Histone complement of a rapidly evolving chordate *Oikopleura dioica*:

Developmental and sex-specific deployment of novel and universal histone variants and their posttranslational modifications.

Alexandra Moosmann

Dissertation for the degree philosophiae doctor (PhD)

at the University of Bergen

22.10.2010

Histone complement of a rapidly evolving chordate *Oikopleura dioica*

Developmental and sex-specific deployment of novel and universal histone variants and their posttranslational modifications

by

Alexandra Moosmann

Thesis submitted in partial fulfilment of the requirements for the degree of
Philosophiae Doctor (PhD)

Department of Molecular Biology

University of Bergen

2010

Acknowledgements

Many people have inspired and supported me during the time of my thesis. I would like to thank all of those who made these years special and memorable.

First and foremost, I would like to express my sincere gratitude to my supervisors Dr. Eric Thompson and Prof. Rein Aasland. Thank you Eric, for introducing me into chromatin research, your guidance and friendly support throughout these years and the always open door to your office. Thank you also, for your faith in my abilities as a scientist and for giving me the freedom to develop my own ideas. And finally, for teaching me that science should always be fun!

Thank you very much Rein, for your continuous interest in this project and stimulating discussions on the nucleosome structure.

Furthermore, I thank our collaborators from the Bergen Center for Computational Science Yvan Strahm and Svenn Helge Grindhaug who established the histone data base. I still owe you a lot of German beer!

I would like to thank Prof. Henk Stunnenberg and his group for the interest in this work and for the collaboration on the mass spectrometry analysis. I further thank the people at the Proteomic Unit of the Bergen University and in particular Hilde Garberg for assistance and advice with MASCOT searches.

Lucas Leclère I thank very much for discussion and advise on gene phylogeny.

I am also very grateful to Martina Raasholm, Jean-Marie Bouquet, Karin Zech and Ugur Bal who performed all the microinjections. The work on histone-GFP fusions would not have been possible without your commitment.

Thank`s to all the present and former members of the Thompson group! Fabio Spada and Cristina Chioda, I thank for their help at the start of the project and for introducing me to confocal microscopy. Phillipe Ganot I thank for teaching me the decisive sex-specific differences of *O. dioica* and other species. Furthermore, I would like to thank Julia Hosp for her ever sunny nature and for several late-evening-lab-favours. I am also very grateful to Coen Campsteijn for his input into this work and for evaluation of the manuscript, but also for giving me motivation and support. Thank you for your contagious "there-is-nothing-we-can't-do-attitude", I have learned a lot from you! Especially, I thank Martina Raasholm. You have made yourself irreplaceable as an excellent co-worker and a great friend outside the lab. Your support, friendship and calm advice meant a lot to me and my time in Bergen would not have been the same without you.

All the people at the Sars Centre: Thank you for the great working environment and a lot of memorable Norwegian Christmas parties.

I thank the administrative staff of the Sars Centre, especially Carol Bruce and Doreen Nyhus for their friendly assistance and all the time they have spent on (maternity-leave)- paper work.

Thank`s to all our Bergen friends for a wonderful time in Norway!

Both of my parents I would like to thank for always supporting every road I considered worth pursuing. Thank you for teaching me that the world is a great place that only awaits discovery!

Above all, my special gratitude goes to my husband Matthias. I would not have been able to complete this work without all the time and love you invest into our family life. Thank you also for your indispensible help during the last sprint to the finish and for being my everlasting hero.

Thank you, Liv and Theo for being the nicest distraction from a PhD thesis I can imagine!

Table of Contents

Abstract

The packaging of DNA into nucleosomes is a fundamentally conserved property of the eukaryotic nucleus which is evident in the conservation of histone sequences. Nevertheless, it is now clear that histone sequence variants have diversified in many species to assume crucial roles in the regulation of gene expression, DNA repair, chromosome segregation and other processes. While considerable data exist on coding sequences of histones and some selected histone variants in a wide variety of organisms, the information available on total histone gene complements is much more limited. *Oikopleura dioica* (Od) is a dioecious marine urochordate that occupies a key phylogenetic position near the invertebrate-vertebrate transition with the smallest genome ever found in a chordate (70 Mb). Its short life cycle is characterized by a developmental switch between mitotic and endocycling cells, making *O. dioica* an attractive model to study the spatial and temporal use of histone variants and posttranslational histone modifications (PTMs) throughout development and in different cell cycle types. We have characterized the complete histone gene complement and the developmental expression of histone genes present in the first assembly of the *O. dioica* draft genome and identified the major Od PTMs by mass-spectrometric analysis. Furthermore, we analyzed the dynamics and distribution of phosphorylated H3 variants during mitosis and meiosis of *O. dioica* and the deposition of the centromeric variant OdCenH3 in mitotic and endocycling cells with respect to centromeric PTMs. The Od histone gene complement displays several features not known from other chordates, including male-specific variants in all of the core histone families, N-terminal H2A.Z splice variants, and a diverse array of H2A variants but absence of the near universal variant H2AX. The results here suggest significant plasticity in histone gene organization, the variation within histone families and the chromosomal distribution of mitotic PTMs within the chordate lineage. This further supports the view that histone gene complements may also evolve adaptively to the specific life history traits, cell cycle regulation and genome architecture of organisms.

1. Introduction

The term “gene” has engaged scientists and philosophers of the past and current century. Genes seem to be the biological legacy we inherit from our parents, a blueprint that determines the identity of each organism. Ten years after the completion of the draft sequence of the human genome, it has become evident that the DNA sequence of chromosomes does not completely reflect the genetic complexity of an organism. Different cell types within one organism all contain the same DNA, but it is the different store of information for the initiation of gene activity that determines their respective differentiation pathways. To preserve this identity, the pattern of gene expression must be maintained and inherited throughout generations of cells. The mechanisms that control gene activity without changing the underlying DNA sequence are generally referred to as “epigenetic” gene regulations.

Cellular DNA of eukaryotes is not “naked” but resides in the cell nucleus in a “chromatinized” form, packed around small histone proteins. Histones aid to compact the large amounts of DNA into chromatin fibers and eventually into chromosomes that fit into the nucleus. In addition to its packing function, chromatin provides an important mechanism that controls DNA access. Nucleosomes are the fundamental units of chromatin; octameric histone complexes consisting of four different types of core histone proteins (H2A, H2B, H3 and H4) and the DNA packaged around them. This structure is stabilized by a fifth type of histone, the linker histone H1 (Fig. 1). While packaging into chromatin is necessary for efficient and accurate segregation of daughter cells during each cell cycle, access to DNA must also be allowed in order for transcription, replication, repair and recombination to take place. Thus, alterations of the nucleosome structure and the chromatin fiber influence these processes by making DNA more or less accessible to nuclear factors. Chromatin based inheritance participates during development when genome information must be used selectively to shape a highly complex organism. One mechanism to “remodel” chromatin is via histone alterations, generated through sequence variants or via the chemical modification of histones. Histone variants incorporated into the nucleosome can confer specific properties to the chromatin fiber that determine a more or less

accessible structure to the DNA interacting factors. Furthermore, the activity of factors involved in nuclear processes can be modulated through their interaction with covalently modified histone tails that protrude from the nucleosome. Certain histone modifications signal gene activation while others serve as a signal for the inactivation of genes. Interestingly, just like some histone variants, many histone modifications are also evolutionarily conserved across species, suggesting organisms share the fundamental mechanisms of gene regulation through chromatin dynamics.

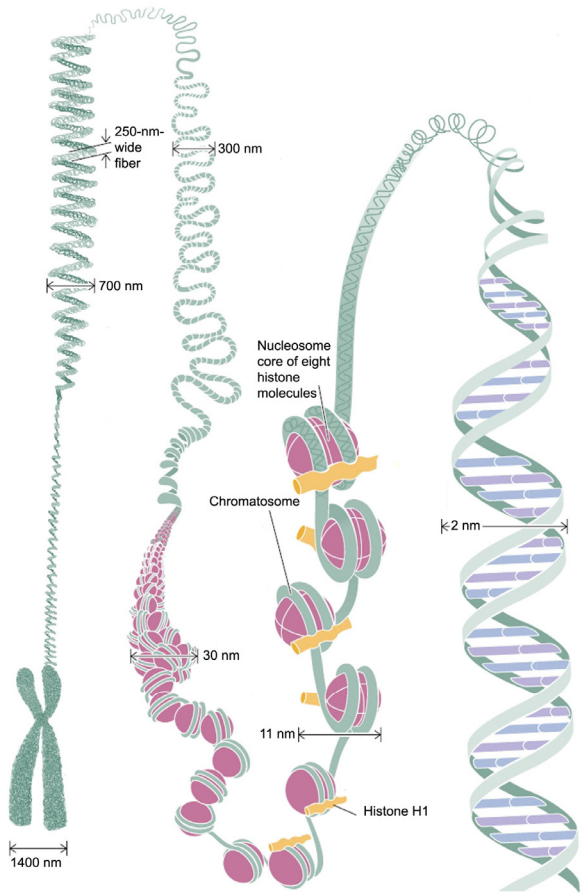

Figure 1 Chromatin structure and levels of organization. The double helix of DNA is wrapped around the nucleosome core consisting of eight histone proteins [$(H3\text{-}H4)_2$ $(H2A\text{-}H2B)_2$)]. With the binding of the linker histone H1, this results in a "beads-on-a-string" structure of 11 nm which condenses further into a helical 30 nm fiber. Association with the nuclear scaffold will further package the chromatin into loops and a structure of 300-700 nm. During the metaphase of the cell cycle, chromatin strands become even more condensed and form metaphase chromosomes, leading to ~500 fold compaction of the DNA. Modified from (Pray, 2008).

1.1. Packaging DNA into the nucleus – basic chromatin elements

All eukaryotic organisms face the problem that they must compact long DNA molecules more than 10 000 fold into a small nucleus, on the order of 10 μm in size. The packaging of DNA into chromatin and chromosomes is a solution that is conserved throughout the animal and plant kingdoms. Likewise, histones are some of the most conserved proteins known with some members of different species being virtually identical in amino acid sequence.

1.1.1. The Histones

Histones were first purified by Albrecht Kossel from bird erythrocytes and sperm in 1884 (Turner, 2001). As they were further characterized during the last century, five histone families were distinguished which are now universally designated H1, H2A, H2B, H3 and H4. Histones are small and highly positively charged proteins spanning a molecular weight range of 10-32 kDa. High content of positive lysine and arginine residues make histones excellent DNA binding proteins and contribute to their tight interaction with negatively charged DNA (Turner, 2001).

Among the five histone families, H2A, H2B, H3 and H4 are the core histones, characterized by a structural domain known as the “histone fold”. This domain comprises a sequence of three alpha helices, one long and two short, with two non-helical loops separating them (Arents et al., 1991). The histone fold domains are structurally important motifs for nucleosome formation. Through the histone fold, core histones interact among themselves and with the nucleosomal DNA (Fig. 2). Histone fold domains further confer interactions with other nuclear factors and nucleosome-nucleosome contacts which explains their extreme degree of amino acid sequence conservation. In contrast to the histone fold domain, the N-terminal and C-terminal tails of histones are more variable in sequence and lengths, depending on the type of histone. They protrude from the nucleosome, are highly flexible and also contain sites for different histone posttranslational modifications. The four H3 and H2B N-terminal tails exit through the minor groove of the DNA superhelix and

contribute to chromatin compaction by attaching to the entering and exiting linker DNA (Hill and Thomas, 1990; Kan et al., 2007; Luger and Richmond, 1998). Histone H2A is unique among the core histones in having both an N-and a C-terminal basic tail (Fig. 2). The H2A C-terminal tail binds the DNA around the dyad axis (Usachenko et al., 1994) whereas the N-terminal tails of H2A contact DNA towards the periphery of the nucleosome (Pruss and Wolffe, 1993),. Furthermore, histone tails play important roles in the assembly of higher order structures and individual histone tails affect the process of chromosome assembly with different efficiencies (de la Barre et al., 2000; Dorigo et al., 2003; Hansen et al., 1998). Among all histone tails, H4 tails mediate the most internucleosomal interactions, followed by the H3, H2A, and H2B tails in decreasing order (Arya and Schlick, 2006). Nevertheless, the role of histone tails remains somewhat enigmatic, given that they do not yield clear electron density maps.

Linker histones are another class of histones that are important for further compaction into the 30 nm fiber and higher order chromatin structures. Structurally, metazoan H1s are divided into three domains: A short, flexible N-terminal, a globular domain containing a winged-helix fold and a long, extremely lysine rich C-terminal tail (Allan et al., 1980).

1.1.2. The first level of folding – nucleosome structure

Octameric nucleosomes form the fundamental unit of chromatin with essentially the same basic structure across different organisms. They exist in all eukaryotes, except dinoflagellates that appear to have lost their bulk histones (Hackett et al., 2005). In the core particle, 147 bp of DNA are wrapped in 1.7 left-handed superhelical turns around the histone octamer consisting of 2 H2A-H2B dimers flanking a H3-H4 tetramer (Luger et al., 1997).

A structurally important motif for the formation of the $(H3\text{-}H4)_2$ $(H2A\text{-}H2B)_2$ octamer are the histone fold domains. They interlock with each other through hydrophobic residues and give rise to another important structure, the “histone handshake” (Arents et al., 1991) (Fig. 2). Assembly of the canonical eukaryotic

nucleosome is thought to begin with an $(H3–H4)_2$ tetramer, held together by a strong 4-helix bundle between the two H3 molecules, consisting of salt bridges and hydrogen bonds (Luger et al., 1997) (Fig. 2). The addition of two H2A–H2B dimers forms an octamer with a left-handed helical ramp that subsequently locks the DNA into a negatively supercoiled configuration. The H2A–H2B dimers also interact with the $(H3–H4)_2$ tetramer through a weak 4-helix bundle between the α2 and α3 helices of H4 and H2B (Fig.2). These contacts involve the docking domain of H2A which interacts with the C-terminal domain in H4 and parts of H3 on one side and a four-helix bundle between the histone fold regions of H4 and H2B. However, the interactions linking H2A-H2B dimers to the $(H3-H4)_2$ tetramer cannot persist in the absence of DNA, and the histone octamer is not stable under physiological ionic conditions

1.1.3. The chromatin fiber and higher order folding

The next step in the packaging of DNA is the binding of the linker histones to the linker DNA between the nucleosomes that will make up the "chromatosome" (Fig. 1). The globular domain of the linker histone binds at the entry-exit site of the nucleosome particle, whereas specific residues of the positively charged C-terminal portion bind both linker DNAs, bringing them together into the so-called "stem" structure (Syed et al., 2010). This interaction further tightens the association of the DNA and the nucleosome and produces a more defined angle of DNA entry and exit with folding to a "beads-on-a-string" structure of 11 nm observable by electron microscopy (Thoma et al., 1979). A current view is that these polynucleosomal arrays are further stabilized by internucleosomal histone interactions and the association of additional proteins including linker H1 histones to condense into a helical 30 nm fiber, which would result in a 40-fold compaction of the linear DNA. The two main models proposed for the 30nm structure are the solenoid (one-start) and the zig-zag (two start) model that depend on the length of the linker DNA. However, the 30 nm fiber has not been crystallized and its very existence is still debated. Alternatively it has been proposed that the nucleosome fiber exists in a disordered state that undergoes dynamic movement to control the accessibility of DNA (Maeshima et al.,

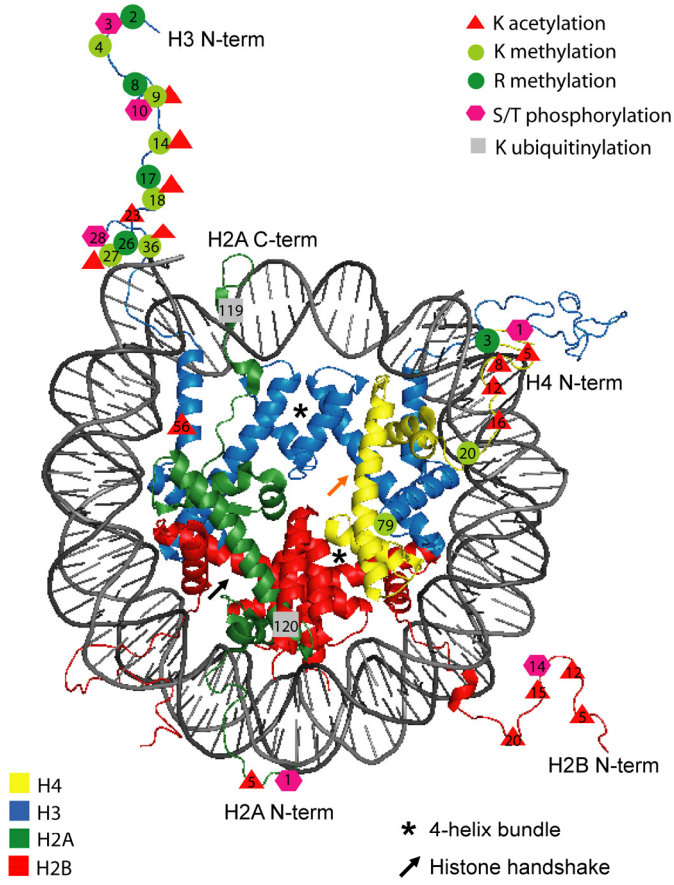

Figure 2 Structure of the nucleosome core particle and histone modifications. A Pymol generated ribbon diagram of an X-ray crystal structure of the *Xenopus laevis* NCP (PDB: 1kx) showing the DNA helix (grey) around the histone core and histone-histone interactions within the core. For easier view, only one of the two H2A (green) and H2B molecules are shown. The histone folds of H2A and H2B and H3 (blue) and H4 (yellow) interact with each other through the "histone handshake motif" (indicated with arrows). The two H3-H4 dimers interact through a 4-helix bundle (asterisks) formed only by the H3 molecules to form the H3-H4 tetramer and each pair of H2A-H2B molecules interacts with the tetramer through a 4-helix bundle between H2B and H4. The major methylation-, acetylation, ubiquitinylation and phosphorylation marks found on the N-terminal tail and histone core of H3, the N-terminal tail of H4 and H2B and the N- and C-terminal tail of H2A are shown as indicated in the legend.

2010). To achieve additional packaging of the chromatin fiber into chromosomes, a current model proposes that the chromatin fiber form loops of 300-700 nm dimensions held together by a proteinaceous structure referred to as the nuclear scaffold (Cremer et al., 1993; Kantidze and Razin, 2009).

1.2. Epigenetic information and histone modifications

The term epigenetics, was originally derived from 'epigenesis' (Gr. *epi*, on top of; *genesis, beginning),* a general theory first expressed by Aristotle to describe the gradual changes during the development of animals [Aristotle's '*On the Generation of Animals*', cited in Jablonka and Lamb (Jablonka and Lamb, 2002)]. The meaning of the word has gradually narrowed and today epigenetics is generally accepted as ''the study of changes in gene function that are mitotically and/or meiotically heritable and that do not entail a change in DNA sequence.'' (Russo et al. 1996). Epigenetic modifications considered crucial for regulating chromatin structure comprise histone variants, posttranslational modifications of histones, and covalent modifications of DNA bases.

The predominant modification in mammalian DNA is methylation of cytosine in a CpG nucleotide context, converting cytosine to 5-methylcytosine by a DNA methyltransferase. Because the promoter regions of silenced genes possess significantly more methylated cytosines in comparison with actively transcribed genes, a current view is that DNA methylation causes gene inactivation. Methylation of cytosine may prevent the binding of specific transcription factors or attract mediators of chromatin remodeling, such as histone-modifying enzymes or other repressors of gene expression (Klose and Bird, 2006).

1.2.1. Posttranslational modifications of histones and the histone code

Histones are subject to a variety of posttranslational modifications (PTMs) that can alter gene expression and chromatin structure. Histone modifications can distinguish

large domains in heterochromatin and euchromatin, sometimes in conjunction with histone variants, and may have an 'indexing' function involved in large scale genome organization. According to the "histone code hypothesis", single or combined marks on histones store and transmit information on the gene expression status through mitosis and subsequent cell generations (Strahl and Allis, 2000). Nevertheless, histone modifications can also directly influence replication, transcription and cell cycle regulation in a transient manner which does not contribute to cellular memory. Hence, more recently it was suggested to restrict the term "histone code" to PTM marks that exert a heritable and therefore epigenetic function (Turner, 2007).

Residues of the N-termini of histones H3 and H4 and the amino and C-termini of histones H2A, H2B and H1, are particularly subject to a variety of PTMs (Fig. 2). Phosphorylation; acetylation and methylation have been studied extensively (Peterson and Laniel, 2004), while relatively little is known about ubiquitination; sumoylation; ADP ribosylation; glycosylation; biotinylation, carbonylation (Bartova et al., 2008; Bergmann, 2010; Fujiki et al., 2010; Garcia-Dominguez and Reyes, 2009; Hassan and Zempleni, 2008) and lysine butyrylation and propionylation (Chen et al., 2007). Here we focus on the first three types of PTMs.

