Molecular basis of the post-meiotic male genome programming

Naghmeh Hoghoughi

Communauté
UNIVERSITÉ Grenoble Alpes

THÈSE

Pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTE UNIVERSITE GRENOBLE ALPES

Spécialité : Chimie Biologie

Arrêté ministériel : 25 mai 2016

Présentée par

Naghmeh HOGHOUGHI

Thèse dirigée par **Dr. Saadi KHOCHBIN, DRCE, CNRS**

Préparée au sein du **laboratoire CRI UGA - INSERM U1209 - CNRS UMR 5309 à l'Institut pour l'Avancée des Biosciences** dans **l'École Doctorale Chimie et Sciences du Vivant**

Base moléculaire de la programmation post-méiotique du génome male

Molecular basis of the post-meiotic male genome programming

Thèse soutenue publiquement le **10 décembre 2019**, devant le jury composé de :

Pr. Claire VOURCH Professeur, Université Grenoble Alpes, Grenoble	Présidente
Dr. Matthieu GERARD Directeur de recherche, CEA, Saclay	Rapporteur
Dr. Nathalie BEAUJEAN Directeur de recherche, INRA, Bron	Rapporteur
Dr. Daniel JOST Chargé de recherche, ENS, Lyon	Examinateur
Dr. Saadi KHOCHBIN Directeur de recherche, CNRS, Grenoble	Directeur de thèse

To My family,

To all my loved ones,

Acknowledgments

There are many people who supported me since the beginning of my journey, whom I would like to thank here.

First, I would like to sincerely acknowledge the member of my jury, Dr. Matthieu Gérard, Dr. Nathalie BEAUJEAN, Dr. Daniel JOST and Pr. Claire VOURCH, who accepted to be part of my jury and evaluate my work.

Saadi, I can write another 200 pages only on how grateful and lucky I feel to get the opportunity to work with you and grow up scientifically and personally beside you. Thank you for your trust, your infinite encouragement, your optimism, your availability, and your valuable scientific advices. Your confidence in me made this work happen, and I will be always grateful for that. Thank you Saadi.

I would like to thank my annual follow-up meeting jury members Pr. Claire VOURCH and Dr. Homeira NAWABI, for their dedication and constructive scientist advices.

I would like to thank all the members of SK team, Sophie R, Sophie B, Sandrine, Thierry, Fayçal, Florent, Katia, Anne-Laure and Michel without whom this work would be impossible. Sophie R, thank you very much for your scientific advices, bioinformatic support and your kindness during these years. Sophie B, thank you for your support from the beginning. You thought me everything at bench, with your golden hands for experiments!! Thank you also for your moral support, who helped me go through difficult moments. Sandrine, thank you for teaching me everything in cell culture and sorry to bother with my non-stop questions. Thierry, thank you so much for your great help with the experiments and helping me with my manuscript. Fayçal and Michel, thank you for helping me prepare presentations whenever I had one, and thank you for your kindness and your support. I would like to especially thank to Domenico, Clovis, Mengqing, Alexandra, Azadeh, best interns Mélanie and Nael and the former PhD students of the lab (Afsaneh and Mahya), which I had the chance to work beside them in the beginning of my PhD. I will be always grateful for our happy hours in the lab, sharing drinks and joyful laughers beside each of you. Thank you, every member of SK team, for making this journey unforgettable for me.

I would also like to thank, my IAB friends, Valentina, Alexandre, Tristan, Estelle, Sara, Lorrie, Marie for all the joyful moments that we had together during experiments or during afterworks and especially their support during these last months. Also, I want to thank former PhD students of IAB, Matteo, Mathieu, Yohan, Anne-Sophie who supported me during the first years of my arrival in IAB, and for all the great moments that we had in the lab or outside.

I wish to thank everyone in IAB who directly or indirectly helped me to solve experiment problems. Thank you very much.

I would like to acknowledge my friends of Master years, the ones that I met during the conference organizations (Especially Léa), with whom I developed true friendship.

I want to especially thank my “adoptive″ family Agnès and Ahmad (Chouanard et Sharonizadeh) with whom I have had the best moments since my arrival in Grenoble.

Many thanks to Delphine who always supported me, thank you for your immense kindness and exceptional moments of laughter and joy.

I want to hugely thank Neda and Flavien, for their immense amount of help and support since my arrival in Grenoble.

Last but not least, I would like to thank my parents. Their infinite support for my journey was essential during these years, they were far but always close to my heart. Thank you for everything.

Abstract

Spermatogenic cell chromatin undergoes a drastic reorganization during spermiogenesis. This phenomenon occurs for two principles reasons: first, to establish a highly compact genome through protamine incorporation, in order to protect the sperm genome before and upon fertilization. Second, in order to successfully transfer the paternal information carried by the sperm genome, to the next generation. Recently, our laboratory identified various actors and molecular mechanisms, which are at the heart of this drastic genome reorganization. A genome-wide histone hyperacetylation initiates the process of histone replacement by transition proteins and protamines. Our laboratory identified NUT as the main player of histone hyperacetylation at the onset of histone-to-protamine transition process. NUT through recruitment of histone acetyltransferase CBP/p300, induces histone H4 hyperacetylation mainly at lysine 5 and lysine 8 residues. Next, first bromodomain of BRDT recognizes histone acetylation and initiates the process of histone removal. Previously identified testis-specific H2A variant, H2A.L.2, forms dimer with histone H2B variant TH2B at the time of histone eviction in elongating spermatids. The dimer H2A.L.2-TH2B then incorporates into the chromatin and opens the nucleosome structure, which allows the invasion of histones by transition proteins as well as pre-protamine processing. Protamines finally displace histone-TPs complexes and re-package the genome. Previously, our laboratory reported that a fraction of H2A.L.2 is associated with pericentric heterochromatin regions in mature spermatozoa. During my PhD, first we discovered the structural and molecular basis of the activation of CBP/p300. We speculate that the same mechanisms could lead to the Nut-p300-dependent induced histone acetylation, leading to the initiation of histone removal in post-meiotic phases of spermatogenesis. Second, we demonstrated that nucleosome dissociation following H2A.L.2 incorporation, provides an opportunity for H2A.L.2-TH2B dimer to associate to DNA in a stable manner, while H3-H4 tetramer is being displaced. This leads to the generation of H2A.L.2-TH2B dimer-based structures, which persist in mature sperm. In addition, we deciphered the molecular basis of the intrinsic ability of H2A.L.2 to target the pericentric heterochromatin regions. We also identified a critical role for RNA in regulating and in controlling the turnover and localization of H2A.L.2 in pericentric heterochromatin regions. In conclusion, our work highlights the presence of dimer-based structures within the sperm chromatin, as well as a specific re-programming of the male pericentric heterochromatin, which might bear crucial information transmittable to the next generation.

chromatin reprogramming, H2A.L.2, spermatogenesis, CBP/p300, subnucleosomal structure

Résumé

La chromatine des cellules spermatogèniques subit une réorganisation radicale pendant la spermiogenèse. Ce phénomène se produit pour deux raisons principales : premièrement, pour établir un génome très compact par l'incorporation de protamine, afin de protéger le génome du sperme avant et pendant la fécondation. Deuxièmement, afin de transférer avec succès l'information paternelle portée par le génome du sperme, à la génération suivante. Récemment, notre laboratoire a identifié différents acteurs et mécanismes moléculaires, qui sont au cœur de cette réorganisation radicale du génome. Une hyperacétylation des histones à l'échelle du génome amorce le processus de remplacement des histones par des protéines de transition et des protamines. Notre laboratoire a identifié NUT comme le principal régulateur de l'hyperacétylation des histones au début du processus de transition histone-protamine. NUT par recrutement d'histone acétyltransférase CBP/p300, induit une hyperacétylation de l'histone H4 principalement au niveau des résidus lysine 5 et lysine 8. Ensuite, le premier bromodomaine de BRDT reconnaît l'acétylation des histones et déclenche le processus d'élimination des histones. Le variant H2A.L.2 spécifique au testicule, forme un dimère avec le variant d'histone H2B, TH2B, au moment de l'éviction des histones dans les spermatides allongées. Le dimère H2A.L.2-TH2B s'incorpore alors dans la chromatine et ouvre la structure du nucléosome. Ceci permet l'invasion des histones par les protéines de transition ainsi que le traitement pré-protaminique. Les protamines déplacent finalement les complexes histone-TPs et compacte le génome. Auparavant, notre laboratoire a rapporté qu'une fraction de H2A.L.2 reste associée à des régions hétérochromatines péricentriques dans les spermatozoïdes matures. Au cours de mon doctorat, nous avons d'abord découvert les bases structurelles et moléculaires de l'activation du CBP/p300. Nous spéculons que les mêmes mécanismes pourraient mener à l'acétylation des histones induite par Nut-p300-dépendante, menant à l'initiation de l'élimination des histones dans les phases post-méiotiques de la spermatogenèse. Nous avons démontré également que la dissociation des nucléosomes à la suite de l'incorporation du H2A.L.2-TH2B permet au dimère H2A.L.2-TH2B de s'associer à l'ADN de façon stable, tandis que le tétramère H3-H4 est déplacé. Ceci conduit à la génération de structures à base de dimères H2A.L.2-TH2B, qui persistent dans le sperme mature. De plus, nous avons déchiffré la base moléculaire de la capacité intrinsèque de H2A.L.2 à cibler les régions hétérochromatines péricentriques. Nous avons également identifié un rôle critique vis-à-vis de l'ARN dans la régulation et le contrôle du renouvellement et de la localisation d'H2A.L.2 dans les régions hétérochromatines péricentriques. En conclusion, nos travaux mettent en évidence la présence de structures à base de dimères dans la chromatine du sperme, ainsi qu'une reprogrammation spécifique de l'hétérochromatine péricentrique masculine. Ceci pourrait porter des informations essentielles transmissibles à la génération suivante.

réorganisation de la chromatine, H2A.L.2, Spermatogénèse, CBP/p300, structures subnucléosomales

Table of Contents

ABBREVIATION

2OG	2 oxo Glutarate
5caC	5-carboxylcytosine
5fC	5-formylcytosine
5hmeC	5 hydroxy methyl Cytosine
5meC	5 methy Cytosine
ACSS2	acyl-CoA synthetase short-chain family member 2
ADP	Adenosine Diphosphate
AM-PD	active-modification / passive dilution
ANP32E	Acidic leucine-rich nuclear phosphoprotein 32 family member E
ASF1	Anti-Silencing Function 1
ATAC	*Drosophila* ADA2A-containing Complex
ATM	ataxia telangiectasia mutated
ATRX	alpha thalesemia/mental retardation syndrome X
AutoN	Auto inhibitory N-tereminal
Bbd	Bar body deficient
Bdf1	Bromodomain containing factor 1
BER	base excision repair
BET	Bromodomain and Extra terminal
BMP4	bone morphogenic protein 4
BRD	Bromodomain
BRDT	Bromodomain Testis Specific Protein
CABIN	Calcineurine Binding Protein
CAF-1	chromatin Assembly Factor-1
CATD	Cenp-A Targeting Domain
CD	Chromo Domain
Cdk-1	cyclin-dependant Kinase 1

CENP-A	Centromeric Protein-A
CHD	chromodomain helicase DNA binding
ChIP-seq	Chromatin ImmunoPrecipitation-Sequencing
Co-REST	corepressor for element-1-silencing transcription factor
CREB	c-AMP Response Elements
CS H1	cleavage stage linker H1
CTD	C Terminal Domain
d.p.p	days post-partum
DAXX	Death-associated protein 6
DBD	DNA Binding Domain
DNMT	DNA Methyl Transferases
DBS	Double Strand Break
DNA	DeoxyriboNucleic Acide
DPPA-3	pluripotency-associated protein 3
EMT	Endothelial-Mesenchimal Transition
endo-siRNA	endogenous small-interacting RNA
ER	Endoplamsic Reticulum
Ezh2	Enhancer of zeste homolog 2
FACT	Facilitates ChromatinTranscription
FISH	Fluorescent In Situ Hybridization
GCN5	General Control Nonderepressible 5
GNATs	GCN5-related N-acetyltransferases
H2A.Bbd	H2A.Bar Body Deficient
H2A.L	H2A.Like
HATs	Histone Acetyl Transferases
HDACs	Histone Deacetylases
HILS	linker histone H1-like protein
HIRA	Histone Regulator A
HJURP	Holliday Junction Recognition Protein
HKMTs	histone lysine methyltransferases
HPG	hypothalamic pituitary gonadal axis

HSF1	Heat Shock Factor 1
HSP	Heat Shock Protein
HSS	HAND-SANT-SLIDE
ICRs	Imprinting Control Regions
INO80	Inositol requiring 80
IR	ionizing radiation
ISWI	Imitation Switch
JHDM	JmjC domain-containing histone demethylation
KMTs	lysine methyltransferases
Lap1	Lack of Acidic Patch
LC	Leydig Cell
LH	luteinizing hormone
LINE-1	Long Interspersed Nuclear Elements-1
LNA	Locked Nucleic Acid
lncRNA	Long non-coding RNA
LSD1	lysine-specific demethylase 1
MEF	Mouse Embryonic Fibroblast
miRNAs	micro RNA
MNase	Micrococcal protease
MSCI	Mouse Sex Chromosome Inactivation
MYST	MOZ, Ybf2/Sas3, Sas2 and Tip60
NAP-1	Nucleosome Assembly Protein-1
NegC	Negative-regulator of Coupling
NMC	NUT midline carcinoma
NUT	Nucler protein in Testis
PADI4	peptidylearginine deiminase 4
PCNA	proliferating cell nuclear antigen
PGC	Primordial Germ Cells
PHC	Pericentric HeteroChromatin
piRNA	piwi-interacting RNA
Plk-1	polo-like kinase 1

PRC1	polycomb repressive complex 1
Prm/PRM	Protamine
PRMTs	protein arginine N-methyl transferases
PTM	Post Translational Modification
RNA	DeoxyRiboNucleic Acid
Rpd3	reduced potassium dependency 3
SAGA/ STAGA	Spt-Ada-Gcn5 Acetyltransferase
SET	Su (var) - Enhancer of zest, Trithorax
SETDB1	SET domain bifurcated 1
SINEs	Small Interspersed Nuclear Elements
Sir2	silent information regulator 2
SncRNA	Small non-coding RNA
SRCAP	SNF2-related CBP activator protein
SSC	spermatogonial stem cells
SWI/SNF	switch/sucrose non-fermentable
TADs	Topologically Associated Domains
TAF-1	TATA-binding protein associated factor-1
TDG	Thymine DNA glycosylase
TET	Ten-Eleven Translocation
TP	Transition Protein
tRNA	Transfer RNA
TSA	Trochostatine A
TSS	Transcriptional Start Sites
UBN1	Ubinuclein-1
UHRF1	Ubiquitin-like, containing PHD and RING finger domains 1
WB	Western Blot
WHD	Wing Helix Domain
ZEB1	Zinc finger E-box binding homeobox 1

Introduction

I. Chromatin Structure and Function

1. General overview of chromatin

1.1. Chromatin structure

1.1.1. The discovery of chromatin

The revolutionary discoveries of many dedicated scientists during several decades shaped our current conception of chromatin. The history of the chromatin started in 1871 when Friedrich Miescher isolated a phosphate-rich molecule in the nucleus of leukocytes, which he named nuclein (Miescher-Rüsch, 1871). About 10 years later in 1882, Walther Flemming observed a dense nuclear structure that absorbed the aniline (basophilic) dyes, which he termed chromatin in reference to Greek word *khroma*, meaning stained. He also suggested that chromatin and nuclein are the same chemical substances (Flemming, 1882) which was also supported by Hertwig, who at the time was working on the egg fertilisation of sea urchins. Pursuing his investigations, Miescher later demonstrated that nuclein is composed of two types of molecules: one basic component of proteic-nature and one phosphate-rich acidic molecule. In 1884, Albrecht Kossel purified an acid-soluble protein fraction from the nuclei of chicken erythrocyte cells, which he called histones (a mix of terms *histos*, "tissue" in Greek and from Greek peptos, *digested*,) (Kossel, 1884). After Miescher's demonstrations regarding the two components of nuclein, in 1889, R. Altmann separated nuclein in two fractions, the proteins and one acidic fraction which he baptised nucleic acid (Altmann, 1889). In the following years Kossel identified four nucleobases (A, T, G, C), the components of nucleotides, which are the basic building blocks of the nucleic acids. In 1919, following the discovery of deoxyribose by P. Leven within the acid nucleic molecule, the name "DNA" was proposed for the first time.

In 1896, Wilson proposed that the acidic-nature fraction of the nuclein is responsible for the genetic inheritance within generations (Wilson, 1900). In 1928, Frederick Griffith discovered if non-pathogenic bacteria are exposed to dead pathogenic bacteria (by heat), in the same medium, they become pathogenic. In 1944, O. Avery, C. MacLeod and M. McCarty, infected mice with modified bacterial strains, which revealed that DNA is the principle heredity

factor in the chromatin (Avery, MacLeod et al., 1944). In 1952, Alfred Hershey and Martha Chase conducted a series of radioactive labelling experiments in bacteriophages and then infected bacterial cells with these bacteriophages. They observed that when ^{35}S radioactive is incorporated into the protein component of bacteriophages, it was not transmitted to the infected bacterial cells. In contrast, ^{32}P, which is mostly incorporated into the acid nucleic component of bacteriophages, was transmitted to the infected bacterial cells. Finally, in 1952, thanks to A. Hershey and M. Chase landmark discoveries, the whole scientific community was convinced that DNA is the genetic material (Hershey & Chase, 1952).

James Watson and Francis Crick proposed the double-helix structure of the DNA in 1953. This revelation resulted in the comprehension of DNA as a macromolecule, which encode the genetic information (Watson & Crick, 1953) (Franklin & Gosling, 1953). In 1961, MW. Nirenberg revealed that the genetic information coded by the DNA macromolecule, is decrypted into amino acid, which allowed the revelation of the genetic code (Nirenberg & Matthaei, 1961) (Figure 1).

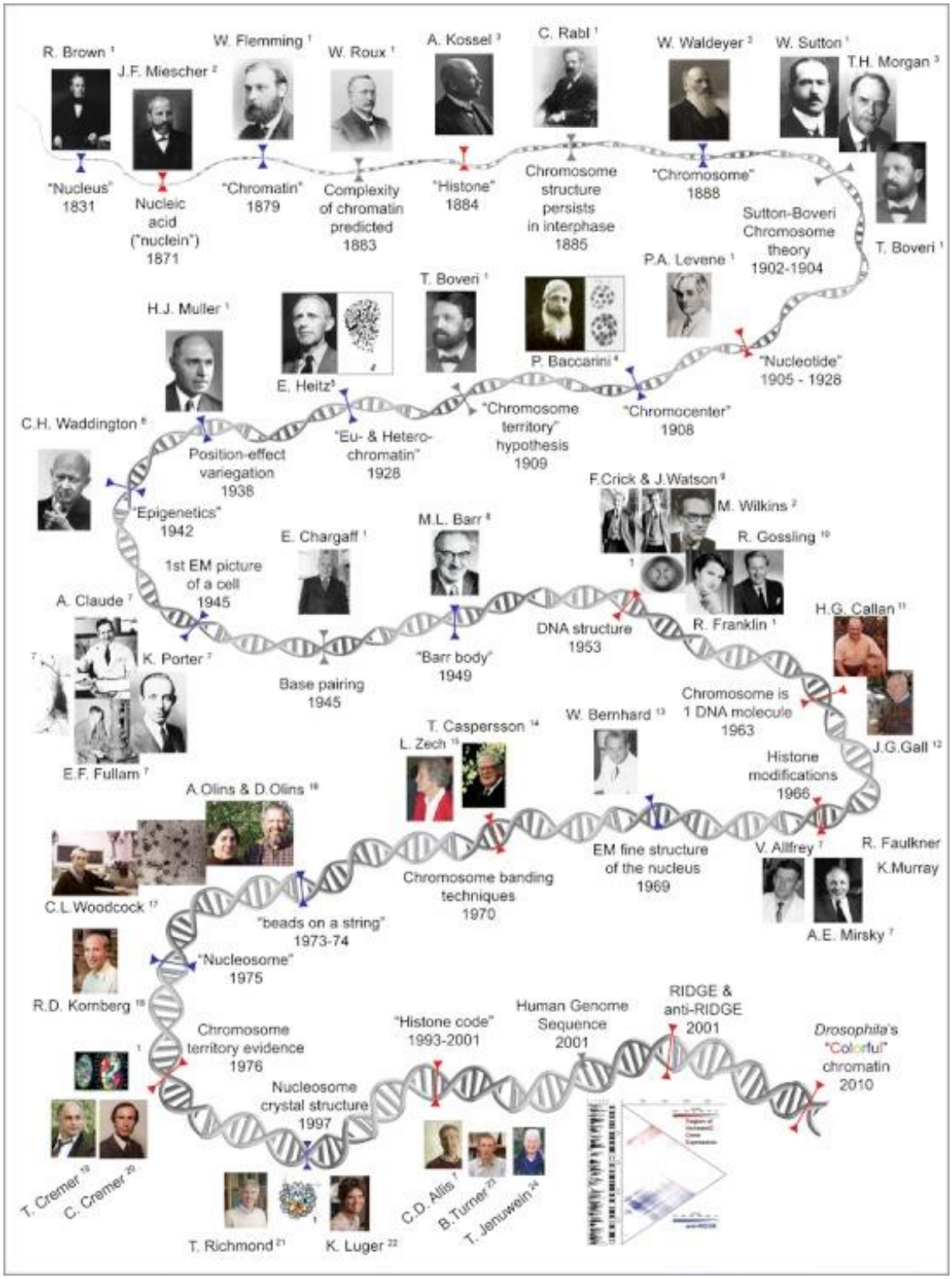


Figure 1. Timetable collection of landmark discoveries and concepts (gray) on molecular (red) and cellular (blue) aspects of chromatin. Taken from (Jost, Bertulat et al., 2012)

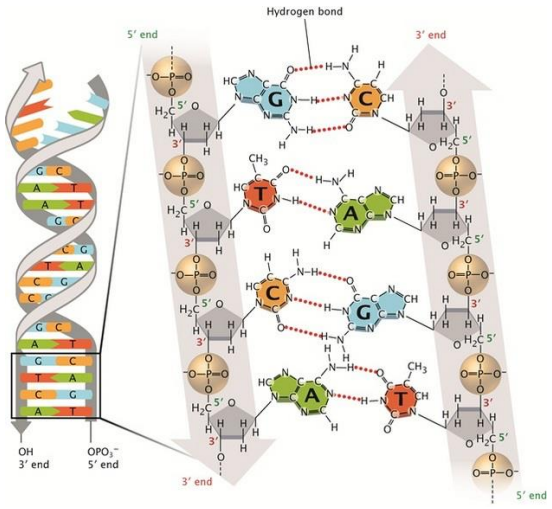
1.1.2. Nucleosome: the basic unit of chromatin

In eukaryotic cells, the genetic material is packed into a highly compacted nucleoprotein structure called chromatin inside the nucleus. Chromatin is composed of DNA and small basic proteins called histones. DNA molecules inside human cells with over two meters in length are composed of four nucleotides Thymine (T), Guanine (G), Cytosine (C) and Adenine (A). Each of these nucleotides contains a sugar (deoxyribose), a phosphate group and a nitrogenous base. Guanin-Cytosine and Thymine-Adenine through their nitrogenous bases are capable of base pairing and maintain the double helix structure of DNA (Figure 2A) (Alberts, Johnson et al., 2002). On the other hand, histones are the main protein components of chromatin, around which DNA is wrapped.

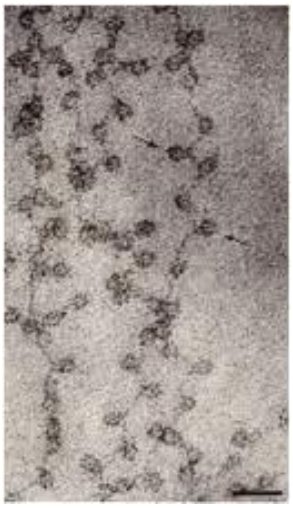
In 1970, over a decade many studies have been carried out to better understand the chromatin organization. In 1973, after an accidental usage of a DNA nuclease on the chromatin, Hewish and Bergoyne identified the basic repeating structural and functional sub-unit of the chromatin (Hewish & Burgoyne, 1973). Furthermore, in 1974, the electron microscopy images of the chromatin revealed that the chromatin structure appears as “beads on a string” (Figure 2B) (Olins & Olins, 1974), (Woodcock, Safer et al., 1976).

Finally, in 1975, Thomas and Kornberg following an experiment through which they chemically crosslinked histones to chromatin, demonstrated that approximately 1.7 left-handed superhelical round or 146 base pairs of DNA molecule is wrapped around a histone octamer (Kornberg, 1974, Van Holde, Sahasrabuddhe et al., 1974). These observations conducted to the discovery of the fundamental sub-unit of chromatin, named nucleosome by Pierre Chambon in 1975 (Germond, Bellard et al., 1976). In 1997, the X-ray crystal structure of the nucleosome core particle was established by Karoline Luger (Figure 2C) (Luger, Mäder et al., 1997, Richmond & Davey, 2003).

A



B



C

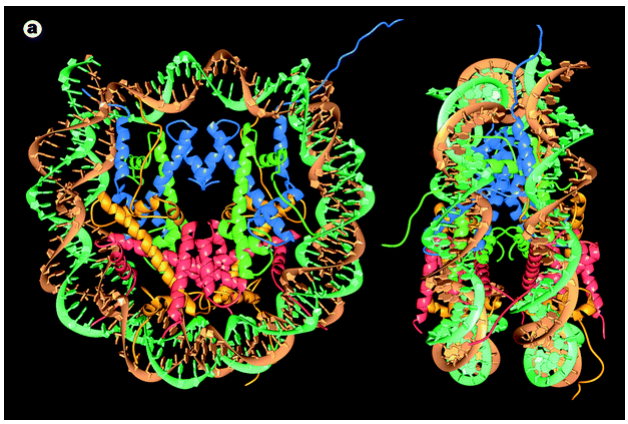


Figure 2 Chromatin composition and structure. A) The double helix model for DNA. Two polynucleotide chains twisted around each other representing the repeating pattern of a 5-carbon sugar and a phosphate group. Each of the four nitrogenous bases are bound to the base on the other strand via hydrogen bonds and are attached to one sugar, taken from © 2008 Nature Education. **B)** The electron microscopic image of beads on a string model of chromatin fiber, taken from (Olins & Olins, 2003). **C)** The top (left) and the later view (right) of the crystal structure of the core nucleosome, bearing histone octamer and 147bp of DNA strand wrapping around, taken from (Luger et al., 1997).

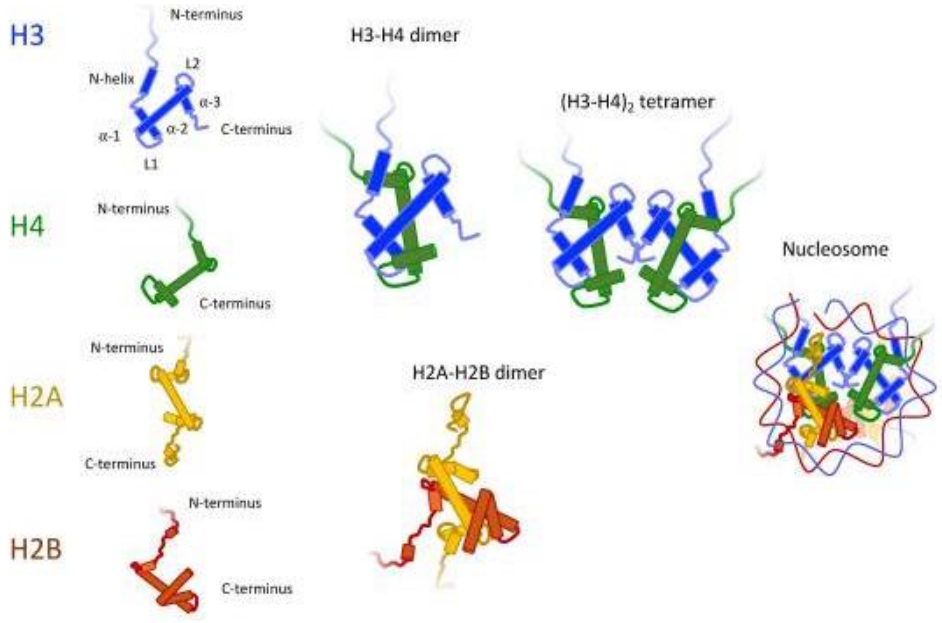
1.1.2.1. Canonical/ Core histones

As mentioned before, histones were discovered by Albrecht Kossel in 1884, through the purification of an acid-soluble protein fraction from the nuclei of chicken erythrocyte cells. The structure of the nucleosome core particle with 7 Å resolution was first established by Richmond in 1984 (Richmond, Finch et al., 1984). The crystal structure of the nucleosome core particle with a resolution of 2.8 Å was obtained by K.Luger in 1997, and later in 2003, the crystal structure of the nucleosome with the resolution of 1.9 Å was proposed by Richmond and Davey (Luger et al., 1997, Richmond & Davey, 2003).

The nucleosome core particle is composed of small basic proteins of canonical histones H2A, H2B, H3 and H4 ranging from 11 to 15 kDa. Inside the nucleosome core particle, one tetramer of H3-H4 interacts with two dimers of H2A-H2B to form an octamer (Figure 3A). The four core histones are very rich in lysine and arginine amino acids, which explain the basic nature of histones. This characteristic leads to the electrostatic interactions and the establishment of hydrogen bonds between positively charged histones and negatively charged deoxyribose sugar and phosphate backbone of DNA (Khorasanizadeh, 2004). In 1994, a phylogenetic analysis of core histones revealed that histones, which form dimers, have co-evolved during evolution and they are conserved from yeast to human (Thatcher & Gorovsky, 1994).

Canonical histones have a main central globular domain that is structured into 3 α helices (α1 - α2 and α3), connected by short loops L1 and L2, which mediate heterodimeric interactions between core histones and lead to the formation of histone fold, (Arents, Burlingame et al., 1991). During dimerization, the loop L1 of one histone aligns with the loop L2 of the other histone and create “handshake” motif. The handshake motif interacts with DNA via L1L2 sites and α1α1 sites of histones (Figure 3B), (McGinty & Tan, 2014, Ramaswamy & Ioshikhes, 2013).

A



B

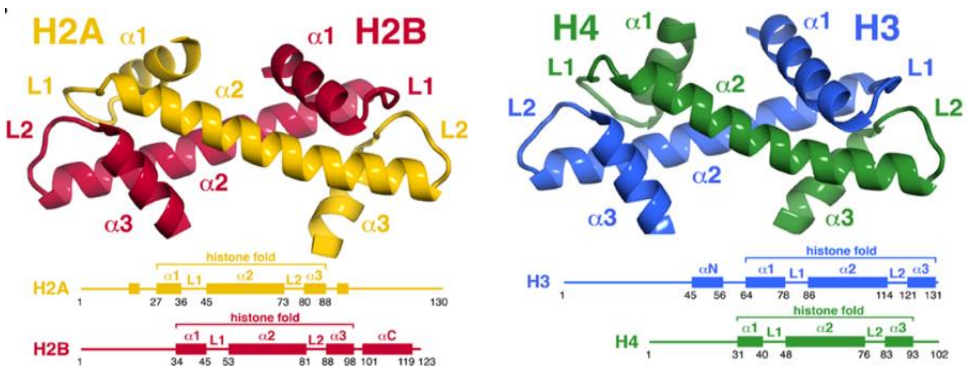


Figure 3. The crystal structure of the histone- fold heterodimers. A) H2A, H2B and H4 histone fold. Schematic representation of α1, α2, α3, N-helix and N/C termini of the four histone core. Two H3-H4 dimers fold to form an (H3-H4)2-heterotetramere. The tetramere is located to the center and is capped by two H2A-H2B dimers within the nucleosome core. Taken from (Pardal, Fernandes-Duarte et al., 2019). **B)** H2A-H2B and H3-H4 dimer's secondary structure (top) and schemes (bottom). Molecular graphics were using PyMOL software (the PyMOL Molecular Graphics System, version 1.6, Schrondinger, LLC- PDB ID 1KX5), taken from (McGinty & Tan, 2014).

In addition to the histone fold domain, core histones contain N-terminal and C-terminal extremities that are localized outside of the nucleosome core (Luger et al., 1997). N-terminal and c-terminal extremities are highly conserved from yeast to human. Flexible tails of the core histones interact with DNA via the minor groove. Since N/C-terminal extremities are exposed outside of the nucleosome, they are highly accessible to PTMs. PTMs may have a role in the structure of the nucleosome but can certainly modify nucleosome-nucleosome interactions and consequently can affect chromatin structure and are important sites of recognition by chromatin regulatory factors (Khorasanizadeh, 2004, Kouzarides, 2007, Strahl & Allis, 2000).

1.1.2.2. Linker histones

Linker histones H1 is known as the fifth histone, which interacts with the internucleosomal (linker) DNA, between two adjacent nucleosomes at the entry and exit site of the nucleosomal DNA. The assembly of one nucleosome with linker histones generate a structure called chromatosome (Simpson, 1978). Linker histones extend up to 20 bp the length of nucleosomal DNA protected from micrococcal nuclease. There are 10 isoforms known today for linker histone H1: (H1.1-H1.5, H1°, H1t, H1T2, HILS, and H1oo). These H1 isoforms could belong to different groups such as, embryonic, replication and differentiation-dependent isoforms. Some of the replication-independent H1 isoforms are as following: the cleavage stage (CS1 H1), the differentiation dependent H1 (H1°) and the testis-specific H1 (H1t) (Khochbin, 2001).

Linker histone H1/H5 family members, contain all a tripartite structure: one conserved globular domain of ~75 residues, an N-terminal tail of ~20-35 residues and a highly basic CTD of ~100 residues (Allan, Hartman et al., 1980). The globular domain is the essential domain for nucleosome recognition. It contacts the nucleosome dyad, and binds both linker DNAs, whereas the lysine-rich N-terminal has little involvement in chromatin binding. The binding of CTD of H1 disturbs the 2-fold symmetry of the nucleosome and confers an asymmetrical conformation to the nucleosome, which seems essential for the formation of higher-order chromatin structures (Figure 4) (Bednar, Garcia-Saez et al., 2017). Furthermore, the highly basic CTD of H1 would neutralize both the electrostatic and the negative charge of linker DNA, resulting in a stronger association between H1 and DNA. In addition, Turner and colleagues

demonstrated that CTD phosphorylation in a cell-cycle dependent manner, is correlated with alterations in chromatin condensation level (Turner, Watson et al., 2018).

It is of note that the acetylation or formylation of three lysines (K52, K73 and K85) within globular domain of H1.0 in human cell lines, and in H1.2-H1.5, in mice would result in the eviction of H1 from nucleosome (Bednar et al., 2017, Wiśniewski, Zougman et al., 2007).

The deletion of linker histones genes did not lead to mortality and showed certain levels of redundancy between different H1 isoforms, as well as a complex regulatory mechanism between linker histones and chromatin structure (Khochbin, 2001, Robinson & Rhodes, 2006).

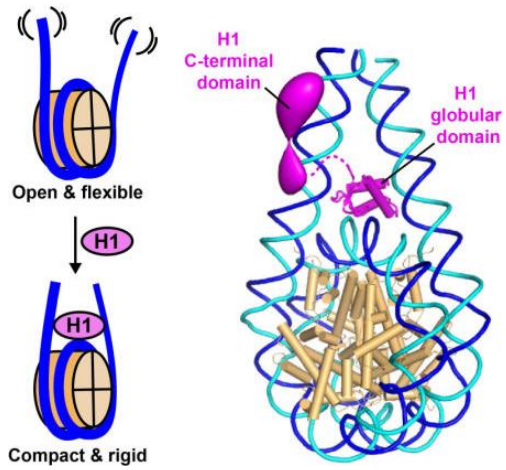


Figure 4. Schematic representation of H1-containing nucleosome. H1 binding through its CTD, disturbs the symmetry of two DNA strands and generates a more compact nucleosome. Taken from (Bednar et al., 2017).

2. Hierarchical level of chromatin organization

Nucleosome as the basic unit and the primary structural basis of chromatin, plays an important role in the compaction and packing of ~ 2 meters of DNA inside the ~10 to 100 μm3 of human cells. Positively-charged histone octamers are considered to mediate the first level of chromatin compaction by neutralizing the negative charge of acidic nature of DNA molecules (Talbert & Henikoff, 2017). Meanwhile, the compaction level that nucleosomes confer to the chromatin is not sufficient for packing DNA inside the nucleus. Observations by electron microscopy showed that nucleosomes, as the first level of chromatin organization,

form an 11nm nucleofilament structure similar to “beads on a string”, in which beads represent nucleosomes and the string represents DNA (Figure 5A), (Felsenfeld & Groudine, 2003, Olins & Olins, 1974, Oudet, Gross-Bellard et al., 1975, Woodcock et al., 1976). Specific interaction between nucleosomes results in the generation of the second level of chromatin compaction, which consists in the formation of a 30nm fiber of packed nucleosomes (Figure 5B). At this stage of compaction, two well-known structural models exist, based on in vitro experiments. The first model, known as interdigitated solenoid or “one-start helix” is characterized by interactions between consecutive nucleosomes (Figure 6A, the left image), whereas in “two-start helix” or zigzag model, interactions occur between alternate nucleosomes (Figure 6A, the right image) (Luger, Dechassa et al., 2012, Robinson & Rhodes, 2006). Recently, the crystal structure of a H1-bound hexanucleosome reveals a flat untwisted, two-start chromatin fiber conformation. Upon a shift in the ionic environment, this flat conformation can switch to a twisted conformation, suggesting a possible chromatin assembly pathway for the 30nm fiber (Figure 6B) (Garcia-Saez, Menoni et al., 2018).

Higher levels of chromatin fiber compaction are less known today due to the limits in the investigation methods and experimental procedures. Today, we believe that the 30 nm fiber is fold up into higher-ordered structured loops or nucleosome arrays that are visible as chromosomal domains in interphase (Figure 5C) (Cremer & Cremer, 2001, Taddei, Hediger et al., 2004). Moreover, mitotic chromosomes represent the highest level of compaction (Figure 5D).

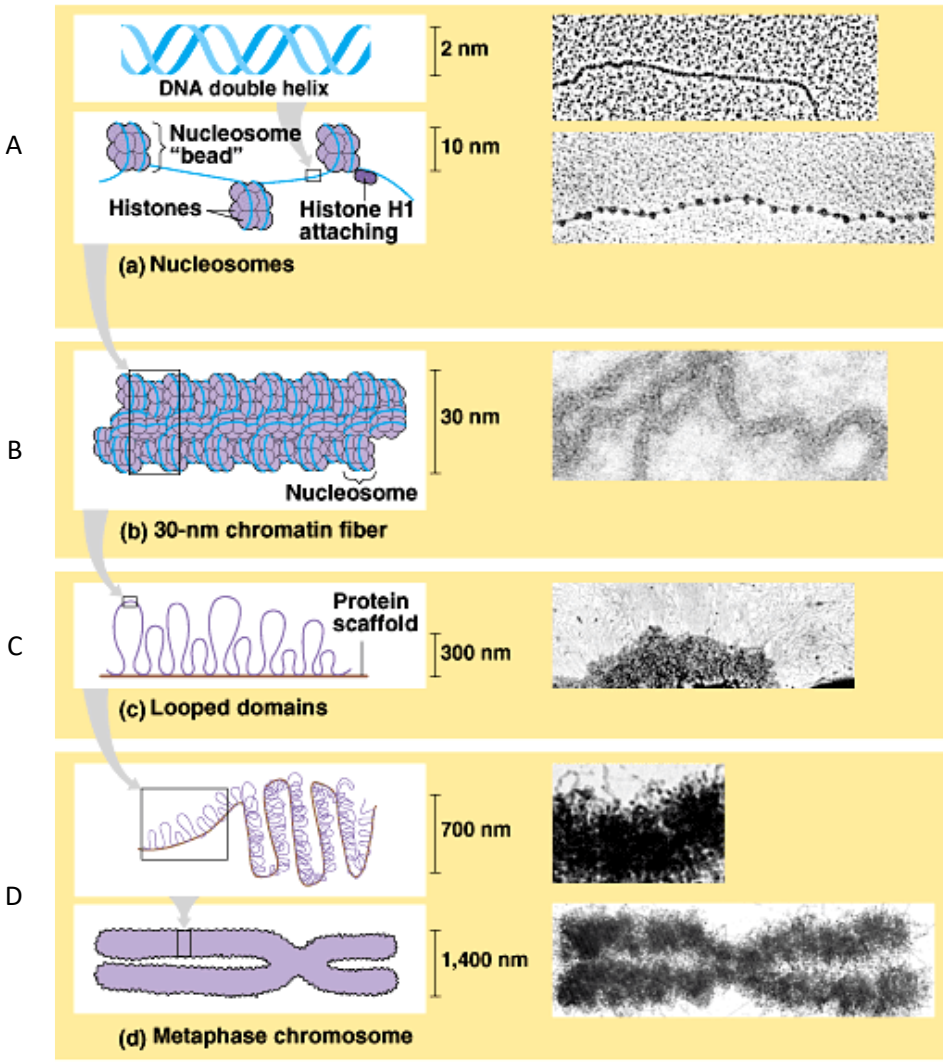


Figure 5. Different level of chromatin organization. A) Nucleosome is the first level of chromatin compaction, corresponding to 10nm fiber. **B)** The 10nm fiber wraps around itself and generates 30nm chromatin fiber. **C)** The chromatin wraps around itself again and generates the 300 nm fiber. **D)** generation of metaphasic chromosome which corresponds to the highest level of chromatin compaction. Taken from Pearson education, 2012.

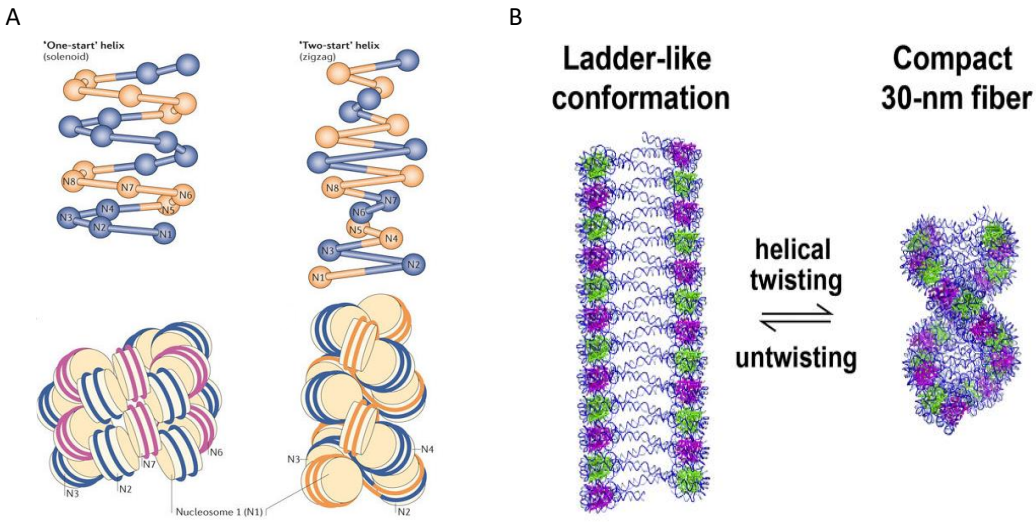


Figure 6. Three models for chromatin secondary structure. A) Two classical models (Selonoid and zigzag), which are characterized by the interaction between consecutive (n, n+1) and alternate (n, n+2) nucleosomes. Taken from (Luger et al., 2012) **B)** Recently proposed model suggesting ionic-dependent conformation between a flat, untwisted, two-start chromatin fiber and a helical twisted 30nm fiber. Taken from (Garcia-Saez et al., 2018).

3. Functional compartments of chromatin

In human, the genetic information is encoded by 25000 genes, which surprisingly represent only 1.5% of total DNA. The integrity of these sequences is essential to assure the normal cell functions. In 1928, Emile Heitz by using a novel in situ method in cytological experiments, noticed that some regions of the cell nucleus are brighter than other regions at interphase. Therefore, he decided to call the brighter regions "euchromatin" meaning the relaxed chromatin and more dense regions "heterochromatin" meaning the more compact chromatin (Figure 7). The chromosomal regions that harbor the heterochromatin are telomers, pericentric and centromeric heterochromatin, called constitutive heterochromatin. Moreover, heterochromatin is subdivided into facultative and constitutive heterochromatin. Facultative heterochromatin is formed at gene-rich regions in order to ensure the proper regulation of genes, while heterochromatin regions remain silent throughout the cell cycle and in MEF cells are associated to subnuclear compartments referred to as chromocenters. Euchromatin is mostly composed of transcriptionally-active genes and is often referred to as protein-coding

region, whereas heterochromatin is referred to as non-coding protein compartment of chromatin. In addition, euchromatin regions replicate during S-phase, while heterochromatin regions are only replicated at the end of the S-phase (Rhind & Gilbert, 2013).

For long time, the scientific community believed that heterochromatin is associated to transcriptionally inactive chromatin. Many scientific evidences are now demonstrating that the heterochromatin regions are transcribed under certain conditions such as under heat shock in human cells (Jolly, Metz et al., 2004). In addition, a strand-specific transcription of pericentric satellites has been proved essential for the early stage of mice embryonic development (Probst, Okamoto et al., 2010).

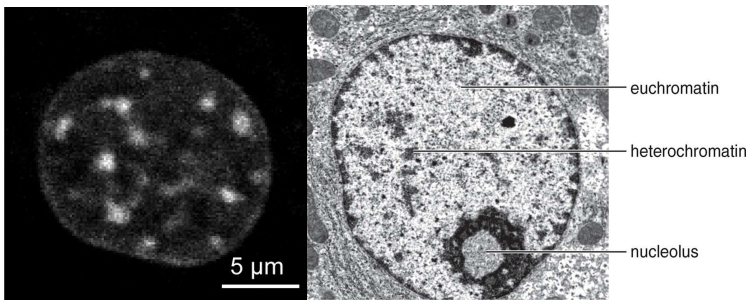


Figure 7. Schematic depiction of heterochromatin and euchromatin. Right image : electrone microscopic representing euchromatin (brighter) and heterochromatin (darker) regions, taken from (site http://schoolbag.info/chemistry/mcat_biochemistry/34.html). Left image: heterochromatin regions fuse together and create chromocenters or Dapi-dense regions in MEF cells (Imai, Nozaki et al., 2017).

II. Epigenome Dynamic and plasticity

1. Definition of Epigenetics

In 1942, the term "Epigenetics" (above the genes) was proposed by Conrad Hal Waddington to consider the differences in phenotype despite the identical genotype of the cells in different tissues (Campbell & Wood, 2019). Epigenetic *stricto sensu* defines all the information, which are inherited from one generation to another. However, a broader definition considers all parameters that are essential for the regulation of chromatin structure and function, independently of the DNA sequence. Epigenome refers to the ensemble of "bookmarks", which accompany DNA or histones with possible action on the structure of chromatin, resulting in a gene-expression landscape. Factors considered in the field of epigenetics include histones, histone variants, non-histone chromatin remodelling complexes, chemical PTMs of histones, DNA methylation and non-coding RNAs. Molecular mechanisms relying on these factors play crucial roles during cell differentiation and cell division. These mechanisms could be involved in transgenerational passage of gene-regulatory information as well as in the impact of environmental signals for transcriptional regulation of the cells.

The latest discoveries demonstrated promising therapeutic potentials of approaches that include the epigenome editing concept (Stricker, Koferle et al., 2017). As one example of this new concept, it is worthy to mention that in the absence of related transcription factor (GATA1), a zinc-finger-LDB1 fusion protein can target the inactive β-globin locus in erythroblast cell line. This results in the formation of a chromatin loop between the promoter and the locus control region and leads to the activation of β-globin transcription (Deng, Rupon et al., 2014). Current advances in epigenetic engineering suggest a more dynamic environment rather than a constant landscape as described by Waddington (Figure 8) (Stricker et al., 2017).

A B

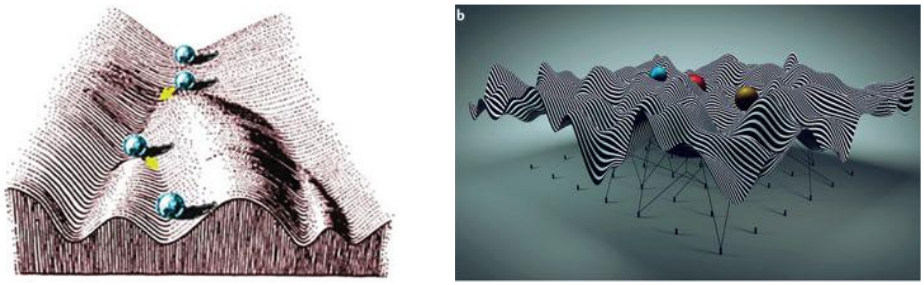


Figure 8. Two models of epigenetic landscape. A) Epigenetic landscape proposed by Waddingtone, a metaphore of epigenome mechanism, taken from (Noble, 2015). **B)** Novel epigenetic landscape model proposed by (Stricker et al., 2017), representing epigenome editing.

In this chapter, all the factors involved in the regulation and modulation of the epigenome in a direct or indirect fashion, are precisely described.

2. DNA methylation

Cytosine methylation under the form of 5-methylcytosine or 5-mC is an inherited, reversible covalent DNA modification found in all the vertebrates, flowering plants, certain fungus and invertebrates, as well as protists and in some bacterial strains. It consists of adding one methyl (CH_3) group on the 5^{th} carbon of the cytosine residue. DNA methylation occurs mostly in CG dinucleotides of the genome and is often associated with transcriptional repression when it occurs on gene expression regulatory elements (Bogdanovic & Lister, 2017). Cytosine methylation is considered as the modification, which plays a functional role in many epigenetic mechanisms such as imprinting, X-chromosome inactivation, and silencing of repetitive DNA.

This modification is added on the cytosine by a family of conserved enzymes called DNMTs. DNMT carboxy-terminal catalytic domain transfers the methyl groupe on the 5^{th} carbon of cytosines in CpG dinucleotides and then locally deforms the DNA in order to have access to the targeted cytosine for the *de novo* deposition of the methyl group (Goll & Bestor, 2005, Rountree, Bachman et al., 2000). It has been demonstrated that DNA methylation is

essential for the survival of the higher eukaryotes since *Dnmt* gene depletion in mice results in embryonic lethality (Li, Bestor et al., 1992, Okano, Takebayashi et al., 1999). Three families of DNMT enzymes (DNMT1, DNMT2 and DNMT3) play an essential role in the establishment and maintenance of DNA methylation.

- DNMT1 is the main enzyme responsible for the propagation of DNA methylation during replication through interaction with components of replication machinery. Beside the capacity of the C-terminal region of this enzyme to maintain methylation pattern during replication, the long non-catalytic N-terminal domain of DNMT1 can distinguish between hemi-methylated and unmethylated DNA and can repress transcription through direct interaction with HDAC2 and its recruitment to transcriptionally active regions (Rountree et al., 2000).
- DNMT2 is the most conserved enzymes of DNMT family throughout evolution (Goll & Bestor, 2005). In many eukaryotic organisms, DNMT2 methylates cytosine at the position 38 of tRNAs. Currently, more than 90 different modifications have been discovered on tRNA molecules, and more importantly, their roles in response to various stimuli such as stress have been highlighted. The methylation at cytosine 38 of tRNA anti-codon loop protects this molecule from stress-induced cleavage, which can in turn modify the dicer-dependent siRNA pathway (Durdevic, Mobin et al., 2013, Schaefer & Lyko, 2010).
- DNMT3A and 3B are two functional cytosine methyltransferases of DNMT3 family enzymes that are implicated in *de novo* methylation of CpG islands, principally during germ cell development and in early embryonic development. These two enzymes transfer the methyl group to hemi-methylated or unmethylated cytosines at CpG islands and are less expressed than DNMT1 enzymes in adult tissues. The *Dnmt3A*-depleted mice die at young adult ages and are sterile. DNMT3 like (DNMT3L) is the third homolog of this family, which is deficient for methyltransferase activity and has been shown important for methylation pattern in male and female germ cells (Goll & Bestor, 2005).
- DNMT3C is a recently identified *de novo* DNA methyltransferase, that is evolved via the duplication of Dnm3B gene in rodents. This enzyme methylates the promoters of retrotransposons in the male germ cells, and is involved in the epigenetic control of these retrotransposons (Barau, Teissandier et al., 2016).

2.1. DNA methylation in coding and non-coding compartments of the genome

In vertebrates, DNA methylation occurs throughout the whole genome but in invertebrates and non-flowering plants, cytosine methylation follows a random pattern. For instance, it mostly occurs at repetitive DNA elements and within many actively transcribed genes. Cytosine methylation at repetitive DNA sequences, as a protective mechanism to inhibit the propagation of these elements throughout the whole genome, could be a good explanation for a role for methylation in order to maintain the genome integrity. In contrast, DNA methylation within highly transcribed genes does not seem to be relevant for gene regulation. However, recent studies propose that DNA methylation might have a role in the suppression of intragenic promoters, counterbalancing their possible activity at the time of chromatin disruption at the time of transcription by RNA pol II. Today, despite all of our advanced discoveries, our insight into the molecular and functional mechanisms underlying this epigenetic mark needs to be improved (Schubeler, 2015).

2.2. DNA methylation and genomic imprinting

In addition to DNA methylation in adult tissues, this modification plays a crucial role in a phenomenon called genomic imprinting (Kelsey & Feil, 2013). Genomic imprinting is associated with a differential regulation of mono-allelic parental genes. In another words, during genomic imprinting independently of paternal or maternal genes, the alleles that are methylated are transcriptionally repressed. Genomic imprinting is an epigenetic process, associated with DNA methylation that does not affect the DNA sequence. DNA imprinting is established in adult germ cells and remains through somatic cell division of an organism.

ICRs are the essential domain, which control the binding of regulatory factors such as the transcription factors. Allele-specific methylation of ICR inhibits the binding of transcription factors to the regulatory regions and promotes gene silencing. ICRs are rich in CpG dinucleotides and they bear the parent-of-origin DNA methylation pattern, which resists throughout the development. Very interestingly, recently the imprinted regulatory non-coding RNAs (ncRNAs) have been shown to become crucial in the tissue-specific expression of imprinted genes. Additionally, the regulatory role of lncRNAs) in imprinted gene expression

has also been shown (Sanli & Feil, 2015). Figure 9 explains the DNA methylation status of male and female genome starting from PGCs throughout early development in mice.

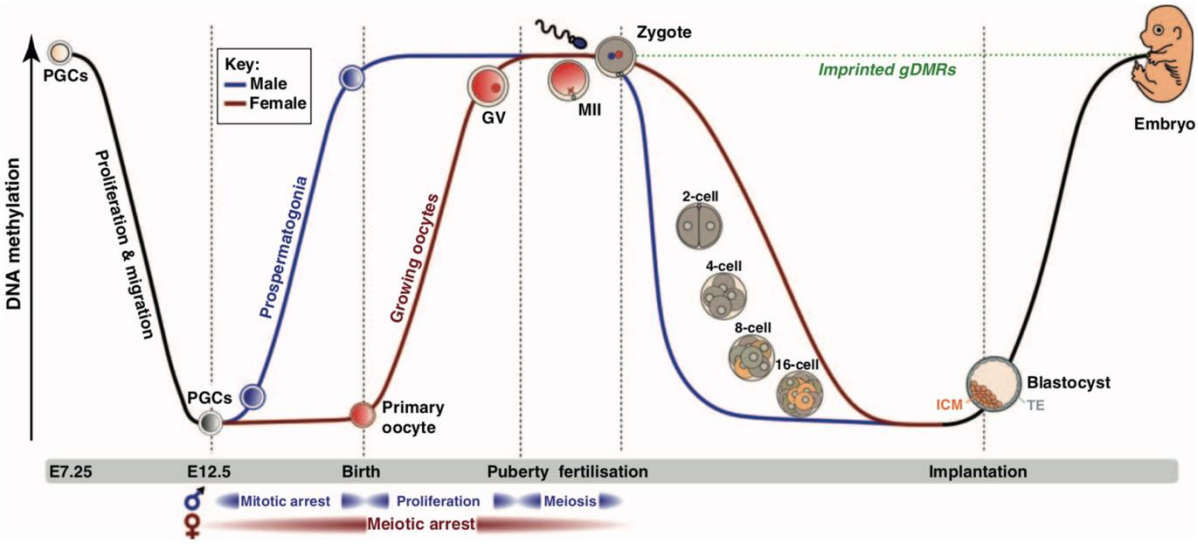


Figure 9. Levels of DNA methylation in PGC, primary oocytes, and after implantation. In the male and female mice, DNA methylation is at its maximum at embryonic stage 12.5. During the mitotic arrest and after birth, DNA methylation increases and reaches its highest level. Therefore, during spermatogenesis and oogenesis, DNA methylation is at its highest level. After implantation, DNA demethylation occurs on both paternal and maternal genome until blastocyst, and again the *de novo* DNA methylation occurs until the embryo. Taken from (Smallwood & Kelsey, 2012).

3. DNA demethylation

DNA demethylation refers to removing or modifying the methyl group from 5meC, which occurs in either an active or a passive way. The latest data on active DNA demethylation involves two TET and TDG groups of enzymes (Wu & Zhang, 2017). In mammals, three subfamilies of TET proteins TET1, TET2, and TET3 have been further characterized. It has been demonstrated that TET1 catalyzes the reaction of 5meC to 5-hmeC *in vitro*. This enzyme has been found dependent to 2OG and Fe (II) and its depletion resulted in a decreased level of 5hmC (Tahiliani, Koh et al., 2009). Furthermore, TET2 and TET3 activity has been also associated to 5mC conversion to 5hmC.

In addition, TET proteins promote the oxidation of 5mC to 5fC and 5caC. In order to complete the active DNA demethylation, TDG recognizes 5caC as the substrate and promotes the TDG-initiated BER, which lead to the incorporation of a non-methylated cytosine into DNA in a replication-independent fashion (He, Li et al., 2011). The unmodified cytosine can be also restored in a replication-dependent way through AM-PD mechanism. Upon the incorporation of the unmodified cytosine into the newly synthesized DNA strand throughout the replication, DNMT1 is recruited to 5mC:C site through UHRF1(Bostick, Kim et al., 2007). Moreover, it has been demonstrated that DNMT1 has much less affinity for the 5hmC, 5acC and 5fC sites (Ji, Lin et al., 2014). This later results in a less methylated cytosine throughout the DNA replication and a dilution of methylated cytosine after each replication.

In mammals, DNA demethylation is implicated in a variety of different mechanisms such as pre-implantation development, PGC development, pluripotency and differentiation, neuronal functions and cancer (Saitou, Kagiwada et al., 2012, Wu & Zhang, 2014, Wu & Zhang, 2017). In the case of pre-implantation development, the parental genome undergoes an epigenetic reprogramming with a highlight of DNA demethylation. At post-fertilization there is 50% decrease in DNA methylation of the paternal genome, which is mainly mediated by replication-dependent dilution and TET-3 action. The maternal genome also undergoes a decrease in DNA methylation but to a less extent compared to the paternal genome. It has been shown that the DPPA-3 or PGC7 is recruited to the maternal genome via H3K9me2, which might protect maternal genome from TET-3 effect (Nakamura, Liu et al., 2012).

4. Histone Post-Translational modifications

4.1. General Overview on histone PTMs

PTMs refer to the addition of one or multiple chemical group(s) to one or multiple amino acid residue(s) through a covalent and enzymatic reaction, which subsequently induce new properties to one given protein. Adding PTMs on histones have been widely studied during last fifty years. These chemical modifications on histones are recognized by various chromatin-binding factors, which may in turn alter the chromatin structure but also constitute a sign-posting system (Ruthenburg, Li et al., 2007). The effect of histone/protein PTM could depend

on two major mechanisms. One mechanism depends on varying the electric charge of the chromatin or on altering the intra-nucleosomal interactions. The second mechanism could depend on engaging extra-nucleosomal chromatin regulators, who specifically bind PTMs and are capable of mediating a change, directly or indirectly in chromatin structure (Kouzarides, 2007, Wysocka, Swigut et al., 2005, Wysocka, Swigut et al., 2006). The outcome of these chromatin structural changes is either an activation or a repression of gene expression.

PTMs mostly occur at C or N-termini amino acid residues of histone tails but they can also affect histone core particles. Furthermore, some of these chemical modifications are also present within the globular domain of core histones, which are essential for histone-DNA and histone-histone interactions (Cosgrove, Boeke et al., 2004, Garcia, Hake et al., 2007, Mersfelder & Parthun, 2006).

There are a number of PTMs including lysine acetylation, lysine and arginine methylation, ubiquitylation, sumoylation, serine and threonine phosphorylation, butyrylation, formylation, propionylation, citrullination, proline isomerization and ADP ribosylation. These modifications are reported on more than 60 different positions of histones (Martin & Zhang, 2007, Ruthenburg et al., 2007). In 2011, Minjia Tan and colleagues through a specific mass-spectrometry based proteomic approach discovered more than 67 new histone PTM/positions. In addition, two new histone modifications, tyrosine hydroxylation (Yoh) and lysine crotonylation has been identified during this study (Tan, Luo et al., 2011). Another PTM (lysine malonylation) was reported the same year also based on mass spectrometry approach (Peng, Lu et al., 2011). A complete summary of all of the PTMs as positions known until today is illustrated in figure 10 (Sabari, Zhang et al., 2017).

Enzymes, which are responsible for dynamic regulation (deposition and removal) of some of these PTMs such as acetylation, methylation, phosphorylation, crotonylation, butyrylation, sumoylation and malonylation, have been further identified, and will be briefly discussed in this section.

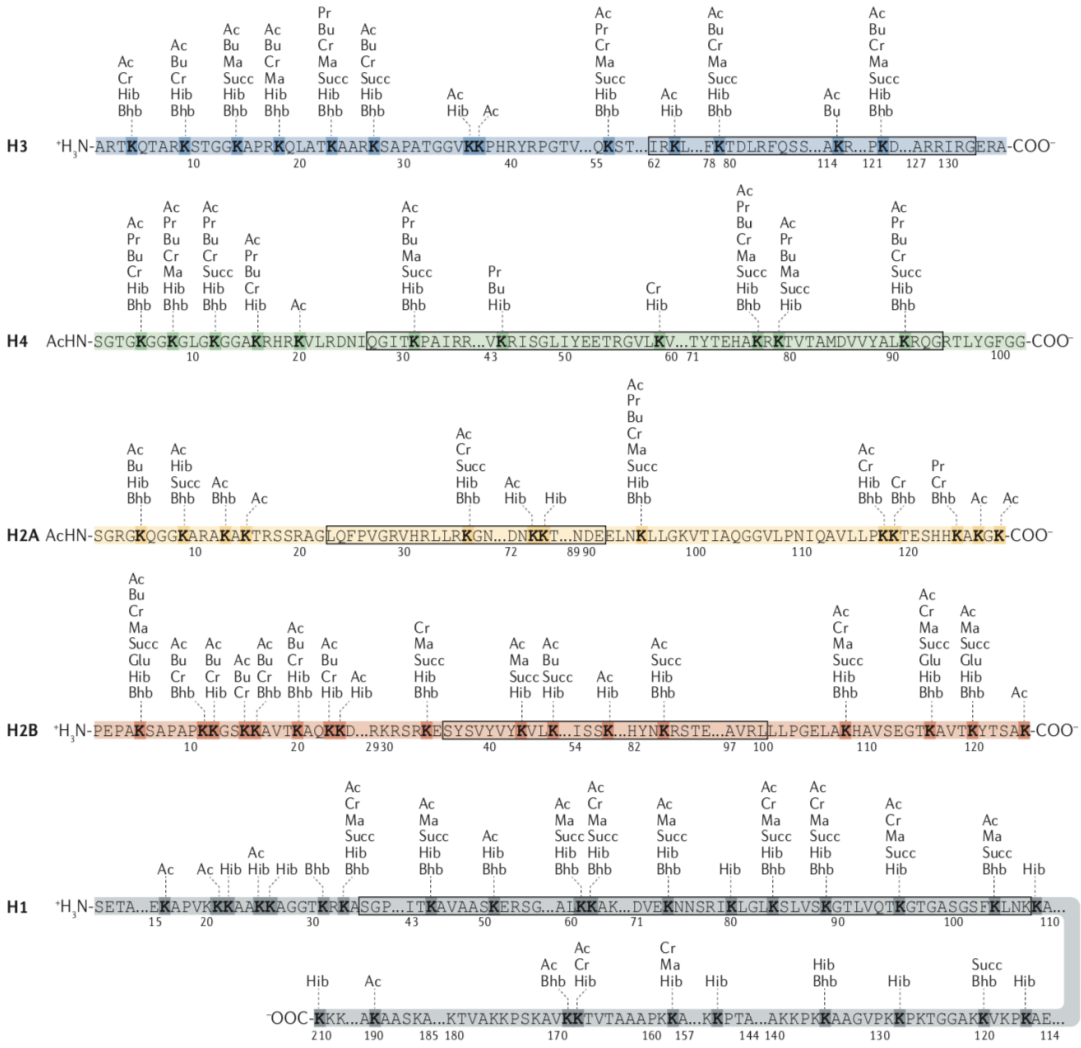


Figure 10. Recent overview of lysine acylations and their distribution on histones. Taken from (Sabari et al., 2017).

4.2. Acetylation

Acetylation as a reversible modification, consists of transferring one acetyl (CH3-CHO) group from the acetyl-coA cofactor to the amine group of lysine residue within histones. Adding this modification neutralizes the positive charge of lysines, alters electrostatic interactions between histones and DNA, which then leads to a more relax, and less compact chromatin (Kouzarides, 2007, Zentner & Henikoff, 2013) and it is important for both transcriptional initiation and elongation. In addition, lysine acetylation on histone N-terminal

tail provides more accessible DNA for binding of transcription factors and is generally associated to an active transcription status of the chromatin. The acetyl group is deposited on histone core and tail lysines by specific enzymes called HATs. This modification is removed from these regions by enzymes called HDACs or sirtuins.

4.2.1. HATs

In 1995, James E. Brownell and D. Allis successfully extracted the type-A histone acetyl transferase from the nuclei of the ciliate, tetrahymena. This led to the identification of the type-A or "nuclear" histone acetyltransferase enzyme, P55, as the homolog of Gcn5 gene in yeast which was at the time a well-known yeast adaptor protein necessary for the activation of a wide variety of transcriptional activators (Brownell & Allis, 1995, Brownell, Zhou et al., 1996, Georgakopoulos & Thireos, 1992). Furthermore, they also identified a highly conserved bromodomain motif at the c-terminal of both P55 and Gcn5. At that time, this motif had been observed in a variety of different proteins that had direct functions in transcriptional activation. Despite all of these observations, a clear vision of the exact role of these catalytic domains and of the conserved bromodomain, and their role in inducing transcriptional activation had not been achieved.

Today, HATs are divided into 6 main families based on the sequence divergence of their HAT domains. These families are 1) HAT1 (KAT1), 2) Gcn5 (KAT2) / (hKAT2B)/ GNATs, 3) MYST, human MOZ (hKAT6A) - yeast Ybf2/Sas3 (ySAS3, KAT6), yeast Sas2 (ySAS2, KAT8) and human Tip60 (hTIP60, KAT5)), 4) p300/CBP (two human paralogs hp300 (KAT3B) and hCBP (KAT3A)) and 5) Rtt109 (yRtt109, KAT11) (identification as a yeast regulator of Ty1 transposition gene product 109). The sixth family of HATs are steroid receptor coactivators such as human ACTR/AIB1, human SRC1, human TAF250, human ATF2 and human CLOCK. Depending on their nuclear or cytoplasmic localization, HATs are divided into two general groups; Type A or the nuclear HATs, which were identified basically as the cofactors of transcription, which also play important role in DNA repair mechanisms and cell cycle progression. Type B HATs acetylate histones inside the cytoplasm. X-ray crystallography structural studies revealed one structurally conserved core region between all of these five HAT families despite a diversity in their sequence. This core region essentially

harbours one 3-stranded β-sheet and one long α-helix parallel which cover one side of the β-sheet and it interact with Acetyl-CoA cofactors (Yuan & Marmorstein, 2013).

4.2.1.1. GCN5 family

GCN5 is one of the most studied HATs conserved from yeast to human. GCN5 and PCAF are acetyltransferase subunits of multi complex mammalian proteins called SAGA (also known as STAGA, TFTC and ATAC). These subunits are responsible for the acetylation of nucleosomal H3 especially on lysines 9/14/18/27 (Guelman, Kozuka et al., 2009, Jin, Yu et al., 2011, Wang, Faiola et al., 2008).

4.2.1.2. MYST family

MYST is a large family of highly conserved proteins from yeast to human, which contain HAT activity. Myst family is composed of five human HATs; MOZ, MORF, Tip60, MOF and HBO1 (Avvakumov, Lalonde et al., 2012). MYST family of proteins are crucial for some cellular activities such as Dosage compensation in *Drosophila,* oncogenic alterations resulting in different types of cancer development such as leukaemia (Carrozza, Utley et al., 2003) and gene silencing in yeast.

Tip60 HAT, in addition to histones, is responsible for acetylation of non-histone protein substrates such as oncogene c-Myc and tumor-suppressor proteins p53 and Rb. Tip60 acetylation of non-histone proteins could have major effects on the stability and function of these substrates. For example, the proteasome machinery recognizes c-terminal acetylation of Rb and degrades this protein. Tip60 also acetylates histones such as H2A at lysine 5, H3 at lysine 14 and H4 at lysines 5/8/12 and 16 (Ikura, Ogryzko et al., 2000).

The Sas2 and Sas3 members of MYST family in yeast are known to play a role in transcriptional silencing. In particular, Sas 2 is recognized to enhance the silencing at telomeric regions (Cai, Jin et al., 2010, Carrozza et al., 2003).

4.2.1.3. CBP/p300 family

CBP and p300 are transcriptional coactivators that harbor acetyltransferase activity via their HAT domain and are mostly expressed in humans and higher eukaryotes.

Rtt109 is a structural homolog of CBP and p300 in yeast and in *Arabidopsis thaliana.* P300 protein is about 300kDa in size and was first identified as a protein, which binds to E1A an adenoviral oncogenic transcription factor. Meanwhile CBP was identified as a protein, which binds to CREB transcription factor. The two proteins are called CBP/p300 because of the high sequence homology between the two proteins (Beverley M. Dancy, 2015).

CBP/p300 interact with many proteins that in most cases target them to enhancers. In vivo, This protein is also implicated in the acetylation of H3K18 that is known as a transcriptionally active mark (Jin et al., 2011). A comprehensive proteomic analysis of p300 acetylation *in vivo* revealed additional preferred acetylation sites on histones. Interestingly, numerous acetylation on H2B N-terminal tail such as H2B K5, K11, K12, K15, K16, K20 and K23, were identified. In addition to N-terminal H2B lysine acetylation, H3K18, H3K27 and H3K36 sites were among the most pronounced acetylation sites on histone H3. These data identified first the unknown p300 acetylation lysine sites and second, new CBP/p300 histone acetylation sites, which could be used as markers for the catalytic activity of this enzyme (Weinert, Narita et al., 2018). Moreover, a genome-wide transcriptomic analyses in mouse early (round) and late (elongating) post-meiotic germ cells, identified a group of genes undergoing a global shut down in spermatocytes and round spermatids but become activated at early elongating stage, right before the global shut down of transcription. Interestingly, the overlap of transcriptome and chip-seq data from CBP/p300 knock down mice, demonstrated the similarity between these two groups and they were involved in metabolic remodeling. These data suggest a CBP/p300-dependant gene-expression programming that mediate specific metabolic state in pre- and post-meiotic spermatids (Boussouar, Goudarzi et al., 2014). CBP/p300 harbor many protein domains that are essential for their implication in various cellular functions (Figure 11) (Dancy & Cole, 2015).

HAT activity of CBP/p300 is essential for gene expression regulation specifically at transcriptional regulatory regions such as promoters and enhancers. Recently, the crystal structure of p300 with propionyl, butyryl and acetyl-CoA has been solved which nicely

describes the transfer of acyl chain to the substrate (histone tails in this context) (Kaczmarska, Ortega et al., 2017). One important factor that regulates the acetyltransferase activity of CBP/p300 protein is its HAT domain auto-acetylation mechanism. It has been shown that the hypo-acetylated form of HAT domain reduces the enzymatic catalytic activity of CBP/p300 (Thompson, Wang et al., 2004). Recently, Ortega and colleagues through a structural analysis revealed the mechanism by which a *trans*-autoacetylation of HAT domains of two adjacent p300 molecules is the consequence of two transcription factors dimerization and their subsequent binding to CBP/p300. This recent finding is explained in detail in the result section of this manuscript (Ortega, Rengachari et al., 2018).

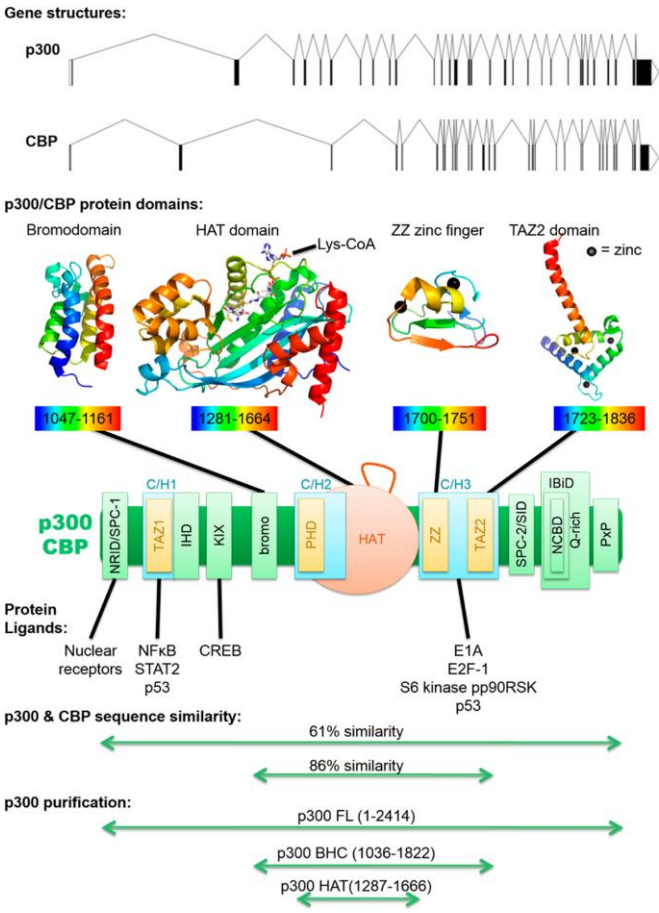


Figure 11. Schematic representation of different domains of p300 (top) and CBP (below). Example of crystal structures for the main domains of CBP/p300 such as bromodomain, catalytic HAT domain, ZZ zinc finger (solution NMR) and TAZ-2 are represented. Taken from (Dancy & Cole, 2015).

