recombinase, a H3.3A or H3.3B functional knock-out mutant is generated (Figure 3-1).

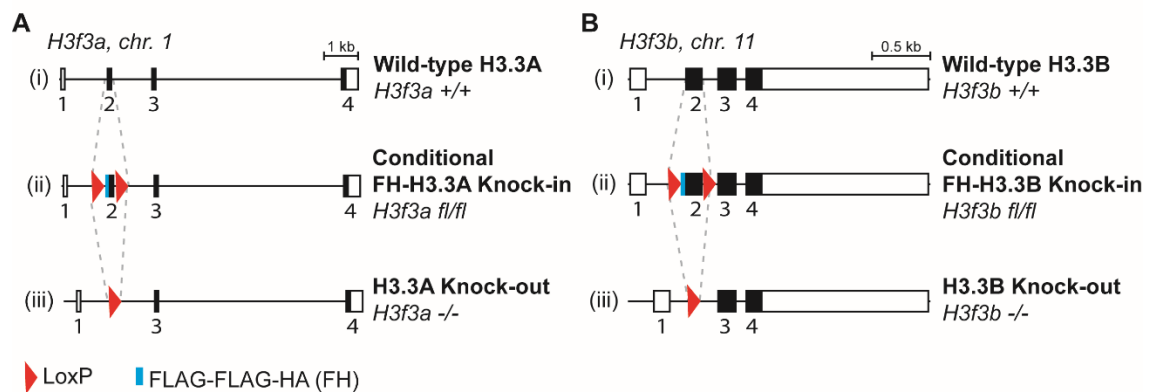


Figure 3-1. Generation of H3.3A and H3.3B mouse models. (i) Wild-type *H3f3a* and *H3f3b* gene structure. The open reading frame is indicated by black boxes (ii) A DNA element encoding the FLAG-FLAG-HA amino acids (blue box) was inserted in frame with the N-terminus of H3.3A and H3.3B. In addition, a LoxP site (fl, red triangle) was inserted on both ends of exon 2. Mouse lines were derived from ES cells carrying the modified *H3f3a* fl/fl and *H3f3b* fl/fl alleles. (iii) Structure of H3.3A-KO and H3.3B-KO, after Cre recombinase expression, which deletes exon 2 and generates loss of function (KO) alleles, *H3f3a* -/- and *H3f3b* -/-. Figure adapted from Ors *et al.*²⁷⁹

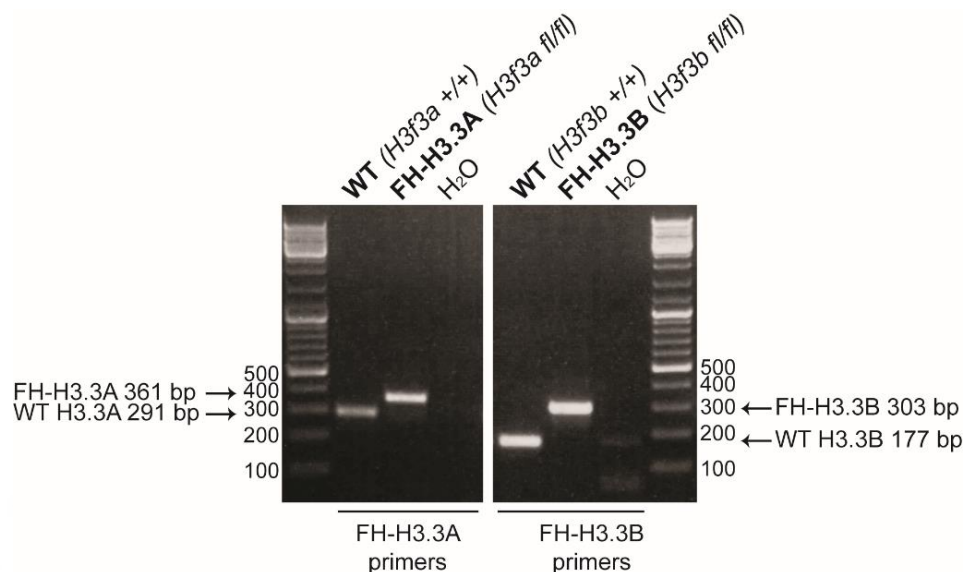


Figure 3-2. Genotype validation of FH-H3.3A and FH-H3.3B mice. Agarose gel electrophoresis of genomic DNA amplicons after PCR with relevant genotyping primers.

Mouse genotypes were validated by PCR on genomic DNA purified from mouse tail biopsies (Figure 3-2 A). Relative mRNA levels for *H3f3a* and *H3f3b* did not show significant difference in livers collected from WT, FH-H3.3A and FH-H3.3B adult mice and across genotypes (Figure 3-3 A). Protein expression of the tagged H3.3A and H3.3B proteins in liver chromatin was verified by Western Blot using an antibody against FLAG, equal loading was confirmed using an antibody against histone H4. The FLAG-FLAG-HA tag added in frame with the N-terminal of H3.3 protein shifted the size of the protein by approximately 8 kDa. Chromatin isolated from WT mice served as control for the specificity of the FLAG antibody against FH-H3.3A or FH-H3.3B (Figure 3-3 B).

H3f3a fl/fl and *H3f3b* fl/fl mice were crossed with Cre deleter mice expressing the Cre recombinase ubiquitously under the actin b promoter. The deletion of the floxed exons was verified by genotyping. The resulting heterozygous *H3f3a* +/- and *H3f3b* +/- mice were further crossed between each other to obtain single H3.3A, single H3.3B and double H3.3 knock-outs. Complete knock-out of H3.3 was verified by PCR (Figure 3-4 A) as well as RT-qPCR for RNA expression (Figure 3-4 B) and Western blot for protein expression (Figure 3-4 C).

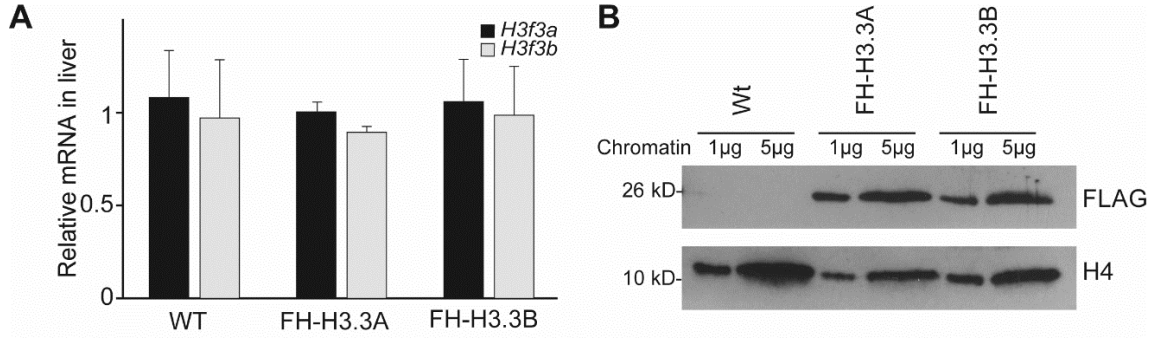


Figure 3-3. FH-H3.3A and FH-H3.3B expression in mouse livers. **A.** Relative H3.3 mRNA expression in WT, FH-H3.3A and FH-H3.3B livers. Quantitative RT-qPCR mRNA profiles normalized to ribosomal protein S9 (Rps9) mRNA and relative to one of the WT livers set at 1.0. Mean relative mRNA levels of 4 samples per genotype $\pm$ standard deviation. Two-way ANOVA analysis resulted in no significant change between genotypes and *H3f3a* and *H3f3b* levels. **B.** Different quantities of liver nuclei extractions from WT, FH-H3.3A and FH-H3.3B mice were analyzed for FLAG expression. Equal loading was verified with histone H4 expression.

Compared to WT liver RNA, H3.3A expression was completely abolished in H3.3A-KO livers while H3.3B RNA expression showed a 1.5-fold increase. Likewise, H3.3B expression was completely abolished in H3.3B-KO livers whereas H3.3A RNA expression showed a near 2-fold increase (Figure 3-4 B). Western blot analysis showed a complete loss of FLAG epitope in H3.3A-KO and H3.3B-KO livers, an antibody against H4 was used as a loading control (Figure 3-4 C). Unfortunately, due to lack of a specific H3.3 antibody, it wasn't possible to determine whether the protein expression of the remaining H3.3 gene varied when one of the H3.3 coding genes was knocked-out.

FH-H3.3A and FH-H3.3B mice were also crossed with the aim of obtaining homozygous mutants for all 4 H3.3 coding alleles. Unfortunately, even though animals with 1 to 3 mutant alleles were easily obtained, animals with 4 mutant alleles were never observed.

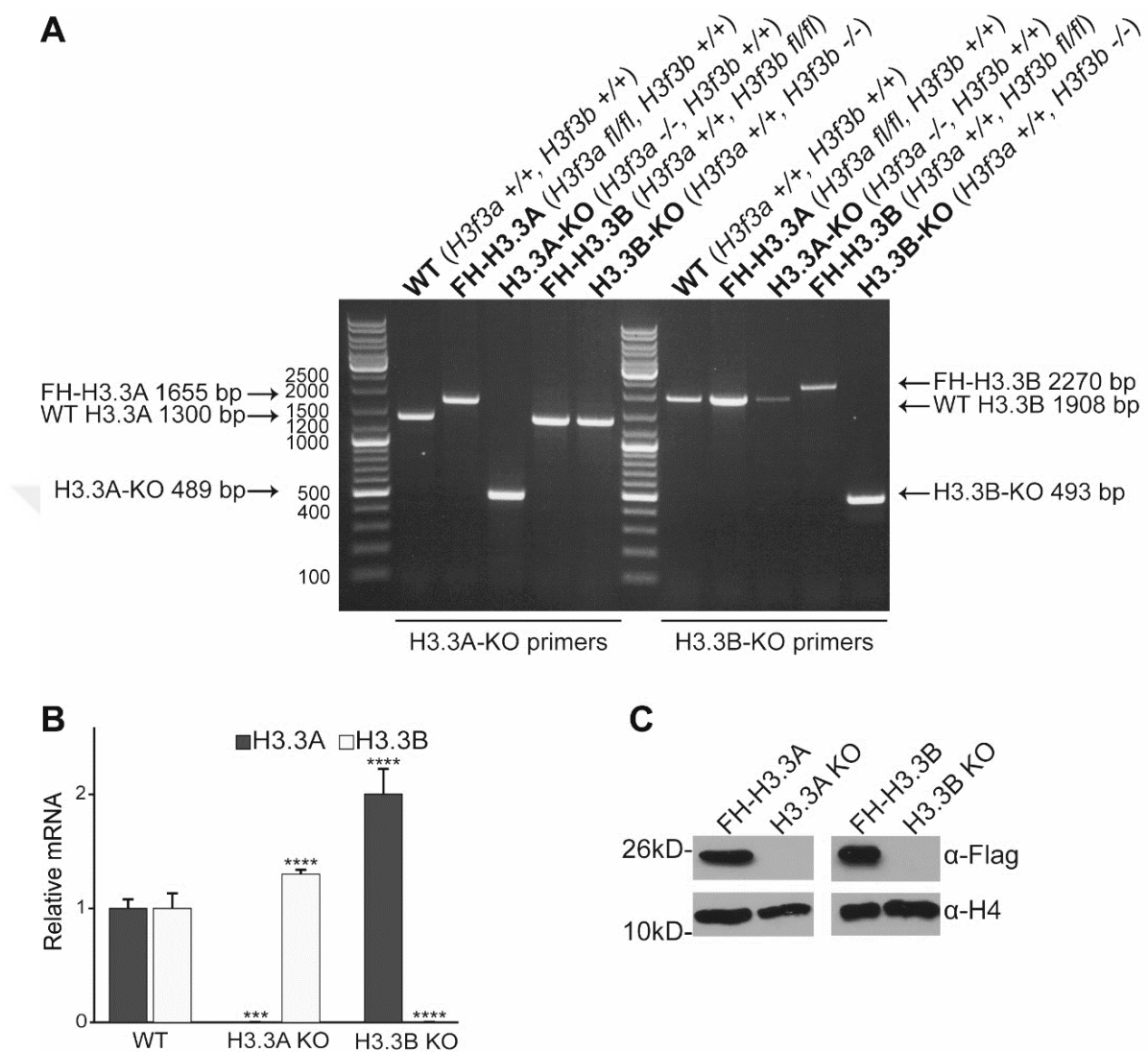


Figure 3-4. H3.3 knock-out validation of WT, FH-H3.3A, H3.3A-KO, FH-H3.3B and H3.3B-KO mice by genotyping, RT-qPCR and Western blot. **A.** Agarose gel electrophoresis of DNA amplicons after PCR with relevant genotyping primers to confirm genotypes of mice used in study. **B.** Relative H3.3 mRNA expression in WT, H3.3A-KO and H3.3B-KO mouse livers. Quantitative RT-qPCR mRNA profiles normalized to ribosomal protein S9 (Rps9) mRNA and relative to the values obtained for the control WT livers set at $1.0 \pm \text{s.e.m.}$ One-way ANOVA, $n=6$, $P < 0.001$ for WT vs H3.3A-KO for H3.3A expression and $P < 0.0001$ for WT vs. H3.3B-KO for H3.3A expression, WT vs. H3.3A-KO and WT vs. H3.3B-KO for H3.3B expression. **C.** Loss of the FLAG epitope in H3.3-KO mouse livers. H4 was used as loading control.

3.1.2. WT, FH-H3.3A, FH-H3.3B, H3.3A-KO and H3.3B-KO mouse lines are phenotypically similar.

For general phenotype assessment, data from 23 pair mated females on average litter size, weaned to born ratio, gender and genotype distribution was collected (Table 3-1 and Table 3-2). Average litter size, weaned to born and female (F) to male (M) ratios showed no significant differences between the different mouse lines (Table 3-1) and genotype distributions were observed at near mendelian ratios (Table 3-2). When observed, mutant mice did not show any difference in behavior compared to their litter mate controls regarding aggressiveness or activity. Although no scientific data was recorded regarding size and weight, it should be noted that KO models seemingly showed more failure to thrive than their control littermates. No other pelage, skeletal or body conformational changes were observed.

Table 3-1. Average litter size, weaned to born and female (F) to male (M) ratios in mouse models. Data observed in 23 pair mated females. The error in average litter size corresponds to standard deviation. One-way ANOVA analysis yielded in no significant differences between mouse lines in F to M and weaned to born ratios ($P > 0.05$).

Strain	Average litter size (born)	F to M ratio	Weaned to born ratio
WT	6.7 ± 1.4	0.84	0.87
FH-H3.3A	6 ± 1.5	0.83	0.84
H3.3A-KO	6.4 ± 2.6	0.93	0.82
FH-H3.3B	6.6 ± 2	0.91	0.87
H3.3B-KO	7.1 ± 1.7	0.86	0.89

Table 3-2. Genotype distribution after heterozygous mating of mouse lines. Data observed in 100 mice born from heterozygous crossings. Chi-square test ($P > 0.05$)

Genotype	+/+	+/- or +/-	fl/fl or -/-
Strain			
FH-H3.3A	0.24	0.56	0.2
H3.3A-KO	0.26	0.5	0.24
FH-H3.3B	0.29	0.51	0.20
H3.3B-KO	0.29	0.5	0.21

Even though the different mouse models showed comparable phenotypes, there was a clear decrease in fertility in single knock-out mice. When homozygous H3.3-KO (*H3f3a*^{-/-} or *H3f3b*^{-/-}) mice were crossed between each other, H3.3A-KO mice seldom yielded viable litters, indicating a subfertility in this mouse-line and with H3.3B-KO mice no litters were observed even after 6 months, indicating clear infertility in this mouse-line.

Unfortunately, due to fertility problems in *H3f3a*^{-/-} and *H3f3b*^{-/-} mice and possible *in utero* lethal effects mentioned above and also described in literature^{186,188,191,252}, efforts to produce viable double H3.3A-KO and H3.3B-KO mice were fruitless.

In summary, a novel knock-in/conditional knock-out mouse line for the study of the H3.3 protein was established and characterized. The second exon of either one of the endogenous H3.3 coding genes *H3f3a* or *H3f3b* was targeted to insert a DNA element encoding the tags FLAG-FLAG-HA in frame with the N-terminus of H3.3. Moreover, LoxP sites were inserted at either end of exon 2 to allow for deletion of the region encompassing the start codon upon Cre recombinase expression.

Mice were characterized on a molecular level by genotype, RNA expression and protein expression assessment. Upon Cre recombinase expression, the FH-H3.3 protein expression was completely lost demonstrating an efficient functional knock-out model. Phenotypic characterization of mice yielded in no significant differences between WT, FH-H3.3A, FH-H3.3B, H3.3A-KO and H3.3B-KO mouse lines. Unfortunately, generation of double H3.3A-KO/ H3.3B-KO mice proved to be impossible due to fertility and viability problems associated with the loss of H3.3.

These mouse models constitute convenient tools for the study of H3.3 for two main reasons; the presence of the tag at the N-terminal site of H3.3 will allow for efficient recognition of H3.3 and the conditional knock-out of H3.3 will help further elucidate the role of this protein in the physiology of the liver and other cellular processes. Furthermore, having separate mouse lines for H3.3A and H3.3B will allow to determine the specific function of each H3.3 coding gene.

3.2. Genome-wide distribution of H3.3 at nucleosome resolution in the liver

Determining the precise genome-wide localization map of H3.3A and H3.3B is one of the first steps in uncovering the functional role of H3.3 as sites of preferential enrichment can be indicative of specific cellular functions.

The ChIP technique determines whether a protein of interest is directly associated with a specific region of the genome. Crosslinked-ChIP (X-ChIP) relies on the use of crosslinking agents like formaldehyde to “freeze” the protein-DNA interactions. The signal obtained from those studies can be tainted by artifacts of looping and protein-protein interactions rather than direct DNA binding. Native-ChIP (N-ChIP) does not involve the use of crosslinking agents and can be used to study proteins closely bound to DNA like histones. In N-ChIP, chromatin is prepared by digesting nuclear preparations with micrococcal nuclease. This DNase primarily digests DNA unbound by proteins, in this case, between nucleosomes. The sequence of the nucleosome bound DNA obtained after immunoprecipitation with an antibody against FH-H3.3 is determined using paired-end deep sequencing and aligned to the mouse genome. This method results in the highly accurate positioning of H3.3 on the whole genome at the nucleosome resolution.

3.2.1. Mono-nucleosome preparation from liver tissue ensures high-resolution profiling of H3.3

In order to map the genome-wide localization of H3.3A and H3.3B in the liver at the resolution level of the nucleosome, we optimized a high-resolution native chromatin immunoprecipitation (N-ChIP) technique. In summary, liver nuclei were isolated from 5 pooled livers per genotype and chromatin was fragmented via micrococcal nuclease digestion (Figure 3-5 A) Mono-nucleosomes were then isolated on a 5-20% sucrose gradient (Figure 3-5 B). FH-H3.3 was immunoprecipitated using an HA antibody coupled to agarose beads. Chromatin from wild-type livers with no HA epitope was used as a negative control. The ChIPs were set in duplicate; one of the duplicates was used for protein extraction and the other for DNA purification. Western blot analysis of input mono-nucleosomes and ChIP samples showed high enrichment of ChIP with FH-H3.3A or FH-H3.3B and the presence of the H4 signal accounted to the integrity of the nucleosomes before and after ChIP. The signals for FLAG and H4 observed in the mono-nucleosomes

also provide evidence for nucleosome incorporation of FH-H3.3 (Figure 3-5 C). The purified ChIP DNA was quantified (Figure 3-5 D). ChIP-qPCR against β -actin (*Actb*) and interferon α receptor 2 (*Ifnar2*) distal promoter regions, previously reported to be enriched with H3.3²²⁰, showed a high enrichment over input for FH-H3.3A and FH-H3.3B compared to WT negative control. Enrichment was similar between FH-H3.3A and FH-H3.3B (Figure 3-5 E).

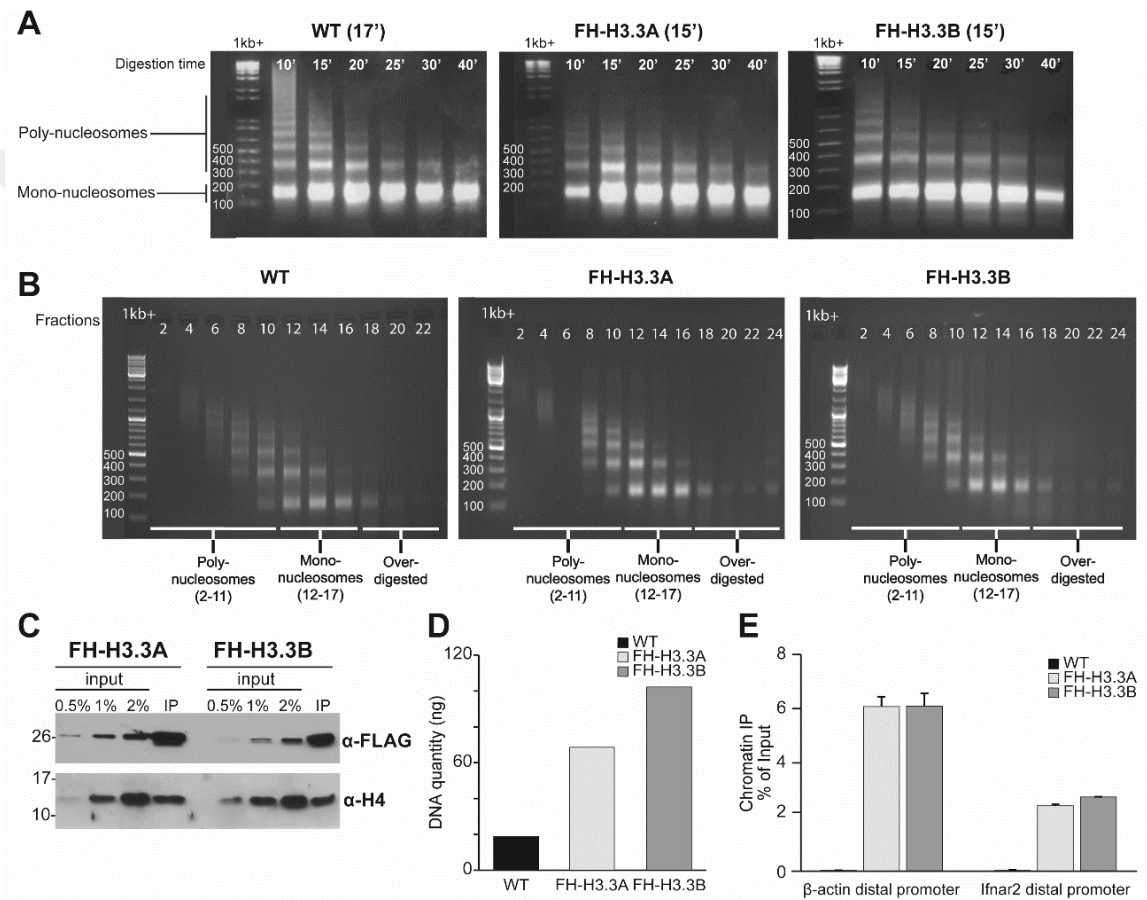


Figure 3-5. Mono-nucleosome preparation for N-ChIP and ChIP validation. Digestion assays for WT, FH-H3.3A and FH-H3.3B mouse liver nuclei. Nuclei isolated from mouse livers were digested for the indicated times and DNA fragments were separated on a 1.5% agarose gel to determine optimal digestion time. **B.** Agarose gel electrophoresis of 5-20% sucrose gradient fractions after separation of isolated nucleosomes following MNase digestion, fractions enriched in mono-nucleosomes were pooled and consequently used in ChIP. **C.** Western blot of FH-H3.3A and FH-H3.3B mono-nucleosomes used as input and IP material from HA-H3.3 ChIP. Membranes were marked with FLAG and H4 antibodies to verify nucleosome integrity before and after immunoprecipitation. **D.** DNA quantity of ChIP DNA (ng) from WT, FH-H3.3A and FH-H3.3B ChIP. **E.** ChIP-qPCR for WT, FH-H3.3A and FH-H3.3B ChIP using primers against β -actin (*Actb*) and Interferon α receptor 2 (*Ifnar2*) distal promoter regions to check for ChIP specificity and efficiency.

The verified ChIP DNA was sent to our collaborator Dr. Thomas Westerling in Dana Farber Cancer Institute, Boston, MA, USA, for library preparation and deep-sequencing. Bioinformatic analysis was undertaken by Dr. Razvan Chereji in NIH, Bethesda, MD, USA and Dr. Christophe Papin in IGBMC, Strasbourg, France.

Sequenced ChIP DNA was aligned to the reference mouse genome (mm10). Data showed high enrichment of H3.3A and H3.3B at promoter regions and near transcription start sites (TSS) but also at gene bodies and intergenic regions (Figure 3-6).

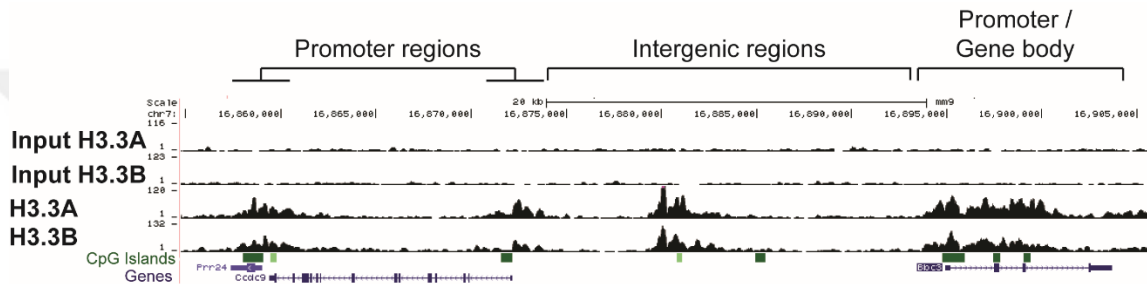


Figure 3-6. ChIP-Seq data reveals enrichment of H3.3. ChIP-Seq profiles across a representative region (100 MB of chromosome 7) in FH-H3.3A and FH-H3.3B mouse livers with HA antibody coupled agarose beads. Y-axis represents the number of reads spanning a genomic position. CpG islands (green) and genes (dark blue) are annotated below the ChIP-Seq panels.

3.2.2. H3.3 is highly enriched at TSS and its enrichment positively correlates with gene expression

Our data analysis showed that H3.3A and H3.3B were both present at the nucleosomes upstream and downstream of TSS and to a lower degree at transcription termination sites (TTS) (Figure 3-7). H3.3 was found to be absent at the TSS itself. It is to be noted that the TSS is often described as a nucleosome free-region that allows easy access to DNA binding transcription factors.

When genes were grouped according to their transcript levels as high, medium, low and non-transcribed genes, H3.3A and H3.3B were found to be enriched at the TSS of both active, highly expressed genes as well as repressed genes with low expression (Figure 3-8). The level of H3.3 enrichment around TSS correlated with gene expression. In actively transcribed genes, higher levels of H3.3 were observed at their promoter regions in comparison to promoter regions of lower expressed genes.

In addition, our data showed H3.3 enrichment at up to 5 nucleosomes downstream of TSS. This level of resolution has not been obtained in other genome-wide H3.3 mapping studies and confirm the high resolution quality of our data^{205,213}.

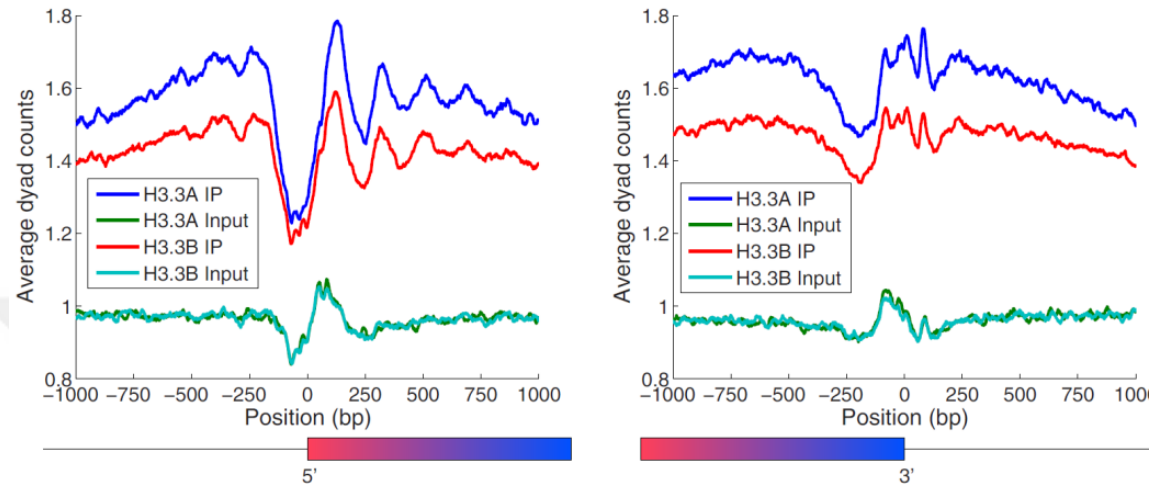


Figure 3-7. H3.3A and H3.3B are enriched around transcription start (TSS) and termination sites (TTS). Profiles of H3.3A (blue) and H3.3B (red) variants across the TSS and TTS. Y-axis represents the average number of tags per genes per 40 000 mapped reads. X-axis represents the position in base pairs relevant to the TSS. Data produced in collaboration with Dr. Razvan Chereji.

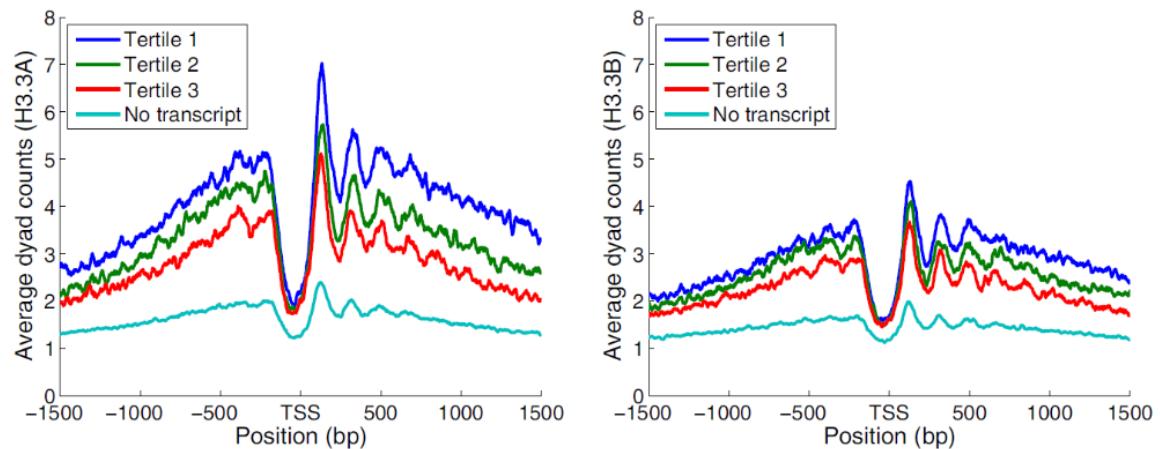


Figure 3-8. H3.3 is present at TSS and positively correlates with gene expression. Profiles of H3.3A(left) and H3.3B (right) across TSS for highly active (Tertile 1, blue), medium expressing (Tertile 2, green), low expressing (Tertile 3, green) and non-transcribed (light blue) genes. Y-axis represents average number of tags per gene per 40 000 reads. X-axis represents the position in base pairs relevant to the TSS. Data produced in collaboration with Dr. Razvan Chereji.

3.2.3. H3.3A and H3.3B have identical enrichment patterns at most genomic sites except at some retroviral repeat elements

H3.3 was found enriched in TSS, CpG islands and enhancers and its enrichment correlated with active transcription markers such as RNA polymerase II (Pol II), H3K4me3 and H3K27ac (Figure 3-9). Regardless of the differences in nucleotide sequence between the *H3f3a* and *H3f3b* genes, their protein products H3.3A and H3.3B have identical amino-acid sequences and H3.3A and H3.3B showed similar deposition patterns at most genomic sites.

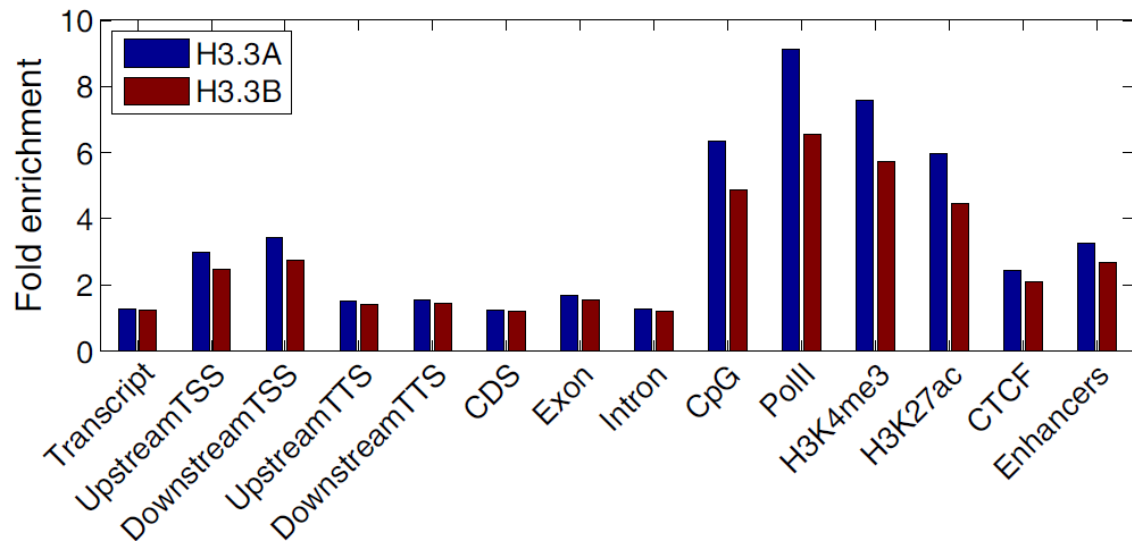


Figure 3-9. H3.3A and H3.3B show similar deposition patterns at most genomic regions. Enriched regions were identified for FH-H3.3A and FH-H3.3B. Y-axis represents the fold enrichment and X-axis corresponds to different categories of gene annotation and active transcription marks. Data produced in collaboration with Dr. Razvan Chereji.

