

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/169128>

Please be advised that this information was generated on 2017-12-07 and may be subject to change.

CHROMATIN REGULATION IN *XENOPUS* EMBRYOS



Ila van Kruijsbergen

Chromatin regulation in *Xenopus* embryos

Ila van Kruijsbergen

The research described in this thesis was performed at the department of Molecular Developmental Biology at the Radboud Institute for Molecular Life Sciences (RIMLS), Radboud University Nijmegen, The Netherlands.

© 2017 by Ila van Kruijsbergen, Maarssen, The Netherlands

Chromatin regulation in *Xenopus* embryos

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op maandag 24 april 2017
om 16.30 uur precies

door

Ila van Kruijsbergen

geboren op 11 april 1989

te Gilze en Rijen

Promotor

prof. dr. G.J.C. Veenstra

Manuscriptcommissie

prof. dr. ir. H.G. Stunnenberg

prof. dr. J.H. Jansen

prof. dr. M.S. van Lohuizen (UvA)

CONTENTS

Chapter 1	page 7
General introduction	
Chapter 2	page 29
Recruiting polycomb to chromatin	
Chapter 3	page 51
Embryonic transcription is controlled by maternally defined chromatin state	
Chapter 4	page 85
Heterochromatic histone modifications at transposons in <i>Xenopus tropicalis</i> embryos	
Chapter 5	page 123
Dynamic Ezh2 and Ep300 recruitment to enhancers during zygotic genome activation, pluripotency and germ layer commitment in <i>Xenopus tropicalis</i>	
Chapter 6	page 159
Summary and discussion	
Chapter 7	page 169
Epilogue	
Nederlandse samenvatting	page 170
Curriculum vitae	page 175
List of publications	page 176
Dankwoord	page 178

CHAPTER 1

Introduction

*During embryogenesis one single cell develops into a complex, multicellular organism. Although many cells in the body are highly specialized, they all contain nearly identical genetic information. The cellular diversification is a result of different gene regulatory networks that arise in different cell types. Strict coordination of transcription regulation is essential to obtain and maintain this complexity during embryogenesis. Transcription regulation is coordinated by the interplay between the transcription machinery and chromatin. Transcription can be activated or repressed by tweaking the accessibility of genes and non-coding elements. The research described in this thesis is focused on how chromatin regulation, at the level of histone modifications, changes during *Xenopus tropicalis* development.*

1. Embryogenesis in *Xenopus*

*The stages of embryonic development in *Xenopus**

Even though embryogenesis is not identical in different vertebrates, most vertebrates go through similar developmental processes: cleavage, gastrulation and organogenesis including neurulation (Fig. 1) (Gilbert 2013). The embryonic stage of *Xenopus* starts with fertilization and ends with the feeding larva. The polarization of the *Xenopus* embryo is already visible in the one-cell stage, since the pigmentation is darker ventrally than dorsally. The maternal load determines the animal-vegetal axis, and the sperm entry point determines the dorsal-ventral axis (Fig. 1A) (De Robertis et al. 2000).

Fertilization is followed by cleavage. During rapid synchronous mitotic cleavage cycles the relatively big one-cell stage embryo is divided into smaller cells, called blastomeres. The blastomeres form a sphere and after the cleavage stage the *Xenopus* embryo acquires an internal cavity during the blastula stage (Gilbert 2013). At this stage the cells of the embryo are still pluripotent. Pluripotent cells can differentiate to all three germ layers: ectoderm, endoderm and mesoderm. Due to cell polarization ectoderm and endoderm are already maternally specified in *Xenopus*. The animal region will give rise to the ectoderm and the vegetal region will give rise to endoderm (Fig. 1B). However, during the blastula stage, cells in the animal pole of the *Xenopus* embryo can still differentiate to all three germ layers when exposed to specific growth factors (Paranjpe and Veenstra 2015; Borchers and Pieler 2010).

The three germ layers are formed at the gastrula stage. In *Xenopus* the mesoderm will be induced at the equatorial region between the animal and the

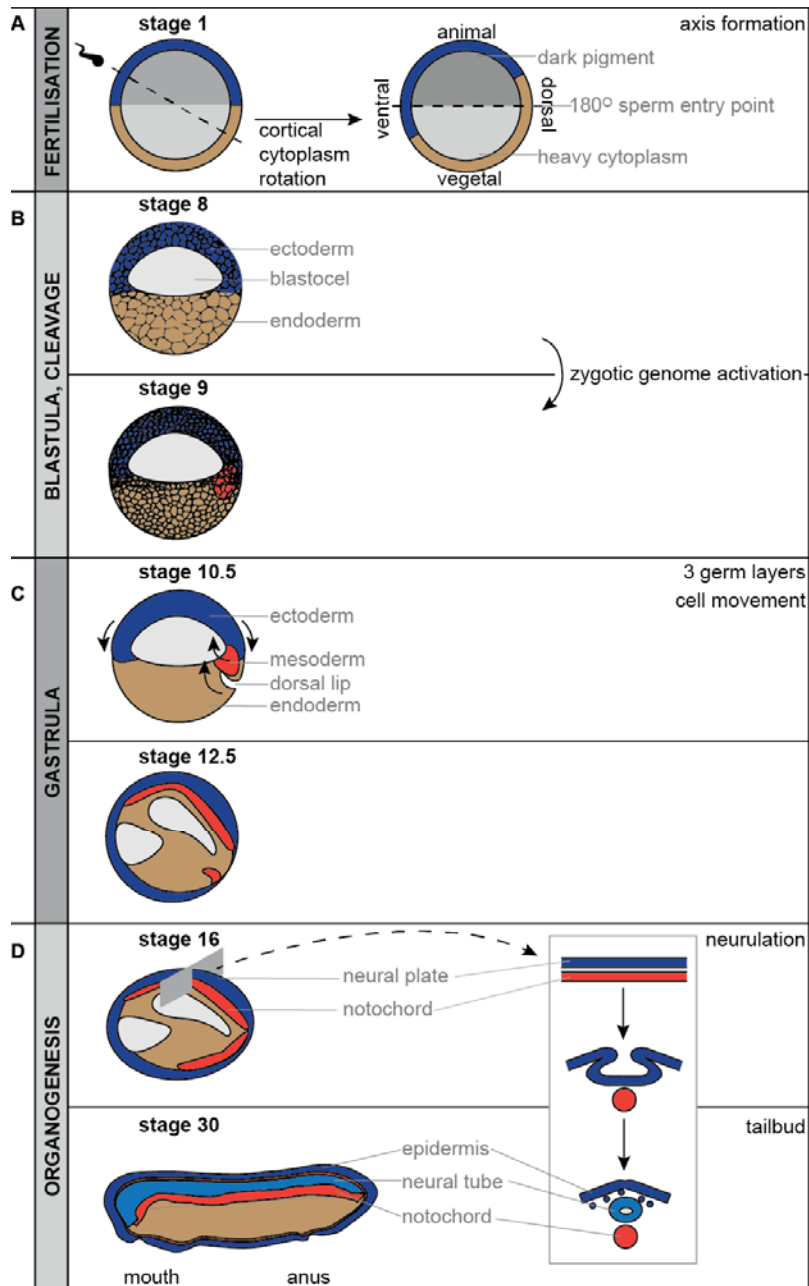


Fig. 1: *Xenopus* embryogenesis. Drawings of cross-sections of *Xenopus* embryos at A) fertilization, B) late cleavage/blastula stage, C) gastrula stage and D) organogenesis stage. A) The animal-vegetal and dorsal-ventral axis are determined by the cytoplasm gradient and sperm entry point respectively. B) Zygotic transcription starts at the end of the cleavage stage. C) The three germ layers are formed at the gastrula stage. Mesoderm moves from the posterior to the anterior side and ectoderm covers the outside of the embryo. D) Neurulation occurs at the start of organogenesis: guided by the notochord the neural tube is formed from the ectoderm. The embryonic stage, but not organogenesis, ends at the tailbud stage.

vegetal pole (Fig. 1C). The cellular diversification is a result of different gene regulatory networks that arise in the three germ layers (Kiecker et al. 2016). During the process of gastrulation the cell division rate slows down and the cells start to move and rearrange. A blastopore is formed by invagination and involution of endodermal and mesodermal cells, while ectodermal cells spread over the outer layer of the embryo. Mesodermal tissue increases and moves from the posterior to the anterior side of the embryo (Winklbauer and Schürfeld 1999).

Gastrulation is followed by neural tube formation, the formation of somites and development of brain, eyes and organs like heart, liver and the pronephros (Gilbert 2013). During the neurula stage the neural tube develops from ectodermal cells that form the neural plate (Fig. 1D). The neural tube will give rise to the spinal cord and the brain. Neural tube formation is directed by the underlying notochord which is derived from the mesoderm. The mesoderm also forms somites at the sides of the neural tube. These segmented structures cause head to tail patterning and give rise to vertebrae and associated muscles. During neurulation neural crest cells differentiate from the epidermis. These cells will migrate away from the neural tube to develop into amongst others neurons and pigment cells (Colas and Schoenwolf 2001).

Neurulation is completed during the tailbud stage and tail formation begins (Fig. 1D). The process of organogenesis, however, is not finished yet at this last embryonic stage. During organogenesis tissues and organs are formed by cell interactions and rearrangements. Many organs are formed from cells that are derived from more than one germ layer (Gilbert 2013). Though, a rough classification can be made: the mesoderm will give rise to the notochord, axial skeleton, cartilage, connective tissue, trunk muscles, kidneys and blood; the endoderm to endocrine glands and the gastrointestinal, respiratory and urinary systems; and the ectoderm to the nervous system, epidermis and pigment (Kiecker et al. 2016). The tailbud stage includes the phylotypic stage. At the phylotypic stage the vertebrate body plan and organs are formed. Different vertebrate embryos are more similar at this stage compared to the stages before and after, both morphologically and on the RNA level (Irie and Kuratani 2011).

The research described in this thesis was performed in *Xenopus tropicalis* embryos at the stages 8, 9, 10.5, 12.5, 16 and 30. These stages represent the early blastula, late blastula, early gastrula, late gastrula, neurula and tailbud stages (Fig. 1).

Embryonic transcription regulation

The different cell types that arise during development are characterized by different biochemical and cellular properties. These properties are determined by the pool of proteins and RNA that is present in the cell. Initially, embryos are largely transcriptionally silent. At the beginning of embryonic development the embryo relies on the proteins and RNA provided by the mother (Lee et al. 2014). The time at which embryonic transcription starts varies between animals. Embryonic transcription already starts after the first cleavage cycle in mice, while it is initiated after twelve cleavage cycles in *Xenopus*. The start of embryonic transcription, zygotic genome activation (ZGA), occurs at mid blastula transition (MBT) and coincides with cell cycle lengthening and cell motility in *Xenopus* (Fig. 1B) (Paranjpe and Veenstra 2015).

Two classical models describe potential causes of the ZGA/MBT in *Xenopus*: ‘the increasing nucleocytoplasmic ratio model’ and ‘the maternal clock model’ (Lee et al. 2014). The first model is based on the increasing number of nuclei relative to cytoplasmic volume. The nucleo-cytoplasmic ratio of the cells in the early embryo increases each cell division. This could result in the dilution of maternally provided repressors of transcription (Newport et al. 1982a; Newport et al. 1982b). It can also lead to the dilution of factors that stimulate replication, such as Cut5, RecQ4, Treslin and Drf1 (Collart et al. 2013). Reduced DNA replication rates could lead to cell cycle lengthening, which can provide a more permissive environment for the transcription regulatory proteins.

The second model is focused on the accumulation of factors that stimulate ZGA, rather than on the dilution of factors that inhibit ZGA. The maternal clock model states that a threshold of maternally provided factors has to be reached to induce ZGA (Howe et al. 1995). Maternal factors could be activated gradually over time, for example via polyadenylation of maternal mRNA. Zygotic transcription in *Xenopus* embryos is preceded by a wave of polyadenylation of maternal transcripts (Collart et al. 2014). Moreover, studies in embryos treated with cordycepin (which blocks polyadenylation) showed that polyadenylation of maternal mRNA is necessary for the activation of many zygotically activated genes (Collart et al. 2014). The two classical models are not necessarily mutually incompatible. The sum of increasing activators and decreasing inhibitors of transcription could regulate ZGA.

Causes of ZGA might also be found in mechanisms that regulate transcription in general: the interplay between the transcriptional machinery and the chromatin (Lee et al. 2014). Transcription factors (TFs) can direct the interaction

between the transcription machinery and the chromatin. TFs are proteins that bind at cis-regulatory elements, such as enhancers and promoters. They recruit the transcriptional machinery to chromatin or they repel it from chromatin. Thereby they control spatiotemporal expression patterns of gene transcription. TFs can stimulate multiple different cell fates depending on their concentrations and binding partners (Heinz et al. 2015).

TFs guide the interactions between the chromatin and the transcription machinery in at least two ways. Firstly, TFs can influence the timing of transcriptional activity by mediating an open, accessible chromatin environment (Zaret and Mango 2016). The chromatin has to be in an open conformation before the transcription machinery or specific chromatin modifying proteins can bind to the DNA. Some TFs, pioneer TFs, can bind to their target sequences in a condensed chromatin context. TFs cannot open chromatin themselves, but enable chromatin remodeling via binding of remodeling proteins (which will be described in the next section).

Secondly, TFs guide cell lineage-specific interactions between the transcription machinery and the chromatin. Each TF recognizes specific DNA sequences (motifs) (Heinz et al. 2015). Transcription complexes and chromatin regulating complexes can bind to TFs, so that TFs can enable their recruitment to specific genomic targets. OCT4 (POU5F1), SOX2 and NANOG in mouse or Oct91/25/60 (Pou5f3), Sox2, Ventx1/2 in *Xenopus* are pluripotency TFs (Zhang and Cui 2014). These pluripotency TFs bind to motifs specific for enhancers and promoters of pluripotency genes and mediate their activation during early developmental stages. It has been shown that the pluripotency TFs are also necessary to induce the ZGA during zebrafish development (Lee et al. 2013; Leichsenring et al. 2013). After the ZGA a new pool of TFs becomes available which can induce transcription of a new selection of genes. After ZGA pluripotency TFs can also be involved in the stimulation of differentiation. Sox2 for example stimulates differentiation towards neuronal cell fate (Zhang and Cui 2014).

TFs influence transcription regulation by initiating chromatin accessibility and by guiding protein recruitment through motif-specific DNA binding. Hereby, TFs influence the interactions between the transcriptional machinery and the chromatin. In the next section I will explain in more detail how the chromatin properties can be regulated and how that relates to (the onset of) transcription.

2. Chromatin regulation

Epigenetic regulation

Upon cell division, many biochemical and cellular properties are passed on to the daughter cells. These characteristics are not transmitted by changes in the DNA sequence, but the memory of the cell is provided by inherited components and by mechanisms that act on the chromatin: epigenetic mechanisms. An epigenetic trait is defined as “stably heritable phenotype resulting from changes in chromosome without alterations in the DNA sequence” (Berger et al. 2009). Epigenetics thus forms a link between genotype and phenotype.

Nucleosomes are the basic units of chromatin. A nucleosome consists of DNA and histones: a 145-147 bp long stretch of DNA is wrapped around an octamer of two copies of each core histone protein (H2A, H2B, H3 and H4) (Richmond et al. 2016; Luger et al. 1997). Chromatin fibers undergo condensation which facilitates organized packing of DNA in the nucleus (Kornberg 1977). Besides enabling the organized packing of DNA, nucleosomes also have a function in transcription regulation. The transcription is influenced by mechanisms that alter the chromatin characteristics at multiple levels. Chromatin characteristics can be altered at the level of I) global chromatin organization, II) local chromatin organization and III) nucleosomal composition (Fig. 2). Next I will explain various aspects of the chromatin regulation at all three levels.

Regulation of the chromatin

I) The global localization of chromatin within the nucleus is related to transcription regulation. Active chromatin compartments are relatively open, while repressed compartments are more condensed (Lieberman-Aiden et al. 2009). Condensed chromatin domains locate at the periphery of the nucleus and open chromatin more towards the center (Cremer and Cremer 2010). The open chromatin compartments are further partitioned into topologically associated domains (TADs). TADs are self-interacting sub-megabase regions of which the boundaries are associated with the insulator protein CTCF (Dixon et al. 2012; Nora et al. 2012). TADs in which transcription is regulated by the same TFs or repressed by the same epigenetic mechanisms cluster together in the nuclear space (Schoenfelder et al. 2010; Denholtz et al. 2013; de Wit et al. 2013). However, DNA-DNA interactions are more often formed within than outside the domains (Dixon et al. 2012; Nora et al. 2012). The intra-domain interactions are formed between cis-regulatory elements such as enhancers and promoter via

looping mechanisms (Kuznetsova and Stunnenberg 2016).

II) Local chromosomal organization also relates to transcription regulation. Nucleosome remodeling complexes act on nucleosomes to regulate the DNA accessibility. There are four families of chromatin remodeling complexes: switching defective/sucrose nonfermenting or Brg/Brahma-associated factor (SWI/SNF or BAF), imitation switch (ISWI), chromodomain helicase and DNA binding (CHD) and inositol requiring 80 (INO80) (Witkowski and Foulkes 2015). They alter the DNA accessibility by moving, destabilizing, ejecting or restructuring nucleosomes. These processes are ATP-dependent (Hargreaves and Crabtree 2011). Changes in density of nucleosomes or in their position relative to cis-regulatory elements influences transcription.

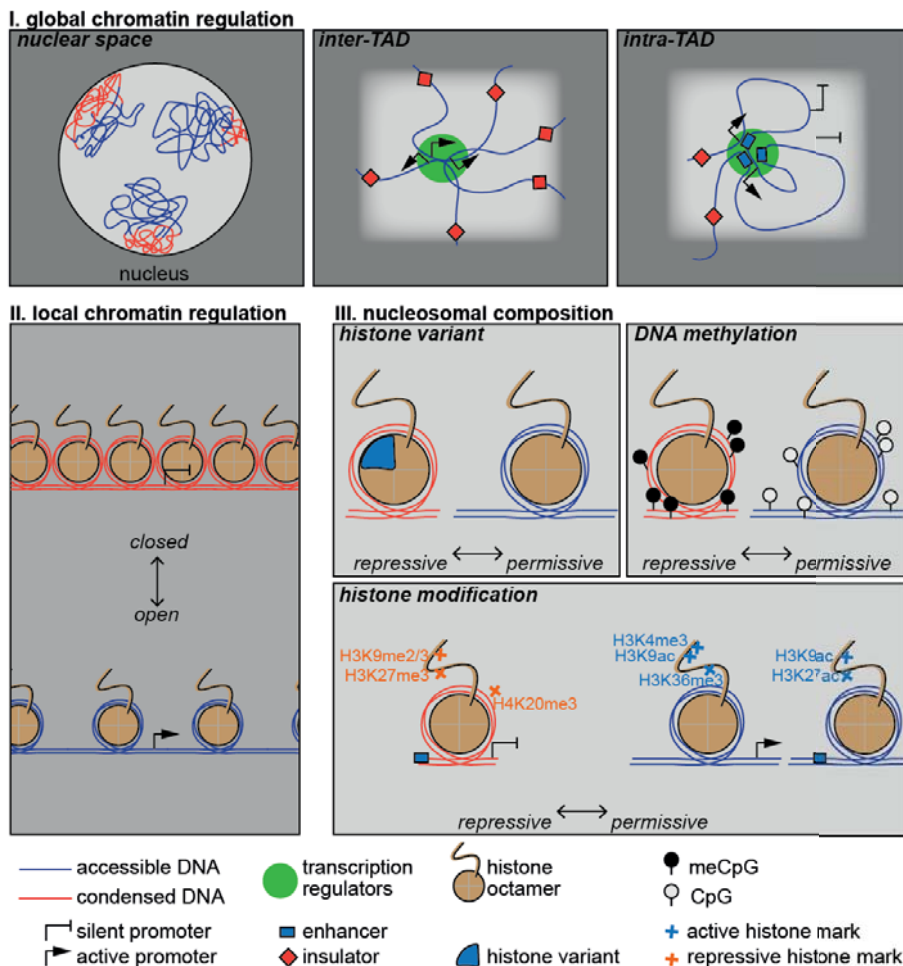


Fig. 2: Chromatin regulation. Chromatin characteristics are regulated at the level of I) global chromatin organization, II) local chromatin organization and III) nucleosomal composition.

III) Lastly, mechanisms that act on the composition or chemical modifications of individual nucleosomes influence transcription. The research described in this thesis is mainly focused on this level of chromatin regulation. A nucleosome can undergo various alterations that can all alter the transcriptional state of the genomic locus where the nucleosome is incorporated. Both components of a nucleosome, the DNA and the histone octamer, are susceptible to transformations.

The nucleotides of the DNA can be modified. The most common DNA modifications occur at cytosines upstream of a guanine (CpG dinucleotides). DNA methyltransferases methylate these cytosines on the fifth position of the cytosine pyrimidine ring (meCpG) (Bird and Southern 1978; Bogdanović and Veenstra 2009). The meCpG modification can be passively removed via dilution through DNA replication and it can be actively removed by hydroxylation. Subsequently, the hydroxyl group can be converted in a formylcytosine, which can be converted in a carboxylcytosine (Tahiliani et al. 2009; Ito et al. 2010; Ito et al. 2011). meCpG can repress promoters via the recruitment of meCpG binding protein complexes (Li et al. 1993; Bogdanović and Veenstra 2009). At enhancers and in gene bodies, however, meCpG is not per se associated with transcriptional repression (Hellman and Chess 2007; Lister et al. 2009; Schmidl et al. 2009; Stadler et al. 2011). Most CpG dinucleotides in vertebrate genomes are methylated (Bird 1986). meCpG-depleted regions have a relatively high CpG density and are mostly promoters. Genes with meCpG-depleted promoters can be active or repressed, depending on the histone modifications that are deposited at these sites (Lee et al. 2006; Bernstein et al. 2006; Mendenhall et al. 2010).

Histones are post-translationally modified by histone modifying enzymes. The recruitment of histone modifying enzymes to chromatin can be initiated or maintained by TFs, pre-existing histone modifications, or a pre-existing transcriptional state. These different mechanism involved in recruiting histone modifiers will be discussed in more detail in chapter 2. Histones proteins have unstructured N-terminal tails that project away from the core of the complex. The core regions and the tails of the histone complex are both susceptible to modifications, but so far the modifications of the tails have been studied more extensively (Lawrence et al. 2016). Histones can undergo various modifications at multiple amino acids. Lysines can for example be acetylated, methylated, sumoylated or ubiquitylated; while arginines can be methylated; and threonines and serines can be phosphorylated (Kouzarides 2007).

Histone modifications can influence the chromatin compaction, since

they can change the charge of histone tails which changes their interaction with the negatively charged DNA (Shogren-Knaak et al. 2006). Furthermore, histone modifications regulate transcription by the recruitment of effector proteins (Vermeulen et al. 2010; Bartke et al. 2010). Via effector proteins histone modifications can stimulate or repress transcription by for example: stabilization of the transcription machinery, compaction of chromatin, and recruitment of chromatin remodeling complexes (Vermeulen et al. 2007; Boros et al. 2014; Chatterjee et al. 2011). Acetylation of histones is generally related to transcription activation. Methylation of histones is related to both activation and repression, depending on the amino acid (Kouzarides 2007). The transcriptional status related to the histone modifications that will come back later in this thesis are listed in Table 1.

The histone complex can not only be altered by post-translational modifications, but also via incorporation of variant histones. The incorporation of these non-canonical histone variants can affect the recruitment of chromatin binding-proteins and the chromatin structure (Talbert and Henikoff 2010). This can result in positive or negative regulation of transcription. MacroH2A1 is for example linked to transcriptional repression, while H3.3 is linked to transcriptional activity (Ahmad et al. 2002; Costanzi and Pehrson 1998).

Chromatin regulation is complex since global chromatin organization, local chromatin organization and nucleosomal composition respond to changes within and between these different levels. Just as transcription regulation, the chromatin structure is highly dynamic during embryogenesis. This will be discussed in the next sections.

Table 1: associations of the histone modifications studied in this thesis

Modification	Related transcriptional state	reference
H3K4me3	transcriptional activation, permissive promoters	(Bernstein et al. 2002; Santos-Rosa et al. 2002; Brinkman et al. 2006; Guenther et al. 2007)
H3K36me3	transcription elongation	(Vakoc et al. 2006; Kizer et al. 2005)
H3K9ac	transcriptional activation	(Spencer et al. 1997)
H3K4me1	permissive enhancers	(Heintzman et al. 2007; Bonn et al. 2012)
H3K27ac	active enhancers	(Bonn et al. 2012; Creyghton et al. 2010; Rada-Iglesias et al. 2011)
H3K27me3	transcriptional repression, enhancers/promoters poised for activation	(Müller et al. 2002; Cao et al. 2002; Czermin et al. 2002; Rada-Iglesias et al. 2011; Voigt et al. 2013)
H3K9me2	transcriptional repression	(Snowden et al. 2002)
H3K9me3	transcriptional repression, centromeres, telomeres, repetitive elements	(Peters et al. 2003; Martens et al. 2005; Schotta et al. 2004; Mikkelsen et al. 2007)
H4K20me3	transcriptional repression, centromeres, telomeres, repetitive elements	(Schotta et al. 2004; Martens et al. 2005; Mikkelsen et al. 2007; Benetti et al. 2007)

Global chromatin organization and development

Global chromatin organization in the nucleus is highly dynamic during early development. The 3-dimensional chromatin organization changes during early mouse development (Ahmed et al. 2010). Chromatin is extensively dispersed in the nucleus of the mouse 1-cell embryo, while they concentrate into compact structures around the nuclear envelope at the 2-cell stage. At the 8-cell stage uncompact chromatin fibers disperse through the nucleus again. They remain dispersed through the nucleus until lineage commitment (Ahmed et al. 2010). Before ZGA pericentromeric and centromeric sequences also cluster together around nucleolar polar bodies, but from 4-cell stage they start to spread throughout the nucleus (Probst et al. 2007; Aguirre-Lavin et al. 2012). Also in *Xenopus* the chromatin interactions with the nuclear lamina are different before and after ZGA (Hair et al. 1998). The chromosome length decreases from blastula stage to tailbud stage too (Micheli 1993). The compaction of different chromosomes in the nuclear space depends on their gene density. Chromosomes with a relatively low gene density are closer to peripheral locations compared to chromosomes with higher gene density in cattle, but only after ZGA (Koehler et al. 2009).

Furthermore, during development DNA-DNA interactions are dynamic as well. Genome wide chromatin interactions within and between TADs transform during lineage specification of human embryonic stem (ES) cells (Dixon et al. 2015).

Local chromatin organization and development

The chromatin is relatively condensed in early mouse zygotes. In late zygotes and at the 2-cell stage, however, the DNA becomes globally permissive. The DNA is more repressed again at ICM when it is only locally permissive (Wu et al. 2016). The accessibility of DNA also decreases after ZGA in *Xenopus* embryos (Hair et al. 1998).

Not only nucleosome density, but also nucleosome positioning changes during early development. Zebrafish studies have shown that the position of the histone complexes with respect to the DNA sequence is dynamic as well (Zhang et al. 2014; Haberle et al. 2014). Whereas nucleosomes are randomly positioned before ZGA, after ZGA they form a well-ordered array. At ZGA nucleosomes precisely position at transcription start sites independently of active transcription or RNA polymerase II binding (Zhang et al. 2014). Furthermore, together with

a switch in nucleosome positioning zebrafish embryos switch their promoter usage at ZGA. During this switch transcription start sites change from A/T-rich to C/G-rich locations (Haberle et al. 2014).

Studies in mouse ES cells have shown that various remodeling complexes are required for proper development. The embryonic specific SWI/SNF complex is for example required for the maintenance of pluripotency in ES cells (Ho et al. 2009). Its subunit BRG1 facilitates the repression of differentiation genes, while it also prevents repression of pluripotency genes (Ho et al. 2011). The remodeling complexes CHD11 and INO80 are necessary to maintain an open chromatin conformation at pluripotency genes in ES cells (Gaspar-Maia et al. 2009; Wang et al. 2014). Remodeler NuRD on the other hand stimulates the exit from pluripotency by suppressing pluripotency genes (Reynolds et al. 2012).

Nucleosome composition and development

Nucleosome characteristics change during development. Specific histone variants are required for proper development. In *Xenopus*, for example, incorporation of H3.3 is necessary for mediating the proper response to mesoderm inducing cues (Lim et al. 2013). Besides via histone variant incorporation, nucleosomes undergo changes during embryogenesis via CpG methylation and histone modifications (Paranjpe and Veenstra 2015).

The dynamics of CpG methylation are different between mammalian and non-mammalian vertebrates. The DNA is CpG hypermethylated in mouse sperm, while it is hypomethylated in mouse oocytes. Upon fertilization DNA is demethylated; active demethylation is more pronounced for paternally derived DNA than for maternally derived DNA (Santos et al. 2002; Mayer et al. 2000; Oswald et al. 2000; Shen et al. 2014). This results in hypomethylated DNA at ZGA. Similar as for mouse, zebrafish sperm DNA is CpG hypermethylated and oocyte DNA is CpG hypomethylated (Potok et al. 2013; Jiang et al. 2013). However, measurements starting in 4-8 cell embryos indicate that general demethylation does not occur in zebrafish. Upon fertilization meCpG levels rather directly increase towards sperm levels (Potok et al. 2013; Jiang et al. 2013). Studies in *Xenopus* (starting from stage 6) also suggest the absence of global demethylation (Veenstra and Wolffe 2001; Bogdanović et al. 2011). So, whereas DNA is CpG hypomethylated during ZGA in mouse it is not in zebrafish and *Xenopus*. DNA methyltransferase knockout studies have shown that re-methylation in the mouse blastocyst is necessary for proper development (Okano et al. 1999; Li et al. 1992). Studies in mouse ES cells indicated that the re-methylation is essential for differentiation

(Tsumura et al. 2006).

meCpG at promoters and gene bodies is associated with gene repression in somatic cells (Bogdanović and Veenstra 2009). During development active CpG demethylation at enhancers also results in a more open chromatin context (Bogdanović et al. 2016). Via this mechanism enhancers of developmental genes can be activated during the phylotypic stage in mouse, zebrafish and *Xenopus*. Remarkably, meCpG is temporally uncoupled from transcriptional repression during the blastula and gastrula stages in *Xenopus* embryos (Bogdanovic et al. 2011). Methylated promoters do drive expression at these early developmental stages. meCpG is also uncoupled from transcription repression during mammalian gametogenesis (Hammoud et al. 2014).

Possibly meCpG has an alternative function during these early stages, such as guiding histone modifying enzymes. Trimethylation on lysine 4 or 27 of histone H3 (H3K4me3 and H3K27me3) occurs at meCpG depleted promoters (Lee et al. 2006; Bernstein et al. 2006; Mendenhall et al. 2010). H3K4me3 marks transcriptionally permissive promoters, while H3K27me3 marks transcriptionally silenced genes (Table 1). In mouse and zebrafish the two histone modifications have been reported to co-occur at genes that are primed for activation (Voigt et al. 2013; Vastenhouw et al. 2010). This may seem different from *Xenopus*, where the co-occurrence of these marks is not quantitatively dominant (Akkers et al. 2009; van Heeringen et al. 2014). However, zebrafish and *Xenopus* embryos are both depleted of H3K27me3 deposition during the early developmental stages (Akkers et al. 2009; Vastenhouw et al. 2010; Lindeman et al. 2011; van Heeringen et al. 2014). In zebrafish and *Xenopus* H3K27me3 marking starts to accumulate after ZGA, while H3K4me3 starts to accumulate already before ZGA (Akkers et al. 2009; Vastenhouw et al. 2010). A temporal hierarchy of H3K4me3 and H3K27me3 deposition is also seen in pre-implantation mouse embryos (X. Liu et al. 2016). The dynamic H3K4me3 deposition observed in pre-ZGA zebrafish, *Xenopus*, and mouse embryos might have an instructive function (Blythe et al. 2010; Lindeman et al. 2011; Dahl et al. 2016).

Histone modifications at enhancers are also remodeled during development. Studies in mouse ES cells and zebrafish embryos showed that enhancers primed for activation carry mono-methylation of histone H3 on lysine 4 (H3K4me1) (Creyghton et al. 2010; Bogdanovic et al. 2012). Concomitant with transcription of their target genes, enhancers are additionally marked by acetylation of histone H3 at lysine 27 (H3K27ac) (Creyghton et al. 2010; Bogdanovic et al. 2012). During zebrafish embryogenesis the shift in H3K27ac marking is accompanied by a shifts in expression of pluripotency genes to tissue-specific genes (Bogdanovic

et al. 2012). Enhancers that are inactive, but poised for activation in human ES cells are marked by H3K4me1 and H3K27me3. The H3K27me3 marking at these enhancers is replaced by H3K27ac when they are activated upon differentiation (Rada-Iglesias et al. 2011). H3K4me1-marked enhancers recruit transcription stimulating as well as repressing complexes in blastula stage *Xenopus* embryos (van Heeringen et al. 2014). At this pluripotent stage RNA polymerase II and subunits of Polycomb Repressive Complex 2 bind at the same set of enhancers. This indicates the presence of a balance of opposing activities that act on enhancers during development.

While genes and enhancers are differentially regulated in different cell types, subtelomeric and pericentric chromosome locations are condensed in all cell types. While the facultative condensed loci are marked by H3K27me3, constitutive heterochromatin are marked by tri- and/or dimethylation of histones H3 and H4 at respectively lysines 9 and 20 (H3K9me2/3, H4K20me3). These histone modifications can recruit proteins that cause condensation (Saksouk et al. 2015). During mouse embryogenesis the zygote loses H4K20me3 at the 2-cell stage (Kourmouli et al. 2004). H3K9me3 is only passively diluted out, and is re-deposited already after the first couple of cleavage cycles (Puschendorf et al. 2008; Liu et al. 2004). Also in *Xenopus* embryos H3K9me3 is more stable than H4K20me3; total H3K9me2/3 levels fluctuate less than 2-fold between the blastula and the tadpole stage, while H4K20me3 increases 5-fold (Schneider et al. 2011).

In summary, global chromatin organization, local chromatin organization and nucleosomal composition are all highly dynamic during embryogenesis. The research described in this thesis will be focused on histone modification dynamics, but it should be realized that all different layers of chromatin regulation are interdependent.

3. Overview of this thesis

The dynamic gene activity during embryogenesis is tightly regulated by the interplay between the transcriptional machinery, TFs and chromatin. Knowledge about each individual component is essential to understand how these components work together to regulate transcription.

Histone modification profiles can be generated by chromatin immunoprecipitation followed by high throughput sequencing (ChIP-seq). This

method relies on antibodies that are used to identify DNA sequences part of nucleosomes carrying a specific histone modification. After DNA-protein fixation and DNA fractionation, DNA is captured with histone modification-specific antibodies. Thousands of cells are required to generate high quality genome-wide binding profiles using a standard CHIP-seq protocol. High cell numbers are easily obtained using *Xenopus tropicalis* embryos, since they develop in eggs outside the mother and many batches consist of over a thousand embryos. *X. tropicalis* are relatively close to humans in gene content and synteny (Hellsten et al. 2010). Different from *X. laevis* and zebrafish, *X. tropicalis* did not undergo genome duplication. Therefore, *X. tropicalis* is a very convenient model system to study histone modifications during vertebrate embryonic development.

The aim of the research presented in this thesis was to characterize the dynamics of diverse histone modifications during embryonic development in *X. tropicalis* embryos. Before we explore the epigenetic signatures during development, first a review about the various factors that influence recruitment of one particular histone modifying complex will be provided. Chapter 2 is focused on the recruitment of Polycomb Repressive Complex 2 (PRC2). This complex is responsible for the deposition of H3K27me₃, a repressive mark which is highly dynamic during development.

We studied histone modification marking during *X. tropicalis* development from the blastula to tailbud stage (Chapter 3). We show that modifications that mark a transcriptionally active state (H3K4me₃, H3K36me₃, H3K9ac) and that mark transcriptionally repressed states (H3K27me₃, H3K9me_{2/3}, H4K20me₃) are dynamic and accumulating after ZGA. Furthermore, we address if the deposition of histone modifications is determined maternally or by newly expressed embryonic factors. We show that H3K4me₃ and H3K27me₃ deposition is mainly independent of embryonic transcription, while it is required for the recruitment of histone H3K27 acetyltransferase Ep300.

We also studied repressive histone modifications at transposable elements (Chapter 4). As will be explained, these parasitic elements can be beneficial for evolution, but are potentially harmful for individuals. Therefore, we examined to which extent repressing histone modifications could be involved in the regulation of transposons during embryogenesis. We also addressed which mechanisms could be involved in the recruitment of the histone modifying enzymes that catalyze these modifications.

Furthermore, we analyzed the characteristics of nucleosomes bound by methyltransferase Ezh2 (Chapter 5). Ezh2 binds at accessible enhancers

and H3K27me3 deposition does not mimic Ezh2 binding during *Xenopus* development. Our research highlights the complex relationships between transcription regulation and the recruitment of activating and repressing proteins to the genome.

Finally, I will discuss the work presented in this thesis (Chapter 6).

References

- Aguirre-Lavin, Tiphaine, Pierre Adenot, Amélie Bonnet-Garnier, Gaétan Lehmann, Renaud Fleurot, Claire Boulesteix, Pascale Debey, and Nathalie Beaujean. 2012. "3D-FISH Analysis of Embryonic Nuclei in Mouse Highlights Several Abrupt Changes of Nuclear Organization during Preimplantation Development." *BMC Developmental Biology* 12. BioMed Central: 30.
- Ahmad, Kami, Steven Henikoff, K. Ahmad, S. Henikoff, A.S. Akhmanova, P.C.T. Bindels, J. Xu, et al. 2002. "The Histone Variant H3.3 Marks Active Chromatin by Replication-Independent Nucleosome Assembly." *Molecular Cell* 9 (6). Elsevier: 1191–1200.
- Ahmed, Kashif, Hesam Dehghani, Peter Rugg-Gunn, Eden Fussner, Janet Rossant, and David P Bazett-Jones. 2010. "Global Chromatin Architecture Reflects Pluripotency and Lineage Commitment in the Early Mouse Embryo." *PLoS One* 5 (5): e10531.
- Akkers, Robert C, Simon J van Heeringen, Ulrike G Jacobi, Eva M Janssen-Megens, Kees-Jan François, Hendrik G Stunnenberg, and Gert Jan C Veenstra. 2009. "A Hierarchy of H3K4me3 and H3K27me3 Acquisition in Spatial Gene Regulation in *Xenopus* Embryos." *Developmental Cell* 17 (3): 425–34.
- Bartke, Till, Michiel Vermeulen, Blerta Xhemalce, Samuel C Robson, Matthias Mann, and Tony Kouzarides. 2010. "Nucleosome-Interacting Proteins Regulated by DNA and Histone Methylation." *Cell* 143 (3): 470–84.
- Benetti, Roberta, Susana Gonzalo, Isabel Jaco, Gunnar Schotta, Peter Klatt, Thomas Jenuwein, and María A Blasco. 2007. "Suv4-20h Deficiency Results in Telomere Elongation and Derepression of Telomere Recombination." *The Journal of Cell Biology* 178 (6). The Rockefeller University Press: 925–36.
- Berger, Shelley L, Tony Kouzarides, Ramin Shiekhattar, and Ali Shilatifard. 2009. "An Operational Definition of Epigenetics." *Genes & Development* 23 (7): 781–83.
- Bernstein, Bradley E, Emily L Humphrey, Rachel L Erlich, Robert Schneider, Peter Bouman, Jun S Liu, Tony Kouzarides, and Stuart L Schreiber. 2002. "Methylation of Histone H3 Lys 4 in Coding Regions of Active Genes." *Proceedings of the National Academy of Sciences of the United States of America* 99 (13): 8695–8700.
- Bernstein, Bradley E, Tarjei S Mikkelsen, Xiaohui Xie, Michael Kamal, Dana J Huebert, James Cuff, Ben Fry, et al. 2006. "A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells." *Cell* 125 (2): 315–26.
- Bird, A P. 1986. "CpG-Rich Islands and the Function of DNA Methylation." *Nature* 321 (6067): 209–13.
- Bird, Adrian P. 1978. "Use of Restriction Enzymes to Study Eukaryotic DNA Methylation: II. The Symmetry of Methylated Sites Supports Semi-Conservative Copying of the Methylation Pattern." *Journal of Molecular Biology* 118 (1). Academic Press: 49–60.
- Bird, Adrian P., and Edwin M. Southern. 1978. "Use of Restriction Enzymes to Study Eukaryotic DNA Methylation: I. The Methylation Pattern in Ribosomal DNA from *Xenopus laevis*." *Journal of Molecular Biology* 118 (1). Academic Press: 27–47.
- Blythe, Shelby A, Sang-wook Cha, Emmanuel Tadjuidje, Janet Heasman, and Peter S Klein. 2010. "Beta-Catenin Primes Organizer Gene Expression by Recruiting a Histone H3 Arginine Methyltransferase, Prmt2." *Developmental Cell* 19 (2): 220–31.
- Bogdanovic, Ozren, Ana Fernandez-Minan, Juan J Tena, Elisa de Lacalle-Mustienes, Carmen Hidalgo, Ila van Kruijsbergen, Simon J van Heeringen, Gert Jan C Veenstra, and Jose Luis Gomez-Skarmeta. 2012. "Dynamics of Enhancer Chromatin Signatures Mark the Transition from Pluripotency to Cell Specification during Embryogenesis." *Genome Research*, May.
- Bogdanovic, Ozren, Steven W Long, Simon J van Heeringen, Arie B Brinkman, Jose Luis Gómez-Skarmeta, Hendrik G Stunnenberg, Peter L Jones, and Gert Jan C Veenstra. 2011. "Temporal Uncoupling of the DNA Methylome and Transcriptional Repression during Embryogenesis." *Genome Research* 21 (8): 1313–27.
- Bogdanović, Ozren, Arne H Smits, Elisa de la Calle Mustienes, Juan J Tena, Ethan Ford, Ruth Williams, Upeka Senanayake, et al. 2016. "Active DNA Demethylation at Enhancers during the Vertebrate Phylogenetic Period." *Nature Genetics*, February.
- Bogdanović, Ozren, and Gert Jan C Veenstra. 2009. "DNA Methylation and Methyl-CpG Binding Proteins: Developmental Requirements and Function." *Chromosoma* 118: 549–65.
- Bonn, Stefan, Robert P Zinzen, Charles Girardot, E Hilary Gustafson, Alexis Perez-Gonzalez, Nicolas Delhomme, Yad Ghavi-Helm, Bartek Wilczyński, Andrew Riddell, and Eileen E M Furlong. 2012. "Tissue-Specific Analysis of Chromatin State Identifies Temporal Signatures of Enhancer Activity during Embryonic Development." *Nature Genetics* 44 (2): 148–56. doi:10.1038/ng.1064.
- Borchers, Annette, and Tomas Pieler. 2010. "Programming Pluripotent Precursor Cells Derived from *Xenopus* Embryos to Generate Specific Tissues and Organs." *Genes* 1 (3): 413–26.
- Boros, Joanna, Nausica Arnoult, Vincent Stroobant, Jean-François Collet, and Anabelle Decottignies. 2014. "Polycomb Repressive Complex 2 and H3K27me3 Cooperate with H3K9 Methylation to Maintain Heterochromatin Protein 1α at Chromatin." *Molecular and Cellular Biology* 34 (19): 3662–74.
- Brinkman, Arie B, Thijs Roelofsen, Sebastiaan W C Pennings, Joost H A Martens, Thomas Jenuwein, and

- Hendrik G Stunnenberg. 2006. "Histone Modification Patterns Associated with the Human X Chromosome." *EMBO Reports* 7 (6): 628–34. doi:10.1038/sj.embor.7400686.
- Cao, Ru, Liangju Wang, Hengbin Wang, Li Xia, Hediye Erdjument-Bromage, Paul Tempst, Richard S. Jones, and Yi Zhang. 2002. "Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing." *Science* 298.
- Chatterjee, Nilanjana, Divya Sinha, Mekonnen Lemma-Dechassa, Song Tan, Michael A Shogren-Knaak, and Blaine Bartholomew. 2011. "Histone H3 Tail Acetylation Modulates ATP-Dependent Remodeling through Multiple Mechanisms." *Nucleic Acids Research* 39 (19): 8378–91.
- Colas, J F, and G C Schoenwolf. 2001. "Towards a Cellular and Molecular Understanding of Neurulation." *Developmental Dynamics* 221 (2): 117–45.
- Collart, Clara, George E Allen, Charles R Bradshaw, James C Smith, and Philip Zegerman. 2013. "Titration of Four Replication Factors Is Essential for the *Xenopus Laevis* Midblastula Transition." *Science* 341 (6148): 893–96.
- Collart, Clara, Nick D L Owens, Leena Bhaw-Rosun, Brook Cooper, Elena De Domenico, Ilya Patrushev, Abdul K Sesay, James N Smith, James C Smith, and Michael J Gilchrist. 2014. "High-Resolution Analysis of Gene Activity during the *Xenopus* Mid-Blastula Transition." *Development (Cambridge, England)* 141 (9): 1927–39.
- Costanzi, C, and J R Pehrson. 1998. "Histone macroH2A1 Is Concentrated in the Inactive X Chromosome of Female Mammals." *Nature* 393 (6685): 599–601.
- Cremer, Thomas, and Marion Cremer. 2010. "Chromosome Territories." *Cold Spring Harbor Perspectives in Biology* 2 (3): a003889.
- Creyghton, Menno P, Albert W Cheng, G Grant Welstead, Tristan Kooistra, Bryce W Carey, eveline J Steine, Jacob Hanna, et al. 2010. "Histone H3K27ac Separates Active from Poised Enhancers and Predicts Developmental State." *PNAS* 107 (50): 21931–36.
- Czermin, Birgit, Raffaella Melfi, Donna McCabe, Volker Seitz, Axel Imhof, and Vincenzo Pirrotta. 2002. "Drosophila Enhancer of Zeste/ESC Complexes Have a Histone H3 Methyltransferase Activity That Marks Chromosomal Polycomb Sites." *Cell* 111 (2): 185–96.
- Dahl, John Arne, Inkyung Jung, Håvard Aanes, Gareth D Greggains, Adeel Manaf, Mads Lerdrup, Guoqiang Li, et al. 2016. "Broad Histone H3K4me3 Domains in Mouse Oocytes Modulate Maternal-to-Zygotic Transition." *Nature* 537 (7621): 548–52.
- De Robertis, E M, J Larraín, M Oelgeschläger, and O Wessely. 2000. "The Establishment of Spemann's Organizer and Patterning of the Vertebrate Embryo." *Nature Reviews. Genetics* 1 (3). NIH Public Access: 171–81.
- de Wit, Elzo, Britta A. M. Bouwman, Yun Zhu, Petra Klous, Erik Splinter, Marjon J. A. M. Verstegen, Peter H. L. Krijger, et al. 2013. "The Pluripotent Genome in Three Dimensions Is Shaped around Pluripotency Factors." *Nature* 501 (7466). *Nature Research*: 227–31.
- Denholtz, Matthew, Giancarlo Bonora, Constantinos Chronis, Erik Splinter, Wouter de Laat, Jason Ernst, Matteo Pellegrini, and Kathrin Plath. 2013. "Long-Range Chromatin Contacts in Embryonic Stem Cells Reveal a Role for Pluripotency Factors and Polycomb Proteins in Genome Organization." *Cell Stem Cell* 13 (5): 602–16.
- Dixon, Jesse R., Siddarth Selvaraj, Feng Yue, Audrey Kim, Yan Li, Yin Shen, Ming Hu, Jun S. Liu, and Bing Ren. 2012. "Topological Domains in Mammalian Genomes Identified by Analysis of Chromatin Interactions." *Nature* 485 (7398). *Nature Publishing Group*: 376–80.
- Dixon, Jesse R, Inkyung Jung, Siddarth Selvaraj, Yin Shen, Jessica E Antosiewicz-Bourget, Ah Young Lee, Zhen Ye, et al. 2015. "Chromatin Architecture Reorganization during Stem Cell Differentiation." *Nature* 518 (7539): 331–36.
- Gaspar-Maia, Alexandre, Adi Alajem, Fanny Polesso, Rupa Sridharan, Mike J Mason, Amy Heidersbach, João Ramalho-Santos, et al. 2009. "Chd1 Regulates Open Chromatin and Pluripotency of Embryonic Stem Cells." *Nature* 460 (7257): 863–68.
- Gilbert, Scott F. 2013. *Developmental Biology*. Tenth. Sunderland: Sinauer Associates. <http://www.sinauer.com/developmental-biology-663.html>.
- Guenther, Matthew G, Stuart S Levine, Laurie A Boyer, Rudolf Jaenisch, and Richard A Young. 2007. "A Chromatin Landmark and Transcription Initiation at Most Promoters in Human Cells." *Cell* 130 (1): 77–88.
- Haberle, Vanja, Nan Li, Yavor Hadzhiev, Charles Plessy, Christopher Previti, Chirag Nepal, Jochen Gehrig, et al. 2014. "Two Independent Transcription Initiation Codes Overlap on Vertebrate Core Promoters." *Nature* 507 (7492). *Europe PMC Funders*: 381–85.
- Hair, A, M N Prioleau, Y Vassetzky, and M Méchali. 1998. "Control of Gene Expression in *Xenopus* Early Development." *Developmental Genetics* 22 (2): 122–31.
- Hammoud, Saher Sue, Diana H.P. Low, Chongil Yi, Douglas T. Carrell, Ernesto Guccione, Bradley R. Cairns, D. Ballow, et al. 2014. "Chromatin and Transcription Signatures of Mammalian Adult Germline Stem Cells and Spermatogenesis." *Cell Stem Cell* 15 (2). Elsevier: 239–53.
- Hargreaves, Diana C, and Gerald R Crabtree. 2011. "ATP-Dependent Chromatin Remodeling: Genetics, Genomics and Mechanisms." *Cell Research* 21 (3). *Nature Publishing Group*: 396–420.
- Heintzman, Nathaniel D, Rhona K Stuart, Gary Hon, Yutao Fu, Christina W Ching, R David Hawkins, Leah O Barrera, et al. 2007. "Distinct and Predictive Chromatin Signatures of Transcriptional Promoters and En-

- hancers in the Human Genome." *Nature Genetics* 39 (3). Nature Publishing Group: 311–18.
- Heinz, Sven, Casey E Romanoski, Christopher Benner, and Christopher K Glass. 2015. "The Selection and Function of Cell Type-Specific Enhancers." *Nature Reviews. Molecular Cell Biology* 16 (3): 144–54.
- Hellman, Asaf, and Andrew Chess. 2007. "Gene Body-Specific Methylation on the Active X Chromosome." *Science* 315 (5815): 1141–43.
- Hellsten, Uffe, Richard M Harland, Michael J Gilchrist, David Hendrix, Jerzy Jurka, Vladimir Kapitonov, Ivan Ovcharenko, et al. 2010. "The Genome of the Western Clawed Frog *Xenopus Tropicalis*." *Science (New York, N.Y.)* 328 (5978): 633–36.
- Ho, Lena, Erik L Miller, Jehnna L Ronan, Wen Qi Ho, Raja Jothi, and Gerald R Crabtree. 2011. "esBAF Facilitates Pluripotency by Conditioning the Genome for LIF/STAT3 Signalling and by Regulating Polycomb Function." *Nature Cell Biology* 13 (8): 903–13.
- Ho, Lena, Jehnna L Ronan, Jiang Wu, Brett T Staahl, Lei Chen, Ann Kuo, Julie Lessard, Alexey I Nesvizhskii, Jeff Ranish, and Gerald R Crabtree. 2009. "An Embryonic Stem Cell Chromatin Remodeling Complex, esBAF, Is Essential for Embryonic Stem Cell Self-Renewal and Pluripotency." *Proceedings of the National Academy of Sciences of the United States of America* 106 (13): 5181–86.
- Howe, J A, M Howell, T Hunt, and J W Newport. 1995. "Identification of a Developmental Timer Regulating the Stability of Embryonic Cyclin A and a New Somatic A-Type Cyclin at Gastrulation." *Genes & Development* 9 (10): 1164–76.
- Irie, Naoki, and Shigeru Kuratani. 2011. "Comparative Transcriptome Analysis Reveals Vertebrate Phylotypic Period during Organogenesis." *Nature Communications* 2: 248.
- Ito, Shinsuke, Ana C D'Alessio, Olena V Taranova, Kwonho Hong, Lawrence C Sowers, and Yi Zhang. 2010. "Role of Tet Proteins in 5mC to 5hmC Conversion, ES-Cell Self-Renewal and Inner Cell Mass Specification." *Nature* 466 (7310): 1129–33.
- Ito, Shinsuke, Li Shen, Qing Dai, Susan C Wu, Leonard B Collins, James A Swenberg, Chuan He, and Yi Zhang. 2011. "Tet Proteins Can Convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine." *Science* 333 (6047): 1300–1303.
- Jiang, Lan, Jing Zhang, Jing-Jing Wang, Lu Wang, Li Zhang, Guoqiang Li, Xiaodan Yang, et al. 2013. "Sperm, but Not Oocyte, DNA Methylome Is Inherited by Zebrafish Early Embryos." *Cell* 153 (4). NIH Public Access: 773–84.
- Kiecker, Clemens, Thomas Bates, and Esther Bell. 2016. "Molecular Specification of Germ Layers in Vertebrate Embryos." *Cellular and Molecular Life Sciences* 73 (5): 923–47.
- Kizer, Kelby O, Hemali P Phatnani, Yoichiro Shibata, Hana Hall, Arno L Greenleaf, and Brian D Strahl. 2005. "A Novel Domain in Set2 Mediates RNA Polymerase II Interaction and Couples Histone H3 K36 Methylation with Transcript Elongation." *Molecular and Cellular Biology* 25 (8): 3305–16.
- Koehler, Daniela, Valeri Zakhartchenko, Lutz Froenicke, Gary Stone, Roscoe Stanyon, Eckhard Wolf, Thomas Cremer, and Alessandro Brero. 2009. "Changes of Higher Order Chromatin Arrangements during Major Genome Activation in Bovine Preimplantation Embryos." *Experimental Cell Research* 315 (12): 2053–63.
- Kornberg, R D. 1977. "Structure of Chromatin." *Annual Review of Biochemistry* 46: 931–54.
- Kourmouli, Niki, Peter Jeppesen, Shantha Mahadevaiah, Paul Burgoyne, Rong Wu, David M Gilbert, Silvia Bongiorno, et al. 2004. "Heterochromatin and Tri-Methylated Lysine 20 of Histone H4 in Animals." *Journal of Cell Science* 117 (Pt 12). The Company of Biologists Ltd: 2491–2501.
- Kouzarides, Tony. 2007. "Chromatin Modifications and Their Function." *Cell* 128: 693–705.
- Kuznetsova, Tatyana, and Hendrik G. Stunnenberg. 2016. "Dynamic Chromatin Organization: Role in Development and Disease." *The International Journal of Biochemistry & Cell Biology* 76: 119–22.
- Lawrence, Moyra, Sylvain Daujat, Robert Schneider, K. Bloom, A. Joglekar, G. Hadlaczyk, et al., et al. 2016. "Lateral Thinking: How Histone Modifications Regulate Gene Expression." *Trends in Genetics* 32 (1). Elsevier: 42–56.
- Lee, Miler T, Ashley R Bonneau, and Antonio J Giraldez. 2014. "Zygotic Genome Activation during the Maternal-to-Zygotic Transition." *Annual Review of Cell and Developmental Biology* 30: 581–613.
- Lee, Miler T, Ashley R Bonneau, Carter M Takacs, Ariel A Bazzini, Kate R DiVito, Elizabeth S Fleming, and Antonio J Giraldez. 2013. "Nanog, Pou5f1 and SoxB1 Activate Zygotic Gene Expression during the Maternal-to-Zygotic Transition." *Nature* 503 (7476). NIH Public Access: 360–64.
- Lee, Tong Ihn, Richard G Jenner, Laurie A Boyer, Matthew G Guenther, Stuart S Levine, Roshan M Kumar, Brett Chevalier, et al. 2006. "Control of Developmental Regulators by Polycomb in Human Embryonic Stem Cells." *Cell* 125 (2): 301–13.
- Leichsenring, Manuel, Julia Maes, Rebecca Mössner, Wolfgang Driever, and Daria Nischtchouk. 2013. "Pou5f1 Transcription Factor Controls Zygotic Gene Activation in Vertebrates." *Science (New York, N.Y.)* 341 (6149): 1005–9.
- Li, E, C Beard, and R Jaenisch. 1993. "Role for DNA Methylation in Genomic Imprinting." *Nature* 366 (6453): 362–65.