4.2.2. HDACs

Histone deacetylases are enzymes responsible for removing the acetyl group from histones, therefore neutralizing the effect of HATs on gene transcriptional regulation. HDACs have been found in plants, animals and fungi, as well as in bacteria and they have crucial importance for many cellular functions, such as chromatin assembly, DNA repair and recombination. Here chromatin-related functions of this group of enzymes are further discussed. A phylogenetic analysis divided HDACs into four groups: Group 1 encompasses all the protein homologues of Rpd3 (in yeast), which corresponds to HDAC1, 2, 3 and 8 in human. Group 2 contains homologous proteins of Hda1 (in yeast), which corresponds to HDAC 4, 5, 6, 9 and 10 in human. Homologous proteins of Sir2 (in yeast) compose the third group of HDACs, which correspond to the sirtuine protein family of 1 to 7 in human. The last group includes HDAC11 (Yang & Seto, 2007).

Rpd3 deacetylates four core histones and Hda1 deacetylates H3 and H2B. These two enzymes are respectively recruited to chromatin through DNA-binding transcription regulator and DNA-binding repressor factors. Deletion of Rpd3 and Hda1 is associated with hyperacetylation of histones. Moreover, HDACs may induce transcriptional repression by recruiting repressors either in a direct or indirect manner. For instance, HDAC3 is responsible for deacetylation of histone H4 on lysine 5, which in turn promotes the recruitment of a mammalian repressor SMRT (Shahbazian & Grunstein, 2007).

HDAC6 discovered in our laboratory (Verdel & Khochbin, 1999) was found to be cytoplasmic (Verdel, Curtet et al., 2000) and associated to microtubules and tubulin deacetylation (Hubbert, Guardiola et al., 2002, Matsuyama, Shimazu et al., 2002, Zhang, Li et al., 2003). In addition, our laboratory discovered the ubiquitin-binding activity of HDAC6 (Seigneurin-Berny, Verdel et al., 2001) as well as its biochemical and structural properties (Boyault, Gilquin et al., 2006). A role for HDAC6 in cellular stress response has been shown by several groups. This activity of HDAC6 relays on two major mechanisms; first by managing protein aggregates via aggresome formation and the autophagic degradation of the aggregates and second through triggering the dissociation of the repressive HDAC6/HSF1 complex and activating HSF1, which subsequently induces the expression of HSPs (Boyault, Zhang et al., 2007, Kawaguchi, Kovacs et al., 2003). Moreover, in 2009, Wang and colleagues revealed the

chromatin targeting of HDAC6 to active genes via its interaction with RNA Pol II (Wang, Zang et al., 2009).

4.3. Methylation

Methylation consists of a reversible reaction of transferring one methyl (CH3) group from S-adenosylmethionine to the amine (NH2) group of basic histone residues such as lysine and arginine. Since methylation of histone does not affect the electrostatic charge of amino acids, the effect of this modification on dynamic of chromatin is less drastic than acetylation. Histone methylation generally plays crucial roles in gene expression through chromatin compaction or the recruitment of active or repressive regulatory complexes to specific chromatin regions. In general, lysines can be mono (me1), di (me2) or trimethylated (me3). It has been reported that arginines are mono- or di-methylated on their guanidinyl group. Histone H3 and histone H4 methylation are among the most studied histone methylations (Di Lorenzo & Bedford, 2011, Greer & Shi, 2012, Martin & Zhang, 2005).

The biological consequences of lysine methylation are different, depending on which lysine residue is methylated and the number of methylations within the lysine residue. For instance, H3K4, H3K36 and H3K79 are associated to transcriptionally active sites or euchromatin regions, whereas H3K9, H3K27 and H4K20 are related to constitutive or facultative heterochromatin or transcriptionally inactive chromatin (Nielsen, Oulad-Abdelghani et al., 2001). More specifically, H3K9 and H3K27 methylation is associated respectively to constitutive and facultative heterochromatin formation, mostly by engagement of proteins involved in heterochromatin formation and gene silencing.

Furthermore, it has been shown that the nuclear hormone receptor co-activator CARM1 can methylate arginine residues of histone H3. In addition, CARM1-mediated methylation at arginine 17 of H3 is associated with activation of estrogen receptor-regulated PS2 gene (Bauer, Daujat et al., 2002).

4.3.1. Histone methyltransferases

Histone methylation is mediated by enzymes called HKMTs. Specific categories of HMTs are known, such as KMTs, which are responsible for lysine methylation and PRMTs, which methylate arginine residues of histones.

H3K9 HKMTs are among the most studied methyltransferases. This family contains four principal enzymes: Suv39h1/h2, SETDB1/ESET, G9a and GLP.

Suv39 was first identified in *Drosophila* as the suppressor of position effect variegation *SU (VAR) 39* and its mammalian homologs are Suv39h1 and Suv39h2 (Rea, Eisenhaber et al., 2000). This protein contains two functionally distinct domains, the N-terminal CD and the C-terminal SET domain. Most of HMTs contain SET catalytic domain at their C-terminal protein part, with the exception of DOT-1 like proteins. Suv39h1 prefers mono or di-methylated H3K9 as the substrate and catalyze the tri methylation of H3K9, which is essentially involved to the establishment of pericentric heterochromatin (Peters, Kubicek et al., 2003).

SETDB1 is important for early development and di-methylates H3K9 in the absence of its cofactor hAM. Meanwhile, in the presence of its cofactor it catalyzes the tri-methylation of H3K9 (Dodge, Kang et al., 2004).

On the other hand, Ezh2, which is the catalytic subunit of PRC1 catalyzes the tri-methylation of H3K27, which is involved in X-chromosome inactivation (Plath, Fang et al., 2003, Trojer & Reinberg, 2007).

4.3.2. Histone demethylases

The first histone demethylase identified LSD1, revealed specific activity towards methylated K4 of histone H3 and its function is evolutionary conserved from *Schizosacharomyces pombe* to mammals. LSD1 belongs to Co-REST complex and contributes to its transcriptional repression activity (Shi, Lan et al., 2004).

The family of JmjC domain-containing proteins contain a lysine demethylase activity. JHDM1, JHDM2A, JHDM3A, JMJD2A and JMJD3 are different enzymes of this family.

Depending on demethylation of different histone lysine residues, these different members are involved either in transcriptional activation or repression.

For instance, JHDM1 has been shown in vivo to reduce the di-methylation of H3 K36, which is conserved from yeast to human while JMJD3 catalyses the removal of tri-methylation on H3 lysine 27 (Chen, Zang et al., 2006, Tsukada, Fang et al., 2006).

Among the most studied arginine demethylases we can mention PADI4 enzyme, which transforms methylated arginines to citrulines, an event which is associated to transcriptional repression (Christophorou, Castelo-Branco et al., 2014).

4.4. Other histone Lysine acylations

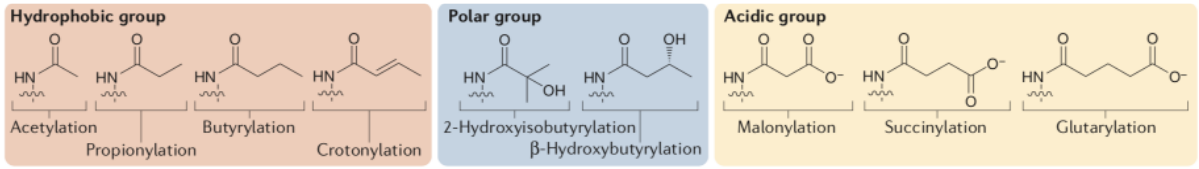
The majority of recently identified histone PTMs are functional non-acetyl small chain lysine acylations (mentioned above), that has been demonstrated to play active roles in gene expression and chromatin structure, through different mechanisms. These modifications differ in hydrocarbon chain length, charge and hydrophobicity (Sabari et al., 2017). Acylation is divided into three groups: hydrophobic group, polar group and acidic group (Figure 12A). Hydrophobic group is composed of acetylation, propionylation, butyrylation and crotonylation modifications. These modifications increase the hydrophobicity via their longer hydrocarbon chain. The polar group encompasses β-hydroxybutyryl and 2-hydroxyisobutyryl modifications, which promote histone Lysine interactions with other proteins via hydrogen bonds. The acidic group contains malonyl, succinyl and glutaryl modifications, which change the charge at Lysine residues from +1 to -1.

As discussed previously, described HATs, such as CBP/p300, GNAT and Myst members are the most well-known acyltransferases identified. There are data suggesting that these HATs could also use various Acyl-CoAs (Figure 12B). Therefore, different acyl-CoAs could compete with each other for binding to different acyltransferases. This could constitute a mechanism for differential protein acylations. Therefore, the more abundant acyl-CoA wins the competition and confers its effects. Short chain fatty acids are the major source of short chain acyl-CoA production in cells. Acetate via the enzymatic activity of the ACSS2, may produce Ac-CoA, which is used by acyltransferases in order to induce histone lysine acylations. However, there are no evidences for a direct synthesis of acyl-CoAs by ACSS2. The only

available data is based on ACSS2 knock-down which results in a decrease in histone crotonylation (Sabari, Tang et al., 2015).

Lysine acylations are generally known as the hallmarks of active genes. These PTMs neutralize the positive charge of lysines and could induce structural changes to the chromatin, which lead to the gene expression and probably the recruitment of other proteins. The proportional differences of acyl-CoAs can lead to functional changes and may have consequences in terms of chromatin regulations (Goudarzi, Zhang et al., 2016, Tan et al., 2011, Xie, Zhang et al., 2016).

A



B

Enzyme class	Lysine acylations							
	Kpr	Kbu	Kcr	Kma	Ksucc	Kglu	Khib	Kbhb
Writers								
p300/CBP	+	+	+	NA	+	+	NA	+
GNATs	+	–/+	–	NA	NA	NA	NA	NA
MYSTs	+	–	–	NA	NA	NA	NA	NA
Erasers								
NAD^+-dependent sirtuins	SIRT1–3	SIRT1–3	SIRT1–3	SIRT5	SIRT3–5,7	SIRT5	NA	NA
Zn^{2+}-dependent HDACs	NA	NA	HDAC3	NA	NA	NA	NA	NA

Figure 12: Different groups of lysine acylations and various HATs and their potential binding to different lysine acylations. A) Representation of 3 groups of short-chain acylations on the amine group of lysines, regarding to their chemical property. **B)** Different types of HATs/HDACs and their different combinations with various lysine acylations. Taken from (Sabari et al., 2017).

5. Histone variants

Canonical histones have non-allelic isoforms with variable sequence homologies referred to as “histone variants”. Most of the canonical histone genes are expressed during S phase and are assembled into nucleosomes behind the replication fork (Replication-dependent). In contrast, most of the histone variants are incorporated into the chromatin throughout the cell cycle therefore they are called Replication-independent histones. The mRNAs encoding canonical histones lack introns and, they form a stem-loop sequence at their 3’ end, which is crucial for their processing and results in a regulated stability of the corresponding mRNAs. In contrast, some genes encoding histone variants can harbor one or several introns and the corresponding mRNAs could be polyadenylated at their 3’ end.

During evolution, histone variants have gained different functions despite the high sequence conservation among these proteins. Histone variants can induce structural changes within the octamer core particle and alter DNA wrapping around histones, causing structural transformation of the nucleosomes and changing their dynamics. In general, histone variants are implicated in numerous cellular functions such as transcriptional regulation, DNA repair, chromosome segregation, facilitated process of DNA re-packaging such as histone disassembly and replacement by protamines in the context of germ cells (Hoghoughi, Barral et al., 2018, Marzluff, Gongidi et al., 2002, Talbert & Henikoff, 2017). In mammals, H2A, H2B, H3 canonical histones and linker histone H1 possess a number of variants (Figure 13). However, no variant for H4 has been identified in higher eukaryotes in contrast to the lower eukaryotes such as *trypanosome brucei*, where it is probably enriched at RNA PolII transcription termination sites (Siegel, Hekstra et al., 2009).

In contrast to the homogenous distribution of canonical histones, histone variants are distributed in a locus-specific and heterogenous fashion throughout the chromatin. They are also implicated in the genome reorganization during early developmental stages, ES cells and lineage commitment, somatic cell reprogramming, cancer (Buschbeck & Hake, 2017) and the epigenetic reprogramming during later stages of spermatogenesis, which is discussed in detail in the next section.

In recent years the role of histone variants in modulating the epigenetic status of the chromatin has attracted attention. Histone variants could confer stable key characteristics to

the chromatin and therefore change the epigenome plasticity. This specific particularity of histone variants could facilitate the emergence of cancer cells (Wang, Chuffart et al., 2019).

Various histone variants of H2A, H2B, H3, H1 linker histone, and their implication in cellular mechanisms such as transcription and specifically in chromatin function are further highlighted in this section.

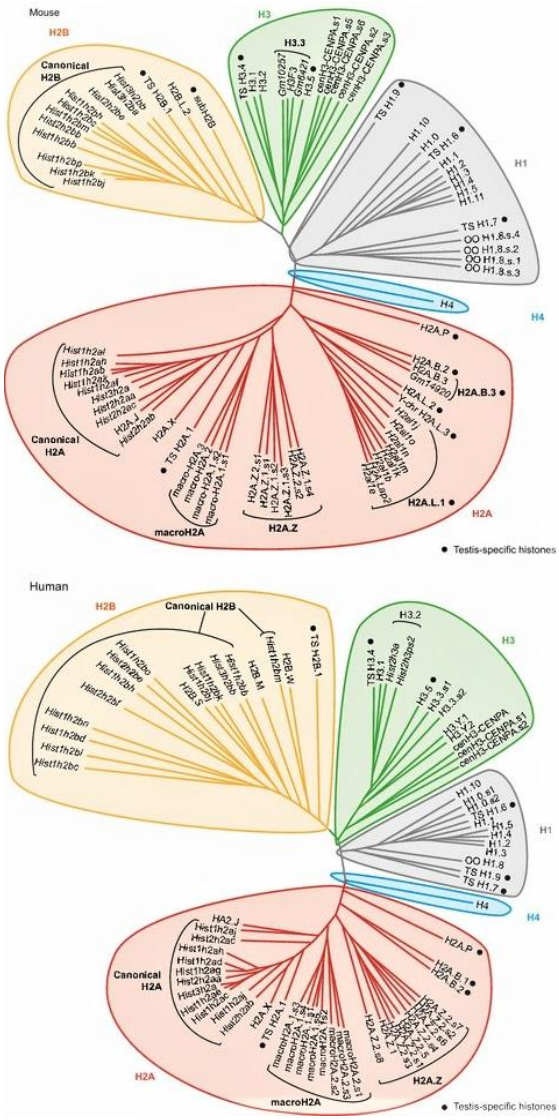


Figure 13. The phylogenic tree of different histone variants in mouse and human. Taken from (El Kennani, Adrait et al., 2017).

5.1. Variants of linker histone H1

In mouse and human, around 15 isoforms of linker histone H1 has been identified (Figure 13). The N- and C- terminal part of H1 histones present more diversity in term of sequence and length among subtypes, whereas the central globular domains are more conserved among isoforms and during evolution. Histones H1 variants are remarkably expressed during specific timing of embryonic development and cell differentiation. One of the well-characterized linker histone H1 is CS H1 in (sea urchin), B4 in Xenopus and H100 in mouse. For instance, the cleavage stage linker histone B4 is expressed during oogenesis of *Xenopus Laevis* and constitutes the majority of linker histone during early development. Other well-characterized linker histone H1s, are differentiated-dependent linker H1(H1(0) and H5). H1(0) is largely expressed in vertebrates, whereas H5 is mostly expressed in bird's erythrocytes. Testis-specific linker histone H1s are expressed in vertebrates and invertebrates in different spermatogenic cell types. The functional role of each of these variants still remains obscure (Khochbin, 2001, Khochbin & Wolffe, 1994).

5.2. Variants of histone H2A

In higher eukaryotes, eight variants of histone H2A has been identified up today: H2A.X, H2A.Z.1, H2A.Z.2.1, H2A.Z.2.2, H2A.Bbd/H2A.B, macroH2A1.1, macroH2A.1.2 and macroH2A2 are among the most studied in terms of structure and function (Buschbeck & Hake, 2017).

In this section, some of the most prominent characteristics of the variants of H2A, H2B and H3 are described.

5.2.1. H2A.Z: Controlling transcriptional regulation

The 2.6 Å crystal structure of the nucleosome containing histone variant H2A.Z, unraveled the structural alterations of H2A.Z nucleosomes compared to canonical H2A-containing nucleosomes. These alterations involve modified interactions between H2A.Z/H2B dimer and H3-H4 tetramer, as well as the interaction between the two H2A.Z/H2B dimers

themselves. Despite these differences, the H2A.Z/H2B–DNA interactions, which are mediated through H2A.Z L1 and H2B L2 loops, do not seem to change, compared to H2A nucleosomes (Suto, Clarkson et al., 2000).

The implication of H2A.Z histone variant in gene regulation, transcriptional control and the regulation of centromeric heterochromatin have been highlighted in many studies. Recent studies in different organisms discuss the differential abundance of H2A.Z throughout the enhancers and promoters. For instance, in yeast H2A.Z is present on transcriptional start sites of active and inactive genes, while in *Drosophila* as well as in plants, its presence at promoters is correlated with a reduced gene expression. Meanwhile, RNA POL II- coupled remodeler, FACT depletes H2A.Z from coding regions, which is correlated with increased local transcription (Hu, Shen et al., 2011, Jeronimo, Watanabe et al., 2015, Kusch, Mei et al., 2014).

About 15 years ago, it has been demonstrated that components of Tip60 and SRCAP complexes combine with sub-unit of SWR1 and NuA4 complexes and promote the link between acetylation and H2A.Z deposition on the chromatin (Raisner, Hartley et al., 2005).

Recent studies revealed that the deposition of H2A.Z via p400/Tip60 complex occurs upon the induction of androgen receptor expression. Active transcriptional states of genes are also correlated with active PTM marks on H2A.Z such as H2A.Zac. Additionally, Ubiquitination of H2A.Z is associated with genome transcriptional repression and its de-ubiquitination is associated with genome transcriptional activation (Giaimo, Ferrante et al., 2019, Ito, Kayukawa et al., 2018).

5.2.2. H2A.X and DNA damage repair

Cells are regularly exposed to various types of stresses such as oxidative stress, IR, apoptosis-inducing signaling and meiotic homologous recombination, which cause DNA double-strand break and jeopardize genome integrity. H2A.X is conserved in all eukaryotes and contain 20 amino acids more than conventional H2A at its C-terminal end, which harbors the conserved motif of SQE (Li, Eirin-Lopez et al., 2005, Talbert & Henikoff, 2010). H2A.X phosphorylation of serine 139 (known as γ-H2A.X) has been for long time referred to as a marker of DNA damage. H2A.X-deficient mouse embryonic cells demonstrate high levels of genomic instability. In addition, spermatogenesis is defective in H2A.X-deficient mice. Three

types of kinases are responsible for the phosphorylation of H2A.X; ATM, DNA-PK and ATR (Xiao, Li et al., 2009).

The precise molecular mechanisms demonstrating the role of H2A.X in response to DNA damage has not yet completely understood. Recently, histone chaperon FACT has been identified as being responsible for the deposition of *de novo* H2A.X at DNA repair sites, at the same time as DNA repair machinery activation. Concomitantly to the action of FACT, histone chaperone ANP32E removes H2A.Z from DNA damage sites inside the chromatin (Figure 14) (Piquet, Le Parc et al., 2018).

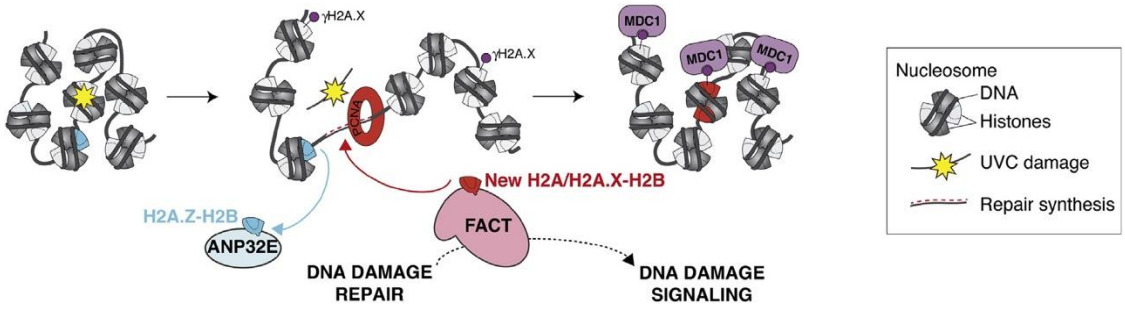


Figure 14: Coordinated action of FACT and ANP32E for chromatin reshaping after DNA damage induction. Taken from (Piquet et al., 2018).

In addition to DNA damage repair, recently H2A.X is considered as a regulator of epithelial-mesenchymal transition in human colon adenocarcinoma cells via inducing the activation of EMT transcription factors (Slug & ZEB1). Recently, by using super light-resolution microscopy, the structural basic units read by the DNA repair machinery known as phosphor-H2A nano foci, have been identified. These regions measure 200nm of lateral diameter and harbor 75Kb of chromatin domains (Natale, Rapp et al., 2017, Weyemi, Redon et al., 2016).

5.2.3. MacroH2A and transcriptional repression

H2A histone variant macroH2A was first identified in rat liver nucleosomes. It harbors a large C-terminal (20KD) of unknown function (Chadwick & Willard, 2001, Pehrson & Fried, 1992). Its N-terminal has 60% identity to the conventional H2A sequence and the rest of the protein shares no sequence similarity to conventional H2A. MacroH2A contains two allelic isoforms named mH2A1 and mH2A2, which share 80% sequence homology. The RNA of mH2A1gene undergoes a differential alternative splicing, which leads to the production of mH2A1.1 and mH2A1.2, which are expressed in certain tissues (Chadwick & Willard, 2001).

MacroH2A is enriched on the inactive X-chromosome of female mammals and is associated with the transcriptional repression of the corresponding X chromosomes in female mammals known as dosage compensation. Dosage compensation allows the comparable level of X-linked gene expression between male and female. The inactive X chromosome is densely stained and localized at the periphery of the nucleus and is referred to as the Bar body. Furthermore, macroH2A is highly associated with heterochromatin foci in senescent or quiescent human cells driven by HIRA and ASF1a chromatin regulators (Zhang, Poustovoitov et al., 2005).

Apart from its association with X chromosome inactivation, macroH2A is known as a transcriptional repressor. One of the mechanisms through which macroH2A represses transcription is through the inactivation of p300 HAT, which blocks post-translational modifications of histones which are essential for transcription initiation. These observations have been confirmed *in vivo* since the nucleosomes containing the mH2A are less acetylated than the ones with conventional H2A (Chakravarthy, Gundimella et al., 2005). Moreover, mH2A blocks the action of chromatin remodeling complexes, which are necessary for nucleosomal mobility during transcription. The L1 loop of mH2A, localized between the two first α helices of mH2A, is responsible for this blockage (Doyen, An et al., 2006).

5.2.4. H2A.Bbd and transcriptional active chromatin domains

Mammalian-specific H2A variant H2A.Bbd has been first identified following its exclusion from the Bar body foci inside human cells nucleus. This variant is highly expressed in adult testis, to a lower extent in brain tissues and in embryonic stem cells. H2A.Bbd is slightly shorter than H2A and shares only 48% homology with conventional H2A. The most sequence variations exist in the docking domain of H2A.Bbd, which is responsible for $(H2A\text{-}H2B)_2$ interactions with $(H3\text{-}H4)_2$ within the core nucleosome (Bao, Konesky et al., 2004). Acid extraction of histones from H2A.Bbd transfected cells following separation on sucrose gradient and ultra-centrifugation, revealed the incorporation of H2A.Bbd into the nucleosome. Immunofluorescent analysis demonstrated the co-localization of H2A.Bbd with acetylated H4 (Eirin-Lopez, Ishibashi et al., 2008). The fact that H2A.Bbd is excluded from Bar body along with its colocalization with acetylated H4 revealed its association with transcriptionally active chromatin regions (Chadwick & Willard, 2001, Tolstorukov, Goldman et al., 2012).

Nucleosome reconstitution studies revealed H2A.Bbd ability to induce more unstable nucleosomes and these nucleosomes harbor 118bp of DNA wrap around each nucleosome (Bao et al., 2004, Gautier, Abbott et al., 2004). *In vivo* studies revealed the presence of H2A.Bbd at DNA replication and repair sites, whereas *in vitro* nucleosome reconstitution revealed the formation of organized hexasomes with shorter wrap of DNA. These studies suggested the formation of a transient H2A.Bbd-dependant intermediate organized chromatin structure after DNA replication or DNA repair (Arimura, Kimura et al., 2013).

Later studies involving genome-wide mapping of nucleosome bearing H2A.Bbd, revealed their enrichment at transcriptionally active genes. Moreover, depletion of H2A.Bbd reveals its role in mRNA splicing and highlights this histone variant's role in transcription and mRNA processing (Tolstorukov et al., 2012). Recent investigations on testis-specific mouse orthologue of variant H2A.Bbd referred to as H2A.Lap1 also revealed this variant incorporation into the nucleosome, its localization to the gene body of active genes and its association with RNA processing factors such as spliceosome.

5.3. Variants of histone H3

Five distinct isoforms of histone H3 variants are present in mammals; H3.1, H3.2, H3.3, H3t and CENP-A. H3.2 and H3.3 differ only in four amino acids and H3t is only expressed in testis. CENP-A is mostly known as a component of centromeres and it plays essential roles in maintaining genome integrity. These more or less minimal differences between each H3 variants represent significant functional consequences at chromatin level. Histone variant H3t will be further discussed in the next chapter and the rest are discussed here in the following section.

5.3.1. CENP-A: centromeric nucleosome

Centromeres are particular chromatin domains present throughout the cell cycle and provide a platform for the transient assembly of kinetochore complex during mitosis. Kinetochores themselves ensure the attachment of spindle microtubules and proper chromosome segregation (Foltz, Jansen et al., 2009). In human, centromeres harbour up to 5.10^6 bp of centric DNA and they are composed of DNA repetitive sequences referred to as α-satellites, minor satellites in mice. These repeats differ in sequence and length among different species (Muller & Almouzni, 2017).

Mammalian centromeres harbor a centromere-specific histone H3 variant known as *Centromeric-Protein A* (CENP-A). Cenp-A knock out induces embryonic lethality in mice, which indicates its crucial role in centromere functions. Crystal structure of CENP-A containing nucleosomes revealed that DNA wraps only 121 bp around the histone octamer of two H2A, two H2B, two CENP-A and two H4 (Tachiwana, Kagawa et al., 2011). During G1 phase, this centromeric variant is incorporated into the nucleosome through histone chaperone HJURP via its affinity for CATD domain in CENP-A sequence (Foltz et al., 2009). Recent studies using crystal structure, electron cryo-microscopy and by replacing αN-helix of H3 into CENP-A, revealed CENP-A DNA ends are more flexible than histone H3. The chimeric CENP-A-αN-helix H3 protein creates less flexible nucleosome which result in the recruitment of H1, whereas CENP-A nucleosome provides less stable nucleosome, which excludes histone H1s and is crucial for the fidelity of mitosis. In addition, it has been shown that CENP-A is

among the histone variants which remain in the mature spermatozoa which highlights its crucial role probably for the establishment of the functional male chromatin in post-fertilization stage (Roulland, Ouararhni et al., 2016). The expression of Cenp-A is ubiquitous and its presence is essential for the recruitment of centromeric proteins and kinetochores to ensure genome integrity and the proper functioning of mitosis. Old Cenp-A variants are necessary to direct new Cenp-A deposition (McKinley & Cheeseman, 2016). Initially, electron microscopy analyses and an antibody recognizing cenp-A, allowed the detection of cenp-A during spermatogenesis (Biggiogera, Muller et al., 1992). Like somatic cells, during spermatogenesis, Cenp-A ensures the correct kinetochore assembly and chromosome segregation in mitosis as well as in meiosis. Evidences showed that, Cenp-A is quantitatively retained in mature bovine spermatozoa nucleus in specific foci (Palmer, O'Day et al., 1990).

5.3.2. H3.1, H3.2 and H3.3

As mentioned above, in mammals, in addition to CENP-A, 3 different isoforms of H3 are incorporated into nucleosome throughout the cell cycle. H3.1 and H3.2 are different only in one amino acid at the position of 96 in the protein sequence and are expressed during the S phase. In contrast H3.3 is expressed in S, G1 and G2 phases and it has a constant expression throughout the differentiation (Ahmad & Henikoff, 2002). In terms of protein sequence, H3.3 differs in 4 amino acids compared to H3.2 and 5 amino acids with H3.1. Two genes *H3.3A* and *H3.3B* are responsible for H3.3 expressions. H3.1/2 are deposited into the nucleosome through CAF-1 histone chaperone, whereas H3.3 is incorporated into the nucleosome of transcriptionally active regions, through histone chaperone HIRA (Tagami, Ray-Gallet et al., 2004). Recently in mESC, one mutation at serine 31 of H3.3, enhances the p300 activity at enhancer regions, suggesting a role for nucleosomes as novel cofactors of p300 (Martire, Gogate et al., 2019). Furthermore, DAXX-ATRX complex has been identified responsible for the deposition of H3.3 into PHC (Figure 15) (Muller & Almouzni, 2017). Interestingly, recent studies reveal the cooperation between H3.3 and H3K9me3 within the same nucleosome, which highlights its role in the maintenance of heterochromatin at repeat regions (Elsasser, Noh et al., 2015).

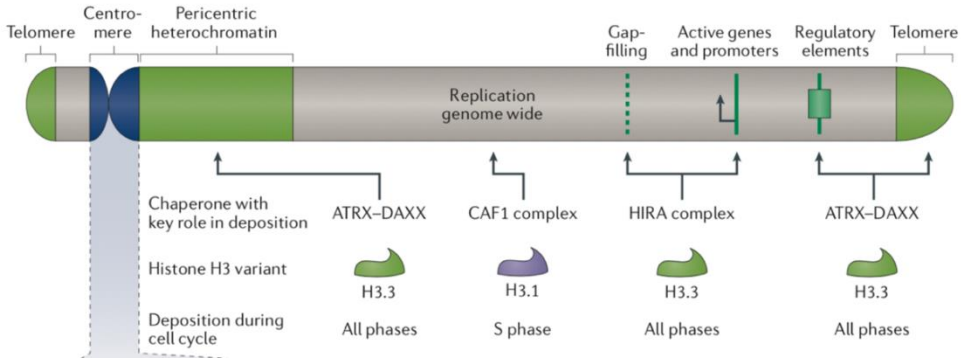


Figure 15: Different H3 histone variants with their associated chaperones. Taken from (Muller & Almouzni, 2017).

III. Dynamic of nucleosomal assembly and histone turnover

Nucleosome assembly is a crucial process for maintaining the epigenetic stability of chromatin following cellular processes such as DNA replication, DNA repair and gene transcription. During DNA replication and transcription, nucleosomes are removed, displaced and modified through proteins complexes called histone chaperones. Since histones are very basic proteins, they are accompanied by chaperones in order to prevent interactions with other cellular components, as well as protein aggregation, and to facilitate their assembly. Histone chaperone complexes assemble nucleosomes in either replication-dependent or independent manner (Burgess & Zhang, 2013). In the following section more details on the role of most studied histone chaperones, in *de novo* histone deposition during DNA replication, DNA damage repair, active gene transcription and heterochromatin and centromere maintenance as well as their role as chromatin remodelers are further discussed.

1. Histone chaperones

1.1. CAF-1 in DNA replication and DNA repair

In human, protein complex CAF-1 contains 3 subunits: p48, p60 and p150. The largest subunit of CAF-1 protein complex, p150, mediates the replication- dependent H3.1/H3.2-H4 nucleosome assembly (Hammond, Stromme et al., 2017). P150 subunit brings newly synthesized H3 and H4 inside the nucleus. The crystal structure of yeast CAF-1 subunit Cac1C, reveals the formation of a WHD at its C terminus. Structural evidences are in favor of the fact that this domain consolidates CAF-1 binding to PCNA at DNA replication site (Zhang, Gao et al., 2016). The link between CAF-1 and PCNA reveals that these factors cooperate with each other in order to connect DNA replication with nucleosome assembly and maintain the epigenetic inheritance during chromatin replication (Figure 16a) (Shibahara & Stillman, 1999, Zhang, Shibahara et al., 2000). Other investigations aiming to understand the mechanisms by which heterochromatin organization reshapes after DNA replication, revealed the association of p150 subunit of CAF-1 to HP-1. This study demonstrated that CAF-1 interaction with HP-1 enhances the targeting of HP-1 to the pericentric heterochromatin regions during replication (Quivy, Roche et al., 2004).

In vivo studies revealed the recruitment of the p150 subunit of CAF-1 in UV damaged sites. In human, at sites of DNA damage, CAF-1 is responsible for de novo deposition of H3.1. However, CAF-1 nucleosome assembly of H3.1 has been observed after DBS, independently of PCNA action (Figure 16b) (Brachet, Beneut et al., 2015, Li & Tyler, 2016).

1.2. Role of HIRA in histone deposition in active genes

Protein purification of non-nucleosomal H3.1 and H3.3 containing complex, revealed the specificity of HIRA for H3.3 (Tagami et al., 2004). Histone chaperone HIRA is responsible for replication-independent deposition of H3.3-H4 in gene bodies, telomere and regulatory elements. CHIP-seq analysis revealed the highest rate of H3.3 at active gene promoters and enhancers (Figure 16c) (Hammond et al., 2017). Histone chaperon complex containing HIRA contains also two other components UBN1 and CABIN (Loyola & Almouzni, 2004). Moreover, structural and mutational studies found UBN1 specificity for H3.3 through the

conserved Hpc-2 related domain of UBN1 (Ricketts, Frederick et al., 2015). Furthermore, HIRA accumulates at DNA damaged sites and is responsible for H3.3 deposition prior to DNA repair and plays crucial role in transcription recovery and reactivation subsequent to DNA repair (Adam, Polo et al., 2013). In addition, HIRA plays essential roles in other cellular events such as nucleosome assembly post-fertilization, in male pronucleus. In *Drosophila*, a point mutation in *Hira* gene influences the deposition of H3.3 into the chromatin of the male pronucleus after fertilization (Loppin, Bonnefoy et al., 2005).

1.3. Asf1 H3-H4 chaperone

In cytoplasm, ASF1, a conserved eukaryotic histone chaperon, binds H3.1/H3.2-H4 and H3.3-H4 and transfers them inside the nucleus before passing them over to CAF-1 or HIRA, respectively (Green, Antczak et al., 2005, Groth, Ray-Gallet et al., 2005, Mello, Sillje et al., 2002, Tang, Poustovoitov et al., 2006, Tyler, Adams et al., 1999). ASF1a and ASF1b are the two paralogues of ASF1 protein in vertebrates. ASF1a has a preference to interact with HIRA and ASF1b interacts with CAF-1. ASF1a is ubiquitously expressed in every tissue and throughout the cell cycle but the ASF1b is expressed in S phase and in tissues with high rate of proliferation (Abascal, Corpet et al., 2013).

1.4. Role of DAXX in H3.3-H4 deposition at heterochromatin regions

The *in vivo* deposition of H3.3-H4 dimer in heterochromatin loci such as pericentromeres and telomeres, is mediated by histone chaperone DAXX (Figure 16d). This chaperone, in cooperation with chromatin remodeling factor ATRX, ensures the replication-independent of H3.3-H4 dimer assembly within the nucleosome. ATRX via its ADD (ATRX-DNMT3-DNMT3L) recognizes the H3K9me3, which directs DAXX-ATRX chaperone to the heterochromatin sites (Dhayalan, Tamas et al., 2011, Eustermann, Yang et al., 2011, Iwase, Xiang et al., 2011). The crystal structure of histone-binding domain of DAXX, demonstrated that this domain wraps around the H3.3-H4 dimer in order to ensure the H3.3 specific recognition. In addition, functional studies revealed the participation of Gly 90 amino acid in

H3.3 and Glu 225 in DAXX in the specific recognition between H3.3 and DAXX (Elsasser, Huang et al., 2012).

1.5. HJURP Cenp-A-H4 deposition at centromeric regions

Centromere maintenance throughout the cell cycle requires the Cenp-A deposition via histone chaperone HJURP. HJURP has been identified as the histone chaperone responsible for the deposition of Cenp-A-H4 dimer to the centromere nucleosomes. Identification of the interacting partners with pre-deposited Cenp-A, revealed the specificity of HJURP for Cenp-A centromeric histone variant (Dunleavy, Roche et al., 2009). In human cells, newly-synthesized Cenp-A is incorporated into the chromatin in a replication-uncoupled manner during early G1 phase. HJURP homologue protein with the same function in fission yeast is known as Scm3 and in *Drosophila* is called CAL1 (Chen, Dechassa et al., 2014, Williams, Hayashi et al., 2009). HJURP/Scm3 recruitment to the centromeres is mediated through the interaction with protein complex Mis18. Mis18 complex in vertebrates contains three paralogs: Mis18α and Mis18β and Mis18BP1 (Fujita, Hayashi et al., 2007, Maddox, Hyndman et al., 2007). HJURP via its centromere targeting domain, interacts with C-terminal coiled-coil of Mis18α-β (Nardi, Zasadzinska et al., 2016, Wang, Liu et al., 2014). The proper functioning and recruitment of HJURP and Mis18 is mediated through the phosphorylation of Cdk-1and Plk-1kinases on the consensus T40 and S110 within the Mis18BP1 residues 20-130 (Figure 16e) (Pan, Klare et al., 2017, Spiller, Medina-Pritchard et al., 2017, Stankovic, Guo et al., 2017).

1.6. NAP-1 histone chaperone family

Nap-1 histone chaperon family, are capable to reconstitute nucleosome from free histones and DNA in vitro. In mammals, five genes (NAP1L1-NAP1L5) have been identified. Nap-1 is the most studied members of the NAP histone chaperone family and it is highly conserved among eukaryotes. The 6.7 Å crystal structure of Nap1 in complex with H2A-H2B, revealed the interaction of NAP1 through its C-terminal acidic surface with H2A from the

dimer H2A-H2B (Figure 16c) (Aguilar-Gurrieri, Larabi et al., 2016). However, Nap-1 and Vps75, a member of the Nap-1 family, can interact with histone H3 and H4 in a tetrameric conformation (Bowman, Ward et al., 2011). Throughout the spermatogenesis, histone chaperone NAP1L4 interacts with testis-specific histone variant H3t (Tachiwana, Osakabe et al., 2008). NAP1L4 is also recognized as the histone chaperone of the H2A testis-specific histone variant called H2A.L.2 (Barral, Morozumi et al., 2017).

1.7. FACT: The facilitator of transcription

Histone chaperon complex FACT was first identified through a promoting factor, which facilitates the action of RNA Polymerase II, in vitro (Belotserkovskaya, Oh et al., 2003). Upon transcription, FACT interacts with nucleosomal H2A-H2B dimer and facilitate DNA unwinding, thus RNA Pol II can move forward easily and continues the transcription (Figure 16c) (Hsieh, Kulaeva et al., 2013). FACT is a heterodimeric protein, composed of two subunits Spt16 and SSRP and it is expressed mostly in differentiated and stem cells. Loss of FACT in *Drosophila* is associated with higher occurrence of transcription-associated histone marks such as H3K4me3 and H3K36me3. In addition, recent evidences suggest that FACT play crucial role in keeping RNA Pol II at promoter-proximal pausing, a mechanism which leads to gene activation (Tettey, Gao et al., 2019).

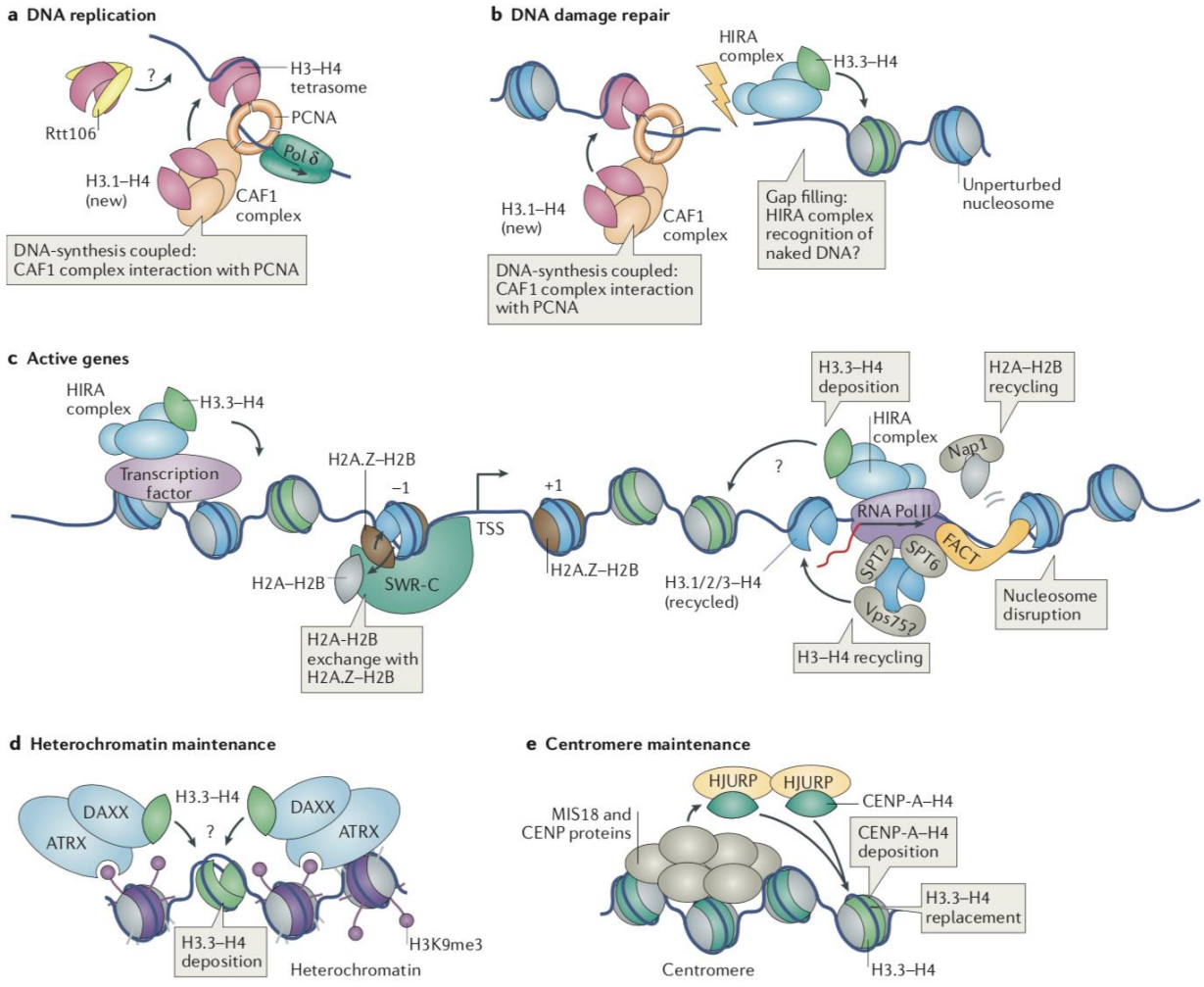


Figure 16: Histone-specific chaperones and their mechanism of function. Taken from (Hammond et al., 2017).

2. ATP-dependent chromatin remodeling complexes

Chromatin remodeling is indispensable for numerous fundamental chromatin processes such as DNA replication, transcription and DNA repair. These factors control the dynamic of chromatin and mediate nucleosome sliding, defining nucleosome position on the chromatin and nucleosome assembly and disassembly. For instance, upon transcription, the remodeling factors act on nucleosomes present on regulatory regions of chromatin such as promoters and enhancers and promote transcription factor binding to these sites. Chromatin remodeling involves ATP-dependent remodeling complexes, which are divided into four family of

enzymes: ISWI, CHD, INO80, and SWI/SNF (Clapier, Iwasa et al., 2017). All of these complexes contain one ATPase domain, which has the mission to dissociate DNA-nucleosome interaction and translocate DNA along the histone surface. The crucial role of each of these remodeling factors is briefly discussed in this section.

2.1. ISWI and CHD in nucleosome assembly

In addition to histone assembly behind the replisome machinery, ISWI and CHD remodeling subfamily enzymes promote the histone assembly into an octamer form and then accommodate octameric nucleosomes within the chromatin with specific defined distances from each other (Figure 17A) (Fei, Torigoe et al., 2015, Varga-Weisz, Wilm et al., 1997). These activities are conducted mainly by the ATPase subunit within each remodeler. The ATPase subunit of ISWI and CHD are very similar whereas the main differences between these two are in their C and N-terminal part. ISWI contains HSS domain at its C-terminal (Grune, Brzeski et al., 2003), while CHD contains a NegC domain and a DBD containing only a SANT and SLIDE domain (Hauk, McKnight et al., 2010). HSS domain in ISWI binds to H3 C- and N-terminal histone tails and linker DNA at both sides of the nucleosome (Boyer, Latek et al., 2004, Dang & Bartholomew, 2007). ATPase subunit in ISWI is flanked with AutoN and NegC domains, which are negative regulators of ATPase domain and DNA translocation. CHD contains two chromodomains at its N-terminal, which recognize histone methylation sites (Kunert & Brehm, 2009).

2.2. SWI/SNF in chromatin access

The accessibility of transcription factors and other major chromatin regulators to key regulatory chromatin regions such as promoters and enhancers is vital for the proper functioning of chromatin-associated mechanisms. SWI/SNF remodeling complexes render chromatin more accessible through mechanisms such as nucleosome sliding and complete or partial nucleosome eviction from the chromatin (Figure 17B) (Clapier et al., 2017). The ATPase domain of SWI/SNF contains two lobes. At the N-terminal of this protein there is a HAS domain and an adjacent post-HAS domain. One AT-hook and a bromodomain exist at

the C-terminal part of SWI/SNF. All of these domains flanking the ATPase region, are believed to control the activity of the ATPase subunit.

2.3. INO80 in Nucleosome editing

In mammals, INO80 (Swr1 in yeast) subfamily of remodelers are mostly implicated in nucleosome editing processes. They can take out one or more histones from the nucleosome and replace it with either a canonical or a histone variant. Similar to other remodeling factors, the ATPase domain of this protein is also composed of two lobes but with a larger insertion part between two lobes. At the N-terminal of this protein, there is a HAS domain that is responsible for nucleation of actin and actin-related proteins. INO80 subfamily proteins contain SWR1C, p400 and Snf2-related CBP activator, which replaces H2A-H2B dimers with H2A.Z histone variant in nucleosomes containing H2A.Z-H2B dimers (Clapier et al., 2017) (Figure 17C).

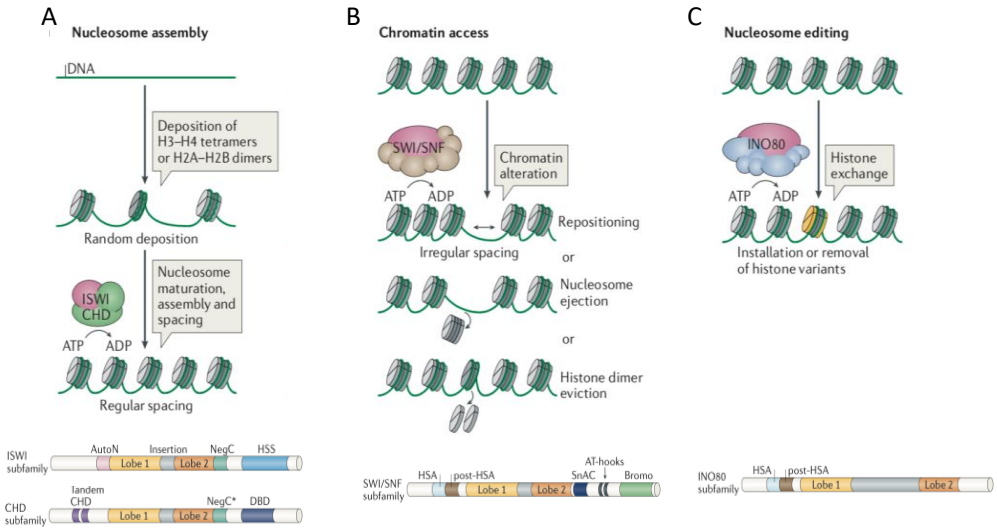


Figure 17: Histone chaperones in chromatin assembly, access and editing. Taken from (Clapier et al., 2017).

IV. Mammalian Spermatogenesis

1. General overview of spermatogenesis

Spermatogenesis is a multi-step process which leads to the generation of male gametes referred to as mature spermatozoa. Mature spermatozoa bear the mission to leave the organism and deliver the male genome to the oocyte. Upon the fertilization, a new developmental programming involving both the male and female genome takes place, which ensure the correct development of zygote. In order to fulfil its mission and for a safe journey aiming at delivering male genome inside the oocyte, the genome of mature sperm needs to be highly compacted and protected from the exterior harsh environment. In human, spermatogenesis is separated into two phases: the first phase begins during embryonic period and the second phase, is initiated in puberty and lasts throughout the male adulthood life. Between the first and the second phase, spermatogonial stem cells go through a period of testicular quiescence during which they experience a developmental arrest. In contrast in mice, spermatogonial stem cells continue their development without passing through a quiescent phase.

In rodents, the embryonic stage leads to the development of SSCs and the development of gonocytes. Around E6.25 day in the proximal epiblast, embryonic cells under the control of BMP4 and BMP8b adopt the germ cell fate by formation of PGC precursors. At E7.5, these precursor cells start to migrate toward the genital ridge, and meanwhile undergo the genome-wide DNA demethylation. At E12.5, sex determination occurs through the expression of Sry gene, which is a Y-linked gene encoding a transcription factor responsible for the initiation of testis development (Manku & Culty, 2015). At E12.5, these cells form testicular cords and they are called gonocytes, which are englobed by Sertoli cells. Within the testicular cords, a mix of quiescent and proliferative active PGCs co-exists in neonatal testis in mice (Culty, 2013).

During the second phase, the proliferative portion of PGCs undergoes several rounds of spermatogonial mitotic divisions, which lead to their differentiation into primary spermatocytes. These cells go through two successive meiotic divisions. The first meiosis generates secondary spermatocytes, which themselves go through the second meiosis and produce haploid spermatids. These cells undergo a maturation process involving dramatic chromatin reprogramming processes, which lead to the production of mature spermatozoa and

the packaging of the male genome (Figure 18). In this section, the step by step spermatogenic cells differentiation process in addition to general structure of testis will be discussed.

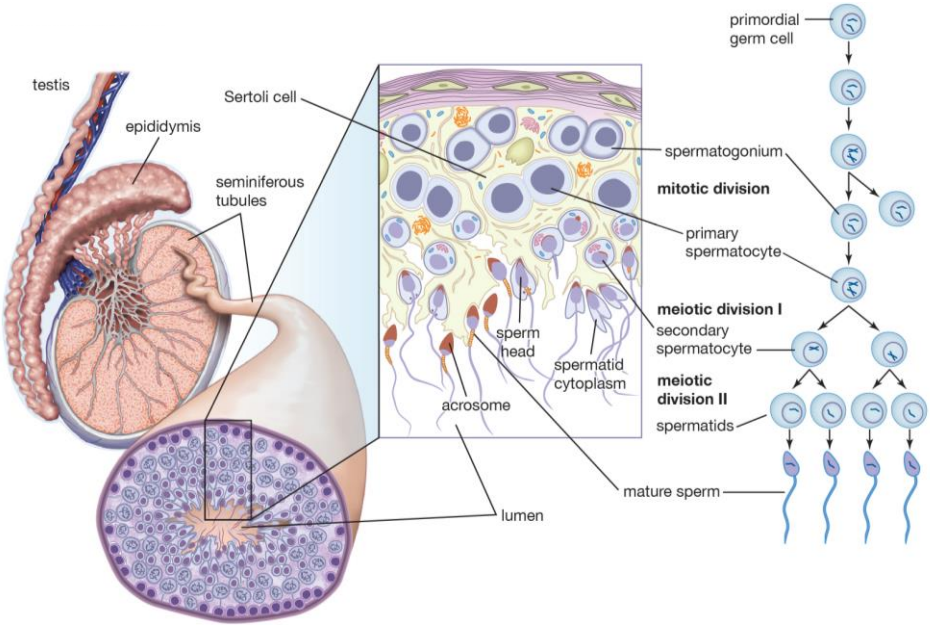


Figure 18. Schematic representation of spermatogenesis and genome compaction of mature sperm in testis. Taken from www.britannica.com

1.1. General organization of adult testis

The mammalian male reproductive system is composed of different organs and glands such as testes, extra-testicular ducts and sex glands. Mature mammalian testis, the male reproductive gland, is wrapped inside a layer of connective tissue called Tunica albuginea. The two male gonads, the testes, possess oval structures, and are localized in and protected by the scrotum, which is an extension of abdominal wall. In terms of structure, tunica albuginea is divided into numerous compartments/lobes harboring seminiferous tubules, which are the primary place of the spermatozoa production. Seminiferous tubules possess a helicoidal structure and their both extremities are open to the *rete testis*, a place which delivers the spermatozoa to the efferent duct which is the initial part of the epididymis (Figure 19A).

Seminiferous tubules are separated by interstitial (or intertubular) tissues which mainly contain Leydig cells, blood and lymphatique vessels, macrophages and mesenchymal tissues (Jan, Hamer et al., 2012). These tubules contain both germinal spermatogenic cells and somatic cells referred to as Sertoli cells. Sertoli cells are attached together through tight junctions and generate the blood-testis barrier. This barrier divides seminiferous epithelium into two segments: basal and adluminal compartments, which limit the intercellular transit of liquids and substances due to their high metabolic demand and high immunogenicity nature of these cells (Neto, Bach et al., 2016). Despite the non-dividing nature of Sertoli cells, these cells are essential for the nutrition of developing spermatogenic cells. They are also involved in the phagocytosis of the residual body produced during post-meiotic stage of spermatogenesis. After pituitary LH stimulation, leydig cells produce steroid hormones mainly androgens including testosterone, which are key factors for the regulation and development of spermatogenesis. Leydig cells are located between blood vessels and seminiferous tubules in testis. Three waves of Leydig cells differentiation occurs in humans. The first wave starts between 8 and 18 weeks of gestation as the result of the sex-determining region (SRY) gene - silencing. The second wave starts 2-3 months after birth due to an increase of the LH hormone. The third wave of maturation is initiated during puberty after the maturation of HPG and adult LCs remain active during adulthood.

Spermatogenic cells which constitute the epithelium of seminiferous tubules contain spermatogonia, spermatocytes and spermatids. In addition, collagen fibers, fibroblasts, and contractile cells referred to as myoid cells, are the main components of the wall of seminiferous tubules (Figure 19B).

Spermatozoa which are produced in the testes will finish their process of maturation by passing through epididymis, which is a long tube connecting each testis to the vas deferens, a place where sperms are concentrated and stored (Figure 19A). Epididymis is composed of three parts: 1. *Caput* (head), 2. *Corpus* (body) and 3. *Cauda* (tail). Immature spermatozoa exit from the *rete testis* and enter into *caput* and pass through the *corpus* and become mature inside the *cauda* before its entrance inside the vas deferens, which conducts the transport of sperms from epididymis to the ejaculatory ducts.

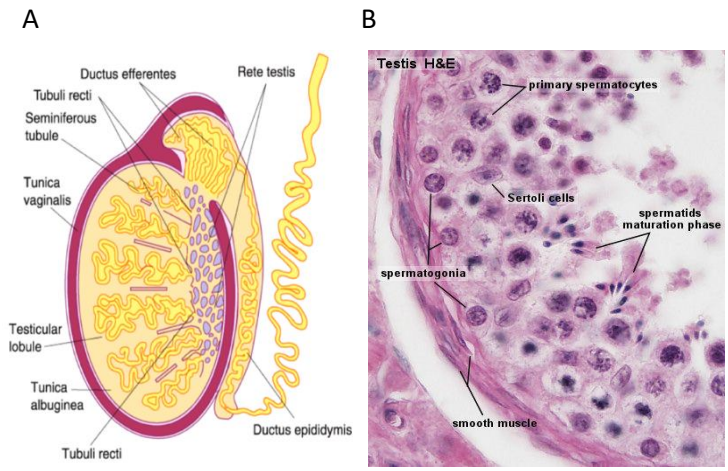


Figure 19: A) Schematic representation of seminiferous tubules, Vas deferens, ductus epididymis and Rete testis, taken from https://www.slideshare.net/bubbosahaljasim1/male-reproductive-system-lab. **B)** Haematoxylin and Eosin staining of one cross section slide of tubule, demonstrating different patterns of staining in different stages of germ cells, adapted from https://embryology.med.unsw.edu.au/embryology/index.php/Sertoli_cell.

1.2. Three phases of spermatogenesis

Spermatogenesis takes place roughly during three major phases: mitotic (also known as proliferative phase), meiotic and post-meiotic. Each of these phases are specifically discussed in this section.

1.2.1. Mitotic phase

During mitotic phase, germinal stem cells called spermatogonia rapidly proliferate and become mature. Spermatogonia are localized at the proximity of seminiferous epithelium, are in direct contact with basal lamina of seminiferous tubule walls and are located outside of blood-testis barrier. One portion of spermatogonial stem cells referred to as type A, maintain the pool of adult stem cells within the testis. Another portion of spermatogonial stem cells proliferate by mitosis and differentiate from type A to type B and establish the pool of progenitor germ cells, which then after undergo several rounds of mitotic divisions. Type B spermatogonia contain several nucleoli and are transcriptionally-active in order to provide the

proteins which are necessary for the next mitotic division. These divisions generate the first cells for the meiotic phase, which migrate through the lumen of seminiferous tubules.

1.2.2. Meiotic phase

In human, meiosis is only initiated at the onset of puberty in post-natal life. Meiosis refers to a specialize process which has the goal to maintain the same chromosome number post-fertilization from one generation to another. In order to fulfil this aim, meiosis involves two rounds of cell division during two steps known as meiosis I and meiosis II. Each step is divided into five consecutive stages: prophase, metaphase, anaphase, telophase and cytokinesis (Figure 20).

During prophase I, the pairs of homologous chromosomes gather to form tetrads containing four chromatids, therefore the recombination can occur between any of the four chromatids. In contrast to mitosis, meiosis is a much longer process due to its long prophase stage, which is divided into five stages: leptotene, zygotene, pachytene, diplotene and diakinesis. DNA duplication is initiated in the beginning of the leptotene stage referred to as preleptotene. During leptotene, progressive condensation of chromosomes occurs and the synaptonemal complexes are assembled. In zygotene, homologous chromosomes match and line up in the middle of the cells through the assembly of the central element of synaptonemal complex. Homologous recombination and crossing over between the same chromosomal pairs take place in pachytene stage (Figure 20). Moreover, non-sister chromatids of homologous chromosomes exchange homology segments with each other. During diplotene stage, the synaptonemal complexes degrade and homologous chromosomes separate from each other. The nuclear membrane disappears and is degraded into vesicles during the diakinesis and two homologous chromosomes line up on the metaphase plate and meiotic spindle is formed. During metaphase stage, the homologous chromosomes line up against each other on the metaphase plate and the sister chromatids are attached to kinetochores for segregation. During metaphase I, sex chromosomes (X, Y) go through a recombination due to the similar parts of the chromosome between them. This recombination is essential for ensuring the proper segregation of sex chromosomes. In anaphase I, the microtubules become shorter and pull away two homologous chromosomes to the opposite poles, which is followed by the end of meiosis

I and generation of two daughter cells each containing one pair of homologous chromosomes. The second meiosis is similar to the mitotic division. During the course of the second meiosis, sister chromatids separate through the action of kinetochore shortening in anaphase II leading to the generation of four haploid spermatids.

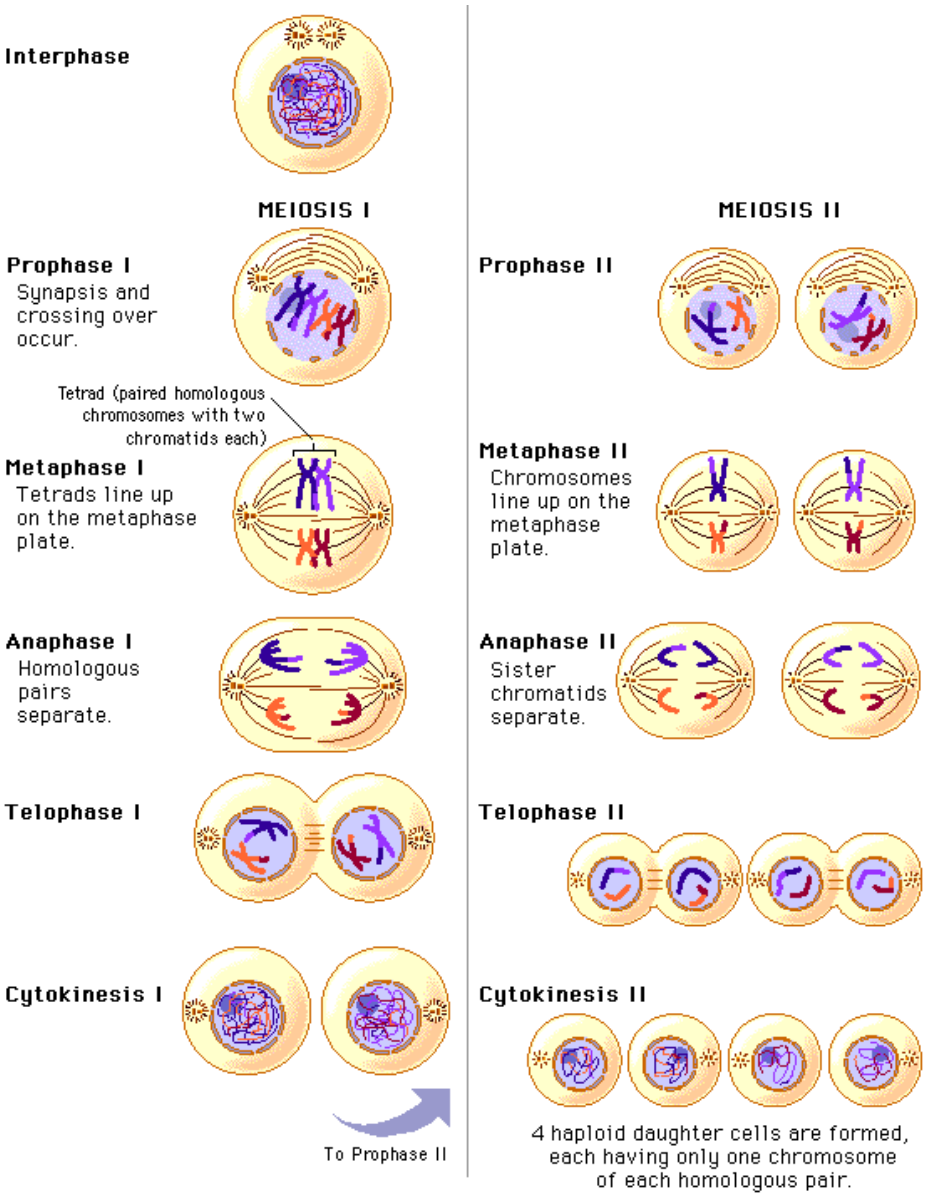


Figure 20. Schematic representation of meiosis. Taken from Pearson education.

1.2.3. Post-Meiotic phase (Spermiogenesis):

After consecutive meiotic divisions, haploid round spermatids undergo the final differentiation and maturation processes, which mainly involve three major steps:

I) Golgi apparatus differentiates and generates granules, which later fuse together and generate the acrosome. Acrosome contains the enzymes necessary for the penetration of spermatozoa inside the oocytes through the lysis of the zona pellucida, which is a glycoprotein layer surrounding mammalian oocyte. Acrosome forms at the anterior/head pole of the nucleus and it densifies and continues up to the posterior pole of the nucleus allowing the detection of different stages of maturation.

II) The formation of the flagellum or tail is initiated by the development of an axoneme from one of the two centrioles. The flagellum development induces the spermatid's plasma membrane to extend. This extension creates a thick midpiece containing mitochondria, which provide the energy necessary for the sperm motility (Figure 21). The excess of cytoplasm is discarded through residual body (RB) formation, RBs are then eliminated through phagocytosis by Sertoli cells.

III) The nucleus of the haploid spermatids elongates and the chromatin condensates at the maximal level. This condensation is characterized by the complete reorganisation of the chromatin structure. During this reorganisation, spermatid chromatin changes from nucleosome-based structure to a nucleoprotamin chromatin. This new genome-packaging structure protects the sperm genome from the environmental dangers on their ways through the female genital tract.

During post-meiotic differentiation of spermatids, these cells undergo major morphological transformations involving genome-wide chromatin compaction transitional states. Early post-meiotic spermatids are referred to as round spermatids. These cells possess a specific transcriptional programme, which directs to the progression of round spermatids toward the generation of mature spermatozoa. The nuclei of these cells become elongated which is accompanied by a gradual shutdown of transcription. In order to maintain the expression of essential proteins such as transition proteins TPs and PRMs in condensing spermatids and accomplish the final differentiation of spermatozoa, the mRNAs encoding the essential late-

expressing proteins are stored throughout the post-meiotic phase and are translated in later stages (Rathke, Baarends et al., 2014).

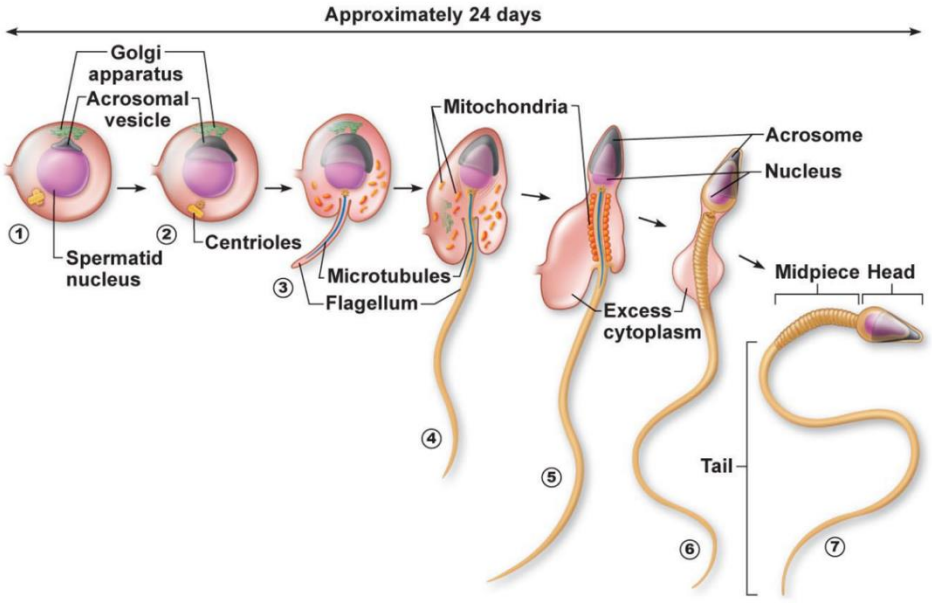


Figure 21. Spermatid maturation involving chromatin condensation, gradual development of the flagellum and acrosome formation during post-meiotic phase. Taken from Pearson education, 2013.

1.3. Final maturation of spermatozoa in epididymis

The epididymal environment is crucial for the maturation and storage of spermatozoa. Since sperm cells transcriptional and translational programme is completely turned off, their functional maturation is completely dependent on the specialized epididymal environment. Epithelial cells, referred to as peritubular interstitium, constitute around 80% of intraluminal epididymis, contain secretory apparatus such as ER, Golgi and secretory granules. These cells bear high exocytotic activity, more specifically near caput and corpus and endocytotic activity near distal segments such as cauda. The maturation of sperm cells and the intercellular communication between sperm and epithelial cells is highly buffered through the secretory and absorptive activity of the different populations of epithelial cells within epididymis. Other elements which regulate the intraluminal environment are the vesicles referred to as

epididymosomes, which are responsible for the delivery of the macromolecules necessary for the sperm cells maturation (Zhou, De Iuliis et al., 2018).

In addition, these vesicles contain a variety of sncRNA including miRNA and tRNA fragments, which are delivered to maturing sperm cells. Moreover, sperm cells can transfer these sncRNAs to the oocytes upon fertilization. It has been also demonstrated that, these sncRNAs regulate the epigenetic control of the embryo through targeting of specific subset of genes (Sharma, Conine et al., 2016).

1.4. Stages of spermatogenesis

In mice, male germ cell differentiation is classified into 12 different stages, corresponding to one cycle of maturation of spermatogonia to spermatozoa. This classification was mainly established on the basis of the acid-Schiff technique (PAS-staining) of testis paraffin cross-sections. One stage is defined through the association of one group of 3 to 4 different cell types, corresponding to four generations of germ cells on one cross section of the seminiferous tubule (Figure 22A). Moreover, the kinetics of germ cell differentiation is specifically regulated in a way that the cells in the same stages of differentiation are constantly related (Meistrich & Hess, 2013). In mice, each 8.6 days, the spermatogonia cells differentiation is reinitiated in a coordinated and irreversible manner, in order to ensure the constant production of the sperm cells. In addition, spermatogonial sperm cell differentiation and migration both occur in a centripetal manner in order to guarantee the depot of sperm cells in the lumen of seminiferous tubules.

In mice, the whole process of spermatogenesis takes 35 days as follows: 13 days for the genesis of primary spermatocytes, one day for the production of secondary spermatocytes and 13 days for spermatids maturation. Upon birth, the seminiferous tubules contain mainly spermatogonia. After 11 to 14 days, spermatocytes start to appear and spermatids are generated around the 21st day after birth.

Germ cells in different part of seminiferous tubules are not in the same stage of differentiation, which ensure the constant production of spermatozoa. Different stages of germ cells are placed next to each other in one cross-section of the seminiferous tubule. Type A spermatogonia are not very abundant and are distributed in almost every stage. During stages

I-VIII the number of differentiating spermatogonia or preleptotene spermatocytes increase. The intermediate spermatogonia are mostly found in stages II to IV and after differentiation and becoming type B spermatogonia, they are mostly found in stages V-VI. Pre-leptotene (Pl), leptotene (l), zygotene (Z) and pachytene (P) are respectively abundant in stages VII-VIII, IX-X, X-XII and I-X. Diplotene spermatocytes are only found in stage XI. Secondary spermatocytes are present in XII stage. Haploid spermatids differentiation itself is separated into 16 stages. During stages 1-7, spermatids develop acrosome granules and these granules slowly come to the surface of the nucleus. Stage 8 corresponds to elongating spermatids in which the acrosome covers half of the nucleus surface and the nucleus starts to elongate. Stage 10 to 12 corresponds to condensing spermatids with a high level of chromatin condensation and acrosomal coverage is more limited on the nucleus surface (Figure 22B).

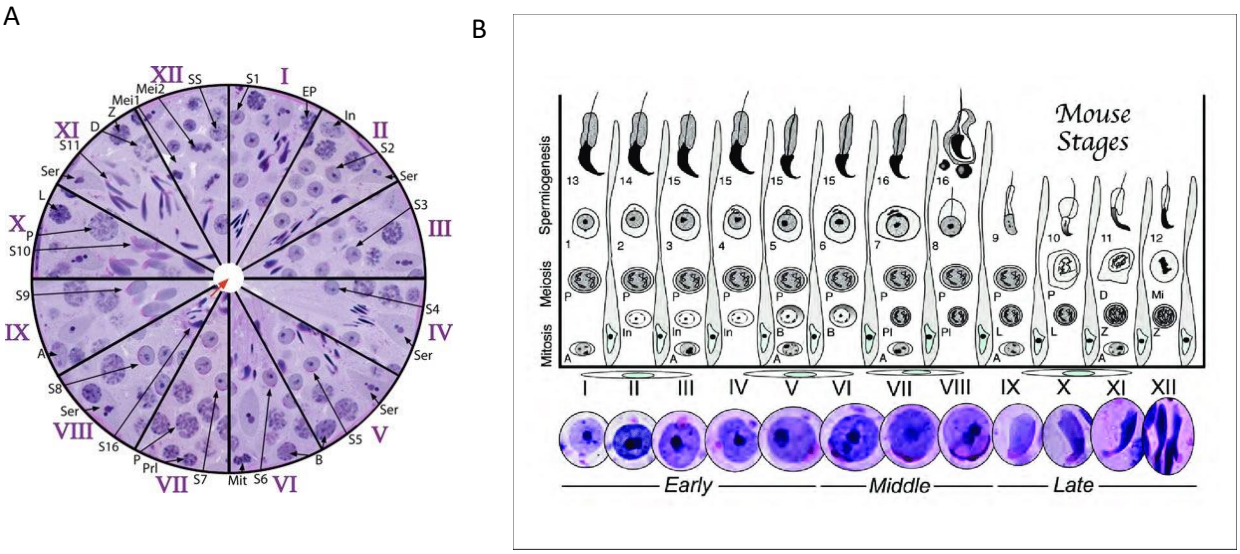


Figure 22. Step by step recognition of spermatogenic cell stages, inside of seminiferous tubules in mouse. Taken from (Meistrich & Hess, 2013), Methods Mol Biol.

V. Chromatin organization during mammalian spermatogenesis

1. General overview of chromatin organization during spermatogenesis

Germinal cells bear the responsibility to transmit their genome to the offspring. In order to fulfil this mission, germinal cells undergo a drastic genome reorganization during the post-meiotic stage of spermatogenesis to reach the maximum level of their genome protection. In rhesus monkey, three-dimensional higher-order chromatin structure analyses revealed that the TADs are mostly depleted in pachytene spermatocytes. At this stage, only small intermediary local genome organizing compartments are observed that switch between transcribing and non-transcribing regions (Wang, Wang et al., 2019).