1.2.1.1. Histone acetylation

Acetyltransferases catalyze the addition of acetyl groups to either the ε- (side chain acetylation) or the α-amino group of specific lysines in histones and other proteins (Turner, 2001). Histone ε-acetylation is normally associated with transcription and euchromatin. Acetylations of several lysine residues within the N-termini of H3 (K9, K14, K18, K23) and H4 (K5, K8, K12, K16) are hallmarks of transcription and localize to the promoters of active genes in yeast, *Drosophila*, human and mouse (Rando, 2007). Not surprising, histone acetylation is a highly dynamic modification and the removal of acetyl groups is catalyzed by deacetylases, some of which are transcriptional repressors (Taunton et al., 1996). The role of histone acetylation in transcriptional activation has been explained by two different but not mutually exclusive models. In the first, acetylation of core histone lysine residues directly inhibits chromatin condensation by neutralizing the positive charge of lysine residues

and thereby disrupting interaction with the negatively charged DNA backbone. Consequently, relaxation of chromatin makes the transcriptional site available to the transcription machinery (Choi and Howe, 2009). In the second model, covalently modified core histone amino termini serve as a signal for the binding of trans-acting factors. Effector proteins bind acetylated histones through specific bromodomain-motifs and enhance the binding of the RNA polymerase complex and related factors. They may also recruit further structure remodeling or chemical modification activities (Ruthenburg et al., 2007). For instance, loss of deacetylases in yeast results in the transcription initiation of cryptic promoters, supporting the view that histone acetylation enhances access to DNA (Carrozza et al., 2005). Furthermore, studies on H4K16 acetylation suggest that histone acetylation may disrupt intranucleosomal interactions and regulate compaction of the 30 nm chromatin fiber (Shogren-Knaak et al., 2006).

1.2.1.2. Histone methylation

Histone methylation occurs on the nitrogens in arginine and lysine side chains through the activity of methyltransferases that catalyze the addition of a methyl group from S-adenosyl methionine (SAM). Arginine residues can be modified by one or two methyl groups in either a symmetric or asymmetric conformation whereas lysine residues can be mono-, di- or trimethylated. Arginine methylation has been associated with transcriptional activation whereas lysine methylation has been linked to both activation and repression. For instance, methylated H3K4, H3K36 and H3K79 are considered to be marks for transcriptionally potentiated chromatin structures while methylated H3K9, H3K27 and H4K20 mark silent chromatin (Peters and Schubeler, 2005).

Histone methylation and in particular trimethylation, was long regarded as irreversible because of the high thermodynamic stability of the N–CH3 bond. The recent identification of several demethylases showed that methylation can also be reversed without the exchange of bulk histones (Cloos et al., 2008). Although methylation of lysine and arginine residues does not influence the net charge of the affected residues, it increases the hydrophobicity and the steric bulk of the amino acid

side chain (Honda et al., 1975). Indeed, it was recently shown that trimethylation of H4 at lysine 20 enhanced the ability of nucleosomal arrays to fold and condense *in vitro*, demonstrating that histone methylation can indeed affect higher order chromatin structure directly (Lu et al., 2008). Nevertheless, it is now clear that histone lysine methylation also creates binding sites that can alter the local properties of chromatin for transcription. Methylated lysines recruit proteins through binding to small domains such as PhD fingers and chromodomains, whereas methylated arginines recruit tudordomains (Bannister et al., 2001; Boisvert et al., 2005; Shi et al., 2007). H3 trimethylated at Lys 9 for example recruits HP1 (heterochromatin protein 1) that binds H3K9me_3 directly through its chromodomain. HP1 mediates silencing through the interaction with the H3K9-specific methyltranserase, which then in turn cooperates with histone deacetylaeses to establish long term transcriptional repression (Fischle et al., 2005).

1.2.1.3. Histone phosphorylation

In vivo, proteins are phosphorylated at the side chains of serine, threonine and, rarely, tyrosine, residues by substituting a phosphate for a hydroxyl group to give an O-phosphate linkage. Phosphorylation is enacted by protein kinases using nucleotide triphosphates (ATP, GTP, cyclic AMP) as PO_4 donors (Turner, 2001), and reversed by phosphatases. Phosphorylation plays important roles in a wide range of cellular processes such as enzyme activation and inhibition and protein degradation. In histones, posttranslational phosphorylation of linker H1 histones and histone H3 have been most extensively studied. Both types of histone phosphorylations are prevalent during mitosis.

Phosphorylation is the most extensively studied modification of linker histones but its exact role in chromatin structure remains controversial. As it peaks during the G2 cell cycle phase and mitosis it seems to invoke condensation of chromatin. Contradictorily, it is believed to control chromatin decondensation and DNA replication during S-phase (Baatout and Derradji, 2006). Core histones H3 variants have conserved residues within their N-terminal tail, namely Thr3, Ser10, Thr11 and Ser28 which are phosphorylated in a wide range of organisms during mitosis. While

Ser10 and Ser28 phosphorylation (H3S10P and H3S28P) appear to have an additional role in the transcriptional activation of genes in interphase nuclei, the other residues have thus far only been described as a mitotic mark. However, across different species there seems to be neither a consistency in the spatial distributions of these phosphorylation marks nor in their mechanistic function. While in some organisms, specific H3 residues are phosphorylated at centromeres, the same phosphorylation signal is excluded from the centromeric site in others (Perez-Cadahia et al., 2009). Ser10 phosphorylation has been linked to chromosomes condensation in *Tetrahymena* (Wei et al., 1999) but it dispensable for it in yeast (Hsu et al., 2000).

Possibly there is little functional conservation of the specific phosphorylated residues among different species. A recently proposed idea is that the connotation of a specific H3 phosphorylation mark may only become meaningful in combination with other histone modifications which would infer a lineage-specific "histone code" (Cerutti and Casas-Mollano, 2009). A prominent example of the interplay of H3S10P with other modifications is "cross talk" with the di- and trimethylation of H3K9. H3S10P has been shown to promote dissociation of H3K9me_3-bound HP1 from chromosomes without an alteration in H3K9me_3 levels as cells enter mitosis (Fischle et al., 2005; Hirota et al., 2005). The ejection of HP1 at these sites may allow the dynamic rearrangements of chromatin higher-order structure required for mitotic chromosome condensation.

1.2.2. The writers and erasers of histone modification marks

As mentioned above, histone modifying enzymes catalyze the addition or removal of an array of covalent modifications of histones. These modifications need to be set and erased in a dynamic fashion for efficient switching between gene expression stages and different chromatin structures. An overview of several histone modifying and de-modifying enzymes and their modes of action is given in Fig. 3. A table of the modifying and demodifying enzymes identified in the genome of *Oikopleura dioica,* is provided in the appendix (appendix, table A1A and A1B).

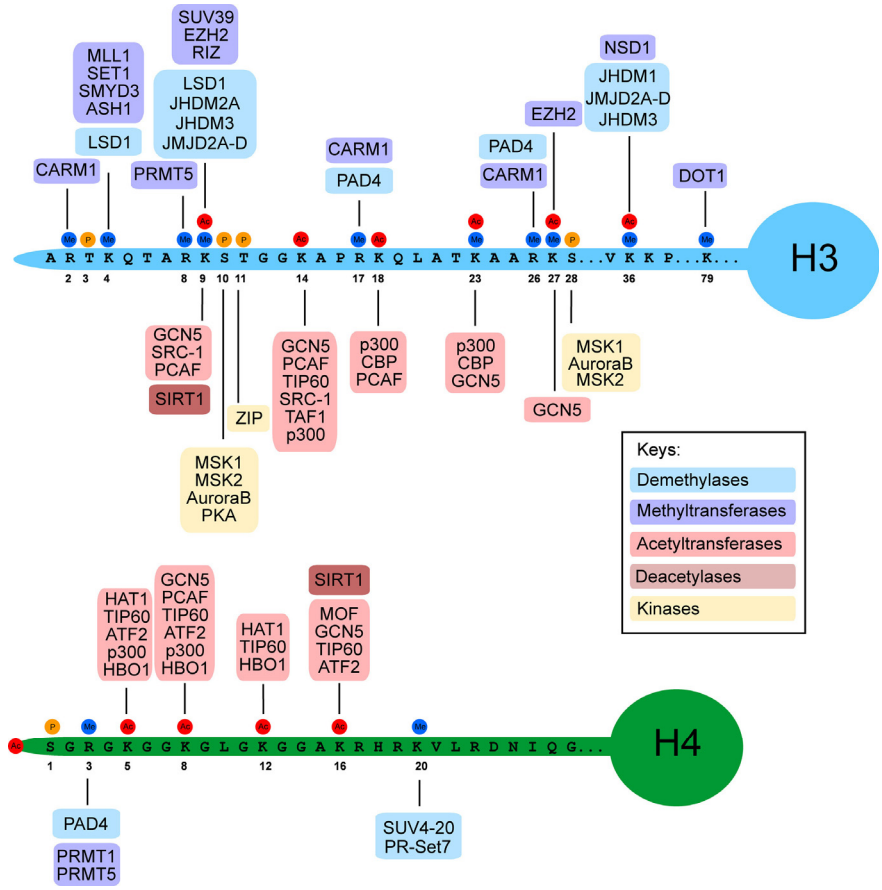

Figure 3 Selection of modifying and demodifying enzymes acting on H3 and H4. A selected panel of H3 and H4 modifying and demodifying enzymes is shown. Methyltransferases and demethylases (see enzyme keys) add and remove histone methylation marks on lysines and arginines (blue) repectively, while acetyltransferases and deacetylases add and remove acetylations from lysines (Siegel et al.). Kinases phosphorylate histones at serines, threonines and tyrosines (yellow) but not much is known about histone phosphatases, the enzymes responsible for histone dephosphorylation. Modified from (Bjerga, 2009).

1.2.2.1. Acetyltransferases and deacetylates

Lysine acetyltransferases (KATs) catalyze the addition of acetyl groups from the co-substrate Ac-CoA (Lee et al., 2007) and are evolutionary conserved across the animal kingdom from yeast to human. The first KAT identified was the N-acetyltransferase Gcn5, cloned from yeast and *Tetrahymena* (Brownell et al., 1996; Kleff et al., 1995).

Today, five families of acetyltransferases are known, the GCN5 class (GNATs), the CBP/p300-family , the MYST proteins, TAF1 family and the nuclear receptor co-activator family (Marmorstein and Roth, 2001). Structurally, KAT proteins share a structurally conserved central core region for Ac-CoA cofactor binding despite a significant divergence in sequence between the different KAT families (Marmorstein and Trievel, 2009). Most KATs have broad substrate specificity and all families acetylate H3. While the GNAT family members act primarily on H3 and H4, CBP/p300 proteins are able to acetylate all other core histones. MYST enzymes such as Tip 60, MOF, HBO and MOZ are known to act on H3, H4 and H2A (Roth et al., 2001).

The "erasers" of acetylation marks are the Histone deacetylases (HDACs) which are divided into two superfamilies, the classical HDACs and the sirtuins (de Ruijter et al., 2003). Remarkably, none of the deacetylases contain "reader domains" such as bromodomains, PHD finger, tudordomains or chromodomains and the sequence similarity among family members is very low. However, HDACs have been found in complexes with reader domain-containing proteins such as the nucleosome remodeling and deacetylase complex (NuRD) (Cunliffe, 2008).

1.2.2.2. Methyltransferases and demethylases

Methyltransferases catalyze the transfer of methyl groups from S-adenosyl methionine to lysines (KMT) or arginines (PRMT) in histones and exist in all eukaryotes studied to date. They are divided into eight distinct families: The TRX, EZ, SUV3-9, ASH1, SET8, SUV4-20, SMYD, RMT, SET7/9 and PRDM subfamilies. All family members contain a catalytic SET domain, which was first recognized as a conservative sequence in the three *Drosophila* genes, suppressor of position effect variegation (Suv39) (Tschiersch et al., 1994), the Polycomb-group chromatin regulator Enhancer of zeste (E(z)) (Jones and Gelbart, 1993) and the trithorax-group chromatin regulator trithorax (Trx) (Akasaka et al., 1996). The exception to the rule is the DOT1 family. Their members contain no SET domain and are structurally unrelated to SET-domain proteins (Dillon et al., 2005). The Dot1 methyltransferase is an evolutionarily conserved protein that was originally identified

in *S. cerevisiae* as a disruptor of telomeric silencing (Singer et al., 1998). Dot1 specifically methylates Lys79 in the globular region of H3 and has also been shown to play an important role in heterochromatin formation and the embryonic development of mammals (Jones et al., 2008). The remarkable substrate specificity of KMT proteins is probably achieved through the recognition of the histone residues flanking the methylated lysine residue (Park et al., 2002; Qian and Zhou, 2006). Two more substrate-specific methyltransferases are the KMT proteins Suv4-20 and SET8/PR-Set7 which both methylate the N-terminal tail of H4 at lysine 20 (H4K20). While in yeast a single enzyme Set9, mediates all stages of H4K20 methylation (Sanders et al., 2004) two enzymes, PR-Set7 and Suv4-20, exist in human that are responsible for monomethylation and di-or trimethylation of H4K20 respectively ($H4K20me_{2/3}$) (Yang et al., 2008). The strict monomethylation product specificity of PR-Set7 is thought to result from the presence of a tyrosine residue in the active PR-Set7 site which may hydrogen bond to the monomethyl-lysine product and inhibit further methylation (Collins et al., 2005; Couture et al., 2005). While H4K20 methylation in general has been associated with heterochromatin, the subcellular localization of the three different H4K20 methylation stages and their genomic distribution differs, suggesting that they are functionally distinct (Yang et al., 2008).

It is now clear that histone methylation is also reversible. The removal of methyl marks is catalyzed by demethylases which are grouped into three major families, the LSD1 family, the Jumonji C familiy (JMJC) (Mosammaparast and Shi, 2010) and the PAD familiy (Wang et al., 2004). Several lysine demethylases (KDMs) identified so far are also known to be substrate-specific including the two JMJC familiy members JHDM1 that demethylates H3K36 and JHDM1 that demethylates H3K9 (Tsukada et al., 2006; Yamane et al., 2006).

1.2.2.3. Kinases and phosphatases regulate histone phosphorylation

Protein kinases, the enzymes that catalyze the linkage of a phosphate group to the side chains of serine, threonine or tyrosine residues, belong to a very large and divergent family of enzymes, constituting approximately 2% of the human proteome

(Johnson and Hunter, 2005). Eukaryotic cells contain hundreds of different protein kinases that are organized into a complex network of signaling pathways to coordinate cellular activities such as cell cycle regulation, cell movement, transcription, apoptosis and DNA repair. The 518 kinases identified in human are divided into 18 major groups consisting of more than 90 subfamilies (Manning et al., 2002). However, most protein kinases belong to a single superfamily containing a eukaryotic protein kinase (ePK) catalytic domain.

Linker histone H1 phosphorylation is attributable to the action of growth-associated, cyclin dependant kinases (CDKs) that phosphorylate H1s in a cell cycle regulated fashion (Hale et al., 2006).

The phosphoinositide 3-kinase related kinases (PIKKs) comprise a family of proteins that play central roles in stress-induced signaling pathways. Upon DNA damage, the PIKK kinases DNA-PK, ATM and Rad3 related (ATR) and ataxia–telangiectasia mutated (ATM) phosphorylate the histone variant H2AX to γH2AX, which is one of the earliest marks of DNA damage. DNA repair-related proteins subsequently congregate at the γH2AX foci during the repair process and γH2AX is an important element in proper damage response foci formation by enhancing the retention of repair factors after their initial recruitment (Celeste et al., 2003).

The phosphorylation of the four characteristic phosphor-residues within the N-terminal tail of H3, Thr3, Ser10, Thr11 and Ser28 is tightly regulated by site-specific kinases. Several studies have identified the members of the Aurora kinase family Aurora A and Aurora B as the enzymes responsible for the phosphorylation of Ser10 and Ser28 during the mitotic phase of the cell cycle. While *in vitro* results from *Xenopus* suggest that Aurora A and Aurora B both phosphorylate Ser10 and Ser28 (Murnion et al., 2001), data from human cells implicate Aurora B as a potential mitotic Ser10 and Ser28 kinase (Prigent and Dimitrov, 2003). Mammalian Aurora kinases are normally regulated in a cell cycle dependant manner with peak expression profile at the gap 2 (G2)-mitosis transition phase (Kimura et al., 1999). Interestingly, overexpression of Aurora kinases has been observed in many human cancer cell lines,

suggesting an important role for the Aurora enzymes in the regulation of cell proliferation. The protein kinase responsible for the H3 phosphorylation at Thr3 has only recently identified in human cells as the haploid germ cell specific nuclear protein kinase (haspin) (Dai et al., 2006). Haspin phosphorylates histone H3 at Thr3 during mitosis, particularly at the inner centromeres and appears to play a role in chromatid cohesion (Higgins, 2010). In most organisms the kinase responsible for Thr3 phosphorylation has not been identified but genes encoding haspin homologs are present in all major eukaryotic phyla (Higgins, 2003). This suggests that this kinase may be involved in essential functions for eukaryotes. Mitotic H3 phosphorylation of Thr11 is regulated by the death-associated protein-like (DAP) kinase although the role for this phosphorylation mark is not clear (Preuss et al., 2003). Thr11 phosphorylation also occurs in plants but the responsible kinase still remains to be identified.

After chromosome segregation all mitotic phosphorylation marks on H3 decrease and mitotic H3 kinase activity is reversed by specific phosphatases such as PP1, a type 1 protein phosphatase (Hsu et al., 2000). In mammalian cells, PP1 dephosphorylates H3 at Ser10 and Ser28, whereby the Ser28 residue seems more sensitive to PP1 activity (Goto et al., 2002). Aurora B is activated upon PP1 inhibition and human Aurora B kinase forms a complex with PP1 to regulate the spatiotemporal features of H3 phosphorylation during mitosis (Sugiyama et al., 2002).

1.3. Replicating epigenetic histone marks

An increasing body of data has shown that histone modifications and some histone variants such as H2A.Z and CenH3 can serve as epigenetic memory marks (Bernad et al., 2009; Brickner et al., 2007). Nevertheless, a prerequisite for histone based information to serve an epigenetic function is the recycling of modified histones and histone variants, so they may serve as a blueprint for the newly synthesized histones through cell divisions. Consequently, one of the key questions in chromatin research is how histones are deposited and transmitted to the two daughter strands during DNA replication.

Evidence that the pre-existing nucleosomes ahead of the replication fork are disrupted and segregate to the daughter strands came from *in vivo* density and radiolabeling experiments revealing that "hybrid nucleosomes" exist on nascent DNA and are mainly constituted of an old $(H3\text{-}H4)_2$ tetramer with one new and one old H2A-H2B dimer or a new H3-H4 tetramer with two old H2A-H2B dimers (Annunziato et al., 1982; Jackson and Chalkley, 1981a, b). The assembly of both parental and newly synthesized histones occurs without particular preference for either the leading or lagging DNA strand and occurs in a semi-conservative manner (Fig. 4) (Jackson, 1988; Sogo et al., 1986). Studies investigating the fate of new histones and the exchange of histone-GFP fusions in living cells showed that H2A-H2B and H1 have a high exchange rate during replication, but the majority of H3 and H4 remain permanently bound (Kimura and Cook, 2001; Misteli et al., 2000). Due to the less dynamic behavior of H3-H4, the research of the past years has been focused on the role of H3-H4 as putative vehicles for histone marks. The recent discovery that the initial substrate for H3-H4 *de novo* assembly are H3-H4 dimers and not tetramers (Benson et al., 2006; Tagami et al., 2004) further spurred the idea that "mixed" $(H3\text{-}H4)_2$ tetramers consisting of new and old H3-H4 dimers could provide a mechanism to transfer information via H3 and H4 histones onto daughter strands. Recent evidence suggests the chaperone anti silencing function 1 (Asf1) may be involved in tetramer splitting and the management of parental histones during replication. Histones in the Asf1-H3/H4-MCM (Mini chromosome maintenance) helicase complex display a parental signature of posttranslational modifications, suggesting that the histones bound to Asf1 have been removed from the parental chromatin (Groth et al., 2007). The existence of mixed tetramers *in vivo*, however, has only very recently been confirmed in human cell lines (Xu et al., 2010). Here, a combined approach of SILAC (Stable Isotope Labeling with Amino acids in Cell culture)-labeled and flag-tagged histones was employed. Histones were purified and subsequently analyzed by quantitative mass-spectrometry to distinguish old from new histones. Intriguingly, significant amounts of tetramers containing the variant H3.3 were split during replication, while tetramers consisting of canonical H3.1 histones were maintained, suggesting that both types of parental histone segregation exist *in*

vivo. The results support the idea that the H3.1 modifications of large heterochromatic regions are maintained by copying them from the neighboring, pre-existing nucleosome. Furthermore, it raises the interesting possibility that H3-H4 tetramer splitting, if it occurs, may depend on the type of histone variant or the chromatin region.