- Li, E, T H Bestor, and R Jaenisch. 1992. "Targeted Mutation of the DNA Methyltransferase Gene Results in Embryonic Lethality." *Cell* 69 (6): 915–26.
- Lieberman-Aiden, Erez, Nynke L van Berkum, Louise Williams, Maxim Imakaev, Tobias Ragoczy, Agnes Telling, Ido Amit, et al. 2009. "Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome." *Science* 326 (5950): 289–93.
- Lim, Chin Yan, Bruno Reversade, Barbara B Knowles, Davor Solter, K. Ahmad, S. Henikoff, J. Bednar, et al. 2013. "Optimal Histone H3 to Linker Histone H1 Chromatin Ratio Is Vital for Mesodermal Competence in *Xenopus*." *Development* 140 (4). Oxford University Press for The Company of Biologists Limited: 853–60.
- Lindeman, Leif C, Ingrid S Andersen, Andrew H Reiner, Nan Li, Havard Aanes, Olga Østrup, Cecilia Winata, et al. 2011. "Prepatterning of Developmental Gene Expression by Modified Histones before Zygotic Genome Activation." *Developmental Cell* 21: 993–1004.
- Lister, Ryan, Mattia Pelizzola, Robert H Downen, R David Hawkins, Gary Hon, Julian Tonti-Filippini, Joseph R Nery, et al. 2009. "Human DNA Methylomes at Base Resolution Show Widespread Epigenomic Differences." *Nature* 462 (7271): 315–22.
- Liu, Honglin, Jin-Moon Kim, and Fugaku Aoki. 2004. "Regulation of Histone H3 Lysine 9 Methylation in Oocytes and Early Pre-Implantation Embryos." *Development* 131 (10): 2269–80.
- Liu, Xiaoyu, Chenfei Wang, Wenqiang Liu, Jingyi Li, Chong Li, Xiaochen Kou, Jiayu Chen, et al. 2016. "Distinct Features of H3K4me3 and H3K27me3 Chromatin Domains in Pre-Implantation Embryos." *Nature*, September.
- Luger, Karolin, Timothy J. Richmond, Armin W. Mäder, Robin K. Richmond, and David F. Sargent. 1997. "Crystal Structure of the Nucleosome Core Particle at 2.8 Å Resolution." *Nature* 389 (6648).
- Martens, Joost H A, Roderick J O'Sullivan, Ulrich Braunschweig, Susanne Opravil, Martin Radolf, Peter Steinlein, and Thomas Jenuwein. 2005. "The Profile of Repeat-Associated Histone Lysine Methylation States in the Mouse Epigenome." *The EMBO Journal* 24 (4): 800–812.
- Mayer, Wolfgang, Alain Niveleau, Jörn Walter, Reinald Fundele, and Thomas Haaf. 2000. "Embryogenesis: Demethylation of the Zygotic Paternal Genome." *Nature* 403 (6769). Nature Publishing Group: 501–2.
- Mendenhall, Eric M, Richard P Koche, Thanh Truong, Vicky W Zhou, Biju Issac, Andrew S Chi, Manching Ku, and Bradley E Bernstein. 2010. "GC-Rich Sequence Elements Recruit PRC2 in Mammalian ES Cells." *PLoS Genetics* 6 (12): 1–10.
- Mikkelsen, Tarjei S, Manching Ku, David B Jaffe, Biju Issac, Erez Lieberman, Georgia Giannoukos, Pablo Alvarez, et al. 2007. "Genome-Wide Maps of Chromatin State in Pluripotent and Lineage-Committed Cells." *Nature* 448 (7153). NIH Public Access: 553–60.
- Müller, Jürg, Craig M. Hart, Nicole J. Francis, Marcus L. Vargas, Aditya Sengupta, Brigitte Wild, Ellen L. Miller, Michael B. O'Connor, Robert E. Kingston, and Jeffrey A. Simon. 2002. "Histone Methyltransferase Activity of a *Drosophila* Polycomb Group Repressor Complex." *Cell* 111 (2): 197–208.
- Newport, J, M Kirschner, E.D. Adamson, H.R. Woodland, E.M. de Robertis, K. Nishikura, E.M. de Robertis, et al. 1982. "A Major Developmental Transition in Early *Xenopus* Embryos: II. Control of the Onset of Transcription." *Cell* 30 (3). Elsevier: 687–96.
- Newport, J, M Kirschner, R. Bachvarova, E.H. Davidson, J.E.M. Ballantine, H.R. Woodland, E.A. Sturgess, et al. 1982. "A Major Developmental Transition in Early *Xenopus* Embryos: I. Characterization and Timing of Cellular Changes at the Midblastula Stage." *Cell* 30 (3). Elsevier: 675–86.
- Nora, Elphège P, Bryan R Lajoie, Edda G Schulz, Luca Giorgetti, Ikuhiro Okamoto, Nicolas Servant, Tristan Pilot, et al. 2012. "Spatial Partitioning of the Regulatory Landscape of the X-Inactivation Centre." *Nature* 485 (7398). NIH Public Access: 381–85.
- Okano, Masaki, Daphne W Bell, Daniel A Haber, En Li, T.H Bestor, T.H Bestor, A.P Laudano, et al. 1999. "DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development." *Cell* 99 (3). Elsevier: 247–57.
- Oswald, J, S Engemann, N Lane, W Mayer, A Olek, R Fundele, W Dean, et al. 2000. "Active Demethylation of the Paternal Genome in the Mouse Zygote." *Current Biology* 10 (8). Elsevier: 475–78.
- Paranjpe, Sarita S, and Gert Jan C Veenstra. 2015. "Establishing Pluripotency in Early Development." *Biochimica et Biophysica Acta* 1849 (6): 626–36.
- Peters, Antoine H F M, Stefan Kubicek, Karl Mechtler, Roderick J O'Sullivan, Alwin A H A Derijck, Laura Perez-Burgos, Alexander Kohlmaier, et al. 2003. "Partitioning and Plasticity of Repressive Histone Methylation States in Mammalian Chromatin." *Molecular Cell* 12 (6): 1577–89.
- Potok, Magdalena E, David A Nix, Timothy J Parnell, and Bradley R Cairns. 2013. "Reprogramming the Maternal Zebrafish Genome after Fertilization to Match the Paternal Methylation Pattern." *Cell* 153 (4). NIH Public Access: 759–72.
- Probst, Aline V., Fátima Santos, Wolf Reik, Geneviève Almouzni, and Wendy Dean. 2007. "Structural Differences in Centromeric Heterochromatin Are Spatially Reconciled on Fertilisation in the Mouse Zygote." *Chromosoma* 116 (4). Springer-Verlag: 403–15.
- Puschendorf, Mareike, Rémi Terranova, Erwin Boutsma, Xiaohong Mao, Kyo-ichi Isono, Urszula Brykczynska, Carolin Kolb, et al. 2008. "PRC1 and Suv39h Specify Parental Asymmetry at Constitutive Heterochromatin in

- Early Mouse Embryos." *Nature Genetics* 40 (4): 411–20.
- Rada-Iglesias, Alvaro, Ruchi Bajpai, Tomek Swigut, Samantha a Brugmann, Ryan a Flynn, and Joanna Wysocka. 2011. "A Unique Chromatin Signature Uncovers Early Developmental Enhancers in Humans." *Nature* 470.
- Reynolds, Nicola, Paulina Latos, Antony Hynes-Allen, Remco Loos, Donna Leaford, Aoife O'Shaughnessy, Olukunbi Mosaku, et al. 2012. "NuRD Suppresses Pluripotency Gene Expression to Promote Transcriptional Heterogeneity and Lineage Commitment." *Cell Stem Cell* 10 (5). Elsevier: 583–94.
- Richmond, T J, J T Finch, B Rushton, D Rhodes, and A Klug. 2016. "Structure of the Nucleosome Core Particle at 7 Å Resolution." *Nature* 311 (5986): 532–37.
- Saksouk, Nehmé, Elisabeth Simboeck, and Jérôme Déjardin. 2015. "Constitutive Heterochromatin Formation and Transcription in Mammals." *Epigenetics & Chromatin* 8. BioMed Central: 3.
- Santos, Fátima, Brian Hendrich, Wolf Reik, and Wendy Dean. 2002. "Dynamic Reprogramming of DNA Methylation in the Early Mouse Embryo." *Developmental Biology* 241 (1). Academic Press: 172–82.
- Santos-Rosa, Helena, Robert Schneider, Andrew J Bannister, Julia Sherriff, Bradley E Bernstein, N C Tolga Emre, Stuart L Schreiber, Jane Mellor, and Tony Kouzarides. 2002. "Active Genes Are Tri-Methylated at K4 of Histone H3." *Nature* 419 (6905): 407–11.
- Schmidl, Christian, Maja Klug, Tina J Boeld, Reinhard Andreesen, Petra Hoffmann, Matthias Edinger, and Michael Rehli. 2009. "Lineage-Specific DNA Methylation in T Cells Correlates with Histone Methylation and Enhancer Activity." *Genome Research* 19 (7): 1165–74.
- Schneider, Tobias D, Jose M Arteaga-Salas, Edith Mentele, Robert David, Dario Nicetto, Axel Imhof, and Ralph A W Rupp. 2011. "Stage-Specific Histone Modification Profiles Reveal Global Transitions in the Xenopus Embryonic Epigenome." *PLoS One* 6 (7): e22548.
- Schoenfelder, Stefan, Tom Sexton, Lyubomira Chakalova, Nathan F Cope, Alice Horton, Simon Andrews, Sreenivasulu Kurukuti, et al. 2010. "Preferential Associations between Co-Regulated Genes Reveal a Transcriptional Interactome in Erythroid Cells." *Nature Genetics* 42 (1): 53–61.
- Schotta, Gunnar, Monika Lachner, Kavitha Sarma, Anja Ebert, Roopsha Sengupta, Gunter Reuter, Danny Reinberg, and Thomas Jenuwein. 2004. "A Silencing Pathway to Induce H3-K9 and H4-K20 Trimethylation at Constitutive Heterochromatin." *Genes & Development* 18 (11): 1251–62.
- Shen, Li, Azusa Inoue, Jin He, Yuting Liu, Falong Lu, and Yi Zhang. 2014. "Tet3 and DNA Replication Mediate Demethylation of Both the Maternal and Paternal Genomes in Mouse Zygotes." *Cell Stem Cell* 15 (4). NIH Public Access: 459–70.
- Shogren-Knaak, Michael, Haruhiko Ishii, Jian-Min Sun, Michael J Pazin, James R Davie, and Craig L Peterson. 2006. "Histone H4-K16 Acetylation Controls Chromatin Structure and Protein Interactions." *Science* 311 (5762): 844–47.
- Snowden, Andrew W, Philip D Gregory, Casey C Case, and Carl O Pabo. 2002. "Gene-Specific Targeting of H3K9 Methylation Is Sufficient for Initiating Repression in Vivo." *Current Biology* 12 (24): 2159–66.
- Spencer, T E, G Jenster, M M Burcin, C D Allis, J Zhou, C A Mizzen, N J McKenna, et al. 1997. "Steroid Receptor Coactivator-1 Is a Histone Acetyltransferase." *Nature* 389 (6647): 194–98.
- Stadler, Michael B, Rabih Murr, Lukas Burger, Robert Ivanek, Florian Lienert, Anne Schöler, Erik van Nimwegen, et al. 2011. "DNA-Binding Factors Shape the Mouse Methylome at Distal Regulatory Regions." *Nature* 480 (7378): 490–95.
- Tahiliani, Mamta, Kian Peng Koh, Yinghua Shen, William A Pastor, Hozefa Bandukwala, Yevgeny Brudno, Suneet Agarwal, et al. 2009. "Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1." *Science* 324 (5929): 930–35.
- Talbert, Paul B., and Steven Henikoff. 2010. "Histone Variants — Ancient Wrap Artists of the Epigenome." *Nature Reviews Molecular Cell Biology* 11 (4). Nature Publishing Group: 264–75.
- Tsumura, Akiko, Tomohiro Hayakawa, Yuichi Kumaki, Shin-ichiro Takebayashi, Morito Sakaue, Chisa Matsuoka, Kunitada Shimotohno, et al. 2006. "Maintenance of Self-Renewal Ability of Mouse Embryonic Stem Cells in the Absence of DNA Methyltransferases Dnmt1, Dnmt3a and Dnmt3b." *Genes to Cells* 11 (7): 805–14.
- Vakoc, Christopher R, Mira M Sachdeva, Hongxin Wang, and Gerd A Blobel. 2006. "Profile of Histone Lysine Methylation across Transcribed Mammalian Chromatin." *Molecular and Cellular Biology* 26 (24): 9185–95.
- van Heeringen, Simon J, Robert C Akkers, Ila van Kruijsbergen, M Asif Arif, Lars L P Hanssen, Nilofar Sharifi, and Gert Jan C Veenstra. 2014. "Principles of Nucleation of H3K27 Methylation during Embryonic Development." *Genome Research* 24 (3): 401–10.
- Vastenhouw, Nadine L, Yong Zhang, Ian G Woods, Farhad Imam, Aviv Regev, X Shirley Liu, John Rinn, and Alexander F Schier. 2010. "Chromatin Signature of Embryonic Pluripotency Is Established during Genome Activation." *Nature* 464 (7290).
- Veenstra, Gert Jan C., and Alan P. Wolffe. 2001. "Constitutive Genomic Methylation during Embryonic Development of Xenopus." *Biochimica et Biophysica Acta* 1521 (1): 39–44.
- Vermeulen, Michiel, H Christian Eberl, Filomena Matarese, Hendrik Marks, Sergei Denissov, Falk Butter, Kenneth K Lee, et al. 2010. "Quantitative Interaction Proteomics and Genome-Wide Profiling of Epigenetic Histone Marks and Their Readers." *Cell* 142 (6). Elsevier: 967–80.

- Vermeulen, Michiel, Klaas W Mulder, Sergei Denissov, W W M Pim Pijnappel, Frederik M A van Schaik, Radhika A Varier, Marijke P A Baltissen, Henk G Stunnenberg, Matthias Mann, and Marc Timmers. 2007. "Selective Anchoring of TFIIID to Nucleosomes by Trimethylation of Histone H3 Lysine 4." *Cell* 131 (1): 58–69.
- Voigt, Philipp, Wee-Wei Tee, and Danny Reinberg. 2013. "A Double Take on Bivalent Promoters." *Genes & Development* 27 (12): 1318–38.
- Wang, Li, Ying Du, James M Ward, Takashi Shimbo, Brad Lackford, Xiaofeng Zheng, Yi-liang Miao, et al. 2014. "INO80 Facilitates Pluripotency Gene Activation in Embryonic Stem Cell Self-Renewal, Reprogramming, and Blastocyst Development." *Cell Stem Cell* 14 (5). NIH Public Access: 575–91.
- Winklbauer, R, and M Schürfeld. 1999. "Vegetal Rotation, a New Gastrulation Movement Involved in the Internalization of the Mesoderm and Endoderm in *Xenopus*." *Development* 126 (16): 3703–13.
- Witkowski, Leora, and William D Foulkes. 2015. "In Brief: Picturing the Complex World of Chromatin Remodelling Families." *The Journal of Pathology* 237 (4): 403–6. doi:10.1002/path.4585.
- Wu, Jingyi, Bo Huang, He Chen, Qiangzong Yin, Yang Liu, Yunlong Xiang, Bingjie Zhang, et al. 2016. "The Landscape of Accessible Chromatin in Mammalian Preimplantation Embryos." *Nature* 534 (7609): 652–57.
- Zaret, Kenneth S, and Susan E Mango. 2016. "Pioneer Transcription Factors, Chromatin Dynamics, and Cell Fate Control." *Current Opinion in Genetics & Development* 37: 76–81.
- Zhang, Shuchen, and Wei Cui. 2014. "Sox2, a Key Factor in the Regulation of Pluripotency and Neural Differentiation." *World Journal of Stem Cells* 6 (3): 305–11.
- Zhang, Y., N. L. Vastenhouw, J. Feng, K. Fu, C. Wang, Y. Ge, A. Pauli, P. van Hummelen, A. F. Schier, and X. S. Liu. 2014. "Canonical Nucleosome Organization at Promoters Forms during Genome Activation." *Genome Research* 24 (2). Cold Spring Harbor Laboratory Press: 260–66.

CHAPTER 2

Recruiting polycomb to chromatin

Ila van Kruijsbergen, Saartje Hontelez, Gert Jan C. Veenstra

The International Journal of Biochemistry & Cell Biology (2015) 67: 177-87

IvK wrote the review, with help from SH and GJCV.

Abstract

Polycomb group (PcG) proteins are key regulators in establishing a transcriptional repressive state. Polycomb Repressive Complex 2 (PRC2), one of the two major PcG protein complexes, is essential for proper differentiation and maintenance of cellular identity. Multiple factors are involved in recruiting PRC2 to its genomic targets. In this review we will discuss the role of DNA sequence, transcription factors, pre-existing histone modifications, and RNA in guiding PRC2 towards specific genomic loci. The DNA sequence itself influences the DNA methylation state, which is an important determinant of PRC2 recruitment. Other histone modifications are also important for PRC2 binding as PRC2 can respond to different cellular states via crosstalk between histone modifications. Additionally, PRC2 might be able to sense the transcriptional status of genes by binding to nascent RNA, which could also guide the complex to chromatin. In this review we will discuss how all these molecular aspects define a local chromatin state which controls accurate, cell-type specific epigenetic silencing by PRC2.

1. Introduction: Role of Polycomb in development

The role of polycombgroup (PcG) proteins as repressors of early developmental genes was first described in *Drosophila melanogaster*. PcG proteins were shown to control segmentation during early embryogenesis by maintaining temporal and spatial repression of *Hox* genes (Lewis 1978; Duncan 1982). In mouse, various knockout studies have demonstrated a similar role for PcG proteins in the maintenance of a repressive transcriptional state (reviewed in: Aloia et al. 2013; Signolet and Hendrich 2015). PcG proteins can form different multi-subunit protein complexes, of which Polycomb Repressive Complex 1 and 2 (PRC1 and PRC2) have been characterized most extensively (see Box 1). Both PRC complexes are histone modifiers. PRC2 catalyses mono-, di-, and trimethylation of histone H3 on lysine K27 (H3K27me1/2/3) by its subunit Ezh2, and PRC1 catalyses monoubiquitylation of histone H2A on lysine 119 (H2AK119ub1) by its subunit Ring1 (Czermin et al. 2002; Kuzmichev et al. 2002; Müller et al. 2002; de Napoles et al. 2004; Pengelly et al. 2013).

Post-translational modifications can regulate transcription, because they can function as a docking site or modulate the affinity of nuclear proteins (Musselman et al. 2012b). In this way PcG proteins can limit the accessibility of DNA for the transcription machinery by compacting chromatin (reviewed in: Di Croce & Helin, 2013; Schwartz & Pirrotta, 2013). Besides altering the

BOX1: complex compositions

PcG proteins contribute to two major protein complexes: Polycomb Repressive Complex (PRC) 1 and PRC2. PRC1 has multiple complex compositions, each with its own properties as reviewed by (reviewed in: Turner & Bracken 2013; Di Croce & Helin 2013). There are two major PRC1 complexes, each containing different core subunits: (i) Cbx, Phc, Ring, and Pcgf, or (ii) Rybp, Ring, and Pcgf. Each of these subunits have different paralogs (Turner and Bracken 2013). The catalytic subunit of PRC1 can be either Ring1a or Ring1b, which monoubiquitylate histone H2A on lysine 119 (H2AK119) (de Napoles et al. 2004), however, their activity depends on the complex composition (Turner and Bracken 2013).

The core components of PRC2 are Enhancer of zeste (Ezh2), Embryonic ectoderm development (Eed), and Suppressor of zeste 12 (Suz12). These subunits exist as monomers in the complex in a 1:1:1 stoichiometry (Smits et al. 2013; Xu et al. 2015), and comprise the minimal composition necessary for catalytic activity of Ezh2, resulting in mono-, di-, or trimethylation of H3K27 (Cao and Zhang 2004; Pasini et al. 2004; Nekrasov et al. 2005). Non-core PRC2 proteins such as RbAp48/46, PCL1/2/3, AEBP2, Jarid2, c17orf96 and C10orf12 can be substoichiometrically present in the complex (Smits et al. 2013), and can increase the catalytic activity (e.g. RbAp46/48 and AEBP2) or the binding and targeting of PRC2 (e.g. Jarid2 and PCL) (reviewed in: Vizán et al. 2015).

Ezh2 is the only PRC2 core subunit known to have a paralog, namely Ezh1. Expression of Ezh2 and Ezh1 is dissimilar, and are found in complexes with distinct composition and function. Ezh2 generally forms a core together with both Eed and Suz12, whereas Ezh1 has been found alone or in a complex together with Suz12 (Xu et al. 2015). Although both molecules show a partial redundancy in catalytic activity and localization, Ezh2 is generally believed to deploy di- and tri-methylation of H3K27 on repressed genomic loci, whereas Ezh1 is more associated with monomethylation of H3K27 on regions with active transcription (Mousavi et al. 2012; Xu et al. 2015). During cell differentiation, the ratio between Ezh1 and Ezh2 containing PRC2 changes, with Ezh2 levels decreasing and Ezh1 levels increasing upon differentiation (Margueron et al. 2008; Mousavi et al. 2012; Xu et al. 2015). To date, most studies on PRC2 focused on the Ezh2 containing variant and its function in transcriptional silencing.

accessibility of chromatin PcG proteins can as well mediate epigenetic repression by counteracting activating histone modifications (figure 1A,B). In contrast to PcG proteins some of the Trithorax Group (TrxG) proteins catalyse trimethylation of histone H3 on lysine K4 (H3K4me3) and lysine K36 (H3K36me3) at genes that are transcriptionally active. Various studies have highlighted that PcG proteins antagonize transcriptional activation by TrxG proteins (reviewed in: Steffen & Ringrose 2014). PcG proteins also counteract activating histone modifications at regulatory elements across the genome. Methylation of H3K27 prevents acetylation of this lysine (H3K27ac), a modification which is enriched at active enhancer regions (Ferrari et al. 2014).

These biochemical mechanisms via which PcG proteins mediate transcription silencing have been extensively studied. At the same time, how PRC complexes are directed to their genomic targets remains an important question. This review is focussed on the several aspects that affect the recruitment of PRC2 to its genomic targets: DNA sequence, transcription factors, pre-existing histone modifications, and RNA. First we will briefly summarize recent findings on polycomb mediated transcriptional regulation. After that we will discuss in more detail the recent findings on PRC2 recruitment.

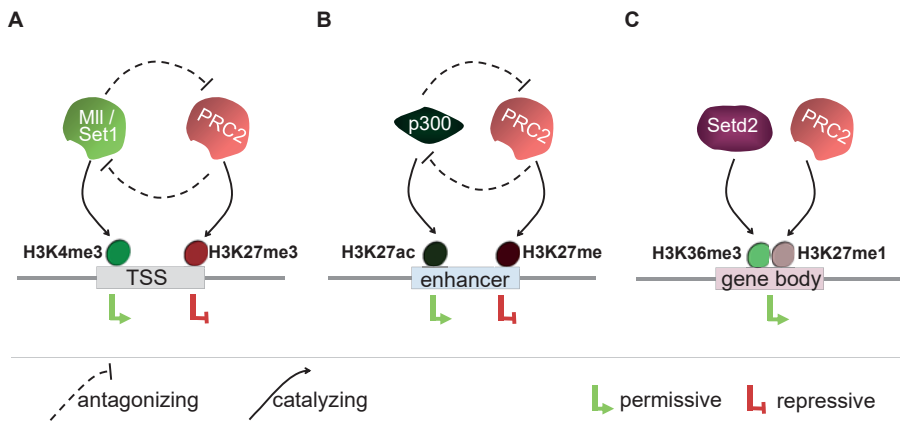


Figure 1. Roles of PRC2. The activity of PRC2 is different at functionally distinct genomic regions. A) PRC2 inhibits gene activation by trimethylation of H3K27 at transcription start sites (TSSs), which prevents Mll or Set1-mediated trimethylation of H3K4 at the TSS. B) Methylation of H3K27 by PRC2 on enhancers prevents activation by antagonizing acetylation of this substrate by p300. C) Upon transcription, monomethylation of H3K27 by PRC2 co-occurs with H3K36me3 deposition by Setd2.

2. Sequential polycomb action: a paradigm under pressure

Trimethylated H3K27 can serve as a docking site for PRC1 component PC (Cbx in mammals) (Cao et al. 2002). In the absence of enzymatically active PRC2, H3K27 cannot be trimethylated and PRC1 binding is lost (Cao et al. 2002; Wang et al. 2004; Boyer et al. 2006). These observations gave rise to the sequential or hierarchical model, which postulates that once PRC2 is recruited and trimethylates H3K27, PRC1 is recruited by virtue of the affinity of its Cbx subunit for this methylated residue. However, not all recent findings fit the classical sequential model, suggesting alternative mechanisms for the establishment of polycomb-mediated regulation of transcription.

The classical model predicts co-occurrence of PRC1 and PRC2 subunits on genomic loci, however, genome-wide profiling studies in embryonic stem cells (ESCs) showed that PRC1 and PRC2 proteins share only a subset of binding sites (Boyer et al. 2006; Ku et al. 2008; Blackledge et al. 2014). Early ChIP-on-chip assays in mouse ESCs indicated that merely 25% of all PcG enriched transcription start sites (TSS) were occupied by all four proteins that were profiled: PRC1 components Phc1 and Rnf2, and PRC2 components Eed and Suz12 (Boyer et al. 2006). More recently, ChIP-sequencing assays on Ring1b and Ezh2 binding showed that almost 90% of the Ring1b binding sites were also occupied by Ezh2, whereas only 50% of the Ezh2 binding sites bound Ring1b as well (Ku et al. 2008). A stronger, but still not perfect overlap for Ezh2 at Ring1b targets was found by

Blackledge and colleagues. In their study Ring1b and Ezh2 shared about 80% of their targets (Blackledge et al. 2014). These findings show that PRC1 and PRC2 do not always bind the same regions, contrary to what may be expected on basis of the classical model of PRC2 and PRC1 action.

Independent functions and recruitment mechanisms for PRC1 and PRC2 have been identified. Genomic and proteomic analysis of PRC1 complexes identified six major groups, containing distinct subunits and differing in genomic binding, of which only a small subset co-localized with H3K27me3 (Gao et al. 2012). Furthermore, it is demonstrated that PRC1 recruitment is not solely dependent on H3K27me3, as it can still deposit H2AK119ub and repress gene transcription in PRC2-deficient mouse ESCs (Tavares et al. 2012). Although PRC2 can still be involved in recruiting PRC1 to shared binding sites, recent studies showed that PRC1 can also be involved in the recruitment of PRC2 (Blackledge et al. 2014; Cooper et al. 2014; Kalb et al. 2014). Knockdown of PRC1 not only resulted in a loss of H2AK119ub, but also in reduced PRC2 binding (Blackledge et al. 2014). The role of H2AK119ub in PRC2 recruitment will be further discussed in sections 3. and 5.2. of this review. These findings suggest that the order of events can be bidirectional rather than unidirectional as described in the classical model.

Another caveat in the classical model is that it only focusses on the H3K27 trimethylation by PRC2, even though PRC2 also catalyses mono- and dimethylation of H3K27 (Ferrari et al. 2014). In the past, genome-wide studies in murine ESCs identified PcG proteins and H3K27me3 in the vicinity of the transcription start site (TSS, figure 1A) of genes, many of which encode transcription factors with important functions in development (Bernstein et al. 2006; Boyer et al. 2006). More recently, Ferrari and colleagues characterized the distribution of H3K27me1 and H3K27me2 in mouse ESCs, and found them to be located at functionally distinct genomic regions. H3K27me1 is mainly enriched in the bodies of actively transcribed genes (figure 1C), whereas H3K27me2 was broadly distributed throughout the genome, covering approximately 70% of all histones. Genes and enhancers covered with H3K27me2 were deprived of marks associated with genomic activation, and associated with low expression levels (Ferrari et al. 2014).

However, H3K27me2 is not highly abundant throughout *Xenopus* development. Mass spectrometry (MS) based analysis showed that H3K27me2 levels rose from 3% in blastula stage to 15% in tadpoles (Schneider et al. 2011). Furthermore, culture conditions might influence dimethylation levels. When ESCs are cultured in 2i medium instead of serum, trimethylation levels of H3K27 reduce dramatically (Marks et al. 2012). However, even if H3K27me2 is not

generally distributed throughout the whole genome PRC2 can also counteract acetylation of H3K27 at enhancers by trimethylation (Pinello et al. 2014; Abou El Hassan et al. 2015).

The picture that now emerges constitutes complementing biochemical PRC1 and PRC2 activities, but also shows previously unknown roles in the regulation of transcription. In the next paragraphs we will discuss the molecular determinants involved in recruiting PRC2 to its genomic targets.

3. Sequence context of PRC2 action: Genetic prerogative or epigenetic consequence?

CpG dinucleotide density and its methylation status are good predictors of mammalian PRC2 recruitment. Analysing the DNA underlying PRC2-bound loci for sequence features in mammals revealed an enriched representation of CpG dense regions (Lee et al. 2006). CpG richness is a feature that is also found at the TSS of genes marked by H3K4me3 (Bernstein et al. 2006). Indeed, insertion of CpG-rich elements was sufficient for the recruitment of PRC2 and deposition of H3K27me3, as well as H3K4me3, to exogenous loci in mouse ESCs (Mendenhall et al. 2010). Vice versa, a comparative study of mouse and human ESCs showed that loss of CpG-rich elements resulted in loss of H3K27me3 deposition at these regions (Lynch et al. 2012).

CpG dinucleotides can be subjected to methylation, which prevents them from binding PRC2 (Bartke et al. 2010). Mass spectrometry (MS) based analysis showed that incorporation of methylated CpG DNA in nucleosomes antagonized the binding of PRC2 subunit Eed (Bartke et al. 2010). Indeed, mutual exclusion of CpG-island (CGI) methylation and H3K27me3 deposition was demonstrated in vertebrate genomes (Bogdanovic et al. 2011; Lynch et al. 2012). At loci with low CpG dinucleotide density, however, DNA methylation and H3K27me3 were found to co-occur (Brinkman et al. 2012). Not only CpG density, but also G+C richness is a property of methylation-free regions. Deposition of either H3K4me3 or H3K27me3 is the default chromatin state at these loci, as was shown by integration of artificial CGI-like DNA sequences into the genome of ESCs (Wachter et al. 2014). CpG-richness at promoters is particularly prevalent in mammals. In non-mammalian vertebrates relatively few CpG dinucleotides overlap with gene promoters. Even so, promoters in non-mammalian vertebrates contain non-methylated clusters of CpGs, called non-methylated islands (NMI), which are highly conserved across species (Long et al. 2013b). In *Xenopus* embryos, trimethylation of either H3K27 or H3K4 is closely associated with the presence

of NMIs (van Heeringen et al. 2014). During gastrulation H3K27 trimethylation is acquired in pre-existing hypomethylated regions in *Xenopus*. These studies show conserved PRC2 recruitment to hypomethylated regions in vertebrates.

DNA binding proteins that direct PRC2 toward NMIs might operate via PRC1 (Farcas et al. 2012). Unmethylated CxxC domains can be recognized by Zinc finger (ZF)-CxxC domain proteins, such as KDM2B (Long et al. 2013a). Affinity purification of KDM2B from ESCs followed by MS revealed that it forms a complex with the PRC1 subunit Ring1b. Recruitment of KDM2B to promoters leads to H2AK119ub deposition, followed by PRC2 binding and H3K27me3-mediated silencing (Farcas et al. 2012). Removal of the ZF-CxxC domain of KDM2B resulted in loss of Ring1b binding at roughly half of the Ring1b binding sites in mouse ESCs. In addition, KDM2B binding sites showed reduced levels of ubiquitinated H2AK119 and Suz12 recruitment in KDM2B deficient cells. Targeted KDM2B binding induced local enrichment of Ring1b, H2AK119ub, Ezh2, and H3K27me3, independent of its demethylase activity. Hence, KDM2B mediates

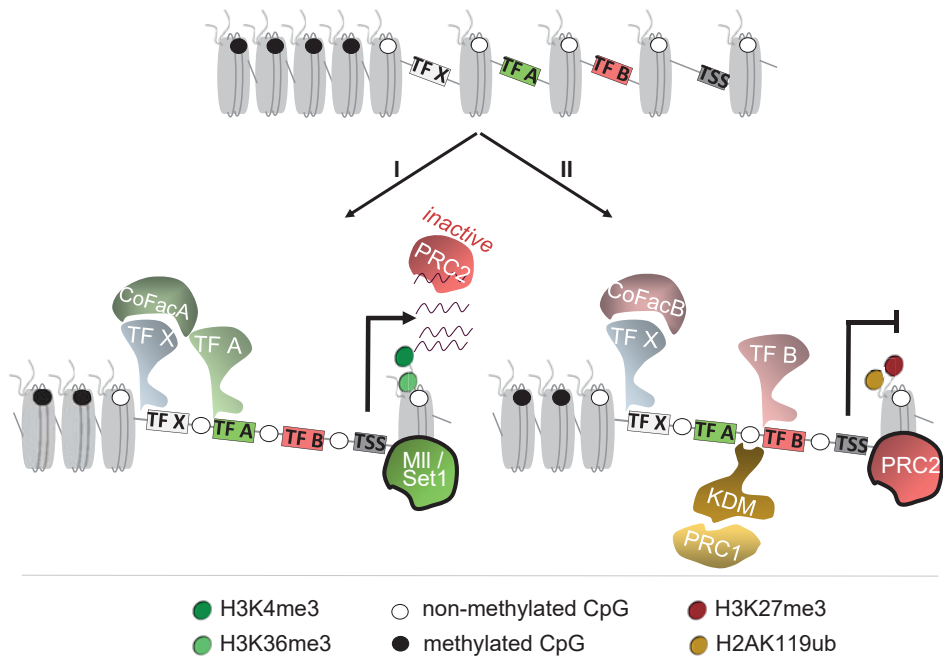


Figure 2. Sequence context of PRC2 action. Non methylated Islands (NMIs) are susceptible for gene activation by TrxG proteins (for example MII, Set1) or repression by PcG proteins. (I) MII catalyses H3K4me3 in the presence of transcription factors (TF) that facilitate binding of MII, creating a permissive state for transcription. PRC2 might recognize these actively transcribed regions by binding nascent RNA, but is antagonized by MII. (II) In the absence of transcription activating factors, PRC2 can bind at NMIs via positioning by TFs or their cofactors (CoFac). Zinc Finger-CxxC domain proteins (KDM) that bind PRC1, can also stimulate PRC2 recruitment by providing a docking site, H2AK119ub.

PCR1 recruitment to NMIs, and is required for PRC2-catalyzed trimethylation of H3K27 at these loci (Blackledge et al. 2014). PRC1-independent recruitment of PRC2 to unmethylated DNA might also occur via PRC2-accessory proteins with DNA binding capacity, such as Jarid2. Jarid2 was shown to co-occur with PRC2 genome-wide, and motif analysis in ESCs showed that Jarid2-PRC2 bound loci were enriched for both CCG-repeats and GA-rich regions (Peng et al. 2009).

Computational analyses to identify sequences that recruit PRC2 suggest a central role for NMIs (figure 2). A Support Vector Machine trained on a subset of sequences underlying H3K27me3 domains, accurately predicted H3K27me3 status of unknown sequences in a cross-species analysis in frog, zebrafish, and human, CpG-density differences between mammals and other vertebrates notwithstanding (van Heeringen et al. 2014). This pan-vertebrate sequence conservation within NMIs suggests that additional genetic factors determine when and where NMIs become marked by H3K27me3 or by H3K4me3. The next section will further discuss the role of specific sequence properties and transcription factor (TF) binding sites in PRC2 recruitment.

4. Interplay of transcription factor binding and PRC2 recruitment

4.1. PcG response elements in Drosophila

The first evidence for motif-specific PRC2 recruitment was found in *Drosophila*. Within the Bithorax complex, a cluster of three homeotic genes which are important in segmental development, specific DNA regulatory elements to which PcG proteins are recruited were identified (Simon et al. 1993). Insertion of these PcG response elements (PREs) in a reporter plasmid resulted in repression of transcription in a PcG-dependent manner (Simon et al. 1993). The first sequence-specific DNA-binding protein that was shown to mediate PcG recruitment to PREs was Pleiohomeotic (Pho). Pho was shown to bind a 17 bp sequence located within a 176 bp fragment located upstream of the engrailed locus, which was previously linked to PcG mediated silencing in transgenic flies. This 17 bp PRE was highly conserved and essential, but not sufficient for the PcG mediated silencing (Brown et al. 1998). Following this discovery multiple more PREs were found in *Drosophila* and these PREs contained binding motifs for various TFs (like Gaga, Pho, and Zeste binding motifs) (reviewed in: Kassis and Brown 2013).

Locations of PREs throughout the genome were computationally predicted based on diverse TF binding motifs that were enriched in experimentally confirmed PREs (Ringrose et al. 2003). However, two independent genome-wide assays proved that PRC2 and PRC1 bind to some, but not the majority of these predicted PREs in *Drosophila* (Schwartz et al. 2006; Tolhuis et al. 2006). Genome-wide studies that characterized the binding sites of various sequence-specific DNA-binding proteins have shown co-occupancy of multiple TFs, suggesting a cooperative recruitment of PcG components in *Drosophila*. However, many of the putative PcG recruiters (TFs like Pho and Gaga) were not solely enriched at PcG binding sites, but also at the H3K4me3-associated TrxG binding sites (Schuettengruber et al. 2009). These results imply that different factors work together to recruit PcG proteins or that these TFs have another function besides PcG repression.

Recently, a study on the function and evolution of PREs shed new light on the functionality, specificity, and cooperativity of PcG recruiters (Schuettengruber et al. 2014). Comparing H3K27 methylation, PH (PRC1) binding, and DNA sequence in five different *Drosophila* species showed that, despite variations in the underlying sequence, PcG domains were highly conserved in syntenic regions. Unexpectedly, not the DNA sequence, but the TF binding itself was highly conserved, with both Pho and Dorsal Switch Protein (Dsp1) binding to low specificity sites at the PcG domains. Cooperative binding sites for Pho and Dsp1 showed the highest overlap with PcG domains, and prediction of Pho binding was more accurate as a function of PH binding and Pho motifs, compared to TF motifs alone. This suggests a bidirectional interaction between PcG proteins and other proteins, stabilizing the PcG domains (Schuettengruber et al. 2014).

4.2. PcG and transcription factor motifs in vertebrates

PRE-like mechanisms of PRC2 recruitment have been elusive in vertebrates as no clear ortholog to any of the *Drosophila* PRC2-recruiting factors has been found. However, a variety of TFs influence PRC2 recruitment in vertebrates. The first H3K27me3 and PcG profiling studies in ESCs already suggested a possible relation between PcG proteins and TFs, based on the co-localization of PcG components with pluripotency factors Oct4, Sox2 and Nanog (Bernstein et al. 2006; Boyer et al. 2006; Lee et al. 2006). More recent studies suggest that the correlation between DNA sequence and histone modifications might be the result of TF-mediated recruitment of histone modifiers (figure 2) (Benveniste et al. 2014). Analyses of TF binding from genome-wide profiling studies in

H1 cells, K562 cells, and GM12878 cells demonstrated that TF binding more accurately predicted the presence of H3K4me1, H3K4me3, H3K9ac, H3K27ac or H3K27me3 at promoters and enhancers, compared to the DNA sequence itself. This indicates that TFs might form a link between specific DNA sequences and the histone modifiers (Benveniste et al. 2014).

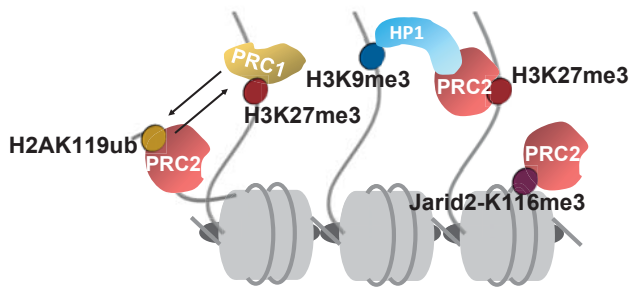
Conversely, deletion of motifs for transcription activators from NMIs was found to be sufficient for PRC2 recruitment and H3K27me3 deposition in ESCs (Mendenhall et al. 2010). Minimal DNA sequence elements capable of autonomously recruiting PRC2 were recently defined by using iterative genome editing in mouse ESCs. This demonstrated the influence of surrounding sequences on PRC2 recruitment, as an active enhancer-promoter sequence surrounding CG-rich sequences was shown to prevent PRC2 recruitment and trimethylation of H3K27 at these loci (Jermann et al. 2014). Jermann and colleagues proposed that CGIs bind PRC2 by default, provided that they are devoid of DNA methylation and are not transcriptionally active. Inhibition of RNA polymerase II was indeed sufficient to obtain Suz12 binding and trimethylation of H3K27me3 in mouse ESCs (Riising et al. 2014). Sites with increased H3K27me3 upon transcriptional inhibition were found to be ectopic CpG targets in other, differentiated tissues. A genetic-default model for PRC2 action was also suggested by Van Heeringen and colleagues, based on the observation that the pan-vertebrate conserved DNA sequence signatures of H3K27me3 are linked to a propensity for H3K27me3 across different cell types. This suggests that methylation of H3K27 is default at these regions and is actively prevented by cell type-specific factors (van Heeringen et al. 2014).

Besides the absence of particular transcription activators, PRC2 recruitment correlates also with the presence of specific TF motifs. Distinctive motif contributions were identified when comparing Ezh2-positive and -negative NMIs in ESCs. Ezh2-negative NMIs were marked by H3K4me3, and showed strong enrichment for motifs of transcriptional activators like NFY, Myc, and Ets1. In contrast, Ezh2-positive NMIs were mostly H3K27me3 enriched, and were associated with motifs for TFs that are known to be expressed in ESCs: NESF/REST, Cux1, and NFκB (Ku et al. 2008). In *Xenopus*, NMIs that gain H3K4me3 are enriched for motifs that bind housekeeping TFs. NMIs that gain H3K27me3, on the other hand, generally contain motifs for developmental regulators, like Sox and homeobox TFs (van Heeringen et al. 2014). Binding sites that were predicted to recruit PcG components in motif analyses, such as for Rest and Runx1, induced ectopic H3K27 methylation. Furthermore, their respective TFs were shown to physically interact with PcG proteins (Dietrich et al. 2012; Yu et al. 2012; Arnold

et al. 2013). For example, regions that obtained H3K27me3 during neurogenesis were enriched for a specific set of motifs, among which binding sites for Rest and Snail. Insertion of Rest and Snail motifs was sufficient to ectopically induce H3K27 methylation in mouse ESCs (Arnold et al. 2013). More recently, a study in *Xenopus* showed that Snail2 cooperates with PRC2 via Ezh2 binding, which is important in modulating the expression of neural crest genes. Co-occupancy of Snail2 and Ezh2 was shown to be important for maintenance of H3K27me3 levels and expansion of the neural crest domain (Tien et al. 2015).

However, TFs can also be involved in both transcriptional activation or repression depending on the environmental context, which comprises CpG density and available co-factors (Arnold et al. 2013; Pinello et al. 2014). For example, Rest binding during neurogenesis was shown to increase trimethylation of H3K27 at CpG-rich loci, but to decrease trimethylation of H3K27 at CpG poor loci upon

A Recruiting histone modifications



B Inactivating histone modifications

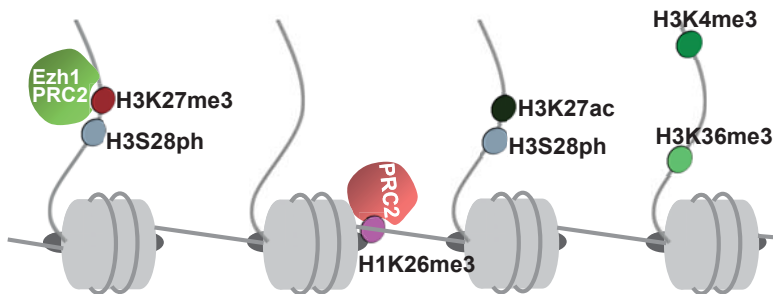


Figure 3. PRC2 guidance by modified histones. A) Multiple posttranslational modifications stimulate the recruitment of PRC2. PRC2 can bind to H3K27me3 and H2AK119ub. Binding to these marks or to trimethylated Jarid2-K119 stimulates its activity. On heterochromatic regions, PRC2 binding to H3K27me3 and HP1 binding to H3K9me3 cooperate to facilitate formation and maintenance of heterochromatic state. B) Histone modifications that inactivate PRC2 are H3K27ac, H3K4me3 and H3K36me3. These modifications inactivate PRC2 when they are located on the same histone tail as where the complex is located. H1K26me3 inactivates PRC2 after binding the complex. When H3S28ph is positioned next to H3K27me3, Ezh2 is repelled and exchanged for Ezh1.

differentiation (Arnold et al. 2013). Environmental effects could also be a result of differential co-factor binding, which has been suggested to contribute to cell type-specific PcG recruitment (figure 2). A recent analysis of H3K27me3 profiles in 19 different cell lines identified regions with variable H3K27me3 deposition across cell-lines, so called high plasticity regions (HPRs). HPRs were found at both CGIs surrounding TSSs as well as distal elements. Motif analysis yielded 41 cell-type specific associations between TF motifs and distal HPRs. Genome-wide binding profiles showed that binding of these TFs was indeed enriched at HPRs. Tal1 binding correlated with HPRs in primary human erythroid progenitor cells, however, its capacity to recruit PRC2 was found to be determined by co-factor binding, rather than Tal1 binding itself. Inactive, H3K27me3 marked enhancers were generally occupied by Tal1-GFI1B, whereas Tal1-Gata1 was found at active, H3K27ac marked enhancers (Pinello et al. 2014).

These studies highlight the complex relationships between the binding of sequence-specific activators and repressors and the recruitment of PRC2 but fall short of establishing that PRE-like mechanisms of PRC2 recruitment also exist in vertebrates. TFs and cofactors can be used to separate NMIs targeted for transcription activation or repression. In addition to DNA binding factors, pre-existing histone modifications and chromatin structure are also important factors in proper PRC2 targeting, as is discussed in the next section.

5. Responsive PRC2 binding: management by modified histones

5.1. Nucleosome density

Chromatin structure can direct PRC2 binding in two ways, namely by nucleosome density and by crosstalk with histone modifications (figure 3). Binding sites for PcG and TrxG proteins have a relatively high histone replacement rate and a low nucleosome occupancy, as was shown at the homeotic gene clusters in fly (Mito et al. 2007). Contradictory, PRC2 binding and activity was increased when comparing dinucleosomes with mononucleosomes (Martin et al. 2006). Despite the relatively high histone replacement rate for PcG proteins in fly, nucleosome turnover rate is higher in regions occupied by TrxG proteins compared to regions bound by PcG proteins (Deal et al. 2010).

Despite the diminished nucleosome density at CGIs prior to PRC2 recruitment, nucleosome compaction seems to increase at these loci just before PRC2 binding (Yuan et al. 2012). Yuan and colleagues tested whether the density of the substrate chromatin could regulate PRC2. They found that preventing

transcription activation for the gene *AYP26a1* in mouse ESCs by withdrawal of retinoic acid resulted in increased nucleosome density prior to H3K27me3 deposition (Yuan et al. 2012). CGIs that became PRC2 targets upon transcription inhibition in mouse ESCs, also showed lower nucleosome density prior to PRC2 binding, compared to CGIs that did not recruit PRC2 (Riising et al. 2014). Recently Tee and colleagues described how altering the chromatin accessibility upon Erk1/2 binding can stimulate PRC2 recruitment in ESCs (Tee et al. 2014). These studies indicate that PcG targets have a relatively low nucleosome density, which already becomes denser just before binding of the complex.

5.2. Stimulating PRC2 binding

Pre-existing histone modifications such as H3K27me3, H2AK119ub, and H3K9me3 can facilitate PRC2 recruitment (figure 3A). These epigenetic marks are partially transmitted during cell proliferation, and reconstituted by means of positive feedback. For example, PRC2 was shown to bind to its own catalytic product, H3K27me3, by the aromatic cage of Eed (Margueron et al. 2009; Xu et al. 2010). Eed was shown to recognize trimethylated histone peptides, with a particularly high affinity for H3K27me3, H1K26me3, and H3K9me3 (Xu et al. 2010). Furthermore, Eed binding to H3K27me3 results in allosteric activation of the complex and propagation of the mark, as was shown in vitro and in *Drosophila* (Margueron et al. 2009; Xu et al. 2010). In the absence of pre-existing H3K27me3, methylated Jarid2 was suggested to facilitate PRC2 recruitment. Interestingly, methylation of Jarid2 at lysine K116 is mediated by PRC2 itself. Jarid2-K116me3 is recognized by Eed, which in turn triggers an allosteric activation of PRC2's enzymatic activity. Jarid2-K116me3, but not unmethylated Jarid2, was found to have a higher affinity for Eed compared to H3K27me3. Knockdown of Jarid2, or introduction of a methylation-deficient Jarid2 had no consequences for ESCs, but caused disturbed H3K27me3 patterns in differentiated embryoid bodies. This suggests that pre-existing H3K27me3 accounts for the maintenance of H3K27me3 during cell division, whereas the nucleation of new domains during cell differentiation is dependent on Jarid2-K116me3 (Sanulli et al. 2015).

H3K27me3 can also serve as a docking site for PRC1 component Cbx (Cao et al. 2002; Wang et al. 2004; Boyer et al. 2006). The Ring1 subunit of PRC1 can catalyse H2AK119 ubiquitylation (de Napoles et al. 2004), which in turn can serve as a docking site for PRC2 (Blackledge et al. 2014; Cooper et al. 2014; Kalb et al. 2014). PRC2 components were strongly enriched in affinity pull downs with either H2AK118ub or H2AK119ub using *Drosophila* or mouse ESC

nuclear extracts, respectively. These studies demonstrate that ubiquitinated H2A serves as a binding site for Jarid2–Aebp2–containing PRC2 and promotes H3K27 trimethylation (Kalb et al. 2014). Binding of a MBD-Ring1b/Pcgf4 fusion protein to densely CpG methylated DNA resulted in H2AK119ub deposition in mouse. This was sufficient to establish H3K27me3 at paternal pericentric heterochromatin (PCH) domains (Cooper et al. 2014). In a separate study, Tet-repressor fusion proteins were used to recruit PRC1 to a Tet-operator site that was introduced in the mouse genome. The Tet-repressor was fused to Pcgf 1, 2, 3, 4, or 5, which are known to be present in different PRC1 complexes. Although Ring1b was recruited with every Pcgf fusion variant, profound ubiquitylation of H2AK119 only occurred in the presence of Pcgf1, 3, and 5. Fusion proteins that could mediate H2AK119ub enrichment, also recruited catalytically active PRC2 to the site (Blackledge et al. 2014). These studies suggest that PRC2 and PRC1 positively influence each other's recruitment.

Methylated H3K9 is also associated with recruitment of PRC2. Proteome analysis in mouse ESCs uncovered that H3K9me3 and H3K27me3 are rarely found on the same peptide, but do co-occur in an asymmetric composition on different histone H3 tails (Voigt et al. 2012; Sidoli et al. 2014). Eed has strong affinity for H3K9me3, however, in vitro methylation assays showed that the binding of PRC2 to H3K9me3 substrates does not change the methyltransferase activity of Ezh2 (Xu et al. 2010). In HeLa and mouse ES cells, PRC2 and H3K9 methyltransferase G9a/GLP were shown to have a physical interaction, and genome-wide profiling of G9a/GLP binding revealed 25% overlap with PRC2 loci. H3K27me3 methylation at these shared binding sites was decreased in G9a and/or GLP deficient cells, independent of the derepression of these targets. Binding of G9a, but not of a G9a catalytically dead mutant, to an artificial docking site resulted in Ezh2 recruitment and trimethylation of H3K27. In addition, disturbed Ezh2 binding in G9a mutants ESCs could be rescued by wild type G9a, but not by a G9a catalytically inactive protein (Mozzetta et al. 2014).

Another way by which methylation of H3K9 recruits PRC2 is via the structural adaptor protein HP1 (Boros et al. 2014). In a pulldown experiment with H3 tail peptides methylated at H3K9 and/or H3K27, H3K27me3 was found to increase H3K9me3 dependent HP1 binding. Knockdown of Ezh2 in human fibrosarcoma cells caused proteasomal degradation of HP1, and overexpression of H3K27me2/3 demethylase resulted in removal of HP1 from chromatin, both independent of changes for H3K9me3. Hence PRC2 and H3K27me3 cooperate with H3K9me3 to facilitate heterochromatin formation and maintenance, by stabilizing HP1 binding (Boros et al. 2014).

5.3. PRC2 blockers

Histone modifications associated with transcription activation, such as H1K26me3, H3K27ac, H3S28ph, H3K36me3, and H3K4me3, are thought to inhibit PRC2 recruitment (figure 3B). PRC2 can be diverted from its target sites, via docking of the complex to H1K26me3 substrates. H1K26me3 competes with H3K27me3 and H3K9me3 for binding in the aromatic cage of Eed. Docking to trimethylated H1K26, however, decreases the enzymatic activity of PRC2 (Xu et al. 2010).

Acetylation of H3K27 and methylation of the same residue are mutually exclusive but the two modifications could occur at separate histone H3 tails within the same nucleosome. However, H3K27me_{2/3} containing nucleosomes that also contain H3K27ac could hardly be detected by MS on mononucleosomes from mouse ESCs, mouse embryonic fibroblasts, and HeLa cells (Voigt et al. 2012). Genome-wide profiling in *Drosophila* embryos and mouse ESCs revealed that acetylation and methylation of H3K27 are inversely related; H3K27me₃ was found to increase at loci where H3K27ac was decreased, and vice versa (Tie et al. 2009; Pasini et al. 2010). It was shown in mouse ESCs that NuRD-dependent deacetylation of H3K27 indeed led to recruitment of catalytically active PRC2 (Reynolds et al. 2011). In *Drosophila* embryos several histone modifying enzymes are in proximity to nascent DNA already 5 minutes after replication, including the ortholog of Ezh2 (E(z)), the H3K27 acetyltransferase CPB, and H3K27 demethylase UTX. Acetylation of H3K27 was achieved within 10 minutes after replication. In contrast, H3K27me₃ could not be detected until one hour after replication (Petruk et al. 2013). The balance between acetylation and methylation of H3K27 changed upon treatment with inhibitors for CPB or UTX, showing trimethylation of H3K27 15 minutes after replication, together with a decreased acetylation of H3K27. This suggests that acetylation and demethylation of H3K27 are important to prevent aberrant deposition and accumulation of H3K27me₃ (Petruk et al. 2013).

Acetylation of H3K27 might be facilitated by phosphorylation of the flanking serine residue S28. Targeting the H3S28 phosphatase Msk1 to the endogenous promoter of α -globulin in HEK293 cells resulted in transcription activation of the gene. At the α -globulin promoter both H3S28Ph and H3K27ac levels were increased and present on the same histone tail, while H3K27me₃ levels were decreased (Lau and Cheung 2011). In HeLa cells stress activation led to increased phosphorylation of H3S28 on histone tails that were also trimethylated on H3K27, resulting in decreased binding of Cbx8 and Suz12 (Gehani et al. 2010). A separate study on PRC2 binding at the myogenin promoter during skeletal

muscle cell differentiation showed that increased Msk1 and H3S28ph binding during transcriptional activation resulted in displacement of Ezh2, but not Ezh1, at the promoter (see Box 1) (Stojic et al. 2011). Similar results were obtained in affinity-purification experiments from extracts of differentiated myotubes using histone H3 tail peptides that were unmodified, or modified with K27me3 or K27me3/S28ph. Ezh1 bound with comparable affinity to both K27me3 and K27me3/S28ph modified peptides, whereas Ezh2 binding was significantly weakened in the presence of S28ph (Stojic et al. 2011).

In the fly, trimethylated H3K4 and H3K36, catalysed by Trx and Ash respectively (Mll and Setd2 in mammals), antagonize PcG-mediated silencing. Affinity assays showed that the binding of Su(z)12 in complex with Nurf55 (Suz12 and Rbbp4/RbAp48, Rbbp7/RbAp46 in mammals) to H3 peptides could significantly be reduced if the H3 peptides were methylated on lysine K4. In absence of Nurf55, H3-Su(z)12 binding was not affected, however, H3K4me3 and H3K36me3 did inhibit the catalytic activity of PRC2. Inhibition of di- and trimethylation by PRC2 was observed on H3 tails also trimethylated on K4 or K36, but not when these modifications were present on separate peptides (Schmitges et al. 2011). Though, in vivo trimethylation of H3K4 and H3K36 is rarely detected on H3 tails that are also tri-methylated for H3K27 (Sidoli et al. 2014; Yuan et al. 2011).

However, co-occurrence of H3K27me3 and H3K4me3 on different H3 tails in the same nucleosome has been reported (Voigt et al. 2012). MS on H3K4me3-containing mononucleosomes showed the presence of H3K27me3 and H3K4me3 within the same nucleosome, which was higher in mouse ESCs (approximately 15% of H3K4me3-containing nucleosomes) compared to mouse embryonic fibroblasts (Voigt et al. 2012). In *Drosophila* and *Xenopus*, significant co-occurrence of H3K27me3 and H3K4me3 within the same nucleosomal DNA population could not be detected (Akkers et al. 2009; Schuettengruber et al. 2009; Gan et al. 2010). In addition, when ESCs were cultured in 2i medium instead of serum, trimethylation levels of H3K27, and consequently the H3K27me3/K4me3 bivalent state, reduced dramatically (Marks et al. 2012). However, various studies showed that PRC2 can be recruited to actively transcribed genes via Polycomb-like (PCL) proteins which can bind to H3K36me3 (Ballaré et al. 2012; Musselman et al. 2012a; Cai et al. 2013). PCL protein Phf19 not only interacts with PRC2 but also interacts with H3K36me3 demethylase NO66; therefore, PCL proteins might recruit PRC2 to set up repression (Brien et al. 2012).

6. RNA regulated recruitment

Despite the repressive effect of H3K36me3 and H3K4me3 on PcG mediated silencing, PRC2 recruitment has also been positively associated with active transcription. Highly expressed genes showed monomethylated H3K27, which was dependent on H3K36me3, whereas lowly expressed genes accumulated dimethylation at H3K27 throughout the gene bodies (Ferrari et al. 2014). Knockdown of H3K36 methyltransferase Setd2 resulted in a loss of both H3K36me3 and H3K27me1, in addition to accumulation of H3K27me2 at these intergenic regions. Loss of PRC2 reduced accumulation of both H3K27me1 and H3K27me2, but not of H3K36me3. Furthermore, Eed deletion led to transcriptional upregulation of H3K27me2-marked genes and downregulation of H3K27me1-marked genes. MS data on H3K36me3 purified histones confirmed the presence of both K27me1 and K36me3 on the same H3 peptide (Ferrari et al. 2014). These results indicate that the methylation state of H3K36 regulates PRC2 action and subsequently determines methylation of H3K27.

These results suggest a role for PRC2 in actively transcribed genes, even though the presence of stable PRC2-binding could not be detected at these regions. One way by which PRC2 could be recruited to active genes is through interaction with RNA molecules. Multiple studies have reported binding of specific RNAs to PRC2, including non-coding (nc, lnc) RNAs such as Xist repA ncRNA in X-chromosome silencing (Zhao et al. 2008; da Rocha et al. 2014), and HOTAIR ncRNA in silencing of hox genes in human (Rinn et al. 2007; Tsai et al. 2010). In addition, lncRNAs were recently shown to function as scaffolds, stabilizing the binding between various PRC2 subunits such as Ezh2 and Jarid2 (Kaneko et al. 2014a).

In addition to sequence specific RNA-binding, PRC2 was also reported to bind RNA molecules in a nonselective manner. RNA immunoprecipitation in ESCs showed PRC2 to associate with thousands of different RNA molecules (Zhao et al. 2010; Kaneko et al. 2013). Quantitative electrophoretic mobility shift assays (EMSA) of reconstituted human PRC2 with various RNA molecules revealed that PRC2 binding is size-dependent rather than sequence-dependent, with lower affinity for shorter RNA molecules (Davidovich et al. 2013). The majority of the PRC2-bound RNA sequences corresponded to the 5'- regions of genes that were transcriptionally active. ChIP-sequencing data from various mouse cell lines revealed that the genes belonging to these PRC2 bound-RNAs were positively associated with Ezh2 recruitment and trimethylation of H3K4 and H3K36, but were depleted of H3K27me3 (Davidovich et al. 2013; Kaneko et al. 2013). Interestingly, H3K27me3 on Ezh2-RNA genes was more pronounced in differentiated mouse embryonic fibroblasts, as compared to pluripotent ESCs (Kaneko et al. 2013). RNA binding was shown to suppress the histone methyltransferase activity of Ezh2,

although the RNA binding affinity of Ezh2 was reduced when bound to other PRC2 subunits (Cifuentes-Rojas et al. 2014). Di- and trimethylation of H3K27 on Ezh2-RNA genes could be induced by CRISPR-mediated truncation of the 5'-end these genes (Kaneko et al. 2014b). Together, these studies support a model in which PRC2 uses RNA binding to scan the genome, sensing the transcriptional activity of genes and deploying or redistributing the complex accordingly (figure 2B).

7. Conclusion and perspective

A growing body of evidence indicates that RNA transcripts, pre-existing histone modifications, and transcription factors together define a local chromatin state which controls accurate, cell-type specific epigenetic silencing by PRC2. Genetic sequence sets the fate for potential PRC2 targets, but the timing of stable PRC2-binding at these loci is influenced by TFs. Forming complexes with the different Ezh paralogs can result in different outcomes with respect to PRC2's function in transcription regulation. This suggests that lineage-specific TFs are involved in determining the transcriptional output of potential PRC2 targets by modulating both the complex composition and the recruitment of the complex. Exactly which TFs are involved in regulating the expression of PcG target genes and in guiding of PcG proteins towards their targets remains one of the key questions to be addressed. Further studies are needed to uncover how TFs and their co-factors influence PRC2 regulation.

PRC2 also senses pre-existing histone modifications and binds to nascent RNA molecules, so that the complex can respond appropriately to different cellular states. The exact order of molecular events that specify these cellular states and their interplay remain to be elucidated. Resolving these molecular mechanisms will be both important and rewarding, as PcG mediated transcriptional repression is essential for maintenance of cellular identity.

8. Acknowledgements

This work has been supported by the US National Institutes of Health (NICHHD, grant R01HD069344).