In order to assess the enrichment of H3.3 at repetitive sequences, nucleosome sequences were remapped to an “artificial chromosome” consisting of concatenated repeats of 20kbp of Repbase motifs²⁸³. Mapped reads were normalized and computed to the number of reads that mapped to each repeated motif per 1M of total mapped nucleosomes. All samples had about 40M mapped nucleosomes. Enrichment of each variant (IP/Input) was computed and the regions were sorted according to the maximum enrichment between H3.3A and H3.3B. Interestingly, at some repeat sequences, comprised mainly of endogenous retrovirus repeats (ERVs), a differential enrichment between H3.3A and H3.3B was observed. The first 30 regions with the highest differential enrichment are shown in Figure 3-10

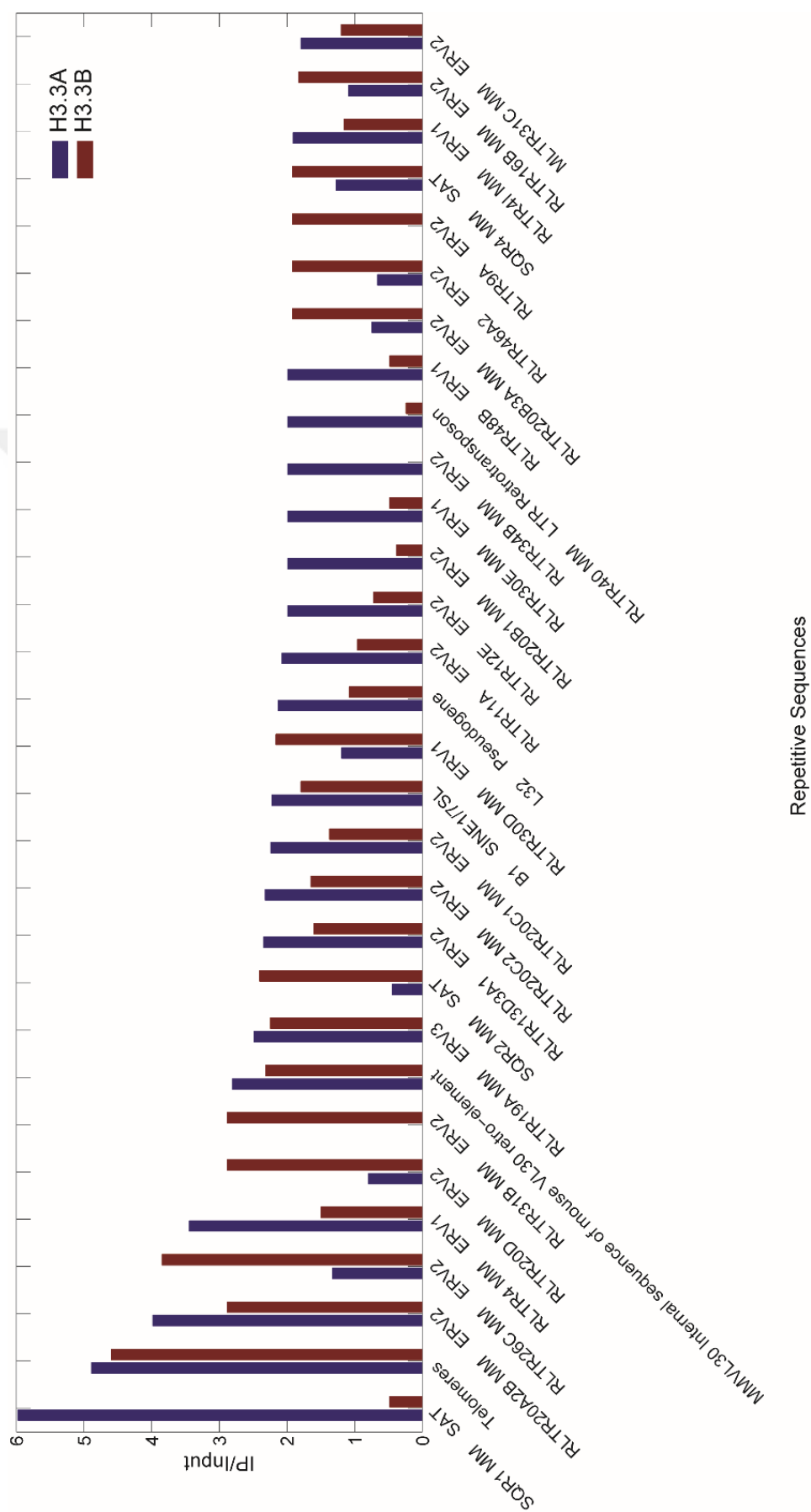


Figure 3-10. H3.3A and H3.3B enrichment differs at repetitive sequences. Reads from FH-H3.3 were aligned to a library of consensus repetitive sequences ²⁸³
The first 30 regions with the highest enrichment are represented as fold enrichment over input. Data produced in collaboration with Dr. Razvan Chereji.

3.2.4. H2A.Z is co-localized with H3.3 around the TSS

In order to map H2A.Z at the nucleosome level, we performed a cross-linked ChIP coupled to micrococcal nuclease digestion for higher resolution in WT mouse livers. Genome-wide H2A.Z mapping studies in mouse liver tissue showed an H2A.Z enrichment at promoter regions and CpG islands. Consistent with current literature, H2A.Z was found enriched both upstream and downstream of TSS ^{131,159,284}.

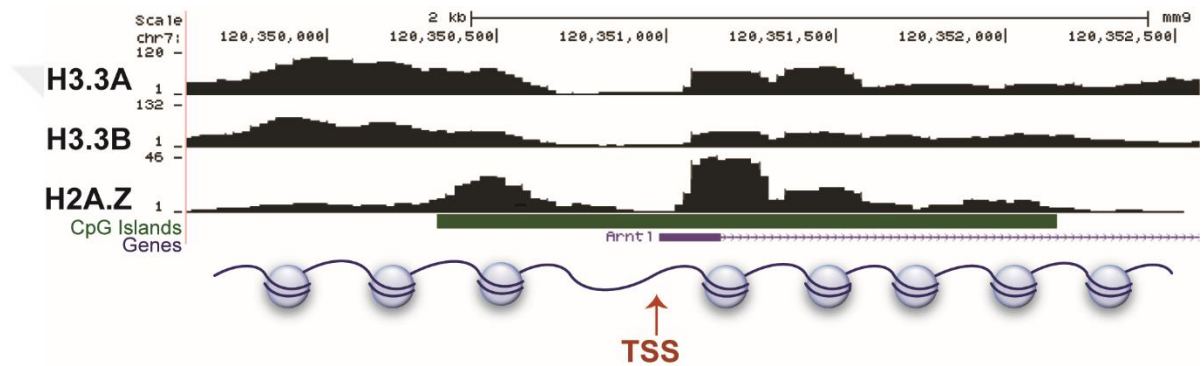


Figure 3-11. Genome-wide enrichment pattern of FH-H3.3A, FH-H3.3B and H2A.Z at TSS. ChIP-Seq profiles across a representative TSS (*Arnt1* gene) in mouse livers with HA antibody coupled agarose beads for FH-H3.3A and FH-H3.3B and H2A.Z antibody for WT. Y-axis represents the number of reads spanning a genomic position. CpG islands (green) and genes (dark blue) are annotated below the ChIP-Seq panels. A schematic representation of nucleosomes distribution is also presented.

As for H3.3, H2A.Z enrichment around the TSS positively correlated with gene expression levels. In actively transcribed genes, higher levels of H2A.Z enrichment were observed at their promoter regions in comparison to the promoter regions of lower expressed genes. However, whereas H3.3 enrichment extended to the gene body and transcription termination sites (TTS) in highly expressed genes, H2A.Z remained solely enriched at promoter regions regardless of the genes' transcription levels (Figure 3-12).

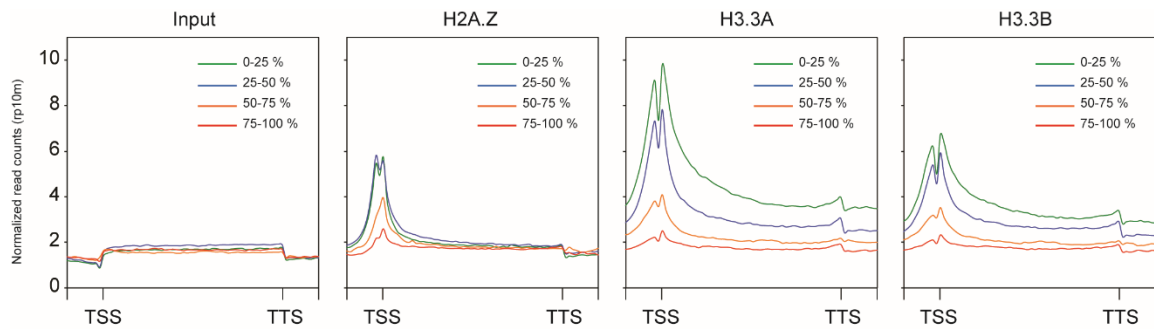


Figure 3-12. Normalized densities of H2A.Z, H3.3A and H3.3B within gene bodies expressed at different levels in mouse liver. Genes were sorted in quartiles according to their expression level. Data produced in collaboration with Dr. Christophe Papin.

In summary, high nucleosome resolution genome-wide mapping studies of H3.3 in the mouse liver, showed that H3.3 was found highly enriched at promoter regions but also at gene bodies, intergenic regions and TTS. Although H3.3 was not localized at the TSS itself, clear nucleosomal peaks up to 5 nucleosomes downstream and 3 nucleosomes upstream of TSS were observed. H3.3 enrichment around the TSS positively correlated with gene expression. In fact, H3.3 was also enriched in gene bodies of highly expressed genes. Similarly, H2A.Z was found enriched at promoter regions and this enrichment positively correlated with higher levels of expression. However, H2A.Z enrichment did not extend to the gene bodies or TTS of highly expressed genes.

Lastly, at most genomic sites, there were no differences between H3.3A and H3.3B enrichments. However, at some repeat sequences there were significant differences in enrichment between H3.3A and H3.3B.

In concordance with current literature, these results suggested an important role for H3.3 in transcription regulation. The H3.3A and H3.3B conditional knock-out mice developed in our laboratory placed us in a unique position to study this effect directly.

3.3. Effect of H3.3 loss on transcriptome

The results of the genome-wide mapping studies for H3.3 suggested an important role for H3.3 in transcription. The correlation of H3.3 localization with active transcription markers and highly expressed genes led to the hypothesis that in the absence of H3.3, the transcriptome would be heavily impacted. The effect of H3.3 on transcription was studied using RNA-Seq technology. Total transcriptome analysis by RNA-Seq, allows to determine and quantify all transcripts present in a given sample, including non-coding RNA with very high precision.

3.3.1. Knock-out of a single H3.3 coding gene does not affect liver transcriptome

In order to determine the effect of H3.3 loss on the transcription, RNA was isolated from validated WT, H3.3A-KO and H3.3B-KO liver tissues for total RNA sequencing. Library construction and sequencing was performed at the IGBMC Microarray and Sequencing platform (Strasbourg, France) and bioinformatic analysis was undertaken by Dr. Christophe Papin.

Bioinformatic analyses confirmed the positive relationship between high gene expression and high H3.3 and H2A.Z enrichment at TSS (Figure 3-13). However, H3.3A-KO and H3.3B-KO liver transcriptomes did not show much difference compared to the WT liver transcriptome. There was a total of 7 deregulated genes in H3.3A-KO livers and 11 deregulated genes in H3.3B-KO livers when compared to WT liver transcriptomes (Figure 3-14A). Given the enrichment variation between H3.3A and H3.3B at repetitive sequences and a recent study suggesting a role for H3.3 in repression of transcription from retroviral elements in ESCs²²⁴, a knock-out of either one of the genes could have resulted in variation in expression of repeat sequences. However, once again the single knock-out did not result in any significant differences in the expression of repeat families when compared to the WT control livers (Figure 3-14B). Transcript levels of H3.3B increased significantly when H3.3A was knocked-out, likewise, when H3.3B was knocked-out, H3.3A transcript levels increased (Figure 3-4A). This increase in transcript levels of the remaining H3.3 coding gene in single KO mice, could account for a compensation effect for the loss of either one of the H3.3 proteins. Accordingly, the necessity to generate double H3.3-KO mice where both H3.3 coding genes would be targeted for knock-out became evident. However, as it was underlined in chapter 3.1, efforts to produce double knock-out mice were unsuccessful.

To circumvent this problem, mouse embryonic fibroblasts were generated from FH-H3.3A and FH-H3.3B mouse embryos as detailed in section 2.4.2. and a new cell model in which expression from both H3.3 coding genes was suppressed by combining the Cre/Lox knock-out system with RNA interference based knock-down technology was developed as described in the following section 3.3.2.

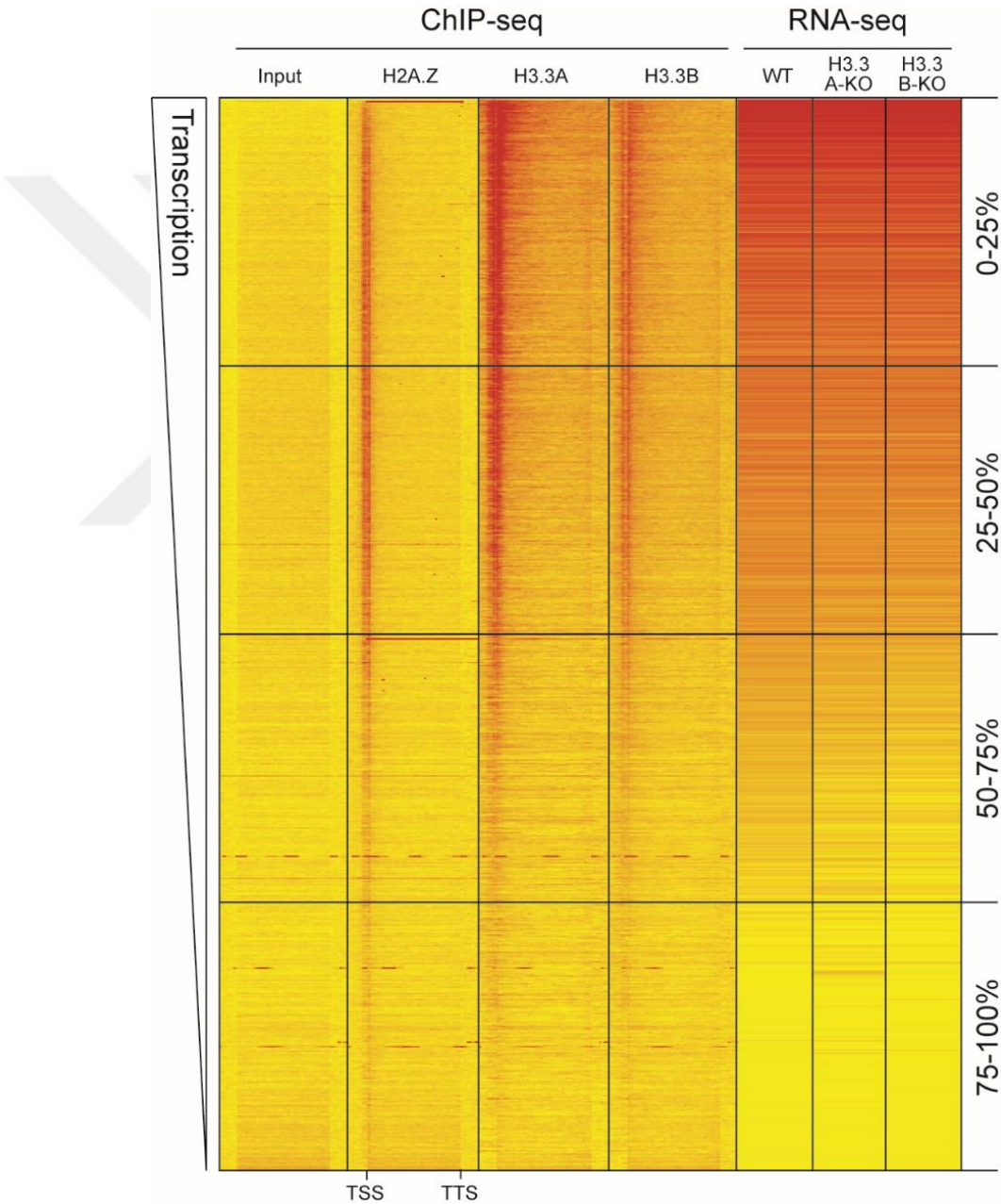
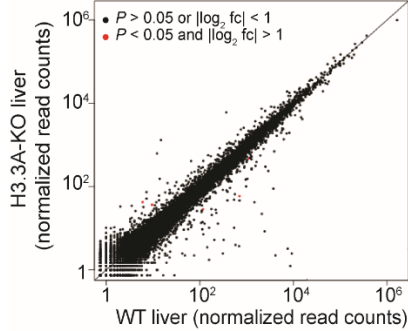


Figure 3-13. H3.3A, H3.3B and H2A.Z enrichment at TSS correlates with transcription levels. Heatmaps of H2A.Z, FH-H3.3A, FH-H3.3B localization and WT, H3.3A-KO, H3.3B-KO RNA levels within gene bodies in mouse liver. Genes were sorted according to their expression levels.

A. Global gene expression

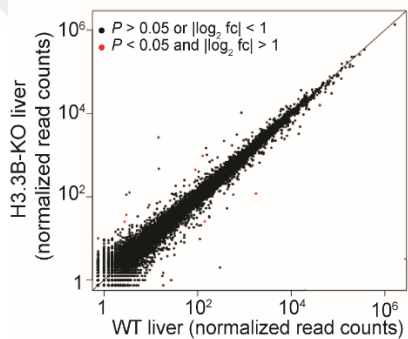
Only 7 genes significantly deregulated
in absence of H3.3A



Deregulated genes in H3.3A-KO
($P < 0.05$ and $|\log_2(KO/WT)| > 1$)

Genes	Log2 (KO/WT)
H3f3a	-3,25
RP23-381B19.7	2,45
Eif5a13-ps	3,12
Dct	-1,70
Osbp2	-2,47
Plekha6	-1,03
Cdc42ep5	1,62

Only 11 genes significantly deregulated
in absence of H3.3B



Deregulated genes in H3.3B-KO
($P < 0.05$ and $|\log_2(KO/WT)| > 1$)

Genes	Log2 (KO/WT)
H3f3b	-3,58
Cyp26b1	1,98
Gm15431	2,57
Serpina9	-1,96
Gm11645	-2,30
Cyp26a1	2,04
Gm10499	2,06
Gm15610	1,98
Gdf15	1,41
Derl3	-1,21
Gm12750	1,95

B. Transcription of DNA repeats

No repeat family significantly mis-regulated
in absence of H3.3A or H3.3B

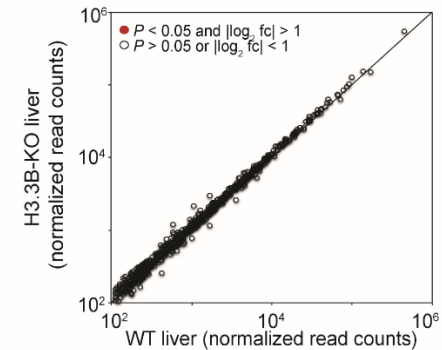
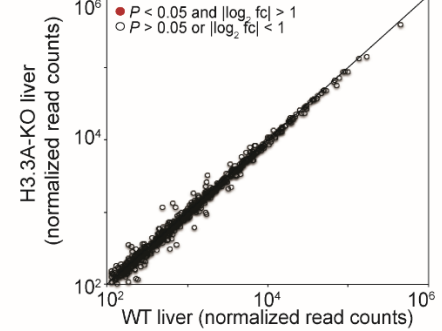


Figure 3-14. Effect of single H3.3 loss on the transcriptome of adult mouse livers. . A. Scatter plots comparing global gene expression profiles of control WT and H3.3A-KO and WT and H3.3B-KO livers. Red dots indicate differentially expressed genes ($P < 0.05$ and $|\log_2 fc| > 1$). **B.** Scatter plots comparing global transcription of repeat families in H3.3A-KO and H3.3B-KO livers compared to control WT livers. Red dots indicate differentially expressed genes ($P < 0.05$ and $|\log_2 fc| > 1$). Data produced in collaboration with Dr. Christophe Papin.

3.3.2. H3.3 depleted mouse embryonic fibroblasts

Problems in obtaining double H3.3-KO mice lead us to develop a strategy based on mouse embryonic fibroblasts (MEFs), a more easy-of-use model where one of the endogenous H3.3 coding genes can be knocked-out by Cre expression and the expression of the remaining gene knocked-down using RNA interference technology (Figure 3-15).

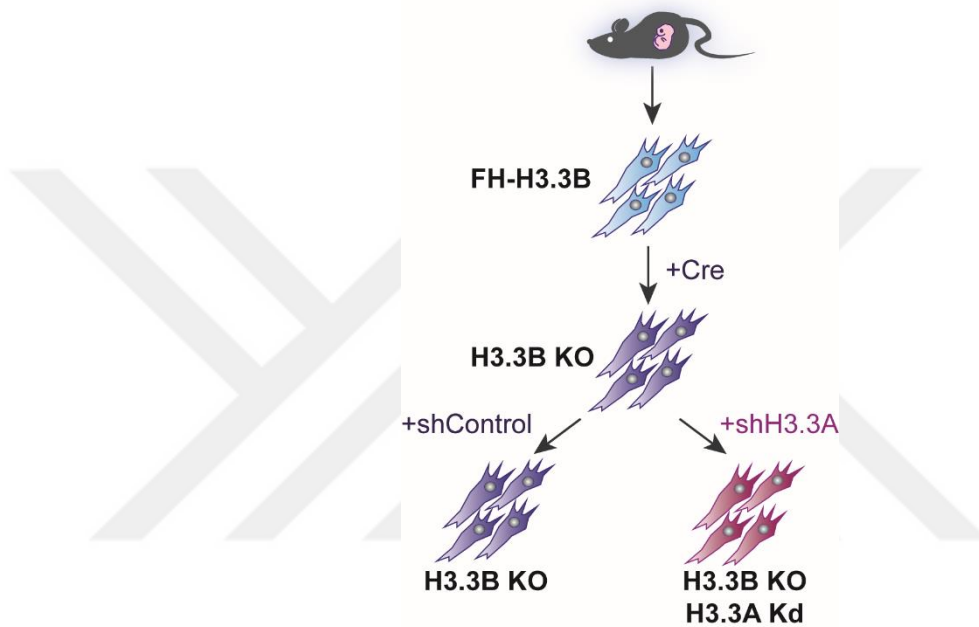


Figure 3-15. Generation of H3.3B-KO and H3.3B-KO / H3.3A-Kd MEFs. Mouse embryonic fibroblasts (MEFs) were isolated from *H3f3b* fl/fl embryos at 13.5 dpf. Transformed MEFs were infected with Cre recombinase expressing adenovirus to generate loss of function alleles, *H3f3b* ^{-/-}. H3.3B-KO MEFs were further infected with shRNA targeting the H3.3A transcript (shH3.3A) to produce double H3.3 deficient MEFs. H3.3B-KO MEFs were also infected with an empty lentiviral vector (shControl) as an infection control. Reproduced from Ors *et al.* ²⁷⁹

Homozygous *H3f3a* fl/fl and *H3f3b* fl/fl MEFs were isolated from FH-H3.3A or FH-H3.3B mouse embryos. Stable cell lines obtained with the serial passage method ²⁷⁷ were further infected with a Cre recombinase expressing Adenovirus and the endogenous *H3f3a* or *H3f3b* genes were knocked-out efficiently, verified by genotyping, RT-qPCR and Western blot (Figure 3-16).

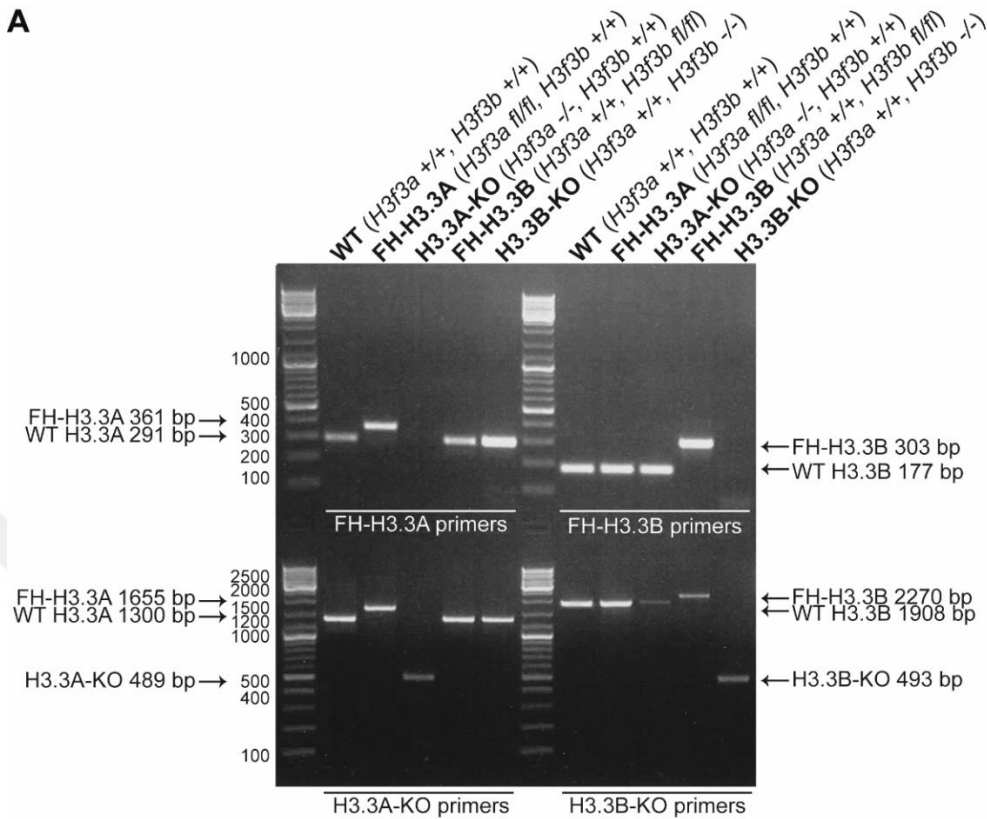
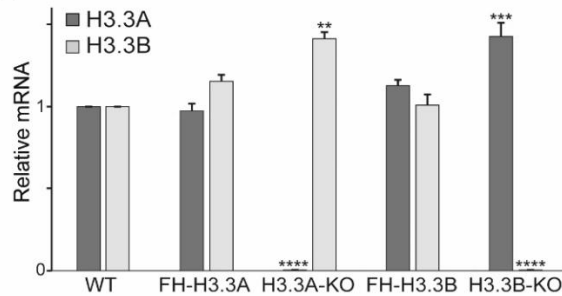
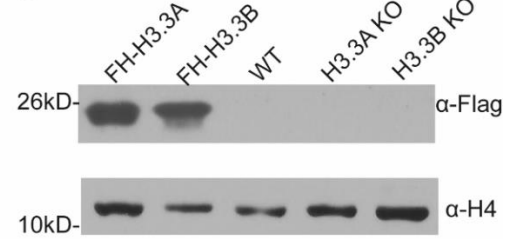
A**B****C**

Figure 3-16. Validation of WT, FH-H3.3A, H3.3A-KO, FH-H3.3B and H3.3B-KO MEFs by genotyping, RT-qPCR and Western blot. A. Agarose gel electrophoresis of DNA amplicons after PCR with relevant genotyping primers to confirm genotypes of MEFs used in study. **B.** Relative H3.3 mRNA expression in WT, FH-H3.3A, H3.3A-KO, FH-H3.3B and H3.3B-KO MEFs. Quantitative RT-qPCR mRNA profiles normalized to ribosomal protein S9 (Rps9) mRNA and relative to the values obtained for the control WT livers set at $1.0 \pm \text{s.e.m.}$ One-way ANOVA, $n=3$, for H3.3A expression: $P < 0.0001$ for WT vs. H3.3A-KO and $P < 0.001$ for WT vs. H3.3B-KO, for H3.3B expression: $P < 0.01$ for WT vs. H3.3A-KO and $P < 0.0001$ for WT vs. H3.3B-KO. One-way ANOVA, $n=6$, no significant differences in H3.3A or H3.3B expression between WT, FH-H3.3A and FH-H3.3B. **C.** Loss of the FLAG epitope in H3.3-KO mouse livers. H4 was used as loading control. C. Loss of the FLAG epitope in H3.3-KO MEFs after Cre expression. H4 was used as loading control. Figure B reproduced from Ors *et al.* ²⁷⁹

RT-qPCR experiments, showed a complete loss of H3.3A or H3.3B as it was the case in mouse livers. Again, as it was the case in livers, H3.3B expression in H3.3A-KO MEFs and H3.3A expression in H3.3B-KO MEFs was up to 1.4-fold higher than in control MEFs. There was no significant difference in H3.3 expression between WT, FH-H3.3A and FH-H3.3B MEFs (Figure 3-16 B). In Western blot, the FLAG epitope was completely lost in H3.3A-KO and H3.3B-KO MEFs (Figure 3-16 C) confirming the functional knock-out of H3.3 proteins.

Although the ultimate target cell lines were H3.3A-KO or H3.3B-KO MEFs, shRNA knock-down optimization studies were undertaken in FH-H3.3A and FH-H3.3B MEFs. The reason being that in these cell lines, variation of FH-H3.3A or FH-H3.3B expression can be specifically monitored thanks to the presence of the FLAG epitope.

In order to determine the shRNA with the best knock-down efficiency, FH-H3.3A and FH-H3.3B MEFs were transduced with five different shRNA against H3.3A and H3.3B. Knock-down efficiency was assessed by RT-qPCR (Figure 3-17 A) and Western blot (Figure 3-17 B). According to these results, the most efficient shRNA was the one against H3.3A with 80-90% efficiency.

The H3.3B-KO cell line was then stably transduced with anti-H3.3A shRNA (shH3.3A) or a control shRNA (shControl). Compared to the FH-H3.3B control cell line, H3.3B-KO transduced with the shControl had no H3.3B RNA expression and one-fold increased H3.3A RNA expression. In H3.3B-KO transduced with shH3.3A, H3.3A expression decreased by almost 90% at 7 days after transduction (Figure 3-18 A). All cells displayed a fibroblast like appearance (Figure 3-18 C). However, H3.3B-KO / H3.3A-Kd MEFs showed a significantly slower proliferation rate as shown by the increase in their doubling time compared to the control FH-H3.3B cells (Figure 3-18 B). For Figure 3-18 C, even though cells were plated at same seeding density and pictures taken after the same time in culture, FH-H3.3B MEFs had more densely covered the culture plate compared to H3.3B-KO and even more so compared to H3.3B-KO / H3.3A-Kd MEFs.

With a near complete depletion of H3.3, this cell model constituted a valuable tool to study the involvement of H3.3 in transcription.

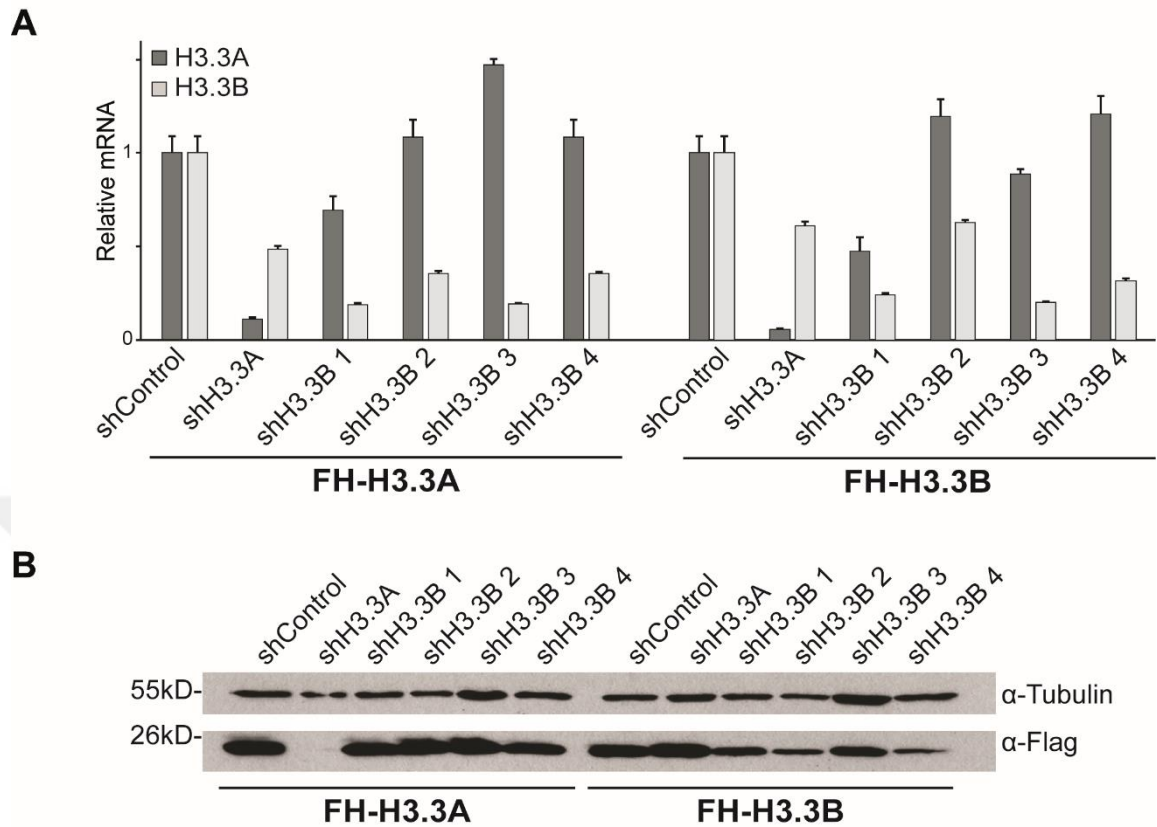


Figure 3-17. H3.3 specific shRNA selection and validation of knock-down efficiency. **A.** Relative H3.3 mRNA expression in FH-H3.3A and FH-H3.3B MEFs transduced with different shRNA against H3.3. Quantitative RT-qPCR mRNA profiles normalized to ribosomal protein S9 (Rps9) mRNA and relative to the values obtained for the control WT livers set at $1.0 \pm \text{s.e.m.}$ **B.** Western blot analysis of whole cell extracts from FH-H3.3A and FH-H3.3B MEFs transduced with different shRNA against H3.3. Expression of FLAG epitope was assessed to determine knock-down efficiency. Tubulin was used as loading control.

It should be noted that, mice with double heterozygous $H3f3a^{+/-} / H3f3b^{+/-}$ genotypes were also crossed with the aim of obtaining double knock-out MEFs before the onset of embryonic lethality. However, out of the 37 embryos collected at 13,5 dpf from 6 different matings, none of them were double knock-outs and only 3 of them were deleted for 3 alleles all $H3f3a^{+/-} / H3f3b^{-/-}$. MEFs isolated from embryos missing 3 alleles coding for H3.3, died rapidly after 2-3 passages without giving possibility to cell transformation. Given that mating of single heterozygous mice resulted in genotypes at expected mendelian ratios, these results indicated that embryonic lethality most likely occurred before 13.5 dpf and that a single H3.3 coding allele was not always enough to compensate for the loss of the three others.