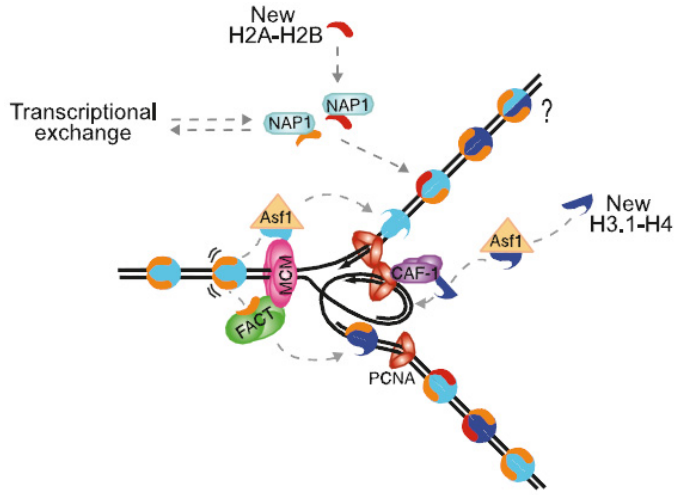

Figure 4 Model of the nucleosome disruption and restoration at the replication fork. The MCM helicase complex mediates unwinding of the DNA as part of the replication fork progression complex. Old H2A-H2B dimers are removed by the histone chaperone complex FACT, allowing binding of the chaperone Asf1 that triggers the disruption of the H3-H4 tetramer. On nascent DNA, nucleosome assembly occurs in a stepwise fashion, with the addition of $(H3\text{-}H4)_2$ by the chaperone CAF-1 via recruitment to PCNA and the addition of two H2A-H2B dimers by the chaperone NAP1. The question mark indicates that the semi-conservative distribution of parental and new histones may result in mixed $(H3\text{-}H4)_2$ containing one old and new H3-H4 dimer or either new or old $(H3\text{-}H4)_2$. From (Groth, 2009).

1.4. Histone variants

Specific properties of histone proteins can alter underlying nucleosome dynamics. In addition to histone modifications, this includes variation in the amino acid sequence

of the conserved histone proteins. With the exception of histone H4, several sequence variants of each histone family have been reported for most organisms. These histone variants are able to replace the canonical histones, and alter the composition of individual nucleosomes, whereby the term “histone variant” usually refers to non-allelic sequence variants. Histone variants can be classified according to their expression patterns as either replication dependent (RD or canonical histones) or replication independent (RI or replacement) variants that replace the major histones outside of S-phase. RD histone genes are intron-less and contain a stem-loop (Black et al.) structure within the 3`UTR of the (RD) histone mRNA instead of a polyadenylation signal (Marzluff and Duronio, 2002). In contrast, replacement variants are usually encoded by single orphan genes that contain introns. They are expressed at a constant level throughout the cell cycle and transcribed through polyadenylated mRNAs (Malik and Henikoff, 2003).

1.4.1. Histone variant families

1.4.1.1. Linker histone variants

H1 variants form a complex family of related proteins with distinct species, tissue and developmental specificity. Multiple linker histone family members are expressed in animals and plants and up to 11 linker histone subtypes have been identified in human and mouse. Due to the heterogeneity of this histone family, different H1 subtypes are usually classified by their mode of expression, i.e. as being S-phase dependant (H1.1-H1.5 in human) or replacement linker histones (H1x and H1.0 in human). H1x is ubiquitously expressed while H1.0 exists mainly in terminally differentiated cells (Happel and Doenecke, 2009). Additionally, vertebrates usually have germ cell-specific H1s, such as the testes-specific variants H1t, H1T2, Hils1 and the oocyte-specific H1Foo of human. Knock out studies suggest, that individual H1s might be partially redundant, at least in their ability to compact chromatin globally (Izzo et al., 2008). Nevertheless, H1 variants differ in their biophysical properties, their association with repressed or active chromatin and their ability to increase or decrease transcription when overexpressed. Beyond their function in gene regulation,

H1 variants may also be implicated in other biological processes such as DNA repair (Hashimoto et al., 2007).

1.4.1.2. H3 variants

Because of the strong 4-helix bundle between the two H3 molecules, H3 variants have a key structural role in organizing the nucleosome and are extremely conserved in their primary sequence. Several universal H3 variants have been identified in metazoans that exhibit distinct roles and modes of assembly into chromatin. The replication-dependent H3.1 and H3.2 differ by only a single amino acid change at position 96 where serine is replaced by cysteine in H3.1. H3.1 is mammalian-specific but the function of Cys96 in H3.1 is not known. It has been proposed that Cys96 may form intermolecular disulfide bonds and thereby facilitate the formation of heterochromatin (Hake and Allis, 2006). While H3.1 and H3.2 are incorporated through bulk chromatin assembly, the replacement variant H3.3 can undergo RD or RI assembly and deposition occurs primarily in the transcribed regions of euchromatin. Additionally, H3.3 has shown to contain marks associated with transcriptional activity (Johnson et al., 2004; McKittrick et al., 2004). Four, highly conserved amino acid changes distinguish H3.3 from H3.1/H3.2. Three of these substitutions lie within the alpha 2 helix and exclude H3.2 from RI deposition in *Drosophila* (Ahmad and Henikoff, 2002). Additionally, H3.3 contains a substitution of one Ala to Ser residue within the N-terminal tail at position 31. Ser31 has been shown to constitute a H3.3-specific phosphorylation mark in the mitosis and meiosis of mammals and *Drosophila* (Hake et al., 2005; Sakai et al., 2009).

Ranging from yeast to mammals, all organisms additionally express a centromere-specific H3 variant, now commonly called CenH3, which is the least conserved H3 family member. CenH3 is essential for recruitment of kinetochore components and accounts for the specialized higher order folding of centromeric chromatin through distinct changes in the structure of nucleosomes. Replacement of canonical H3 by CenH3 therefore determines centromere identity, inheritance and function (Torras-Llort et al., 2009).

The different spatial and temporal deposition for the H3 variants and their association with different PTM marks (Hake et al., 2006) have lead to the hypothesis that they define a persistent epigenetic cellular memory by indexing the genome into functionally separate domains such as euchromatin, facultative heterochromatin or constitutive heterochromatin (Hake and Allis, 2006).

1.4.1.3. H4 histones

Histone H4 proteins are the most conserved histone proteins and all vertebrates share one H4 that is 100% identical in its amino acid sequence. So far no amino acid sequence variants have been reported in any multicellular organism and diversified H4s are so far only known from trypanosomes and ciliates (Hayashi et al., 1984; Lowell et al., 2005). Why H4s are so invariable is not fully understood. It is speculated that since H4 makes contact with all other histones in the octamer, variation in its sequence is less well tolerated (Bernstein and Hake, 2006).

1.4.1.4. H2A variants

The H2A histone family is the most diverse with the greatest number of variants. To date 4 H2A variants have been characterized in mammals that can substitute canonical H2A, but most vertebrate genomes encode an even higher number of H2A sequence isoforms with unknown functions. In addition to the canonical H2A.1, the variants H2AX, H2AZ, macroH2A and H2ABbd have been described in different organisms. The functions of H2As are very diverse and their deposition may alter the structure of the chromatin fiber and access to the DNA significantly.

H2AX and H2AZ are of nearly universal occurrence and present in most animal species. H2AX plays an important role in the maintenance of genome integrity of eukaryotic genomes by participating in the repair of double stranded DNA-breaks (DSB). H2AX has arisen multiple times during evolution but similar evolutionary constrains have led to the convergent acquisition of the H2AX-specific phosphorylation motif SQE/DΦ (in which Φ represents a hydrophobic residue) within the C-termal tail (Malik and Henikoff, 2003). As a response to DSBs, this motif becomes rapidly phosphorylated by the phosphoinositide 3-kinase-like kinases, ataxia telangiectasia mutated (ATM), ataxia telangiectasia and RAD3-related (ATR) and the

DNA-dependent protein kinase (DNA-PK) creating the γ-H2AX form. γ-H2AX then helps to recruit and/or retain DNA repair proteins, histone modifying enzymes and chromatin remodeling complexes (Pinto and Flaus, 2010). H2AX is also essential for condensation and silencing of the male sex chromosome in mice although it is not crucial for meiotic homologous recombination (Fernandez-Capetillo et al., 2003). In mammals, H2AX represents only 10% of the total H2A pool while H2AX is the major H2A-component in *S. cerevisiae* and entirely replaces the canonical H2A in fungi, Giardia spp. and Cryptosporidium spp. (Talbert and Henikoff, 2010). However, the nematode *Caenorhabditis elegans* and protozoan parasites such as *Plasmodium* and *Trypanosomes* have no H2AX. The variant H2A.Z has been one of the most extensively studied histone variants in the recent years. H2A.Z diverged from canonical H2A before the diversification of modern eukaryotes. H2A.Z is only 60% identical to canonical H2A, but its amino acid sequence is highly conserved between different organisms. H2A.Z histones are essential for the viability in a range of species, including *Tetrahymena*, *Drosophila, Xenopus* and mice (Clarkson et al., 1999; Liu et al., 1996; Ridgway et al., 2004; van Daal and Elgin, 1992), yet the role for this variant remains controversial and its functions may also have some species-specificity. H2A.Z has been implicated in many diverse biological processes, such as gene activation, chromosome segregation, heterochromatin silencing, and progression through the cell cycle (Zlatanova and Thakar, 2008). Recent studies on the genome-wide deposition of H2A.Z point towards a function for H2A.Z in the establishment and maintenance of chromatin boundaries that define promoter elements and those that demarcate genes (Jin et al., 2009). The H2A.Z of *Drosophila* (H2AvD) is an exception in that the H2AX phosphorylation-motif feature is merged with the H2A.Z variant (van Daal et al., 1988).

Macro H2As (mH2As) are characterized by a histone fold domain followed by a large, non-histone domain that includes a basic region and a macro domain. The macrodomain reduces transcription factor access and represses transcriptional activation mediated by the histone acetyltransferase p300 (Doyen et al., 2006). Isoforms of mH2A are enriched in the Xi chromosomes of females mammals

(Chadwick and Willard, 2002) and it has been proposed that mH2A interacts with the to facilitate the silencing of one X chromosome in mammals to achieve dosage compensation (Nesterova et al., 2002). Initially mH2A was thought to be vertebrate-specific but more recently mH2A genes have also been identified in other animal phyla such as cnidarians, echinoderms and cephalochordates (Talbert and Henikoff, 2010).

1.4.1.5. H2B variants

In comparison with members of the H1, H2A and H3 families, H2B histones lack specialized replication-independent variants in the somatic cell lineage. Most organisms possess genes for several H2B isoforms but the variability observed in H2Bs seems to occur exclusively in the male germ line of vertebrates and invertebrate organisms and the pollen of plants with largely unknown functions (Aul and Oko, 2002; Marzluff et al., 2006; Ueda and Tanaka, 1995). So far two testes-specific variants have been described in human, including the testes-specific TSH2B (Zalensky et al., 2002), and H2BFW (Churikov et al., 2004). TSH2B is present during male meiosis but also retained in human sperm and enriched in genes for spermatogenesis, suggesting that it might promote their transcription or prevent their packaging in protamines (Hammoud et al., 2009). In contrast to other testes-specific histone variants, the assembly of H2BFWT into nucleosomes is not associated with an increased instability but allows the chromatin fibers to resist chromatin compaction. It was found that H2BFWT is unable to recruit chromosome condensation factors or assist in mitotic chromosome assembly (Boulard et al., 2006).

In addition to universal histone variants, the histone complements of most metazoans also include lineage-specific histone isoforms that exhibit lineage specific substitutions in their amino acid sequence.

1.4.2. Lineage-specific histone variants within the male germline

The majority of lineage-specific variants described for metazoans are specific to the male germline. Chromatin undergoes extensive reorganization and remodeling during

spermatogenesis (Gaucher et al., 2010) and testes-specific histone variants are thought to participate in the genome-wide displacement of the canonical histones to facilitate reprogramming of the male genome. Histone subtypes exclusive to the male germline have been identified in all of the core histone families except H4. Spermatogenesis-specific H2A variants have been described in human and mice. The variant H2A.Bbd was initially identified as being largely excluded from the inactive, female X chromosome of humans (Chadwick and Willard, 2001) and has recently been found to be involved in the spermiogenesis of mice (Ishibashi et al., 2010). A study of the mouse orthologs of H2A.Bbd, H2AL1/H2AL2 (Syed et al., 2009) and another testes-specific variant H2AL3 showed that they become specifically associated with pericentric regions after male meiosis, just before and during the synthesis and assembly of protamines (Govin et al., 2007). While H2B variants are subject to rapid diversification primarily in the male germline of most metazoans (see 1.3.1.5), testes-specific H3s have so far only been described in rat, human and few plant species (Okada et al., 2005; Trostle-Weige et al., 1984; Witt et al., 1996).

1.4.3. Histone variants -The structure behind the function

Chromatin higher order structure is affected by amino acid variations on the histone octamer surface and in the histone tails that mediate nucleosome – nucleosome interactions. Thus, it may be plausible to assume that the incorporation of histone variants modulates the structure of the nucleosome. Crystallographic data is now available for most of the universal histone variants and various techniques have been established to study the dynamic behavior of histone variants *in vitro* and *in vivo.*

The structural effects of H2A variants are of particular interest since they have diverged significantly in amino acid sequence from canonical H2A and appear to designate the nucleosome for a diverse set of functions. Many studies have attempted to identify certain characteristics of H2A.Z that might directly affect the behavior of nucleosomes but data on the biochemical and physical properties of the H2A.Z-nucleosome have not yet provided clear answers on its stability or how H2A.Z performs its many functions (Zlatanova and Thakar, 2008). Some recent studies

suggest that acetylation of H2A.Z leads to destabilization (Abbott et al., 2001; Thambirajah et al., 2006) and alters the conformation of the nucleosomal core particle (NCP) when it is reconstituted with other acetylated core histones (Ishibashi et al., 2010). A common model is that the acetylation of N- or C-terminal histone lysine residues neutralizes the positive charge of histone tails, thereby leading to a relaxation of densely packed chromatin (Roth et al., 2001). Like the H2A.Z-nucleosome, NCPs containing the variant H3.3 show only a subtle destabilization but co-existence of both variants within a nucleosome affects nucleosome position (Thakar et al., 2009).

Destabilization of the nucleosome also appears to be a common theme in histone variants specific to the male germline and has also been reported for the testes-specific TSH2B variant (Li et al., 2005). Another H2A variant that has shown to lead to a less stable nucleosome is the variant H2A.Bbd. H2A.Bbd appears to be a highly specialized, variant which is expressed in mammalian testes and only 48% identical to H2A.1. H2A.Bbd displays some unusual structural properties leading to a more relaxed nucleosome that protects only 130bp DNA. However, its effect on the nucleosome structure are not yet clear, since it also appears to repress transcription and inhibits remodeling by the, SWI/SNF complex (Bao et al., 2004; Doyen et al., 2006; Menoni et al., 2007).

In contrast, NCPs containing macroH2A are stabilized as revealed by sedimentation analysis and salt dissociation experiments (Abbott et al., 2001; Pehrson and Fried, 1992). This data further supports the view that macroH2A is a “chromatin silencer” that interferes with transcription factor binding and the SWI/SNF remodeling complex (Angelov et al., 2003). However, several PTMs have been described for mH2A, suggesting roles beyond heterochromatinization (Thambirajah et al., 2009).

So far there is little structural information for the H2AX variants. Experiments mimicking the phosphorylation of the SQE/DΦ motif by a replacement of Ser with Glu, resulted in a more relaxed chromatin structure that protects less DNA from

MNase digest which may facilitate the function of H2AX in DNA repair (Fink et al., 2007). An increase in stability has also been reported for nucleosomes containing the testes-specific H3 variant of mammals (Tachiwana et al., 2010).

Despite the high sequence divergence that exists in some histone variants and their significant functional consequences, crystallographic analyses suggest that the effects of most histone variants on the overall structure of the nucleosome core particle (NCP) are surprisingly subtle. CenH3, however, appears to be a histone variant that may indeed induce more drastic structural alterations of the NCP. The actual composition and structure of the CenH3-containing nucleosomes is still a matter of debate and might also be species-specific. While *in vitro* reconstitution and affinity purification experiments suggest that human and fly CenH3-nucleosomes form "canonical" $(CenH3/H4/H2A/H2B)_2$ octamers (Blower et al., 2002; Foltz et al., 2006; Yoda et al., 2000), intranucleosomal cross-linking experiments and atomic-force microscopy measurements, led to the proposition of a "half-nucleosome" model for *Drosophila*, suggesting that CenH3-nucleosomes exist as (CenH3/H4/ H2A/H2B) tetramers, rather than as octamers (Dalal et al., 2007). The latter idea and results from supercoiling assays, further led to the proposal that CenH3-nucleosomes may direct the DNA into a right handed superhelix, introducing positive superhelical turns which may in turn provide a single uncondensed location in the condensed mitotic chromosomes that remains accessible for kinetochore binding proteins (Furuyama and Henikoff, 2009). Moreover, recent results from *S. cerevisiae* suggest, that the single CenH3-nucleosome (Cse4 in yeast) lacks the H2A/H2B dimers. Instead, Cse4-nucleosomes contain the non-histone protein Scm3, suggesting that Cse4-nucleosomes are composed of $(CenH3Cse4/H4/Scme3)_2$ hexamers (Mizuguchi et al., 2007). Nevertheless, despite these different models, it is now widely accepted that CenH3 nucleosomes must adopt highly unconventional arrangements on the DNA that presents centromeric chromatin to the poleward face of the condensing nucleosome.

1.4.4. Mixing and matching histone variants

The increasing data on variant nucleosomes indicates that different histone variants most likely co-exist *in vivo* (Fig. 5). Nucleosomes containing several histone variants may provide an additional degree of freedom for altering the composition and thus the stability of the nucleosome.

H3.3 and H2A.Z are considered as "active" replacement variants that co-exist in the NCPs of active promoter, enhancer and insulator regions as revealed by genome wide studies (Jin et al., 2009; Mito et al., 2005). In support for this view is the fact that H2A.Z/H3.3-double variant nucleosome are highly unstable and appear to mark chromatin regions "in flux" with a high turnover (Jin and Felsenfeld 2007, Jin et al. 2009). Another example for a variant that might preferable be incorporated within a H3.3-nucleosome is H2A.Bbd. Assembly and disassembly of H2A.Bbd-nucleosomes is accomplished more efficiently *in vitro* when NCPs contain H3.3 instead of canonical H3 (Okuwaki et al., 2005). The distinct structural properties of centromeric chromatin containing the variant CenH3 may further be enhanced by the incorporation of other variants. In human cell lines, the centromeric CenH3-nucleosomes purified from human cell lines are enriched in macroH2A and H2A.Z (Foltz et al., 2006). Furthermore, the synchronized expression of testes-specific H3-, H2A- and H2B- variants during the spermatogenesis of different organisms suggest that these variants may potentially be part of the same NCP. The testes-specific H2A variants of mice (H2AL1/L2) for example dimerize more efficiently with TH2B than with the canonical H2B, suggesting that "male nucleosomes" do exist *in vivo* (Govin et al., 2007).

1.4.5. The evolution of histone variability

Although chromatin is considered to be eukaryotic, the evolutionary origins of eukaryotic core histones already lead back to the histone proteins found in the Archaea. Similar to eukaryotic histones, archeal histones comprise a single histone fold domain containing three alpha helices and two intervening loops but their lack the C- and the N-terminal tail (Sandman et al., 1990). In halobacteria however,

unusual “doublet” histones were identified consisting of an end-to-end duplication of the histone fold which may have led to a subfunctionalization of the N- and the C-termini and eventually, to the evolution of eukaryotic histones (Malik and Henikoff, 2003).

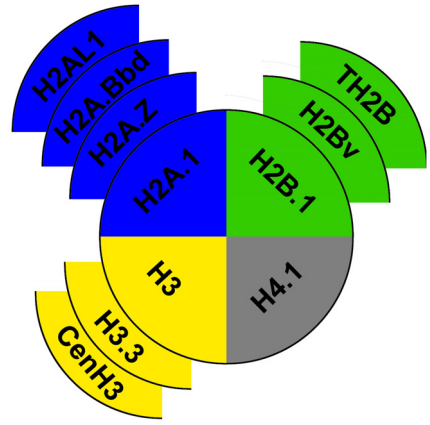

Figure 5 Variant nucleosomes. The figure shows potential combinatorial possibilities of histone variants within one histone octamer. One nucleosome may contain only one type of variant as either mixed dimers (e.g. H2A.Z/H2A.1) or alternatively as a dimer of the same variant (e.g. H2A.Z/ H2A.Z). Variant nucleosomes may also be composed of multiple variants from different families as for example $(H2A.Z/H2Bv/H3/H4.1)_2$ or $(H2A.Z/H2B.1/H3.3/H4.1)_2$. instead of the conventional $(H2A.1/H2B.1/H3/H4.1)_2$. The cartoon is based on the nucleosome code hypothesis (Bernstein and Hake, 2006).

Because of the extreme sequence conservation among the canonical histone proteins, histones were long viewed as slowly evolving proteins that lack any diversification. With the increasing number of sequenced genomes it is now clear that histone variants have diversified in many species to assume crucial functions in gene regulation, DNA repair, chromosome segregation and other processes. While “universal” histone variants are found in nearly all eukaryotes and reflect ancient functions common to eukaryotic cells, lineage-specific variants have specialized to the unique biology of their host organism.

Analyses on histone complements show that histone gene organization differs considerably among species. The majority of the human-, mouse, sea urchin- and *Drosophila* histone genes are organized as large clusters (Marzluff et al., 2002; Marzluff et al., 2006; Nagel and Grossbach, 2000) and in *Drosophila* and sea urchin, these genes form arrays with several hundred genes tandemly reiterated. In contrast, *C. elegans* has a much smaller number of histone genes that are dispersed in small groups throughout the genome (Roberts et al., 1987). Likewise, histone genes are not physically linked in *O. dioica*, as shown by a first analysis of the histone complement by Chioda and colleagues (Chioda et al., 2002), suggesting that there is no correlation between histone organization and phylogenetic position.