References

- Abou El Hassan M, Yu T, Song L, Bremner R. Polycomb Repressive Complex 2 Confers BRG1 Dependency on the CIITA Locus. *J Immunol*. 2015 Apr 10.
- Akkers RC, van Heeringen SJ, Jacobi UG, Janssen-Megens EM, François K-J, Stunnenberg HG, et al. A hierarchy of H3K4me3 and H3K27me3 acquisition in spatial gene regulation in *Xenopus* embryos. *Dev Cell*. 2009 Sep;17(3):425–34.
- Aloia L, Di Stefano B, Di Croce L. Polycomb complexes in stem cells and embryonic development. *Development*. 2013 Jun 15;140(12):2525–34.
- Arnold P, Schöler A, Pachkov M, Balwierz PJ, Jørgensen H, Stadler MB, et al. Modeling of epigenome dynamics identifies transcription factors that mediate Polycomb targeting. *Genome Res*. 2013 Jan 1;23(1):60–73.
- Ballaré C, Lange M, Lapinaite A, Martin GM, Morey L, Pascual G, et al. Phf19 links methylated Lys36 of histone H3 to regulation of Polycomb activity. *Nat Struct Mol Biol*. 2012 Dec;19(12):1257–65.
- Bartke T, Vermeulen M, Xhemalce B, Robson SC, Mann M, Kouzarides T. Nucleosome-interacting proteins regulated by DNA and histone methylation. *Cell*. 2010 Oct 29;143(3):470–84.
- Benveniste D, Sonntag H-J, Sanguinetti G, Sproul D. Transcription factor binding predicts histone modifications in human cell lines. *Proc Natl Acad Sci U S A*. 2014 Sep 16;111(37):13367–72.
- Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, et al. A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell*. 2006;125:315–26.
- Blackledge NP, Farcas AM, Kondo T, King HW, McGouran JF, Hanssen LLP, et al. Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. *Cell*. 2014 Jun 5;157(6):1445–59.
- Bogdanovic O, Long SW, van Heeringen SJ, Brinkman AB, Gómez-Skarmeta JL, Stunnenberg HG, et al. Temporal uncoupling of the DNA methylome and transcriptional repression during embryogenesis. *Genome Res*. 2011 Aug 1;21(8):1313–27.
- Boros J, Arnoult N, Stroobant V, Collet J-F, Decottignies A. Polycomb repressive complex 2 and H3K27me3 cooperate with H3K9 methylation to maintain heterochromatin protein 1 α at chromatin. *Mol Cell Biol*. 2014 Oct 1;34(19):3662–74.
- Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, et al. Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature*. 2006 May 18;441(7091):349–53.
- Brien GL, Gambero G, O'Connell DJ, Jerman E, Turner SA, Egan CM, et al. Polycomb PHF19 binds H3K36me3 and recruits PRC2 and demethylase NO66 to embryonic stem cell genes during differentiation. *Nat Struct Mol Biol*. 2012 Dec;19(12):1273–81.
- Brinkman AB, Gu H, Bartels SJJ, Zhang Y, Matarese F, Simmer F, et al. Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. *Genome Res*. 2012 Jun;22(6):1128–38.
- Brown JL, Mucci D, Whiteley M, Dirksen M-L, Kassis JA. The *Drosophila* Polycomb Group Gene pleiohomeotic Encodes a DNA Binding Protein with Homology to the Transcription Factor YY1. *Mol Cell*. 1998 Jun;1(7):1057–64.
- Cai L, Rothbart SB, Lu R, Xu B, Chen W-Y, Tripathy A, et al. An H3K36 methylation-engaging Tudor motif of polycomb-like proteins mediates PRC2 complex targeting. *Mol Cell*. 2013 Mar 7;49(3):571–82.
- Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, et al. Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing. *Science*. 2002 Nov 1;298(5595).
- Cao R, Zhang Y. SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. *Mol Cell*. 2004 Jul 2;15(1):57–67.
- Cifuentes-Rojas C, Hernandez AJ, Sarma K, Lee JT. Regulatory interactions between RNA and polycomb repressive complex 2. *Mol Cell*. 2014 Jul 17;55(2):171–85.
- Cooper S, Dienstbier M, Hassan R, Schermelleh L, Sharif J, Blackledge NP, et al. Targeting polycomb to pericentric heterochromatin in embryonic stem cells reveals a role for H2AK119u1 in PRC2 recruitment. *Cell Rep*. 2014 Jun 12;7(5):1456–70.
- Di Croce L, Helin K. Transcriptional regulation by Polycomb group proteins. *Nat Struct Mol Biol*. 2013 Oct;20(10):1147–55.
- Czermin B, Melfi R, McCabe D, Seitz V, Imhof A, Pirrotta V. *Drosophila* Enhancer of Zeste/ESC Complexes Have a Histone H3 Methyltransferase Activity that Marks Chromosomal Polycomb Sites. *Cell*. 2002 Oct;111(2):185–96.
- Davidovich C, Zheng L, Goodrich KJ, Cech TR. Promiscuous RNA binding by Polycomb repressive complex 2. *Nat Struct Mol Biol*. 2013 Nov;20(11):1250–7.
- Deal RB, Henikoff JG, Henikoff S. Genome-wide kinetics of nucleosome turnover determined by metabolic labeling of histones. *Science*. 2010 May 28;328(5982):1161–4.
- Dietrich N, Lerdrup M, Landt E, Agrawal-Singh S, Bak M, Tommerup N, et al. REST-mediated recruitment of

- polycomb repressor complexes in mammalian cells. *PLoS Genet.* 2012 Jan;8(3):e1002494.
- Duncan IM. Polycomblike: a gene that appears to be required for the normal expression of the bithorax and antennapedia gene complexes of *Drosophila melanogaster*. *Genetics.* 1982 Sep;102(1):49–70.
- Farcas AM, Blackledge NP, Suddbery I, Long HK, McGouran JF, Rose NR, et al. KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands. *Elife.* 2012 Jan 18;1:e00205.
- Ferrari KJ, Scelfo A, Jammula S, Cuomo A, Barozzi I, Stützer A, et al. Polycomb-Dependent H3K27me1 and H3K27me2 Regulate Active Transcription and Enhancer Fidelity. *Mol Cell.* 2014 Jan 9;53(1):49–62.
- Gan Q, Schones DE, Ho Eun S, Wei G, Cui K, Zhao K, et al. Monovalent and unpoised status of most genes in undifferentiated cell-enriched *Drosophila* testis. *Genome Biol.* 2010 Jan;11(4):R42.
- Gao Z, Zhang J, Bonasio R, Strino F, Sawai A, Parisi F, et al. PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes. *Mol Cell.* 2012 Feb 10;45(3):344–56.
- Gehani SS, Agrawal-Singh S, Dietrich N, Christophersen NS, Helin K, Hansen K. Polycomb group protein displacement and gene activation through MSK-dependent H3K27me3S28 phosphorylation. *Mol Cell.* 2010 Sep 24;39(6):886–900.
- Van Heeringen SJ, Akkers RC, van Kruijsbergen I, Arif MA, Hanssen LLP, Sharifi N, et al. Principles of nucleation of H3K27 methylation during embryonic development. *Genome Res.* 2014 Mar 1;24(3):401–10.
- Jermann P, Hoerner L, Burger L, Schübeler D. Short sequences can efficiently recruit histone H3 lysine 27 trimethylation in the absence of enhancer activity and DNA methylation. *Proc Natl Acad Sci U S A.* 2014 Aug 19;111(33):E3415–21.
- Kalb R, Latwiel S, Baymaz HI, Jansen PWTC, Müller CW, Vermeulen M, et al. Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression. *Nat Struct Mol Biol.* 2014 Jun;21(6):569–71.
- Kaneko S, Bonasio R, Saldaña-Meyer R, Yoshida T, Son J, Nishino K, et al. Interactions between JARID2 and noncoding RNAs regulate PRC2 recruitment to chromatin. *Mol Cell.* 2014a Jan 23;53(2):290–300.
- Kaneko S, Son J, Bonasio R, Shen SS, Reinberg D. Nascent RNA interaction keeps PRC2 activity poised and in check. *Genes Dev.* 2014b Oct 15;28(18):1983–8.
- Kaneko S, Son J, Shen SS, Reinberg D, Bonasio R. PRC2 binds active promoters and contacts nascent RNAs in embryonic stem cells. *Nat Struct Mol Biol.* Nature Publishing Group; 2013 Nov 20;20(11):1258–64.
- Kassis JA, Brown JL. Polycomb group response elements in *Drosophila* and vertebrates. *Adv Genet.* 2013 Jan;81:83–118.
- Ku M, Koche RP, Rheinbay E, Mendenhall EM, Endoh M, Mikkelsen TS, et al. Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. van Steensel B, editor. *PLoS Genet.* 2008 Oct;4(10):e1000242.
- Kuzmichev A, Nishioka K, Erdjument-Bromage H, Tempst P, Reinberg D. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev.* 2002 Nov 15;16(22):2893–905.
- Lau PNI, Cheung P. Histone code pathway involving H3 S28 phosphorylation and K27 acetylation activates transcription and antagonizes polycomb silencing. *Proc Natl Acad Sci U S A.* 2011 Feb 15;108(7):2801–6.
- Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, et al. Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell.* 2006 Apr 21;125(2):301–13.
- Lewis EB. A gene complex controlling segmentation in *Drosophila*. *Nature.* 1978 Dec 7;276(5688):565–70.
- Long HK, Blackledge NP, Klose RJ. ZF-CxxC domain-containing proteins, CpG islands and the chromatin connection. *Biochem Soc Trans.* 2013a Jun 1;41(3):727–40.
- Long HK, Sims D, Heger A, Blackledge NP, Kutter C, Wright ML, et al. Epigenetic conservation at gene regulatory elements revealed by non-methylated DNA profiling in seven vertebrates. *Elife.* 2013b Jan 26;2:e00348.
- Lynch MD, Smith AJH, De Gobbi M, Flenley M, Hughes JR, Vernimmen D, et al. An interspecies analysis reveals a key role for unmethylated CpG dinucleotides in vertebrate Polycomb complex recruitment. *EMBO J.* 2012 Jan 18;31(2):317–29.
- Margueron R, Justin N, Ohno K, Sharpe ML, Son J, Drury WJ, et al. Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature.* 2009 Oct 8;461(7265):762–7.
- Margueron R, Li G, Sarma K, Blais A, Zavadil J, Woodcock CL, et al. Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. *Mol Cell.* 2008 Nov 21;32(4):503–18.
- Marks H, Kalkan T, Menafra R, Denissov S, Jones K, Hofmeister H, et al. The transcriptional and epigenomic foundations of ground state pluripotency. *Cell.* 2012 Apr 27;149(3):590–604.
- Martin C, Cao R, Zhang Y. Substrate preferences of the EZH2 histone methyltransferase complex. *J Biol Chem.* 2006 Mar 31;281(13):8365–70.
- Mendenhall EM, Koche RP, Truong T, Zhou VW, Issac B, Chi AS, et al. GC-Rich Sequence Elements Recruit PRC2 in Mammalian ES Cells. *PLoS Genet.* 2010;6(12):1–10.
- Mito Y, Henikoff JG, Henikoff S. Histone replacement marks the boundaries of cis-regulatory domains. *Science.* 2007 Mar 9;315(5817):1408–11.

- Mousavi K, Zare H, Wang AH, Sartorelli V. Polycomb protein Ezh1 promotes RNA polymerase II elongation. *Mol Cell*. 2012 Jan 27;45(2):255–62.
- Mozzetta C, Pontis J, Fritsch L, Robin P, Portoso M, Proux C, et al. The histone H3 lysine 9 methyltransferases G9a and GLP regulate polycomb repressive complex 2-mediated gene silencing. *Mol Cell*. 2014 Jan 23;53(2):277–89.
- Müller J, Hart CM, Francis NJ, Vargas ML, Sengupta A, Wild B, et al. Histone Methyltransferase Activity of a Drosophila Polycomb Group Repressor Complex. *Cell*. 2002 Oct;111(2):197–208.
- Musselman CA, Avvakumov N, Watanabe R, Abraham CG, Lalonde M-E, Hong Z, et al. Molecular basis for H3K36me3 recognition by the Tudor domain of PHF1. *Nat Struct Mol Biol*. 2012a Dec;19(12):1266–72.
- Musselman CA, Lalonde M-E, Côté J, Kutateladze TG. Perceiving the epigenetic landscape through histone readers. *Nat Struct Mol Biol*. 2012b Dec;19(12):1218–27.
- De Napoles M, Mermoud JE, Wakao R, Tang YA, Appanah R, et al. Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev Cell*. 2004 Nov;7(5):663–76.
- Nekrasov M, Wild B, Müller J. Nucleosome binding and histone methyltransferase activity of Drosophila PRC2. *EMBO Rep*. 2005 Apr;6(4):348–53.
- Pasini D, Bracken AP, Jensen MR, Lazzarini Denchi E, Helin K. Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *EMBO J*. 2004 Oct 13;23(20):4061–71.
- Pasini D, Malatesta M, Jung HR, Walfridsson J, Willer A, Olsson L, et al. Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes. *Nucleic Acids Res*. 2010 Aug;38(15):4958–69.
- Peng JC, Valouev A, Swigut T, Zhang J, Zhao Y, Sidow A, et al. Jarid2/Jumonji coordinates control of PRC2 enzymatic activity and target gene occupancy in pluripotent cells. *Cell*. 2009 Dec 24;139(7):1290–302.
- Pengelly AR, Copur Ö, Jäckle H, Herzig A, Müller J. A histone mutant reproduces the phenotype caused by loss of histone-modifying factor Polycomb. *Science*. 2013 Feb 8;339(6120):698–9.
- Petruk S, Black KL, Kovermann SK, Brock HW, Mazo A. Stepwise histone modifications are mediated by multiple enzymes that rapidly associate with nascent DNA during replication. *Nat Commun*. 2013 Jan;4:2841.
- Pinello L, Xu J, Orkin SH, Yuan G-C. Analysis of chromatin-state plasticity identifies cell-type-specific regulators of H3K27me3 patterns. *Proc Natl Acad Sci U S A*. 2014 Jan 21;111(3):E344–53.
- Reynolds N, Salmon-Divon M, Dvinge H, Hynes-allen A, Balasooriya G, Leaford D, et al. NuRD-mediated deacetylation of H3K27 facilitates recruitment of Polycomb Repressive Complex 2 to direct gene repression. *EMBO J*. 2011 Nov;1–13.
- Riising EM, Comet I, Leblanc B, Wu X, Johansen JV, Helin K. Gene silencing triggers polycomb repressive complex 2 recruitment to CpG islands genome wide. *Mol Cell*. 2014 Aug 7;55(3):347–60.
- Ringrose L, Rehmsmeier M, Dura J-M, Paro R. Genome-wide prediction of Polycomb/Trithorax response elements in *Drosophila melanogaster*. *Dev Cell*. 2003 Nov;5(5):759–71.
- Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Bruggmann S a, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell*. 2007 Jun 29;129(7):1311–23.
- Da Rocha ST, Boeva V, Escamilla-Del-Arenal M, Ancelin K, Granier C, Matias NR, et al. Jarid2 Is Implicated in the Initial Xist-Induced Targeting of PRC2 to the Inactive X Chromosome. *Mol Cell*. 2014 Jan 23;53(2):301–16.
- Sanulli S, Justin N, Teissandier A, Ancelin K, Portoso M, Caron M, et al. Jarid2 Methylation via the PRC2 Complex Regulates H3K27me3 Deposition during Cell Differentiation. *Mol Cell*. 2015 Jan;
- Schmitges FW, Prusty AB, Faty M, Stützer A, Lingaraju GM, Aiwezian J, et al. Histone methylation by PRC2 is inhibited by active chromatin marks. *Mol Cell*. 2011 May 6;42(3):330–41.
- Schneider TD, Arteaga-Salas JM, Mentele E, David R, Nicetto D, Imhof A, et al. Stage-specific histone modification profiles reveal global transitions in the *Xenopus* embryonic epigenome. *PLoS One*. 2011 Jan;6(7):e22548.
- Schuettengruber B, Ganapathi M, Leblanc B, Portoso M, Jaschek R, Tolhuis B, et al. Functional anatomy of polycomb and trithorax chromatin landscapes in *Drosophila* embryos. *PLoS Biol*. 2009 Jan 13;7(1):e13.
- Schuettengruber B, Oded Elkayam N, Sexton T, Entrevan M, Stern S, Thomas A, et al. Cooperativity, specificity, and evolutionary stability of Polycomb targeting in *Drosophila*. *Cell Rep*. 2014 Oct 9;9(1):219–33.
- Schwartz YB, Kahn TG, Nix DA, Li X, Bourgon R, Biggin M, et al. Genome-wide analysis of Polycomb targets in *Drosophila melanogaster*. *Nat Genet*. 2006;38(6):700–5.
- Schwartz YB, Pirrotta V. A new world of Polycombs: unexpected partnerships and emerging functions. *Nat Rev Genet*. 2013 Dec;14(12):853–64.
- Sidoli S, Schwämmle V, Ruminowicz C, Hansen TA, Wu X, Helin K, et al. Middle-down hybrid chromatography/tandem mass spectrometry workflow for characterization of combinatorial post-translational modifications in histones. *Proteomics*. 2014 Oct;14(19):2200–11.
- Signolet J, Hendrich B. The function of chromatin modifiers in lineage commitment and cell fate specification. *FEBS J*. 2015 May;282(9).

- Simon J, Chiang A, Bender W, Shimell MJ, O'Connor M. Elements of the drosophila bithorax complex that mediate repression by polycomb group products. 1993. p. *Developmental biology* 158.
- Smits AH, Jansen PWTC, Poser I, Hyman AA, Vermeulen M. Stoichiometry of chromatin-associated protein complexes revealed by label-free quantitative mass spectrometry-based proteomics. *Nucleic Acids Res.* 2013 Jan 7;41(1):e28.
- Steffen PA, Ringrose L. What are memories made of? How Polycomb and Trithorax proteins mediate epigenetic memory. *Nat Rev Mol Cell Biol.* 2014 May;15(5):340–56.
- Stojic L, Jasencakova Z, Prezioso C, Stützer A, Bodega B, Pasini D, et al. Chromatin regulated interchange between polycomb repressive complex 2 (PRC2)-Ezh2 and PRC2-Ezh1 complexes controls myogenin activation in skeletal muscle cells. *Epigenetics Chromatin.* 2011 Jan;4:16.
- Tavares L, Dimitrova E, Oxley D, Webster J, Poot R, Demmers J, et al. RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3. *Cell.* 2012 Feb 17;148(4):664–78.
- Tee W-W, Shen SS, Oksuz O, Narendra V, Reinberg D. Erk1/2 activity promotes chromatin features and RNAPII phosphorylation at developmental promoters in mouse ESCs. *Cell.* 2014 Feb 13;156(4):678–90.
- Tie F, Banerjee R, Stratton CA, Prasad-Sinha J, Stepanik V, Zlobin A, et al. CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing. *Development.* 2009 Sep;136(18):3131–41.
- Tien C-L, Jones A, Wang H, Gerigk M, Nozell S, Chang C. Snail2/Slug cooperates with Polycomb repressive complex 2 (PRC2) to regulate neural crest development. *Development.* 2015 Jan 23;142(4):722–31.
- Tolhuis B, Wit E De, Muijters I, Teunissen H, Talhout W, Steensel B Van, et al. Genome-wide profiling of PRC1 and PRC2 Polycomb chromatin binding in Drosophila melanogaster. *Nat Genet.* 2006;38(6):694–700.
- Tsai M-C, Manor O, Wan Y, Mosammamaparast N, Wang JK, Lan F, et al. Long noncoding RNA as modular scaffold of histone modification complexes. *Science.* 2010 Aug 6;329(5992):689–93.
- Turner SA, Bracken AP. A “complex” issue: deciphering the role of variant PRC1 in ESCs. *Cell Stem Cell.* 2013 Feb 7;12(2):145–6.
- Vizán P, Beringer M, Ballaré C, Di Croce L. Role of PRC2-associated factors in stem cells and disease. *FEBS J.* 2015 May;282(9):1723–35.
- Voigt P, LeRoy G, Drury WJ, Zee BM, Son J, Beck DB, et al. Asymmetrically modified nucleosomes. *Cell.* 2012 Sep 28;151(1):181–93.
- Wachter E, Quante T, Merusi C, Arczewska A, Stewart F, Webb S, et al. Synthetic CpG islands reveal DNA sequence determinants of chromatin structure. *Elife.* 2014 Sep 26;3:e03397.
- Wang L, Brown JL, Cao R, Zhang Y, Kassis JA, Jones RS. Hierarchical recruitment of polycomb group silencing complexes. *Mol Cell.* 2004 Jun 4;14(5):637–46.
- Xu C, Bian C, Yang W, Galka M, Ouyang H, Chen C, et al. Binding of different histone marks differentially regulates the activity and specificity of polycomb repressive complex 2 (PRC2). *PNAS.* 2010;107(45):19266–71.
- Xu J, Shao Z, Li D, Xie H, Kim W, Huang J, et al. Developmental Control of Polycomb Subunit Composition by GATA Factors Mediates a Switch to Non-Canonical Functions. *Mol Cell.* 2015 Jan.
- Yu M, Mazor T, Huang H, Huang H-T, Kathrein KL, Woo AJ, et al. Direct recruitment of polycomb repressive complex 1 to chromatin by core binding transcription factors. *Mol Cell.* 2012 Feb 10;45(3):330–43.
- Yuan W, Wu T, Fu H, Dai C, Wu H, Liu N, et al. Dense chromatin activates Polycomb repressive complex 2 to regulate H3 lysine 27 methylation. *Science.* 2012 Aug 24;337(6097):971–5.
- Yuan W, Xu M, Huang C, Liu N, Chen S, Zhu B. H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *J Biol Chem.* 2011 Mar 11;286(10):7983–9.
- Zhao J, Ohsumi TK, Kung JT, Ogawa Y, Grau DJ, Sarma K, et al. Genome-wide identification of polycomb-associated RNAs by RIP-seq. *Mol Cell.* 2010 Dec 22;40(6):939–53.
- Zhao J, Sun BK, Erwin J a, Song J-J, Lee JT. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science.* 2008 Oct 31;322:750–6.

CHAPTER 3

Embryonic transcription is controlled by maternally defined chromatin state

Saartje Hontelez*, Ila van Kruijsbergen*, Georgios Georgiou*, Simon J. van Heeringen, Ozren Bogdanovic, Ryan Lister, Gert Jan C. Veenstra

*These authors contributed equally to this work

Nature Communications (2015) 6:10148

Experiments were designed by SH, IvK and GJCV. ChIP-seq data production was done by SH and IvK. Bisulfite sequencing was done by OB and RL. GG performed ChrommHMM analysis. All other analyses were done by SH, with support from GG and SJvH. SH wrote the paper. SH, IvK and GG contributed equally to the study. All authors discussed the results and commented on the manuscript.

Abstract

Histone modifying enzymes are required for cell identity and lineage commitment, however little is known about the regulatory origins of the epigenome during embryonic development. Here we generate a comprehensive set of epigenome reference maps, which we use to determine the extent to which maternal factors shape chromatin state in Xenopus embryos. Using α -amanitin to inhibit zygotic transcription, we find that the majority of H3K4me3 and H3K27me3-enriched regions form a maternally defined epigenetic regulatory space with an underlying logic of hypomethylated islands. This maternal regulatory space extends to a substantial proportion of neurula stage-activated promoters. In contrast, p300-recruitment to distal regulatory regions requires embryonic transcription at most loci. The results show that H3K4me3 and H3K27me3 are part of a regulatory space that exerts an extended maternal control well into post-gastrulation development, and highlight the combinatorial action of maternal and zygotic factors through proximal and distal regulatory sequences.

Introduction

During early embryonic development cells differentiate, acquiring specific transcription and protein expression profiles. Histone modifications can control the activity of genes through regulatory elements in a cell-type specific fashion^{1, 2, 3, 4}. Recent advances have been made in the annotation of functional genomic elements of mammalian cells, *Drosophila* and *Caenorhabditis* through genome-wide profiling of chromatin marks^{5, 6}. Immediately after fertilization, the embryonic genome is transcriptionally silent, and zygotic genome activation (ZGA) occurs after a number of mitotic cycles⁷. In *Drosophila* and zebrafish (*Danio rerio*) ZGA starts after 8 and 9 mitotic cycles, respectively, in mammals transcription starts at the two-cell stage^{8, 9}, whereas in *Xenopus* this happens after the first 12 cleavages at the mid-blastula transition (MBT)^{10, 11, 12}. Permissive H3K4me3 and repressive H3K27me3 histone modifications emerge during blastula and gastrula stages^{13, 14, 15, 16}. To date, little is known about the origin and specification of the epigenome in embryonic development of vertebrates, which is essential for understanding physiological cell lineage commitment and differentiation.

To explore the developmental origins of epigenetic regulation we have generated epigenome reference maps during early development of *Xenopus tropicalis* embryos and assessed the need for embryonic transcription in their acquisition. We find a hierarchical appearance of histone modifications, with

a priority for promoter marks which are deposited hours before transcription activation on regions with hypomethylated DNA. Surprisingly, the promoter H3K4me3 and the Polycomb H3K27me3 modifications are largely maternally defined (MaD), providing maternal epigenetic control of gene activation that extends well into neurula and tailbud stages. By contrast, p300 recruitment to distal regulatory elements is largely under the control of zygotic factors. Moreover, this maternal-proximal and zygotic-distal dichotomy of gene regulatory sequences also differentiates between early and late Wnt signalling target genes, suggesting that different levels of permissiveness are involved in temporal target gene selection.

Progressive specification of chromatin state

We have performed ChIP-sequencing of eight histone modifications, RNA polymerase II (RNAPII) and the enhancer protein p300 at five stages of development: blastula (st. 9), gastrula (st. 10.5, 12.5), neurula (st. 16) and tailbud (st. 30). These experiments allow identification of enhancers (H3K4me1, p300)^{17, 18, 19, 20}, promoters (H3K4me3, H3K9ac)^{14, 21, 22, 23}, transcribed regions (H3K36me3, RNAPII)²² and repressed and heterochromatic domains (H3K27me3, H3K9me2, H3K9me3, H4K20me3)^{1, 14, 24, 25}. In addition we generated pre-MBT (st. 8) maps for three histone modifications (H3K4me3, H3K9ac, H3K27me3) and single-base resolution DNA methylome maps using whole genome bisulfite sequencing of blastula and gastrula (st. 9 and 10.5) embryos (Fig. 1, Supplementary Fig. 1). Our data set consists of 2.7 billion aligned sequence reads representing the most comprehensive set of epigenome reference maps of vertebrate embryos to date. Using a Hidden Markov Model approach²⁶ we have identified 19 chromatin states based on co-occurring ChIP signals (Fig. 2a). This analysis identifies combinations of ChIP signals at specific genomic sequences without distinguishing between overlapping histone modifications that result from regional or cell type specificity and co-occurrence in the same cells¹⁴. Seven main groups were recognized, namely (i) Polycomb (H3K27me3, deposited by PRC2), (ii) poised enhancers, (iii) p300-bound enhancers, (iv) transcribed regions, (v) promoters, (vi) heterochromatin and (vii) unmodified regions (Fig. 2a, Supplementary Fig. 2). Alluvial plots of state coverage per stage show that all states increase in coverage during development, except for the unmodified state (Fig. 2b, Supplementary Fig. 2a). Unmodified regions decrease in coverage during development, however, even at tailbud stage 67% of the total epigenome remains naive for the modifications and bound proteins in our data set (Supplementary Fig. 2b). Promoter coverage remains relatively constant

during development from blastula to tailbud stages, in contrast to the Polycomb state which increases in coverage during gastrulation. P300-bound enhancers are highly dynamic during development (Fig. 2b). Global enrichment levels of modified regions show similar dynamics, and reveal a priority for promoter marking at or before the blastula stage, followed by enhancer activation and heterochromatic repression during late blastula and gastrulation stages (Supplementary Fig. 3a-b). A detailed time course between fertilization and early gastrulation shows that both H3K4me3 and H3K9ac emerge hours before the start of embryonic transcription (Supplementary Fig. 3c). We and others have previously reported that H3K4me3 is acquired during blastula stages¹⁴. Indeed, H3K4me3 and H3K9ac levels increase strongly before the MBT, well before embryonic transcription starts. This however raises the question to what extent histone modifications are regulated by maternal or embryonic factors.

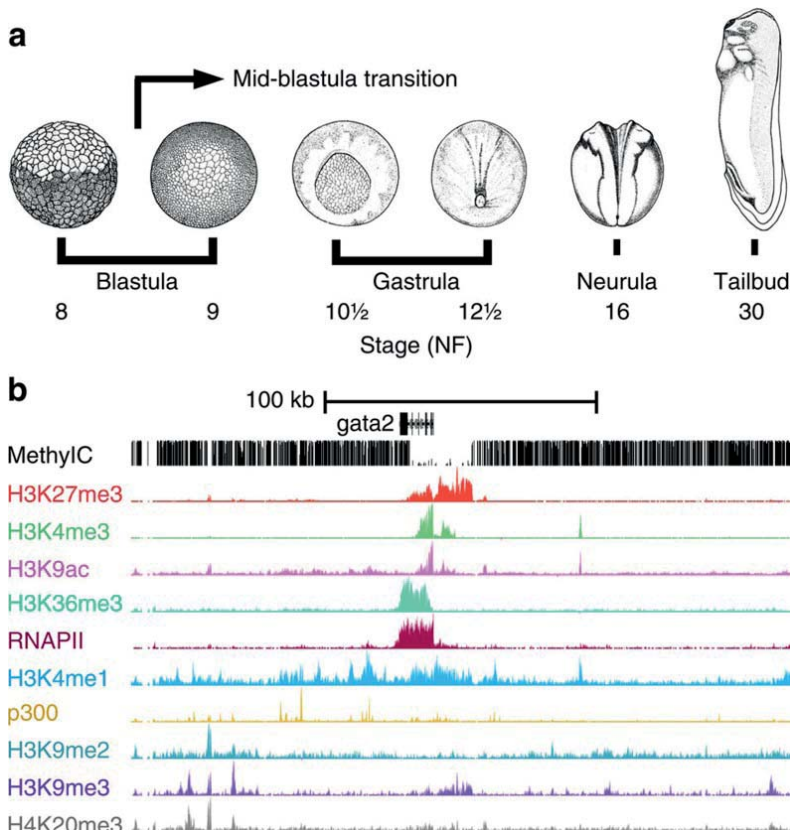


Figure 1. Reference epigenome maps of *Xenopus tropicalis* development. (a) Genome-wide profiles were generated for stages 8 and 9 (blastula, before and after MBT), 10.5 and 12.5 (gastrula), 16 (neurula) and 30 (tailbud). Adapted from Tan, M.H. et al. *Genome Res.* 23, 201–216 (2013), under a Creative Commons License (Attribution-NonCommercial 3.0 Unported License), as described at <http://creativecommons.org/licenses/by/3.0/>. (b) *Gata2* locus with late gastrula (stage 10.5) methylC-seq, ChIP-seq enrichment of histone modifications, RNAPII and p300 (cf. Supplementary Fig. 1, 2).

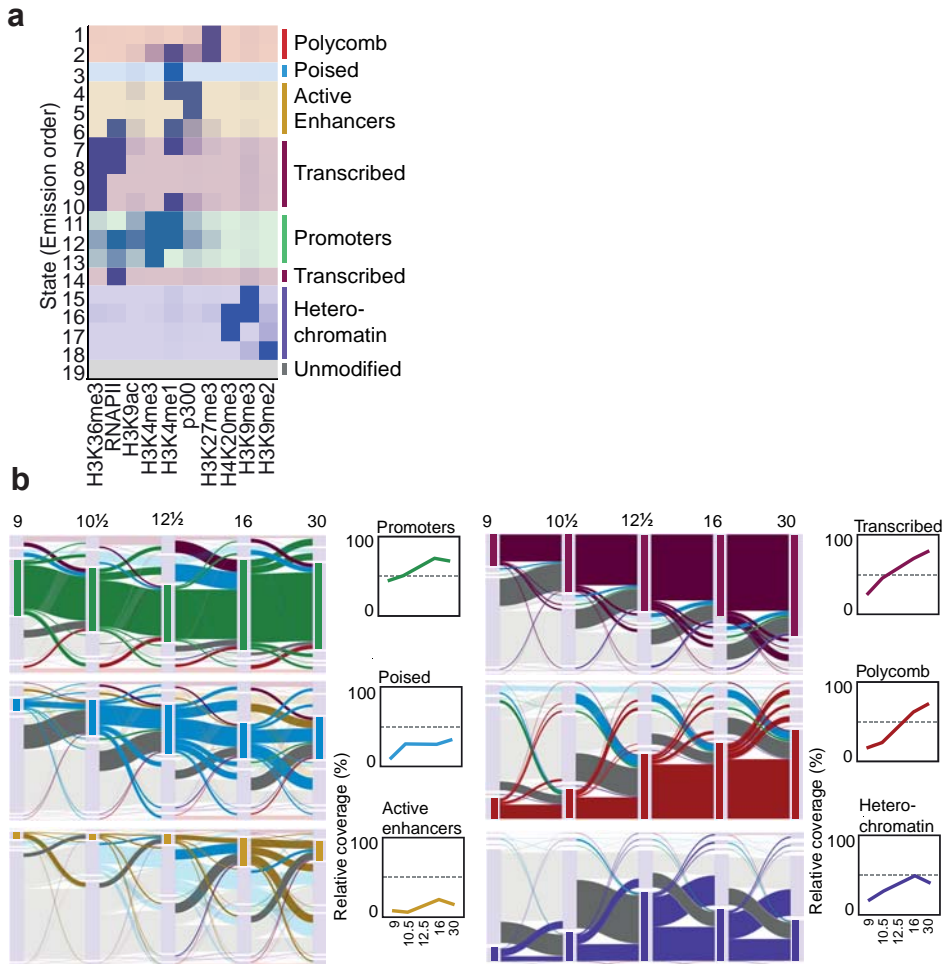


Figure 2. Chromatin state dynamics. (a) Emission states (same for all developmental stages) of the hidden Markov model, identifying the 19 most prevalent combinations of histone modifications and bound proteins. From top to bottom: Polycomb (red), Poised enhancers and promoters (blue), Active Enhancers (gold), Transcribed (dark magenta), Promoter (green), Heterochromatin (purple) and unmodified (grey). (b) Alluvial plots of chromatin state coverage during development. Each plot shows the transitions (to and from the highlighted group of chromatin states) across developmental stages (st. 9-30). The height represents the base pair coverage of the chromatin state relative to the modified genome. The “modified genome” has a chromatin state other than unmodified in any of the stages 9-30. From top to bottom left: promoters (green), poised (blue), p300-bound enhancers (gold). From top to bottom right: transcribed (dark magenta), Polycomb (red) and heterochromatin (purple). Line plots: Chromatin state coverage per stage as a percentage of the modified genome.

Maternal and zygotic epigenetic regulation

To determine the maternal and zygotic contributions to chromatin state, we used α -amanitin to block embryonic transcription (Fig. 3a). Alpha-amanitin blocks the translocation of RNA polymerase II (RNAPII) on DNA, thereby preventing transcript elongation²⁷. It is therefore expected that injection of α -amanitin

into embryos will stall RNAPII, immobilizing it on DNA after its recruitment to pre-initiation complexes. Indeed, both RNAPII elongation and embryonic transcription were effectively blocked in α -amanitin-injected embryos (Fig. 3b-c, Supplementary Fig. 4a). New transcription is necessary for gastrulation^{11, 28, 29}, but α -amanitin injected embryos survive to the equivalent of stage 11 control embryos. ChIP-sequencing of replicates of α -amanitin-injected and control embryos (stage 11) revealed that the majority of H3K4me3 (86%) and H3K27me3 (90%) regions are consistently modified with these modifications independently of embryonic transcription (Fig. 3d, Supplementary Fig. 4b-c). This is especially surprising given the temporal hierarchy of H3K27me3 and H3K4me3, and the relatively late acquisition of H3K27me3 (Fig. 2b). By contrast, only 15% of the p300-bound regions recruit p300 independently of active transcription (Fig. 3d). This suggests that the promoter-permissive H3K4me3 mark and the Polycomb-repressive H3K27me3 mark are mostly controlled by maternal factors (MaD, maternally defined), whereas p300 binding to regulatory regions is largely zygotically defined (ZyD). Regions with MaD H3K4me3 and H3K27me3 acquire these modifications more robustly and also earlier during development

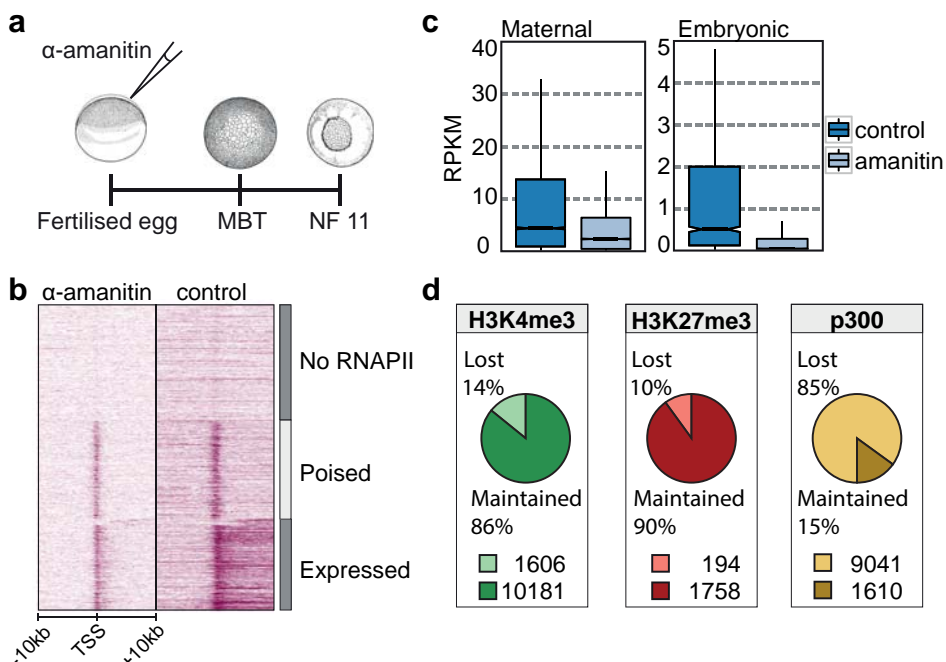


Figure 3. Developmental acquisition of chromatin states. (a) Inhibition of embryonic transcription with α -amanitin (b) RNAPII on the TSS of genes in control and α -amanitin injected embryos (stage 11). (c) Box plots showing RNA expression levels (RPKM) of maternal and embryonic transcribed genes in control and α -amanitin injected embryos (stage 11). Box: 25th (bottom), 50th (internal band), 75th (top) percentiles. Whiskers: 1.5 * interquartile range of the lower and upper quartiles, respectively. (d) ChIP-sequencing on chromatin of α -amanitin injected and control embryos reveals maternal and zygotic origins of H3K4me3, H3K27me3 or p300 binding. Data from two biological replicates, see Supplementary Fig. 4.

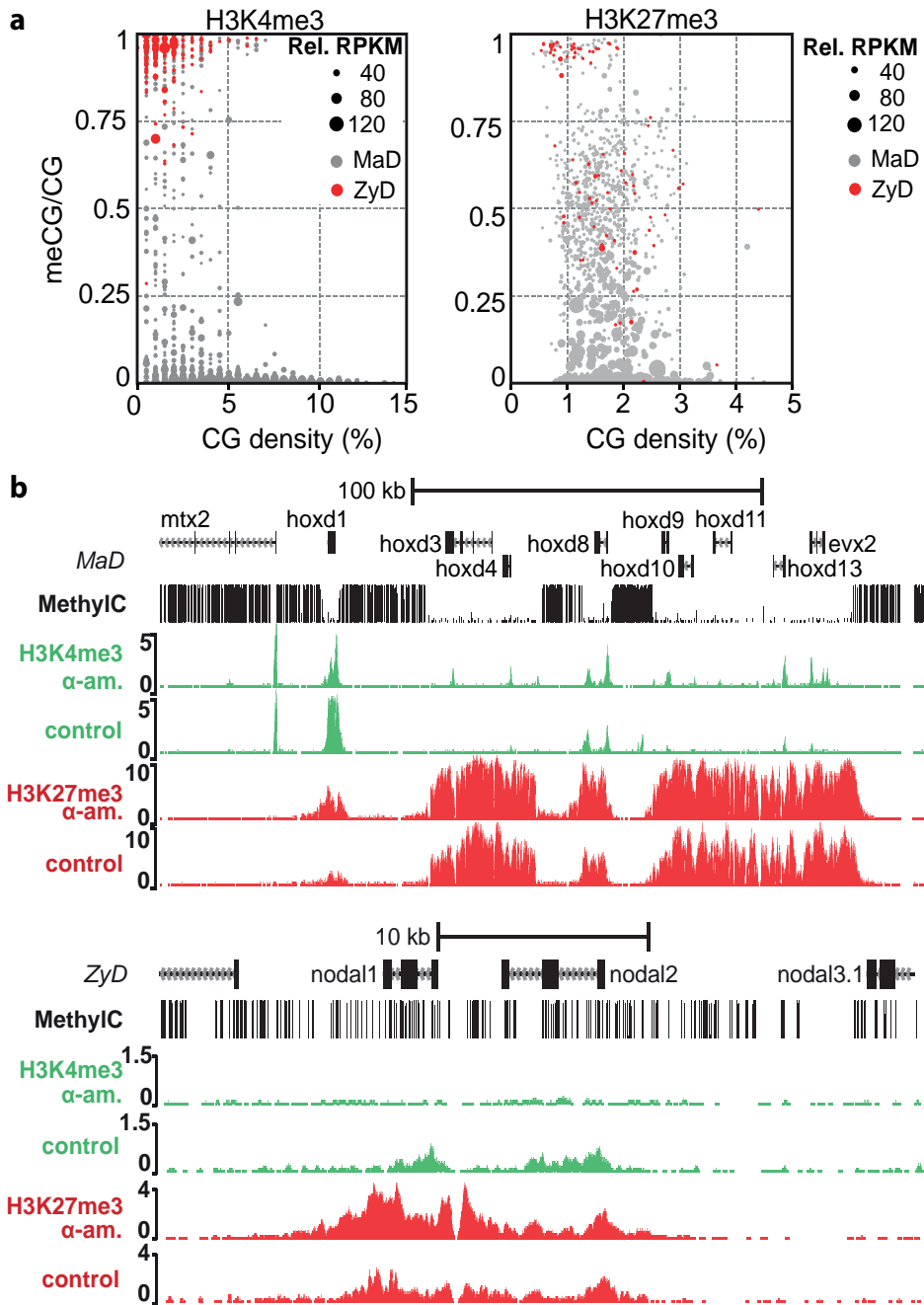
compared to ZyD regions (Supplementary Fig. 4d). By contrast, ZyD p300-bound regions show more robust p300 recruitment during gastrulation compared to p300 MaD regions. These data show a pervasive maternal influence on the developmental acquisition of key histone modifications.

DNA methylation logic of maternal control

Trimethylation of H3K4 and H3K27 has been associated with CpG density and a lack of DNA methylation. The Set1 and related MLL complexes are responsible for H3K4me3. Set1 is recruited to hypomethylated CpG domains via the Cxxc1 protein (Cfp1)^{30,31,32}. In the absence of H3K4me3, PRC2 binding to hypomethylated CpGs results in H3K27me3 and inhibition of gene activation^{13, 33}. Using our whole genome bisulfite sequencing data we determined that MaD H3K4me3 promoters are predominantly hypomethylated (Fig. 4a, Supplementary Fig. 5a, Supplementary Data 1). Conversely, promoters decorated with ZyD H3K4me3 almost exclusively have highly methylated promoters. Demethylation of ZyD promoters was not detected, and methylation levels of MaD and ZyD regions were similar in stage 9 and stage 10.5 (Supplementary Fig. 5a-b). In addition, H3K4me3 often extends asymmetrically from promoters into gene bodies (+1-2 kb from transcription start site (TSS)) (Supplementary Fig. 5c), likely representing the second and third nucleosomes that are trimethylated via RNAPII-recruited Set1 in actively transcribed genes³⁴. Concordantly, α -amanitin reduces H3K4me3 at downstream positions. Interestingly, we also find poised enhancers that gain H3K4me3 in α -amanitin injected embryos and which exhibit intermediate to high levels of DNA methylation (Supplementary Fig. 5d-e).

The majority of promoters with ZyD H3K27me3 shows intermediate to high levels of DNA methylation (Fig. 4a, Supplementary Fig. 5a, Supplementary Data 1). Some of the MaD H3K27me3 regions are methylated, but the highly enriched H3K27me3 domains (larger dots) are almost exclusively both maternally defined and hypomethylated. This is illustrated by the *hoxd* cluster which harbours a large hypomethylated domain with MaD H3K4me3 and H3K27me3 (Fig. 4b). There are also examples of reciprocal changes of H3K4 and H3K27 methylation, for example at the hypermethylated promoters of *nodal1* and *nodal2*.

ZyD p300-bound regions are generally hypermethylated, whereas MaD p300-bound regions show a variable degree of DNA methylation (Supplementary Fig. 5e). However, promoters that overlap with MaD p300 peaks are hypomethylated in 77% of the cases, whereas 96% of the promoters that are associated with ZyD p300 peaks are hypermethylated (Supplementary Fig. 5f), showing that p300-



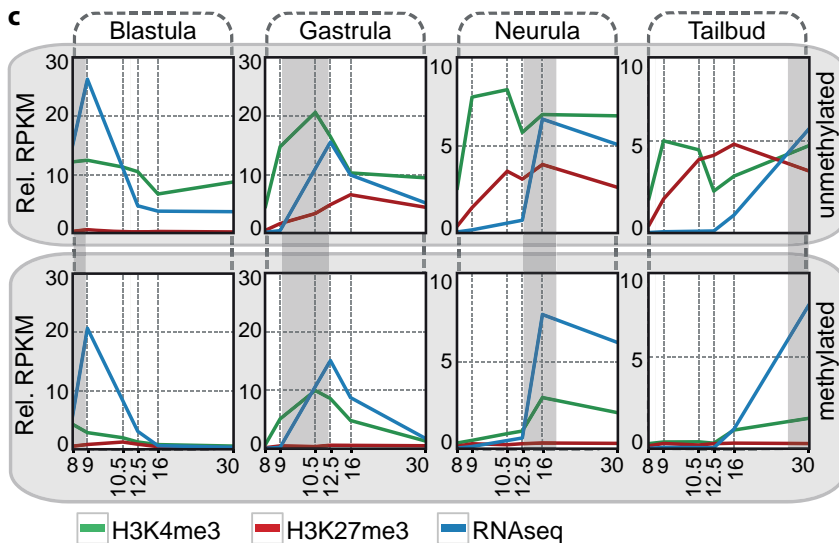


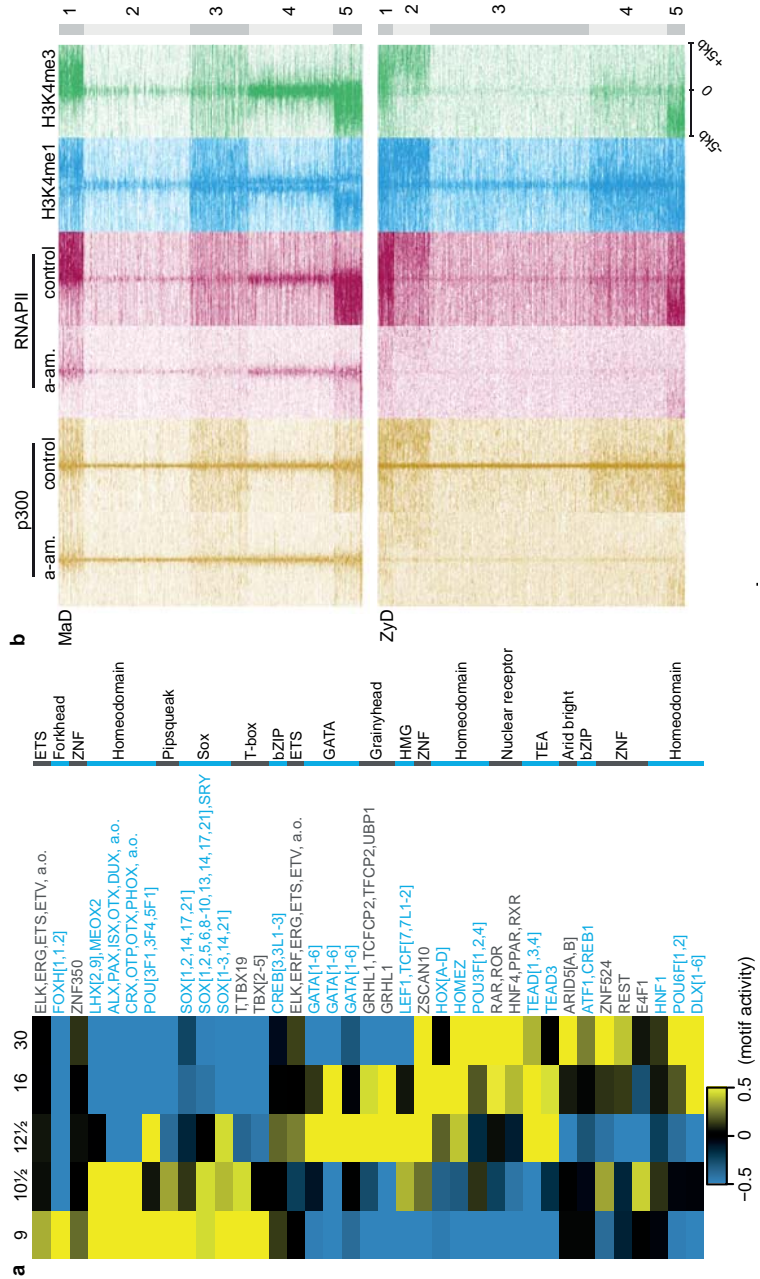
Figure 4. DNA methylation logic of maternally versus zygotically defined H3K4me3 and H3K27me3. (a) CpG density and methylation at stage 9 of promoters (H3K4me3: + 100 bp from TSS; H3K27me3: + 2.5 kb from TSS) that contain a zygotic defined (ZyD, lost in α -amanitin treated embryos, red) or maternal defined (MaD, maintained in α -amanitin treated embryos, grey) peak for H3K4me3 (left) or H3K27me3 (right) after inhibition of embryonic transcription. The size of the dot indicates the relative RPKM of the histone modification (background corrected). (b) *Hoxd* (MaD) and *nodal1, -2* (ZyD) loci with stage 9 methylC-seq, H3K4me3 and H3K27me3 in control and α -amanitin injected embryos. (c) Developmental profiles of H3K4me3 and H3K27me3 (median background corrected RPKM) at genes without detectable maternal mRNA do correlate with activation for methylated promoters (lower panels) but not for hypomethylated CpG island promoters (upper panels).

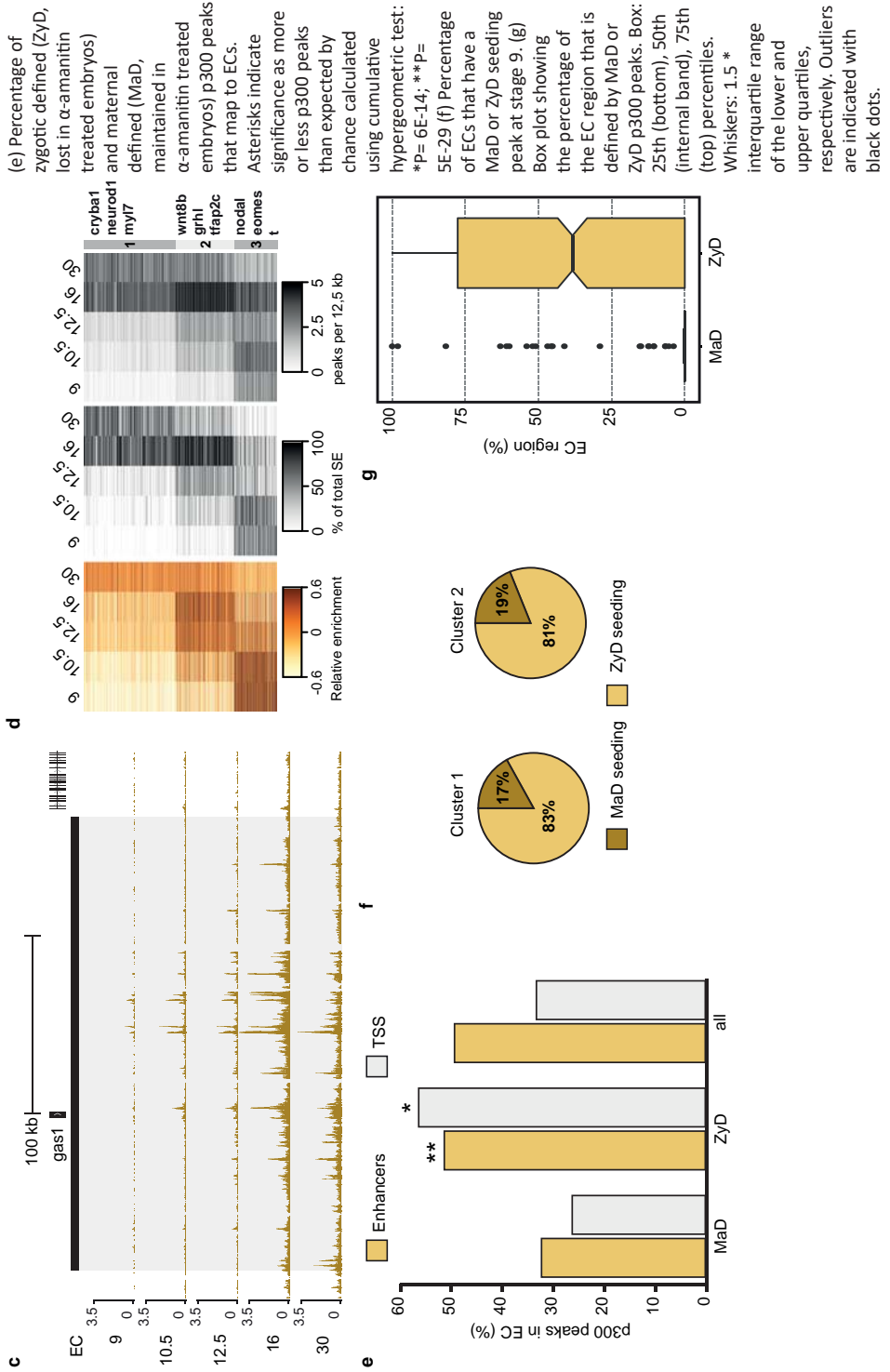
recruiting hypomethylated promoters tend to be under complete maternal control, for both H3K4 methylation and p300 recruitment.

To further explore the relationships between DNA methylation, histone modifications and developmental activation of transcription we determined correlations with different measures of gene activity such as RNA-seq and ChIP-seq of RNAPII and H3K36me3 (Supplementary Fig. 6). We find that H3K36me3 and RNAPII in gene bodies correlate well with each other but less with transcript levels (RNA-seq), presumably due to the effects of RNA stability. A much lower correlation was found between either measure of gene activity and the promoter marks H3K4me3 and H3K9ac, especially at early stages. In part this may be caused by time delays of transcriptional activation relative to acquisition of permissive histone modifications^{14, 15}. It raises the question to what extent a lack of DNA methylation at promoters, which is associated with MaD H3K4me3, uncouples promoter marking and transcriptional activation. Therefore, we grouped transcribed genes without detectable maternal mRNA³⁵ based on the stage of maximum expression and DNA methylation (Fig. 4c). We find that developmentally activated promoters with hypomethylated CpG-

Figure 5. Zygotically recruited p300 enhancer clusters (EC) domains.

(a) Modelled transcription factor motif activity to p300 enrichment (see Methods). Activity reflects modelled contributions in p300 peak RPKM. (b) Heatmaps of Mad (upper panel) and Zyd (lower panel) p300 binding sites in α -amanitin treated embryos. (c) Developmental increase in genomic coverage of the gas1 EC by acquisition of p300 binding at enhancers. (d) EC dynamics of p300 enrichment (left panel), percentage of total EC region identified in each stage based on stage-dependent p300 binding (middle panel) and number of p300 peaks (per 12,5 kb) in EC.





islands are trimethylated at H3K4 or H3K27 early on, irrespective of the time of transcriptional activation. By contrast, methylated promoters show a much closer relation between H3K4me3 and gene expression. Although H3K4me3 is known to stabilize the transcription initiation factor Taf3 (a subunit of TFIID) and can also interact with the chromatin remodeller Chd1^{36, 37, 38}, hypomethylated promoters gain H3K4me3 autonomously with their hypomethylated CpG island status, independent of embryonic transcription.

ZyD p300-bound domains shape enhancer clusters

P300 can be recruited by transcription factors that bind to regulatory elements. We therefore modelled transcription factor motif contributions to p300 binding across multiple developmental stages (see Methods). The results predict specific transcription factors to recruit p300 in a stage-specific fashion (Fig. 5a). Clustering of MaD and ZyD p300-bound regions with H3K4me3, H3K4me1 and RNAPII data revealed that ZyD p300 is recruited to distal regulatory sequences that lose both p300 and RNAPII binding in the presence of α -amanitin, whereas MaD p300-binding mostly includes promoter-proximal regions that are H3K4me3-decorated and recruit RNAPII in the presence of α -amanitin but without elongating (Fig. 5b). Indeed, MaD p300 regions are enriched for promoter-related motifs (Supplementary Fig. 7). Although some ZyD p300-bound regions overlap with annotated transcription start sites (Supplementary Fig. 5f), most of these sequences are decorated with H3K4me1 in the absence of H3K4me3, suggesting they correspond to distal regulatory sequences (Fig. 5b). Both MaD and ZyD p300-bound regulatory regions recruit embryonically regulated transcription factors such as Otx2, Gsc, Smad2/3, Foxh1, T (Xbra), Vegt and Eomes (Supplementary Fig. 8)^{39, 40, 41}, suggesting that multiple transcription factors contribute to p300 recruitment.

Large enhancer clusters (ECs) are thought to improve the stability of enhancer-promoter interactions, are associated with genes coding for developmental regulators, and have been implicated in cell differentiation^{42, 43, 44}. During development the cluster size of p300-bound enhancers grows dynamically by p300-seeding of individual enhancers (Fig. 5c-d, see Methods). Histone modifications and transcript levels of EC-associated genes are developmental stage-specific, confirming the association of ECs with developmental genes (Supplementary Fig. 9, Supplementary Data 2). Analysis of the percentage of the total EC regions identified in each stage show that most p300-bound ECs increase in genomic coverage during development by newly gained p300 binding

at enhancers (EC clusters 1 and 2), whereas a group of early ECs (EC cluster 3) decrease in coverage as a result of the decreasing number of p300 peaks that contribute to the EC.

We next examined how MaD and ZyD p300-bound regions contribute to p300-bound ECs. Approximately 50% of all ZyD p300-bound enhancers are located in ECs at stage 11. Among MaD p300-bound enhancers this fraction is much reduced (Fig. 5e). Similarly, a much larger fraction of ZyD p300-bound promoters is found in ECs compared to MaD p300-bound promoters. Up to 20% of the developmental ECs that are seeded at stage 9 have a MaD p300 seeding site (Fig. 5f). However, very few ECs can be called based on MaD p300, showing that formation of p300-bound enhancer clusters requires embryonic transcription (Fig. 5g).

Extended maternal epigenetic control

We next examined the extent to which the MaD epigenome is maintained during development. Genes were grouped based on MaD or ZyD trimethylation of H3K4 and H3K27 in the promoter (Supplementary Data 3, see Methods). For p300 we counted the total number of MaD and ZyD peaks in the cis-regulatory landscapes of genes (Fig. 6a). Remarkably, MaD H3K4me3-regulated genes represent the majority of all H3K4me3-enriched genes in both early and late developmental stages. Even at neurula and tailbud stages only a small fraction of the H3K4me3-decorated genes are ZyD. Similarly, maternal control of H3K27me3 also extends late into development, albeit to a smaller degree. After gastrulation, the number of MaD H3K27me3 regulated genes slightly decreases, whereas ZyD increases. However, also at neurula stage more than 50% of the Polycomb (PRC2) -regulated genes are under MaD H3K27me3 control. By contrast, p300 in cis-regulatory regions of genes is almost exclusively ZyD in all stages (Fig. 6a).

Many genes may maintain MaD H3K4me3 because they are constitutively expressed throughout development. We therefore analysed the regulation of genes that are exclusively embryonically transcribed. We find that 487 of 983 (49,5%) genes which are expressed between blastula and tailbud stages but not expressed in oocytes or before the MBT, feature a MaD H3K4me3 promoter (Supplementary Fig. 10a). Most of the MaD H3K4me3 genes that are modified by PRC2 exhibit MaD H3K27me3. When separating embryonic transcripts based on developmental activation, we find MaD H3K4me3 for 58% of the gastrula genes and up to 74% of the neurula expressed genes (Fig. 6b, Supplementary Fig. 10b). In most cases MaD H3K4me3-regulated genes also have MaD H3K27me3 control.

This indicates an important role for the MaD epigenome in the regulation of embryonic transcripts.

To explore the distinctions between expression inside and outside the maternal regulatory space, we analysed Wnt signalling targets. Early Wnt/beta-catenin signalling serves to specify dorsal fates following fertilization, leading to organizer gene expression. This has been shown to depend on Prmt2-mediated promoter poising before the MBT⁴⁵. Indeed, we find that seven of eight early

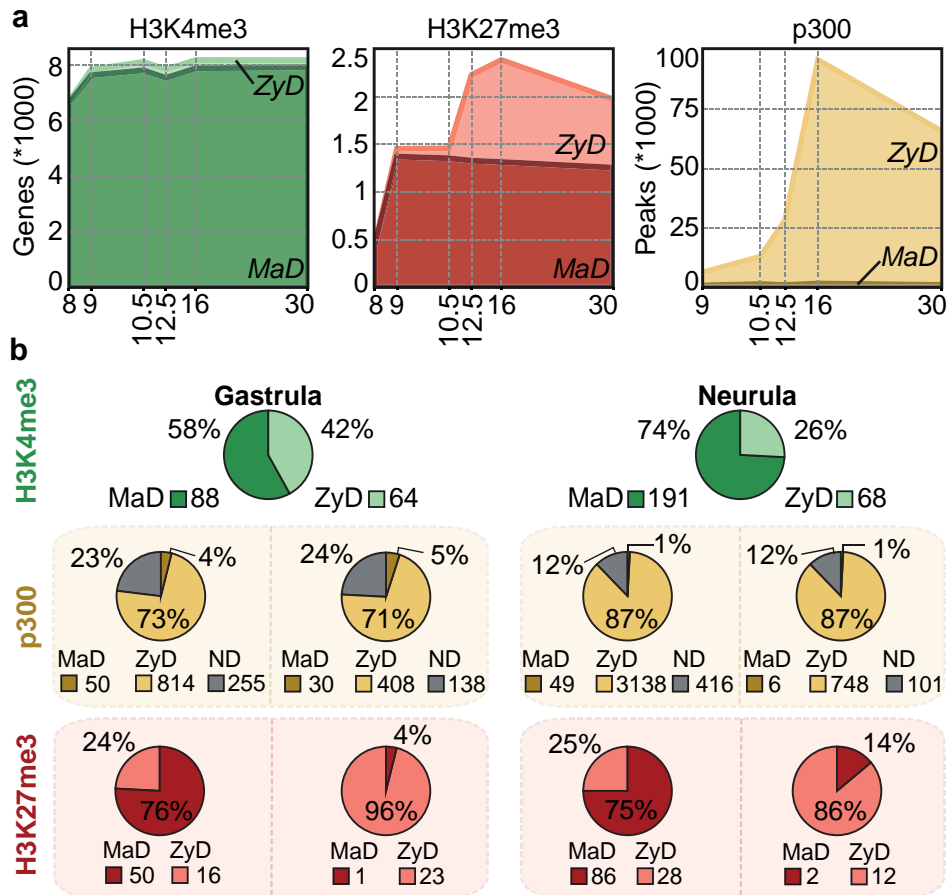


Figure 6. Maternal epigenetic control extends beyond gastrulation. Maternally defined (MaD) peaks emerge at or before stage 11 independent of embryonic transcription. Zygotically defined (ZyD) peaks appear before stage 11 and are lost in α -amanitin treated embryos, or emerge at or after stage 12. Not determined (ND) peaks are not consistently detected in replicate 1 and 2 and generally have low enrichment values. (a) Total number of genes with a MaD or ZyD peak in their promoter (H3K4me3 and H3K27me3), or total number of MaD and ZyD peaks per GREAT region (p300). ND peaks are not shown. (b) MaD and ZyD regulation of gastrula and neurula expressed genes. The pie charts show the number genes with a MaD or ZyD peak in their promoter (H3K4me3 and H3K27me3) or the number of MaD, ZyD and ND peaks per cis-regulatory region (p300). The H3K27me3 and p300 pie charts represent: Gastrula expressed genes with a MaD (far left) or ZyD (middle left) H3K4me3 peak; neurula expressed genes with a MaD (middle right) or ZyD (far right) H3K4me3 peak.

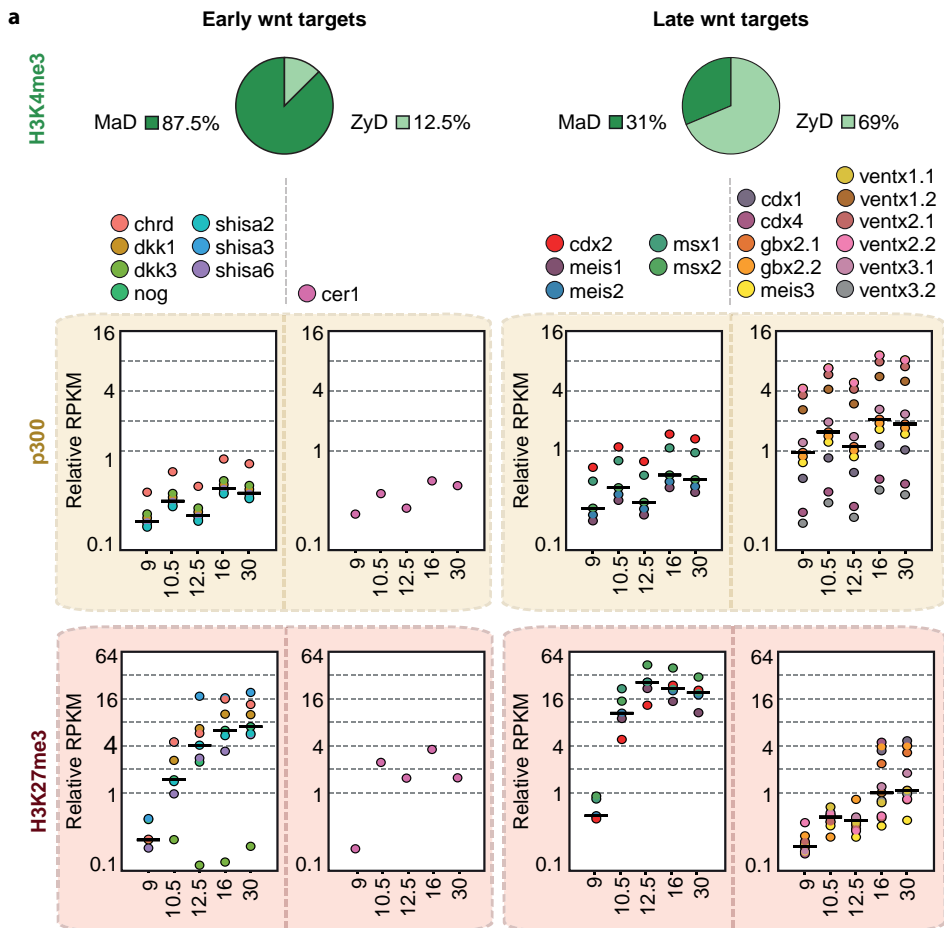
Wnt/beta-catenin targets have a hypomethylated island promoter marked with MaD H3K4me3 (Fig. 7a, Supplementary Fig. 10c). Wnt signalling also plays an important role after the MBT, when it ventralises and patterns mesoderm. The majority of these later targets turn out to have a methylated promoter with ZyD H3K4me3. Notably, these ZyD H3K4me3 late Wnt targets are associated with high binding of p300 in their locus; many of the p300 binding events happen at distal regulatory regions. In contrast, MaD H3K4me3 Wnt targets have less p300 binding but are marked with H3K27me3 (Fig. 7a-b). These results illustrate the dichotomy in proximal and distal regulation that is associated with transcriptional activation of maternal and zygotic Wnt target genes, which is paradigmatic of the distinctive maternal and zygotic epigenetic programs that are orchestrated by DNA methylation and exert a long-lasting influence in development (Fig. 8).