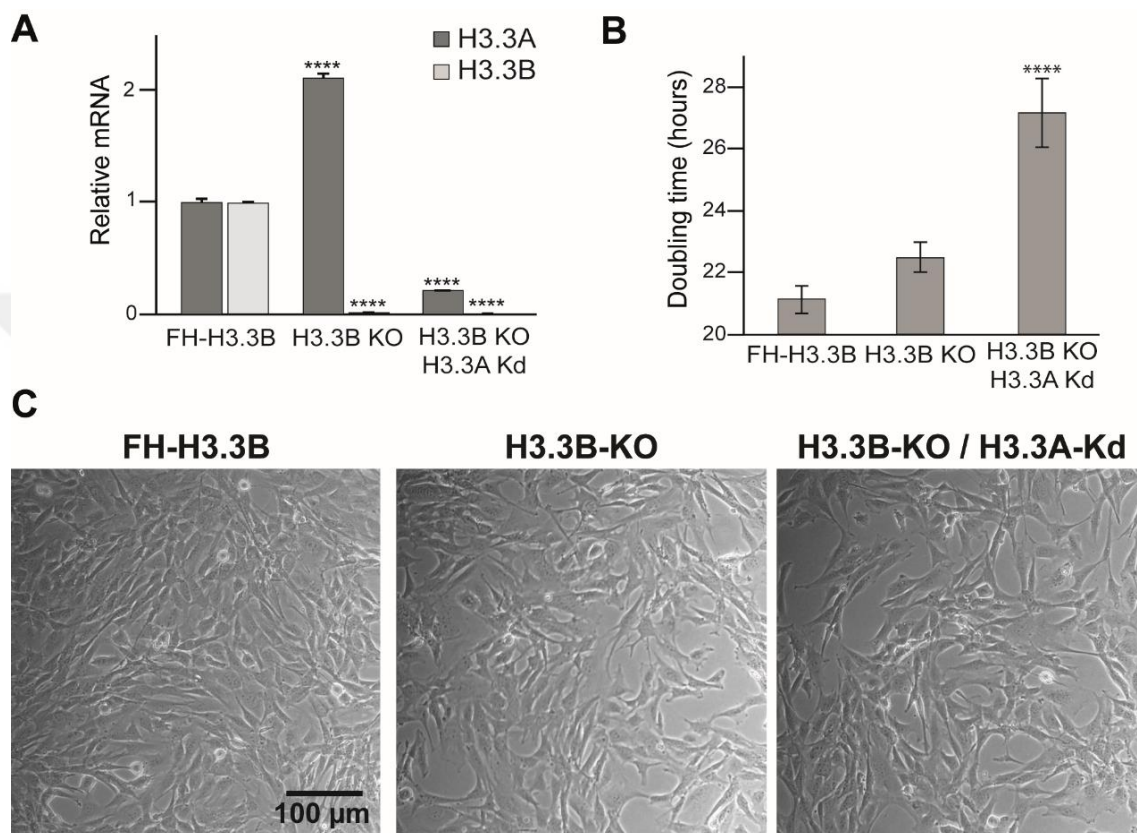


Figure 3-18. H3.3 expression and doubling times in H3.3B-KO / H3.3A-Kd MEF model. **A.** Relative H3.3 RNA expression in MEFs. Quantitative RT-qPCR mRNA profiles normalized to ribosomal protein S9 (Rps9) mRNA and relative to the values obtained for the control FH-H3.3B MEFs set at $1.0 \pm \text{s.e.m.}$ One-way ANOVA, $n=4$, $P < 0.001$ for FH-H3.3B vs. H3.3B-KO and $P < 0.01$ for FH-H3.3B vs. H3.3B-KO / H3.3A-Kd. **B.** Average doubling times in MEFs. 25,000 cells of each cell type were plated and counted at different times to calculate the doubling time. Error bars represent standard deviation, one-way ANOVA, $P < 0.001$ for FH-H3.3B vs. H3.3B-KO / H3.3A-Kd, $n=4$. **C.** Representative images of FH-H3.3B, H3.3B-KO and H3.3B-KO / H3.3A-Kd MEFs. Equal numbers of cells were seeded for each group at the beginning. A and B reproduced from Ors *et al.*²⁷⁹

3.3.3. H3.3 depletion in MEFs has some, yet minimal effect on the transcriptome

Total RNA samples from biological replicates of H3.3B-KO cells transduced with shControl or shH3.3A were sequenced and their transcriptomes compared. As verified by RT-qPCR experiments (Figure 3-18 A), H3.3A expression was significantly reduced in the H3.3B-KO/H3.3A-Kd MEFs cells (Figure 3-19 A). RNA-Seq analysis showed that a substantial part of the transcriptome did not vary when H3.3 was depleted. Only about 4% of the transcribed genome (800 genes) showed significant deregulation, with 449 significantly upregulated and 401 significantly downregulated genes (Figure 3-19 B).

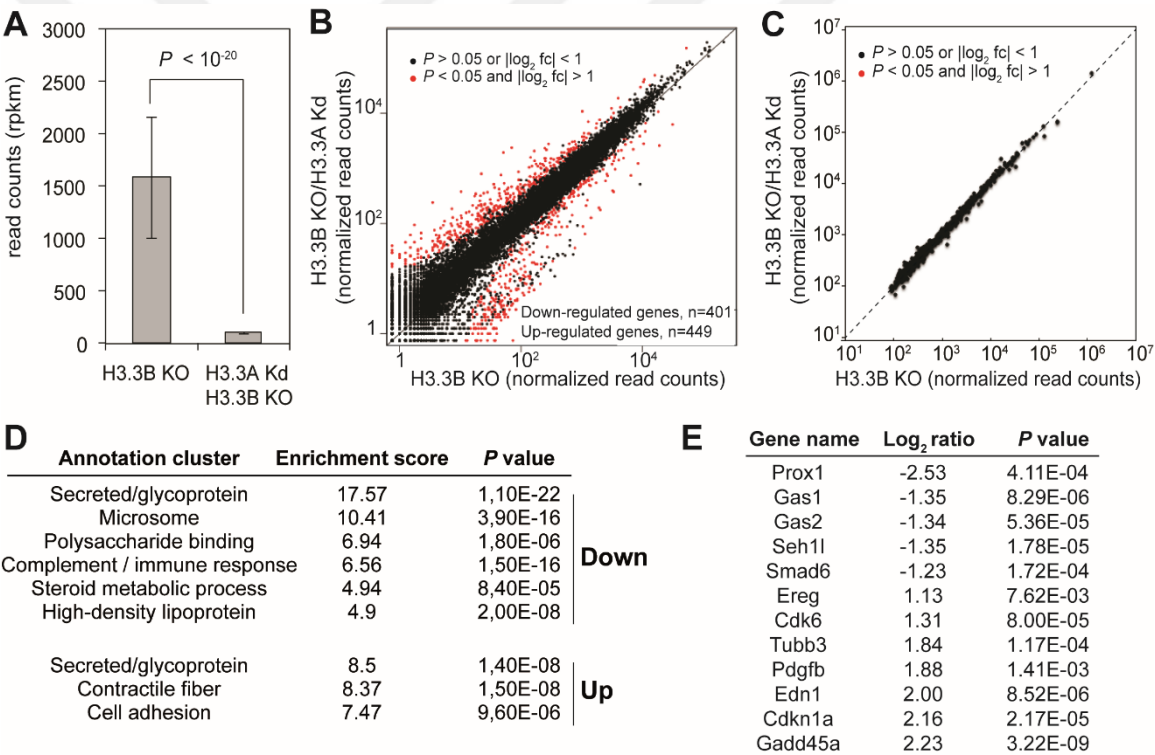


Figure 3-19. Effect of H3.3 loss on MEF transcriptome. **A.** Bar graph representing H3.3A expression in control H3.3B-KO and H3.3B-KO / H3.3A-Kd MEFs. **B.** Scatter plots comparing gene expression profiles of control (shControl) and H3.3B-KO / H3.3A-Kd (shH3.3A) MEFs. Red dots indicate differentially expressed genes ($P < 0.05$ and $|\log_2 \text{fc}| > 1$). **C.** Scatter plots comparing global transcription of repeat families in H3.3B-KO and H3.3B-KO / H3.3A-Kd MEFs. **D.** Functional annotation clustering of differentially expressed genes. **E.** List of the differentially expressed genes implicated in mitosis. Data produced in collaboration with Dr. Christophe Papin. Reproduced from Ors *et al.*²⁷⁹

A recent study showed an implication for H3.3 in the maintenance of the silent state of endogenous retroviruses (ERVs) ²²⁴. It would therefore be expected for repeat transcripts to be affected in the absence of H3.3. However, our analysis yielded in no significant differences in global transcription of DNA repeats such as retro-elements (including long terminal repeats (LTR), ERV, long (LINE) and short (SINE) interspersed nuclear elements) or tandem repeats (including major satellites, telomeres and microsatellites), which account for more than 40% of the mouse genome (Figure 3-19 C).

Functional clustering of differentially expressed genes indicated significant downregulation of genes implicated in lipid and sterol processing while factors involved in cell adhesion and motility were upregulated (Figure 3-19 D). Among the differentially regulated genes there were also a number of genes implicated in cell cycle progression (Figure 3-19 E).

Over the past few years, two other groups published studies related to the role of H3.3 in transcription ^{187,188,285}. Similarly to ours, these studies showed minimal impact on global transcription levels upon H3.3 depletion. Interestingly, when expression profiles of H3.3 null embryos in a p53-null background ¹⁸⁷ and H3.3B-KO / H3.3A-Kd ESCs ²⁸⁵ were compared to the expression profiles obtained in H3.3B-KO / H3.3A-Kd MEFs, there was little to no overlap between the three datasets. Essentially, there were only 9 deregulated genes in common with H3.3 depleted embryos and 6 deregulated genes in common with H3.3 depleted ESCs (Figure 3-20).

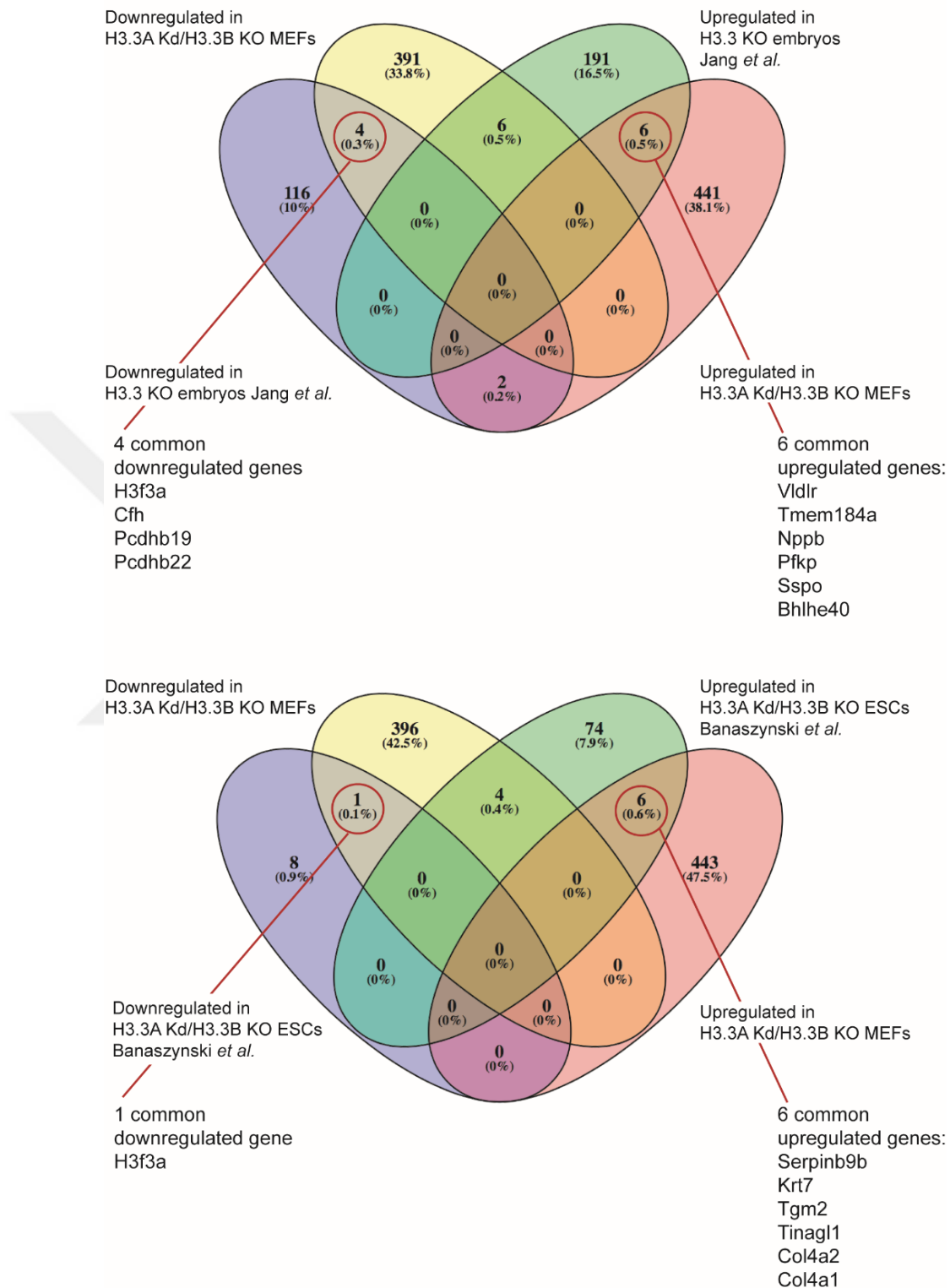


Figure 3-20. Comparative analysis of the differentially expressed genes in H3.3 deficient embryos and in H3.3 depleted ESCs. Differentially expressed genes in H3.3-null embryos¹⁸⁷ and H3.3-deficient ESCs²⁸⁵ were compared to analyze overlap between affected transcripts. Lists of genes that show similar patterns of up or downregulation in transcription are indicated. Data produced in collaboration with Dr. Christophe Papin. Reproduced from Ors *et al.*²⁷⁹

In summary, the knockout of only one of the genes coding for H3.3 had no effect on the liver transcriptome. The increase of H3.3 expression from the remaining H3.3 coding gene likely compensated for the single gene knock-out. It was therefore necessary to generate a model where both of the genes could be efficiently silenced. Although double H3.3 coding gene knock-out mouse models proved impossible to generate, mouse embryonic fibroblasts were effectively derived and immortalized from FH-H3.3 mice. These MEFs served to develop an effective cell model where most H3.3 expression was successfully suppressed. Even though the developed knock-out/knock-down strategy effectively depleted H3.3 in MEFs, the effect of loss of H3.3 was minimal on global transcription and non-existent for transcription of repeat families. However, the H3.3B-KO / H3.3A-Kd cells showed significant proliferation deficiency when compared to control FH-H3.3B cells. Additionally, genes involved in cell cycle were found deregulated in H3.3 deficient MEFs. Thus, it became of interest to focus studies in determining the role of H3.3 in mitotic progression.

3.4. H3.3 involvement in mitotic progression

The decrease in proliferation observed in H3.3 deficient MEFs lead us to study the role of H3.3 during mitosis.

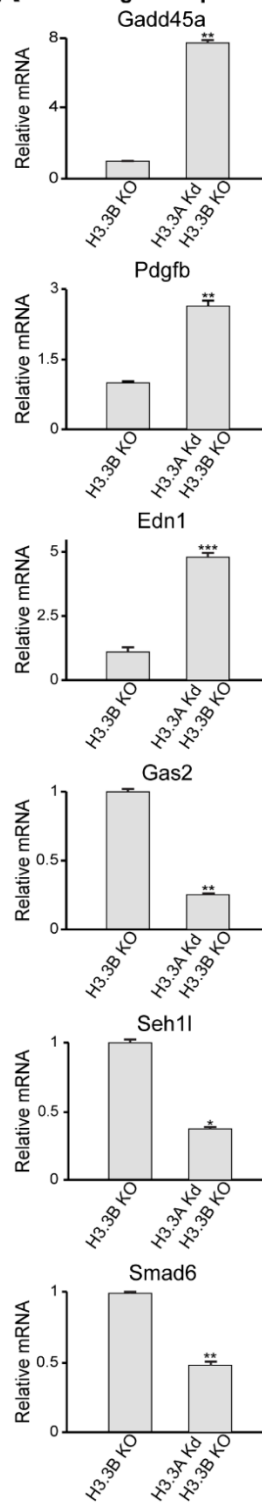
3.4.1. H3.3 and H2A.Z are present at TSS of deregulated genes involved in cell cycle

Among the deregulated genes from the RNA-Seq study, a sub-group of genes was implicated in cell cycle progression. Six of these genes were selected and the difference in gene expression levels between H3.3B-KO and H3.3B-KO / H3.3A-Kd MEFs was validated by RT-qPCR (Figure 3-21 A). The genes selected mostly showed a 2-3-fold change except for *Gadd45a* which showed a nearly 7 times increase in H3.3B-KO / H3.3A-Kd cells compared to H3.3B-KO cells. A native ChIP experiment with an anti-HA antibody was performed on chromatin isolated from FH-H3.3B MEFs to determine H3.3 localization on the studied genes. All of the candidate genes showed a high H3.3 enrichment relative to input at their TSS regions (Figure 3-21 B). Chromatin isolated from WT MEFs expressing no HA epitope was used as negative control and presented no enrichment. These results suggest that even though H3.3 localization is thought to be heavily implicated in transcription, when removed transcription of target genes is only mildly affected.

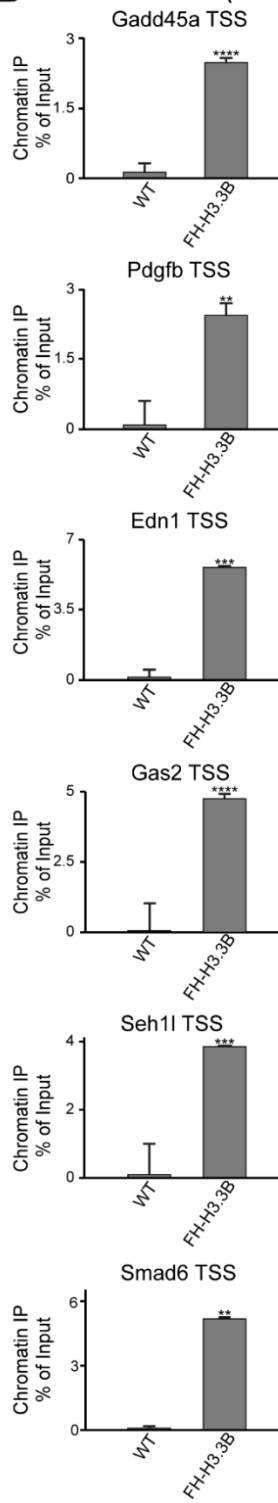
3.4.2. H3.3 depletion does not affect H2A.Z enrichment at promoter regions

The presence of dual H3.3-H2A.Z variant containing nucleosomes at promoter regions is thought to positively regulate transcription^{60,131,159,231}. The effect of H3.3 loss on H2A.Z localization at promoter regions was studied by performing cross-linked ChIP experiment in WT, H3.3B-KO and H3.3B-KO / H3.3A-Kd MEFs. Mock IgG was used as negative control for the IP. Regardless of the transcriptional deregulation observed with the examined genes, H2A.Z enrichment at their TSS did not differ in single H3.3B-KO or H3.3 depleted H3.3B-KO / H3.3A-Kd MEFs when compared to H2A.Z enrichment in WT MEFs (Figure 3-21 C).

A Relative gene expression



B FH-H3.3B ChIP (α-HA)



C H2A.Z ChIP

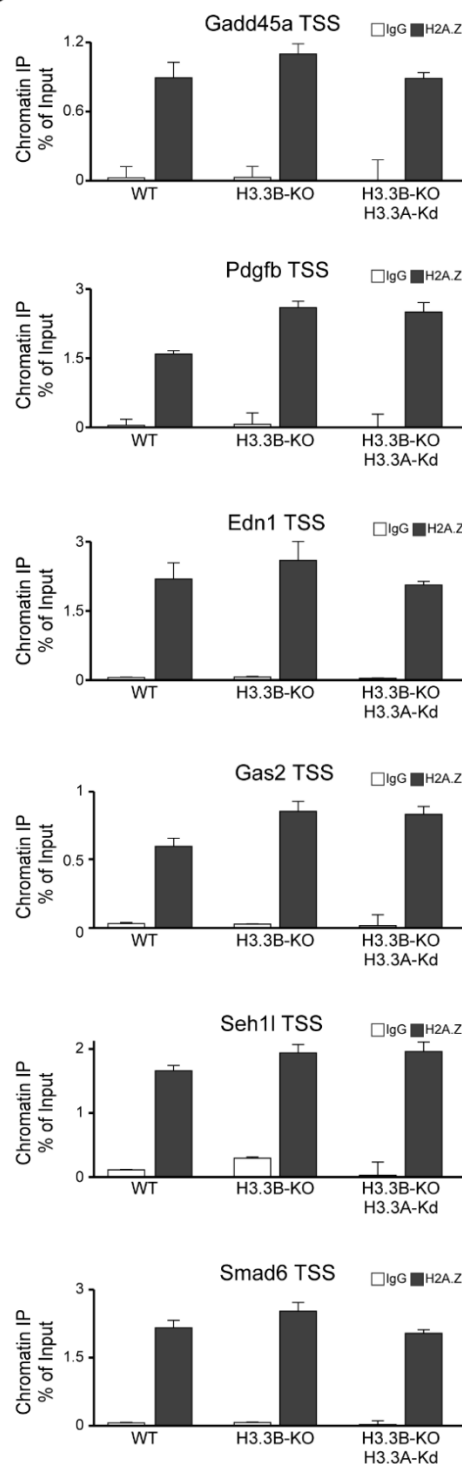


Figure 3-21. H3.3 and H2A.Z enrichment at transcription start sites (TSS) of mitotic genes showing differential mRNA expression. A. Quantitative PCR assay for gene expression in H3.3B-KO and H3.3B-KO / H3.3A-Kd MEFs. RNA values are normalized to ribosomal protein S9 (Rps9) mRNA and are relative to H3.3B-KO value set at $1.0 \pm \text{s.e.m}$, two-tailed t-test, $P < 0.05$ for *Sehl1*, $P < 0.01$ for *Gadd45a*, *Pdgfb*, *Gas2*, *Smad 6*, $P < 0.001$ for *Edn1*, $n=4$ **B.** H3.3B enrichment at TSS of differentially transcribed genes by ChIP-quantitative PCR assay (ChIP-qPCR). Values are expressed in % of enrichment relative to input $\pm \text{s.e.m}$. Chromatin from wild-type (WT) MEFs containing no HA-tag were used as a negative IP control. The TSS of examined genes were significantly enriched in H3.3B compared to the negative control, one-tailed t-test, $n=4$, $P < 0.01$ for *Pdgfb* and *Smad6*, $P < 0.001$ for *Edn1* and *Sehl1*, and $P < 0.0001$ for *Gadd45a* and *Gas2*. **C.** H2A.Z enrichment at TSS of differentially transcribed genes. ChIP-qPCR assay. Values are expressed in enrichment relative to input $\pm \text{s.e.m}$. Rabbit IgG was used as the control. The TSS of examined genes were significantly enriched in H2A.Z compared to the negative control one-tailed t-test, $P < 0.05$, $n=3$. ChIP-qPCR graphs are representative of 3 separate experiments. H2A.Z enrichment at studied TSS did not vary significantly between samples (One-way ANOVA test, $P > 0.05$, $n=3$). Taken from Ors *et al.*²⁷⁹

3.4.3. H3.3 depletion results in defective mitotic progression

In addition to the increase in cell doubling time, H3.3B-KO / H3.3A-Kd cells exhibited increased number of defects in their nuclear structure. Using immunofluorescent staining of nuclear envelope protein lamin and DNA, nuclear and mitotic defects were quantified (Figure 3-22 A to C). While even the knock-out of the single H3.3B coding gene led to a doubling of nuclear defects observed at interphase, H3.3B-KO / H3.3A-Kd cells had nearly 3 times the number of micronuclei or polylobed nuclei compared to control FH-H3.3B cells (Figure 3-22 B). In addition to interphase defects, H3.3 depleted cells displayed significantly higher percentages of mitotic defects such as chromosome misalignment at metaphase plate, lagging chromosomes and chromatin bridges during early and late anaphase (Figure 3-22 C). Together these defects accounted for a clear chromosome structure dysregulation. Furthermore, flow cytometric analysis of cell cycle of WT, FH-H3.3B, H3.3B-KO and H3.3B-KO / H3.3A-Kd MEFs, showed a near 50% decrease in number of cells in the G2/M phase of the cell cycle in H3.3 depleted cells with an increased time of residence in G0/G1 phase (Figure 3-22 D, Appendix A).

In summary, although both H3.3 and H2A.Z are present around the TSS of several genes implicated in cell cycle progression, near complete depletion of H3.3 does not have the expected dramatic effect on their gene transcription. Furthermore, H2A.Z enrichment is not impacted by this depletion. However, the decrease in cell doubling time coincident with the increase of aberrant nuclear and mitotic chromosome organization in H3.3 depleted cells underlines an important yet under-investigated role for H3.3 in chromosome segregation, nuclear structure and genome integrity.

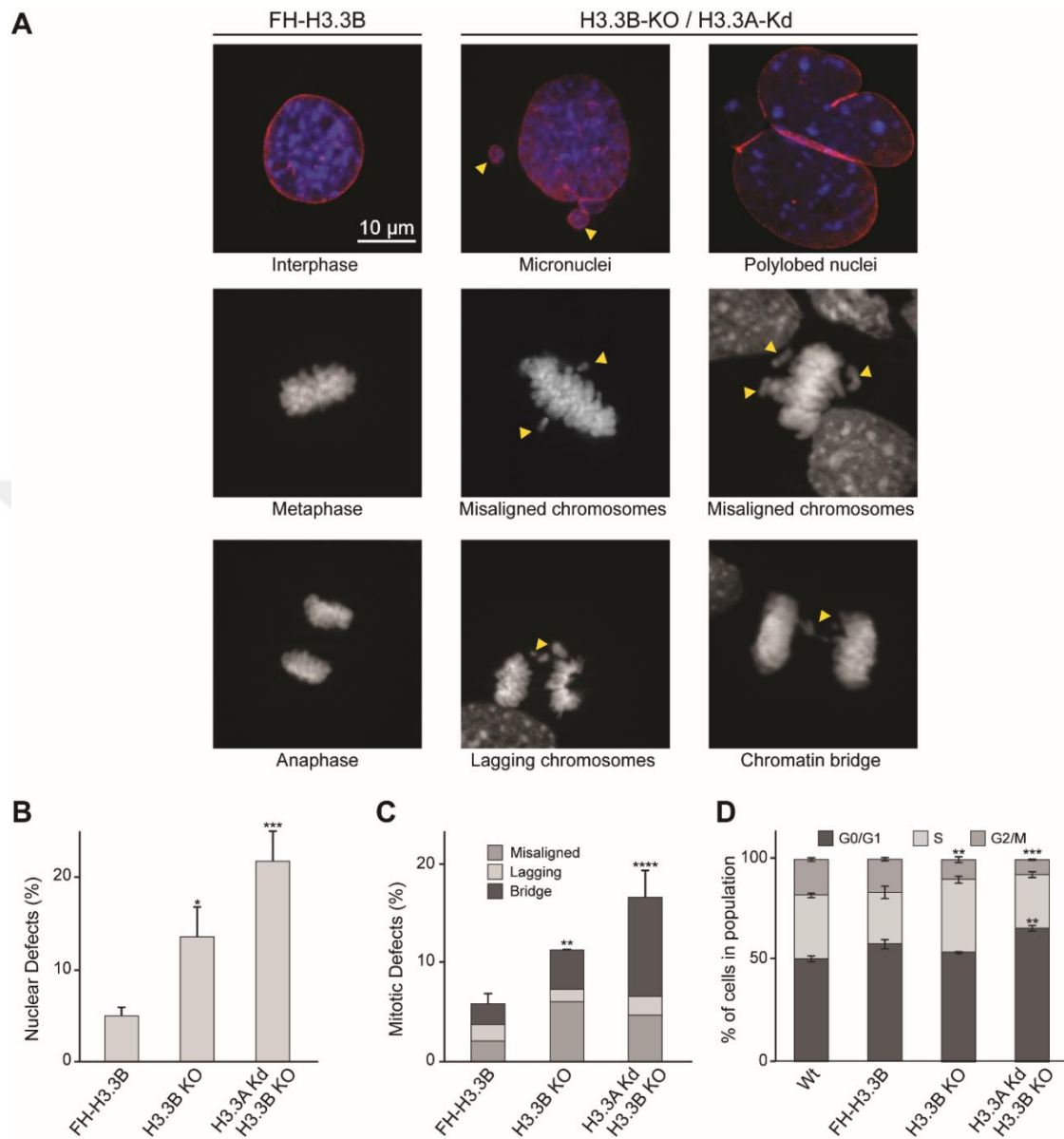


Figure 3-22. Mitotic defects in H3.3 deficient MEFs. **A.** Representative images of control FH-H3.3B nuclei and nuclear abnormalities with micronuclei and polylobed nuclei as well as mitotic abnormalities with chromatin bridges, misaligned and lagging chromosomes observed in H3.3 deficient MEFs. Quantification of **B.** nuclear and **C.** mitotic abnormalities in control and H3.3 deficient MEFs. Numbers are representative of 3 separate experiments. Hundred twenty-five mitotic events were scored for each line. **D.** Cell cycle analysis by Flow cytometry in wildtype (WT), FH-H3.3B cells, H3.3B-KO and double H3.3B-KO / H3.3A-Kd cells treated with propidium iodide (PI). Figure reproduced from Ors *et al.*²⁷⁹

Chapter 4. Discussion

Chromatin is the form of structural organization of genetic information in eukaryotic nuclei. The basic repeating unit of chromatin is the nucleosome that consists of a double strand DNA molecule wrapped around an octamer of histones ^{1,2}. Although initially, the sole function of chromatin was believed to be the compaction of the large DNA molecule into the small nuclear volume, over the years, scientific studies have uncovered that chromatin, through its restructuration, is involved in all cellular processes that require access to DNA, including replication, transcription, repair and cell division. To allow for these processes to progress properly, chromatin needs to be highly dynamic and flexible. In the heart of the main mechanisms for introducing such variation, lies histones, the very basic small proteins that constitute the core octamer of nucleosomes ⁵⁹. Through their post-translational modification or their replacement by variants, histones are central actors in epigenetic regulation ⁶¹.

Histone variants are isoforms of their replication dependent (RD) counterparts, that differ in their gene locus, structure and expression as well as in their time and site of incorporation into chromatin. Their incorporation into chromatin gives way for new functional and structural properties. H3.3 is an evolutionarily well conserved histone H3 variant that differs by only 4 or 5 amino-acids from the RD histones H3.1 or H3.2. The differing amino-acids mainly lie in the globular domain of the protein and are implicated in its recognition by specific chaperones HIRA and DAXX/ATRX that allow for non-RD incorporation into chromatin. The only specific amino-acid that's found in the N-terminal tail of H3.3 and is accessible in the context of the nucleosome is the serine residue at position 31 that also constitutes an additional phosphorylation site. H3.3 is coded by two genes *H3f3a* and *H3f3b*, found in two different loci and differing in their nucleotide sequences but coding for the same amino-acid sequence.

Mutations in histone variants or changes in their expression have been implicated in numerous diseases including cancer. Discovering the precise mechanisms of action and functions of these variants is crucial in understanding the development of a diseased state and providing elements for detection and treatment. The work presented in this study aims to uncover the functional role of H3.3 in different cellular processes.

As H3.3 has a replication independent and heterogenous incorporation into chromatin, determining the genome-wide enrichment map of H3.3 would provide a basis to hypothesize on its function. Chromatin immunoprecipitation coupled to next generation deep sequencing constitutes a very potent tool to determine precisely where specific proteins bind to the genome. However, the few amino-acid difference between H3.3 and RD H3 variants constitutes an obstacle for the production of specific chromatin immunoprecipitation (ChIP) grade antibodies needed for accurate genome-wide mapping of H3.3. The existing mapping studies relied either on the ectopic expression of H3.3^{159,219,220,222} or the replacement of only one of the alleles coding for *H3f3b* by a tagged H3.3 protein²⁰⁵ in cultured cell lines. There were no studies targeting all four H3.3 coding alleles or conducted in mouse models expressing tagged H3.3 proteins.

Additionally, in all of these studies, the tag was placed at C-terminal of H3.3 and although initial investigations on the use of these tags reported no appreciable differences in incorporation between the tagged and untagged H3.3¹⁶, the central location of the C-terminal region of H3.3 in the nucleosome¹ could alter the structure of the nucleosomes and affect its function. Also, a tag at the C-terminal tail could be less accessible to antibodies favoring off-target binding to histones not incorporated into nucleosomes.

In our study, these problems were addressed by the generation of novel knock-in/conditional knock-out mouse lines producing either a H3.3A or a H3.3B protein with an N-terminal FLAG-FLAG-HA (FH-H3.3). Additionally, the tagged protein could be conditionally inactivated by Cre recombinase expression, resulting in the functional knock-out of H3.3A or H3.3B. As the N-terminal tail of H3.3 protrudes outside of the nucleosome, it's unlikely to alter nucleosome core structure¹. Furthermore, the common modifications taking place on the N-terminal tail remain distant from the tag and are once again believed to be not affected by it¹⁶⁶. This study showed that FH-H3.3A or FH-H3.3B producing mice were viable, showed no observable differences from their wild-type littermates and the FH-H3.3 protein produced in these mice was efficiently incorporated into chromatin. Surprisingly however, efforts to produce mice tagged at all four H3.3 coding genes were fruitless. As a result, it is not possible to ascertain whether the N-terminal tag on H3.3 is completely inconsequential to the structure of the nucleosome. Nonetheless, the mouse models established in this study constitute invaluable tools for the specific study of H3.3 incorporation dynamics and function in mice.

Previous genome-wide localization studies showed that H3.3B was enriched near transcription start sites as well as pericentromeric sites, telomeres and retroviral elements^{201,205,216,219,220,224}. Thanks to the mouse models generated in our study, the relative contribution of H3.3A and H3.3B to H3.3 incorporation in the adult mouse liver genome could be determined.

Combining a high-resolution native chromatin immunoprecipitation technique with high precision paired-end deep sequencing, genome-wide H3.3A and H3.3B distribution was mapped in the adult mouse liver. Similar H3.3 enrichment profiles were observed between H3.3A and H3.3B around transcription start and termination sites (TSS and TTS). H3.3 was enriched up to three nucleosomes upstream and five nucleosomes downstream of TSS and to a lower degree at TTS. Furthermore, H3.3 enrichment positively correlated with high gene expression levels and markers of active transcription such as Pol II, histone marks associated with active transcription. No H3.3 was present at the TSS itself. H3.3 has been described to have a high turnover rate at promoter regions^{213,215}. Such high levels of nucleosome exchange at promoter regions would lead to transient exposition of transcription factor binding sites on DNA facilitating transcription. Furthermore, H2A.Z showed enrichment patterns similar to H3.3 at these regions even though H2A.Z did not extend as far as H3.3 downstream of TSS. These two histone variants are colocalized at nucleosomes of promoter regions to form less stable but dynamic nucleosome structures that can be easily replaced by transcription machinery¹⁵⁹. Taken together, this data points to a functional role for H3.3 in active transcription. As a result, upon deletion of H3.3, a downregulation of highly expressed genes would be expected.