The drastic genome-wide reorganization is mainly characterized by the replacement of the majority of histones with highly basic non-histone proteins called TPs and Prms (Figure 23). In contrast to the Prms, TPs play a crucial but transient role in the process of histone replacement and do not persist inside the mature sperms. Recent works from our laboratory showed that in contrast to the general belief, TPs and Prms act together and not successively (Barral et al., 2017).

It is also important to note that in certain species of vertebrates, birds and fish, histones are directly replaced by Prms or Prm-like proteins (Kurtz, Martinez-Soler et al., 2007, Oliva & Dixon, 1991). Moreover, spermatogenic cells also contain many testis-specific histone variants, which replace canonical histones during meiosis and post-meiotic spermatids upon the onset of genome reorganization and direct the first steps toward histone replacement. Therefore, in mammals the drastic genome reorganization in haploid spermatids is characterized by three major events; 1) canonical histone replacement by testis-specific histone variants creating less stable nucleosome formation (Govin, Caron et al., 2004, Montellier, Boussouar et al., 2013) (Barral et al., 2017), 2), a wave of genome-wide histone hyperacetylation at the beginning of elongating spermatids (Boussouar, Rousseaux et al., 2008), and 3) histone removal and replacement by TPs and Prms (Gaucher, Reynoird et al., 2010, Goudarzi, Shiota et al., 2014).

Testis-specific variants of canonical histones H2A, H2B, H3 and H1 are expressed in testis. Few years ago in our laboratory, several mouse H2A and H2B testis-specific variants

(H2A.L.1, H2A.L.2, H2A.L.3, H2B.L.1 and H2B.L.2) have been identified through a proteomic analysis on isolated condensing spermatids (Govin, Escoffier et al., 2007). Functional and structural analyses of testis-specific variants in the nucleosomes revealed their association with the generation of unstable nucleosomes. Previously, we proposed that the drastic reorganization of the genome is a multi-step process in which histone variants destabilize the nucleosomes and prepare the nucleosome for invasion by TPs and Prms. Recently, our lab identified the essential role of H2A.L.2 in destabilizing nucleosomes, which lead to the nucleosome invasion by TPs, which in turn create a platform for Prms maturation and correct incorporation (Barral et al., 2017). This work highlighted the concomitant action of histone variant, TPs and Prms in the process of histone to Prm exchange (Figure 23). Interestingly, despite the global replacement of histones by Prms, in the mouse mature sperms about1% of histones remain in the mature sperm and they are not replaced by Prms (Erkek, Hisano et al., 2013). However other works propose histone retention in mouse up to 8-10% (Jung, Sauria et al., 2017).

Four types of Tps (TP1, TP2, TP3 and TP4) are described in the literature in different species. In mammals, among these four types, TP1 and TP2 contain 55% and 40% of the total nuclear spermatid's TPs, respectively. TP1 is a revolutionary conserved, small basic protein of 54 amino acids which is very rich in arginines, lysines and serines. TP2 is two times bigger than TP1 and contains two zinc finger domains and its c-terminal is enriched in basic residues (Kimmins & Sassone-Corsi, 2005). Surprisingly, histone-Prm exchange occurs in mice lacking either TP1 or TP2 or both of these proteins (Shirley, Hayashi et al., 2004, Zhao, Shirley et al., 2004a, Zhao, Shirley et al., 2004b). The fertility rate of TP1-null mutant mice is decreased up to 50% and TP2-null mutant mice pups are smaller in size. 80% of the mature sperm cells in TP1 and TP2-null double mutant mice are dead in the epididymis and few spermatids among the 20% alive sperms, are mobile (Shirley et al., 2004). Therefore, it is evident that TPs do not have a direct role in the histone removal process. the chromatin compaction is defective in mice lacking both TPs and these mice are infertile (Shirley et al., 2004, Zhao et al., 2004b). Interestingly, H2A.L.2 knock out mice revealed a defect in chromatin compaction as well as infertility, similar to both TP1 and TP2- deficient mice (Barral et al., 2017). Hence, these evidences reveal the crucial role of histone variants and TPs in the process of histone to protamine exchange in haploid spermatids.

Prms are small arginine-rich proteins which are responsible for the high chromatin compaction in condensing spermatids. Prms are able to pack the spermatids DNA in a volume less than 5% of a somatic cell nucleus (Wyrobek, Meistrich et al., 1976). Prms harbor structural properties that confer specific capacities. For instance, the central portion is highly rich in arginine which generates hydrogen bonds with DNA through the amine group of the arginine residues, facilitating the Prm-DNA interaction. Protamine's two terminal extremities are highly rich in cysteine, which become phosphorylated through their serine and threonine residues. After binding to DNA, they lose the phosphate group and cysteine residues are oxidized, leading to the generation of disulfide bridges that strengthen the interaction between Prms and confer stability to the highly compacted chromatin (Balhorn, 2007).

In general, vertebrates possess from one to 15 protamine genes per haploid genome, which create clusters on the same chromosome. In mammals, Prm1 and Prm2 are the most studied. Most mammals contain Prm1, whereas Prm2 is only found in most rodents, primates and few other placental mammals. In mice, Prm1 and Prm2 gene transcription initiate in stage 7 round spermatids, at the same time as TP encoding genes (Mali, Kaipia et al., 1989). Meanwhile, Prm1 and Prm2 proteins are only expressed in elongating and condensing spermatids, respectively, corresponding to stages 14 to 16. The delay in Prm2 protein expression is due to its post-translational regulation. In contrast to Prm1, which is directly translated from the mRNA, Prm2 is synthesized as a precursor that undergoes proteolytic processing after DNA binding. In an attempt to understand if the two Prms provide redundancy, one of the two Prms genes was mutated. This study revealed that Prm haploinsufficiency results in mice sterility and both Prms are necessary for the proper sperm function and processing of Prm2 (Cho, Willis et al., 2001).

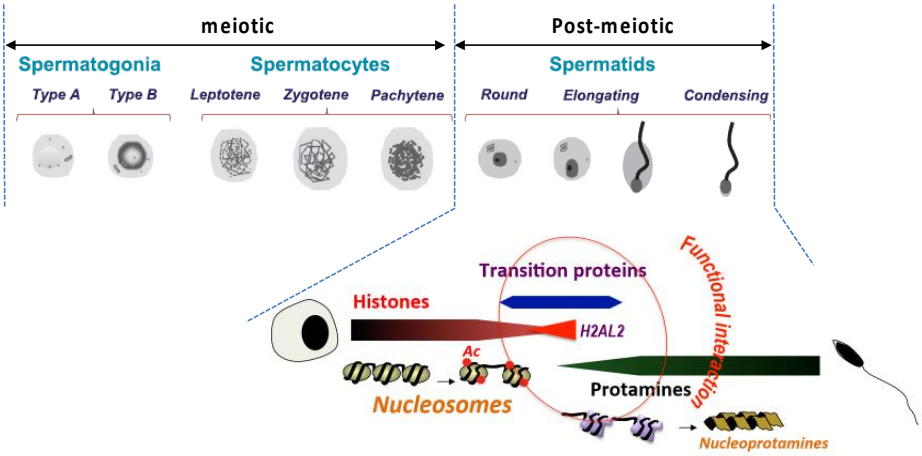


Figure 23. Molecular mechanisms underlying histone-to-protamine exchange. The representation of the concomitant action of the genome-wide histone hyper acetylation, H2A replacement by H2A.L.2, appearance of TPs and Prms initiated in elongating spermatids.

It is noteworthy to mention that, in addition to the fact that TPs and Prms compete for binding to DNA and that they are the driving force for histone removal, it is also relevant to take into account the possibility of TP/Prm-associated chaperones, as well as histone chaperones, that are responsible for TPs and Prms deposition and histone removal in haploid spermatids.

The mechanisms underlying degradation of histones are not clearly defined. Some data suggest histone degradation through specific enzymes such as proteases or the ubiquitin-proteasome system (Faulkner & Bhatnagar, 1987, Marushige, Marushige et al., 1976). In one of these studies, the invalidation of *RNF8*, a gene encoding a histone E3 ubiquitin ligase, caused a defective histone removal. Furthermore, they observed a decrease in the H4K16 acetylation level in *RNF8*$^{-/-}$ mice, which could be explained by the downregulation of the mof HAT (Lu, Wu et al., 2010). Surprisingly, another study evidenced no change in either H4K16ac level or in histone-Prm exchange process (Sin, Barski et al., 2012). Therefore, we can speculate that histone hyperacetylation is a consequence of the decrease in HDACs level in elongating spermatids (Goudarzi et al., 2014). Moreover, It has been shown that proteasomes containing the activator PA200/Blm 10 specifically target the core histone and mediate their degradation in an ubiquitin-independent manner (Qian, Pang et al., 2013).

In the following section, the role of major actors in histone-to-protamin exchange process are further discussed. Precise histone variants implications during histone removal and Prm assembly will also be further discussed in the final chapter.

2. Major actors involved in histone-to-protamine exchange during spermiogenesis

2.1. NUT

Other testis-specific protein involved in the process of histone-to-protamine exchange, is NUT. NUT is involved in a highly aggressive type of carcinoma known as NMC. *NUT* gene is located on the chromosome 15q14 and is fused to ubiquitously expressed, double bromodomain-containing BRD4 (67% of cases) or BRD3 (25% of cases) proteins (French, Miyoshi et al., 2003, Wu & Chiang, 2007). Thus, in the context of NMC, a chromosomal translocation t (15;19) (q13; p13), leads to the generation of a fusion protein. Further studies also in the context of NMC, revealed the strong interaction between the NUT (through the N-terminal F1c region containing amino acids 346-593) of this fusion protein and CBP/p300. In fact, this interaction strongly enhanced CBP/p300 HAT activity and led to the propagation of transcriptionally-inactive hyperacetylated foci (Figure 24, lower panel). Furthermore, by using patient-derived cells, our laboratory could show that the trapping p300 into BRD4-NUT foci, led to p53 inactivation and was found as the major oncogenic mechanism involved in NMC (Reynoird, Schwartz et al., 2010). This study highlighted the off-context role of a testis-specific factor and its contribution to malignant cell transformation. Further experiments revealed that BRD4-NUT-induced histone hyperacetylated chromatin domains are controlled through HDACs (Schwartz, Hofer et al., 2011) and are also limited to the TAD domains (Alekseyenko, Walsh et al., 2015).

The physiological role of NUT in the context of spermatogenesis has been recently identified. The expression of NUT is restricted to the late round and the beginning of elongating spermatids. *NUT*$^{-/-}$ mice show a major arrest in histone-to-protamine exchange process which led to the arrest of spermatogenesis. In addition, these mice contained smaller testes; they were sterile and lacked mature spermatozoa inside the cauda epididymis. Comparative mass spectrometry and western blotting analysis of histone PTMs on the late round and early

elongating *NUT*$^{-/-}$ and WT mice, showed a decrease in histone H2A and H4 acetylation, specifically on H4K5 and H4K8 positions. Transcriptomic analysis on *NUT*$^{-/-}$ and WT mice, revealed the implication of NUT on the expression of a subset of genes expressed in spermatids. In this study, NUT association with p300 had also been observed in the context of spermatids similar to the NMC context (Reynoird et al., 2010). Besides, ectopic BRD4-NUT interaction with CBP/p300 (as the only HAT that interacts with BRD4-NUT) had been also observed in the NMC context (Alekseyenko, Walsh et al., 2017). Interestingly, comparative transcriptomic analysis of p300 (Boussouar et al., 2014) and NUT-associated gene expression in spermatids (Shiota, Barral et al., 2018), demonstrated the expression of the same gene population is controlled by both NUT and p300. These experiments along with many others, revealed that in the context of spermatids, NUT interaction with p300 enhances p300 HAT acetylation activity especially on H4K5 and H4K8 residues, providing acetylated sites for Brdt first bromodomain, which later mediate the genome-wide histone removal (Goudarzi et al., 2016, Shiota et al., 2018).

2.2. BRDT

In many different species such as the fly, trout, rooster, rat, mouse and human, the genome-wide wave of histone hyperacetylation is associated with the histone removal process in elongating spermatids (Christensen & Dixon, 1982, Grimes & Henderson, 1984, Hazzouri, Pivot-Pajot et al., 2000). Therefore, many studies focused on finding the relation between these two events. Early *In vitro* studies revealed that, once incubated with Prms, histone removal is facilitated after histone hyperacetylation (Marushige et al., 1976, Oliva, Bazett-Jones et al., 1987, Oliva & Mezquita, 1986). In somatic cells histone assembly and removal are linked to specific factors and machineries (Burgess & Zhang, 2013, Ray-Gallet & Almouzni, 2010). In mouse, our laboratory has particularly studied Brdt, which is a member of the BET sub-family of double bromodomain-containing proteins. BET proteins are a class of transcriptional regulators, which have been found in animals, plants and fungi. In mammals, this family of proteins contain Brd2, Brd3 and Brd4, in *Drosophila* Fsh and in yeast Bdf1p and Bdf2p. In plants, BET family of proteins contain one bromodomain (the domain that specifically recognize the acetylation mark) and in human, mouse and fungi, BET harbor two bromodomains (BD1 and BD2) spaced ~ 150 residues apart at its N-terminal end of the protein.

BET family of proteins also harbor an ET domain, which bears a protein-protein interaction motif. Brd2 and Brd3-associated chromatin is highly enriched at H4K5, H4K12 and H3K14 acetylation. H3K9 di-methylation as a marker of facultative heterochromatin, also has been associated to regions that Brd2 and Brd3 were located (LeRoy, Rickards et al., 2008). Thus, on one hand, BET proteins interact with chromatin through their bromodomains and on the other hand they attract different protein complexes to the active chromatin through their ET domain (Florence & Faller, 2001).

Brdt/BRDT is the mammalian homolog of TAF-1 in human and its expression is limited to male germ cells, precisely in pachytene-diplotene spermatocytes and during later stages all through spermiogenesis. Testis-specific expression of BRDT was also confirmed in mouse testis, after cloning the murine cDNA of this protein. It has been shown that, BRDT specifically binds the hyperacetylated N-terminal tail of histone H4. Interestingly, *in vivo* and *in vitro* studies revealed BRDT's high capacity in chromatin remodeling only after hyperacetylation induction via TSA. In addition, in this study, the authors demonstrated that both BRDT bromodomains as well as the domains flanking bromodomains are necessary for the proper action of BRDT (Pivot-Pajot, Caron et al., 2003). In 2009, the crystal structure of BD1 bound to diacetylated H4 tail, revealed the cooperation between acetylated lysines (H4 K5 and K8) for binding to one common pocket. Surprisingly, BD1 was not able to bind mono-acetylated H3 and H4 peptides. Interestingly, a later study showed that *in vitro* BD1 can also bind nucleosomes bearing H4K5acK8ac, while BD2 could not bind nucleosomes at all (Miller et al., 2016). Functional studied using different mouse models showed that the binding H4 di-acetylated K5 and K8 by Brdt's BD1 mediates the histone eviction in elongating spermatids (Gaucher, Boussouar et al., 2012).

These studies highlighted for the first time, the multi-acetylated ligand recognition of one pocket (here BD1) and its functional implications. These discoveries can have possible implications for the understanding of the functions and properties of other bromodomain-containing factors such as BET family members, TAF-1 and CBP/p300 (Moriniere, Rousseaux et al., 2009). Furthermore, *In vivo* analysis in $Brdt^{\Delta BD1/\Delta BD1}$ mice, revealed a defective spermatogenesis and abnormalities at the onset of elongating spermatids and histone-protamine exchange (Shang, Nickerson et al., 2007). In 2012, Gaucher et al, through a series of functional studies of $Brdt^{-/-}$, $Brdt^{\Delta BD1/\Delta BD1}$ and a dominant negative *Brdt* expressing mouse model, presented the first complete molecular analyses of this protein in its physiological context

during mouse spermatogenesis. RNA extracted from prepubertal and adult mice testes, revealed the first *Brdt* mRNA accumulation at 10-12 d.p.p. and the protein production at 12 d.p.p. *Brdt*$^{-/-}$ mice showed a complete absence of post-meiotic cells but no defect in earlier cell stages in adult mice testes. In addition, transcriptomic analyses on *Brdt*$^{-/-}$ mice in pachytene/diplotene stage, revealed the downregulation of two-third and up-regulation of one-third of the genes. Interestingly, ChIP-seq analysis revealed BRDT localization at the promoter of subset of genes with hyper acetylated TSSs. Moreover, despite the synthesis of TPs and Prms in *Brdt*$^{\Delta BD1/\Delta BD1}$ mice, they remained inside the cytoplasm, following a complete arrest of spermatogenesis (Gaucher et al., 2012). All these data, point to BRDT as a master regulator of meiotic and post-meiotic male germ cell gene expression as well as an as well as an important mediator of histone-to-protamine exchange process.

2.3. Histone variants in mammalian spermatogenesis

As previously mentioned, during histone-to-protamine transition process in spermatids, canonical histone replacement by histone variants is one of the major steps towards genome-wide histone removal. It is therefore, crucial to understand the precise functional and molecular mechanisms of various histone variants expressed in this process. In this chapter, in the context of spermatogenesis, different characteristics and the role of various mammalian histone variants, with a precise focus on murine variants, are discussed (Figure 24).

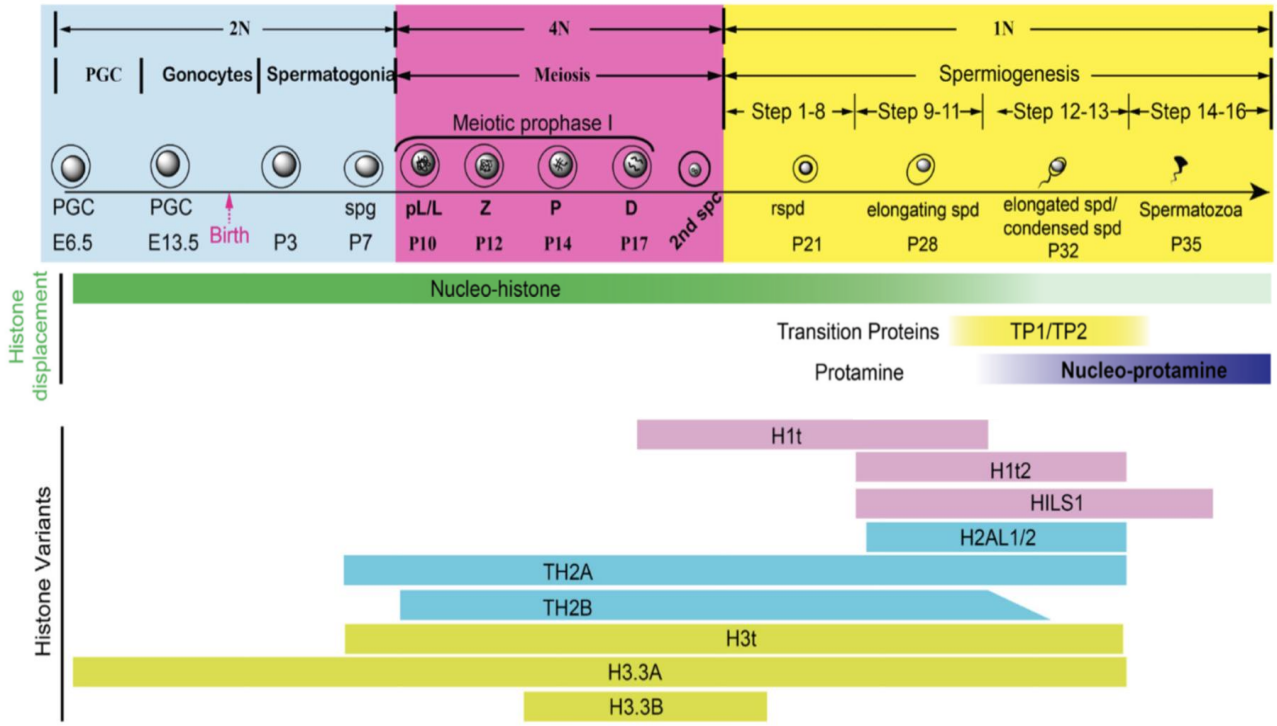


Figure 24. Recapitulative of different stages of spermatogenesis with color-coded representation of testis-specific histone variants in each stage. Taken from (Bao & Bedford, 2016)

2.3.1. Histone H1 variants

Among linker histone H1 variants, H1t, H1T2 and HILS1 are specifically expressed in testis (Hergeth & Schneider, 2015).

2.3.1.1. H1t

H1t is specifically synthesized during prophase of the first meiosis, from mid to late pachytene spermatocytes, and remains in the nucleus until chromatin reorganization occurs during elongating stages and is lost after the initiation of chromatin compaction in spermatids (Wolfe & Grimes, 2003). H1t expression had been observed in rat, mouse, hamster, rabbit, monkey, bull and human (Grimes, Wolfe et al., 1990). In contrast to canonical H1, H1t binding capacity to chromatin is weaker, due to the absence of several repetitive DNA-binding motifs (S/TPKK- AAKK/AKKA-PKAAKAP- K/RPK/R) at its C-terminus (Khadake & Rao, 1995). These observations are indicative of the fact that H1t might induce less stability to the chromatin fiber compared to canonical linker histone H1 (De Lucia, Faraone-Mennella et al., 1994, Khadake & Rao, 1995). In an attempt to investigate the role of H1t in spermatocytes, round and elongating spermatids, *H1t* gene was inactivated. These mice did not show any specific phenotype and the mice were fertile without any phenotypic defect in spermatogenesis. Further studies, demonstrated that the absence of H1t is to certain degree compensated by other linker H1s, leading to proper spermatogenesis and therefore fertilization (Drabent, Kardalinou et al., 1991, Govin et al., 2004, Lin, Sirotkin et al., 2000).

2.3.1.2. H1T2

H1T2 is expressed in round and stages 9-14 elongating/condensing spermatids and is specifically localized at apical pole within a cap-like structure at the inner periphery of the nuclear membrane under the acrosome. This observation was further detailed after H1T2 was observed in total nuclear extracts in which the majority of the protein was in the chromatin fraction. In addition, immunodetection analyses showed that the localization of this protein is asymmetrical within the chromatin domains in both round and elongating spermatids. In

contrast to round spermatids where both H1t and H1T2 are expressed, they only share partial expression in elongating spermatids, since H1t disappears upon the initiation of elongation. Meanwhile, H1T2 expression is concomitant with TPs and Prms expression at the time of histone protamine transition. *H1T2*$^{-/-}$ mice revealed abnormal elongation and chromatin condensation as well as a strong impairment of fertility. Despite a decrease in litter volume, the pups did not show any defect and were normal and viable. Furthermore, the fertility capacity of *H1T2+/-* heterozygous mice was comparable to WT. Thus, the asymmetrical organisation of H1T2 in spermatids is essential for proper chromatin condensation during spermiogenesis (Martianov, Brancorsini et al., 2005, Tanaka, Iguchi et al., 2005).

2.3.1.3. HILS1

Spermatid-specific HILS1 is expressed in both mouse and human and *HILS1* genes do not harbour introns and are located in the intron 8 of the alpha-sarcoglycan protein-coding gene. HILS1 counts for 10% of total chromatin protein in mouse spermatids. Protein alignment revealed that human and mouse HILS1 share 50% amino acid identity, in which 51% cover their globular domain. C-terminal and globular domain of human and mouse HILS1 are similar in size, while, human, HILS1 has a longer N-terminal region. Moreover, 20% of amino acids of this linker histone variant are strongly basic, such as lysine or arginine and its globular domain contain a winged-helix structure similar to H1 (Yan, Ma et al., 2003). Similar to H1T, HILS1 is only detected in spermatids between steps 9-15, in elongating and condensing spermatids nucleus, at the same time as the histone-to-protamine transition (Bao & Bedford, 2016, Iguchi, Tanaka et al., 2004, Yan et al., 2003). The fact that the three types of H1 variants are expressed in different spermatids with little overlap in their expression timing, suggest the sequential action of linker histones in chromatin organization during spermiogenesis (Govin et al., 2004). Mass spectrometry analyses led to the identification of 15 novel PTMs on N-, C-termini and globular domain of HILS1, including, lysine acetylation, serine/threonine and tyrosine phosphorylation sites (Mishra, Gupta et al., 2015). Meanwhile, the molecular significance underlying these PTMs is yet to be studied. Recent study on oligo-nucleosomes prepared on a sucrose gradient followed by ultracentrifugation, revealed that adding HILS1 on oligo-nucleosomes does not lead to chromatin compaction. In addition, this study demonstrated

the association of HILS1 with intergenic repetitive LINE-1 elements in spermatids (Mishra, Shalini et al., 2018).

2.3.2. Histone H2B variants

Up today, most of the known H2B histone variants are exclusively expressed in male germ cells.

2.3.2.1. TH2B

TH2B or (TSH2B according to the new nomenclature) (El Kennani et al., 2017), was first identified in mature testis extract of several mammals, based on the biochemical fractionation and electrophoresis analyses (Branson, Grimes et al., 1975, Shires, Carpenter et al., 1975). In contrast to other histones, TH2B is conserved among species (Zalensky, Siino et al., 2002). Mass spectrometry analyses revealed that TH2B contains three additional phosphorylation sites and two (S11 and T22) at its N-terminal tail compared to somatic H2B (Lu, Xie et al., 2009). Gene invalidation analyses of *Th2b* without disturbing *Th2a* gene, was a challenge for decades for functional studies (Choi & Chae, 1993, Huh, Hwang et al., 1991). My host laboratory could design a strategy to specifically silence *Th2b* gene without affecting *Th2a* gene Montellier et al. (2013).

The first functional studies on TH2B in its physiological context, during spermatogenic cell differentiation were carried out mainly through the generation of different mouse models (Montellier et al., 2013). In this study, in one construct, three consecutive affinity tags were introduced to the C-terminus of TH2B. In another mouse strain, the expression of TH2B was completely aborted. Results obtained from these mice revealed that from the beginning of meiosis, TH2B massively replaces H2B. Moreover, TH2B is expressed during the whole spermatogenesis process and its accumulation specifically starts in early spermatocytes, suggesting its wide-range of action during spermatogenesis. However, the complete absence of TH2B showed no defect during spermatogenesis and the mutant mice were fertile. However, in this model, a compensatory replacement of H2B was observed. Quantitative and

comparative proteomic approaches demonstrated that in the absence of TH2B, histones gain some novel PTMs located at strategic positions, capable of inducing the nucleosome instability and therefore mimicking the role of TH2B in its absence. For instance, among these PTMs, the methylation of H4R34, H4R55, H4R67 and H2BR72, residues that are implicated in histone-DNA interaction, could potentially induce nucleosome stability.

Interestingly, in TH2B mutant mice containing three affinity tags at the C-terminus, no spermatozoa production was observed. In these mice spermatogenesis was blocked at the beginning of elongating stage. Our hypothesis, based on the data obtained is that the presence of Tag at the C-terminus of TH2B, somehow affects the exchange of TH2B-TH2A by TH2B-H2A.L.2, recently shown to be necessary for the invasion of nucleosomes by TPs and Prms and the final Prm assembly and genome compaction (Barral et al., 2017).

Double *Th2b*$^{-/-}$ and *Th2a*$^{-/-}$ mice were infertile. In these mice, the absence of TH2B is compensated by the increase and accumulation of H2B, but no H2A accumulation has been observed. Therefore, probably an under-dosage of H2A in the nucleosome, disturb the proper chromatin organisation and hence prevents one to conclude on the specific role of TH2B and TH2A (Shinagawa, Huynh et al., 2015).

2.3.2.2. H2BFWT/H2B.W

Nucleosome reconstitution analyses showed that H2BFWT can replace the conventional H2B in the nucleosome despite its low homology (45%) with this conventional histone. However, the presence of H2BFWT in the nucleosome does not influence the structure of the nucleosome in terms of stability (Boulard, Gautier et al., 2006). H2BFWT has been associated to the telomeric chromatin in human mature spermatozoa (Churikov, Siino et al., 2004). Additionally, SWI/SNF appeared to be capable of remodeling H2BFWT nucleosome. The same study revealed that the N-terminus of H2BFWT shares few homologies with the N-terminus of conventional H2B. Additionally, H2BFWT-containing nucleosomes were defective for the recruitment of chromosome compaction factors and participate to the assembly of mitotic chromosomes (Boulard et al., 2006).

2.3.2.3. SubH2Bv/H2B.L.1

SubH2Bv (H2B.L.1 according to the recent nomenclature) is a H2B variant first identified in bull, which is specifically expressed in male germ cells. This variant is located in the sub-acrosome of primates, rodents and marsupials as well as in bovids (Talbert, Ahmad et al., 2012). Moreover, it is located at the sub-acrosomal region in spermatozoa which is mostly implicated in the assembly of acrosome (Aul & Oko, 2002). In mice, SubH2Bv orthologs (H2B.L.1) were identified through proteomic approach in my host laboratory (Govin et al., 2007).

2.3.3. Histone H3 variants

In mammals**,** H3.1, H3.2, H3.3 and CENP-A are variants of canonical histone H3, that are ubiquitously expressed in every tissue. H3t is the testis-specific H3 variant.

2.3.3.1. Testis-specific H3

Testis-specific H3 was first reported about 40 years ago on the basis of peptide analysis (Franklin & Zweidler, 1977). H3T also known as H3.4 (Witt, Albig et al., 1996), is exclusively expressed in human, rat and mouse testis and at a lower extent in some somatic cells, specifically under pathological situations (Govin, Caron et al., 2005, Tachiwana et al., 2008, Trostle-Weige, Meistrich et al., 1984). H3t synthesis is initiated in spermatogonia and continues in spermatocytes and early spermatids (Trostle-Weige et al., 1984). FISH experiments revealed H3t gene localization on the chromosome 1 (1q42), which differed from the previously reported histone cluster localization on chromosome 6 (Albig, Ebentheuer et al., 1996). Only 30 years after its discovery, the possible functions of this variant were proposed in human and mouse. In human, *in vitro* and *in vivo* analyses of H3T compared to conventional H3.1 revealed that H3T confers instability to the nucleosome. Crystal structures of H3T-containing nucleosome demonstrated Met71 and Val111 residues at the both ends of H3T α2 helix are responsible for inducing unstable nucleosome (Tachiwana, Kagawa et al., 2010). Other set of *in silico* approaches in mouse, revealed the replication-dependent H3t

incorporation in chromatin (Maehara, Harada et al., 2015). Finally, in 2017, functional studies were performed on H3t-deficient mice in order to reveal the role of this histone variant. H3t-deficiency in these mice led to infertility and azoospermia with smaller testes size, suggesting that this variant is necessary for the development of early stages of spermatogenesis and its significant role in chromatin organization during these stages. However in this case there is no compensation of other H3 genes and hence it is not possible to distinguish between histone H3 under-dosage and the specific function of H3t (Hoghoughi et al., 2018) In contrast to H3T in human, the crystal structure of H3t in mice demonstrated that Met71 and Val111 residues are not conserved in H3T mouse homologue. Furthermore, they revealed that H42 side chain of the H3t is less basic compared to equivalent R42 side chain of H3T in human, leading to a weaker electrostatic interaction with DNA entry-exit regions compared to H3.1. In addition, H3t in its histone fold shares a chaperone recognition motif (SAVM) with H3.1, H3.2 and H3T (Ueda, Harada et al., 2017).

2.3.3.2. H3.3

H3.3 is a replication-independent variant of the histone H3 gene family. In mice, H3.3 gene is encoded by two different genes, *H3f3a* and *H3f3b*, which generate the same H3.3 protein. The invalidation of each of these genes in mice models, revealed different consequences in these mice. *H3f3a*$^{-/-}$ mice lived until adulthood and male mice showed a decrease in the fertility rate and dysmorphic spermatozoa. Meanwhile, *H3f3b* mutants had growth-deficiency and died at birth (Tang, Jacobs et al., 2015). Other studies revealed, *H3f3b* disruption leads to abnormal chromatin organization, which inhibits proper protamine incorporation into the chromatin (Yuen, Bush et al., 2014). In somatic cells, H3.3 incorporation into chromatin is associated with a role of this specific histone variant at promoters and enhancers of active genes (Chen, Zhao et al., 2013). Interestingly, immunofluorescence analyses showed the disappearance of H3.1 and H3.2 in sex chromosomes during spermatogenesis. Meanwhile, measuring signal intensity by using an antibody which recognize the C-ter of H3.1, H3.2 and Not H3.3 variant, revealed the *de novo* H3.3 nucleosome incorporation on sex-chromosomes during MSCI (van der Heijden, Derijck et al., 2007). The functional significance of this phenomenon remains obscure. In mice, H3.3 massive incorporation into paternal genome is essential for the pericentric heterochromatin *de novo*

assembly during post-fertilization, which also involves the polycomb complex. The disruption of H3.3K27 accumulation at this step, leads to aberrant heterochromatin transcripts accumulation, perturbation of the proper chromosome segregation and a developmental arrest (Santenard, Ziegler-Birling et al., 2010). A combination of H3.3 and H2A.Z in nucleosomes of promoters and enhancers in stem cells, has been demonstrated to be an efficient combination for inducing chromatin instability and promoting transcriptional activity (Chen et al., 2013, Jin & Felsenfeld, 2007).

2.3.4. Histone H2A variants

Various histone H2A variants are expressed in testis. Among these H2A.X, H2A.Z, macroH2A and H2A.Bbd are ubiquitously expressed in tissues and TH2A, H2A.L.1, H2A.L.2, H2A.L.3 and H2A.B.3 are uniquely expressed in testis. Testis-specific H2A variants are further discussed in this section.

2.3.4.1. TH2A

TH2A is incorporated into the chromatin during meiosis. It exists at its highest level in pachytene spermatocytes and it gradually disappears around the end of elongating spermatids (Hoghoughi et al., 2018). MNase digestion of pachytene chromatin revealed, TH2A incorporation induce nucleosome instability (Rao & Rao, 1987). TP incorporation into the chromatin was deficient in *TH2A*$^{-/-}$ and *TH2B*$^{-/-}$ mice mutants and defective spermatids were generated in testes (Shinagawa et al., 2015). Structural analysis of TH2A-containing nucleosomes revealed a dramatic defect in histone-dimer DNA contacts due to the modified residues at N- and C- terminal ends of histone variants (Padavattan, Thiruselvam et al., 2017). Recently, mass spectrometry analyses in mice, identified TH2A-Thr 127 phosphorylation at the time of chromatin condensation in spermiogenesis and during early embryonic mitosis. Upon chromatin condensation, TH2A-Thr 127 phosphorylation is co-localized with H3.3 variant at TSSs of the genome and is removed immediately after fertilization. Interestingly, phosphorylated TH2A gets accumulated on maternal and paternal pericentromeric heterochromatin regions upon the early embryo divisions (Hada, Masuda et al., 2017).

2.3.4.2. H2A.L.1, H2A.L.2 and H2A.L.3

In my host laboratory, more than 10 years ago, a proteomic approach on condensing spermatids, allowed the discovery of mouse testis-specific H2A.Like 1, H2A.Like 2 and H2A.Like 3 histone variants. Among these 3 variants H2A.L.2 shares 70% homology with the conventional H2A. The antibody made against the newly-identified variants H2A.L.2 was not able to detect H2A.L.1 (Barral et al., 2017) . On the other hand, western blot analyses revealed the absence of H2A.L.3 either in germ cells or in whole testis. Thus, further investigations were carried out on H2A.L.1/2 variants. About 10 years later, the spatio-temporal H2A.L.2 and TPs co-expression in elongating spermatids were demonstrated (Barral et al., 2017). IF, WB and IH analyses showed that these variants first appear during step 9 in elongating spermatids and their expression increases in condensing spermatids. DNA FISH analyses using pericentric heterochromatin-specific probes, revealed the preferential H2A.L.2 localization to dapi-dense regions of chromatin, which here represent pericentric heterochromatin. IF analyses on de-condensed spermatozoa confirmed that H2A.L.2 persists in mature spermatozoa and is not replaced with protamines. However, it rapidly disappears from paternal heterochromatin after fertilization (Wu, Caron et al., 2008). Interestingly, H2A.L.2 showed preferential dimerization with TH2B in nucleosomes, which leads to nucleosome-instability (Govin et al., 2007).

Nucleosome H2A.L.2 incorporation is associated with the generation of unstable nucleosomes. Conventional H2A at its C-terminal harbor a “docking domain” which is responsible for H2A-H3 interaction within the nucleosome. In addition, H2A, harbors six acidic residues at the surface of the nucleosome referred to as “acidic patch”, which interact with basic residues at N-terminal of the H4 of the adjacent nucleosome and hence can mediate a chromatin compaction (Chodaparambil, Barbera et al., 2007, Zhou, Fan et al., 2007). On the other hand, the docking domain of H2A.L.2 is shorter and has different protein sequence compared to H2A (Syed, Boulard et al., 2009).

Molecular investigations on the role of H2A.L.2 involving *H2A.L.2*$^{-/-}$ mice, unraveled the importance of this histone variant in the process of histone-to-protamine transition. H2A.L.2 incorporation into nucleosome induces instability leading to a more “open” nucleosome structure. This allows nucleosome invasion by TPs, which themselves create a platform for in-coming pre-Prm2 and allowing the proper processing of this Prm. Mature Prms then win the competition with TPs/histones for DNA-binding. In the absence of H2A.L.2, TPs

accumulate on the nucleosomes and their incorporation is disrupted due to the “close” state of nucleosomes. This TPs accumulation on the nucleosomes, leads to a defective pre-Prm2 processing, pre-Prm2 accumulation on the nucleosomes and defective Prm incorporation inside the nucleosome. Electron microscopy analyses revealed that H2A.L.2 mutant mice have a defect in chromatin compaction and they are infertile (Barral et al., 2017).

2.3.4.3. H2A.Lap 1-4

An *in silico* study led to the identification of four H2A variant referred to as H2A.Lap1-4 (Lack of Acidic Patch). As mentioned previously, the acidic patch refers to a series of six acidic residues present on the surface of the nucleosome and allow the inter-nucleosomal interaction in order to facilitate chromatin compaction (Soboleva, Nekrasov et al., 2011). Further analysis revealed that some of these variants had been already identified in my host laboratory (Govin et al., 2007). Therefore, H2A.L.2 is H2A.Lap3; H2A.L.1 is H2A.Lap2; H2A.L.3 is H2A.Lap4 and H2A.Lap1 is now known as H2A.B.3 (El Kennani et al., 2017).

2.3.4.4. H2A.B.3

H2A.B.3 or (H2A.Lap1) mRNA expression in initiated at day 19 ppt at pachytene spermatocyte stage and is highly enriched in round spermatids. The incorporation of H2A.B.3 as well as its human homolog H2A.Bbd, into the nucleosome is associated with the creation of unstable nucleosomes, harboring only 120bp of DNA (Montel, Menoni et al., 2009, Soboleva et al., 2011). Lack of acidic patch at the surface of the nucleosomes bearing H2A.B.3, weaken its capacity of mediating chromatin compaction. In elongating spermatids, this variant is located in cytoplasm probably because if its replacement by H2A.L.2, before genome-wide histone-protamine exchange. Interestingly, in round spermatids, H2A.B.3 is mainly located in euchromatin, but it is also located at inactive heterochromatic X and Y chromosomes. CHIP-seq analyses revealed the enrichment of H2A.B.3 at the TSS of highly active genes in late pachytene and round spermatids, suggesting its role in gene activation. Furthermore, H2A.B.3 is linked to the genes which escape X chromosome inactivation in late round spermatids (Soboleva et al., 2011).

In addition to TSS regions of active genes, ChIP-seq analyses revealed the association of H2A.B.3 with intron-exon boundary within the gene body of active genes in testis as well as in brain. In contrast to previous data, these ChIP-seq analyses revealed the negative correlation between H2A.Z and transcription, and demonstrated the intron-exon boundary sites are depleted from H2A.Z. Interestingly, RNA-binding gel-shift assays and RNA pull down analyses, revealed that H2A.B.3 was associated with RNA, through its 6 arginine-containing RNA binding motif located at the N-terminal of the protein. The link between H2A.B.3 and RNA is highly correlated with gene transcription. RNA IP analyses further revealed the association between H2A.B.3 and the intron-exon boundaries and the splicing speckles (Soboleva, Parker et al., 2017).

Various combinations of H2A, H2A.Z and H2A.B.3 have been proposed near or within the TSS regions of active genes. Recent work did not allow to link H2A.Z-containing nucleosomes to the gene transcription (Soboleva et al., 2017). In addition, H2A.Z genome mapping with H2A.B.3 revealed a link between these two H2A variants. This H2A variants combination could be defined as a mark of specific pattern of gene expression among a subset of genes. Additionally, these histone variants could directly regulate RNA, RNA processing and splicing.

3. Paternal transgenerational epigenetic inheritance

Transmission of Genomic DNA from one generation to another defines the basis of transgenerational inheritance, which is essential for the maintenance and transfer of our genetic heritage to other generations. Meanwhile, epigenetic inheritance through sperm, such as transmission of covalent DNA modifications, persistence of nucleosomes, specific histone variants and histone marks, chromatin architecture and small RNAs, is essential and greatly contributes to the completion of this transmission. Here, the transmission of some specific histones (nucleosome retention) and small RNAs through sperm, and their potential role in early embryo development, are further discussed.

3.1. Nucleosome retention in mature sperm

During spermiogenesis, male genome is packed with Prms, but to some extent, certain fractions of histones remain in the mature sperm and therefore they are transmitted to the next generation and could provide a possibility for epigenetic inheritance. About 10 years ago, we thought histones compose 10-15% of the male genome in mature human sperm (van der Heijden, Ramos et al., 2008), in mice only 1% (or 8-10%) of the mature sperm genome is composed of histones (Erkek et al., 2013, Jung et al., 2017, Yoshida, Muratani et al., 2018). The early works showed that in human nucleosomes remaining in sperm genome are highly enriched at promoters of genes that are involved in embryonic development such as *HOX* gene clusters as well as imprinted genes and microRNA clusters (Hammoud, Nix et al., 2009). CHIP-seq experiments in human and mouse spermatozoa, revealed the enrichment of H3K27me3, a modification related to transcriptional repression is enriched near TSS of regulatory genes with developmental function in early embryo. These observations suggested the propagation of this repressive mark through polycomb complex 2, in order to maintain the state of totipotency in early embryo stages of development. Additionally, these experiments also detected H3K4me2 enrichment in testis-specific and housekeeping genes, but their functional significance remains to be investigated (Brykczynska, Hisano et al., 2010). Further investigations in mouse sperm, highlighted the specific localization of residual histones in unmethylated CpG-rich sequences. In addition, Chip-Seq and RNA-seq analyses demonstrated also a high enrichment of H3.3 at TSS regions in mouse sperm suggesting a role for histone-variants in these cells (Erkek et al., 2013).

Other investigations involving high throughput sequencing of the MNase digested mono-nucleosomes in human and bovine sperm, surprisingly showed nucleosome retention at intergenic and gene-poor regions, specifically in centromere repeats and retrotransposons such as LINEs and SINEs (Samans, Yang et al., 2014). In this study they focused on octamer nucleosomes and 146bp length of DNA. Later in the same year, another study by using an advanced nucleosome mapping technique (paired-end deep sequencing of an entire MNase digestion ladder) (Henikoff, Belsky et al., 2011), confirmed the nucleosome enrichment in gene poor regions. In contrast to the first reported works, this study shows that nucleosomes are generally depleted from the developmental promoters (Carone, Hung et al., 2014). A recent study, by examining the quality of sperms obtained from swim-up technique (the experiment

in which cauda epididymis is maintained in a 5cm tube in appropriate medium at 37°C, in order for spermatozoa to swim up), demonstrated that the swim up fraction contains the "histone replacement-uncompleted sperms" (Yoshida et al., 2018). This observation suggested that most of the nucleosomes found in all of the reported analyses are not yet completely replaced by protamines. Therefore, these authors developed a method in which they isolated only the fraction of sperms with complete histone replacement. Chip-Seq analyses on the histone replacement completed sperm cells, seem to confirm the early reports and point to histone retention on the promoters of specific genes, important for early embryo development.

3.2. Non-coding RNA populations

In addition to DNA provided by both male and female gametes, other information and factors are also transferred to the embryo during fertilization and could contribute to embryogenesis. Among these factors, different types of RNAs are transmitted from sperm cell to the offspring. Small RNAs including miRNAs, endo-siRNAs and piRNAs are crucial for mammalian development. Expression timing of each of these RNAs are different throughout the development. For instance, piRNAs are expressed during spermatogenesis. Endo siRNAs are crucial for oocyte meiosis and miRNAs are ubiquitously expressed in somatic tissues. Additionally, in mammals, sperm carries long as well as small non-coding RNAs such as piRNAs and microRNAs and they have been recently identified as the post-transcriptional regulators during development (Figure 24) (Suh & Blelloch, 2011). miRNAs have been shown to be crucial for the spermatogenesis (Bouhallier, Allioli et al., 2010, Yan, Dong Xda et al., 2009). For instance, Dicer mice mutants revealed a defect in spermatozoa production (Hayashi, Chuva de Sousa Lopes et al., 2008). On the other hand, piRNAs are also essential for spermatogenesis. Based on their timing of expression and the piwi proteins that they are associated, piRNAs have been separated into two groups: repetitive and non-repetitive (Aravin, Gaidatzis et al., 2006). Repetitive class of piRNAs are expressed before meiotic pachytene and the second non-repetitive class of piRNAs are expressed during the pachytene stage (Aravin & Bourc'his, 2008). Repetitive piRNAs have been shown to be associated with the repression of transposable elements in spermatogenesis and they enhance *de novo* DNA demethylation of the transposons (Aravin & Bourc'his, 2008, Malone & Hannon, 2009).

Interestingly, a recent study by characterizing the developmental dynamics of small RNAs in multiple purified sperm populations from testis and post-testicular maturation, demonstrated that first, the majority of RNA contents of mature sperm cells are tRNAs. Second, *in vitro* and *in vivo* tRNAs tracking from epididymis to sperm cells, unraveled that these tRNAs, which are generated in somatic cells of epididymis are transferred to the germ cells through the epididymosomes from epididymis to mature sperm. During the maturation of sperm cells, a drastic change of small RNAs content from piRNAs to tRNAs takes place between testicular sperms and the epididymis sperm cells (Sharma, Sun et al., 2018). These studies highlight the importance of sperm small RNA content for early development and also shed light on how environmental information is transferred from somatic to germ cells.

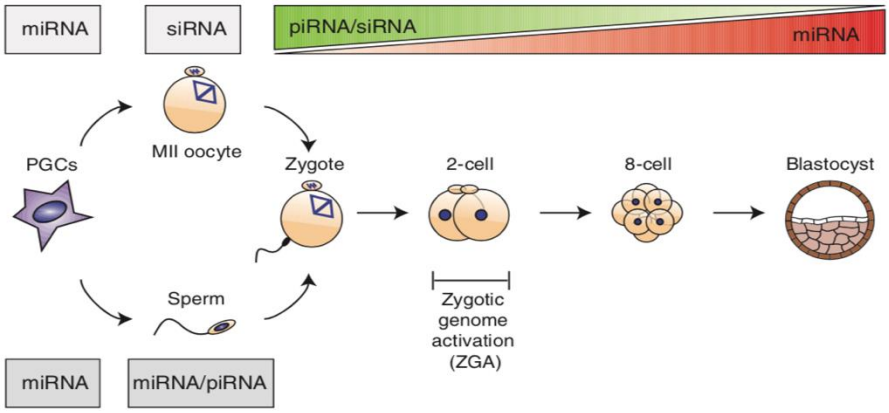


Figure 24: Schematic representation of differential small RNAs in sperm and during early embryonic development. Taken from (Suh & Blelloch, 2011).

Upon fertilization the maternal and paternal chromatin undergo a series of rearrangements for the zygote to totipotent. Paternal chromatin undergoes a genome-wide protamine-to-histone exchange as well as an active DNA demethylation, which is visible by immunofluorescence and in every loci (Beaujean, Taylor et al., 2004, Mayer, Niveleau et al., 2000). However, keeping a memory of the parental origin of the genome, is essential for imprinting and dosage compensation (Surani, 2001). Additionally, the parental memory is a key to defining functional regions of chromatin, for instance constitutive heterochromatin in embryo. These functional regions are located near the centromeres and are essential for chromosome segregation (Probst, Dunleavy et al., 2009). During early embryonic

development, paternal pericentric domains lack pericentric marks such as H3K9me3 and HP1α. In fission yeast and in plants the transcription of pericentric heterochromatin has been shown essential for the chromatin formation and transcriptionally silent state (Verdel, Jia et al., 2004, Zaratiegui, Irvine et al., 2007). Various non-coding RNAs have been shown to be implicated in silencing of chromatin domains during dosage compensation and genomic imprinting. 3D FISH analyses demonstrated that maternal pericentric regions bear silencing marks such as H3K9me3, H4K20me3 and HP1beta, whereas paternal pericentric regions only bear HP1beta mark (Probst, Santos et al., 2007). RNA FISH analyses, on zygote, 2-cell, 4-cell and 8-cell stage embryo, revealed an expression burst of forward and reverse pericentric transcripts from 2-cell until 4-cell stage followed by a subsequent decrease after the formation of chromocenters. Interestingly, an attempt to inhibit the pericentric major satellite by DNA-LNA gapmers injection caused developmental arrest in early embryo stages. Therefore, the combined expression of forward and reverse major transcripts during early embryonic development, suggest an important role for the regulation and establishment of pericentric heterochromatin domains of paternal chromatin for ensuring the subsequent proper cell divisions (Probst et al., 2010).

Although many studies in mammals, have unraveled the importance of small and long non-coding RNAs in early embryonic development, the questions still remain regarding the transmission of pericentric satellites and their concrete functional and mechanistic importance for offspring, is yet to be defined.

Thesis Objectives

Mammalian Spermatogenesis represents an exceptional model for studying chromatin organisation and genome reprogramming. Understanding the specific step-by-step molecular mechanisms underlying this extreme reprogramming, could greatly advance our knowledge, first, on the spermatogenic cells' chromatin reprogramming and second, on the general comprehension of chromatin remodeling in other physiopathological contexts. During the last decade, my host laboratory developed tools and suitable animal models, which led to the identification of testis-specific factors that are at the heart of this genome-wide chromatin reprogramming.

Data from our laboratory show that a wave of genome-wide histone hyperacetylation initiates this reprogramming in post-meiotic (elongating) spermatids. Following this genome-wide histone hyperacetylation, during the final stages of spermatogenesis, an almost complete histone eviction and their replacement by Prms occur. My host laboratory recently identified the essential role of two testis-specific factors NUT and histone variant H2A.L.2 in the genome-wide histone hyperacetylation and in histone eviction in elongating and condensing spermatids, respectively. Additionally, the functional interaction between NUT and p300 enhances the HAT activity of this enzyme leading to H4 histone hyperacetylation and preparation for histone replacement (Shiota et al., 2018). The incorporation of H2A.L.2 subsequently to the function of NUT and p300, prepares nucleosomes for TP invasion and proper Prm processing and assembly (Barral et al., 2017). As observed previously (Govin et al., 2007), a fraction of H2A.L.2 remains associated to the pericentric heterochromatin regions in the mature spermatozoa and is not replaced by Prms. Interestingly, these two last studies found the presence of sub-nucleosomal structures after extensive MNase digestion during nucleosome-to-nucleoprotamine transformation.

My work aimed first:

- To investigate the detailed mechanistic insight into the activation of p300 with the aim to better understand its specific activity in histone hyperacetylation in elongating spermatids.

and Second:

- To characterize the impact of H2A.L.2-containing sub-nucleosomal structures on the final spermatozoa chromatin organisation, and decipher the mechanisms

underlying the specific H2A.L.2's association with the pericentric heterochromatin regions in the mature spermatozoa

Results

I. The molecular basis of the p300 acetyltransferase activation

As previously reported by my host laboratory (Shiota et al., 2018), NUT enhances the HAT activity of p300 especially on histone H4 lysines 5 and 8 residues in post-meiotic spermatids. Moreover, histone hyperacetylation is associated with the genome-wide histone removal in elongating spermatids. Therefore, deciphering the molecular mechanisms of p300 activation is the first step towards the understanding the regulatory mechanisms underlying the genome-wide histone hyperacetylation.

This work reports for the first time, the link between transcription factors, p300 activation and chromatin modification. Here we show that transcription factor dimerization leads to the trans-autoacetylation of a lysine-rich auto-inhibitory loop (AIL), within the HAT catalytic domain of p300. This AIL serves as a pseudosubstrate inhibiting the access of substrates to the p300 catalytic site. Once p300 is activated (AIL acetylated) and histones acetylated, it recognizes acetylation on adjacent histones and propagate histone acetylation.

This work also illustrates how transcription factors, the proteins that recognize a specific DNA sequence and determine where and when genes are regulated, control histone acetylation. Histone acetylation is specifically initiated at the level of a DNA sequence recognized by transcription factors, which explains why the DNA of active genes is also specifically associated with acetylated histones. However, in the case of spermatogenesis p300 activation is ensured by Nut, which is not a transcription factor. Taken together, these investigations bring important indications on how Nut may activate histone acetylation in spermatids and how this enzyme can ensure histone acetylation propagation, a phenomenon, which leads to H2A.L.2 incorporation and nucleosome transformation.

During my PhD, I participated to the functional experiments of this project mostly reported in Figure 4 of the article and some supplementary figures.

ARTICLE

https://doi.org/10.1038/s41586-018-0621-1

Transcription factor dimerization activates the p300 acetyltransferase

Esther Ortega[1], Srinivasan Rengachari[1,5], Ziad Ibrahim[1,2], Naghmeh Hoghoughi[3], Jonathan Gaucher[1,6], Alex S. Holehouse[4], Saadi Khochbin[3] & Daniel Panne[1,2]*

The transcriptional co-activator p300 is a histone acetyltransferase (HAT) that is typically recruited to transcriptional enhancers and regulates gene expression by acetylating chromatin. Here we show that the activation of p300 directly depends on the activation and oligomerization status of transcription factor ligands. Using two model transcription factors, IRF3 and STAT1, we demonstrate that transcription factor dimerization enables the *trans*-autoacetylation of p300 in a highly conserved and intrinsically disordered autoinhibitory lysine-rich loop, resulting in p300 activation. We describe a crystal structure of p300 in which the autoinhibitory loop invades the active site of a neighbouring HAT domain, revealing a snapshot of a *trans*-autoacetylation reaction intermediate. Substrate access to the active site involves the rearrangement of an autoinhibitory RING domain. Our data explain how cellular signalling and the activation and dimerization of transcription factors control the activation of p300, and therefore explain why gene transcription is associated with chromatin acetylation.

Signals that emanate from cellular receptors ultimately lead to changes in gene expression that drive cellular change and organismal development. Gene expression is typically controlled through the coordinated activity of DNA-binding transcription factors, chromatin regulators and the general transcription machinery. For instance, in the innate immune system, pattern recognition receptors recognize and engage with various pathogen-associated molecular patterns[1], and subsequently bind to adaptor proteins such as STING (stimulator of interferon genes). These adaptor proteins engage the latent DNA-binding transcription factor interferon (IFN) regulatory factor 3 (IRF3) and enable recruitment and activation of the non-canonical IκB kinase TBK1[1]. TBK1 then phosphorylates IRF3 in a C-terminal motif, resulting in the removal of autoinhibition, dimerization and adaptor displacement[2,3]. Activated IRF3 dimers bind to p300/CBP (where CBP is CREB-binding protein; p300 and CBP are also known as KAT3B and KAT3A, respectively) to stimulate chromatin acetylation and gene expression of the antiviral type I IFNs IFNα and IFNβ[3–5]. Type I IFNs are secreted and bind to specific cell-surface IFN receptors, which results in the activation of Janus kinase–signal transducers and activators of transcription (JAK–STAT) signalling[6]. The activated, tyrosine-phosphorylated STATs then dimerize, translocate to the nucleus and bind to p300/CBP to stimulate the transcription of IFN-stimulated genes[7].

p300/CBP are known to interact with more than 400 binding partners including the basal transcription machinery[8]. The large protein interactome of p300/CBP results in near-universal recruitment of these HATs to enhancers, and p300 occupancy has been used to identify enhancers genome-wide[9,10]. p300/CBP catalyses the acetylation of histone H3K27 to form H3K27ac, a modification that is considered an 'activation' mark[11]. However, recruitment of p300/CBP does not always correlate with gene activation and is occasionally associated with repression[12–16]. A large number of chromatin regions that bind p300/CBP therefore do not contain this canonical H3K27ac modification, which indicates that HAT activity at such sites is blocked[15,17]. Therefore, it is a major challenge to understand the mechanism that enables switching between inactive and active states of p300/CBP on enhancers, and to causally link cellular signalling to the recruitment of p300/CBP, the regulation of HAT activity and the establishment of repressed, poised and active chromatin.

Here we have investigated how the activation and oligomerization status of p300 transcription factor ligands such as IRF3 and STAT1 affects the catalytic activity of p300. We found that the kinase-activated and dimeric, but not the inactive or monomeric, variants of these transcription factors support robust p300 HAT activation. We demonstrate that transcription factor dimerization enables p300 *trans*-autoacetylation in a lysine-rich, intrinsically disordered autoinhibitory loop (AIL) in the HAT domain that serves as a 'pseudosubstrate' and is important for regulating the HAT activity of p300[18]. A crystal structure of the core domain of p300 provides a snapshot of a potential *trans*-autoacetylation reaction intermediate in which the AIL projects into the active site of a neighbouring p300 molecule. As HAT activation is closely linked to transcription factor activation, these results causally relate cellular signalling to the activation and DNA targeting of a chromatin modifier and provide mechanistic insights into the long-standing and general correlation between an active, acetylated chromatin structure and gene transcription.

Transcription factor dimerization activates p300

To explore whether p300 is activated by signal-dependent IRF3 dimerization, we produced three recombinant IRF3 species: inactive monomers (IRF3); active, TBK1-phosphorylated IRF3 dimers (pIRF3); and a truncation mutant that lacked the C-terminal autoinhibitory element (IRF3ΔC) (Extended Data Fig. 1a, b). Truncation of the C-terminal autoinhibitory element allows for p300/CBP binding but abolishes IRF3 dimerization[19]. We confirmed the oligomerization status by gel filtration chromatography (Extended Data Fig. 1b), and investigated the effect of IRF3 activation and oligomerization status on the autoacetylation of p300s in the presence of [^{14}C]acetyl coenzyme A

[1]European Molecular Biology Laboratory, Grenoble, France. [2]Leicester Institute of Structural and Chemical Biology, Department of Molecular and Cell Biology, University of Leicester, Leicester, UK. [3]CNRS UMR 5309, INSERM U1209, Université Grenoble Alpes, Institute for Advanced Biosciences, Grenoble, France. [4]Department of Biomedical Engineering and Center for Biological Systems Engineering, Washington University in St. Louis, St. Louis, MO, USA. [5]Present address: Max Planck Institute for Biophysical Chemistry, Department of Molecular Biology, Göttingen, Germany. [6]Present address: Université Grenoble Alpes, INSERM U1042, HP2 Laboratory, Grenoble, France. *e-mail: daniel.panne@le.ac.uk

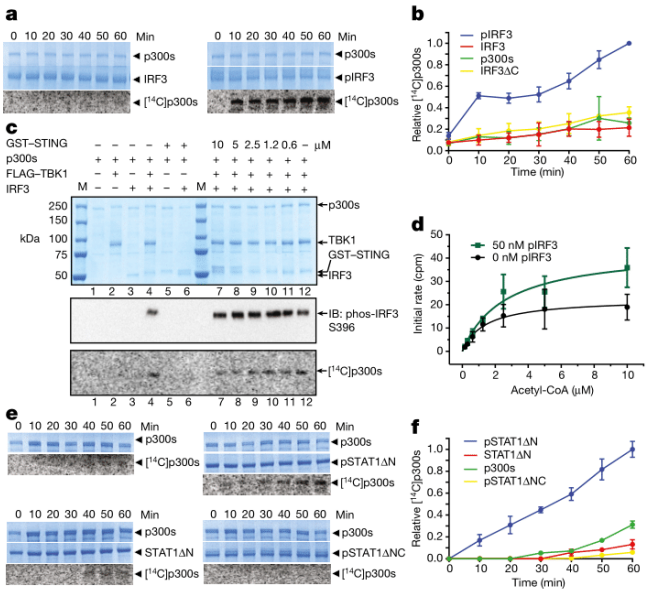


Fig. 1 | Transcription factor dimerization activates p300. a, p300s was incubated for the indicated times in the presence or absence of inactive, monomeric IRF3 or TBK1-phosphorylated, dimeric pIRF3. Samples were analysed by SDS–PAGE followed by Coomassie staining and autoradiography. Representative data of three independent experiments are shown. **b**, Quantification of the autoacetylation of p300s. **c**, p300 is activated by TBK1-mediated IRF3 phosphorylation. p300s was incubated with recombinant GST–STING, TBK1 and IRF3 in the presence of ATP and [^{14}C]acetyl-CoA. Top, Coomassie-stained SDS–PAGE gel. Middle, analysis of IRF3 phosphorylation on S396 using immunoblotting. Bottom, autoradiography. Representative data of three independent experiments are shown. **d**, HAT scintillation proximity assay. The degree of histone H4 substrate acetylation was quantified using scintillation counting. **e**, As in **a** but using inactive, monomeric STAT1ΔN or activated, dimeric pSTAT1ΔN. Activated, dimeric pSTAT1ΔNC that lacks the C-terminal TAD did not stimulate the autoacetylation of p300s. Samples were analysed as in **a**. Representative data of three independent experiments are shown. **f**, Quantification of the autoacetylation of p300s. Intensity values were normalized by dividing by the maximum autoacetylation signal obtained after 60 min. Error bars shown in **b**, **d** and **f**: three independent experiments were performed, data are mean ± s.d. Data analysis and plotting was performed with GraphPad Prism 7.0. For gel source data, see Supplementary Fig. 1.

(acetyl-CoA). p300s is a short p300 construct that spans from the TAZ1 to the NCBD/IBiD domains and contains a deletion of the flexible N- and C-terminal regions (Extended Data Fig. 4a). p300s autoacetylated slowly in the absence of IRF3 (Extended Data Fig. 1c). The inclusion of inactive, monomeric IRF3 or IRF3ΔC did not modify HAT activity (Fig. 1a, Extended Data Fig. 1c). By contrast, the inclusion of active, TBK1-phosphorylated IRF3 dimers (pIRF3) resulted in a rapid burst of autoacetylation followed by a gradual increase of acetylated p300s (Fig. 1a, b). As IRF3ΔC did not support p300 HAT activation, IRF3 dimerization and not solely p300 binding is essential for HAT activation.

p300 HAT activation was directly dependent on the TBK1-mediated phosphorylation of IRF3 on Ser396, a critical residue for IRF3 activation and dimerization[2,3]. Only when both TBK1 and IRF3 were included in the reaction did we observe p300 activation (Fig. 1c, lane 4). We observed only a modest stimulatory effect of the adaptor protein STING (Fig. 1c, lanes 7–12), probably owing to the relatively large amounts of TBK1, which is already active and phosphorylates IRF3 even in the absence of STING[20]. We conclude that IRF3 phosphorylation by TBK1 and its dimerization are required for p300 HAT activation.

To analyse the effect of pIRF3 on p300 activation and acetylation of the histone substrate, we established a scintillation proximity HAT assay. We incubated saturating amounts of a biotinylated histone H4 substrate peptide with p300s in the absence or the presence of equimolar pIRF3 and increasing concentrations of [^{3}H]acetyl-CoA (Fig. 1d). pIRF3 stimulated p300 histone-substrate acetylation, as determined by the increased rate of H4 acetylation obtained in the presence of pIRF3 ($V_{max} = 43.8 \pm 5.3$ cpm min^{-1} as compared to $V_{max} = 22.5 \pm 2.8$ cpm min^{-1} in the absence of pIRF3). These data indicate that pIRF3 not only stimulates p300 autoacetylation and activation, but also stimulates more efficient acetylation of the histone substrate.

We also investigated the effect of STAT1 on p300 activation. STATs are activated in response to cytokine receptor engagement and Janus kinase activation[21]. JAK-mediated phosphorylation of STAT1 on Tyr701 induces dimerization and translocation to the nucleus, where STAT1 binds to DNA elements to regulate gene expression. STAT1 contains a C-terminal transactivation domain (TAD) through which it interacts with p300/CBP[7]. A naturally occurring splice variant, STAT1β, lacks the TAD and acts in a dominant negative manner[22]. Structures of the active, STAT1 Tyr701-phosphorylated dimer bound to DNA, as well as the STAT1 TAD bound to the TAZ2 domain of CBP, have been determined previously[23,24].

To understand the effect of STAT1 activation and oligomerization status on p300 activity, we produced STAT1ΔN that lacked the N-domain and STAT1ΔNC lacking the N-domain and the TAD as non-phosphorylated monomers or as Tyr701-phosphorylated dimers (Extended Data Fig. 1e–h). We found that p300s autoacetylated slowly in the absence of STAT1, and that the addition of non-phosphorylated, monomeric STAT1ΔN did not stimulate p300s autoacetylation beyond background levels (Fig. 1e, f). By contrast, the addition of Tyr701-phosphorylated STAT1ΔN (pSTAT1ΔN) dimers to p300s resulted in a rapid increase of p300 autoacetylation. Activation required the C-terminal TAD of STAT1, as the addition of a Tyr701-phosphorylated STAT1 dimer (pSTAT1ΔNC) that lacked the TAD did not stimulate p300 autoacetylation (Fig. 1e, f).

STAT1 dimerization, and not solely interaction with the TAZ2 domain, is required for the activation of p300. Unphosphorylated, monomeric STAT1ΔN, which contains the TAD and is able to interact with the TAZ2 domain of CBP, did not stimulate p300 activity. However, stimulation with STAT1 was not as potent compared with that of IRF3, possibly because our STAT1 preparation is unphosphorylated on Ser727, which is required for maximal gene activation[25]. Together, our data are consistent with a model in which the AIL peptide serves as an intramolecular 'pseudosubstrate' and a competitive HAT inhibitor[18]. Dimeric ligands such as pIRF3 and pSTAT1 allow p300 activation by bringing two molecules together to enable *trans*-autoacetylation of the AIL, which in turn relieves autoinhibition and enables more efficient entry of substrates into the HAT active site.

Structure of p300 adopts an AIL–swap conformation

To further understand the role of the AIL in the regulation of these structural transitions, we crystallized the hypoacetylated form of the catalytic core of p300 comprising the bromo-RING-PHD-HAT domains (BRP-HAT) that contained the AIL. Crystals were obtained using a similar protocol as published previously[26]. Crystals diffracted to a minimal Bragg spacing of 3.1 Å and we determined the structure by molecular replacement (Extended Data Table 1). The crystal form contained four p300 molecules in the asymmetric unit (Extended Data Fig. 2). Comparison with our previous structure[26] showed that the bromo-PHD-HAT domains overlay well on each other with a root-mean-square deviation (r.m.s.d.) of approximately 1 Å. However, the RING domains were not visible in the initial electron density map. Anomalous difference density maps showed a density peak for the zinc atom of the RING domain, but it was not at the expected location. Manual repositioning enabled the correct placement of the RING domains into the new position and the refinement of the structure (Fig. 2a, Extended Data Fig. 3).

The p300 molecules show an antiparallel arrangement of the BRP-HAT domains (Extended Data Fig. 2a). In this configuration, the HAT

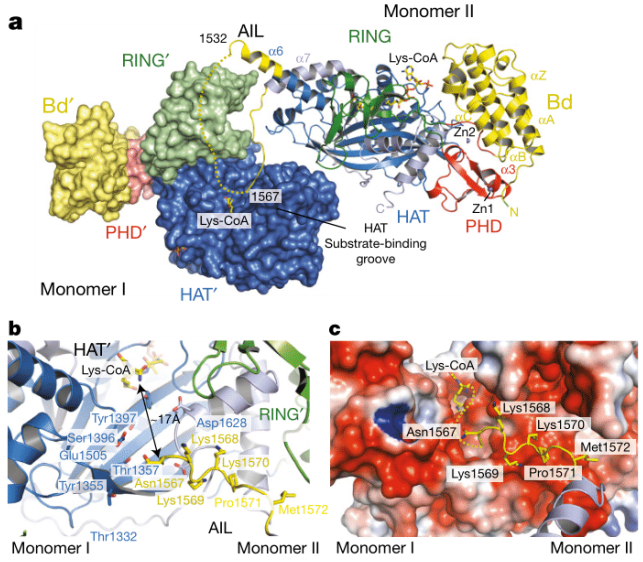


Fig. 2 | The structure of p300 adopts a AIL-swap conformation. **a**, Monomer I is surface-rendered and monomer II is shown as a cartoon. The AIL loop from monomer II is shown in yellow. The AIL lies near the HAT substrate-binding groove of monomer I. A disordered segment of the AIL is shown as a dotted line. **b**, Close-up view of the residues of the AIL loop from monomer II and residues of monomer I in the substrate-binding pocket. **c**, Binding of the positively charged AIL is mediated by interactions with negatively charged residues in the substrate-binding pocket of the HAT.

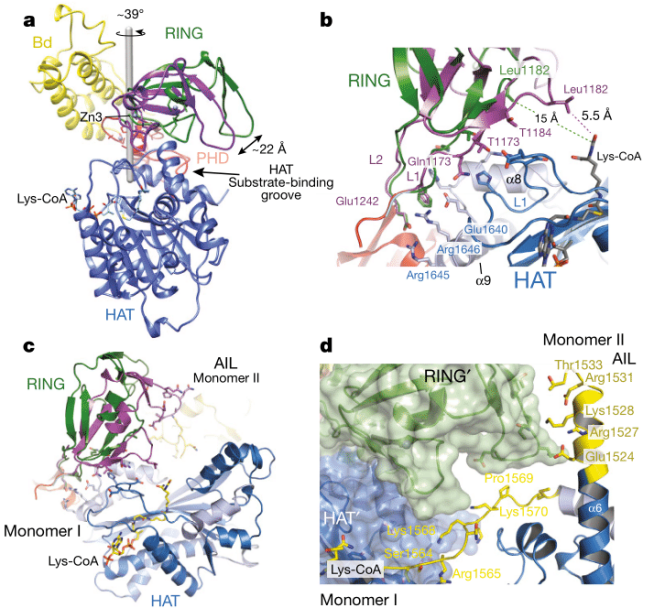


Fig. 3 | Structural rearrangement of the RING domain. **a**, The RING domain (green) rotates by approximately 39°, resulting in a displacement of about 22 Å away from the active site. The rotation axis is indicated as a grey rod. **b**, In the loop-swap conformation, residues in the RING–HAT interface are disrupted resulting in a more open HAT active site. Leu1182 is positioned approximately 15 Å away from the Lys-CoA inhibitor in the loop-swap conformation (green) but within 5.5 Å in the absence of the loop swap (magenta). **c**, Repositioning of the RING domain enables the AIL from monomer II to approach the active site of HAT on monomer I. **d**, Details of the interaction surface of the AIL from monomer II with the RING domain of monomer I.

domains from two neighbouring molecules are closely apposed (Fig. 2a). In all protomers, AIL residues 1520–1532 adopt a helical extension of α6 which packs against the outwardly rotated RING domain of the neighbouring protomer (Fig. 2a). In monomer II, residues 1566–1581 extend away from the HAT domain and associate with the substrate-binding pocket of the HAT domain in monomer I, approximately 17 Å away from the lysine substrate binding tunnel (Fig. 2b). The remainder of the AIL (residues 1532–1564) is disordered. In this conformation, positively charged residues K1568, K1569 and K1570 project towards the highly electronegative substrate-binding pocket of the HAT domain in monomer I (Fig. 2c). Analysis by size-exclusion chromatography–multi-angle laser-light scattering (SEC–MALS) revealed that p300 is monomeric at low micromolar concentrations (see Extended Data Fig. 6), which suggests that the AIL loop-swapped interactions do not appear to mediate the formation of stable dimers, but may instead constitute more transient self-associations. Although the AIL is clearly flexible and the electron density over the exchanged region is not visible in all protomers (Fig. 2b, c, Extended Data Fig. 2b, c), this arrangement supports the interpretation that, at high concentrations and when in close proximity to each other, two p300 monomers can engage each other through an AIL loop-swap.

Structural rearrangement of the RING domain

We previously proposed that active-site restriction by the RING domain is a negative regulatory mechanism for HAT activity[26]. A restricted active site is predicted to reduce the probability of substrates engaging with the active site by random diffusion, and could thus be important in enabling the regulation of acetylation by substrate recruitment. Consistent with this model, mutations that map to the structural framework that holds the RING domain in place result in HAT activation in cells[26]. In our current structure the RING domain rotates by around 39° away from the HAT active site, which results in an overall displacement by around 22 Å as compared to the previously determined structure that lacks the AIL (Fig. 3a). The axis of rotation is located perpendicular to the flexible loops L1 and L2 that connect the RING to the PHD domain.

The inward-rotated conformation (magenta in Fig. 3a) is stabilized by interactions between Glu1242 of the RING domain and Arg1645 and Arg1646 of helix α9 of the HAT domain. In addition, Gln1173, Thr1174 and Thr1184 of the RING domain pack against the unusually long loop (L1) in the HAT domain that covers the CoA portion of the Lys-CoA inhibitor. As a result, Leu1182 resides within about 5.5 Å of the lysine moiety of Lys-CoA (Fig. 3b). This inward conformation of the RING domain thus restricts substrate access to the HAT domain: the incoming AIL from the neighbouring p300 monomer II would clash with the RING domain in the inward conformation (Fig. 3c).

In the outward-rotated conformation, the interactions that attach the RING domain to the HAT domain are mostly disrupted (Fig. 3b). Leu1182 is positioned around 15 Å away from the substrate-binding site and the RING domain is cradled by the AIL extension of helix α6 of the neighbouring p300 molecule (monomer II residues 1524–1533; Fig. 3d). Despite shape complementarity, with a small buried surface area of about 320 $Å^2$, the interface is predominantly polar, which is uncharacteristic of a typical protein–protein interface. However, this interaction could help to stabilize an outward-rotated conformation of the RING domain and a more open active site of the HAT, apparently to enable access of the AIL and *trans*-acetylation.