Discussion

The H3K4me3 modification poises promoters for transcription initiation by stabilizing Taf3/TFIID binding^{36, 37}. Promoter H3K4 methylation based on an underlying DNA methylation logic driven by maternal factors at the blastula stage sets the stage for a default program of gene expression. Most constitutively expressed house-keeping genes are within this maternal regulatory space, as well a subset of developmentally regulated genes. Remarkably, many late expressed genes have hypomethylated promoters and are already poised for activation by H3K4me3 during early blastula stages. H3K4me3 is not sufficient for gene transcription and additional embryonic factors are required for activation in many cases. Genes with MaD H3K4me3 generally have fewer p300-bound enhancers associated with them, suggesting they are regulated by promoter-proximal elements. This further underscores the permissive nature of this regulation, as opposed to zygotically regulated events at both promoters (H3K4me3) and enhancers (recruitment of p300). The H3K27me3 modification is gradually acquired between blastula and gastrula stages on spatially regulated genes, repressing lineage-specific genes in other lineages^{13, 14}. The acquisition of this modification in the absence of transcription indicates that it is uncoupled from the inductive events of the early embryo, suggesting a default maternal response to a lack of transcriptional activation. The results indicate that maternal factors set permissions and time-dependent constraints on a subset of genes with reduced CpG methylation at their promoter. These permissions and constraints are likely to channel embryonic cell fates into a limited number of directions by controlling hierarchical developmental progression by master regulators. Previously we observed that DNA methylation does not lead to transcriptional

repression in early embryos, whereas it does in oocytes and late embryos⁴⁶. The observations described here suggest a new role of DNA methylation in defining a maternal-embryonic program of gene expression. In zebrafish, the maternal methylome is reprogrammed between fertilization and ZGA, to match the paternal methylome. This also occurs in maternal-haploid fish, and appears to align with CG content^{47, 48}, suggesting an intrinsic maternal mechanism that sets the stage for the MaD epigenome.

Gene expression outside maternal regulatory space could be mediated by p300-associated enhancers, most of which require new transcription for recruitment of p300. Promoter and enhancer activation in the ZyD regulatory space likely involves binding of specific factors. Indeed, we find that both MaD and ZyD p300-bound regulatory regions recruit embryonically regulated transcription factors. Enhancers often contain binding sites for many different



proteins, which can play different roles in opening up chromatin, recruitment of co-activators and establishing looping interactions with promoters. Future experiments will shed light on the maternal-zygotic hierarchy and the regulatory transitions underlying these events and the roles of maternal and zygotic pioneer factors. We find that ZyD p300-bound enhancers shape enhancer clusters. These form dense hubs of regulatory activity, and EC p300 binding is generally correlated with the expression of the associated genes. The work reported here suggests that recruitment of p300 to “seeding” enhancers precedes establishing cluster-wide activity of the local enhancer landscape. Future work will also need to address to which extent seeding causes relaxation and opening of the local chromatin and activity of neighbouring enhancers.

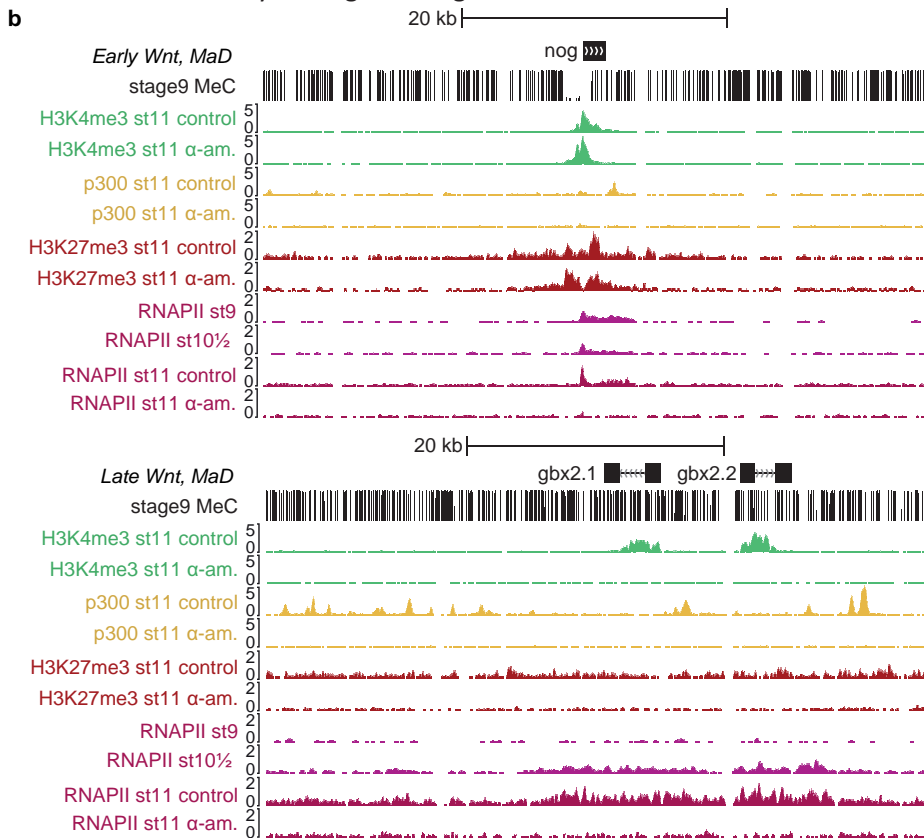


Figure 7 Maternal and Zygotic regulatory space separates early and late Wnt target genes. (a) The number of genes with MaD or ZyD H3K4me3 (pie charts) and relative RPKM (dot plots, horizontal line: median) of p300 in cis-regulatory regions of genes and H3K27me3 on promoters (+ 2.5 kb from TSS) at different developmental stages that have maternally or zygotically defined H3K4me3 at the promoter. Early targets *sia1* and *sia2* are not included, these genes lose H3K4me3 after stage 9 and cannot be assigned to MaD or ZyD space based on our stage 11 α -amanitin data. H3K4me3 on these genes is acquired at stage 8, before embryonic transcription. (b) Browser views of the early Wnt target *nog* (*noggin*) and the late Wnt targets *gbx2.1* and *gbx2.2* with ChIP-seq enrichment of H3K4me3, p300 and RNAPII on control and α -amanitin injected embryos and RNAPII on stage 9 and 10.5.

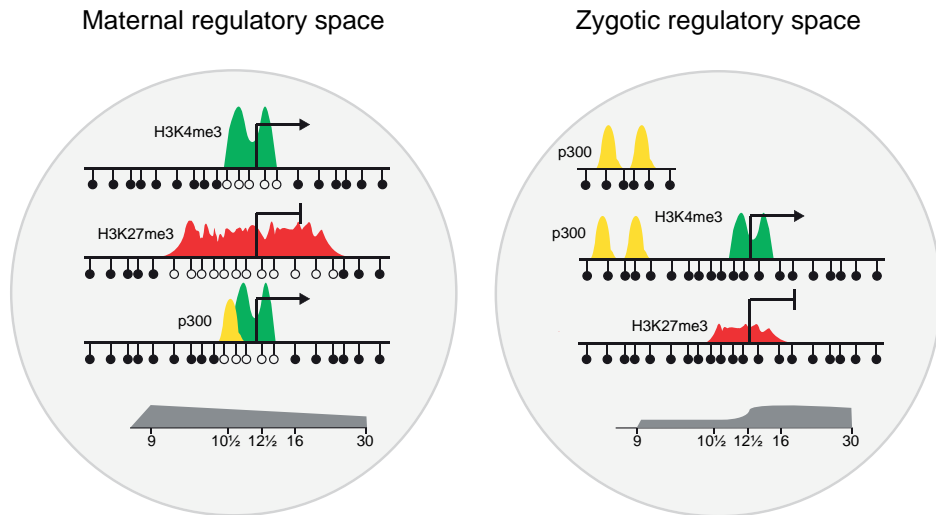


Figure 8. Model of maternal and zygotic regulatory space. This shows the segregation of maternal regulatory space, which contains hypomethylated promoters that are mainly controlled by maternal factors, and zygotic regulatory space, which includes methylated promoters and enhancers that are under zygotic control. Most p300-bound enhancers are in zygotic space, however, they can regulate promoters in both maternal and zygotic space, crossing the regulatory space border. This may contribute to varying degrees of permissiveness to transcriptional activation. Maternal regulatory space extends well into neurula and tailbud stages and includes many embryonic genes which are activated at specific stages of development. Zygotic regulatory space requires zygotic transcription, is established from the mid-blastula stage onwards but increases in relative contribution during development.

Key proteins of the molecular machinery involved in DNA methylation (Dnmt3a, Tet2), H3K4me3 (Mll1-4, Kdm5b/c), H3K27me3 (Ezh2, Eed, Kdm6a/b) and enhancer histone acetylation (p300) are not only highly conserved between species but also frequently mutated in cancer^{49, 50, 51}. Moreover cancer-specific hypermethylated regions tend to correspond to Polycomb-regulated loci in embryonic stem cells and DNA methylation may restrict H3K27 methylation globally^{52, 53}. In addition, the sequence signatures of hypomethylated regions that acquire H3K4me3 or H3K27me3 are conserved between fish, frogs and humans¹³. These observations suggest that the molecular mechanisms that orchestrate the maternal and zygotic regulatory space are conserved. One key difference between mammals and non-mammalian vertebrates is the specification of extra-embryonic lineages between zygotic genome activation and the blastocyst stage in mammals¹⁰, so it is likely that the way this plays out for specific genes differs between species. In summary, our results provide an unprecedented view of the far reach of maternal factors in zygotic life through chromatin state. The dichotomy of maternal promoter-based and embryonic enhancer regulation demarcates an epigenetic maternal-to-zygotic transition that is maternal-permissive to the expression of some embryonic genes and restrictive to others. This highlights the combinatorial interplay of maternal and

zygotic factors through distinct mechanisms.

Methods

Animal procedures

Xenopus tropicalis embryos were obtained by in vitro fertilisation, dejellied in 3% cysteine and collected at the indicated stage. Fertilised eggs were injected with 2.3 nl of 2.67 ng/ μ l α -amanitin and developed until the control embryos reached mid-gastrulation (stage 11). Animal use was conducted under the DEC permission (Dutch animal experimentation committee) RU-DEC 2012-116 and 2014-122 to G.J.C.V..

ChIP-sequencing and RNA-sequencing

Chromatin for chromatin-immunoprecipitation (ChIP) was prepared as previously described 54, 55, with minor modifications. Antibody was incubated with chromatin overnight, followed by incubation with Dynabeads[®] Protein G for 1 hour. The following antibodies were used: anti-H3K4me1 (Abcam ab8895, 1 μ g /15 embryo equivalents (eeq)), anti-H3K4me3 (Abcam ab8580, 1 μ g /15 eeq), anti-H3K9ac (Upstate/Millipore 06-942, 1 μ g/15 eeq), anti-H3K36me3 (Abcam ab9050, 1 μ g /15 eeq), anti-H3K27me3 (Upstate/Millipore 07-449, 1 μ g/15 eeq), anti-H3K9me2 (Diagenode C15410060, 1 μ g /15 eeq), anti-H3K9me3 (Abcam ab8898, 2 μ g /15 eeq), anti-H4K20me3 (Abcam ab9053, 2 μ g /15 eeq), anti-p300 (Santa Cruz sc-585, 1 μ g /15 eeq), and anti-RNAPII (Diagenode C15200004, 1 μ g /15 eeq). For all ChIP-seq samples of the epigenome reference maps and RNAPII ChIP-seq samples of the α -amanitin experiments three biological replicates of different chromatin isolations of 45 embryos were pooled. Two biological replicates for H3K4me3 (α -amanitin injected: resp. 90 and 56 embryo equivalents (eeq); control: resp. 45 and 67 eeq), H3K27me3 (α -amanitin injected: resp. 90 and 180 eeq; control: resp. 45 and 202 eeq) and p300 (α -amanitin injected: resp. 112 and 56 eeq; control: resp. 112 and 67 eeq) ChIP-seq samples of the α -amanitin experiments were generated. For RNA-seq samples of the α -amanitin experiments RNA from 5 embryos from one biological replicate was isolated and depleted of ribosomal RNA as previously described³⁵. Samples were subjected to a qPCR quality check pre- and post-preparation. Libraries were prepared with the Kapa Hyper Prep kit (Kapa Biosystems), and sequencing was done on the Illumina HiSeq2000 platform. Reads were mapped to the reference *Xenopus tropicalis* genome JGI7.1, using STAR (RNA-seq) or BWA (ChIP-seq) allowing one mismatch.

MethylC-seq

Genomic DNA from *Xenopus* embryos stage 9 and 10.5 was obtained as described before 56. MethylC-seq library generation was performed as described previously 57. Library amplification was performed with KAPA HiFi HotStart Uracil+ DNA polymerase (Kapa Biosystems, Woburn, MA), using 6 cycles of amplification. Single-read MethylC-seq libraries were processed and aligned as described previously⁵⁸.

Quantitative PCR (qPCR)

PCR reactions were performed on a CFX96 Touch™ Real-Time PCR Detection System (BioRad) using iQ Custom SYBR Green Supermix (BioRad). We performed RNA expression PCR (RT-qPCR) and ChIP-qPCR for H3K4me3 and H3K9ac on promoters of *odc1*, *eef1a10*, *rnf146*, *tor1a*, *zic1*, *cdc14b*, *eomes*, *xrcc1*, *drosha*, *gdf3*, *t*, *tbx2*, *fastkd3*, *gs17* (see Supplementary Methods for primer sequences). ChIP-qPCR enrichment over background was calculated using the average of 5 negative loci.

Detection of enriched regions

We used MACS259 with standard settings and a q-value of 0.05. Fragment size was determined using phantompeakqualtools⁶⁰. Broad settings (--BROAD) were used for H3K4me1, H3K36me3, H3K27me3, H3K9me2, H3K9me3, H4K20me3 and RNAPII. Broad and narrow peaks were merged for H3K4me3. For H3K9ac narrow peaks were used. For p300 broad peaks were used in the ChomHMM analysis, narrow p300 peaks were used for super-enhancer and MaD versus ZyD analyses. All peaks were called relative to an input control track. Peaks that showed at least 75% overlap with 1 kb regions that have more than 65 input reads, and peaks that have a ChIP-seq RPKM higher than the 95 percentile of random background regions are excluded from further analysis. Only scaffolds 1-10 (the chromosome-sized scaffolds) were included in the analysis. Relative RPKM was calculated by dividing the ChIP-seq RPKM of the peaks by the ChIP-seq RPKM of the 95 percentile of random background regions.

We used MAnorm⁶¹ to determine differentially enriched regions in α -amanitin and control embryos. We used merged peak sets of replicate 1 replicate 2 and stage 10.5 to avoid bias caused by peak calling. Lost, gained and unchanged peaks per biological replicate were determined using the following parameters:

lost peaks have M-values greater than 1 and a -log base 10(p-value) greater than 5 (for H3K27me3) or 1.3 (for H3K4me3 and p300) and have a relative RPKM (background corrected) greater than 1 in stage 11 control (no cut-off was used for st.11 control of H3K27me3 rep.1), stage 10.5 (H3K4me3 and p300) or stage 12 (H3K27me3); increased peaks have M-values smaller than -1 and a -log base 10(p-value) greater than 5 (H3K27me3) or 1.3 (H3K4me3 and p300) and have a rel. RPKM greater than 1 in stage 11 α -amanitin, stage 10.5 (H3K4me3 and p300) or stage 12 (H3K27me3); unchanged peaks are neither gained nor lost and have a rel. RPKM greater than 1 in stage 11 control (no cut-off was used for st.11 control of H3K27me3 rep.1), stage 11 α -amanitin, stage 10.5 (H3K4me3 and p300) or stage 12 (H3K27me3). Maintained peaks are peaks that are not lost and have a rel. RPKM greater than 1 in stage 11 control (no cut-off was used for st.11 control of H3K27me3 rep.1), stage 11 α -amanitin, stage 10.5 (H3K4me3 and p300) or stage 12 (H3K27me3). Common lost, gained, unbiased and maintained peaks are present in both replicates. All other peaks are considered not defined (ND). Replicate-specific peaks were only used for Supplementary Fig. 4b, for all other figures the common peaks were used.

DNA methylation levels in Supplementary Fig. 4d was calculated using previously published Bio-CAP data⁶². Bio-CAP RPKM levels of stage 11-12 were calculated for H3K4me3, H3K27me3 and p300 peaks, and corrected for Input values. For Fig. 4c genes were considered “hypomethylated” if the Bio-CAP/Input ratio on the promoter (+ 1 kb from TSS) was higher than 1.

RNA expression analysis was performed as previously published³⁵. Embryonic transcripts were separated based on the clustering of maximum expression levels per stage in Fig. 3d of Paranjpe et al.³⁵ (cluster 1 = blastula, cluster 5 = gastrula, clusters 3 and 4 = neurula, clusters 2 and 6 = tailbud).

Enhancer clusters were called as previously described⁴³. Enhancer Clusters are called per stage and merged to determine the total Enhancer Cluster region. Percentage of the EC region is calculated relative to the total Enhancer Cluster region.

MaD and ZyD classification

Maternally defined (MaD) peaks emerge at or before stage 11 and are also acquired in α -amanitin treated embryos in both replicates. Zygotically defined (ZyD) peaks appear at or before stage 11 and are lost in α -amanitin treated embryos in both replicates, or emerge after stage 11. To classify MaD and ZyD

H3K4me3 genes we ran MAnorm on promoters (+ 250 bp from TSS) only, using similar restrictions as described in Detection of enriched regions. MaD H3K4me3 genes have a maintained promoter in both replicates, ZyD H3K4me3 genes have a lost promoter H3K4me3 peak in both α -amanitin replicates, or a peak that emerges after stage 11. MaD H3K27me3 genes have at least one MaD peak in the vicinity of their promoter (+ 2.5 kb from TSS). ZyD H3K27me3 genes have at least one ZyD peak in their promoter and lack a MaD peak. Not defined (ND) peaks or genes do not meet the criteria for neither MaD nor ZyD. For p300 the total number of ZyD and MaD peaks was counted in GREAT⁶³ regions of genes.

ChomHMM analysis

Chromatin states were discovered and characterized using ChromHMM v1.1026, an implementation of a hidden Markov model. As input we used the enriched regions from ten tracks (H3K27me3, H3K36me3, H3K4me1, H3K4me3, H3K9ac, H3K9me2, H3K9me3, H4K20me3, p300 and RNAPII) across five developmental stages. We trained and ran the model with a range of states, and determined the 19 emission states model as the optimal number of states that could sufficiently capture the biological variation in co-occurrence of chromatin marks. We subsequently classified the states into 7 main groups based on the presence and absence of specific chromatin marks.

The segmentation files of the 7 main groups per stage were binned in 200 base pairs intervals. An $m \times n$ matrix was created, where m corresponds to the 200 base-pair intervals and n to the developmental stages (9-30). Each element $a(i,j)$ represents the chromatin state of interval i at stage j . For each chromatin group occurrences were counted per stage n . The changes between stage n and $n+1$ were plotted using Sankey diagrams (<https://github.com/tamc/Sankey>), a flow diagram closely related to alluvial diagrams.

Motif analyses

For the prediction of motif contribution to p300 recruitment (Fig. 5a) we have implemented the ISMARA method developed by Balwierz et al.⁶⁴. This method uses motif activity response analysis to determine the transcription factors that drive the observed changes in chromatin state across samples. As input we used the number of known motifs found per p300 binding site and the RPKM of the p300 peaks per developmental stage. The model infers the unknown motif activities from the equation in which the changes in signal levels are explained

with the number of binding sites and the unknown motif activities. Motifs that showed a z-score activity that was higher than 13 are shown in Fig. 5a. Enriched motifs (Supplementary Fig. 7) were detected with *gimme diff*, a tool from the *GimmeMotifs* package 65. The vertebrate motifs used in this script were obtained from CIS-BP (<http://cisbp.ccb.utoronto.ca/>)66 and clustered using *gimme cluster* from *GimmeMotifs*. The motifs are available at <http://dx.doi.org/10.6084/m9.figshare.1555851> (Van Heeringen, Simon J. (2015): Vertebrate motif clusters v3.0. figshare.).

Generation of plots and heatmaps

All heatmaps were generated using *fluff* (<http://simonvh.github.com/fluff>)13 or *gplots* (<http://cran.r-project.org/web/packages/gplots/index.html>). For all heatmap clustering, the Euclidean distance metric was used. Other plots were generated using *ggplot2* (<http://ggplot2.org/>).

Data accessibility

The data generated for this work have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE67974. Visualization tracks are available at the authors' web site (<http://www.ncmls.nl/gertjanveenstra>).

Additional Information ChIP-seq and bisulfite-seq data is available via the GEO database through GEO Series accession number GSE67974. Visualization tracks and a data track hub are available at the authors' web site (<http://www.ncmls.nl/gertjanveenstra>). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to G.J.C.V. (g.veenstra@science.ru.nl).

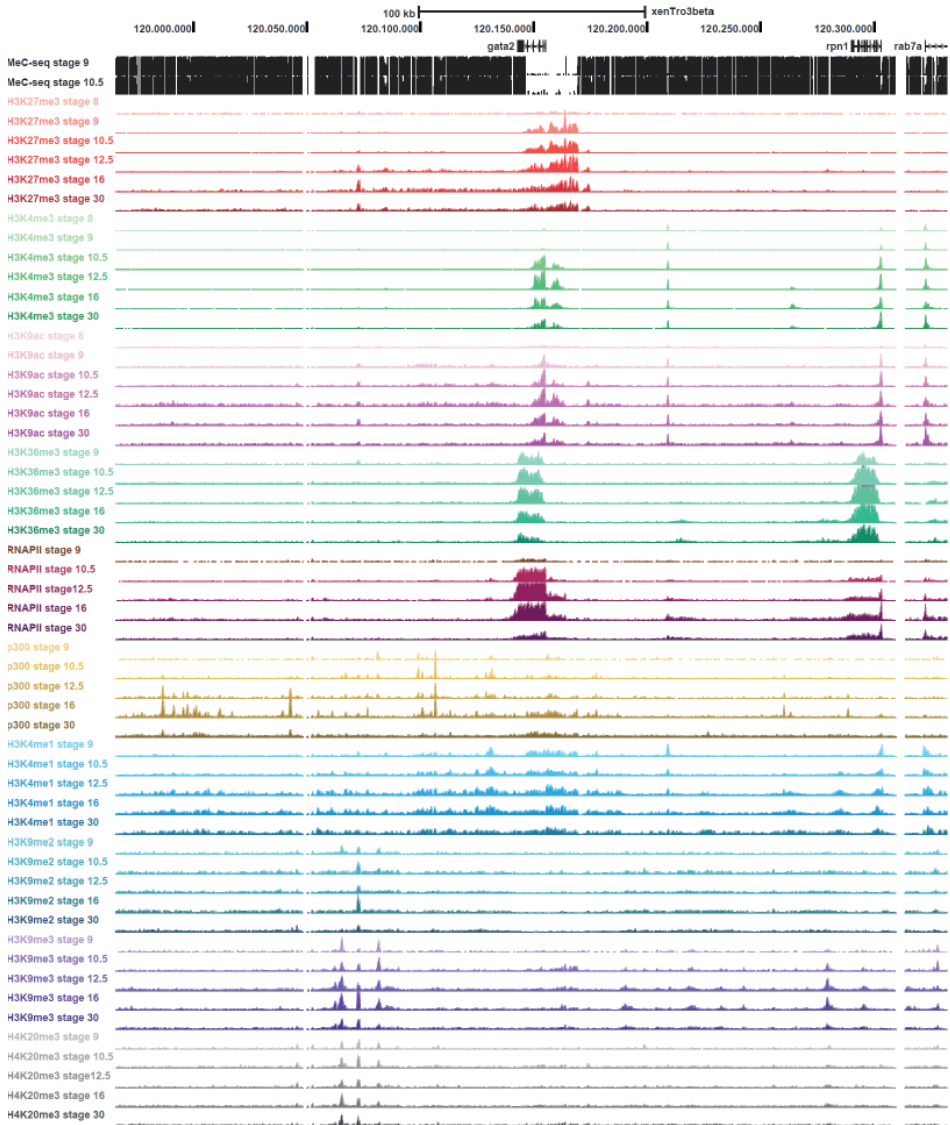
Supplementary Information accompanies this paper at <http://www.nature.com/naturecommunications>

Acknowledgements

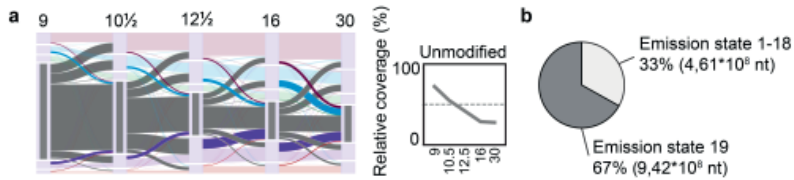
This work is supported by the US National Institutes of Health (NICHD, grant R01HD069344). Part of this work was carried out on the Dutch national

e-infrastructure with the support of SURF Foundation. SJvH is supported by the Netherlands Organization for Scientific research (NWO-ALW, grant 863.12.002). O.B. is supported by an Australian Research Council Discovery Early Career Researcher Award - DECRA (DE140101962).

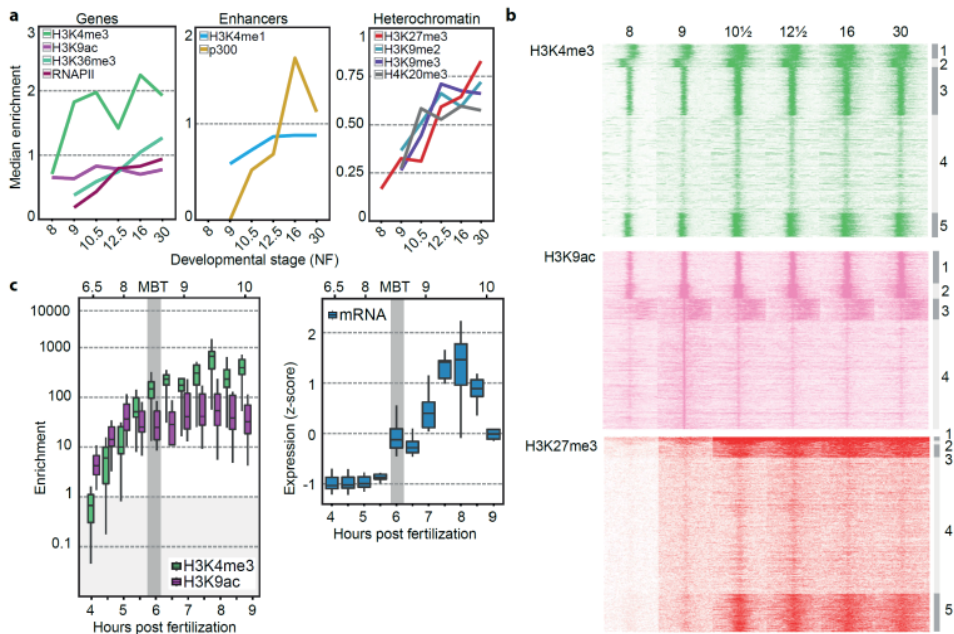
Supplementary Figures



Supplementary Figure 1. *Gata2* locus. ChIP-seq enrichment of histone modifications, RNAPII and p300 for stages 9- 30. The heterochromatin tracks (H3K9me2, H3K9me3 and H4K20me3) are shown including non-unique sequence reads, identifying repetitive regions enriched for these modifications. ChIP-sequencing on stage 8 (blastula, pre-MBT) was done for histone modifications H3K4me3, H3K9ac and H3K27me3.

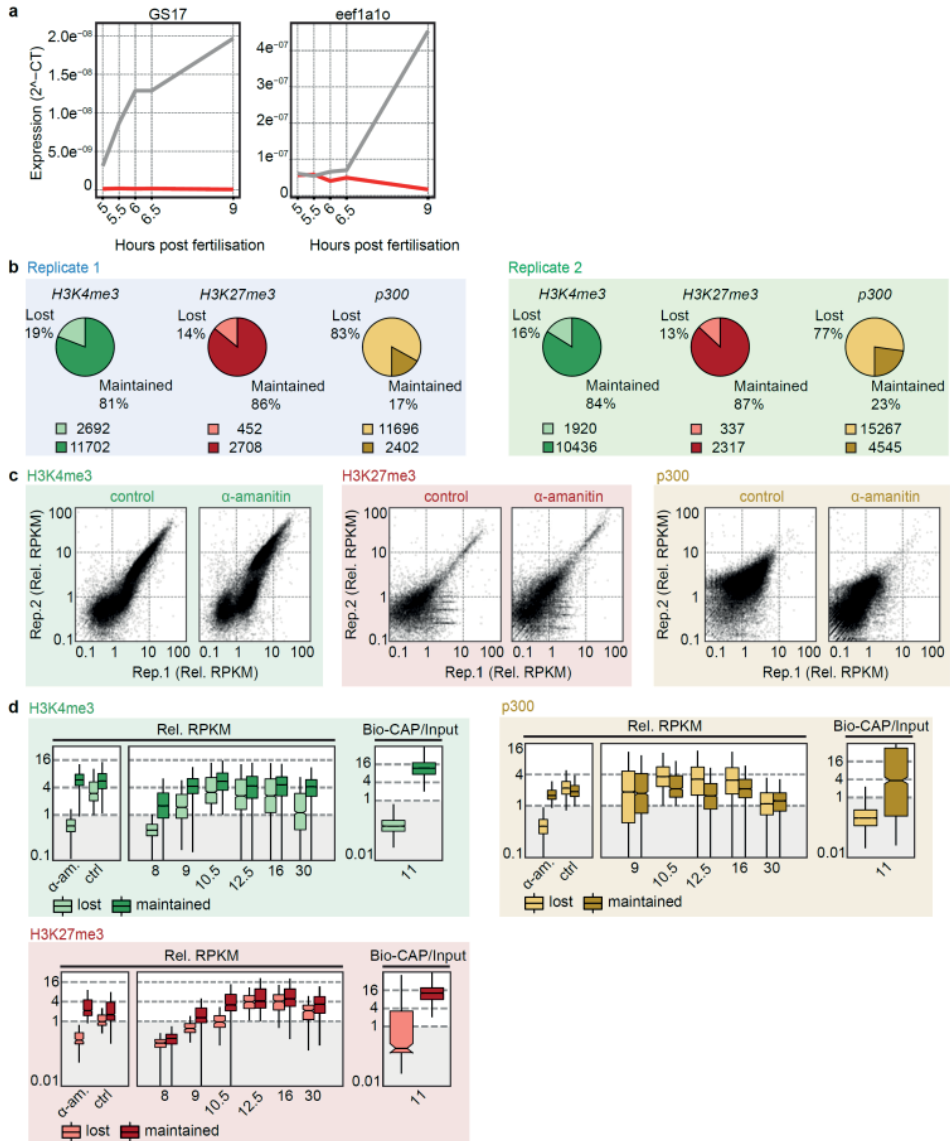


Supplementary Figure 2. Unmodified state coverage. (a) Alluvial plots of unmodified state (grey) coverage during development. The height represents the fraction of the modified genome that contributes to the same or a different chromatin state. The line plots shows coverage of the unmodified state per stage as a percentage of the sum of all regions that are state 1-18 at any stage. (b) Absolute nucleotide coverage of emission state 19 and states 1-18 at stage 30. It should be noted that ‘unmodified’ specifically refers to the examined histone modifications and that this state shows abundant DNA methylation.



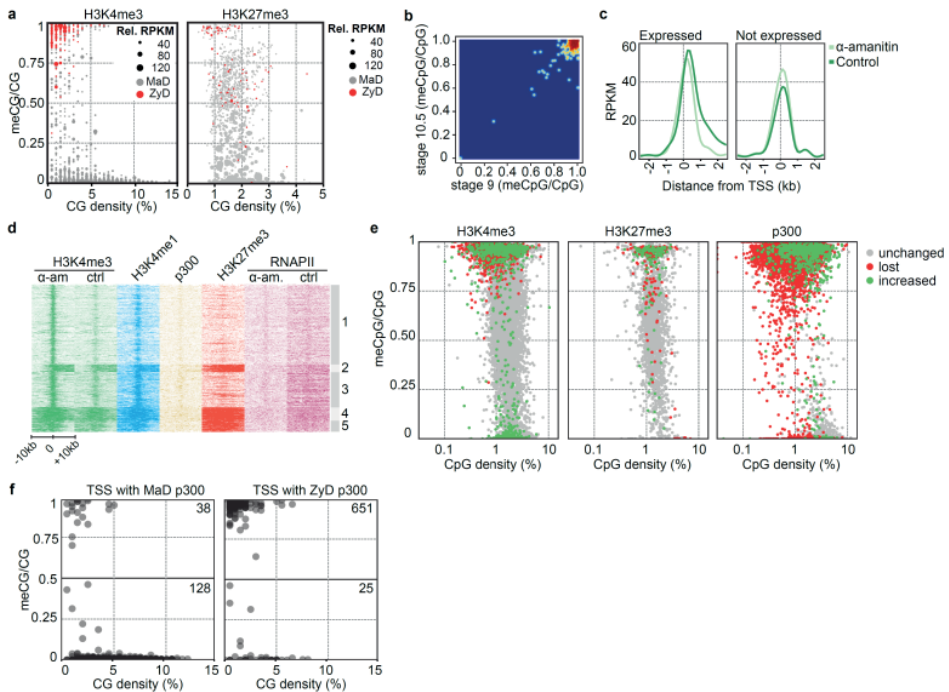
Supplementary Figure 3. Progressive specification of the epigenome. (a) Median enrichment of chromatin marks during development. (b) RPKM levels of H3K4me3, H3K9ac and H3K27me3 stage 8-30 on stage 9 peaks. Most stage 9 H3K4me3 and H3K9ac peaks show already significant enrichment at stage 8, whereas H3K27me3 markedly increases in late blastula and early gastrula embryos. (c) Detailed time series from 4 to 9 hours post fertilization (13 genes, average values of two biological replicates, see Methods). Left panel: Box plot of ChIP-qPCR for H3K9ac (pink) and H3K4me3 (green). Right panel: Box plot of RNA expression (RT-qPCR). Box: 25th (bottom), 50th (internal band), 75th (top) percentiles. Whiskers: 1.5 * interquartile range of the lower and upper quartiles, respectively.

Embryonic transcription is controlled by maternally defined chromatin state

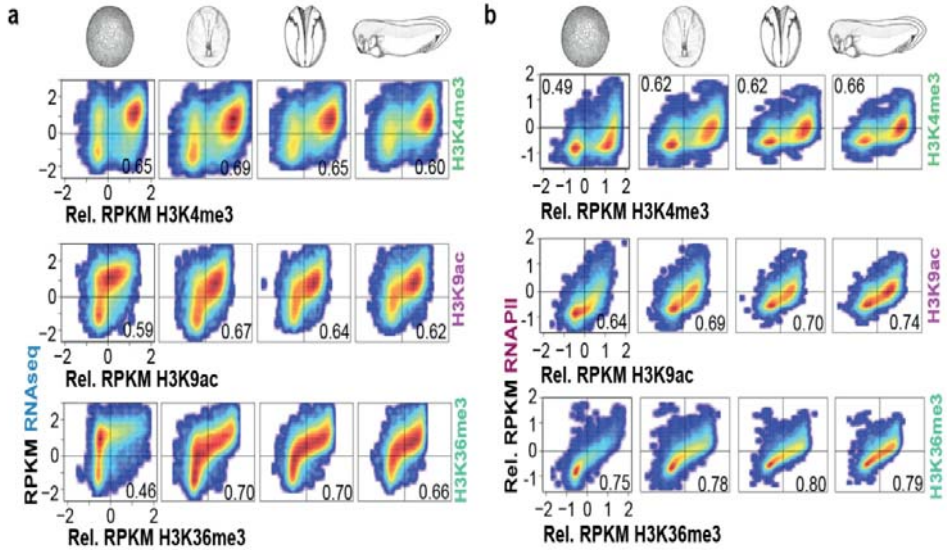


Supplementary Figure 4. Maternal and zygotic acquisition of chromatin state. (a) RNA expression (RT-qPCR) of *gs17* (embryonic transcript), *eef1a1o* (maternal transcript, induced at MBT) in α -amanitin and control embryos. (b) Lost and Maintained peaks of H3K4me3, H3K27me3 and p300 in replicate 1 (left, blue background) and replicate 2 (right, green background). Pie charts representing percentage and numbers of lost and maintained peaks per replicate. (c) Scatter plots with relative RPKM (background corrected) of replicate 1 (x-axis) and replicate 2 (y-axis) on peaks that are lost or maintained in both experiments. (d) Left and middle panels show box plots of relative RPKM (background corrected) of regions with MaD or ZyD H3K4me3, H3K27me3 or p300-binding. Right panels show box plots of input corrected RPKM of previously profiled Bio-CAP data representing hypomethylated DNA domains 61. MaD trimethylation of H3K4 and H3K27 is detected almost exclusively on Bio-CAP-enriched regions indicating clusters of hypomethylated CpGs. Box: 25th (bottom), 50th (internal band), 75th (top) percentiles. Whiskers: 1.5 * interquartile range of the lower and upper quartiles, respectively.

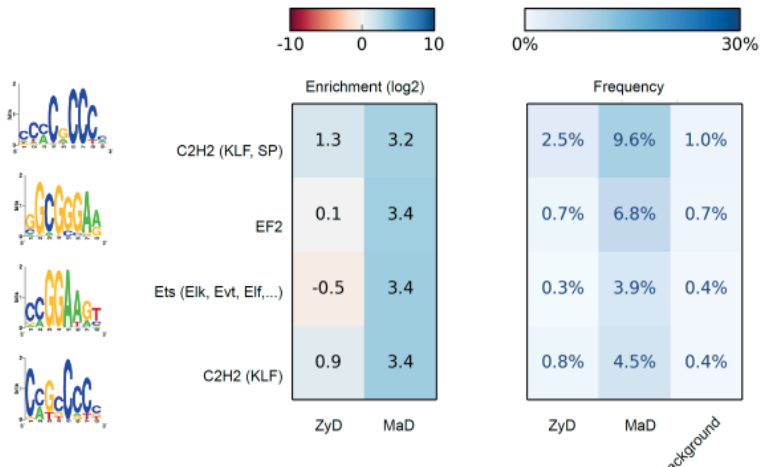
Chromatin regulation in *Xenopus* embryos



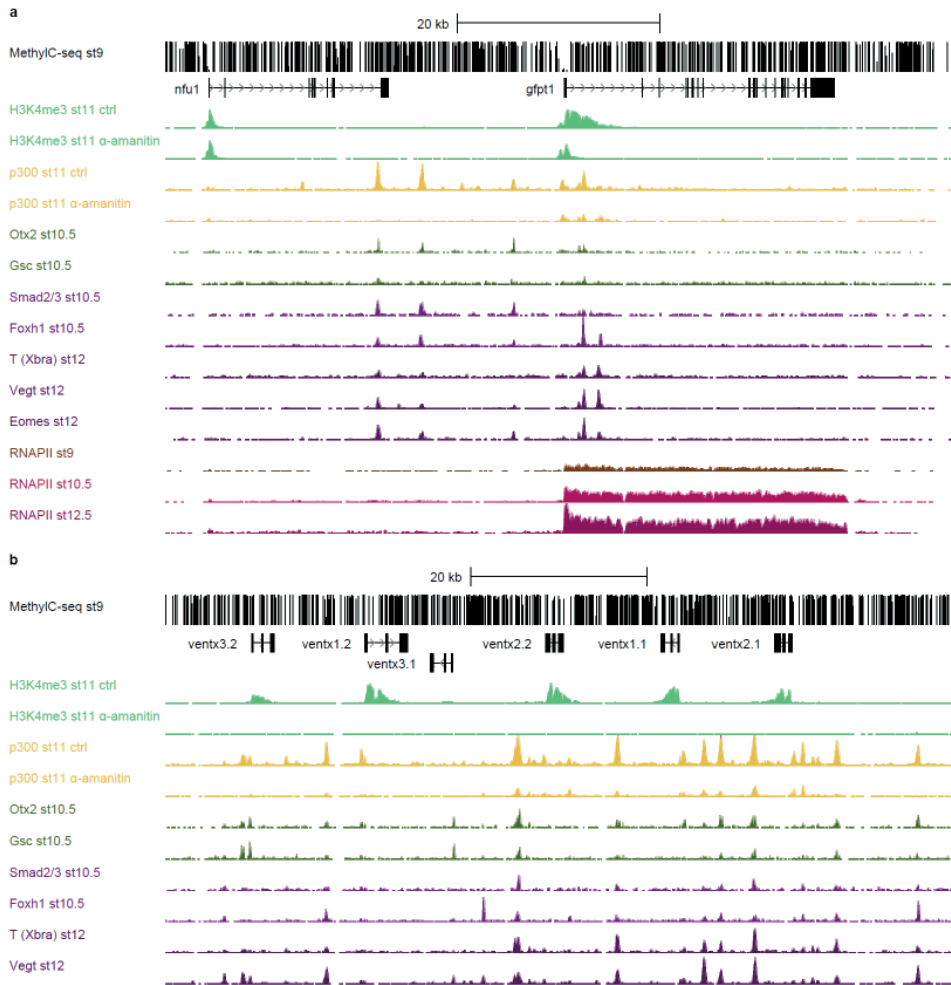
Supplementary Figure 5. Methylation logic for maternal and zygotic defined chromatin state. (a) CpG density and methylation at stage 10.5 of promoters (H3K4me3: + 100 bp from TSS; H3K27me3: + 2.5 kb from TSS) that contain a zygotic defined (ZyD, lost in α -amanitin treated embryos, red) or maternal defined (MaD, maintained in α -amanitin treated embryos, grey) peak for H3K4me3 (left) or H3K27me3 (right) after inhibition of embryonic transcription. The size of the dot indicates the relative RPKM (background corrected). (b) Density heatmap of DNA methylation stage 9 (x-axis) and stage 10.5 (y-axis) on ZyD promoters (+ 100 bp from TSS). (c) Mean relative RPKM of stage 11 α -amanitin and control H3K4me3 on promoters of stage 10.5 expressed (left) and not expressed genes (right). (d) Heatmap representation of regions with increased H3K4me3 deposition in α -amanitin treated embryos. (e) CG density and methylation (stage 9) on lost, increased and unchanged H3K4me3 (left), H3K27me3 (middle) or p300 (right) peaks. For the purpose of simplicity, unchanged and increased peaks are collectively referred to as MaD in the rest of this article. (f) CG density and methylation on promoters (+ 100 bp from TSS) that overlap with MaD (left) or ZyD (right) p300-bound peaks. The values in the middle and top corners indicate the number of promoters with meCG/CG ratio above or below 0.5.



Supplementary Figure 6. Correlation of chromatin marks and transcription. Density correlation plots of relative RPKM (background corrected) for H3K4me3 and H3K9ac (+ 1 kb from TSS) and H3K36me3 (genes bodies) with (a) RNAseq (exons) or (b) RNA polymerase II (gene bodies).

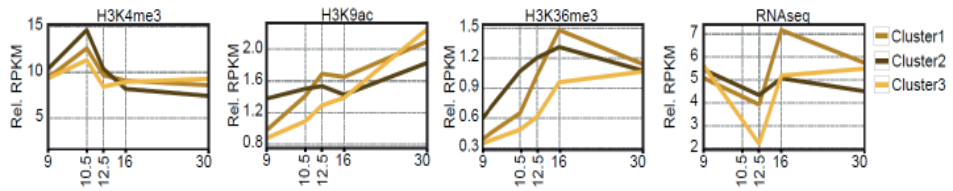


Supplementary Figure 7. MaD p300 regions are enriched for promoter related motif sequences. Motif enrichment and frequency in MaD and ZyD p300-bound regions.

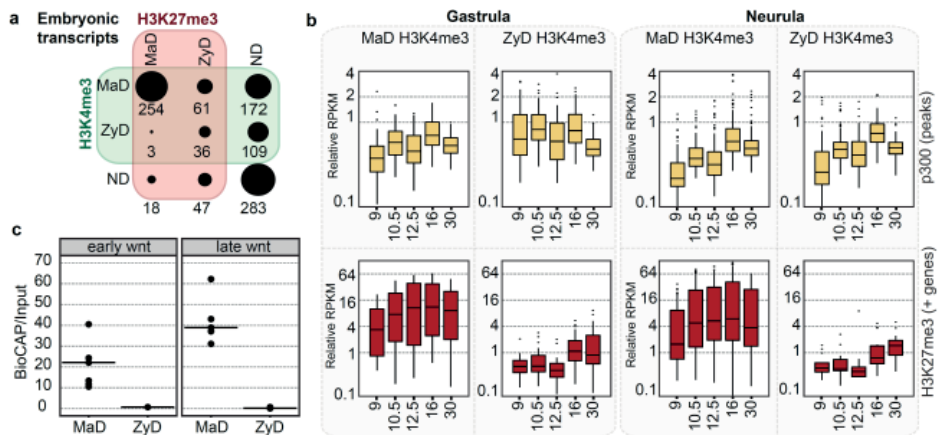


Supplementary Figure 8. MaD and ZyD p300 bound regions recruit embryonically regulated transcription factors. *Gfp1* (a) and *ventx* (b) locus with stage 9 MethylC-seq and ChIP-seq enrichment of H3K4me3 and p300 on control and α -amanitin injected embryos, transcription factors Otx2, Gsc, Smad2/3, Foxh1, T (Xbra), Vegt, Eomes and RNAPII on stage 9 10.5 and 12.5.

Embryonic transcription is controlled by maternally defined chromatin state



Supplementary Figure 9. Histone modifications and transcript levels of EC-associated genes. Median relative RPKM (background corrected) of H3K4me3 and H3K9ac (+ 1 kb from TSS), H3K36me3 (gene bodies) and RNAseq (exons) for genes near ECs per heatmap cluster (Figure 3d).



Supplementary Figure 10. Maternal and zygotic control of embryonic transcripts. Maternally defined (MaD) peaks emerge at or before stage 11 independent of embryonic transcription. Zygotically defined (ZyD) peaks appear before stage 11 and are lost in α -amanitin treated embryos, or emerge at or after stage 12. Not determined (ND) peaks are not detected in stage 11 control embryos. (a) Maternal and zygotic control of H3K4me3 and H3K27me3 on promoters of embryonic transcripts (total number of transcripts: 983). (b) Box plots of p300 RPKM (background corrected) in GREAT regions of genes, or H3K27me3 RPKM (background corrected) in promoters of genes with at least one H3K27me3 peak in their promoter (+ 2.5 kb from TSS). Box: 25th (bottom), 50th (internal band), 75th (top) percentiles. Whiskers: 1.5 * interquartile range of the lower and upper quartiles, respectively. Outliers are indicated with black dots. (c) BioCAP enrichment (RPKM BioCAP/ RPKM Input) as a measure for hypomethylated DNA domains on the promoters (+ 1 kb from TSS) of early and late Wnt target genes.

References

1. Mikkelsen TS, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448, 553-560 (2007).
2. Strahl BD, Allis CD. The language of covalent histone modifications. *Nature* 403, 41-45 (2000).
3. Lee JS, Smith E, Shilatifard A. The language of histone crosstalk. *Cell* 142, 682-685 (2010).
4. Jenuwein T, Allis CD. Translating the histone code. *Science* 293, 1074-1080 (2001).
5. Gerstein MB, et al. Integrative analysis of the *Caenorhabditis elegans* genome by the modENCODE project. *Science* 330, 1775-1787 (2010).
6. Roy S, et al. Identification of functional elements and regulatory circuits by *Drosophila* modENCODE. *Science* 330, 1787-1797 (2010).
7. Tadros W, Lipshitz HD. The maternal-to-zygotic transition: a play in two acts. *Development* 136, 3033-3042 (2009).
8. Kane DA, Kimmel CB. The zebrafish midblastula transition. *Development* 119, 447-456 (1993).
9. O'Farrell PH, Stumpff J, Su TT. Embryonic cleavage cycles: how is a mouse like a fly? *Current biology* : CB 14, R35-45 (2004).
10. Paranjpe SS, Veenstra GJ. Establishing pluripotency in early development. *Biochimica et biophysica acta*, (2015).
11. Newport J, Kirschner M. A major developmental transition in early *Xenopus* embryos: I. characterization and timing of cellular changes at the midblastula stage. *Cell* 30, 675-686 (1982).
12. Newport J, Kirschner M. A major developmental transition in early *Xenopus* embryos: II. Control of the onset of transcription. *Cell* 30, 687-696 (1982).
13. van Heeringen SJ, et al. Principles of nucleation of H3K27 methylation during embryonic development. *Genome research* 24, 401-410 (2014).
14. Akkers RC, et al. A hierarchy of H3K4me3 and H3K27me3 acquisition in spatial gene regulation in *Xenopus* embryos. *Developmental cell* 17, 425-434 (2009).
15. Vastenhouw NL, et al. Chromatin signature of embryonic pluripotency is established during genome activation. *Nature* 464, 922-926 (2010).
16. Lindeman LC, et al. Prepatterning of developmental gene expression by modified histones before zygotic genome activation. *Developmental cell* 21, 993-1004 (2011).
17. Kim TK, et al. Widespread transcription at neuronal activity-regulated enhancers. *Nature* 465, 182-187 (2010).
18. Visel A, et al. ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* 457, 854-858 (2009).
19. Heintzman ND, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* 459, 108-112 (2009).
20. Heintzman ND, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature genetics* 39, 311-318 (2007).
21. Roh TY, Wei G, Farrell CM, Zhao K. Genome-wide prediction of conserved and nonconserved enhancers by histone acetylation patterns. *Genome research* 17, 74-81 (2007).
22. Bernstein BE, et al. Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* 120, 169-181 (2005).
23. Santos-Rosa H, et al. Active genes are tri-methylated at K4 of histone H3. *Nature* 419, 407-411 (2002).
24. Barski A, et al. High-resolution profiling of histone methylations in the human genome. *Cell* 129, 823-837 (2007).
25. Schotta G, et al. A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes & development* 18, 1251-1262 (2004).
26. Ernst J, Kellis M. ChromHMM: automating chromatin-state discovery and characterization. *Nature methods* 9, 215-216 (2012).
27. Chafin DR, Guo H, Price DH. Action of alpha-amanitin during pyrophosphorolysis and elongation by RNA polymerase II. *The Journal of biological chemistry* 270, 19114-19119 (1995).
28. Sible JC, Anderson JA, Lewellyn AL, Maller JL. Zygotic transcription is required to block a maternal program of apoptosis in *Xenopus* embryos. *Developmental biology* 189, 335-346 (1997).
29. Skirkanich J, Luxardi G, Yang J, Kodjabachian L, Klein PS. An essential role for transcription before the MBT in *Xenopus laevis*. *Developmental biology* 357, 478-491 (2011).
30. Clouaire T, et al. Cfp1 integrates both CpG content and gene activity for accurate H3K4me3 deposition in embryonic stem cells. *Genes & development* 26, 1714-1728 (2012).
31. Clouaire T, Webb S, Bird A. Cfp1 is required for gene expression-dependent H3K4 trimethylation and H3K9

- acetylation in embryonic stem cells. *Genome biology* 15, 451 (2014).
32. Thomson JP, et al. CpG islands influence chromatin structure via the CpG-binding protein Cfp1. *Nature* 464, 1082-1086 (2010).
 33. van Kruijsbergen I, Hontelez S, Veenstra GJ. Recruiting polycomb to chromatin. *The international journal of biochemistry & cell biology* 67, 177-187 (2015).
 34. Ng HH, Robert F, Young RA, Struhl K. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Molecular cell* 11, 709-719 (2003).
 35. Paranjpe SS, Jacobi UG, van Heeringen SJ, Veenstra GJ. A genome-wide survey of maternal and embryonic transcripts during *Xenopus tropicalis* development. *BMC genomics* 14, 762 (2013).
 36. Lauberth SM, et al. H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. *Cell* 152, 1021-1036 (2013).
 37. Vermeulen M, et al. Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* 131, 58-69 (2007).
 38. Sims RJ, 3rd, et al. Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. *Molecular cell* 28, 665-676 (2007).
 39. Gentsch GE, et al. In vivo T-box transcription factor profiling reveals joint regulation of embryonic neurodermal bipotency. *Cell reports* 4, 1185-1196 (2013).
 40. Yasuoka Y, et al. Occupancy of tissue-specific cis-regulatory modules by Otx2 and TLE/Groucho for embryonic head specification. *Nature communications* 5, 4322 (2014).
 41. Chiu WT, et al. Genome-wide view of TGFbeta/Foxh1 regulation of the early mesendoderm program. *Development* 141, 4537-4547 (2014).
 42. Pott S, Lieb JD. What are super-enhancers? *Nature genetics* 47, 8-12 (2015).
 43. Whyte WA, et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* 153, 307-319 (2013).
 44. Hnisz D, et al. Super-enhancers in the control of cell identity and disease. *Cell* 155, 934-947 (2013).
 45. Blythe SA, Cha SW, Tadjuidje E, Heasman J, Klein PS. beta-Catenin primes organizer gene expression by recruiting a histone H3 arginine 8 methyltransferase, Prmt2. *Developmental cell* 19, 220-231 (2010).
 46. Bogdanovic O, et al. Temporal uncoupling of the DNA methylome and transcriptional repression during embryogenesis. *Genome research* 21, 1313-1327 (2011).
 47. Potok ME, Nix DA, Parnell TJ, Cairns BR. Reprogramming the maternal zebrafish genome after fertilization to match the paternal methylation pattern. *Cell* 153, 759-772 (2013).
 48. Jiang L, et al. Sperm, but not oocyte, DNA methylome is inherited by zebrafish early embryos. *Cell* 153, 773-784 (2013).
 49. Lawrence MS, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* 505, 495-501 (2014).
 50. Huether R, et al. The landscape of somatic mutations in epigenetic regulators across 1,000 paediatric cancer genomes. *Nature communications* 5, 3630 (2014).
 51. Wu G, et al. The genomic landscape of diffuse intrinsic pontine glioma and pediatric non-brainstem high-grade glioma. *Nature genetics* 46, 444-450 (2014).
 52. Simmer F, et al. Comparative genome-wide DNA methylation analysis of colorectal tumor and matched normal tissues. *Epigenetics* 7, 1355-1367 (2012).
 53. Brinkman AB, et al. Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. *Genome research* 22, 1128-1138 (2012).
 54. Jallow Z, Jacobi UG, Weeks DL, Dawid IB, Veenstra GJ. Specialized and redundant roles of TBP and a vertebrate-specific TBP paralog in embryonic gene regulation in *Xenopus*. *Proceedings of the National Academy of Sciences of the United States of America* 101, 13525-13530 (2004).
 55. Akkers RC, Jacobi UG, Veenstra GJ. Chromatin immunoprecipitation analysis of *Xenopus* embryos. *Methods in molecular biology* 917, 279-292 (2012).
 56. Bogdanovic O, Fernandez-Minan A, Tena JJ, de la Calle-Mustienes E, Gomez-Skarmeta JL. The developmental epigenomics toolbox: ChIP-seq and MethylCap-seq profiling of early zebrafish embryos. *Methods* 62, 207-215 (2013).
 57. Lister R, et al. Global epigenomic reconfiguration during mammalian brain development. *Science* 341, 1237905 (2013).
 58. Lister R, et al. Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* 471, 68-73 (2011).
 59. Zhang Y, et al. Model-based analysis of ChIP-Seq (MACS). *Genome biology* 9, R137 (2008).
 60. Kharchenko PV, Tolstorukov MY, Park PJ. Design and analysis of ChIP-seq experiments for DNA-binding proteins. *Nature biotechnology* 26, 1351-1359 (2008).

61. Shao Z, Zhang Y, Yuan GC, Orkin SH, Waxman DJ. MANorm: a robust model for quantitative comparison of ChIP-Seq data sets. *Genome biology* 13, R16 (2012).
62. Long HK, et al. Epigenetic conservation at gene regulatory elements revealed by non-methylated DNA profiling in seven vertebrates. *eLife* 2, e00348 (2013).
63. McLean CY, et al. GREAT improves functional interpretation of cis-regulatory regions. *Nature biotechnology* 28, 495-501 (2010).
64. Balwiercz PJ, Pachkov M, Arnold P, Gruber AJ, Zavolan M, van Nimwegen E. ISMARA: automated modeling of genomic signals as a democracy of regulatory motifs. *Genome research* 24, 869-884 (2014).
65. van Heeringen SJ, Veenstra GJ. GimmeMotifs: a de novo motif prediction pipeline for ChIP-sequencing experiments. *Bioinformatics* 27, 270-271 (2011).
66. Weirauch MT, et al. Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* 158, 1431-1443 (2014).

CHAPTER 4

Heterochromatic histone modifications at transposons in *Xenopus tropicalis* embryos

Ila van Kruijsbergen*, Saartje Hontelez*, Dei M. Elurbe, Simon J. van Heeringen, Martijn A. Huynen, Gert Jan C. Veenstra

*These authors contributed equally to this work

Developmental Biology (2016) In Press

lvK, SH and GJCV designed the study. CHIP data production was done by lvK and SH. The intra-subfamily divergence was determined by DME, with support from MAH. All other analyses were done by lvK with support from SJvH. All authors discussed the results. lvK wrote the manuscript with help and critical review of SH and GJCV.

Summary

*Transposable elements are parasitic genomic elements that can be deleterious for host gene function and genome integrity. Heterochromatic histone modifications are involved in the repression of transposons. However, it remains unknown how these histone modifications mark different types of transposons during embryonic development. Here we document the variety of heterochromatic epigenetic signatures at parasitic elements during development in *Xenopus tropicalis*, using genome-wide ChIP-sequencing data and ChIP-qPCR analysis. We show that specific subsets of transposons in various families and subfamilies are marked by different combinations of the heterochromatic histone modifications H4K20me3, H3K9me2/3 and H3K27me3. Many DNA transposons are marked at the blastula stage already, whereas at retrotransposons the histone modifications generally accumulate at the gastrula stage or later. Furthermore, transposons marked by H3K9me3 and H4K20me3 are more prominent in gene deserts. Using intra-subfamily divergence as a proxy for age, we show that relatively young DNA transposons are preferentially marked by early embryonic H4K20me3 and H3K27me3. In contrast, relatively young retrotransposons are marked by increasing H3K9me3 and H4K20me3 during development, and are also linked to piRNA-sized small non-coding RNAs. Our results implicate distinct repression mechanisms that operate in a transposon-selective and developmental stage-specific fashion.*

Introduction

Large parts of the human, mouse and frog genomes (resp. 46%, 37% and 35%) consist of repetitive and transposable elements (TEs) (Lander et al. 2001; Waterston et al. 2002; Hellsten et al. 2010). Different TEs can be distinguished based on their evolutionary origin and functional properties (Koonin et al. 2015; Muñoz-López & García-Pérez 2010). DNA transposons and retrotransposons represent two main classes with different replication strategies. DNA transposons move through the genome via cut and paste or rolling cycle mechanisms, while retrotransposons require RNA intermediates for a copy-paste mode of amplification. Depending on the transcriptional regulatory elements that they are equipped with, retrotransposons are divided further into subclasses: long terminal repeats (LTR), long interspersed nuclear elements (LINE) and short interspersed nuclear elements (SINE). Finally, within a subclass TEs are grouped into unique families and subfamilies based on their evolutionary origin (Bao et al. 2015). *Xenopus* does contain all main repeat families that are found in

mammals. However, whereas most TEs in mammals fall in the retrotransposon class, 70% of all TEs in *Xenopus* are DNA transposons (Hellsten et al. 2010).

It has been shown in human, mouse and frogs that TEs are actively and dynamically transcribed during early embryogenesis (Grau et al. 2014; Levin & Moran 2011). TEs contain functional elements, such as enhancers, promoters, polyadenylation signals, insulators and transcription factor binding sites. As a consequence TE spreading has had a major influence on host evolution (Friedli & Trono 2015). However, TEs also form a threat for the integrity of the host genome, since insertion can cause loss-of-function mutations. Insertion sites of retro- and DNA transposons in the *Xenopus* genome negatively correlate with the location of exons (Shen et al. 2013). This implies that TEs are more often found outside than inside coding regions, since gene disruption has a negative influence of survival of the zygote. Besides the loss-of-function mutations, also the introduction of cis-regulatory elements upon TE transposition can perturb host gene regulation (Friedli & Trono 2015). Therefore hosts have developed repressive defense mechanisms that restrain TE proliferation to some extent.