However, the underlying mechanisms that drive histone variant evolution have only marginally been explored in higher metazoans and with the completion of genome sequencing in these organisms; it has become clear that the number of histone isoforms is larger than previously identified. Furthermore, comprehensive developmental expression data is only available for sea urchin histone genes and thus information on the function of histone variants in metazoans, mainly comes from studies of universal histone variants described in the previous chapters.

There is recent evidence that histone variants may indeed evolve adaptively to specific life history traits of organisms. The H2A variants of rotifers are an interesting example of histone variants that might have adapted to specific environmental conditions. Bdelloid rotifers are asexual freshwater invertebrates and are known for their extraordinary ability to survive desiccation and high dosages of ionizing radiation that cause hundreds of DNA double-strand breaks (DSBs) per genome (Gladyshev and Meselson, 2008). Remarkably, no H2AX, H2A.Z or canonical H2As exists in these animals, but H2A variants have extended C-terminal tails. The function of these variants is unknown, but it is speculated that they are part of the adaptations that have evolved to survive desiccation and the attendant burden of DNA damage (Van Doninck et al., 2009). The diversified H4 and H3 variants of ciliates may likewise reflect a histone complement adaptation to the specific life history traits of an organism. Ciliates are single-celled alveolates that have two types of nuclei. A

diploid, mitotic micronucleus gives rise to a second, polyploid macronucelus through differential DNA elimination, amplification and fragmentation. The fact that the expression of the two *Euplotes crassus* variants H3(P) and H2A(P) is synchronized with these processes and exclusive to the developing macronucleus suggests that they may have specialized to facilitate these processes (Ghosh and Klobutcher, 2000; Talbert and Henikoff, 2010). Diversified forms of each of the core histones also exist in trypanosomes that lack the typical RNAPol II promoters of eukaryotes. Instead, termination and initiation of transcription occurs at strand switch regions (SSR) which separate the polycistronic trypanosome transcription units. SSRs are punctuated with specific sets of histone variants, depending on whether they initiate or terminate transcription, suggesting an ancestral mode of transcriptional regulation based on histone variants and not on transcription factors (Siegel et al., 2009).

The fact that histone variants may evolve adaptively to the life history traits or genome features of organisms also raises the possibility that the same may be true for the panel of histone modifications. Indeed, there is some evidence that the repertoire of PTMs becomes more complex from unicellular organisms to mammals. Assessment of the PTM profiles of *Tetrahymena*, Yeast and *Plasmodium* revealed a general predominance for modifications associated with a transcriptionally active state as well as novel modifications (Garcia et al., 2007; Trelle et al., 2009).

The idea that organisms may have evolved species-specific strategies to adapt their chromatin interface is intriguing but further metazoan models will be required to complement the current knowledge on histone variability and PTM profiles across species. Ideally, these may include other chordates with genome features and biological traits distinct from those of mammals. Rapidly evolving, marine urochordates, that bridge the plyogenetic gap of the invertebrate-vertebrate transition, present an opportunity to investigate whether specific life history strategies, phylogenetic position or underlying genome features correlate with the histone variability and histone modifications of organisms. The genome of the dioecious larvacean *Oikopleura dioica* has undergone extreme compaction which makes *O. dioica* an exciting new chordate model to further investigate these questions.

1.5. The urochordate *Oikopleura dioica*

Appendicularians (larvaceans) are marine zooplankton that represent the second most abundant component of marine meso-zooplankton after copepods (Fenaux, 1998). Serving as food for other zooplankton and fishes, appendicularians have ecological importance in the marine food web. *Oikopleura dioica* populations respond very quickly and opportunistically to algal blooms with growth rates exceeding those of copepods (Nakamura et al., 1997; Troedsson et al., 2002) and the species *O. dioica* has been the topic of many ecological studies. More recently, the sequencing of the *O. dioica* genome and improvements of the laboratory culture of this animal have also provided the basis for a number of insightful studies on the evolution and cellular biology of this organism, revealing its potential as an experimental model.

1.5.1. Ecological impact and *O. dioica* phylogeny

Within the chordate phylum, appendicularia is one of three classes belonging to the urochordates (tunicates), which are now considered to be the closest living relatives of vertebrates (Fig. 6) (Delsuc et al., 2006). Urochordates exhibit a simplified chordate body plan, characterized by a tail with a notochord and dorsal neural tube, an endostyle and gill slits. Appendicularians are also called larvaceans because they retain the larval tail throughout their entire life. While the two tunicate classes of ascidiacea and thaliacea eventually resorb their tadpole tail and undergo a metamorphosis that involves major changes in the body plan, the metamorphosis of the larvaceans consists of a simple “tail shift” with respect to the trunk axis, orienting the distal end of tail in the same direction as the mouth (Fig. 7). The species *O. dioica* was first described in 1821 by Chamisso and Eysenhart (Flood and Deibel, 1998). The animal (Gr. *oikos* house), lives and filter feeds on particles in a gelatinous house, which is secreted from polyploid epidermis cells of the trunk also referred to as the oikoplastic epithelium (Fenaux, 1985; Spada et al., 2001; Thompson et al., 2001). During the growth of the animal, houses are repeatedly discarded and renewed and constitute a main component of the marine snow and the carbon cycling of marine ecosystems (Alldredge, 1976; Robison et al., 2005).

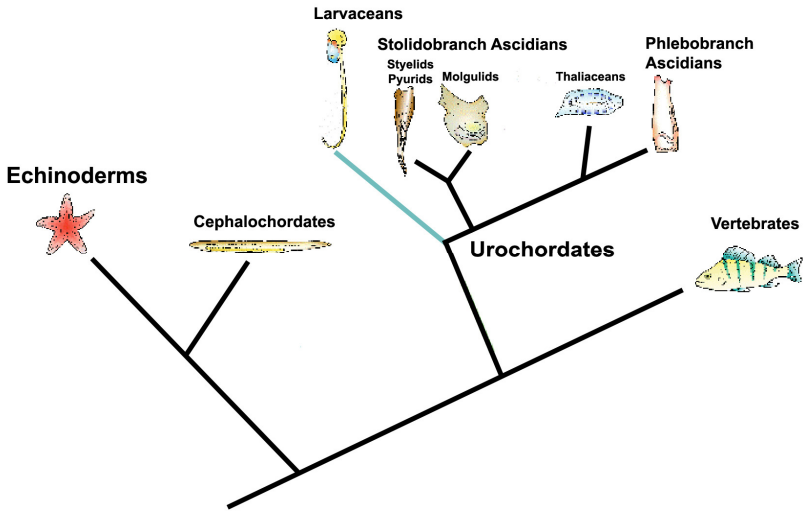

Figure 6 Phylogeny of the chordates. According to the new phylogeny of chordates (Delsuc et al., 2006), tunicates (urochordates) are the new sister group of vertebrates. Modified from a drawing of Billie Swalla, University of Washington.

1.5.2. *O. dioica* development and life cycle

O. dioica is now cultured and bred under laboratory conditions in Norway, USA, France and Japan with a very short life cycle (6 days at 15°C) (Bouquet et al., 2009). Schematically, the life cycle of *O. dioica* can be characterized by a developmental switch between two different cell cycle types, mitotic and endoreduplicative (Fig. 7).

Oocytes and sperm are released into the sea water and following fertilization, the first division is already observed after 30 min. Early embryonic development is highly determinant and a fixed number of cells is established through rapid mitotic divisions. Organogenesis commences approximately at 2.5 h post-fertilization (pf), 30 min before the tadpole hatches. Organogenesis is completed at metamorphosis (tail shift) which occurs 9-11 h pf, when cells in most tissues exit mitosis and increase their nuclear volume by endoreduplication, making the animal grow about 10-fold in

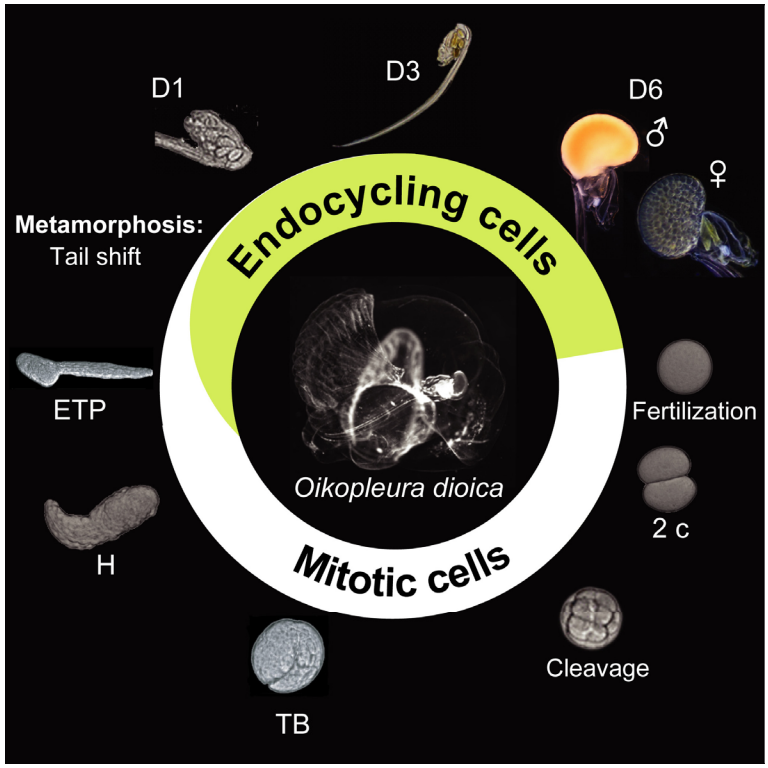

Figure 7 Life cycle of Oikopleura dioica. The life cycle begins with rapid mitotic divisions, 30 min pf (2c). 2.5 h pf, the tadpole is already visible in a tail bud stage (TB). The tadpole hatches 3 – 3.5 h pf (H) at the onset of organogenesis when the majority of cells still divide mitotically. At early tadpole stages (ETP 5-7 h pf) cells of the oikoplastic epithelium gradually shift from mitosis to endocycling. At approximately 12 h pf, embryos undergo a metamorphosis (tail shift) and juvenile day 1 (D1) animals will start to inflate their first house. Gonads start to differentiate at D4 and by D6, mature oocytes and sperm are clearly visible in females and males respectively. Adult animals die shortly after releasing their mature gametes into the sea water.

size until day 6 (D6). In the oikoplastic epithelium of adult animals that secretes the mucous house, this results in different cell field-specific ploidy levels (Ganot and Thompson, 2002). Only few tissues such as the gut epithelium and the gonad continue mitotic divisions. The subsequent stages are referred to as day1-day6 (D1-D6), whereby females and males can only be distinguished from D4 onwards.

Contrary to all other tunicate species which are hermaphroditic, *O. dioica* has separate male and female sexes. The female and male gonads of *O. dioica* show similarities to the syncitium of *Drosophila,* where proliferating germline nuclei occupy a common cytoplasm. The ovary, in which the germline compartment has been termed the "coenocyst", gives rise to two types of nuclei: meiotic nuclei and endocycling nurse nuclei (Ganot et al., 2007a; Ganot et al., 2007b). At the end of the life cycle, female gonads are filled with metaphase I arrested oocytes and male and female gametes are released into the sea water. The underlying cellular processes that make *O. dioica* particularly adept in producing a highly variable number of oocytes have recently been addressed in studies on the ovary (suppl. paper I).

Endoreduplication, or endocycling, is common in protists, plants and animals including arthropods, molluscs and mammals. In endocycling cells, S-phases alternate with distinct gap-phases, but there is no cell-division. Endocycling cell cycles vary among species in that some retain hallmarks of mitosis while other examples lack mitotic remnants including chromosome condensation, nuclear envelope breakdown and reorganization of microtubules (Edgar and Orr-Weaver, 2001). Endocycling is usually found in cell cycle types with high secretory activities or reserved for nutrient uptake and storage with high metabolic activity. Some of the best-studied examples include *Drosophila* follicle and nurse cells, rodent Trophoblasts and plant endosperm that can reach polyploidy levels of 24 000 C (DNA contents as a multiple of the haploid genome) (Traas et al., 1998). A common hypothesis is that increasing DNA content by endocycling sustains the mass production of proteins. Additionally, the steady increase in DNA content also provides a strategy for growth. *C. elegans* and *Drosophila* larvae are examples where overall growth is mainly driven by endocycling (Lozano et al., 2006).

1.5.3. *O. dioca* as a model for chromatin studies

Many recent advances in our understanding of the molecular and cellular mechanisms that govern development stem from model organisms with short life cycles in which genetic approaches are feasible, such as *Drosophila* and *C. elegans*. Several traits

make *O. dioica* an interesting model for ecology, developmental cell biology, gene expression studies and evolution. These include i) its key phylogenetic position near the invertebrate-vertebrate transition, ii) the extremely compacted genome and high gene density, iii) its short life cycle and the availability of long-term cultures, iv) the invariant cell fate and v) its transparency throughout the entire life cycle. For the study of chromatin dynamics, the transition between different cell cycle types and its extremely compacted genome architecture is of particular interest. *O. dioica* has the smallest genome ever found in a chordate of only 72 Mb (Seo et al., 2001) and the first assembly of the draft genome is now available at Genoscope (http://www.genoscope.cns.fr/externe/GenomeBrowser/Oikopleura/). Despite the tiny genome size, the number of genes is estimated as 18000 with a high gene density of one gene per 4–5 kb. Intergenic sequences are very short, and 62% of introns are smaller than 50 bp. Indeed, the extreme compaction of the chordate genome in the larvacean lineage of the urochordates means that regulatory regions have been strongly compressed. Often these are in the order of a single nucleosome or less. Furthermore, 25% of the *O. dioica* genes are transcribed as operons and there is complex interlacing of transcriptional units with very different expression profiles. An example is given in Fig. 8, showing the tiling array expression data of the polo-like kinase 1 (Plk1) gene locus. Male-specific, female specific and embryonic gene expression are specifically regulated within less than 6 kb by very short, bidirectional promoters with the smallest being only 100 bp in size. In such a context, interesting questions regarding genome wide active and repressive chromatin marks arise. Are promoters and enhancers defined by similar histone modification combinations and histone variants as described for other species? How significant are histone modifications and histone variants in indexing regulatory regions of such a small size? Furthermore, the cell cycle regulated acquisition and deposition of some histone modifications and histone variants suggests that not all chromatin marks serve as an "information" code about transcriptional activity. Instead, some marks might function as determinants that facilitate chromatin condensation and confer topological changes specific to mitotic chromosomes. *O. dioica* also provides the opportunity to study such a "structure based chromatin code" and the chromatin dynamics associated with

the rapid transitions from mitosis to endocycling and meiosis that occur during the *O. dioica* development.

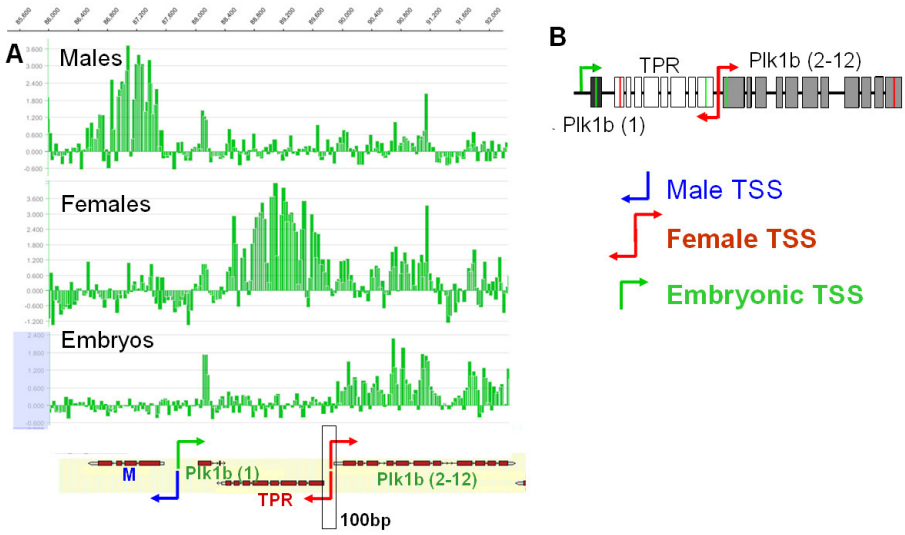

Figure 8 The Polo-like kinase-1b locus in *O. dioica*. The male-specific, female-specific and embryonic gene expression are specifically regulated within less than 6 kb by very short promoters. A) Tiling array expression data of the Plk1b-locus in D6 females, D6 males and 7 h embryos. Green bars show the intensity of probe signal obtained after hybridization with cDNA. Gene annotations of the Plk1b genes and exons (indicated in brackets), the female-specific TPR gene and a male-specific gene (M) upstream of Plk1b are shown below. Arrows indicate promoters and direction of transcription. B) Detailed exon-intron organization of the Plk1-locus. The Plk1b-gene is expressed in somatic tissues throughout development and in late animals predominantly in females. Expression of the first Plk1b-exon is driven from a 350 bp, bidirectional promoter region that also regulates the male-specific gene upstream of Plk1b. In mature females, exons 2-12 of Plk1b are expressed from a second bidirectional 100 bp-promoter located within the first intron of the Plk1b gene. The first Plk1-intron contains a female-specific TPR gene in opposite orientation which is expressed from the same bidirectional promoter region.

2. Aims of study

Histone variants participate in important cellular processes, including DNA repair, cell division, transcriptional repression and the activation of genes. Despite their central roles in these essential functions, and generally high degree of sequence conservation across evolution, histone complements among different species vary in their divergence, the number of lineage-specific variants and partly, in the absence or presence of the universal variants H2A.Z, H2AX, MacroH2A, CenH3 and H3.3. Histone complements may therefore reflect specific life history traits and genome features of different organisms and reveal relevant information about the function and evolution of histone variants. The switch between mitosis and endocycling during the *O. dioica* development and the absence of mitosis in most tissues of the later stages, raises for example intriguing questions with respect to the centromeric histone variant CenH3. Do endocycling chromosomes still need to incorporate this variant? Is a possible absence of CenH3-chromatin achieved through the regulation of CenH3 expression or a lack of the CenH3 deposition machinery? Furthermore, the important phylogenetic position of *O. dioica* provides an interesting opportunity to compare the diversity of histone variants found in the chordate lineage and to study their function with respect to the specific life history traits of this dioecious urochordate. Moreover, the possible lineage specific function that has been proposed for mitotic histone H3 phosphorylation at different residues, make *O. dioica* an interesting model to study the degree of conservation or specialization of histone modifications within the chordates.

More specifically the aims of this work have been:

1) To identify the full complement of histone genes that are present in first assembly of the *Oikopleura dioica* genome and to characterize their expression throughout development.

2) To establish a public data base for *O. dioica* histone sequences- and expression data.

3) To purify *O. dioica* histones for subsequent mass spectrometry analysis for the identification of the histone modifications present in *O. dioica* and to verify the expression of histone variants at the protein level.

4) To establish a histone-eGFP fusion expression system for *in vivo* studies on the centromeric variant CenH3 in mitotic and endocycling cells.

5) To study the temporal and spatial distribution of histone H3 phosphorylation during, mitosis, meiosis and endocycling in *O. dioica* cells.

3. List of papers

Paper I

Moosmann A, Campsteijn C, Jansen PW, Schmid M, Stunnenberg H, Thompson EM.

High diversity of developmental stage-specific histone variants in the urochordate *Oikopleura dioica.* Manuscript in preparation.

Paper II

Schulmeister A, Schmid M, Thompson EM. 2007. **Phosphorylation of the histone H3.3 variant in mitosis and meiosis of the urochordate *Oikopleura dioica.*** *Chromosome Res.* **15**: 189 – 201.

Paper III

Moosmann A, Schmid M, Bouquet JM, Campsteijn C, Zech K, Bal U, Thompson EM **Chromatin signatures at the *Oikopleura dioica* centromere.** Manuscript in preparation.

4. Summary of results

4.1. High diversity of developmental stage-specific histone variants in the urochordate *Oikopleura dioica* (Paper I)

With the first assembly of the *O. dioica* draft genome available at Genoscope, this work characterizes the complete histone complement of *O. dioica* including the number of histone genes, histone gene organization and all histone sequence variants. We have previously reported that, similar to the histone gene complements of other higher eukaryotes, the canonical histone genes of *O. dioica* are organized as clusters and share divergent promoters (Chioda et al., 2002). In contrast to the histone complements of most higher metazoans, the number of histone genes is relatively low, lacking the organization in large clusters or tandem arrays that have for example been found in mammals and *Drosophila,* respectively. In this work, we have identified 47 histone genes in the *O. dioica* genome encoding 31 different histone proteins (6 histone H3s, 5 histone H1s, 2 histone H4s, 7 histone H2Bs and 11 histone H2As) dispersed in small clusters of 2 – 5 genes throughout the genome.

Developmental and sex-specific expression of some histone genes has been reported from other organisms (Marzluff et al., 2006; Wolfe and Grimes, 1991) but has rarely been comprehensive. We investigated the developmental stage-specific expression of histone genes throughout *O. dioica* development by quantitative reverse transcriptase-polymerase chain reaction (quantitative RT-PCR). *O. dioica* histone genes are co-regulated in clusters and genes belonging to the same locus mostly showed the same expression profile throughout the development. Moreover, many histone genes were indeed exclusively expressed at distinct developmental stages and histone gene clusters could be assigned to seven expression patterns: 1) throughout the entire life cycle, 2) exclusively during organogenesis, 3) male-specific expression in mature D6 animals, 4) expression peaking at metamorphosis, 5) expression that primarily takes place in D3/D4 adults, 6) expression peaking in early tadpoles and male D5 animals and 7) transcripts that are predominantly present from 1h pf to D4.