To study the role of H3.3 in transcription, RNA-Seq on total RNA was used to determine the effect of single H3.3 knock-out in mouse livers on the global transcriptome. Although a clear positive correlation between H3.3 enrichment at promoters and expression levels was observed, the deletion of only one of the genes coding for H3.3 in the organism produced no significant deregulation in global gene expression in the liver. As both H3.3 genes code for the exact same protein with a basally equal level of expression in the liver and since the depletion of one H3.3 coding gene leads to an increase of expression of the other, it is likely that the product of one gene can compensate for the loss of the other, indicating an imperative need to produce models where both H3.3 coding genes are suppressed. Unfortunately, double knock-out mice were not viable and transcriptomic

analyses could not be undertaken. Instead, a new cellular model essentially depleted of all H3.3 expression was developed using transformed mouse embryonic fibroblasts isolated from FH-H3.3 mice, combined with an RNA interference strategy. As these cells were transformed using the serial passage method ²⁷⁷, they contain multiple genetic mutations involved in the regulation pathways of cell survival, apoptosis and proliferation which promote cell-growth. However, these mutations are likely to be heterogenous and the precise pathway targeted in the immortalized MEFs is not known. For the sake of scientific reproducibility and accuracy, it is therefore important to use simultaneously developed cells and compare the knock-out MEFs to their parent MEFs. In the present study, FH-H3.3B MEFs were used to produce H3.3B-KO MEFs which in turn were used to knock-down H3.3A expression. Since, H3.3B-KO MEFs had an almost 2-fold increase of the *H3f3a* mRNA compared to FH-H3.3B MEFs it is likely that H3.3A compensates for the loss of H3.3B as observed in the livers. Therefore, in this study, the transcriptomes of H3.3B-KO MEFs to H3.3B-KO/H3.3A-Kd MEFs were compared to determine the effect of complete H3.3 depletion on transcription in MEFs.

Surprisingly, regardless of a small subset of deregulated genes, the overall effect of H3.3 depletion on the transcriptome of MEFs remained restricted. Also, there were as many upregulated genes as downregulated genes questioning the role of H3.3 as an activator of transcription. ESCs where H3.3 expression was depleted with a combined knock-out/knock-down strategy similar to the present study, remained functionally pluripotent ²⁸⁵ further casting doubt on the role of H3.3 on active transcription. Moreover, even though such unimpressive effects on transcription were observed in other studies using H3.3 depleted embryos ¹⁸⁷ or ESCs ²⁸⁵, there was no overlap between the sets of affected transcripts. The MEFs used in our study were isolated from embryos at 13.5 dpf and the embryos used in the above-mentioned study were at 10.5 dpf and of p53-null background ¹⁸⁷. Thus, this observation could in part be explained by the fact that transcription profiles vary according to tissue types and developmental stages. However, even though the samples used in these studies came from different models at different developmental stages, the lack of complete overlap together with the limited global effect of H3.3 depletions, challenges the idea of a prominent role for H3.3 in transcription.

Aside from the H3.3 specific serine residue at position 31, H3.3 is modified at the same sites as the replication dependent H3.2 or H3.1. Notwithstanding, H3.3 is more enriched in

histone marks of active transcription such as K9, K14, K18 and K37 acetylation and K36, K79 methylations^{166,225,226}. Additionally, H3.3 is deprived of histone marks associated with repression of transcription such as K27 methylation. However, in the absence of H3.3, H3.1 or H3.2 could be differentially modified through compensatory mechanisms to minimize the drastic effect of the loss of H3.3 modifications on chromatin. It is also possible that even though H3.3 expression was essentially depleted in MEFs, the remaining activity is sufficient to partially maintain the transcriptional function of H3.3. In this case, the minimum levels of H3.3 required to maintain transcriptional function would be lower than those required to maintain genomic stability linked functions. In addition, uncharacterized H3 genes, in most part similar in sequence and distribution to H3.3, were recently identified in the mouse genome²⁸⁶. The expression of such genes may account for some compensation for H3.3 depletion in our models.

Additionally, although H3.3 and H2A.Z were both found enriched at the promoter regions of selected genes, the depletion of H3.3 had no effect on H2A.Z presence at these regions. As other studies have described these two variants to co-exist in the same nucleosome at promoter regions and play an important role in promoter structure¹⁵⁹, it would be interesting to investigate the effect of H2A.Z depletion at these same sites on H3.3 incorporation and gene transcription to evaluate whether it's actually the H2A.Z variant that is indispensable for transcription regulation and H3.3 enrichment. In the previously mentioned study in ESCs²⁸⁵, nucleosome turnover rates decrease upon H3.3 depletion, adding to the hypothesis that double H3.3/H2A.Z variant containing nucleosomes allowed the establishment of a chromatin structure favorable to dynamic exchange of effectors¹⁵⁹. Yet it appeared, that abolishing H3.3 from these structures was not sufficient to severely affect transcription. Corroborating existing research on H3.3 depletion in ESCs and mouse embryos^{187,188,204,285}, our data suggested that H3.3 plays only a facilitating role in the regulation of transcription in adult livers.

Although H3.3A and H3.3B show similar enrichment profiles at various genomic sites in the adult mouse liver, they appear to differ at some endogenous retroviral repetitive sequences. As these elements constitute a high proportion of the mouse genome and may contain regulatory sequences that could affect expression of nearby genes, they are usually silenced and restricted to heterochromatin^{240,241}. The preferential enrichment of either one of the H3.3 coding genes might account for the chronological order of deposition into

chromatin during development that would persist in differentiated liver tissue as H3.3 turnover in heterochromatic sites is very slow^{213–215}. Given the preferential enrichment at certain repeats and the recent implication of H3.3 in the silencing of retroviral elements in ESCs²²⁴, we expected to observe an upregulation of retroviral transcripts in H3.3 depleted MEFs and even in single knock-out livers. However, in both models, an impact on ERV transcription could not be observed. The role for H3.3 in silencing ERVs might therefore be restricted to the pluripotent cell state and be replaced by other mechanisms, namely H3K9me3, H4K20me3 modifications and DNA methylation^{226,242,287}, in differentiated cells. Although H3.3K9me3 has been linked to heterochromatin formation in telomeres²⁴², a coincident study in H3K9me3 mediated ERV silencing in ESCs, attributed this function to DAXX/ATRX interaction with methyltransferases independently of H3.3 incorporation suggesting that H3.3 is not essential to maintain heterochromatin state mediated by this type of modification at all ERVs^{288,289}.

On the other hand, H3.3 implication in the maintenance of genomic integrity was clearly observed in this study as H3.3 depleted MEFs showed a severe increase in mitotic defects such as misaligned or lagging chromosomes and chromatin bridges. As a result, the nuclear matrix structure of interphase cells was dramatically altered. This was demonstrated by the coincident increase in micro and polylobed nuclei in H3.3-depleted cells. In a physiologically normal cell, such mitotic defects would result in the activation of the spindle attachment checkpoint (SAC) and consequent cell cycle arrest through tumor suppressor pathways such as the p53 pathway. Evading such growth suppressor mechanisms is a hallmark of cancer and cancer cells often present aneuploidy resulting from mitotic defects and chromosome instability²⁹⁰ as observed in H3.3 depleted cells. A similar study reporting the same type of mitotic defects was conducted in H3.3 depleted MEFs in a p53-null background suggesting that p53 mediated cell cycle arrest and apoptosis can explain the lethality of non-transformed H3.3 depleted cells¹⁸⁷. Of note, the precise pathway targeted in the transformed MEFs used in our study is not known. Nonetheless, a deregulation of genes involved in cell-cycle regulation was observed in transformed H3.3 depleted cells including p53 targets: *Cdkn1a* which codes for p21 involved in growth arrest and *Gadd45a* involved in DNA repair²⁹¹. While p53 can be activated during mitosis in response to DNA damage or prolonged stay in mitosis^{292–294}, it was recently shown that chromosome missegregation can also trigger p53 mediated cell cycle arrest through the

extension of H3.3 specific S31 phosphorylation from pericentromeric chromatin to the rest of the chromosome, providing a lead on the mechanism behind H3.3 implication in these observations ²⁴⁹. The close proximity of the phosphorylation sites on the N-terminal tail of H3.3 (S10, S28, S31) to other common modification sites such as the K9, K27 and K36 residues also suggest a possible mechanism through alteration of signals conveyed by the modification of these residues. This line of research is further supported by the recent discovery of point-mutations in the coding DNA sequence (cds) of H3.3 coding genes resulting in K27M/I, G34R/W/L and K36M substitutions, in various cancers ^{249,274}.

In addition to pericentromeric chromatin, H3.3 was described to be enriched at telomeric repeats ^{205,295}. Our study found that H3.3A and H3.3B were similarly enriched at telomeres in livers as well as in MEFs. In H3.3 depleted ESCs and MEFs, strong abnormalities in telomere structure and maintenance were observed ^{187,188,285}.

Thus, our data combined with current state of literature, highlight a very important role for H3.3 at maintaining genomic stability at pericentromeric and telomeric chromatin.

The importance of having two genes that code for the same protein is inferred by the conservation of this property in mammals and birds ¹³⁹. This redundancy may be the result of the higher demand for histones in larger genomes as it is the case for the replication-dependent histones. However, the preferential expression of H3.3 coding genes at different developmental stages and tissue types points to a function specific regulation of their expression. Given the divergence in their non-coding sequences, this differential regulation could take place both at the transcriptional and post-transcriptional levels. The promoter of *H3f3a* contains elements characteristic of a basally expressed gene while the promoter of *H3f3b* contains responsive elements indicative of an inducible gene¹⁴. Furthermore, their 3'UTR regions contain ARE elements that could influence post-transcriptional mRNA stability ^{194,196}. The family of AUF1 proteins that interact with these ARE elements are known to have tissue specific expression ¹⁹⁷. Over the past decade, an increasing number of non-coding RNAs were identified to play an important role in regulating protein expression. Additionally, many of these non-coding RNAs are tissue-specific ⁶⁹⁻⁷¹. Recently, AUF1 proteins were also shown to interact with non-coding RNAs ¹⁹⁸. The two H3.3 coding genes also differ in their 5' UTR and intronic sequences and the different secondary structures formed by the 5'UTRs could be interacting with different RNA binding proteins. In light

of recent discoveries in non-coding RNA mediated regulation mechanisms, the regulatory sequences of the two H3.3 coding genes should be analyzed for interaction with non-coding RNAs with the aim of discovering new elements to explain their differential expression.

In contrast to other studies ^{186,188}, the single H3.3A-KO or H3.3B-KO mice produced from the FH-H3.3 mice in this study, were viable (as for Jang et al. ¹⁸⁷) and did not present any discernable phenotypes apart in fertility. In single knock-out mice and MEFs, an increase at the mRNA levels of the remaining transcript was observed suggesting a compensatory regulation between the two genes. In some tissues like the brain, transcripts from *H3f3a* account for the majority of H3.3 expression, suggesting different functional consequences of H3.3 coding gene depletion depending on the studied tissue. This difference can be illustrated by the fact that mutations found in *H3F3A* drive childhood brain tumors whereas mutations in *H3F3B* have been associated with chondroblastoma and bone tumors. However in the liver, both transcripts are present at similar levels ^{186,187} and the respective encoded protein products are enriched at similar levels at most genomic locations pointing to a similar function in adult liver tissue.

H3.3 implication in fertility was further substantiated by this study, as H3.3A-KO mice were sub-fertile, while H3.3B-KO mice were infertile. This effect has previously been reported in other H3.3 knock-out mouse models ^{186,188,191,251,296} except for the most recent by Jang et al. which reported no fertility defects ¹⁸⁷. While this discrepancy observed in growth deficiency and fertility across different H3.3-KO mouse models can be accounted for by the different genetic backgrounds ²⁹⁷, all of the studies where both H3.3 coding genes are inactivated, including ours, result in non-viable embryos during development underlining the fundamental importance of H3.3 in mouse embryogenesis.

In conclusion, the data presented in this study represent the first H3.3A or H3.3B specific genome-wide mapping study in the adult mouse liver at nucleosome resolution and provided additional understanding of H3.3 involvement in transcription and mitotic progression. Novel knock-in/ conditional knock-out mouse models were generated and a high resolution native chromatin immunoprecipitation method was optimized and combined to next generation paired-end sequencing to determine the relative enrichment patterns of FH-H3.3A and FH-H3.3BA in the mouse liver. Although, FH-H3.3A and FH-

H3.3B showed similar enrichment patterns at most genomic sites, in concordance with previous studies, a subset of endogenous retroviral repeat elements was differentially enriched. As the knock-out of both H3.3 coding genes was embryonically lethal, a new cell model combining mouse embryonic fibroblasts and RNA interference was developed. Regardless of a clear enrichment of H3.3 around TSS and a positive correlation with active transcription, RNA-Seq analysis showed that H3.3 depletion had only minimal effect on the global transcriptome and no effect on retroviral transcript levels. Instead, an increase in mitotic and nuclear abnormalities underlined an important role for H3.3 in the maintenance of genomic integrity during mitotic progression. Figure 4-1 depicts a graphical abstract of the main points addressed in this study.

ROLE OF HISTONE VARIANT H3.3 IN TRANSCRIPTION AND MITOTIC PROGRESSION

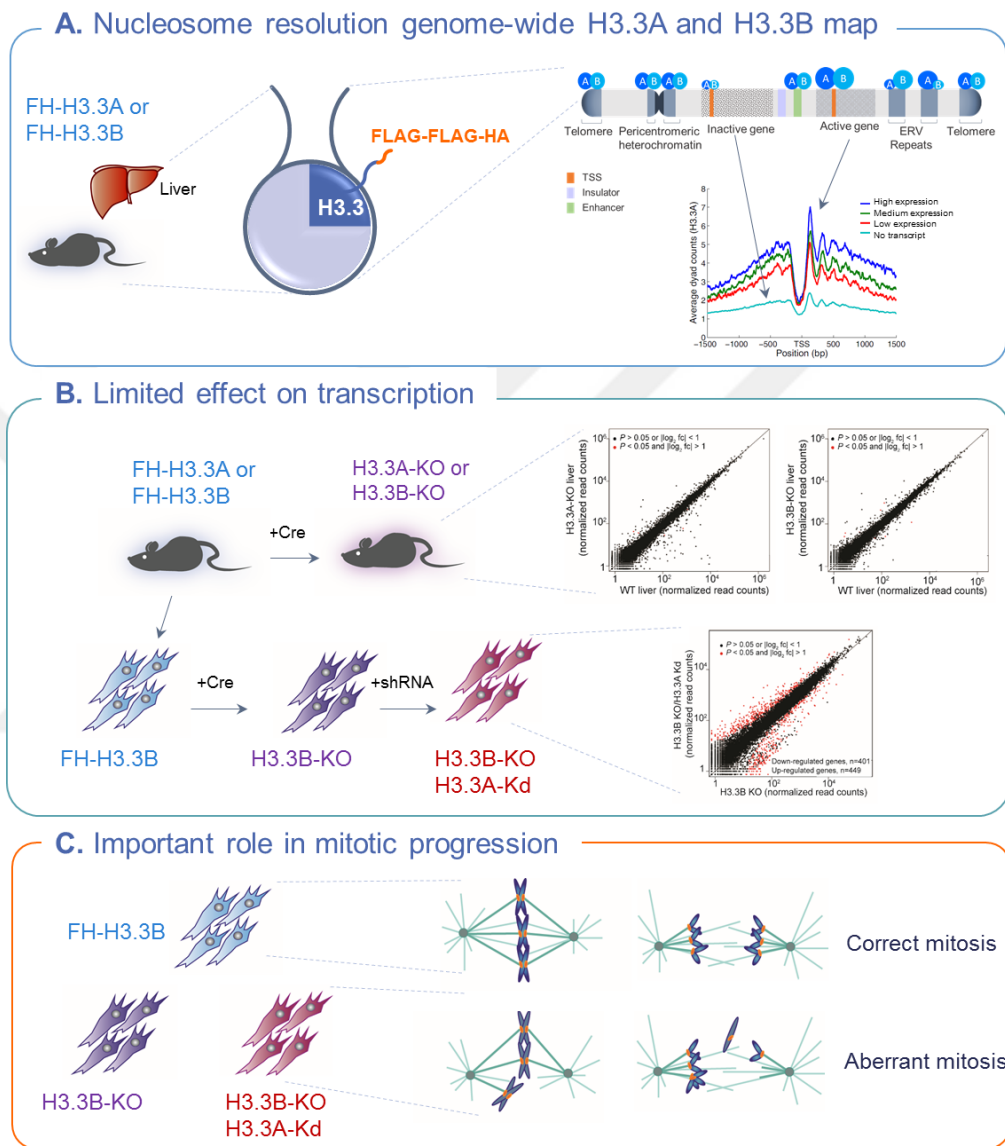


Figure 4-1. Graphical abstract of thesis. Main conclusions drawn from the study of H3.3 function in the context of this thesis. **A.** Genome-wide mapping of H3.3A and H3.3B in the liver was determined at nucleosome resolution using novel knock-in/ conditional knock-out mouse lines. H3.3A and H3.3B are equally enriched at telomeres, pericentromeric chromatin and around TSS of both active and inactive genes (Section 3.1 and 3.2). A preferential enrichment for H3.3A or H3.3B is observed at ERV repeat elements. H3.3 enrichment positively correlates with expression levels (Section 3.2). **B.** The depletion of a single H3.3 coding gene does not affect transcription levels in the adult mouse liver. The effect of complete H3.3 depletion in MEFs on transcription is higher than in livers all the while remaining limited as seen in the scatter plots (Section 3.3). **C.** H3.3B-KO and H3.3B-KO / H3.3A-Kd MEFs are characterized by an increase in mitotic abnormalities (Section 3.4).

Chapter 5. Perspectives

Complete H3.3 depletion in MEFs

This study relies on the use of a combined knock-out / knock-down approach to obtain H3.3 depleted cells. Although 90% of H3.3 expression is lost in these cells, the use of a model allowing a complete H3.3 depletion would provide results untainted with residual H3.3 expression. For this, an H3.3A knock-out mouse model developed by Tang et al.²⁵² was acquired and crossed with the FH-H3.3B mouse model. MEFs were isolated and transformed from these mice with the aim of inducing Cre mediated complete knock-out of H3.3.

Implication of the N-terminal tail of H3.3 and its phosphorylation in mitotic regulation

An important question that this study raises is: what is the precise mechanism by which H3.3 loss in mitotic progression is mediated? Previous work in Stefan Dimitrov's research group has shown an implication of phosphorylation of N-terminal histone tail in mediating function^{100,298}. The N-terminal tail of H3.3 has 3 phosphorylation sites. Two of them, S10 and S28 are common to replication-dependent H3 variants and have been shown to be phosphorylated during mitosis and involved in chromosome condensation. The third one is H3.3 specific and has been reported to be phosphorylated at pericentromeric chromatin during mitosis and implicated in p53 mediated cell-cycle arrest in response to chromosome missegregation^{202,248,250}. The implication of H3.3 tail phosphorylation can be addressed by expressing different exogenous molecular constructs where the N-terminal H3.3 phosphorylation sites are mutated into residues that cannot be phosphorylated, in H3.3 depleted MEFs and observing the effect on mitotic abnormalities. Preliminary results indicate that while the construct coding for a wildtype H3.3 establishes a normal mitotic progression, the expression of H3.3 constructs lacking all three phosphorylation sites or the integral N-terminal tail fail to rescue the abnormal phenotype to different extents. This work can help answer whether phosphorylation of the N-terminal tail of H3.3 is important for mitotic progression.

H3.3 implication in liver regeneration

As in all cancers, alteration of epigenetic regulation mechanisms is implicated in liver tumorigenesis ²⁹⁹. In cancers in general, early detection is crucial for higher life expectancy. However, liver cancer is generally diagnosed late and is resistant to most common cancer treatments ³⁰⁰. Eighty percent of all liver cancers are hepatocellular carcinomas (HCC). Most frequent reasons for HCC are hepatitis B and C infections, cirrhosis, alcohol consumption and chronic liver damage due to toxin exposure ³⁰⁰. When faced with such damages a healthy liver is able to activate its regeneration mechanism and renew itself ³⁰¹. Deregulation in liver regeneration pathways may lead to cancer. Elucidating liver development and damage response mechanisms will surely lead to new treatment methods.

Over the last years, research in the epigenetics field has increased and the role of epigenetic control mechanisms in the initiation and development of cancer has been described. Research has demonstrated that histone variants play important roles in some types of cancer. However, the role of histone variants in liver cancer has not yet been fully elucidated. The present study and others have linked H3.3 histone variant function to transcription as well as maintenance of genome stability ^{58,159,186,187,220,224,279,285,302}. There is no account for this variant's role in transformation of liver tissue.

The lethality of the knock-out of both genes coding for H3.3 in the whole organism constitutes a challenge for studying H3.3 implication in liver physiology. Targeted depletion of H3.3 can be established by crossing FH-H3.3 mice with TTR-Cre-ER mice ³⁰³. These mice express the Cre recombinase under a tamoxifen inducible liver specific transthyretin promoter. This model would allow the establishment of single and double H3.3 knock-outs specifically in the liver after development and study their implication in the dynamic process of liver regeneration providing great insight to the differential and combined roles of H3.3 coding proteins in liver physiology and transformation.

References

1. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251–260 (1997).
2. Alberts, B. *et al.* *Molecular Biology of the Cell*. (2008).
3. Harp, J. M., Hanson, B. L., Timm, D. E. & Bunick, G. J. Asymmetries in the nucleosome core particle at 2.5 Å resolution. *Acta Crystallogr. Sect. D* **56**, 1513–1534 (2000).
4. Ellis, R. J. Assembly chaperones: a perspective. *Philos. Trans. R. Soc. B Biol. Sci.* **368**, (2013).
5. Weber, C. M. & Henikoff, S. Histone variants: dynamic punctuation in transcription. *Genes Dev.* **28**, 672–82 (2014).
6. Kimmins, S. & Sassone-Corsi, P. Chromatin remodelling and epigenetic features of germ cells. *Nature* **434**, 583–589 (2005).
7. Arents, G., Burlingame, R. W., Wangt, B.-C., Love, W. E. & Moudrianakis, E. N. The nucleosomal core histone octamer at 3.1 Å resolution: A tripartite protein assembly and a left-handed superhelix. *Biochemistry* **88**, 10148–10152 (1991).
8. Arents, G. & Moudrianakis, E. N. The histone fold: a ubiquitous architectural motif utilized in DNA compaction and protein dimerization. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 11170–4 (1995).
9. Happel, N. & Doenecke, D. Histone H1 and its isoforms: Contribution to chromatin structure and function. *Gene* **431**, 1–12 (2009).
10. Marzluff, W. F., Gongidi, P., Woods, K. R., Jin, J. & Maltais, L. J. The human and mouse replication-dependent histone genes. *Genomics* **80**, 487–498 (2002).
11. Jaeger, S., Barends, S., Giegé, R., Eriani, G. & Martin, F. Expression of metazoan replication-dependent histone genes. *Biochimie* **87**, 827–834 (2005).
12. Osley, M. A. The regulation of histone synthesis in the cell cycle. *Annu. Rev. Biochem.* **60**, 827–61 (1991).
13. Pusarla, R. H. & Bhargava, P. Histones in functional diversification: Core histone variants. *FEBS J.* **272**, 5149–5168 (2005).
14. Frank, D., Doenecke, D. & Albig, W. Differential expression of human replication and cell cycle dependent H3 histone genes. *Gene* **312**, 135–143 (2003).
15. Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W. & Richmond, T. J. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.* **319**, 1097–1113 (2002).

16. Tagami, H., Ray-Gallet, D., Almouzni, G. & Nakatani, Y. Histone H3.1 and H3.3 Complexes Mediate Nucleosome Assembly Pathways Dependent or Independent of DNA Synthesis. *Cell* **116**, 51–61 (2004).
17. van Holde, K. & Zlatanova, J. The nucleosome core particle: does it have structural and physiologic relevance? *Bioessays* **21**, 776–780 (1999).
18. Syed, S. H. *et al.* Single-base resolution mapping of H1–nucleosome interactions and 3D organization of the nucleosome. *Proc. Natl. Acad. Sci.* **107**, 9620–9625 (2010).
19. Flanagan, T. W. & Brown, D. T. Molecular dynamics of histone H1. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1859**, 468–475 (2016).
20. Zhou, B.-R. *et al.* Structural Mechanisms of Nucleosome Recognition by Linker Histones. *Mol. Cell* **59**, 628–638 (2015).
21. Schrödinger, LLC. *The {PyMOL} Molecular Graphics System, Version~1.8.* (2015).
22. Woodcock, C. L. & Dimitrov, S. Higher-order structure of chromatin and nucleosomes. *Curr. Opin. Genet. Dev.* **11**, 130–135 (2001).
23. Allis, C. D., Jenuwein, T. & Reinberg, D. *Epigenetics.* Cold Spring Harbor Laboratory Press (Cold Spring Harbor Laboratory Press, 2007).
24. Felsenfeld, G. & Groudine, M. Controlling the double helix. *Nature* **421**, 448–453 (2003).
25. Robinson, P. J. J., Fairall, L., Huynh, V. A. T. & Rhodes, D. EM measurements define the dimensions of the ‘30-nm’ chromatin fiber: Evidence for a compact, interdigitated structure. *Proc. Natl. Acad. Sci.* **103**, 6506–6511 (2006).
26. Finch, J. T. & Klug, A. Solenoidal model for superstructure in chromatin. *Proc. Natl. Acad. Sci. U. S. A.* **73**, 1897–1901 (1976).
27. Woodcock, C. L., Frado, L. L. & Rattner, J. B. The higher-order structure of chromatin: evidence for a helical ribbon arrangement. *J. Cell Biol.* **99**, 42 LP-52 (1984).
28. Bednar, J. *et al.* Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 14173–8 (1998).
29. Dorigo, B. *et al.* Nucleosome Arrays Reveal the Two-Start Organization of the Chromatin Fiber. *Science (80-.).* **306**, 1571–1573 (2004).
30. Hansen, J. C. Conformational Dynamics of the Chromatin Fiber in Solution: Determinants, Mechanisms, and Functions. *Annu. Rev. Biophys. Biomol. Struct.* **31**, 361–392 (2002).
31. Grigoryev, S. A., Arya, G., Correll, S., Woodcock, C. L. & Schlick, T. Evidence for heteromorphic chromatin fibers from analysis of nucleosome interactions. *Proc. Natl. Acad. Sci.* **106**, 13317–13322 (2009).
32. Li, G. & Reinberg, D. Chromatin higher-order structures and gene regulation. *Curr. Opin. Genet. Dev.* **21**, 175–186 (2011).

33. Woodcock, C. L. & Ghosh, R. P. Chromatin Higher-order Structure and Dynamics. *Cold Spring Harb. Perspect. Biol.* **2**, a000596 (2010).
34. Joti, Y. *et al.* Chromosomes without a 30-nm chromatin fiber. *Nucleus* **3**, 404–410 (2012).
35. Razin, S. V & Gavrilov, A. A. Chromatin without the 30-nm fiber. *Epigenetics* **9**, 653–657 (2014).
36. Cleveland, D. W., Mao, Y. & Sullivan, K. F. Centromeres and Kinetochores: From Epigenetics to Mitotic Checkpoint Signaling. *Cell* **112**, 407–421 (2003).
37. Pluta, A. F., Mackay, A. M., Ainsztein, A. M., Goldberg, I. G. & Earnshaw, W. C. Centromere - Hub of chromosomal activities. *Science* (80-.). **270**, 1591–1594 (1995).
38. Blackburn, E. H. & Szostak, J. W. The Molecular Structure of Centromeres and Telomeres. *Annu. Rev. Biochem.* **53**, 163–194 (1984).
39. Blackburn, E. H. Structure and function of telomeres. *Nature* **350**, 569–573 (1991).
40. Fritz, A. J. *et al.* Chromosomes at Work: Organization of Chromosome Territories in the Interphase Nucleus. *J. Cell. Biochem.* **117**, 9–19 (2016).
41. Pines, J. Cubism and the cell cycle: the many faces of the APC/C. *Nat Rev Mol Cell Biol* **12**, 427–438 (2011).
42. Almagro, S., Rivelino, D., Hirano, T., Houchmandzadeh, B. & Dimitrov, S. The Mitotic Chromosome Is an Assembly of Rigid Elastic Axes Organized by Structural Maintenance of Chromosomes (SMC) Proteins and Surrounded by a Soft Chromatin Envelope. *J. Biol. Chem.* **279**, 5118–5126 (2004).
43. Kireeva, N., Lakonishok, M., Kireev, I., Hirano, T. & Belmont, A. S. Visualization of early chromosome condensation. *J. Cell Biol.* **166**, 775 LP-785 (2004).
44. Ono, T., Fang, Y., Spector, D. L. & Hirano, T. Spatial and Temporal Regulation of Condensins I and II in Mitotic Chromosome Assembly in Human Cells. *Mol. Biol. Cell* **15**, 3296–3308 (2004).
45. Chen, D. *et al.* Condensed mitotic chromatin is accessible to transcription factors and chromatin structural proteins. *J. Cell Biol.* **168**, 41–54 (2005).
46. Parsons, G. G. & Spencer, C. a. Mitotic repression of RNA polymerase II transcription is accompanied by release of transcription elongation complexes. *Mol. Cell. Biol.* **17**, 5791–802 (1997).
47. Dephoure, N. *et al.* A quantitative atlas of mitotic phosphorylation. *Proc. Natl. Acad. Sci.* **105**, 10762–10767 (2008).
48. Wurzenberger, C. & Gerlich, D. W. Phosphatases : providing safe passage. *Nat. Publ. Gr.* **12**, 469–482 (2011).
49. Andrews, P. D., Knatko, E., Moore, W. J. & Swedlow, J. R. Mitotic mechanics: the auroras come into view. *Curr. Opin. Cell Biol.* **15**, 672–683 (2003).

50. Santamaria, A. *et al.* The Plk1-dependent Phosphoproteome of the Early Mitotic Spindle. *Mol. Cell. Proteomics* **10**, (2011).
51. Cremer, T. & Cremer, M. Chromosome Territories. *Cold Spring Harb. Perspect. Biol.* **2**, (2010).
52. Trojer, P. & Reinberg, D. Facultative Heterochromatin: Is There a Distinctive Molecular Signature? *Mol. Cell* **28**, 1–13 (2007).
53. Grewal, S. I. S. & Elgin, S. C. R. Heterochromatin: new possibilities for the inheritance of structure. *Curr. Opin. Genet. Dev.* **12**, 178–187 (2002).
54. Workman, J. L. Nucleosome displacement in transcription. *Genes Dev.* **20**, 2009–2017 (2006).
55. Han, M. & Grunstein, M. Nucleosome loss activates yeast downstream promoters in vivo. *Cell* **55**, 1137–1145 (1988).
56. Laybourn, P. J. & Kadonaga, J. T. Role of nucleosomal cores and histone H1 in regulation of transcription by RNA polymerase II. *Science (80-.)*. **254**, 238 LP-245 (1991).
57. Waddington, C. H. The Epigenotype. *Int. J. Epidemiol.* **41**, 10–13 (2012).
58. Goldberg, A. D., Allis, C. D. & Bernstein, E. Epigenetics: A Landscape Takes Shape. *Cell* **128**, 635–638 (2007).
59. Henikoff, S. & Ahmad, K. Assembly of Variant Histones Into Chromatin. *Annu. Rev. Cell Dev. Biol.* **21**, 133–153 (2005).
60. Henikoff, S. Nucleosome destabilization in the epigenetic regulation of gene expression. *Nat Rev Genet* **9**, 15–26 (2008).
61. Jaenisch, R. & Bird, A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.* **33 Suppl**, 245–254 (2003).
62. Holliday, R. & Pugh, J. E. DNA modification mechanisms and gene activity during development. *Science (80-.)*. **187**, 226–232 (1975).
63. Bird, A. DNA methylation patterns and epigenetic memory DNA methylation patterns and epigenetic memory. *Genes Dev.* 6–21 (2002). doi:10.1101/gad.947102
64. Antequera, F. & Bird, A. in *DNA Methylation: Molecular Biology and Biological Significance* (eds. Jost, J.-P. & Saluz, H.-P.) 169–185 (Birkhäuser Basel, 1993). doi:10.1007/978-3-0348-9118-9_8
65. Ng, H. H. & Bird, A. DNA methylation and chromatin modification. *Curr. Opin. Genet. Dev.* **9**, 158–163 (1999).
66. Goll, M. G. & Bestor, T. H. Eukaryotic Cytosine Methyltransferases. *Annu. Rev. Biochem.* **74**, 481–514 (2005).
67. Saksouk, N., Simboeck, E. & Déjardin, J. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics Chromatin* **8**, 3 (2015).

68. Zaratiegui, M., Irvine, D. V. & Martienssen, R. A. Noncoding RNAs and Gene Silencing. *Cell* **128**, 763–776 (2007).
69. Wei, J.-W., Huang, K., Yang, C. & Kang, C.-S. Non-coding RNAs as regulators in epigenetics (Review). *Oncol. Rep.* 3–9 (2016). doi:10.3892/or.2016.5236
70. Khalil, A. M. *et al.* Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc. Natl. Acad. Sci.* **106**, 11667–11672 (2009).
71. Guttman, M. *et al.* Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**, 223–227 (2009).
72. Bernstein, E. & Allis, C. D. RNA meets chromatin review.pdf. 1635–1655 (2005). doi:10.1101/gad.1324305
73. Wang, K. C. & Chang, H. Y. Molecular mechanisms of long noncoding RNAs. *Mol. Cell* **43**, 904–914 (2011).
74. Robert Finestra, T. & Gribnau, J. X chromosome inactivation: silencing, topology and reactivation. *Curr. Opin. Cell Biol.* **46**, 54–61 (2017).
75. Rinn, J. L. *et al.* Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Non-Coding RNAs. *Cell* **129**, 1311–1323 (2007).
76. Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res.* **21**, 381–395 (2011).
77. Lawrence, M., Daujat, S. & Schneider, R. Lateral Thinking: How Histone Modifications Regulate Gene Expression. *Trends Genet.* **32**, 42–56 (2016).
78. Tan, M. *et al.* Identification of 67 Histone Marks and Histone Lysine Crotonylation as a New Type of Histone Modification. *Cell* **146**, 1016–1028 (2011).
79. Strahl, B. D. & Allis, C. D. The language of covalent histone modifications. *Nature* **403**, 41–45 (2000).
80. Tropberger, P. & Schneider, R. Scratching the (lateral) surface of chromatin regulation by histone modifications. *Nat Struct Mol Biol* **20**, 657–661 (2013).
81. Chatterjee, N. *et al.* Histone Acetylation near the Nucleosome Dyad Axis Enhances Nucleosome Disassembly by RSC and SWI/SNF. *Mol. Cell. Biol.* **35**, 4083–4092 (2015).
82. Sadakierska-Chudy, A. & Filip, M. A Comprehensive View of the Epigenetic Landscape. Part II: Histone Post-translational Modification, Nucleosome Level, and Chromatin Regulation by ncRNAs. *Neurotox. Res.* **27**, 172–197 (2015).
83. Bhaumik, S. R., Smith, E. & Shilatifard, A. Covalent modifications of histones during development and disease pathogenesis. *Nat Struct Mol Biol* **14**, 1008–1016 (2007).
84. Koch, C. M. *et al.* The landscape of histone modifications across 1% of the human genome in five human cell lines. *Genome Res.* **17**, 691–707 (2007).