Various epigenetic modifications are involved in transcriptional repression. Besides DNA methylation (mainly on CpG dinucleotides, meCG), multiple histone modifications are involved in epigenetic silencing, such as histone H3 di- and trimethylation of lysine K9 (H3K9me_{2/3}), histone H3 trimethylation of lysine K27 (H3K27me₃) and histone H4 trimethylation of lysine K20 (H4K20me₃) (Lu et al. 2008; Fischle et al. 2003; Schultz et al. 2002; Jacobs & Khorasanizadeh 2002). Whereas chromatin decorated with H3K9me_{2/3} and H4K20me₃ is often referred to as constitutive heterochromatin, chromatin with H3K27me₃ is known as facultative heterochromatin, because it is found on genic regions in a cell type-specific manner (Trojer & Reinberg 2007). The repressive histone modifications can mediate transcriptional repression by chromosomal condensation via the recruitment of effector proteins such as Heterochromatin protein 1 (HP1) (Jacobs & Khorasanizadeh 2002; Lu et al. 2008; Schultz et al. 2002; van Kruijsbergen et al. 2015).

DNA methylation (meCG) is necessary for repressing distinct TEs in committed cells. For example, in mouse fibroblasts repression of LINE and ERV retrotransposons is dependent on meCG (Bulut-Karslioglu et al. 2014), but repression of Alu SINE retrotransposons rather depends on H3K9me₃ (Varshney et al. 2015). While needed in differentiated cells, meCG is dispensable for repression of LINE and ERV in mouse embryonic stem (ES) cells (Bulut-Karslioglu

et al. 2014; Matsui et al. 2010; Karimi et al. 2011; Hutnick et al. 2010; Martens et al. 2005).

A substantial part of TEs in pre-gastrula E6.25 mouse embryos is marked with H3K9me2 and/or H3K27me3 (Zylicz et al. 2015). However, less than 1% of the marked TEs showed increased expression in mouse embryos and ES cells lacking the G9a and EZH2 methyltransferases responsible for these modifications (Dong et al. 2008; Leeb et al. 2010; Maksakova et al. 2013; Zylicz et al. 2015). In contrast, the transcription of TEs marked with H3K9me3 increased in mouse ES cells deficient for the H3K9me3 methyltransferases SETDB1 or SUV39H1,2 (Wolf & Goff 2009; Matsui et al. 2010; Karimi et al. 2011; Bulut-Karslioglu et al. 2014). H4K20me3 was found to function downstream of H3K9me3 in silencing retrotransposons (Matsui et al. 2010). However, H4K20me3 can also function independently of H3K9me3, as occurs at DNA transposon family Charlie in mouse ES cells and at the retrotransposon family IAP in quiescent cells (Bierhoff et al. 2014; Martens et al. 2005). All together these studies have demonstrated that heterochromatic histone modifications are needed to restrain TEs.

Targeting of histone modifiers to TEs can be achieved via small non-coding RNAs such as piRNAs. piRNAs are derived from the transcribed RNA intermediates of retrotransposons, which are bound by Argonaute proteins (Vagin et al. 2006; Kalmykova et al. 2005). These Argonaute protein-containing complexes interact with H3K9 methyltransferases (SUV39 and SETDB) and are targeted to genomic DNA by the piRNAs (Klenov et al. 2011; Wang & Elgin 2011; Sienski et al. 2012; Le Thomas et al. 2013; Rozhkov et al. 2013; Sienski et al. 2015). A second mechanism to recruit histone modifiers involves DNA-binding zinc-finger proteins. The zinc-finger protein KRAB-ZFP interacts with TRIM28/KAP1 and the H3K9 methyltransferase SETDB1/ESET (Wolf & Goff 2009; Frieze et al. 2010; Rowe et al. 2010). A third way to recruit histone modifiers involves transcription factors. H3K9me2 catalyzing enzymes G9a/GLP are guided to their genomic targets via interaction with DNA-binding proteins, such as REST and SNAIL1 (Dong et al. 2012; Roopra et al. 2004).

piRNA and KAP1 driven recruitment mechanisms change over evolutionary time (Castro-Diaz et al. 2014; Pezic et al. 2014). piRNA dependent H3K9me3 deposition occurs at full length LINEs, but not at degraded LINEs in germ cells (Pezic et al. 2014). Zinc-finger proteins bind to specific DNA motifs and co-evolve with TEs (Jacobs et al. 2014). KAP1 also binds more often at relatively young subfamilies of LINE L1 in human and mouse ES cells (Castro-Diaz et al. 2014). The

youngest LINE L1s, however, are not bound by KAP1 either and are silenced via a meCG dependent mechanism (Castro-Diaz et al. 2014).

It has been shown that H3K9me3 is required for repression of retrotransposons in gastrulating *Xenopus* embryos (Herberg et al. 2015). Furthermore, also in frogs small RNAs are involved in repression of TEs by mediating deposition of H3K9me3 and H4K20me3 (Faunes et al. 2012; Harding et al. 2014). However, it is not known which TEs are repressed by histone modifications and what the dynamics are during development. Here we report on the histone modifications present on all TEs of different classes and families using genome-wide ChIP-sequencing data and ChIP-qPCR analysis. We show that specific subsets of TEs in different families and subfamilies are marked by different combinations of heterochromatic histone modifications and that these patterns have different developmental dynamics and are more prominent in gene deserts. We show that early embryonic marking of retrotransposons is linked to small RNAs. Using intra-subfamily divergence as a proxy for age, we show that relatively young DNA transposon subfamilies are marked by early embryonic H4K20me3 and H3K27me3, while at relatively young retrotransposon subfamilies H3K9me3 and H4K20me3 accumulate during development. Our study implicates dynamic repression mechanisms that operate in a developmental stage-specific and a TE-selective fashion.

Materials and methods

Animal procedures

Xenopus tropicalis embryos were obtained by in vitro fertilization. Embryos grew at 23°C in 10% Marc's Modified Ringer's solution (MMR) (88 mM NaCl; 2 mM KCl; 2 mM CaCl₂; 1 mM MgCl₂; 5 mM HEPES, pH 7.4) and were dejellied in 10% MMR + 3% cysteine (pH 7.8-8). Animal use was conducted under the DEC permission (Dutch Animal Experimentation Committee) RU-DEC 2012-116 and 2014-122 to G.J.C.V.

ChIP-qPCR

Embryos were fixed in 1% formaldehyde, methanol-free (Thermo Scientific #28906) for 30 minutes. Formaldehyde was quenched with 125 mM glycine in

25% MMR. Embryos were homogenized in a low-salt buffer (20 mM Tris, pH 8; 70 mM KCl; 1 mM EDTA; 10% glycerol; 5 mM DTT; 0.125% Igepal; cOmplete Protease Inhibitor Cocktail (Roche #04693132001)) (300 embryos/2 mL) and sonicated until DNA fragments had a size of 0.2-2 kb. Yolk was removed by spinning it down. For each ChIP, chromatin extract from 15 embryo equivalents was two-fold diluted with IP buffer (50 mM Tris, pH 8; 100 mM NaCl; 2 mM EDTA; 1 mM DTT; 1% Igepal; Complete Protease Inhibitor Cocktail) for overnight incubation with the antibody: anti-H3K9me2 (Diagenode C15410060, 1 µg); anti-H3K9me3 (Abcam ab8898, 2 µg); anti-H4K20me3 (Abcam ab9053, 2 µg). DNA bound by antibody was captured using 1/10 volume of Dynabeads Protein G during a 1 hour incubation. The beads were washed with ChIP1 buffer (IP buffer + 0.1% deoxycholate), ChIP2 buffer (ChIP1 buffer + 400 mM NaCl), ChIP3 buffer (ChIP1 buffer + 250 mM LiCl), ChIP1 buffer and TE buffer (10 mM Tris-Cl, pH 7.5; 1 mM EDTA). Chromatin was eluted from the beads in 0.1 M NaHCO₃, pH 8.8 + 1% SDS. NaCl (final 25 mM) and 5 µg proteinase K were added for reversal at 65°C. DNA was purified by phenol:chloroform:isoamyl alcohol (25:24:1) extraction. DNA was precipitated by adding 1/10 volume NaOAc (3M, pH 5.2), 2.5 volumes ethanol and glycogen (2 µg/µL) at -20°C. DNA was washed in 70% ethanol and dissolved in TE. qPCR was carried out with iQ Custom SYBR Green Super mix (BioRad) on a CFX96 Real-Time PCR Detection System (BioRad) using an annealing temperature of 60°C. Primers for qPCR (Supplementary table 5) were designed in Primer3 (v.0.4.0) and purchased from Biolegio (Nijmegen, The Netherlands).

Data analysis

Mapping

We used RepeatMasker (version open-4.0.3) to identify repeats in *Xenopus tropicalis* genome JGI7.1 using all frog repeats in the included RepBase repeat library (release 20130422). ChIP data (Hontelez et al. 2015) (read length = 42 bp, Supplementary table 6) was mapped to the reference *Xenopus tropicalis* genome JGI7.1 using BWA version 0.6.1-r104 (Li & Durbin 2009). Duplicates were marked using bamUtil version 1.0.2 (<http://genome.sph.umich.edu/wiki/BamUtil>). RNA-seq reads (Collart et al. 2014) (Ribozero, 4 hpf & 7 hpf, read length = 60 bp) were mapped with GSNAP (version 2012-07-20) (Wu & Nacu 2010). After adapter clipping with fastx_clipper (part of FASTX Toolkit 0.0.13.2) (http://hannonlab.cshl.edu/fastx_toolkit/), small RNA reads (Harding et al. 2014) were mapped with BWA using standard settings (read length in Fig. 4A). MethylC-seq

reads (read length > 105 bp) (Bogdanović et al. 2016) were mapped to in silico bisulfite- converted reference genome JGI7.1 with Bowtie alignment algorithm as described previously (Bogdanović et al. 2016), but with allowing two mismatches in the seed. For this mapping we excluded multi-mappers (reads mapping to multiple locations), after which >90% of all genomic C was covered.

Quantification of reads

Duplicate reads were removed from the ChIP-seq data (samtools view -F 1024). After eliminating duplicate reads (but not multimapped reads) Reads Per Kilobase of transcript per Million mapped reads (RPKM) was calculated for all annotated transposons using peakstats.py (version 2.1) (<http://dx.doi.org/10.5281/zenodo.50023>). Enrichment for the histone modifications was calculated by dividing the RPKM of ChIP-seq data by the RPKM of an input control track (non-ChIPed, stage 9, all reads excluding duplicates). This division over the input control track corrects for inaccuracies in the genome assembly, as repetitive DNA is hard to assemble into larger contigs. The calculation of enrichment values relative to stage 30 input DNA gave similar results, which is expected because embryos of different stages contain the same genomic DNA (Supplementary Fig. 1).

Peakstats.py (using `--remove_dup` option to eliminate duplicated reads) was also used to calculate RPKM for RNA and small RNA at each annotated transposon location.

We calculated the fraction meCG ($((\text{sum of all Cs})/(\text{sum of all C+Ts}))$) for each annotated transposon for which C+T coverage > 4 using bedtools map (version 2.20.1), output sum (Quinlan & Hall 2010).

Next, subfamily annotation according to Repbase was used to calculate median RPKM (for small RNA), average RPKM (RNA), or median enrichment (over input for ChIP or fraction for meCG) for each transposon subfamily (Supplementary table 1). Subfamily sizes are included in Supplementary table 4: 'number of sequences in subfamily library'.

Blast alignment

After aligning the Repbase derived *Xenopus* repeat library to the *Xenopus tropicalis* genome (assembly JGI7.1) we generated subfamily specific sequence

libraries (with a maximum of 1000, evenly distributed along the chromosomes). For each sequence forming part of each of these libraries, we performed a BLAST search against its whole subfamily and summed the result of the obtained bitscore divided by its length $\sum_{i=1}^n (\text{Bitscore}[i]/\text{Length}[i])$. This quantity that scales linearly with the number of homologs of a sequence and its level of similarity serves as an approximation of the recent activity of the gene family.

Generation of plots

Plots were generated in R. PAM clustering of subfamily median enrichment for H4K20me3, H3K9me2/3 and H3K27me3 was done using the package “cluster”, version 2.0.3. Heatmaps of the clustering were visualized with “heatmap.2()” using the “gplots” package, version 2.17.0 (<http://cran.r-project.org/web/packages/gplots/index.html>). All other plots were generated with the package “ggplot2”, version 2.1.0 (<http://ggplot2.org>).

Results

Majority of transposons not decorated by repressive histone modifications

We have generated genome-wide epigenome reference maps, including the constitutive heterochromatin modifications H4K20me3, H3K9me2 and H3K9me3 (Hontelez et al. 2015). These marks were profiled for *Xenopus tropicalis* embryos in developmental stages 9, 10.5, 12.5, 16 and 30, which represent blastula, early gastrula, late gastrula, neurula and organogenesis. We observed that most enrichment for these heterochromatin modifications is found in the non-unique portion of the genome and that they show different dynamics during development (Fig. 1A, B). We validated peaks sets of the ChIP-sequencing data using ChIP-qPCR, and also checked whether quantitative differences in ChIP-sequencing signals across developmental stages could be reproduced by ChIP-qPCR (Supplementary Fig. 2). The heterochromatin-marked repeats can be found at intronic regions (Fig. 1A), as well as in relatively close proximity to genes, for example downstream of *noggin* (Fig. 1B). Given that 35% of the *X.tropicalis* genome consists of TEs and that some TEs obviously are not enriched for heterochromatic histone modifications (Fig. 1A, B), we wondered what the distribution of histone modifications over distinct TEs is.

Transposable elements can be classified into DNA transposons and retrotransposons, with subclasses, families and subfamilies for each (Fig.

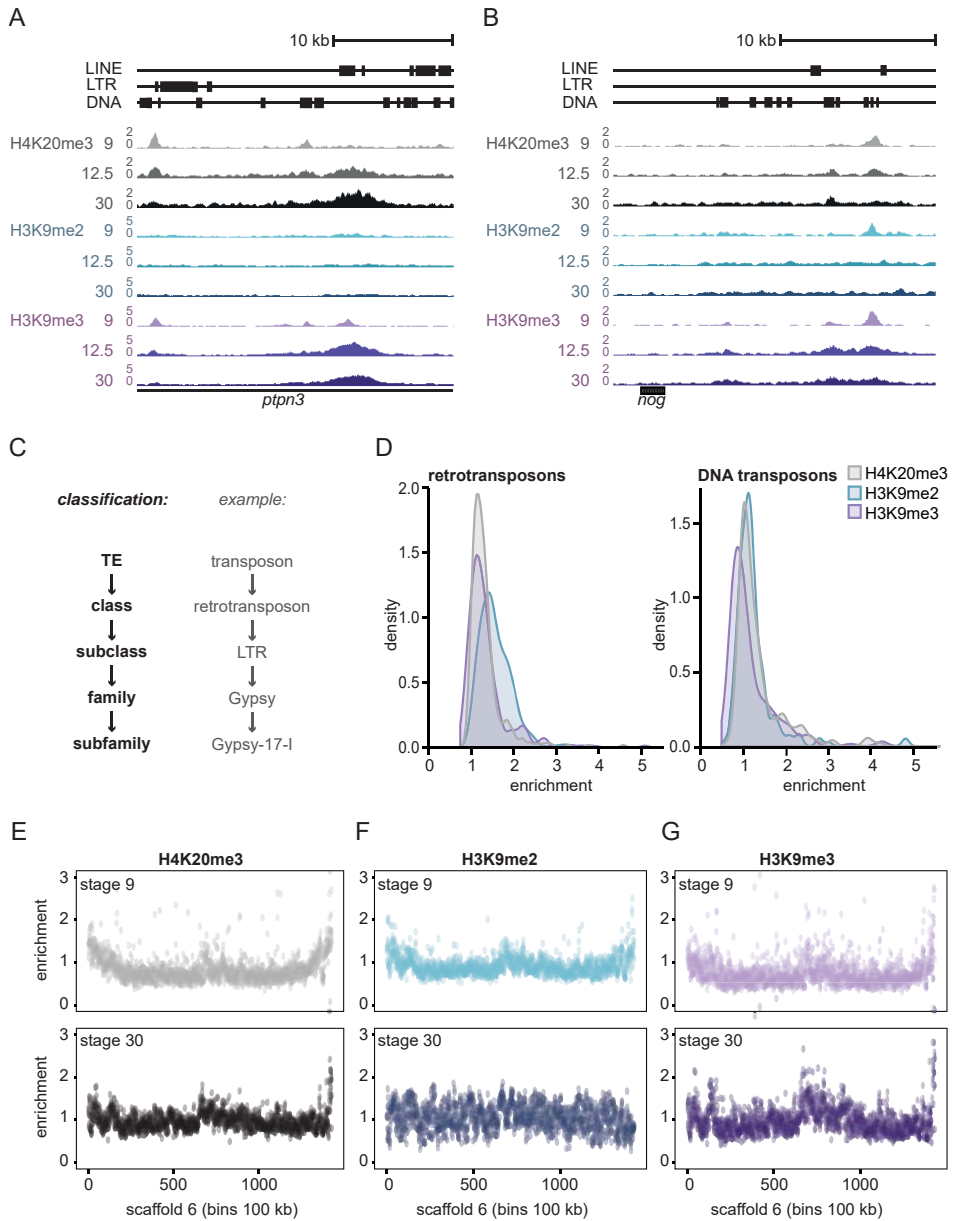


Fig. 1: A subset of transposons acquires H4K20me3, H3K9me2 or H3K9me3. A, B) Genome browser views of the first intron of *ptpn3* (panel A) and downstream of *nog* (panel B), showing H4K20me3 (grey), H3K9me2 (blue) and H3K9me3 (purple) signal (RPKM) at TEs in developmental stages 9, 12.5 and 30 (top to bottom). C) TEs can be classified in DNA and retrotransposons, with subclasses, families and subfamilies. D) Distribution of histone modification enrichment at retrotransposon (left) and DNA transposon (right) subfamilies. Median H4K20me3, H3K9me2 and H3K9me3 enrichment over input was calculated for all TE subfamilies. For each subfamily we used the stage (9, 10.5, 12.5, 16 or 30) with maximal enrichment. E, F, G) Chromosome scale enrichment of histone modifications at Scaffold 6. Histone modification RPKM enrichment over input DNA of E) H4K20me3, F) H3K9me2 and G) H3K9me3 was calculated in bins of 100kb. Top: stage 9. Bottom: stage 30.

1C, Supplementary table 1). In our analysis we included ChIP-sequencing reads that can map to multiple locations to identify the genomic locations of repeats enriched for heterochromatic histone modifications. To identify the TE subfamilies that are enriched for these histone modifications, we calculated the median Reads Per Kilobase per Million mapped reads (RPKM) for each TE subfamily. The enrichment was calculated by dividing the subfamily median RPKM by the median RPKM of the input track (Supplementary table 1, see Methods). Most retro- and DNA transposon subfamilies were not or barely enriched for H4K20me3, H3K9me2 or H3K9me3 (Fig. 1D).

Heterochromatic modifications are expected to be enriched at subtelomeric and pericentric chromatin. This is also observed for most *Xenopus* chromosome-sized scaffolds, showing moderate enrichment (~1.5-2.5-fold) of H4K20me3 and H3K9me3 throughout development and of H3K9me2 at the blastula stage, in large 100 kb bins of genomic sequence (Fig. 1E, F, G, Supplementary Fig. 3).

Dynamics and co-occurrence

To gain insight into the global dynamics and co-occurrence of histone modifications at TEs, a clustering analysis of retro- and DNA transposon subfamilies was performed, using the subfamily median enrichments for H4K20me3, H3K9me2/3 and H3K27me3 (Fig. 2A, B, Supplementary table 2, 3). Furthermore, we also quantified the behavior of the histone modifications for the most strongly enriched retro- and DNA transposons (>2-fold compared to background) (Fig. 2C, D).

The majority of enriched retrotransposons subfamilies gained H3K9me3 and H4K20me3 between blastula (stage 9) and early gastrula (stage 10.5) (Fig. 2A, clusters 1, 2, 4). In contrast, cluster 5 of the DNA transposon subfamilies was strongly enriched mainly for H4K20me3 and H3K27me3 at the blastula stage, but subsequently lost these histone modifications later in development (Fig. 2B). These differences were also reflected in the enrichment of all retrotransposons and DNA transposons that were at least two-fold enriched over input for one of the marks in one stage of development (Fig. 2C, D). Despite the divergent heterochromatic signatures of retro- and DNA transposons, for both TE classes the most strongly enriched clusters also had the lowest transcript levels (Supplementary Fig. 4) (Collart et al. 2014). Their distinct dynamic patterns, which can also be observed at individual loci (Fig. 1A), suggest that retro- and DNA transposons interact differently with the host during the course of embryonic development.

At both retrotransposons and DNA transposons, the levels of H4K20me3 and H3K9me3 were more similar to each other than to H3K9me2 (Fig. 2A, B

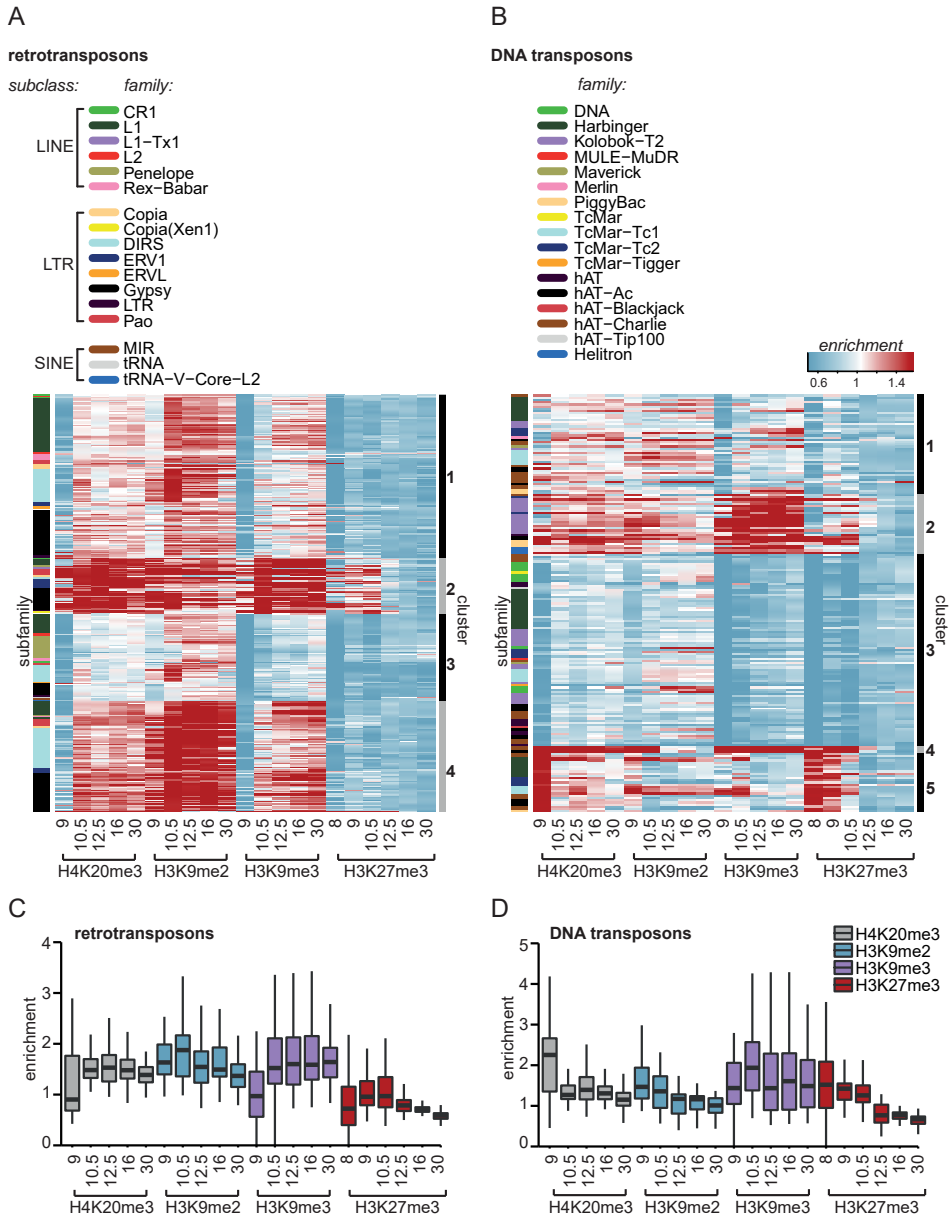


Fig. 2: Variation in repressive histone modification dynamics at retro- and DNA transposons. A) Retro- and B) DNA transposon subfamilies were clustered (PAM, cluster-bar right of heatmap) based on median enrichment over input for H4K20me3, H3K9me2/3 and H3K27me3 during stages 9, 10.5, 12.5, 16 and 30. The family to which a subfamily belongs is depicted in the left side-bar. C, D) The general dynamics of median enrichment of repressive histone marks were plotted for only enriched (>2-fold over input) C) retro- and D) DNA transposon subfamilies. The upper and lower hinges correspond to the 25th and 75th percentiles and the horizontal line in between represents the median.

and Supplementary Fig. 5 and 6). Even though subfamilies decorated with one of the three constitutive heterochromatin marks were often also marked by the other two marks, the strength of the enrichment for the three marks was different. For example, in cluster 1 of retrotransposons H3K9me2 was relatively abundant compared to H4K20me3 and H3K9me3 (Fig. 2A), whereas for the DNA transposons in cluster 2 H3K9me3 marking was relatively strong compared to H4K20me3 and H3K9me2 (Fig. 2B).

Interestingly, the facultative heterochromatin mark H3K27me3 co-occurred with constitutive heterochromatin marks during late blastula and early gastrula (Fig. 2A, B). The correlation of H3K27me3 was highest with H4K20me3 on DNA transposons, but only during these early stages (DNA transposons Spearman rho 0.9 and 0.1, retrotransposons 0.8 and 0.1 at resp. stages 9 and 30). A substantial number of DNA transposon subfamilies was already marked with H3K27me3 at stage 8, before the onset of zygotic genome activation (Fig. 2B), which is different from the H3K27me3 dynamics observed on genic regions, where H3K27me3 starts to accumulate between blastula and gastrula stages (Hontelez et al. 2015; van Heeringen et al. 2014). To examine the occurrence of the other modifications at early blastula stages, we performed ChIP-qPCR for H3K9me2/3 and H4K20me3. At one gypsy and two helitron elements the H3K9me2 modification was found at stage 8, and for one harbinger element H3K9me2/3 and H4K20me3 were already enriched before stage 9; for the most part however, these histone modifications strongly increased between stages 8 and 9 (Supplementary Fig. 7).

Whereas retro- and DNA transposons had different dynamic patterns, the various TE subclasses, families and subfamilies could not be distinguished by heterochromatic histone modification signatures (Fig. 2A, B; subfamily indicated by color code in side-bar). For example, subfamilies belonging to two of the largest retrotransposon families L1 (subclass LINE) and Gypsy (subclass LTR) were assigned to all four clusters (Fig. 2A). Similarly, subfamilies belonging to the same family of DNA transposons were spread over all clusters (Fig. 2B). These data document the variation in heterochromatic histone modifications within TE classes and families.

Gene density and repressive histone modifications

Given the variability in histone marking between TE subfamilies, we wondered whether the genomic environment could be involved in the observed differences. Cis-regulatory elements within TEs can potentially perturb

transcription regulation when located near or within the genes (Casa & Gabellini 2012). Hence, regulation of TEs in a gene-dense context might be different from TEs located in gene-poor loci.

Gene densities (genes/Mbp) were calculated for annotated TEs by counting the number of genes within 1 megabase of the centre of the TE. Median gene density was calculated for all TE subfamilies (Supplementary table 1)

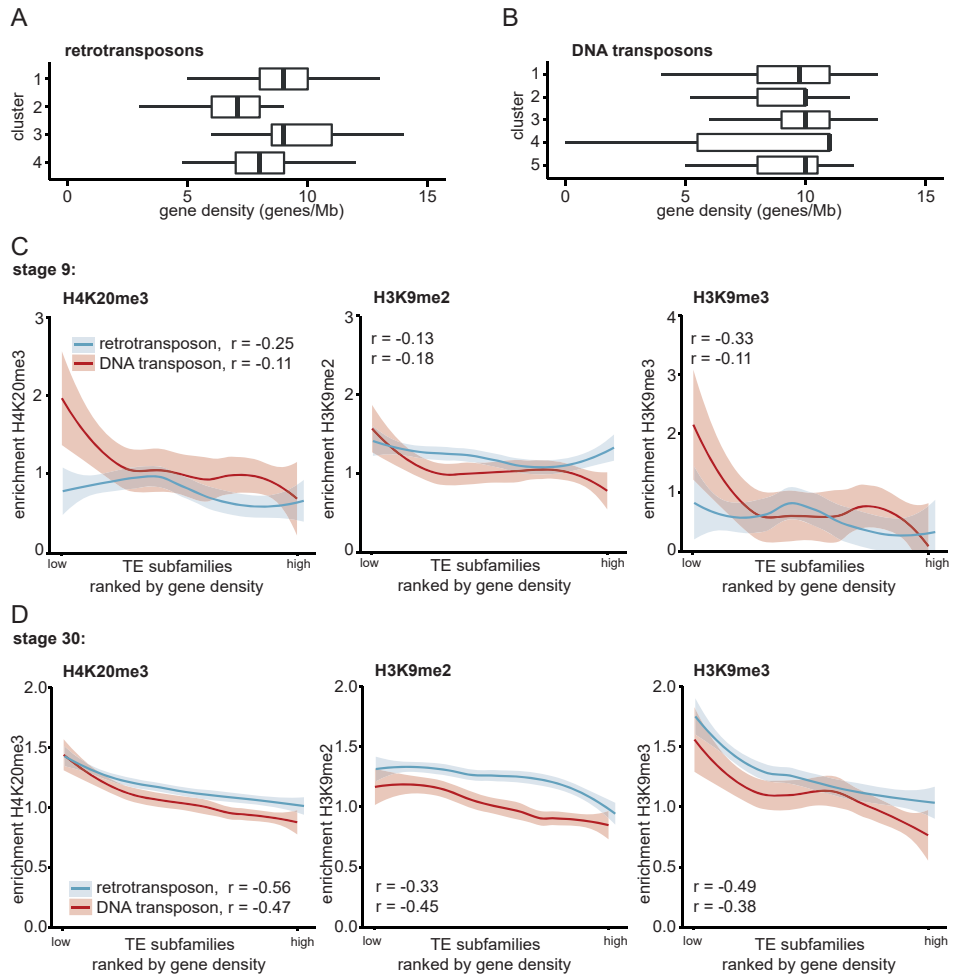


Fig. 3: Strongest enrichment for H4K20me3 and H3K9me3 at transposons in gene desert. For each TE subfamily the median gene density was calculated 500.000 bp up- and downstream of the center of each TE. A) Retro- and B) DNA transposon subfamilies were grouped according to clusters determined in Fig. 2A, B. Median gene densities for the subfamilies were plotted for each cluster. The left and right hinges correspond to the 25th and 75th percentiles and the vertical line in between represents the median. C, D) All TE subfamilies were ranked by gene density and median H4K20me3 (left), H3K9me2 (middle) and H3K9me3 (right) enrichments during stage C) 9 and D) 30 were plotted for retro- (blue) and DNA transposon (red) subfamilies. Loess was used as smoothing method, with a 0.95 confidence interval. Spearman correlation between histone mark and gene density was determined (left corner).

and subfamilies were grouped according to the clusters of Fig. 2. Strikingly, subfamilies from the retrotransposon clusters (clusters 2 and 4) with the highest enrichment for H4K20me3 and H3K9me3 were located at relatively gene-poor regions (Fig. 3A). The gene densities of subfamilies from the DNA transposon clusters, however, were more similar to each other (Fig. 3B).

To probe these relationships further, we analysed the correlation between gene density and repressive histone modifications. TE subfamilies were ranked by gene density and the subfamily median enrichment for the histone modifications in stage 9 and 30 were plotted (Fig. 3C, D respectively). At stage 30 gene density and constitutive heterochromatin marks were inversely correlated for both retro- and DNA transposons, which was more pronounced for H4K20me3 and H3K9me3 compared to H3K9me2 (Fig. 3D). Furthermore, these inverse patterns are less evident earlier in development (Fig. 3C). No inverse patterns with gene density could be observed for the facultative heterochromatin mark H3K27me3, the enhancer-binding protein p300, or the active promoter mark H3K4me3 (Supplementary Fig. 8).

Together these results imply that TEs with heterochromatic marks are more likely to be located in gene deserts, most likely reflecting a TE integration bias (Shen et al. 2013).

H4K20me3, H3K9me3 and small RNAs co-occur at retrotransposons

Upon zygotic genome activation, Polycomb Repressive Complex 2 (PRC2) catalyzes H3K27me3 deposition on genic regions lacking DNA methylation (Hontelez et al. 2015). On the other hand, DNA methylation can stimulate H3K9me3 deposition in genic regions by recruiting histone H3K9 methyltransferase SETDB1 (Matsumura et al. 2015). Small RNAs can also direct deposition of heterochromatin modifications, since it guides protein complexes that recruit H3K9 methyltransferases (Klenov et al. 2011; Wang & Elgin 2011; Sienski et al. 2012; Le Thomas et al. 2013; Rozhkov et al. 2013). We examined to which extent DNA methylation and small RNAs could be involved in depositing heterochromatin modifications at TEs.

We used recently published bisulfite sequencing profiles (Bogdanović et al. 2016) to analyze the DNA methylation status of TEs (Supplementary table 1). Median DNA methylation at stage 9 was calculated for all TE subfamilies, and subfamilies were grouped according to the clusters described in Fig. 2 (Supplementary Fig. 9). High DNA methylation levels were observed for

all subfamily clusters (methylation fraction > 0.9). This indicates that DNA methylation status does not distinguish between TEs that do or do not obtain repressive histone modifications. Furthermore, it indicates that H3K27me3 is not deposited according to DNA methylation logic at TEs, in contrast to what has been observed for genic regions (Hontelez et al. 2015).

To assess the relationship between small RNAs and histone modifications we used a small RNA data set of stage 8, 10 and 18 embryos, size-selected for a range of 18-30 nucleotides (Harding et al. 2014). Small RNAs with a length of 28-29 nucleotides, most likely piRNA (Lim & Kai 2015), were most abundant among small RNA reads mapping to TEs (Fig. 4A, Supplementary Fig. 10). Approximately five-fold more small RNA reads mapped to retrotransposons than to DNA transposons (Fig. 4A), despite the fact that the sequence coverage of DNA transposons in the collection of repetitive elements exceeds that of retrotransposons by a factor of 2.5. Small RNA reads mainly mapped to approximately a quarter of all TE subfamilies (Supplementary Fig. 11, Supplementary table 1). Although small RNA coverage of TE subfamilies changed to some extent during development, their ranking for small RNA coverage was largely stable (Supplementary Fig. 11, Supplementary table 1).

Next, we compared the small RNA data with our repressive histone modification profiles. Median small RNA coverage (RPKM in stage 8) was analyzed for the different clusters of TE subfamilies (cf. Fig. 2). Interestingly, small RNAs were most abundant for the retrotransposon cluster with the strongest enrichment for H4K20me3 and H3K9me3 (Fig. 4B, C). Moreover, the clusters with most small RNA transcripts corresponded with the clusters lowest in gene density (Fig. 3A, B, Fig. 4B, C). To probe these relationships further, we analysed the correlation between small RNA and repressive histone modifications. TE subfamilies were sorted based on abundance of small RNA during stage 8. Subsequently subfamily median enrichments for the histone modifications were plotted. At stage 9 H4K20me3 and H3K9me3 were predominantly present at retrotransposon subfamilies that exhibited relatively strong small RNA coverage (Fig. 4D, E, to the right of red line, median RPKM >0). Retrotransposons with less small RNA did obtain these repressive histone modifications later in development (Fig. 4D, E, left of red vertical line). By contrast, the correlations between H4K20me3 or H3K9me3 and small RNA were less pronounced for DNA transposons (Supplementary Fig. 12). Furthermore, the pattern of H3K9me2 enrichment was different from the other modifications as it was not correlated with small RNA (Supplementary Fig. 12, 13).

We found similar trends when comparing small RNA occurrence with

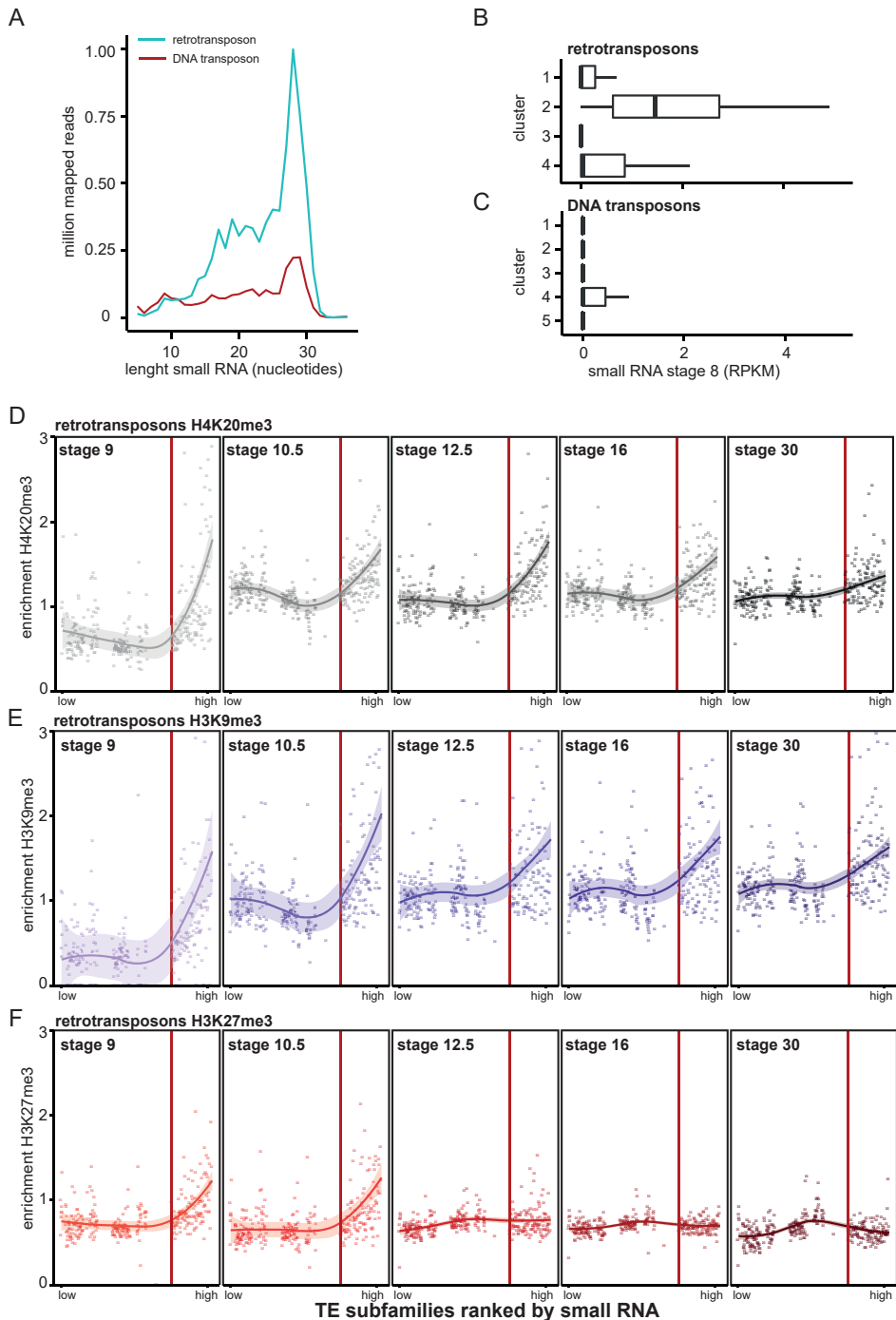


Fig. 4: Small RNA aligns to retrotransposon subfamilies that obtain H4K20me3, H3K9me3 and H3K27me3 first. A) Stage 8 small RNA reads that aligned to retro- (blue) and DNA transposons (red) were split based on nucleotide length and counted. B, C) Median small RNA (RPKM stage 8) was calculated for each TE subfamily

H3K27me3 enrichment at TEs in early developmental stages (Fig. 4F, Supplementary Fig. 12). At stage 9, H3K27me3 is predominantly present at retrotransposon subfamilies that exhibit relatively strong small RNA coverage. However, TE subfamilies earliest enriched for H4K20me3 and H3K9me3 remained marked with these modifications, while H3K27me3 diminished during development.

These results show that the presence of small RNA in the embryo is linked to heterochromatic histone modifications at retrotransposons.

Epigenetic variation within TE families derived from age of TE

TE subfamilies that belong to the same TE family do have variable epigenetic signatures (sidebar Fig. 2A, B). It has been reported that both small RNA-dependent and KAP1-dependent recruitment of histone modifying enzymes occur at relatively young TEs (Pezic et al. 2014; Castro-Diaz et al. 2014). We therefore asked if TE age provides an explanation for the variation in epigenetic regulation within TEs families.

For each subfamily we performed a sequence alignment for all individual TEs within the subfamily. Parasitic elements erode by mutations over evolutionary time. Therefore, we expected young TE subfamilies to have higher alignment scores than old subfamilies. We used the alignment bitscore as a proxy for age (Supplementary table 4). Bitscores were corrected for TE fragment length (Supplementary Fig. 14).

Alignment scores were analyzed for the clusters of TE subfamilies (cf. Fig. 2A, B) (Fig. 5A, B). Among the retrotransposon clusters, the cluster most strongly enriched for H4K20me3 and H3K9me3 (cluster 2) had the highest alignment score (Fig. 5A). This indicates that this cluster contained the youngest retrotransposon subfamilies. Among the DNA transposons, the cluster that increased in H3K9me3 binding (cluster 2) also had a higher alignment score than the clusters that lacked heterochromatic histone modifications (cluster 1, 3) (Fig. 5B). DNA transposon subfamilies with the highest alignment scores were in cluster 5. This indicates that the DNA transposons that lost repressive marks from blastula to tailbud

and plotted according to B) retro- and C) DNA transposon clusters determined in Fig. 2A, B. The left and right hinges correspond to the 25th and 75th percentiles and the vertical line in between represents the median. D, E, F) All TE subfamilies were ranked by amount of small RNA (stage 8) and median D) H4K20me3 E) H3K9me3 and F) H3K27me3 enrichments were visualized for retrotransposons. Histone marks were plotted for stage 9, 10.5, 12.5, 16 and 30 (left to right). The line was plotted using Loess smoothing method, with a 0.95 confidence interval (line shade). All subfamilies left of the red line have a median small RNA RPKM of zero.

(cluster 5) were younger than the DNA transposons that acquired H3K9me3 during these developmental stages (cluster 2).

These results indicate that recruitment of histone modifying enzymes occurs at relatively young retro- and DNA transposons. Furthermore, our study implies that different repression mechanisms operate at young relative to older DNA transposons.

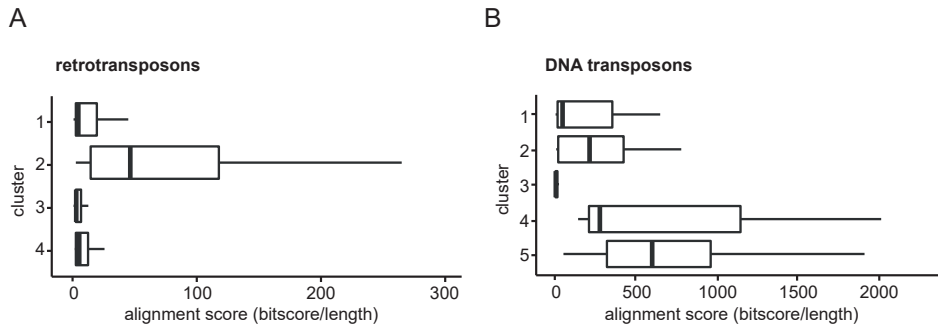


Fig. 5: Transposon clusters enriched with H3K9me3, H4K20me3 and H3K27me3 contain relatively young TEs. A, B) TEs belonging to the same TE subfamily were pairwise aligned to each other using BLAST. The median alignment score was determined for each TE subfamily by dividing the bitscores by the length of the aligned sequences. A) Retro- and B) DNA transposons were grouped according to clusters determined in Fig. 2A, B. The median alignment scores of the subfamilies were plotted for each cluster. The left and right hinges correspond to the 25th and 75th percentiles and the vertical line in between represents the median. The alignment scores of retro- and DNA transposons cannot be directly compared to each other, because of differences in library sizes (Supplementary Fig. 14A, B).

Discussion

TEs can form a threat for the integrity of the host genome (Friedli & Trono 2015). To protect host genome integrity, inserted TEs should be tightly regulated during embryogenesis. In most cases the repression of TEs is independent of genes, but in some cases regulatory sequences of protein-coding genes are regulated by heterochromatic histone modifications. For example, H4K20me3 and the Kmt5b and -5c (Suv4-20h1 and -h2) methyl transferases play a role in the repression of the mouse *Oct4* (*Pou5f1*) and *Xenopus oct25* (*pou5f3.2*) gene and the regulated exit from pluripotency during lineage commitment (Nicetto et al. 2013). In some cases TEs and their active repression may have been co-opted to stably repress genes during development. Our study indicates that TE subfamilies acquiring heterochromatic histone modifications are relatively young. Our data documents the variation in heterochromatic histone modifications within TE classes and families. Not only genomic location and evolutionary age are linked to these differences, but also small RNA abundance.

Epigenetic dynamics

Retrotransposons that gained H3K9me3 during *Xenopus* development also gained H4K20me3. This similar behavior of modifications at retrotransposons is different from what has been observed in mouse embryos, where H3K9me3 appears a few cleavage cycles before H4K20me3 (Burton & Torres-Padilla 2014). Furthermore, H3K9me3 is never completely removed during mouse early development, since de novo H3K9me3 deposition already occurs before parentally inherited H3K9me3 is fully diluted out (Burton & Torres-Padilla 2014). We, however, could barely detect H3K9me3 before zygotic genome activation, which suggests that the mark is being removed more drastically than what has been described for mice. This seems to be true for many histone modifications, which appear to be largely absent before the mid-blastula transition, to be re-established in the blastula or at subsequent stages (Akkers et al. 2009; Hontelez et al. 2015; van Heeringen et al. 2014).

The histone modification differences between frog and mice may originate from the difference in the duration of the embryonic cell division cycles. Before gastrulation, cell cycles in *Xenopus* are completed in 30 minutes, while mouse embryos reach the two cell stage 24 hours after fertilization (O'Farrell et al. 2004). From work with *Drosophila* embryos it is known that histone modifications are re-established on newly replicated chromatin with a delay depending on the histone modification (Petruk et al. 2013; Petruk et al. 2012), so it is possible that the developmental acquisition of these modifications depends on cell cycle lengthening.

The dynamics of heterochromatic histone modifications at retro- and DNA transposons are different. While most retrotransposons in our study gained heterochromatic marks, a substantial subset of DNA transposon subfamilies lost them during development. Interestingly, this inversed dynamics at the two TE classes might be linked to the disparate risk of TE amplification. Retrotransposons use RNA intermediates; therefore amplification of retrotransposons is more likely to take place after zygotic genome activation. DNA transposons, however, can only increase in copy number if their transposition occurs during S phase of the cell cycle. This can occur if a transposon cut from one of the newly synthesised daughter strands is pasted into a region that has not replicated yet, resulting in an additional copy in one of the two daughter cells. Therefore, DNA transposons may form a bigger threat for the integrity of the host genome when genome

replication occurs more frequently, which is before zygotic genome activation. The repressive histone modification dynamics (Fig. 2A, B) may therefore partly reflect amplification risk at different times of development as well as the activity of the mechanisms that generate the defensive histone modification response.

Recruiting histone modifiers

piRNAs bind to Argonaute complexes and can mediate silencing of TEs by recruiting H3K9 methyltransferases SUV3-9 and SETDB1 (Klenov et al. 2011; Wang & Elgin 2011; Sienski et al. 2012; Le Thomas et al. 2013; Rozhkov et al. 2013). Once H3K9me3 is deposited, HP1 can bind to this mark and recruit SUV4-20, which catalyzes H4K20me3 (Jacobs & Khorasanizadeh 2002). In our study 28 bp long small RNAs mostly aligned to retrotransposon clusters (cluster 2 and 4) which were also marked by H3K9me3 and H4K20me3. Therefore, it is likely that during *Xenopus* embryogenesis piRNAs are involved in guiding H3K9 and H4K20 methyltransferases to retrotransposons.

In contrast to the H3K9me3-enriched retrotransposons, the DNA transposon cluster (cluster 2) that gained substantial H3K9me3 marking during development was not enriched for small RNAs. This suggests that DNA transposons recruit H3K9 methyltransferases in a piRNA independent way, for example via KRAB domain-containing zinc-finger proteins, which bind TRIM28/KAP1 and SETDB1 (Friedte et al. 2010).

We found another intriguing set of DNA transposon subfamilies (cluster 5) that was heavily marked with H4K20me3 independent of H3K9me3 deposition. We could not identify a link between small RNAs and these TEs. At ERVs in mouse ES cells, H3K9me3 binding by HP1 protein was important for spreading of H4K20me3, but was dispensable for initial H4K20me3 deposition (Maksakova et al. 2011). So, piRNA and H3K9me3-HP1 independent recruitment mechanisms can guide H4K20 methyltransferases to TEs, for example by long non-coding RNA. Recently, it was shown that SUV4-20H2 is recruited to IAP retrotransposons by long non-coding RNA in quiescent and terminally differentiated cells (Bierhoff et al. 2014). This process was also independent from H3K9me3 and HP1.

Retrotransposons minimally enriched for H3K9me3 often did obtain H3K9me2. SUV3-9 (KMT1A) can catalyze both methylation forms, whereas methyltransferase G9a (EHMT2) can catalyze dimethylation, but not

trimethylation of H3K9 (Rice et al. 2003). Therefore, recruitment of G9a rather than Suv3-9 could result in selective dimethylation of H3K9. Furthermore, the low correlation between piRNA and H3K9me2 indicates that it is unlikely that piRNA-dependent mechanisms (involving Suv3-9 or Setdb1) are involved in the recruitment of H3K9 di-methyltransferases. Additionally, H3K9me2-mediated TE silencing can also be achieved independent of KAP1 at MERVLs in mouse ES cells (Maksakova et al. 2013). Instead of piRNA or KAP1-driven mechanisms, DNA-binding proteins might be involved in guiding H3K9 dimethyltransferases towards TEs. It has been shown, for example, that REST, SNAIL1 and JARID2 recruit G9a/GLP (Shirato et al. 2009; Roopra et al. 2004; Dong et al. 2012).

Concluding remarks

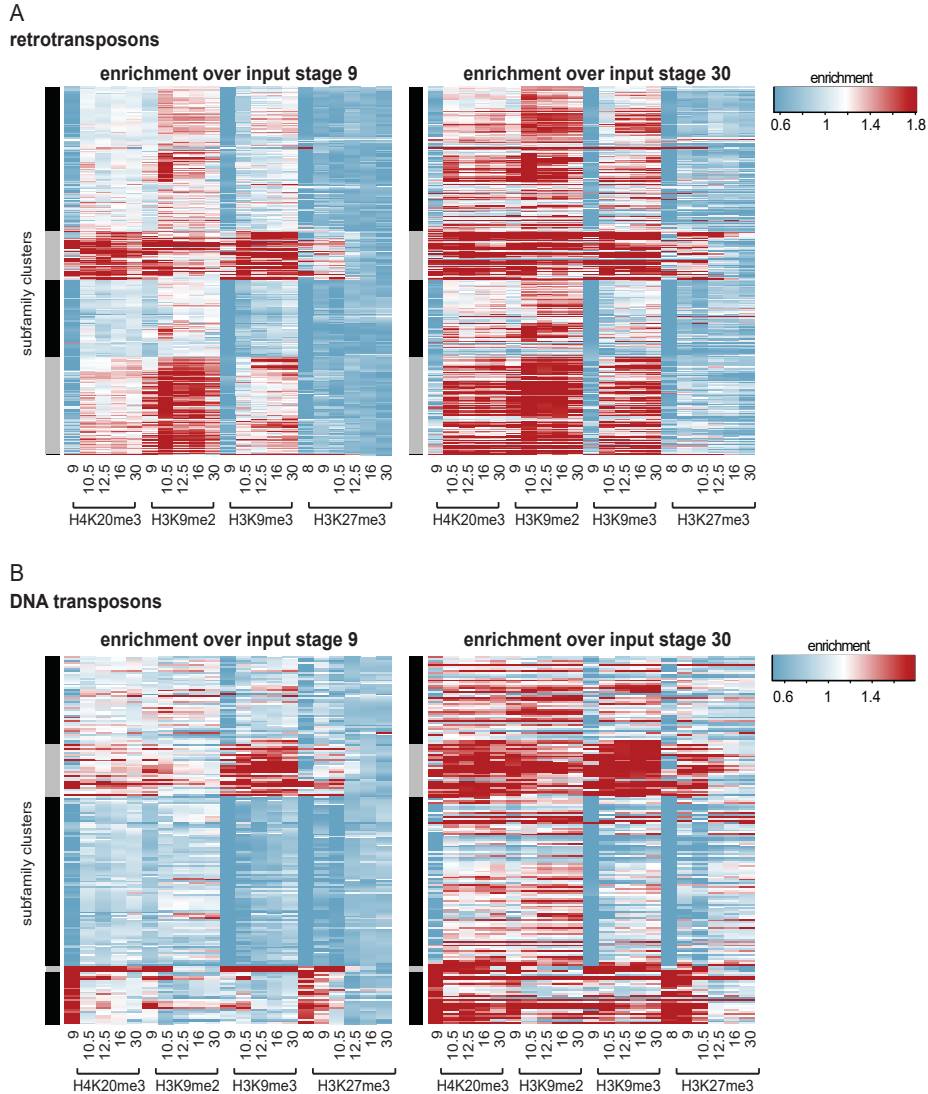
Overall, our study shows that epigenetic regulation is variable between TE subfamilies. The epigenetic variation we described can very well rely on different strategies to recruit histone modifiers. Therefore it will be interesting to decipher the interplay of the factors involved in different recruitment mechanisms like piRNA, KAP1 and DNA-binding proteins.

Acknowledgements

We thank Ozren Bogdanović for his help with the bisulfite-seq analysis. We thank Georgios Georgiou for bioinformatics support. This work has been supported by the US National Institutes of Health (NICHD, grant R01HD069344). S.J.v.H. was supported by the Netherlands Organization for Scientific Research (NWO--ALW) [863.12.002].

Supplementary tables accompany this paper at <https://www.journals.elsevier.com/developmental-biology/>

Supplementary Figures

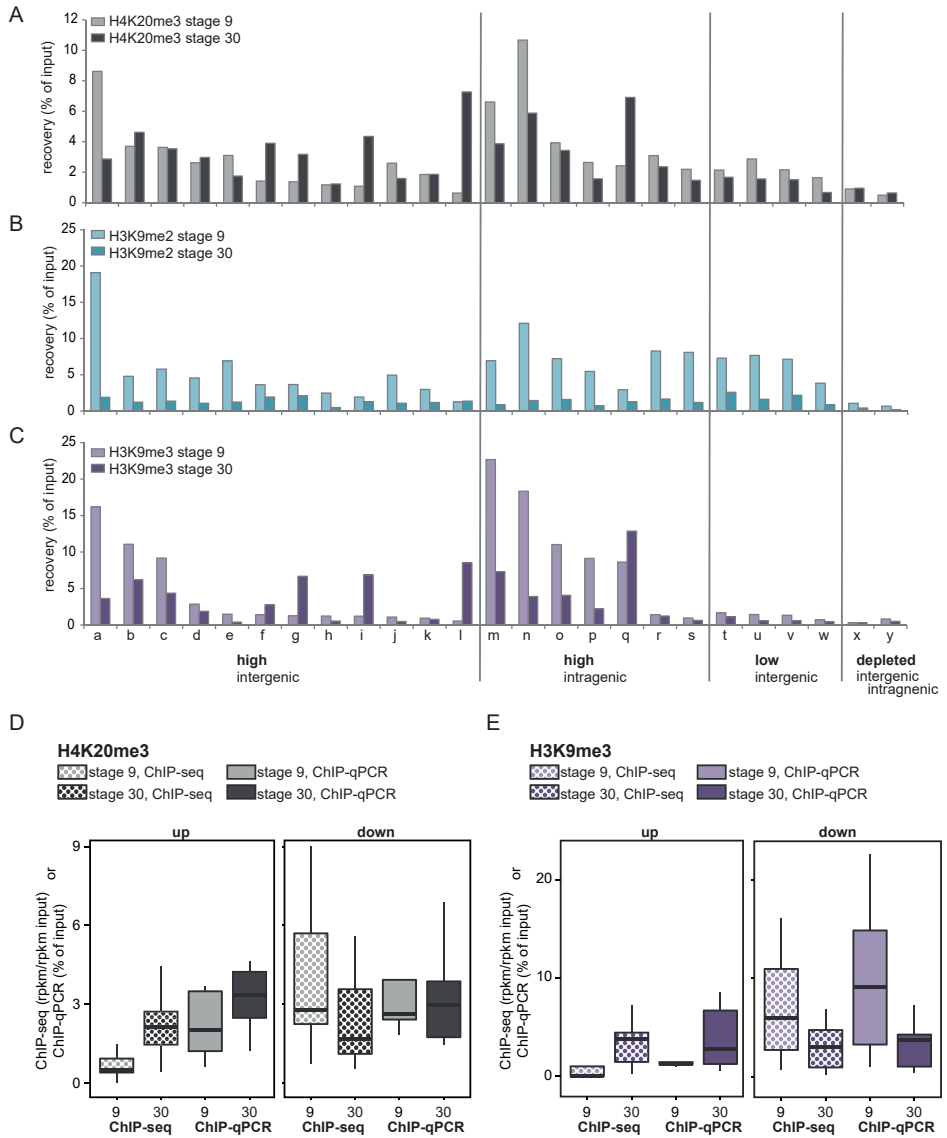


Supplemental Fig. 1: Histone modification enrichment over input control stage 9 vs. over input control stage 30.

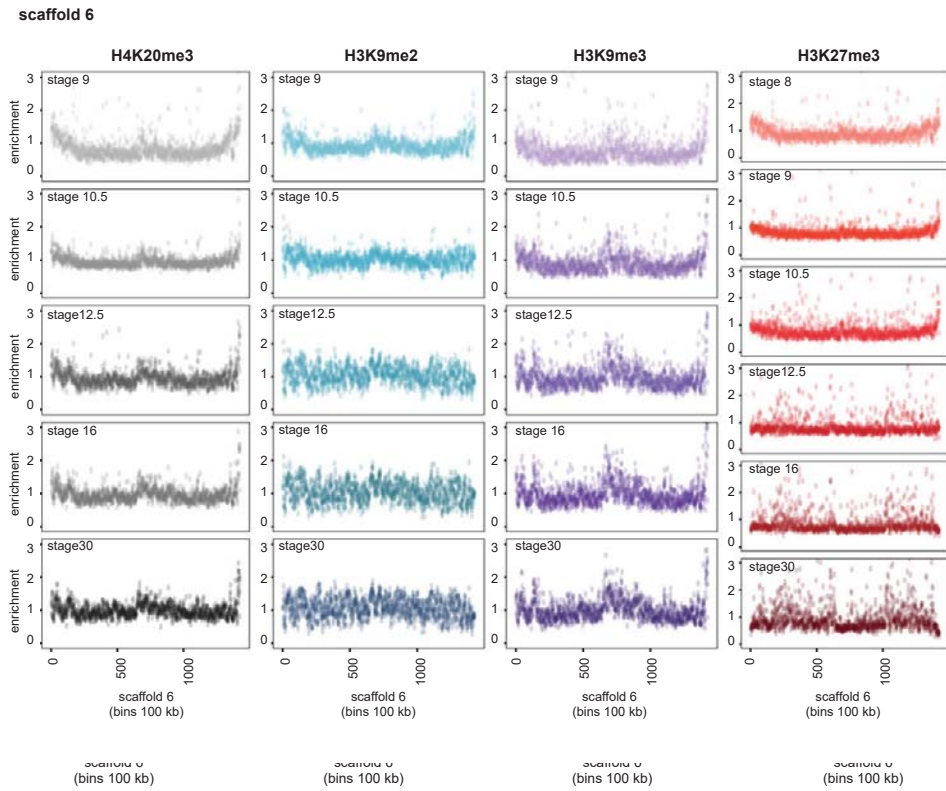
The input DNA was derived from sonicated chromatin, which has been decrosslinked, purified and sequenced. While there could be (subtle) differences in physical properties during sonication depending on chromatin structure specific to a developmental stage, the input sample essentially represents genomic DNA. As embryos of all stages contain the same genomic DNA, the stage should not have a major impact. We verified this comparing an input control processed from stage 9 embryos and from stage 30 embryos.

We plotted transposon subfamily enrichment of heterochromatin modifications over both stage 9 (left) and stage 30 (right) input in the order that was determined by PAM clustering of enrichment over stage 9. The results were very similar; the values were only a bit higher across the board for stage 30. The trends were clearly the same (compare left and right panels for: A) retrotransposons and B) DNA transposons).

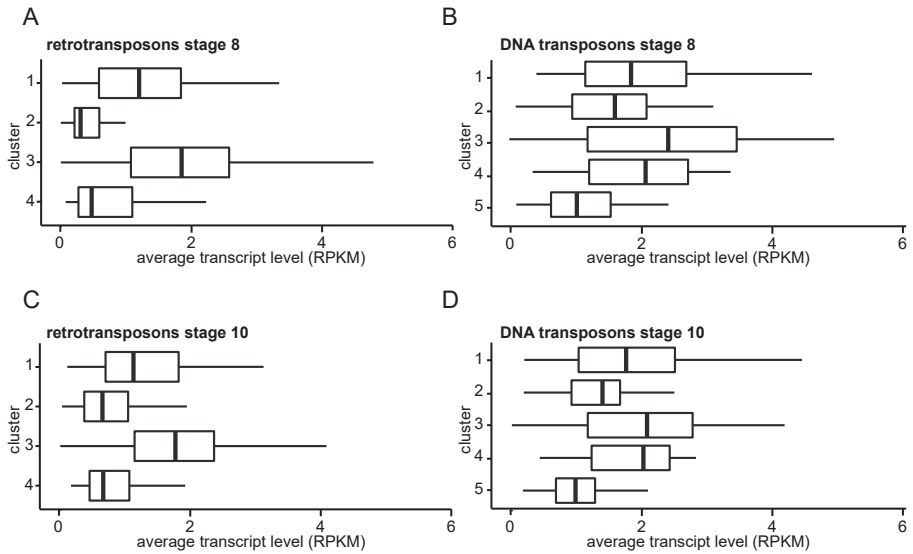
We used a single input controls derived from stage 9 embryos for the final analysis.