We have summarized all the cDNA, genomic sequences and expression in the "*Oikopleura* histone database" (http://apps.cbu.uib.no/oikohistonedb/. Login: *oiko*, password: *JasVopoi*) where all the sequence data is also available for download.

Comparison of the *O. dioica* histone protein sequences with the histone variants identified in other organisms and urochordates revealed that despite the low number of histone genes, the *O. dioica* genome encodes a very diverse array of different histone isoforms. Most universal histone variants are present including the variants H3.3, CenH3 and H2A.Z but *O. dioica* lacks the variants macroH2A and H2AX. Additionally, *O. dioica* expresses 17 different core histones that appear to be specific to the appendicularian lineage of which 15 are significantly enriched in the male gonad. These also include one H4 sequence variant (H4t), exclusively expressed in the male gonad of maturing animals, which is the first H4 variant reported thus far for higher metzoans. In most animals, histones are removed by protamines during the condensation phase of the spermatid nuclei but a certain percentage of the DNA remains nucleosome bound in mature sperm (Gatewood et al., 1990). These observations have led to the idea that canonical histones and histone variants retained in sperm could mark the paternal genome and transfer epigenetic information to the zygote. Considering the large number of male-specific histone isoforms in *O. dioica*, we therefore asked if these were still present in the chromatin of mature sperm and subjected *O. dioica* sperm histones to mass-spectrometric analyses. Interestingly, we found several canonical histones and nearly all of the male-specific ones to be retained in *O. dioica* sperm, supporting the idea that they may serve functions beyond spermatogenesis. *O. dioica* lineage-specific histone proteins further include two very divergent H2A variants (H2A.3 and H2A.4) of which H2A.3 appeared to be enriched in the female gonad. Finally, we found that *O. dioica* uses alternative splicing to expand its repertoire of histone variants. Within the genomic sequence of the OdH2A.Z gene, several putative alternative splice sites exist suggesting that *O. dioica* is able to modify the number of GK-residue motifs within its H2A.Z N-terminal tail by alternative splicing. We were able to confirm the expression of two H2A.Z splice variants, that lack one (H2A.Zb) and two N-terminal lysine residues (H2A.Zc) by rapid amplification of cDNA ends (RACE) and cloning with specific primer pairs.

A current view is that histone modification patterns become more complex from unicellular eukaryotes to mammals (Garcia et al., 2007) Its highly compact and gene-dense genome with short promoters and the extensive use of cellular endocycles for growth, make *O. dioica* an interesting candidate to study histone posttranslational modifications (PTMs) with respect to current paradigms about their function in higher eukaryotes. Purifying *O. dioica* histones from adult animals and their subsequent mass-spectrometric analysis identified many PTMs that are conserved marks for transcriptional activity or repression in a wide range of organisms. These included several acetylated lysine residues within N-terminal tails of histone H3 and H4 as well as the methylation of H3 at lysine 4 (H3K4), H3K36 and H3K79 which have previously been associated with transcriptionally competent chromatin. Furthermore, *O. dioca* histones carried histones marks that typically mark silent chromatin including the methylation of H3K9, H3K27 and H4K20. We found all of these modifications to occur in *O. dioica*. Previously it has been reported that these "silent" marks are reduced in unicellular organisms (Garcia et al., 2007; Trelle et al., 2009), which has been explained by the fact that in contrast to higher vertebrates, the majority of the genome in unicellular eukaryotes is transcriptionally competent (Garcia et al., 2007). Considering the high gene-density of the *O. dioica* genome (one gene per 4–5 kb) and the short intergenic regions, we also expect the proportion of the *O. dioica* genome that is permanently silenced to be relatively low compared to that of human. However, our analysis revealed no obvious lack of "silent" heterochromatic marks in *O. dioica*, since methylation of H3K27 and H4K20 were both detected by our mass spectrometry analysis and the methylation of H3K9 was confirmed by immunofluorescence staining (Paper III). Thus, we conclude that while there are some differences in the panel of histone sequence variants deployed in *O. dioica*, the spectrum of histone modifications and histone modifying- and demodifying enzymes (appendix, table A1A and A1B) is generally very similar to that found in other chordates.

4.2. Phosphorylation of H3 variants in mitosis and meiosis of the urochordate *Oikopleura dioica* (Paper I and paper II)

H3 phosphorylation is considered as a hallmark of mitosis, yet its precise role during chromosome condensation and segregation is not clear. While the H3 phosphorylation marks themselves are conserved, their spatial and temporal distributions vary widely across species (Perez-Cadahia et al., 2009). We assessed the spatial and temporal patterns of canonical H3 phosphorylated at residues Thr3 (H3T3P), Ser10 (H3S10P) and Ser28 (H3S28P) and of the replacement variant H3.3 phosphorylated at Ser31 (H3.3S31P) in mitotic and meiotic chromosomes of *O. dioica* by immunofluorescence staining. With the exception of H3T3P, the distribution of these marks in the mitosis of *O. dioica* deviated considerably from the reported signals of other animals and plant species. H3T3P and H3S28P were both significantly enriched at the inner core of centromeres in the mitotic chromosomes of *O. dioica*, whereas H3S10P and H3.3S31P were more widely distributed throughout prophase, prometaphase and metaphase chromosomes (paper III). Furthermore we observed different temporal kinetics for these PTMs. H3T3P, H3S10P and H3.3S31P commenced at the onset of prophase followed by H3S28- phosphorylation in late prometaphase. While signals for H3S28P and H3S10P could still be observed during anaphase (paper III), no signal for H3T3P and H3.3S31P was observed during this stage.

We also identified H3.3S31P as a modification during oogenic meiosis in the dioecious *O. dioica*. The female and male gonads of *O. dioica* show similarities to the syncitium of *Drosophila,* where proliferating germline nuclei occupy a common cytoplasm (Ganot et al., 2007a; Ganot et al., 2007b). H3.3S31P initiated together with H3S28P in all meiotic nuclei in late diplotene, after H3S10P. However, H3.3S31P was retained only on the subset of meiotic nuclei that seeded maturing oocytes that resumed meiosis after prophase I arrest (paper II). The functional role(s) of this mark in meiotic resumption and oocyte maturation remain to be determined.

4.3. Chromatin signatures at the *Oikopleura dioica* centromere (paper III)

Centromeric chromatin is packaged by specialized nucleosomes containing the centromere-specific histone H3 variant CenH3. CenH3 is considered as the determinant for centromere identity, function and the epigenetic inheritance of centromeres but how CenH3 is guided specifically to centromeres is not fully understood. In this work we have established a histone-GFP fusion expression system to study the localization of CenH3 in endocycling and mitotic cells of *O. dioica in vivo*. Microinjection of CenH3-GFP cmRNA showed that OdCenH3 was readily deposited at the centromeres of mitotically dividing chromosomes. In contrast we did not observe a "reloading" of OdCenH3-GFP at the centromeres of endocycling cells and CenH3 transcription appeared to decrease in developmental stages dominated by endocycling. This suggests that CenH3 is no longer required during the endocycling cell cycle of *O. dioica* when mitosis is completely bypassed.

Centromeric histone modifications might be one mechanism that guides the specific deposition of CenH3, which led us to investigate PTMs at the centromeres of mitotic *O. dioica* chromosomes by immunofluorescence. OdCenH3-eGFP nucleosomes faced out towards the kinetochore, while H3 nucleosomes phosphorylated at H3T3 and H3S28 were located at the interface between sister-chromatids. H3K4 dimethylation was absent from centromeres and we observed no clear enrichment of H3K9 trimethylation at the peri-centromeric regions and no exclusion from centromeres, which is in contrast to what has been reported from other organisms. Testing a mutual dependency between H3S28P and CenH3 deposition, we used the Aurora phosphatase PP1 inhibitor Calyculin A to hyperphosphorylate *O. dioica* chromosomes. Results from these experiments and the overexpression of CenH3 showed that CenH3 deposition in mitotic cells occurred independently of H3S28P but the localization of the CenH3 variant appeared to restrict the H3S28P mark to centromeres.

5. General discussion

A unicellular parasite invading its host, a fertilized oocyte developing into a human, temperature drops and salinity increases, as diverse as the challenges for different organisms may seem to quickly adapt to environmental changes and to develop into different shapes and sizes, they all master them through a "high fidelity", temporal and tissue-specific regulation of their genes. One mechanism that regulates gene activity by controlling the accessibility to the DNA is the packaging of this genetic information with the help of histone proteins into chromatin. While the DNA sequence constitutes the general information potential of a cell, chromatin dynamics may be viewed as part of the "instruction manual" that can determine variable read outs of the same DNA sequences at different times, tissues or in different organisms. Histone variability, by chemical modification and/or through histone amino acid changes, is one factor that can profoundly alter chromatin structure as well as chromatin "interface" and thereby influence the read out of different genes. How universal are the "chromatin manuals" across different species and how much do they correlate with specific life history strategies or underlying genome features are interesting questions that have begun to be investigated. Nevertheless, due to the large genome sizes of most multicellular organisms and the largely repetitive organization of histone genes, functional studies of histone subtypes have thus far focused on few divergent universal histone variants and the picture of the sequence variability that exists within the histone complement of even established models is most likely not yet complete. With this work, we would like to introduce the small yet diverse histone complement of the chordate *O. dioica* (http://apps.cbu.uib.no/oikohistonedb) as an exciting model to study histone modifications and the function of histone variants in the context of the animal's development and gametogenesis.

5.1. The histone complement of *Oikopleura dioica* – small but diverse

To date only a few histone complements have been analyzed completely but it is already apparent that the genomic organization of histone genes varies significantly between organisms. While in yeast there are two copies of each of the core histones, the majority of the human-, mouse-, sea urchin- and *Drosophila* histone genes are organized as large clusters (Marzluff et al., 2002; Nagel and Grossbach, 2000). In *Drosophila* and sea urchin, these genes form arrays with several hundred genes tandemly reiterated, while the histone genes of C. elegans cluster in small groups which are dispersed throughout the genome (Roberts et al., 1987) Why the organization of histone complements differ to such an extent in different species is not known. It has been proposed that tandem repeats of histone genes correlate with rapid embryonic development in organisms but the small and dispersed histone cluster of the rapidly developing *O. dioica* provides another example in opposition to this idea (Chioda et al., 2002). Our comparison of the histone complements of *O. dioica* and that of the two ascidian species *C. savigny* and *C. intestinalis* further showed that within the urochordates- and even within the ascidian lineage, histone gene organization differs significantly. While at least 130 histone gene copies are present in the *C. savigny* genome (estimated 180 Mb), partly organized in large clusters, we find only very few (27) histone genes in the *C. intestinalis* genome (156 Mb) that are mostly interspersed with non-histone genes. The fact that the *O. dioica* histone gene complement rather resembles the organization of the *C. elegans* histone complement than the complement of either of the two other urochordate species further supports the view that the differences in histone gene complements can not be attributed to phylogenetic position or genome size.

Similar to *C. elegans, O. dioica* also belongs to a very rapidly evolving branch and shows a number of specializations, including a rapid embryonic development, a very compacted genome with very short promoters and intergenic regions, transcription as polycistrons (Seo et al., 2001), and the extensive use of the endoreduplicative cell cycle (Ganot and Thompson, 2002). What is the “chromatin

toolbox" of such a rapidly evolving and specialized planktonic chordate? In *O. dioica*, 31 different histone proteins are encoded in only 47 histone genes and most of the universal histone variants are also present including the variants CenH3, H3.3, H3.2 and H2A.Z (paper I). The mass spectrometry analysis conducted in this work and previously performed immunofluorescence stainings further showed that *O. dioica* histones also carry all of the major histone methylation-, acetylation- and phosphorylation marks reported from other organisms (Spada et al., 2005a; Spada et al., 2005b). Additionally, the most commonly known "writers" and "erasers" of histone modifications are present within the *O. dioica* genome assemblage (appendix, table A1A and A1B). This indicates that despite fundamental differences in life histories, genome sizes and different types of cell cycle regulation, organisms share the fundamental mechanisms to regulate their chromatin structure and interface.

5.2. A histone variant complement linked to small genome size and endocycling?

Most multicellular organisms appear to express a minimum set of histone variants including H3.2, H3.3, CenH3, H2A.Z, H2AX and MacroH2A (mH2A) in addition to their canonical histones. However, more recent data from "exotic" model organisms have raised interesting questions on how histone complements might evolve adaptively to the specific life history traits, the genome architecture and cell cycle regulations of different species. Trypanosomes for example lack the typical RNA polymerase II promoter elements and mark the transcription start and stop sites of their polycistronic transcription units with specific sets of diversified H2Bv, H3v and H4v variants (Siegel et al., 2009). In bdelloid rotifers no H2AX, H2A.Z or canonical H2As exists. Instead these animals have H2A variants with extended C-terminal tails that may be part of an adaptation mechanism to survive desiccation and an extreme burden of DNA damage that these animals are able to endure (Van Doninck et al., 2009). Some indications suggest that the histone complement of *O. dioica* has also undergone some specializations, particularly within the family of H2A histones. These include the deployment of two very divergent H2A variants (H2A.3 and

H2A.4) and H2A.Z splice variants that differ in their number of acetylatable lysines within their N-terminal tail. Moreover, *O. dioica* appears to have lost the variant H2AX and no mH2A was found.

The absence of the variant H2AX raises interesting questions about genome integrity and the pathway for double stranded break repair (DSB) in our model chordate. H2AX is present in most animals species, including the urochordate *Ciona* but is absent in some protozoan parasites and *C. elegans*. Rapid phosphorylation of H2AX (γ-H2AX) occurs during both of the two alternative DNA repair pathways, the nonhomologous end-joining (NHEJ) pathway and the alternative homologous recombination (HR) pathway that compete for DSB in eukaryotic cells (Shrivastav et al., 2008). However, more recently it was shown that this variant is neither essential for the HR pathway (Yuan and Chen, 2010), nor for meiotic recombination (Fernandez-Capetillo et al., 2003) which could mean that HR is the DSB-repair pathway of choice in *O. dioica*. In support for this idea is the observation that, as in *C. elegans,* the PIKK kinase DNA-PK, one of the key proteins in NHEJ, appears to be absent in *O. dioica*. While during mitosis, larger genomes of higher eukaryotes generally present a challenge to locating a homologous template for HR repair this may not be an obstacle in the small genome of *O. dioica* where, at least in the endocycling tissues, several hundred copies of each locus are available for recombination.

Another interesting question is whether the switch to endocycling and the absence of mitosis changes the evolutionary pressure on the conserved histone structure, allowing for more diverse sequence variants to occur. One example showing that polyploidy presents opportunities for specialized adaptation of histones are the variant H3(P) and H2A(P) of the ciliate *Euplotes crassus* which are both exclusively expressed in the developing, polytene chromosome stage of the macronucleus (Ghosh and Klobutcher, 2000). However, lineage-specific somatic H2A variants that are not testes-specific have so far not been described for chordates. The replacement H2A.3 and H2A.4 variants of *O. dioica* exhibit only 55 % sequence identity to their canonical counterpart and are therefore an interesting exception. We

found both variants predominantly expressed during the endocycling stages and OdH2A.3 appeared to be enriched in the female gonads. Interestingly, a similar divergent H2A variant is also present in the *C. elegans* and *D. melanogaster* genomes. Each of these organisms also employs endocycles in growth. However, whether these variants are expressed in the endocycling tissues of these animals is not known.

An intriguing finding is the fact that *O. dioica* appears to reduce the number of acetylatable lysines within the variant H2A.Z through the translation of different splice variants. Acetylation of H2A.Z is generally associated with a less stable nucleosome structure and occurs at the promoter of transcriptionally active genes (Bruce et al., 2005; Ishibashi et al., 2009). Studies on H2A.Z (hv1) acetylation in the ciliate *Tetrahymena* revealed that acetylation of the N-terminus works to modulate an essential charge patch that, in contrast to a site-specific histone code, is used to alter the charge of the N-terminal domain (Ren and Gorovsky, 2001, 2003). Neutralizing the charge of the highly positive, N-terminal H2A.Z tail by acetylation may lower affinity for DNA and facilitate access for essential competing factors at promoters or prevent higher order folding. Reducing the number of the 6 N-terminal Lys residues present in *O. dioica* H2A.Z would similarly result in a significant loss of positive charge and create different H2A.Z variants with gradually differing affinities for DNA. Why would such a regulation be beneficial over a regulation by acetylation? If indeed the function of H2A.Z is to set up a chromatin architecture that is compatible with gene regulation, H2A.Z variants may provide a more “static” mechanism to configure genes for a specific level of transcriptional activity. In the endocycling cells of the *O. dioica* epithelium, where predominantly genes for house production are expressed, once established, gene activity patterns are essentially maintained for the rest of the animal’s life. Thus, in mitotically inactive, differentiated cells the deposition of H2A.Z splice variants could facilitate the establishment of more permanent marks to modulate transcriptional activity. Furthermore, our discovery that putative alternative splice sites exist in the H2A.Z genes of a wide range of organisms raises the interesting possibility that differentially charged H2A.Z splice variants exist in other species.

5.3. Histone variants of the male germline in *O. doica* – evolution in action

Another interesting feature of the histone complement of *O. dioica* is the notable high number of histone sequence variants that were either exclusive to or significantly enriched in the male germline, including three different H3 proteins and, remarkably, a variant of the H4 family. Testes-specific H3 variants appear to be rare in animals and have only been reported for rat and human (Trostle-Weige et al., 1984; Witt et al., 1996), while no histone variants seem to exist in the H4 family of other multicellular organisms. The function of male germline-specific histones is not well understood and may vary, depending on the exact timing of their expression during spermatogenesis. Thus far, the replacement of canonical histones with male-specific variants has been associated with the erasure and re-establishment of the paternal epigenetic state and the chromatin organization of late spermatid condensation (Gaucher et al., 2010).

A notable feature of "male" histone variants across different species is their divergence in amino acid sequence compared to their canonical counterparts and the fact that they have diversified more rapidly within different lineages than somatic histones. This is particularly apparent in our model. Among all of the different core histone sequence variants, 15 of 31 were significantly enriched in the male gonad of *O. dioica* and they appear to have diverged to a greater extent from the canonical histones than what has been described for the testes-specific histones of vertebrates (paper I). The rapid evolution of histone H2B variants, which are all restricted to the male germinal cell lineage, has been explained as a consequence of neofunctionalization and subfunctionalization events after gene duplication (Gonzalez-Romero et al., 2010). However, why male germline histone variants diverge so rapidly and appear to "escape" the usually high selective pressure on the histone amino acid sequence is not clear. Are the structural/functional constraints during spermatogenesis more relaxed or does the sequence variability in these histone variants reflect specificity? Some data suggest it could potentially be a combination of the two that drives histone evolution in the male germline. At least for the testes

H2B variants there is support that a selection mechanism exists that maintains a biased amino acid composition in these variants which might be related to their role in the reorganization of chromatin during spermatogenesis (Gonzalez-Romero et al., 2010). Nevertheless, this may not explain why such variants did not arise in other tissues where chromatin undergoes extensive remodeling. So far, three conclusions about the properties of testes-specific histones can be drawn from the few studies on mammalian variants that exist to date: i) They are expressed at specific time windows during spermatogenesis (Gaucher et al., 2010), ii) they can be deposited at specific chromatin regions (Govin et al., 2007) and iii) they appear to destabilize the structure of the nucleosome (Gautier et al., 2004; Li et al., 2005; Tachiwana et al., 2010). Likewise, the male gonad-specific histones of *O. dioica* were expressed at specific time windows during male maturation and some of their predicted structural properties indicate a destabilizing role for these histones in the "male nucleosome" (paper I). The co-activation of several male-specific histone clusters at a precise time in the urochordate could be an indication that nucleosomes are evicted and replaced on a large scale. The exact onset of spermiogenesis in *O. dioica* is not known but the late expression of some of the male histone clusters suggests that the incorporation of these histones, as in other species, may occur after the completion of meiosis when transcription is very low and processes such as mitosis and DNA replication are absent. While the absence of these processes may lead to more relaxed evolutionary constraints on the nucleosome and chromatin structure, the synchronous and large scale eviction of previous histones at the same time might select for "destabilizing" histones. Several evolutionary "routes" might lead to histones forming less stable nucleosomes which could explain the lineage-specific diversification of these histones. In support for partly relaxed evolutionary constraints on histones during spermatogenesis is also the notion that we found possible polymorphisms within the N-terminal region of the male gonad-specific H3 variants of *O. dioica* (e.g. Ile to either Leu or Phe in H3t.3 at position 62), indicating that male histone variants appear to diversify very rapidly in *O. dioica*. Another conclusion that can be drawn from these observations is that even within this rapidly evolving urochordate, the amino acid sequence of the universal somatic variants and canonical histones remain

extremely conserved, demonstrating the high evolutionary constraints that preserve these histone structures within the chordates.