85. Hong, L., Schroth, G. P., Matthews, H. R., Yau, P. & Bradbury, E. M. Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 'tail' to DNA. *J. Biol. Chem.* **268**, 305–314 (1993).
86. Katan-Khaykovich, Y. & Struhl, K. Dynamics of global histone acetylation and deacetylation in vivo: rapid restoration of normal histone acetylation status upon removal of activators and repressors. *Genes Dev.* **16**, 743–752 (2002).
87. Martin, A. M., Pouchnik, D. J., Walker, J. L. & Wyrick, J. J. Redundant Roles for Histone H3 N-Terminal Lysine Residues in Subtelomeric Gene Repression in *Saccharomyces cerevisiae*. *Genetics* **167**, 1123 LP-1132 (2004).
88. Dion, M. F., Altschuler, S. J., Wu, L. F. & Rando, O. J. Genomic characterization reveals a simple histone H4 acetylation code. *Proc. Natl. Acad. Sci. United States Am.* **102**, 5501–5506 (2005).
89. Shogren-Knaak, M. *et al.* Histone H4-K16 Acetylation Controls Chromatin Structure and Protein Interactions. *Science* (80-.). **311**, 844 LP-847 (2006).
90. Dorigo, B., Schalch, T., Bystricky, K. & Richmond, T. J. Chromatin Fiber Folding: Requirement for the Histone H4 N-terminal Tail. *J. Mol. Biol.* **327**, 85–96 (2003).
91. Bannister, A. J. *et al.* Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120–124 (2001).
92. Canzio, D. *et al.* Chromodomain-Mediated Oligomerization of HP1 Suggests a Nucleosome-Bridging Mechanism for Heterochromatin Assembly. *Mol. Cell* **41**, 67–81 (2011).
93. Fischle, W. *et al.* Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes Dev.* **17**, 1870–1881 (2003).
94. Botuyan, M. V. *et al.* Structural Basis for the Methylation State-Specific Recognition of Histone H4-K20 by 53BP1 and Crb2 in DNA Repair. *Cell* **127**, 1361–1373 (2006).
95. Wang, Z. *et al.* Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet* **40**, 897–903 (2008).
96. Banerjee, T. & Chakravarti, D. A Peek into the Complex Realm of Histone Phosphorylation. *Mol. Cell. Biol.* **31**, 4858–4873 (2011).
97. Wang, F. *et al.* Histone H3 Thr-3 phosphorylation by Haspin positions Aurora B at centromeres in mitosis. *Science* **330**, 231–235 (2010).
98. Paull, T. T. *et al.* A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage. *Curr. Biol.* **10**, 886–895 (2000).
99. Wei, Y., Mizzen, C. A., Cook, R. G., Gorovsky, M. A. & Allis, C. D. Phosphorylation of histone H3 at serine 10 is correlated with chromosome condensation during mitosis and meiosis in *Tetrahymena*. *Proc. Natl. Acad. Sci.* **95**, 7480–7484 (1998).

100. de la Barre, A. *et al.* Core histone N-termini play an essential role in mitotic chromosome condensation. *EMBO J.* **19**, 379 LP-391 (2000).
101. Hans, F. & Dimitrov, S. Histone H3 phosphorylation and cell division. *Oncogene* **20**, 3021–3027 (2001).
102. Wilkins, B. J. *et al.* A Cascade of Histone Modifications Induces Chromatin Condensation in Mitosis. *Science (80-.)*. **343**, 77 LP-80 (2014).
103. Fischle, W. *et al.* Regulation of HP1 – chromatin binding by histone H3 methylation and phosphorylation. **438**, (2005).
104. Lo, W.-S. *et al.* Phosphorylation of Serine 10 in Histone H3 Is Functionally Linked In Vitro and In Vivo to Gcn5-Mediated Acetylation at Lysine 14. *Mol. Cell* **5**, 917–926 (2017).
105. Cheung, P. *et al.* Synergistic Coupling of Histone H3 Phosphorylation and Acetylation in Response to Epidermal Growth Factor Stimulation. *Mol. Cell* **5**, 905–915 (2017).
106. Lau, P. N. I. & Cheung, P. Histone code pathway involving H3S28 phosphorylation and K27 acetylation activates transcription and antagonizes polycomb silencing. *Proc. Natl. Acad. Sci.* **108**, 2801–2806 (2011).
107. Varier, R. A. *et al.* A phospho/methyl switch at histone H3 regulates TFIID association with mitotic chromosomes. *EMBO J.* **29**, 3967 LP-3978 (2010).
108. Smith, C. L. & Peterson, C. L. in (ed. Biology, B. T.-C. T. in D.) **Volume 65**, 115–148 (Academic Press, 2004).
109. Venkatesh, S. & Workman, J. L. Histone exchange, chromatin structure and the regulation of transcription. *Nat Rev Mol Cell Biol* **16**, 178–189 (2015).
110. Lusser, A. & Kadonaga, J. T. Chromatin remodeling by ATP-dependent molecular machines. *BioEssays* **25**, 1192–1200 (2003).
111. Clapier, C. R. & Cairns, B. R. The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.* **78**, 273–304 (2009).
112. Cairns, B. R., Kim, Y. J., Sayre, M. H., Laurent, B. C. & Kornberg, R. D. A multisubunit complex containing the SWI1/ADR6, SWI2/SNF2, SWI3, SNF5, and SNF6 gene products isolated from yeast. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 1950–1954 (1994).
113. Cote, J., Quinn, J., Workman, J. L. & Peterson, C. L. Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. *Science (80-.)*. **265**, 53 LP-60 (1994).
114. Mohrmann, L. & Verrijzer, C. P. Composition and functional specificity of SWI2/SNF2 class chromatin remodeling complexes. *Biochim. Biophys. Acta - Gene Struct. Expr.* **1681**, 59–73 (2005).
115. Bao, Y. & Shen, X. INO80 subfamily of chromatin remodeling complexes. *Mutat. Res. Mol. Mech. Mutagen.* **618**, 18–29 (2007).

116. Mizuguchi, G., Xiao, H., Wisniewski, J., Smith, M. M. & Wu, C. Nonhistone Scm3 and Histones CenH3-H4 Assemble the Core of Centromere-Specific Nucleosomes. *Cell* **129**, 1153–1164 (2007).
117. van Attikum, H., Fritsch, O., Hohn, B. & Gasser, S. M. Recruitment of the INO80 Complex by H2A Phosphorylation Links ATP-Dependent Chromatin Remodeling with DNA Double-Strand Break Repair. *Cell* **119**, 777–788 (2004).
118. Morrison, A. J. *et al.* INO80 and γ -H2AX Interaction Links ATP-Dependent Chromatin Remodeling to DNA Damage Repair. *Cell* **119**, 767–775 (2004).
119. Marfella, C. G. A. & Imbalzano, A. N. The Chd family of chromatin remodelers. *Mutat. Res. Mol. Mech. Mutagen.* **618**, 30–40 (2007).
120. Murawska, M. & Brehm, A. CHD chromatin remodelers and the transcription cycle. *Transcription* **2**, 244–253 (2011).
121. Khattak, S., Lee, B. R., Cho, S. H., Ahnn, J. & Spoerel, N. A. Genetic characterization of Drosophila Mi-2 ATPase. *Gene* **293**, 107–114 (2002).
122. Kehle, J. *et al.* dMi-2, a Hunchback-Interacting Protein That Functions in Polycomb Repression. *Science* (80-.). **282**, 1897 LP-1900 (1998).
123. Corona, D. F. V & Tamkun, J. W. Multiple roles for ISWI in transcription, chromosome organization and DNA replication. *Biochim. Biophys. Acta - Gene Struct. Expr.* **1677**, 113–119 (2004).
124. Mattioli, F., D'Arcy, S. & Luger, K. The right place at the right time: chaperoning core histone variants. *EMBO Rep.* **16**, 1454 LP-1466 (2015).
125. Laskey, R. A., Honda, B. M., Mills, A. D. & Finch, J. T. Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA. *Nature* **275**, 416–420 (1978).
126. Burgess, R. J. & Zhang, Z. Histone chaperones in nucleosome assembly and human disease. *Nat Struct Mol Biol* **20**, 14–22 (2013).
127. Hondele, M. & Ladurner, A. G. The chaperone–histone partnership: for the greater good of histone traffic and chromatin plasticity. *Curr. Opin. Struct. Biol.* **21**, 698–708 (2011).
128. Winkler, D. D. & Luger, K. The Histone Chaperone FACT: Structural Insights and Mechanisms for Nucleosome Reorganization. *J. Biol. Chem.* **286**, 18369–18374 (2011).
129. Foltz, D. R. *et al.* Centromere-Specific Assembly of CENP-A Nucleosomes Is Mediated by HJURP. *Cell* **137**, 472–484 (2009).
130. Cook, A. J. L., Gurard-Levin, Z. A., Vassias, I. & Almouzni, G. A Specific Function for the Histone Chaperone NASP to Fine-Tune a Reservoir of Soluble H3-H4 in the Histone Supply Chain. *Mol. Cell* **44**, 918–927 (2011).

131. Obri, A. *et al.* ANP32E is a histone chaperone that removes H2A.Z from chromatin. *Nature* **505**, 648–53 (2014).
132. Henikoff, S., Furuyama, T. & Ahmad, K. Histone variants, nucleosome assembly and epigenetic inheritance. *Trends Genet.* **20**, 320–326 (2004).
133. Talbert, P. B. & Henikoff, S. Histone variants — ancient wrap artists of the epigenome. **11**, (2010).
134. Li, B., Carey, M. & Workman, J. L. The Role of Chromatin during Transcription. *Cell* **128**, 707–719 (2007).
135. Kamakaka, R. T., Biggins, S., Kamakaka, R. T. & Biggins, S. Histone variants: Deviants? *Genes Dev.* **19**, 295–310 (2005).
136. Boulard, M., Bouvet, P., Kundu, T. K. & Dimitrov, S. Histone variant nucleosomes: structure, function and implication in disease. *Subcell. Biochem.* **41**, 71–89 (2007).
137. Murr, R. *Interplay between different epigenetic modifications and mechanisms. Advances in Genetics* **70**, (Elsevier Inc., 2010).
138. McKittrick, E., Gafken, P. R., Ahmad, K. & Henikoff, S. Histone H3.3 is enriched in covalent modifications associated with active chromatin. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 1525–30 (2004).
139. Waterborg, J. H. Evolution of histone H3: emergence of variants and conservation of post-translational modification sites. *Biochem Cell Biol* **90**, 79–95 (2012).
140. Maze, I., Noh, K.-M., Soshnev, A. A. & Allis, C. D. Every amino acid matters: essential contributions of histone variants to mammalian development and disease. *Nat Rev Genet* **15**, 259–271 (2014).
141. Cheema, M. S. & Ausio, J. The structural determinants behind the epigenetic role of histone variants. *Genes (Basel)*. **6**, 685–713 (2015).
142. Clarkson, M. J., Wells, J. R. E., Gibson, F., Saint, R. & Tremethick, D. J. Regions of variant histone His2AvD required for *Drosophila* development. *Nature* **399**, 694–697 (1999).
143. Chakravarthy, S. *et al.* Structural Characterization of the Histone Variant macroH2A. *Mol. Cell. Biol.* **25**, 7616–7624 (2005).
144. Chakravarthy, S. & Luger, K. The Histone Variant Macro-H2A Preferentially Forms ‘Hybrid Nucleosomes’. *J. Biol. Chem.* **281**, 25522–25531 (2006).
145. Costanzi, C. & Pehrson, J. R. Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature* **393**, 599–601 (1998).
146. Angelov, D. *et al.* The Histone Variant MacroH2A Interferes with Transcription Factor Binding and SWI/SNF Nucleosome Remodeling. *Mol. Cell* **11**, 1033–1041 (2003).
147. Creppe, C. *et al.* MacroH2A1 Regulates the Balance between Self-Renewal and Differentiation Commitment in Embryonic and Adult Stem Cells. *Mol. Cell. Biol.* **32**, 1442–1452 (2012).

148. Buschbeck, M. *et al.* The histone variant macroH2A is an epigenetic regulator of key developmental genes. *Nat Struct Mol Biol* **16**, 1074–1079 (2009).
149. Doyen, C.-M. *et al.* Mechanism of Polymerase II Transcription Repression by the Histone Variant macroH2A. *Mol. Cell. Biol.* **26**, 1156–1164 (2006).
150. Gautier, T. *et al.* Histone variant H2ABbd confers lower stability to the nucleosome. *EMBO Rep.* **5**, 715 LP-720 (2004).
151. Rogakou, E. P., Pilch, D. R., Orr, A. H., Ivanova, V. S. & Bonner, W. M. DNA Double-stranded Breaks Induce Histone H2AX Phosphorylation on Serine 139. *J. Biol. Chem.* **273**, 5858–5868 (1998).
152. van Attikum, H., Fritsch, O. & Gasser, S. M. Distinct roles for SWR1 and INO80 chromatin remodeling complexes at chromosomal double-strand breaks. *EMBO J.* **26**, 4113 LP-4125 (2007).
153. Faast, R. *et al.* Histone variant H2A.Z is required for early mammalian development. *Curr. Biol.* **11**, 1183–1187 (2001).
154. Guillemette, B. *et al.* Variant Histone H2A.Z Is Globally Localized to the Promoters of Inactive Yeast Genes and Regulates Nucleosome Positioning. *PLOS Biol.* **3**, e384 (2005).
155. Raisner, R. M. *et al.* Histone Variant H2A.Z Marks the 5' Ends of Both Active and Inactive Genes in Euchromatin. *Cell* **123**, 233–248 (2005).
156. Barski, A. *et al.* High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell* **129**, 823–837 (2007).
157. Sarcinella, E., Zuzarte, P. C., Lau, P. N. I., Draker, R. & Cheung, P. Monoubiquitylation of H2A.Z distinguishes its association with euchromatin or facultative heterochromatin. *Mol. Cell. Biol.* **27**, 6457–68 (2007).
158. Greaves, I. K., Rangasamy, D., Ridgway, P. & Tremethick, D. J. H2A.Z contributes to the unique 3D structure of the centromere. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 525–530 (2007).
159. Jin, C. *et al.* H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nat. Genet.* **41**, 941–5 (2009).
160. Li, Z. *et al.* Foxa2 and H2A.Z Mediate Nucleosome Depletion during Embryonic Stem Cell Differentiation. *Cell* **151**, 1608–1616 (2012).
161. Xu, Y. *et al.* Histone H2A.Z Controls a Critical Chromatin Remodeling Step Required for DNA Double-Strand Break Repair. *Mol. Cell* **48**, 723–733 (2012).
162. Gévry, N., Chan, H. M., Laflamme, L., Livingston, D. M. & Gaudreau, L. p21 transcription is regulated by differential localization of histone H2A.Z. *Genes Dev.* **21**, 1869–1881 (2007).

163. Li, A. *et al.* Characterization of Nucleosomes Consisting of the Human Testis/Sperm-Specific Histone H2B Variant (hTSH2B). *Biochemistry* **44**, 2529–2535 (2005).
164. Montellier, E. *et al.* Chromatin-to-nucleoprotamine transition is controlled by the histone H2B variant TH2B. *Genes Dev.* **27**, 1680–1692 (2013).
165. Churikov, D. *et al.* Novel human testis-specific histone H2B encoded by the interrupted gene on the X chromosome. *Genomics* **84**, 745–756 (2004).
166. Hake, S. B. *et al.* Expression patterns and post-translational modifications associated with mammalian histone H3 variants. *J. Biol. Chem.* **281**, 559–568 (2006).
167. Elsaesser, S. J., Goldberg, A. D. & Allis, C. D. New functions for an old variant : no substitute for histone H3.3. *Curr. Opin. Genet. Dev.* **20**, 110–117 (2010).
168. Filipescu, D. & Szenker, E. Developmental roles of histone H3 variants and their chaperones. **29**, (2013).
169. Malik, H. S. & Henikoff, S. Phylogenomics of the nucleosome. *Nat Struct Mol Biol* **10**, 882–891 (2003).
170. Witt, O., Albig, W. & Doenecke, D. Testis-Specific Expression of a Novel Human H3 Histone Gene. *Exp. Cell Res.* **229**, 301–306 (1996).
171. Tachiwana, H. *et al.* Structural basis of instability of the nucleosome containing a testis-specific histone variant, human H3T. *Proc. Natl. Acad. Sci.* **107**, 10454–10459 (2010).
172. Wiedemann, S. M. *et al.* Identification and characterization of two novel primate-specific histone H3 variants, H3.X and H3.Y. *J. Cell Biol.* **190**, 777 LP-791 (2010).
173. Kujirai, T. *et al.* Structure and function of human histone H3.Y nucleosome. *Nucleic Acids Res.* **44**, 6127–6141 (2016).
174. Schenk, R., Jenke, A., Zilbauer, M., Wirth, S. & Postberg, J. H3.5 is a novel hominid-specific histone H3 variant that is specifically expressed in the seminiferous tubules of human testes. *Chromosoma* **120**, 275–285 (2011).
175. Urahama, T. *et al.* Histone H3.5 forms an unstable nucleosome and accumulates around transcription start sites in human testis. *Epigenetics Chromatin* **9**, 2 (2016).
176. Wysocka, J. *et al.* A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **442**, 86–90 (2006).
177. Heard, E. Delving into the diversity of facultative heterochromatin: the epigenetics of the inactive X chromosome. *Curr. Opin. Genet. Dev.* **15**, 482–489 (2005).
178. Postberg, J., Forcob, S., Chang, W. & Lipps, H. J. The evolutionary history of histone H3 suggests a deep eukaryotic root of chromatin modifying mechanisms. *BMC Evol. Biol.* **10**, 259 (2010).
179. Thatcher, T. H. & Gorovsky, M. A. Phylogenetic analysis of the core histones H2A, H2B, H3, and H4. *Nucleic Acids Res.* **22**, 174–179 (1994).

180. Kent, W. J. *et al.* The Human Genome Browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).
181. Marzluff, W. F. & Duronio, R. J. Histone mRNA expression: multiple levels of cell cycle regulation and important developmental consequences William F Marzluff*. 1–8 (2002). doi:10.1016/S0955067402003873
182. Bramlage, B., Kosciessa, U. & Doenecke, D. Differential expression of the murine histone genes H3.3A and H3.3B. *Differentiation* **62**, 13–20 (1997).
183. Wells, D., Hoffman, D. & Kedes, L. Unusual structure, evolutionary conservation of non-coding sequences and numerous pseudogenes characterize the human H3.3 histone multigene family. *Nucleic Acids Res.* **15**, 2871–2889 (1987).
184. Wells, D. & Kedes, L. Structure of a human histone cDNA: evidence that basally expressed histone genes have intervening sequences and encode polyadenylylated mRNAs. *Proc. Natl. Acad. Sci. U. S. A.* **82**, 2834–8 (1985).
185. Akhmanova, A. S. *et al.* Structure and expression of histone H3.3 genes in *Drosophila melanogaster* and *Drosophila hydei*. *Genome* **38**, 586–600 (1995).
186. Tang, M. C. W. *et al.* Contribution of the Two Genes Encoding Histone Variant H3.3 to Viability and Fertility in Mice. *PLoS Genet.* **11**, 1–23 (2015).
187. Jang, C., Shibata, Y., Starmer, J., Yee, D. & Magnuson, T. Histone H3.3 maintains genome integrity during mammalian development. *Genes Dev.* **1**, 1377–1392 (2015).
188. Bush, K. M. *et al.* Endogenous mammalian histone H3.3 exhibits chromatin-related functions during development. *Epigenetics Chromatin* **6**, 7 (2013).
189. Turner, J. & Crossley, M. Mammalian Krüppel-like transcription factors: more than just a pretty finger. *Trends Biochem. Sci.* **24**, 236–240 (1999).
190. Witt, O., Albig, W. & Doenecke, D. Transcriptional regulation of the human replacement histone gene H3.3B. *FEBS Lett.* **408**, 255–260 (1997).
191. Yuen, B. T. K., Bush, K. M., Barrilleaux, B. L., Cotterman, R. & Knoepfler, P. S. Histone H3.3 regulates dynamic chromatin states during spermatogenesis. *Development* 1–12 (2014). doi:10.1242/dev.106450
192. Mazumder, B., Seshadri, V. & Fox, P. L. Translational control by the 3'-UTR: the ends specify the means. *Trends Biochem. Sci.* **28**, 91–98 (2003).
193. Feng, R. *et al.* Regulation of the expression of histone H3.3 by differential polyadenylation. *Genome* **48**, 503–510 (2005).
194. Pulcrano, G. *et al.* PLAUF binding to the 3'UTR of the H3.3 histone transcript affects mRNA stability. *Gene* **406**, 124–133 (2007).
195. Fucci, L., Aniello, F., Branno, M., Biffali, E. & Geraci, G. Isolation of a new H3.3 histone variant cDNA of *P. lividus* sea urchin: Sequence and embryonic expression. *Biochim. Biophys. Acta - Gene Struct. Expr.* **1219**, 539–542 (1994).

196. Bevilacqua, A., Ceriani, M. C., Capaccioli, S. & Nicolin, A. Post-transcriptional regulation of gene expression by degradation of messenger RNAs. *J. Cell. Physiol.* **195**, 356–372 (2003).
197. Lu, J.-Y. & Schneider, R. J. Tissue Distribution of AU-rich mRNA-binding Proteins Involved in Regulation of mRNA Decay. *J. Biol. Chem.* **279**, 12974–12979 (2004).
198. White, E. J. F., Matsangos, A. E. & Wilson, G. M. AUF1 regulation of coding and noncoding RNA. *Wiley Interdiscip. Rev. RNA* **8**, e1393--n/a (2017).
199. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, (1997).
200. Tachiwana, H. *et al.* Structures of human nucleosomes containing major histone H3 variants. *Acta Crystallogr. Sect. D* **67**, 578–583 (2011).
201. Ahmad, K. & Henikoff, S. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol. Cell* **9**, 1191–1200 (2002).
202. Hake, S. B. *et al.* 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**, 6344–6349 (2005).
203. Jacob, Y. *et al.* Selective Methylation of Histone H3 Variant H3.1 Regulates Heterochromatin Replication. *Science (80-.).* **343**, 1249 LP-1253 (2014).
204. Wong, L. H. *et al.* ATRX interacts with H3.3 in maintaining telomere structural integrity in pluripotent embryonic stem cells. 351–360 (2010). doi:10.1101/gr.101477.109.20
205. Goldberg, A. D. *et al.* Distinct Factors Control Histone Variant H3.3 Localization at Specific Genomic Regions. *Cell* **140**, 678–691 (2010).
206. Drané, P., Ouararhni, K., Depaux, A., Shuaib, M. & Hamiche, A. The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. *Genes Dev.* **24**, 1253–1265 (2010).
207. Ray-Gallet, D. *et al.* Dynamics of Histone H3 Deposition In Vivo Reveal a Nucleosome Gap-Filling Mechanism for H3 . 3 to Maintain Chromatin Integrity. *Mol. Cell* **44**, 928–941 (2011).
208. Tang, Y. *et al.* Identification of an Ubinuclein 1 region required for stability and function of the human HIRA/UBN1/CABIN/ASF1a histone H3.3 chaperone complex. **51**, 2366–2377 (2013).
209. Ray-Gallet, D. *et al.* HIRA Is Critical for a Nucleosome Assembly Pathway Independent of DNA Synthesis. *Mol. Cell* **9**, 1091–1100 (2002).
210. Elsaesser, S. J. & Allis, C. D. HIRA and Daxx Constitute Two Independent Histone H3.3-Containing Predeposition Complexes. *Cold Spring Harb. Symp. Quant. Biol.* **LXXV**, 27–34 (2010).
211. Ricketts, D. M. *et al.* Ubinuclein-1 confers histone H3.3-specific-binding by the HIRA histone chaperone complex. *Nat. Commun.* **6**, 7711 (2015).

212. Drane, P. *et al.* The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. *Genes Dev.* **24**, 1253–1265 (2010).
213. Kraushaar, D. C. *et al.* Genome-wide incorporation dynamics reveal distinct categories of turnover for the histone variant H3.3. *Genome Biol.* **14**, R121 (2013).
214. Huang, C. *et al.* H3.3-H4 Tetramer Splitting Events Feature Cell-Type Specific Enhancers. *PLOS Genet.* **9**, e1003558 (2013).
215. Deal, R. B., Henikoff, J. G. & Henikoff, S. Genome-Wide Kinetics of Nucleosome Turnover Determined by Metabolic Labeling of Histones. *Science (80-.).* **328**, 1161 LP-1164 (2010).
216. Mito, Y., Henikoff, J. G. & Henikoff, S. Genome-scale profiling of histone H3.3 replacement patterns. *Nat Genet* **37**, 1090–1097 (2005).
217. Wirbelauer, C., Bell, O. & Schübeler, D. 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**, 1761–1766 (2005).
218. Mito, Y., Henikoff, J. G. & Henikoff, S. Histone Replacement Marks the Boundaries of cis-Regulatory Domains. *Science (80-.).* **315**, 1408 LP-1411 (2007).
219. Chow, C. *et al.* Variant histone H3.3 marks promoters of transcriptionally active genes during mammalian cell division. *EMBO Rep.* **6**, 354–360 (2005).
220. Jin, C. & Felsenfeld, G. Distribution of histone H3.3 in hematopoietic cell lineages. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 574–579 (2006).
221. Sutcliffe, E. L. *et al.* Dynamic Histone Variant Exchange Accompanies Gene Induction in T Cells. *Mol. Cell. Biol.* **29**, 1972–1986 (2009).
222. Tamura, T. *et al.* Inducible Deposition of the Histone Variant H3.3 in Interferon-stimulated Genes. *J. Biol. Chem.* **284**, 12217–12225 (2009).
223. Wong, L. H. *et al.* Histone H3.3 incorporation provides a unique and functionally essential telomeric chromatin in embryonic stem cells. *Genome Res.* **19**, 404–414 (2009).
224. Elsässer, S. J., Noh, K.-M., Diaz, N., Allis, C. D. & Banaszynski, L. A. Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells. *Nature* **522**, 240–4 (2015).
225. Delbarre, E. *et al.* Chromatin Environment of Histone Variant H3.3 Revealed by Quantitative Imaging and Genome-scale Chromatin and DNA Immunoprecipitation. *Mol. Biol. Cell* **21**, 1872–1884 (2010).
226. Lund, E. G., Collas, P. & Delbarre, E. Transcription outcome of promoters enriched in histone variant H3.3 defined by positioning of H3.3 and local chromatin marks. *Biochem. Biophys. Res. Commun.* **460**, 348–353 (2015).

227. Loyola, A., Bonaldi, T., Roche, D., Imhof, A. & Almouzni, G. PTMs on H3 Variants before Chromatin Assembly Potentiate Their Final Epigenetic State. *Mol. Cell* **24**, 309–316 (2006).
228. Johnson, L. *et al.* Mass spectrometry analysis of Arabidopsis histone H3 reveals distinct combinations of post-translational modifications . *Nucleic Acids Res.* **32**, 6511–6518 (2004).
229. Lin, C.-J., Conti, M. & Ramalho-Santos, M. Histone variant H3.3 maintains a decondensed chromatin state essential for mouse preimplantation development. *Development* **140**, 3624 LP-3634 (2013).
230. Braunschweig, U., Hogan, G. J., Pagie, L. & van Steensel, B. Histone H1 binding is inhibited by histone variant H3.3. *EMBO J.* **28**, 3635 LP-3645 (2009).
231. Jin, C. & Felsenfeld, G. Nucleosome stability mediated by histone variants H3.3 and H2A.Z. *Genes Dev.* **21**, 1519–1529 (2007).
232. Dunleavy, E. M., Almouzni, G. & Karpen, G. H. H3.3 is deposited at centromeres in S phase as a placeholder for newly assembled CENP-A in G₁ phase. *Nucleus* **2**, 146–157 (2011).
233. Yukawa, M. *et al.* Genome-wide analysis of the chromatin composition of histone H2A and H3 variants in mouse embryonic stem cells. *PLoS One* **9**, (2014).
234. Thakar, A. *et al.* H2A.Z and H3.3 histone variants affect nucleosome structure: Biochemical and biophysical studies. *Biochemistry* **48**, 10852–10857 (2009).
235. Bönisch, C. & Hake, S. B. Histone H2A variants in nucleosomes and chromatin: more or less stable? *Nucleic Acids Res.* **40**, 10719–10741 (2012).
236. Rhee, H. S., Bataille, A. R., Zhang, L. & Pugh, B. F. Article Subnucleosomal Structures and Nucleosome Asymmetry across a Genome. *Cell* **159**, 1377–1388 (2014).
237. Voon, H. P. J. *et al.* ATRX Plays a Key Role in Maintaining Silencing at Interstitial Heterochromatic Loci and Imprinted Genes. *Cell Rep.* **11**, 405–418 (2015).
238. Blasco, M. A. The epigenetic regulation of mammalian telomeres. *Nat Rev Genet* **8**, 299–309 (2007).
239. Voon, H. P. J. & Wong, L. H. New players in heterochromatin silencing: Histone variant H3.3 and the ATRX/DAXX chaperone. *Nucleic Acids Res.* **44**, 1496–1501 (2015).
240. Stocking, C. & Kozak, C. A. Endogenous retroviruses: Murine endogenous retroviruses. *Cell. Mol. Life Sci.* **65**, 3383–3398 (2008).
241. Leung, D. C. & Lorincz, M. C. Silencing of endogenous retroviruses: when and why do histone marks predominate? *Trends Biochem. Sci.* **37**, 127–133 (2012).
242. Udugama, M. *et al.* Histone variant H3.3 provides the heterochromatic H3 lysine 9 tri-methylation mark at telomeres. *Nucleic Acids Res.* **43**, 10227–10237 (2015).

243. Pérez-Cadahía, B., Drohic, B. & Davie, J. R. H3 phosphorylation: dual role in mitosis and interphase This paper is one of a selection of papers published in this Special Issue entitled 30th Annual International Asilomar Chromatin and Chromosomes Conference and has undergone the Journal's usual peer review. *Biochem. Cell Biol.* **87**, 695–709 (2009).
244. Van Hooser, A., Goodrich, D. W., Allis, C. D., Brinkley, B. R. & Mancini, M. A. Histone H3 phosphorylation is required for the initiation, but not maintenance, of mammalian chromosome condensation. *J. Cell Sci.* **111**, 3497 LP-3506 (1998).
245. Hsu, J.-Y. *et al.* Mitotic Phosphorylation of Histone H3 Is Governed by Ipl1/aurora Kinase and Glc7/PP1 Phosphatase in Budding Yeast and Nematodes. *Cell* **102**, 279–291 (2000).
246. Garcia, B. A. *et al.* Modifications of Human Histone H3 Variants during Mitosis. *Biochemistry* **44**, 13202–13213 (2005).
247. Hendzel, M. *et al.* Mitosis-specific phosphorylation of histone H3 initiates primarily within pericentromeric heterochromatin during G2 and spreads in an ordered fashion coincident with mitotic chromosome condensation. *Chromosoma* **106**, 348–60 (1997).
248. Schulmeister, A., Schmid, M. & Thompson, E. M. Phosphorylation of the histone H3.3 variant in mitosis and meiosis of the urochordate *Oikopleura dioica*. *Chromosom. Res.* **15**, 189 (2007).
249. Hinchcliffe, E. H. *et al.* Chromosome missegregation during anaphase triggers p53 cell cycle arrest through histone H3.3 Ser31 phosphorylation. *Nat. Cell Biol. advance on*, 668–675 (2016).
250. Chang, F. T. M. *et al.* CHK1-driven histone H3.3 serine 31 phosphorylation is important for chromatin maintenance and cell survival in human ALT cancer cells. *Nucleic Acids Res* **43**, 2603–2614 (2015).
251. Couldrey, C., Carlton, M., Nolan, P., Colledge, W. & Evans, M. 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.* **13**, 2489–95 (1999).
252. Tang, M. C. W., Jacobs, S. A., Wong, L. H. & Mann, J. R. Conditional allelic replacement applied to genes encoding the histone variant H3.3 in the mouse. *Genesis* **51**, 142–146 (2013).
253. Inoue, A. & Zhang, Y. Nucleosome assembly is required for nuclear pore complex assembly in mouse zygotes. *Nat. Struct. Mol. Biol.* **21**, 609–16 (2014).
254. Duarte, L. F. *et al.* Histone H3.3 and its proteolytically processed form drive a cellular senescence programme. *Nat. Commun.* **5**, 5210 (2014).
255. Urban, M. K. & Zweidler, A. Changes in nucleosomal core histone variants during chicken development and maturation. *Dev. Biol.* **95**, 421–428 (1983).
256. Piña, B. & Suau, P. Changes in histones H2A and H3 variant composition in differentiating and mature rat brain cortical neurons. *Dev. Biol.* **123**, 51–58 (1987).

257. Hake, S. B. & Allis, C. D. Histone H3 variants and their potential role in indexing mammalian genomes: the 'H3 barcode hypothesis'. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 6428–6435 (2006).
258. Bano, D., Piazzesi, A., Salomoni, P. & Nicotera, P. The histone variant H3.3 claims its place in the crowded scene of epigenetics. *Aging (Albany. NY).* **9**, (2017).
259. Portela, A. & Esteller, M. Epigenetic modifications and human disease. *Nat. Biotechnol.* **28**, 1057–1068 (2010).
260. Zink, L. M. & Hake, S. B. Histone variants: Nuclear function and disease. *Curr. Opin. Genet. Dev.* **37**, 82–89 (2016).
261. Schwartzenuber, J. *et al.* Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature* **482**, 226–231 (2012).
262. Wu, G. *et al.* Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. *Nat Genet* **44**, 251–253 (2012).
263. Castel, D. *et al.* Histone H3F3A and HIST1H3B K27M mutations define two subgroups of diffuse intrinsic pontine gliomas with different prognosis and phenotypes. *Acta Neuropathol.* **130**, 815–827 (2015).
264. Kallappagoudar, S., Yadav, R. K., Lowe, B. R. & Partridge, J. F. Histone H3 mutations—a special role for H3.3 in tumorigenesis? *Chromosoma* **124**, 177–189 (2015).
265. Bender, S. *et al.* Reduced H3K27me3 and DNA Hypomethylation Are Major Drivers of Gene Expression in K27M Mutant Pediatric High-Grade Gliomas. *Cancer Cell* **24**, 660–672 (2013).
266. Lewis, P. W. *et al.* Inhibition of PRC2 Activity by a Gain-of-Function H3 Mutation Found in Pediatric Glioblastoma. *Science (80-.).* **340**, 857 LP-861 (2013).
267. Park, S.-M. *et al.* Histone variant H3F3A promotes lung cancer cell migration through intronic regulation. *Nat. Commun.* **7**, 12914 (2016).
268. Graber, M., Schweinfest, C., Reed, C., Papas, T. & Baron, P. Isolation of differentially expressed genes in carcinoma of the esophagus. *Ann. Surg. Oncol.* **2**, 192–7 (1996).
269. Behjati, S. *et al.* Distinct H3F3A and H3F3B driver mutations define chondroblastoma and giant cell tumor of bone. *Nat Genet* **45**, 1479–1482 (2013).
270. Presneau, N., Shen, Z., Provencher, D., Mes-Masson, A.-M. & Tonin, P. N. Identification of novel variant, 1484delG in the 3'UTR of H3F3B, a member of the histone 3B replacement family, in ovarian tumors. *Int. J. Oncol.* **26**, 1621–1627 (2005).
271. Ayoubi, H. A., Mahjoubi, F. & Mirzaei, R. Investigation of the human H3.3B (H3F3B) gene expression as a novel marker in patients with colorectal cancer. *J. Gastrointest. Oncol. Vol 8, No 1 (February 2017) J. Gastrointest. Oncol.* (2017).