Regulation of HAT activity by flanking domains

To systematically analyse the flanking domains, we generated a series of p300 constructs (Extended Data Fig. 4a) and analysed the effect on HAT activity in vitro and in cells. Overexpression of p300 generally resulted in hyperacetylated, active p300 variants (Extended Data Fig. 4b, c), which probably masks the functional role of structural elements potentially involved in autoinhibition of deacetylated p300. Deletion of the RING domain did not considerably alter autoacetylation or histone acetylation (Extended Data Fig. 5a). This deletion did

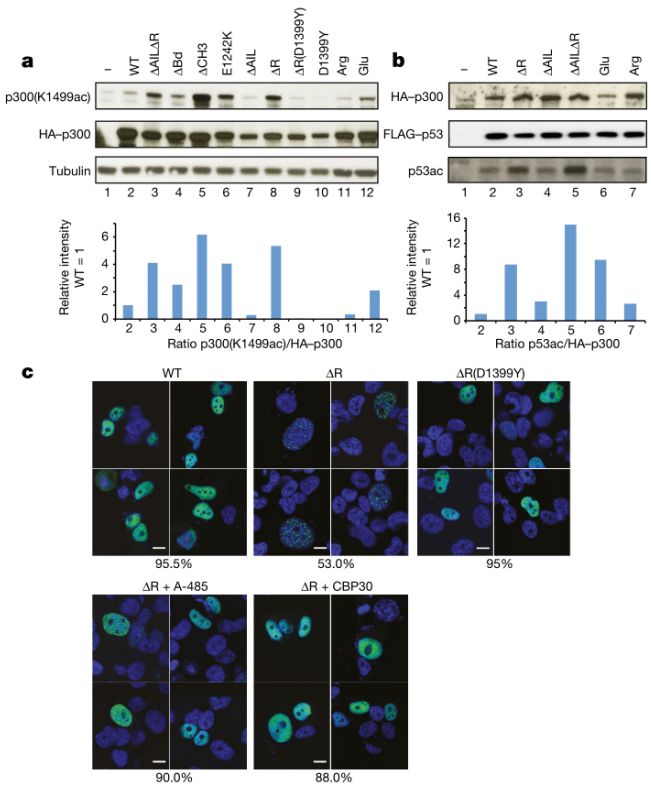


Fig. 4 | Regulation of HAT activity by flanking domains. a, Top, indicated variants of p300 were transiently co-transfected with p53 in COS cells and samples were analysed by western blotting using the indicated antibodies. Bottom, quantification of the p300(K1499ac) signal. WT, wild type. **b**, Top, analysis of p53 acetylation. Bottom, quantification of the p53 acetylation signal. Representative data of three independent experiments are shown. For details of the mutants, see Extended Data Fig. 4a. Arg and Glu indicate proteins in which lysine amino acids in the AIL segment spanning amino acids 1546–1570 were mutated to arginine or glutamate, respectively. **c**, Top, H1299 cells were transfected with the indicated constructs and analysed by immunofluorescence using anti-HA for p300 (green); cell nuclei were stained with Hoechst (blue). Bottom, cells were treated with the A-485 HAT or the CBP30 bromodomain inhibitor. The percentage of cells showing the indicated phenotype ($n = 200$ cells) is indicated below each panel. Scale bars, 10 μm. For gel source data, see Supplementary Fig. 1.

not adversely affect the structural integrity of p300, as shown by a crystal structure of the BΔRP module containing this deletion (Extended Data Fig. 5c).

Deletion of the AIL (ΔAIL) in all constructs resulted in decreased histone acetylation, but bromodomain deletion (ΔBd) did not affect HAT function (Extended Data Fig. 5a, b). Together, our results are consistent with previous observations of CBP that—at least in the active, hyperacetylated state of the enzyme—RING deletion does not substantially affect HAT activity and that the p300 AIL positively contributes to substrate acetylation[27]. We next introduced mutations into full-length p300 and monitored their effect on p300 autoacetylation and p53 acetylation upon transient co-overexpression in cells. Deletion of the RING (ΔR) and CH3 domains resulted in markedly increased p300 autoacetylation and p53 acetylation, but deletion of the bromodomain or AIL had no major effect (Extended Data Fig. 5e). As expected, introduction of the catalytic mutants D1399Y or Y1467F abolished p300 autoacetylation or p53 acetylation (Extended Data Fig. 5e). Immunofluorescence analysis showed that wild-type p300 as well as a ΔBd and ΔAIL deletion were uniformly distributed in the nucleus, but that the HAT-activating p300 variants ΔR and ΔCH3 formed nuclear foci that co-localized with p53 (Extended Data Fig. 5d). To validate these results, we analysed and confirmed the phenotype of p300 mutants and p53 acetylation in another cell line (Fig. 4a, b). In addition, we analysed p300 variants in which eleven lysine amino acids (spanning amino acids 1546–1570 of the AIL) were mutated to arginine or glutamate, and found reduced or slightly increased p300 autoacetylation or p53 acetylation levels, respectively (Fig. 4a, b).

As we observed the formation of nuclear foci only with HAT-activating variants, we proposed that hyperacetylation drives p300 to form biomolecular condensates in cells. Accordingly, the introduction of a HAT-inactivating D1399Y mutation into p300 ΔRING, treatment with the p300 HAT inhibitor A-485[28], or treatment with the p300/CBP bromodomain inhibitor CBP30 greatly reduced foci formation (Fig. 4c). We therefore conclude that HAT activation drives biomolecular condensation of p300 in cells, apparently through substrate engagement by the bromodomain.

Regulation of p300 by the AIL and RING domain

We next sought to understand how the highly conserved and intrinsically disordered AIL segment contributes to the regulation of the catalytic function of p300. The AIL spans amino acid residues 1532–1567 and is positively charged in the deacetylated state, with an estimated isoelectric point (pI) of 10.9, and net charge of +7 at neutral pH. By contrast, upon autoacetylation of residues spanning Lys1542–1560[29], we estimate a pI of 3.5 and a net charge of −2. As the proximal substrate-binding groove of p300 is largely acidic (Fig. 2c), we proposed—consistent with earlier predictions[30]—that a deacetylated AIL would engage the substrate binding site through electrostatic interactions, presumably to prevent access of positively charged lysine-containing substrates. Given the disordered nature of the AIL, this proposed interaction is expected to be highly dynamic[27].

We tested this hypothesis through all-atom Monte Carlo simulations[31]. To make this approach tractable, our simulations held the backbone dihedral angles associated with the folded domains fixed, but all other degrees of freedom, including all backbone and side chain dihedral angles in the AIL, were fully sampled. As a result, these simulations should be seen to assess how the AIL interacts with the remainder of p300 given the observed crystal structure. Simulations were performed on the AIL in the deacetylated and acetylated states in the context of the p300 monomer. These simulations enabled us to investigate how acetylation influenced the conformation and intramolecular interactions of the AIL.

Simulations of the deacetylated AIL revealed the presence of extensive yet highly degenerate electrostatic interactions between the AIL and the RING domain and between the AIL and the HAT substrate-binding site. These interactions were quantifiable in terms of the normalized distances between pairs of amino acid residues (Fig. 5a, Supplementary Video 1). Lysine residues in the AIL dynamically associate through long-range electrostatic interactions with acidic residues (E1334, E1442, E1505, D1622, D1625 and D1628) in the p300 HAT substrate-binding pocket (Fig. 5c). The importance of these residues for substrate acetylation has been shown previously[32], and nuclear magnetic resonance data for CBP confirm that the AIL is intrinsically disordered in the deacetylated state[27].

By contrast, in the acetylated state we found no interactions between the AIL and the substrate binding site (Fig. 5b and Supplementary Video 2). The acetylated AIL essentially behaved like a self-avoiding random coil without any strong biases for interaction with itself or with the surrounding folded domains, including the bromodomain. It has been proposed that the AIL of CBP, when acetylated on K1596 (K1558 in p300), engages the bromodomain intramolecularly, thus competing with histone binding and negatively regulating substrate acetylation[27]. Isothermal calorimetry experiments showed the highest binding affinity for multiacetylated peptides, including the diacetylated histone peptides H3(K14ac/K18ac) and H4(K12ac/K16ac), generally following the pattern KacNNNKac (Extended Data Table 2). Monoacetylated peptides typically had weaker binding affinity. A crystal structure of the H4(K12ac/K16ac) peptide bound to BΔRP (Extended Data Fig. 5c)

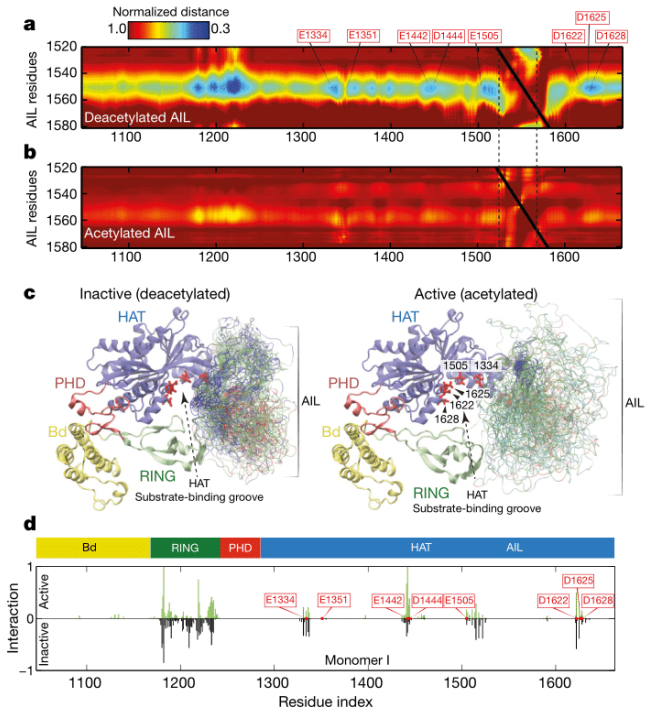


Fig. 5 | Acetylation of the AIL regulates dynamic interaction with the substrate-binding pocket of p300. a, Normalized distances between the AIL and residues in the inactive monomer. Inter-residue distances were normalized by the distances expected if the AIL behaved as a self-avoiding random coil. Electrostatic interaction mediated by conserved lysine residues between K1542 and K1560 of the AIL and aspartic acid or glutamic acid residues around the active site of the HAT domain (E1334, E1351, E1442, D1444, E1505, D1622, D1625 and D1628). The extensive contacts between the AIL and the RING domain originate in part from the proximity of the RING domain to the AIL in its inhibitory conformation. **b**, Normalized distances between the AIL and all residues in the active (acetylated) monomer. After acetylation, lysine-mediated electrostatic interactions are lost. **c**, Representative conformations with the AIL shown as an ensemble for the inactive deacetylated monomer (left) and the active acetylated monomer (right). The C_α atoms of residues in the AIL are coloured according to charge: blue, positive; red, negative; green, non-charged. The HAT substrate-binding groove is more exposed in the active acetylated state, due to both the relative position of the RING domain and the lack of preferential interactions by the AIL. **d**, Intermolecular interactions in the loop-swapped dimer between the AIL of one HAT and the adjacent subunit of the other. The adjacent subunit is either in the active (top) or inactive (bottom) conformation. In the active state, the AIL is able to directly engage with residues E1442 and E1444 from the adjacent HAT substrate-binding groove, which suggests that certain orientations of the RING domain can sterically hinder access to the AIL.

confirmed the acetyllysine-specific binding mode. However, a AIL peptide acetylated on the three lysines K1549, K1558 and K1560—corresponding to some of the most highly acetylated residues in the AIL[29]—failed to bind to the BRP module. Thus our interpretation is that the multiacetylated AIL is not a substrate for the bromodomain, presumably because of suboptimal spacing or sequence environment of the acetylated lysine sites of the AIL.

To understand how the RING domain influences the ability of substrates—including the AIL—to enter the active site of an adjacent p300 molecule, we performed simulations of the AIL in the context of the loop-swapped dimer, using a harmonic potential to maintain the AIL in the active site in order to assess potential intermolecular interactions (Fig. 5d, Extended Data Fig. 6a). In the active RING conformation, the AIL is able to engage the substrate binding site. However, in the inactive conformation, the frequency of contacts between the AIL and the acidic active site residues E1442 and D1444—residues proximal to the lysine substrate binding tunnel—was reduced by 70–75% (Fig. 5d). Thus, in the inactive conformation, the RING domain at least partially reduces catalytic activity by limiting accessibility of the active site to the AIL and other substrates.

One prediction from our models is that the deacetylated form of p300 adopts a more compact conformation, owing to dynamic engagement of the AIL with the HAT substrate-binding site, whereas the acetylated form adopts a more 'open' conformation (Fig. 5d). To test this possibility, we produced hypo- and hyperacetylated p300 variants (Extended Data Fig. 6e–g) and analysed the preparations by SEC–MALLS. All preparations were monomeric at the concentration tested (2 mg ml^{-1}) (Extended Data Fig. 6b–d, Extended Data Table 3). Hyperacetylation of p300 BRP-HAT resulted in a small decrease in the elution volume, which is indicative of a larger hydrodynamic radius (Extended Data Fig. 6b). A similar result was obtained upon comparison of hyper- and hypoacetylated BRP-HAT-CH3 (Extended Data Fig. 6c). By contrast, a variant that lacks the AIL showed no change in the elution volume upon hyperacetylation (Extended Data Fig. 6c). Our data are therefore consistent with the model that the catalytic p300 'core' adopts a compact conformation in the hypoacetylated state, with autoacetylation resulting in a more extended conformation.

Discussion

Our findings provide detailed mechanistic insights into how cellular signalling controls the activity of a chromatin regulator. We propose a multi-step process for p300 HAT activation and signal transmission to chromatin (Extended Data Fig. 7a–d). In the basal state, the deacetylated AIL is expected to maintain an overall positively charged environment in close proximity to the active site of the enzyme, thus preventing access of positively charged lysine-rich substrates. Direct access to the CoA-binding tunnel and autoacetylation of the AIL *in cis* appears to be prohibited, in part due to the positioning of the RING domain (Fig. 5d).

Cellular signalling initiates phosphorylation of transcription factors, such as IRF3 or STAT1, which results in their activation and dimerization. The activated, dimeric transcription factors are in their DNA-binding-competent conformation and can engage two molecules of p300 in the nucleus, thus increasing the likelihood of AIL disengagement from its inhibitory position *in cis* and of its capture *in trans* by a second p300 molecule. Association of two p300 molecules does not necessarily require precise stereospecific interactions between the structured domains, because acetylation at several lysines in the AIL indicates a series of possible conformations in such transiently associating dimers. We predict that regulated oligomerization uncouples recruitment from HAT activation, which could explain why not all p300/CBP recruitment events result in chromatin acetylation and gene activation[12–17,33].

It has been proposed that enhancer RNA interacts with the AIL to regulate CBP HAT activity[34]. We have attempted to reproduce these results using Klf6, one of the most potent enhancer RNAs reported[34]. We could not detect p300 HAT activation using up to equimolar amounts of Klf6 (Extended Data Fig. 7e, g). We note that, in a previous study, CBP was purified in buffer containing EDTA; this is detrimental to the structure of p300/CBP owing to the presence of multiple zinc-binding domains[35]. When unfolded by incubation with EDTA, CBP and p300 have a high tendency to aggregate and to form non-specific interactions[35]. Paradoxically, as the HAT domain is not affected, inclusion of EDTA can have an 'activating' effect in biochemical assays, apparently due to such non-specific aggregation (Extended Data Fig. 7f). The detrimental effects of EDTA on the structure and function of p300/CBP need to be taken into account in the interpretation of such data.

The ability of certain histone-modifying enzymes to bind to the post-translational modification (PTM) they generate has led to models in which such enzymes might propagate modified chromatin domains by a positive-feedback loop[36]. According to this view histone PTMs and other chromatin modifications form an additional,

DNA-sequence-independent layer of the genome, which is read out by enzymes that recognize these modifications to 'epigenetically' regulate genomic function[36]. An alternative view proposes that histone PTMs ultimately depend on DNA-sequence-dependent recruitment of chromatin modifiers, and so do not necessarily form an independent 'epigenetic' layer of the genome[8,37–39]. The controversy has arisen because it has been difficult to disentangle, for most chromatin regulators, the relative contributions of DNA targeting and histone PTM substrate engagement to the overall chromatin-modification reaction.

We show that regulation of p300 is linked to the activation and oligomerization status of transcription factor ligands, and therefore conclude that specificity for p300-mediated chromatin acetylation arises mainly through transcription-factor-mediated and DNA-sequence-dependent genome targeting. The next question is how the bromodomain contributes to p300 function. Although it is clear that the bromodomain can engage acetylated histone peptides and bind to hyperacetylated chromatin[26,40], deletion or mutation of the bromodomain has no apparent effect on substrate acetylation[26,41], has only minimal effects in a haematopoiesis model system[42], and bromodomain inhibition does not adversely affect genome targeting of CBP[43].

We favour a model in which DNA binding provides the lead anchoring mechanism: local hyperacetylation increases the binding valency by enabling bromodomain substrate engagement, which further helps to compartmentalize the biochemical reaction and contributes to signal maintenance[40]. p300 HAT-activating mutants form biomolecular condensates in cells when transiently overexpressed (Fig. 4c, Extended Data Fig. 5d). Treatment with a HAT inhibitor or bromodomain inhibitor greatly reduces the formation of condensates, which indicates that hyperacetylation and bromodomain–substrate engagement are critical in driving assembly. The formation of condensates, possibly through phase-separation, may provide a mechanism to enable signal integration on enhancers and transcriptional control[44]. It will be critical to disentangle cause–effect relationships of DNA targeting, chromatin modification and histone PTM substrate engagement of other chromatin regulators[45–47].

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41586-018-0621-1.

Received: 14 February 2018; Accepted: 15 August 2018;
Published online 15 October 2018.

1. Chen, Q., Sun, L. & Chen, Z. J. Regulation and function of the cGAS–STING pathway of cytosolic DNA sensing. *Nat. Immunol.* **17**, 1142–1149 (2016).
2. Panne, D., McWhirter, S. M., Maniatis, T. & Harrison, S. C. Interferon regulatory factor 3 is regulated by a dual phosphorylation-dependent switch. *J. Biol. Chem.* **282**, 22816–22822 (2007).
3. Zhao, B. et al. Structural basis for concerted recruitment and activation of IRF-3 by innate immune adaptor proteins. *Proc. Natl Acad. Sci. USA* **113**, E3403–E3412 (2016).
4. Parekh, B. S. & Maniatis, T. Virus infection leads to localized hyperacetylation of histones H3 and H4 at the IFN-β promoter. *Mol. Cell* **3**, 125–129 (1999).
5. Panne, D., Maniatis, T. & Harrison, S. C. An atomic model of the interferon-beta enhanceosome. *Cell* **129**, 1111–1123 (2007).
6. Stark, G. R. & Darnell, J. E., Jr. The JAK-STAT pathway at twenty. *Immunity* **36**, 503–514 (2012).
7. Zhang, J. J. et al. Two contact regions between Stat1 and CBP/p300 in interferon gamma signaling. *Proc. Natl Acad. Sci. USA* **93**, 15092–15096 (1996).
8. Bedford, D. C. & Brindle, P. K. Is histone acetylation the most important physiological function for CBP and p300? *Aging* **4**, 247–255 (2012).
9. Heintzman, N. D. et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* **39**, 311–318 (2007).
10. Visel, A. et al. ChIP–seq accurately predicts tissue-specific activity of enhancers. *Nature* **457**, 854–858 (2009).
11. Jin, Q. et al. Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. *EMBO J.* **30**, 249–262 (2011).
12. Bedford, D. C., Kasper, L. H., Fukuyama, T. & Brindle, P. K. Target gene context influences the transcriptional requirement for the KAT3 family of CBP and p300 histone acetyltransferases. *Epigenetics* **5**, 9–15 (2010).
13. Zhao, L. et al. Integrated genome-wide chromatin occupancy and expression analyses identify key myeloid pro-differentiation transcription factors repressed by Myb. *Nucleic Acids Res.* **39**, 4664–4679 (2011).
14. Waltzer, L. & Bienz, M. *Drosophila* CBP represses the transcription factor TCF to antagonize Wingless signalling. *Nature* **395**, 521–525 (1998).
15. Holmqvist, P. H. & Mannervik, M. Genomic occupancy of the transcriptional co-activators p300 and CBP. *Transcription* **4**, 18–23 (2013).
16. Kasper, L. H., Qu, C., Obenauer, J. C., McGoldrick, D. J. & Brindle, P. K. Genome-wide and single-cell analyses reveal a context dependent relationship between CBP recruitment and gene expression. *Nucleic Acids Res.* **42**, 11363–11382 (2014).
17. Rada-Iglesias, A. et al. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**, 279–283 (2011).
18. Thompson, P. R. et al. Regulation of the p300 HAT domain via a novel activation loop. *Nat. Struct. Mol. Biol.* **11**, 308–315 (2004).
19. Qin, B. Y. et al. Crystal structure of IRF-3 in complex with CBP. *Structure* **13**, 1269–1277 (2005).
20. Larabi, A. et al. Crystal structure and mechanism of activation of TANK-binding kinase 1. *Cell Rep.* **3**, 734–746 (2013).
21. Levy, D. E. & Darnell, J. E., Jr. Stats: transcriptional control and biological impact. *Nat. Rev. Mol. Cell Biol.* **3**, 651–662 (2002).
22. Shuai, K., Stark, G. R., Kerr, I. M. & Darnell, J. E., Jr. A single phosphotyrosine residue of Stat91 required for gene activation by interferon-gamma. *Science* **261**, 1744–1746 (1993).
23. Chen, X. et al. Crystal structure of a tyrosine phosphorylated STAT-1 dimer bound to DNA. *Cell* **93**, 827–839 (1998).
24. Wojciak, J. M., Martinez-Yamout, M. A., Dyson, H. J. & Wright, P. E. Structural basis for recruitment of CBP/p300 coactivators by STAT1 and STAT2 transactivation domains. *EMBO J.* **28**, 948–958 (2009).
25. Darnell, J. E., Jr. STATs and gene regulation. *Science* **277**, 1630–1635 (1997).
26. Delvecchio, M., Gaucher, J., Aguilar-Gurrieri, C., Ortega, E. & Panne, D. Structure of the p300 catalytic core and implications for chromatin targeting and HAT regulation. *Nat. Struct. Mol. Biol.* **20**, 1040–1046 (2013).
27. Park, S. et al. Role of the CBP catalytic core in intramolecular SUMOylation and control of histone H3 acetylation. *Proc. Natl Acad. Sci. USA* **114**, E5335–E5342 (2017).
28. Lasko, L. M. et al. Discovery of a selective catalytic p300/CBP inhibitor that targets lineage-specific tumours. *Nature* **550**, 128–132 (2017).
29. Karanam, B., Jiang, L., Wang, L., Kelleher, N. L. & Cole, P. A. Kinetic and mass spectrometric analysis of p300 histone acetyltransferase domain autoacetylation. *J. Biol. Chem.* **281**, 40292–40301 (2006).
30. Karanam, B. et al. Multiple roles for acetylation in the interaction of p300 HAT with ATF-2. *Biochemistry* **46**, 8207–8216 (2007).
31. Vitalis, A. & Pappu, R. V. ABSINTH: a new continuum solvation model for simulations of polypeptides in aqueous solutions. *J. Comput. Chem.* **30**, 673–699 (2009).
32. Liu, X. et al. The structural basis of protein acetylation by the p300/CBP transcriptional coactivator. *Nature* **451**, 846–850 (2008).
33. Soutoglou, E. et al. Transcription factor-dependent regulation of CBP and P/CAF histone acetyltransferase activity. *EMBO J.* **20**, 1984–1992 (2001).
34. Bose, D. A. et al. RNA binding to CBP stimulates histone acetylation and transcription. *Cell* **168**, 135–149.e122 (2017).
35. Matt, T., Martinez-Yamout, M. A., Dyson, H. J. & Wright, P. E. The CBP/p300 TAZ1 domain in its native state is not a binding partner of MDM2. *Biochem. J.* **381**, 685–691 (2004).
36. Allis, C. D. & Jenuwein, T. The molecular hallmarks of epigenetic control. *Nat. Rev. Genet.* **17**, 487–500 (2016).
37. Ptashne, M. Epigenetics: core misconcept. *Proc. Natl Acad. Sci. USA* **110**, 7101–7103 (2013).
38. Rando, O. J. Combinatorial complexity in chromatin structure and function: revisiting the histone code. *Curr. Opin. Genet. Dev.* **22**, 148–155 (2012).
39. Henikoff, S. & Shilatifard, A. Histone modification: cause or cog? *Trends Genet.* **27**, 389–396 (2011).
40. Nguyen, U. T. et al. Accelerated chromatin biochemistry using DNA-barcoded nucleosome libraries. *Nat. Methods* **11**, 834–840 (2014).
41. Rack, J. G. M. et al. The PHD finger of p300 influences its ability to acetylate histone and non-histone targets. *J. Mol. Biol.* **426**, 3960–3972 (2014).
42. Kimbrel, E. A. et al. Systematic in vivo structure–function analysis of p300 in hematopoiesis. *Blood* **114**, 4804–4812 (2009).
43. Picaud, S. et al. Generation of a selective small molecule inhibitor of the CBP/p300 bromodomain for leukemia therapy. *Cancer Res.* **75**, 5106–5119 (2015).
44. Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K. & Sharp, P. A. A phase separation model for transcriptional control. *Cell* **169**, 13–23 (2017).
45. Coleman, R. T. & Struhl, G. Causal role for inheritance of H3K27me3 in maintaining the OFF state of a *Drosophila* HOX gene. *Science* **356**, eaai8236 (2017).
46. Laprell, F., Finkl, K. & Müller, J. Propagation of polycomb-repressed chromatin requires sequence-specific recruitment to DNA. *Science* **356**, 85–88 (2017).
47. Wang, X. & Moazed, D. DNA sequence-dependent epigenetic inheritance of gene silencing and histone H3K9 methylation. *Science* **356**, 88–91 (2017).

Acknowledgements This work was supported by grant 16-0280 from Worldwide Cancer Research. E.O. was supported by an EMBL Interdisciplinary Postdoctoral (EIPOD) fellowship. S.R. was supported by the Fondation ARC pour la recherche sur le Cancer and by the Fondation FINOVI. A.S.H. is a postdoctoral fellow in the laboratory of R.V. Pappu at Washington University in St. Louis. The computational work was supported by the Human Frontiers

Science Program (grant RGP0034/2017 to R.V. Pappu) and the St Jude Collaborative Research Consortium on Membraneless Organelles (to R.V. Pappu). We thank the staff at the European Synchrotron Radiation Facility (ESRF) beamlines ID29; L. Signor for mass spectroscopy analysis; R. Vance for the plasmid encoding GST-STING; and P. Cole for the A-485 inhibitor. S.K. and D.P. were supported by ANR Episperm3 program. S.K. received additional support from Fondation ARC Canc'air project (RAC16042CLA), Plan Cancer (CH7-INS15B66 and ASC16012CSA) and the Université Grenoble Alpes ANR-15-IDEX-02 LIFE and IDEX SYMER.

Reviewer information *Nature* thanks L. Chen, V. Hilser and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Author contributions E.O. designed and performed most experiments, analysed and validated the data and revised the draft with assistance from S.R., Z.I., N.H. and J.G. A.S.H. performed computational modelling and revised the draft. S.K. provided supervision, funding acquisition and commented on the draft. D.P. was involved in conceptualization, supervision, project administration, funding acquisition and wrote the original and revised drafts.

Competing interests The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-018-0621-1.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-018-0621-1.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to D.P.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized, and the investigators were not blinded to allocation during experiments and outcome assessment.

Constructs. For cell-free protein expression, cDNA of p300 (NCBI reference sequence: NM_001429.3) variants were cloned into the pIVEX2.4d vector (Roche) with a N-terminal 6× His tag and a C-terminal Flag tag. In the ΔR constructs, the RING domain encompassing residues 1169–1241 was replaced by a glycine amino acid residue linker. In the ΔAIL constructs, loop amino acid residues comprising residues 1520–1581 were replaced by the flexible linker sequence SGGSG. For *Escherichia coli* expression, cDNA encoding residues 1048–1282, for the BRP or BΔRP were cloned into the vector pETM-33 (EMBL) with a tobacco etch virus (TEV)-cleavable N-terminal glutathione *S*-transferase (GST) tag. p300 BRP_HAT variants were cloned into pFASTBAC1 (Thermo Fisher) and expressed in insect cells as shown previously[26]. p300s constructs, spanning amino acid residues 324–2094, were cloned into pFASTBAC1 vector with an N-terminal Flag tag. Haemagglutinin (HA)-tagged full-length p300 variants were cloned into pcDNA3.1 (Thermo Fisher). Point mutations were introduced by QuikChange mutagenesis (Agilent). Point mutations and nucleotide deletions carried out in p300 full-length (1–2414) or p300s (324–2094) were done through transfer vectors as described previously[26]. STAT1ΔN (136–748), STAT1ΔNC (136–713) and IRF3ΔC (1–382) with a C-terminal intein tag were cloned into the pTXB1 vector (New England Biolabs) using the restriction enzymes NdeI (STAT1) or NcoI (IRF3) and SpeI. IRF3 (1–427) with an N-terminal His-tag cleavable by TEV protease was cloned using the restriction enzymes NcoI and XhoI into the vector pETM-11 (EMBL). All constructs were confirmed by DNA sequencing.

Expression and purification. Expression and purification of Flag-tagged p300s constructs was done as described previously[2]. This method enables purification of p300s variants that are already pre-acetylated. Expression and purification of p300 BRP_HAT and SIRT2 were performed as described previously[26]. TBK1 was expressed in insect cells and purified as described previously[20]. Cell-free protein synthesis was done in a 50 μl reaction volume. In brief, 10 μg ml^{-1} of His-p300 variants in pIVEX2.4d were added to a reaction mixture containing 1 mM amino acid mix, 0.8 mM rNTPs (guanosine-, uracil-, and cytidine-5′-triphosphate ribonucleotides), 1.2 mM adenosine-5′-triphosphate, 55 mM HEPES, pH 7.5, 68 μM folinic acid, 0.64 mM cyclic adenosine monophosphate, 3.4 mM dithiothreitol (DTT), 27.5 mM ammonium acetate, 2 mM spermidine, 5 μM $ZnCl_2$, 80 mM creatine phosphate, 208 mM potassium glutamate, 16 mM magnesium acetate, 250 μg ml^{-1} creatine kinase, 27 μg ml^{-1} T7 RNA polymerase, 0.175 μg ml^{-1} tRNA and 67 μl ml^{-1} S30 *E. coli* bacterial extract. Incubation was carried out at 22 °C with agitation for 16 h. Proteins were purified using Ni-NTA chromatography (IMAC Sepharose 6 FF, GE healthcare) in buffer 1 (20 mM TRIS, pH 8.0, 300 mM NaCl, 1 mM DTT, 5 μM $ZnCl_2$) containing Complete Protease Inhibitors EDTA-Free (Roche). The resin was washed with 20 column volumes of buffer 1 and the protein eluted with 5 column volumes of buffer 1 containing 300 mM Imidazole. The protein was concentrated in a pre-washed Amicon Ultra 0.5 ml Ultracel 10K centrifugal filter (molecular weight cut off 10 kDa; EMD Millipore). The protein was buffer-exchanged into buffer 1 using 0.5 ml Zeba Spin desalting columns (molecular weight cut off 7 kDa; Thermo Scientific), flash-frozen in liquid nitrogen and stored at −80 °C.

For expression of GST–BRP and GST–BΔRP fusion proteins in *E. coli* BL21 (DE3), LB medium enriched with 100 μM $ZnCl_2$ was used. Cell pellets were resuspended in buffer 1 containing Complete Protease Inhibitors EDTA-Free (Roche) and lysed using a microfluidizer (Microfluidics Corp.). The lysate was clarified by centrifugation for 30 min at 39,000*g* in a JA-25.5 rotor (Beckman) and applied to a Glutathione Sepharose 4 Fast Flow resin according to instructions from the manufacturer (GE Healthcare). The resin was washed with buffer 1 and incubated with His-tagged TEV protease (1:100 w/w) for 14–16 h at 4 °C. Subtractive Ni-NTA chromatography (IMAC Sepharose 6 FF, GE Healthcare) was then used to remove the residual His-tag and TEV protease. The untagged protein was further purified by gel filtration on a High Load 16/60 Superdex 75 column (GE Healthcare) equilibrated in 20 mM HEPES, pH 7.5, 300 mM NaCl, 0.5 mM TCEP and 5 μM $ZnCl_2$. The final protein was concentrated to 15 mg ml^{-1} in a prewashed Amicon Ultra-15 Centrifugal filter (molecular weight cut off 10 kDa; EMD Millipore), flash-frozen in liquid nitrogen and stored at −80 °C.

The expression and purification of non-phosphorylated STAT1 variants (STAT1ΔN, STAT1ΔNC) and IRF3ΔC (1–382) was done in *E. coli* using the IMPACT expression system (New England Biolabs). For the expression of Y701 phosphorylated variants (pSTAT1ΔN, pSTAT1ΔNC), proteins were co-expressed with Elk receptor tyrosine kinase domain in *E. coli* BL21(DE3) TKB1 cells (Agilent). Cells were collected by centrifugation and resuspended in buffer 2 (20 mM HEPES pH 7.5, 500 mM NaCl). The cells were lysed in a microfluidizer (Microfluidics Corp.) and the soluble fraction was obtained by centrifugation for 30 min at 39,000*g* in a JA-25.5 rotor (Beckman). The supernatant was first passed over chitin beads (New England Biolabs) and washed with buffer 2 for 10 column volumes. The protein was cleaved at 4 °C for 16 h in buffer 2 containing 50 mM DTT, eluted and further purified by gel filtration on a High Load 16/60 Superdex 200 column (GE Healthcare) equilibrated in buffer 2.

GST–STING, comprising the soluble cytoplasmic domain spanning amino acids 138–378, was expressed in *E. coli* BL21(DE3) at 37 °C for 3 h. The cells were collected by centrifugation and resuspended in buffer 3 (20 mM TRIS, pH 8.0, 300 mM NaCl, 1 mM DTT) containing Complete Protease Inhibitors EDTA-Free (Roche). The cells were lysed in a microfluidiser (Microfluidics Corp.) and the soluble fraction was obtained by centrifugation as above. The supernatant was passed over equilibrated Glutathione Sepharose 4 Fast Flow resin according to instructions by the manufacturer (GE Healthcare). The resin was washed with buffer 3 and eluted with 10 mM reduced Glutathione in buffer 3. The protein was further purified by gel filtration on a High Load 16/60 Superdex 200 column (GE Healthcare) equilibrated in 20 mM HEPES, pH 7.5, 300 mM NaCl, 0.5 mM TCEP. The final protein was concentrated to 16 mg ml^{-1} in a prewashed Amicon Ultra-15 centrifugal filter (molecular weight cut off 30 kDa; EMD Millipore), flash-frozen in liquid nitrogen and stored at −80 °C.

IRF3 was expressed in *E. coli* BL21(DE3) at 18 °C for 16 h. The cells were collected by centrifugation and resuspended in buffer 2 containing 10 mM imidazole. The cells were lysed in a microfluidiser (Microfluidics Corp.) and the soluble fraction was obtained by centrifugation as above. The supernatant was passed over Ni^{2+}-conjugated IMAC Sepharose resin (GE Healthcare) and washed with buffer 2 containing 20 mM imidazole. The protein was eluted in buffer 2 containing 500 mM imidazole and was further purified by gel filtration on a High Load 16/60 Superdex 200 column in buffer 2 containing 0.5 mM TCEP. IRF3 was phosphorylated in vitro at a 1:10 molar ratio TBK1:IRF3 (1 mg ml^{-1}) in presence of 5 mM $MgCl_2$ and 1 mM ATP. The reaction was incubated at 30 °C for 1 h and then for an additional 10 h at 21 °C. Phosphorylated IRF3 was further purified by size-exclusion chromatography on a Superdex S200 16/60 column (GE Healthcare) in 20 mM HEPES, pH 7.5, 300 mM NaCl, 0.5 mM TCEP. The production of recombinant histones was done following standard procedures[48].

Crystallization and structure determination. The p300 BRP_HAT construct comprising the AIL and the mutation Y1467F was deacetylated as described previously[26]. The protein at 4.5 mg ml^{-1} was incubated with a threefold molar excess of the bi-substrate inhibitor Lys-CoA[32] before crystallization. Crystals in the $P2_1$ space group were grown by hanging-drop vapour diffusion at 4 °C by mixing equal volumes of protein and crystallization solution containing 100 mM HEPES, pH 7.5, 18–22% polyethylene glycol 3350 and 0.2 M NaCl. Crystals were cryoprotected in 20–25% ethylene glycol and drop-frozen in liquid nitrogen. We collected native diffraction data to a minimum Bragg spacing of 3.1 Å resolution at the ESRF on beamline ID29 under a nitrogen gas stream at 100 K, at a wavelength of 1.282 Å. We processed the data with XDS (Extended Data Table 1). The structure of the p300 BRP_HAT was determined by molecular replacement using Phaser. There are four copies in the asymmetric unit and the RING domains were initially not visible in the electron density map and are partially disordered. Inspection of an anomalous difference map indicated peak density for the zinc ions and enabled positioning of the RING domain in the outward-rotated conformation. A final model was produced by iterative rounds of manual model building in Coot and refinement using PHENIX. The final model contains residues 1045 –1664 with a deletion of residues 1534–1567 and was refined to a 3.1 Å resolution with an R_{work} and an R_{free} of 19% and 26%, respectively (Extended Data Table 1). Analysis of the refined structure by MolProbity showed that there are no residues in disallowed regions of the Ramachandran plot. The MolProbity all-atom-clash score was 1.91, placing the structure in the 100th percentile among structures refined at 3.1 Å resolution ($n = 2,108$).

The BΔRP construct at 15 mg ml^{-1} was mixed with 2 mM of a 11-mer histone peptide H4 (10–20) GLGKacGGAKacRHR (only the underlined amino acid sequence is visible in the electron density map) containing two acetylated lysine residues at K12 and K16, H4(K12ac/K16ac). Crystals in the $P2_12_12_1$ space group were grown by hanging-drop vapour diffusion at 21 °C by mixing equal volumes of protein and crystallization solution containing 1.6 M ammonium sulfate and 100 mM bicine at pH 9.0. Crystals were cryoprotected in 20% ethylene glycol and drop-frozen in liquid nitrogen. We collected native diffraction data to a minimum Bragg spacing of 2.5 Å resolution at the ESRF on beamline ID29 under a nitrogen gas stream at 100 K, at a wavelength of 1.0 Å (Extended Data Table 1). Data processing, molecular replacement and refinement were performed as indicated above. The final model contains two copies of the BΔRP module corresponding to residues 1049–1279 of p300 in the asymmetric unit. As expected, replacement of the RING domain residues 1169–1241 by a single glycine amino acid linker did not adversely affect the remainder of the BP module. Analysis of the refined structure by MolProbity showed that there are no residues in disallowed regions of the Ramachandran plot. The MolProbity all-atom-clash score was 0.97, placing the structure in the 100th percentile ($n = 6,960$).

Monte Carlo simulations. All-atom Monte Carlo simulations were performed using ABSINTH implicit solvent model and version 2 of the CAMPARI Monte Carlo simulation engine (http://campari.sourceforge.net)[31]. The initial AIL loop was constructed using MODELLER, and the complete set of backbone and side chain torsional angles were sampled for the AIL for which electron density was missing. Simulation analysis was performed with MDTraj and CTraj (http://pappulab.wustl.edu/CTraj.html)[49]. The backbone degrees of freedom of the folded domains were not sampled, while all amino acid side chains were fully sampled. CAMPARI simulations explore conformational space through perturbation to the torsional angles (as opposed to Cartesian positions, as is typical for molecular dynamics). Consequently, a fully closed loop represents a major sampling challenge. To address this, we severed the covalent backbone bond between the N-terminal part of the AIL loop and the folded domain, and replaced this bond with a strong harmonic potential that recapitulates the distances constraint associated with the covalent bond. This allows moves to fully rotate the chain and markedly improves the efficiency of conformational sampling.

We generated 5,000 independent non-overlapping starting configurations and used a clustering approach to identify the most distinct 200 conformations. These were used as the starting conformations for full simulations. We ran 200 independent simulations of the deacetylated and acetylated p300 in the monomeric form, and 200 independent simulations of the loop-swapped p300 dimer in the active and inactive form (800 simulations total). Analysis was performed after an initial equilibration. Dimer simulations applied a harmonic potential between residue 1550 from the AIL and residue 1442 from the other monomer to maintain the AIL in the active site. This enabled us to directly compare active-site accessibility of the AIL. For monomer simulations, no restraints were applied.

Each residue on the folded structure was evaluated for contacts with any residue in the AIL, and these contacts were summed to give an effective contact score. In this manner, the residues on the folded structure that most frequently interacted with any residue on the AIL were directly identified. Interaction was primarily electrostatic in nature, with residues E1334, E1444, E1505, D1622, D1625 and D1628 engaging in direct interactions. There are also extensive interactions between the AIL and the RING domain, although we cannot rule out that these interactions are driven by the harmonic potential applied to pull E1442 towards the active site. As might be expected, the AIL–RING interactions differed between the active and inactive conformations.

To assess interactions between the AIL and the folded domains in the monomer simulations, scaling map analysis was performed. In this analysis, a simulation of the AIL as a true self-avoid random coil is performed to generate a reference state, and then the mean inter-residue distances obtained in the full simulations are normalized by the distances obtained from this reference. The self-avoiding random coil simulations are performed using an identical protocol to the full simulations, with the exception that the only contribution to inter-atomic interactions comes from the repulsive part of the Lennard-Jones potential, meaning no attractive inter-atomic interactions or solvation effects are experienced. This ensures we generate a sequence and structure-specific self-avoiding random coil ensemble that provides a true reference state. Extensive details on the technical aspects associated with the generation of this reference state have been described previously[50]. The scaling maps enable us to easily identify local regions that engage in interactions that cause deviations from self-avoiding random coil behaviour.

HAT assays. The standard autoacetylation HAT assay was performed using [^{14}C] acetyl-CoA (Perkin-Elmer). Autoacetylation of p300 was quantified by autoradiography after SDS–PAGE gel analysis. The p300s preparations were equilibrated in 1× HAT buffer (25 mM TRIS-HCl, pH 7.5, 100 mM NaCl, 1 mM DTT, 10% glycerol and 1× Complete EDTA-free protease inhibitor (Roche)) for 10 min at 30 °C before initiation of the reaction by the addition of 200 μM [^{14}C]acetyl-CoA for the indicated time points. For experiments containing IRF3 STAT1 or the enhancer RNA (eRNA) Klf6, autoacetylation assays were performed at a fixed equimolar concentration (2 μM) of p300s and the indicated transcription factor or Klf6. Assays were performed in triplicate with different batches of proteins and on different days. At the indicated time point, 5 μl of the reaction was quenched by addition of 5 μl of 2× SDS gel loading buffer followed by analysis on a 4–20% SDS–PAGE gel. Experiments shown in Fig. 1c were performed in 1× kinase buffer (20 mM HEPES, pH 7.5, 250 mM NaCl, 20 mM β-glycerol phosphate, 1 mM sodium vanadate, 10 mM $MgCl_2$, 1 mM DTT, 1 mM ATP and a mix of 20 μM [^{14}C]acetyl-CoA and 80 μM cold acetyl-CoA (A2056, Sigma). 1 μM of p300s was incubated in the presence or absence of 1 μM IRF3, 2 μM TBK1 and 1 μM STING (lanes 1–6) or increasing amounts of STING as indicated (lanes 7–12). The gels were analysed by western blotting as indicated below or were fixed for 30 min in a solution containing 3% glycerol, 10% glacial acetic acid, 20% ethanol (v/v/v) in water. The gels were soaked for 5 min in a solution containing 1% glycerol, 5% PEG8000 in water and were dried for 30 min using a Bio-Rad Gel Dryer and the radioactivity quantified on a phosphorimage analyser (Typhoon, GE Healthcare) followed by analysis using ImageJ 1.8.0_112[51].

A p300 HAT scintillation proximity assay was designed similar to that described previously[28]. In brief, as a substrate we used a synthetic histone H4 peptide containing 15 amino acids derived from the N terminus of human H4 that was chemically attached to biotin with an amino hexanoic linker (Biotin-C6- GRGKGGKGLGKGGAK) (from peptid.de). The synthetic peptide was re-suspended in water and adjusted to pH 7.0 with concentrated NaOH.

A typical reaction contained p300s (50 nM), 12.5 μM biotinylated H4 peptide, acetyl-CoA (0.1 μM to 10 μM set at around 10 times the apparent K_m) in 20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 5 μM $ZnCl_2$, 0.01% Tween-20 and 0.1% BSA (w/v). For reactions containing pIRF3, 50 nM was added. 20 μl of a 2× reaction mixture containing p300s, H4 peptide with and without pIRF3 was pre-incubated at 30 °C for 5 min. The reaction was initiated by the addition of 20 μl of 2× acetyl-CoA containing a 1:3 mix of tritiated [^{3}H]acetyl-CoA (PerkinElmer; NET290050UC) with cold acetyl-CoA. For example, for 10 μM final acetyl-CoA concentration, a mix of 5 μM [^{3}H]acetyl-CoA and 15 μM cold acetyl-CoA (A2056, Sigma) was used. The reaction was quenched at the indicated time points by delivering 40 μl of the reaction mix into 120 μl of 0.5 M HCl in a FlashPlate Plus Streptavidin 96-well scintillant-coated microplate (Perkin Elmer, SMP103001PK). The plate was incubated for 1 h, and light emission was counted in a MicroBeta2 Scintillation Counter (Perkin Elmer) at 1 min per well in the top count mode. Counts per minute (cpm) were plotted against acetyl-CoA concentration. Typical progress curves are shown in Extended Data Fig. 1d. The initial rate was estimated by linear regression during the first 10 min of the reaction and plotted against acetyl-CoA concentration. All data were analysed using GraphPad Prism 7.0.

For the results shown in Fig. 4, acetylation reactions were performed in acetylation reaction buffer HAT (25 mM Tris-HCl, pH 7.5, 100 mM NaCl, 0.1 mM EDTA, 1 mM DTT, 10% glycerol and 1× complete EDTA-free protease inhibitor (Roche)) with 50 μM acetyl CoA (Sigma), 100 ng ml^{-1} trichostatin A and 2 μg of purified histone octamer. Reactions were incubated for 30 min at 30 °C and stopped by the addition of 3× SDS gel loading buffer, then used for Coomassie staining and immunoblotting.

Multi angle laser light scattering-size exclusion chromatography. Before SEC–MALLS runs, p300 variants were acetylated and deacetylated using p300 HAT or SIRT2 as described previously[26]. The reactions were analysed by liquid chromatography–mass spectrometry as described previously[52]. Size-exclusion chromatography was performed at a flow rate of 0.5 ml min^{-1} on a Superdex 200 Increase 10/300 GL column equilibrated in SEC–MALLS buffer (20 mM HEPES, 300 mM NaCl, 5 μM $ZnCl_2$, 0.5 mM TCEP) at 21 °C. A 50 μl sample of p300 at 2 mg ml^{-1} was injected onto the column and multi angle laser light scattering was recorded with a laser emitting at 690 nm using a DAWN-EOS detector (Wyatt Technology Corp.). The refractive index was measured using a RI2000 detector (Schambeck SFD). The molecular weight was calculated from differential refractive index measurements across the centre of the elution peaks using the Debye model for protein using ASTRA software version 6.0.5.3.

In vitro eRNA transcription. eKlf6 eRNA corresponding to 496 nucleotides of the sense strand of human chr13:5802100-5802596 (ref. [34]) was produced by in vitro transcription from a pMA plasmid containing a eKlf6 insert synthesized by GeneArt Gene Synthesis (Thermo Fisher). pMA_Klf6 plasmid (50 μg) was linearized with 80 U of KpnI-HF in a final volume of 100 μl and incubated at 37 °C for 14–16 h. The in vitro transcription reaction was done in a final volume of 1 ml, using 1× T7 buffer, T7 RNA Polymerase and 1 U of RNaseOUT Recombinant Ribonuclease Inhibitor (Thermo Fisher). After incubation for 2 h at 37 °C, 0.5 U of TURBO DNase (2 U μl^{-1}) (Thermo Fisher) and 1 μM $CaCl_2$ was added to the reaction and incubated for 30 min at 37 °C. Following DNaseI treatment, 2 μl of a 30 mg ml^{-1} stock of proteinase K powder (Thermo Fisher), dissolved in proteinase K buffer (10 mM TRIS pH 7.5, 1 mM $CaCl_2$ and 40% glycerol), was added and incubated for 45 min at 37 °C. Buffer was exchanged into 20 mM HEPES, pH 7.5, 300 mM NaCl, 0.5 mM TCEP using Amicon Ultra-0.5 ml centrifugal filters (molecular weight cut off 3 kDa, EMD Millipore). To further purify the RNA, 3 volumes of TRIzol (Thermo Fisher) was added to the RNA sample, followed by isopropanol precipitation. Purified Kfl6 RNA was resuspended in 20 mM HEPES, pH 7.5, 300 mM NaCl and 0.5mM TCEP. RNA was quantified using a Nanodrop spectrophotometer (Thermo Fisher). The quality of Klf6 was assessed by agarose gel electrophoresis in 1× TBE buffer or using denaturing 6M urea 14% PAGE (Extended Data Fig. 7c).

Immunoblotting, immunofluorescence and antibodies. We have used the cell lines COS and H1299. They are not on the list of commonly misidentified cell lines maintained by the International Cell Line Authentication Committee. COS cells were purchased from ATCC, product reference ATCC-CR1-1651, lot no. 4171903. H1299 cells were authenticated on December 6th 2016 by LGC Standards: Cell Line Authentication Service. Cell lines were authenticated using short tandem repeat (STR) analysis as described in 2012 in ANSI Standard (ASN-0002) Human cell lines were authenticated using standardization of STR Profiling by the ATCC Standards Development Organization as described in ref. [53]. Upon the receipt of COS cells

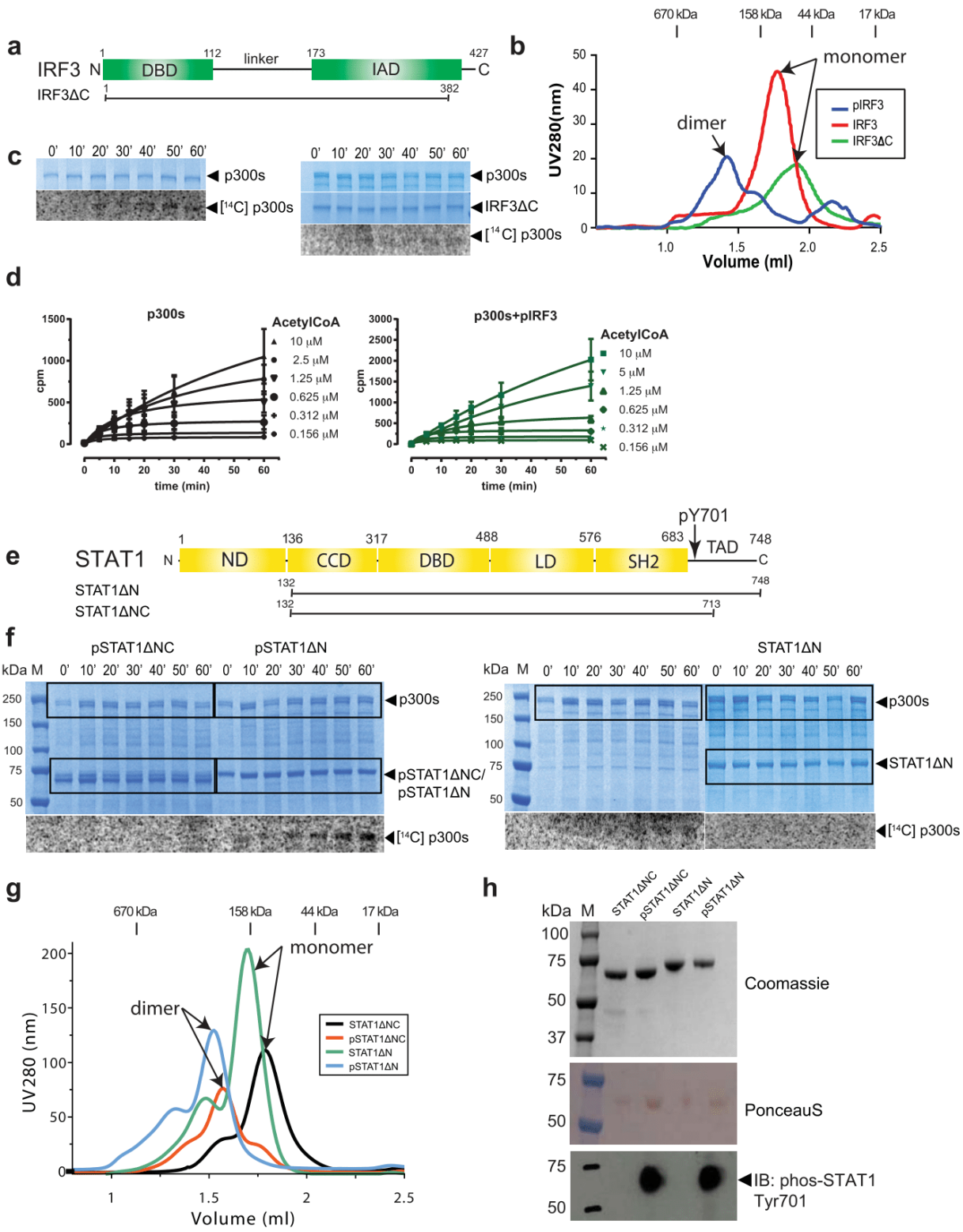
and after the authentication of H1299 cells, they were expanded. A mycoplasma contamination test (MycoAlert Mycoplasma Detection Kit, Lonza cat no. LT07-418) was performed and the mycoplasma-free cells were frozen and kept in liquid nitrogen. After thawing they were kept in culture for 30 passages with a mycoplasma contamination test after 15 passages. For immunoblotting, proteins were separated on 4–12% Bis-Tris SDS–PAGE gel (NuPAGE precast gel, Thermo Fisher) and transferred onto a nitrocellulose membrane (Hybond C+, GE Healthcare). Membranes were blocked with 5% skim milk in PBST buffer (PBS, 0.1% Tween-20) and probed with anti-p300(K1499ac) rabbit polyclonal antibody (1:2,500 dilution; Cell Signaling, 4771), anti-Kac rabbit polyclonal antibody (1:2,500 dilution; Cell Signaling, 9441), anti-Flag mouse monoclonal antibody (1:2,500 dilution; Sigma, F1804), anti-HA rabbit polyclonal antibody (1:2,500 dilution, Abcam, ab9110). For the detection of STAT1 or IRF3 phosphorylation, the membrane was blocked with 5% milk in PBST followed by overnight incubation at 4 °C with anti-phospho-Stat1 (Tyr701) rabbit monoclonal antibody (1:2,500 dilution; Cell Signaling no. 9171) in PBST buffer containing 5% BSA. For detection of IRF3 S396 phosphorylation, we used anti-phospho IRF3 S396 rabbit monoclonal antibody (1:2.500 dilution; Cell Signaling no. 4947). Incubations were performed as above. Membranes were washed extensively in PBST buffer before and after incubation with anti-rabbit or anti-mouse HRP-conjugated secondary antibody (1:10,000 dilution; GE Healthcare, NA934 or NA931), and protein bands were visualized on film after the ECL reaction (ECL Prime, GE Healthcare). Immunofluorescence was done as described previously[26] or as follows: 24 h post-transfection, cells were treated with a DMSO control or with 2.5 μM (final concentration) of HAT inhibitor A-485 (P. Cole, Harvard Medical School) or p300/CBP Bromodomain inhibitor (CBP30, Sigma, no. SML1133) dissolved in DMSO. The final DMSO concentration in all assays was 0.25%. After 24 h, cells were rinsed once with RNase-free PBS 1X. Next, cells were permeabilized in freshly prepared 0.2% Triton (Sigma) buffer for 5 min and were then fixed in freshly prepared 4% formalin solution (Sigma) for 10 min at room temperature. After three washes in RNase-free 1X PBS at room temperature, the cells were incubated in 5% skim milk in 1X PBS for 30 min and then probed with anti-HA high-affinity monoclonal antibody (1:100 dilution; Roche Applied Science, cat. no. 11867423001) overnight at 4 °C. Cells were washed extensively with 1X PBS before and after incubation with Alexa Fluor 488-conjugated secondary antibody (1:500 dilution; Invitrogen, no. A-11006) for 1 h at 37 °C. Cells were counterstained with Hoechst (250 ng ml^{-1}) and examined under a confocal laser scanning microscope (LSM510, Zeiss).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

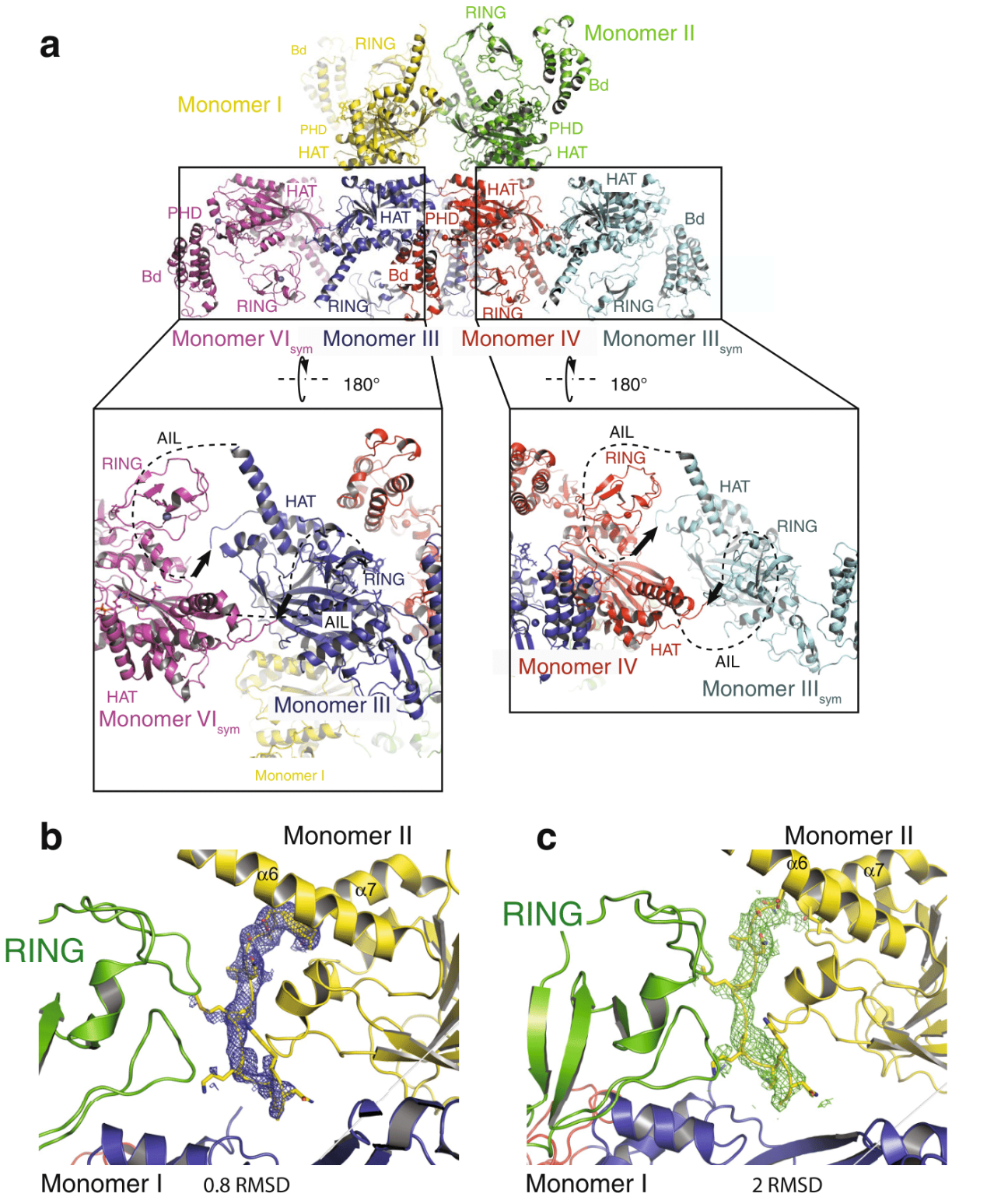
Coordinates for the p300 core structure and BΔRP bound to a diacetylated histone H4 peptide are available from the Protein Data Bank (PDB) under accession numbers 6GYR and 6GYT, respectively. Source data are available for Fig. 1b, f and Extended Data Fig. 1d. Figure 1d shows the initial velocities from reactions shown in Extended Data Fig. 1d.

48. Luger, K., Rechsteiner, T. J. & Richmond, T. J. Preparation of nucleosome core particle from recombinant histones. *Methods Enzymol.* **304**, 3–19 (1999).
49. McGibbon, R. T. et al. MDTraj: a modern open library for the analysis of molecular dynamics trajectories. *Biophys. J.* **109**, 1528–1532 (2015).
50. Holehouse, A. S., Garai, K., Lyle, N., Vitalis, A. & Pappu, R. V. Quantitative assessments of the distinct contributions of polypeptide backbone amides versus side chain groups to chain expansion via chemical denaturation. *J. Am. Chem. Soc.* **137**, 2984–2995 (2015).
51. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).
52. Kaczmarska, Z. et al. Structure of p300 in complex with acyl-CoA variants. *Nat. Chem. Biol.* **13**, 21–29 (2017).
53. Capes-Davis, A. et al. Match criteria for human cell line authentication: where do we draw the line? *Int. J. Cancer* **132**, 2510–2519 (2013).
54. Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).

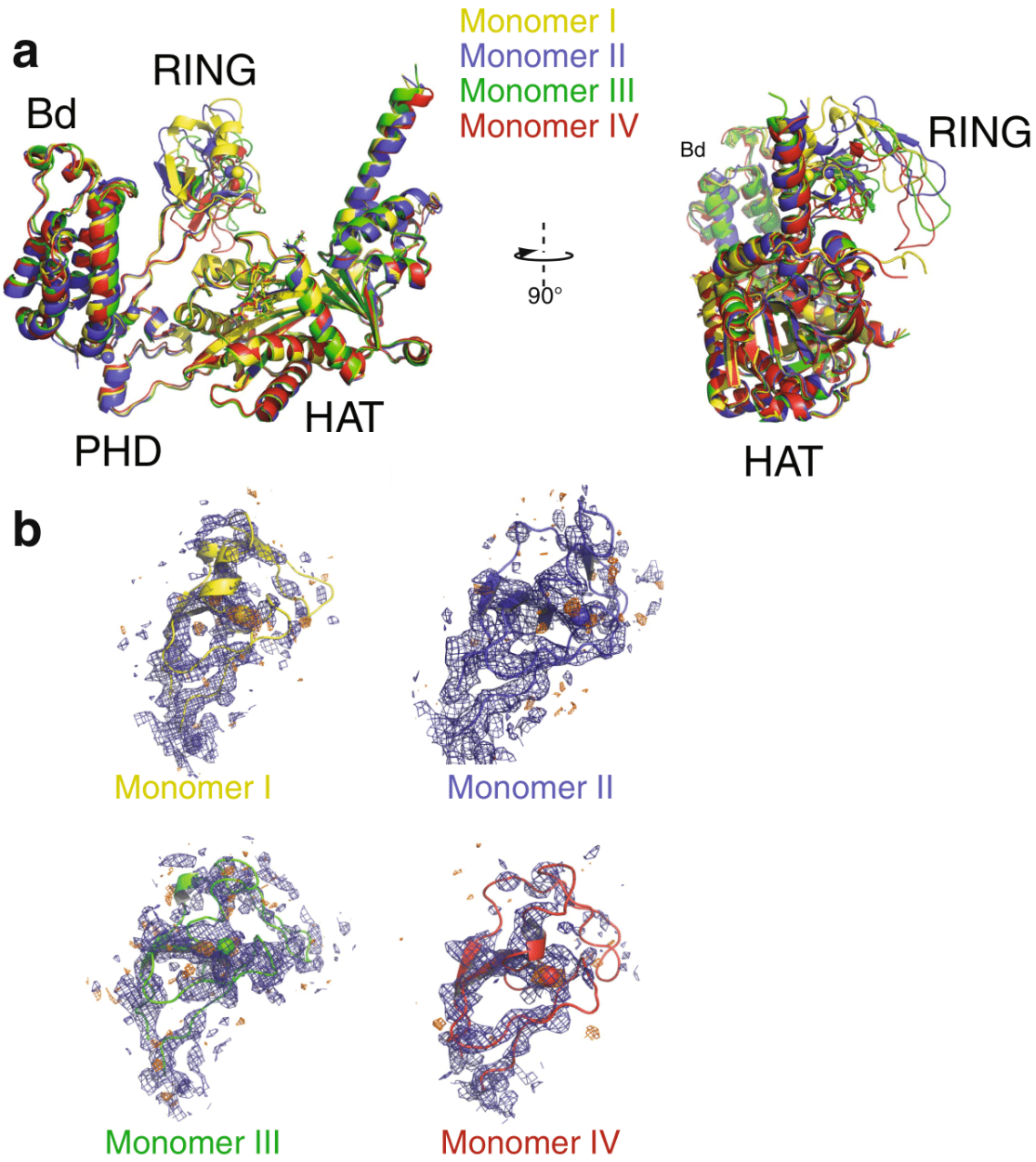


Extended Data Fig. 1 | See next page for caption.

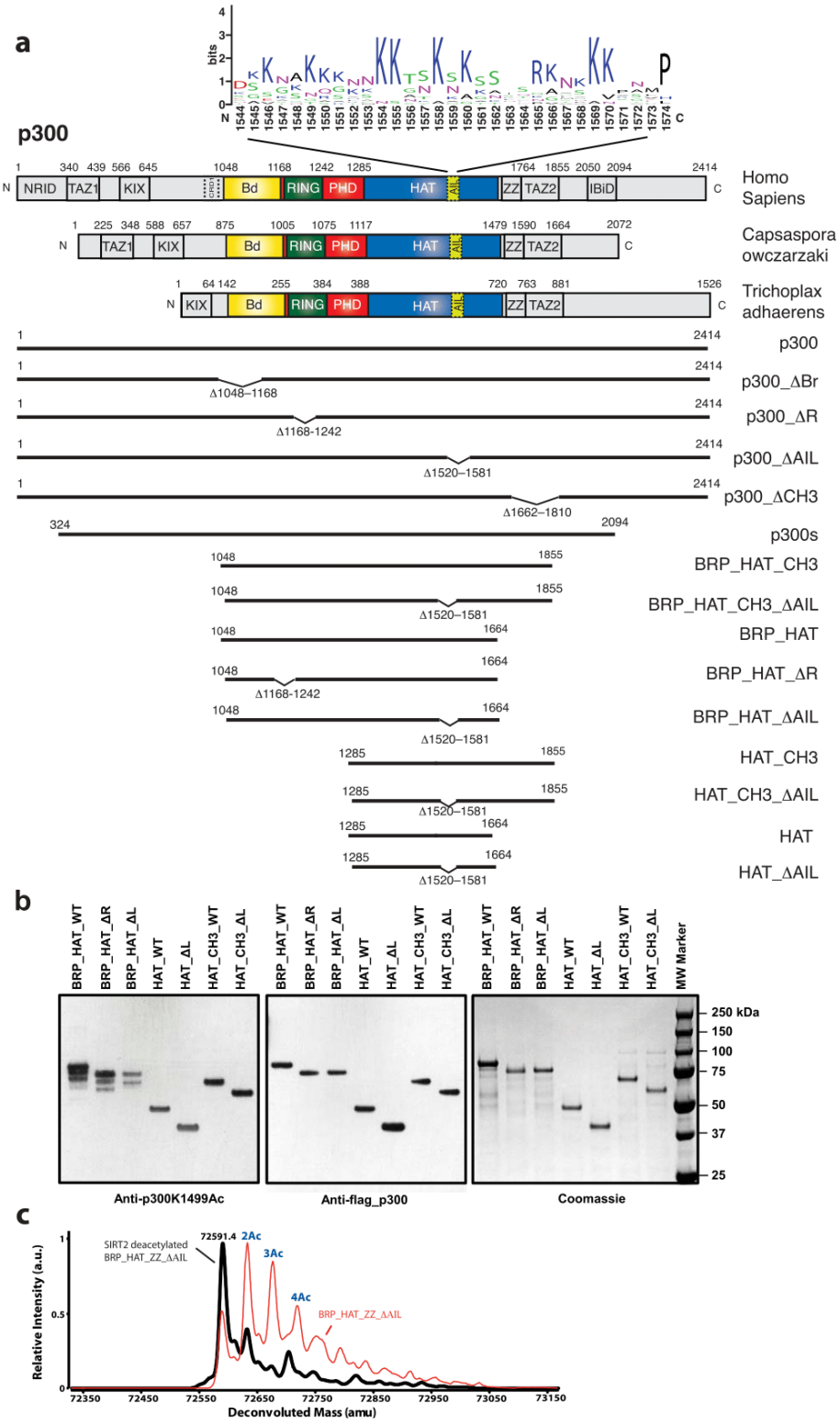
Extended Data Fig. 1 | The effect of IRF3 or STAT1 activation and oligomerization on p300 autoacetylation. a, The domain structure of IRF3. The truncation construct used is shown at the bottom. **b**, Size-exclusion chromatography of IRF3 variants. Red, unphosphorylated IRF3; blue, phosphorylated pIRF3; green, C-terminally truncated IRF3ΔC. Representative data of three independent experiments are shown. **c**, A constant amount of p300s (2 μM) was incubated alone or in the presence of C-terminally truncated IRF3ΔC (2 μM) for the indicated time points. Samples were analysed by SDS–PAGE followed by Coomassie staining and autoradiography. **d**, Progress curves of HAT scintillation proximity assay. Histone H4 substrate acetylation in the presence (green) or absence (black) of pIRF3 and varying concentrations of [^{3}H]acetyl-CoA. The degree of histone H4 substrate acetylation at different time points and the initial velocity (cpm min^{-1}) at the indicated acetyl-CoA concentrations were determined and plotted in Fig. 1e. Three independent experiments were performed and the mean value and error bars representing the standard deviation are shown. **e**, The domain structure of STAT1. The truncation constructs used are shown at the bottom, and the Tyr701 phosphorylation site is indicated. **f**, Uncropped images of SDS–PAGE gels shown in Fig. 1d. The ^{14}C autoacetylation signal of p300s is shown at the bottom. **g**, Size-exclusion chromatography of STAT1 variants. Black, STAT1ΔNC; green, STAT1ΔN; red, Y701-phosphorylated pSTAT1ΔNC; blue, Y701-phosphorylated pSTAT1ΔN. **h**, SDS–PAGE analysis of STAT1 variants and analysis by western blotting. Top, Coomassie staining of SDS–PAGE gel; middle, PonceauS staining; bottom, western blot using anti-Phospho-Stat1 (Tyr701). Representative data of three independent experiments are shown. For gel source data, see Supplementary Fig. 1.



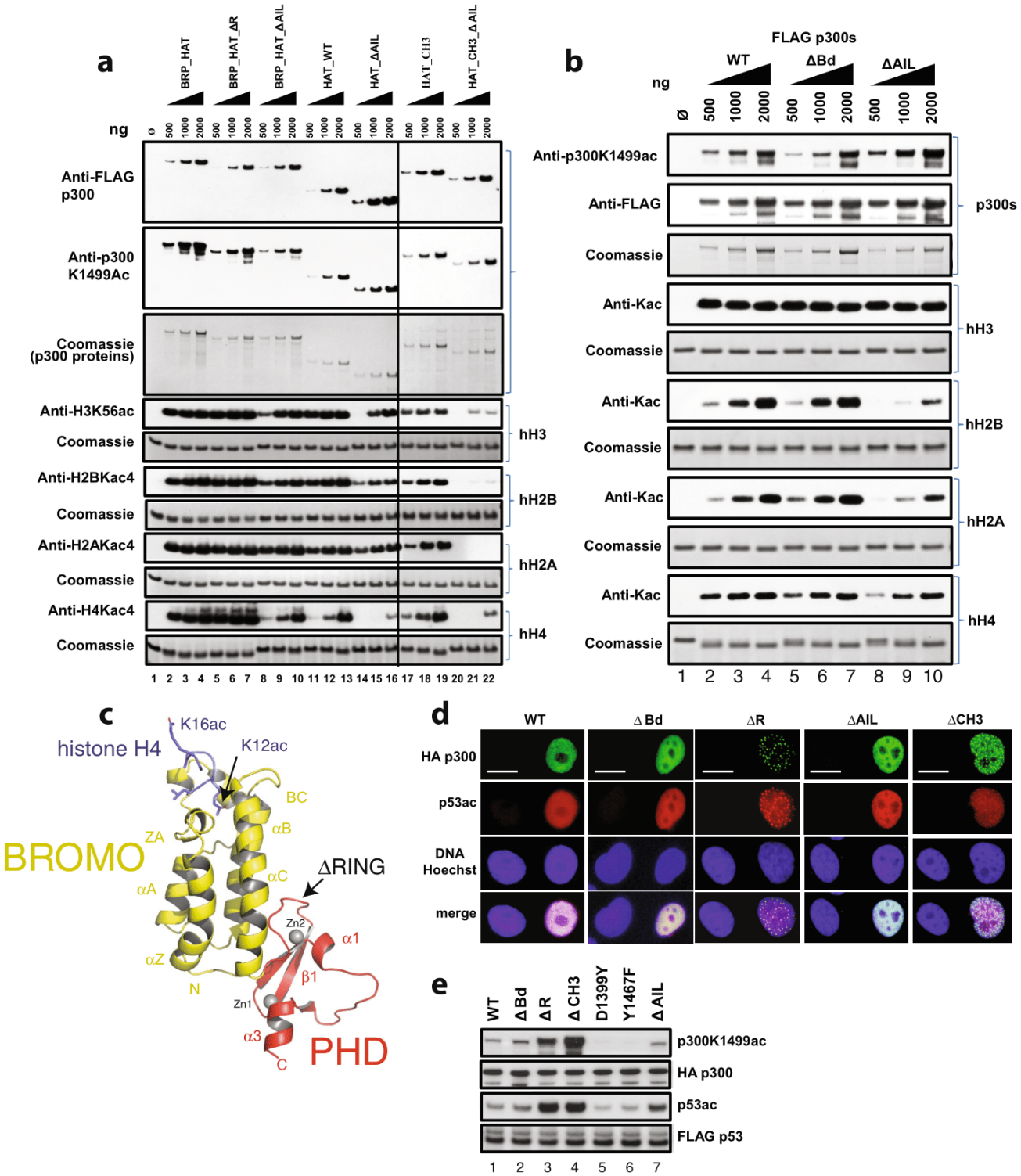
Extended Data Fig. 2 | Crystal packing of the p300 core molecule. **a**, There are four p300 molecules (monomers I–IV) in the asymmetric crystallographic unit. The four molecules show an antiparallel arrangement of the BRP-HAT domains. As a result, HAT domains from monomers I and II are closely apposed. Monomers III and IV engage monomer IV_{sym} and monomer III_{sym}, respectively, of a neighbouring crystallographic unit, showing that all promoters are in a AIL loop-swap conformation. Black arrows indicate the direction of the AIL. The disordered segment of the AIL is shown as a black dotted line. **b**, **c**, Electron density of the AIL. $2F_o - F_c$ (**b**) and $F_o - F_c$ (**c**) difference density omit maps contoured at 0.8 and 2.0 r.m.s.d., respectively. Coloured as in Fig. 3.