The simplest explanation of why *O. dioica* appears to be the only multicellular organism with a male-specific H4 variant might be that they have not yet been identified. Indeed, testes-specific expression of H4 genes uncoupled from replication in meiosis also occurs in mammals (Wolfe and Grimes, 1991) and genes encoding for putative H4 sequence variants can be found in the *Xenopus tropicalis*, *C. savigny* and *C. elegans* genomes through inspection of genome sequence data, although expression of such loci in these organisms has not been confirmed. Interestingly, the *X. tropicalis* H4 gene (Ensemble, scaffold_188 1555479 to 1555724) also contains a residue change from Ala to Thr at position 69 (the substitution in *O. dioica* comprises a change from Ala to Ser). Based on the histone structure, Ser69 resides within the alpha 2 helix and is not solvent accessable but faces towards the alpha 3 of H4. It is not clear whether substitutions at this position can invoke any structural changes within the nucleosome. However, the identification and expression analysis of further H4 genes encoding for H4 sequence isoforms in other organisms could be an interesting goal for future studies.

5.4. Histone variants of the male germline – are they epigenetic marks?

The fact that some of the male germline-specific and canonical histone variants are not replaced by protamines but are retained within the mature sperm in some organisms has further raised the interesting possibility that they may mark the paternal genome and transfer epigenetic information to the zygote. A recent analysis of human sperm chromatin has for example shown that nucleosomes retained in sperm are significantly enriched at many loci important for embryo development, including genes of key embryonic transcription factors and signaling pathway proteins (Hammoud et al., 2009). We also found most male gonad-specific variants and canonical histones to be present in the mature sperm of *O. dioica*, suggesting that beyond a possible function in remodeling and structural organization during

spermatogenesis, they could potentially play a role in the epigenetic control of embryogenesis (paper I).

Although there are few studies on the histone variants specific to the male germline, there is evidence from other histone variants that may indeed serve an epigenetic function. The variant CenH3 is a prominent example as it is retained at the centromeres of mammalian sperm chromatin, suggesting that it organizes the centromeres of the paternal genome during early embryogenesis (Palmer et al., 1990). The genome wide distribution of H3.3 and nuclear transplantation experiments in *Xenopus* support the idea that the distribution of H3.3 reflects patterns of active chromatin and that H3.3 might function as a potential epigenetic mark through which active gene states can be inherited (Mito et al., 2005; Ng and Gurdon, 2008; Schwartz and Ahmad, 2005; Wirbelauer et al., 2005). H3.3 also appears to play a role in the genome-wide remodeling of the decondensing male pronucleus in *C. elegans*, flies and mice (Orsi et al., 2009) and has been reported to replace canonical H3 in a locus-specific manner (Santenard submitted, cited in Santenard and Torres-Padilla, 2009). This suggests that H3.3 could establish epigenetic signatures in the male pronucleus after fertilization. How specific incorporation of H3.3 is achieved at those loci may provide a potential link to the function of sperm-specific histone variants. As male germline-specific histone H2A, H2B and H3 variants appear to generally cause a destabilization of the nucleosome (Gautier et al., 2004; Li et al., 2005; Tachiwana et al., 2010), they could potentially "label" distinct loci for replacement by H3.3 and facilitate histone eviction at those sites as the sperm chromatin decondenses. However, no functional studies of testes-specific histone variants have yet been performed and it remains to be seen whether their presence in the male pronucleus is essential.

5.5. Centromeric chromatin signatures in the presence and absence of mitosis

The switch from mitosis to endoreduplication in *O. dioica* also provides an interesting opportunity to study the cell cycle regulated deposition of histones. The

variant CenH3 determines the function and epigenetic inheritance of centromeres and malfunctioning CenH3 deposition usually results in lethality (Torras-Llort et al., 2009). However, how the variant is targeted specifically to centromeres in different organisms is not well understood, since the incorporation of CenH3 can occur independently of the underlying DNA sequence. One important mechanism thought to contribute to CenH3 incorporation at centromeres, is the specific cell-cycle timing of CenH3 deposition which occurs during late mitosis in human and flies (Hemmerich et al., 2008; Jansen et al., 2007; Schuh et al., 2007). This suggests that the mitotic phase of the cell cycle is crucial for CenH3 deposition in some organisms and led us to hypothesize that during endocycling, CenH3-chromatin is constantly diluted and gaps are not replenished with new CenH3 due to the absence of mitotic stages. Our data suggest that this may indeed be the case and even when a GFP-tagged CenH3 protein is abundant, we do not see a renewal of CenH3 at the centromere of endocycling nuclei (paper III). However, it is not clear whether this is due to the lack of the deposition machinery because the mitotic cell cycle phase is bypassed or there is no requirement for CenH3 due to an underreplication of centromeric DNA. Endocycling cell cycles vary among species in that some retain hallmarks of mitosis while other examples lack mitotic remnants including chromosome condensation, nuclear envelope breakdown and reorganization of microtubules (Edgar and Orr-Weaver, 2001). Consequently, the centromere dynamics during endocycles in different species are variable. In *Arabidopsis* the organization of endoreduplicated sister centromeres for example even depends on the type of endoreduplicating tissue (Fang and Spector, 2005). While in root epidermal cells, CenH3-tagged centromere foci are normally bigger, and some foci with irregular shapes are composed of two to three clustered fluorescent spots, the CenH3 foci of the larger leaf epidermal pavement cells CenH3 foci become smaller and continuously increases in number, indicating that endoreduplicated sister centromeres in these cells are more disassociated than those in the root epidermal cells. Studying CenH3 deposition in the endocycling tissues of different species could therefore provide important information about the underlying mechanisms essential for CenH3 deposition.

The centromeric and pericentromeric chromatin of different species also contains distinct histone modification marks that might play a role in centromere determination and function. H3S28 and H3T3 phosphorylation (H3S28P/T3P) are interesting candidates in *O. dioica* since we found these modifications to be particularly enriched at centromeres (paper III). Putative ectopic CenH3 foci resulting from CenH3-overexpression are likewise Ser28 phosphorylated in *O. dioica*. However, hyperphosphorylation of Ser28 throughout the chromosome did not lead to an obvious spreading of CenH3 localization, although the number of foci observed exceeded that of diploid cells. This suggests at least that H3S28P alone may not determine CenH3 deposition.

The function of centromeric H3 phosphorylation is not well studied and H3S28 phosphorylation at centromeres has not been reported elsewhere. It is possible that H3S28P also plays a role in sister chromatid cohesion as has been reported for H3T3P which is enriched at the centromeres of human cell lines (Dai et al., 2005; Higgins, 2010). However, H3TSP and H3S28P at centromeres exhibited slightly different temporal kinetics in *O. dioica*, suggesting that their functions may not entirely overlap. Alternatively, it has been suggested that mitotic PTMs might also function as a "ready" production label or as a "licensing system" that distinguishes chromosomes that have successfully passed through the metaphase-anaphase checkpoint (Hans and Dimitrov, 2001). The dynamics of H3.3S31P (paper II) and H3S28P (suppl. paper I) during female meiosis in *O. dioica* for example provide indirect support for the ready production label hypothesis as these marks are retained only on those nuclei of the coenocyst that seed oocytes and complete meiosis. The remaining unselected meiotic nuclei lose these labels and fail to complete meiosis. Thus, one might also envision a possible role for H3S28P in mitotic checkpoint regulation in *O. dioica*.

5.6. Plasticity in histone modifications across species – a lineage-specific "mitotic histone code"?

Several observations have recently lead to the proposal that transient mitotic histone marks may constitute a lineage-specific "mitotic histone code". These include the fact that the temporal and spatial patterns of mitotic histone modifications, particularly H3 phosphorylations, vary widely across species. H3 phosphorylation has been associated with mitotic chromosome condensation, sister chromatid cohesion and, partly, gene activation in interphase nuclei (Perez-Cadahia et al., 2009) but the subject is still a matter of debate. The data on H3 phosphorylation available so far suggests that phosphorylation of the same residues may not have the same function in different species (Cerutti and Casas-Mollano, 2009). Our observations in the metaphase chromosomes of *O. dioica* confirm that there is indeed very little conservation in the localization of H3 phosphorylation marks (paper II and paper III) between species. Another conclusion that can be drawn from our data is that this is not only true for H3 phosphorylation but also for other cell cycle regulated marks such as the trimethylation of H3K9 ($H3K9me_3$) and dimethylation of H3K4 ($H3K4me_2$). The "active" $H3K4me_2$ mark co-localizes with the CenH3-chromatin of metaphase chromosomes in *Drosophila* and human (Sullivan and Karpen, 2004) while the heterochromatic $H3K9me_3$ mark is enriched at pericentromeres, but is excluded from centromeric chromatin in these species (Peters et al., 2003). Both marks are associated with opposing function with respect to transcriptional activation and it has been proposed that they may determine the chromatin borders of centromeres. However, the broad distribution observed for both marks and the exclusion of $H3K4me_2$ from centromeres in the mitotic chromosomes *O. dioica* suggest that the function of these marks, at least during mitosis, may not be conserved (paper III). More recently it has become clear that different H3 phosphorylation marks may not necessarily act as "point" signals but rather depend on co-existing modifications on the same H3 tail that could confer distinct "read outs" of these combinational marks. A recent report for example suggests that H3T3P in human cells is always found within a combinatorial modification pattern with $H3K4me_3$ and

H3R8me$_2$ in mitotic cells (Markaki et al., 2009). Considering that neither H3K4me$_2$ (paper III) nor H3K4me$_3$ (Coen Campsteijn, personal communication) did extend into the H3T3P positive centromeres of *O. dioica*, raises the interesting possibility is that combinational marks may also vary among organisms and create lineage-specific read outs. In support of this view is the notion that the role for the dual H3S10P/H3K9me$_3$ in the mitotic release of HP1 does not seem to have been conserved in the plant lineage. H3S10P, in combination to H3K9me$_3$ in mammals, is believed to provide a "mitotic binary switch", which determines recruitment or eviction of heterochromatin protein 1 (HP1) from mitotic chromatin, whereby the addition of Ser10 phosphorylation ejects the HP1 that is bound to H3K9me$_3$ (Dormann et al., 2006; Hirota et al., 2005). The plant homolog of HP1 however, does not bind H3K9me$_3$ but instead binds trimethylated H3K27 (Turck et al., 2007) raising some doubts about a conserved role of H3K9 methylation in HP1 recruitment. It thus seems that the chromosomal localization, the timing and putative functions of H3 phosphorylation and other transient mitotic marks, with the possible exception of H3T3P, may have diverged substantially and a lineage specific mitotic histone code exists.

5.7. Future goals and perspectives

O. dioica is an attractive new model system for a variety of chromatin studies. We have now established a histone-eGFP fusion expression system to study the distribution of histone variants in embryos and young adults and future studies will address the localization of other histone variants including the divergent H2A.3/.4 histones of *O. dioica.*

Chromatin immunoprecipitation (ChIP) and ChIP-chip experiments are already successfully underway to determine the distribution of chromatin marks over the compacted regulatory regions and open reading frames of *O. dioica* genes and to study the distribution of these marks on a genome-wide scale. We further plan to extend these experiments to ChIP-seq studies in combination with *O. dioica* histone variant specific antibodies. An antibody against the H4t variant monomethtylated at

Lys20 used in ChIP-chip and ChIP-seq on mature sperm chromatin will for example enable us to determine whether this modified variant localizes to developmentally important loci.

Additionally, the establishment of ChIP-chip experiments presents an opportunity to study intriguing epigenetic questions in *O. dioica*. We have developed large scale experimental set up for "density stress conditions" that induces growth arrest at D3/D4 (prior to gonad maturation) in our model. Animals with stalled growth can be kept under these conditions for several days/weeks, which can prolong the normal life span 3-fold (Coen Campsteijn and David Osborne, personal communication). Animals "released" from these conditions will mature and spawn normally. In the rapidly evolving urochordate, the genome-wide expression patterns of subsequent generations derived from "stressed" parents may reveal whether environmental conditions can affect the epigenetic state of the germline.

To further test the interplay between the different mitotic chromatin marks we have generated an H3.2S28A-mutant expression vector (paper III, Fig. S2B). Cells that exhibited a high expression for the tagged H3.2S28A mutant showed no H3S28P signal (appendix, Fig. A1), suggesting that the Ser28-mutated H3.2 protein replaces endogenous H3.2 and is able to drown out the H3S28P signal. With this construct we plan to test the effect of H3S28P abrogation on CenH3 deposition and the spatial distribution of other histone modifications during mitosis. Additional clonings are currently underway to generate a H3.2S28E plasmid where the Ser28 is replaced by the residue glutamate, which "mimics" a phosphorylated Ser28. We expect that deposition of this mutant will mimic a Ser28 phosphorylation throughout the entire chromosome, thereby abrogating the centromeric restriction of H3S28P in a similar way as the PP1 inhibitor Calyculin A causes spreading of H3S28P. This construct however, will allow us to dissect the causative effects of H3S28P more directly. Incubating embryos in different chemicals provides another powerful tool to study the function of H3S28 phosphorylation in *O. dioica*. Our preliminary experiments with an Aurora B kinase inhibitor suggest for example that endocycling can be induced in the proliferating *O. dioica* cells (not shown).

Finally, the female and male gonads of *O. dioica* present some exquisite tools to study histone variant deposition and the function of histone modifications. Recombinant fusion proteins are now successfully expressed from cmRNA and plasmids when injected into the male and female gonad and we plan to inject the H3S28A and the H3S28E mutant constructs into the gonads of females. This experimental set up may provide direct evidence whether H3S28P is required for progression through meiosis after fertilization or whether maturation can be "forced" upon non-H3S28P-labeled oocytes with the overexpression of the H3S28E construct. Furthermore generating new constructs by mutating the H4t variant-specific residues will provide clues whether these residue determine a specific localization of this variant in male germline chromatin.

The results presented in this thesis provide new insights into the diversity of histone variants and their developmental-specific regulation that exists among the chordates. Histone variants assume crucial roles in gene expression, centromere function, DNA repair, and sperm compaction and the causal relationships between dysregulation of some histone variants and tumorigenesis and infertility are already well established. Novel histone variants are continuously being discovered in mammals and other species and their studies provide an important contribution to understand the function behind histone variability and their implications in diseases. With this work we have developed new tools to study the deposition and the function of novel and universal histone variants and their histone modifications in the different cell cycles of the urochordate *O. dioica*.

6. Appendix

Table A1A. Modifying enzymes responsible for histone methylation, acetylation and phosphorylation in *O. dioica*.

Protein Family	Modifying enzyme	Histone substrate (residue numbers of mammalian histones)	Isoforms identified in *O. dioica*
Acetyltransferases			
GNAT	GCN5	H3, H4	Not found*
	PCAF	H3, H4K8	1
	HAT1	H4K12, H4K5	1
CBP/p300	CBP/P300	H3, H4	1
Myst	Myst3	H3K14	1
	ESA1/Myst1/Kat5	H3, H4, H2A	2
	Tip60	H3K14	1
BET	BRDT (testes)		2
BR140	Brpf1		2
Nuclear receptor coactivator family	SRC1		Not found
Nuclear receptor coactivator family	ACTR		Not found
Methyltransferases			
Suv3-9	Suv39	H3K9, H4K79	2
RMT type I	Carm1	H3R2, H3R17, H3R26	1
Dot1	Dot1	H3K79	1
TRX	MLL4	H3K4	2
	Set1	H3K4	1
	SETDB1	H3K4	1
SET3/4	MLL5	H3K4	1
RMT type II	PMRT5	H3R8, H4R3	1
EZ	EZH2	H3K27, H3K9	1
Set8	Set7/8	$H4K20Me_1$	1
Suv4-20	Suv4-20	$H4K20Me_2/Me_3$	2
	SETMar		1
Set	Set2/ NSD1/WHSC1	H3K36, H4K20	2
	EHMT1a	H3K9	2
SMYD	SMYD5		1
	SMYD3	H3K4	2
Kinases			
	Haspin	H3T3	1
Aurora	Aurora A	H3S10, H3S28	1
	Aurora B	H3S10, H3S28	1
PIKK	ATM/ATR	H2AX	1
PIKK	DNA-PK	H2AX	Not found*

*Enzymes not found in *O. dioica* but identified in *Ciona. Ciona* protein sequences were used as queries to search the Od genome.

Table A1B Demodifying *O. dioica* enzymes removing histone methylation-, acetylation- and phosphorylation.

Protein Family	Demodifying enzyme	Residue targets identified so far	Isoforms identified in *O. dioica*
Deacetylases			
Sirtuins	SIRT1	H3K9, H4K16	1
	SIRT1	H3K9, H4K16	1
	SIRT2	H4K16	1
	SIRT5		1
	SIRT4		1
	SIRT6		1
	SIRT7		1
HDAC	HDAC8		2
	HDAC7		1
	HDAC1		1
	HDAC3		1
Demethylases			
JMjC	JHDM1	H3K36	1
	UTX/UTY	H3 K27	3
	JMJD3	H3K9, H3K36	1
	KDM7A	H3K9, H3K27	1
LSD	LSD1	H3K4, H3K9	2
PAD4	PAD4	H3R26, H3R17, H4R3	1

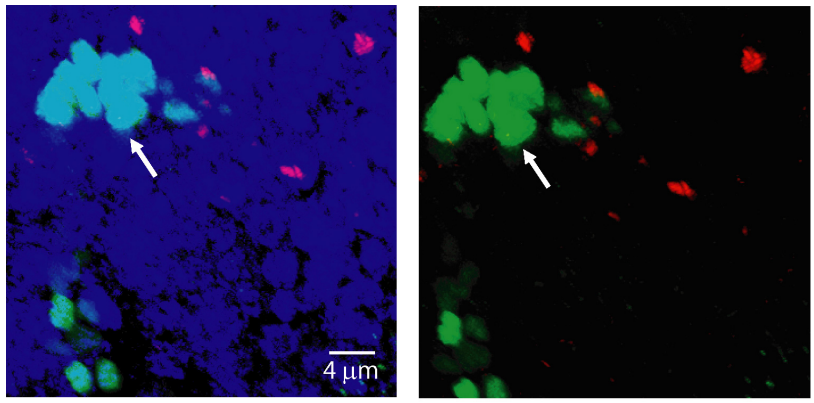

Figure A1 Overexpression of a H3.2-DsRed fusion protein containing a mutation at Ser28 to Ala28. Immunostaining of embryos with an anti DsRed- (green) and anti H3S28P antibody after microinjection of the pH3A28-DsRed expression plasmid. Mitotic cells that were successfully transfected expressed a mutated form of canonical H3.2 histone where Ser28 had been replaced by Ala28. The strength of H3.2A28-DsRed signal varied between transfected cells but cells with a high enrichment for the mutated H3.2 never showed a H3S28P signal.

7. References

Abbott, D.W., Ivanova, V.S., Wang, X., Bonner, W.M., and Ausio, J. (2001). Characterization of the stability and folding of H2A.Z chromatin particles: implications for transcriptional activation. J Biol Chem *276*, 41945-41949.

Ahmad, K., and Henikoff, S. (2002). Histone H3 variants specify modes of chromatin assembly. Proc Natl Acad Sci U S A *99 Suppl 4*, 16477-16484.

Akasaka, T., Kanno, M., Balling, R., Mieza, M.A., Taniguchi, M., and Koseki, H. (1996). A role for mel-18, a Polycomb group-related vertebrate gene, during theanteroposterior specification of the axial skeleton. Development *122*, 1513-1522.

Allan, J., Hartman, P.G., Crane-Robinson, C., and Aviles, F.X. (1980). The structure of histone H1 and its location in chromatin. Nature *288*, 675-679.

Alldredge, A. (1976). Discarded Appendicularian houses as sources of food, surface habitats, and particulate organic matter in planktonic environments. Limnol Oceanogr *21*, 14–23.

Angelov, D., Molla, A., Perche, P.Y., Hans, F., Cote, J., Khochbin, S., Bouvet, P., and Dimitrov, S. (2003). The histone variant macroH2A interferes with transcription factor binding and SWI/SNF nucleosome remodeling. Mol Cell *11*, 1033-1041.

Annunziato, A.T., Schindler, R.K., Riggs, M.G., and Seale, R.L. (1982). Association of newly synthesized histones with replicating and nonreplicating regions of chromatin. J Biol Chem *257*, 8507-8515.

Arents, G., Burlingame, R.W., Wang, B.C., Love, W.E., and Moudrianakis, E.N. (1991). The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-handed superhelix. Proc Natl Acad Sci U S A *88*, 10148-10152.

Arya, G., and Schlick, T. (2006). Role of histone tails in chromatin folding revealed by a mesoscopic oligonucleosome model. Proc Natl Acad Sci U S A *103*, 16236-16241.

Aul, R.B., and Oko, R.J. (2002). The major subacrosomal occupant of bull spermatozoa is a novel histone H2B. Dev Biol *242*, 376-387.

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

Bao, Y., Konesky, K., Park, Y.J., Rosu, S., Dyer, P.N., Rangasamy, D., Tremethick, D.J., Laybourn, P.J., and Luger, K. (2004). Nucleosomes containing the histone variant H2A.Bbd organize only 118 base pairs of DNA. EMBO J *23*, 3314-3324.

Bartova, E., Krejci, J., Harnicarova, A., Galiova, G., and Kozubek, S. (2008). Histone modifications and nuclear architecture: a review. J Histochem Cytochem *56*, 711-721.

Benson, L.J., Gu, Y., Yakovleva, T., Tong, K., Barrows, C., Strack, C.L., Cook, R.G., Mizzen, C.A., and Annunziato, A.T. (2006). Modifications of H3 and H4 during chromatin replication, nucleosome assembly, and histone exchange. J Biol Chem *281*, 9287-9296.

Bergmann, A. (2010). The role of ubiquitylation for the control of cell death in Drosophila. Cell Death Differ *17*, 61-67.