272. Heaphy, C. M. *et al.* Altered Telomeres in Tumors with ATRX and DAXX Mutations. *Science* (80-.). **333**, 425 LP-425 (2011).
273. Jiao, Y. *et al.* DAXX/ATRX, MEN1, and mTOR Pathway Genes Are Frequently Altered in Pancreatic Neuroendocrine Tumors. *Science* (80-.). **331**, 1199 LP-1203 (2011).
274. Buschbeck, M. & Hake, S. B. Variants of core histones and their roles in cell fate decisions, development and cancer. *Nat Rev Mol Cell Biol* **advance on**, (2017).
275. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
276. Ye, J. *et al.* Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* **13**, 1–11 (2012).
277. Xu, J. Preparation, Culture, and Immortalization of Mouse Embryonic Fibroblasts. *Curr. Protoc. Mol. Biol.* **Chapter 28**, 1–8 (2005).
278. Rosenbloom, K. R. *et al.* The UCSC Genome Browser database: 2015 update. *Nucleic Acids Res.* **43**, D670–D681 (2015).
279. Ors, A. *et al.* Histone H3.3 regulates mitotic progression in mouse embryonic fibroblasts. *Biochem. Cell Biol.* (2017). doi:10.1139/bcb-2016-0190
280. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
281. Ye, T. *et al.* seqMINER: an integrated ChIP-seq data interpretation platform. *Nucleic Acids Res.* **39**, e35–e35 (2011).
282. Zhang, Y. *et al.* Model-based Analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
283. Jurka, J. *et al.* Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet. Genome Res.* **110**, 462–467 (2005).
284. Bargaje, R. *et al.* Proximity of H2A.Z containing nucleosome to the transcription start site influences gene expression levels in the mammalian liver and brain. *Nucleic Acids Res.* **40**, 8965–8978 (2012).
285. Banaszynski, L. A. *et al.* Hira-Dependent Histone H3.3 Deposition Facilitates PRC2 Recruitment at Developmental Loci in ES Cells. **5**, 107–120 (2013).
286. Maehara, K. *et al.* Tissue-specific expression of histone H3 variants diversified after species separation. *Epigenetics Chromatin* **8**, 35 (2015).
287. Mikkelsen, T. S. *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560 (2007).
288. Sadic, D. *et al.* Atrx promotes heterochromatin formation at retrotransposons. *EMBO Rep.* **16**, 836–850 (2015).

289. Groh, S. & Schotta, G. Silencing of endogenous retroviruses by heterochromatin. *Cell. Mol. Life Sci.* 1–11 (2017). doi:10.1007/s00018-017-2454-8
290. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: The next generation. *Cell* **144**, 646–674 (2011).
291. Smith, M. L. *et al.* p53-Mediated DNA Repair Responses to UV Radiation: Studies of Mouse Cells Lacking p53, p21, and/or gadd45 Genes. *Mol. Cell. Biol.* **20**, 3705–3714 (2000).
292. Hayashi, M. T., Cesare, A. J., Fitzpatrick, J. A. J., Lazzerini-Denchi, E. & Karlseder, J. A telomere-dependent DNA damage checkpoint induced by prolonged mitotic arrest. *Nat Struct Mol Biol* **19**, 387–394 (2012).
293. Uetake, Y. & Sluder, G. Prolonged Prometaphase Blocks Daughter Cell Proliferation Despite Normal Completion of Mitosis. *Curr. Biol.* **20**, 1666–1671 (2010).
294. Janssen, A., van der Burg, M., Szuhai, K., Kops, G. J. P. L. & Medema, R. H. Chromosome Segregation Errors as a Cause of DNA Damage and Structural Chromosome Aberrations. *Science* (80-.). **333**, 1895 LP-1898 (2011).
295. Ivanauskiene, K. *et al.* The PML-associated protein DEK regulates the balance of H3.3 loading on chromatin and is important for telomere integrity. *Genome Res.* **24**, 1584–1594 (2014).
296. Hödl, M. & Basler, K. Transcription in the absence of histone H3.2 and H3K4 methylation. *Curr. Biol.* **22**, 2253–2257 (2012).
297. Montagutelli, X. Effect of the Genetic Background on the Phenotype of Mouse Mutations. *J. Am. Soc. Nephrol.* **11**, S101–S105 (2000).
298. Goutte-Gattat, D. *et al.* Phosphorylation of the CENP-A amino-terminus in mitotic centromeric chromatin is required for kinetochore function. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 8579–84 (2013).
299. Ozen, C. *et al.* Genetics and epigenetics of liver cancer. *N. Biotechnol.* **30**, 381–384 (2013).
300. American Cancer Society. Global Cancer Facts & Figures 3rd Edition. *Am. Cancer Soc.* 1–64 (2015). doi:10.1002/ijc.27711
301. Fausto, N. Liver regeneration. *J. Hepatol.* **32**, 19–31 (2000).
302. Pogribny, I. P. I. P. *et al.* Role of epigenetic aberrations in the development and progression of human hepatocellular carcinoma. *Cancer Lett.* **342**, 223–230 (2014).
303. Tannour-Louet, M., Porteu, A., Vaulont, S., Kahn, A. & Vasseur-Cognet, M. A tamoxifen-inducible chimeric Cre recombinase specifically effective in the fetal and adult mouse liver. *Hepatology* **35**, 1072–1081 (2002).

Appendices

Appendix A – Flow cytometry analysis of the cell cycle by DNA content (PI incorporation)

A representative cell cycle analysis by flow cytometry using PI staining is shown in Figure A. All other samples used in this study were analyzed similarly.

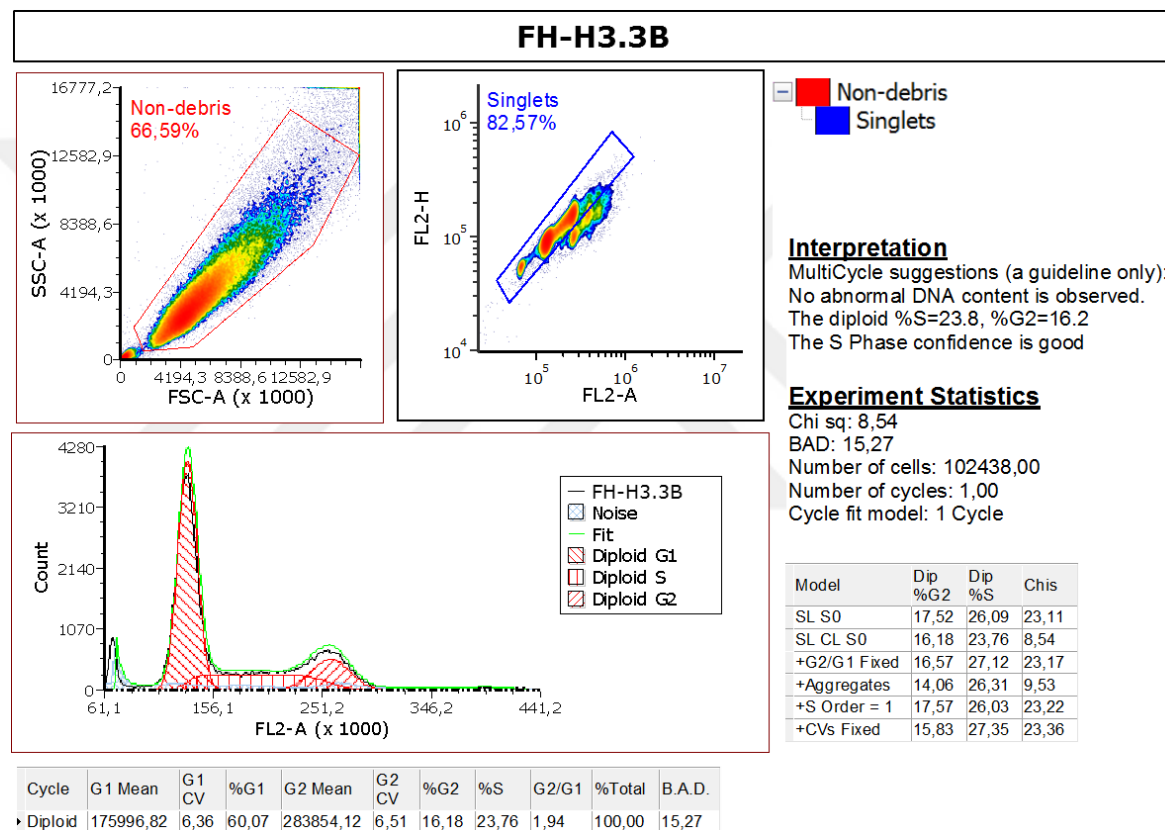


Figure A. Cell cycle analysis by flow cytometry of FH-H3.3B MEFs. Single cells (Gate1: Non-debris) were identified by measuring forward scatter (FSC) and side scatter (SSC). Pulse processing using pulse area (FL2-A) vs. pulse height (FL2-H) was used to exclude cell doublets (Gate2: Singlets). PI (FL2-A) was measured at 605nm to determine DNA content. Multicycle AV module embedded in the FCS Express software was used to fit the Gaussian curve of DNA quantity to the cell cycle model SL CL 50 which has the lowest χ^2 value among the 6 models proposed by the software.

Appendix B – Copyright Permissions

- Copyright permission for “Figure 1-4. The eukaryotic cell cycle.” adapted from Pines, 2011 ⁴¹

License Number	4105320861045
License date	May 10, 2017
Licensed Content Publisher	Nature Publishing Group
Licensed Content Publication	Nature Reviews Molecular Cell Biology
Licensed Content Title	Cubism and the cell cycle: the many faces of the APC/C
Licensed Content Author	Jonathon Pines
Licensed Content Date	Jul 1, 2011
Licensed Content Volume	12
Licensed Content Issue	7
Type of Use	reuse in a dissertation / thesis
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	No
Figures	Figure1. The cell cycle
Author of this NPG article	No
Your reference number	
Title of your thesis / dissertation	ROLE OF HISTONE VARIANT H3.3 IN TRANSCRIPTION AND MITOTIC PROGRESSION
Expected completion date	Apr 2017
Estimated size (number of pages)	150
Requestor Location	Aysegul Ors Üniversiteler Mahallesi Bilkent University Department of MBG Main campus, SB Building, 2nd floor Ankara, 06800 Turkey Attn: Aysegul Ors
Billing Type	Invoice
Billing Address	Aysegul Ors Üniversiteler Mahallesi Bilkent University Department of MBG Main campus, SB Building, 2nd floor Ankara, Turkey 06800 Attn: Aysegul Ors
Total	0.00 USD

- Copyright permission for “Figure 1-6. Human core and linker histone variants.” taken from Maze et al., 2014

License Number	4078481485279
License date	Mar 29, 2017
Licensed Content Publisher	Nature Publishing Group
Licensed Content Publication	Nature Reviews Genetics
Licensed Content Title	Every amino acid matters: essential contributions of histone variants to mammalian development and disease
Licensed Content Author	Ian Maze, Kyung-Min Noh, Alexey A. Soshnev, C. David Allis
Licensed Content Date	Mar 11, 2014
Licensed Content Volume	15
Licensed Content Issue	4
Type of Use	reuse in a dissertation / thesis
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Figures	Figure 1 Human core and linker histone variants
Author of this NPG article	no
Your reference number	
Title of your thesis / dissertation	ROLE OF HISTONE VARIANT H3.3 IN TRANSCRIPTION AND MITOTIC PROGRESSION
Expected completion date	Apr 2017
Estimated size (number of pages)	150
Requestor Location	Aysegul Ors Üniversiteler Mahallesi Bilkent University Department of MBG Main campus, SB Building, 2nd floor Ankara, 06800 Turkey Attn: Aysegul Ors
Billing Type	Invoice
Billing Address	Aysegul Ors Üniversiteler Mahallesi Bilkent University Department of MBG Main campus, SB Building, 2nd floor Ankara, Turkey 06800 Attn: Aysegul Ors
Total	0.00 USD

- Copyright permission for “Figure 1-12. Contribution of histone mutations and deregulations in their expression to tumorigenesis in humans.” taken from Zink and Hake, 2016²⁶⁰

License Number	4078530031512
License date	Mar 29, 2017
Licensed Content Publisher	Elsevier
Licensed Content Publication	Current Opinion in Genetics & Development
Licensed Content Title	Histone variants: nuclear function and disease
Licensed Content Author	Lisa-Maria Zink,Sandra B Hake
Licensed Content Date	April 2016
Licensed Content Volume	37
Licensed Content Issue	n/a
Licensed Content Pages	8
Start Page	82
End Page	89
Type of Use	reuse in a thesis/dissertation
Intended publisher of new work	Other
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Original figure numbers	Figure 2. Deregulation and mutation of histone variants contribute to various human tumors.
Title of your thesis/dissertation	ROLE OF HISTONE VARIANT H3.3 IN TRANSCRIPTION AND MITOTIC PROGRESSION
Expected completion date	Apr 2017
Estimated size (number of pages)	150
Elsevier VAT number	GB 494 6272 12
Requestor Location	Aysegul Ors Üniversiteler Mahallesi Bilkent University Department of MBG Main campus, SB Building, 2nd floor Ankara, 06800 Turkey Attn: Aysegul Ors
Total	0.0 USD

- Copyright permission for “Table 1-1. Point mutations in H3.3 and its chaperones observed in human cancer” taken from Buschbeck and Hake, 2017 ²⁷⁴

License Number	4081100644032
License date	Apr 02, 2017
Licensed Content Publisher	Nature Publishing Group
Licensed Content Publication	Nature Reviews Molecular Cell Biology
Licensed Content Title	Variants of core histones and their roles in cell fate decisions, development and cancer
Licensed Content Author	Marcus Buschbeck, Sandra B. Hake
Licensed Content Date	Feb 1, 2017
Type of Use	reuse in a dissertation / thesis
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	No
Figures	Table 3: Point mutations in histones and their chaperones in human cancer
Author of this NPG article	No
Title of your thesis / dissertation	ROLE OF HISTONE VARIANT H3.3 IN TRANSCRIPTION AND MITOTIC PROGRESSION
Expected completion date	Apr 2017
Estimated size (number of pages)	150
Requestor Location	Aysegul Ors Üniversiteler Mahallesi Bilkent University Department of MBG Main campus, SB Building, 2nd floor Ankara, 06800 Turkey Attn: Aysegul Ors
Billing Type	Invoice
Billing Address	Aysegul Ors Üniversiteler Mahallesi Bilkent University Department of MBG Main campus, SB Building, 2nd floor Ankara, Turkey 06800 Attn: Aysegul Ors
Total	0.00 USD

Articles

1. **Ors, A.** *et al.* Histone H3.3 regulates mitotic progression in mouse embryonic fibroblasts. *Biochemistry and Cell Biology*. (2017). doi:10.1139/bcb-2016-0190
2. Ozen, C., Yildiz, G., Dagcan, A.T., Cevik, D., **Ors, A.**, Keles, U., Topel, H., Ozturk, M. Genetics and epigenetics of liver cancer. *New Biotechnology* 30, 381–384 (2013).



Histone H3.3 regulates mitotic progression in mouse embryonic fibroblasts

Aysegul Ors^{1,2*}, Christophe Papin^{3*}, Bertrand Favier⁴, Yohan Roulland², Defne Dalkara², Mehmet Ozturk⁵, Ali Hamiche^{3#} Stefan Dimitrov^{2#} and Kiran Padmanabhan^{6#}.

¹ Department of Molecular Biology and Genetics, Faculty of Science, Bilkent University, 06800 Ankara, Turkey

² Université Grenoble Alpes, Institute for Advanced Biosciences, INSERM U1209/ CNRS 5309, 38700 La Tronche, France

³ Université de Strasbourg, Institut de Génétique et Biologie Moléculaire et Cellulaire (IGBMC), CNRS, INSERM, Equipe labélisée Ligue contre le Cancer, 1 rue Laurent Fries, B.P. 10142, 67404 Illkirch Cedex, France.

⁴ Université de Grenoble Alpes, Team GREPI, Etablissement Français du Sang, EA 7408, BP35, 38701 La Tronche, France.

⁵ Izmir Biomedicine and Genome Center, Faculty of Medicine, Dokuz Eylül University, Izmir, Turkey.

⁶ Institut Génomique Fonctionnelle de Lyon (IGFL), Ecole Normale Supérieure de Lyon, UMR 5242, 46 Allée d'Italie, 69364 Lyon Cedex 07, France.

^{*}, equal contribution

[#], corresponding authors: kiran.padmanabhan@ens-lyon.fr, Stefan.dimitrov@univ-grenoble-alpes.fr, hamiche@igbmc.fr

Abstract

H3.3 is a histone variant, which marks transcription start sites as well as telomeres and heterochromatic sites on the genome. H3.3 presence is thought to positively correlate with transcriptional status of its target genes. Using a conditional genetic strategy against H3.3B combined with short hairpin RNAs against H3.3A, we essentially depleted all H3.3 gene expression in mouse embryonic fibroblasts. Following nearly complete loss of H3.3 in cells, our transcriptomic analyses show very little impact on global gene expression as well as on histone variant H2A.Z localization. Instead, fibroblasts display slower cell growth and an increase in cell death coincident with large-scale chromosome misalignment in mitosis and large polylobed or micronuclei in interphase cells. Thus we conclude that H3.3 may additionally have an important under-explored role in chromosome segregation, nuclear structure and maintenance of genome integrity.

Keywords

H3.3, transcription, Mouse embryonic fibroblasts, RNA-seq, mitosis

Introduction

The replacement of canonical histones, core constituents of chromatin, by their variants drives chromatin dynamics and allows for functional and structural regulation of key cellular mechanisms (Boulard et al. 2007, Henikoff 2008). In contrast with replication dependent (RD) histones, the deposition of replication independent (RI) histones occurs in a cell cycle independent manner (Henikoff et al. 2004). There are RI variants for all the RD histones except for histone H4 (Hake et al. 2005, Pusarla and Bhargava 2005). The histone H3 variant H3.3 differs from its RD counterparts H3.1 and H3.2 by only 5 or 4 amino-acids respectively (Szenker et al. 2011). Murine H3.3 is coded by two genes;

H3f3a on chromosome 1 and *H3f3b* on chromosome 17, that despite variable mRNA sequences encode identical proteins (Krimer et al. 1993).

H3.3 containing nucleosomes, especially when associated with the histone H2A variant H2A.Z, may be promoting a less stable chromatin structure favorable to an active transcription state (Jin and Felsenfeld 2007, Jin et al. 2009). Accordingly, genome-wide localization studies have mapped H2A.Z and H3.3 at promoter regions and transcription start sites (TSS) and H3.3 density positively correlates with active transcription. However, H3.3 was also found associated with repressed genes and heterochromatic sites such as pericentric sites, telomeres and retroviral elements (Bargaje et al., 2012; Drané et al., 2010; Elsässer et al., 2015; Goldberg et al., 2010; Kraushaar et al., 2013). The differential deposition may reflect distinct targeting chaperone complexes. While the histone chaperone HIRA/ Ubinuclein (UBN1) targets H3.3 localization to TSS, the death domain associated protein (DAXX) together with its chromatin remodeling partner alpha thalassemia/mental retardation syndrome X-Linked protein (ATRX) are responsible for H3.3 deposition at pericentric heterochromatin and telomeres (Drané et al. 2010, Goldberg et al. 2010, Szenker et al. 2011, Daniel Ricketts et al. 2015).

Due to its presence at promoters and its higher enrichment at TSS and gene bodies of highly expressed genes (Goldberg et al. 2010), H3.3 has been strongly associated with a role in active transcription. It would thus be expected that transcription levels be highly impacted in the absence of H3.3. Intriguingly however, studies in the developing embryo showed that H3.3 depletion appears to have limited effect on gene transcription, while instead it seems to play an important role in the maintenance of genomic integrity (Bush et al. 2013, Jang et al. 2015). There is an increase in mitotic defects and consequent aneuploidy as well as DNA damage in H3.3B knock-out mouse embryonic fibroblasts (MEF) (Bush et al. 2013) and H3.3 depleted embryonic stem cells

(ESCs) (Banaszynski et al. 2013), indicative of chromosome structure dysfunction. Heterochromatin at telomeric, centromere, and pericentromeric repeat sequences presented a more open structure in the absence of H3.3 indicating a role for H3.3 in chromatin compaction (Jang et al. 2015).

Using a conditional gene-targeting strategy, we knocked out the histone *H3f3b* gene in MEFs and then using short hairpin RNAs against H3.3A mRNA in these cells, we derived cell lines that were essentially completely depleted of H3.3 expression. Deep RNA sequencing identified a set of nearly 800 genes that were mildly either up- or down-regulated in H3.3 depleted fibroblasts. Our results indicate that the loss of H3.3 has minimal impact on H2A.Z localization at the TSS and overall transcriptional rates. Importantly, H3.3 knockout cells display serious defects in mitotic progression including chromatin bridges in anaphase and misaligned chromosomes in metaphase and limits cell proliferation.

Materials and methods

Mouse strains

The FH-H3.3Bb mutant mouse line was established at the Phenomin-iCS (Phenomin - Institut Clinique de la Souris-, Illkirch, France; <http://www.ics-mci.fr/en/>). The targeting vector was constructed as follows. A 0.5 kb fragment encompassing exon 2 was amplified by PCR (from 129S2/SvPas ES cells genomic DNA) and subcloned in an iCS proprietary vector. This iCS vector contains a LoxP site as well as a floxed and flipped Neomycin resistance cassette. A DNA element encoding the FLAG-FLAG-HA amino acids was inserted in frame with the N-terminus of H3.3B. A 4.5 kb fragment (corresponding to the 5' homology arm) and 3.5 kb fragment (corresponding to the 3' homology arms) were amplified by PCR and subcloned. The linearized construct was electroporated in

129S2/SvPas mouse embryonic stem (ES) cells. After selection, targeted clones were identified by PCR using external primers and further confirmed by Southern blot with 5' and 3' external probes. 2 positive ES clones were injected into C57BL/6N blastocysts, and the male chimaeras derived gave germline transmission. Mice were housed in the Plateforme de haute Technologie Animale (PHTA, Grenoble, France) mouse facility (agreement number C 38 516 10001, registered protocol n° 321 at ethical committee C2EA-12).

Cell culture

Mouse embryonic fibroblasts (MEFs) were derived from E13.5 embryos. Heads and internal organs were removed and the torso was minced into chunks of tissue. Cells were cultured in DMEM high glucose, sodium pyruvate, Glutamax (Gibco) with 10% FBS and penicillin/streptomycin in a humidified incubator at 37°C and a 5% CO₂ atmosphere. Cells were maintained in culture using the 3T3 protocol (Xu 2005). *H3f3b*^{fl/fl} MEFs were infected with adenovirus expressing Cre recombinase (Ad-CMV-iCre, Vector Biolabs, Philadelphia, PA) to disrupt the endogenous *H3f3b* allele. Virus was diluted in serum-free DMEM at a multiplicity of infection of 500. Infection medium was replaced with fresh complete medium the next day and cells were analyzed for knock-out efficiency after 3 days.

Virus production and infections

293T cells were cotransfected with pLP1, pLP2, pLP3 and pLKO.1-shRNA vectors at a ratio of 1:0.5:0.6:2 respectively using Lipofectamine 2000 reagent (Invitrogen). shRNA against H3.3A mRNA (Dharmacon GE, TRCN0000012026) to knock-down H3.3A or a scrambled control shRNA as control (shControl) was used. Transfection medium was

replaced the next day with fresh complete medium. Viral supernatant was collected 48h post transfection and filtered through 0.45um filters. H3.3B KO MEFs at ~70% confluence were infected with the viral supernatant. Infection medium was replaced the next day and cells were selected with 3µg/ml puromycin. Cells were analyzed for knock-down efficiency 3 days later.

qPCR

For gene expression analysis, total RNA was extracted with TRIzol (Life Technologies) and reverse-transcribed with Superscript II (Invitrogen) and random hexamer mix. Ribosomal protein S9 was used as the reference gene. Takara SYBR qPCR Premix Ex Taq (Tli RNaseH Plus) and LightCycler 480 (Roche) real-time system were used. qPCR cycling conditions were as follows: (3min 95°C [10s 95°C, 30s 60°C]x40).

The gene specific primers used were: Rps9_F 5'TTGTCGCAAAACCTATGTGACC3', Rps9_R 5'GCCGCCTTACGGATCTTGG3', H3f3a_F 5'ACAAAAGCCGCTCGCAAGAG3', H3f3a_R 5'ATTCTCGCACCAGACGCTG3', H3f3b_F 5'TGGCTCTGAGAGAGATCCGTCGTT, H3f3b_R 5'GGATGTCTTTGGGCATGATGGTGAC3', Gadd45a_F 5'TGCTGCTACTGGAGAACGAC3', Gadd45a_R 5'TCCATGTAGCGACTTTCCCG3', PDGFb_F 5'GAGTCGGCATGAATCGCTG3', PDGFb_R 5'GCCCCATCTTCATCTACGGA3', Edn1_F 5'CCCACTCTTCTGACCCCTTT3', Edn1_R 5'GGCTCTGCACTCCATTCTCA3', Gas2_F 5'GCCGAGATTTGGGAGTTGAT3', Gas2_R 5'GCTTTATCAGACCAGGAGGC3', Seh1l_F 5'ATGACGGCTGTGTTAGGTTGT3', Seh1l_R 5'TACTCAGCTGTGCTTTCTGCT3', Smad6_F 5'GCCACTGGATCTGTCCGATT3', Smad6_R 5'GGTCGTACACCGCATAGAGG3',

Western blot

Cells were plated in 6-well plates and collected in 200µl of 2x Laemmli sample buffer (4% SDS, 20% Glycerol, 125mM Tris-HCl pH 6.8, 10% β-mercaptoethanol, 0.02% bromophenol blue). After brief sonication to fragment DNA, 20µl of sample was loaded and separated on 15% SDS-PAGE. Proteins were detected using anti-FLAG (1:2000, Sigma-Aldrich F3165) and anti-H4 (1:5000, Abcam ab10158)

RNA-seq

After isolation of total cellular RNA from subconfluent MEFs using TRIzol reagent, libraries of template molecules suitable for strand specific high throughput DNA sequencing were created using “TruSeq Stranded Total RNA with Ribo-Zero Gold Prep Kit” (# RS-122-2301, Illumina). Briefly, starting with 300 ng of total RNA, the cytoplasmic and mitochondrial ribosomal RNA (rRNA) were removed using biotinylated, target-specific oligos combined with Ribo-Zero rRNA removal beads. Following purification, the RNA was fragmented into small pieces using divalent cations under elevated temperature. The cleaved RNA fragments were copied into first strand cDNA using reverse transcriptase and random primers, followed by second strand cDNA synthesis using DNA Polymerase I and RNase H. The double stranded cDNA fragments were blunted using T4 DNA polymerase, Klenow DNA polymerase and T4 PNK. A single ‘A’ nucleotide was added to the 3’ ends of the blunt DNA fragments using a Klenow fragment (3’ to 5’ exo minus) enzyme. The cDNA fragments were ligated to double stranded adapters using T4 DNA Ligase. The ligated products were enriched by PCR amplification (30s at 98°C; [10s at 98°C, 30s at 60°C, 30s at 72°C] x 12 cycles; 5min at 72°C). Then surplus PCR primers were removed by purification using AMPure XP beads (Agencourt

Biosciences Corporation). Final cDNA libraries were checked for quality and quantified using 2100 Bioanalyzer (Agilent). The libraries were loaded in the flow cell at a concentration of 7pM, and clusters were generated in the Cbot and sequenced in the Illumina Hiseq 2500 as single-end 50 base reads following Illumina's instructions. Image analysis and base calling were performed using RTA 1.17.20 and CASAVA 1.8.2. Reads were mapped onto the mm9 assembly of the mouse genome by using Tophat (Trapnell et al. 2009) and the bowtie aligner (Langmead et al. 2009). Quantification of gene expression was performed using HTSeq (Anders et al. 2015) and gene annotations from Ensembl release 67. Read counts have been normalized across libraries with the statistical method proposed by Anders and Huber (Anders et al. 2010) and implemented in the DESeq Bioconductor library. Resulting p-values were adjusted for multiple testing by using the Benjamini and Hochberg method (Hochberg and Benjamin 1990).

The RNA-seq datasets (raw data as well as processed expression datasets) obtained in MEFs have been deposited in the Gene Expression Omnibus (GEO) under the accession number GSE84308.

Repeat analysis

Repeat analyses of RNA-seq datasets were performed as follows. Reads were aligned to repetitive elements in two passes. In the first pass, reads were aligned to the non-masked mouse reference genome (NCBI37/mm9) using BWA v0.6.2 (Li and Durbin 2009). Positions of the reads uniquely mapped to the mouse genome were cross-compared with the positions of the repeats extracted from UCSC (RMSK table in UCSC database for mouse genome mm9) and reads overlapping a repeat sequence were annotated with the repeat family. In the second pass, reads not mapped or multi-mapped

to the mouse genome in the previous pass were aligned to RepBase v18.07 (Jurka et al. 2005) repeat sequences for rodent. Reads mapped to a unique repeat family were annotated with their corresponding family name. Finally, we summed up the read counts per repeat family of the two annotation steps. Data were normalized based upon library size. Difference of repeat read counts between samples was expressed as the $\log_2$ -ratio (shH3.3A / shControl). The statistical significance of the difference between samples was assessed using the Bioconductor package DESeq. Processed datasets were restricted to repeat families with more than 100 mapped reads per RNA sample to avoid over- or underestimating fold enrichments due to low sequence representation.

Chromatin immunoprecipitation

ChIP experiments were performed from 15 cm dishes of subconfluent MEFs.

For H2A.Z ChIP, cells were crosslinked with 1% paraformaldehyde for 7min at room temperature. The reaction was stopped by adding glycine to a final concentration of 0.125 M for 5 min. Input chromatin was diluted 1:10 for a final ChIP buffer composition of 20mM Tris, pH 8.0, 150mM NaCl, 2mM EDTA, 0.1% SDS, 1% Triton-X. 5 μ g of anti-IgG control (Abcam ab46540) or polyclonal anti-H2A.Z antibodies were added and incubated over-night on a rotary shaker at 4°C. A mix of 8 μ l of magnetic protein A and 8 μ l magnetic protein G beads (Dynabeads, Lifetechnology) were washed in ChIP buffer, resuspended in the original volume and added to the ChIP samples for 4-6 hours (rotary shaker, 4 °C). The magnetic beads were collected on a magnetic rack and washed for 5min with ChIP buffer, 5min with Wash buffer II (20mM Tris, pH 8.0, 500mM NaCl, 2mM EDTA, 0.1% SDS, 1% Triton-X) and 5min with Wash Buffer III (10mM Tris-HCl, pH 8.0, 1mM EDTA, 0.25M LiCl, 1% NP-40, 1% Sodium-deoxycholate (Na-DOC)). Immunoprecipitated

material was eluted twice in 100µl of elution buffer (SDS 1%, 0.1 M NaHCO₃) for 15 min at room temperature and crosslinks were reversed by incubating at 65 °C overnight.

For HA-H3.3 ChIP, 100µg of nuclei were digested with 2U of Micrococcal nuclease S7 (Roche 10107921001) in digestion buffer (10mM Tris-HCl, pH 7.5, 3mM CaCl₂, 1x Protease inhibitors) for 12min at 37°C. 20µg of input chromatin was diluted in native chip buffer at a final concentration of 10mM Tris-HCl pH 7.5, 80mM NaCl, 1mM EDTA, 0.5% Triton-X. 20µl of anti-HA affinity matrix (Roche 118150160001) was added and incubated at 4°C overnight. Chromatin isolated from wild-type MEFs with no HA epitope was used as negative control. Beads-immunoprecipitated material were washed on rotary shaker with 10mM Tris-HCl pH 7.5, 80mM-150mM NaCl, 1mM EDTA, 0.5% Triton-X prior to elution and qPCR analysis.

Primers used for TSS analysis by qPCR were: Gadd45a-TSS_F 5'TTTCCGCTCAACTCTGCCTT3', Gadd45a-TSS_R 5'ACTCTGCACTGCTGCCTC3', Pdgfb-TSS_F 5'AGCTCTGCGCTTTCTGATCT3', Pdgfb-TSS_R 5'GATGGTTCGTCTTCACTCGC3', Edn1-TSS_F 5'AACTAATCTGGTTCCTCCGCC3', Edn1-TSS_R 5'GAGGTGGGGCTGATCATTGT3', Gas2-TSS_F 5'GTTACTAGAAAGCTCATGCCACT3', Gas2-TSS_R 5'CCCAAACACTAAGCTAAGACAGA3', Seh1l-TSS_F 5'TCATCACTGACTGCTGCTTC3', Seh1l-TSS_R 5'CTTAGGAATGATGGGGACGC3', Smad6-TSS_F 5'ATATCCTTCTGGGTCTTGCCA3', Smad6-TSS_R 5'GCTCAAGGGTGTGTCAGCAAAA3'

Immunofluorescent staining

Cells were fixed in formalin solution (Sigma-Aldrich) for 15min at 37°C, permeabilized with 0.2% Triton-X and incubated with lamin-B antibody (Santa-Cruz sc-6217) diluted

1/300 in 10% goat serum-PBS. Anti-goat IgG coupled with Cyanine 3 (Jackson 705-165-147) was used as secondary antibody. DNA was stained with the Hoechst 33342 (Invitrogen H3570) intercalating dye. All microscopy was performed on fixed cells with a Zeiss Axio Imager Z1 microscope with a Plan-Apochromat x63 objective. Z-stacks images were acquired with a Zeiss AxioCam camera piloted with the Zeiss Axiovision 4.8.10 software. All image treatment was performed using Fiji (ImageJ2-rc14) (Schindelin et al. 2012, 2015).

FACS analysis

Cells were trypsinized and fixed in cold 70% ethanol overnight. After wash with PBS, cells were stained with propidium iodide (PI) solution containing 5µg/µl PI (Sigma Aldrich) and 200µg/mL RNase A at 37°C for 30minutes. Approximately 100,000 cells per condition were analyzed using Accuri C6 (Becton Dickinson). The percentage of cells in each phase of the cell cycle was determined using ModFit 4.1 software.