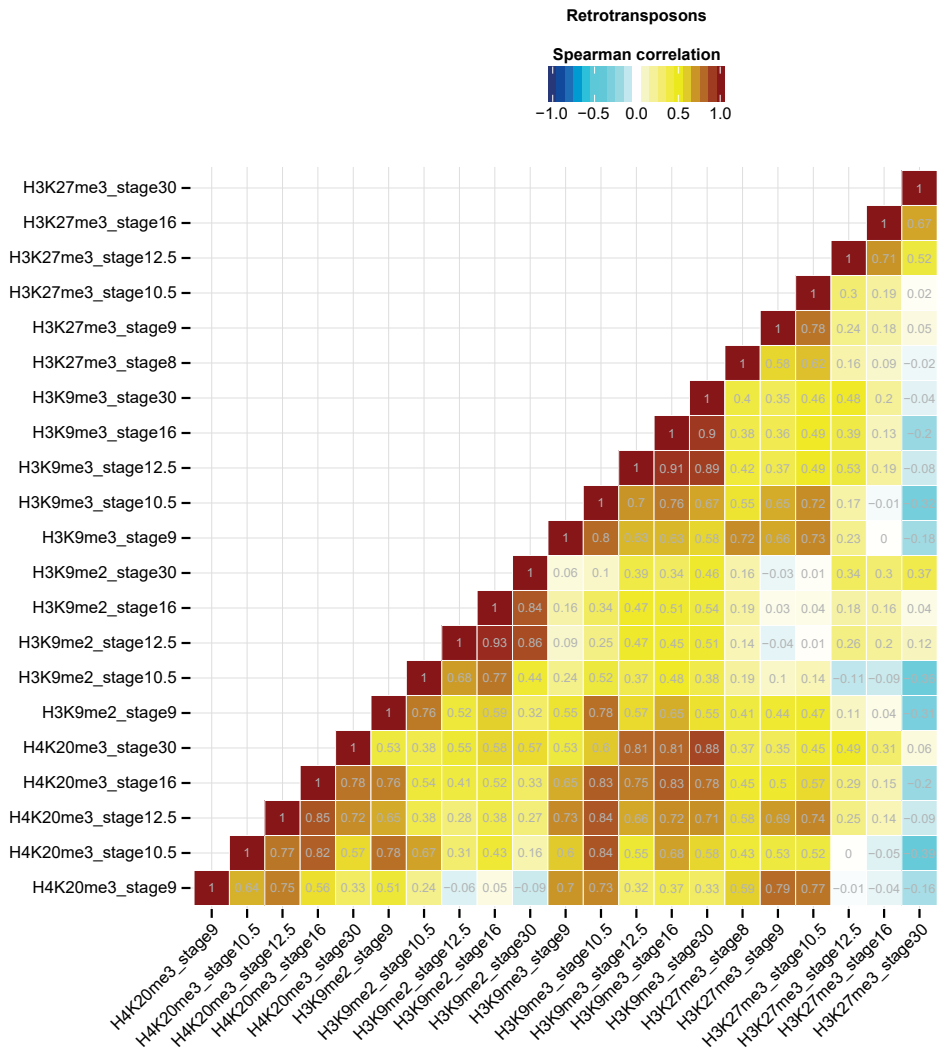
Supplemental Fig. 2: Comparison of ChIP sequencing and ChIP-qPCR. A, B, C) ChIP-qPCR for H4K20me3, H3K9me2 and H3K9me3 was performed at stage 9 and 30. Primers were designed for inter- and intragenic loci which had high, low, or no coverage in ChIP-seq for H4K20me3 and/or H3K9me3 (as indicated on x-axis). Recovery from input (%) was plotted. D, E) Dynamics of H4K20me3 and H3K9me3 between stage 9 and 30 were compared between ChIP-seq and ChIP-qPCR. Locus 'a' to 's' (panel A, C) were grouped by up- or downregulation according to the ChIP-seq data, after which boxplots were generated. Enrichment over input track was plotted for ChIP-seq and recovery of input was plotted for ChIP-qPCR.



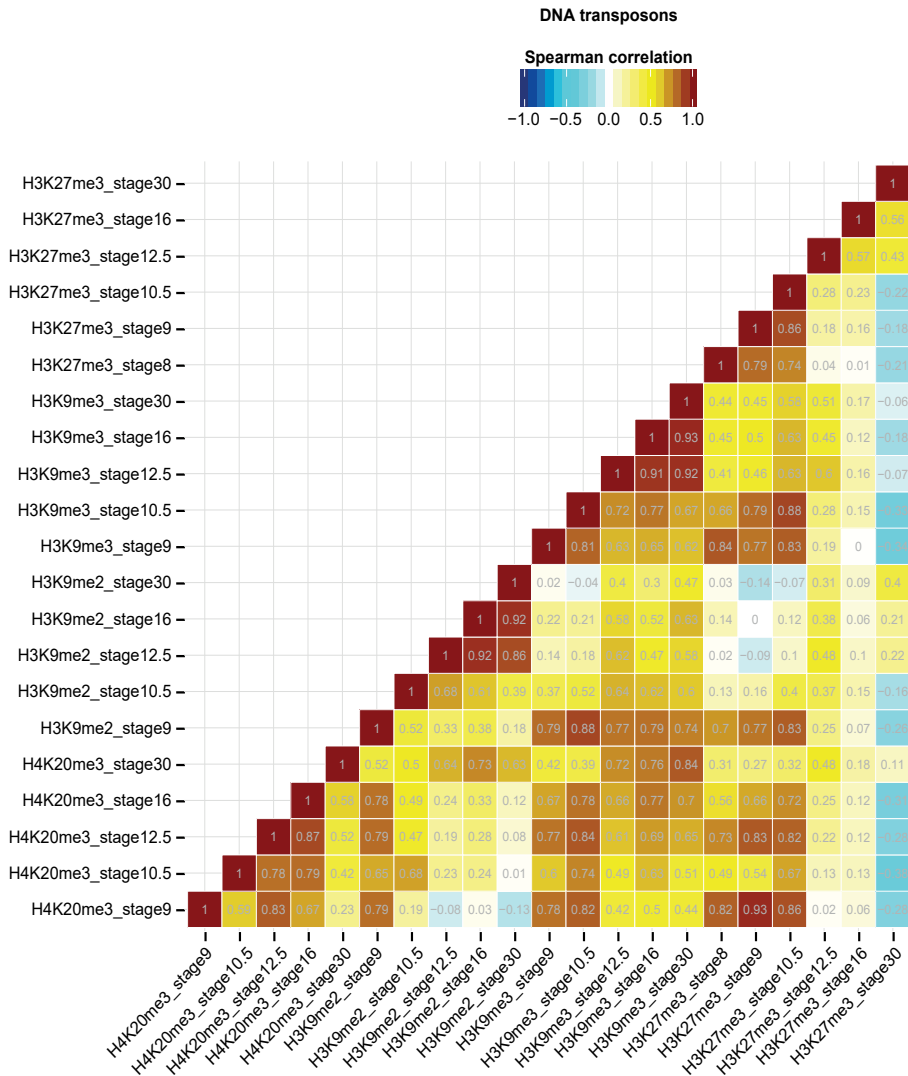
Supplemental Fig. 3: Distribution of heterochromatic histone modifications over scaffold 6. Each scaffold was divided in bins of 100 kb. RPKM of the heterochromatic histone modification was calculated for each bin. Enrichment of H4K20me3, H3K9me2, H3K9me3 and H3K27me3 (left to right) at stage(8,) 9, 10.5, 12.5, 16 and 30 (top to bottom) over the input track was visualized for each bin.



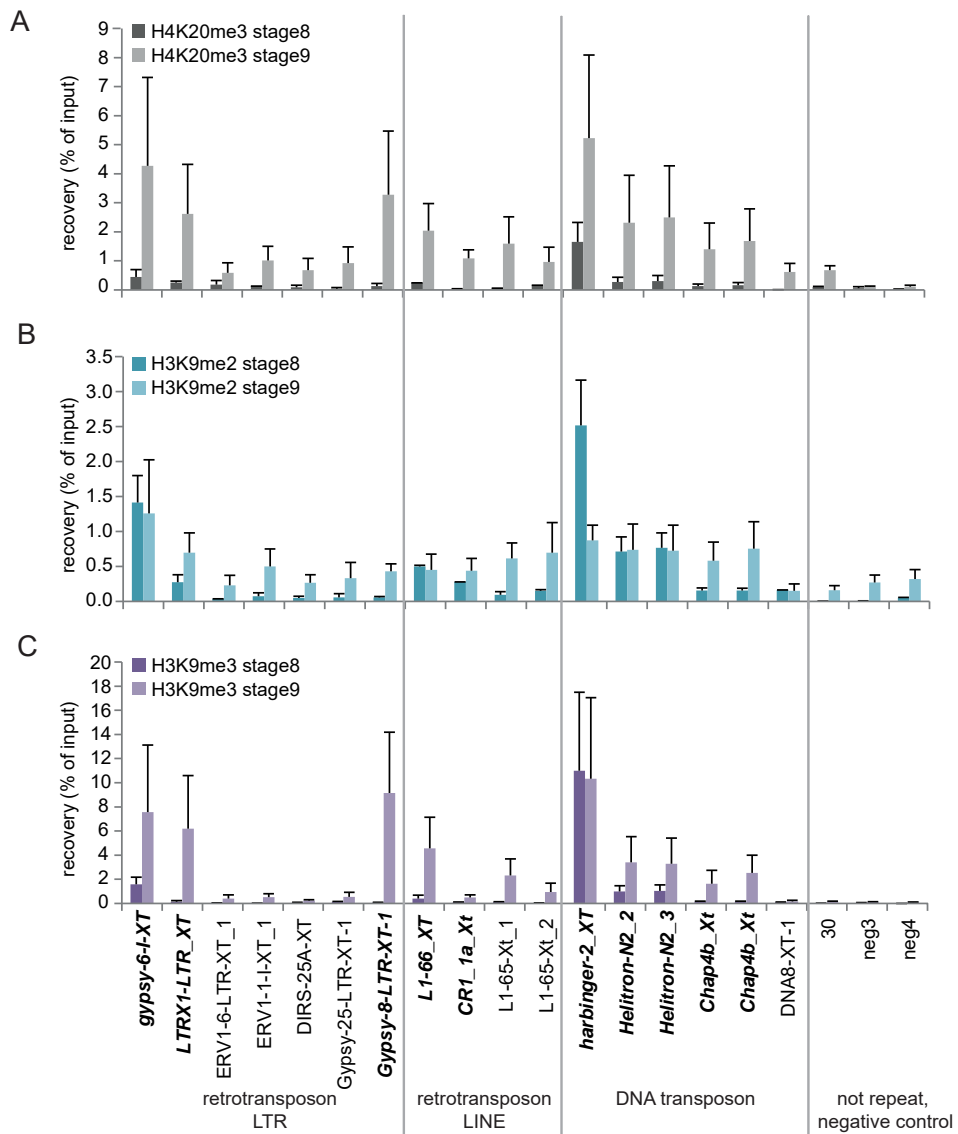
Supplemental Fig. 4: Relatively low transcript levels for transposon clusters with most abundant heterochromatin marking. For each TE subfamily the medium transcript abundance (RPKM) was calculated for embryos in A, B) stage 8 and C, D) stage 10. A, C) Retro- and B, D) DNA transposons were grouped according to clusters as determined in Fig. 2A, B. Average transcript levels for the subfamilies were plotted for each clusters: the left and right hinges correspond to the cluster's 25th and 75th percentiles and the vertical line in between represents the cluster's median RPKM.



Supplemental Fig. 5: Correlation plot of histone modifications at retrotransposons. The median enrichments of each heterochromatic histone modification for each retrotransposon subfamily were used to calculate spearman correlation, as indicated by color and values.



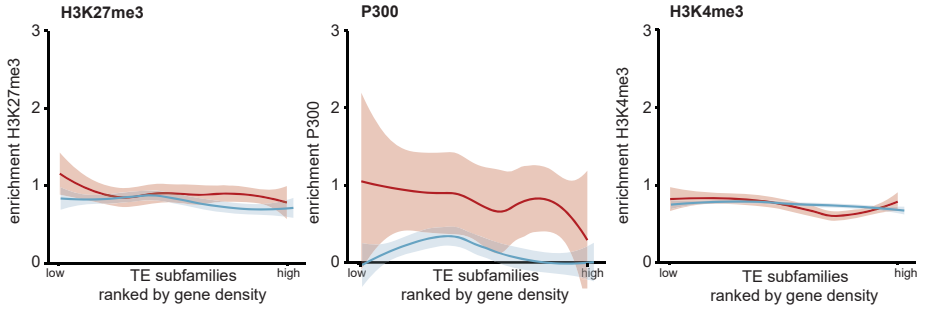
Supplemental Fig. 6: Correlation plot of histone modifications at DNA transposons. The median enrichments of each heterochromatic histone modification for each DNA transposon subfamily were used to calculate spearman correlation, as indicated by color and values.



Supplemental Fig. 7: H4K20me3 and H3K9me3 acquired in stage 9. ChIP against A) H4K20me3, B) H3K9me2 and C) H3K9me3 were performed in stage 8 (dark bars) and in stage 9 (light bars). Recovery of input was analyzed by qPCR for various transposon types, as indicated on the x-axis. Transposons written bold+italic are the individual transposons enriched for at least one heterochromatic mark in the stage 9 ChIP-seq data. n=2, average +SEM.

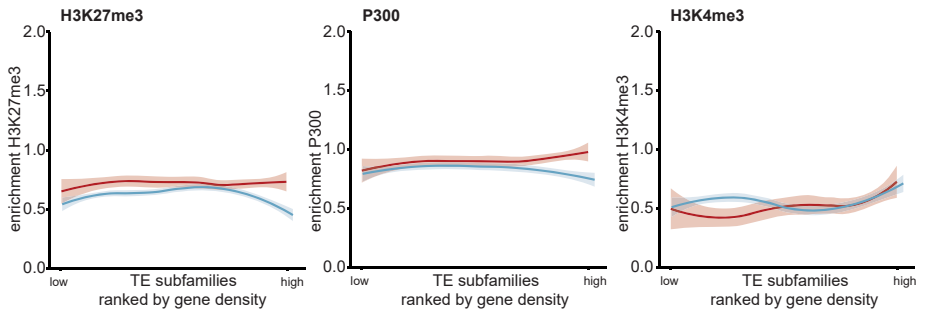
A

stage 9:

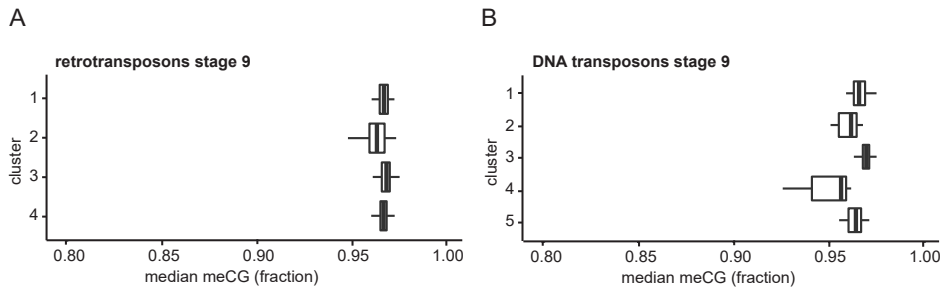


B

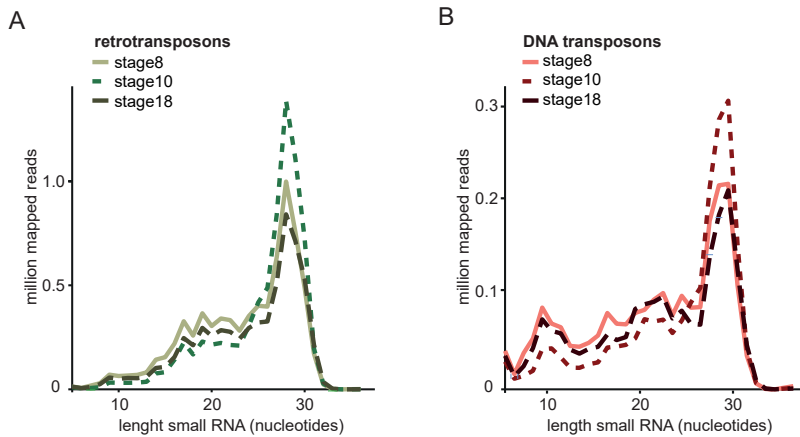
stage 30:



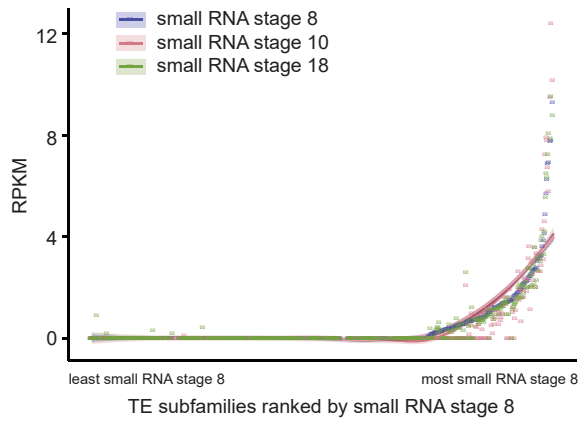
Supplemental Fig. 8: H3K27me3, P300 and H3K4me3 at transposons do not anti-correlate with gene density. For each TE subfamily the median gene density was calculated 500.000 bp up- and downstream of the center of each TE. All TE subfamilies were ranked by gene density and median H4K27me3 (left), P300 (middle) and H3K4me3 (right) enrichments during stage A) 9 and B) 30 were plotted for retro- (blue) and DNA transposon (red) subfamilies. Loess was used as smoothing method, with a 0.95 confidence interval.



Supplemental Fig. 9: Transposons are hypermethylated during blastula stage. Median methylation fraction (=meCG/(meCG+CG)) during stage 9 was calculated for each subfamily. meCG fraction was plotted for A) retro- and B) DNA transposon subfamilies grouped according to clusters determined in Fig. 2A, B. The left and right hinges correspond to the 25th and 75th percentiles and the vertical line in between represents the median.

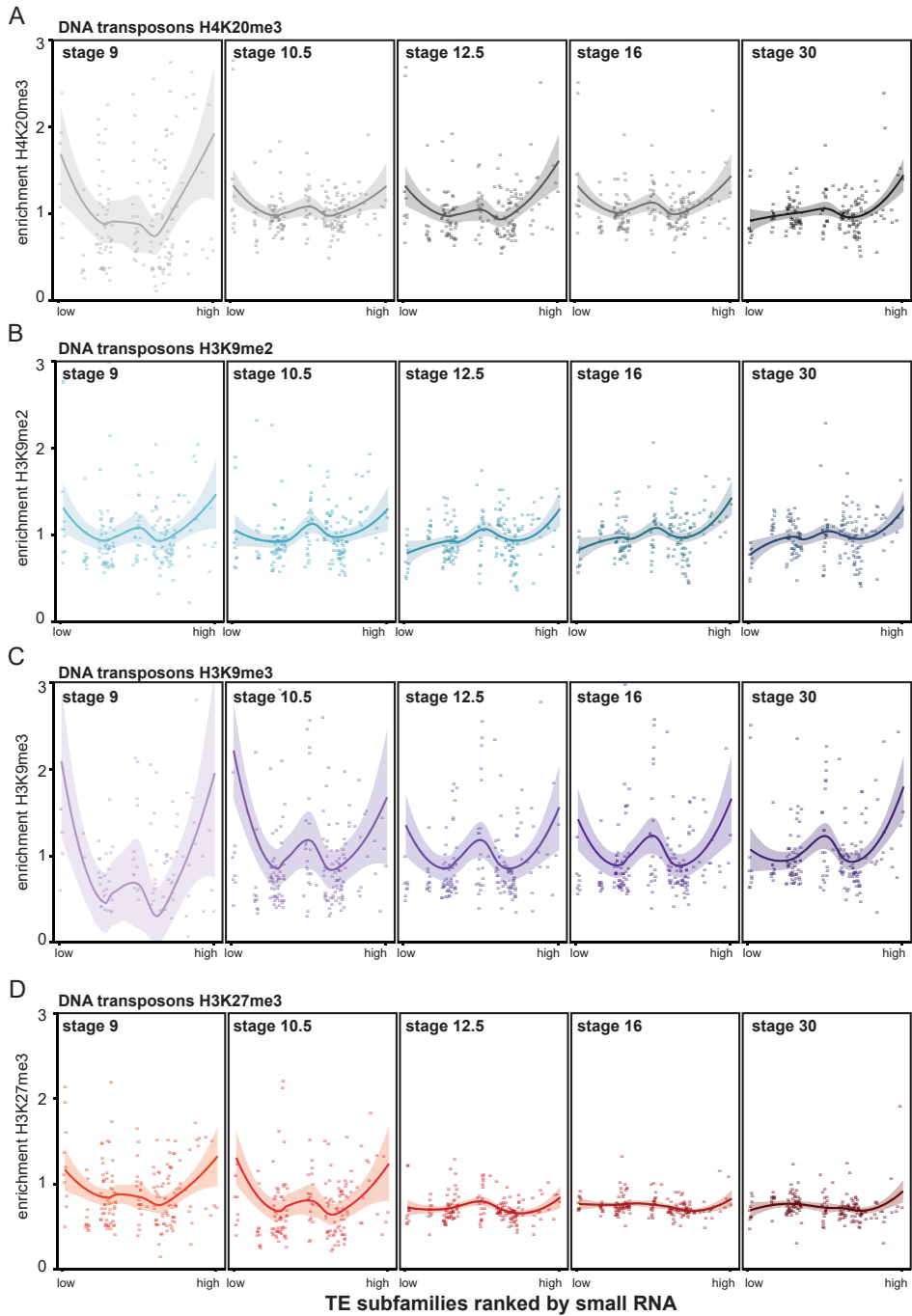


Supplemental Fig.10: The majority of small RNA mapping to transposons is 28 nucleotides long. Small RNA reads (stage 8, 10 and 18) that aligned to A) retro- and B) DNA transposons were split based on nucleotide length (x-axis) and counted (y-axis).

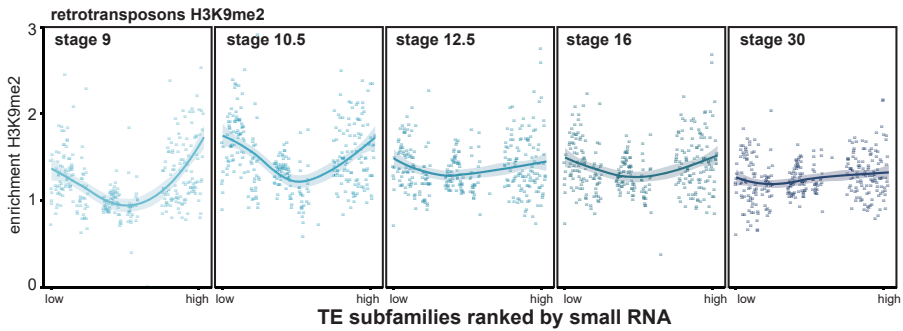


Supplemental Fig.11: The transposons subfamily ranking for small RNA is largely stable during development. Median small RNA levels (RPKM) at stage 8, 10 and 18 were calculated for each TE subfamily. All TE subfamilies were ranked by amount of small RNA at stage 8 and small RNA levels for all three stages were plotted in this order. Loess was used as smoothing method, with a 0.95 confidence interval (line shade).

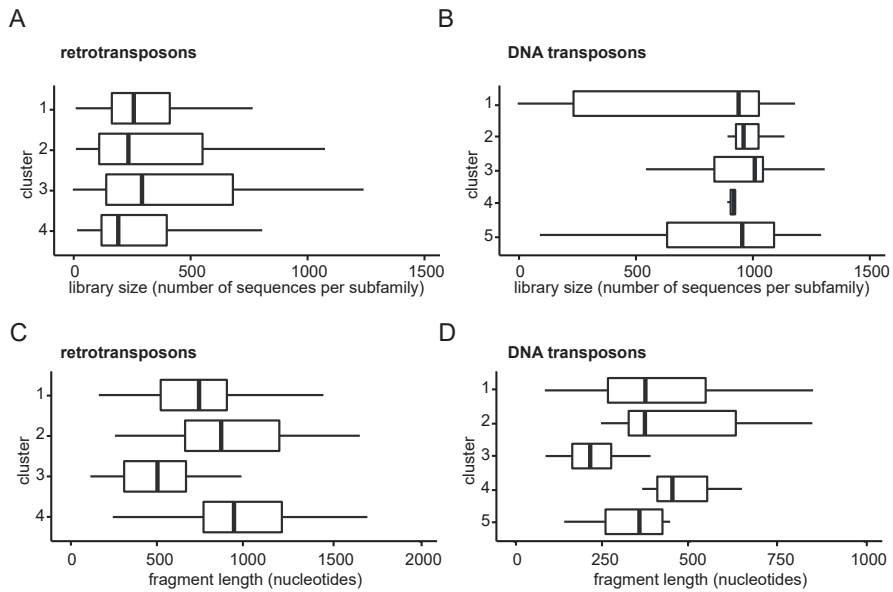
Chromatin regulation in Xenopus embryos



Supplemental Fig. 12: Small RNA does not correlate with heterochromatic histone modifications at DNA transposon subfamilies. All TE subfamilies were ranked by amount of small RNA and in this order median A) H4K20me3, B) H3K9me2, C) H3K9me3 and D) H3K27me3 enrichments were visualized for DNA transposons. Histone marks were plotted for stage 9, 10.5, 12.5, 16 and 30 (left to right). The line was plotted using Loess smoothing method, with a 0.95 confidence interval (line shade).



Supplemental Fig. 13: Small RNA does not correlate with H3K9me2 at retrotransposons. All TE subfamilies were ranked by amount of small RNA and in this order median H3K9me2 enrichment for each retrotransposon subfamily was visualized. Histone marks were plotted for stage 9, 10.5, 12.5, 16 and 30 (left to right). The line was plotted using Loess smoothing method, with a 0.95 confidence interval (line shade).



Supplemental Fig. 14: Library sizes and fragment lengths of retro- and DNA transposon clusters. The individual sequences per TE subfamily were counted. These library sizes were plotted for the A) retro- and B) DNA transposon clusters as determined in Fig. 2A, B. The median length of the sequences belonging to each TE subfamily was determined and these median fragment lengths were also plotted for the C) retro- and D) DNA transposon clusters as determined in Fig. 2A, B. The left and right hinges correspond to the 25th and 75th percentiles and the vertical line in between represents the median.

References

- Akkers, R.C. et al., 2009. A hierarchy of H3K4me3 and H3K2me3 acquisition in spatial gene regulation in *Xenopus* embryos. *Developmental Cell*, 17(3), pp.425–434.
- Bao, W., Kojima, K.K. & Kohany, O., 2015. Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mobile DNA*, 6(1), p.11.
- Bierhoff, H. et al., 2014. Quiescence-induced lncRNAs trigger H4K20 trimethylation and transcriptional silencing. *Molecular cell*, 54(4), pp.675–82.
- Bogdanović, O. et al., 2016. Active DNA demethylation at enhancers during the vertebrate phylotypic period. *Nature genetics*.
- Bulut-Karslioglu, A. et al., 2014. Suv39h-dependent H3K9me3 marks intact retrotransposons and silences LINE elements in mouse embryonic stem cells. *Molecular cell*, 55(2), pp.277–90.
- Burton, A. & Torres-Padilla, M.-E., 2014. Chromatin dynamics in the regulation of cell fate allocation during early embryogenesis. *Nature reviews. Molecular cell biology*, 15(11), pp.723–34.
- Casa, V. & Gabellini, D., 2012. A repetitive elements perspective in Polycomb epigenetics. *Frontiers in genetics*, 3, p.199.
- Castro-Diaz, N. et al., 2014. Evolutionally dynamic L1 regulation in embryonic stem cells. *Genes & development*, 28(13), pp.1397–409.
- Collart, C. et al., 2014. High-resolution analysis of gene activity during the *Xenopus* mid-blastula transition. *Development*, 141(9), pp.1927–39.
- Dong, C. et al., 2012. G9a interacts with Snail and is critical for Snail-mediated E-cadherin repression in human breast cancer. *The Journal of clinical investigation*, 122(4), pp.1469–86.
- Dong, K.B. et al., 2008. DNA methylation in ES cells requires the lysine methyltransferase G9a but not its catalytic activity. *The EMBO journal*, 27(20), pp.2691–701.
- Faunes, F. et al., 2012. Characterization of small RNAs in *Xenopus tropicalis* gastrulae. *Genesis*, 50(3), pp.260–70.
- Fischle, W. et al., 2003. Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes & development*, 17(15), pp.1870–81.
- Friedli, M. & Trono, D., 2015. The Developmental Control of Transposable Elements and the Evolution of Higher Species. *Annual review of cell and developmental biology*.
- Frietze, S. et al., 2010. ZNF274 recruits the histone methyltransferase SETDB1 to the 3' ends of ZNF genes. *PLoS one*, 5(12), p.e15082.
- Grau, J.H. et al., 2014. LTR retroelements are intrinsic components of transcriptional networks in frogs. *BMC genomics*, 15, p.626.
- Harding, J.L. et al., 2014. Small RNA profiling of *Xenopus* embryos reveals novel miRNAs and a new class of small RNAs derived from intronic transposable elements. *Genome research*, 24(1), pp.96–106.
- van Heeringen, S.J. et al., 2014. Principles of nucleation of H3K27 methylation during embryonic development. *Genome research*, 24(3), pp.401–10.
- Hellsten, U. et al., 2010. The genome of the Western clawed frog *Xenopus tropicalis*. *Science*, 328(5978), pp.633–6.
- Herberg, S. et al., 2015. Histone H3 lysine 9 trimethylation is required for suppressing the expression of an embryonically activated retrotransposon in *Xenopus laevis*.
- Hontelez, S. et al., 2015. Embryonic transcription is controlled by maternally defined chromatin state. *Nature Communications*, 6, p.10148.
- Hutnick, L.K. et al., 2010. Repression of retrotransposal elements in mouse embryonic stem cells is primarily mediated by a DNA methylation-independent mechanism. *The Journal of biological chemistry*, 285(27), pp.21082–91.
- Jacobs, F.M.J. et al., 2014. An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons. *Nature*, 516(7530), pp.242–5.
- Jacobs, S.A. & Khorasanizadeh, S., 2002. Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science*, 295(5562), pp.2080–3.
- Kalmykova, A.I., Klenov, M.S. & Gvozdev, V.A., 2005. Argonaute protein Piwi controls mobilization of retrotransposons in the *Drosophila* male germline. *Nucleic acids research*, 33(6), pp.2052–9.
- Karimi, M.M. et al., 2011. DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. *Cell stem cell*, 8(6), pp.676–87.
- Klenov, M.S. et al., 2011. Separation of stem cell maintenance and transposon silencing functions of Piwi protein. *Proceedings of the National Academy of Sciences of the United States of America*, 108(46), pp.18760–5.
- Koonin, E. V., Krupovic, M. & Yutin, N., 2015. Evolution of double-stranded DNA viruses of eukaryotes: from bacteriophages to transposons to giant viruses. *Annals of the New York Academy of Sciences*, 1341,

pp.10–24.

van Kruisbergen, I., Hontelez, S. & Veenstra, G.J.C., 2015. Recruiting polycomb to chromatin. The international journal of biochemistry & cell biology, 67, pp.177–87.

Lander, E.S. et al., 2001. Initial sequencing and analysis of the human genome. Nature, 409(6822), pp.860–921.

Leeb, M. et al., 2010. Polycomb complexes act redundantly to repress genomic repeats and genes. Genes & development, 24(3), pp.265–76.

Levin, H.L. & Moran, J. V., 2011. Dynamic interactions between transposable elements and their hosts. Nature Reviews Genetics, 12(9), pp.615–627.

Li, H. & Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics, 25(14), pp.1754–60.

Lim, R.S.M. & Kai, T., 2015. A piece of the pi(e): the diverse roles of animal piRNAs and their PIWI partners. Seminars in Cell & Developmental Biology, 47–48, pp.17–31.

Lu, X. et al., 2008. The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure. Nature structural & molecular biology, 15(10), pp.1122–4.

Maksakova, I.A. et al., 2013. Distinct roles of KAP1, HP1 and G9a/GLP in silencing of the two-cell-specific retrotransposon MERVL in mouse ES cells. Epigenetics & chromatin, 6(1), p.15.

Maksakova, I.A. et al., 2011. H3K9me3-binding proteins are dispensable for SETDB1/H3K9me3-dependent retroviral silencing. Epigenetics & chromatin, 4(1), p.12.

Martens, J.H.A. et al., 2005. The profile of repeat-associated histone lysine methylation states in the mouse epigenome. The EMBO journal, 24(4), pp.800–12.

Matsui, T. et al., 2010. Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. Nature, 464(7290), pp.927–31.

Matsumura, Y. et al., 2015. H3K4/H3K9me3 Bivalent Chromatin Domains Targeted by Lineage-Specific DNA Methylation Pauses Adipocyte Differentiation. Molecular cell, 60(4), pp.584–96.

Muñoz-López, M. & García-Pérez, J.L., 2010. DNA transposons: nature and applications in genomics. Current genomics, 11(2), pp.115–28.

Nicetto, D. et al., 2013. Suv4-20h histone methyltransferases promote neuroectodermal differentiation by silencing the pluripotency-associated Oct-25 gene. PLoS genetics, 9(1), p.e1003188.

O'Farrell, P.H., Stumpff, J. & Su, T.T., 2004. Embryonic cleavage cycles: how is a mouse like a fly? Current biology: CB, 14(1), pp.R35–45.

Petruk, S. et al., 2013. Stepwise histone modifications are mediated by multiple enzymes that rapidly associate with nascent DNA during replication. Nature communications, 4, p.2841.

Petruk, S. et al., 2012. TrxG and PcG proteins but not methylated histones remain associated with DNA through replication. Cell, 150(5), pp.922–33.

Pezić, D. et al., 2014. piRNA pathway targets active LINE1 elements to establish the repressive H3K9me3 mark in germ cells. Genes & development, 28(13), pp.1410–28.

Quinlan, A.R. & Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics, 26(6), pp.841–2.

Rice, J.C. et al., 2003. Histone Methyltransferases Direct Different Degrees of Methylation to Define Distinct Chromatin Domains. Molecular Cell, 12(6), pp.1591–1598.

Roopra, A. et al., 2004. Localized domains of G9a-mediated histone methylation are required for silencing of neuronal genes. Molecular cell, 14(6), pp.727–38.

Rowe, H.M. et al., 2010. KAP1 controls endogenous retroviruses in embryonic stem cells. Nature, 463(7278), pp.237–40.

Rozhkov, N. V, Hammell, M. & Hannon, G.J., 2013. Multiple roles for Piwi in silencing Drosophila transposons. Genes & development, 27(4), pp.400–12.

Schultz, D.C. et al., 2002. SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. Genes & development, 16(8), pp.919–32.

Shen, J.J. et al., 2013. Genomic dynamics of transposable elements in the western clawed frog (*Silurana tropicalis*). Genome biology and evolution, 5(5), pp.998–1009.

Shirato, H. et al., 2009. A jumonji (Jarid2) protein complex represses cyclin D1 expression by methylation of histone H3-K9. The Journal of biological chemistry, 284(2), pp.733–9.

Sienski, G. et al., 2015. Silencio/CG9754 connects the Piwi-piRNA complex to the cellular heterochromatin machinery. Genes & development, 29(21), pp.2258–71.

Sienski, G., Dönertas, D. & Brennecke, J., 2012. Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. Cell, 151(5), pp.964–80.

Le Thomas, A. et al., 2013. Piwi induces piRNA-guided transcriptional silencing and establishment of a repres-

- sive chromatin state. *Genes & development*, 27(4), pp.390–9.
- Trojer, P. & Reinberg, D., 2007. Facultative heterochromatin: is there a distinctive molecular signature? *Molecular cell*, 28(1), pp.1–13.
- Vagin, V. V et al., 2006. A distinct small RNA pathway silences selfish genetic elements in the germline. *Science*, 313(5785), pp.320–4.
- Varshney, D. et al., 2015. SINE transcription by RNA polymerase III is suppressed by histone methylation but not by DNA methylation. *Nature communications*, 6, p.6569.
- Wang, S.H. & Elgin, S.C.R., 2011. Drosophila Piwi functions downstream of piRNA production mediating a chromatin-based transposon silencing mechanism in female germ line. *Proceedings of the National Academy of Sciences of the United States of America*, 108(52), pp.21164–9.
- Waterston, R.H. et al., 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature*, 420(6915), pp.520–62.
- Wolf, D. & Goff, S.P., 2009. Embryonic stem cells use ZFP809 to silence retroviral DNAs. *Nature*, 458(7242), pp.1201–4.
- Wu, T.D. & Nacu, S., 2010. Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics*, 26(7), pp.873–81.
- Zylicz, J.J. et al., 2015. Chromatin dynamics and the role of G9a in gene regulation and enhancer silencing during early mouse development. *eLife*, 4, p.e09571.

CHAPTER 5

Dynamic Ezh2 and Ep300 recruitment to enhancers during zygotic genome activation, pluripotency and germ layer commitment in *Xenopus tropicalis*

Ila van Kruijsbergen, Saartje Hontelez, Ann Rose Bright, Jin Cho, Ken Cho, Simon J. van Heeringen, Gert Jan C. Veenstra

lvK and GJCV designed the study. CHIP data production was done by lvK and SH, with exception of the Foxh1 ChIP-seq data which were produced by JC and KC. Sequential ChIPs were performed by lvK. ATAC data were produced by ARB. α -amanitin experiments were performed by lvK and SH. All analyses were done by lvK, with support from SJvH. lvK wrote the manuscript with help from GJCV.

Abstract

*Strict coordination of gene regulation is required for successful embryogenesis. Genomic profiles of histone modifications and histone modifiers can provide insight into how this is achieved. Tri-methylation of histone H3 on lysine K27 (H3K27me3) marks transcriptionally silenced genes with functions in developmental regulation. H3K27me3 is deposited by Polycomb Repressive Complex 2 (PRC2) protein Enhancer of zeste homolog 2 (Ezh2). The catalytic activity of PRC2 as well as its recruitment towards genes can be influenced by the transcriptional status of a genic locus. How this relates to PRC2-RNA interactions and the binding of transcription factors is less clear. Ezh2 does not only bind at developmental genes, but also at enhancers marked with acetylation of histone H3 on lysine K27 (H3K27ac) and the Ezh2 antagonist Ep300 which catalyzes this modification. The characteristics of Ezh2-bound enhancers during embryogenesis remain unknown. In this study we characterize Ezh2 binding sites in the blastula and early gastrula stages in *X. tropicalis*. We show that all accessible enhancers are bound by Ezh2 and Ep300 during these early embryonic stages. Furthermore, the binding dynamics of Ezh2 and Ep300 are similar. The Ezh2-bound enhancers are generally H3K27me3-depleted. Moreover, H3K27me3 deposition is often initiated at sites distal from Ezh2 recruitment. We compared Ezh2 and Ep300 recruitment to binding of Sox2 and Foxh1, two transcription factors with binding motifs enriched at Ezh2-bound enhancers. Similar to Ep300, Sox2 binds at the same enhancers as Ezh2, however, with a slight delay. Foxh1 only binds at Ezh2 bound enhancers during the blastula stages. Furthermore, Foxh1 has additional binding sites at less accessible chromatin locations. We also investigated the influence of zygotic transcription on Ezh2 binding and catalytic activity. Zygotic transcription is not required for Ezh2 recruitment to enhancers and is not required to keep enhancers free of H3K27me3 marking. Our study reveals that Ezh2 binding is a general property of accessible enhancers during early embryogenesis in *X. tropicalis*. Our study highlights the complex relationships between regulation of gene expression and the recruitment of activating and repressing proteins to the genome.*

Introduction

Enhancer of zeste homolog 2 (Ezh2) is a protein methyltransferase that is conserved between plants, insects, nematodes and vertebrates (Barnett et al. 2001; Czermin et al. 2002). The *Drosophila* ortholog (*E*)z was identified first (Kalisch and Rasmuson 1974). *Drosophilae* carrying a *zeste* mutation have

repressed eye pigmentation and this phenotype is enhanced in the presence of *E(z)* mutations. Stronger repression of eye pigmentation is not the only morphological change in the *E(z)* mutants. *E(z)* mutations also cause homeotic transformations, which are transformations of one body part into another body part as a result of ectopic expression of segmentation identity genes (Jones and Gelbart 1990). Proteins coded by genes of which mutations result in homeotic transformations are collectively called Polycomb group proteins and therefore *E(z)* has been classified as a PcG protein (Jones and Gelbart 1990).

PcG proteins are important for patterning and cell fate maintenance in vertebrates as well. EZH2 knock-out mice die before completing gastrulation, but conditional EZH2 depletion in mouse embryos showed that the protein is necessary for maintaining *Hox*-cluster gene expression in the developing limb (Wyngaarden et al. 2011; O'Carroll et al. 2001). *Ezh2* is dispensable for gastrulation and tissue specification in *Ezh2*-depleted zebrafish, but also in zebrafish the protein is necessary for tissue maintenance in at least myocardial tissues, liver and pancreas (San et al. 2016).

PcG proteins function in retaining cell fate by depositing epigenetic modifications that preserve gene repression. PcG proteins form two main complexes, Polycomb Repressive Complex (PRC) 1 and PRC2 (Di Croce and Helin 2013). *Ezh2* is the catalytic subunit of PRC2 which catalyzes tri-methylation (me3) of histone H3 on lysine K27 (H3K27) (Kuzmichev et al. 2002). Functional PRC2 contains at least two other core components, Embryonic ectoderm development (EED) and Suppressor of zeste 12 (SUZ12), and can contain optional subunits such as RBAP48/46, PCL1/2/3, AEBP2 and JARID2 (Cao and Zhang 2004; Pasini et al. 2004; Smits et al. 2013). PRC1 contains the catalytic subunit RING1A/B that catalyzes mono-ubiquitination of histone H2A on lysine 119 (H2A119ub) (de Napoles et al. 2004). Transcriptionally inactive regions of the genome interact with each other in the nucleus (Denholtz et al. 2013; Lieberman-Aiden et al. 2009; Joshi et al. 2015; Schoenfelder et al. 2015). This conformation which reduces accessibility for the transcription machinery is stabilized by H3K27me3 and H2A119ub (Aranda et al. 2015). Moreover, PcG proteins might interfere with RNA elongation by antagonizing the phosphorylation of S2 RNAPII (Brookes et al. 2012).

Inactive genes and enhancers with functions in developmental regulation are catalytic targets for PcG proteins (Bernstein et al. 2006; Boyer et al. 2006; Bonn et al. 2012; Rada-Iglesias et al. 2011). PRC2 is recruited to genomic targets via multiple strategies (van Kruijsbergen et al. 2015). Proteins that bind specific DNA sequences, such as TFs and non-methylated CpG binding proteins, can recruit

PcG proteins (Schuettengruber et al. 2014; Benveniste et al. 2014; Farcas et al. 2012). Pre-existing histone modifications can also guide PcG proteins. PRC1 binds H3K27me3 and PRC2 binds H2A119ub, which provides positive allosteric regulation (Blackledge et al. 2014; Cao et al. 2002). Negative allosteric regulation is provided by histone modifications marking active genomic regions, such as H3K27 acetylation (ac) at enhancers and histone H3 lysine K4 (H3K4) me3 at promoters (Petruk et al. 2013; Schmitges et al. 2011). Additionally, PRC2 might sense the transcriptional state of genes via the binding of RNA. The effect of PRC2-RNA interactions on PRC2 association with chromatin is under debate: 1) promiscuous binding of nascent RNA could stimulate recruitment of PRC2 to genes that escaped repression (Davidovich and Cech 2015; Davidovich et al. 2013); 2) RNA binding might not affect Ezh2 binding to chromatin, but inhibit methyltransferase activity (Kaneko et al. 2013; Kaneko, Son, et al. 2014; Kaneko, Bonasio, et al. 2014); 3) PRC2-RNA interaction can antagonize PRC2-chromatin interactions (Beltran et al. 2016). We have previously shown that H3K27me3 deposition at promoters in *Xenopus* embryos is largely unaffected when zygotic transcription is inhibited (Hontelez et al. 2015).

Although the role of EZH2 in mediating transcriptional repression via H3K27me3 deposition at genic regions is well-studied, much less is known about potential other roles of EZH2. Besides H3K27, EZH2 can also methylate non-histone targets such as TFs. Via TF methylation EZH2 acts as a co-activator of STAT3 and AR and as a co-repressor of GATA4 (Xu et al. 2012; He et al. 2012; Kim et al. 2013). Remarkably, Ezh2 binding is very prominent at enhancers in *Xenopus tropicalis* at the blastula stage (van Heeringen et al. 2014).

It remains unclear how repressors (such as Ezh2) and activators (such as TFs and H3K27 acetyltransferase Ep300) interact at enhancers to regulate gene expression. The *X. tropicalis* embryo is an attractive model to study the interplay between Ezh2 and Ep300, since Ezh2 is strongly present at enhancers in this model. In the current study we characterize chromosomal Ezh2 binding sites in early *X. tropicalis* embryos. We show that Ezh2 dynamically binds at all accessible enhancers at the blastula and gastrula stages. Interestingly, H3K27me3 is barely detected at these enhancers, but rather starts to be deposited at Ezh2-free regions adjacent to the Ezh2 binding sites. Zygotic transcription was not required for the binding of Ezh2 at enhancers and neither for the H3K27me3-free status of these enhancers. Furthermore, we show that the binding dynamics of Ezh2 is similar to that of Ep300 and Sox2 at the blastula and gastrula stages, while the Foxh1 binding only partly overlaps the Ezh2 binding at the blastula stages. Our

results highlight the complex relationship between regulation of gene expression and genomic recruitment of activating and repressing proteins.

Materials and methods

Animal procedures

Xenopus tropicalis embryos were obtained by in vitro fertilization (IVF). Males were primed with 100 U human chorionic gonadotropin (hCG) 1-2 days before IVF. Females were primed with 15 U hCG 1-2 days before IVF and boosted with 150 U hCG 5 hours before IVF. Testis were isolated fresh before fertilization. Each testis was crunched in 500 μ L Leibovitz's L-15 medium + 10% fetal calf serum. Eggs were obtained and fertilized by adding 200-500 μ L sperm containing medium. 10% Marc's Modified Ringer's solution (MMR) (88 mM NaCl; 2 mM KCl; 2 mM CaCl₂; 1 mM MgCl₂; 5 mM HEPES, pH 7.4) was added to the fertilized eggs 3 minutes after IVF. The embryos were cultured in 10% MMR at 23°C. The jelly coats were removed in 10% MMR + 3% cysteine (pH 7.8-8). Animal use was conducted under the DEC permission (Dutch Animal Experimentation Committee) RU-DEC 2012-116 and 2014-122 to Gert Jan C. Veenstra.

Fixation and sonication

The embryos were fixed in 1% formaldehyde, methanol-free (Thermo Scientific #28906) for 30 minutes. The formaldehyde was quenched with 125 mM glycine in 25% MMR for 30 minutes. The embryos were washed in 25% MMR for 2x 15 minutes and homogenized on ice in a low-salt buffer (20 mM Tris, pH 8; 70 mM KCl; 1 mM EDTA; 10% glycerol; 5 mM DTT; 0.125% Igepal; cOmplete Protease Inhibitor Cocktail (Roche #04693132001)) (6.67 μ L/embryo). The lysate was sonicated until DNA fragments had a size of 0.2-2 kb. Yolk was removed by spinning it down.

Chromatin Immunoprecipitation (ChIP)

After fixation and sonication the snap-frozen chromatin extract from 15 embryo equivalents was two-fold diluted with IP buffer (50 mM Tris, pH 8; 100 mM NaCl; 2 mM EDTA; 1 mM DTT; 1% Igepal; Complete Protease Inhibitor Cocktail) for overnight incubation at 4°C with the antibody: 5 μ g anti-Ezh2 (Abcam; ab3748; lot GR154179-1), 1 μ g anti-Ep300 C-20 (Santa Cruz Biotechnology; sc-

585; #A0914), 4 μ L anti-Sox2 (R&D Systems; MAB2018; lot KGQ0313121 clone 245610), anti-Foxh1 (custom, Cho lab), anti-Ring1b (Cell Signaling; D22F2 XP(R), batch 03/2014), 5 μ L anti-H3K27me2 (Cell Signaling; D18C8; #9728, lot 09/2014). Antibody bound DNA was captured using 1/10 volume of Dynabeads Protein G during a 1 hour incubation at 4°C. The beads were washed with 900 μ L ChIP1 buffer (IP buffer + 0.1% deoxycholate), ChIP2 buffer (ChIP1 buffer + 400 mM NaCl), ChIP3 buffer (ChIP1 buffer + 250 mM LiCl), ChIP1 buffer and 500 μ L TE buffer (10 mM Tris-Cl, pH 7.5; 1 mM EDTA) subsequently, each for 5 minutes at 4°C. Chromatin was eluted from the beads in 0.1 M NaHCO₃, pH 8.8 + 1% SDS at room temperature. NaCl (final 25 mM) and 5 μ g proteinase K were added for reversal at 65°C. DNA that would be prepared for sequencing was purified with a QIAquick PCR clean-up kit (Qiagen). DNA that would be analysed by qPCR instead of sequencing was purified by phenol:chloroform:isoamyl alcohol (25:24:1) extraction; precipitated by adding 1/10 volume NaOAc (3M, pH 5.2), 2.5 volumes ethanol and glycogen (2 μ g/ μ L) at -20°C; washed in 70% ethanol; and dissolved in TE.

Sequential ChIP

The ChIP protocol was adapted for the sequential ChIP. The antibodies were crosslinked to the beads prior to the first ChIP. Dynabeads Protein G beads were resuspended in 2 volumes PBS (137 mM NaCl; 2.7 mM KCl; 4.3 mM Na₂HPO₄; 1.47 KH₂PO₄; pH 7.4). Prior to crosslinking the antibodies bound to the beads during a 2 hours incubation at room temperature. The antibody-bound beads were washed 0.2 M sodium borate pH 9 and resuspended in 10 volumes 0.2 M sodium borate pH 9 freshly supplemented with 5.2 mg/mL dimethyl pimelimidate. The antibodies crosslinked to the beads during a 30 minutes incubation at room temperature. The reaction was quenched by washing the beads for 1 hour with 10 volumes 0.2 M ethanolamine pH 8. The beads were washed for 30 seconds with 10 volumes glycine pH 3. After a last wash with PBS the first ChIP was started according to the protocol described above. The eluate of the first ChIP was diluted 6 times in IP buffer for the reChIP which was performed as a normal ChIP.

qPCR

qPCR was carried out in reaction volumes of 25 μ L with iQ Custom SYBR Green Super mix (BioRad) on a CFX96 Real-Time PCR Detection System (BioRad) using

an annealing temperature of 60°C. Primers (Table 1) for qPCR were designed in Primer3 (v.0.4.0) and purchased from Biolegio (Nijmegen, The Netherlands).

Table 1: qPCR primers

Name	Forward primer (5'-3')	Reverse primer (5'-3')
bmp7.2	TCTTCTGCCCTCTCCTGTA	TGTTGCTACCCAAAACACCA
cnn1	GCATCTTGCTTGTCTTTGC	GAGGGCGTGTGTGTGTTAGA
pdgfa1	CGGATATCTGTAGGGCGAAA	TCCCCAATAATTTGCTGAA
map3k5	CTCCGCAAAGTCTGCCTTC	ATACTGCCCTGGGAAAGGTT
anp32c	GCAGGGGTTACTGAAAACCA	ATGGGCAGGCTTCCTTAGAT
ttc4	TGATTTTTTACATATGCAGGCA	GTGTACATGCAGGTGGCAGTG
a	GCGAGCGAGAGACAGCTCTA	AAGCTTGATTTGGGGGAGTT
b	CAAATCAATTGGAATGTCGC	GGCTCCACAGTTTCAGAACA
c	TTAGCCTCGGGGTAGTTGAG	GGCTACAAAGTTGCAAAGG
d	CTCCGCAAAGTCTGCCTTC	ATACTGCCCTGGGAAAGGTT
e	GGGACTTTGTGTTGAGCCT	TAGCTGCTCCTCCTCCAAAC
f	TCTTCTGCCCTCTCCTGTA	TGTTGCTACCCAAAACACCA
g	CGGATATCTGTAGGGCGAAA	TCCCCAATAATTTGCTGAA
h	CTAGCGATCTCCACCCTTG	ACCTTTAGCTGGGAGGAAA
i	CACAGTAAAAGTCTGCCTTA	AATTAGGCAGGGCATTGTGT
j	TTTATGCGCCTGCTCCTATT	CTTCAAGTAGCCCCATTGA
k	TTCAGTGCTGGGAGTTTAC	AGTTGGCTTACCTACCGGA
l	TTGACCCGGTTTTATCCTGT	TAAACACCTGCCAGCGTATG
m	TGATTTTTTACATATGCAGGCA	GTGTACATGCAGGTGGCAGTG
n	CTCCTCCCTGGAAGTGCTAC	AGACACCCCTACCTCACAC

Sequencing library preparation

After quality control by qPCR the outputs of the ChIPs from three biological replicates were pooled for library construction, which was started with a maximum of 5 ng DNA. The libraries were prepared with the KAPA Hyper Prep Kit (KAPABiosystems) according to the manufacturers protocol. The DNA fragments were subjected to end repair and A-tailing after which NEXTflex ChIP-seq Barcodes (final concentration 28 nM, Bioo Scientific) were ligated. Agencourt AMPure XP beads were used for a post-ligation clean-up after which the library was amplified by PCR. The PCR products were purified on QIAquick MinElute columns using the QIAquick PCR purification kit (Qiagen). Fragments of 300 bp were selected for sequencing using an E-gel SizeSelect 2% (Invitrogen).

Assay for Transposase Accessible Chromatin (ATAC)

Embryos (~50,000 cells) were lysed in 10 μ L ice cold lysis buffer (10 mM Tris-HCl, pH 7.4; 10 mM NaCl; 3 mM MgCl₂; 0.1% IGEPAL CA-630). Next, 25 μ L 2x TD Buffer (Illumina Cat #FC-121-1030), 2.5 μ L Tn5 Transposase (Illumina Cat #FC-121-1030) and 12.5 μ L water were added to the embryo lysate for the transposition reaction at 37°C for 30 minutes. The DNA was extracted in 10 μ L Elution buffer (10 mM Tris, pH 8) using a MinElute Kit (Qiagen). The fragmented DNA was amplified in a PCR reaction by mixing the DNA with 9.7 μ L water, 2.5 μ L 25 μ M Customized Nextera PCR Primer 1, 2.5 μ L 25 μ M Customized Nextera PCR Primer2, 0.3 μ L 100x SYBR Green I (Invitrogen Cat #S-7563), and 25 μ L NEBNext High-Fidelity 2x PCR Master Mix (New England Labs Cat #M0541) and running the following program: 72°C for 5 minutes; 98°C for 30 seconds; 5x (98°C for 10 seconds + 63°C for 30 seconds + 72°C for 1 minute); hold at 4°C until qPCR was performed to determine how many more amplification cycles can be performed without saturating the reaction. Therefore, 5 μ L from the PCR mix that had been amplified for five times was mixed with: 4.44 μ L water, 0.25 μ L 25 μ M Customized Nextera PCR primer 1, 0.25 μ L 25 μ M Customized Nextera PCR primer 2, 0.06 μ L 100x SYBR Green I, and 5 μ L NEBNext High-Fidelity 2x PCR Master Mix. The qPCR ran: 98°C for 30 seconds; 20x (98°C for 10 seconds + 63°C for 30 seconds + 72°C for 1 minute). After determining how many more amplification cycles could be run without saturating the PCR reaction the remaining 45 μ L from the PCR mix that had been amplified for five times ran the additional cycles: x times (98°C for 10 seconds + 63°C for 30 seconds + 72°C for 1 minute). The DNA library was purified with a QIAquick PCR clean-up kit (Qiagen).

Data analysis

Mapping

ChIP data was mapped to the *Xenopus tropicalis* genome JGI7.1 using BWA (version 0.6.1-104) (Li and Durbin 2009). Duplicate reads and reads mapping to multiple loci were removed (samtools (version 1.2) view -F 1024; grep (version 2.21) XT:A:U) (Li et al. 2009; Kernighan and Pike 1984).

Peak calling

Peaks were called relative to an input control track (stage 9 embryos, fixed and sonicated). We used macs2 (version 2.0.10.20130306) (Zhang et al. 2008).

The broad (--broad) setting was used for all data except for the ATAC data. The q-value (-q) used for the H3K27me3 data was 0.0005; for the ATAC data 0.001; for the Foxh1, Ep300 and Sox2 data 0.01; and for all other data 0.05.

We merged peaks (in different peak sets) that were maximally 10 bp apart with bedtools merge (version v2.20.1) (-d 10). Bedtools intersect (version v2.20.1) was used to calculate overlap between peak sets and bedtools subtract (version v2.20.1) was used to subtract peak sets (Quinlan and Hall 2010).

Quantification of reads

ChIP-seq data was quantified by calculating Reads Per Kilobase per Million mapped reads (RPKM) using peakstats.py (version 2.1) (options: -b 1 --rpkm --remove_dup --unique) (<http://dx.doi.org/10.5281/zenodo.50023>). For the analysis in Fig. 3 RPKM was calculated differently as described below in 'Heatmap (2)'.

Bisulfite-seq data was quantified using bedtools map (version v2.20.1). We calculated the fraction meCG ((sum of all Cs)/(sum of all C+Ts)) for Ezh2 peaks for which C+T coverage > 4 using output sum (Quinlan and Hall 2010).

Gene ontology analysis

Gene ontology analysis was performed with DAVID Bioinformatics Resources 6.7 (Huang et al. 2009a; Huang et al. 2009b). The gene regulatory domains were determined by the 'basal plus extension' method as used by the GREAT tool (McLean et al. 2010). The gene was extended 5 kb upstream and 1 kb downstream of the TSS, after which it was extended further up to the nearest gene with a maximum of 1 Mb in both directions.

Motif analysis

GimmeMotifs (van Heeringen and Veenstra 2011) (v0.8.6) was used to identify which motifs from the Weirauch motif database (Weirauch et al. 2014) were enriched in the sequences of each cluster. First, background files were created with bedtools (v2.20.1) getfasta and shuffle. These background files were used to calculate the threshold (FDR 0.01) for each motif in the database using gimme threshold. The motif enrichment was finally calculated for the sequences within

the Ezh2 clusters by gimme scan using these motif specific threshold values.

Generation of plots

Heatmap (1)

Clustering and heatmaps (Fig. 1, 5) were generated using fluff heatmap v2.0.1 (Georgiou and van Heeringen 2016). We used the method (-C) kmeans for the Ezh2 data and the method (-C) none for all other data for which we used the Ezh2-cluster file as an input. We used the option -r to plot RPKM.

Heatmap (2)

Clustering of gene regulatory domains (as described in ‘Gene ontology analysis’) (Fig. 3) was done with heatmap.2() using R (version 3.2.0) package gplots (version 2.17.0) (R Development Core Team 2008) (<http://cran.r-project.org/web/packages/gplots/index.html>).

We divided the gene regulatory domains over three groups, namely 1) stage 8 H3K27me3 peaks (ENS), 2) stage 8, 9 and 10.5 H3K27me3 peaks merged and 3) the H3K27me3-free surrounding region by performing the bedtools sub-commands as described above in ‘Peak calling’. The regions that obtained an Ezh2 peaks during stage 8, 9, and/or 10.5 were selected within these three groups and H3K27me3 and Ezh2 RPKM were calculated at these loci. All aligned Ezh2 and H3K27me3 reads were counted at the Ezh2 peak locations within the three groups using bedtools multicov (version v2.20.1). The length covered by the peak locations within the three groups was calculated using GNU Awk (version 4.0.2). RPKM was calculated by dividing [the aligned reads at the peak locations] by [(the length of the peak locations divided by 1000) multiplied by [the total number of mapped reads of a ChIP-seq sample multiplied by 1000000]].

We clustered all regulatory domains that had an H3K27me3 peak at stage 8 and at which H3K27me3 deposition increased towards stage 10.5 with heatmap.2(). We performed Pearson clustering of the gene regulatory domains based on the Ezh2 binding using an agglomerative method by applying the options “distfun=function(x) as.dist((1-cor(t(x)))/2)” and “hclust=function(x) hclust(x,method=“ward.D2)””.

Correlation plot

A correlation plot was generated in R (version 3.2.0) with the package `corrplot` for all motifs that were at least five times enriched over background in at least one of the clusters (Murdoch and Chow 1996; Friendly 2002).

Box plots, bar plots, density plots

All other plots were generated in R (version 3.2.0) with the package `ggplot2` (version 2.1.0) (Wickham 2009).

Results

Ezh2 binds at active, accessible enhancers

In this study we aimed to characterize chromosomal Ezh2 binding sites during early developmental stages in *X. tropicalis*. Therefore we generated genome-wide Ezh2 binding profiles for *X. tropicalis* embryos in blastula (stage 8 and 9) and gastrula (stage 10.5) stages. This developmental window comprises the stages of zygotic genome activation, pluripotency and the formation of the three germ layers. The catalytic activity of Ezh2 is highly dynamic during this period; H3K27me3 deposition increases strongly at genic regions from late blastula onwards (Hontelez et al. 2015; van Heeringen et al. 2014). Furthermore, enhancer activity is also highly dynamic during blastula and gastrula stages (Hontelez et al. 2015).