Bernad, R., Sanchez, P., and Losada, A. (2009). Epigenetic specification of centromeres by CENP-A. Exp Cell Res *315*, 3233-3241.

Bernstein, E., and Hake, S.B. (2006). The nucleosome: a little variation goes a long way. Biochem Cell Biol *84*, 505-517.

Bjerga, G.E.K. (2009). The duality of p300: The lysine binding domains communicate with the acetyltransferase domain (University of Bergen).

Black, J.C., Mosley, A., Kitada, T., Washburn, M., and Carey, M. (2008). The SIRT2 deacetylase regulates autoacetylation of p300. Mol Cell *32*, 449-455.

Blower, M.D., Sullivan, B.A., and Karpen, G.H. (2002). Conserved organization of centromeric chromatin in flies and humans. Dev Cell *2*, 319-330.

Boisvert, F.M., Chenard, C.A., and Richard, S. (2005). Protein interfaces in signaling regulated by arginine methylation. Sci STKE *2005*, re2.

Boulard, M., Gautier, T., Mbele, G.O., Gerson, V., Hamiche, A., Angelov, D., Bouvet, P., and 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*, 1518-1526.

Bouquet, J.M., Spriet, E., Troedsson, C., Ottera, H., Chourrout, D., and Thompson, E.M. (2009). Culture optimization for the emergent zooplanktonic model organism Oikopleura dioica. J Plankton Res *31*, 359-370.

Brickner, D.G., Cajigas, I., Fondufe-Mittendorf, Y., Ahmed, S., Lee, P.C., Widom, J., and Brickner, J.H. (2007). H2A.Z-mediated localization of genes at the nuclear periphery confers epigenetic memory of previous transcriptional state. PLoS Biol *5*, e81.

Brownell, J.E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D.G., Roth, S.Y., and Allis, C.D. (1996). Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. Cell *84*, 843-851.

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

Baatout, S., and Derradji, H. (2006). About histone H1 phosphorylation during mitosis. Cell Biochem Funct *24*, 93-94.

Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.J., Anderson, S., Yates, J., Washburn, M.P*., et al.* (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell *123*, 581-592.

Celeste, A., Fernandez-Capetillo, O., Kruhlak, M.J., Pilch, D.R., Staudt, D.W., Lee, A., Bonner, R.F., Bonner, W.M., and Nussenzweig, A. (2003). Histone H2AX phosphorylation is dispensable for the initial recognition of DNA breaks. Nat Cell Biol *5*, 675-679.

Cerutti, H., and Casas-Mollano, J.A. (2009). Histone H3 phosphorylation: universal code or lineage specific dialects? Epigenetics *4*, 71-75.

Chadwick, B.P., and Willard, H.F. (2001). A novel chromatin protein, distantly related to histone H2A, is largely excluded from the inactive X chromosome. J Cell Biol *152*, 375-384.

Chadwick, B.P., and Willard, H.F. (2002). Cell cycle-dependent localization of macroH2A in chromatin of the inactive X chromosome. J Cell Biol *157*, 1113-1123.

Chen, Y., Sprung, R., Tang, Y., Ball, H., Sangras, B., Kim, S.C., Falck, J.R., Peng, J., Gu, W., and Zhao, Y. (2007). Lysine propionylation and butyrylation are novel post-translational modifications in histones. Mol Cell Proteomics *6*, 812-819.

Chioda, M., Eskeland, R., and Thompson, E.M. (2002). Histone gene complement, variant expression, and mRNA processing in a urochordate Oikopleura dioica that undergoes extensive polyploidization. Mol Biol Evol *19*, 2247-2260.

Choi, J.K., and Howe, L.J. (2009). Histone acetylation: truth of consequences? Biochem Cell Biol *87*, 139-150.

Churikov, D., Zalenskaya, I.A., and Zalensky, A.O. (2004). Male germline-specific histones in mouse and man. Cytogenet Genome Res *105*, 203-214.

Clarkson, M.J., Wells, J.R., Gibson, F., Saint, R., and Tremethick, D.J. (1999). Regions of variant histone His2AvD required for Drosophila development. Nature *399*, 694-697.

Cloos, P.A., Christensen, J., Agger, K., and Helin, K. (2008). Erasing the methyl mark: histone demethylases at the center of cellular differentiation and disease. Genes Dev *22*, 1115-1140.

Collins, R.E., Tachibana, M., Tamaru, H., Smith, K.M., Jia, D., Zhang, X., Selker, E.U., Shinkai, Y., and Cheng, X. (2005). In vitro and in vivo analyses of a Phe/Tyr switch controlling product specificity of histone lysine methyltransferases. J Biol Chem *280*, 5563-5570.

Couture, J.F., Collazo, E., Brunzelle, J.S., and Trievel, R.C. (2005). Structural and functional analysis of SET8, a histone H4 Lys-20 methyltransferase. Genes Dev *19*, 1455-1465.

Cremer, T., Kurz, A., Zirbel, R., Dietzel, S., Rinke, B., Schrock, E., Speicher, M.R., Mathieu, U., Jauch, A., Emmerich, P*., et al.* (1993). Role of chromosome territories in the functional compartmentalization of the cell nucleus. Cold Spring Harb Symp Quant Biol *58*, 777-792.

Cunliffe, V.T. (2008). Eloquent silence: developmental functions of Class I histone deacetylases. Curr Opin Genet Dev *18*, 404-410.

Dai, J., Sullivan, B.A., and Higgins, J.M. (2006). Regulation of mitotic chromosome cohesion by Haspin and Aurora B. Dev Cell *11*, 741-750.

Dai, J., Sultan, S., Taylor, S.S., and Higgins, J.M. (2005). The kinase haspin is required for mitotic histone H3 Thr 3 phosphorylation and normal metaphase chromosome alignment. Genes Dev *19*, 472-488.

Dalal, Y., Wang, H., Lindsay, S., and Henikoff, S. (2007). Tetrameric structure of centromeric nucleosomes in interphase Drosophila cells. PLoS Biol *5*, e218.

de la Barre, A.E., Gerson, V., Gout, S., Creaven, M., Allis, C.D., and Dimitrov, S. (2000). Core histone N-termini play an essential role in mitotic chromosome condensation. EMBO J *19*, 379-391.

de Ruijter, A.J., van Gennip, A.H., Caron, H.N., Kemp, S., and van Kuilenburg, A.B. (2003). Histone deacetylases (HDACs): characterization of the classical HDAC family. Biochem J *370*, 737-749.

Delsuc, F., Brinkmann, H., Chourrout, D., and Philippe, H. (2006). Tunicates and not cephalochordates are the closest living relatives of vertebrates. Nature *439*, 965-968.

Dillon, S.C., Zhang, X., Trievel, R.C., and Cheng, X. (2005). The SET-domain protein superfamily: protein lysine methyltransferases. Genome Biol *6*, 227.

Dorigo, B., Schalch, T., Bystricky, K., and Richmond, T.J. (2003). Chromatin fiber folding: requirement for the histone H4 N-terminal tail. J Mol Biol *327*, 85-96.

Dormann, H.L., Tseng, B.S., Allis, C.D., Funabiki, H., and Fischle, W. (2006). Dynamic regulation of effector protein binding to histone modifications: the biology of HP1 switching. Cell Cycle *5*, 2842-2851.

Doyen, C.M., An, W., Angelov, D., Bondarenko, V., Mietton, F., Studitsky, V.M., Hamiche, A., Roeder, R.G., Bouvet, P., and Dimitrov, S. (2006). Mechanism of polymerase II transcription repression by the histone variant macroH2A. Mol Cell Biol *26*, 1156-1164.

Edgar, B.A., and Orr-Weaver, T.L. (2001). Endoreplication cell cycles: more for less. Cell *105*, 297-306.

Fang, Y., and Spector, D.L. (2005). Centromere positioning and dynamics in living Arabidopsis plants. Mol Biol Cell *16*, 5710-5718.

Fenaux, R. (1985). Rhythm of secretion of oikopleurid‘s houses. Bull Mar Sci *37*, 498–503.

Fenaux, R. (1998). Anatomy and functional morphology of the Appendicularia. In Bone, Q., The Biology of Pelagic Tunicates (New York, Oxford University Press).

Fernandez-Capetillo, O., Mahadevaiah, S.K., Celeste, A., Romanienko, P.J., Camerini-Otero, R.D., Bonner, W.M., Manova, K., Burgoyne, P., and Nussenzweig, A. (2003). H2AX is required for chromatin remodeling and inactivation of sex chromosomes in male mouse meiosis. Dev Cell *4*, 497-508.

Fink, M., Imholz, D., and Thoma, F. (2007). Contribution of the serine 129 of histone H2A to chromatin structure. Mol Cell Biol *27*, 3589-3600.

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

Flood, P.R., and Deibel, D. (1998). The appendicularian house. In Bone, Q., The Biology of Pelagic Tunicates (New York, Oxford University Press).

Foltz, D.R., Jansen, L.E., Black, B.E., Bailey, A.O., Yates, J.R., 3rd, and Cleveland, D.W. (2006). The human CENP-A centromeric nucleosome-associated complex. Nat Cell Biol *8*, 458-469.

Fujiki, R., Chikanishi, T., Hashiba, W., and Kato, S. (2010). [Role of nuclear O-glycosylation in epigenetic regulation]. Tanpakushitsu Kakusan Koso *55*, 61-68.

Furuyama, T., and Henikoff, S. (2009). Centromeric nucleosomes induce positive DNA supercoils. Cell *138*, 104-113.

Ganot, P., Bouquet, J.M., Kallesoe, T., and Thompson, E.M. (2007a). The Oikopleura coenocyst, a unique chordate germ cell permitting rapid, extensive modulation of oocyte production. Dev Biol *302*, 591-600.

Ganot, P., Kallesoe, T., and Thompson, E.M. (2007b). The cytoskeleton organizes germ nuclei with divergent fates and asynchronous cycles in a common cytoplasm during oogenesis in the chordate Oikopleura. Dev Biol *302*, 577-590.

Ganot, P., and Thompson, E.M. (2002). Patterning through differential endoreduplication in epithelial organogenesis of the chordate, Oikopleura dioica. Dev Biol *252*, 59-71.

Garcia-Dominguez, M., and Reyes, J.C. (2009). SUMO association with repressor complexes, emerging routes for transcriptional control. Biochim Biophys Acta *1789*, 451-459.

Garcia, B.A., Hake, S.B., Diaz, R.L., Kauer, M., Morris, S.A., Recht, J., Shabanowitz, J., Mishra, N., Strahl, B.D., Allis, C.D.*, et al.* (2007). Organismal differences in post-translational modifications in histones H3 and H4. J Biol Chem *282*, 7641-7655.

Gatewood, J.M., Cook, G.R., Balhorn, R., Schmid, C.W., and Bradbury, E.M. (1990). Isolation of four core histones from human sperm chromatin representing a minor subset of somatic histones. J Biol Chem *265*, 20662-20666.

Gaucher, J., Reynoird, N., Montellier, E., Boussouar, F., Rousseaux, S., and Khochbin, S. (2010). From meiosis to postmeiotic events: the secrets of histone disappearance. FEBS J *277*, 599-604.

Gautier, T., Abbott, D.W., Molla, A., Verdel, A., Ausio, J., and Dimitrov, S. (2004). Histone variant H2ABbd confers lower stability to the nucleosome. EMBO Rep *5*, 715-720.

Ghosh, S., and Klobutcher, L.A. (2000). A development-specific histone H3 localizes to the developing macronucleus of Euplotes. Genesis *26*, 179-188.

Gladyshev, E., and Meselson, M. (2008). Extreme resistance of bdelloid rotifers to ionizing radiation. Proc Natl Acad Sci U S A 105, 5139-5144.

Gonzalez-Romero, R., Rivera-Casas, C., Ausio, J., Mendez, J., and Eirin-Lopez, J.M. (2010). Birth-and-death long-term evolution promotes histone H2B Variant diversification in the male germinal cell line. Mol Biol Evol 27, 1802-1812.

Goto, H., Yasui, Y., Nigg, E.A., and Inagaki, M. (2002). Aurora-B phosphorylates Histone H3 at serine28 with regard to the mitotic chromosome condensation. Genes Cells *7*, 11-17.

Govin, J., Escoffier, E., Rousseaux, S., Kuhn, L., Ferro, M., Thevenon, J., Catena, R., Davidson, I., Garin, J., Khochbin, S.*, et al.* (2007). Pericentric heterochromatin reprogramming by new histone variants during mouse spermiogenesis. J Cell Biol *176*, 283-294.

Groth, A. (2009). Replicating chromatin: a tale of histones. Biochem Cell Biol *87*, 51-63.

Groth, A., Rocha, W., Verreault, A., and Almouzni, G. (2007). Chromatin challenges during DNA replication and repair. Cell *128*, 721-733.

Hackett, J.D., Scheetz, T.E., Yoon, H.S., Soares, M.B., Bonaldo, M.F., Casavant, T.L., and Bhattacharya, D. (2005). Insights into a dinoflagellate genome through expressed sequence tag analysis. BMC Genomics *6*, 80.

Hake, S.B., and Allis, C.D. (2006). Histone H3 variants and their potential role in indexing mammalian genomes: The "H3 barcode hypothesis". Proc Natl Acad Sci U S A *103*, 6428-6435.

Hake, S.B., Garcia, B.A., Duncan, E.M., Kauer, M., Dellaire, G., Shabanowitz, J., Bazett-Jones, D.P., Allis, C.D., and Hunt, D.F. (2006). Expression patterns and post-translational modifications associated with mammalian histone H3 variants. J Biol Chem *281*, 559-568.

Hake, S.B., Garcia, B.A., Kauer, M., Baker, S.P., Shabanowitz, J., Hunt, D.F., and Allis, C.D. (2005). Serine 31 phosphorylation of histone variant H3.3 is specific to regions bordering centromeres in metaphase chromosomes. Proc Natl Acad Sci U S A *102*, 6344-6349.

Hale, T.K., Contreras, A., Morrison, A.J., and Herrera, R.E. (2006). Phosphorylation of the linker histone H1 by CDK regulates its binding to HP1alpha. Mol Cell *22*, 693-699.

Hammoud, S.S., Nix, D.A., Zhang, H., Purwar, J., Carrell, D.T., and Cairns, B.R. (2009). Distinctive chromatin in human sperm packages genes for embryo development. Nature *460*, 473-478.

Hans, F., and Dimitrov, S. (2001). Histone H3 phosphorylation and cell division. Oncogene *20*, 3021-3027.

Hansen, J.C., Tse, C., and Wolffe, A.P. (1998). Structure and function of the core histone N-termini: more than meets the eye. Biochemistry *37*, 17637-17641.

Happel, N., and Doenecke, D. (2009). Histone H1 and its isoforms: contribution to chromatin structure and function. Gene *431*, 1-12.

Hashimoto, H., Sonoda, E., Takami, Y., Kimura, H., Nakayama, T., Tachibana, M., Takeda, S., and Shinkai, Y. (2007). Histone H1 variant, H1R is involved in DNA damage response. DNA Repair (Amst) *6*, 1584-1595.

Hassan, Y.I., and Zempleni, J. (2008). A novel, enigmatic histone modification: biotinylation of histones by holocarboxylase synthetase. Nutr Rev *66*, 721-725.

Hayashi, H., Nomoto, M., and Iwai, K. (1984). Tetrahymena histone H4. Complete amino acid sequences of two variants. J Biochem *96*, 1449-1456.

Hemmerich, P., Weidtkamp-Peters, S., Hoischen, C., Schmiedeberg, L., Erliandri, I., and Diekmann, S. (2008). Dynamics of inner kinetochore assembly and maintenance in living cells. J Cell Biol *180*, 1101-1114.

Higgins, J.M. (2003). Structure, function and evolution of haspin and haspin-related proteins, a distinctive group of eukaryotic protein kinases. Cell Mol Life Sci *60*, 446-462.

Higgins, J.M. (2010). Haspin: a newly discovered regulator of mitotic chromosome behavior. Chromosoma *119*, 137-147.

Hill, C.S., and Thomas, J.O. (1990). Core histone-DNA interactions in sea urchin sperm chromatin. The N-terminal tail of H2B interacts with linker DNA. Eur J Biochem *187*, 145-153.

Hirota, T., Lipp, J.J., Toh, B.H., and Peters, J.M. (2005). Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. Nature *438*, 1176-1180.

Honda, B.M., Candido, P.M., and Dixon, G.H. (1975). Histone methylation. Its occurrence in different cell types and relation to histone H4 metabolism in developing trout testis. J Biol Chem *250*, 8686-8689.

Hsu, J.Y., Sun, Z.W., Li, X., Reuben, M., Tatchell, K., Bishop, D.K., Grushcow, J.M., Brame, C.J., Caldwell, J.A., Hunt, D.F.*, et al.* (2000). Mitotic phosphorylation of histone H3 is governed by Ipl1/aurora kinase and Glc7/PP1 phosphatase in budding yeast and nematodes. Cell *102*, 279-291.

Ishibashi, T., Dryhurst, D., Rose, K.L., Shabanowitz, J., Hunt, D.F., and Ausio, J. (2009). Acetylation of vertebrate H2A.Z and its effect on the structure of the nucleosome. Biochemistry *48*, 5007-5017.

Ishibashi, T., Li, A., Eirin-Lopez, J.M., Zhao, M., Missiaen, K., Abbott, D.W., Meistrich, M., Hendzel, M.J., and Ausio, J. (2010). H2A.Bbd: an X-chromosome-encoded histone involved in mammalian spermiogenesis. Nucleic Acids Res *38*, 1780-1789.

Izzo, A., Kamieniarz, K., and Schneider, R. (2008). The histone H1 family: specific members, specific functions? Biol Chem *389*, 333-343.

Jablonka, E., and Lamb, M.J. (2002). The changing concept of epigenetics. Ann N Y Acad Sci *981*, 82-96.

Jackson, V. (1988). Deposition of newly synthesized histones: hybrid nucleosomes are not tandemly arranged on daughter DNA strands. Biochemistry *27*, 2109-2120.

Jackson, V., and Chalkley, R. (1981a). A new method for the isolation of replicative chromatin: selective deposition of histone on both new and old DNA. Cell *23*, 121-134.

Jackson, V., and Chalkley, R. (1981b). A reevaluation of new histone deposition on replicating chromatin. J Biol Chem *256*, 5095-5103.

Jansen, L.E., Black, B.E., Foltz, D.R., and Cleveland, D.W. (2007). Propagation of centromeric chromatin requires exit from mitosis. J Cell Biol *176*, 795-805.

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.

Johnson, L., Mollah, S., Garcia, B.A., Muratore, T.L., Shabanowitz, J., Hunt, D.F., and Jacobsen, S.E. (2004). Mass spectrometry analysis of Arabidopsis histone H3 reveals distinct combinations of post-translational modifications. Nucleic Acids Res *32*, 6511-6518.

Johnson, S.A., and Hunter, T. (2005). Kinomics: methods for deciphering the kinome. Nat Methods *2*, 17-25.

Jones, B., Su, H., Bhat, A., Lei, H., Bajko, J., Hevi, S., Baltus, G.A., Kadam, S., Zhai, H., Valdez, R.*, et al.* (2008). The histone H3K79 methyltransferase Dot1L is essential for mammalian development and heterochromatin structure. PLoS Genet *4*, e1000190.

Jones, R.S., and Gelbart, W.M. (1993). The Drosophila Polycomb-group gene Enhancer of zeste contains a region with sequence similarity to trithorax. Mol Cell Biol *13*, 6357-6366.

Kan, P.Y., Lu, X., Hansen, J.C., and Hayes, J.J. (2007). The H3 tail domain participates in multiple interactions during folding and self-association of nucleosome arrays. Mol Cell Biol *27*, 2084-2091.

Kantidze, O.L., and Razin, S.V. (2009). Chromatin loops, illegitimate recombination, and genome evolution. Bioessays *31*, 278-286.

Kimura, H., and Cook, P.R. (2001). Kinetics of core histones in living human cells: little exchange of H3 and H4 and some rapid exchange of H2B. J Cell Biol *153*, 1341-1353.

Kimura, M., Matsuda, Y., Yoshioka, T., and Okano, Y. (1999). Cell cycle-dependent expression and centrosome localization of a third human aurora/Ipl1-related protein kinase, AIK3. J Biol Chem *274*, 7334-7340.

Kleff, S., Andrulis, E.D., Anderson, C.W., and Sternglanz, R. (1995). Identification of a gene encoding a yeast histone H4 acetyltransferase. J Biol Chem *270*, 24674-24677.

Klose, R.J., and Bird, A.P. (2006). Genomic DNA methylation: the mark and its mediators. Trends Biochem Sci *31*, 89-97.

Lee, E.R., Murdoch, F.E., and Fritsch, M.K. (2007). High histone acetylation and decreased polycomb repressive complex 2 member levels regulate gene specific transcriptional changes during early embryonic stem cell differentiation induced by retinoic acid. Stem Cells *25*, 2191-2199.

Li, A., Maffey, A.H., Abbott, W.D., Conde e Silva, N., Prunell, A., Siino, J., Churikov, D., Zalensky, A.O., and Ausio, J. (2005). Characterization of nucleosomes consisting of the human testis/sperm-specific histone H2B variant (hTSH2B). Biochemistry *44*, 2529-2535.