Results

A combined genetic and shRNA strategy allows nearly complete depletion of H3.3 expression in mouse embryonic fibroblasts

Figure 1A describes the strategy used to generate the transgenic knock-in epitope tagged FLAG-FLAG-HA-H3.3B (FH-H3.3B) mice line. Pregnant transgenic knock-in animals were sacrificed on day 13.5 post fertilization and isolated embryos were used to derive mouse embryonic fibroblasts (MEFs) using the 3T3 protocol (Xu 2005) (**Figure 1B**). Cells were treated with a Cre expressing adenovirus to generate H3.3B knockout (KO) cells in culture. Single cells were selected and clonally expanded and genotyped to confirm the deletion of the *H3f3b* gene. Loss of H3.3B expression was verified by qPCR

as well as by immunoblotting for the Flag epitope (**Figures 1C, 1D**). H3.3B KO MEFs were then subsequently stably transfected with a scrambled control shRNA (shControl) or an shRNA targeting the coding sequence of H3.3A mRNA (shH3.3A). H3.3B knockout (H3.3B KO) cells transduced with control shRNA (shControl) displayed a fold increase in H3.3A mRNA levels while cells treated with shRNA against H3.3A (shH3.3A) showed a near complete knockdown (Kd) of H3.3A mRNA expression (**Figure 1C**). All cells displayed fibroblast like behavior and H3.3 depleted cell lines (H3.3B KO / H3.3A Kd) were slower to grow in culture as shown by the increase in their doubling time (**Figure 1E**).

H3.3 depletion has a mild effect on the global transcriptome

We then isolated total RNA from control shRNA treated H3.3B KO cells and from H3.3B KO / H3.3A Kd MEFs and carried out Ribozero RNA-seq analysis. RNA-seq analysis was carried out on pooled datasets from biological duplicate experiments. As seen in qPCR experiments (**Figure 1C**), H3.3A expression was significantly reduced in the H3.3B KO / H3.3A Kd MEFs cells (**Figure 2A**). A large fraction of the expressed genome does not change at the transcriptomic level in H3.3 depleted cells. Around 4% of transcribed genes in MEFs (800 genes in total) are mildly misregulated in H3.3 depleted cells with ~400 of them displaying upregulated expression and another 400 with downregulated expression with respect to the control cells ($P < 0.05$ cutoff) (**Figure 2B**).

H3.3 was recently implicated in maintenance of the silent state of endogenous retroviruses (ERVs) in the mouse genome (Elsässer et al. 2015). However, no significant changes could be seen in the global transcription of DNA repeats such as retro-elements (including long terminal repeats (LTR), ERV, long (LINE) and short (SINE) interspersed

nuclear elements) or tandem repeats (including major satellites, telomeres and microsatellites), which account for more than 40% of the mouse genome (**Figure 2C**). Functional clustering of differentially expressed genes indicated significant downregulation of genes implicated in lipid and sterol processing while factors involved in cell adhesion and motility were upregulated (**Figure 2D**). There was also dysregulation in the expression of genes involved in cell cycle progression (**Figure 2E**). We compared our gene expression profile to that of the H3.3 null embryos in a *p53*-null background (Jang et al. 2015) and to H3.3B KO/H3.3A Kd ESCs (Banaszynski et al. 2013). Even though all three studies indicate very little impact on global transcription, an overlap between the datasets was essentially non-existent and limited to 9 genes in comparison to H3.3 depleted embryos and 6 genes in comparison to H3.3 depleted ESCs (**Figure 2F**).

H2A.Z variant dynamics are not affected at differentially expressed genes in H3.3 depleted cells

We then selected 6 genes that were either up- or down-regulated in H3.3B-KO compared to H3.3B KO / H3.3A Kd fibroblasts and validated their gene expression levels by RT-qPCR analysis (**Figure 3A**). While the change in gene expression was about 50% for the selected genes (as seen in the RNA-seq analysis), with the exception of *Gadd45a* which was upregulated nearly 7 times, most other genes (*Pdgfb*, *Edn1*) showed a change in expression of 2-3 fold. To determine the extent to which H3.3 is present at the transcription start sites (TSS) of these 6 candidate genes, we isolated chromatin from FH-H3.3BMEFs and performed native chromatin immunoprecipitation (ChIP) experiments using an antibody against the HA tag. ChIP on isolated mononucleosomes indicated that H3.3 was highly enriched relative to input DNA at all of the selected candidate TSS

regions (**Figure 3B**). Mononucleosomes from wild-type mouse fibroblasts, used as a control for ChIP to rule out non-specific interaction with the antibody, showed essentially no enrichment at the TSS. Thus, while the presence of H3.3 at the transcription start sites is thought to positively correlate with transcription, its depletion from cells has a muted impact on overall transcription rates.

Histone variant H2A.Z also marks the TSS of genes and the presence of dual H3.3-H2A.Z variant nucleosomes is thought to positively regulate inducible transcription (Jin and Felsenfeld 2007, Henikoff 2008, Jin et al. 2009, Obri et al. 2014). We then tested if the deposition of the histone variant H2A.Z at the TSS of these selected candidate genes was affected by the depletion of H3.3 from the fibroblasts. Crosslinked and sonicated chromatin was generated from either WT, H3.3B KO cells or from H3.3 depleted (H3.3B KO / H3.3A Kd) cells and H2A.Z deposition was determined relative to a control IgG (**Figure 3C**). At all the TSS examined, loss of H3.3 did not have a significant impact on H2A.Z deposition regardless of the transcriptional changes of the corresponding genes.

H3.3 depletion results in defective mitotic progression

Following the shRNA mediated depletion of H3.3A and in stark comparison to control cells, we observed an increase in the rate of cell doubling time in culture as well as increased rate of cell death. Depletion of H3.3 from fibroblasts resulted in cells with larger nuclei, increased number of micronuclei and polylobed nuclear structures in interphase cells (**Figures 4A and 4B**). Loss of H3.3B alone led to a doubling of the number of interphase nuclear defects while complete depletion of H3.3 led to a near 3 time increase of polylobed and micronuclei appearance (**Figure 4C**). While there was nearly a 50% reduction in the number of cells entering mitosis in H3.3 depleted fibroblasts, in addition, these cells displayed marked defects in chromosome alignment

at the metaphase plate, lagging chromosomes in anaphase as well as telophase bridges, indicating chromosome structure dysregulation (**Figures 4B, 4D and 4E**). Finally, FACS analyses indicated that as compared to wild-type or control FH-H3.3B cells, H3.3B KO and H3.3A Kd/H3.3B KO depleted cells showed progressively increased residence time in G0/G1 phase of the cell cycle and relatively fewer cells in S-phase or mitosis (**Figure 4E**).

Discussion and conclusion

Recent research by Jang et al. on H3.3 function in the developing embryo have surprisingly shown that the loss of H3.3 function has very little impact on global transcription, with most affected genes being up- or downregulated by not much greater than 2-fold (Jang et al. 2015). In this study, we extend our analysis to mouse embryonic fibroblasts and intriguingly observe very similar results to the study by Jang et al. - *i.e.* little to no effect on global transcription. Using a novel transgenic 'conditional knock-in/knockout' mouse model for H3.3B, we generated mouse fibroblasts that yielded H3.3B knockout cells upon Cre expression. Further expression of a specific and efficient shRNA against H3.3A resulted in almost complete depletion of H3.3 expression from fibroblasts. We then performed genome-wide transcriptome analysis and found very few genes (~4% of transcribed genes) showing significant changes in gene expression. While this result is similar to what was described in embryos, the ensemble of the affected transcriptome shows little to no overlap between the two experimental studies. Furthermore, our results suggest that the changes at the transcriptional level in H3.3 knockout cells do not impact histone variant H2A.Z presence at the TSS of the selected genes. Thus, despite the positive correlation between the transcriptional rate of genes and H3.3 accumulation at their corresponding TSS, the global effect on transcription observed upon H3.3 depletion

is suggestive of a facilitative role rather than that of an essential positive regulatory factor. This also correlates with what has been observed in the past in mouse embryos as well as stem cells (Wong et al. 2010, Banaszynski et al. 2013, Jang et al. 2015). Recent data has demonstrated a silencing role for H3.3 by localization at endogenous retroviral elements (ERVs) in ESCs (Elsässer et al. 2015). Our study has not revealed an impact on ERV transcription upon H3.3 depletion in MEFs. The role of H3.3 in retrotransposon silencing may therefore be specific to pluripotent cells and replaced by other mechanisms during differentiation.

Thus, our study supports an understudied key role of H3.3 in maintenance of genome integrity. H3.3 knockout ES cells and mouse fibroblasts show dramatic defects in mitosis with 3-4 times increase in the number of defects such as lagging chromosomes, anaphase bridges while the total number of cells entering mitosis is much lower in H3.3 null cells. We also observed significant changes to the nuclear matrix structure in interphase cells with the appearance of many polylobed and micronuclei in H3.3 depleted cells. Unlike Jang et al, wherein the impact of H3.3 loss was studied in a p53 null background, our cells are completely depleted for H3.3B and H3.3A is knocked-down with greater than 90% efficiency. However, the similarity of the results indicates a key role for H3.3 in regulating faithful chromosome segregation during mitosis and perhaps in maintaining nuclear architecture in interphase.

In conclusion, we depleted H3.3 expression in mouse embryonic fibroblasts combining genetic and shRNA strategies. This near complete depletion of H3.3 from mammalian fibroblasts affected transcription at a handful of genes, while global transcription rates were altered only about 2-fold with no effect seen at all on expression of retroviral repeat elements. Instead, we showed that H3.3 plays an important role in

faithful completion of the cellular mitotic program and maintaining genomic integrity.

Abbreviations

TSS: transcription start sites **MEF:** mouse embryonic fibroblasts **ESC:** embryonic stem cells **KO:** knock-out **Kd:** knock-down **ERV:** endogenous retroviral elements **LTR:** long terminal repeats **LINE and SINE:** long and short interspersed nuclear element **RNA-seq:** ribonucleic acid sequencing **RT-qPCR:** reverse transcription quantitative polymerase chain reaction **rRNA:** ribosomal RNA **s.e.m.:** standard error of the mean **PI:** propidium iodide **WT:** wild-type

Acknowledgements

This work was supported by institutional funds from the Université de Strasbourg-UDS, the Université de Grenoble Alpes-UGA, the Centre National de la Recherche Scientifique-CNRS, the Institut National de la Santé et de la Recherche Médicale- INSERM (Plan Cancer), and by grants from : the Institut National du Cancer-INCA (INCa_4496, INCa_4454 and INCa PLBIO15-245), the Fondation pour la Recherche Médicale-FRM (DEP20131128521), the Université de Strasbourg Institut d'Etudes Avancées (USIAS-2015-42) (A.H), the Association pour la Recherche sur le Cancer, La Ligue Nationale contre le Cancer (Équipe labellisée to A.H). K.P. was supported by a fellowship from Ligue Comité de l'Isère (R15026CC), and the ATIP-AVENIR installation grant. A.O. was awarded the European Molecular Biology Organization (EMBO) short-term fellowship, Scientific and Technological Research Council of Turkey (TUBITAK) 2214/A doctoral research grant and French Ministry of Foreign Affairs scholarship. Sequencing was performed by the IGBMC Microarray and Sequencing platform, a member of the 'France Génomique' consortium (ANR-10-INBS-0009).

References

Anders, S., Huber, W., Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M., Snyder, M., Mortazavi, A., Williams, B., McCue, K., Schaeffer, L., Wold, B., Robertson, G., Hirst, M., Bainbridge, M., Bilenky, M., Zhao, Y., Zeng, T., Euskirchen, G., Bernier, B., Varhol, R., Delaney, A., Thiessen, N., Griffith, O., He, A., Marra, M., Snyder, M., Jones, S., Licatalosi, D., Mele, A., Fak, J., Ule, J., Kayikci, M., Chi, S., Clark, T., Schweitzer, A., Blume, J., Wang, X., Darnell, J., Darnell, R., Smith, A., Heisler, L., Mellor, J., Kaper, F., Thompson, M., Chee, M., Roth, F., Giaever, G., Nislow, C., Marioni, J., Mason, C., Mane, S., Stephens, M., Gilad, Y., Wang, L., Feng, Z., Wang, X., Wang, X., Zhang, X., Robinson, M., Smyth, G., Whitaker, L., Robinson, M., McCarthy, D., Smyth, G., Robinson, M., Smyth, G., Cameron, A., Trivedi, P., Robinson, M., Oshlack, A., Loader, C., McCullagh, P., Nelder, J., Agresti, A., Engström, P., Tommei, D., Stricker, S., Smith, A., Pollard, S., Bertone, P., Morrissy, A., Morin, R., Delaney, A., Zeng, T., McDonald, H., Jones, S., Zhao, Y., Hirst, M., Marra, M., Kasowski, M., Grubert, F., Heffelfinger, C., Hariharan, M., Asabere, A., Waszak, S., Habegger, L., Rozowsky, J., Shi, M., Urban, A., Hong, M., Karczewski, K., Huber, W., Weissman, S., Gerstein, M., Korb, J., Snyder, M., Benjamini, Y., Hochberg, Y., Bullard, J., Purdom, E., Hansen, K., Dudoit, S., Bloom, J., Khan, Z., Kruglyak, L., Singh, M., Caudy, A., Smyth, G., Smyth, G., Lönnstedt, I., Speed, T., Gentleman, R., Carey, V., Bates, D., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C., Maechler, M., Rossini, A., Sawitzki, G., Smith, C., Smyth, G., Tierney, L., Yang, J., Zhang, J., Bliss, C., Fisher, R., Clark, S., Perry, J., Lawless, J., Saha, K., Paul,

- S., Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. 2010. Differential expression analysis for sequence count data. *Genome Biol.* **11**(10): R106. doi:10.1186/gb-2010-11-10-r106.
- Anders, S., Pyl, P.T., and Huber, W. 2015. HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**(2): 166–169. doi:10.1093/bioinformatics/btu638.
- Banaszynski, L.A., Wen, D., Dewell, S., Whitcomb, S.J., Lin, M., Diaz, N., Chapgier, A., Goldberg, A.D., Canaani, E., Rafii, S., Zheng, D., Elsa, S.J., and Allis, C.D. 2013. Facilitates PRC2 Recruitment at Developmental Loci in ES Cells. **5**: 107–120. doi:10.1016/j.cell.2013.08.061.
- Bargaje, R., Alam, M.P., Patowary, A., Sarkar, M., Ali, T., Gupta, S., Garg, M., Singh, M., Purkanti, R., Scaria, V., Sivasubbu, S., Brahmachari, V., and Beena Pillai. 2012. Proximity of H2A.Z containing nucleosome to the transcription start site influences gene expression levels in the mammalian liver and brain. *Nucleic Acids Res.* **40**(18): 8965–8978. doi:10.1093/nar/gks665.
- Boulard, M., Bouvet, P., Kundu, T.K., and Dimitrov, S. 2007. No TitleHistone variant nucleosomes: structure, function and implication in disease. *Subcell. Biochem.* **41**: 71–89.
- Bush, K.M., Yuen, B.T., Barrilleaux, B.L., Riggs, J.W., O'Geen, H., Cotterman, R.F., and Knoepfler, P.S. 2013. Endogenous mammalian histone H3.3 exhibits chromatin-related functions during development. *Epigenetics Chromatin* **6**(1): 7. doi:10.1186/1756-8935-6-7.
- Daniel Ricketts, M., Frederick, B., Hoff, H., Tang, Y., Schultz, D.C., Singh Rai, T., Grazia Vizioli, M., Adams, P.D., and Marmorstein, R. 2015. Ubinuclein-1 confers histone H3.3-specific-binding by the HIRA histone chaperone complex. *Nat.*

- Commun. **6**: 7711. doi:10.1038/ncomms8711.
- Dran??, P., Ouararhni, K., Depaux, A., Shuaib, M., and Hamiche, A. 2010. The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. *Genes Dev.* **24**(12): 1253–1265. doi:10.1101/gad.566910.
- Elsässer, S.J., Noh, K.-M., Diaz, N., Allis, C.D., and Banaszynski, L.A. 2015. Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells. *Nature* **522**(7555): 240–4. doi:10.1038/nature14345.
- Goldberg, A.D., Banaszynski, L.A., Noh, K.M., Lewis, P.W., Elsaesser, S.J., Stadler, S., Dewell, S., Law, M., Guo, X., Li, X., Wen, D., Chappier, A., DeKolver, R.C., Miller, J.C., Lee, Y.L., Boydston, E.A., Holmes, M.C., Gregory, P.D., Greally, J.M., Rafii, S., Yang, C., Scambler, P.J., Garrick, D., Gibbons, R.J., Higgs, D.R., Cristea, I.M., Urnov, F.D., Zheng, D., and Allis, C.D. 2010. Distinct Factors Control Histone Variant H3.3 Localization at Specific Genomic Regions. *Cell* **140**(5): 678–691. Elsevier Ltd. doi:10.1016/j.cell.2010.01.003.
- Hake, S.B., Garcia, B. a, Kauer, M., Baker, S.P., Shabanowitz, J., Hunt, D.F., and Allis, C.D. 2005. Serine 31 phosphorylation of histone variant H3.3 is specific to regions bordering centromeres in metaphase chromosomes. *Proc. Natl. Acad. Sci. U. S. A.* **102**(18): 6344–6349. doi:10.1073/pnas.0502413102.
- Henikoff, S. 2008. Nucleosome destabilization in the epigenetic regulation of gene expression. *Nat Rev Genet* **9**(1): 15–26. doi:nrg2206 [pii]\n10.1038/nrg2206.
- Henikoff, S., Furuyama, T., and Ahmad, K. 2004. Histone variants, nucleosome assembly and epigenetic inheritance. *Trends Genet.* **20**(7): 320–326. doi:10.1016/j.tig.2004.05.004.
- Hochberg, Y., and Benjamin, Y. 1990. More Powerful Procedures for Multiple

- Significance Testing. **9**(July 1988): 811–818.
- Jang, C., Shibata, Y., Starmer, J., Yee, D., and Magnuson, T. 2015. Histone H3 . 3 maintains genome integrity during mammalian development. *Genes Dev.* **1**: 1377–1392. doi:10.1101/gad.264150.115.GENES.
- Jin, C., and Felsenfeld, G. 2007. Nucleosome stability mediated by histone variants H3.3 and H2A.Z. *Genes Dev.* **21**(12): 1519–1529. doi:10.1101/gad.1547707.
- Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K., and Felsenfeld, G. 2009. H3.3/H2A.Z double variant-containing nucleosomes mark “nucleosome-free regions” of active promoters and other regulatory regions. *Nat. Genet.* **41**(8): 941–5. Nature Publishing Group. doi:10.1038/ng.409.
- Jurka, J., Kapitonov, V. V., Pavlicek, A., Klonowski, P., Kohany, O., and Walichiewicz, J. 2005. Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet. Genome Res.* **110**(1–4): 462–467. doi:10.1159/000084979.
- Kraushaar, D.C., Jin, W., Maunakea, A., Abraham, B., Ha, M., and Zhao, K. 2013. Genome-wide incorporation dynamics reveal distinct categories of turnover for the histone variant H3.3. *Genome Biol.* **14**(10): R121. doi:10.1186/gb-2013-14-10-r121.
- Krimer, D.B., Cheng, G., and Skoultchi, A.I. 1993. Induction of H3.3 replacement histone mRNAs during the precommitment period of murine erythroleukemia cell differentiation. *Nucleic Acids Res* **21**(12): 2873–2879. doi:10.1093/nar/21.12.2873.
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**(3): R25. doi:10.1186/gb-2009-10-3-r25.
- Li, H., and Durbin, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**(14): 1754–1760. doi:10.1093/bioinformatics/btp324.

- Obri, A., Ouararhni, K., Papin, C., Diebold, M.-L., Padmanabhan, K., Marek, M., Stoll, I., Roy, L., Reilly, P.T., Mak, T.W., Dimitrov, S., Romier, C., and Hamiche, A. 2014. ANP32E is a histone chaperone that removes H2A.Z from chromatin. *Nature* **505**(7485): 648–53. Nature Publishing Group. doi:10.1038/nature12922.
- Pusarla, R.H., and Bhargava, P. 2005. Histones in functional diversification: Core histone variants. *FEBS J.* **272**(20): 5149–5168. doi:10.1111/j.1742-4658.2005.04930.x.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P., and Cardona, A. 2012. Fiji: an open-source platform for biological-image analysis. *Nat Meth* **9**(7): 676–682. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved. Available from <http://dx.doi.org/10.1038/nmeth.2019>.
- Schindelin, J., Rueden, C.T., Hiner, M.C., and Eliceiri, K.W. 2015. The ImageJ ecosystem: An open platform for biomedical image analysis. *Mol. Reprod. Dev.* **82**(7–8): 518–529. doi:10.1002/mrd.22489.
- Szenker, E., Ray-Gallet, D., and Almouzni, G. 2011. The double face of the histone variant H3.3. *Cell Res* **21**(3): 421–434. Nature Publishing Group. doi:10.1038/cr.2011.14.
- Trapnell, C., Pachter, L., and Salzberg, S.L. 2009. TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics* **25**(9): 1105–1111. doi:10.1093/bioinformatics/btp120.
- Wong, L.H., Mcghie, J.D., Sim, M., Anderson, M.A., Ahn, S., Hannan, R.D., George, A.J., Morgan, K.A., Mann, J.R., and Choo, K.H.A. 2010. ATRX interacts with H3 . 3 in maintaining telomere structural integrity in pluripotent embryonic stem cells. :

351–360. doi:10.1101/gr.101477.109.20.

Xu, J. 2005. Preparation, Culture, and Immortalization of Mouse Embryonic Fibroblasts.

Curr. Protoc. Mol. Biol. **Chapter 28**: 1–8. doi:10.1002/0471142727.mb2801s70.

Figure Legends

Figure 1. Generation of H3.3 MEF models and H3.3 expression.

A. (i) Wild-type *H3f3b* gene structure. The open reading frame is indicated by black boxes (ii) A DNA element encoding the FLAG-FLAG-HA amino acids was inserted in frame with the N-terminus of H3.3B. In addition, a loxP site was inserted on both ends of Exon 2. A mouse line was derived from ES cells carrying the modified *H3f3b^{fl/fl}* allele (iii) Structure of the H3.3B KO, after Cre recombinase expression, which deletes exon 2 and generates loss of function (KO) allele, *H3f3b^{-/-}*. **B.** FH-H3.3B Mouse embryonic fibroblasts (MEFs) isolated from *H3f3b^{fl/fl}* embryos at 13.5 dpf. Immortalized MEFs were infected with Cre recombinase expressing adenovirus to generate loss of function allele. *H3f3b^{-/-}* MEFs were further infected with either a scrambled control shRNA (shControl) or one targeting the H3.3A mRNA (shH3.3A) to produce H3.3B knockout (KO) / H3.3A Knockdown (Kd) MEFs. **C.** Relative H3.3 RNA expression in MEFs. Quantitative RT-qPCR mRNA profiles normalized to ribosomal protein S9 (Rps9) mRNA and relative to the values obtained for the control FH-H3.3B MEFs set at $1.0 \pm \text{s.e.m.}$ One-way ANOVA, $P < 0.001$ for FH-H3.3B vs H3.3B KO and $P < 0.01$ for FH-H3.3B vs. H3.3B KO / H3.3A Kd, $n=4$. **D.** Western blot analysis of the loss of FH-H3.3B expression after Cre expression. **E.** Average doubling times in MEFs. 25,000 cells of each cell type were plated and counted at different times to calculate the doubling time. Error bars represent standard deviation, one-way ANOVA, $P < 0.001$ for FH-H3.3B vs. H3.3B KO / H3.3A Kd, $n=4$.

Figure 2. Genome-wide transcriptome analysis of MEFs in the absence of H3.3.

A. Bar graph representing H3.3A expression in control H3.3BKO and H3.3A Kd/H3.3B KO / H3.3A Kd MEFs. **B.** Scatter plots comparing gene expression profiles of control (shControl) and H3.3B KO / H3.3A Kd (shH3.3A) MEFs. Red dots indicate differentially expressed genes ($P < 0.05$ and $|\log_2 \text{fc}| > 1$). **C.** Scatter plots comparing global transcription of repeat families in H3.3B KO and H3.3B KO / H3.3A Kd MEFs. **D.** Functional annotation clustering of differentially expressed genes. **E.** List of the differentially expressed genes implicated in mitosis. **F.** Comparative analysis of the differentially expressed genes in H3.3 deficient embryos (Jang et al. 2015) and in H3.3 depleted ESCs (Banaszynski et al. 2013). Lists of genes that show similar patterns of up or downregulation in transcription are indicated.

Figure 3. H3.3 and H2A.Z enrichment at transcription start sites (TSS) of genes showing differential mRNA expression.

A. Quantitative PCR assay for gene expression in H3.3B KO and H3.3B KO / H3.3A Kd MEFs. RNA values are normalized to ribosomal protein S9 (Rps9) mRNA and are relative to H3.3B KO value set at $1.0 \pm \text{s.e.m.}$, two-tailed t-test, $P < 0.05$ for *Sehl1*, $P < 0.01$ for *Gadd45a*, *Pdgfb*, *Gas2*, *Smad 6*, $P < 0.001$ for *Edn1*, $n=4$. **B.** H3.3B enrichment at TSS of differentially transcribed genes by ChIP-quantitative PCR assay (ChIP-qPCR). Values are expressed in % of enrichment relative to input $\pm \text{s.e.m.}$. Chromatin from wild-type (WT) MEFs containing no HA-tag were used as a negative IP control. The TSS of examined genes were significantly enriched in H3.3B compared to the negative control, one-tailed t-test, $n=4$, $P < 0.01$ for *Pdgfb* and *Smad6*, $P < 0.001$ for *Edn1* and *Sehl1*, and $P < 0.0001$ for *Gadd45a* and *Gas2*. **C.** H2A.Z enrichment at TSS of differentially transcribed genes. ChIP-qPCR assay. Values are expressed in enrichment relative to input $\pm \text{s.e.m.}$. Rabbit

IgG was used as the control. The TSS of examined genes were significantly enriched in H2A.Z compared to the negative control, one-tailed t-test, $P < 0.05$, $n=3$. ChIP-qPCR graphs are representative of 3 separate experiments. H2A.Z enrichment at studied TSS did not vary significantly between samples (One-way ANOVA test, $P > 0.05$, $n=3$).

Figure 4. Mitotic defects in H3.3 deficient MEFs.

A. Representative images of control nuclei. **B.** Representative image of nuclear abnormalities with micronuclei and polylobed nuclei and of mitotic abnormalities with chromatin bridges, misaligned and lagging chromosomes observed in H3.3 deficient MEFs. Defects are indicated by yellow arrows. **C.** Nuclear abnormalities in control FH-H3.3B and H3.3 deficient MEFs. One-way ANOVA, $P < 0.05$ for FH-H3.3B vs. H3.3B KO and $P < 0.001$ for FH-H3.3B vs. H3.3B KO / H3.3A Kd MEFs, $n=4$. **D.** Mitotic abnormalities in control FH-H3.3B and H3.3 deficient MEFs. 125 mitotic events were scored for each line and experiment. Error bars represent the standard deviation between experiments on the total number of abnormalities. One-way ANOVA $P < 0.01$ for FH-H3.3B vs. H3.3B KO and $P < 0.0001$ for FH-H3.3B vs. H3.3B KO / H3.3A Kd MEFs, $n=3$. **E.** Flow cytometric analysis of cell cycle for wild-type (WT), FH-H3.3B, H3.3B KO and double H3.3B KO / H3.3A Kd cells treated with propidium iodide (PI). Approximately 100,000 cells per condition were analyzed per experiment. Error bars represent the standard deviation between experiments (One-way ANOVA, $P < 0.01$ for WT vs. H3.3B KO / H3.3A Kd, $n=3$).