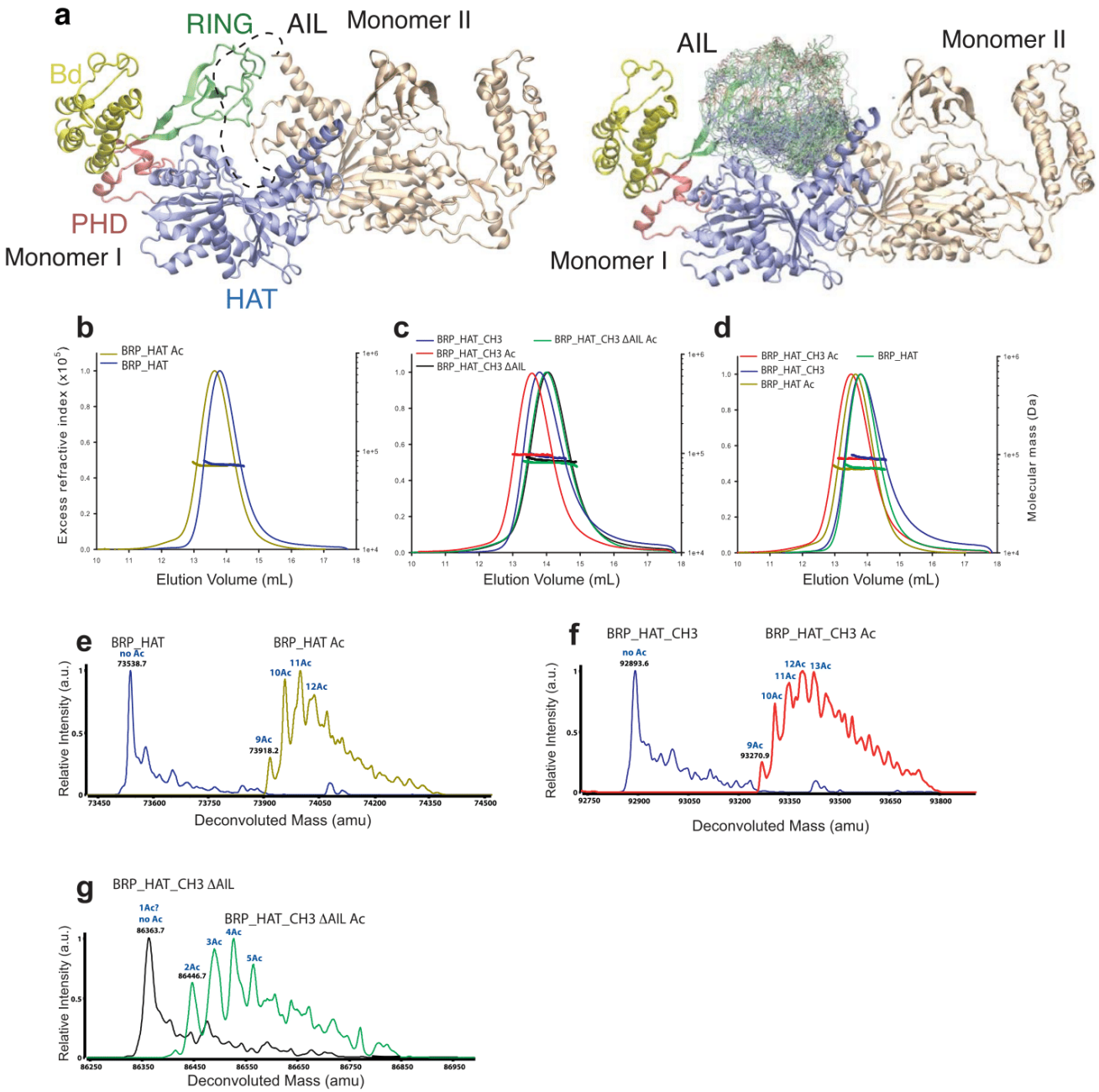
Extended Data Fig. 3 | Structural analysis of the RING domains. **a**, Superposition of the four p300 molecules (monomers I–IV) in the asymmetric crystallographic unit. Whereas the bromodomains (Bd), PHD and HAT domains superpose with a r.m.s.d. of approximately 0.9 Å, the RING domains adopt multiple conformations. **b**, $2F_o - F_c$ (blue mesh) and anomalous difference Fourier maps (orange mesh) for the four RING domains contoured around 1σ and 2.5σ, respectively.



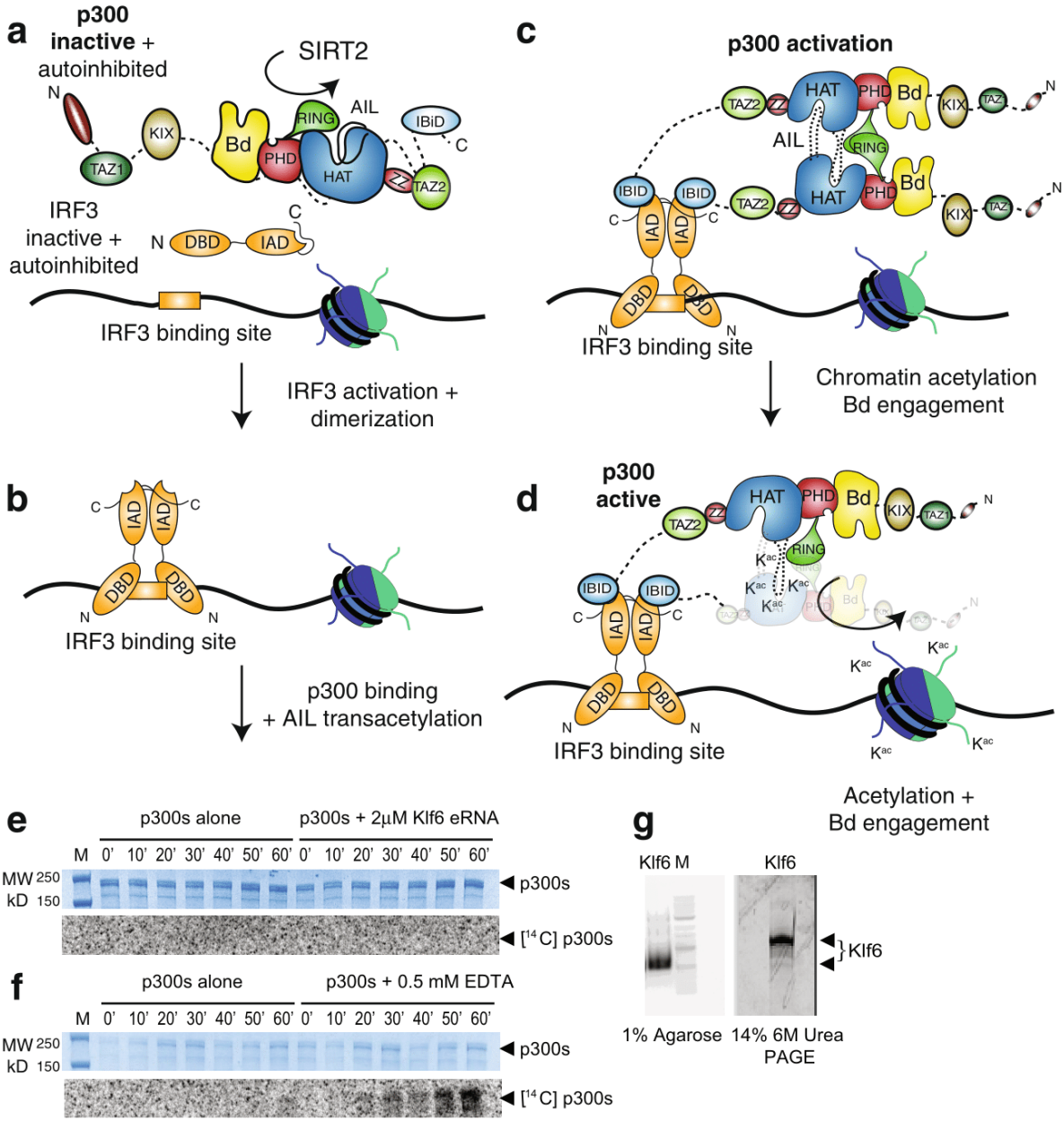
Extended Data Fig. 4 | Regulation of HAT activity by flanking domains. **a**, The domain structure of p300. Sequence conservation of the AIL is shown using WebLogo[54]. The constructs used are shown. **b**, Analysis of in vitro expression of the indicated p300 variants. Purified proteins were analysed for autoacetylation by immunoblotting with anti-p300(K1499ac) antibody (left), anti-Flag antibody (middle) and Coomassie staining (right). Representative data of three independent experiments are shown. **c**, Representative mass spectrometric analysis of BRP_HAT_ZZ_ΔAIL after in vitro expression (red) and after SIRT2 mediated deacetylation (black).



Extended Data Fig. 5 | Regulation of HAT activity by flanking domains. **a**, The AIL contributes to histone substrate acetylation of activated p300. The details of the constructs used are indicated in Extended Data Fig. 4. Defined amounts of p300 variants were incubated with acetyl-CoA and the indicated histones before SDS–PAGE analysis, followed by Coomassie staining and western blotting with the indicated antibodies. **b**, The indicated amounts of purified p300s variants were incubated with histone octamers as in **a**, followed by SDS–PAGE and immunoblot analysis with the indicated antibodies. Anti-Kac, pan-acetyl-lysine antibody. Representative data of three independent experiments are shown. **c**, Crystal structure of the H4(K12ac/K16ac) peptide bound to the BΔRP module containing an in-frame RING deletion. Amino acid residues 1169–1241 were replaced by a single glycine residue. The deletion removes the RING domain (black arrow) and does not adversely affect the structural integrity of the BΔRP module. **d**, **e**, Indicated variants of p300 were co-expressed with p53 in H1299 cells and analysed by immunofluorescence with the indicated antibodies (**d**) or by western blotting (**e**). Representative data of three independent experiments are shown. Scale bars, 10 μm.



Extended Data Fig. 6 | Autoacetylation changes the hydrodynamic properties of p300. a, Simulations of the AIL in the context of the loop-swapped dimer. Left, cartoon of the trajectory of the AIL (dashed line). Right, representative conformations with the AIL C_α backbone atoms are coloured according to charge. **b**, SEC–MALLS analysis of deacetylated (blue) and acetylated (yellow) p300 core. Note the decrease in elution volume upon acetylation. **c**, SEC-MALLS analysis of deacetylated (blue), acetylated (red) BRP_HAT_CH3 and deacetylated (black) and acetylated (green) BRP_HAT_CH3 ΔAIL. There is no increase in elution volume upon acetylation of the ΔAIL construct. **d**, Comparison of acetylated and deacetylated BRP_HAT and BRP_HAT_CH3. The deacetylated BRP_HAT (green) and deacetylated BRP_HAT_CH3 (blue) elute at the same position, which is indicative of a similar hydrodynamic radius. The acetylated BRP_HAT (yellow) and BRP_HAT_CH3 (red) elute at a larger elution volume. The normalized refractive index is plotted as a function of elution volume from an S200 column coupled to a MALLS detector. Calculated molecular masses are plotted as a function of volume for each eluted peak. The experiment was carried out at least three times with similar results. One representative example of each sample is shown. **e**, Mass spectrometry analysis (electrospray ionization) of the BRP_HAT before (blue) and after (yellow) autoacetylation. The molecular mass and the number of acetylation events are indicated. **f**, Mass spectrometry analysis of BRP_HAT_CH3 before (blue) and after (red) autoacetylation. **g**, Mass spectrometry analysis of BRP_HAT_CH3_ΔAIL before (black) and after (green) autoacetylation.



Extended Data Fig. 7 | Molecular model and controls showing that p300 acetyltransferase activity is not stimulated by eRNA. a, p300 is maintained in the inactive state by deacetylases such as SIRT2. IRF3 is autoinhibited by a C-terminal segment in the IAD domain. **b**, TBK1 phosphorylation activates and dimerizes IRF3. The activated IRF3 dimer engages the IBID domain of p300. **c**, Recruitment of two molecules of p300 results in *trans*-autoacetylation in the AIL loop and HAT activation. **d**, Activated p300 can acetylate chromatin and engage acetylated substrates via the bromodomain. **e**, A constant amount of p300s (2 μM) was incubated in [^{14}C]acetyl-CoA alone or in the presence of 2 μM Klf6 eRNA for the indicated time points. Samples were analysed by SDS–PAGE followed by Coomassie staining (top) and autoradiography (bottom). **f**, As in **e** but in the presence of 0.5 mM EDTA. The experiment was carried out at least twice with consistency. One representative example is shown. **g**, Quality control of Klf6 RNA. 3 μg Klf6 was deposited on a 1% agarose gel or a 14% 6 M urea PAGE gel and detected by SYBR Safe stain. M, 100-bp DNA ladder.

Extended Data Table 1 | Data collection, phasing and refinement statistics

	BRP-HAT	BΔRP
Data collection		
Space group	$P2_1$	$P2_12_12_1$
Cell dimensions		
a, *b*, *c* (Å)	100.7, 146.6, 116.3	49.6, 83.7, 165.6
α, β, γ (°)	90, 91.7, 90	90, 90, 90
Resolution (Å)	50–3.10*	42.7–2.50*
No. reflections	106462	23996
R_{sym} or R_{merge}	8.7 (89.7)*	6.3 (141.3)*
$I / \sigma I$	7.48 (0.7)*	7.6 (1.0)*
Completeness (%)	99.0 (94.0)*	97.29 (92.2)*
Redundancy	1.9 (1.9)*	9.1 (3.6)*
Refinement		
Resolution (Å)	50-3.1	42.7-2.5
R_{work} / R_{free}	0.19 / 0.26	0.24 / 0.27
No. atoms	19370	2769
Protein	19100	2693
Lys-CoA ligand	256	–
Zinc	14	4
H4 K12AcK16Ac	–	76
B-factors (mean; $Å^2$)		
Proteins	95.0	64.0
RING domains	164.8	
Lys-CoA ligand	69.2	–
H4 K12AcK16Ac	–	60.2
R.m.s deviations		
Bond lengths (Å)	0.013	0.004
Bond angles (°)	1.73	0.68

*Data from one crystal. Values in parentheses are for highest-resolution shell.

Extended Data Table 2 | Thermodynamic analysis of the interaction between p300 BRP and histone peptides by ITC

Protein	Peptide (residues)	K_d (μM)	N*
BRP	H3 unmodified (1-20)	--	No binding
	H3 K4me1 (1-20)	--	No binding
	H3 K4me3 (1-20)	--	No binding
	H3 K9ac (1-20)	--	No binding
	H3 K14ac (1-20)	1761 ± 356	1.13
	H3 K18ac (1-20)	large	1.01
	H3 K9acK14ac (1-20)	578 ± 47	1.06
	H3 K14acK18ac (1-20)	104 ± 27	1.03
	H3 K23ac (11-30)	--	No binding
	H3 K18acK23ac (11-30)	--	No binding
	H3 S10pho (1-20)	--	No binding
	H3 T11pho (1-20)	--	No binding
BRP	H4 unmodified (4-24)	--	No binding
	H4 K5ac (4-24)	--	No binding
	H4 K8ac (4-24)	58 ± 12	1.04
	H4 K5acK8ac (4-24)	828 ± 10	1.03
	H4 K8acK12ac (4-24)	90 ± 18	1.01
	H4 K12ac (4-24)	71 ± 14	1.05
	H4 K16ac (4-24)	--	No binding
	H4 K12K16diac (4-24)	25 ± 5	0.99
	H4 K20ac (1-24)	305 ± 6	1.00
	H4 K16acK20ac (1-24)	205 ± 10	1.23
	H4 K5K8K12K16K20penta-ac (1-24)	38 ± 4	1.07
	H4 S1pho K5K8K12K16K20penta-ac (1-24)	54 ± 2	1.12
BRP_N1132A	H4 K20ac (1-24)	--	No binding
	H4 K16acK20ac (1-24)	--	No binding
	H4 K5K8K12K16K20penta-ac (1-24)	--	No binding
	H4 S1pho K5K8K12K16K20penta-ac (1-24)	--	No binding
BRP	AIL (1545-1562)	--	No binding
BRP	AIL K1549K1558K1560tri-ac (1545-1562)	--	No binding

Mean and s.d. were determined from experiments performed in triplicate. Horizontal lines separate experiments involving different histone peptides or different protein constructs.
Histone peptide sequences: H3 (1–20) ARTKQTQRKSTGGKAPRKQL, H3 (11–30) TGGKAPRKQLATKASRKSAP, H4 (4–24) GKGGKGLGKGGAKRHRKVLRD.
AIL (Autoinhibitory loop peptide) SKNAKKKNNKKTSKNKSS (1545–1562). No binding was detected to non-acetylated peptides, or with a construct containing a mutation that abolishes acetyllysine binding (N1132A).
*Binding stoichiometry.

Extended Data Table 3 | Summary of SEC-MALLS and mass spectrometry experiments

Sample	MM_{MS} Da	MM_{th} Da	Acetylation level	MM_{SLS} Da (2 mg·ml^{-1})
BRP_HAT	73538	73538	~0	73380 ± 1.5%
BRP_HAT Acetyl	73918	73538	>8	71690 ± 1.6%
BRP_HAT_CH3	92893	92891	~0	92810 ± 2.0%
BRP_HAT_CH3 Acetyl	93270	92891	>9	90700 ± 2.0%
BRP_HAT_CH3 ΔAIL	86363	86362	~0	84650 ± 2.2%
BRP_HAT_CH3 ΔAIL Acetyl	86446	86362	>2	80450 ± 1.7%

Column labelling is as follows: MM_{MS}; molar masses determined by mass spectrometry; MM_{th}, the theoretical molar mass calculated from the appropriate primary sequences, acetylation levels were estimated based on the mass differences as compared to the non-acetylated sample; MM_{SLS}, molar masses determined by SEC-MALLS at a concentration of 2 mg ml^{-1}. All p300 constructs contained the mutation Y1467F. The experiment was carried out at least three times with consistency. Results from one representative example are shown. The mass and errors reported for SEC-MALLS are the weight average molar mass and residual standard deviations of the observed data from the fitted values calculated using ASTRA.

II. Histone H2A-H2B variants dimer-based sperm genome organization

Naghmeh Hoghoughi[1], Yuich Morozumi[1], Sophie Barral[1], Yohann Couté[2], Sandrine Curtet[1], Florent Chuffart[1], Guillaume Charbonnier[3], Denis Puthier[3], Thierry Buchou[1], Sophie Rousseaux[1], Hitoshi Kurumizaka[4], Saadi Khochbin[1*]

1 – CNRS UMR 5309; Inserm, U1209; Université Grenoble Alpes; Institut Albert Bonniot, Grenoble, F-38700 France

2- Université Grenoble Alpes; Inserm U1038; CEA, BIG-BGE, Grenoble, France.

3- TGML, platform IbiSA; Aix Marseille Univ, Inserm U1090, TAGC, Marseille, France,

4- Laboratory of Structural Biology, Graduate School of Advanced Science and Engineering; Research Institute for Science and Engineering; Institute for Medical-oriented Structural Biology, Waseda University, 2-2 Wakamatsu-cho, Shinjuku-ku, Tokyo 162-8480, Japan

Correspondence:

Saadi Khochbin: saadi.khochbin@univ-grenoble-alpes.fr

Summary

A genome-wide uncharacterized nucleosome disassembly accompanies histone-to-protamine replacement in maturing spermatozoa. Here we show that the nucleosome dissociation process in haploid male germ cells creates an opportunity for the testis-specific H2A.L.2-TH2B dimer to strongly interact with DNA, while H3-H4 tetramer is being displaced, generating a new dimer-based transitional structural unit, which persists in mature spermatozoa. Additionally, H2A.L.2 possesses the intrinsic ability to specifically target the pericentric regions of the genome leading to a stable maintenance of H2A.L.2-TH2B on these regions in spermatozoa. Our functional investigations indicate a critical role for RNA in the dynamic turnover of H2A.L.2 and in controlling its pericentric localization. These data not only highlight the unexpected ability of H2A-H2B species to organize the genome in the absence of histone H3-H4 tetramer but also reveal the heterogeneous nature of the histone-based genome organization of the male genome. This work has general conceptual implications in the understanding of the molecular basis of genome bookmarking as well as in the debated transmission of positional information to the zygote.

Introduction

The transformation of the universal basis of the genome organization in eukaryotes into a new protamine-based DNA packaging system during spermatogenesis remains one of the most fascinating open questions in biology. In mammals a near genome-wide histone removal occurs during the post-meiotic stages of spermatogenesis in specific cell types known as elongating and condensing spermatids. Histone removal is also proceeded by a genome-wide histone acetylation, especially on H4 lysine 5 and lysine 8 (Shiota et al., 2018), creating a specific binding site for the first bromodomain of the testis-specific BET factor, Brdt (Miller, Simon et al., 2016, Moriniere et al., 2009, Sasaki, Ito et al., 2009)We previously showed that Brdt is required in turn for the removal of histones and loading of nucleosomes by Transition Proteins (TPs) and protamines (PRMs) (Barral et al., 2017). Our work on the involvement of histone variants in this process also led us to the discovery of a series of testis-specific H2A and H2B histone variants as potentially interesting candidates in the control of post-meiotic male genome organization (Govin et al., 2007). One of these variants, which we named H2A.L.2, displays the remarkable property to be expressed during late spermatogenesis at the time of histone removal (Govin et al., 2007) and therefore appeared as a particularly interesting candidate for further studies. We demonstrated that this histone H2A variant dimerizes with the testis-specific H2B variants, TH2B (Govin et al., 2007; Montellier et al., 2013) and is co-expressed with TPs (Barral et al., 2017). We also observed that the transformation of nucleosomes into nucleoprotamins involves the appearance of transitional states, characterized by a shorter DNA fragment protected against MNase digestion (Govin et al., 2007; Barral et al., 2017). These studies questioned the dogma of the successive replacement of histones, first by TPs, themselves later replaced by PRMs. Indeed our *in vitro* and *in vivo* data strongly suggested that H2A.L.2-TH2B incorporation opens the nucleosomes, allowing TPs to be loaded. TPs do not displace the full histone complements from DNA but buffer and control the action of PRMs, which are the true histone displacers (Barral et al., 2017).

Although these investigations brought important clarifications on the molecular basis of the transformation of the nucleosomes into nucleoprotamines, the actual impact of these transitional states on the final organization of the male genome in spermatozoa has remained an open question.

Here we precisely address the issue of the fate of transitional states observed in condensing spermatids, during the establishment of the final structure of the mature male genome in maturing spermatozoa.

The most remarkable finding reported here is that H2A.L.2-TH2B presents the unique property to strongly interact with DNA, hence offering a resistance to the pressure of TP-PRM in histone displacement. We also show that, within a spermatozoa population, in a high proportion of the genomic positions associated with nucleosomes, sub-nucleosome particles are also present, suggesting a cell-to-cell variability in terms of full nucleosome retention. Our interpretation is that, in genomic positions escaping nucleosomes removal, the pressure from protamines would lead to a phenomenon of variegation in term of nucleosome integrity in any given position. Under these conditions, the capability of H2A.L.2-TH2B to strongly interact with DNA ensures a histone-based genome organization in the absence of full nucleosomes. We also demonstrate that H2A.L.2 has the intrinsic ability to target pericentric heterochromatin, differentiating these regions from the rest of the genome in mature spermatozoa. Finally, our findings shed new light on the real nature of the male genome organization in spermatozoa and bring new arguments feeding the debates on nucleosome retention in mature spermatozoa.

Results

H2A.L.2-TH2B dimer based spermatogenic cell genome organization

Previously we showed that extensive MNAse digestion of chromatin from condensing spermatids, in addition to the surviving nucleosomes, releases sub-nucleosomal fragments. These fragments correspond to dissociating nucleosomes, which we named “small structures” (Barral et al., 2017, Govin et al., 2007, Montellier et al., 2013).

In order to have unbiased information on the nature of the histones composing these structures, we set up two independent purification procedures based on the capture of DNA through hydroxy-apatite or the capture of basic proteins using heparin-coated beads. After extensive digestion of condensing spermatids chromatin by MNase, the released nucleosomes and the small structures were independently captured by hydroxy-apatite and heparin and the associated histone complement was identified (Fig. 1A). The same procedure was applied to nucleosomes from a mixture of total spermatogenic cells, and used as a reference for relative quantification of the identified histones in the surviving nucleosomes or in the small structures in the condensing spermatids.

The results obtained demonstrated a depletion of H3 and H4, as well as an enrichment of H2A.L.2/L1 species in the small structure (Fig. 1B).

These experiments demonstrate that during the invasion of nucleosomes by TP/PRMs and the subsequent histone displacement, histone octamer dissociates into TH2B-H2A.L.2 dimers and H3-H4 tetramers. Unexpectedly, under these conditions, TH2B-H2A.L.2 dimers remain bound to DNA and therefore captured by hydroxyapatite, while H3 and H4 are displaced (Fig. 1C). The depletion of H3 and H4 was also observed following the heparin capture of the small structures suggesting that these histones are not present in these fractions even under free forms.

These observations prompted us to propose the hypothesis that H2A.L.2-TH2B dimers may provide a basis for a yet uncharacterized male genome specific organizational unit, independent of H3 and H4. To this end, we decided to consider the biochemical properties of H2A.L.2-TH2B histone dimers compared to canonical H2A-H2B dimers.

H2A.L.2-TH2B dimers present enhanced DNA-binding compared to canonical H2A-H2B.

H2A.L.2 and TH2B were produced in bacteria, purified and used in different combinations to generate nucleosomes along with purified H3 and H4.

Thermal stability assays showed that H2A.L.2-TH2B dimer-containing nucleosomes are remarkably more stable than nucleosomes bearing H2A-H2B. H2A.L.2 and TH2B shared the ability to independently increase the stability of the corresponding nucleosomes but to a lesser extent compared to the H2A.L.2-TH2B (Fig. 2A).

In order to better understand the basis of the activity of the H2A.L.2-TH2B dimer, we decided to study the DNA-binding activity of different combinations of H2A-H2B dimers independently.

The binding of histone dimers with DNA was monitored by the formation of slow migrating bands in a gel-shift assay on a native gel. This experiment revealed a clearly enhanced DNA-binding property of the H2A.L.2-TH2B dimer compared to the other combinations of the considered H2A-H2B species (Fig. 2B). In all cases, compared to H2A-H2B, H2A.L.2-TH2B generated more intense shifted bands. In these sets of experiments, H2A-TH2B and H2A.L.2-H2B dimer combinations also resulted in a shifted band intensity which was intermediary between those observed when H2A-H2B (lowest) and H2A.L.2-TH2B (highest) were assayed (Fig. 2B).

These experiments revealed new properties of H2A.L.2-TH2B containing nucleosomes, which were found more stable than canonical nucleosomes. This result was unexpected since we previously showed that H2A.L.2-TH2B containing nucleosomes are in an open configuration ready to load TPs and that H2A.L.2-TH2B containing octamers are very unstable and break into dimers and tetramers (Barral et al., 2017). A stronger interaction of H2A.L.2-TH2B with DNA, as shown here, nicely explains this apparent discrepancy regarding the property of H2A.L.2-TH2B containing nucleosomes. These observations are indeed fully supported by a recent investigation considering the structures of partially unwrapped transient nucleosome intermediates which revealed that rearrangements in the H2A–H2B dimer strengthen interactions with the unwrapping DNA and promote nucleosome stability (Bilokapic, Strauss et al., 2018). Under these conditions, the replacement of H2A–H2B dimers by H2A.L.2-TH2B

dimers should confer a greater stability to the corresponding nucleosome because of the enhanced DNA-binding activity of this dimer.

Finally, we also incubated H2A.L.2-TH2B or H2A-H2B dimer-DNA complexes with purified TP1 and TP2 to mimic small structures in the context of condensing spermatids.

Figure 2C and 2D show that under these conditions, TPs do not strongly interact with naked DNA (lanes 1 - 4), but efficiently interact with dimers-DNA complexes (lanes 5 – 12). These experiments suggest that, compared to naked DNA, TP binding is favored in the presence of histone dimer-DNA complexes.

This stronger interaction of H2A.L.2-TH2B with DNA allows for the establishment of more stable histone-TP-DNA complexes as can be judged by a clear shift of H2A.L.2-TH2B-DNA complexes compared to that of H2A-H2B-DNA complexes (compare lanes 5 – 7 to lanes 10 – 12), which tend to dissociate in the presence of TPs (lanes 5 – 7, the smear above the naked DNA).

This mechanism could be important to enable TPs to dissociate nucleosomes, i. e., dissociating histone dimers from the H3-H4 tetramer (Fig. 1D), rather than displacing the whole histone octamer from DNA.

H2A.L.2 has the intrinsic property to target pericentric heterochromatin

Previously, we showed that H2A.L.2 is expressed at the same time as transition proteins and is required for the loading of TPs onto nucleosomes (Barral et al., 2017) However, considering spermatogenic cell stages that follow histone removal, a fraction of H2A.L.2 survives in restricted regions of the nucleus in condensing spermatids as well as in epididymal spermatozoa (Fig. 3A). Careful consideration of the sub-nuclear domains associated with the surviving H2A.L.2 reveals that they largely correspond to the DAPI-bright regions that are known to be mostly enriched in the pericentric major satellite DNA (Fig. 3A, condensing spermatids).

These observations indicate that, although upon its synthesis H2A.L.2 is incorporated in a genome-wide manner (Fig. 3A. elongating spermatids) and plays a role in the genome wide opening of nucleosomes and in the loading of TPs onto nucleosomes and the subsequent

nucleosome dismantlement by PRMs (Barral et al., 2017), the pericentric regions present a particular status. Indeed, in condensing spermatids, H2A.L.2 could be observed in the DAPI-bright pericentric regions, in contrast with its disappearance from the rest of the genome, a situation which persists in mature spermatozoa (Fig. 3A, condensing spermatids and spermatozoa).

However, a co-detection of H2A.L.2 and PRM indicates that the pericentric regions do not exclusively contain these histone-based structures and that they are also largely protaminized (Fig. 3A, spermatozoa).

In order to better understand the nature and origin of these structures in late spermatids and in maturing spermatozoa, we decided to deeply characterize their molecular properties and investigate the mechanisms directing their formation.

For this purpose, we used somatic cells, a totally irrelevant system, to know whether the pericentric localization of H2A.L.2 is specific to the context of spermatogenic cells or if it could also be observed in an unrelated context upon its ectopic expression in somatic cells.

To this end, we chose NIH3T3 cells, which display clear DAPI-bright pericentric heterochromatin foci (chromocenters), to ectopically express H2A.L.2. In parallel, we expressed H2A.B.3, a short H2A histone variant, which is structurally close to H2A.L.2, but in contrast to H2A.L.2, is expressed in meiotic and early post-meiotic cells and disappears later when H2A.L.2 is expressed (Soboleva et al., 2011). We also used the replication-dependent H2A as another control.

Figure 3B shows that H2A.L.2 is mostly observed in the pericentric regions in the completely out-of-context situation of somatic cells. Under the same conditions, despite important sequence and structural similarities between H2A.L.2 and H2A.B.3, the latter was unable to concentrate in the pericentric regions and showed a similar pattern of nuclear distribution as H2A.

Based on this experiment, we conclude that H2A.L.2 has the intrinsic property to target the pericentric heterochromatin and that no other testis-specific factor is required for this particular genomic targeting of H2A.L.2.

An additional control consisted in the expression of a chimeric protein bearing a major satellite-specific zinc finger fused to the VP16 activation domain (Ptvlp64) (gift from Maria-Elena

Toress Padillas). The expression of this protein leads to a complete disorganization of the chromocenters, as can be judged by the dispersion of the DAPI-bright regions in the nucleus. Interestingly, under these conditions the co-localization of H2A.L.2 and the DAPI-dense regions of the genome was completely lost (Fig. 3B). Since under these conditions the structural and the "epigenetic" nature of the heterochromatin is largely destroyed by the targeted activity of VP16, these observations exclude the specific recognition of the major satellite sequence by H2A.L.2, as the basis for the localized accumulation of this variant in the pericentric regions.

Pericentric localization of H2A.L.2 is controlled by its N-terminal RNA-binding motif.

H2A.L.2 belongs to a group of short H2A variants that all lack the typical H2A acidic patch, and present a shorter C-terminal docking domain than the canonical H2As (Molaro, Young et al., 2018, Soboleva et al., 2011). Additionally, H2A.L.2 shares with H2A.B.3 a remarkable enrichment in arginine residues (Fig. 4A), which, in H2A.B.3, was shown to be involved in binding RNA and controlling its intracellular localization (Anuar, Kurscheid et al., 2019).

In order to test the role of this putative RNA-binding domain at the N-terminal of H2A.L.2, we replaced this region by the corresponding region of H2A.

The ectopic expression of this mutant of H2A.L.2 (H2A.L.2-H2Anter) in NIH3T3 cells showed an enhanced localization of the protein in the pericentric regions, which was unexpected. Indeed, compared to wild-type H2A.L.2, the localization of the protein in the nucleoli as well the nuclear genomic background levels significantly decreased, leading to enhanced chromocenter localization of this H2A.L.2 mutant (Fig.4A).

These data indicate that the N-ter putative RNA binding domain of H2A.L.2 could be an important element in controlling the nuclear localization of the protein and that, in our test model of NIH3T3 cells, it could act by preventing the localization of H2A.L.2 in the heterochromatin regions.

The dynamic turnover of H2A.L.2 is controlled by its N-terminal RNA binding domain

To better characterize the role of this N-terminal domain in the nuclear localization of H2A.L.2, we also set up a FRAP approach to measure and compare the stability of the association between the wild-type or mutant form of H2A.L.2 with the pericentric regions. After photo-beaching, the recovery of wild-type H2A.L.2 at the pericentric regions was much faster than that of the mutant protein (Fig. 4B). An interpretation of this observation is that the RNA binding domain of H2A.L.2 would keep a fraction of the protein in the extra-chromocenter regions of the genome and hence would guarantee dynamic exchanges of the protein between the pericentric heterochromatin and the rest of the genome. This dynamics of H2A.L.2 localization would be lost upon the replacement of H2A.L.2 N-terminal by that of H2A.

A prediction of this model is that the ectopic expression of RNA in our test NIH3T3 cells stably expressing wild-type or mutant H2A.L.2, should result in a diffuse nuclear accumulation of wild-type H2A.L.2 but should not affect the localization of the mutant H2A.L.2. In order to test this hypothesis, we cloned a fragment of satellite DNA in an expression vector and transfected NIH3T3 cells stably expressing either wild-type or the N-terminal mutant of H2A.L.2. Figure 4C shows that, as predicted, the ectopic RNA expression in these cells enhance the diffuse nuclear distribution only of wild-type H2A.L.2.

Enhanced major satellite RNA expression and localization is associated with H2A.L.2 pericentric localization during late spermatogenesis

Our investigation of H2A.L.2 nuclear localization in somatic NIH3T3 cells, a context irrelevant to spermatogenesis, showed that the dynamics and efficiency of H2A.L.2 localization at the chromocenter regions depends on opposing properties of this H2A variant: its RNA-binding activity and its intrinsic ability to target pericentric heterochromatin. However, in the context of spermatogenesis, these two properties of H2A.L.2 could cooperate to stabilize the association of H2A.L.2 with pericentric regions. Indeed, an enhanced expression of major satellite RNA and its association with these regions could stabilize H2A.L.2 at the pericentric regions during the final phases of spermatogenesis.

We used an RNA-seq approach to compare the transcriptome of spermatogenic cells at spermatocyte, round spermatids and condensing spermatids. To test the role of H2A.L.2 itself in inducing the expression of the major satellite RNA, we also considered the transcriptome of the corresponding H2A.L.2 KO cells. Figure 5A shows that in both wild-type and H2A.L.2 KO cells the expression of sense (forward) and anti-sense (reverse) satellite RNA significantly increased in condensing spermatids. These results showed that H2A.L.2 incorporation at the pericentric region is not needed for the activation of satellite RNA expression, since the increase in major satellite RNAs was also observed in the absence of H2A.L.2.

Accordingly, in NIH3T3 cells, we also observed that the pericentric localization of H2A.L.2 did not induce the expression of these regions, while the expression of Ptvlp64, capable of targeting pericentric regions, is capable of efficiently inducing pericentric RNA expression as can be judged by the accumulation of both forward and reverse major satellite transcripts (Fig. 5B).

An increased targeted accumulation of satellite RNA at the pericentric regions could re-enforce the intrinsic ability of H2A.L.2 to associate with these regions.

Accordingly, we set up a FISH-based detection of major satellite RNA on the epididymal maturing spermatozoa from wild-type and H2A.L.2 KO mice. Figure 5C shows that in both cases a localized accumulation of sense and anti-sense major satellite RNAs could be observed.

Considering all the data on the major satellite RNA expression and pericentric H2A.L.2 localization, it appears that when in contrast to NIH3T3 cells, RNA accumulates at the pericentric regions, the combined intrinsic ability of H2A.L.2 to target the pericentric regions as well as its RNA-binding activity cooperate to target H2A.L.2 to and stabilize its localization in these regions (Fig. 5D).

Alternative nucleosomal and sub-nucleosomal structures in the organization of spermatozoa genome

Our data highlighted the occurrence of H3-H4 less H2A.L.2-TH2B dimer-based genome organization in condensing spermatids. Although our subsequent investigations pointed to pericentric regions as a major genomic location of H2A.L.2-containing genomic regions, we

cannot exclude the presence of non-nucleosomal genome-organizing structures in other genomic regions.

In fact, considering in situ visualization experiments, the particular nature of the pericentric heterochromatin organization, including the repetitive nature of the associated DNA sequences, could generate the visual impression of H2A.L.2 accumulation on these regions while minimizing the presence of H2A.L.2 and sub-nucleosomal structures on other genomic regions in mature spermatozoa.

In order to tackle this issue, we decided to identify all genomic regions respectively enriched in the nucleosomal and small structures in condensing spermatids as well as in maturing spermatozoa. The question was whether the nucleosomal regions and small structures mapped in condensing spermatids would respectively remain nucleosomal or non-nucleosomal in the spermatozoa.

The nuclei of maturing spermatozoa obtained by the swim-up protocol were decondensed and digested to different extents by MNase. Figure 6A shows that MNase digestion of the spermatozoa chromatin releases non-nucleosomal DNA, which is sensitive to the extent of MNase digestion, as well as nucleosomal fragments, which present greater stability towards MNase digestion. After 4 min of MNase digestion we obtain heterogeneous sub-nucleosomal DNA fragments, whereas longer MNase digestion (12 min) resulted in more defined sub-populations of DNA fragments. We sequenced the nucleosomal as well as the sub-nucleosomal fragments obtained after both MNase digestion times (as indicated in Fig. 6A).

The DNA fragment populations released after the short MNase digestion very likely also contained protaminized DNA of heterogeneous sizes (Fig. 6A, 4 min), whereas a longer MNase digestion released more structured and MNase-resistant genomic regions (Fig. 6A, 12 min). In order to better characterize the genome positioning of these structures within the maturing sperm genome, we compared their genomic positions with those of the strongly localized 4000 nucleosomes, defined previously round spermatids (Barral et al., 2017).

The genomic regions corresponding to nucleosomes with strong positioning in round spermatids were identified (Fig. 6B, panel 2) and the presence of nucleosomes was then investigated in the preceding stage, spermatocytes (Fig. 6B, panel 1) as well as in the following stages, condensing spermatids (Fig. 6B, panel 3 and 7).

Considering each specific nucleosome-rich region of the genome in round spermatids, condensing spermatids are heterogeneous in terms of nucleosomes or small structures, since some cells retain the positioned nucleosomes, while others lose these nucleosomes but gain defined sub-nucleosomal structures (small structures) on the same precise regions (Fig. 6B, compare panels 3 and 7). Interestingly, small structures could appear anywhere within each nucleosome-rich region as can be attested by the diagonal shape of their alignment when they are placed from the 5' to the 3' ends of the corresponding nucleosome position. Remarkably, in most of the considered genomic regions, the same exact positions can be occupied by nucleosomes and small structures (Fig. 6B, panel 3 and 7). Considering condensing spermatids, the presence of both nucleosomes and small structure on each given genomic position could be due to two alternative possibilities. First, it could be due to the heterogeneity of condensing spermatids in terms of stages (step 12 to 16), and nucleosomes could be present in early condensing spermatids, while they would be replaced by small structures in late condensing spermatids. Second this heterogeneity could be a characteristic of all genomic regions in post-meiotic cells and remain so in sperm cells. Our observation of a similar heterogeneity in maturing spermatozoa (Fig. 6B, panels 4, 5, 6, 8 and 9) strongly suggest that the latter hypothesis is correct: for a given position, we may have either full nucleosomes or sub-nucleosomal structures. Indeed, considering each genomic position, condensing spermatids and maturing spermatozoa are heterogeneous populations, with each cell containing either a nucleosome or a small structure.

Based on these data, we propose that, as soon as in the early condensing spermatids, a proportion of nucleosomes resists removal and remain intact in some cells, whereas in other cells they become transformed into small structures and remain under this state in mature spermatozoa.

This heterogeneity between cells was observed here for about 2/3 of the 4000 regions corresponding to strongly positioned nucleosomes in round spermatids. However, approximately 1/3rd of these well-positioned nucleosome-rich regions fully retain nucleosomes in condensing spermatids and later in mature spermatozoa. These are nucleosomes that fully escape the transformation into small structures.

Overall these studies reveal the unexpected nature of spermatozoa genome organization and give compelling evidence that the notion of nucleosome removal and retention in spermatozoa

needs to be revised. Indeed, although some nucleosomes remain intact in some cells, in other cells only H2A-H2B dimer-based structures are observed at these positions.

Discussion

During post-meiotic re-organization of the male genome, the invasion of H2A.L.2-containing nucleosomes by protamines, modulated by a simultaneous incorporation of TPs leads to the generation of transitional genome organizational states, which are characterized by an altered genomic DNA protection against MNase (Barral et al., 2017). Here we could capture and purify these transitional states thanks to hydroxyapatite, which binds DNA, or heparin, which binds histones. A comprehensive and unbiased identification of the histone complement associated with these sub-nucleosomal DNA fragments, clearly shows the presence of a H2A.L.2-TH2B dimer bound to DNA and the displacement of the H3-H4 tetramer.

Subsequent biochemical investigations demonstrated a particular and unexpected role for H2A.L.2-TH2B in a nucleosome. Previously, we had shown that, in agreement with other published data on the short H2A variants (Arimura et al., 2013, Bao et al., 2004, Doyen et al., 2006, Soboleva et al., 2011, Syed et al., 2009), H2A.L.2 generates open nucleosomes protecting only about 130 bp of DNA against MNase digestion (Barral et al., 2017). We also observed that, compared to canonical nucleosomes, this H2A.L.2-TH2B dimer-containing open nucleosome is a better "receptor" for TPs. However, further investigations reported here suggest that these H2A.L.2-TH2B dimer-containing nucleosomes are also more stable than canonical nucleosomes. The fact that, compared to H2A-H2B dimers, H2A.L.2-TH2B dimers show the remarkable property to strongly interact with DNA, gives a key to the observed increased stability of the corresponding nucleosomes under thermal stability assays. Indeed, a recent report demonstrates that partial nucleosome unwrapping induces a rearrangement of the H2A-H2B dimer, strengthening its interaction with DNA, stabilizing the nucleosome (Bilokapic et al., 2018) Hence the basic capacity of the H2A.L.2-TH2B dimer to strongly interact with DNA perfectly explains the greater thermal stability of the H2A.L.2-TH2B-containing nucleosomes in the light of the data reported by Bilokapic and colleagues.

Based on these data sets, we can now portray the nucleosomes that undergo histone-to-protamine replacement. The incorporation of H2A.L.2-TH2B dimers into the nucleosomes creates open nucleosomes and relaxed chromatin. TPs, together with PRMs, can be loaded onto these nucleosomes, which start unwrapping the nucleosomal DNA and dissociating the octamers. Indeed, we previously showed that H2A.L.2-containing octamers are extremely unstable (Barral et al., 2017). However, under this condition, the enhanced interaction of

H2A.L.2-TH2B with DNA, similarly to what was reported by Bilokapic and colleagues (Bilokapic et al., 2018), also maintains H2A.L.2-TH2B strongly associated with DNA, while H3-H4 are being displaced by the TP-PRM complex. This stable H2A.L.2-TH2B dimer DNA binding therefore forms the basis of a new DNA-packaging structure, which persists in mature spermatozoa. Accordingly, the MNase digestion of chromatin from both condensing spermatids and spermatozoa and the mapping of sub-nucleosomal DNA fragments in these cells revealed that a large proportion of genomic regions associated with small structures remains sub-nucleosomal in spermatozoa, strongly suggesting that these H3-H4 tetramer-less H2A.L.2-TH2B dimer based structures create yet uncharacterized spermatozoa genome organizational units.

An important message of this work is also that these small structures and nucleosomes could actually be positioned on the same exact region in different cells. This observation reveals the occurrence of a cell-to-cell heterogeneity of genome organization on given positions. Indeed, our data clearly show that both nucleosomal and sub-nucleosomal structures could be present at the same position, indicating that in some cells a given position is nucleosomal whereas in other cells it is occupied by small structures.

This situation appears very similar to the so-called “position effect variegation” in somatic cells, where the heterochromatin boundary shows a cell-to-cell variability affecting the expression of the nearby genes. In the present situation, surviving nucleosomes adjacent to protaminized regions are submitted to the pressure of the neighbouring protamines, which promote nucleosomal dismantlement, but, even in cases of successful nucleosome dissociation, full displacement is simultaneously opposed by the intrinsic ability of H2A.L.2-TH2B dimers to strongly interact with DNA. This situation leads to a variegation effect and a cell-to-cell variability in nucleosome retention affecting well-positioned nucleosomal genomic regions, with nucleosomes fully retained in some cells, and occupied by smaller sub-nucleosomal structures in other cells.

This finding has important implications on the controversial data published by different groups regarding the specific positions of retained nucleosomes in mature spermatozoa (Brykczynska et al., 2010, Carone et al., 2014, Erkek et al., 2013, Samans et al., 2014, Yamaguchi, Hada et al., 2018, Yoshida et al., 2018). Indeed, the preferential genomic positions of retained nucleosomes in sperm cells is still a subject of debates. Our observations suggest that the same positions could be occupied by nucleosomal as well as sub-nucleosomal structures and that the

state of this heterogeneity within a spermatozoa population could impact the outcome of MNase digestion of their nuclei. Our data also show that sub-nucleosomal structures are more sensitive to MNase digestion than nucleosomes which explains why the extent of MNase digestion could affect the outcome of the studies, as proposed by Saitou and Kurimoto (Saitou & Kurimoto, 2014) and Yoshida and colleagues (Yoshida et al., 2018). To avoid this problem these authors proposed an MNase-free method for mapping mature spermatozoa histone H3-containing genomic regions.

Our microscopy investigations also revealed that pericentric regions display a specific organization characterized by the preferential retention of H2A.L.2, which persists from condensing spermatids to the final stage of mature spermatozoa. Due to the repetitive nature of the pericentric genomic regions, it is difficult to quantitatively estimate the amounts of H2A.L.2 present on these regions compared to the rest of the genome, and therefore our conclusions should mostly be based in situ data. However, the specific structure of these regions and various parameters such antibody accessibility could also enhance this preferential localization of the signal. Indeed, it has been shown that the reported almost exclusive enrichment of macroH2A histone variants on the inactive X chromosome observed by microscopy, was exaggeratedly enhanced compared to macroH2A enrichment on the X chromosome measured by ChIP-seq (Mietton, Sengupta et al., 2009, Perche, Vourc'h et al., 2000). Therefore, despite the *in situ* preferential detection of H2A.L.2 over the pericentric regions, we also expect its presence on other genomic regions, in agreement with the detection of small structures on unique genomic sequences in mature spermatozoa reported here. However, the preferential accumulation of H2A.L.2 in the pericentric regions, is strongly supported by the remarkable property of H2A.L.2, which is capable of targeting pericentric regions upon its ectopic expression in the completely unrelated context of somatic cells. Inspired by published work on another member of testis-specific short H2A variant expressed in spermatocytes and in round spermatids, H2A.B.3 (Soboleva et al., 2017) with a characterized RNA-binding motif, we observe here that H2A.L.2 targeting of the pericentric heterochromatin can be controlled by RNA. Indeed, in somatic cells, the replacement of H2A.L.2 putative RNA-binding domain enhances its pericentric localization and stabilizes its association with these regions. However, in late spermatids and in spermatozoa, as opposed to somatic cells, the pericentric regions not only become expressed but also the corresponding RNAs remain associated with these regions. Based on our detailed analysis of H2A.L.2 nuclear localization

in somatic cells, we can propose that, in sperm cells, both the intrinsic ability of H2A.L.2 to target pericentric regions and its RNA binding activity cooperate to stabilize this variant in the pericentric regions.

Additionally, the co-immunodetection of PRM and H2A.L.2 shows that, despite pericentric regions being preferentially associated with H2A.L.2, these regions are also protaminized, suggesting that, even within these regions, a mixture of different types of DNA-packaging structures should be present. This particular organization of the pericentric regions involving H2A.L.2 could have important implications for the structuration of the male heterochromatin post-fertilization (Borsos & Torres-Padilla, 2016, Probst et al., 2010), but this question, similarly to the issue of position-specific nucleosome retention discussed by many groups (see the references above), remains open.

The investigations reported here shed new light on the organization of the male genome and in particular on a decade of debates over the existence and nature of genomic regions preferentially associated with nucleosome retention in mature spermatozoa.

The appearance of such properties associated with late expressing H2A variant, H2A.L.2 supports the importance of histone retention in the organization of the final male genome and epigenome and the transmission of a positional information to the zygote.

Figure. 1

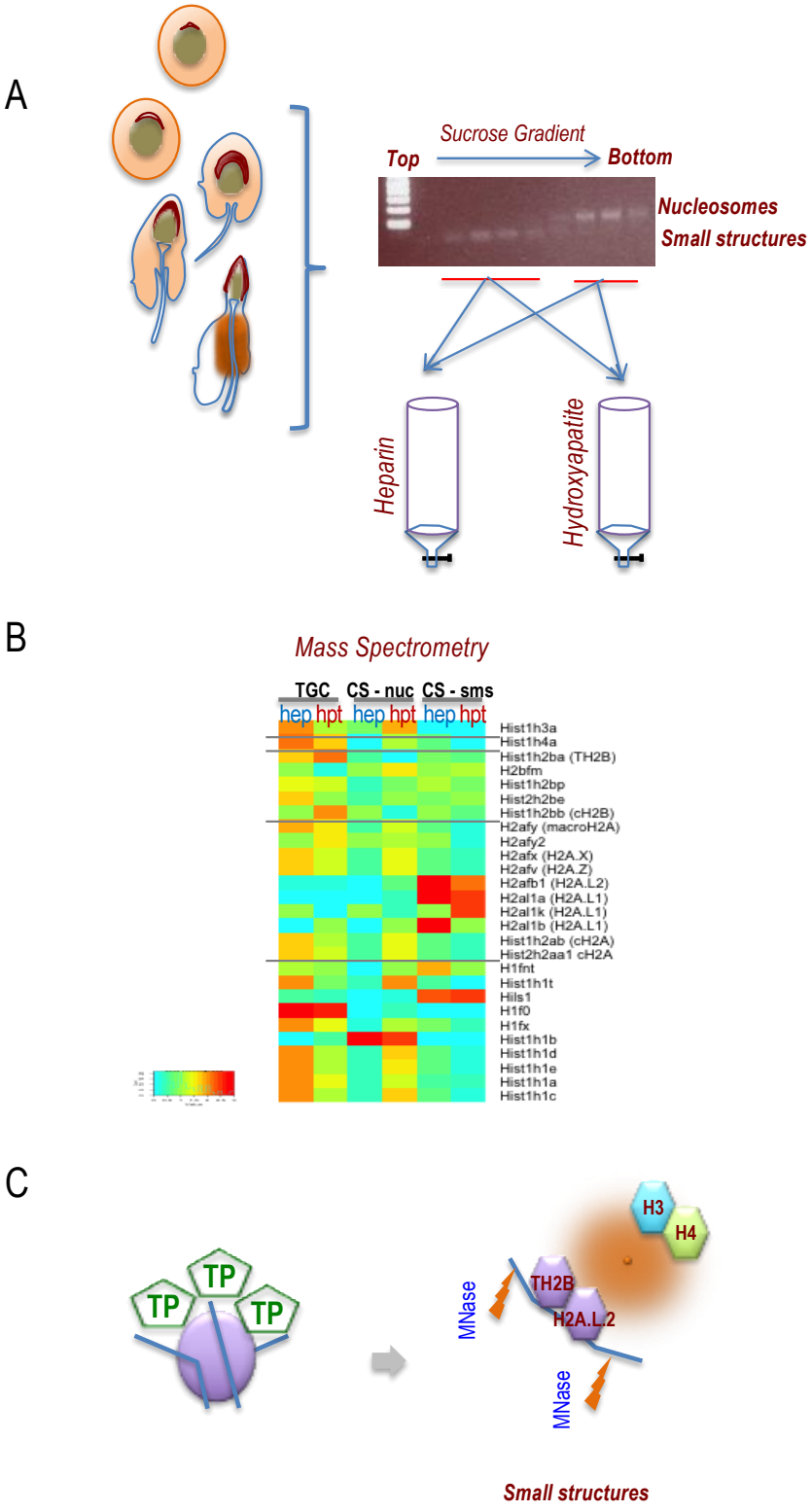

A
Sucrose Gradient
Top
Bottom
Nucleosomes
Small structures
Heparin
Hydroxyapatite
B
Mass Spectrometry
TGC
CS - nuc
CS - sms
hep hpt hep hpt hep hpt
Hist1h3a
Hist1h4a
Hist1h2ba (TH2B)
H2bfm
Hist1h2bp
Hist2h2be
Hist1h2bb (cH2B)
H2afy (macroH2A)
H2afy2
H2afx (H2A.X)
H2afv (H2A.Z)
H2afb1 (H2A.L2)
H2al1a (H2A.L1)
H2al1k (H2A.L1)
H2al1b (H2A.L1)
Hist1h2ab (cH2A)
Hist2h2aa1 cH2A
H1fnt
Hist1h1t
Hils1
H1f0
H1fx
Hist1h1b
Hist1h1d
Hist1h1e
Hist1h1a
Hist1h1c
C
TP
TP
TP
H3
H4
TH2B
H2A.L.2
MNase
MNase
Small structures

Figure. 2

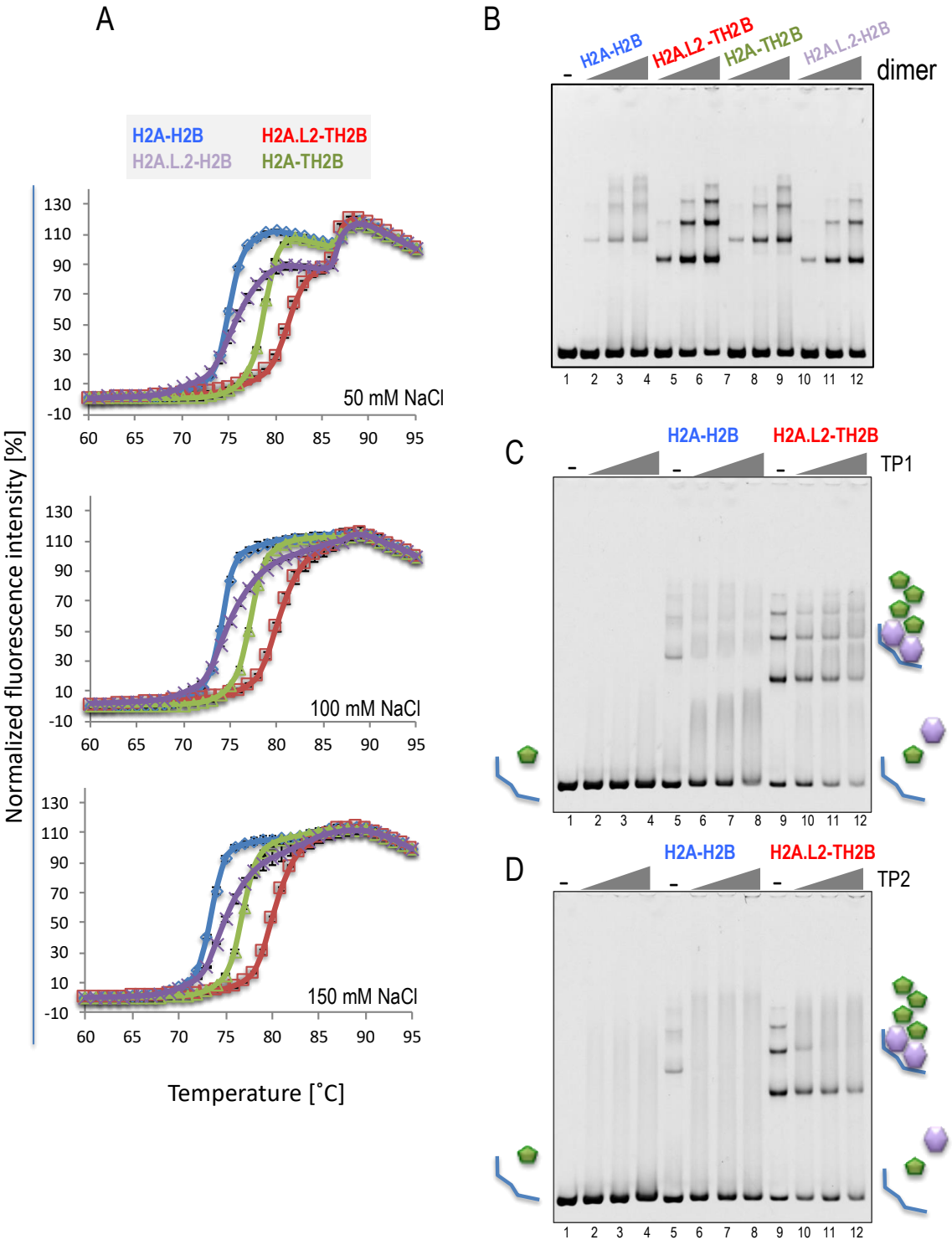

A
B
C
D
H2A-H2B
H2A.L2-TH2B
H2A.L.2-H2B
H2A-TH2B
dimer
TP1
TP2
50 mM NaCl
100 mM NaCl
150 mM NaCl
Normalized fluorescence intensity [%]
Temperature [˚C]

Figure. 3

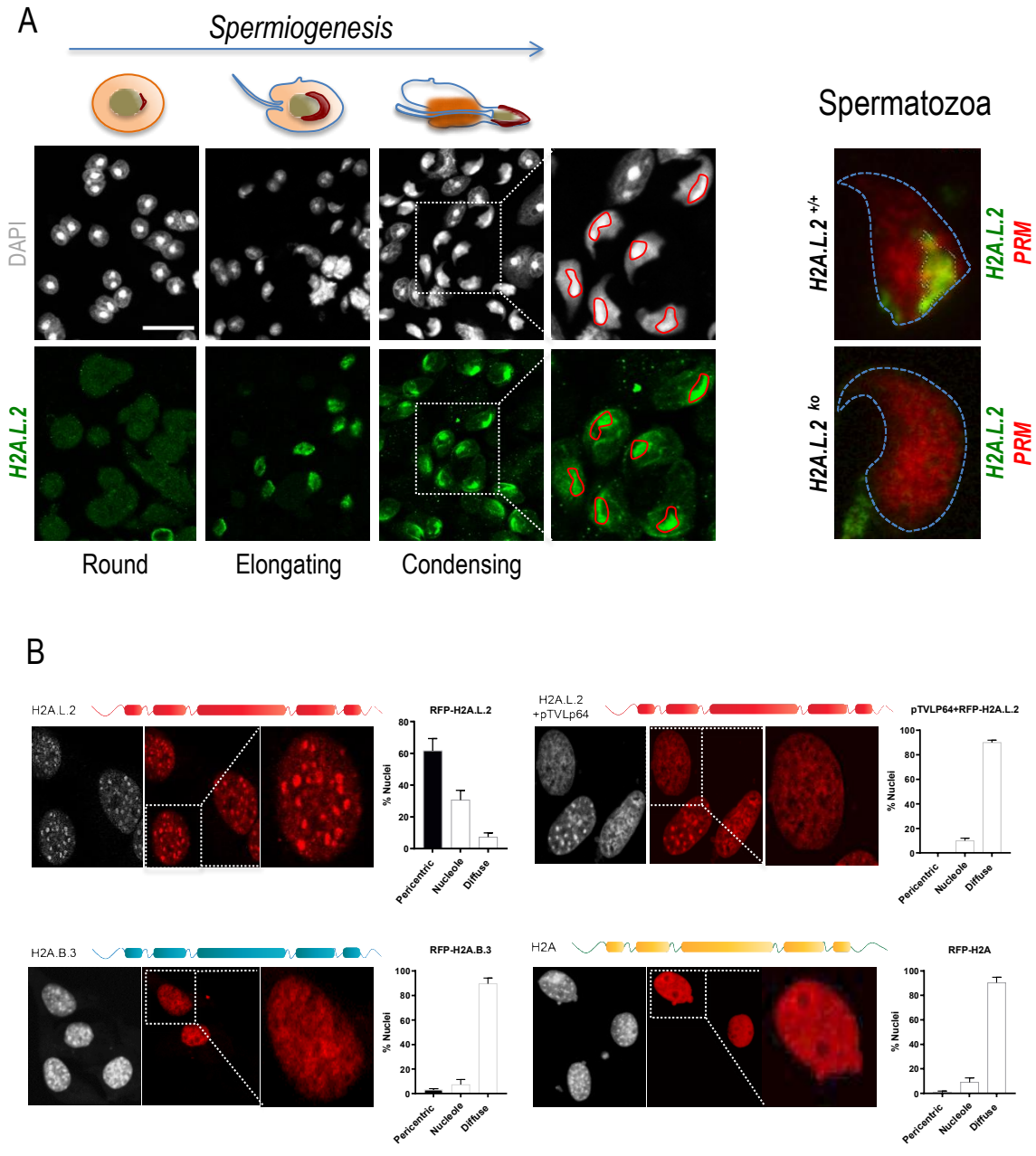

A
Spermiogenesis
Spermatozoa
DAPI
H2A.L.2
H2A.L.2 +/+
H2A.L.2 ko
H2A.L.2 PRM
H2A.L.2 PRM
Round
Elongating
Condensing
B
H2A.L.2
RFP-H2A.L.2
H2A.L.2 +pTVLp64
pTVLP64+RFP-H2A.L.2
H2A.B.3
RFP-H2A.B.3
H2A
RFP-H2A
% Nuclei
Pericentric
Nucleole
Diffuse

Figure 4.

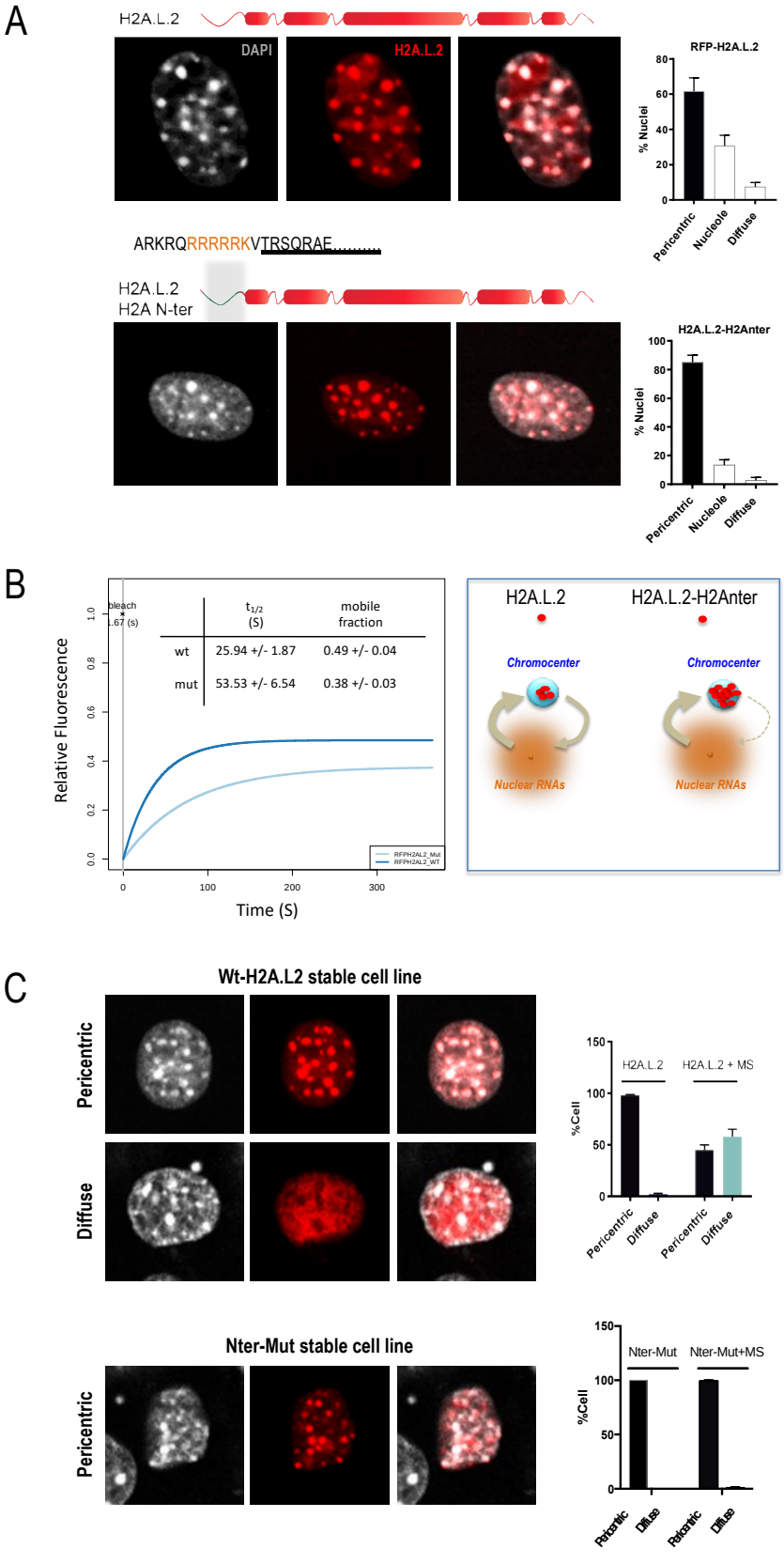

A
H2A.L.2
DAPI
H2A.L.2
RFP-H2A.L.2
% Nuclei
Pericentric
Nucleole
Diffuse
ARKRQRRRRRKVTRSQRAE.....
H2A.L.2
H2A N-ter
H2A.L.2-H2Anter
B
bleach
1.67 (s)
t1/2 (S)
mobile fraction
wt
25.94 +/- 1.87
0.49 +/- 0.04
mut
53.53 +/- 6.54
0.38 +/- 0.03
Relative Fluorescence
Time (S)
H2A.L.2
H2A.L.2-H2Anter
Chromocenter
Nuclear RNAs
C
Wt-H2A.L2 stable cell line
Pericentric
Diffuse
%Cell
H2A.L.2
H2A.L.2 + MS
Nter-Mut stable cell line
Nter-Mut
Nter-Mut+MS

Figure. 5

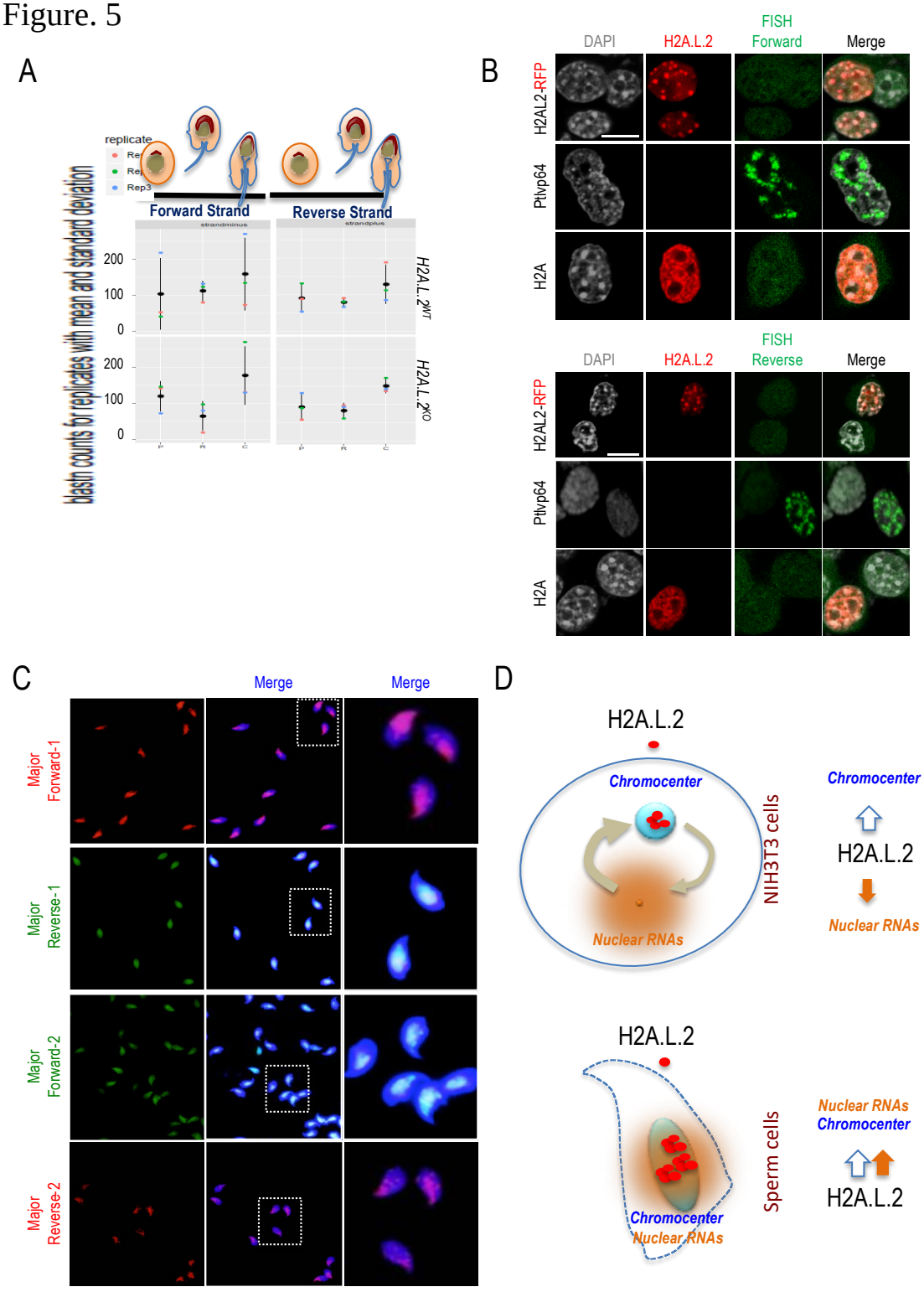
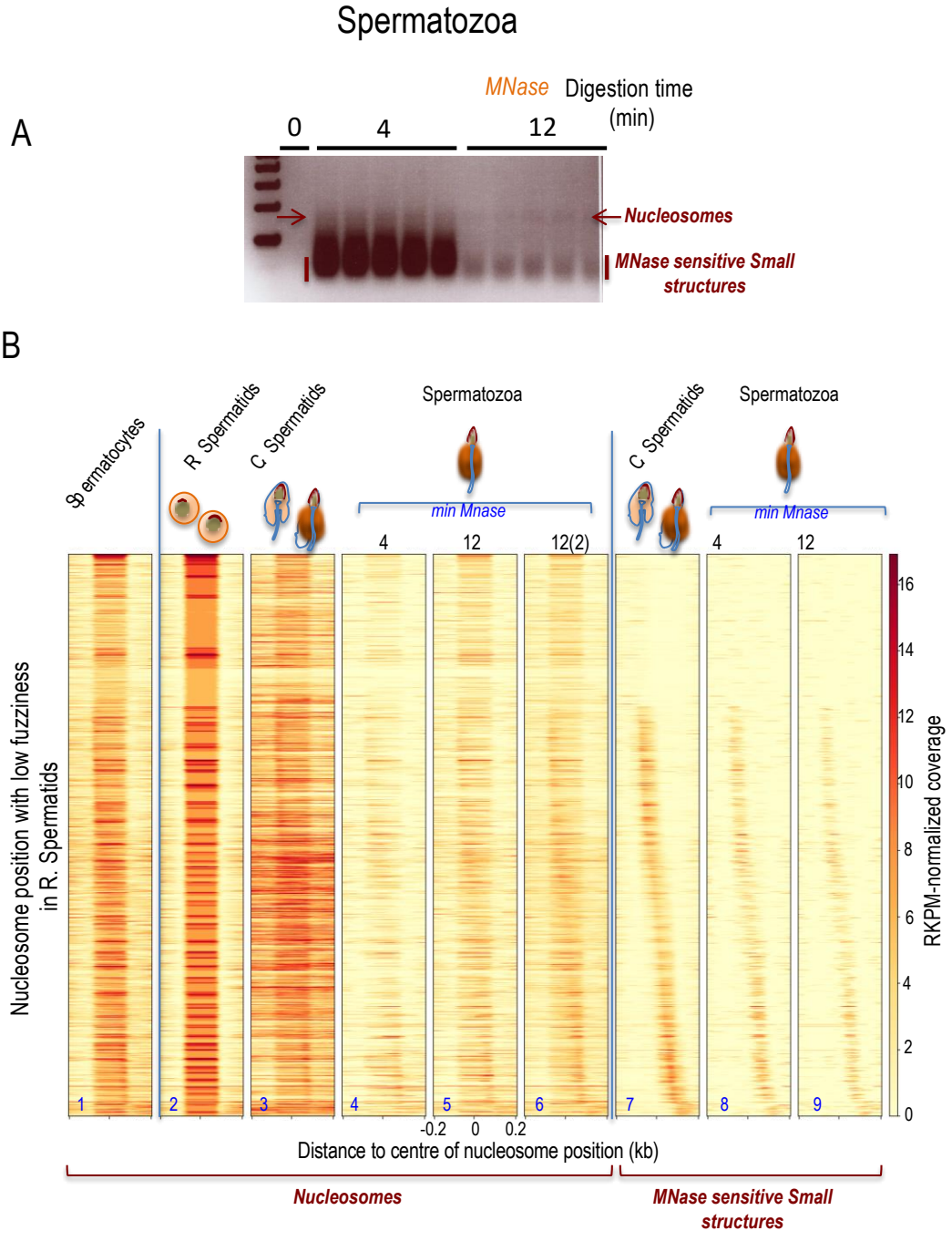

A
B
C
D
replicate
Rep3
Forward Strand
Reverse Strand
blastn counts for replicates with mean and standard deviation
H2A.L.2WT
H2A.L.2KO
DAPI
H2A.L.2
FISH
Forward
Reverse
Merge
H2AL2-RFP
Pttvp64
H2A
Major
Forward-1
Reverse-1
Forward-2
Reverse-2
H2A.L.2
Chromocenter
Nuclear RNAs
NIH3T3 cells
Sperm cells

Figure. 6


Spermatozoa
A
MNase Digestion time (min)
0
4
12
Nucleosomes
MNase sensitive Small structures
B
Spermatocytes
R Spermatids
C Spermatids
Spermatozoa
min Mnase
4
12
12(2)
C Spermatids
Spermatozoa
min Mnase
4
12
Nucleosome position with low fuzziness in R. Spermatids
RKPM-normalized coverage
16
14
12
10
8
6
4
2
0
1
2
3
4
5
6
7
8
9
-0.2
0
0.2
Distance to centre of nucleosome position (kb)
Nucleosomes
MNase sensitive Small structures

Legend to the Figures

Figure 1: Characterization of histone complement associated with surviving nucleosomes and small structures in condensing spermatids

A) Nuclei from condensing spermatids were digested by MNase to release the surviving nucleosomes and small structures. The released nucleoproteic structures were then separated on a sucrose gradient and different fractions analysed on an agarose gel to visualize the corresponding DNA fragment sizes.