We identified 31.602 Ezh2 binding sites in stage 8, 9 and 10.5 embryos using the MACS2 peak calling algorithm. These Ezh2 binding sites were also enriched for PRC2-associated protein Jarid2 and the PRC1 subunit Ring1b (Supplementary Fig. 1A), suggesting that both polycomb repressive complexes are recruited to these sequences. Clustering analysis (Euclidean, k-means) identified sequences with different Ezh2 binding dynamics (Fig. 1A: clusters 1-4). One of the clusters (cluster 1) contained loci that were marked by Ezh2 in stage 8 but gradually lost Ezh2 during stage 9 and 10.5. In contrast, Ezh2 binding increased during these stages in clusters 2, 3 and 4. During stage 9 and 10.5 Ezh2 levels were highest in cluster 3 (average RPKM 9 in cluster 3 compared to RPKM 4 and 5 in clusters 2 and 4). Ezh2 binding site in cluster 4 differed from other binding sites, because of the presence of groups of adjacent peaks (examples in Fig. 1B).

We analyzed Ep300-occupancy at the Ezh2 binding sites to study if there was a relation between Ezh2 binding and enhancer activity. We could hardly

detect Ep300 binding at stage 8. The majority (87%) of all Ep300 binding sites identified at stage 9 and 10.5 using MACS2 were also present in our Ezh2 peak-set. Furthermore, Ep300 showed the same binding pattern as Ezh2 when ordered according to Ezh2 clustering (Fig. 1A). At stage 8 Ep300 binding was most abundant in cluster 1, similar to Ezh2 binding at this stage. Additionally this cluster also lost Ep300 binding at stage 10.5. The cluster with the strongest Ezh2 recruitment in stage 9 and 10.5 (cluster 3) was the cluster with the strongest Ep300 binding during these stages as well. Furthermore, cluster 4 sequences generally featured multiple Ep300 binding sites, similar to Ezh2 peaks (examples

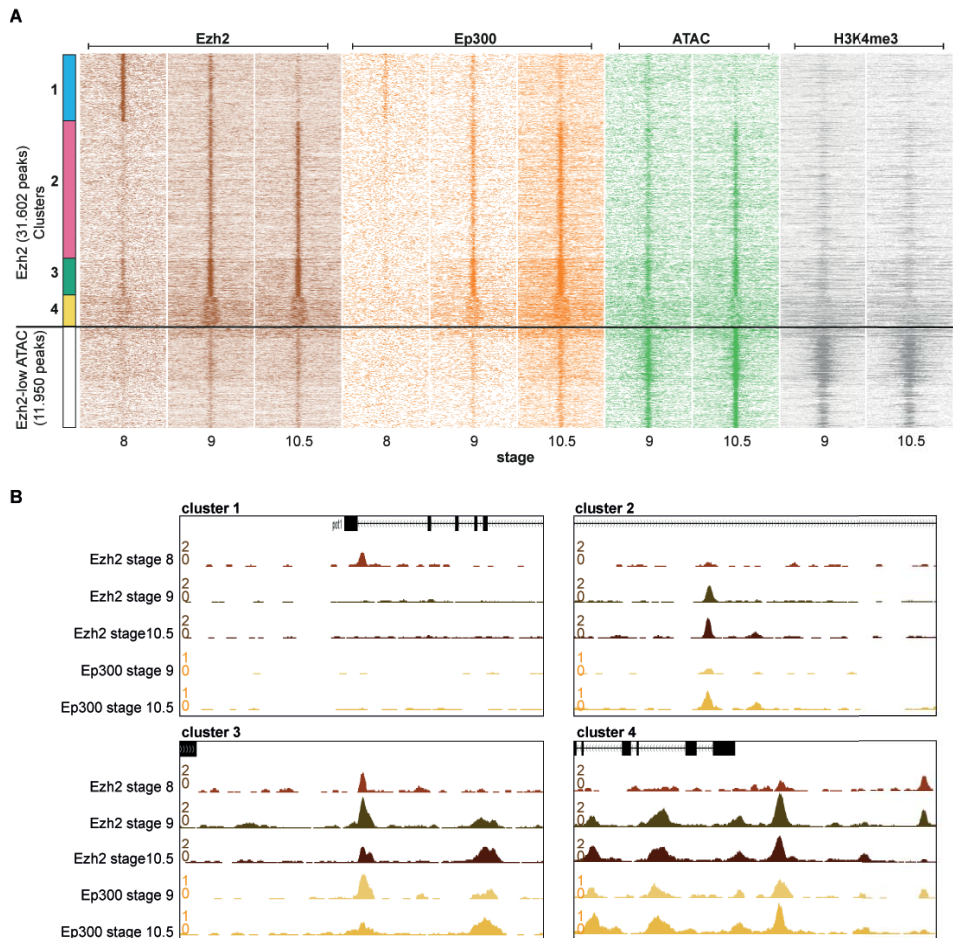


Fig. 1: Ezh2 binds at all accessible enhancers during blastula and gastrula stages. A) Stage 8, 9 and 10.5 Ezh2 ChIP-seq peaks (macs2 callpeak) were clustered. Ezh2, Ep300, ATAC and H3K4me3 (RPKM stage (8), 9, 10.5) were visualized 5 kb up- and downstream of Ezh2 peak locations (cluster 1-4) and 5 kb up- and downstream of ATAC peak locations which were absent in the Ezh2 peak-set (Ezh2-low ATAC peaks). B) Genome browser views showing examples of enhancers for each cluster. Locations in the reference *X. tropicalis* genome JGI7.1 are scaffold3 99044494-99054494; scaffold6 124132203-124142203; scaffold7 31990242-32000242; scaffold5 73327458-73337458.

in Fig. 1B).

Sequential ChIPs were performed to address the possibility if Ezh2 and Ep300 could bind enhancers together at the same time in the same cells. Enhancer sequences could be captured with an Ep300 antibody when re-ChIPping Ezh2 ChIP-captured chromatin (Supplementary Fig. 2A, B). Starting with the Ezh2 ChIP, enhancers were on average five-fold enriched in an Ep300 re-ChIP compared to a mock re-ChIP. Remarkably, enhancer sequences could hardly be captured when using the output of an Ep300 ChIP for an Ezh2 re-ChIP (Supplementary Fig. 2C, D). This could mean that a large fraction of Ezh2-associated chromatin is bound by Ep300, but conversely only a small fraction of Ep300-bound chromatin is also bound by Ezh2; this can be explained by a relatively low residency time of Ezh2 at enhancers compared to Ep300.

To gain insight in the chromatin accessibility of Ezh2 binding sites we performed ATAC sequencing for stage 9 and 10.5. Strikingly, similar to Ep300 recruitment, the ATAC-seq signals were comparable to the levels of Ezh2 binding in clusters 1-4 (Fig. 1A). The ATAC signal was the strongest in the cluster that also had most Ezh2 bound (cluster 3). Additionally, similar to Ezh2, ATAC reads mapped to a relatively broad cluster of peaks in cluster 4. Interestingly, unlike Ep300 a considerable fraction of ATAC peaks was not present in the Ezh2 peak set (Ezh2-low) (Fig. 1A). Moreover, the ATAC-signal was more prominent for these Ezh2-low ATAC peaks.

In order to distinguish enhancers and promoters we examined H3K4me3 enrichment at the various clusters (Fig. 1A). H3K4me3 was most prominent at the loci with the strongest ATAC signal and with relatively low Ezh2 and Ep300 binding. This indicated that the sites with open chromatin but low Ezh2 represented accessible promoters, while clusters 1-4 comprised enhancers. Our data shows that Ezh2 binding is a general property of accessible enhancers during and following ZGA.

Ep300 recruitment is accompanied with H3K27 acetylation at Ezh2 marked enhancers

Ezh2 catalyzes H3K27me3 at genes, many of which encode TFs with important functions in development (Boyer et al. 2006; Bernstein et al. 2006). However, most Ezh2-marked enhancers are not enriched for H3K27me3 in *X. tropicalis* at the blastula stage (van Heeringen et al. 2014). We asked if the Ezh2-bound enhancers remained depleted of H3K27me3 during development and therefore

we quantified the levels of H3K27me3 at the four Ezh2 clusters (Fig. 1A) during blastula, gastrula, neurula and tailbud stages. We found that H3K27me3 marking remained negligible at all four Ezh2-clusters (Fig. 2A, Supplementary Fig. 1B). Next we investigated which other epigenetic modifications did mark the Ezh2-

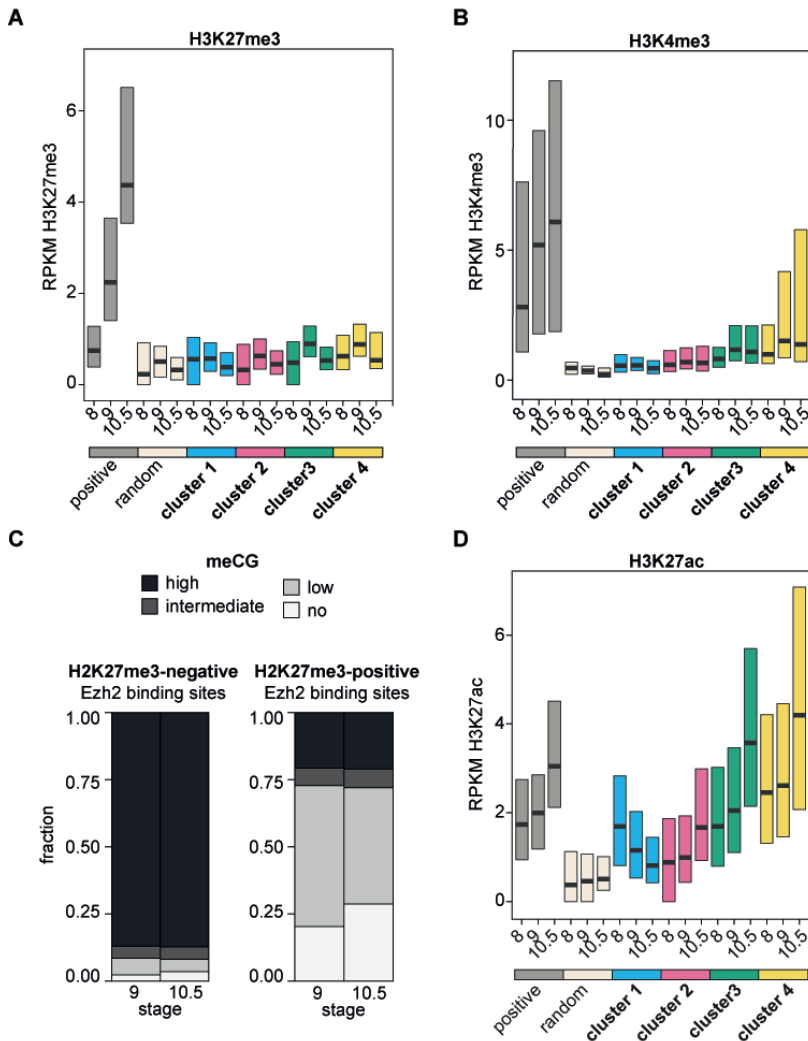


Fig. 2: Ezh2 bound enhancers are marked by hypermethylated CpGs and H3K27ac. A, B) H3K27me3 and H3K4me3 deposition (RPKM) were measured by ChIP-seq at stage 8, 9 and 10.5 for each Ezh2 cluster according to Fig. 1A. The upper and lower hinges correspond to the 25th and 75th percentiles and the horizontal line in between represents the median. The occurrence of both modifications is also shown for positive control regions which have a peak for the marks during at least one of the stages, as well as for random genomic regions. C) DNA methylation was measured by bisulfite sequencing in stage 9 and 10.5. Ezh2 binding sites were divided over two groups: H3K27me3-positive Ezh2 peaks ($n=1.831$) and H3K27me3-negative Ezh2 peaks ($n=31.088$). The CpG methylation ratio (meCG/allCG) was calculated, and visualised as the fraction of Ezh2 peaks with high (meCG/all CG >0.8), intermediate ($0.2 < \text{meCG/all CG} < 0.8$), low ($0 < \text{meCG/all CG} < 0.2$), no (meCG/all CG=0) methylation. D) H3K27ac binding (RPKM) was determined by ChIP-seq at stage 8, 9 and 10.5 for each Ezh2 cluster according to Fig. 1A. Visualisation as in Fig. 2A, B.

bound enhancers, since H3K27me3 did not.

Besides H3K27me3, Ezh2 can also catalyze dimethylation of H3K27 (H3K27me2). It has been shown that H3K27me2 is present on approximately 70% of the total histone H3 in mouse embryonic stem cells (ESCs), where it presumably prevents acetylation and activity of enhancers (Ferrari et al. 2014). We did not detect H3K27me2 at Ezh2 binding sites in *X. tropicalis* embryos at stage 10.5 (Supplementary Fig. 1C). The H3K27me2 modification is also not as abundant in *Xenopus* embryos as in mouse ESCs, since this modification is present on only 2% and 5% of the total histone H3 during blastula and gastrula (Schneider et al. 2011).

Histone methylation by Ezh2 is inhibited by the presence of H3K4me3 on the same histone tail (Schmitges et al. 2011). In accordance with Fig. 1A H3K4me3 was hardly enriched in the Ezh2 clusters compared to random background (Fig. 1A, 2B, Supplementary Fig. 1C).

H3K27me3 is acquired at DNA methylation-free domains in *X. tropicalis* embryos (van Heeringen et al. 2014). Not only H3K27me3 marking, but also SUZ12 binding increases at poised enhancers upon DNA demethylation in Dnmt KO mouse ESCs (King et al. 2016). This suggests a direct antagonism between DNA methylation and the PRC2 complex. However, the majority of Ezh2-bound enhancers was hypermethylated in *X. tropicalis* embryos (Fig. 2C). We asked whether the H3K27me3-depleted Ezh2 peaks could be distinguished by their CpG methylation status. Ezh2 peaks were divided into two groups based on the presence or absence of H3K27me3 and the two groups had an average CpG content of 2.3 and 1.5 CpGs per 100 bp respectively (Supplementary Fig. 3). As expected H3K27me3-positive Ezh2 peaks were hypomethylated (Fig. 2C). CpGs in H3K27me3-depleted Ezh2 peaks, on the other hand, were highly methylated (Fig. 2C).

H3K27me3 and acetylation of H3K27 (H3K27ac) are mutually exclusive. Moreover, Ezh2 binding sites were also more stably bound by H3K27 acetyltransferase Ep300, so abundant H3K27 acetylation could potentially prevent H3K27me3 deposition at enhancers. H3K27ac data of *X. tropicalis* at stage 8, 9 and 10.5 (Gupta et al. 2014) were used to visualize H3K27ac levels for the Ezh2 clusters (Fig. 2D, supplementary Fig. 1C). H3K27ac levels were higher than the background for all four Ezh2 clusters (Fig. 2D). Cluster 1 was the only cluster with decreasing H3K27ac levels from stage 8 to stage 10.5, similar to what was observed for the Ezh2 binding dynamics (Fig. 2D). Furthermore, the clusters in which the Ezh2 peaks were strongest (cluster 3) and broadest (cluster

4) correspondingly had the highest H3K27ac levels.

Our analysis shows that on epigenetic level Ezh2 bound enhancers are marked by hypermethylated CpGs and H3K27ac in early *X. tropicalis* embryos.

Locations where H3K27me3 is initially deposited are depleted of Ezh2 binding

X. tropicalis blastula and gastrula embryos are useful to study the dynamics of H3K27me3 deposition, since during these stages initial H3K27me3 nucleation and spreading of the modification in larger domains occurs (Akkers et al. 2009; Hontelez et al. 2015; van Heeringen et al. 2014). Ezh2-bound enhancers were not marked by H3K27me3 (Fig. 2A). We asked to what extent H3K27me3 early nucleation sites (ENS) were bound by Ezh2.

We selected genes that featured an H3K27me3 ENS at stage 8 (Fig. 3A). We analyzed the Ezh2 binding and H3K27me3 deposition at the regulatory regions of all genes that contained an ENS in their vicinity (159 genes, methods). For this analysis we divided the gene regulatory domain into three areas: 1) the ENS itself, 2) the entire H3K27me3-domain and 3) the H3K27me3-free surrounding region (the 'GREAT' regions, see methods; for example the regulatory domain of *gata3* in Fig. 3A).

The genes were clustered (Pearson, hierarchical) according to the Ezh2 binding dynamics at the three different sites within their regulatory domain (Fig. 3B). Remarkably, the Ezh2 binding at ENS was relatively low, despite the strong H3K27me3 deposition at these ENS (Fig. 3B, C). Stable Ezh2 binding was more prominent at the H3K27me3-domain than at the ENS itself (Fig. 3B, C). Moreover, we also frequently detected strong Ezh2 binding in the H3K27me3-free surrounding of the genes (Fig. 3B, C).

With the exception of one group of genes (Fig. 3B top rows (I), 3C top panel), which contains genes of which the GREAT region was (almost) completely marked by H3K27me3, Ezh2 binding is more prominently observed in the H3K27me3-free surrounding region compared to the early nucleation sites (Fig. 3B, C). Moreover, Ezh2 binding was frequently stronger or earlier present in the H3K27me3-free gene surrounding than within the H3K27me3-domain itself (Fig. 3B, 3C middle panel). The H3K27me3-free surrounding bound Ezh2 at the early blastula stage (Fig. 3B middle rows (II), 3C middle panel), or after ZGA (Fig. 3B bottom rows (III), 3C bottom panel).

Ezh2 recruitment to the enhancers may mediate repression at adjacent promoters. It has been reported that a subset of Polycomb targets cluster together in the nucleus (Denholtz et al. 2013; Joshi et al. 2015). It is therefore possible that looping of Ezh2-bound enhancers to Polycomb targets is involved in methylation of H3K27. The absence of Ezh2 at the ENS could fit with a model in which Ezh2 is brought into proximity of promoters by enhancers looping to these promoters.

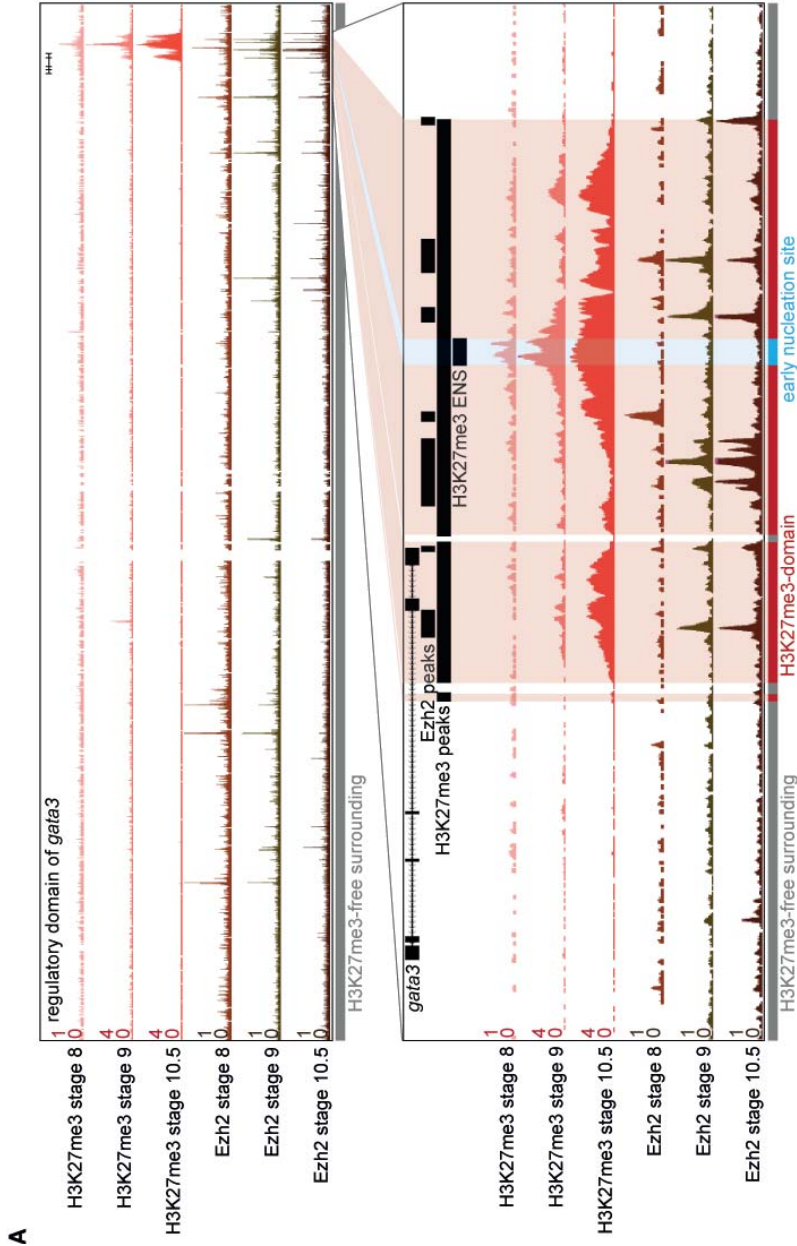
Ezh2 binding and H3K27 methyltransferase activity at enhancers are independent of transcription

Ezh2 does not only bind DNA, but it also binds to long non-coding (lnc) RNA and mRNA (Davidovich et al. 2013; Khalil et al. 2009; Zhao et al. 2010). Whether PCR2-RNA interactions stimulate or reduce PCR2-chromatin interactions remains under debate (Davidovich and Cech 2015; Beltran et al. 2016). Active enhancers do not merely activate the transcription of target genes, but are also transcribed themselves, producing a special type of lnc RNA, named enhancer RNA (eRNA) (Kim et al. 2010). Three possible roles of eRNA have been considered: 1) eRNA is a noisy side-effect from active RNAPII which is nearby; 2) the process of transcription of the enhancer is required for activating transcription of the target gene; 3) the eRNA itself is necessary for activation of the target gene (Lam et al. 2014). These contradictory views about the role of PRC2-RNA interactions and about the role of eRNA motivated us to investigate the influence of zygotic transcription on Ezh2 binding at enhancers.

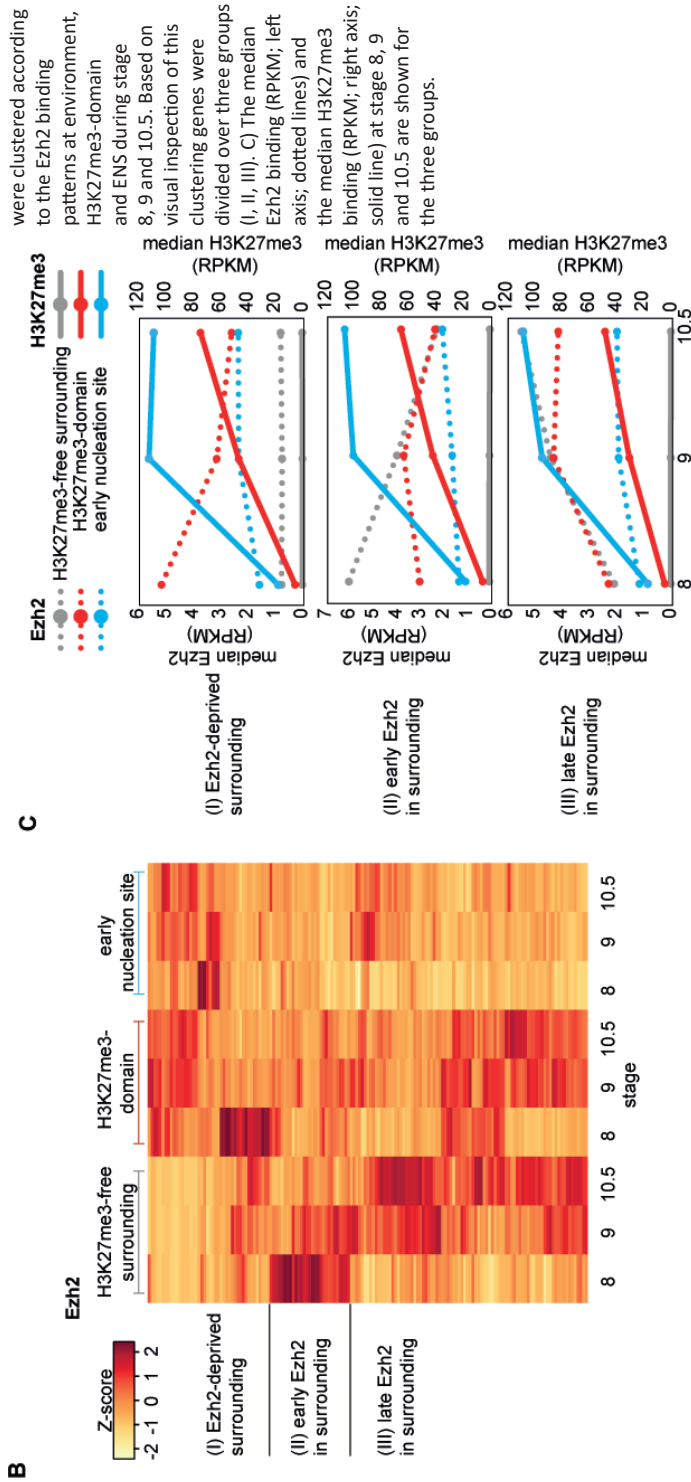
First we measured if RNAPII was present at Ezh2 bound enhancers in the clusters determined in Fig. 1A. The blastomeres of the *Xenopus* embryo become highly transcriptionally active at stage 8.5 (Newport et al. 1982), therefore we analyzed RNAPII binding during stage 9 and 10.5. At stage 9 RNAPII levels were higher than background for all Ezh2 clusters (Fig. 4A). The decrease of RNAPII binding from stage 9 to 10.5 was most substantial for cluster 1, similar to what was observed for Ezh2 binding (Fig. 4A). The presence of RNAPII indicates that transcription could occur on Ezh2-bound enhancers.

We blocked zygotic transcription by micro-injection of embryos with α -amanitin. The α -amanitin treated embryos could not gastrulate and died at stage 11. Ezh2 ChIP was performed before gastrulation at stage 9 and Ezh2 binding was measured by qPCR at four intragenic enhancers and at two negative loci. Ezh2 binding was comparable for α -amanitin treated embryos and non-injected embryos (Fig. 4B). The α -amanitin treatment inhibits transcription elongation by

Fig. 3: Loci where H3K27me3 deposition starts are depleted of Ezh2 binding. A) Genome browser view showing H3K27me3 (red) and Ezh2 (brown) ChIP-seq profiles at stage 8, 9 and 10.5 in the regulatory domain of *gata3*. Ezh2 and H3K27me3 peaks were called and stage 8 H3K27me3 peaks were named H3K27me3 early nucleation sites (ENS) (lower panel: top black bars). For all genes that contained an ENS in their regulatory domain we defined ENS (stage 8 H3K27me3-peaks), H3K27me3-domain (stage 8+9+10.5 H3K27me3 peaks) and H3K27me3-free surrounding ('gene regulatory domain' minus 'H3K27me3-domain'), as indicated in the bottom colour-bar by subsequently blue, red and grey. B) Ezh2 binding at Ezh2 peaks in H3K27me3-free surrounding, H3K27me3-domain and ENS were quantified. Genes (#159)



Dynamic Ezh2 and Ep300 recruitment to enhancers during zygotic genome activation, pluripotency and germ layer commitment in *Xenopus tropicalis*



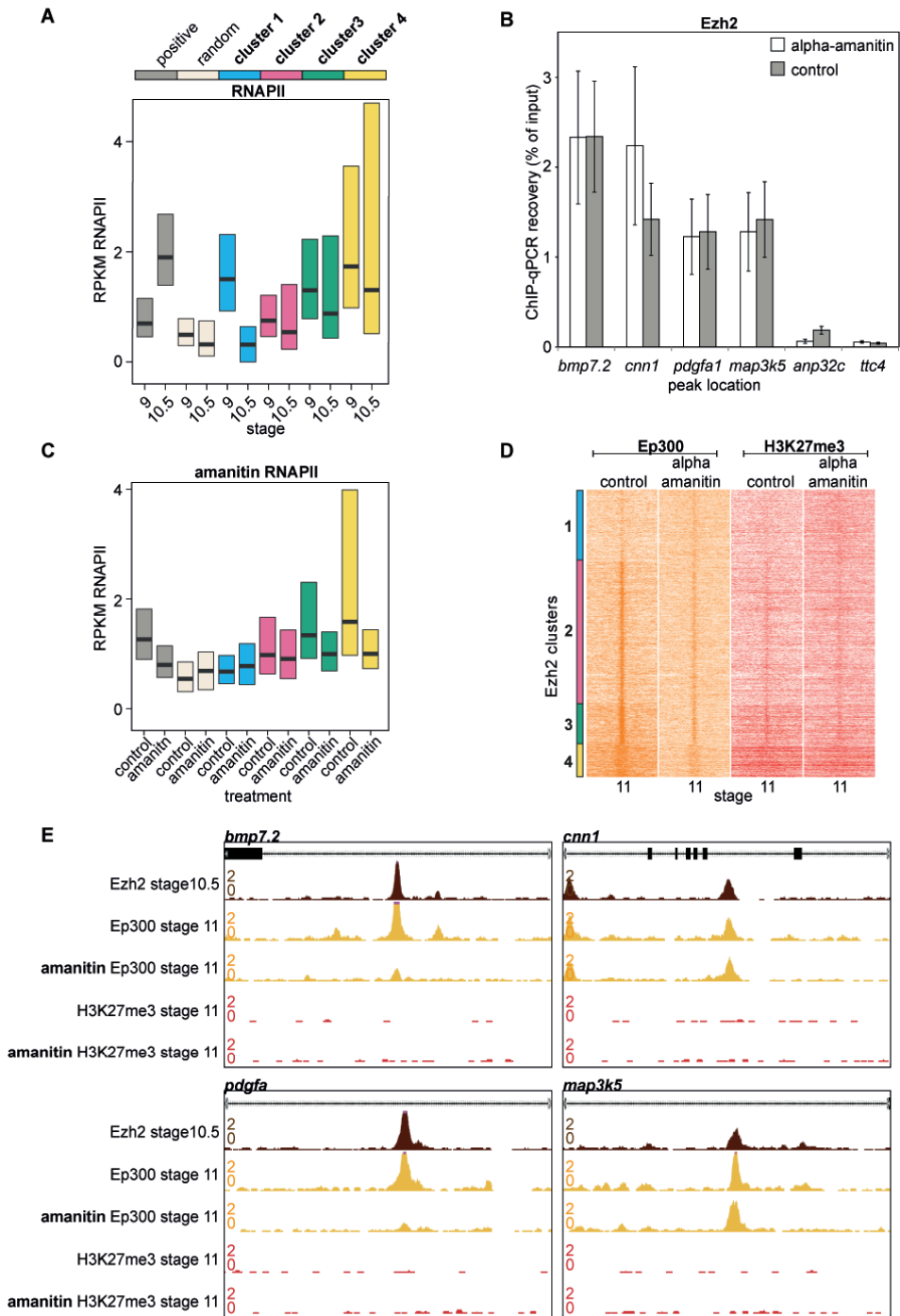


Fig. 4: Ezh2 recruitment and H3K27me3 deposition at enhancers are independent of zygotic transcription. A) RNAPII binding (RPKM) was measured by ChIP-seq at stage 9 and 10.5 for each Ezh2 cluster according to Fig. 1A. Visualisation as in Fig. 2A, B. B) Non-injected embryos and embryos injected with α -amanitin were

blocking the translocation of RNAPII on DNA (Chafin et al. 1995; Hontelez et al. 2015), but it does not prevent RNAPII binding at the enhancers (Fig. 4C). RNAPII binding is decreased, but not depleted at clusters 3 and 4 (Fig. 4C). The results suggest however that new transcription is not required for recruiting Ezh2 to the enhancers.

Next, we analyzed the influence of transcription inhibition on the H3K27me3 deposition at enhancers. We recently showed that H3K27me3 deposition at promoters is largely maternally determined, while Ep300 binding largely depends on zygotic transcription in *X. tropicalis* at stage 11 (Hontelez et al. 2015). Now, we analyzed the influence of transcription inhibition on H3K27me3 deposition at the Ezh2 clusters as determined in Fig. 1A. H3K27me3 deposition did not increase upon α -amanitin treatment (Fig. 4D). This suggests that the catalytic activity of Ezh2 at enhancers is not inhibited by the interaction of PRC2 with embryonic transcripts in this system.

Reduced Ep300 binding in α -amanitin treated embryos was not accompanied by an increase in H3K27me3 deposition (Fig. 4D). Two of the locations that we analyzed by Ezh2 ChIP-qPCR (Fig. 4B) lost Ep300 binding and two kept Ep300 binding upon transcription inhibition (Fig. 4E). All four enhancers remained depleted for H3K27me3 upon α -treatment (Fig. 4E). These data suggest that, if H3K27 methylation is dependent on the balance between Ezh2 and Ep300 recruitment, the reduction of Ep300 was not sufficient for gaining H3K27me3.

Ezh2 binds at pluripotency enhancers during blastula and gastrula

TFs can be involved in recruiting PRC2 to genomic targets (Benveniste et al. 2014; Schuettengruber et al. 2014; van Kruijsbergen et al. 2015). TFs also control the activity of enhancers; variation in TF concentrations or TF combinations results in dynamic enhancer activities (Heinz et al. 2015). We aimed to identify which TFs were recruited to enhancers that bound Ezh2 around ZGA. Since most TFs bind to specific DNA sequences we performed motif analysis on the sequences of the Ezh2 peaks in the clusters determined in Fig. 1A.

We identified various motifs that were enriched in the sequences of the

subjected to Ezh2 ChIP-qPCR in stage 9. n=4, average recovery +/- Std dev. C) For each Ezh2 cluster according to Fig. 1A RNAPII binding (RPKM) was measured by ChIP-seq at stage 11 in non-treated or α -amanitin treated embryos. Visualisation as in Fig. 2A, B. D) Ep300 and H3K27me3 binding of non-treated and α -amanitin treated embryos (RPKM, stage 11) were visualized 5 kb up- and downstream of Ezh2 peak locations according to the clustering in Fig. 1A. E) Genome browser views of Ezh2 binding at stage 10.5 (brown) at the locations analysed by ChIP-qPCR in Fig. 4B. Ep300 (yellow) and H3K27me3 (red) binding (RPKM) at stage 11 for non-treated and α -amanitin treated embryos are also shown.

Ezh2 clusters (Fig. 5A). Cluster 1 was more than five times enriched for diverse homeodomain_POU motifs. These motifs can be bound by proteins that contain a bipartite DNA binding domain composed of a POU-specific subdomain and a

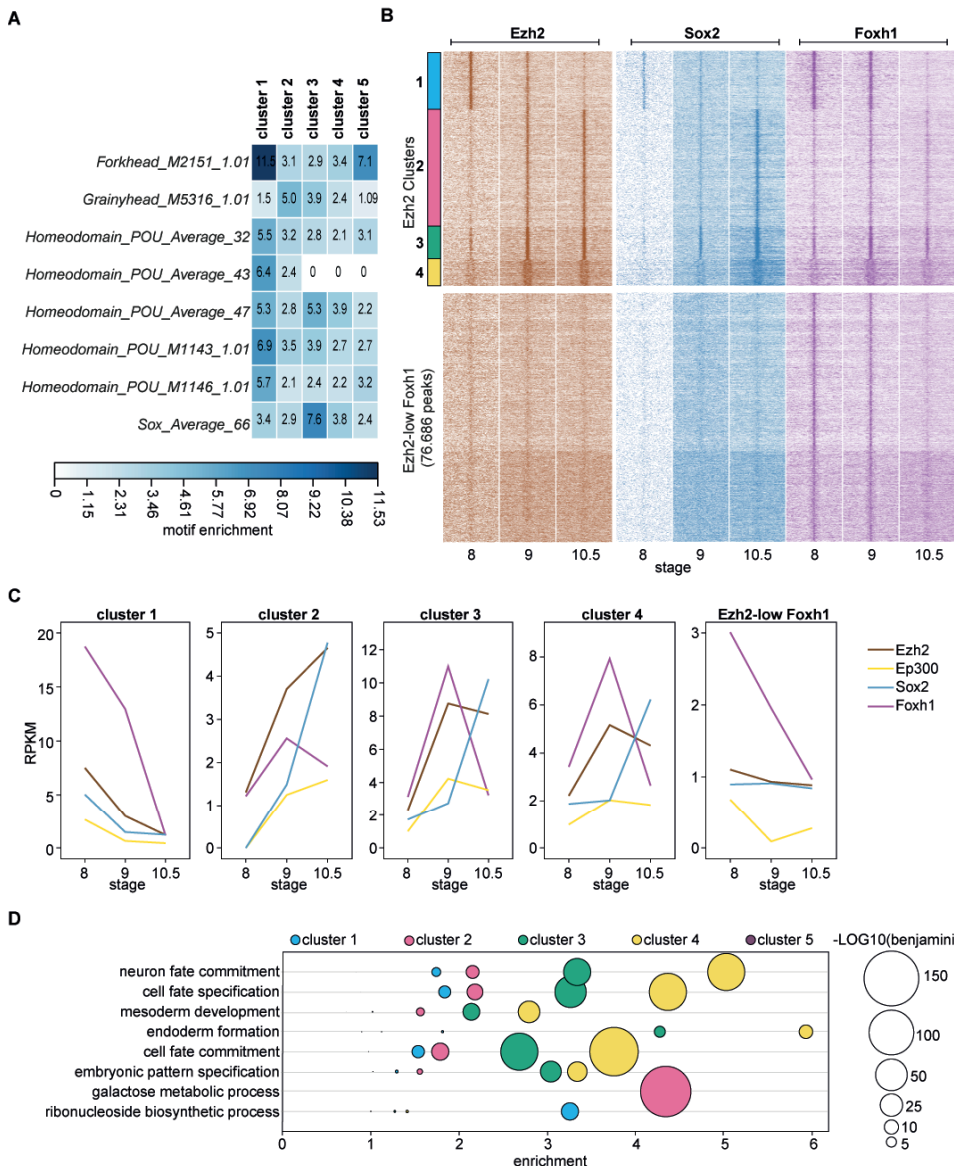


Fig. 5: Ezh2 binds at Foxh1 marked enhancers that are bound by Sox2. A) Motif analysis was done for all Ezh2 clusters (according to Fig. 1A) and for the Ezh2-depleted Foxh1 peaks. Motifs are depicted if they are more than 5-fold enriched over background. B) Ezh2, Sox2 and Foxh1 ChIP-seq data (stage 8, 9, 10.5) were plotted in a heatmap according to clustering on Ezh2 peaks (cluster 1-4, Fig. 1A) and for all Foxh1 binding sites free of Ezh2 (x-axis 10 kb). C) The average Ezh2, Ep300, Sox2 and Foxh1 binding at stage 8, 9 and 10.5 was quantified for the five clusters. D) Gene ontology analysis (DAVID) was performed for genes in the GREAT regions of the five clusters.

POU homeodomain, such as homologs of the pluripotency TF Oct4 (Ryan and Rosenfeld 1997). Furthermore, cluster 1 was more than 11 times enriched for a Foxh1 motif. The Foxh1 and homeodomain_POU motifs were also enriched in the other Ezh2 clusters, albeit to a lesser extent (Fig. 5A). Cluster 2 was most enriched for a Grainyhead motif. A Sox motif was enriched in all clusters, but most strongly in cluster 3 (over seven times). This analysis suggests that Ezh2 bound enhancers might be co-occupied by pluripotency TFs during blastula and gastrula stages.

Next, we generated Foxh1 and Sox2 binding profiles for stage 8, 9 and 10.5 to confirm that these TFs indeed bound at the same sites as Ezh2. Only 30% of the Foxh1 peaks was marked by Ezh2 (Fig. 5B). Sox2 and Ep300 did not bind at Ezh2-depleted Foxh1 binding sites either (Fig. 5B, C). Foxh1 dynamics and Ezh2 dynamics were different. Foxh1 binding was lost in all clusters between late blastula and gastrula (Fig. 5B, C). 80% of all Sox2 peaks coincided with Ezh2 peaks. Sox2, Ezh2 and Ep300 binding had similar dynamics in cluster 1 (Fig. 5B, C). Cluster 1 lost all three proteins after stage 8. The binding dynamics of Ezh2 and Ep300 were also similar to each other in cluster 2, 3 and 4, but the binding dynamics of Sox2 was slightly different at these clusters. Between stage 9 and 10.5 Sox2 binding increased more than Ezh2 and Ep300 binding did in these three clusters. Moreover, in cluster 3 and 4 Ezh2 and Ep300 binding were relatively stable between stage 9 and 10.5, while Sox2 recruitment had just started at this period (Fig. 5C). This suggests that Sox2 binds enhancers with a delay with respect to Ezh2 and Ep300.

We performed a gene ontology analysis to get an impression of which genes could be regulated by Ezh2-bound enhancers (Fig. 5D). Sox2 is a regulator of both pluripotency and neural differentiation (Zhang and Cui 2014). Foxh1 is a TF that regulates mesendodermal development by mediating both activation and repression of Nodal target genes (Reid et al. 2016). The clusters in which Ezh2, Ep300 and Sox2 peaks were strongest (cluster 3) and broadest (cluster 4) during late blastula and gastrula were enriched for terms related to (neuronal) cell fate specification and endoderm and mesoderm development. The cluster that recruited Ezh2, Ep300 and Sox2 during early blastula (cluster 1) was not enriched for these terms, but for terms related to biosynthetic processes. Regulation of active metabolic genes by Ezh2 has also been described in mouse ESCs (Brookes et al. 2012). Cluster 2, most enriched for the Grainyhead motif, was also more enriched for metabolic terms than for differentiation related terms.

Together our results emphasize the dynamic relationships between transcriptional repression and activation.

Discussion

EZH2 transiently localizes at sites of ongoing DNA replication in human fibroblasts (Hansen et al. 2008). It has been proposed that PRC2 transiently scans the whole genome by accompanying the replication fork, but that it only stably binds upon recognition of H3K27me3. Via this mechanism the repressive histone modification can be transmitted during cell division (Hansen et al. 2008). However, PRC1 and PRC2 can also bind at H3K27me3-free loci. Ring1b for example binds at enhancers in mouse embryonic stem cells (ESCs) and neuronal progenitor cells (Kloet et al. 2016), very similar to the binding of Ring1b in *X.tropicalis* embryos (Fig. S1A); EZH2 binds at promoters and in gene bodies in human prostate adenocarcinoma derived cells (Xu et al. 2012; Xu et al. 2015).

Ezh2 might bind enhancers and deposit H3K27me3 at these loci at a low rate, in competition with more prevalent acetyltransferases and counteracted by histone demethylases (Petruk et al. 2013; Tie et al. 2009). However, the α -amanitin-induced loss of Ep300 at enhancers was not sufficient to alter the Ezh2 binding and H3K27me3. This could be due to a number of things, including redundancy between Ep300 and other acetyltransferases (e.g. Crebbp/CBP) and the activity of histone demethylases.

Ezh2 as methyltransferase of non-histone targets

We showed that Ezh2 binds at enhancers which are also bound by TFs like Sox2 and Foxh1. Activity and stability of (pluripotency) TFs can be regulated by post-translational modifications (Carr et al. 2015; Cai et al. 2012). The activity and stability of SOX2, for example, are regulated by a methylation-phosphorylation switch at lysine K119. SOX2-K119 methylation by SET7 negatively influences these (Fang et al. 2014). Methyltransferase G9A, on the other hand, stabilizes SOX2 (Lee et al. 2015).

EZH2 has also been shown to function as a co-activator for AR, STAT3, and E2F1 (Xu et al. 2016; Xu et al. 2012; Kim et al. 2013). Catalytically active EZH2 is necessary for AR binding at specific enhancers that are active and free of H3K27me3 in human prostate adenocarcinoma derived cells (Xu et al. 2012). Similarly, AR is necessary for EZH2 binding at these locations, which could indicate that the proteins stabilize each other's binding. EED and SUZ12 are not necessary for stabilization of AR binding (Xu et al. 2012). STAT3 is activated by EZH2, however, in a PRC2-context. Besides EZH2, EED and SUZ12 are required for activating STAT3 in glioblastoma stem-like cells (Kim et al. 2013).

Besides its role as a co-activator, methylation by Ezh2 can also negatively influence TFs. This has been reported for GATA4 and ROR α (Lee et al. 2012; He et al. 2012). Transcriptional activity of GATA4 is inhibited by EZH2 in cardiac muscle-derived cells and mouse embryo hearts. GATA4 methylation by EZH2 inhibits GATA4-EP300 interaction (He et al. 2012). A study in mouse embryonic fibroblasts, human embryonic kidney cells, and human breast tumor tissues showed that methylation by EZH2 can also stimulate protein degradation (Lee et al. 2012). EZH2 can mono-methylate the nuclear receptor ROR α at lysine K38, which facilitates the degradation of ROR α via interaction with the E3 ubiquitin ligase complex DCAF1/DDB1/CUL4 (Lee et al. 2012).

These studies show that Ezh2 regulates TF activity independent of histone modifications. Ezh2 might bind at enhancers to function as co-activator or co-repressor rather than to catalyze methylation of H3K27.

Ezh2 as mediator of looping of cis-regulatory elements

Condensed and open chromatin domains are spatially segregated in the nucleus (Lieberman-Aiden et al. 2009). This larger nuclear structure is further organized into regions, topological domains (TADs), the boundaries of which are associated with insulator proteins (Nora et al. 2012; Dixon et al. 2012). TADs marked with H3K27me3 and PcG proteins cluster together and TADs bound by pluripotency factors cluster together within the larger nuclear structures in ESCs (Denholtz et al. 2013). PcG proteins are involved in mediating interactions between TADs, because the loss of PRC2-protein EED results in reduced interactions between TADs (Denholtz et al. 2013). PRC1 subunit RING1B and PRC2 subunit EED are necessary to maintain promoter-promoter interactions, but not promoter-enhancer interactions of H3K27me3 marked genes located in different TADs in ESCs (Schoenfelder et al. 2015; Joshi et al. 2015). However, chromosomal interactions occur more frequently within TADs than between TADs (Nora et al. 2012; Dixon et al. 2012).

EP300 and pluripotency TFs, which we detected at Ezh2 binding sites, are both involved in mediating looping interactions over shorter distances between promoters and enhancers (Fang et al. 2014; de Wit et al. 2013). Several studies have reported that PcG proteins have a role in establishing chromatin interactions within actively transcribed regions as well. Ring1b generated interest, since it co-immunoprecipitated with Cohesin; this protein was known to be important for stabilizing promoter-enhancer loops in nuclear extracts of *Drosophila* (Strübbe et al. 2011). While their binding is mutually antagonistic at silenced genes,

Ring1b and Cohesin indeed co-bind at genes that are transcriptionally active and depleted of H3K27me3 in *Drosophila* (Schaaf et al. 2013).

3D-FISH experiments at the *Meis2* locus in mouse embryos revealed that RING1B looping to promoters was not only necessary for establishing tissue specific silencing of the gene, but also for tissue specific activation of the gene (Kondo et al. 2014). Interaction of the *Meis2* promoter with a region downstream of the poly-A site at the 3'-end of the gene that was strongly bound by RING1B was necessary to mediate the promoter-enhancer interaction. First a bipartite structure between this RING1B-bound region and the promoter was formed. Next, the enhancer joined to form an intermediate tripartite interaction. This was a prerequisite for the bipartite interaction between promoter and enhancer (Kondo et al. 2014).

We showed that the chromosomal locations where H3K27me3 deposition starts are depleted of Ezh2 in *X. tropicalis*. Ezh2 rather bound around these early H3K27me3 nucleation sites. An enhancer-like looping mechanism is a plausible explanation for how distantly bound Ezh2 could regulate adjacent genes. Loop-formation involving Ezh2 might not only facilitate tissue-specific silencing, but also tissue specific activation of genes.

Future directions

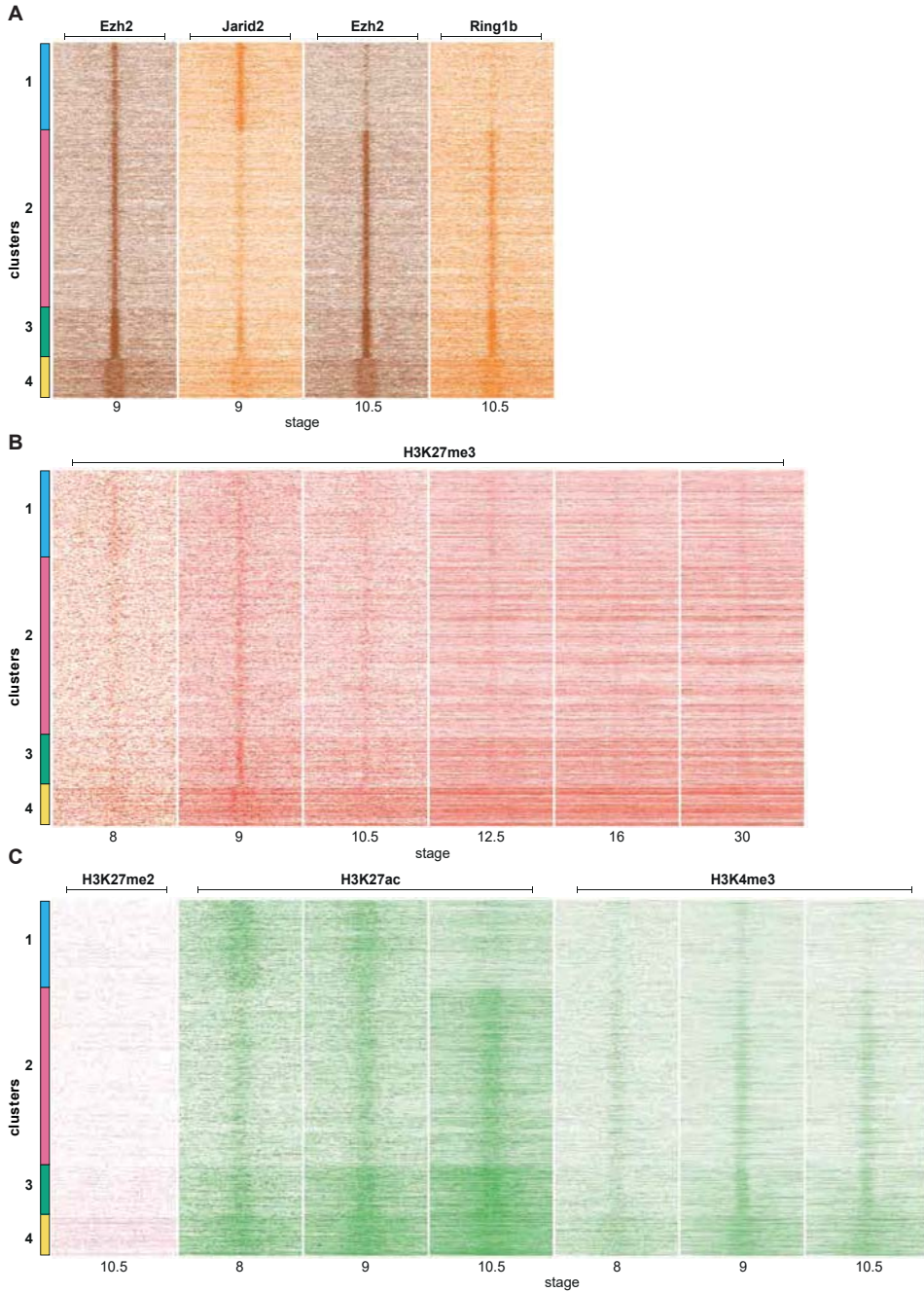
With the transition from late blastula to early gastrula, Foxh1 binding strongly decreased in all Ezh2 clusters while Ezh2 binding was relatively stable (Fig. 5B, C). Furthermore, all Ezh2-positive Foxh1 binding sites were located in accessible chromatin, while Ezh2-depleted Foxh1 binding sites were located in a more compact chromatin context. This suggests that Foxh1 might function as a pioneer factor (Iwafuchi-Doi et al. 2016; Zaret and Mango 2016). It would be interesting to investigate if Foxh1 is one of the TFs required to open gene regulatory regions to make chromatin accessible for the transcription machinery.

A highly conserved protein such as Ezh2 may very well have developed multiple biological functions. Future studies should focus on the role of Ezh2 as regulator of TFs as well as on the role of chromatin conformation in targeting of Ezh2 to Polycomb-repressed genes.

Pluripotency factors are necessary to induce ZGA in zebrafish embryos (Lee et al. 2013; Leichsenring et al. 2013). In *Xenopus* a major increase of zygotic transcription occurs between stage 8 and 9, the period in which Ezh2 binding at

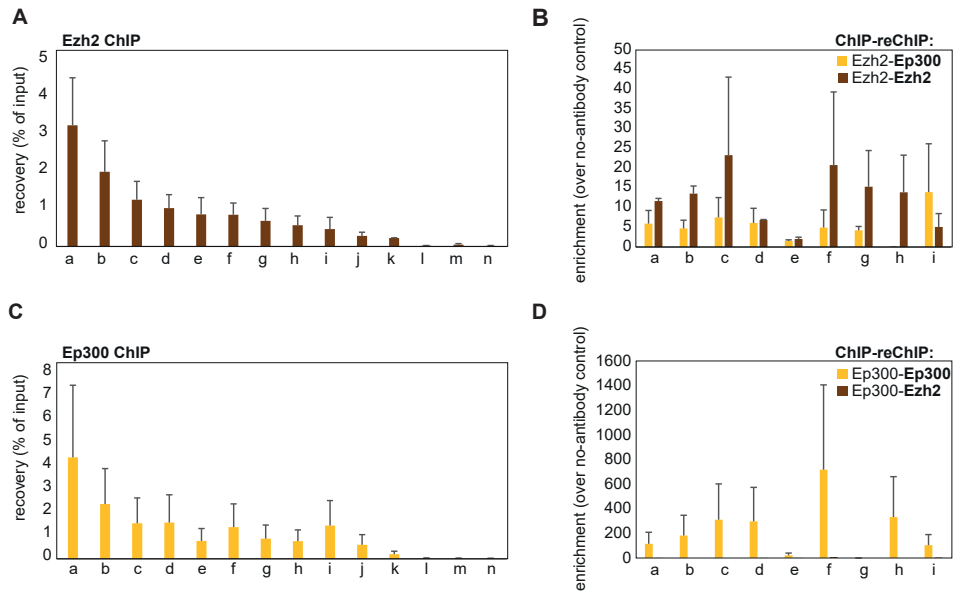
enhancers also increases. Our observation that Ezh2 co-binds enhancers with pluripotency factors reveals a potential link between Ezh2 and ZGA, which will need more work to elucidate fully.

Supplementary Figures

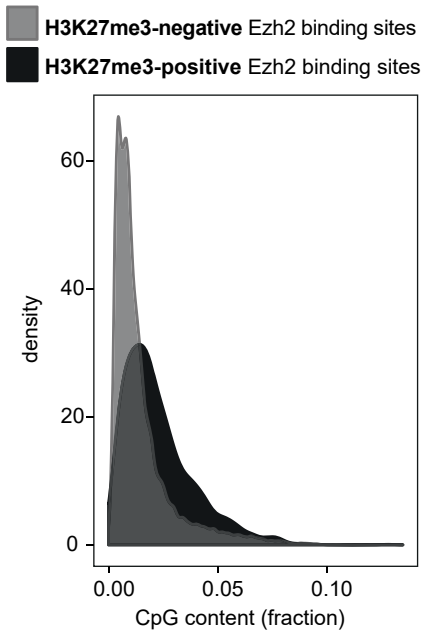


Dynamic Ezh2 and Ep300 recruitment to enhancers during zygotic genome activation, pluripotency and germ layer commitment in Xenopus tropicalis

Sup. Fig. 1: Ezh2 binds in a polycomb-context but H3K27 remains depleted of methylation. Binding (RPKM) of various proteins and histone modifications were visualized 5 kb up- and downstream of Ezh2 peak locations according to the clustering in Fig. 1A. A) Heatmap of Ezh2 at stage 9, Jarid2 at stage 9, Ezh2 at stage 10.5 and Ring1b at stage 10.5. B) Heatmap of H3K27me3 at stage 8, 9, 10.5, 12.5, 16 and 30. C) Heatmap of H3K27me2 at stage 10.5, H3K27ac at stage 8, 9, 10.5 and H3K4me3 at stage 8, 9 and 10.5.



Sup. Fig. 2: Ezh2-P300 and P300-Ezh2 ChIP-reChIP. A) Ezh2 ChIP-qPCR was performed. Recovery of input was plotted. B) The output of the 1st ChIP was subjected to a reChIP using an Ezh2 antibody, an Ep300 antibody, or no antibody. Enrichment over the no antibody control reChIP was plotted. C, D) similar as Sup. Fig. 2A, B, but starting the 1st ChIP with an Ep300 antibody. n=2, average + std dev.



Sup. Fig. 3: CpG content. Ezh2 binding sites were divided over two groups: H3K27me3-positive Ezh2 peaks and H3K27me3-negative Ezh2 peaks. CpG content (CpG/bp) was calculated and visualised in a density plot.

References

- Akkers, Robert C, Simon J van Heeringen, Ulrike G Jacobi, Eva M Janssen-Megens, Kees-Jan François, Hendrik G Stunnenberg, and Gert Jan C Veenstra. 2009. "A Hierarchy of H3K4me3 and H3K27me3 Acquisition in Spatial Gene Regulation in *Xenopus* Embryos." *Developmental Cell* 17 (3): 425–34.
- Aranda, Sergi, Gloria Mas, and Luciano Di Croce. 2015. "Regulation of Gene Transcription by Polycomb Proteins." *Science Advances* 1 (11): e1500737. doi:10.1126/sciadv.1500737.
- Barnett, M.W., R.A. Seville, S. Nijjar, R.W. Old, and E.A. Jones. 2001. "Xenopus Enhancer of Zeste (XEZ); an Anteriorly Restricted Polycomb Gene with a Role in Neural Patterning." *Mechanisms of Development* 102 (1): 157–67.
- Beltran, Manuel, Christopher M Yates, Lenka Skalska, Marcus Dawson, Filipa P Reis, Keijo Viiri, Cynthia L Fisher, et al. 2016. "The Interaction of PRC2 with RNA or Chromatin Is Mutually Antagonistic." *Genome Research*, May.
- Benveniste, Dan, Hans-Joachim Sonntag, Guido Sanguinetti, and Duncan Sproul. 2014. "Transcription Factor Binding Predicts Histone Modifications in Human Cell Lines." *Proceedings of the National Academy of Sciences of the United States of America* 111 (37): 13367–72.
- Bernstein, Bradley E, Tarjei S Mikkelsen, Xiaohui Xie, Michael Kamal, Dana J Huebert, James Cuff, Ben Fry, et al. 2006. "A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells." *Cell* 125 (2): 315–26.
- Blackledge, Neil P, Anca M Farcas, Takashi Kondo, Hamish W King, Joanna F McGouran, Lars L P Hanssen, Shinsuke Ito, et al. 2014. "Variant PRC1 Complex-Dependent H2A Ubiquitylation Drives PRC2 Recruitment and Polycomb Domain Formation." *Cell* 157 (6): 1445–59.
- Bonn, Stefan, Robert P Zinzen, Charles Girardot, E Hilary Gustafson, Alexis Perez-Gonzalez, Nicolas Delhomme, Yad Ghavi-Helm, Bartek Wilczyński, Andrew Riddell, and Eileen E M Furlong. 2012. "Tissue-Specific Analysis of Chromatin State Identifies Temporal Signatures of Enhancer Activity during Embryonic Development." *Nature Genetics* 44 (2): 148–56. doi:10.1038/ng.1064.
- Boyer, Laurie A, Kathrin Plath, Julia Zeitlinger, Tobias Brambrink, Lea A Medeiros, Tong Ihn Lee, Stuart S Levine, et al. 2006. "Polycomb Complexes Repress Developmental Regulators in Murine Embryonic Stem Cells." *Nature* 441 (7091): 349–53.
- Brookes, Emily, Inês de Santiago, Daniel Hebenstreit, Kelly J Morris, Tom Carroll, Sheila Q Xie, Julie K Stock, et al. 2012. "Polycomb Associates Genome-Wide with a Specific RNA Polymerase II Variant, and Regulates Metabolic Genes in ESCs." *Cell Stem Cell* 10 (2). Elsevier: 157–70.
- Cai, Ning, Mo Li, Jing Qu, Guang-Hui Liu, and Juan Carlos Izpisua Belmonte. 2012. "Post-Translational Modulation of Pluripotency." *Journal of Molecular Cell Biology* 4 (4): 262–65.
- Cao, Ru, Liangjun Wang, Hengbin Wang, Li Xia, Hediye Erdjument-Bromage, Paul Tempst, Richard S Jones, and Yi Zhang. 2002. "Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing." *Science* 298 (5595): 1039–43.
- Cao, Ru, and Yi Zhang. 2004. "SUZ12 Is Required for Both the Histone Methyltransferase Activity and the Silencing Function of the EED-EZH2 Complex." *Molecular Cell* 15 (1): 57–67.
- Carr, Simon M, A Poppy Roworth, Cheryl Chan, and Nicholas B La Thangue. 2015. "Post-Translational Control of Transcription Factors: Methylation Ranks Highly." *The FEBS Journal* 282 (23): 4450–65.
- Chafin, D R, H Guo, and D H Price. 1995. "Action of Alpha-Amanitin during Pyrophosphorolysis and Elongation by RNA Polymerase II." *The Journal of Biological Chemistry* 270 (32): 19114–19.
- Czeremila, Birgit, Raffaella Melfi, Donna McCabe, Volker Seitz, Axel Imhof, and Vincenzo Pirrotta. 2002. "Drosophila Enhancer of Zeste/ESC Complexes Have a Histone H3 Methyltransferase Activity That Marks Chromosomal Polycomb Sites." *Cell* 111 (2): 185–96.
- Davidovich, Chen, and Thomas R Cech. 2015. "The Recruitment of Chromatin Modifiers by Long Noncoding RNAs: Lessons from PRC2." *RNA* 21 (12): 2007–22.
- Davidovich, Chen, Leon Zheng, Karen J Goodrich, and Thomas R Cech. 2013. "Promiscuous RNA Binding by Polycomb Repressive Complex 2." *Nature Structural & Molecular Biology* 20 (11): 1250–57.
- de Napoles, Mariana, Jacqueline E Mermoud, Rika Wakao, Y Amy Tang, Mitusuhiko Endoh, Ruth Appanah, Tatyana B Nesterova, et al. 2004. "Polycomb Group Proteins Ring1A/B Link Ubiquitylation of Histone H2A to Heritable Gene Silencing and X Inactivation." *Developmental Cell* 7 (5): 663–76.
- de Wit, Elzo, Britta A. M. Bouwman, Yun Zhu, Petra Klous, Erik Splinter, Marjon J. A. M. Verstegen, Peter H. L. Krijger, et al. 2013. "The Pluripotent Genome in Three Dimensions Is Shaped around Pluripotency Factors." *Nature* 501 (7466): 227–31.
- Denholtz, Matthew, Giancarlo Bonora, Constantinos Chronis, Erik Splinter, Wouter de Laat, Jason Ernst, Matteo Pellegrini, and Kathrin Plath. 2013. "Long-Range Chromatin Contacts in Embryonic Stem Cells Reveal a Role for Pluripotency Factors and Polycomb Proteins in Genome Organization." *Cell Stem Cell* 13 (5): 602–16.
- Di Croce, Luciano, and Kristian Helin. 2013. "Transcriptional Regulation by Polycomb Group Proteins." *Nature Structural & Molecular Biology* 20 (10): 1147–55.