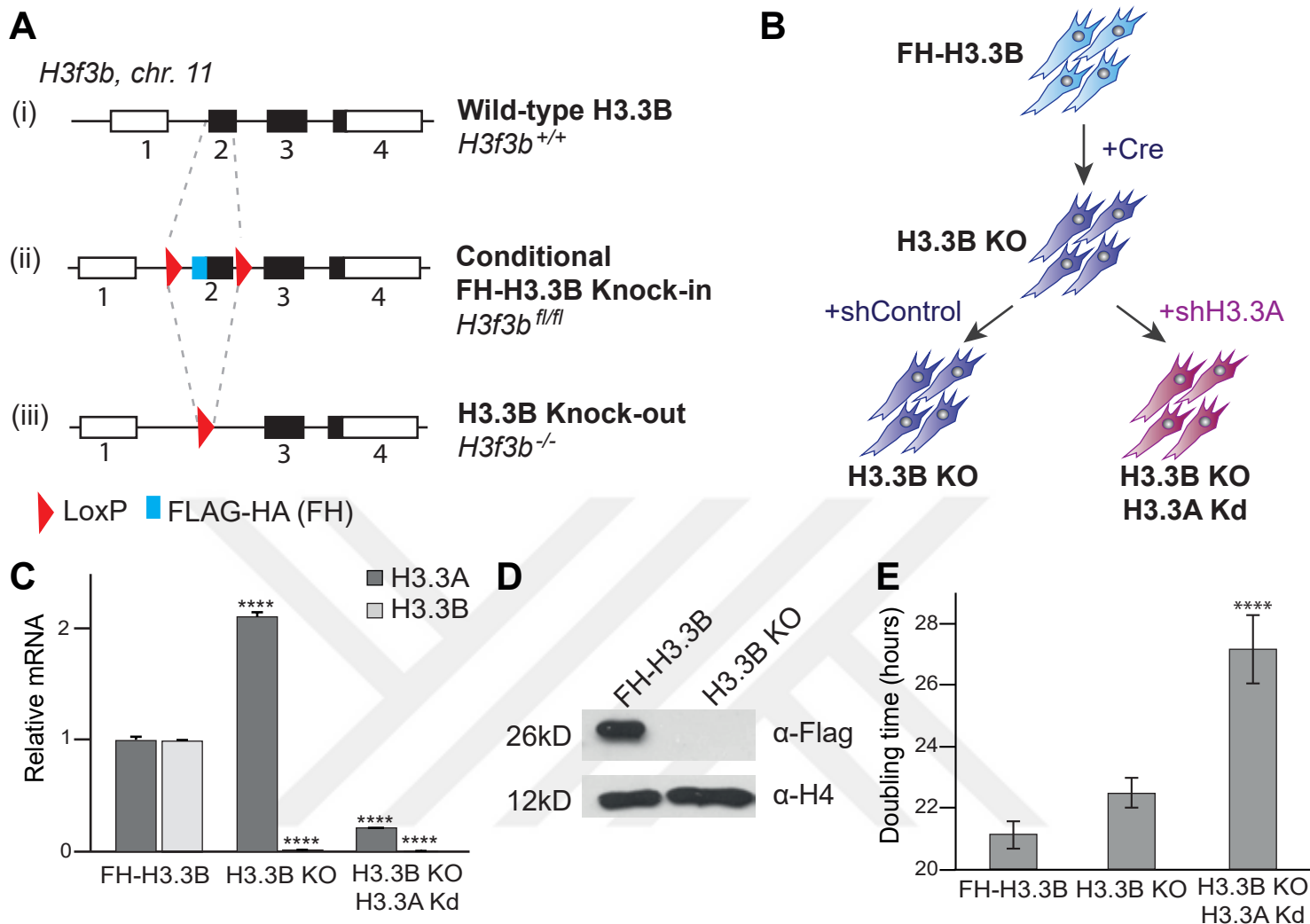
FIGURE 1

FIGURE 2

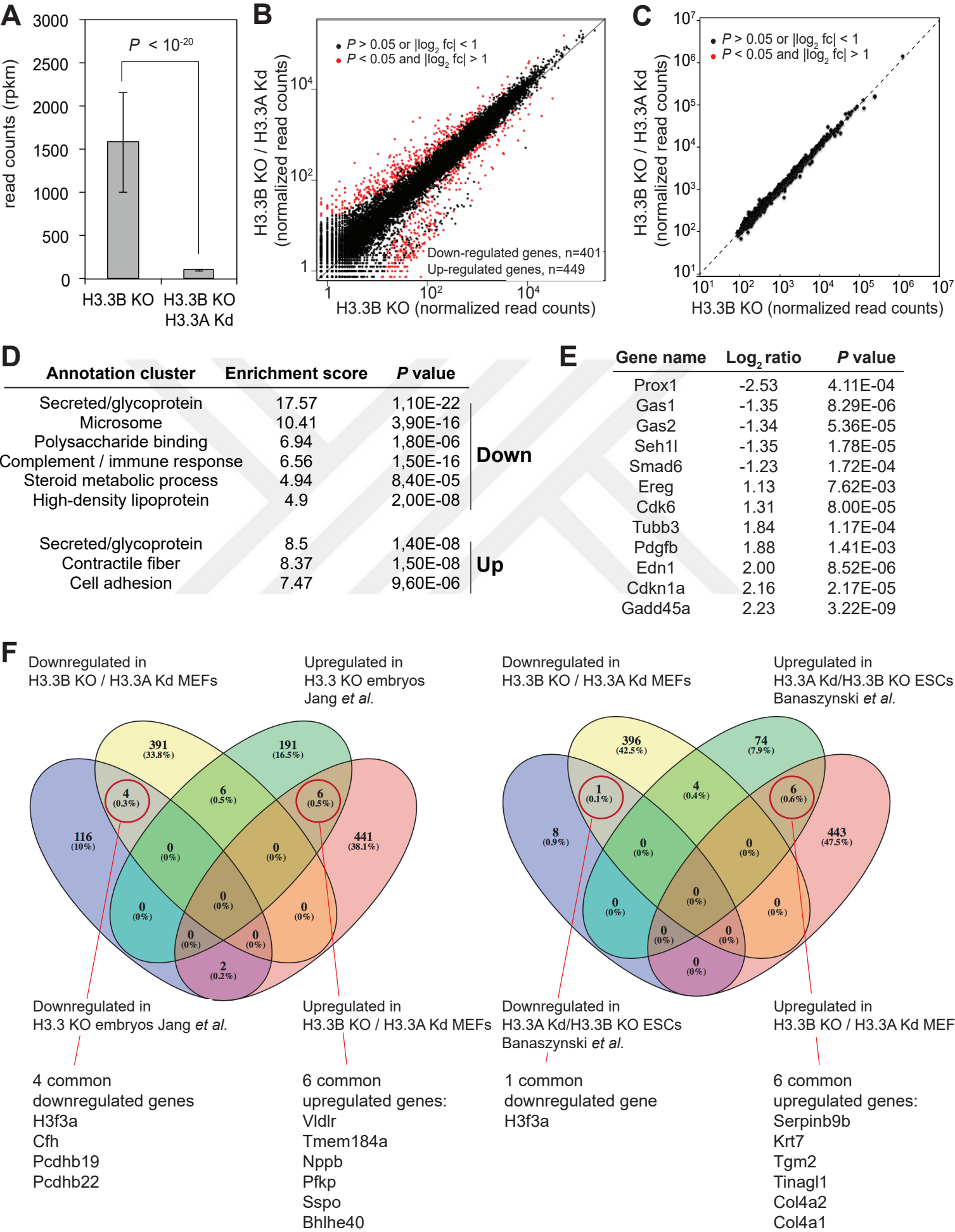


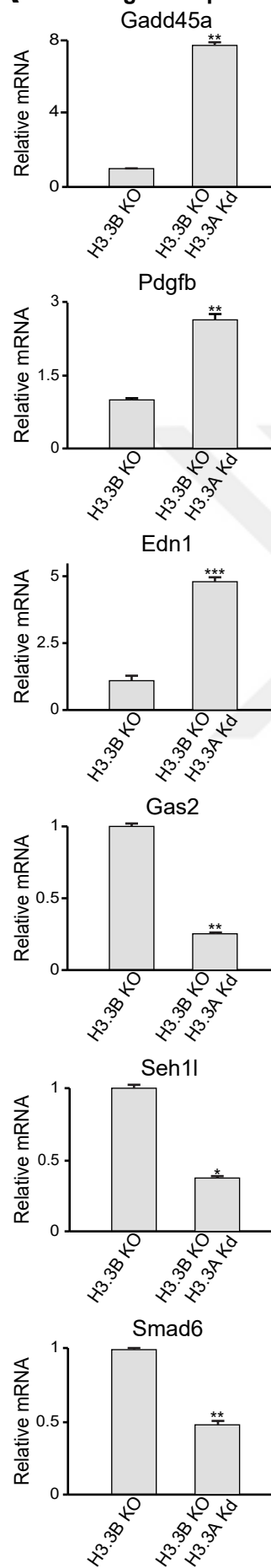
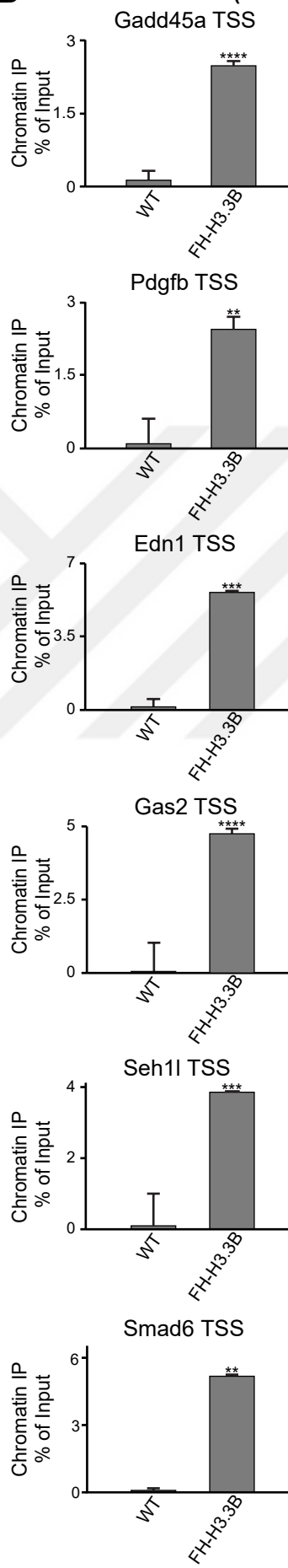
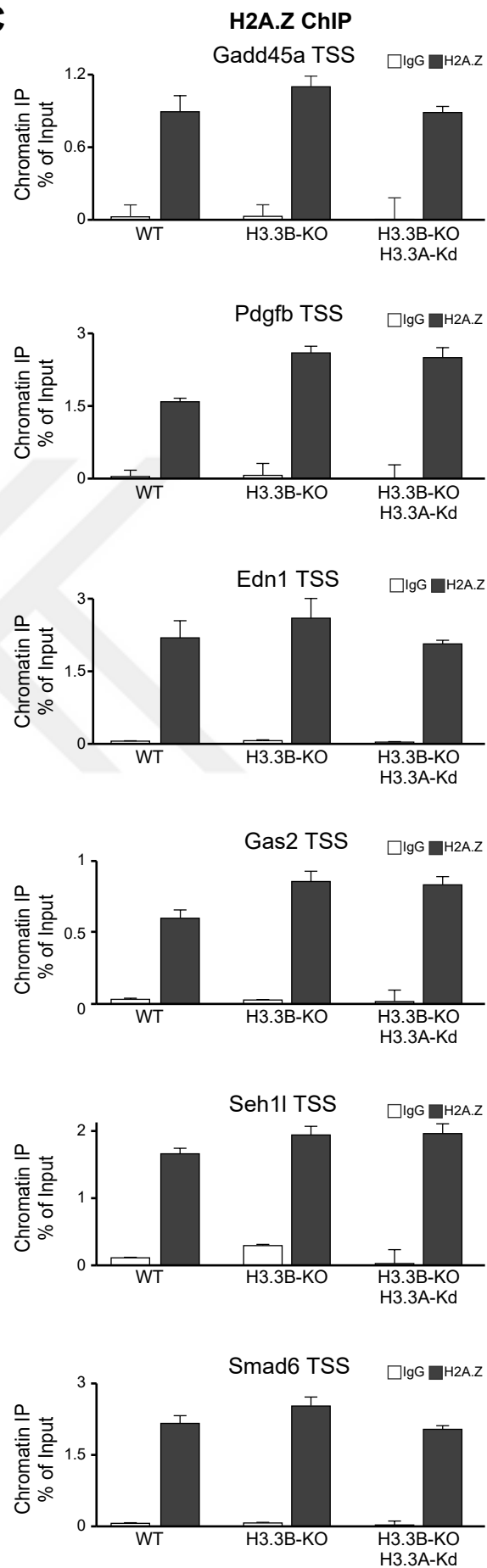
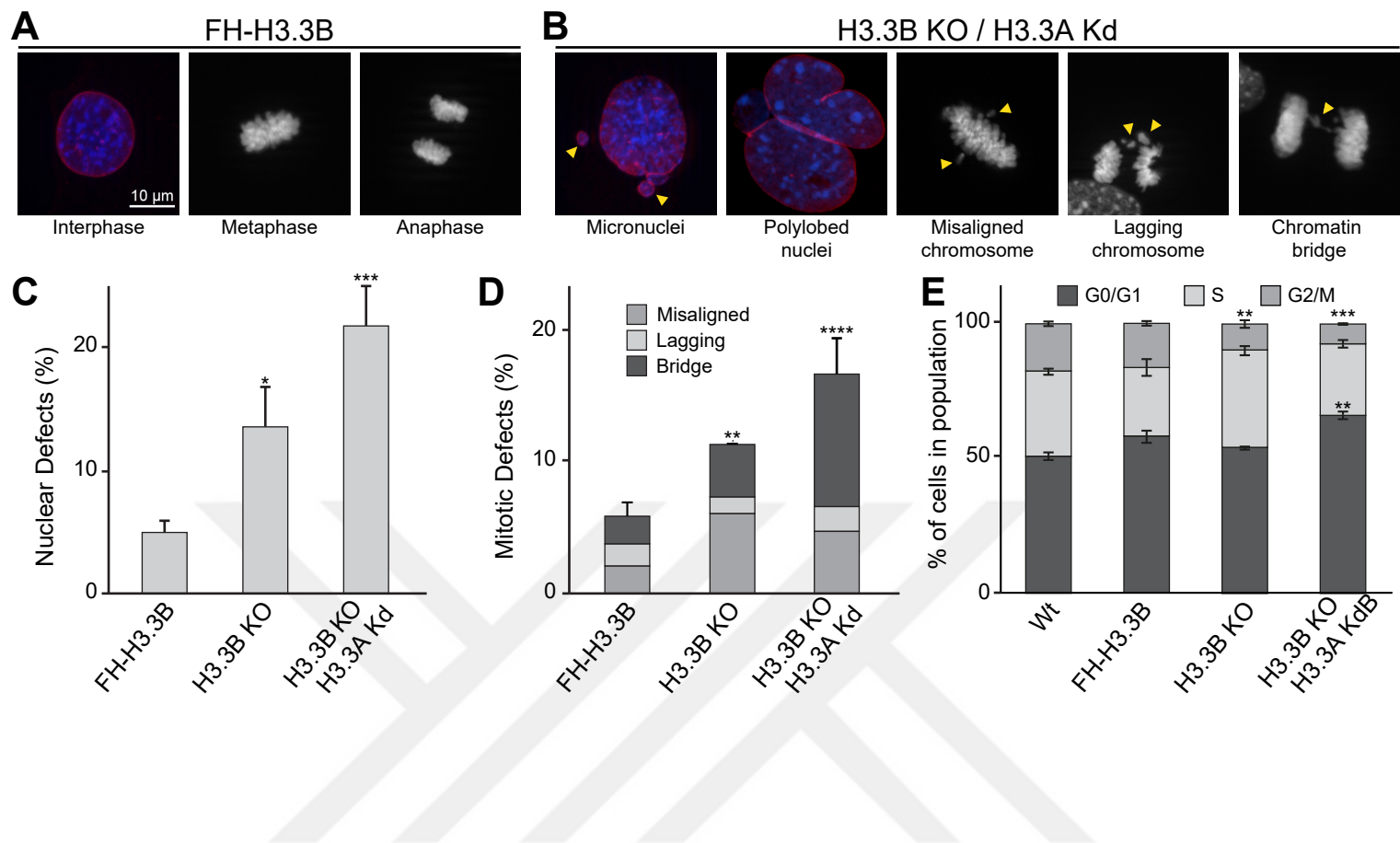
FIGURE 3**A Relative gene expression****B FH-H3.3B ChIP (α-HA)****C**

FIGURE 4





Genetics and epigenetics of liver cancer

Cigdem Ozen¹, Gokhan Yildiz^{1,2}, Alper Tunga Dagcan¹, Dilek Cevik^{1,2}, Aysegul Ors^{1,2}, Umur Keles¹, Hande Topel¹ and Mehmet Ozturk^{1,2}

¹ Bilkent University, BilGen Genetics and Biotechnology Center, Department of Molecular Biology and Genetics, 06800 Ankara, Turkey

² Université Joseph Fourier – Grenoble 1, INSERM Institut Albert Bonniot, U823, Site Santé-BP 170, 38042 Grenoble Cedex 9, France

Hepatocellular carcinoma (HCC) represents a major form of primary liver cancer in adults. Chronic infections with hepatitis B (HBV) and C (HCV) viruses and alcohol abuse are the major factors leading to HCC. This deadly cancer affects more than 500,000 people worldwide and it is quite resistant to conventional chemo- and radiotherapy. Genetic and epigenetic studies on HCC may help to understand better its mechanisms and provide new tools for early diagnosis and therapy. Recent literature on whole genome analysis of HCC indicated a high number of mutated genes in addition to well-known genes such as *TP53*, *CTNNB1*, *AXIN1* and *CDKN2A*, but their frequencies are much lower. Apart from *CTNNB1* mutations, most of the other mutations appear to result in loss-of-function. Thus, HCC-associated mutations cannot be easily targeted for therapy. Epigenetic aberrations that appear to occur quite frequently may serve as new targets. Global DNA hypomethylation, promoter methylation, aberrant expression of non-coding RNAs and dysregulated expression of other epigenetic regulatory genes such as *EZH2* are the best-known epigenetic abnormalities. Future research in this direction may help to identify novel biomarkers and therapeutic targets for HCC.

Introduction

The most frequent primary liver cancers are hepatocellular carcinoma (HCC) and cholangiocarcinoma in adults, and hepatoblastoma in children. More than 80% of liver tumours are HCCs [1]. This review will focus primarily on HCC, one of the most frequent cancers worldwide with more than 500,000 new cases observed each year. Almost the same number of deaths is observed because of this cancer could not be easily treated. The most efficient treatment for HCC is liver transplantation, provided that it is detected early enough. Surgical removal and chemo-embolisation of tumour nodules are other alternatives. These tumours are usually resistant to chemo- or radiotherapy [1–3]. Targeted therapy of HCC is in its infancy. The only clinically relevant drug is a kinase inhibitor, Sorafenib, has only a modest effect on patient survival [4].

The aetiology of HCC is well known. Chronic liver injury associated primarily with hepatitis B (HBV) and C (HCV) virus infection

constitutes the most important cause of HCC. Other factors, such as alcohol abuse and dietary exposure to aflatoxins, are also established causes, but their contribution to the disease aetiology is much less than the contributions of viral agents. The unprecedented increase in obesity rates in both developed and developing countries is a rising concern for HCC risk that may account for the unexpected increase in HCC incidence in the Western world [1].

Molecular mechanisms of hepatocellular carcinogenesis remain ill-defined, mainly due to disease heterogeneity. The heterogeneity of agents that cause chronic liver injury (HBV, HCV, aflatoxins and alcohol) and the ways they interact with the host DNA and epigenetic players are the most probable parameters contributing to HCC heterogeneity.

Chromosomal aberrations and hepatitis B virus integration into the host genome

Chromosomal aberrations such as deletions and copy number gains are frequent in HCC. Initial studies identified that HCC

Corresponding author: Ozturk, M. (ozturk@fen.bilkent.edu.tr)

harbours multiple chromosomal abnormalities, predominantly losses, with increased chromosomal instability in tumours associated with HBV infection. Common alterations include gain of chromosomes 1q, 8q and 17q, and loss of 4q [5]. Recently, data from whole genome analysis techniques showed that chromosomes 1q, 5, 6p, 7, 8q, 17q and 20 display chromosomal gains, while 1p, 4q, 6q, 8p, 13q, 16, 17p and 21 exhibit losses in HCC [6].

In addition, HBV DNA is often integrated into the host genome in patients with HBV-related HCCs [7]. This integration may have *cis* and *trans* effects. Viral DNA integration into or near gene sequences may alter gene expression as well as gene integrity. In addition, integrated viral DNA may encode wild-type or truncated viral proteins acting in *trans* on the host genome, either by deregulating gene expression or by interacting with host proteins [8]. Recently reported whole genome studies indicated that the viral integration is associated with breakpoints within the HBV genome that primarily localised to the downstream region of the *HBX* gene. HBV genome integration was observed within or upstream of the *TERT* (telomerase reverse transcriptase) gene in four HBV-related HCCs. However, HBV integration sites within the same or different tumours did not show specific patterns, suggesting that the virus does not target specific host sequences. [9]. Based on these findings, it is highly probable that landscape changes in the structural integrity of chromosomes, as well as random but multiple integrations of HBV genomes into host genomes, cause high levels of instability in the chromosomal integrity of HCC. Some of these aberrations may hit crucial genes such as *TERT*, which may directly contribute to tumour development by inappropriate activation or inactivation of the genes themselves. In addition, the integration of viral enhancer sequences in the vicinity of crucial genes may lead to aberrant gene expression in HCC.

Gene mutations

Since the discovery of *TP53* as the first mutated gene in HCC over 20 years ago [10] and until very recently, only four genes were known to display frequent alterations in liver cancers. While *TP53*, *CTNNB1* (encoding β -catenin) and *AXIN1* genes usually display point mutations and small deletions, *CDKN2A* (encoding p16^{INK4a}) undergoes homozygous deletions and epigenetic silencing [11,12].

During the past two years, the first reports of whole-genome or exome sequencing data for HCC have appeared [6,9,13]. This is the beginning of a new era of HCC genetics, because of the fact that these new techniques will allow the visualisation of the mutational landscape of HCC. Figure 1 shows a summary of primary findings gathered by ourselves from two recently published reports [6,9]. Each study first analysed a small set of tumours ($n = 20\text{--}25$) for a genome-wide search of somatically mutated genes; significantly mutated genes were then further tested for mutations using a larger set of tumours ($n > 100$).

A close examination of the data of Fig. 1 indicates that *TP53* and *CTNNB1* represent the two most frequently mutated genes. A second group of genes (*AXIN1* and *ARID1A*) was found to present less frequent mutations, but still present in more than 10% of HCC samples studied. The third group is the largest with 22 genes displaying recurrent mutations in less than 10% of tumours. Guichard et al. [6] reported that Wnt/ β -catenin, p53, PI3K/Ras signalling, oxidative, endoplasmic reticulum stress pathways and chromatin remodelling were frequently affected by these mutations.

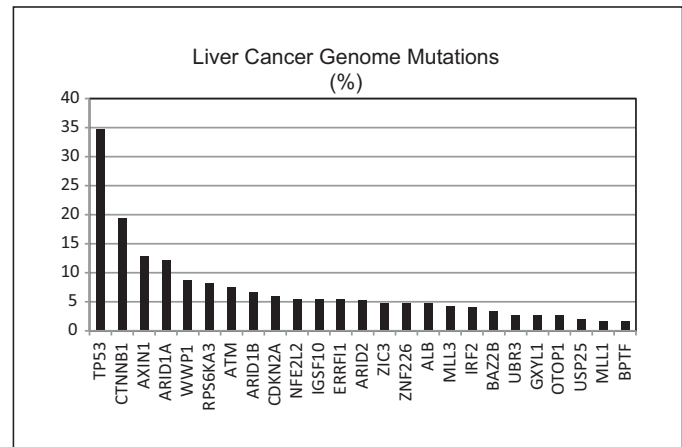


FIGURE 1

Most frequently mutated genes in hepatocellular carcinoma.

Whole genome sequencing allowed the detection of recurrent somatic mutations in several genes annotated as associated with chromatin regulation, such as *ARID1A*, *ARID1B*, *ARID2*, *MLL*, *MLL3*, *BAZ2B*, *BRD8*, *BPTF*, *BRE* and *HIST1H4B*. Notably, 14 out of the 27 tumours (52%) had either somatic point mutations or indels in at least one of these chromatin regulators. In both sets of experiments (whole genome sequencing and the validation sets), the number of indels in chromatin regulator genes was significantly higher than those in genes belonging to the other categories. This suggests that loss-of-function mutations are enriched in these chromatin-regulator genes in HCC genomes [9].

As shown in Table 1, the frequent mutations that identified so far in HCC are likely to result in loss of function with the notable exception of *CTNNB1* mutations. It will be interesting to study why loss of function rather than gain of function of crucial genes is associated with HCC. By contrast, this pattern of mutation does not offer a broad spectrum of therapeutic intervention applications. Cancer cells can easily be targeted by blocking genes that are aberrantly overactive in these cells. The restoration of a lost gene activity to achieve a therapeutic intervention is difficult to achieve. Thus, although the genome-wide analyses have been very helpful in establishing the list of a large set of mutated genes in HCC, this will most probably serve diagnostic needs while the chance of their therapeutic use is more limited.

Epigenetic deregulation

Epigenetic regulation of gene expression involves DNA methylation, post-translational histone modifications, chromatin changes and non-coding RNAs that are often affected in cancer cells [14,15]. The role of epigenetic deregulation in HCC is being increasingly recognised [16]. In addition to changes in DNA methylation, microRNA expression, mutations affecting epigenetic regulatory genes have recently been discovered in HCC [6,9,13].

HCC cells display global hypomethylation as well as promoter hypermethylation of a large set of genes [17]. Promoter hypermethylation appears to affect mainly tumour suppressor and anti-proliferative genes resulting in downregulation of gene expression (Fig. 2). Aberrations in microRNA expression have also been observed with several of them being linked to metabolic and phenotypic changes in HCC cells [14,18–20].

TABLE 1
Most frequent gene mutations in hepatocellular carcinoma are predicted to lead to a loss-of-function

Genes	% mutation rates	Protein function	Known/expected outcome
TP53	35	DNA damage response, other	Loss-of-function
CTNNB1	19	Positive regulator of Wnt signalling	Gain-of-function
AXIN1	13	Negative regulator of Wnt signalling	Loss-of-function
ARID1A	12	Chromatin remodelling	Loss-of-function
WWP1	9	E3 ubiquitin ligase	Loss-of-function?
RPS6KA3	8	Ribosomal protein S6 kinase	?
ATM	8	DNA damage response	Loss-of-function?
ARID1B	7	Chromatin remodelling	Loss-of-function?
CDKN2A	6	Positive regulator of senescence	Loss-of-function
NFE2L2	5	Redox homeostasis?	?
IGSF10	5	?	Loss-of-function
ERRFI1	5	EGFR/ERB2 kinase inhibitor	Loss-of-function
ARID2	5	Chromatin remodelling	Loss-of-function?

Several genes encoding epigenetic regulatory proteins are involved in hepatocellular malignancy. The *EZH2* (KMT6) encodes the catalytic component of the Polycomb Repressive Complex 2 (PRC2), creating the transcriptionally repressive H2K27Me3 histone mark which results in transcriptional silencing [21]. *EZH2* is over-expressed in HCC and mostly associated with the progression and aggressive biological behaviour of HCC [22,23]. *EZH2* protein silences Wnt pathway antagonists and constitutively activates Wnt/ β -catenin signalling causing cell proliferation in HCC cells [24]. *EZH2* also exerts a prometastatic function through epigenetic silencing of multiple tumour suppressor miRNAs including miR-139-5p, miR-125b, miR-101, let-7c and miR-200b [25]. Yang et al. identified an lncRNA called lncRNA-HEIH (High Expression in HCC) that associates with *EZH2* to repress *EZH2* target genes such

as p16^{Ink4a} and p21^{Cip1} in HBV-related HCC [26]. BMI1 is another PRC2 member overexpressed in HCC. Effendi et al. determined that BMI1 is upregulated in early and well-differentiated HCC and this expression correlates with ABCB1 expression [27]. Expression of histone deacetylases (HDACs) is deregulated in different cancers [28], and some of them are also deregulated in HCC. HDACs-1, -2 and -3 are over-expressed in HCC [29,30]. LC3B-II-induced inactivation of HDAC1 caused regression of HCC cell proliferation and triggered caspase independent autophagy. p21^{Cip1} and p27^{Kip1} were selectively induced while cyclin D1 and CDK2 were suppressed by inactivation of HDAC1. As a result, HDAC1 inactivation resulted in hypophosphorylation of pRb in the G1/S checkpoint to inactivate E2F/DP1 transcriptional activity. Also, p21^(WAF1/Cip1) transcriptional activity was suppressed by

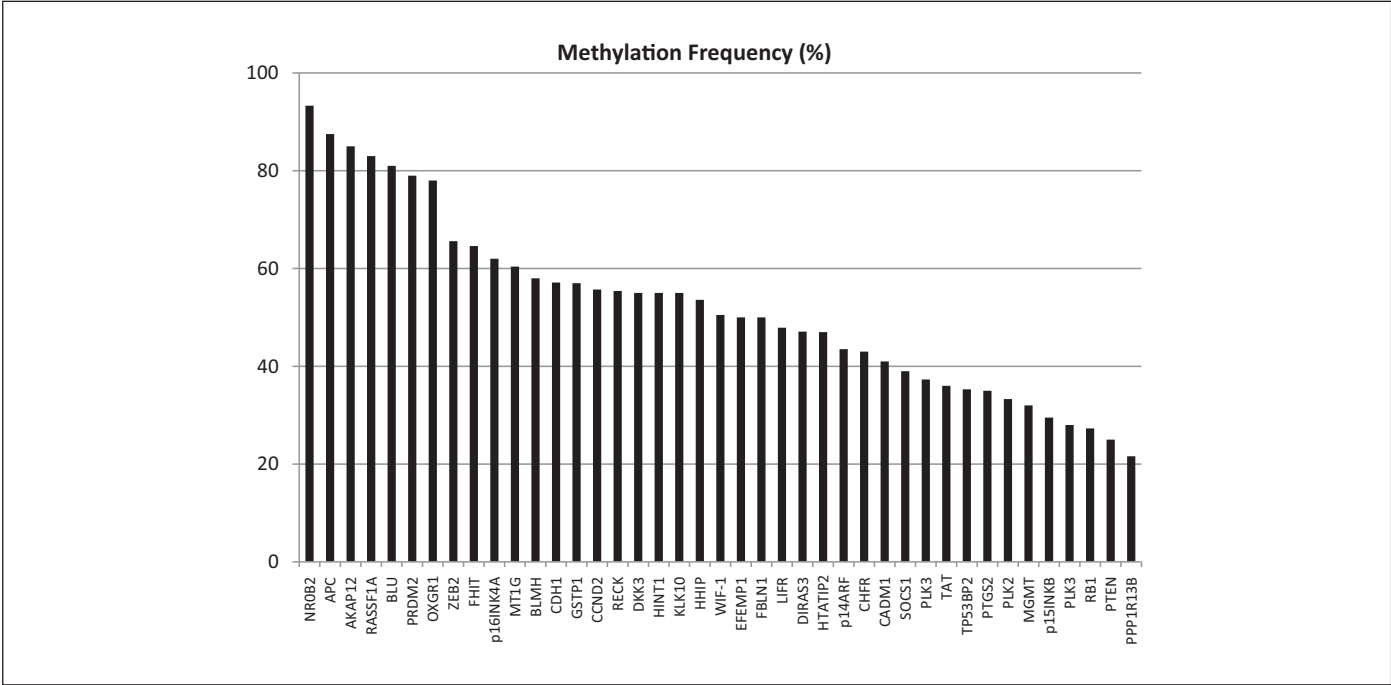


FIGURE 2
The frequency of promoter methylation in hepatocellular carcinoma.

HDAC1 by interaction with an Sp1-binding site in the p21^(WAF1/Cip1) promoter [31]. HDAC4 also suppresses the promoter activity of miR-200a and its expression and interacts with Sp1 in the miR-200a promoter to attenuate histone H3 acetylation levels. miR-200a represses HDAC4 expression through targeting the 3'-untranslated region of messenger RNA of HDAC4. In this respect, miR-200a has an ability to induce its own transcription and increase the levels of histone H3 acetylation at its promoter. Furthermore, miR-200a induces up-regulation of the levels of total acetyl-histone H3 and histone H3 acetylation in the p21^{Cip1} promoter [32].

DNA methylating enzymes *DNMT1*, *DNMT3A* and *DNMT3B* are over-expressed in HCC compared to noncancerous liver samples [33,34]. Finally, *CENPA* expression was found to be significantly elevated in HCC tissues, and a positive correlation exists between CENP-A expression and HBx COOH mutations in HCC tissues. HBx mutant increases the expression of *CENPA* mRNA [35].

Future perspectives

Recent advances in genome sequencing technologies will change radically our capabilities for fine mapping of hepatocellular cancer genomes. It is expected that patient tumours will be fully analysed in a short time at a moderate cost. Therefore, the genomic and

epigenomic status of the patient's own tumour will be a crucial element for decision making in terms of disease prognosis, therapeutic choices and prediction of patient survival. However, most of the known mutations observed in HCC are associated with a loss of function. Apparently, targetable genes found in other cancers such as growth factor receptors and intracellular protein kinases are not mutated at significant levels in HCC. Therefore, we need to find other targets for the treatment of liver cancers. Epigenetic characterisation of HCC has allowed the discovery of many epigenetic players in this disease. However, these studies are far from being complete. The rarity of targetable mutations in HCC justifies a systematic study of epigenetic changes to identify new targets for the therapy of this disease.

Acknowledgements

The research study is supported by grants from TÜBİTAK (109S191 and 111T558) with additional support from State Planning Office (DPT-KANİLTEK Project), Turkish Academy of Sciences, Institut National de Cancer and La Ligue Nationale Contre le Cancer in France (Equipe labélisée). C.O., G.Y. and D.C. received fellowships from Turkish Academy of Sciences (C.O.), TÜBİTAK (G.Y., D.C.) and EMBO (G.Y.).

References

- [1] El-Serag HB. Hepatocellular carcinoma. *New England Journal of Medicine* 2011;365:1118–27.
- [2] El-Serag HB. Epidemiology of viral hepatitis and hepatocellular carcinoma. *Gastroenterology* 2012;142:1264–1273.e1.
- [3] El-Serag HB, Marrero JA, Rudolph L, Reddy KR. Diagnosis and treatment of hepatocellular carcinoma. *Gastroenterology* 2008;134:1752–63.
- [4] Forner A, Llovet JM, Bruix J. Hepatocellular carcinoma. *Lancet* 2012;379:1245–55.
- [5] Buendia MA. Genetic alterations in hepatoblastoma and hepatocellular carcinoma: common and distinctive aspects. *Medical and Pediatric Oncology* 2002;39:530–5.
- [6] Guichard C, Amaddeo G, Imbeaud S, Ladeiro Y, Pelletier L, Maad IB, et al. Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. *Nature Genetics* 2012;44:694–8.
- [7] Brechot C, Pourcel C, Louise A, Rain B, Tiollais P. Presence of integrated hepatitis B virus DNA sequences in cellular DNA of human hepatocellular carcinoma. *Nature* 1980;286:533–5.
- [8] Brechot C, Gozuacik D, Murakami Y, Paterlini-Brechot P. Molecular bases for the development of hepatitis B virus (HBV)-related hepatocellular carcinoma (HCC). *Seminars in Cell Biology* 2000;10:211–31.
- [9] Fujimoto A, Totoki Y, Abe T, Borojevich KA, Hosoda F, Nguyen HH, et al. Whole-genome sequencing of liver cancers identifies biological influences on mutation patterns and recurrent mutations in chromatin regulators. *Nature Genetics* 2012;44:760–4.
- [10] Bressan B, Galvin KM, Liang TJ, et al. Abnormal structure and expression of p53 gene in human hepatocellular carcinoma. *Proceedings of the National Academy of Sciences of the United States of America* 1990;87:1973–7.
- [11] Ozturk M. Genetic aspects of hepatocellular carcinogenesis. *Seminars in Liver Disease* 1999;19:235–42.
- [12] Ozturk M, Arslan-Ergul A, Bagislar S, Senturk S, Yuzugullu H. Senescence and immortality in hepatocellular carcinoma. *Cancer Letters* 2009;286:103–13.
- [13] Huang J, Deng Q, Wang Q, Li KY, Dai JH, Li N, et al. Exome sequencing of hepatitis B virus-associated hepatocellular carcinoma. *Nature Genetics* 2012;44:1117–21.
- [14] Sandoval J, Esteller M. Cancer epigenomics: beyond genomics. *Current Opinion in Genetics and Development* 2012;22:50–5.
- [15] Rodriguez-Paredes M, Esteller M. Cancer epigenetics reaches mainstream oncology. *Nature Medicine* 2011;17:330–9.
- [16] Pogribny IP, Rusyn I. Role of epigenetic aberrations in the development and progression of human hepatocellular carcinoma. *Cancer Letters* 2012 [Epub ahead of print].
- [17] Sceusi EL, Loose DS, Wray CJ. Clinical implications of DNA methylation in hepatocellular carcinoma. *HPB (Oxford)* 2011;13:369–76.
- [18] Burchard J, Zhang C, Liu AM, Poon RT, Lee NP, Wong KF, et al. microRNA-122 as a regulator of mitochondrial metabolic gene network in hepatocellular carcinoma. *Molecular Systems Biology* 2010;6:402.
- [19] Lachenmayer A, Alsinet C, Savic R, Cabellos L, Toffanin S, Hoshida Y, et al. Wnt-pathway activation in two molecular classes of hepatocellular carcinoma and experimental modulation by sorafenib. *Clinical Cancer Research* 2012;18:4997–5007.
- [20] Murakami Y, Yasuda T, Saigo K, Urashima T, Toyoda H, Okanoue T, et al. Comprehensive analysis of microRNA expression patterns in hepatocellular carcinoma and non-tumorous tissues. *Oncogene* 2006;25:2537–45.
- [21] Cao R, Wang L, Wang H, Xia L, Xia L, Erdjument-Bromage H, Tempst P, et al. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* 2002;298:1039–43.
- [22] Cai MY, Tong ZT, Zheng F, Liao YJ, Wang Y, Rao HL, et al. EZH2 protein: a promising immunomarker for the detection of hepatocellular carcinomas in liver needle biopsies. *Gut* 2011;60:967–76.
- [23] Sasaki M, Ikeda H, Itatsu K, Yamaguchi Y, Sawada S, Minato H, et al. The overexpression of polycomb group proteins Bmi1 and EZH2 is associated with the progression and aggressive biological behavior of hepatocellular carcinoma. *Laboratory Investigation* 2008;88:873–82.
- [24] Cheng AS, Lau SS, Chen Y, Kondo Y, Li MS, Feng H, et al. EZH2-mediated concordant repression of Wnt antagonists promotes beta-catenin-dependent hepatocarcinogenesis. *Cancer Research* 2011;71:4028–39.
- [25] Au SL, Wong CC, Lee JM, Fan DN, Tsang FH, Ng IO, et al. Enhancer of zeste homolog 2 epigenetically silences multiple tumour suppressor microRNAs to promote liver cancer metastasis. *Hepatology* 2012;56:622–31.
- [26] Yang F, Zhang L, Huo XS, Yuan JH, Xu D, Yuan SX, et al. Long noncoding RNA high expression in hepatocellular carcinoma facilitates tumour growth through enhancer of zeste homolog 2 in humans. *Hepatology* 2011;54:1679–89.
- [27] Effendi K, Mori T, Komuta M, Masugi Y, Du W, Sakamoto M. Bmi-1 gene is upregulated in early-stage hepatocellular carcinoma and correlates with ATP-binding cassette transporter B1 expression. *Cancer Science* 2010;101:666–72.
- [28] Weichert W. HDAC expression and clinical prognosis in human malignancies. *Cancer Letters* 2009;280:168–76.
- [29] Wu LM, Yang Z, Zhou L, Zhang F, Xie HY, Feng XW, et al. Identification of histone deacetylase 3 as a biomarker for tumour recurrence following liver transplantation in HBV-associated hepatocellular carcinoma. *PLoS ONE* 2010;5:e14460.
- [30] Quint K, Agaimy A, Di Fazio P, Montalbano R, Steindorf C, Jung R, et al. Clinical significance of histone deacetylases 1, 2, 3, and 7: HDAC2 is an independent predictor of survival in HCC. *Virchows Archiv* 2011;459:129–39.
- [31] Xie HJ, Noh JH, Kim JK, Jung KH, Eun JW, Bae HJ, et al. HDAC1 inactivation induces mitotic defect and caspase-independent autophagic cell death in liver cancer. *PLoS ONE* 2012;7:e34265.
- [32] Yuan JH, Yang F, Chen BF, Lu Z, Huo XS, Zhou WP, et al. The histone deacetylase 4/SP1/microRNA-200a regulatory network contributes to aberrant histone acetylation in hepatocellular carcinoma. *Hepatology* 2011;54:2025–35.
- [33] Choi MS, Shim YH, Hwa JY, Lee SK, Ro JY, Kim JS, et al. Expression of DNA methyltransferases in multistep hepatocarcinogenesis. *Human Pathology* 2003;34:11–7.
- [34] Saito Y, Kanai Y, Sakamoto M, Saito H, Ishii H, Hirohashi S. Expression of mRNA for DNA methyltransferases and methyl-CpG-binding proteins and DNA methylation status on CpG islands and pericentromeric satellite regions during human hepatocarcinogenesis. *Hepatology* 2001;33:561–8.
- [35] Li Y, Zhu Z, Zhang S, Yu D, Yu H, Liu L, et al. ShRNA-targeted centromere protein A inhibits hepatocellular carcinoma growth. *PLoS ONE* 2011;6:e17794.