B) Nucleosomal (nuc) and subnucleosomal (smst) fractions were obtained by MNase digestion of condensing spermatids (CS) as described in A. Nucleosomal fractions of total male germ cells (TGC) were also prepared. The indicated fractions were pooled and applied to either a hydroxyapatite colon (hpt) or to an agarose-heparin (hep) colon. The retained materials were then eluted and subjected to mass spectrometry identification of the proteins and quantification estimation by using specific peptide counts. The histones were specifically considered. The variations in the relative amounts of each histone species in the different samples were calculated after normalizing all peptide counts values with the mean peptide counts value in all three samples. The obtained relative amounts of individual histones were colour-coded as indicated and represented on the heatmap.

C) Schematic representation of the situation of histones octamer in the small structures as can be judged from the proteomics experiments.

Figure 2: Enhanced DNA-binding by H2A.L.2-TH2B histone dimer

A) Graphical representation of thermal stability assay of the nucleosomes containing the indicated combination of histone dimers in the presence of 50, 100 and 150 mM NaCl. The intensity of SYPRO orange fluorescence at each temperature was normalized relative to that at 95˚C, and averages of three independent experiments are plotted against temperature ranging from 60˚C to 95˚C, with standard deviation values.

B) The 145-bp Widom 601 DNA (20 ng/µl) was incubated with the indicated combinations of histone dimers and the samples were then analysed by 6% native polyacrylamide gel

electrophoresis (PAGE). The concentrations of histone dimers were 0 µM (lane 1), 0.3 µM (lanes 2, 5, 8 and 11), 0.6 µM (lanse 3, 6, 9 and 12) and 0.9 µM (lanes 4, 7, 10 and 13).

C and D) The 145-bp Widom 601 DNA (20 ng/µl) and the indicated combinations of histone dimers (0.9 µM) were incubated with either TP1 (C) or TP2 (D) and the samples were analysed by 6% PAGE. The concentrations of TP1 or TP2 were 0 µM (lanes 1, 5 and 9), 0.6 µM (lanes 2, 6 and 10), 0.9 µM (lane 3, 7 and 11) and 1.2 µM (lanes 4, 8 and 12).

Figure 3: Pericentric localization of H2A.L.2 in condensing spermatids

A) The intranuclear localization of H2A.L.2 was monitored by immunofluorescence during three successive stages of post-meiotic spermatogenesis. From left to right, round, elongating and condensing spermatids are shown after DAPI-staining (upper panels) and H2A.L.2 immunostaining (lower panels). The indicated squares are magnified and shown on the right. The DAPI-bright pericentric regions are delimited by a red line. The right panels also show the co-immunodetection of H2A.L.2 and PRM in the epididymal spermatozoa.

B) The indicated RFP-tagged H2As were expressed in NIH3T3 cells and the intranuclear localization of the proteins visualized. In one experiment Ptvlp64 synthetic protein corresponding to the major satellite-specific zinc finger fused to the VP16 activation domain was co-expressed with RFP-H2A.L.2. The histograms on the right represent the proportions of cells with the indicated intranuclear localization of the considered histones.

Figure 4: RNA-controlled dynamic pericentric localization of H2A.L.2

A) NIH3T3 cell lines expressing RFP-H2A.L.2 (upper scheme) or a mutated version of RFP-H2A.L.2 with the replacement of its N-terminal sequence with that of H2A, named H2A.L.2-H2Anter (lower scheme), were established and the intracellular localization of the expressed proteins monitored. The histograms represent the proportions of cells showing the indicated intranuclear localizations of the considered proteins.

B) The two cell lines described above were used to monitor protein dynamics by Fluorescence Recovery After Photo-bleaching (FRAP). Ten RFP-H2A.L.2 chromocenter foci (wild-type or mutated) were photo-bleached and the recovery of fluorescence monitored. The ten datasets of

each experiment were individually fitted and used to calculate the half-life ($t_{1/2}$) of fluorescence recovery and the mobile fractions (indicated in the inset). The mean and standard deviation measured at successive time points from the ten datasets are plotted. The scheme on the right depicts the working hypothesis on the role of the H2A.L.2 putative RNA-binding domain in the dynamics of H2A.L.2 as measured by FRAP. Nuclear RNAs ensure the existence of a pool of H2A.L.2 out of chromocenters, capable of supplying the chromocenters after photo-bleaching due to the intrinsic capacity of H2A.L.2 to target the pericentric heterochromatin. H2A.L.2-H2Anter, unable to interact with RNA, is directly targeted to the pericentric regions leaving the cell devoid of any pool of immediately accessible H2A.L.2 to re-populate the chromocenters after photo-bleaching.

C) A fragment of major satellite DNA was cloned in an expression vector under the CMV promoter and expressed in the above-mentioned stably transfected cells lines. The respective proportions of nuclei corresponding to each of the two categories of intranuclear distributions of signal for H2A.L.2 (upper panel) and H2A.L.2-H2Anter (lower panel) are represented as box plots.

Figure 5: H2A.L.2 pericentric chromatin targeting does not induce RNA expression, but the localized expression of pericentric RNA is associated with the pericentric targeting of H2A.L.2 in spermatozoa.

A) Total RNAs from the indicated fractionated spermatogenic cells, spermatocytes, round and condensing spermatids, as well as spermatozoa, were sequenced and the normalized total numbers of reads corresponding to the major satellite sequences (sense and anti-sense) were measured and are represented as histograms.

B) NIH3T3 cells were transfected with RFP-H2A.L.2 or RFP-H2A expressing vectors, or by Ptvlp64, and the expression of major satellite RNA was visualized by FISH with fluorescent probes detecting major satellite RNAs with both sense (upper panel) and anti-sense (lower panel) orientations.

C) The same FISH probes as in “B” were used to detect the expression of major satellite RNA in epididymal spermatozoa. The right panels show magnifications of the indicated fields.

D) Schematic representation of our working hypothesis regarding the respective roles of RNAs in the control of H2A.L.2 intranuclear localization in somatic cells and in spermatozoa. In NIH3T3 cells, the presence of RNA that do not accumulate at chromocenters results in the control of the pericentric accumulation of H2A.L.2, which therefore remains moderate. In condensing spermatids and in spermatozoa the localized accumulation of major satellite RNA cooperates with the intrinsic property of H2A.L.2 to target pericentric heterochromatin to stably target H2A.L.2 at the pericentric regions.

Figure 6: Nucleosomal position variegation in mature spermatozoa

A) Epididymal spermatozoa were purified using the swim-up protocol and the nuclei were decompacted and digested by MNase for 4 and 12 min. The arrows indicate the released nucleosomal fragments used for sequencing. The vertical lateral bars indicate the sub-nucleosomal fragments chosen for sequencing. Nucleosomal and small structure DNA from condensing spermatids were obtained as shown in Figure 1A and sequenced. The MNase sequencing experiments of round spermatids, spermatocytes and condensing spermatids were from Barral and colleagues (Barral et al., 2017).

B) 4000 genomic regions centred around nucleosomes positioned with the lowest fuzziness in round spermatids +/- 200 base pairs were selected and the aligned reads corresponding to nucleosomes and small structures at the same positions are represented in the other samples as indicated. 12(2) corresponds to nucleosomes released after 12 min of MNase digestion sequenced from an independent experiment and compared to the results of the first experiment.

The panels show heatmaps of RPKM-normalized coverage on these selected genomic regions for the indicated stages of spermatogenic cell samples. The rows are ordered according to the condensing spermatids small structure track, using the 5′ to 3′ position of the maximum signal within the −200/+200 bp region around the reference nucleosome centre.

Material and Methods

Spermatogenic Cell Fraction Purification

Cell fraction purification was performed as previously described in (Barral et al., 2017).

Mass Spectrometry-Based Proteomic

AnalysesPreparation and In-Gel Digestion of Proteins

Sample preparation and gel digestion of proteins were performed as previously described (Barral et al., 2017).

Spermatogenic Cell Chromatin Preparation and Separation of Nucleosomal and Sub-nucleosomal Particles on a Sucrose Gradient

Steps 12–16 Condensing Spermatids Preparation

Nucleosomal and subnucleosomal fractions were obtained as previously reported (Barral et al., 2017)

MNase Digestion of Chromatin from Step 12–16 Condensing Spermatids

Chromatin purified from condensing spermatids was digested using MNase for 4 and 12 minutes, as previously reported (Barral et al., 2017).

RNA Extraction from spermatids and spermatozoa

Spermatids were purified on BSA gradient as previously described (Barral et al., 2017). Spermatozoa were collected in 250 µl of PBS and obtained via swim-up as previously described in (Erkek et al., 2013). RNA was extracted using Trizol reagent. Cells were centrifuged for 5 min at 1500 rpm at 4°C. 750µl of Trizol was added to the pellet and resuspended. The suspension was passed through G26 needle for 10 times and incubated for 5minutes in RT. After 5seconds centrifugation (to settle foam), 200µl of chloroform was added to the samples, shaked gently by hand for 15seconds and incubated for 15 minutes in RT. Samples were centrifuged at 12000 rcf for 15min at 4°C and the aqueous (upper) phase was collected. 500µl of chloroform was added to the samples, they were incubated for 3 minutes at RT and centrifuged at 12000rcf at 4°C for 15minutes. 10µl of Glycogen was added to the aqueous phase and tubes were inverted 4 times. 500µl of Isopropanol (500µl/1ml Trizol) was added to the samples and they were gently inverted 4 times following by an incubation for

10minutes at RT and centrifugation of 12000rcf for 15min at 4°C. The supernatant was discarded, 1ml of 75% ethanol was added to the samples following a vortex step in order to detach the pellet, and they were centrifuged at 7500 rcf at 4°C for 5min. Pellets were dried out and resuspended in 20µl of (RNase free) H2O and the RNA concentration were measured by Thermo Scientific™ Spectrophotomètres NanoDrop™ 2000 / 2000c.

Purification of the nucleosomes using hydroxyapatite

Chromatin from condensing spermatids was purified as previously described (Barral et al., 2017). 12.5µg of chromatin was incubated with 80mg hydroxyapatite resin and completed to 1.4mL with wash buffer (5mM $NaPO_4$, 100mM NaCl, 1mM EDTA). The mix was then incubated for 10min at 4°C. The mix of chromatin/hydroxyapatite slurry was then transferred to a micro-spin column and spinned for 1min at 600g at 4°C. The flow-through was discarded and the column was washed 3 times with 200µL of wash buffer. 50µL of elution buffer (500mM $NaPO_4$, 100mM NaCl, 1mM EDTA) was added to the pellet and they were spinned for 1 min at 600g at 4°C.

Purification of the nucleosomes using heparin

300µL of heparin beads were washed with 1mL of wash buffer (100mM $NaPO_4$ and 100mM NaCl) and centrifuged for 1min at 600g in 4°C. 12.5 µg of chromatin was added to heparin beads and the volume was completed upto 1.4mL with wash buffer. The mix was incubated 15min at 4°C. After two consecutive wash steps, beads were eluted in 50µL loading buffer and were boiled at 95°C for 5min and spinned for 1min at 600g in 4°C and the supernatant was collected.

Cell culture and transfection

$1.5x10^5$ NIH 3T3 (Mouse embryonic fibroblasts) cells were cultured in 2-well glass cover lab-tek containers in DMEM 1g/L D-glucose supplemented with 10% newborn Calf serum NBCS + 2% L-glutamine and 1% penicillin/Streptomycin. Cells were maintained at 37°C in a

humidified incubator at 5% CO_2. Cells were transfected with 1µg/µL of each constructs (Table 1, material and method) for 24h, using lipofectamine 2000 and optimum.

RNA Fluorescence *In Situ* Hybridization

Preparation of Probes

0.05mM of FITC-labelled forward strand and Cy3-labelled reverse strand probes (EXIQON) were used for RNA FISH (Table 2). The probes were precipitated in 1 µg of salmon sperm DNA (Invitrogen), 5µl of sodium-acetate (NaAc) and 100 µL of Ethanol (100%) per 150µL reaction. The precipitation mix was kept at -80°C for 30 min. Samples were centrifuged by 20min at 12000 rpm at 4°C. The precipitated DNA was washed twice in 70% ethanol and then air-dried. For each probe, pellet was resuspended thoroughly in 10µL of hybridization buffer (20% formamide, 10% dextran sulfate, SSC2X). The probe was denatured for 5min at 75°C and were kept on ice during cell preparation.

Hybridization

24 hours post- transfection, cells were rinsed once with RNase-free PBS 1X. Next, cells were permeabilized in freshly made 0.5% Saponine / 0.5% Triton (SIGMA) buffer for 5min on ice and were fixed in freshly made 4% formaline solution (SIGMA) for 10 min at room temperature. Three washes in RNase-free PBS1X were carried out at room temperature. Prior to FISH, cells were rinsed twice in ethanol 70%, dehydrated in 90% and 100% ETOH 5min for each at room temperature. Cells were then air-dried, denatured probes were deposited on the plastic slides and 22×22mm RNase-free glass coverslips were placed on the plastic slides. Slides were incubated in a humidified chamber at 37°C. 24 hours post-incubation, coverslips were gently removed. Three washes were carried out using formamide 15%/2X SSC and 1XSSC, respectively, for 5 min each at room temperature. Cells were washed once in PBS1X for 5min at room temperature. DNA was stained using Hoechst (1/500) and slides were mounted using Dako, as previously described.

Immunofluorescence on germ cells

Mature epididymal sperm cells were obtained by swip-up method and were dried out on glass cover slides. Immunofluorescence was performed as described previously (Barral et al., 2017).

Microscope Analysis and Image Processing

We acquired brightfield and fluorescent images of spermatogenic and NIH 3T3 cells under a Zeiss inverted microscope, objective 63X. We used Adobe Photo-shop CS3 and ImageJ for further processing.

DNA sequencing analyses

Paired fastq files were trimmed using fastx_trimmer (with options -l 30 -Q33, meaning 5' trimming keeping 30bp-length fragments). The trimmed fastq files were aligned on the USCS mm10 genome using bowtie2 (with the following options: –end-to-end, –no-mixed, –no-discordant). The bam file corresponding to the MNase sample in round spermatids was filtered according to MAPQ (mapping quality) score greater that 30 and peak calling was done on the resulting filtered bam file using danpos2 tool (with options -m 1, for paired end). All bam files were converted into bigWig files (with the options: –extendReads, –binSize 4, –minMappingQuality 30, –normalizeUsing RPKM). The heatmap was obtained using deepTools computeMatrix on 400 bp genomic regions centred on the 4000 less fuzzy peaks (options: reference-point, –referencePoint TSS, –binSize 10, –beforeRegionStartLength 200, –afterRegionStartLength 200, –sortRegions keep). The resulting matrix was sorted using a custom R script that sorts features according to the small structure peak positions in condensing spermatids, before applying deepTools plotHeatmap (with options –colorMap YlOrRd –sortRegions no).

Purification of histones and TPs.

All mouse histones (H2A, H2A.L.2, H2B, TH2B, H3.1 and H4) and TPs (TP1 and TP2) were produced in *Escherichia coli* cells, and purified by methods described previously (Barral et al., 2017, Kujirai, Arimura et al., 2018).

Preparation of histone complexes

Histone complexes (H2A-H2B, H2A-TH2B, H2A.L.2-H2B, H2A.L.2-TH2B and H3.1-H4) were prepared as reported previously (Arimura et al., 2013). The purified histones were mixed at a 1:1 molar ratio in denaturing buffer [20 mM Tris-HCl (pH7.5), 7 M guanidine-HCl and 20 mM 2-mercaptoethanol]. For the nucleosome reconstitution, the mixtures were dialyzed against buffer A [10 mM Tris–HCl buffer (pH 7.5), 2M NaCl, 1 mM EDTA and 5 mM 2-mercaptoethanol]. The resulting histone complexes were then purified by Superdex200 (GE Healthcare Biosciences) gel filtration chromatography in buffer A. For the preparation of histone dimers for *in vitro* assays, the mixtures were dialyzed against buffer B [20 mM Tris–HCl buffer (pH 7.5), 1 mM PMSF, 1 mM EDTA, 5% glycerol and 5 mM 2-mercaptoethanol], containing 2 M NaCl. The samples were further dialyzed against buffer B, containing 1 M, 0.5 M and 0.1 M NaCl, in a stepwise manner, to decrease the NaCl concentration. The resulting histone complexes were then purified by Superdex200 gel filtration chromatography in buffer B, containing 0.1 M NaCl.

Reconstitution of nucleosomes

All nucleosomes containing the 145-bp 601L DNA fragment, a palindromic derivative of the left half of the Widom 601 sequence (Chua, Vasudevan et al., 2012) were reconstituted and purified as described previously (Barral et al., 2017).

Thermal stability assay

The thermal stability assay was conducted as described previously (Arimura, Shirayama et al., 2014, Taguchi, Horikoshi et al., 2014). Briefly, purified nucleosomes were mixed with SYPRO Orange dye (Sigma-Aldrich) in 19 µl of reaction buffer [20 mM Tris-HCl (pH 7.5) and 1 mM

DTT], containing the indicated amounts of NaCl. The intensity of SYPRO Orange fluorescence was monitored with a StepOnePlus RealTime PCR system (Applied Biosystems), using a temperature gradient from 26˚C to 95˚C, in steps of 1˚C/min.

DNA binding assay of histone dimers

The 145-bp Widom 601 DNA (20 ng/µl) was mixed with the indicated amounts of histone dimers in 10 µl of reaction buffer, containing 27 mM Tris-HCl (pH 7.5), 50 mM NaCl, 0.31 mM EDTA, 0.3 mM PMSF, 1.5% glycerol, 1.5 mM 2-mercaptoethanol and 1 mM DTT, and incubated at 37˚C for 10 min. The reaction mixtures were then separated by 6% non-denaturing PAGE in 0.5 × TBE buffer (45 mM Tris base, 45 mM boric acid and 1 mM EDTA) and the bands were visualized by ethidium bromide staining.

For TPs binding experiments, the indicated amounts of either TP1 or TP2 were mixed with the 145-bp Widom 601 DNA (20 ng/µl) and histone dimers (0.9 µM) in 10 µl of reaction buffer [29 mM Tris-HCl (pH 7.5), 40 mM NaCl, 11 µM $ZnCl_2$, 0.41 mM EDTA, 0.3 mM PMSF, 1.5% glycerol, 1.5 mM 2-mercaptoethanol and 1.1 mM DTT]. After an incubation at 37˚C for 10 min, the reaction mixtures were analysed as described above.

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-H2A.L.2	(Barral et al., 2017)	N/A
Mouse monoclonal anti-Protamine P2	Briar Patch Biosciences	Cat#HUP2B
Bacterial Strains		
Escherichia Coli-One shot Top10 competent cells	Invitrogen	Cat#404003
SURE, competent cells	Agilent Technologies	Cat#200238
Chemicals, Peptides, and Recombinant Proteins		
TRIzol	Life Technologies	cat#15596026
Lipofectamine 2000	Invitrogen	cat#11668-019
OPTI-MEM	Gibco	cat#31985-062
Heparine-sepharose	Abcam	cat#ab193268
Hydroxyapatite	Sigma	cat#289396-25G
Calf serum NBCS	TheroFischer Scientific	cat#16010159
Deposited Data		
Raw mass spectrometry proteomics	This paper	N/A
MNase-seq accession ID	This paper	N/A
RNA-seq accession	This paper	N/A
Experimental Models: Cell Lines		
NIH 3t3 Mouse Embryonic Fibroblast	ATCC	ATCC® CRL-1658
Experimental Models: Organisms/Strains		
Mouse: B6.126/SvPas-*h2a.l.2 neo/neo*	(Barral et al., 2017)	N/A
Oligonucleotides		
Probes: Forward1: FITC-TCTTGCCATATTCCACGTCC	(Probst et al., 2010)	N/A
Forward 2: Cy3-GATTTCGTCATTTTTCAAGT	(Probst et al., 2010)	N/A
Reverse 1: CY3-GCGAGGAAAACTGAAAAAGG	(Probst et al., 2010)	N/A
Reverse 2: FITC-GCGAGAAAACTGAAAATCAC	(Probst et al., 2010)	N/A
Recombinant DNA		
ptag-RFP H2A.L.2	This paper	N/A
ptag-RFP H2A	This paper	N/A
pTVLP-64-ms-Ca B15	Maria-Elena Torres Padillas	N/A
ptag-RFP H2A.B3	Genscript	N/A
Ptag-RFP H2A.L.2-H2Anter	This paper	N/A
Mm Histone H2A	Hitoshi Kurumizaka	N/A
Mm Histone H2B	Hitoshi Kurumizaka	N/A
Mm histone H3	Hitoshi Kurumizaka	N/A
Mm Histone H4	Hitoshi Kurumizaka	N/A
Mm TP1	(Barral et al., 2017)	N/A
Mm TP2	(Barral et al., 2017)	N/A
Mm Histone H2A.L.2	(Barral et al., 2017)	N/A
Mm Histone TH2B	(Barral et al., 2017)	N/A
Other		
Mass spectrometer	Thermo Scientific	LTQ-Orbitrap Velos Pro

Table I. Material and method

Discussion and Perspectives

I. p300 activation via transcription factor dimerization

CBP/p300 has been associated with transcriptional activation for many years. This concept has always considered transcription factor phosphorylation such as IRF3 and STAT, via tyrosine kinases such as TBK1 and JAK (Stark & Darnell, 2012). Following phosphorylation, these transcription factors are activated. They dimerize and translocate to the nucleus and bind to CBP/p300 (Zhang, Vinkemeier et al., 1996). Interestingly, it has been shown that transcription factor dimerization is a crucial step for CBP/p300 activation. How the association of these dimerized-transcription factors can induce the activation of CBP/p300, was still an issue that needed to be addressed.

In 2018, in collaboration with researchers at EMBL Grenoble Outstation, we could define the crystal structure of its activated state and identified the molecular mechanisms underlying the activation of CBP/p300. We demonstrated that the dimerization of transcription factors induces p300 auto-acetylation of a lysine-rich auto-inhibitory loop in the HAT domain. This auto-inhibitory loop (AIL) serves as a "pseudo-substrate" which projects into the active site of the neighboring p300 and regulates the HAT activity of this protein. Once the HAT domain is activated, it induces lysine acetylation on histones. These acetylated histones will be then recognized by the bromodomain of p300. These findings explain for the first time the cellular signaling from transcription factors to the chromatin modifying factors (Figure 25). Moreover, p300 hyperactivity, for instance by HAT-activating mutations, leads to the generation of p300 condensates within the nucleus, which are highly decreased upon HAT or bromodomain inhibitor treatments. This observation suggests the critical role of bromodomain in the propagation of chromatin modification via p300 activity.

Phase-separated multi molecular assemblies are believed to provide separate cellular compartments for biochemical reactions within the cell. Recently, a model has been proposed that presents a hypothetical role of the phase-separated compartments within the cell. It describes these compartments as membrane-less organelles, which contribute to the formation of super-enhancers and enable the super enhancers to activate multiple genes simultaneously. Additionally, according to this model, phase-separated compartments provide higher density of actors critical for transcriptional regulation such as transcription factors and enhancers, which mediate gene transcription through changing the valency of important transcriptional regulators (Hnisz, Shrinivas et al., 2017). Therefore, In the context of this work, the generation

of these condensates in p300 HAT-activating mutants, might strengthen signal integration on enhancers and promote gene transcription.

In the context of spermatogenesis, the collaboration between NUT and p300 in round spermatids, has been associated to histone hyperacetylation and initiation of histone removal (Shiota et al., 2018). Based on our discoveries on the mechanisms of p300 activation, we believe that NUT is also activating p300 through the same scheme (Figure 25). However, in the future, it is critical to test this hypothesis in the context of spermatid histone hyperacetylation.

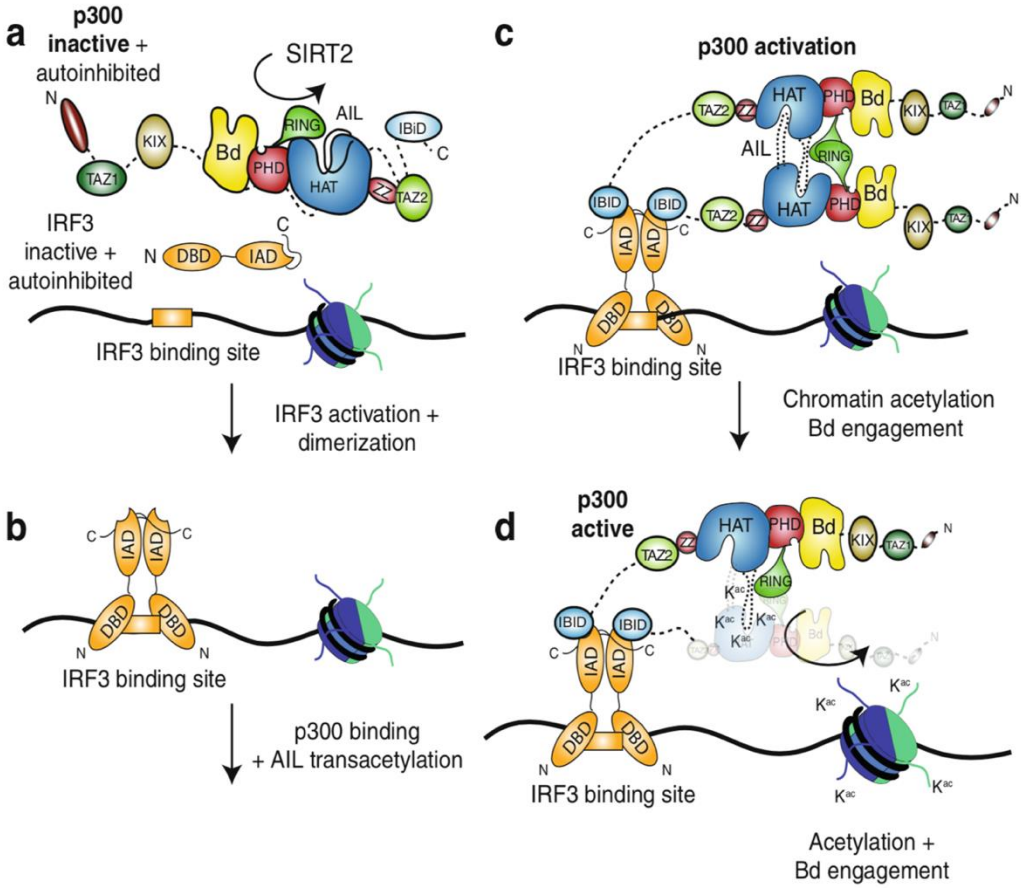


Figure 25. Detailed step by step molecular activation process of p300. Taken from (Ortega et al., 2018).

II. Nucleosome retention in mammalian sperm

In 2009, Hammoud and colleagues by analyzing spermatozoa from a single donor and from a pool of sperm donors through mainly H3K4me2/3 and H3K27me3 ChIP experiments, revealed the first evidences of the specifically - positioned nucleosome retention in the human sperm. Their findings highlighted nucleosome retention at promoters of genes, which are essential for early embryo development as well as genes with particular biological relevance at these early stages of embryonic development. Additionally, they also found an enrichment of H3K4me3 at regions encoding many non-coding RNAs and miRNAs, which are expressed during development (Hammoud et al., 2009).

Since this study, other research programs have been carried out first, in order to investigate nucleosome retention in other species and second, to identify the functional significance of these remaining nucleosomes in the mature sperm. In 2013, Erkek and colleagues performed a series of comparative genome-wide H3K4me3, H3K27me3, H3.1/H3.2 and H3.3 ChIP following high-throughput sequencing analysis using human and mouse spermatids and mature sperm. These authors reported the enrichment of nucleosomes at hypomethylated CpG-islands. In addition, they reported the enrichment of H3.3 on most of the CpG-islands of mouse sperm compared to H3.1/H3.2, which their enrichment was limited only to some CGIs. Since CpG islands are often associated with promoter regions and the fact that nucleosomes are retained on hypomethylated CGIs, they suggested that the nucleosome retention might be associated with a paternal epigenetic information (Erkek et al., 2013). In 2007, Govin and colleagues (Govin et al., 2007) and in 2013, Meyer Ficca and colleagues through FISH analysis reported the association of histones with repetitive sperm sequences in mice (Meyer-Ficca, Lonchar et al., 2013).

Following these observations and since there was no solid evidence of paternal activation of genes bearing nucleosomes during early development, in 2014, Samans and colleagues decided to revisit the mapping of nucleosome preserved sites and their biological importance for post-fertilization stages in human and bovine samples. For this aim, regardless of histone modifications, they prepared 146bp DNA fraction corresponding to mononucleosomes and these sequences were subjected to high-throughput sequencing followed by the analysis of nucleosome binding sites on the human and bovine sperm genome. In contrast to the findings from Hammoud and colleagues, these authors reported nucleosome retention on repetitive

DNA in spermatozoa. Additionally, they reported a differential nucleosome retention depending on the nature of repetitive DNA elements. For instance, they found a significant enrichment of nucleosome-binding sites on LINE1 and SINE, a massive depletion of nucleosome enrichment on DNA transposons, LTRs and low-complexity repeats (LCRs) in both human and bovine sperm. Moreover, they reported an important nucleosome depletion within both human and bovine exons, introns, promoters, TSSs, 3'-UTR and 5'-UTRs. Further analysis of nucleosome-associated promoters revealed the enrichment of the corresponding genes in factors important for RNA and protein processing. The remaining genes, both in human and in bovine sperms were nucleosome-free and were occupied by protamines. Among the nucleosome-free genes, *HOX* gene clusters were identified in both human and bovine sperm, in contrast to previously reported data from Hammoud and colleagues.

Furthermore, after the confrontation of nucleosomal site position in human sperm with sperm DNA methylation data, these authors reported the hypomethylation of the majority of promoters and exons in nucleosome-preserved regions compared to nucleosome-depleted genes. Moreover, when they analyzed the DNA methylation of the CpG islands at a single gene-locus level, as previously described in the literature, it appeared that the maternally expressed genes harboured hypermethylated CpG islands at their promoters, whereas paternally expressed genes harboured hypomethylated CpG promoter regions in sperm (Samans et al., 2014). It is important to note that in Samans and colleague's study, the data on the repetitive elements were later criticized (Royo, Stadler et al., 2016).

The same year, Carone and colleagues by characterizing the MNase digested chromatin of mouse embryonic stem and sperm cells, revealed the profile of the chromatin architecture and nucleosome positions in these cells. In this study, they cross-linked cells with formaldehyde. These cells were lysed, subjected to MNase digestion and the released DNA was sequenced (paired-end sequencing). Interestingly, in this study they also sequenced various DNA size released after MNase digestion, which included not only nucleosome footprints of ~120-150bp corresponding to mononucleosome (147bp), but also shorter fragments of (<80bp). In sperm, their analysis revealed the preferential histone retention over gene-poor regions. Additionally, fluorescence microscopy also confirmed the localization of the majority of nucleosomes at distinct chromocenter regions occupied by repetitive elements. However, these authors also observed the presence of the nucleosome fragments on CpG-rich promoters, but these fractions represented only a small percentage of all nucleosomes retained in sperm and in fact most

promoters were found to be free of nucleosomes. Furthermore, they reported the enrichment of CTCF factor at MNase footprints in mature sperms (Carone et al., 2014). These observations were in agreement with the Samans study but in complete disagreement with earlier reported data.

To explain all these opposing data, it has been speculated that these differences might be due to the differential conditions of MNase digestion, since the enrichment of mononucleosomes was observed after applying high MNase concentration (Saitou & Kurimoto, 2014). Additionally, these conditions could have destroyed all of the sub-nucleosomal histone-preserved regions, which we are reporting here.

In 2018, Yoshida and colleagues developed a protocol in order to analyze histone retention and distribution in mouse sperm. They first purified histone replacement-completed sperm (HRCS) fraction (corresponding to cells with completed process of histone-to-protamine exchange) from epididymis, caput and cauda. Then, after using DTT and heparin, the solubilization of histones were performed by sonication from cross-linked DNA without MNase digestion. This protocol consists of a mild sonication in order to remove sperm tail followed by two rounds of percol isolation in order to remove the low-density sperm (corresponding to histone replacement-uncompleted sperm (HRunCS)). H3 ChIP-seq analysis on HRCS fractions revealed the enrichment of nucleosomes at promoters, which they correlated with the possible transcriptional regulatory role of nucleosome-preserved regions during early embryonic development. Further analyses are needed to explore whether residual histones affect gene expression after fertilization.

Compared to swim-up sperms as well as total cells obtained from cauda epididymis, histone H3 was enriched 5 times less in HRCS fraction. This data confirms first the heterogeneity of sperm population, which should be taken into consideration for future nucleosome mapping analyses, and second they could explain to some extend the controversial data obtained from previous studies.

III. Characterization of sub-nucleosomal structures in mouse condensing spermatids

In 2007, in my host laboratory, sub-nucleosomal structures were first identified through extensive MNase digestion of purified chromatin from mouse condensing spermatids. Due to the high level of chromatin compaction in condensing spermatids and its resistance to MNase digestion, chromatin was first partially decompacted using a detergent. After centrifugation, a soluble extract and non-soluble pellet were obtained. Non-soluble fraction was then subjected to extensive MNase digestion. Two categories of MNase-digested profile (corresponding to ~150bp) and MNase sensitive or small structures, corresponding to ~75bp fragments, were obtained. Both DNA fragments were isolated from agarose gel and used as probes for DNA FISH. Interestingly, both nucleosome-related DNA and small structures were associated with pericentric heterochromatin regions in round and condensing spermatids. In order to identify proteins involved in genome organization in late spermatids, basic proteins were isolated after acid extraction and they were subjected to proteomic analyses. Five new proteins sharing sequence similarities with H2A and H2B were identified, which were called respectively, H2A.L.1 (for H2A.Like 1), H2A.L.2, H2A.L.3, H2B.L.1 and H2B.L.2. According to the latest nomenclature, these proteins are referred to as H2A.Lap2, H2A.Lap3/H2afb1/H2A.B.1, Y-chr H2A.L.3, subH2B/H2B.L.1 H2B.L.2 respectively (El Kennani et al., 2017). Further analysis revealed the localization of H2A.L.1/L.2 in the pericentric heterochromatin regions of late spermatids, suggesting a role for these proteins in organizing pericentric heterochromatin. Interestingly, IP analyses demonstrated the dimerization preference of H2AL.1/L.2 for TH2B rather than H2B (Govin et al., 2007). All of these findings encouraged further characterizations of the MNase-sensitive small structures and their protein content as well as identifying the role of H2A.L.1/L.2 in the pericentric genome organization of late spermatids.

In 2017, in order to characterize the role of H2A.L.2 in chromatin organization in late spermatids, a series of experiments were carried out. The generation of knock out mouse models as well as H2A.L.2-specific antibody, hugely participated to the identification of H2A.L.2's role in late spermatids genome organization. Different experimental approaches revealed that H2A.L.2 incorporation is an important step in the nucleosome-to-nucleoprotamine transformation.

Briefly, in condensing spermatids, after loading H2A.L.2-TH2B dimer onto the nucleosomes, they open the nucleosome structure, which allows the invasion of nucleosomes by TPs. Nucleosome-TP interact with the incoming Prms, and mediate the pre-PRM-2 processing. Prms bind to DNA in competition with histone-TP complexes, leading to chromatin compaction via PRMs (Figure 26) (Barral et al., 2017).

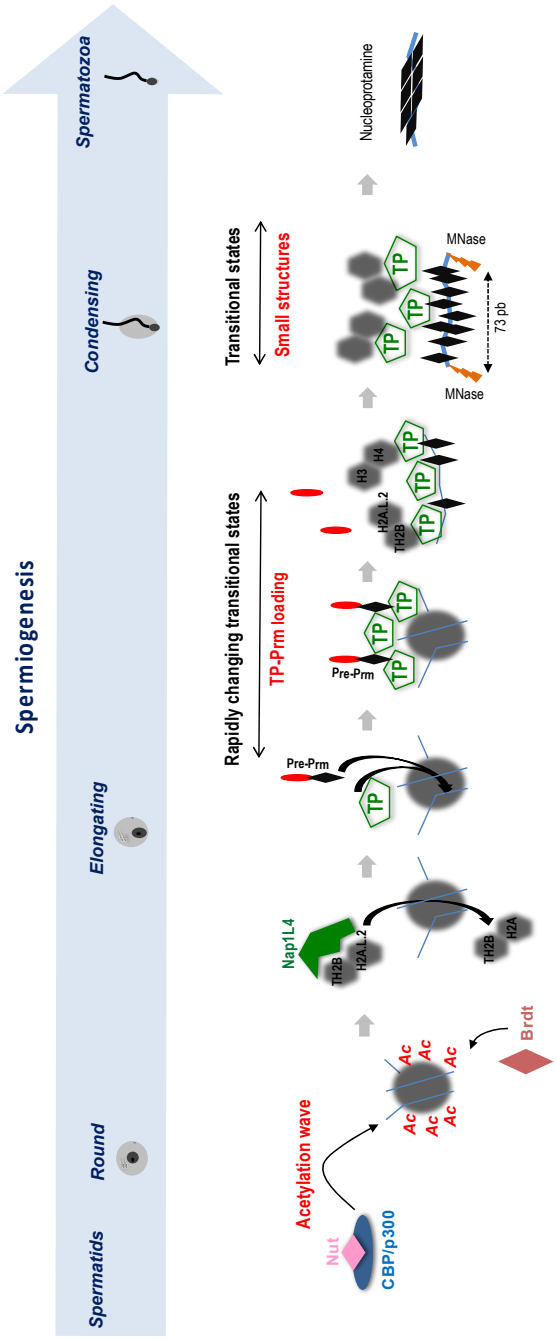


Figure 26: Schematic representation of step by step molecular mechanisms involved in histone hyperacetylation and histone eviction. Taken from Sophie Barral, PhD manuscript, 2018.

IV. H2A.L.2-TH2B dimer-based histone retention in mouse sperm

Previously, we demonstrated that H2A.L.2 nucleosomal incorporation induces nucleosome relaxation (Barral et al., 2017, Montellier et al., 2013), similar to CENP-A nucleosomes, which provide an open structure, which favors the loading of non-nucleosomal proteins to ensure the subsequent microtubules assembly. Similarly, in the case of H2A.L.2 incorporation, the subsequent nucleosomal relaxation, allows TPs and Prms loading onto H2A.L.2-containing nucleosomes in elongating and condensing spermatids, establishing a functional interconnection between histones, TPs and Prms.

In this study, extensive MNase digestion of the nuclei of condensing spermatids, generated mono-nucleosomal (~150bp) and sub-nucleosomal fractions (73bp).

The enrichment of histone complement associated with these DNA fragments, clearly revealed strong presence of dimer H2A.L.2-TH2B and depletion of H3 and H4 in the sub-nucleosomal structures. Thermal stability assay experiment as well as other *in vitro* biochemical assays, revealed a higher stability, enhanced DNA-binding property and more stable histone-DNA-TP complexes, for H2A.L.2-TH2B dimers compared to conventional H2A-H2B dimers. A recent study demonstrated that the partial unwrapping of nucleosome (similar to that obtained during thermal stability assays), changes H2A-H2B conformation, leading to a more stable binding of the dimer to DNA (Bilokapic et al., 2018). These observations in addition to our current data, nicely explain the retention of dimer-based structures at the time of nucleosome dismantlement, rather than total histone complement removal.

Very interestingly, paired-end sequencing of fragments corresponding to nucleosomal and sub-nucleosomal structures in mature sperm, revealed the presence of both nucleosomes and sub-nucleosomal structures at the same positions as previously observed in condensing spermatids (Barral et al., 2017). These observations suggest the presence of stable H2A.L.2-TH2B-DNA complexes, which resist displacement at the time of histone eviction in condensing spermatids. This also suggests a role for TPs in dissociating H2A.L.2-TH2B dimer from H3-H4 tetramer, rather than the whole octamer displacement. Moreover, the fact that sub-nucleosomal structures are present at the same positions as nucleosomes, highlights the

presence of a heterogeneous population of sperm cells. We believe that an interpretation of these data could be that in some cells some nucleosomes resist the pressure of Prms, in others the dimer-based sub-nucleosomal structures appear and in the third group histones are completely replaced by Prms.

Our current observations could also explain the disagreement regarding the subject of histone retention in mature sperm (Erkek et al., 2013, Hammoud et al., 2009, Samans et al., 2014, Yoshida et al., 2018).

In addition to our work on the spermatozoa genome organization, our "off-context" ectopic expression of H2A.L.2 in NIH3T3 cells, also revealed the intrinsic ability of H2A.L.2 to target the pericentric regions of the genome. This observation suggests that no testis-specific factor is required for this targeting. Of course, H2A.L.2 ChIP-seq analyses are required to confirm our *in situ* and fluorescent observation regarding the pericentric targeting of H2A.L.2. However, the repeating nature of pericentric satellites, makes these analyses complicated.

Based on the previous reports, today we know that the mammalian sperm genome is partially packed by histones (in mouse retained at about ~1% compared to somatic cells (Erkek et al., 2013). Histone-preserved-regions have been proposed to potentially play a role in early embryonic gene expression (Arpanahi, Brinkworth et al., 2009, Brykczynska et al., 2010, Hammoud et al., 2009), or might be crucial for the cytosine methylation inheritance in early developmental stages (Nakamura et al., 2012). Therefore, we can also assume that our dimer-based structures might also mark some of these regions, which could be important for early embryonic development.

V. Intrinsic property of H2A.L.2 for pericentric heterochromatin

Previous work in my host laboratory revealed the pericentric localization of H2A.L.2 in late spermatids. In addition, recent data from our laboratory demonstrated that upon H2A.L.2 incorporation, nucleosomes become relaxed, which lead to nucleosome-to-nucleosprotamin transformation. These observations indicate that despite nucleosome removal upon genome-wide H2A.L.2 incorporation, H2A.L.2 remains at pericentric regions, which could be an indication of a differential programming of pericentric regions in mature sperm.

In the present work, in addition to late spermatids and mature sperms, we unraveled the pericenetric localization of H2A.L.2 in mouse embryonic fibroblast NIH 3T3 cells in a complete out-of-context situation. Interestingly, under the same conditions, H2A.B.3 was unable to localize at pericentric regions upon its ectopic expression. We conclude that H2A.L.2 harbours an intrinsic property to localize at the pericentric regions and that no testis-specific factor is required for this specific targeting.

H2A.W, a recently identified histone variant in *Arabidopsis thaliana*, has also been associated with heterochromatin regions (Yelagandula, Stroud et al., 2014). The authors of this work reported that, H2A.W acts in synergy with heterochromatin marks such as H3K9me2 and DNA methylation, in order to maintain the integrity of these regions (Yelagandula et al., 2014). Besides, these authors reported that *in vitro*, H2A.W enhances chromatin fiber-to-fiber interaction via its conserved C-terminal motif. In our study, we also generated multiple truncated versions of H2A.L.2 by replacing its C-terminal motif with canonical H2A and H2A.B.3. These mutations revealed partial mis-localization of H2A.L.2 from pericentric heterochromatin once ectopically expressed in 3T3 cells (data not shown). Therefore, we concluded that the C-terminal region of H2A.L.2 is partially responsible for the pericentric localization of this protein.

Interestingly, the extermination of the published data on H2A.W *in vivo* showed that in K.O. plants, the absence of H2A.W is associated with heterochromatin decondensation (Yelagandula et al., 2014). Similarly, our previous data also revealed that the absence of H2A.L.2 in mouse sperm could lead to the same observation since we detected a general genome decompaction that should also include the pericentric regions.

These data together suggest the specific and particular roles of H2A variants in the organization of pericentric regions.

VI. H2A.L.2 turnover regulation and stabilization via RNA in NIH 3T3 and mouse sperm

N-terminal of H2A.L.2 contains a 7 arginines motif. A similar region in H2A.B.3 histone variant has been characterized as a RNA-binding involved in the interaction of this histone with several nuclear factors involved in RNA processing (Anuar et al., 2019, Soboleva et al., 2017). Surprisingly, ectopic expression of the truncated mutant of H2A.L.2 in which the 7-arginine motif at H2A.L.2 N-terminal was replaced by the canonical H2A counterpart in NIH 3T3 cells, revealed an enhanced localization of the protein in pericentric regions. This observation suggested that this putative RNA-binding motif has a role in the control of the pericentric localization of H2A.L.2.

In order to directly confirm that the N-terminal arginine motif is the RNA binding domain of H2A.L.2, RNA pull-down experiments are ongoing. However, we obtained additional data supporting a role for this motif and for RNA in the pericentric targeting of H2A.L.2. Indeed, the ectopic expression of major satellite transcripts, induced mis-localization of the wild-type H2A.L.2 from pericentric regions but not of the mutated H2A.L.2 lacking this RNA-binding motif. Additionally, FRAP analyses of both wild-type and mutated H2A.L.2 demonstrated that the recovery of wild-type H2A.L.2 at the pericentric regions is much faster than that of the mutant protein. In other words, the 7-arginine motif of H2A.L.2 regulates its turnover via interaction with different types of RNA. The FRAP analyses also revealed that H2A.L.2 interaction with RNA is not specific to major satellite transcripts, since we also showed that these cells do not express satellite RNAs. Therefore, we can propose that, RNA has an impact on the stabilization of H2A.L.2, and in general plays a role in its turnover and dynamics. In the context of mature sperm, our RNA-seq data in addition to RNA FISH analyses, revealed the expression of major satellites RNA in post-meiotic spermatogenic cells including mature spermatozoa. Considering all of these data, we can conclude that in the late spermatogenic cells, major satellite RNAs expression and localized accumulation control and regulate the turnover and stability of H2A.L.2 at pericentric heterochromatin regions.

A role for RNA in the stabilization of chromatin-binding factors has been evidenced for various regulatory factors. A recent study demonstrated that RNA binding of the insulator-binding factor and the higher-order chromatin organizer, CTCF factor, is crucial for its association to chromatin, suggesting the role of RNA in the stability and maintenance of CTCF factor at its cognate DNA binding sites (Saldana-Meyer, Rodriguez-Hernaez et al., 2019). Interestingly, another study also revealed that the major satellite repeat transcripts mediate the stabilization of Suv39h1 and Suv39h2 on the pericentric heterochromatin (Velazquez Camacho, Galan et al., 2017). These observations highlight the importance of RNAs in the stability and function of regulatory elements but to the best of our knowledge, this RNA-mediated chromatin binding stability has not been shown for histones. Therefore, our finding on the role of RNA is the first example of such a role for RNA in the control of histone dynamics.

General Conclusion

The present study shed lights on two important steps in the genome reorganization during late spermatogenesis. First, we discovered that p300 activation requires p300-dimerization. Therefore, this finding will guide us in deciphering the molecular basis of Nut-mediated p300-dependent general histone hyperacetylation in elongating spermatids. Additionally, for the first time, we demonstrated an unknown novel function for the testis-specific H2A.L.2 histone variant in the generation of dimer-based sub-nucleosomal structures in mature spermatozoa, we also defined the ability of RNA to control the dynamics of H2A.L.2 and to direct its final genomic localization in mature spermatozoa .

As previously reported, NUT enhances the HAT activity of p300 for H4 acetylation, especially on lysines 5 and 8, in condensing spermatids, in addition to its canonical role in H3 acetylation in spermatogonia and in spermatocytes. Here, thanks to crystal structures and a series of functional studies, we propose a multi-step process in which the HAT activity of p300 is ensured by the activation and oligomerization of transcriptional co activators ligands. Briefly, following the cellular signaling, transcription factors, such as IRF3 and STAT1, become phosphorylated and dimerized, leading to their activation. These transcription factors are then targeted to DNA, and engage two molecules of p300, leading to the removal of AIL *in cis* from its inhibitory position. This results in activation of p300 HAT. Additionally, activated p300 can recognize acetylated chromatin via its bromodomain and propagates histone acetylation on chromatin (Figure 25). This study reveals a chain of events from DNA targeting by transcription factors to chromatin remodeling.

We speculate that this mechanism could activate p300 in early spermatogenic cells, spermatogonia, spermatocytes and round spermatids. However, during late spermatogenesis, with the decrease in gene expression, Nut is induced and activates p300 in a transcription factor-independent manner. More specifically, Nut orients this p300 enhanced activity towards H4, specially K5 and K8 to generate binding sites for Brdt, which mediates in turn histone removal. Unpublished data from my laboratory indicates that this chain of events leads to the replacement of TH2B-H2A by TH2B-H2A.L.2.

We then focused our attention on H2A.L.2. Here, we report for the first time that not only H2A.L.2 is important for nucleosome invasion by TPs and Prms, but also it could constitute the basis of a new mature sperm genome organization. At the onset of histone eviction process, the incorporation of H2A.L.2 into the nucleosome, induces an open nucleosome, which renders nucleosomes prone to accommodate TPs and Prms, leading to

histone eviction. Our current observations reveal that H2A.L.2-TH2B dimers have a strong affinity to bind to DNA and processes the ability to generate stable dimer-based structures, which resist the TP-Prm pressure for DNA removal. Therefore, we propose, the existence of a heterogeneous pool of mature sperm cells, harbouring nucleoprotamine, H2A.L.2-TH2B-PRM dimer-based and full Prm-compacted chromatin structures (Figure 27).

Of course, further studies are required to evaluate the role of these structures in the transmission of specific male epigenetic information to the zygote.

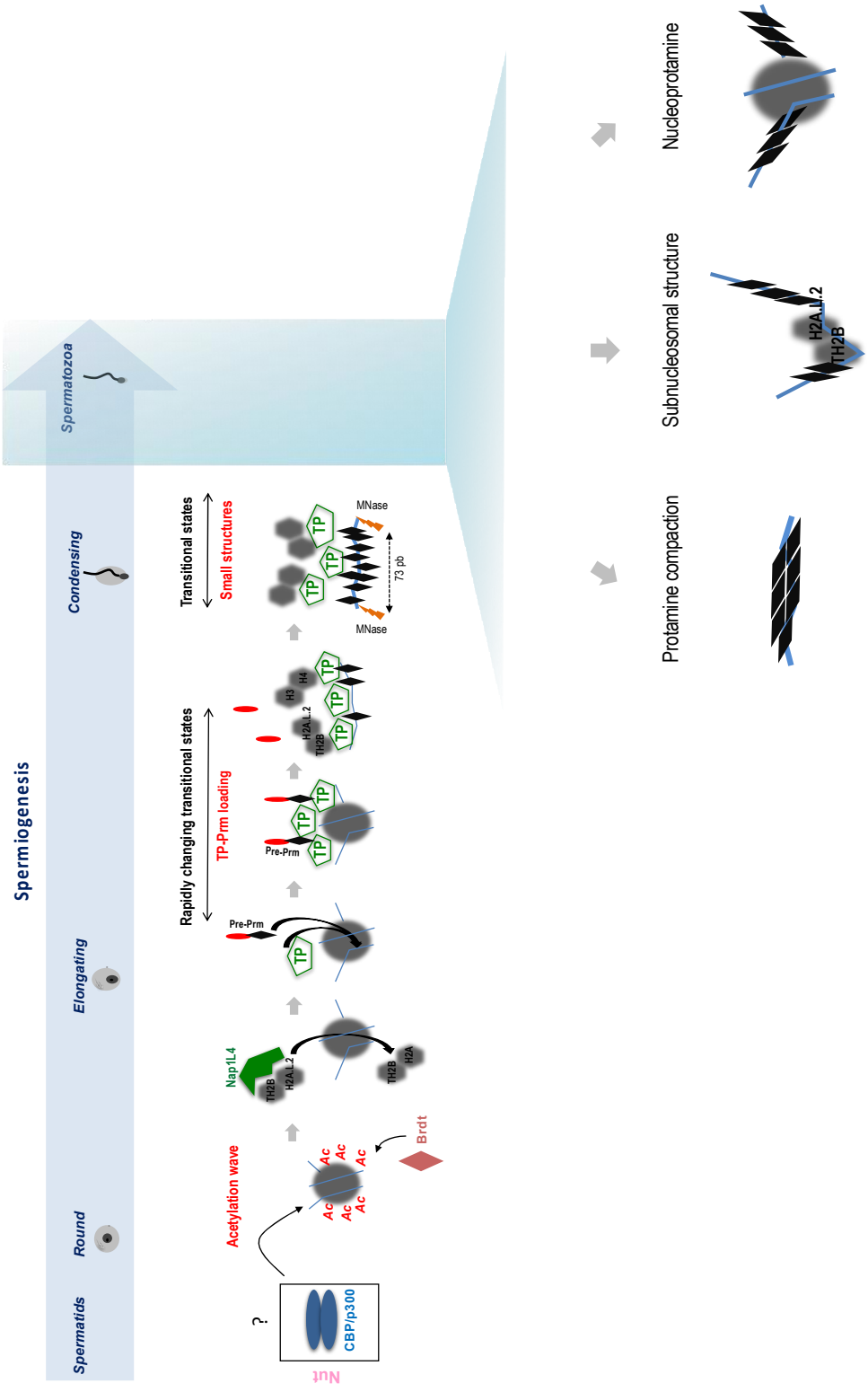


Figure 27. Schematic summary of our current observations. Atomic level molecular mechanisms underlying Nut and p300 interaction are still unknown. Additionally, heterogenous genome-compacted pool of mature sperm cells is highlighted on the spermatozoa.

References

Abascal F, Corpet A, Gurard-Levin ZA, Juan D, Ochsenbein F, Rico D, Valencia A, Almouzni G (2013) Subfunctionalization via adaptive evolution influenced by genomic context: the case of histone chaperones ASF1a and ASF1b. *Mol Biol Evol* 30: 1853-66
Adam S, Polo SE, Almouzni G (2013) Transcription recovery after DNA damage requires chromatin priming by the H3.3 histone chaperone HIRA. *Cell* 155: 94-106
Aguilar-Gurrieri C, Larabi A, Vinayachandran V, Patel NA, Yen K, Reja R, Ebong IO, Schoehn G, Robinson CV, Pugh BF, Panne D (2016) Structural evidence for Nap1-dependent H2A-H2B deposition and nucleosome assembly. *EMBO J* 35: 1465-82
Ahmad K, Henikoff S (2002) Histone H3 variants specify modes of chromatin assembly. *Proc Natl Acad Sci U S A* 99 Suppl 4: 16477-84
Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P (2002) The structure and function of DNA. In *Molecular Biology of the Cell 4th edition*, Garland Science
Albig W, Ebentheuer J, Klobeck G, Kunz J, Doenecke D (1996) A solitary human H3 histone gene on chromosome 1. *Hum Genet* 97: 486-91
Alekseyenko AA, Walsh EM, Wang X, Grayson AR, Hsi PT, Kharchenko PV, Kuroda MI, French CA (2015) The oncogenic BRD4-NUT chromatin regulator drives aberrant transcription within large topological domains. *Genes Dev* 29: 1507-23
Alekseyenko AA, Walsh EM, Zee BM, Pakozdi T, Hsi P, Lemieux ME, Dal Cin P, Ince TA, Kharchenko PV, Kuroda MI, French CA (2017) Ectopic protein interactions within BRD4-chromatin complexes drive oncogenic megadomain formation in NUT midline carcinoma. *Proc Natl Acad Sci U S A* 114: E4184-E4192
Allan J, Hartman P, Crane-Robinson C, Aviles F (1980) The structure of histone H1 and its location in chromatin. *Nature* 288: 675
Altmann R (1889) Ueber nucleinsäuren. *Arch f Anatomie u Physiol* 1: 524-536
Anuar ND, Kurscheid S, Field M, Zhang L, Rebar E, Gregory P, Buchou T, Bowles J, Koopman P, Tremethick DJ, Soboleva TA (2019) Gene editing of the multi-copy H2A.B gene and its importance for fertility. *Genome Biol* 20: 23
Aravin A, Gaidatzis D, Pfeffer S, Lagos-Quintana M, Landgraf P, Iovino N, Morris P, Brownstein MJ, Kuramochi-Miyagawa S, Nakano T, Chien M, Russo JJ, Ju J, Sheridan R, Sander C, Zavolan M, Tuschl T (2006) A novel class of small RNAs bind to MILI protein in mouse testes. *Nature* 442: 203-7
Aravin AA, Bourc'his D (2008) Small RNA guides for de novo DNA methylation in mammalian germ cells. *Genes Dev* 22: 970-5
Arents G, Burlingame RW, Wang B-C, Love WE, Moudrianakis EN (1991) The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-handed superhelix. *Proceedings of the National Academy of Sciences* 88: 10148-10152
Arimura Y, Kimura H, Oda T, Sato K, Osakabe A, Tachiwana H, Sato Y, Kinugasa Y, Ikura T, Sugiyama M, Sato M, Kurumizaka H (2013) Structural basis of a nucleosome containing histone H2A.B/H2A.Bbd that transiently associates with reorganized chromatin. *Sci Rep* 3: 3510
Arimura Y, Shirayama K, Horikoshi N, Fujita R, Taguchi H, Kagawa W, Fukagawa T, Almouzni G, Kurumizaka H (2014) Crystal structure and stable property of the cancer-associated heterotypic nucleosome containing CENP-A and H3.3. *Sci Rep* 4: 7115
Arpanahi A, Brinkworth M, Iles D, Krawetz SA, Paradowska A, Platts AE, Saida M, Steger K, Tedder P, Miller D (2009) Endonuclease-sensitive regions of human spermatozoal chromatin are highly enriched in promoter and CTCF binding sequences. *Genome Res* 19: 1338-49

Aul RB, Oko RJ (2002) The major subacrosomal occupant of bull spermatozoa is a novel histone H2B variant associated with the forming acrosome during spermiogenesis. *Dev Biol* 242: 376-87
Avery OT, MacLeod CM, McCarty M (1944) Studies on the chemical nature of the substance inducing transformation of pneumococcal types: induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. *Journal of experimental medicine* 79: 137-158
Avvakumov N, Lalonde ME, Saksouk N, Paquet E, Glass KC, Landry AJ, Doyon Y, Cayrou C, Robitaille GA, Richard DE, Yang XJ, Kutateladze TG, Cote J (2012) Conserved molecular interactions within the HBO1 acetyltransferase complexes regulate cell proliferation. *Mol Cell Biol* 32: 689-703
Balhorn R (2007) The protamine family of sperm nuclear proteins. *Genome Biol* 8: 227
Bao J, Bedford MT (2016) Epigenetic regulation of the histone-to-protamine transition during spermiogenesis. *Reproduction* 151: R55-70
Bao Y, Konesky K, Park YJ, Rosu S, Dyer PN, Rangasamy D, Tremethick DJ, Laybourn PJ, Luger K (2004) Nucleosomes containing the histone variant H2A.Bbd organize only 118 base pairs of DNA. *EMBO J* 23: 3314-24
Barau J, Teissandier A, Zamudio N, Roy S, Nalesso V, Herault Y, Guillou F, Bourc'his D (2016) The DNA methyltransferase DNMT3C protects male germ cells from transposon activity. *Science* 354: 909-912
Barral S, Morozumi Y, Tanaka H, Montellier E, Govin J, de Dieuleveult M, Charbonnier G, Coute Y, Puthier D, Buchou T, Boussouar F, Urahama T, Fenaille F, Curtet S, Hery P, Fernandez-Nunez N, Shiota H, Gerard M, Rousseaux S, Kurumizaka H et al. (2017) Histone Variant H2A.L.2 Guides Transition Protein-Dependent Protamine Assembly in Male Germ Cells. *Mol Cell* 66: 89-101 e8
Bauer UM, Daujat S, Nielsen SJ, Nightingale K, Kouzarides T (2002) Methylation at arginine 17 of histone H3 is linked to gene activation. *EMBO Rep* 3: 39-44
Beaujean N, Taylor JE, McGarry M, Gardner JO, Wilmut I, Loi P, Ptak G, Galli C, Lazzari G, Bird A, Young LE, Meehan RR (2004) The effect of interspecific oocytes on demethylation of sperm DNA. *Proc Natl Acad Sci U S A* 101: 7636-40
Bednar J, Garcia-Saez I, Boopathi R, Cutter AR, Papai G, Reymer A, Syed SH, Lone IN, Tonchev O, Crucifix C (2017) Structure and dynamics of a 197 bp nucleosome in complex with linker histone H1. *Molecular cell* 66: 384-397. e8
Belotserkovskaya R, Oh S, Bondarenko VA, Orphanides G, Studitsky VM, Reinberg D (2003) FACT facilitates transcription-dependent nucleosome alteration. *Science* 301: 1090-3
Biggiogera M, Muller S, Courtens JL, Fakan S, Romanini MG (1992) Immunoelectron microscopical distribution of histones H2B and H3 and protamines in the course of mouse spermiogenesis. *Microsc Res Tech* 20: 259-67
Bilokapic S, Strauss M, Halic M (2018) Histone octamer rearranges to adapt to DNA unwrapping. *Nat Struct Mol Biol* 25: 101-108
Bogdanovic O, Lister R (2017) DNA methylation and the preservation of cell identity. *Curr Opin Genet Dev* 46: 9-14
Borsos M, Torres-Padilla ME (2016) Building up the nucleus: nuclear organization in the establishment of totipotency and pluripotency during mammalian development. *Genes Dev* 30: 611-21
Bostick M, Kim JK, Esteve PO, Clark A, Pradhan S, Jacobsen SE (2007) UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* 317: 1760-4

Bouhallier F, Allioli N, Lavial F, Chalmel F, Perrard MH, Durand P, Samarut J, Pain B, Rouault JP (2010) Role of miR-34c microRNA in the late steps of spermatogenesis. *RNA* 16: 720-31
Boulard M, Gautier T, Mbele GO, Gerson V, Hamiche A, Angelov D, Bouvet P, Dimitrov S (2006) The NH2 tail of the novel histone variant H2BFWT exhibits properties distinct from conventional H2B with respect to the assembly of mitotic chromosomes. *Mol Cell Biol* 26: 1518-26
Boussouar F, Goudarzi A, Buchou T, Shiota H, Barral S, Debernardi A, Guardiola P, Brindle P, Martinez G, Arnoult C, Khochbin S, Rousseaux S (2014) A specific CBP/p300-dependent gene expression programme drives the metabolic remodelling in late stages of spermatogenesis. *Andrology* 2: 351-9
Boussouar F, Rousseaux S, Khochbin S (2008) A new insight into male genome reprogramming by histone variants and histone code. *Cell Cycle* 7: 3499-502
Bowman A, Ward R, Wiechens N, Singh V, El-Mkami H, Norman DG, Owen-Hughes T (2011) The histone chaperones Nap1 and Vps75 bind histones H3 and H4 in a tetrameric conformation. *Mol Cell* 41: 398-408
Boyault C, Gilquin B, Zhang Y, Rybin V, Garman E, Meyer-Klaucke W, Matthias P, Muller CW, Khochbin S (2006) HDAC6-p97/VCP controlled polyubiquitin chain turnover. *EMBO J* 25: 3357-66
Boyault C, Zhang Y, Fritah S, Caron C, Gilquin B, Kwon SH, Garrido C, Yao TP, Vourc'h C, Matthias P, Khochbin S (2007) HDAC6 controls major cell response pathways to cytotoxic accumulation of protein aggregates. *Genes Dev* 21: 2172-81
Boyer LA, Latek RR, Peterson CL (2004) The SANT domain: a unique histone-tail-binding module? *Nat Rev Mol Cell Biol* 5: 158-63
Brachet E, Beneut C, Serrentino ME, Borde V (2015) The CAF-1 and Hir Histone Chaperones Associate with Sites of Meiotic Double-Strand Breaks in Budding Yeast. *PLoS One* 10: e0125965
Branson RE, Grimes SR, Jr., Yonuschot G, Irvin JL (1975) The histones of rat testis. *Arch Biochem Biophys* 168: 403-12
Brownell JE, Allis CD (1995) An activity gel assay detects a single, catalytically active histone acetyltransferase subunit in Tetrahymena macronuclei. *Proc Natl Acad Sci U S A* 92: 6364-8
Brownell JE, Zhou J, Ranalli T, Kobayashi R, Edmondson DG, Roth SY, Allis CD (1996) Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* 84: 843-51
Brykczynska U, Hisano M, Erkek S, Ramos L, Oakeley EJ, Roloff TC, Beisel C, Schubeler D, Stadler MB, Peters AH (2010) Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. *Nat Struct Mol Biol* 17: 679-87
Burgess RJ, Zhang Z (2013) Histone chaperones in nucleosome assembly and human disease. *Nat Struct Mol Biol* 20: 14-22
Buschbeck M, Hake SB (2017) Variants of core histones and their roles in cell fate decisions, development and cancer. *Nat Rev Mol Cell Biol* 18: 299-314
Cai Y, Jin J, Swanson SK, Cole MD, Choi SH, Florens L, Washburn MP, Conaway JW, Conaway RC (2010) Subunit composition and substrate specificity of a MOF-containing histone acetyltransferase distinct from the male-specific lethal (MSL) complex. *J Biol Chem* 285: 4268-72
Campbell RR, Wood MA (2019) How the epigenome integrates information and reshapes the synapse. *Nat Rev Neurosci* 20: 133-147

Carone BR, Hung JH, Hainer SJ, Chou MT, Carone DM, Weng Z, Fazzio TG, Rando OJ (2014) High-resolution mapping of chromatin packaging in mouse embryonic stem cells and sperm. *Dev Cell* 30: 11-22
Carrozza MJ, Utley RT, Workman JL, Cote J (2003) The diverse functions of histone acetyltransferase complexes. *Trends Genet* 19: 321-9
Chadwick BP, Willard HF (2001) Histone H2A variants and the inactive X chromosome: identification of a second macroH2A variant. *Hum Mol Genet* 10: 1101-13
Chakravarthy S, Gundimella SK, Caron C, Perche PY, Pehrson JR, Khochbin S, Luger K (2005) Structural characterization of the histone variant macroH2A. *Mol Cell Biol* 25: 7616-24
Chen CC, Dechassa ML, Bettini E, Ledoux MB, Belisario C, Heun P, Luger K, Mellone BG (2014) CAL1 is the Drosophila CENP-A assembly factor. *J Cell Biol* 204: 313-29
Chen P, Zhao J, Wang Y, Wang M, Long H, Liang D, Huang L, Wen Z, Li W, Li X, Feng H, Zhao H, Zhu P, Li M, Wang QF, Li G (2013) H3.3 actively marks enhancers and primes gene transcription via opening higher-ordered chromatin. *Genes Dev* 27: 2109-24
Chen Z, Zang J, Whetstine J, Hong X, Davrazou F, Kutateladze TG, Simpson M, Mao Q, Pan CH, Dai S, Hagman J, Hansen K, Shi Y, Zhang G (2006) Structural insights into histone demethylation by JMJD2 family members. *Cell* 125: 691-702
Cho C, Willis WD, Goulding EH, Jung-Ha H, Choi YC, Hecht NB, Eddy EM (2001) Haploinsufficiency of protamine-1 or -2 causes infertility in mice. *Nat Genet* 28: 82-6
Chodaparambil JV, Barbera AJ, Lu X, Kaye KM, Hansen JC, Luger K (2007) A charged and contoured surface on the nucleosome regulates chromatin compaction. *Nat Struct Mol Biol* 14: 1105-7
Choi YC, Chae CB (1993) Demethylation of somatic and testis-specific histone H2A and H2B genes in F9 embryonal carcinoma cells. *Mol Cell Biol* 13: 5538-48
Christensen ME, Dixon GH (1982) Hyperacetylation of histone H4 correlates with the terminal, transcriptionally inactive stages of spermatogenesis in rainbow trout. *Dev Biol* 93: 404-15
Christophorou MA, Castelo-Branco G, Halley-Stott RP, Oliveira CS, Loos R, Radzisheuskaya A, Mowen KA, Bertone P, Silva JC, Zernicka-Goetz M, Nielsen ML, Gurdon JB, Kouzarides T (2014) Citrullination regulates pluripotency and histone H1 binding to chromatin. *Nature* 507: 104-8
Chua EY, Vasudevan D, Davey GE, Wu B, Davey CA (2012) The mechanics behind DNA sequence-dependent properties of the nucleosome. *Nucleic Acids Res* 40: 6338-52
Churikov D, Siino J, Svetlova M, Zhang K, Gineitis A, Morton Bradbury E, Zalensky A (2004) Novel human testis-specific histone H2B encoded by the interrupted gene on the X chromosome. *Genomics* 84: 745-56
Clapier CR, Iwasa J, Cairns BR, Peterson CL (2017) Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat Rev Mol Cell Biol* 18: 407-422
Cosgrove MS, Boeke JD, Wolberger C (2004) Regulated nucleosome mobility and the histone code. *Nat Struct Mol Biol* 11: 1037-43
Cremer T, Cremer C (2001) Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet* 2: 292-301
Culty M (2013) Gonocytes, from the fifties to the present: is there a reason to change the name? *Biol Reprod* 89: 46
Dancy BM, Cole PA (2015) Protein lysine acetylation by p300/CBP. *Chem Rev* 115: 2419-52
Dang W, Bartholomew B (2007) Domain architecture of the catalytic subunit in the ISW2-nucleosome complex. *Mol Cell Biol* 27: 8306-17

De Lucia F, Faraone-Mennella MR, D'Erme M, Quesada P, Caiafa P, Farina B (1994) Histone-induced condensation of rat testis chromatin: testis-specific H1t versus somatic H1 variants. *Biochem Biophys Res Commun* 198: 32-9
Deng W, Rupon JW, Krivega I, Breda L, Motta I, Jahn KS, Reik A, Gregory PD, Rivella S, Dean A, Blobel GA (2014) Reactivation of developmentally silenced globin genes by forced chromatin looping. *Cell* 158: 849-860
Dhayalan A, Tamas R, Bock I, Tattermusch A, Dimitrova E, Kudithipudi S, Ragozin S, Jeltsch A (2011) The ATRX-ADD domain binds to H3 tail peptides and reads the combined methylation state of K4 and K9. *Hum Mol Genet* 20: 2195-203
Di Lorenzo A, Bedford MT (2011) Histone arginine methylation. *FEBS Lett* 585: 2024-31
Dodge JE, Kang YK, Beppu H, Lei H, Li E (2004) Histone H3-K9 methyltransferase ESET is essential for early development. *Mol Cell Biol* 24: 2478-86
Doyen CM, An W, Angelov D, Bondarenko V, Mietton F, Studitsky VM, Hamiche A, Roeder RG, Bouvet P, Dimitrov S (2006) Mechanism of polymerase II transcription repression by the histone variant macroH2A. *Mol Cell Biol* 26: 1156-64
Drabent B, Kardalinou E, Doenecke D (1991) Structure and expression of the human gene encoding testicular H1 histone (H1t). *Gene* 103: 263-8
Dunleavy EM, Roche D, Tagami H, Lacoste N, Ray-Gallet D, Nakamura Y, Daigo Y, Nakatani Y, Almouzni-Pettinotti G (2009) HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. *Cell* 137: 485-97
Durdevic Z, Mobin MB, Hanna K, Lyko F, Schaefer M (2013) The RNA methyltransferase Dnmt2 is required for efficient Dicer-2-dependent siRNA pathway activity in Drosophila. *Cell Rep* 4: 931-7
Eirin-Lopez JM, Ishibashi T, Ausio J (2008) H2A.Bbd: a quickly evolving hypervariable mammalian histone that destabilizes nucleosomes in an acetylation-independent way. *FASEB J* 22: 316-26
El Kennani S, Adrait A, Shaytan AK, Khochbin S, Bruley C, Panchenko AR, Landsman D, Pflieger D, Govin J (2017) MS_HistoneDB, a manually curated resource for proteomic analysis of human and mouse histones. *Epigenetics Chromatin* 10: 2
Elsasser SJ, Huang H, Lewis PW, Chin JW, Allis CD, Patel DJ (2012) DAXX envelops a histone H3.3-H4 dimer for H3.3-specific recognition. *Nature* 491: 560-5
Elsasser SJ, Noh KM, Diaz N, Allis CD, Banaszynski LA (2015) Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells. *Nature* 522: 240-244
Erkek S, Hisano M, Liang CY, Gill M, Murr R, Dieker J, Schubeler D, van der Vlag J, Stadler MB, Peters AH (2013) Molecular determinants of nucleosome retention at CpG-rich sequences in mouse spermatozoa. *Nat Struct Mol Biol* 20: 868-75
Eustermann S, Yang JC, Law MJ, Amos R, Chapman LM, Jelinska C, Garrick D, Clynes D, Gibbons RJ, Rhodes D, Higgs DR, Neuhaus D (2011) Combinatorial readout of histone H3 modifications specifies localization of ATRX to heterochromatin. *Nat Struct Mol Biol* 18: 777-82
Faulkner RD, Bhatnagar YM (1987) A protease activity is associated with testicular chromatin of the mouse. *Biol Reprod* 36: 471-80
Fei J, Torigoe SE, Brown CR, Khuong MT, Kassavetis GA, Boeger H, Kadonaga JT (2015) The prenucleosome, a stable conformational isomer of the nucleosome. *Genes Dev* 29: 2563-75
Felsenfeld G, Groudine M (2003) Controlling the double helix. *Nature* 421: 448-53
Flemming W (1882) *Zellsubstanz, kern und zelltheilung*. Vogel,
Florence B, Faller DV (2001) You bet-cha: a novel family of transcriptional regulators. *Front Biosci* 6: D1008-18