Liu, X., Li, B., and GorovskyMa (1996). Essential and nonessential histone H2A variants in Tetrahymena thermophila. Mol Cell Biol *16*, 4305-4311.

Lowell, J.E., Kaiser, F., Janzen, C.J., and Cross, G.A. (2005). Histone H2AZ dimerizes with a novel variant H2B and is enriched at repetitive DNA in Trypanosoma brucei. J Cell Sci *118*, 5721-5730.

Lozano, E., Saez, A.G., Flemming, A.J., Cunha, A., and Leroi, A.M. (2006). Regulation of growth by ploidy in Caenorhabditis elegans. Curr Biol *16*, 493-498.

Lu, X., Simon, M.D., Chodaparambil, J.V., Hansen, J.C., Shokat, K.M., and Luger, K. (2008). The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure. Nat Struct Mol Biol *15*, 1122-1124.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251-260.

Luger, K., and Richmond, T.J. (1998). The histone tails of the nucleosome. Curr Opin Genet Dev *8*, 140-146.

Maeshima, K., Hihara, S., and Eltsov, M. (2010). Chromatin structure: does the 30-nm fibre exist in vivo? Curr Opin Cell Biol *22*, 291-297.

Malik, H.S., and Henikoff, S. (2003). Phylogenomics of the nucleosome. Nat Struct Biol *10*, 882-891.

Manning, G., Whyte, D.B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002). The protein kinase complement of the human genome. Science *298*, 1912-1934.

Markaki, Y., Christogianni, A., Politou, A.S., and Georgatos, S.D. (2009). Phosphorylation of histone H3 at Thr3 is part of a combinatorial pattern that marks and configures mitotic chromatin. J Cell Sci *122*, 2809-2819.

Marmorstein, R., and Roth, S.Y. (2001). Histone acetyltransferases: function, structure, and catalysis. Curr Opin Genet Dev *11*, 155-161.

Marmorstein, R., and Trievel, R.C. (2009). Histone modifying enzymes: structures, mechanisms, and specificities. Biochim Biophys Acta *1789*, 58-68.

Marzluff, W.F., and Duronio, R.J. (2002). Histone mRNA expression: multiple levels of cell cycle regulation and important developmental consequences. Curr Opin Cell Biol *14*, 692-699.

Marzluff, W.F., Gongidi, P., Woods, K.R., Jin, J., and Maltais, L.J. (2002). The human and mouse replication-dependent histone genes. Genomics *80*, 487-498.

Marzluff, W.F., Sakallah, S., and Kelkar, H. (2006). The sea urchin histone gene complement. Dev Biol *300*, 308-320.

McKittrick, E., Gafken, P.R., Ahmad, K., and Henikoff, S. (2004). Histone H3.3 is enriched in covalent modifications associated with active chromatin. Proc Natl Acad Sci U S A *101*, 1525-1530.

Menoni, H., Gasparutto, D., Hamiche, A., Cadet, J., Dimitrov, S., Bouvet, P., and Angelov, D. (2007). ATP-dependent chromatin remodeling is required for base excision repair in conventional but not in variant H2A.Bbd nucleosomes. Mol Cell Biol *27*, 5949-5956.

Misteli, T., Gunjan, A., Hock, R., Bustin, M., and Brown, D.T. (2000). Dynamic binding of histone H1 to chromatin in living cells. Nature *408*, 877-881.

Mito, Y., Henikoff, J.G., and Henikoff, S. (2005). Genome-scale profiling of histone H3.3 replacement patterns. Nat Genet *37*, 1090-1097.

Mizuguchi, G., Xiao, H., Wisniewski, J., Smith, M.M., and Wu, C. (2007). Nonhistone Scm3 and histones CenH3-H4 assemble the core of centromere-specific nucleosomes. Cell *129*, 1153-1164.

Mosammaparast, N., and Shi, Y. (2010). Reversal of histone methylation: biochemical and molecular mechanisms of histone demethylases. Annu Rev Biochem *79*, 155-179.

Murnion, M.E., Adams, R.R., Callister, D.M., Allis, C.D., Earnshaw, W.C., and Swedlow, J.R. (2001). Chromatin-associated protein phosphatase 1 regulates aurora-B and histone H3 phosphorylation. J Biol Chem *276*, 26656-26665.

Nagel, S., and Grossbach, U. (2000). Histone H1 genes and histone gene clusters in the genus Drosophila. J Mol Evol *51*, 286-298.

Nakamura, Y., Suzuki, K., and Suzuki, S. (1997). Production of Oikopleura dioica (Appendicularia) following a picoplankton “bloom” in a eutrophic coastal area. J Plankton Res *19*, 113–124.

Nesterova, T.B., Mermoud, J.E., Hilton, K., Pehrson, J., Surani, M.A., McLaren, A., and Brockdorff, N. (2002). Xist expression and macroH2A1.2 localisation in mouse primordial and pluripotent embryonic germ cells. Differentiation *69*, 216-225.

Ng, R.K., and Gurdon, J.B. (2008). Epigenetic memory of an active gene state depends on histone H3.3 incorporation into chromatin in the absence of transcription. Nat Cell Biol *10*, 102-109.

Okada, T., Endo, M., Singh, M.B., and Bhalla, P.L. (2005). Analysis of the histone H3 gene family in Arabidopsis and identification of the male-gamete-specific variant AtMGH3. Plant J *44*, 557-568.

Okuwaki, M., Kato, K., Shimahara, H., Tate, S., and Nagata, K. (2005). Assembly and disassembly of nucleosome core particles containing histone variants by human nucleosome assembly protein I. Mol Cell Biol *25*, 10639-10651.

Orsi, G.A., Couble, P., and Loppin, B. (2009). Epigenetic and replacement roles of histone variant H3.3 in reproduction and development. Int J Dev Biol *53*, 231-243.

Palmer, D.K., O'Day, K., and Margolis, R.L. (1990). The centromere specific histone CENP-A is selectively retained in discrete foci in mammalian sperm nuclei. Chromosoma *100*, 32-36.

Park, J.H., Cosgrove, M.S., Youngman, E., Wolberger, C., and Boeke, J.D. (2002). A core nucleosome surface crucial for transcriptional silencing. Nat Genet *32*, 273-279.

Pehrson, J.R., and Fried, V.A. (1992). MacroH2A, a core histone containing a large nonhistone region. Science *257*, 1398-1400.

Perez-Cadahia, B., Drobic, B., and Davie, J.R. (2009). H3 phosphorylation: dual role in mitosis and interphase. Biochem Cell Biol *87*, 695-709.

Peters, A.H., Kubicek, S., Mechtler, K., O'Sullivan, R.J., Derijck, A.A., Perez-Burgos, L., Kohlmaier, A., Opravil, S., Tachibana, M., Shinkai, Y*., et al.* (2003). Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. Mol Cell *12*, 1577-1589.

Peters, A.H., and Schubeler, D. (2005). Methylation of histones: playing memory with DNA. Curr Opin Cell Biol *17*, 230-238.

Peterson, C.L., and Laniel, M.A. (2004). Histones and histone modifications. Curr Biol *14*, R546-551.

Pinto, D.M., and Flaus, A. (2010). Structure and function of histone H2AX. Subcell Biochem *50*, 55-78.

Pray, L. (2008). Eukaryotic genome complexity. Nature Education *1*.

Preuss, U., Bierbaum, H., Buchenau, P., and Scheidtmann, K.H. (2003). DAP-like kinase, a member of the death-associated protein kinase family, associates with centrosomes, centromers, and the contractile ring during mitosis. Eur J Cell Biol *82*, 447-459.

Prigent, C., and Dimitrov, S. (2003). Phosphorylation of serine 10 in histone H3, what for? J Cell Sci *116*, 3677-3685.

Pruss, D., and Wolffe, A.P. (1993). Histone-DNA contacts in a nucleosome core containing a Xenopus 5S rRNA gene. Biochemistry *32*, 6810-6814.

Qian, C., and Zhou, M.M. (2006). SET domain protein lysine methyltransferases: Structure, specificity and catalysis. Cell Mol Life Sci *63*, 2755-2763.

Rando, O.J. (2007). Global patterns of histone modifications. Curr Opin Genet Dev *17*, 94-99.

Ren, Q., and Gorovsky, M.A. (2001). Histone H2A.Z acetylation modulates an essential charge patch. Mol Cell *7*, 1329-1335.

Ren, Q., and Gorovsky, M.A. (2003). The nonessential H2A N-terminal tail can function as an essential charge patch on the H2A.Z variant N-terminal tail. Mol Cell Biol *23*, 2778-2789.

Ridgway, P., Brown, K.D., Rangasamy, D., Svensson, U., and Tremethick, D.J. (2004). Unique residues on the H2A.Z containing nucleosome surface are important for Xenopus laevis development. J Biol Chem *279*, 43815-43820.

Roberts, S.B., Sanicola, M., Emmons, S.W., and Childs, G. (1987). Molecular characterization of the histonegene family of Caenorhabditis elegans. J Mol Biol *196*, 27-38.

Robison, B.H., Reisenbichler, K.R., and Sherlock, R.E. (2005). Giant larvacean houses: rapid carbon transport to the deep sea floor. Science 308, 1609–1611.

Roth, S.Y., Denu, J.M., and Allis, C.D. (2001). Histone acetyltransferases. Annu Rev Biochem 70, 81-120.

Russo, V.E.A., Martienssen, Robert A., Riggs, Arthur D (1996). Epigenetic mechanisms of gene regulation (New York, US, Cold Spring Harbour Laboratory).

Ruthenburg, A.J., Li, H., Patel, D.J., and Allis, C.D. (2007). Multivalent engagement of chromatin modifications by linked binding modules. Nat Rev Mol Cell Biol 8, 983-994.

Sakai, A., Schwartz, B.E., Goldstein, S., and Ahmad, K. (2009). Transcriptional and developmental functions of the H3.3 histone variant in Drosophila. Curr Biol 19, 1816-1820.

Sanders, S.L., Portoso, M., Mata, J., Bahler, J., Allshire, R.C., and Kouzarides, T. (2004). Methylation of histone H4 lysine 20 controls recruitment of Crb2 to sites of DNA damage. Cell 119, 603-614.

Sandman, K., Krzycki, J.A., Dobrinski, B., Lurz, R., and Reeve, J.N. (1990). HMf, a DNA-binding protein isolated from the hyperthermophilic archaeon Methanothermus fervidus, is most closely related to histones. Proc Natl Acad Sci U S A 87, 5788-5791.

Santenard, A., and Torres-Padilla, M.E. (2009). Epigenetic reprogramming in mammalian reproduction: contribution from histone variants. Epigenetics 4, 80-84.

Schuh, M., Lehner, C.F., and Heidmann, S. (2007). Incorporation of Drosophila CID/CENP-A and CENP-C into centromeres during early embryonic anaphase. Curr Biol 17, 237-243.

Schwartz, B.E., and Ahmad, K. (2005). Transcriptional activation triggers deposition and removal of the histone variant H3.3. Genes Dev 19, 804-814.

Seo, H.C., Kube, M., Edvardsen, R.B., Jensen, M.F., Beck, A., Spriet, E., Gorsky, G., Thompson, E.M., Lehrach, H., Reinhardt, R.*, et al.* (2001). Miniature genome in the marine chordate Oikopleura dioica. Science *294*, 2506.

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

Shogren-Knaak, M., Ishii, H., Sun, J.M., Pazin, M.J., Davie, J.R., and Peterson, C.L. (2006). Histone H4-K16 acetylation controls chromatin structure and protein interactions. Science *311*, 844-847.

Shrivastav, M., De Haro, L.P., and Nickoloff, J.A. (2008). Regulation of DNA double-strand break repair pathway choice. Cell Res *18*, 134-147.

Siegel, T.N., Hekstra, D.R., Kemp, L.E., Figueiredo, L.M., Lowell, J.E., Fenyo, D., Wang, X., Dewell, S., and Cross, G.A. (2009). Four histone variants mark the boundaries of polycistronic transcription units in Trypanosoma brucei. Genes Dev *23*, 1063-1076.

Singer, M.S., Kahana, A., Wolf, A.J., Meisinger, L.L., Peterson, S.E., Goggin, C., Mahowald, M., and Gottschling, D.E. (1998). Identification of high-copy disruptors of telomeric silencing in Saccharomyces cerevisiae. Genetics *150*, 613-632.

Sogo, J.M., Stahl, H., Koller, T., and Knippers, R. (1986). Structure of replicating simian virus 40 minichromosomes. The replication fork, core histone segregation and terminal structures. J Mol Biol *189*, 189-204.

Spada, F., Chioda, M., and Thompson, E.M. (2005a). Histone H4 post-translational modifications in chordate mitotic and endoreduplicative cell cycles. J Cell Biochem *95*, 885-901.

Spada, F., Steen, H., Troedsson, C., Kallesoe, T., Spriet, E., Mann, M., and Thompson, E.M. (2001). Molecular patterning of the oikoplastic epithelium of the larvacean tunicate Oikopleura dioica. J Biol Chem *276*, 20624-20632.

Spada, F., Vincent, M., and Thompson, E.M. (2005b). Plasticity of histone modifications across the invertebrate to vertebrate transition: histone H3 lysine 4 trimethylation in heterochromatin. Chromosome Res *13*, 57-72.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41-45.

Sugiyama, K., Sugiura, K., Hara, T., Sugimoto, K., Shima, H., Honda, K., Furukawa, K., Yamashita, S., and Urano, T. (2002). Aurora-B associated protein phosphatases as negative regulators of kinase activation. Oncogene *21*, 3103-3111.

Sullivan, B.A., and Karpen, G.H. (2004). Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. Nat Struct Mol Biol *11*, 1076-1083.

Syed, S.H., Boulard, M., Shukla, M.S., Gautier, T., Travers, A., Bednar, J., Faivre-Moskalenko, C., Dimitrov, S., and Angelov, D. (2009). The incorporation of the novel histone variant H2AL2 confers unusual structural and functional properties of the nucleosome. Nucleic Acids Res *37*, 4684-4695.

Syed, S.H., Goutte-Gattat, D., Becker, N., Meyer, S., Shukla, M.S., Hayes, J.J., Everaers, R., Angelov, D., Bednar, J., and Dimitrov, S. (2010). Single-base resolution mapping of H1-nucleosome interactions and 3D organization of the nucleosome. Proc Natl Acad Sci U S A *107*, 9620-9625.

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.

Tagami, H., Ray-Gallet, D., Almouzni, G., and Nakatani, Y. (2004). Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. Cell *116*, 51-61.

Talbert, P.B., and Henikoff, S. (2010). Histone variants--ancient wrap artists of the epigenome. Nat Rev Mol Cell Biol *11*, 264-275.

Taunton, J., Hassig, C.A., and Schreiber, S.L. (1996). A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. Science *272*, 408-411.

Thakar, A., Gupta, P., Ishibashi, T., Finn, R., Silva-Moreno, B., Uchiyama, S., Fukui, K., Tomschik, M., Ausio, J., and Zlatanova, J. (2009). H2A.Z and H3.3 histone variants affect nucleosome structure: biochemical and biophysical studies. Biochemistry *48*, 10852-10857.

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

Thambirajah, A.A., Li, A., Ishibashi, T., and Ausio, J. (2009). New developments in post-translational modifications and functions of histone H2A variants. Biochem Cell Biol *87*, 7-17.

Thoma, F., Koller, T., and Klug, A. (1979). Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. J Cell Biol *83*, 403-427.

Thompson, E.M., Kallesoe, T., and Spada, F. (2001). Diverse genes expressed in distinct regions of the trunk epithelium define a monolayer cellular template for construction of the oikopleurid house. Dev Biol *238*, 260-273.

Torras-Llort, M., Moreno-Moreno, O., and Azorin, F. (2009). Focus on the centre: the role of chromatin on the regulation of centromere identity and function. EMBO J *28*, 2337-2348.

Trelle, M.B., Salcedo-Amaya, A.M., Cohen, A.M., Stunnenberg, H.G., and Jensen, O.N. (2009). Global histone analysis by mass spectrometry reveals a high content of acetylated lysine residues in the malaria parasite Plasmodium falciparum. J Proteome Res *8*, 3439-3450.

Troedsson, C., Bouquet, J.M., and Aksnes, D.L. (2002). Resource allocation between somatic growth and reproductive output in the pelagic chordate Oikopleura dioica allows opportunistic response to nutritional variation. Mar Ecol Prog Ser *243*, 83–91.

Trostle-Weige, P.K., Meistrich, M.L., Brock, W.A., and Nishioka, K. (1984). Isolation and characterization of TH3, a germ cell-specific variant of histone 3 in rat testis. J Biol Chem *259*, 8769-8776.

Traas, J., Hulskamp, M., Gendreau, E., and Hofte, H. (1998). Endoreduplication and development: rule without dividing? Curr Opin Plant Biol *1*, 498-503.

Tschiersch, B., Hofmann, A., Krauss, V., Dorn, R., Korge, G., and Reuter, G. (1994). The protein encoded by the Drosophila position-effect variegation suppressor gene Su(var)3-9 combines domains of antagonistic regulators of homeotic gene complexes. EMBO J *13*, 3822-3831.

Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M.E., Borchers, C.H., Tempst, P., and Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. Nature *439*, 811-816.

Turck, F., Roudier, F., Farrona, S., Martin-Magniette, M.L., Guillaume, E., Buisine, N., Gagnot, S., Martienssen, R.A., Coupland, G., and Colot, V. (2007). Arabidopsis

TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27. PLoS Genet *3*, e86.

Turner, B.M. (2001). Chromatin and gene regulation, 1 edn (Oxford, Blackwell Science).

Turner, B.M. (2007). Defining an epigenetic code. Nat Cell Biol *9*, 2-6.

Ueda, K., and Tanaka, I. (1995). The appearance of male gamete-specific histones gH2B and gH3 during pollen development in Lilium longiflorum. Dev Biol *169*, 210-217.

Usachenko, S.I., Bavykin, S.G., Gavin, I.M., and Bradbury, E.M. (1994). Rearrangement of the histone H2A C-terminal domain in the nucleosome. Proc Natl Acad Sci U S A *91*, 6845-6849.

Van Doninck, K., Mandigo, M.L., Hur, J.H., Wang, P., Guglielmini, J., Milinkovitch, M.C., Lane, W.S., and Meselson, M. (2009). Phylogenomics of unusual histone H2A Variants in Bdelloid rotifers. PLoS Genet *5*, e1000401.

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

van Daal, A., White, E.M., Gorovsky, M.A., and Elgin, S.C. (1988). Drosophila has a single copy of the gene encoding a highly conserved histone H2A variant of the H2A.F/Z type. Nucleic Acids Res *16*, 7487-7497.

Wang, Y., Wysocka, J., Sayegh, J., Lee, Y.H., Perlin, J.R., Leonelli, L., Sonbuchner, L.S., McDonald, C.H., Cook, R.G., Dou, Y*., et al.* (2004). Human PAD4 regulates histone arginine methylation levels via demethylimination. Science *306*, 279-283.

Wei, Y., Yu, L., Bowen, J., Gorovsky, M.A., and Allis, C.D. (1999). Phosphorylation of histone H3 is required for proper chromosome condensation and segregation. Cell *97*, 99-109.

Wirbelauer, C., Bell, O., and Schubeler, D. (2005). Variant histone H3.3 is deposited at sites of nucleosomal displacement throughout transcribed genes while active histone modifications show a promoter-proximal bias. Genes Dev *19*, 1761-1766.

Witt, O., Albig, W., and Doenecke, D. (1996). Testis-specific expression of a novel human H3 histone gene. Exp Cell Res *229*, 301-306.

Wolfe, S.A., and Grimes, S.R. (1991). Protein-DNA interactions within the rat histone H4t promoter. J Biol Chem *266*, 6637-6643.

Xu, M., Long, C., Chen, X., Huang, C., Chen, S., and Zhu, B. (2010). Partitioning of histone H3-H4 tetramers during DNA replication-dependent chromatin assembly. Science *328*, 94-98.

Yamane, K., Toumazou, C., Tsukada, Y., Erdjument-Bromage, H., Tempst, P., Wong, J., and Zhang, Y. (2006). JHDM2A, a JmjC-containing H3K9 demethylase, facilitates transcription activation by androgen receptor. Cell *125*, 483-495.

Yang, H., Pesavento, J.J., Starnes, T.W., Cryderman, D.E., Wallrath, L.L., Kelleher, N.L., and Mizzen, C.A. (2008). Preferential dimethylation of histone H4 lysine 20 by Suv4-20. J Biol Chem *283*, 12085-12092.

Yoda, K., Ando, S., Morishita, S., Houmura, K., Hashimoto, K., Takeyasu, K., and Okazaki, T. (2000). Human centromere protein A (CENP-A) can replace histone H3 in nucleosome reconstitution in vitro. Proc Natl Acad Sci U S A *97*, 7266-7271.

Yuan, J., and Chen, J. (2010). MRE11-RAD50-NBS1 complex dictates DNA repair independent of H2AX. J Biol Chem *285*, 1097-1104.

Zalensky, A.O., Siino, J.S., Gineitis, A.A., Zalenskaya, I.A., Tomilin, N.V., Yau, P., and Bradbury, E.M. (2002). Human testis/sperm-specific histone H2B (hTSH2B). Molecular cloning and characterization. J Biol Chem *277*, 43474-43480.

Zlatanova, J., and Thakar, A. (2008). H2A.Z: view from the top. Structure *16*, 166-179.

8. Paper I-III and Supplementary Paper I

I

Paper I