Foltz DR, Jansen LE, Bailey AO, Yates JR, 3rd, Bassett EA, Wood S, Black BE, Cleveland DW (2009) Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. *Cell* 137: 472-84
Franklin RE, Gosling RG (1953) Molecular configuration in sodium thymonucleate. *Nature* 171: 740
Franklin SG, Zweidler A (1977) Non-allelic variants of histones 2a, 2b and 3 in mammals. *Nature* 266: 273-5
French CA, Miyoshi I, Kubonishi I, Grier HE, Perez-Atayde AR, Fletcher JA (2003) BRD4-NUT fusion oncogene: a novel mechanism in aggressive carcinoma. *Cancer Res* 63: 304-7
Fujita Y, Hayashi T, Kiyomitsu T, Toyoda Y, Kokubu A, Obuse C, Yanagida M (2007) Priming of centromere for CENP-A recruitment by human hMis18alpha, hMis18beta, and M18BP1. *Dev Cell* 12: 17-30
Garcia BA, Hake SB, Diaz RL, Kauer M, Morris SA, Recht J, Shabanowitz J, Mishra N, Strahl BD, Allis CD, Hunt DF (2007) Organismal differences in post-translational modifications in histones H3 and H4. *J Biol Chem* 282: 7641-55
Garcia-Saez I, Menoni H, Boopathi R, Shukla MS, Soueidan L, Noirclerc-Savoye M, Le Roy A, Skoufias DA, Bednar J, Hamiche A, Angelov D, Petosa C, Dimitrov S (2018) Structure of an H1-Bound 6-Nucleosome Array Reveals an Untwisted Two-Start Chromatin Fiber Conformation. *Mol Cell* 72: 902-915 e7
Gaucher J, Boussouar F, Montellier E, Curtet S, Buchou T, Bertrand S, Hery P, Jounier S, Depaux A, Vitte AL, Guardiola P, Pernet K, Debernardi A, Lopez F, Holota H, Imbert J, Wolgemuth DJ, Gerard M, Rousseaux S, Khochbin S (2012) Bromodomain-dependent stage-specific male genome programming by Brdt. *EMBO J* 31: 3809-20
Gaucher J, Reynoird N, Montellier E, Boussouar F, Rousseaux S, Khochbin S (2010) From meiosis to postmeiotic events: the secrets of histone disappearance. *FEBS J* 277: 599-604
Gautier T, Abbott DW, Molla A, Verdel A, Ausio J, Dimitrov S (2004) Histone variant H2ABbd confers lower stability to the nucleosome. *EMBO Rep* 5: 715-20
Georgakopoulos T, Thireos G (1992) Two distinct yeast transcriptional activators require the function of the GCN5 protein to promote normal levels of transcription. *EMBO J* 11: 4145-52
Germond J-E, Bellard M, Oudet P, Chambon P (1976) Stability of nucleosomes in native and reconstituted chromatins. *Nucleic acids research* 3: 3173-3192
Giaimo BD, Ferrante F, Herchenrother A, Hake SB, Borggrefe T (2019) The histone variant H2A.Z in gene regulation. *Epigenetics Chromatin* 12: 37
Goll MG, Bestor TH (2005) Eukaryotic cytosine methyltransferases. *Annu Rev Biochem* 74: 481-514
Goudarzi A, Shiota H, Rousseaux S, Khochbin S (2014) Genome-scale acetylation-dependent histone eviction during spermatogenesis. *J Mol Biol* 426: 3342-9
Goudarzi A, Zhang D, Huang H, Barral S, Kwon OK, Qi S, Tang Z, Buchou T, Vitte AL, He T, Cheng Z, Montellier E, Gaucher J, Curtet S, Debernardi A, Charbonnier G, Puthier D, Petosa C, Panne D, Rousseaux S et al. (2016) Dynamic Competing Histone H4 K5K8 Acetylation and Butyrylation Are Hallmarks of Highly Active Gene Promoters. *Mol Cell* 62: 169-180
Govin J, Caron C, Lestrat C, Rousseaux S, Khochbin S (2004) The role of histones in chromatin remodelling during mammalian spermiogenesis. *Eur J Biochem* 271: 3459-69
Govin J, Caron C, Rousseaux S, Khochbin S (2005) Testis-specific histone H3 expression in somatic cells. *Trends Biochem Sci* 30: 357-9
Govin J, Escoffier E, Rousseaux S, Kuhn L, Ferro M, Thevenon J, Catena R, Davidson I, Garin J, Khochbin S, Caron C (2007) Pericentric heterochromatin reprogramming by new histone variants during mouse spermiogenesis. *J Cell Biol* 176: 283-94

Green EM, Antczak AJ, Bailey AO, Franco AA, Wu KJ, Yates JR, 3rd, Kaufman PD (2005) Replication-independent histone deposition by the HIR complex and Asf1. *Curr Biol* 15: 2044-9
Greer EL, Shi Y (2012) Histone methylation: a dynamic mark in health, disease and inheritance. *Nat Rev Genet* 13: 343-57
Grimes SR, Jr., Henderson N (1984) Hyperacetylation of histone H4 in rat testis spermatids. *Exp Cell Res* 152: 91-7
Grimes SR, Wolfe SA, Anderson JV, Stein GS, Stein JL (1990) Structural and functional analysis of the rat testis-specific histone H1t gene. *J Cell Biochem* 44: 1-17
Groth A, Ray-Gallet D, Quivy JP, Lukas J, Bartek J, Almouzni G (2005) Human Asf1 regulates the flow of S phase histones during replicational stress. *Mol Cell* 17: 301-11
Grune T, Brzeski J, Eberharter A, Clapier CR, Corona DF, Becker PB, Muller CW (2003) Crystal structure and functional analysis of a nucleosome recognition module of the remodeling factor ISWI. *Mol Cell* 12: 449-60
Guelman S, Kozuka K, Mao Y, Pham V, Solloway MJ, Wang J, Wu J, Lill JR, Zha J (2009) The double-histone-acetyltransferase complex ATAC is essential for mammalian development. *Mol Cell Biol* 29: 1176-88
Hada M, Masuda K, Yamaguchi K, Shirahige K, Okada Y (2017) Identification of a variant-specific phosphorylation of TH2A during spermiogenesis. *Sci Rep* 7: 46228
Hammond CM, Stromme CB, Huang H, Patel DJ, Groth A (2017) Histone chaperone networks shaping chromatin function. *Nat Rev Mol Cell Biol* 18: 141-158
Hammoud SS, Nix DA, Zhang H, Purwar J, Carrell DT, Cairns BR (2009) Distinctive chromatin in human sperm packages genes for embryo development. *Nature* 460: 473-8
Hauk G, McKnight JN, Nodelman IM, Bowman GD (2010) The chromodomains of the Chd1 chromatin remodeler regulate DNA access to the ATPase motor. *Mol Cell* 39: 711-23
Hayashi K, Chuva de Sousa Lopes SM, Kaneda M, Tang F, Hajkova P, Lao K, O'Carroll D, Das PP, Tarakhovsky A, Miska EA, Surani MA (2008) MicroRNA biogenesis is required for mouse primordial germ cell development and spermatogenesis. *PLoS One* 3: e1738
Hazzouri M, Pivot-Pajot C, Faure AK, Usson Y, Pelletier R, Sele B, Khochbin S, Rousseaux S (2000) Regulated hyperacetylation of core histones during mouse spermatogenesis: involvement of histone deacetylases. *Eur J Cell Biol* 79: 950-60
He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L, Sun Y, Li X, Dai Q, Song CX, Zhang K, He C, Xu GL (2011) Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* 333: 1303-7
Henikoff JG, Belsky JA, Krassovsky K, MacAlpine DM, Henikoff S (2011) Epigenome characterization at single base-pair resolution. *Proc Natl Acad Sci U S A* 108: 18318-23
Hergeth SP, Schneider R (2015) The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. *EMBO Rep* 16: 1439-53
Hershey AD, Chase M (1952) Independent functions of viral protein and nucleic acid in growth of bacteriophage. *The Journal of general physiology* 36: 39-56
Hewish DR, Burgoyne LA (1973) Chromatin sub-structure. The digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease. *Biochemical and biophysical research communications* 52: 504-510
Hnisz D, Shrinivas K, Young RA, Chakraborty AK, Sharp PA (2017) A Phase Separation Model for Transcriptional Control. *Cell* 169: 13-23
Hoghoughi N, Barral S, Vargas A, Rousseaux S, Khochbin S (2018) Histone variants: essential actors in male genome programming. *J Biochem* 163: 97-103

Hsieh FK, Kulaeva OI, Patel SS, Dyer PN, Luger K, Reinberg D, Studitsky VM (2013) Histone chaperone FACT action during transcription through chromatin by RNA polymerase II. *Proc Natl Acad Sci U S A* 110: 7654-9
Hu Y, Shen Y, Conde ESN, Zhou DX (2011) The role of histone methylation and H2A.Z occupancy during rapid activation of ethylene responsive genes. *PLoS One* 6: e28224
Hubbert C, Guardiola A, Shao R, Kawaguchi Y, Ito A, Nixon A, Yoshida M, Wang XF, Yao TP (2002) HDAC6 is a microtubule-associated deacetylase. *Nature* 417: 455-8
Huh NE, Hwang IW, Lim K, You KH, Chae CB (1991) Presence of a bi-directional S phase-specific transcription regulatory element in the promoter shared by testis-specific TH2A and TH2B histone genes. *Nucleic Acids Res* 19: 93-8
Iguchi N, Tanaka H, Yamada S, Nishimura H, Nishimune Y (2004) Control of mouse hils1 gene expression during spermatogenesis: identification of regulatory element by transgenic mouse. *Biol Reprod* 70: 1239-45
Ikura T, Ogryzko VV, Grigoriev M, Groisman R, Wang J, Horikoshi M, Scully R, Qin J, Nakatani Y (2000) Involvement of the TIP60 histone acetylase complex in DNA repair and apoptosis. *Cell* 102: 463-73
Imai R, Nozaki T, Tani T, Kaizu K, Hibino K, Ide S, Tamura S, Takahashi K, Shribak M, Maeshima K (2017) Density imaging of heterochromatin in live cells using orientation-independent-DIC microscopy. *Mol Biol Cell* 28: 3349-3359
Ito S, Kayukawa N, Ueda T, Taniguchi H, Morioka Y, Hongo F, Ukimura O (2018) MRGBP promotes AR-mediated transactivation of KLK3 and TMPRSS2 via acetylation of histone H2A.Z in prostate cancer cells. *Biochim Biophys Acta Gene Regul Mech*
Iwase S, Xiang B, Ghosh S, Ren T, Lewis PW, Cochrane JC, Allis CD, Picketts DJ, Patel DJ, Li H, Shi Y (2011) ATRX ADD domain links an atypical histone methylation recognition mechanism to human mental-retardation syndrome. *Nat Struct Mol Biol* 18: 769-76
Jan SZ, Hamer G, Repping S, de Rooij DG, van Pelt AM, Vormer TL (2012) Molecular control of rodent spermatogenesis. *Biochim Biophys Acta* 1822: 1838-50
Jeronimo C, Watanabe S, Kaplan CD, Peterson CL, Robert F (2015) The Histone Chaperones FACT and Spt6 Restrict H2A.Z from Intragenic Locations. *Mol Cell* 58: 1113-23
Ji D, Lin K, Song J, Wang Y (2014) Effects of Tet-induced oxidation products of 5-methylcytosine on Dnmt1- and DNMT3a-mediated cytosine methylation. *Mol Biosyst* 10: 1749-52
Jin C, Felsenfeld G (2007) Nucleosome stability mediated by histone variants H3.3 and H2A.Z. *Genes Dev* 21: 1519-29
Jin Q, Yu LR, Wang L, Zhang Z, Kasper LH, Lee JE, Wang C, Brindle PK, Dent SY, Ge K (2011) Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. *EMBO J* 30: 249-62
Jolly C, Metz A, Govin J, Vigneron M, Turner BM, Khochbin S, Vourc'h C (2004) Stress-induced transcription of satellite III repeats. *J Cell Biol* 164: 25-33
Jost KL, Bertulat B, Cardoso MC (2012) Heterochromatin and gene positioning: inside, outside, any side? *Chromosoma* 121: 555-563
Jung YH, Sauria MEG, Lyu X, Cheema MS, Ausio J, Taylor J, Corces VG (2017) Chromatin States in Mouse Sperm Correlate with Embryonic and Adult Regulatory Landscapes. *Cell Rep* 18: 1366-1382
Kaczmarska Z, Ortega E, Goudarzi A, Huang H, Kim S, Marquez JA, Zhao Y, Khochbin S, Panne D (2017) Structure of p300 in complex with acyl-CoA variants. *Nat Chem Biol* 13: 21-29

Kawaguchi Y, Kovacs JJ, McLaurin A, Vance JM, Ito A, Yao TP (2003) The deacetylase HDAC6 regulates aggresome formation and cell viability in response to misfolded protein stress. *Cell* 115: 727-38
Kelsey G, Feil R (2013) New insights into establishment and maintenance of DNA methylation imprints in mammals. *Philos Trans R Soc Lond B Biol Sci* 368: 20110336
Khadake JR, Rao MR (1995) DNA- and chromatin-condensing properties of rat testes H1a and H1t compared to those of rat liver H1bdec; H1t is a poor condenser of chromatin. *Biochemistry* 34: 15792-801
Khochbin S (2001) Histone H1 diversity: bridging regulatory signals to linker histone function. *Gene* 271: 1-12
Khochbin S, Wolffe AP (1994) Developmentally regulated expression of linker-histone variants in vertebrates. *Eur J Biochem* 225: 501-10
Khorasanizadeh S (2004) The nucleosome: from genomic organization to genomic regulation. *Cell* 116: 259-72
Kimmins S, Sassone-Corsi P (2005) Chromatin remodelling and epigenetic features of germ cells. *Nature* 434: 583-9
Kornberg RD (1974) Chromatin structure: a repeating unit of histones and DNA. *Science* 184: 868-871
Kossel A (1884) Ueber einen peptonartigen Bestandtheil des Zellkerns. *Zeitschrift für physiologische Chemie* 8: 511-515
Kouzarides T (2007) Chromatin modifications and their function. *Cell* 128: 693-705
Kujirai T, Arimura Y, Fujita R, Horikoshi N, Machida S, Kurumizaka H (2018) Methods for Preparing Nucleosomes Containing Histone Variants. *Methods Mol Biol* 1832: 3-20
Kunert N, Brehm A (2009) Novel Mi-2 related ATP-dependent chromatin remodelers. *Epigenetics* 4: 209-11
Kurtz K, Martinez-Soler F, Ausio J, Chiva M (2007) Acetylation of histone H4 in complex structural transitions of spermiogenic chromatin. *J Cell Biochem* 102: 1432-41
Kusch T, Mei A, Nguyen C (2014) Histone H3 lysine 4 trimethylation regulates cotranscriptional H2A variant exchange by Tip60 complexes to maximize gene expression. *Proc Natl Acad Sci U S A* 111: 4850-5
LeRoy G, Rickards B, Flint SJ (2008) The double bromodomain proteins Brd2 and Brd3 couple histone acetylation to transcription. *Mol Cell* 30: 51-60
Li A, Eirin-Lopez JM, Ausio J (2005) H2AX: tailoring histone H2A for chromatin-dependent genomic integrity. *Biochem Cell Biol* 83: 505-15
Li E, Bestor TH, Jaenisch R (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69: 915-26
Li X, Tyler JK (2016) Nucleosome disassembly during human non-homologous end joining followed by concerted HIRA- and CAF-1-dependent reassembly. *Elife* 5
Lin Q, Sirotkin A, Skoultchi AI (2000) Normal spermatogenesis in mice lacking the testis-specific linker histone H1t. *Mol Cell Biol* 20: 2122-8
Loppin B, Bonnefoy E, Anselme C, Laurencon A, Karr TL, Couble P (2005) The histone H3.3 chaperone HIRA is essential for chromatin assembly in the male pronucleus. *Nature* 437: 1386-90
Loyola A, Almouzni G (2004) Histone chaperones, a supporting role in the limelight. *Biochim Biophys Acta* 1677: 3-11
Lu LY, Wu J, Ye L, Gavrilina GB, Saunders TL, Yu X (2010) RNF8-dependent histone modifications regulate nucleosome removal during spermatogenesis. *Dev Cell* 18: 371-84

Lu S, Xie YM, Li X, Luo J, Shi XQ, Hong X, Pan YH, Ma X (2009) Mass spectrometry analysis of dynamic post-translational modifications of TH2B during spermatogenesis. *Mol Hum Reprod* 15: 373-8
Luger K, Dechassa ML, Tremethick DJ (2012) New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? *Nat Rev Mol Cell Biol* 13: 436-47
Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389: 251
Maddox PS, Hyndman F, Monen J, Oegema K, Desai A (2007) Functional genomics identifies a Myb domain-containing protein family required for assembly of CENP-A chromatin. *J Cell Biol* 176: 757-63
Maehara K, Harada A, Sato Y, Matsumoto M, Nakayama KI, Kimura H, Ohkawa Y (2015) Tissue-specific expression of histone H3 variants diversified after species separation. *Epigenetics Chromatin* 8: 35
Mali P, Kaipia A, Kangasniemi M, Toppari J, Sandberg M, Hecht NB, Parvinen M (1989) Stage-specific expression of nucleoprotein mRNAs during rat and mouse spermiogenesis. *Reprod Fertil Dev* 1: 369-82
Malone CD, Hannon GJ (2009) Molecular evolution of piRNA and transposon control pathways in Drosophila. *Cold Spring Harb Symp Quant Biol* 74: 225-34
Manku G, Culty M (2015) Mammalian gonocyte and spermatogonia differentiation: recent advances and remaining challenges. *Reproduction* 149: R139-57
Martianov I, Brancorsini S, Catena R, Gansmuller A, Kotaja N, Parvinen M, Sassone-Corsi P, Davidson I (2005) Polar nuclear localization of H1T2, a histone H1 variant, required for spermatid elongation and DNA condensation during spermiogenesis. *Proc Natl Acad Sci U S A* 102: 2808-13
Martin C, Zhang Y (2005) The diverse functions of histone lysine methylation. *Nat Rev Mol Cell Biol* 6: 838-49
Martin C, Zhang Y (2007) Mechanisms of epigenetic inheritance. *Curr Opin Cell Biol* 19: 266-72
Martire S, Gogate AA, Whitmill A, Tafessu A, Nguyen J, Teng YC, Tastemel M, Banaszynski LA (2019) Phosphorylation of histone H3.3 at serine 31 promotes p300 activity and enhancer acetylation. *Nat Genet* 51: 941-946
Marushige K, Marushige Y, Wong TK (1976) Complete displacement of somatic histones during transformation of spermatid chromatin: a model experiment. *Biochemistry* 15: 2047-53
Marzluff WF, Gongidi P, Woods KR, Jin J, Maltais LJ (2002) The human and mouse replication-dependent histone genes. *Genomics* 80: 487-98
Matsuyama A, Shimazu T, Sumida Y, Saito A, Yoshimatsu Y, Seigneurin-Berny D, Osada H, Komatsu Y, Nishino N, Khochbin S, Horinouchi S, Yoshida M (2002) In vivo destabilization of dynamic microtubules by HDAC6-mediated deacetylation. *EMBO J* 21: 6820-31
Mayer W, Niveleau A, Walter J, Fundele R, Haaf T (2000) Demethylation of the zygotic paternal genome. *Nature* 403: 501-2
McGinty RK, Tan S (2014) Nucleosome structure and function. *Chemical reviews* 115: 2255-2273
McKinley KL, Cheeseman IM (2016) The molecular basis for centromere identity and function. *Nat Rev Mol Cell Biol* 17: 16-29
Meistrich ML, Hess RA (2013) Assessment of spermatogenesis through staging of seminiferous tubules. *Methods Mol Biol* 927: 299-307
Mello JA, Sillje HH, Roche DM, Kirschner DB, Nigg EA, Almouzni G (2002) Human Asf1 and CAF-1 interact and synergize in a repair-coupled nucleosome assembly pathway. *EMBO Rep* 3: 329-34

Mersfelder EL, Parthun MR (2006) The tale beyond the tail: histone core domain modifications and the regulation of chromatin structure. *Nucleic Acids Res* 34: 2653-62
Meyer-Ficca ML, Lonchar JD, Ihara M, Bader JJ, Meyer RG (2013) Alteration of poly(ADP-ribose) metabolism affects murine sperm nuclear architecture by impairing pericentric heterochromatin condensation. *Chromosoma* 122: 319-35
Miescher-Rüsch F (1871) *Ueber die chemische Zusammensetzung der Eiterzellen*.
Mietton F, Sengupta AK, Molla A, Picchi G, Barral S, Heliot L, Grange T, Wutz A, Dimitrov S (2009) Weak but uniform enrichment of the histone variant macroH2A1 along the inactive X chromosome. *Mol Cell Biol* 29: 150-6
Miller TC, Simon B, Rybin V, Grotsch H, Curtet S, Khochbin S, Carlomagno T, Muller CW (2016) A bromodomain-DNA interaction facilitates acetylation-dependent bivalent nucleosome recognition by the BET protein BRDT. *Nat Commun* 7: 13855
Mishra LN, Gupta N, Rao SM (2015) Mapping of post-translational modifications of spermatid-specific linker histone H1-like protein, HILS1. *J Proteomics* 128: 218-30
Mishra LN, Shalini V, Gupta N, Ghosh K, Suthar N, Bhaduri U, Rao MRS (2018) Spermatid-specific linker histone HILS1 is a poor condenser of DNA and chromatin and preferentially associates with LINE-1 elements. *Epigenetics Chromatin* 11: 43
Molaro A, Young JM, Malik HS (2018) Evolutionary origins and diversification of testis-specific short histone H2A variants in mammals. *Genome Res* 28: 460-473
Montel F, Menoni H, Castelnovo M, Bednar J, Dimitrov S, Angelov D, Faivre-Moskalenko C (2009) The dynamics of individual nucleosomes controls the chromatin condensation pathway: direct atomic force microscopy visualization of variant chromatin. *Biophys J* 97: 544-53
Montellier E, Boussouar F, Rousseaux S, Zhang K, Buchou T, Fenaille F, Shiota H, Debernardi A, Hery P, Curtet S, Jamshidikia M, Barral S, Holota H, Bergon A, Lopez F, Guardiola P, Pernet K, Imbert J, Petosa C, Tan M et al. (2013) Chromatin-to-nucleoprotamine transition is controlled by the histone H2B variant TH2B. *Genes Dev* 27: 1680-92
Moriniere J, Rousseaux S, Steuerwald U, Soler-Lopez M, Curtet S, Vitte AL, Govin J, Gaucher J, Sadoul K, Hart DJ, Krijgsveld J, Khochbin S, Muller CW, Petosa C (2009) Cooperative binding of two acetylation marks on a histone tail by a single bromodomain. *Nature* 461: 664-8
Muller S, Almouzni G (2017) Chromatin dynamics during the cell cycle at centromeres. *Nat Rev Genet* 18: 192-208
Nakamura T, Liu YJ, Nakashima H, Umehara H, Inoue K, Matoba S, Tachibana M, Ogura A, Shinkai Y, Nakano T (2012) PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. *Nature* 486: 415-9
Nardi IK, Zasadzinska E, Stellfox ME, Knippler CM, Foltz DR (2016) Licensing of Centromeric Chromatin Assembly through the Mis18alpha-Mis18beta Heterotetramer. *Mol Cell* 61: 774-787
Natale F, Rapp A, Yu W, Maiser A, Harz H, Scholl A, Grulich S, Anton T, Horl D, Chen W, Durante M, Taucher-Scholz G, Leonhardt H, Cardoso MC (2017) Identification of the elementary structural units of the DNA damage response. *Nat Commun* 8: 15760
Neto FT, Bach PV, Najari BB, Li PS, Goldstein M (2016) Spermatogenesis in humans and its affecting factors. *Semin Cell Dev Biol* 59: 10-26
Nielsen AL, Oulad-Abdelghani M, Ortiz JA, Remboutsika E, Chambon P, Losson R (2001) Heterochromatin formation in mammalian cells: interaction between histones and HP1 proteins. *Mol Cell* 7: 729-39
Nirenberg MW, Matthaei JH (1961) The dependence of cell-free protein synthesis in E. coli upon naturally occurring or synthetic polyribonucleotides. *Proceedings of the National Academy of Sciences* 47: 1588-1602

Noble D (2015) Conrad Waddington and the origin of epigenetics. *J Exp Biol* 218: 816-8
Okano M, Takebayashi S, Okumura K, Li E (1999) Assignment of cytosine-5 DNA methyltransferases Dnmt3a and Dnmt3b to mouse chromosome bands 12A2-A3 and 2H1 by in situ hybridization. *Cytogenet Cell Genet* 86: 333-4
Olins AL, Olins DE (1974) Spheroid chromatin units (ν bodies). *Science* 183: 330-332
Olins DE, Olins AL (2003) Chromatin history: our view from the bridge. *Nat Rev Mol Cell Biol* 4: 809-14
Oliva R, Bazett-Jones D, Mezquita C, Dixon GH (1987) Factors affecting nucleosome disassembly by protamines in vitro. Histone hyperacetylation and chromatin structure, time dependence, and the size of the sperm nuclear proteins. *J Biol Chem* 262: 17016-25
Oliva R, Dixon GH (1991) Vertebrate protamine genes and the histone-to-protamine replacement reaction. *Prog Nucleic Acid Res Mol Biol* 40: 25-94
Oliva R, Mezquita C (1986) Marked differences in the ability of distinct protamines to disassemble nucleosomal core particles in vitro. *Biochemistry* 25: 6508-11
Ortega E, Rengachari S, Ibrahim Z, Hoghoughi N, Gaucher J, Holehouse AS, Khochbin S, Panne D (2018) Transcription factor dimerization activates the p300 acetyltransferase. *Nature* 562: 538-544
Oudet P, Gross-Bellard M, Chambon P (1975) Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. *Cell* 4: 281-300
Padavattan S, Thiruselvam V, Shinagawa T, Hasegawa K, Kumasaka T, Ishii S, Kumarevel T (2017) Structural analyses of the nucleosome complexes with human testis-specific histone variants, hTh2a and hTh2b. *Biophys Chem* 221: 41-48
Palmer DK, O'Day K, Margolis RL (1990) The centromere specific histone CENP-A is selectively retained in discrete foci in mammalian sperm nuclei. *Chromosoma* 100: 32-6
Pan D, Klare K, Petrovic A, Take A, Walstein K, Singh P, Rondelet A, Bird AW, Musacchio A (2017) CDK-regulated dimerization of M18BP1 on a Mis18 hexamer is necessary for CENP-A loading. *Elife* 6
Pardal AJ, Fernandes-Duarte F, Bowman AJ (2019) The histone chaperoning pathway: from ribosome to nucleosome. *Essays in biochemistry* 63: 29-43
Pehrson JR, Fried VA (1992) MacroH2A, a core histone containing a large nonhistone region. *Science* 257: 1398-400
Peng C, Lu Z, Xie Z, Cheng Z, Chen Y, Tan M, Luo H, Zhang Y, He W, Yang K, Zwaans BM, Tishkoff D, Ho L, Lombard D, He TC, Dai J, Verdin E, Ye Y, Zhao Y (2011) The first identification of lysine malonylation substrates and its regulatory enzyme. *Mol Cell Proteomics* 10: M111 012658
Perche PY, Vourc'h C, Konecny L, Souchier C, Robert-Nicoud M, Dimitrov S, Khochbin S (2000) Higher concentrations of histone macroH2A in the Barr body are correlated with higher nucleosome density. *Curr Biol* 10: 1531-4
Peters AH, Kubicek S, Mechtler K, O'Sullivan RJ, Derijck AA, Perez-Burgos L, Kohlmaier A, Opravil S, Tachibana M, Shinkai Y, Martens JH, Jenuwein T (2003) Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Mol Cell* 12: 1577-89
Piquet S, Le Parc F, Bai SK, Chevallier O, Adam S, Polo SE (2018) The Histone Chaperone FACT Coordinates H2A.X-Dependent Signaling and Repair of DNA Damage. *Mol Cell* 72: 888-901 e7
Pivot-Pajot C, Caron C, Govin J, Vion A, Rousseaux S, Khochbin S (2003) Acetylation-dependent chromatin reorganization by BRDT, a testis-specific bromodomain-containing protein. *Mol Cell Biol* 23: 5354-65

Plath K, Fang J, Mlynarczyk-Evans SK, Cao R, Worringer KA, Wang H, de la Cruz CC, Otte AP, Panning B, Zhang Y (2003) Role of histone H3 lysine 27 methylation in X inactivation. *Science* 300: 131-5
Probst AV, Dunleavy E, Almouzni G (2009) Epigenetic inheritance during the cell cycle. *Nat Rev Mol Cell Biol* 10: 192-206
Probst AV, Okamoto I, Casanova M, El Marjou F, Le Baccon P, Almouzni G (2010) A strand-specific burst in transcription of pericentric satellites is required for chromocenter formation and early mouse development. *Dev Cell* 19: 625-38
Probst AV, Santos F, Reik W, Almouzni G, Dean W (2007) Structural differences in centromeric heterochromatin are spatially reconciled on fertilisation in the mouse zygote. *Chromosoma* 116: 403-15
Qian MX, Pang Y, Liu CH, Haratake K, Du BY, Ji DY, Wang GF, Zhu QQ, Song W, Yu Y, Zhang XX, Huang HT, Miao S, Chen LB, Zhang ZH, Liang YN, Liu S, Cha H, Yang D, Zhai Y et al. (2013) Acetylation-mediated proteasomal degradation of core histones during DNA repair and spermatogenesis. *Cell* 153: 1012-24
Quivy JP, Roche D, Kirschner D, Tagami H, Nakatani Y, Almouzni G (2004) A CAF-1 dependent pool of HP1 during heterochromatin duplication. *EMBO J* 23: 3516-26
Raisner RM, Hartley PD, Meneghini MD, Bao MZ, Liu CL, Schreiber SL, Rando OJ, Madhani HD (2005) Histone variant H2A.Z marks the 5' ends of both active and inactive genes in euchromatin. *Cell* 123: 233-48
Ramaswamy A, Ioshikhes I (2013) Dynamics of modeled oligonucleosomes and the role of histone variant proteins in nucleosome organization. In *Advances in protein chemistry and structural biology*, pp 119-149. Elsevier
Rao BJ, Rao MR (1987) DNase I site mapping and micrococcal nuclease digestion of pachytene chromatin reveal novel structural features. *J Biol Chem* 262: 4472-6
Rathke C, Baarends WM, Awe S, Renkawitz-Pohl R (2014) Chromatin dynamics during spermiogenesis. *Biochim Biophys Acta* 1839: 155-68
Ray-Gallet D, Almouzni G (2010) Nucleosome dynamics and histone variants. *Essays Biochem* 48: 75-87
Rea S, Eisenhaber F, O'Carroll D, Strahl BD, Sun ZW, Schmid M, Opravil S, Mechtler K, Ponting CP, Allis CD, Jenuwein T (2000) Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406: 593-9
Reynoird N, Schwartz BE, Delvecchio M, Sadoul K, Meyers D, Mukherjee C, Caron C, Kimura H, Rousseaux S, Cole PA, Panne D, French CA, Khochbin S (2010) Oncogenesis by sequestration of CBP/p300 in transcriptionally inactive hyperacetylated chromatin domains. *EMBO J* 29: 2943-52
Rhind N, Gilbert DM (2013) DNA replication timing. *Cold Spring Harb Perspect Biol* 5: a010132
Richmond T, Finch J, Rushton B, Rhodes D, Klug A (1984) Structure of the nucleosome core particle at 7 Å resolution. *Nature* 311: 532
Richmond TJ, Davey CA (2003) The structure of DNA in the nucleosome core. *Nature* 423: 145
Ricketts MD, Frederick B, Hoff H, Tang Y, Schultz DC, Singh Rai T, Grazia Vizioli M, Adams PD, Marmorstein R (2015) Ubinuclein-1 confers histone H3.3-specific-binding by the HIRA histone chaperone complex. *Nat Commun* 6: 7711
Robinson PJ, Rhodes D (2006) Structure of the '30 nm' chromatin fibre: a key role for the linker histone. *Current opinion in structural biology* 16: 336-343
Roulland Y, Ouararhni K, Naidenov M, Ramos L, Shuaib M, Syed SH, Lone IN, Boopathi R, Fontaine E, Papai G, Tachiwana H, Gautier T, Skoufias D, Padmanabhan K, Bednar J,

Kurumizaka H, Schultz P, Angelov D, Hamiche A, Dimitrov S (2016) The Flexible Ends of CENP-A Nucleosome Are Required for Mitotic Fidelity. *Mol Cell* 63: 674-685
Rountree MR, Bachman KE, Baylin SB (2000) DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nat Genet* 25: 269-77
Royo H, Stadler MB, Peters A (2016) Alternative Computational Analysis Shows No Evidence for Nucleosome Enrichment at Repetitive Sequences in Mammalian Spermatozoa. *Dev Cell* 37: 98-104
Ruthenburg AJ, Li H, Patel DJ, Allis CD (2007) Multivalent engagement of chromatin modifications by linked binding modules. *Nat Rev Mol Cell Biol* 8: 983-94
Sabari BR, Tang Z, Huang H, Yong-Gonzalez V, Molina H, Kong HE, Dai L, Shimada M, Cross JR, Zhao Y, Roeder RG, Allis CD (2015) Intracellular crotonyl-CoA stimulates transcription through p300-catalyzed histone crotonylation. *Mol Cell* 58: 203-15
Sabari BR, Zhang D, Allis CD, Zhao Y (2017) Metabolic regulation of gene expression through histone acylations. *Nat Rev Mol Cell Biol* 18: 90-101
Saitou M, Kagiwada S, Kurimoto K (2012) Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells. *Development* 139: 15-31
Saitou M, Kurimoto K (2014) Paternal nucleosomes: are they retained in developmental promoters or gene deserts? *Dev Cell* 30: 6-8
Saldana-Meyer R, Rodriguez-Hernaez J, Escobar T, Nishana M, Jacome-Lopez K, Nora EP, Bruneau BG, Tsirigos A, Furlan-Magaril M, Skok J, Reinberg D (2019) RNA Interactions Are Essential for CTCF-Mediated Genome Organization. *Mol Cell*
Samans B, Yang Y, Krebs S, Sarode GV, Blum H, Reichenbach M, Wolf E, Steger K, Dansranjavin T, Schagdarsurengin U (2014) Uniformity of nucleosome preservation pattern in Mammalian sperm and its connection to repetitive DNA elements. *Dev Cell* 30: 23-35
Sanli I, Feil R (2015) Chromatin mechanisms in the developmental control of imprinted gene expression. *Int J Biochem Cell Biol* 67: 139-47
Santenard A, Ziegler-Birling C, Koch M, Tora L, Bannister AJ, Torres-Padilla ME (2010) Heterochromatin formation in the mouse embryo requires critical residues of the histone variant H3.3. *Nat Cell Biol* 12: 853-62
Sasaki K, Ito T, Nishino N, Khochbin S, Yoshida M (2009) Real-time imaging of histone H4 hyperacetylation in living cells. *Proc Natl Acad Sci U S A* 106: 16257-62
Schaefer M, Lyko F (2010) Solving the Dnmt2 enigma. *Chromosoma* 119: 35-40
Schubeler D (2015) Function and information content of DNA methylation. *Nature* 517: 321-6
Schwartz BE, Hofer MD, Lemieux ME, Bauer DE, Cameron MJ, West NH, Agoston ES, Reynoird N, Khochbin S, Ince TA, Christie A, Janeway KA, Vargas SO, Perez-Atayde AR, Aster JC, Sallan SE, Kung AL, Bradner JE, French CA (2011) Differentiation of NUT midline carcinoma by epigenomic reprogramming. *Cancer Res* 71: 2686-96
Seigneurin-Berny D, Verdel A, Curtet S, Lemercier C, Garin J, Rousseaux S, Khochbin S (2001) Identification of components of the murine histone deacetylase 6 complex: link between acetylation and ubiquitination signaling pathways. *Mol Cell Biol* 21: 8035-44
Shahbazian MD, Grunstein M (2007) Functions of site-specific histone acetylation and deacetylation. *Annu Rev Biochem* 76: 75-100
Shang E, Nickerson HD, Wen D, Wang X, Wolgemuth DJ (2007) The first bromodomain of Brdt, a testis-specific member of the BET sub-family of double-bromodomain-containing proteins, is essential for male germ cell differentiation. *Development* 134: 3507-15
Sharma U, Conine CC, Shea JM, Boskovic A, Derr AG, Bing XY, Belleannee C, Kucukural A, Serra RW, Sun F, Song L, Carone BR, Ricci EP, Li XZ, Fauquier L, Moore MJ, Sullivan

R, Mello CC, Garber M, Rando OJ (2016) Biogenesis and function of tRNA fragments during sperm maturation and fertilization in mammals. *Science* 351: 391-396
Sharma U, Sun F, Conine CC, Reichholf B, Kukreja S, Herzog VA, Ameres SL, Rando OJ (2018) Small RNAs Are Trafficked from the Epididymis to Developing Mammalian Sperm. *Dev Cell* 46: 481-494 e6
Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, Casero RA, Shi Y (2004) Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* 119: 941-53
Shibahara K, Stillman B (1999) Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin. *Cell* 96: 575-85
Shinagawa T, Huynh LM, Takagi T, Tsukamoto D, Tomaru C, Kwak HG, Dohmae N, Noguchi J, Ishii S (2015) Disruption of Th2a and Th2b genes causes defects in spermatogenesis. *Development* 142: 1287-92
Shiota H, Barral S, Buchou T, Tan M, Coute Y, Charbonnier G, Reynoird N, Boussouar F, Gerard M, Zhu M, Bargier L, Puthier D, Chuffart F, Bourova-Flin E, Picaud S, Filippakopoulos P, Goudarzi A, Ibrahim Z, Panne D, Rousseaux S et al. (2018) Nut Directs p300-Dependent, Genome-Wide H4 Hyperacetylation in Male Germ Cells. *Cell Rep* 24: 3477-3487 e6
Shires A, Carpenter MP, Chalkley R (1975) New histones found in mature mammalian testes. *Proc Natl Acad Sci U S A* 72: 2714-8
Shirley CR, Hayashi S, Mounsey S, Yanagimachi R, Meistrich ML (2004) Abnormalities and reduced reproductive potential of sperm from Tnp1- and Tnp2-null double mutant mice. *Biol Reprod* 71: 1220-9
Siegel TN, Hekstra DR, Kemp LE, Figueiredo LM, Lowell JE, Fenyo D, Wang X, Dewell S, Cross GA (2009) Four histone variants mark the boundaries of polycistronic transcription units in Trypanosoma brucei. *Genes Dev* 23: 1063-76
Simpson RT (1978) Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones. *Biochemistry* 17: 5524-5531
Sin HS, Barski A, Zhang F, Kartashov AV, Nussenzweig A, Chen J, Andreassen PR, Namekawa SH (2012) RNF8 regulates active epigenetic modifications and escape gene activation from inactive sex chromosomes in post-meiotic spermatids. *Genes Dev* 26: 2737-48
Smallwood SA, Kelsey G (2012) De novo DNA methylation: a germ cell perspective. *Trends Genet* 28: 33-42
Soboleva TA, Nekrasov M, Pahwa A, Williams R, Huttley GA, Tremethick DJ (2011) A unique H2A histone variant occupies the transcriptional start site of active genes. *Nat Struct Mol Biol* 19: 25-30
Soboleva TA, Parker BJ, Nekrasov M, Hart-Smith G, Tay YJ, Tng WQ, Wilkins M, Ryan D, Tremethick DJ (2017) A new link between transcriptional initiation and pre-mRNA splicing: The RNA binding histone variant H2A.B. *PLoS Genet* 13: e1006633
Spiller F, Medina-Pritchard B, Abad MA, Wear MA, Molina O, Earnshaw WC, Jeyaprakash AA (2017) Molecular basis for Cdk1-regulated timing of Mis18 complex assembly and CENP-A deposition. *EMBO Rep* 18: 894-905
Stankovic A, Guo LY, Mata JF, Bodor DL, Cao XJ, Bailey AO, Shabanowitz J, Hunt DF, Garcia BA, Black BE, Jansen LET (2017) A Dual Inhibitory Mechanism Sufficient to Maintain Cell-Cycle-Restricted CENP-A Assembly. *Mol Cell* 65: 231-246
Stark GR, Darnell JE, Jr. (2012) The JAK-STAT pathway at twenty. *Immunity* 36: 503-14
Strahl BD, Allis CD (2000) The language of covalent histone modifications. *Nature* 403: 41
Stricker SH, Koferle A, Beck S (2017) From profiles to function in epigenomics. *Nat Rev Genet* 18: 51-66
Suh N, Blelloch R (2011) Small RNAs in early mammalian development: from gametes to gastrulation. *Development* 138: 1653-61

Surani MA (2001) Reprogramming of genome function through epigenetic inheritance. *Nature* 414: 122-8
Suto RK, Clarkson MJ, Tremethick DJ, Luger K (2000) Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. *Nat Struct Biol* 7: 1121-4
Syed SH, Boulard M, Shukla MS, Gautier T, Travers A, Bednar J, Faivre-Moskalenko C, Dimitrov S, Angelov D (2009) The incorporation of the novel histone variant H2AL2 confers unusual structural and functional properties of the nucleosome. *Nucleic Acids Res* 37: 4684-95
Tachiwana H, Kagawa W, Osakabe A, Kawaguchi K, Shiga T, Hayashi-Takanaka Y, Kimura H, Kurumizaka H (2010) Structural basis of instability of the nucleosome containing a testis-specific histone variant, human H3T. *Proc Natl Acad Sci U S A* 107: 10454-9
Tachiwana H, Kagawa W, Shiga T, Osakabe A, Miya Y, Saito K, Hayashi-Takanaka Y, Oda T, Sato M, Park SY, Kimura H, Kurumizaka H (2011) Crystal structure of the human centromeric nucleosome containing CENP-A. *Nature* 476: 232-5
Tachiwana H, Osakabe A, Kimura H, Kurumizaka H (2008) Nucleosome formation with the testis-specific histone H3 variant, H3t, by human nucleosome assembly proteins in vitro. *Nucleic Acids Res* 36: 2208-18
Taddei A, Hediger F, Neumann FR, Gasser SM (2004) The function of nuclear architecture: a genetic approach. *Annu Rev Genet* 38: 305-45
Tagami H, Ray-Gallet D, Almouzni G, Nakatani Y (2004) Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* 116: 51-61
Taguchi H, Horikoshi N, Arimura Y, Kurumizaka H (2014) A method for evaluating nucleosome stability with a protein-binding fluorescent dye. *Methods* 70: 119-26
Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, Rao A (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324: 930-5
Talbert PB, Ahmad K, Almouzni G, Ausio J, Berger F, Bhalla PL, Bonner WM, Cande WZ, Chadwick BP, Chan SW, Cross GA, Cui L, Dimitrov SI, Doenecke D, Eirin-Lopez JM, Gorovsky MA, Hake SB, Hamkalo BA, Holec S, Jacobsen SE et al. (2012) A unified phylogeny-based nomenclature for histone variants. *Epigenetics Chromatin* 5: 7
Talbert PB, Henikoff S (2010) Histone variants--ancient wrap artists of the epigenome. *Nat Rev Mol Cell Biol* 11: 264-75
Talbert PB, Henikoff S (2017) Histone variants on the move: substrates for chromatin dynamics. *Nat Rev Mol Cell Biol* 18: 115-126
Tan M, Luo H, Lee S, Jin F, Yang JS, Montellier E, Buchou T, Cheng Z, Rousseaux S, Rajagopal N, Lu Z, Ye Z, Zhu Q, Wysocka J, Ye Y, Khochbin S, Ren B, Zhao Y (2011) Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* 146: 1016-28
Tanaka H, Iguchi N, Isotani A, Kitamura K, Toyama Y, Matsuoka Y, Onishi M, Masai K, Maekawa M, Toshimori K, Okabe M, Nishimune Y (2005) HANP1/H1T2, a novel histone H1-like protein involved in nuclear formation and sperm fertility. *Mol Cell Biol* 25: 7107-19
Tang MC, Jacobs SA, Mattiske DM, Soh YM, Graham AN, Tran A, Lim SL, Hudson DF, Kalitsis P, O'Bryan MK, Wong LH, Mann JR (2015) Contribution of the two genes encoding histone variant h3.3 to viability and fertility in mice. *PLoS Genet* 11: e1004964
Tang Y, Poustovoitov MV, Zhao K, Garfinkel M, Canutescu A, Dunbrack R, Adams PD, Marmorstein R (2006) Structure of a human ASF1a-HIRA complex and insights into specificity of histone chaperone complex assembly. *Nat Struct Mol Biol* 13: 921-9

Tettey TT, Gao X, Shao W, Li H, Story BA, Chitsazan AD, Glaser RL, Goode ZH, Seidel CW, Conaway RC, Zeitlinger J, Blanchette M, Conaway JW (2019) A Role for FACT in RNA Polymerase II Promoter-Proximal Pausing. *Cell Rep* 27: 3770-3779 e7
Thatcher TH, Gorovsky MA (1994) Phylogenetic analysis of the core histones H2A, H2B, H3, and H4. *Nucleic acids research* 22: 174-179
Thompson PR, Wang D, Wang L, Fulco M, Pediconi N, Zhang D, An W, Ge Q, Roeder RG, Wong J, Levrero M, Sartorelli V, Cotter RJ, Cole PA (2004) Regulation of the p300 HAT domain via a novel activation loop. *Nat Struct Mol Biol* 11: 308-15
Tolstorukov MY, Goldman JA, Gilbert C, Ogryzko V, Kingston RE, Park PJ (2012) Histone variant H2A.Bbd is associated with active transcription and mRNA processing in human cells. *Mol Cell* 47: 596-607
Trojer P, Reinberg D (2007) Facultative heterochromatin: is there a distinctive molecular signature? *Mol Cell* 28: 1-13
Trostle-Weige PK, Meistrich ML, Brock WA, Nishioka K (1984) Isolation and characterization of TH3, a germ cell-specific variant of histone 3 in rat testis. *J Biol Chem* 259: 8769-76
Tsukada Y, Fang J, Erdjument-Bromage H, Warren ME, Borchers CH, Tempst P, Zhang Y (2006) Histone demethylation by a family of JmjC domain-containing proteins. *Nature* 439: 811-6
Turner AL, Watson M, Wilkins OG, Cato L, Travers A, Thomas JO, Stott K (2018) Highly disordered histone H1− DNA model complexes and their condensates. *Proceedings of the National Academy of Sciences* 115: 11964-11969
Tyler JK, Adams CR, Chen SR, Kobayashi R, Kamakaka RT, Kadonaga JT (1999) The RCAF complex mediates chromatin assembly during DNA replication and repair. *Nature* 402: 555-60
Ueda J, Harada A, Urahama T, Machida S, Maehara K, Hada M, Makino Y, Nogami J, Horikoshi N, Osakabe A, Taguchi H, Tanaka H, Tachiwana H, Yao T, Yamada M, Iwamoto T, Isotani A, Ikawa M, Tachibana T, Okada Y et al. (2017) Testis-Specific Histone Variant H3t Gene Is Essential for Entry into Spermatogenesis. *Cell Rep* 18: 593-600
van der Heijden GW, Derijck AA, Posfai E, Giele M, Pelczar P, Ramos L, Wansink DG, van der Vlag J, Peters AH, de Boer P (2007) Chromosome-wide nucleosome replacement and H3.3 incorporation during mammalian meiotic sex chromosome inactivation. *Nat Genet* 39: 251-8
van der Heijden GW, Ramos L, Baart EB, van den Berg IM, Derijck AA, van der Vlag J, Martini E, de Boer P (2008) Sperm-derived histones contribute to zygotic chromatin in humans. *BMC Dev Biol* 8: 34
Van Holde K, Sahasrabuddhe CG, Shaw BR, Van Bruggen E, Arnberg AC (1974) Electron microscopy of chromatin subunit particles. *Biochemical and biophysical research communications* 60: 1365-1370
Varga-Weisz PD, Wilm M, Bonte E, Dumas K, Mann M, Becker PB (1997) Chromatin-remodelling factor CHRAC contains the ATPases ISWI and topoisomerase II. *Nature* 388: 598-602
Velazquez Camacho O, Galan C, Swist-Rosowska K, Ching R, Gamalinda M, Karabiber F, De La Rosa-Velazquez I, Engist B, Koschorz B, Shukeir N, Onishi-Seebacher M, van de Nobelen S, Jenuwein T (2017) Major satellite repeat RNA stabilize heterochromatin retention of Suv39h enzymes by RNA-nucleosome association and RNA:DNA hybrid formation. *Elife* 6
Verdel A, Curtet S, Brocard MP, Rousseaux S, Lemercier C, Yoshida M, Khochbin S (2000) Active maintenance of mHDA2/mHDAC6 histone-deacetylase in the cytoplasm. *Curr Biol* 10: 747-9

Verdel A, Jia S, Gerber S, Sugiyama T, Gygi S, Grewal SI, Moazed D (2004) RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* 303: 672-6
Verdel A, Khochbin S (1999) Identification of a new family of higher eukaryotic histone deacetylases. Coordinate expression of differentiation-dependent chromatin modifiers. *J Biol Chem* 274: 2440-5
Wang J, Liu X, Dou Z, Chen L, Jiang H, Fu C, Fu G, Liu D, Zhang J, Zhu T, Fang J, Zang J, Cheng J, Teng M, Ding X, Yao X (2014) Mitotic regulator Mis18beta interacts with and specifies the centromeric assembly of molecular chaperone holliday junction recognition protein (HJURP). *J Biol Chem* 289: 8326-36
Wang T, Chuffart F, Bourova-Flin E, Wang J, Mi J, Rousseaux S, Khochbin S (2019) Histone variants: critical determinants in tumour heterogeneity. *Front Med* 13: 289-297
Wang Y, Wang H, Zhang Y, Du Z, Si W, Fan S, Qin D, Wang M, Duan Y, Li L, Jiao Y, Li Y, Wang Q, Shi Q, Wu X, Xie W (2019) Reprogramming of Meiotic Chromatin Architecture during Spermatogenesis. *Mol Cell* 73: 547-561 e6
Wang YL, Faiola F, Xu M, Pan S, Martinez E (2008) Human ATAC Is a GCN5/PCAF-containing acetylase complex with a novel NC2-like histone fold module that interacts with the TATA-binding protein. *J Biol Chem* 283: 33808-15
Wang Z, Zang C, Cui K, Schones DE, Barski A, Peng W, Zhao K (2009) Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. *Cell* 138: 1019-31
Watson JD, Crick F (1953) A structure for deoxyribose nucleic acid.
Weinert BT, Narita T, Satpathy S, Srinivasan B, Hansen BK, Scholz C, Hamilton WB, Zucconi BE, Wang WW, Liu WR, Brickman JM, Kesicki EA, Lai A, Bromberg KD, Cole PA, Choudhary C (2018) Time-Resolved Analysis Reveals Rapid Dynamics and Broad Scope of the CBP/p300 Acetylome. *Cell* 174: 231-244 e12
Weyemi U, Redon CE, Choudhuri R, Aziz T, Maeda D, Boufraqech M, Parekh PR, Sethi TK, Kasoji M, Abrams N, Merchant A, Rajapakse VN, Bonner WM (2016) The histone variant H2A.X is a regulator of the epithelial-mesenchymal transition. *Nat Commun* 7: 10711
Williams JS, Hayashi T, Yanagida M, Russell P (2009) Fission yeast Scm3 mediates stable assembly of Cnp1/CENP-A into centromeric chromatin. *Mol Cell* 33: 287-98
Wilson EB (1900) *The cell in development and inheritance*. The Macmillan Company,
Wiśniewski JR, Zougman A, Krüger S, Mann M (2007) Mass spectrometric mapping of linker histone H1 variants reveals multiple acetylations, methylations, and phosphorylation as well as differences between cell culture and tissue. *Molecular & Cellular Proteomics* 6: 72-87
Witt O, Albig W, Doenecke D (1996) Testis-specific expression of a novel human H3 histone gene. *Exp Cell Res* 229: 301-6
Wolfe SA, Grimes SR (2003) Transcriptional repression of the testis-specific histone H1t gene mediated by an element upstream of the H1/AC box. *Gene* 308: 129-38
Woodcock C, Safer J, Stanchfield J (1976) Structural repeating units in chromatin: I. Evidence for their general occurrence. *Experimental cell research* 97: 101-110
Wu F, Caron C, De Robertis C, Khochbin S, Rousseaux S (2008) Testis-specific histone variants H2AL1/2 rapidly disappear from paternal heterochromatin after fertilization. *J Reprod Dev* 54: 413-7
Wu H, Zhang Y (2014) Reversing DNA methylation: mechanisms, genomics, and biological functions. *Cell* 156: 45-68
Wu SY, Chiang CM (2007) The double bromodomain-containing chromatin adaptor Brd4 and transcriptional regulation. *J Biol Chem* 282: 13141-5
Wu X, Zhang Y (2017) TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat Rev Genet* 18: 517-534

Wyrobek AJ, Meistrich ML, Furrer R, Bruce WR (1976) Physical characteristics of mouse sperm nuclei. *Biophys J* 16: 811-25
Wysocka J, Swigut T, Milne TA, Dou Y, Zhang X, Burlingame AL, Roeder RG, Brivanlou AH, Allis CD (2005) WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* 121: 859-72
Wysocka J, Swigut T, Xiao H, Milne TA, Kwon SY, Landry J, Kauer M, Tackett AJ, Chait BT, Badenhorst P, Wu C, Allis CD (2006) A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* 442: 86-90
Xiao A, Li H, Shechter D, Ahn SH, Fabrizio LA, Erdjument-Bromage H, Ishibe-Murakami S, Wang B, Tempst P, Hofmann K, Patel DJ, Elledge SJ, Allis CD (2009) WSTF regulates the H2A.X DNA damage response via a novel tyrosine kinase activity. *Nature* 457: 57-62
Xie Z, Zhang D, Chung D, Tang Z, Huang H, Dai L, Qi S, Li J, Colak G, Chen Y, Xia C, Peng C, Ruan H, Kirkey M, Wang D, Jensen LM, Kwon OK, Lee S, Pletcher SD, Tan M et al. (2016) Metabolic Regulation of Gene Expression by Histone Lysine beta-Hydroxybutyrylation. *Mol Cell* 62: 194-206
Yamaguchi K, Hada M, Fukuda Y, Inoue E, Makino Y, Katou Y, Shirahige K, Okada Y (2018) Re-evaluating the Localization of Sperm-Retained Histones Revealed the Modification-Dependent Accumulation in Specific Genome Regions. *Cell Rep* 23: 3920-3932
Yan D, Dong Xda E, Chen X, Wang L, Lu C, Wang J, Qu J, Tu L (2009) MicroRNA-1/206 targets c-Met and inhibits rhabdomyosarcoma development. *J Biol Chem* 284: 29596-604
Yan W, Ma L, Burns KH, Matzuk MM (2003) HILS1 is a spermatid-specific linker histone H1-like protein implicated in chromatin remodeling during mammalian spermiogenesis. *Proc Natl Acad Sci U S A* 100: 10546-51
Yang XJ, Seto E (2007) HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention. *Oncogene* 26: 5310-8
Yelagandula R, Stroud H, Holec S, Zhou K, Feng S, Zhong X, Muthurajan UM, Nie X, Kawashima T, Groth M, Luger K, Jacobsen SE, Berger F (2014) The histone variant H2A.W defines heterochromatin and promotes chromatin condensation in Arabidopsis. *Cell* 158: 98-109
Yoshida K, Muratani M, Araki H, Miura F, Suzuki T, Dohmae N, Katou Y, Shirahige K, Ito T, Ishii S (2018) Mapping of histone-binding sites in histone replacement-completed spermatozoa. *Nat Commun* 9: 3885
Yuan H, Marmorstein R (2013) Histone acetyltransferases: Rising ancient counterparts to protein kinases. *Biopolymers* 99: 98-111
Yuen BT, Bush KM, Barrilleaux BL, Cotterman R, Knoepfler PS (2014) Histone H3.3 regulates dynamic chromatin states during spermatogenesis. *Development* 141: 3483-94
Zalensky AO, Siino JS, Gineitis AA, Zalenskaya IA, Tomilin NV, Yau P, Bradbury EM (2002) Human testis/sperm-specific histone H2B (hTSH2B). Molecular cloning and characterization. *J Biol Chem* 277: 43474-80
Zaratiegui M, Irvine DV, Martienssen RA (2007) Noncoding RNAs and gene silencing. *Cell* 128: 763-76
Zentner GE, Henikoff S (2013) Regulation of nucleosome dynamics by histone modifications. *Nat Struct Mol Biol* 20: 259-66
Zhang JJ, Vinkemeier U, Gu W, Chakravarti D, Horvath CM, Darnell JE, Jr. (1996) Two contact regions between Stat1 and CBP/p300 in interferon gamma signaling. *Proc Natl Acad Sci U S A* 93: 15092-6
Zhang K, Gao Y, Li J, Burgess R, Han J, Liang H, Zhang Z, Liu Y (2016) A DNA binding winged helix domain in CAF-1 functions with PCNA to stabilize CAF-1 at replication forks. *Nucleic Acids Res* 44: 5083-94

Zhang R, Poustovoitov MV, Ye X, Santos HA, Chen W, Daganzo SM, Erzberger JP, Serebriiskii IG, Canutescu AA, Dunbrack RL, Pehrson JR, Berger JM, Kaufman PD, Adams PD (2005) Formation of MacroH2A-containing senescence-associated heterochromatin foci and senescence driven by ASF1a and HIRA. *Dev Cell* 8: 19-30

Zhang Y, Li N, Caron C, Matthias G, Hess D, Khochbin S, Matthias P (2003) HDAC-6 interacts with and deacetylates tubulin and microtubules in vivo. *EMBO J* 22: 1168-79

Zhang Z, Shibahara K, Stillman B (2000) PCNA connects DNA replication to epigenetic inheritance in yeast. *Nature* 408: 221-5

Zhao M, Shirley CR, Hayashi S, Marcon L, Mohapatra B, Suganuma R, Behringer RR, Boissonneault G, Yanagimachi R, Meistrich ML (2004a) Transition nuclear proteins are required for normal chromatin condensation and functional sperm development. *Genesis* 38: 200-13

Zhao M, Shirley CR, Mounsey S, Meistrich ML (2004b) Nucleoprotein transitions during spermiogenesis in mice with transition nuclear protein Tnp1 and Tnp2 mutations. *Biol Reprod* 71: 1016-25

Zhou J, Fan JY, Rangasamy D, Tremethick DJ (2007) The nucleosome surface regulates chromatin compaction and couples it with transcriptional repression. *Nat Struct Mol Biol* 14: 1070-6

Zhou W, De Iuliis GN, Dun MD, Nixon B (2018) Characteristics of the Epididymal Luminal Environment Responsible for Sperm Maturation and Storage. *Front Endocrinol (Lausanne)* 9: 59

ANNEXES

J. Biochem. 2018;163(2):97–103 doi:10.1093/jb/mvx079

JB THE JOURNAL OF BIOCHEMISTRY

JB Special Review—Chromatin Structure and Function: Biological Implications in Epigenetics

Histone variants: essential actors in male genome programming

Received September 6, 2017; accepted October 13, 2017; published online November 20, 2017

Naghmeh Hoghoughi, Sophie Barral, Alexandra Vargas, Sophie Rousseaux and Saadi Khochbin*

CNRS UMR 5309, Inserm, U1209, Université Grenoble Alpes, Institut Albert Bonniot, Grenoble F-38700, France

*Saadi Khochbin, Chromatin and Cell Signaling, Institute for Advanced Bioscience, Domaine de la Merci, La Tronche 38706, France. Tel: +33 4 76 54 95 83, Fax: 33 4 76 54 95 95, email: saadi.khochbin@univ-grenoble-alpes.fr

Prior to its transmission to the offspring, the male genome has to be tightly compacted. A genome-scale histone eviction and the subsequent repackaging of DNA by protamines (Prms) direct this essential genome condensation step. The requirement for male germ cells to undergo such a dramatic and unique genome reorganization explains why these cells express the largest number of histone variants, including many testis-specific ones. Indeed, an open chromatin, nucleosome instability and a facilitated process of histone disassembly are direct consequences of the presence of these histone variants in the chromatin of male germ cells. These histone-induced changes in chromatin first control a stage-specific gene expression program and then directly mediate the histone-to-Prm transition process. This review aims at summarizing and discussing a series of recent functional studies of male germ cell histone variants with a focus on their impact on the process of histone eviction and male genome compaction.

Keywords: H2A.L.2; H3.3; protamines; TH2B; transition proteins.

Spermatogenesis is a differentiation process which produces mature spermatozoa, which will deliver the male genome to the oocytes in the female organism. Spermatogenic cells are continuously produced by an asymmetrical division of progenitor cells, named spermatogonia. These cells on the one hand maintain the stem cell population and on the other hand generate meiotic cells, or spermatocytes, which themselves undergo two successive meiotic divisions. The second meiotic division produces haploid cells known as spermatids. The early post-meiotic spermatids are called round spermatids. These cells activate a specific gene expression program that will direct their transformation into spermatozoa, through a series of reorganizations of their various cellular constituents. These reorganizations mostly take place in the subsequent stages of post-meiotic differentiation, in cells known as elongating and condensing spermatids. Large-scale histone post-translational modifications (PTMs), essentially histone acetylation, signal and regulate the initiation of the process of histone-to-protamine (Prm) replacement in elongating spermatids (*1*–*3*).

Although still largely unexplored, this genome-wide replacement of histones by Prms is the most dramatic of the known processes involved in chromatin remodelling, genome reorganization and reprogramming (*4*). In addition to Prms, early investigations also identified other basic non-histone proteins that appear at the time of histone replacement but, in contrast to Prms, totally disappear afterwards (*5*). These proteins have been named transition proteins (TPs). The early studies of sperm and testis histones in various species also revealed the existence of histones that are only expressed in male germ cells (*6*). Indeed testis-specific histones were among the first histone variants identified. Additionally, spermatogenic cells not only express a large number of various testis-specific histones, but are also characterized by the presence of some histone variants, which almost entirely replace their canonical counterparts (Fig. 1). Therefore, prior to histone replacement, spermatogenic cell chromatin accumulates unique germline histones in replacement of their canonical counterparts. Histone variants, including the testis-specific ones, will eventually be replaced by Prms.

An open question is whether and how this large number of histone variants, individually and collectively, are involved in directing the spermatogenic gene expression program and in preparing chromatin for the final histone-to-Prm transition. Additionally, it is critical to decipher the individual contribution of each of the specific or generalist histone variants to these unique events. Another related question is the interplay between histones, TPs and Prms in the process of histone eviction. Most of these questions have remained obscure until recently. It is only during the last few years, a long time after the discovery of the first testis-specific histone variants that several pilot studies have started to shed light on the role of histone variants in the stepwise programing of the male genome, in the eviction of histones and in the final genome compaction.

Here, we summarize these recent advances in our knowledge of the role of histone variants, specifically testis-specific histone variants, in spermatogenesis, with a specific focus on the process of genome-wide histone eviction and Prm assembly.

TH2B

One of the first histone variants discovered belongs to the H2B family and was named TH2B (TSH2B, according to the new nomenclature). Indeed, the protein was

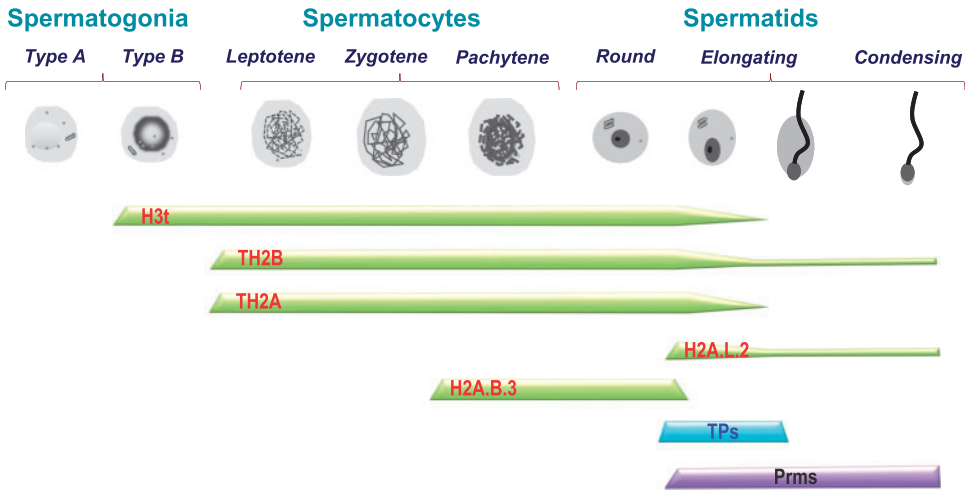


Fig. 1 Differentiation-dependent expression of testis-specific histone variants during spermatogenesis. Detailed studies of the indicated histone variants established the precise timing of their expression during the spermatogenic cells differentiation. A fraction of TH2B and H2A.L.2 escape the general histone-to-Prm transition and remains in the mature sperm nuclei.

reported in 1975 (*7*) as a H2B histone expressed in the testes of several mammalian species.

The first insight into a potential specific function of this H2B variant came from a structural study showing that a histone octamer containing TH2B presents an increased instability compared to a canonical octamer (*8*). Additional studies showed that TH2B is also present during late spermatogenesis and at the time of histone removal, and that it remains in the transitional structures that are generated during the transformation of nucleosomes into nucleoprotamines (*9*).

The first functional studies of TH2B in its physiological setting, during spermatogenic cell differentiation, were published only 38 years after its identification, following the generation of different mouse models (*10*). These studies show that, from the onset of meiosis, TH2B massively replaces H2B, and that most of the genome becomes covered by TH2B-containing nucleosomes. However, surprisingly, in the total absence of TH2B, spermatogenesis seems to occur completely normally, and there is no effect on male fertility.

The investigation of spermatogenesis in TH2B-less mice revealed that, in fact, in the absence of TH2B, a compensatory mechanism leads to the accumulation of H2B, thereby physically compensating the lack of TH2B by an increased production of H2B and allowing the formation of nucleosomes at the genome scale.

This observation raised the question of the functional utility of TH2B during spermatogenesis. Indeed, if H2B can efficiently replace TH2B, then what could be the specific role of TH2B?

The answer to this question came from the comparative quantitative proteomic analysis of histone PTMs between TH2B-less and wild-type spermatogenic cells. These data show that, in the absence of TH2B, nucleosomes undergo an epigenetic reprogramming to become less stable. Indeed, in TH2B-less nucleosomes, histones gain several chemical modifications that, without any exception, are all located at structurally strategic positions and present the potential to weaken the overall stability of the nucleosomes. These are H3K122 and H4K77 crotonylation. These two lysines are in contact with DNA and the addition of a bulky crotonyl group with its charge neutralizing capacity is able to affect nucleosome stability. The other PTMs found in TH2B-less nucleosomes are methylations at H4R35, H4R55, H4R67 and H2BR72. Remarkably all these residues are involved in histone–DNA and histone–histone interactions and their methylation has the potential to weaken these intranucleosomal interactions.

Hence, these investigations indicate that the compensatory mechanism that is activated after the loss of TH2B encompasses not only an over-expression of the canonical H2B, but also specific modifications of histones with nucleosomal destabilizing effects. In other words, the primary function of TH2B could be to generate less stable nucleosomes at the whole genome scale, which would lower the energy necessary for the cells to undertake chromosome-wide (sex chromosome inactivation) and genome-wide histone exchanges in later stages.

The nucleosome destabilizing property of TH2B was later confirmed through a series of structural studies. Compared to H2B, the presence of TH2B in a nucleosome leads to the loss of a hydrogen bound with H4 (*11*). Other structural works considering TH2B and TH2A (an H2A variant co-expressed with TH2B) showed even more dramatic effects on the nucleosomes stability, with significantly reduced histone-DNA contacts and modified histone–histone interactions (*12*, *13*).

TH2B was also found expressed in oocytes and to be assembled again on the paternal genome after fertilization before its total disappearance during later development (Fig. 2) (*10*, *14*, *15*). The use of TH2B–TH2A in cell reprograming showed that their ectopic expression stimulates the reprograming of somatic cells into

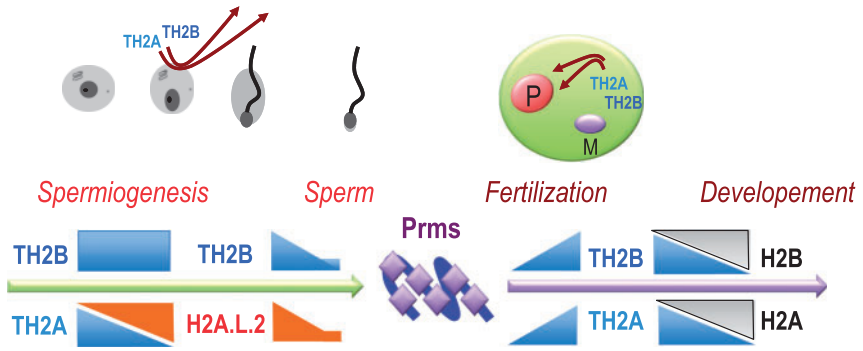


Fig. 2 The packaging cycles of the male genome with TH2A/TH2B. The histone variants TH2A and TH2B become the major histones in spermatogenic cells, until the replacement of TH2A by H2A.L.2, at the time of histone-to-Prm transition. A fraction of TH2B–H2A.L.2 dimer survives histone replacement. After fertilization, the removal of Prms is associated with the re-assembly of maternal TH2A–TH2B on the male genome. These histones will be entirely replaced by their somatic counterparts during early development. The schemes respectively represent maturing spermatids (left) and a fertilized egg with paternal (P) and maternal (M) genomes (right).

induced pluripotent stem cells, a property that could be related to the ability of these histones to create more accessible chromatin (*14*).

In contrast to TH2B-less situation, the knock-out of both TH2B and TH2A leads to an arrest of spermatogenesis during meiosis (*16*). However in this model, although the lack of TH2B induced a compensatory accumulation of H2B, as reported in the TH2B-less model (*10*), no H2A accumulation was observed (*16*). Since TH2A is also the main H2A in spermatogenic cells, in *Th2a*/*Th2b* KO spermatogenic cells, the absence of compensatory gain of H2A expression suggests a significant H2A under-dosage in these cells. The defects observed could therefore be due to an inappropriate chromatin structure and subsequent genome disorganization, rather than to a specific action of TH2A or TH2B (*16*). In agreement with this hypothesis, two recent reports showed that TH2A could be phosphorylated at the TH2A-specific Thr127 and that this phosphorylation is dispensable for normal spermatogenesis and male fertility (*17*, *18*).

The animal models used to decipher TH2B functions also included mice that expressed a modified endogenous TH2B, bearing three tags at its C-terminal (*10*). Remarkably, although in these mice all the investigated events at the cellular and molecular levels, i.e. TH2B-tag assembly, transcription, meiosis, sex-chromosome inactivation, H2A.X phosphorylation and histone exchange, etc., occur normally in spermatogenic cells up to the appearance of elongating spermatids, an arrest of spermatogenesis is observed at the time of histone-to-Prm transition. This study demonstrated that TH2B-tag expressed in female animals could also sustain oocyte maturation and embryonic development with no detectable defects and leading to the generation of healthy adult mice (*10*).

Overall, these data show that the presence of a tag at the C-terminal end of TH2B, interferes very specifically and exclusively with defined and particular events that take place in elongating-condensing spermatids (*10*). Detailed observations indicated, that in these TH2B-tag expressing male germ cells, although TP and Prms are produced, Prms are unable to replace histones leading to the accumulation of transitional structures following nucleosome disassembly. Further investigations are required to decipher the exact role of TH2B, and more specifically of its C-terminal region, in the final histone eviction.

H2A.L.2

In an attempt to identify histones that could facilitate the process of histone-to-Prm transition, surviving histones from spermatids undergoing histone replacement were purified and identified (*9*). This led to the identification of 5 new histone variants, two H2Bs and three H2As, which were respectively named H2B.L.1, H2B.L.2, H2A.L.1, H2A.L.2 and H2A.L.3 [please see their new nomenclature in (*19*, *20*)]. Among these histones, H2A.L.2 presented all the criteria to be considered as a candidate histone variant involved in histone removal. Indeed, the protein is first detected in spermatids at the same time as TPs (*9*, *21*). Additionally, H2A.L.2 along with TH2B was found in the H3/H4-less transitional structures, which were evidenced in condensing spermatids after extensive micrococcal nuclease (MNase) digestion. Finally, a fraction of H2A.L.2 remains in mature spermatozoa associated with the pericentric regions (*9*).

The generation of an H2A.L.2-less mouse model proved its critical role in the process of histone-to-Prm replacement. The molecular studies of spermatogenic cells lacking H2A.L.2 revealed its role in the incorporation of TPs onto chromatin. Additionally, a series of *in vitro* nucleosome reconstitution assays demonstrated that the incorporation of H2A.L.2 can drastically modify the structure of the corresponding nucleosome. A more 'open' H2A.L.2-containing nucleosome is able to load TPs, which invade the nucleosome without releasing the DNA. These TPs in turn buffer the incoming Prms and allow their assembly and the final histone eviction. TPs also mediate the processing of Prm2, which is produced as a pre-Prm2 protein (*21*). In fact the phenotype of H2A.L.2-less spermatogenic cells is very

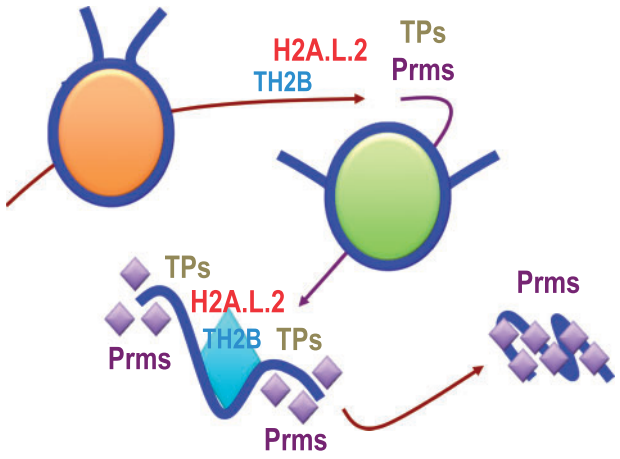


Fig. 3 Coordinate actions of histone variants, TPs and Prms in the process of histone-to-Prm replacement. The exchange of TH2B–TH2A by TH2B–H2A.L.2 opens the nucleosomes, allowing the loading of TPs that in turn direct a controlled replacement of histones, Prm assembly and genome compaction.

close to the phenotype of TP-less cells (*22–24*) suggesting that, in the absence of H2A.L.2, although TPs are produced, they remain non-functional.

Altogether, these observations fully support the idea that the incorporation of H2A.L.2 changes the structure of the nucleosomes allowing chromatin loading of TPs, which in turn buffer and regulate Prm processing and assembly (Fig. 3).

Prms are the true histone displacers. Indeed, they can displace histones even in the absence of TPs (or H2A.L.2). However, although not required for histone removal, TPs are necessary for an efficient Prm-dependent genome compaction.

The major outcome of these investigations is also a new vision of the process of histone-to-Prm transition. Indeed, the general belief was that histones were first replaced by TPs, which were later replaced by Prms to fully compact the genome. The new data described above demonstrate that the full picture of the process requires a histone variant that 'prepares' the nucleosomes to undergo disassembly (Fig. 1). In this context, TPs and Prms work together, not successively, to fully transform the nucleosomes into nucleosprotamine structures (Fig. 3).

H2A.B.3

In an attempt to identify and functionally characterize H2A histone variants that lack the characteristic acidic patch, which is present in most H2As, including canonical H2A, the mouse genome was screened in search for H2A genes with an altered acidic patch (*25*). The acidic patch is typically composed of a series of six acidic amino acids of H2A, present at the surface of a nucleosome, which mediates the nucleosome–nucleosome interaction and hence chromatin compaction (*26*, *27*). This approach led to the identification of four H2A variants with an altered acidic patch, which were named H2A.Lap1-4, 'Lap' standing for Lack of Acidic Patch. A further analysis showed that some of them had already been identified by Govin and colleagues: H2A.L.2 actually is H2A.Lap3, H2A.L.1 is H2A.Lap2 and H2A.L.3 is H2A.Lap4 (*20*). H2A.Lap1, now known as H2A.B.3, was uniquely identified in this study and functionally characterized. This H2A variant shows similar structural features as H2A.L.2 but its expression pattern is drastically different (Fig. 1). H2A.B.3 first appears in late pachytene spermatocytes and particularly accumulates in post-meiotic round spermatids but disappears from the elongating spermatids nucleus (*25*), when H2A.L.2 starts to accumulate (*21*). The protein was then found enriched at the transcriptional start sites (TSS) of highly active genes in spermatocytes as well as in round spermatids. In these latter cells, H2A.B.3 becomes associated with the X-linked genes that escape sex chromosome transcriptional inactivation (*25*). Further studies demonstrated that H2A.B.3 has an RNA-binding motif, binds RNA and a series of RNA processing factors and co-localizes with splicing speckles in highly active nuclear subdomains (*28*). These studies highlight the possibility that H2A.B.3 incorporation at exon–intron boundaries facilitates the recruitment of splicing factors from the splicing speckles.

It is of note that the studies on H2A.B.3 also revealed a specific functional interplay between this histone and the widely-studied H2A.Z in spermatogenic cells. Indeed, several years ago, a microscopic analysis of H2A.Z histone variants revealed a post-meiotic accumulation of H2A.Z on the sex chromosomes (*29*). A high resolution mapping of H2A.Z and H2A.B.3 revealed an interesting relationship between these two H2A variants. The replacement of H2A.Z by H2A.B.3 at the exon–intron boundaries in spermatogenic cells may play a role in the regulation of splicing at active genes. Additionally, in the surround of genes TSSs, a combination of H2A, H2A.Z and H2A.B.3 could define specific gene categories in terms of expression timing and functional classifications of their products (*30*). These studies extend the role of histone variants from a specific action on particular chromatin regions to a possibly direct regulation of RNA processing and splicing.

Testis specific H3

A testis-specific histone H3 was also first reported almost 40 years ago (*31*). However, similarly to TH2B, it is structural biology that brought hints on the possible functions of this variant, first for the human H3T (*32*) and then for the mouse H3t variant (*33*). These studies showed that the alteration of several residues in H3T compared to H3.1, perfectly explains the instability of nucleosomes bearing H3T, observed *in vitro* and *in vivo*. In the case of H3t, only one H3t-specific amino acid, H42, was found to be responsible for creating an 'open' nucleosome with flexible DNA at the entry–exit of the nucleosomes (*33*). The H3t-encoding gene becomes active in differentiating spermatogonia, and its product gradually replaces most of the somatic-type H3s.

In the absence of H3t, a significant reduction of the number of differentiating c-Kit positive spermatogonia was observed, leading to complete male infertility. Further investigations showed a drastic decrease in the number of post-meiotic cells and suggested a possible role for this histone in the entry of cells into meiosis. However, since there was no compensatory increase of H3 gene expression to maintain cellular H3 levels, the observed defects in the absence of H3t could also be explained by histone underdosage.

H3.3

Besides the mouse models harbouring the inactivation of genes encoding testis-specific histone variants, the targeted inactivation of H3.3-encoding genes in the mouse testis was also performed. H3.3 is a replication-independent H3 that is required for chromatin assembly whenever the replication-dependent pathway is not operational, such as post-fertilization chromatin assembly on the paternal genome (*34*–*36*).

In the mouse two independent H3.3 encoding genes, *H3f3a* and *H3f3b*, generate identical H3.3. However individual targeted KO of these genes have different impact on spermatogenesis. The knock-out of *H3f3a* leads to the generation of abnormal spermatozoa (*37*), while the KO of *H3f3b* has a much more dramatic phenotype including a spermatogenesis arrest at the stage of round spermatid (*37*, *38*).

These functional analyses of H3.3 had been awaited since the interesting observation by de Boer's laboratory of the massive replacement of H3 by H3.3 on the sex chromosomes in meiotic cells (*39*). During meiosis the unsynapsed X and Y chromosomes create a specific domain in the nucleus of spermatocytes known as the sex body, which undergoes a total transcriptional inactivation during meiosis known as meiotic sex chromosome inactivation (MSCI).

A consequence of this selective incorporation of H3.3 on the sex chromosomes is the creation of interesting combinations of histone variants. In spermatocytes, the macroH2A histone variant was known to accumulate on the sex body (*40*), while in post-meiotic cells it is H2A.Z that becomes enriched on these chromosomes (*29*). The presence of both H3.3 and macroH2A in spermatocytes and of H3.3 and H2A.Z in round spermatids increases the probability for these combinations to be present within the same nucleosome. While macroH2A is associated with less active and more stable chromatin domains (*41*), the H3.3–H2A.Z combination is known to create a particularly unstable nucleosome (*42*, *43*), which could be important for the post-meiotic reactivation of the sex-linked genes.

The generation of these H3.3 KO models therefore provided the possibility to investigate more specifically the importance of H3.3 on the sex chromosome biology during spermatogenesis. However, at least in spermatocytes, no global defects were observed after a reduction in H3.3 expression (*37*), and the question of the function of H3.3 in MSCI and in the post-meiotic partial reactivation of sex chromosomes at the molecular levels remains open.

Concluding Remarks

In addition to almost all the histone variants present in somatic cells, spermatogenic cells express a number of unique histone variants of the H2A, H2B and H3 types (Fig. 1).

The structural studies of all these testis-specific variants revealed their capacity to induce nucleosome instability. Remarkably, short H2As, H2A.L.2 and H2A.B.3 as well as H3t are able to generate flexible DNA ends at the entry and exit of nucleosomes. Although these variants induce open chromatin features, these histones are expressed at quite different timings during spermatogenesis. H3t is expressed early in differentiating spermatogonia, while H2A.B.3 is expressed in pachytene spermatocytes and to a greater extent in round spermatids and is eventually replaced by H2A.L.2 in elongating spermatids. Therefore, the configuration of chromatin seems to gradually open from the early stage of commitment of cells into meiotic divisions to the late post-meiotic stage of histone-to-Prm transition. Additionally, spermatocytes express TH2B and TH2A that also form unstable and programmable chromatin. TH2B–TH2A remains in spermatogenic cells until the time of histone replacement, when H2A.L.2 accumulates and becomes the privileged partner of TH2B (Fig. 2).

Therefore, due to the action of these histone variants, the chromatin of spermatogenic cells becomes potentially 'prepared' to undergo histone-to-Prm replacement. Consequently, it is possible to speculate that it is the timing of TPs and Prms expression that decides histones replacement. However, the generation of an H2A.L.2-less mouse model demonstrated that the incorporation of this variant is indispensable for TPs to be efficiently loaded onto chromatin and that TPs incorporation controls the efficiency of the Prm-dependent replacement of histones (*21*).

Additionally, a genome-wide occurrence of histone hyperacetylation and of other histone acylations also seems necessary for histone replacement (*1*–*3*). It can therefore be concluded that a change in the structure of nucleosomes due to the incorporation of testis-specific histone variants is only one facet of the mechanisms involved in general histone eviction and that other mechanisms, mostly histone PTM-based events, are equally important.

Therefore, the future challenge will be to decipher the molecular links between histone variants and histone PTMs in the process of coupled histone eviction and Prm assembly.

Funding

This work is supported by ANR Episperm 3 program, by a grant from 'Fondation pour la Recherche Medicale (FRM)' 'analyse bioinformatique pour la recherche en biologie' program as well as INCa libre program (RPT13001CCA). Additional supports were from: Fondation ARC 'Canc'air' project (RAC16042CLA), Plan Cancer (CH7-INS15B66) and Plan Cancer (ASC16012CSA) and the 'Université Grenoble Alpes' ANR-15-IDEX-02.

Conflict of Interest

A.V. is a recipient of a PhD fellowship from CNRS.

References

1. Goudarzi, A., Shiota, H., Rousseaux, S., and Khochbin, S. (2014) Genome-scale acetylation-dependent histone eviction during spermatogenesis. *J. Mol. Biol.* **426**, 3342–3349
2. Goudarzi, A., Zhang, D., Huang, H., Barral, S., Kwon, O.K., Qi, S., Tang, Z., Buchou, T., Vitte, A.L., He, T., Cheng, Z., Montellier, E., Gaucher, J., Curtet, S., Debernardi, A., Charbonnier, G., Puthier, D., Petosa, C., Panne, D., Rousseaux, S., Roeder, R.G., Zhao, Y., and Khochbin, S. (2016) Dynamic competing histone H4 K5K8 acetylation and butyrylation are hallmarks of highly active gene promoters. *Mol. Cell* **62**, 169–180
3. Liu, S., Yu, H., Liu, Y., Liu, X., Zhang, Y., Bu, C., Yuan, S., Chen, Z., Xie, G., Li, W., Xu, B., Yang, J., He, L., Jin, T., Xiong, Y., Sun, L., Liu, X., Han, C., Cheng, Z., Liang, J., and Shang, Y. (2017) Chromodomain protein CDYL acts as a crotonyl-CoA hydratase to regulate histone crotonylation and spermatogenesis. *Mol. Cell* **67**, 853–866
4. Bao, J., and Bedford, M.T. (2016) Epigenetic regulation of the histone-to-protamine transition during spermiogenesis. *Reproduction* **151**, R55–R70
5. Meistrich, M.L., Mohapatra, B., Shirley, C.R., and Zhao, M. (2003) Roles of transition nuclear proteins in spermiogenesis. *Chromosoma* **111**, 483–488
6. Govin, J., Caron, C., Lestrat, C., Rousseaux, S., and Khochbin, S. (2004) The role of histones in chromatin remodelling during mammalian spermiogenesis. *Eur. J. Biochem.* **271**, 3459–3469
7. Shires, A., Carpenter, M.P., and Chalkley, R. (1975) New histones found in mature mammalian testes. *Proc. Natl. Acad. Sci. U S A* **72**, 2714–2718
8. Li, A., Maffey, A.H., Abbott, W.D., Conde e Silva, N., Prunell, A., Siino, J., Churikov, D., Zalensky, A.O., and Ausió, J. (2005) Characterization of nucleosomes consisting of the human testis/sperm-specific histone H2B variant (hTSH2B). *Biochemistry* **44**, 2529–2535
9. Govin, J., Escoffier, E., Rousseaux, S., Kuhn, L., Ferro, M., Thevenon, J., Catena, R., Davidson, I., Garin, J., Khochbin, S., and Caron, C. (2007) Pericentric heterochromatin reprogramming by new histone variants during mouse spermiogenesis. *J. Cell Biol.* **176**, 283–294
10. Montellier, E., Boussouar, F., Rousseaux, S., Zhang, K., Buchou, T., Fenaille, F., Shiota, H., Debernardi, A., Hery, P., Curtet, S., Jamshidikia, M., Barral, S., Holota, H., Bergon, A., Lopez, F., Guardiola, P., Pernet, K., Imbert, J., Petosa, C., Tan, M., Zhao, Y., Gerard, M., and Khochbin, S. (2013) Chromatin-to-nucleoprotamine transition is controlled by the histone H2B variant TH2B. *Genes Dev.* **27**, 1680–1692
11. Urahama, T., Horikoshi, N., Osakabe, A., Tachiwana, H., and Kurumizaka, H. (2014) Structure of human nucleosome containing the testis-specific histone variant TSH2B. *Acta Crystallogr. F Struct. Biol. Commun.* **70**, 444–449
12. Padavattan, S., Shinagawa, T., Hasegawa, K., Kumasaka, T., Ishii, S., and Kumarevel, T. (2015) Structural and functional analyses of nucleosome complexes with mouse histone variants TH2a and TH2b, involved in reprogramming. *Biochem. Biophys. Res. Commun.* **464**, 929–935
13. Padavattan, S., Thiruselvam, V., Shinagawa, T., Hasegawa, K., Kumasaka, T., Ishii, S., and Kumarevel, T. (2017) Structural analyses of the nucleosome complexes with human testis-specific histone variants, hTh2a and hTh2b. *Biophys. Chem.* **221**, 41–48
14. Shinagawa, T., Takagi, T., Tsukamoto, D., Tomaru, C., Huynh, L.M., Sivaraman, P., Kumarevel, T., Inoue, K., Nakato, R., Katou, Y., Sado, T., Takahashi, S., Ogura, A., Shirahige, K., and Ishii, S. (2014) Histone variants enriched in oocytes enhance reprogramming to induced pluripotent stem cells. *Cell Stem Cell* **14**, 217–227
15. Iuso, D., Czernik, M., Toschi, P., Fidanza, A., Zacchini, F., Feil, R., Curtet, S., Buchou, T., Shiota, H., Khochbin, S., Ptak, G.E., and Loi, P. (2015) Exogenous expression of human protamine 1 (hPrm1) remodels fibroblast nuclei into spermatid-like structures. *Cell Rep.* **13**, 1765–1771
16. Shinagawa, T., Huynh, L.M., Takagi, T., Tsukamoto, D., Tomaru, C., Kwak, H.G., Dohmae, N., Noguchi, J., and Ishii, S. (2015) Disruption of Th2a and Th2b genes causes defects in spermatogenesis. *Development* **142**, 1287–1292
17. Hada, M., Kim, J., Inoue, E., Fukuda, Y., Tanaka, H., Watanabe, Y., and Okada, Y. (2017) TH2A is phosphorylated at meiotic centromere by Haspin. *Chromosoma* **126**, 769–780
18. Hada, M., Masuda, K., Yamaguchi, K., Shirahige, K., and Okada, Y. (2017) Identification of a variant-specific phosphorylation of TH2A during spermiogenesis. *Sci. Rep.* **7**, 46228
19. Talbert, P.B., Ahmad, K., Almouzni, G., Ausio, J., Berger, F., Bhalla, P.L., Bonner, W.M., Cande, W.Z., Chadwick, B.P., Chan, S.W., Cross, G.A., Cui, L., Dimitrov, S.I., Doenecke, D., Eirin-Lopez, J.M., Gorovsky, M.A., Hake, S.B., Hamkalo, B.A., Holec, S., Jacobsen, S.E., Kamieniarz, K., Khochbin, S., Ladurner, A.G., Landsman, D., Latham, J.A., Loppin, B., Malik, H.S., Marzluff, W.F., Pehrson, J.R., Postberg, J., Schneider, R., Singh, M.B., Smith, M.M., Thompson, E., Torres-Padilla, M.E., Tremethick, D.J., Turner, B.M., Waterborg, J.H., Wollmann, H., Yelagandula, R., Zhu, B., and Henikoff, S. (2012) A unified phylogeny-based nomenclature for histone variants. *Epigenetics Chromatin* **5**, 7
20. El Kennani, S., Adrait, A., Shaytan, A.K., Khochbin, S., Bruley, C., Panchenko, A.R., Landsman, D., Pflieger, D., and Govin, J. (2017) MS_HistoneDB, a manually curated resource for proteomic analysis of human and mouse histones. *Epigenetics & Chromatin* **10**, 2
21. Barral, S., Morozumi, Y., Tanaka, H., Montellier, E., Govin, J., de Dieuleveult, M., Charbonnier, G., Coute, Y., Puthier, D., Buchou, T., Boussouar, F., Urahama, T., Fenaille, F., Curtet, S., Hery, P., Fernandez-Nunez, N., Shiota, H., Gerard, M., Rousseaux, S., Kurumizaka, H., and Khochbin, S. (2017) Histone variant H2A.L.2 guides transition protein-dependent protamine assembly in male germ cells. *Mol. Cell* **66**, 89–101.e8
22. Zhao, M., Shirley, C.R., Hayashi, S., Marcon, L., Mohapatra, B., Suganuma, R., Behringer, R.R., Boissonneault, G., Yanagimachi, R., and Meistrich, M.L. (2004) Transition nuclear proteins are required for normal chromatin condensation and functional sperm development. *Genesis* **38**, 200–213
23. Zhao, M., Shirley, C.R., Mounsey, S., and Meistrich, M.L. (2004) Nucleoprotein transitions during spermiogenesis in mice with transition nuclear protein Tnp1 and Tnp2 mutations. *Biol. Reprod.* **71**, 1016–1025
24. Shirley, C.R., Hayashi, S., Mounsey, S., Yanagimachi, R., and Meistrich, M.L. (2004) Abnormalities and reduced reproductive potential of sperm from Tnp1- and Tnp2-null double mutant mice. *Biol. Reprod.* **71**, 1220–1229

25. Soboleva, T.A., Nekrasov, M., Pahwa, A., Williams, R., Huttley, G.A., and Tremethick, D.J. (2011) A unique H2A histone variant occupies the transcriptional start site of active genes. *Nat. Struct. Mol. Biol.* **19**, 25–30
26. Chodaparambil, J.V., Barbera, A.J., Lu, X., Kaye, K.M., Hansen, J.C., and Luger, K. (2007) A charged and contoured surface on the nucleosome regulates chromatin compaction. *Nat. Struct. Mol. Biol.* **14**, 1105–1107
27. Zhou, J., Fan, J.Y., Rangasamy, D., and Tremethick, D.J. (2007) The nucleosome surface regulates chromatin compaction and couples it with transcriptional repression. *Nat. Struct. Mol. Biol.* **14**, 1070–1076
28. Soboleva, T.A., Parker, B.J., Nekrasov, M., Hart-Smith, G., Tay, Y.J., Tng, W.-Q., Wilkins, M., Ryan, D., Tremethick, D.J., and Schneider, R. (2017) A new link between transcriptional initiation and pre-mRNA splicing: the RNA binding histone variant H2A.B. *PLoS Genet.* **13**, e1006633
29. Greaves, I.K., Rangasamy, D., Devoy, M., Marshall Graves, J.A., and Tremethick, D.J. (2006) The X and Y chromosomes assemble into H2A.Z-containing [corrected] facultative heterochromatin [corrected] following meiosis. *Mol. Cell. Biol.* **26**, 5394–5405
30. Soboleva, T.A., Nekrasov, M., Ryan, D.P., and Tremethick, D.J. (2014) Histone variants at the transcription start-site. *Trends Genet.* **30**, 199–209
31. Franklin, S.G., and Zweidler, A. (1977) Non-allelic variants of histones 2a, 2b and 3 in mammals. *Nature* **266**, 273–275
32. Tachiwana, H., Kagawa, W., Osakabe, A., Kawaguchi, K., Shiga, T., Hayashi-Takanaka, Y., Kimura, H., and Kurumizaka, H. (2010) Structural basis of instability of the nucleosome containing a testis-specific histone variant, human H3T. *Proc. Natl. Acad. Sci. U S A* **107**, 10454–10459
33. Ueda, J., Harada, A., Urahama, T., Machida, S., Maehara, K., Hada, M., Makino, Y., Nogami, J., Horikoshi, N., Osakabe, A., Taguchi, H., Tanaka, H., Tachiwana, H., Yao, T., Yamada, M., Iwamoto, T., Isotani, A., Ikawa, M., Tachibana, T., Okada, Y., Kimura, H., Ohkawa, Y., Kurumizaka, H., and Yamagata, K. (2017) Testis-specific histone variant H3t gene is essential for entry into spermatogenesis. *Cell Rep.* **18**, 593–600
34. Filipescu, D., Szenker, E., and Almouzni, G. (2013) Developmental roles of histone H3 variants and their chaperones. *Trends Genet.* **29**, 630–640
35. Lin, C.J., Koh, F.M., Wong, P., Conti, M., and Ramalho-Santos, M. (2014) Hira-mediated H3.3 incorporation is required for DNA replication and ribosomal RNA transcription in the mouse zygote. *Dev. Cell.* **30**, 268–279
36. Inoue, A., and Zhang, Y. (2014) Nucleosome assembly is required for nuclear pore complex assembly in mouse zygotes. *Nat. Struct. Mol. Biol.* **21**, 609–616
37. Tang, M.C.W., Jacobs, S.A., Mattiske, D.M., Soh, Y.M., Graham, A.N., Tran, A., Lim, S.L., Hudson, D.F., Kalitsis, P., O'Bryan, M.K., Wong, L.H., Mann, J.R., and Tremethick, D. (2015) Contribution of the two genes encoding histone variant h3.3 to viability and fertility in mice. *PLoS Genet.* **11**, e1004964
38. Yuen, B.T., Bush, K.M., Barrilleaux, B.L., Cotterman, R., and Knoepfler, P.S. (2014) Histone H3.3 regulates dynamic chromatin states during spermatogenesis. *Development* **141**, 3483–3494
39. van der Heijden, G.W., Derijck, A.A., Posfai, E., Giele, M., Pelczar, P., Ramos, L., Wansink, D.G., van der Vlag, J., Peters, A.H., and de Boer, P. (2007) Chromosome-wide nucleosome replacement and H3.3 incorporation during mammalian meiotic sex chromosome inactivation. *Nat. Genet.* **39**, 251–258
40. Hoyer-Fender, S., Costanzi, C., and Pehrson, J.R. (2000) Histone macroH2A1.2 is concentrated in the XY-body by the early pachytene stage of spermatogenesis. *Exp Cell Res.* **258**, 254–260
41. Talbert, P.B., and Henikoff, S. (2017) Histone variants on the move: substrates for chromatin dynamics. *Nat. Rev. Mol. Cell Biol.* **18**, 115–126
42. Jin, C., and Felsenfeld, G. (2007) Nucleosome stability mediated by histone variants H3.3 and H2A.Z. *Genes Dev.* **21**, 1519–1529
43. Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K., and Felsenfeld, G. (2009) H3.3/H2A.Z double variant-containing nucleosomes mark ?nucleosome-free regions? of active promoters and other regulatory regions. *Nat. Genet.* **41**, 941–945

Chapter 16

Characterization of Post-Meiotic Male Germ Cell Genome Organizational States

Jérôme Govin, Sophie Barral, Yuichi Morozumi, Naghmeh Hoghoughi, Thierry Buchou, Sophie Rousseaux, and Saadi Khochbin

Abstract

Dramatic and unique genome reorganizations accompany the differentiation of haploid male germ cells, characterized by a gradual loss of the vast majority of histones leading to a final tight compaction of the genome by protamines. Despite being essential for procreation and the life cycle, the mechanisms driving the transformation of nucleosomes into nucleoprotamines remain poorly understood. To address this issue, our laboratory has developed a number of specific approaches, ranging from the purification of spermatogenic cells at specific stages, the analysis of chromatin transitional states, the functional characterization of histone variants, histone-replacing proteins and their chaperones. This chapter will detail all related relevant techniques with a particular emphasis on methods allowing the functional studies of histone variants and the genome organizational states associated with the studied histones in spermatogenic cells undergoing histone-to-protamine exchange.

Key words Histone, Histone variants, Transition proteins, Protamines, Chaperones, Chromatin, Nucleosome, Spermatogenesis, Germline, Sperm

1 Introduction

During mammalian spermatogenesis, haploid male germ cells are continuously produced following successive meiotic divisions (occurring in spermatocytes). The early post-meiotic cells known as round spermatids undergo radical nuclear shape changes while becoming elongating spermatids, which themselves evolve into condensing spermatids and undergo the final transformations to produce mature spermatozoa. Mature sperm cells are capable of leaving the male reproductive organ and will have to face a challenging environment as they will "swim" toward the oocyte. The preparation of the final genome compaction virtually starts when spermatocytes/spermatid chromatin incorporate specific histone variants that favor the ulterior genome-wide histone eviction [1, 2]. The commitment of progenitor cells into meiosis is associated with

Guillermo A. Orsi and Geneviève Almouzni (eds.), *Histone Variants: Methods and Protocols*, Methods in Molecular Biology, vol. 1832, https://doi.org/10.1007/978-1-4939-8663-7_16,

the accumulation of testis-specific histones of the H2A, H2B, and H3 types with particular structural properties that survive meiosis and are inherited by the post-meiotic spermatids [3–7]. Additionally, haploid cells synthesize their own sets of histone H2A and H2B variants with specific roles in the control of testis-specific gene expression as well as in chromatin remodeling and reorganization and histone-to-protamine replacement [3, 7].

The gradual stage-specific incorporation of these histone variants along with the occurrence of a sharp and dramatic gain of histone posttranslational modifications (PTMs) in elongating spermatids [8–11], mediate the replacement of histones by basic small non-histone proteins known as transition proteins (TPs) and protamines (Prms).

Until recently, the prevailing view was that histones are first replaced by TPs, which are then themselves replaced by Prms [12, 13]. However, a recent functional investigation of a specific H2A histone variant, H2A.L.2, expressed at the time of histone eviction in elongating spermatids showed that an exchange of H2A-H2B histones is actually coupled and concomitant with the action of TPs and Prms, the generation of transitional states and DNA compaction [7]. This particular research on H2A.L.2, and our current understanding of the molecular basis of post-meiotic mouse male germ cell genome reorganization in our laboratory relies on years of pilot studies and development of specific approaches [14–16] that will be detailed here.

Our global approach to the issue of post-meiotic genome reorganization is built first on the design, adaptation, and improvement of techniques to isolate different types of post-meiotic cells [17], and second on the adaptation and development of methods to study different organizational states of chromatin, discover and identify histones, non-histone basic proteins and their chaperones, and generate relevant mouse models for functional molecular studies. Here below we successively describe the techniques for post-meiotic cell fractionation, the identification of basic and acidic proteins (respectively enriched in nucleic acid-binding proteins and chaperones, and in histone and non-histone basic proteins), the identification of histones [18], the study of chromatin organizational states, and the generation of mouse models.

2 Materials

All solutions are stored at room temperature (RT) unless otherwise specified.

2.1 Isolation of Condensing Spermatids

1. Homogenization Medium 1 (HM1): 0.32 M Sucrose, 1.5 mM $CaCl_2$, 10 mM Tris pH 8, 1 mM Dithiothreitol (DTT) and protease inhibitors cOmplete EDTA-free (Roche).

2. Homogenization Medium 2 (HM2): HM1 supplemented with 0.1% Triton.
3. Homogenization Medium 3 (HM3): HM2 supplemented with 0.88 M Sucrose (final concentration).
4. Filter with 100 μm wide pore.
5. Dissection material.
6. Motor-driven Potter-Elvehjem homogenizer.

2.2 Isolation of Basic and Acidic Proteins from Condensing Spermatids

1. Sulfuric acid (H_2SO_4).
2. Trichloro-acetic acid (TCA).
3. Acetone.
4. Acetone supplemented with 0.05% HCl.

2.2.1 Basic Proteins (Histone and Non-histones)

2.2.2 Acidic Proteins (TP Column)

1. TP1 cDNA cloned into pET28 plasmid for production as a recombinant protein (available upon request).
2. BL21 transformation-competent bacteria.
3. Isopropyl-beta-D-thiogalacto-pyranoside (IPTG).
4. Perchloric acid (PCA).
5. Trichloro-acetic acid (TCA).
6. Acetone.
7. Acetone supplemented with 0.05% HCl.
8. CNBr-activated sepharose 4b fast flow (GE Healthcare).
9. Coupling buffer: 0.5 M NaCl, 0.1 M $NaHCO_3$ pH 8.3.
10. Quenching buffer: 0.2 M Glycine pH 8.
11. Storage buffer: 0.1 M Tris pH 8, 0.5 M NaCl, 0.05% NaN_3.
12. Wash buffer 1: 0.1 M Na Acetate pH 4, 0.5 M NaCl.
13. Wash buffer 2: 0.1 M Tris pH 8, 0.5 M NaCl.
14. Extraction buffer 0 (EB0): 50 mM Tris pH 8, 1% NP-40, 0.5% Deoxycholic acid, 0.1% SDS, 1 mM DTT and protease inhibitors cOmplete EDTA-free (Roche).
15. Extraction buffer 1 (EB1): EB0 supplemented with 150 mM NaCl.
16. Extraction buffer 2 (EB2): EB0 supplemented with 1.2 M NaCl.

2.3 MNase Digestion of Condensing Spermatids Nuclei

1. Lysis buffer: 50 mM Tris pH 7.4, 300 mM NaCl, 0.1% NP-40, 0.1% Deoxycholic acid, 1 mM DTT and protease inhibitors cOmplete EDTA-free (Roche).
2. MNase digestion buffer: 10 mM Tris pH 7.5, 10 mM KCl, 1 mM $CaCl_2$.

3. Nuclease S7 Microccocale nuclease.
4. Chromatin Fiber Analysis Mix: 10 mM Tris pH 8, 1 mM EDTA, 0.1 mg/mL Proteinase K, 0.1% SDS.

2.4 Sucrose Gradient Fractionation of Nucleosomal and Sub-Nucleosomal Fragments

1. Gradient sucrose solution: 1 mM Phosphate Buffer pH 7.4 (prepared by mixing of 0.5 M NaH_2PO_4 pH 9 with 0.5 M Na_2HPO_4 pH 4.4 to reach pH 7.4), 80 mM NaCl, 0.2 mM EDTA, 20% Sucrose w/v, cOmplete EDTA-free (Roche), HDAC inhibitors (300 nM Tricostatin A).

2.5 TAP-Tag Histone Purification

1. Total germ cells or purified post-meiotic cells: Prepare as described in Subheading 3.1. They can be flash frozen in liquid nitrogen and stored at −80 °C.
2. Immunoprecipitation (IP) buffer: 20 mM HEPES pH 7.9, 300 mM NaCl, 10% Glycerol, 0.2 mM EDTA, 0.1% NP-40, 1 mM DTT and protease inhibitors cOmplete EDTA-free (Roche). Prepare freshly and cool on ice.
3. 0.45 μm wide pore membrane filters.
4. Anti-FLAG M2 Affinity Agarose Gel (Sigma-Aldrich).
5. Tris-buffered saline (TBS): 20 mM Tris pH 7.4, 150 mM NaCl.
6. 3× FLAG peptide (Sigma-Aldrich): Dissolve 3× FLAG Peptide in TBS at a concentration of 5 mg/mL as a stock solution. Since repeated freeze/thaw cycle is not recommended, make small aliquots of the stock solution and store at −20 °C.
7. Anti-HA antibody.
8. Dynabeads Protein G (Thermo Fisher Scientific).
9. Magnetic separation stand.
10. Conjugation buffer: 20 mM Phosphate buffer pH 7.4 prepared as described above (section 2.4), 150 mM NaCl.
11. Bis(sulfosuccinimidy)suberate (BS3) linker (Sigma-Aldrich, catalogue No. S5799): Dissolve BS3 in conjugation buffer at a concentration of 100 mM just prior to use.

2.6 Proteomic Analysis of Histone Variants

1. MS_HistoneDB files: download from reference 18 (Supp. Files 1 and 2).

3 Methods

A detailed protocol for the fractionation of spermatogenic cells by sedimentation has been published recently [17] (Fig. 1).

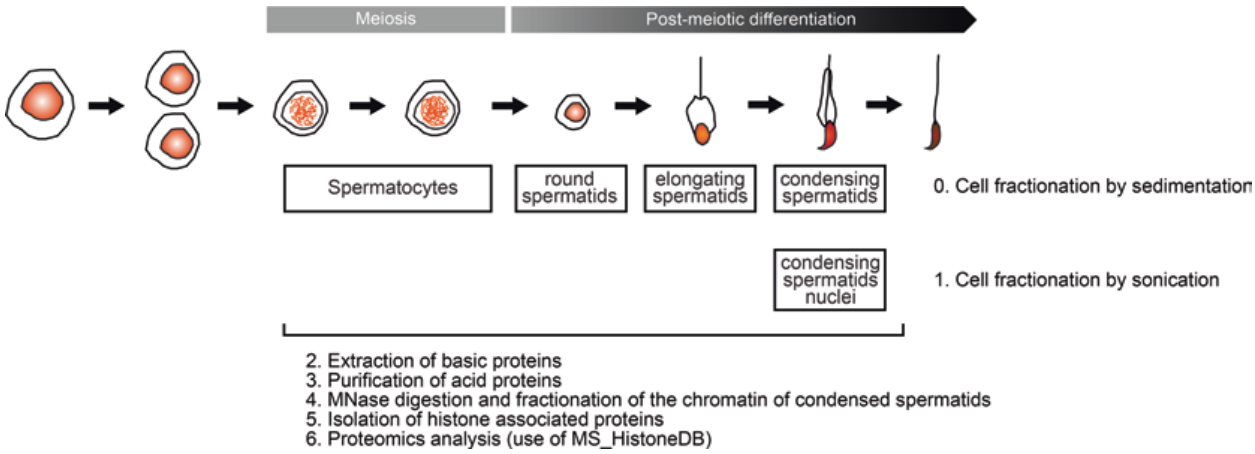


Fig. 1 Summary of the methods used to characterize the different chromatin organizational states during mouse spermatogenesis

3.1 Isolation of Nuclei from Condensing Spermatids

1. Start with 2 testes in cold 1× PBS (*see* **Note 1**).
2. Remove their tunica albuginea using fine tip tweezers.
3. Transfer all testicular material in a potter with 5 mL of Homogenizing Media 1 (HM1).
4. Disrupt mechanically using a motor-driven potter-homogenizer. Perform 10 strokes at 3000 rpm.
5. Complete the volume to 10 mL with Homegenizing Media 2 (HM2).
6. Centrifuge 10 min at 500 × *g* at 4 °C.
7. Resuspend the pellet in 3 mL of cold HM2.
8. Sonicate the tube on ice at 250 J.
9. Centrifuge for 10 min at 500 × *g* at 4 °C.
10. Resuspend the pellet in 3 mL of cold HM2.
11. Sonicate the tube on ice at 250 J.
12. Centrifuge for 10 min at 500 × *g* at 4 °C.
13. Filter the suspension with a 100 µm pore size filter.
14. Add HM2 to obtain a final volume of 10 mL.
15. Centrifuge for 10 min at 500 × *g* at 4 °C and resuspend the pellet in 10 mL of HM2.
16. Centrifuge for 10 min at 500 × *g* at 4 °C and resuspend the pellet in 10 mL of HM3.
17. Centrifuge for 10 min at 500 × *g* at 4 °C and resuspend the pellet in 1 mL of HM1.
18. Transfer in microfuge tubes 1.5 mL.
19. Centrifuge and flash freeze the dry pellets. They can be stored at −80 °C (*see* **Note 2**).

3.2 Isolation of Basic Proteins from Condensing Spermatids Nuclei

1. Resuspend a pellet of condensed spermatids in 500 µL of 0.2 M H_2SO_4. Basic proteins will be solubilized while other proteins will precipitate.
2. Sonicate briefly if necessary.
3. Incubate for 30 min on ice.
4. Centrifuge for 30 min at 21,000 × *g* at 4 °C and transfer the supernatant to a new prechilled tube.
5. Precipitate histones by adding 110 µL of Trichloroacetic acid (TCA) (20% v/v final concentration, *see* **Note 3**).
6. Incubate for 30 min on ice (*see* **Note 4**).
7. Centrifuge for 30 min at 21,000 × *g* at 4 °C. Precipitated histones should be visible on the side of the tube.
8. Discard the supernatant and wash the pellet with 500 µL of Acetone 0.05% HCl.
9. Centrifuge for 10 min at 21,000 × *g* at 4 °C.
10. Wash again in pure Acetone.
11. Centrifuge for 10 min at 21,000 × *g* at 4 °C.
12. Carefully remove the maximum of acetone using P200 and P10 pipettes. Dry the pellet at RT for 10–20 min (*see* **Note 5**).
13. Resuspend the pellet in H_2O (*see* **Note 6**).
14. Centrifuge 10 min at 21,000 × *g* at 4 °C to remove non-soluble proteins. The supernatant contains acid-soluble proteins.

3.3 Isolation of Acidic Proteins from Condensing Spermatids Nuclei

This protocol has been used to identify proteins able to interact with the basic transition protein TP1 [14].

3.3.1 Production and Purification of the Transition Protein 1 (TP1)

1. Transform TP1 cDNA cloned in pET 28 vector in BL21 competent bacteria.
2. The production of TP1 is induced in a 100 mL log phase culture (OD_{600} 0.7) for 1 h at 37 °C with 2 mM IPTG.
3. Wash the pellet in PBS and resuspend in 5 mL of 5% perchloric acid (*see* **Note 7**).
4. Incubate for 30 min incubation on ice.
5. Centrifuge at 21,000 × *g* during 10 min at 4 °C.
6. Add 1.1 mL of trichloroacetic acid (TCA) to reach a final concentration of 20%.
7. Incubate on ice for 30 min.
8. Centrifuge at 21,000 × *g* during 10 min at 4 °C.
9. Discard the supernatant and wash the pellet with 500 µL of 0.05% Acetone HCl.

10. Centrifuge during 10 min at 21,000 × *g* at 4 °C.
11. Wash again in pure Acetone. Centrifuge 10 min at 21,000 × *g* at 4 °C.
12. Carefully remove the maximum of acetone using P200 and P10 pipettes. Dry the pellet at RT for 10–20 min (*see* **Note 5**).
13. Resuspend the pellet in H_2O (*see* **Note 6**).
14. Centrifuge for 10 min at 21,000 × *g* at 4 °C to remove unsoluble proteins. The supernatant contains purified TP1, which is aliquoted and stored at −80 °C. Pure TP1 recombinant protein will be observed on Coomassie-strained SDS-PAGE gel.

3.3.2 TP1 Crosslinking on an Agarose Resin

TP1 is crosslinked to CNBr-activated agarose beads (GE Healthcare) following the manufacturer's instructions.

1. Resuspend 0.6 g of dried resin in 300 mL of 1 mM HCl.
2. Wash with 15 mL of coupling buffer.
3. Resuspend in 6 mL of coupling buffer.
4. Incubate half of the resin for 2 h at RT with 270 μg of TP1 whereas the other half is incubated with a similar amount of BSA.
5. Remove the buffer by centrifugation or filtering and add 2 mL of quenching solution.
6. Incubate for 2 h at RT.
7. Extensively wash the resin sequentially with 2 mL of wash buffers 1 and 2. Store the column at 4 °C in storage buffer.

3.3.3 Purification of TP1 Interacting Proteins

1. Resuspend condensing spermatids in 100 μL of cold Extraction Buffer 1 (EB1).
2. Sonicate briefly to disrupt the cells.
3. Incubate for 15 min at 4 °C.
4. Centrifuge 10 min at 4 °C at 10,000 × *g*.
5. Save the supernatant and resuspend the pellet in 100 μL of cold EB2.
6. Incubate for 15 min at 4 °C.
7. Centrifuge 10 min at 4 °C at 10,000 × *g*.
8. Pool the extracts obtained in EB1 and EB2 and add 700 μL of EB0 to adjust the concentration of NaCl to 150 mM.
9. Wash twice 60 μL of the BSA or TP1 crosslinked resins in 1 mL of EB1.
10. Incubate 200 μL of the protein extract with 60 μL of the resin for 2 h at 4 °C under agitation.
11. Wash three times the resins in 1 mL of EB1.

12. Elute proteins in 15 µL of loading buffer and boil for 5 min.
13. Purified proteins can be analyzed by SDS-PAGE or using standard mass spectrometry procedures (Subheading 3.7).

3.4 MNase Digestion of the Chromatin of Condensing Spermatids

1. Resuspend condensing spermatids (obtained as described in Subheading 3.1, using six testes) in 112.5 µL of Lysis buffer.
2. Incubate for 15 min at 4 °C.
3. Centrifuge for 10 min at 4 °C at 20,000 × *g*.
4. Save the supernatant.
5. Resuspend the pellet in 112.5 µL of lysis buffer.
6. Sonicate the tube on ice at 80 J.
7. Centrifuge for 10 min at 4 °C at 20,000 × *g*.
8. Save the supernatant.
9. Pool both supernatants.
10. Add 112.5 µL of MNase buffer and 11.5 Units of micrococcal nuclease S7.
11. Incubate for 2 min at 37 °C (*see* **Note 8**).
12. Stop the digestion by adding 5 mM EDTA (final concentration) and place on ice.
13. Analyze the length of DNA fragments: digest 5 µL of chromatin fibers with 2.5 µL of Chromatin Fibers Analysis Mix for 15 min at 50 °C. Analyze directly on a 1.5% agarose gel.

3.5 Separation of Nucleosomal Chromatin Fragments

1. Pour the gradient sucrose solution in the tubes and freeze at −20 °C (*see* **Note 9**).
2. The day before use, thaw the tubes slowly by placing them in at 4 °C. The gradient will be formed by differential freezing / thawing of the sucrose solution (*see* **Note 10**).
3. Load slowly the chromatin fibers obtained after **step 13** (Subheading 3.4) on the sucrose gradient. Equilibrate the tubes with EDTA 0.1 mM if necessary.
4. Centrifuge for 16 h at 4 °C in the SW28 rotor at 25,000 rpm (~82,000 × *g*, large tubes), or in the TLS-55 rotor at 32,000 rpm (~68,000 × *g*, small tubes).
5. Collect 500 or 100 µL fractions of the gradients for large or small tubes, respectively.
6. The length of DNA fragments from fractions can be analyzed by digesting 5-10 µL with 2.5 µL of Chromatin Fibers Analysis Mix for 15 min at 50 °C followed by migration on a 1.1% agarose gel.

3.6 Purification of Soluble Histone Associated Proteins in the Male Germline

Histone genes were tagged with a combination of three tags: His, Flag, and HA. The first affinity purification uses the Flag tag since the protein complexes can be recovered using a competitor Flag peptide under non-denaturing condition. The HA tag was used to perform a second affinity purification.

3.6.1 Anti-HA Antibody Cross-Linking to Dynabeads Protein G

Contamination of the purified complex by heavy and light chains of the antibody may interfere with subsequent analyses. For this reason, the antibody is covalently cross-linked to the beads. The following cross-linking protocol is suitable for 5 μg of HA antibody with 50 μL of Dynabeads Protein G (*see* **Note 11**). This procedure can be scaled up or down if necessary. Unless otherwise indicated, perform the whole procedure at RT.

1. Transfer 50 μL of Dynabeads Protein G in a microfuge tube. Place the tube on magnetic separation stand for 1 min and discard the supernatant.
2. Wash the beads with 200 μL of cold PBS.
3. Wash the beads twice with 200 μL of IP buffer.
4. Incubate the beads with 5 μg of the anti-HA antibody in 200 μL of IP buffer. Perform incubation under gentle rotation at 4 °C for 1 h.
5. Discard the supernatant and wash the beads twice with 200 μL of conjugation buffer. During the second wash, transfer the beads to a new microfuge tube.
6. Resuspend the beads in 250 μL of 5 mM BS3 cross-linker diluted in conjugation buffer and gently rotate for 30 min.
7. Add 12.5 μL 1 M Tris pH 7.5 to the beads and incubate for 15 min under rotation (*see* **Note 12**).
8. Wash the beads twice with 200 μL of IP buffer (*see* **Note 13**). During second wash, transfer the beads to a new tube.
9. Add 50 μL of IP buffer to the beads and store on ice until use.

3.6.2 Purification of Soluble Histones Using TAP-Tag

The entire protocol is performed at 4 °C unless otherwise specified.

1. Resuspend the required amount of frozen male germ cell or purified post-meiotic cell pellets prepared using histone TAP-tagged mice (*see* Subheading 3.1) in the appropriate volume of IP buffer (*see* **Notes 14** and **15**).
2. Extract soluble proteins by gently rotating the cell suspension at 4 °C for 30–60 min.
3. During **step 2**, transfer the required amount of Anti-FLAG M2 affinity beads in a new microfuge tube (*see* **Note 16**). Wash the beads once with 1 mL of 1× PBS, and twice with 1 mL of IP buffer. Add an equal volume of IP buffer to the beads to make 50% beads slurry.
4. Recover the clarified lysate by centrifugation at 21,000 × *g* for 10 min, and then filtrate the supernatant through a 0.45 μm wide pore membrane filter (*see* **Note 17**). Transfer 10–20 μL of the supernatant to a new microfuge tube (input fraction) and store at −20 °C.

5. Mix the lysate with the equilibrated Anti-FLAG M2 affinity beads and incubate on a rotator at 4 °C for 4 h (*see* **Note 18**).
6. During **step 5**, prepare the anti-HA antibody cross-linked Dynabeads Protein G (*see* Subheadig 3.5, **step 1** and **Note 11**).
7. Separate the lysate and the beads by centrifugation at 2000 × *g* for 2 min. Remove the supernatant and transfer 10–20 μL of the supernatant to a new microfuge tube (first unbound fraction) and store at −20 °C.
8. Wash the beads in 1 mL of IP buffer. Centrifuge at 2000 × *g* for 2 min and discard the supernatant. Repeat this wash step five times. During the fifth wash, transfer the beads to a new microfuge tube (*see* **Note 19**).
9. Resuspend the beads in an equal volume of 0.3 mg/mL 3× FLAG peptide diluted in IP buffer (*see* **Note 20**). Gently rotate on a rotator at 4 °C for 30 min.
10. Centrifuge the resin at 2000 × *g* for 2 min. Carefully transfer the supernatant to a new microfuge tube.
11. Repeat **steps 9** and **10** again, and combine the elution fractions.
12. To remove the possibly remaining small amount of beads in the combined eluate, spin the sample at 2000 × *g* for 2 min, and carefully recover the supernatant into a new microfuge tube. Transfer 10–20 μL of the supernatant to another microfuge tube (first elution fraction) and store at −20 °C.
13. Mix the clarified elution fraction with the appropriate amount of anti-HA antibody cross-linked Dynabeads Protein G, and incubate with gentle rotation at 4 °C overnight (*see* **Note 21**).
14. Place the tube on a magnetic stand for 1 min. Remove the supernatant and transfer 10–20 μL of the supernatant to a new microfuge tube (second unbound fraction) and store at −20 °C.
15. Add 1 mL of IP buffer to the beads and resuspend gently. After the beads settle down by placing the tube on the stand for 1 min, discard the buffer. Repeat this wash step three times. During the third wash, transfer the beads to a new microfuge tube.
16. Add 10–30 μL of 1× Laemmli sample buffer without reducing agent. Heat the beads at 95 °C for 5 min (*see* **Note 22**).
17. Place the tube on a magnetic stand and transfer the supernatant to a new microfuge tube. Store the sample at −20 °C for further experiments.

18. Analyze the collected samples at each purification step (**steps 4**, **7**, **12**, **14** and **16**) by western blot to estimate the efficiency of TAP-tag purification.

3.7 Proteomic Analysis of Histone Variants

Mass spectrometry is a powerful technique to analyze the composition of chromatin, the diversity of histones present in a specific chromatin fraction and their posttranslational modifications. Several published reviews and protocols already document this type of analysis [19–24]. In brief, histones are extracted and purified from the initial material. They are usually proteolyzed into smaller peptides using a specific enzyme and these peptides are subsequently analyzed by mass spectrometry. Raw data are mainly made of fragmentation spectra acquired on peptides from which it is possible to identify the amino acid sequences using search softwares, such as the widely used Mascot software (Matrix Science). Additional post-analysis treatments enable identification and quantification of the proteins and potentially of their posttranslational modifications present in the initial protein sample [25].

The interpretation of the raw spectra relies on public protein sequence databases, such as SwissProt or UniProtKD. Unfortunately, they are either incomplete or overly redundant with partial annotation, complicating the analysis of histone proteins and their variants. We recently developed a new resource, named MS_HistoneDB, which can be directly used into standard proteomic pipelines and greatly facilitate proteomic approaches dedicated to the analysis of histones [18]. >1000 entries have been collected and manually filtered to create a final list of ~80 unique histone entries, in human and mouse species.

4 Notes

1. Isolation of condensed spermatids is routinely performed on 2 mouse testes but it is possible to modulate the number of testes and adjust the volume of buffers proportionally. The testes can be collected from male mice upon availability in animal facility, briefly washed in cold PBS 1× and stored in microfuge tubes at −80 °C. Male mice have to be 8–9 weeks old in order to produce mature sperm: they reach puberty ~34–38 days after birth and another 28 days are required for the differentiation of sperm.
2. The quality of the preparation can be controlled observing an aliquot of each step with a microscope.
3. Alternatively, acid soluble extract can be buffered by adding 110 µL of Tris 2 M pH 8. The pH of the extract can be verified by pipetting 1 µL on a pH paper. Add more Tris solution if necessary. We recommend to incubate this solution for 30 min

on ice and centrifuge 30 min at 21,000 × *g* at 4 °C to remove any potential precipitate. This option is useful when high concentration of histones is not required.

4. TCA-precipitated histones can be stored overnight at −20 °C. However, a prolonged state in this strong acid could impact posttranslational modifications of histones. Long-term storage should be done on acetone washed pellets.
5. This step should remove the excess of acetone. Extensive drying will result in difficulties to resuspend the pellet in water.
6. Resuspension in water can be difficult. The pellet can be briefly sonicated to facilitate the resuspension of histones. Alternatively, the pellet can be resuspended in SDS-PAGE loading buffer.
7. The purification of TP1 takes advantage of the extreme basic property of this protein, which is soluble at very low pH. Acid soluble proteins are extracted directly from the cell pellet by solubilization in a strong acid (perchloric acid).
8. Conditions of MNase digestion can be optimized by running a time course experiment to obtain nucleosomal fibers of the desired length, evaluated on an agarose as described in Subheading 3.4, **step 13**.
9. Two volumes can be used: large gradients, formed with 16 mL of Sucrose Gradient solution in 16 × 102 mm tubes for SW28 Beckman-Coulter rotor (ref 337986), or small gradients, using 2.3 mL of the same solution in 11 × 34 mm tubes (ref 347357) to be used on a TLS-55 rotor.
10. Ideally in a cold room without any vibration (lab fridges are frequently opened and such vibrations will perturb the quality of the gradient).
11. The affinity of protein G for antibodies varies depending on the host species and antibody subclasses. Dynabeads Protein A is also available if the affinity of the antibody to protein G is not optimal.
12. BS3 crosslinker reacts on primary and secondary amines. Adding 1 M Tris pH 7.5 will provide them in excess and quench the cross-linking reaction.
13. Uncross-linked antibody may remain on the beads. An additional wash with 200 µL of Glycine 100 mM pH 3.0 will remove the uncross-linked antibody. This wash has to be brief to preserve the functionality of the antibody. Then wash the beads twice with 200 µL of IP buffer to remove traces of Glycine acidic solution.
14. The salt concentration of the IP buffer can be adapted depending on the purpose of the experiment. Weakly bound proteins will be lost when using higher salt concentrations.

15. The required amount of cell pellet is dependent on the expression level of the protein of interest and the scale of the experiment. Total male germ cells obtained from two or three mice would be applied to large-scale purification and the analysis of histone-associated proteins by mass spectrometry. In the case of purified post-meiotic cells, larger numbers of mice (more than ten) would be required. The volume of the IP buffer should be 10–20 times the volume of the cell pellet.
16. The required amount of beads varies depending on the scale of the experiment. Generally, 20 μL and 100–200 μL of beads are optimal for small- and large-scale purifications, respectively.
17. Particles (e.g., lipid) present in the cell lysate, which may nonspecifically bind to the beads and cause an increased background, will be removed at that step.
18. Longer time incubation may increase nonspecific binding of background proteins and disrupt weak protein interactions.
19. This step is important to eliminate the contamination of the components that have been absorbed on the surface of the tube wall during the incubation.
20. In the case of less than 50 μL of the beads, it is recommended to use 50 μL of the peptide solution to efficiently recover the eluate.
21. The necessary amount of antibody cross-linked beads depends on the affinity of the antibody toward the protein of interest. In our hands, 10–20 μL of the beads corresponding to 1–2 μg cross-linked antibody is sufficient for large-scale purification if a high-affinity antibody is used.
22. It is not recommended to use 1× Laemmli sample buffer containing reducing agents (DTT or β-mercaptoethanol). It may increase a contamination of the eluate by the light chains of the antibody, released by reducing agents.

Acknowledgments

SK laboratory is supported by a grant from "Foundation pour la Recherche Medicale (FRM)" "analyse bio-informatique pour la recherche en biologie" program as well as by ANR Episperm3 program and by INCa libre program (RPT13001CCA)). Additional supports were from Fondation ARC "Canc'air" project (RAC16042CLA), Plan Cancer (CH7-INS15B66) and Plan Cancer (ASC16012CSA), as well as the "Université Grenoble Alpes" ANR-15-IDEX-02 (LIFE and SYMER).

The group of J. Govin is supported by grants from the Agence Nationale de Recherche [ANR-14-CE16-0027-01 (FungiBET);

ANR-11-PDOC-011-01 (EpiGam); ANR-10-INBS-08 (ProFI); ANR-10-LABX-49-01 (GRAL)] and the EU FP7 Marie Curie Action (Career Integration Grant 304003). Additional supports were from the Univ. of Grenoble Alpes (Agir) and FINOVI from the Région Rhône Alpes (France).

References

1. Govin J, Caron C, Lestrat C et al (2004) The role of histones in chromatin remodelling during mammalian spermiogenesis. Eur J Biochem 271:3459–3469
2. Gaucher J, Reynoird N, Montellier E et al (2010) From meiosis to postmeiotic events: the secrets of histone disappearance. FEBS J 277:599–604
3. Govin J, Escoffier E, Rousseaux S et al (2007) Pericentric heterochromatin reprogramming by new histone variants during mouse spermiogenesis. J Cell Biol 176:283–294
4. Montellier E, Boussouar F, Rousseaux S et al (2013) Chromatin-to-nucleoprotamine transition is controlled by the histone H2B variant TH2B. Genes Dev 27:1680–1692
5. Shinagawa T, Huynh LM, Takagi T et al (2015) Disruption of Th2a and Th2b genes causes defects in spermatogenesis. Development 142(7):1287–1292
6. Ueda J, Harada A, Urahama T et al (2017) Testis-specific histone variant H3t gene is essential for entry into spermatogenesis. Cell Rep 18:593–600
7. Barral S, Morozumi Y, Tanaka H et al (2017) Histone variant H2A.L.2 guides transition protein-dependent protamine assembly in male germ cells. Mol Cell 66:89–101 e8
8. Goudarzi A, Shiota H, Rousseaux S, Khochbin S (2014) Genome-scale acetylation-dependent histone eviction during spermatogenesis. J Mol Biol 426:3342–3349
9. Tan M, Luo H, Lee S et al (2011) Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. Cell 146:1016–1028
10. Dai L, Peng C, Montellier E et al (2014) Lysine 2-hydroxyisobutyrylation is a widely distributed active histone mark. Nat Chem Biol 10:365–370
11. Goudarzi A, Zhang D, Huang H et al (2016) Dynamic competing histone H4 K5K8 acetylation and Butyrylation are hallmarks of highly active gene promoters. Mol Cell 62:169–180
12. Lewis JD, Abbott DW, Ausió J (2003) A haploid affair: core histone transitions during spermatogenesis. Biochem Cell Biol 81:131–140
13. Meistrich ML, Mohapatra B, Shirley CR, Zhao M (2003) Roles of transition nuclear proteins in spermiogenesis. Chromosoma 111:483–488
14. Govin J, Gaucher J, Ferro M et al (2012) Proteomic strategy for the identification of critical actors in reorganization of the post-meiotic male genome. Mol Hum Reprod 18:1–13
15. Rousseaux S, Khochbin S (2012) Combined proteomic and in silico approaches to decipher post-meiotic male genome reprogramming in mice. Syst Biol Reprod Med 58:191–196
16. Rousseaux S, Boussouar F, Gaucher J et al (2011) Molecular models for post-meiotic male genome reprogramming. Syst Biol Reprod Med 57:50–53
17. Buchou T, Tan M, Barral S et al (2017) Purification and analysis of male germ cells from adult mouse testis. Methods Mol Biol 1510:159–168
18. El Kennani S, Adrait A, Shaytan AK et al (2017) MS_HistoneDB, a manually curated resource for proteomic analysis of human and mouse histones. Epigenetics Chromatin 10:2
19. Villar-Garea A, Israel L, Imhof A (2008) Analysis of histone modifications by mass spectrometry. Curr Protoc Protein Sci Chapter 14, Unit 14.10
20. Yuan Z-F, Arnaudo AM, Garcia BA (2014) Mass spectrometric analysis of histone proteoforms. Annu Rev Anal Chem 7:113–128
21. Sidoli S, Bhanu NV, Karch KR et al (2016) Complete workflow for analysis of histone post-translational modifications using bottom-up mass spectrometry: from histone extraction to data analysis. J Vis Exp 17(111):e54112
22. Sabari BR, Zhang D, Allis CD, Zhao Y (2017) Metabolic regulation of gene expression through histone acylations. Nat Rev Mol Cell Biol 18:90–101

23. Zhao Y, Garcia BA (2015) Comprehensive catalog of currently documented histone modifications. Cold Spring Harb Perspect Biol 7:a025064
24. Huang H, Lin S, Garcia BA, Zhao Y (2015) Quantitative proteomic analysis of histone modifications. Chem Rev 115(6):2376–2418 150217145638004
25. Ramus C, Hovasse A, Marcellin M et al (2016) Benchmarking quantitative label-free LC-MS data processing workflows using a complex spiked proteomic standard dataset. J Proteome 132:51–62

Le mystère de la disparition des histones au cours de la spermatogenèse

Sophie Barral, Yuichi Morozumi, Naghmeh Hoghoughi, Sophie Rousseaux, Saadi Khochbin

CNRS UMR 5309 ; Inserm, U1209 ; université Grenoble Alpes ; Institute for Advanced Biosciences, Grenoble, F-38700 France.
saadi.khochbin@univ-grenoble-alpes.fr

L'empaquetage du génome mâle, un processus qui reste mystérieux

Le spermatozoïde est la seule cellule capable de survivre en dehors de l'organisme producteur. Il s'agit d'une cellule hautement spécialisée qui se prépare à quitter l'organe génital mâle et à survivre dans les voies génitales femelles, lieu de la fécondation. La mission du spermatozoïde est de transporter une copie du génome mâle jusqu'à l'ovocyte pour le livrer lors de la fusion gamétique. Le processus qui aboutit à la production des spermatozoïdes est connu sous le nom de spermatogenèse. La différenciation des cellules germinales mâles progénitrices comprend deux étapes principales. La première implique la méiose au cours de laquelle le génome est initialement brassé puis divisé, aboutissant à la production de quatre cellules nommées spermatides, portant chacune une seule copie du génome. Au cours de la deuxième étape, ou spermiogénèse, les gamètes mâles se préparent pour leur futur voyage à l'extérieur de l'organisme de production vers le lieu de fécondation. Lors de cette étape post-méiotique, les gamètes mâles sont le siège d'une véritable métamorphose à l'issue de laquelle ces cellules seront dotées d'un flagelle, indispensable pour la mobilité des futurs spermatozoïdes. Concomitamment, ils subissent une réorganisation spectaculaire et unique de leur génome. Le fait marquant de cette différenciation post-méiotique est la disparition des nucléosomes[1], unités universelles d'organisation et d'empaquetage des génomes chez les eucaryotes.

La disparition des histones, les constituants des nucléosomes, au cours de la spermiogenèse et leur remplacement par des petites protéines basiques non-histones, les protéines de transition et les protamines ont été précédemment décrits [1]. Cette transformation globale de l'organisation du génome dans les cellules germinales mâles a comme objectif la compaction extrême de celui-ci avant le voyage du spermatozoïde afin d'assurer sa protection lors de cette période à risque pour son intégrité.

De nombreux travaux ont désormais montré que toute perturbation affectant le remplacement des histones aboutissait à une compaction défectueuse du génome dans le spermatozoïde et à une infertilité masculine. Le remplacement des histones est donc une étape critique pour la procréation et la perpétuation de l'espèce. Néanmoins, une question fondamentale reste ouverte : nous ne savons pas comment les histones sont dissociées de l'ADN et disparaissent, ni comment cette dissociation des histones est couplée à l'assemblage des protéines non-histones et à l'empaquetage du génome. En d'autres termes, l'ADN se structure autour des protamines et non plus des histones, mais nous ne comprenons pas comment les nucléosomes sont remplacés par les nucléoprotamines.

La vision classique et largement documentée est que les histones sont initialement remplacées par des protéines dites « de transition » elles-mêmes substituées, à leur tour, par les protamines.

Malgré l'importance de ces questions, nous ne disposons d'aucune information quant à la compréhension des mécanismes qui régissent ces bouleversements dans l'organisation du génome. Or la connaissance des mécanismes sous-jacents est susceptible d'influencer non seulement la biologie de la chromatine et l'épigénétique, mais aussi de nous éclairer sur les éléments déterminants d'une fertilité masculine optimale.

Le remplacement des histones : par qui et comment ?

Malgré l'identification et la caractérisation des protéines non-histones qui remplacent les histones au cours de la spermiogénèse, la relation fonctionnelle entre histones, protéines de transition et protamines est longtemps restée obscure. L'observation de l'apparition successive des protéines de transition puis des protamines avait initialement suggéré une action séquentielle de celles-ci dans la réorganisation du génome. Lors d'une première étape, les histones étaient remplacées par les protéines de transition et dans une deuxième phase, les protamines chassaient les protéines de transition et dirigeaient la compaction finale du génome [2,3]. Cependant, des observations ont semé le doute sur

[1] Le nucléosome est composé d'un fragment d'ADN (environ 147 paires de bases) enroulé autour d'un octamère d'histones.

m/s n° 6-7, vol. 33, juin-juillet 2017
DOI : 10.1051/medsci/20173306010

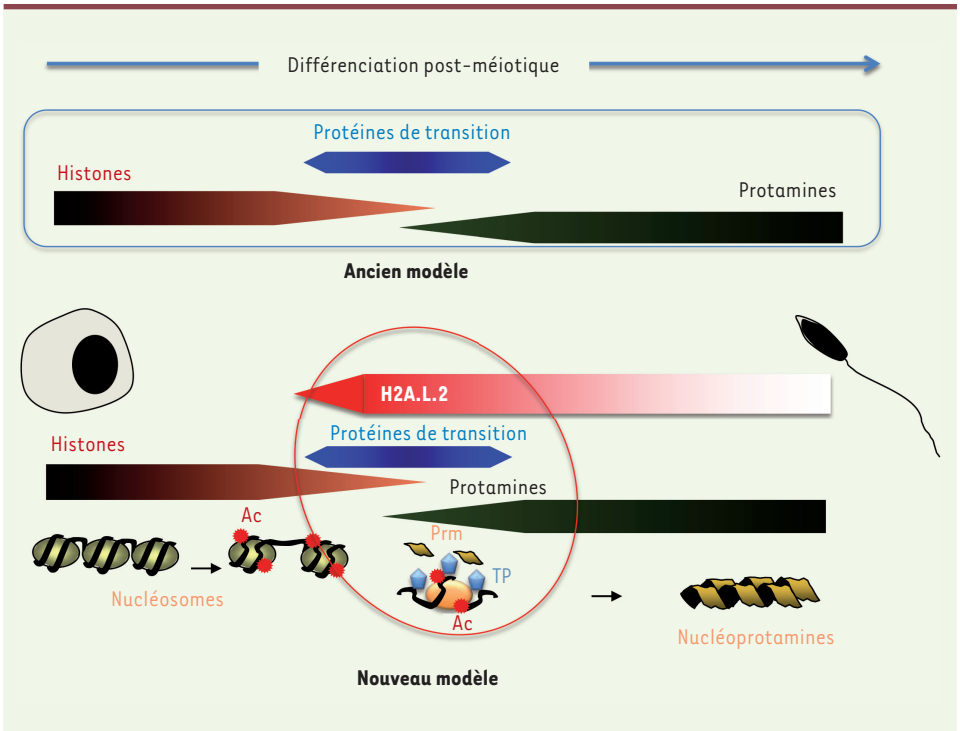


Figure 1. ***Nouveau modèle résumant les étapes impliquées dans le remplacement des histones au cours de la spermatogenèse.*** Le schéma général du remplacement des histones largement répandu proposait l'action, par étapes successives et distinctes, des protéines de transition (TP) remplaçant les histones, puis des protamines (Prm) se substituant aux protéines de transition (modèle ancien, partie supérieure). Nos travaux [8] mettent en évidence une étape supplémentaire qui correspond à la synthèse concomitante du variant d'histone H2A.L.2 et des protéines de transition. Leur action simultanée facilite l'incorporation des protamines. Ce nouveau modèle est présenté dans la partie inférieure de la figure. Ac : acétylation.

cette hypothèse du remplacement des histones en deux étapes successives, l'ordre d'apparition des protéines remplaçant les histones ne semblant pas aussi clair qu'initialement pensé [3]. En effet, les protéines de transition, les protamines et les histones peuvent être simultanément détectées dans la même cellule [4]. D'autres expériences, notamment la génération de souris déficientes (*double Knock Out*) pour les deux gènes codant les protéines de transition, ont rendu la situation encore plus confuse. Ces travaux ont montré qu'en l'absence totale des protéines de transition, le remplacement des histones se produisait [4-6]. Cette dernière observation pose donc la question essentielle du rôle des protéines de transition dans le processus de remplacement des histones. Quel est leur rôle ? Les expériences de double extinction des gènes les codant ont montré que ces protéines sont nécessaires pour une compaction efficace du génome par les protamines ainsi que pour la maturation de la protamine 2 (Prm2). La protamine 2 est synthétisée sous la forme d'une pré-protéine qui subit un clivage protéolytique dans sa partie N-terminale lors de son assemblage. L'accumulation du précurseur pré-Prm2 en l'absence des protéines de transition indique un rôle pour ces protéines dans le contrôle de l'assemblage des protamines. Cependant, aucune donnée publiée ne laissait entrevoir quelle pouvait être la relation fonctionnelle entre les protéines de transition et les protamines.

L'invasion du nucléosome : une histone jouant le rôle de « cheval de Troie »

La découverte en 2007 [7], d'une série de variants d'histones de type H2B et H2A exprimés spécifiquement dans les cellules spermatogéniques a permis de reconsidérer le remplacement des histones sous un nouveau jour. En effet, parmi ces histones, un variant de H2A nommé H2A.L.2 est exprimé tardivement, pratiquement en même temps que les protéines de transition, et de surcroît dans les mêmes cellules que celles-ci [8]. Ces observations ont désigné H2A.L.2 comme un facteur particulièrement intéressant, car il s'agit d'une histone synthétisée au moment même du remplacement général des histones. Cette situation inattendue suggérait que cette histone pouvait jouer un rôle dans la transformation des nucléosomes et plus précisément, qu'elle pouvait être un nouvel acteur dans le jeu d'interactions entre les protéines de transition et les protamines. L'observation de la spermatogenèse des souris invalidées pour le gène codant H2A.L.2 a considérablement renforcé cette hypothèse. Ces souris ont en effet exactement le même phénotype que celles doublement déficientes pour les gènes codant les protéines de transition, à savoir une perturbation de la maturation de la protéine pré-Prm2, une compaction défectueuse du génome dans les spermatozoïdes et une infertilité des mâles. Ces observations suggèrent qu'en l'absence de l'histone H2A.L.2, malgré une synthèse normale, les protéines de transition restent non-fonctionnelles. Ainsi, l'absence du variant H2A.L.2 serait donc fonctionnellement équivalente à la carence en protéines de transition. Autrement dit, l'histone H2A.L.2 serait nécessaire aux protéines de transition pour assurer leur fonction.

Une série d'expériences visant à vérifier cette hypothèse a mis en évidence le

mécanisme par lequel l'histone H2A.L.2 facilite l'action des protéines de transition : l'incorporation de l'histone H2A.L.2 ouvre le nucléosome, ce qui permet aux protéines de transition d'être chargées sur la chromatine et d'interagir avec les protamines tout en assurant la maturation du précurseur pré-Prm2. En l'absence de H2A.L.2, les protéines de transition ne peuvent pas entrer efficacement en interaction avec la chromatine et s'agrègent donc en piégeant une fraction des protamines qui ne peuvent pas être maturées ni assemblées. La conséquence de cette situation est un sous-dosage de protamines au niveau de l'ADN et une compaction défectueuse du génome [8]. L'histone H2A.L.2 se comporte donc comme un cheval de Troie : une fois incorporée au sein du nucléosome, ce variant d'histone ouvre la structure et permet l'invasion du nucléosome par des protéines non-histones.

Conclusion

L'étude fonctionnelle de l'histone H2A.L.2 a permis de mieux comprendre le phénomène mystérieux de la disparition à grande échelle des histones, qui survient au cours de la phase post-méiotique de la spermatogenèse. La stratégie utilisée par la nature est de favoriser un changement structural du nucléosome et une déstabilisation de celui-ci à l'échelle du génome par la synthèse d'un variant d'histone de type H2A produit en même temps que les protéines spécialisées qui remplacent les histones. Nous savons néanmoins que cette étape à elle seule n'est pas suffisante pour assurer cette transformation du nucléosome. Le remplacement se produit dans un contexte où le nucléosome doit préalablement être hyper-acétylé [9,10], (➜) un événement qui pourrait agir en favorisant l'échange des molécules H2A déjà présentes dans le nucléosome avec le variant H2A.L.2 néo-synthétisé, par un accroissement de la dynamique chromatinienne.

(➜) Voir la Nouvelle de S. Rousseaux *et al.*, *m/s* n° 2, février 2010, page 130

En conclusion, ces nouvelles données incitent à reconsidérer le scénario classique impliquant l'action successive des protéines de transition puis des protamines. En effet, l'incorporation de l'histone H2A.L.2 accélère la transformation des nucléosomes en induisant l'action presque simultanée des protéines de transition et des protamines. Ainsi, l'incorporation du variant d'histone H2A.L.2 initie la transformation des nucléosomes et correspond à une nouvelle étape, antérieure au remplacement des histones *(Figure 1)*. ◊

The mystery of histone disappearance during spermatogenesis

LIENS D'INTÉRÊT

Les auteurs déclarent n'avoir aucun lien d'intérêt concernant les données publiées dans cet article.

RÉFÉRENCES

1. Gaucher J, Reynoird N, Montellier E, *et al.* From meiosis to postmeiotic events: the secrets of histone disappearance. *FEBS J* 2010 ; 277 : 599-604.
2. Lewis JD, Abbott DW, Ausio J. A haploid affair: core histone transitions during spermatogenesis. *Biochem Cell Biol* 2003 ; 81 : 131-40.
3. Meistrich ML, Mohapatra B, Shirley CR, *et al.* Roles of transition nuclear proteins in spermiogenesis. *Chromosoma* 2003 ; 111 : 483-8.
4. Zhao M, Shirley CR, Mounsey S, *et al.* Nucleoprotein transitions during spermiogenesis in mice with transition nuclear protein Tnp1 and Tnp2 mutations. *Biol Reprod* 2004 ; 71 : 1016-25.
5. Shirley CR, Hayashi S, Mounsey S, *et al.* Abnormalities and reduced reproductive potential of sperm from Tnp1- and Tnp2-null double mutant mice. *Biol Reprod* 2004 ; 71 : 1220-9.
6. Zhao M, Shirley CR, Hayashi S, *et al.* Transition nuclear proteins are required for normal chromatin condensation and functional sperm development. *Genesis* 2004 ; 38 : 200-13.
7. Govin J, Escoffier E, Rousseaux S, *et al.* Pericentric heterochromatin reprogramming by new histone variants during mouse spermiogenesis. *J Cell Biol* 2007 ; 176 : 283-94.
8. Barral S, Morozumi Y, Tanaka H, *et al.* Histone variant H2A.L.2 guides transition protein-dependent protamine assembly in male germ cells. *Mol Cell* 2017 ; 66 : 89-101.
9. Goudarzi A, Shiota H, Rousseaux S, *et al.* Genome-scale acetylation-dependent histone eviction during spermatogenesis. *J Mol Biol* 2014 ; 426 : 3342-9.
10. Rousseaux S, Petosa C, Müller CW, *et al.* Du nouveau dans la compréhension de la reprogrammation postméiotique du génome mâle. *Med Sci (Paris)* 2010 ; 26 : 130-2