- Dixon, Jesse R., Siddarth Selvaraj, Feng Yue, Audrey Kim, Yan Li, Yin Shen, Ming Hu, Jun S. Liu, and Bing Ren. 2012. "Topological Domains in Mammalian Genomes Identified by Analysis of Chromatin Interactions." *Nature* 485 (7398): 376–80.
- Fang, Fang, Yifeng Xu, Kai-Khen Chew, Xi Chen, Huck-Hui Ng, and Paul Matsudaira. 2014. "Coactivators p300 and CBP Maintain the Identity of Mouse Embryonic Stem Cells by Mediating Long-Range Chromatin Structure." *Stem Cells* 32 (7): 1805–16.
- Fang, Lan, Ling Zhang, Wei Wei, Xueling Jin, Ping Wang, Yufeng Tong, Jiwen Li, et al. 2014. "A Methylation-Phosphorylation Switch Determines Sox2 Stability and Function in ESC Maintenance or Differentiation." *Molecular Cell* 55 (4): 537–51.
- Farcas, Anca M, Neil P Blackledge, Ian Sudbery, Hannah K Long, Joanna F McGouran, Nathan R Rose, Sheena Lee, et al. 2012. "KDM2B Links the Polycomb Repressive Complex 1 (PRC1) to Recognition of CpG Islands." *eLife* 1 (January) e00205.
- Ferrari, Karin J, Andrea Scelfo, Sriganesh Jammula, Alessandro Cuomo, Iros Barozzi, Alexandra Stützer, Wolfgang Fischle, Tiziana Bonaldi, and Diego Pasini. 2014. "Polycomb-Dependent H3K27me1 and H3K27me2 Regulate Active Transcription and Enhancer Fidelity." *Molecular Cell* 53 (1): 49–62.
- Friendly, M. 2002. "Corrgrams: Exploratory Displays for Correlation Matrices." *The American Statistician* 56 (November): 316–24.
- Georgiou, Georgios, and Simon J van Heeringen. 2016. "Fluff: Exploratory Analysis and Visualization of High-Throughput Sequencing Data." *bioRxiv*. Cold Spring Harbor Labs Journals.
- Gupta, Rakhi, Andrea Wills, Duygu Ucar, and Julie Baker. 2014. "Developmental Enhancers Are Marked Independently of Zygotic Nodal Signals in *Xenopus*." *Developmental Biology* 395 (1): 38–49.
- Hansen, Klaus H, Adrian P Bracken, Diego Pasini, Nikolaj Dietrich, Simmi S Gehani, Astrid Monrad, Juri Rappsilber, Mads Lerdrup, and Kristian Helin. 2008. "A Model for Transmission of the H3K27me3 Epigenetic Mark." *Nature Cell Biology* 10 (11): 1291–1300.
- He, Aibin, Xiaohua Shen, Qing Ma, Jingjing Cao, Alexander von Gise, Pingzhu Zhou, Gang Wang, Victor E Marquez, Stuart H Orkin, and William T Pu. 2012. "PRC2 Directly Methylates GATA4 and Represses Its Transcriptional Activity." *Genes & Development* 26 (1): 37–42.
- Heinz, Sven, Casey E Romanoski, Christopher Benner, and Christopher K Glass. 2015. "The Selection and Function of Cell Type-Specific Enhancers." *Nature Reviews. Molecular Cell Biology* 16 (3): 144–54.
- Hontelez, Saartje, Ila van Kruijsbergen, Georgios Georgiou, Simon J. van Heeringen, Ozren Bogdanovic, Ryan Lister, and Gert Jan C. Veenstra. 2015. "Embryonic Transcription Is Controlled by Maternally Defined Chromatin State." *Nature Communications* 6 (December): 10148.
- Huang, Da Wei, Brad T Sherman, and Richard A Lempicki. 2009a. "Systematic and Integrative Analysis of Large Gene Lists Using DAVID Bioinformatics Resources." *Nature Protocols* 4 (1): 44–57.
- . 2009b. "Bioinformatics Enrichment Tools: Paths toward the Comprehensive Functional Analysis of Large Gene Lists." *Nucleic Acids Research* 37 (1): 1–13. doi:10.1093/nar/gkn923.
- Iwafuchi-Doi, Makiko, Kenneth S Zaret, R. C. Addis, J. L. Ifkovits, F. Pinto, L. D. Kellam, P. Estes, et al. 2016. "Cell Fate Control by Pioneer Transcription Factors." *Development* 143 (11): 1833–37.
- Jones, R S, and W M Gelbart. 1990. "Genetic Analysis of the Enhancer of Zeste Locus and Its Role in Gene Regulation in *Drosophila Melanogaster*." *Genetics* 126 (1): 185–99.
- Joshi, Onkar, Shuang-Yin Wang, Tatyana Kuznetsova, Yaser Atlasi, Tianran Peng, Pierre J. Fabre, Ehsan Habibi, et al. 2015. "Dynamic Reorganization of Extremely Long-Range Promoter-Promoter Interactions between Two States of Pluripotency." *Cell Stem Cell* 17 (6): 748–57.
- Kalisch, W.-E., and B. Rasmuson. 1974. "Changes of Zeste Phenotype Induced by Autosomal Mutations in *Drosophila Melanogaster*." *Hereditas* 78 (1): 97–103.
- Kaneko, Syuzo, Roberto Bonasio, Ricardo Saldaña-Meyer, Takahaki Yoshida, Jinsook Son, Koichiro Nishino, Akihiro Umezawa, and Danny Reinberg. 2014. "Interactions between JARID2 and Noncoding RNAs Regulate PRC2 Recruitment to Chromatin." *Molecular Cell* 53 (2): 290–300.
- Kaneko, Syuzo, Jinsook Son, Roberto Bonasio, Steven S Shen, and Danny Reinberg. 2014. "Nascent RNA Interaction Keeps PRC2 Activity Poised and in Check." *Genes & Development* 28 (18): 1983–88.
- Kaneko, Syuzo, Jinsook Son, Steven S Shen, Danny Reinberg, and Roberto Bonasio. 2013. "PRC2 Binds Active Promoters and Contacts Nascent RNAs in Embryonic Stem Cells." *Nature Structural & Molecular Biology* 20 (11): 1258–64.
- Kernighan, Brian W, and Rob Pike. 1984. *THE UNIX PROGRAMMING ENVIRONMENT*. Prentice Hall.
- Khalil, Ahmad M, Mitchell Guttman, Maite Huarte, Manuel Garber, Arjun Raj, Dianali Rivea Morales, Kelly Thomas, et al. 2009. "Many Human Large Intergenic Noncoding RNAs Associate with Chromatin-Modifying Complexes and Affect Gene Expression." *Proceedings of the National Academy of Sciences of the United States of America* 106 (28): 11667–72.
- Kim, Eunhee, Misuk Kim, Dong-Hun Woo, Yongjae Shin, Jihye Shin, Nakho Chang, Young Taek Oh, et al. 2013. "Phosphorylation of EZH2 Activates STAT3 Signaling via STAT3 Methylation and Promotes Tumorigenicity of Glioblastoma Stem-like Cells." *Cancer Cell* 23 (6): 839–52.

Dynamic Ezh2 and Ep300 recruitment to enhancers during zygotic genome activation, pluripotency and germ layer commitment in Xenopus tropicalis

- Kim, Tae-Kyung, Martin Hemberg, Jesse M Gray, Allen M Costa, Daniel M Bear, Jing Wu, David A Harmin, et al. 2010. "Widespread Transcription at Neuronal Activity-Regulated Enhancers." *Nature* 465 (7295): 182–87.
- King, Andrew D, Kevin Huang, Liudmilla Rubbi, Shuo Liu, Cun-Yu Wang, Yinsheng Wang, Matteo Pellegrini, and Guoping Fan. 2016. "Reversible Regulation of Promoter and Enhancer Histone Landscape by DNA Methylation in Mouse Embryonic Stem Cells." *Cell Reports* 17 (1): 289–302.
- Kloet, Susan L, Matthew M Makowski, H Irem Baymaz, Lisa van Voorthuijsen, Ino D Karemaker, Alexandra Santanach, Pascal W T C Jansen, Luciano Di Croce, and Michiel Vermeulen. 2016. "The Dynamic Interactome and Genomic Targets of Polycomb Complexes during Stem-Cell Differentiation." *Nature Structural & Molecular Biology*, June.
- Kondo, Takashi, Kyoichi Isono, Kaori Kondo, Takahiro Endo, Shigeyoshi Itohara, Miguel Vidal, Haruhiko Koseki, et al. 2014. "Polycomb Potentiates meis2 Activation in Midbrain by Mediating Interaction of the Promoter with a Tissue-Specific Enhancer." *Developmental Cell* 28 (1): 94–101.
- Kuzmichev, Andrei, Kenichi Nishioka, Hediye Erdjument-Bromage, Paul Tempst, and Danny Reinberg. 2002. "Histone Methyltransferase Activity Associated with a Human Multiprotein Complex Containing the Enhancer of Zeste Protein." *Genes & Development* 16 (22): 2893–2905.
- Lam, Michael T Y, Wenbo Li, Michael G Rosenfeld, and Christopher K Glass. 2014. "Enhancer RNAs and Regulated Transcriptional Programs." *Trends in Biochemical Sciences* 39 (4): 170–82.
- Lee, Jae-Young, Se-Hwan Lee, Sun-Hee Heo, Kwang-Soo Kim, Changhoon Kim, Dae-Kwan Kim, Jeong-Jae Ko, and Kyung-Soon Park. 2015. "Novel Function of Lysine Methyltransferase G9a in the Regulation of Sox2 Protein Stability." *PLoS One* 10 (10): e0141118.
- Lee, Ji Min, Jason S Lee, Hyunkyung Kim, Kyeongkyu Kim, Hyejin Park, Ji-Young Kim, Seung Hoon Lee, et al. 2012. "EZH2 Generates a Methyl Degron That Is Recognized by the DCAF1/DBB1/CUL4 E3 Ubiquitin Ligase Complex." *Molecular Cell* 48 (4): 572–86.
- Lee, Miler T, Ashley R Bonneau, Carter M Takacs, Ariel A Bazzini, Kate R DiVito, Elizabeth S Fleming, and Antonio J Giraldez. 2013. "Nanog, Pou5f1 and SoxB1 Activate Zygotic Gene Expression during the Maternal-to-Zygotic Transition." *Nature* 503 (7476): 360–64.
- Leichsenring, Manuel, Julia Maes, Rebecca Mössner, Wolfgang Driever, and Daria Nischtchouk. 2013. "Pou5f1 Transcription Factor Controls Zygotic Gene Activation in Vertebrates." *Science* 341 (6149): 1005–9.
- Li, Heng, and Richard Durbin. 2009. "Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform." *Bioinformatics* 25 (14): 1754–60.
- Li, Heng, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, Richard Durbin, and 1000 Genome Project Data Processing Subgroup. 2009. "The Sequence Alignment/Map Format and SAMtools." *Bioinformatics* 25 (16): 2078–79.
- Lieberman-Aiden, Erez, Nynke L van Berkum, Louise Williams, Maxim Imakaev, Tobias Ragoczy, Agnes Telling, Ido Amit, et al. 2009. "Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome." *Science* 326 (5950): 289–93.
- McLean, Cory Y, Dave Bristor, Michael Hiller, Shoa L Clarke, Bruce T Schaar, Craig B Lowe, Aaron M Wenger, and Gill Bejerano. 2010. "GREAT Improves Functional Interpretation of Cis-Regulatory Regions." *Nature Biotechnology* 28 (5): 495–501.
- Murdoch, D. J., and E. D. Chow. 1996. "A Graphical Display of Large Correlation Matrices." *The American Statistician*. Taylor & Francis Group.
- Newport, John, Marc Kirschner, R. Bachvarova, E.H. Davidson, J.E.M. Ballantine, H.R. Woodland, E.A. Sturgess, et al. 1982. "A Major Developmental Transition in Early *Xenopus* Embryos: I. Characterization and Timing of Cellular Changes at the Midblastula Stage." *Cell* 30 (3): 675–86.
- Nora, Elphège P, Bryan R Lajoie, Edda G Schulz, Luca Giorgetti, Ikuhiro Okamoto, Nicolas Servant, Tristan Piolot, et al. 2012. "Spatial Partitioning of the Regulatory Landscape of the X-Inactivation Centre." *Nature* 485 (7398): 381–85.
- O'Carroll, D, S Erhardt, M Pagani, S C Barton, M A Surani, and T Jenuwein. 2001. "The Polycomb-Group Gene *Ezh2* Is Required for Early Mouse Development." *Molecular and Cellular Biology* 21 (13): 4330–36.
- Pasini, Diego, Adrian P Bracken, Michael R Jensen, Eros Lazzerini Denchi, and Kristian Helin. 2004. "Suz12 Is Essential for Mouse Development and for EZH2 Histone Methyltransferase Activity." *The EMBO Journal* 23 (20): 4061–71.
- Petruk, Svetlana, Kathryn L Black, Sina K Kovermann, Hugh W Brock, and Alexander Mazo. 2013. "Stepwise Histone Modifications Are Mediated by Multiple Enzymes That Rapidly Associate with Nascent DNA during Replication." *Nature Communications* 4 (January): 2841.
- Quinlan, Aaron R, and Ira M Hall. 2010. "BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features." *Bioinformatics* 26 (6): 841–42.
- Rada-Iglesias, Alvaro, Ruchi Bajpai, Tomek Swigut, Samantha a Brugmann, Ryan a Flynn, and Joanna Wysocka. 2011. "A Unique Chromatin Signature Uncovers Early Developmental Enhancers in Humans." *Nature* 470 (February): 279–83.
- R Development Core Team. 2008. R: A Language and Environment for Statistical Computing. R Foundation for

Statistical Computing. <https://www.r-project.org/>.

Reid, Christine D., Aaron B. Steiner, Sergey Yaklichkin, Qun Lu, Shouwen Wang, Morgan Hennessy, and Daniel S. Kessler. 2016. "FoxH1 Mediates a Grg4 and Smad2 Dependent Transcriptional Switch in Nodal Signaling during *Xenopus* Mesoderm Development." *Developmental Biology* 414 (1): 34–44.

Ryan, A K, and M G Rosenfeld. 1997. "POU Domain Family Values: Flexibility, Partnerships, and Developmental Codes." *Genes & Development* 11 (10): 1207–25.

San, Bilge, Naomi D Chrispijn, Nadine Wittkopp, Simon J van Heeringen, Anne K Lagendijk, Marco Aben, Jeroen Bakkers, René F Ketting, and Leonie M Kamminga. 2016. "Normal Formation of a Vertebrate Body Plan and Loss of Tissue Maintenance in the Absence of *ezh2*." *Scientific Reports* 6: 24658.

Schaaf, Cheri A, Ziva Misulovin, Maria Gause, Amanda Koenig, David W Gohara, Audrey Watson, and Dale Dorsett. 2013. "Cohesin and Polycomb Proteins Functionally Interact to Control Transcription at Silenced and Active Genes." *PLoS Genetics* 9 (6): e1003560.

Schmitges, Frank W, Archana B Prusty, Mahamadou Faty, Alexandra Stützer, Gondichatnahalli M Lingaraju, Jonathan Aiwazian, Ragna Sack, et al. 2011. "Histone Methylation by PRC2 Is Inhibited by Active Chromatin Marks." *Molecular Cell* 42 (3): 330–41.

Schneider, Tobias D, Jose M Arteaga-salas, Edith Mentele, Robert David, Dario Nicetto, Axel Imhof, and Ralph A W Rupp. 2011. "Stage-Specific Histone Modification Profiles Reveal Global Transitions in the *Xenopus* Embryonic Epigenome." *PLoS One* 6 (7).

Schoenfelder, Stefan, Robert Sugar, Andrew Dimond, Biola-Maria Javierre, Harry Armstrong, Borbala Mifsud, Emilia Dimitrova, et al. 2015. "Polycomb Repressive Complex PRC1 Spatially Constrains the Mouse Embryonic Stem Cell Genome." *Nature Genetics* 47 (10): 1179–86.

Schuettengruber, Bernd, Noa Oded Elkayam, Tom Sexton, Marianne Entrevan, Shani Stern, Aubin Thomas, Eitan Yaffe, Hugues Parrinello, Amos Tanay, and Giacomo Cavalli. 2014. "Cooperativity, Specificity, and Evolutionary Stability of Polycomb Targeting in *Drosophila*." *Cell Reports* 9 (1): 219–33.

Smits, Arne H, Pascal W T C Jansen, Ina Poser, Anthony A Hyman, and Michiel Vermeulen. 2013. "Stoichiometry of Chromatin-Associated Protein Complexes Revealed by Label-Free Quantitative Mass Spectrometry-Based Proteomics." *Nucleic Acids Research* 41 (1): e28.

Strübbe, Gero, Christian Popp, Alexander Schmidt, Andrea Pauli, Leonie Ringrose, Christian Beisel, and Renato Paro. 2011. "Polycomb Purification by in Vivo Biotinylation Tagging Reveals Cohesin and Trithorax Group Proteins as Interaction Partners." *Proceedings of the National Academy of Sciences of the United States of America* 108 (14): 5572–77.

Tie, Feng, Rakhee Banerjee, Carl A Stratton, Jayashree Prasad-Sinha, Vincent Stepanik, Andrei Zlobin, Manuel O Diaz, Peter C Scacheri, and Peter J Harte. 2009. "CBP-Mediated Acetylation of Histone H3 Lysine 27 Antagonizes *Drosophila* Polycomb Silencing." *Development* 136 (18): 3131–41.

van Heeringen, Simon J, Robert C Akkers, Ila van Kruijsbergen, M Asif Arif, Lars L P Hanssen, Nilofar Sharifi, and Gert Jan C Veenstra. 2014. "Principles of Nucleation of H3K27 Methylation during Embryonic Development." *Genome Research* 24 (3): 401–10.

van Heeringen, Simon J, and Gert Jan C Veenstra. 2011. "GimmeMotifs: A de Novo Motif Prediction Pipeline for ChIP-Sequencing Experiments." *Bioinformatics* 27 (2): 270–71.

van Kruijsbergen, Ila, Saartje Hontelez, and Gert Jan C Veenstra. 2015. "Recruiting Polycomb to Chromatin." *The International Journal of Biochemistry & Cell Biology* 67 (May): 177–87.

Weirauch, Matthew T, Ally Yang, Mihai Albu, Atina G Cote, Alejandro Montenegro-Montero, Philipp Drewe, Hamed S Najafabadi, et al. 2014. "Determination and Inference of Eukaryotic Transcription Factor Sequence Specificity." *Cell* 158 (6): 1431–43.

Wickham, Hadley. 2009. "ggplot2: Elegant Graphics for Data Analysis." Springer-Verlag New York. <http://ggplot2.org/>.

Wyngaarden, Laurie A, Paul Delgado-Olguin, I-hsin Su, Benoit G Bruneau, and Sevan Hopyan. 2011. "Ezh2 Regulates Anteroposterior Axis Specification and Proximodistal Axis Elongation in the Developing Limb."

Xu, Han, Kexin Xu, Housheng H He, Chongzhi Zang, Chen-Hao Chen, Yiwen Chen, Qian Qin, et al. 2016. "Integrative Analysis Reveals the Transcriptional Collaboration between EZH2 and E2F1 in the Regulation of Cancer-Related Gene Expression." *Molecular Cancer Research* 14 (2): 163–72.

Xu, Jian, Zhen Shao, Dan Li, Huafeng Xie, Woojin Kim, Jialiang Huang, Jordan E. Taylor, et al. 2015. "Developmental Control of Polycomb Subunit Composition by GATA Factors Mediates a Switch to Non-Canonical Functions." *Molecular Cell*, January.

Xu, Kexin, Zhenhua Jeremy Wu, Anna C Groner, Housheng Hansen He, Changmeng Cai, Rosina T Lis, Xiaoqiu Wu, et al. 2012. "EZH2 Oncogenic Activity in Castration-Resistant Prostate Cancer Cells Is Polycomb-Independent." *Science* 338 (6113): 1465–69.

Zaret, Kenneth S, and Susan E Mango. 2016. "Pioneer Transcription Factors, Chromatin Dynamics, and Cell Fate Control." *Current Opinion in Genetics & Development* 37: 76–81.

Zhang, Shuchen, and Wei Cui. 2014. "Sox2, a Key Factor in the Regulation of Pluripotency and Neural Differentiation." *World Journal of Stem Cells* 6 (3): 305–11.

Dynamic Ezh2 and Ep300 recruitment to enhancers during zygotic genome activation, pluripotency and germ layer commitment in Xenopus tropicalis

Zhang, Yong, Tao Liu, Clifford A Meyer, Jérôme Eeckhoute, David S Johnson, Bradley E Bernstein, Chad Nusbaum, et al. 2008. "Model-Based Analysis of ChIP-Seq (MACS)." *Genome Biology* 9 (9): R137.

Zhao, Jing, Toshiro K Ohsumi, Johnny T Kung, Yuya Ogawa, Daniel J Grau, Kavitha Sarma, Ji Joon Song, Robert E Kingston, Mark Borowsky, and Jeannie T Lee. 2010. "Genome-Wide Identification of Polycomb-Associated RNAs by RIP-Seq." *Molecular Cell* 40 (6): 939–53.

CHAPTER 6

Summary and discussion

The work described in this thesis involved the deposition of histone modifications during development in *Xenopus tropicalis*. We generated genome-wide maps for the binding of various histone modifications throughout development. The deposition of permissive marks H3K4me3 and H3K9ac preceded the deposition of repressive marks H3K27me3, H3K9me2/3 and H4K20me3 as well as the binding of active enhancer mark Ep300 (Chapter 3). We and others found that permissive marks H3K4me3 and H3K9ac are deposited before transcription starts and they could therefore have a guiding function for the transcriptional machinery (Akkers et al. 2009; Vastenhouw et al. 2010). Next, we analyzed the epigenetic maps from three different angles.

Firstly, we asked what the origin of the histone modifications is; are they predetermined by maternal factors or are they zygotically derived (Chapter 3)? We expected that the deposition of H3K4me3 was determined by maternal factors, since this permissive mark was already deposited before ZGA. We have shown that zygotic transcription was indeed largely dispensable for the deposition of H3K4me3. H3K27me3 deposition was also independent of zygotic transcription, which was surprising since H3K27me3 emerged mostly after ZGA. Different from H3K4me3 and H3K27me3 the majority of the binding of H3K27 acetyltransferase Ep300 did depend on zygotic transcription. Together, the first section of our study emphasized the combinatorial action of maternal and zygote derived factors through epigenetic regulation of proximal and distal regulatory sequences.

Secondly, we asked how histone modifications associated with transcriptional repression behaved at transposable elements (Chapter 4). These parasitic repetitive elements can be beneficial for evolution as well as harmful for the zygote (Friedli and Trono 2015). Therefore we expected that proper regulation of transposons during embryogenesis is essential and dynamic. Our study revealed that DNA transposons and retrotransposons are marked by different patterns of repressive histone modifications: a subset of DNA transposons was marked with H4K20me3 and H3K27me3 during early development, while a subset of retrotransposons acquired H4K20me3 and H3K9me3 later in development. Retrotransposon subfamilies that obtained repressive marks could be distinguished from the subfamilies that were not marked with repressive modifications by the presence of small RNA. Furthermore, by using intra-subfamily divergence as a proxy for age, we showed that the transposon subfamilies marked by repressive modifications were relatively young. Accordingly, other studies revealed that the recruitment of enzymes that catalyze the constitutive repressive modifications as well as the occurrence of piRNA are also related with the age of transposons

(Castro-Diaz et al. 2014; Pezic et al. 2014). This second part of our study provides a compendium of the dynamic behavior of repressive epigenetic modifications during vertebrate embryogenesis, which could be a starting point for future research into the (the evolution of) mechanisms of transposon regulation.

Lastly, we characterized Ezh2-bound enhancers (Chapter 5). Looking at our data sets we noticed something remarkable: the overlap between this H3K27me3 methyltransferase and H3K27me3 profiles was rather low. Moreover, the Ezh2 binding was similar to the Ep300 binding pattern. We revealed that Ezh2, just as Ep300, bound at practically all accessible enhancers during early *Xenopus* development. Acetylation was more abundant than methylation of H3K27 at enhancers, but H3K27me3 marking of enhancers did not increase upon α -amanitin induced removal of Ep300. This last part of our study highlighted the complicated relationships between transcriptional activators and repressors.

As summarized above, our work provided new insights about epigenetic regulation during vertebrate embryogenesis. PRC2 was a common theme running through all chapters in this thesis. Therefore, I will now elaborate on the recruitment and (catalytic) activity of this complex.

Ezh2 recruitment

In chapter 2 we have reviewed how multiple players have a role in the recruitment of PRC2 to chromatin. We and others reported that the deposition of H3K4me3 and H3K27me3 correlates with meCpG-depleted regions (Chapter 3) (Bartke et al. 2010; van Heeringen et al. 2014; Wachter et al. 2014). These hypomethylated PRC2 targets are found at transcription start sites and have a relatively high CpG density (Lee et al. 2006; Mendenhall et al. 2010; Lynch et al. 2012). We showed that Ezh2 also binds at enhancers which have a relatively low CpG density and are hypermethylated (Chapter 5).

In Chapter 5 we showed that Ezh2 binds at all accessible enhancers. The enhancers bound by Ezh2 during blastula and gastrula stages were enriched for TF motifs, such as homeodomain_POU, Grainyhead, Foxh1 and Sox motifs. Not all TFs can stably bind DNA in a condensed chromatin context. The TFs that do stably bind to condensed DNA are called pioneer factors. Upon binding these pioneer factors can mediate a more accessible chromatin environment by recruiting nucleosome remodelers (Zaret and Mango 2016). In our study we compared Foxh1 binding with Ezh2 binding. Different from Ezh2, as well as from Ep300 and Sox2, Foxh1 did not only bind at accessible chromatin, but also

at closed chromatin. Possibly Foxh1 functions as a pioneer factor. It would be interesting to study which remodelers are recruited to Foxh1 binding sites and if they are necessary for Ezh2 recruitment to these loci. Furthermore, it would be interesting to study what determines if a Foxh1 binding site becomes assessable or remains closed.

If it holds true that an open chromatin state is sufficient to recruit Ezh2 it is interesting to consider the overall open chromatin conformation during early embryogenesis (see introduction). Ezh2 could potentially randomly bind throughout the whole genome at the start of embryogenesis, since the DNA is generally accessible at that period (Hair et al. 1998; Wu et al. 2016). H3K27me3 deposition starts to take place around ZGA in *Xenopus*. This could be caused by the general condensation of the chromatin which blocks access of Ezh2 to large parts of the genome. The limited portion of the chromatin that remains accessible for Ezh2 upon the condensation ends up being bound by relatively more PRC2 per binding site.

Interestingly, the effects of chromatin condensation might unite the two classical models for what triggers ZGA: the nucleocytoplasmic ratio model and the maternal clock model (Newport et al. 1982a, 1982b; Howe et al. 1995). Condensation reduces the number of potential binding sites for transcription activating complexes as well as for transcription repressive complexes. This could lead to increased binding per binding site for both types of complexes. Binding of repressors such as Ezh2 can increase, but also binding of activators such as Ep300 and Sox2 can increase. Interestingly we found that Ezh2 binding sites largely overlap with Sox2 binding sites. The pluripotency TFs, among which Sox2, have already been shown to be essential factors for the onset of ZGA in zebrafish (Lee et al. 2013; Leichsenring et al. 2013). We highlighted that the interplay between repressors and activators in regulating transcription is highly complex (Chapter 5).

With respect to the recruitment of Ezh2 during development I suggest to focus future studies on nucleosome positioning and accessibility (described as 'local chromatin regulation' in the introduction). It would be interesting to study the influence of the local chromatin regulation on the interplay between (pioneer) TFs and Ezh2.

H3K27me3-H3K4me3 equilibrium

Once Ezh2 binds in an meCpG-depleted region it can catalyze H3K27me3,

however, the region can also remain free of H3K27me3 as we described in chapter 5. What determines if Ezh2 binding also results in H3K27me3-marked repression? It has been proposed that H3K27me3 deposition is prevented by active transcription. Studies revealed that the removal of transcription activation indeed was sufficient to trigger H3K27me3 deposition at meCpG-depleted CpG-rich sequences (Mendenhall et al. 2010; Arnold et al. 2013; Jermann et al. 2014; Riising et al. 2014). It remains to be determined, however, which factors involved in transcriptional activity are exactly responsible for preventing H3K27me3 deposition.

It has been shown that PRC2 specifically binds nascent transcripts. It has been proposed that PRC2 can sense the transcriptional state of a gene via RNA binding (Davidovich et al. 2013; Kaneko et al. 2013). However, in our study transcription inhibition left H3K4me3 and H3K27me3 marking largely undisturbed. Since α -amanitin does not block RNAPII binding, but only RNAPII translocation over the DNA and thereby transcription elongation, short nascent transcripts could potentially still be formed. It has been shown, however, that longer RNA sequences bind PRC2 more efficiently than shorter sequences (Davidovich et al. 2013). Our study suggests that H3K27me3 deposition during embryogenesis is not influenced by the process of transcription itself. It should be noted that RNA was still present in the embryos despite transcription inhibition, since the α -amanitin treatment does not target maternally derived transcripts.

Transcription inhibition in mouse ES cells, by blocking elongation or by stimulating RNAPII degradation, was sufficient to induce PRC2 binding and H3K27me3 deposition at C+G rich loci (Riising et al. 2014). It is possible that the increasing methylation of H3K27 in the study by Riising and colleagues was caused by processes coupled to transcription rather than by transcription itself, such as a changing nucleosome occupancy. The nucleosomal structure is, for example, destabilized by the FACT (facilitates chromatin transcription) complex during RNAPII-driven transcription elongation (Belotserkovskaya et al. 2003). It has been shown that compaction of histones precedes PRC2 activity and is sufficient to induce H3K27me3 deposition in mouse ES cells (Yuan et al. 2012).

Although H3K4me3 and H3K27me3 are both deposited at loci with a relatively high histone replacement rate and low nucleosome occupancy, H3K4me3 marked regions have a slightly higher replacement rate and lower density than H3K27me3 marked regions (Mito et al. 2007; Deal et al. 2010). Ratios of H3K27me3 and H3K4me3 occupancy might be disturbed if transcription inhibition leads to higher nucleosome condensation as a result of the absence of transcription-coupled destabilization of nucleosomes. The balance between

H3K27me3 and H3K4me3 remained undisturbed upon transcription inhibition in our study, while it was disturbed in studies in mouse ES cells (Riising et al. 2014). This difference might be caused by the distinct timing of chromatin condensation and nucleosome positioning observed in *Xenopus* and mice during early development (as described in the introduction).

Inhibition of transcription in mouse ES cells by the deletion of binding motifs for transcription activators changed the levels of H3K4me3 and H3K27me3 marking; H3K27me3 deposition increased, while H3K4me3 deposition was relatively low (Mendenhall et al. 2010; Jermann et al. 2014). The presence of H3K4me3 inhibits H3K27me3 deposition on the same histone tail (Schmitges et al. 2011). This suggests that the presence of H3K4me3 might be sufficient to minimize H3K27me3 marking. However, in zebrafish the two modifications do co-occur at bivalent domains (Vastenhouw et al. 2010). In *Xenopus* and pre-implantation mouse embryos, though, such bivalency is relatively infrequent and unstable (Akkers et al. 2009; Liu et al. 2016). It remains to be determined what the decisive component is that can shift the equilibrium towards H3K4me3 marked activation or H3K27me3 marked repression.

Parasitic targets of Ezh2

We showed that during embryogenesis Ezh2 is also catalyzing H3K27me3 at transposable elements. Studies about the function of H3K27me3 in silencing transposons remain relatively scarce. First studies about epigenetic pathways involved in the silencing of the parasitic part of the genome were focused on meCpG. Later, studies pointed out that during early development histone modifications H3K9me2/3 and H4K20me3 were more important for the silencing of transposons than meCpG related pathways (Martens et al. 2005; Hutnick et al. 2010; Matsui et al. 2010; Karimi et al. 2011; Bulut-Karslioglu et al. 2014). The shift from H3K9me2/3 and H4K20me3 related repression towards meCpG mediated silencing shows that upon differentiation different strategies develop to protect the host genome.

Our study (chapter 4) indicates that there could be an extra transition in repression regulation before the transition from H3K9me2/3 and H4K20me3 to meCpG. We found various transposons marked by H3K27me3 during blastula at stage 8, the earliest stage that we analysed. H3K27me3 was accompanied by other repressive modifications at these particular transposon subfamilies, while normally constitutive and facultative repressive marks are enriched at distinct regions. H3K27me3 was already enriched at these parasitic elements before

ZGA, while we could hardly detect the other repressive marks that early.

Knock out of a PRC2 core component only led to the increase of a small minority of some transposable elements in mouse ES cells grown in serum (Leeb et al. 2010). However, recently a study showed that PRC2 has a substantial role in transposon silencing in hypomethylated ES cells (Walter et al. 2016). ES cells cultured in serum are heavily methylated, while ES cells cultured in a medium containing two small kinase inhibitors and vitamin C (2i+vitC) are hypomethylated. Walter and colleagues revealed that when transferred from serum to 2i+vitC not only meCpG reduced at transposons, but that also H3K9me3 and/or H3K27me3 increased at the same transposons. The deposition of H3K9me3, H3K27me3 or both prevented transcriptional activity of transposons when DNA hypomethylation was induced. H3K9me3 changes occurred earlier than H3K27me3 (Walter et al. 2016). H3K27me3 might have a function in maintaining a heterochromatin environment in the absence of other epigenetic modifications.

Such a temporary function for H3K27me3 is also seen at pericentric heterochromatin in mice (Puschendorf et al. 2008). H3K27me3 is transiently enriched until morula at pericentric heterochromatin, regions normally marked by H3K9me3 instead. We merely detected H3K27me3 at transposable elements during the earliest developmental stages. So, albeit only for a subset of the transposons, H3K27me3 could be important for early silencing processes. We found that the parasitic elements were already heavily methylated at CpGs in stage 9. We showed that H3K27me3 deposition in *Xenopus* is associated with hypomethylated CpGs (van Heeringen et al. 2014). Possibly this could be a reason why H3K27me3 is not very abundant at most transposons in our study. It would be interesting to study meCpG and H3K27me3 deposition in younger embryos to determine if H3K27me3 is more frequently deposited at transposons during the first cleavage cycles.

H3K27me3-independent Ezh2 activity

We showed that Ezh2 binds at enhancers which are generally H3K27me3-depleted. Moreover, H3K27me3 deposition is often initiated at sites distal from where Ezh2 is recruited to (Chapter 5). PRC2 might be brought into close proximity of unmethylated polycomb targets via enhancer-promoter loops. It would be interesting to study if loop-formation is required to engage the H3K27 methylation by PRC2.

As discussed in chapter 5 Ezh2 could have different functions besides catalysing H3K27me3. We proposed that the PRC2 subunit might be involved in the formation of enhancer-promoter loops. Similarly, a study in mouse embryos revealed that a transient interaction with PRC1 subunit RING1B is necessary to mediate promoter-enhancer interactions at the *Meis2* locus (Kondo et al. 2014). I suggest to carry out chromatin conformation capture related studies to clarify if transient Ezh2 binding is involved in a similar mechanism as Ring1b. The transient binding behaviour could explain why Ezh2 binding at (H3K27me3-depleted) enhancers is hardly observed in other studies.

Concluding remarks and future perspective

We generated reference chromatin state maps at multiple stages of embryogenesis and with these we provide a valuable resource for future developmental, genetic and systems biology studies in *X. tropicalis*. With these maps we are able to localize enhancers, promoters and heterochromatin. This information provides a solid basis for future studies in the *Xenopus* community.

Improved ChIP-seq methods will make it possible to perform ChIP-seq experiments with less starting material. Various pre-implantation mouse ChIP-seq studies have already been published last summer; genome wide histone modification binding sites were successfully generated starting with only several hundred cells (Dahl et al. 2016; Liu et al. 2016; Zhang et al. 2016). Although it is becoming more accessible to perform ChIP-seq studies in mammalian embryos, I expect that *Xenopus* as well as zebrafish will stay important model systems for molecular developmental biology. I expect this because of practical considerations concerning the production and monitoring of the embryos. Furthermore, despite the differences in the early embryogenesis of mammalian and non-mammalian vertebrates they do share molecular principles. For example, the temporal hierarchy of K4 and K27 methylation that we observed in *X. tropicalis* (Akkers et al. 2009; van Heeringen et al. 2014) has now also been observed in pre-implantation mouse embryos (Liu et al. 2016).

The integration of data on the dynamics of histone modifications, TF binding, nucleosome positioning, chromosomal interactions and transcription will provide new insights in embryonic transcription regulation. Studies in the field of molecular developmental biology will be accelerated by the breakthrough of DNA editing by CRISPR-Cas9 (Jinek et al. 2012). Molecular developmental biology is facing exciting times!

References

- Akkers RC, Heeringen SJ Van, Jacobi UG, Janssen-megens EM, François K, Stunnenberg HG, et al. A hierarchy of H3K4me3 and H3K2me3 acquisition in spatial gene regulation in *Xenopus* embryos. *Dev Cell* 2009;17(3):425–34.
- Arnold P, Schöler A, Pachkov M, Balwierz PJ, Jørgensen H, Stadler MB, et al. Modeling of epigenome dynamics identifies transcription factors that mediate Polycomb targeting. *Genome Res* 2013 Jan 1;23(1):60–73.
- Bartke T, Vermeulen M, Xhemalce B, Robson SC, Mann M, Kouzarides T. Nucleosome-interacting proteins regulated by DNA and histone methylation. *Cell* 2010 Oct 29 ;143(3):470–84.
- Belotserkovskaya R, Oh S, Bondarenko VA, Orphanides G, Studitsky VM, Reinberg D. FACT Facilitates Transcription-Dependent Nucleosome Alteration. *Science* 2003;301(5636).
- Bulut-Karslioglu A, De La Rosa-Velázquez IA, Ramirez F, Barenboim M, Onishi-Seebacher M, Arand J, et al. Suv39h-dependent H3K9me3 marks intact retrotransposons and silences LINE elements in mouse embryonic stem cells. *Mol Cell* 2014 Jul 17;55(2):277–90.
- Castro-Diaz N, Ecco G, Coluccio A, Kapopoulou A, Yazdanpanah B, Friedli M, et al. Evolutionally dynamic L1 regulation in embryonic stem cells. *Genes Dev* 2014 Jul 1;28(13):1397–409.
- Dahl JA, Jung I, Aanes H, Greggains GD, Manaf A, Lerdrup M, et al. Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. *Nature* 2016 Sep 14;537(7621):548–52.
- Davidovich C, Zheng L, Goodrich KJ, Cech TR. Promiscuous RNA binding by Polycomb repressive complex 2. *Nat Struct Mol Biol* 2013 Nov;20(11):1250–7.
- Deal RB, Henikoff JG, Henikoff S. Genome-wide kinetics of nucleosome turnover determined by metabolic labeling of histones. *Science* 2010 May 28;328(5982):1161–4.
- Friedli M, Trono D. The Developmental Control of Transposable Elements and the Evolution of Higher Species. *Annu Rev Cell Dev Biol* 2015 Sep 17.
- Hair A, Prioleau MN, Vassetzky Y, Méchali M. Control of gene expression in *Xenopus* early development. *Dev Genet* 1998;22(2):122–31.
- van Heeringen SJ, Akkers RC, van Kruijsbergen I, Arif MA, Hanssen LLP, Sharifi N, et al. Principles of nucleation of H3K27 methylation during embryonic development. *Genome Res* 2014 Mar 1;24(3):401–10.
- Howe JA, Howell M, Hunt T, Newport JW. Identification of a developmental timer regulating the stability of embryonic cyclin A and a new somatic A-type cyclin at gastrulation. *Genes Dev* 1995 May 15;9(10):1164–76.
- Hutnick LK, Huang X, Loo T-C, Ma Z, Fan G. Repression of retrotransposal elements in mouse embryonic stem cells is primarily mediated by a DNA methylation-independent mechanism. *J Biol Chem* 2010 Jul 2;285(27):21082–91.
- Jermann P, Hoerner L, Burger L, Schübeler D. Short sequences can efficiently recruit histone H3 lysine 27 trimethylation in the absence of enhancer activity and DNA methylation. *Proc Natl Acad Sci U S A* 2014 Aug 19;111(33):E3415–21.
- Jinek K, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 2012 Aug 17;337(6096):816–21.
- Kaneko S, Son J, Shen SS, Reinberg D, Bonasio R. PRC2 binds active promoters and contacts nascent RNAs in embryonic stem cells. *Nat Struct Mol Biol* 2013 Nov 20;20(11):1258–64.
- Karimi MM, Goyal P, Maksakova IA, Bilenky M, Leung D, Tang JX, et al. DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. *Cell Stem Cell* 2011 Jun 3;8(6):676–87.
- Kondo T, Isono K, Kondo K, Endo TA, Itohara S, Vidal M, et al. Polycomb potentiates *meis2* activation in midbrain by mediating interaction of the promoter with a tissue-specific enhancer. *Dev Cell* 2014 Jan 13;28(1):94–101.
- Lee MT, Bonneau AR, Takacs CM, Bazzini AA, DiVito KR, Fleming ES, et al. Nanog, Pou5f1 and SoxB1 activate zygotic gene expression during the maternal-to-zygotic transition. *Nature* 2013 Nov 21;503(7476):360–4.
- Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, et al. Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* 2006 Apr 21;125(2):301–13.
- Leeb M, Pasini D, Novatchkova M, Jaritz M, Helin K, Wutz A. Polycomb complexes act redundantly to repress genomic repeats and genes. *Genes Dev* 2010 Feb 1;24(3):265–76.
- Leichsenring M, Maes J, Mössner R, Driever W, Onichtchouk D. Pou5f1 transcription factor controls zygotic gene activation in vertebrates. *Science* 2013 Aug 30;341(6149):1005–9.
- Liu X, Wang C, Liu W, Li J, Li C, Kou X, et al. Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. *Nature* 2016 Sep 14.
- Lynch MD, Smith AJH, De Gobbi M, Flenley M, Hughes JR, Vernimmen D, et al. An interspecies analysis reveals a key role for unmethylated CpG dinucleotides in vertebrate Polycomb complex recruitment. *EMBO J* 2012 Jan 18;31(2):317–29.
- Martens JHA, O'Sullivan RJ, Braunschweig U, Opravil S, Radolf M, Steinlein P, et al. The profile of repeat-asso-

- ciated histone lysine methylation states in the mouse epigenome. *EMBO J* 2005 Feb 23;24(4):800–12.
- Matsui T, Leung D, Miyashita H, Maksakova IA, Miyachi H, Kimura H, et al. Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. *Nature* 2010 Apr 8;464(7290):927–31.
- Mendenhall EM, Koche RP, Truong T, Zhou VW, Issac B, Chi AS, et al. GC-Rich Sequence Elements Recruit PRC2 in Mammalian ES Cells. *PLoS Genet.* 2010;6(12):1–10.
- Mito Y, Henikoff JG, Henikoff S. Histone replacement marks the boundaries of cis-regulatory domains. *Science* 2007 Mar 9;315(5817):1408–11.
- Newport J, Kirschner M, Bachvarova R, Davidson EH, Ballantine JEM, Woodland HR, et al. A major developmental transition in early xenopus embryos: I. characterization and timing of cellular changes at the midblastula stage. *Cell* 1982a Oct;30(3):675–86.
- Newport J, Kirschner M, Bachvarova R, Davidson EH, Ballantine JEM, Woodland HR, et al. A major developmental transition in early Xenopus embryos: I. characterization and timing of cellular changes at the midblastula stage. *Cell* 1982b Oct;30(3):675–86.
- Pezic D, Manakov SA, Sachidanandam R, Aravin AA. piRNA pathway targets active LINE1 elements to establish the repressive H3K9me3 mark in germ cells. *Genes Dev* 2014 Jul 1;28(13):1410–28.
- Puschendorf M, Terranova R, Boutsma E, Mao X, Isono K, Brykczynska U, et al. PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos. *Nat Genet* 2008 Apr;40(4):411–20.
- Riising EM, Comet I, Leblanc B, Wu X, Johansen JV, Helin K. Gene silencing triggers polycomb repressive complex 2 recruitment to CpG islands genome wide. *Mol Cell* 2014 Aug 7;55(3):347–60.
- Schmitges FW, Prusty AB, Faty M, Stützer A, Lingaraju GM, Aiwazian J, et al. Histone methylation by PRC2 is inhibited by active chromatin marks. *Mol Cell* 2011 May 6;42(3):330–41.
- Vastenhouw NL, Zhang Y, Woods IG, Imam F, Regev A, Liu XS, et al. Chromatin signature of embryonic pluripotency is established during genome activation. *Nature* 2010;464(7290):922–6.
- Wachter E, Quante T, Merusi C, Arczewska A, Stewart F, Webb S, et al. Synthetic CpG islands reveal DNA sequence determinants of chromatin structure. *Elife* 2014 Sep 26;3:e03397.
- Walter M, Teissandier A, Pérez-Palacios R, Bourc'his D. An epigenetic switch ensures transposon repression upon dynamic loss of DNA methylation in embryonic stem cells. *Elife* 2016 Jan 27; 5.
- Wu J, Huang B, Chen H, Yin Q, Liu Y, Xiang Y, et al. The landscape of accessible chromatin in mammalian preimplantation embryos. *Nature* 2016 Jun 30;534(7609):652–7.
- Yuan W, Wu T, Fu H, Dai C, Wu H, Liu N, et al. Dense chromatin activates Polycomb repressive complex 2 to regulate H3 lysine 27 methylation. *Science* 2012 Aug 24; 337(6097):971–5.
- Zaret KS, Mango SE. Pioneer transcription factors, chromatin dynamics, and cell fate control. *Curr Opin Genet Dev.* 2016;37:76–81.
- Zhang B, Zheng H, Huang B, Li W, Xiang Y, Peng X, et al. Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. *Nature* 2016 Sep 14;537(7621):553–7.

CHAPTER 7

Epilogue

Nederlandse samenvatting

Tijdens de ontwikkeling van een embryo wordt uit één enkele cel een complex organisme gevormd. Alle cellen die gevormd worden vanuit die ene cel bevatten zo goed als hetzelfde genetische materiaal, maar toch ontstaan verschillende celtypes. Hoewel de genen in ons DNA voor al onze eiwitten coderen, wordt in elke cel maar een deel van die informatie gebruikt. Op deze manier worden specifieke cellen gevormd. Het is een intrigerende vraag hoe dit in een embryo gereguleerd wordt op een plaats en tijd afhankelijke manier.

Transcriptieregulatie

Om van DNA tot eiwitten te komen, moeten genen worden gekopieerd naar RNA (transcriptie) dat vervolgens vertaald kan worden tot een eiwit. DNA bevat niet alleen genen, maar ook sequenties die de transcriptie van genen reguleren (cis-elementen) door bijvoorbeeld als aanlegplaats te functioneren voor eiwitten betrokken bij transcriptie. Enhancers (versterkers) zijn cis-elementen die op deze manier de transcriptie van genen kunnen stimuleren. Een vereiste voor het plaatsvinden van transcriptie is dat de transcriptiemachinerie (de eiwitten die transcriptie bewerkstelligen) de genen kunnen vinden en binden. De vindbaarheid en de bereikbaarheid van genen en cis-elementen verschilt van celtype tot celtype.

Er zijn verschillende eiwit(-complexen) betrokken bij het regelen van de vindbaarheid en bereikbaarheid van DNA, waaronder histonen (H). DNA zweeft niet als losse slierten rond in de celkern, maar het zit om histoncomplexen heen gewikkeld. Een histoncomplex en het DNA wat eromheen gewikkeld is, vormen samen een nucleosoom. Door nucleosomen iets uit elkaar te trekken, wordt het DNA dat zich daar bevindt beter vindbaar voor de transcriptiemachinerie. Onder bepaalde omstandigheden kunnen er bovendien kleine veranderingen worden aangebracht aan de histonen (histonmodificaties). Zo kunnen lysines (K) gemethyleerd (me) of geacetyleerd (ac) worden. Deze histonmodificaties kunnen de vindbaarheid en bereikbaarheid van het DNA ook beïnvloeden. De modificaties trekken bijvoorbeeld eiwitten aan die betrokken zijn bij de transcriptie of het compact maken of stoten ze juist af.

Histonmodificaties in *Xenopus tropicalis* embryo's

Met het onderzoek omschreven in dit proefschrift zochten we het antwoord op de vraag hoe verschillende histonmodificaties veranderen tijdens de embryo-

nale ontwikkeling van de klauwkikker, *Xenopus tropicalis*. Om deze vraag te beantwoorden hebben we de DNA-bindingspatronen van verschillende histonmodificaties in kaart gebracht op verschillende tijdstippen van de embryogenese (variërend van blastula tot gastrula). We laten zien dat histonmodificaties die geassocieerd zijn met actieve en toegankelijke genen en enhancers (H3K4me3 en H3K9ac) DNA markeren voordat modificaties die geassocieerd zijn met inactief en ontoegankelijk DNA (H3K27me3, H3K9me2/3 en H4K20me3) aanwezig zijn.

Vervolgens hebben we onze DNA-histonmodificatie kaarten geanalyseerd vanuit drie hoeken. Ten eerste bestudeerden we wie er verantwoordelijk is voor de eerste histonmodificaties in een embryo: de moeder of het embryo zelf? Ten tweede hebben we gekeken naar hoe histonmodificaties die geassocieerd zijn met inactief en ontoegankelijk DNA zich gedragen op transposons. Tot slot hebben we de nucleosoom-eigenschappen van Ezh2 bindende sequenties bestudeerd.

1. Wie is er verantwoordelijk voor de eerste histonmodificaties in een embryo: de moeder of het embryo zelf?

Tijdens de eerste paar cel delingen vindt er nog geen transcriptie plaats in embryo's. Het moment waarop de transcriptie in een embryo start heet Zygotische Genoom Activatie (ZGA). Alle processen die plaatsvinden voor de ZGA worden geregeld door eiwitten die aan de eicel zijn meegegeven door de moeder. Door embryo's te injecteren met de gifstof α -amanitine kan de transcriptie die normaal met ZGA op gang komt, worden voorkomen.

Onder normale omstandigheden markeert H3K4me3 toegankelijke stukken DNA nog voor de transcriptie in het embryo begint. Wij tonen aan dat inhibitie van de transcriptiemachinerie van het embryo de markering met H3K4me3 niet doet veranderen: in embryo's die behandeld zijn met α -amanitine wijkt het H3K4me3 patroon nauwelijks af van niet behandelde embryo's. Verrassender is onze vinding dat H3K27me3 bindingspatronen ook grotendeels gelijk zijn in α -amanitine behandelde en onbehandelde embryo's. Dit is verrassend, omdat deze histonmodificatie die in verband wordt gebracht met inactieve genen pas sterk opkomt ná ZGA. Ons onderzoek toont aan dat de invloed van de moeder op de transcriptieregulatie in een embryo tot ver na de ZGA rijkt. Deze langdurende invloed van de moeder is echter niet aan de orde voor alle histonmodificaties. Ep300 is het eiwit dat verantwoordelijk is voor de histonmodificatie H3K27ac bij actieve enhancers. We laten zien dat Ep300 binding verdwijnt in embryo's waarin de transcriptie wordt voorkomen met α -amanitine. De in het embryo aanwezige eiwitten die zijn meegegeven door de moeder kunnen Ep300 niet naar het DNA sturen.

2. Hoe ziet de bindingsdynamiek van histonmodificaties geassocieerd met transcriptioneel inactief DNA eruit op transposons?

Ongeveer de helft van ons genoom bestaat uit transposons. Hetzelfde geldt voor klauwkickers. Transposons zijn DNA elementen die zich als parasieten in het genoom van de gastheer gedragen. Deze stukken DNA hebben zich in het DNA van de gastheer gesetteld en kunnen zich verspreiden via een knip-en-plak mechanisme (DNA transposons) of door een kopieer-en-plak mechanisme (retrotransposons). De verspreiding van transposons in hun gastheer DNA kan om verschillende redenen gevaarlijk zijn voor de gastheer. Een transposon kan zich bijvoorbeeld invoegen in een gen van de gastheer, waardoor het gen verstoort wordt. Ook door in de nabijheid van genen te landen, kan een transposon de transcriptie verstoren met de cis-elementen die het bevat.

In ons onderzoek laten wij zien dat H3K27me₃, H3K9me_{2/3} en H4K20me₃ (modificaties geassocieerd met transcriptioneel inactief DNA) verrijkt zijn op een deel van de retro- en een deel van de DNA transposons. De dynamiek en combinaties van deze modificaties is verschillend voor retro- en DNA transposons: DNA transposons zijn rond de ZGA sterk verrijkt met H4K20me₃ en H3K27me₃ en deze markeringen nemen af na de ZGA, terwijl retrotransposons verrijkt zijn met H4K20me₃ en H3K9me₃ en deze markeringen nemen toe vanaf ZGA. We laten zien dat de set van retrotransposons die gemarkeerd wordt door deze histonmodificaties dezelfde set is als voor welke small RNAs aanwezig zijn in de embryo's. Verder laten we zien dat zowel retro- als DNA transposons die verrijkt zijn met repressieve histonmodificaties kunnen worden onderscheiden van de transposons die niet verrijkt zijn voor deze modificaties op basis van hun leeftijd: de transposons die wel gemarkeerd zijn met de H3K27me₃, H4K20me₃ en H3K9me_{2/3} zijn relatief nieuw in het gastheer genoom.

3. Wat zijn de karakteristieken van Ezh2 gebonden enhancers?

Ezh2 is het eiwit dat verantwoordelijk is voor de histonmodificatie H3K27me₃. We zien dat Ezh2 bindingspatronen vergelijkbaar zijn met Ep300 bindingspatronen. Onze analyse toont aan dat Ezh2, net als Ep300, verrijkt is op alle toegankelijke enhancers in *Xenopus tropicalis* embryo's. Hoewel Ezh2 bindt aan toegankelijke enhancers zien we nauwelijks H3K27me₃ op deze locaties, terwijl ze wel verrijkt zijn voor H3K27ac. Ook als Ep300 binding wordt voorkomen door embryo's te behandelen met α -amanitine neemt H3K27me₃ bij enhancers niet toe. Onze analyses benadrukken de gecompliceerde relaties tussen transcriptie-activatoren, zoals Ep300, en -repressoren, zoals Ezh2.

Kortom, ten eerste laten we zien dat embryonale transcriptie gecontroleerd wordt door een chromatine staat die door de moeder wordt gedefinieerd. Ten tweede omschrijven we de dynamiek van histonmodificaties die geassocieerd worden met inactief DNA op transposons. Ten derde laten we zien dat Ezh2 en Ep300 beide dynamisch binden aan alle toegankelijke enhancers rond de ZGA. Samengenomen geeft het onderzoek omschreven in dit proefschrift verschillende nieuwe inzichten in de regulatie van histonmodificaties tijdens de embryogenese in gewervelden.

Curriculum vitae

Ila van Kruijsbergen was born on April 11, 1989 in Gilze en Rijen, The Netherlands. Ila completed high school (Mgr. Frencken College, Oosterhout NB, The Netherlands) in 2007 after which she moved to Nijmegen to study Molecular Life Sciences at the Radboud University Nijmegen, The Netherlands. As part of the Bachelor's program (2007-2010) Ila spent a semester at Washington and Jefferson College (Washington PA, USA) and she worked in the lab of Joost Hoenderop (Nijmegen Centre for Molecular Life Sciences (NCMLS), Department of Physiology, Radboud University Nijmegen Medical Centre) where she studied *Omeprazol induced hypomagnesemia in mice*. Her Master's program (2010-2012) included two research internships: Ila studied *polycomb response elements in Xenopus* in the lab of Gert Jan C. Veenstra (NCMLS, Department of Molecular Biology, Faculty of Science, Radboud University Nijmegen); and she studied *enhancer regulation by histone H3 acetylation at lysine K27 during early developmental in zebrafish* in the lab of José Luis Gómez-Skarmeta (Centro Andaluz de Biología del Desarrollo. CSIC/Universidad Pablo de Olavide, Seville, Spain). After receiving her Master (cum laude) Ila joined the lab of Gert Jan C. Veenstra in 2012 (Radboud Institute for Molecular Life Sciences (RIMLS), Department of Molecular Developmental Biology, Faculty of Science, Radboud University Nijmegen) for her PhD studies concerning *chromatin regulation in Xenopus embryos*. Since September 2016 Ila has continued her research in the field of transcription regulation in the lab of Fred van Leeuwen (Netherlands Cancer Institute, Department of Gene Regulation, Amsterdam, The Netherlands).

List of publications

Session AM*, Uno Y*, Kwon T*, Chapman JA, Toyoda A, Takahashi S, Fukui A, Hikosaka A, Suzuki A, Kondo M, van Heeringen SJ, Quigley I, Heinz S, Ogino H, Ochi H, Hellsten U, Lyons JB, Simakov O, Putnam N, Stites J, Kuroki Y, Tanaka T, Michiue T, Watanabe M, Bogdanovic O, Lister R, Georgiou G, Paranjpe SS, **van Kruijsbergen I**, Shu S, Carlson J, Kinoshita T, Ohta Y, Mawaribuchi S, Jenkins J, Grimwood J, Schmutz J, Mitros T, Mozaffari SV, Suzuki Y, Haramoto Y, Yamamoto TS, Takagi C, Heald R, Miller K, Haudenschild C, Kitzman J, Nakayama T, Izutsu Y, Robert J, Fortriede J, Burns K, Lotay V, Karimi K, Yasuoka Y, Dichmann DS, Flajnik MF, Houston DW, Shendure J, DuPasquier L, Vize PD, Zorn AM, Ito M, Marcotte EM, Wallingford JB, Ito Y, Asashima M, Ueno N, Matsuda Y, Veenstra GJ, Fujiyama A, Harland RM, Taira M, Rokhsar DS. (2016) *Genome evolution in the allotetraploid frog *Xenopus laevis**. *Nature*. Oct 20;538(7625):336-343.

van Kruijsbergen I*, Hontelez S*, Elurbe DM, van Heeringen SJ, Huynen MA, Veenstra GJ. (2016) *Heterochromatic histone modifications at transposons in *Xenopus tropicalis* embryos*. *Dev Biol*. 2016 Sep 14. pii: S0012-1606(16)30221-4.

Hatch VL, Marin-Barba M, Moxon S, Ford CT, Ward NJ, Tomlinson ML, Desanlis I, Hendry AE, Hontelez S, **van Kruijsbergen I**, Veenstra GJ, Münsterberg AE, Wheeler GN. (2016) *The positive transcriptional elongation factor (P-TEFb) is required for neural crest specification*. *Dev Biol*. Aug 15;416(2):361-72.

Gazdag E*, Jacobi UG*, **van Kruijsbergen I**, Weeks DL, Veenstra GJ (2016) *Activation of a T-box-Otx2-Gsc gene network independent of TBP and TBP-related factors*. *Development*. Apr 15; 143(8): 1340–1350.

Bogdanović O, Smits AH, de la Calle Mustienes E, Tena JJ, Ford E, Williams R, Senanayake U, Schultz MD, Hontelez S, **van Kruijsbergen I**, Rayon T, Gnerlich F, Carell T, Veenstra GJ, Manzanares M, Sauka-Spengler T, Ecker JR, Vermeulen M, Gómez-Skarmeta JL, Lister R (2016) *Active DNA demethylation at enhancers during the vertebrate phylotypic period*. *Nat Genet*. Feb 29

Hontelez S*, **van Kruijsbergen I***, Georgiou G*, van Heeringen S, Bogdanovic O, Lister R, Veenstra GJ (2015) *Embryonic transcription is controlled by maternally defined chromatin state*. *Nat Commun*. Dec 18;6:10148

van Kruijsbergen I, Hontelez S, Veenstra GJ (2015) *Recruiting polycomb to chromatin*. Int J Biochem Cell Biol. Oct;67:177-87

van Heeringen SJ, Akkers RC, **van Kruijsbergen I**, Arif MA, Hanssen LL, Sharifi N, Veenstra GJ (2014) *Principles of nucleation of H3K27 methylation during embryonic development*. Genome Res. Mar;24(3):401-10

Lameris AL1, Hess MW, **van Kruijsbergen I**, Hoenderop JG, Bindels RJ (2013) *Omeprazole enhances the colonic expression of the Mg(2+) transporter TRPM6*. Pflugers Arch. Nov;465(11):1613-20.

Bogdanovic O1, Fernandez-Miñán A, Tena JJ, de la Calle-Mustienes E, Hidalgo C, **van Kruijsbergen I**, van Heeringen SJ, Veenstra GJ, Gómez-Skarmeta JL (2012) *Dynamics of enhancer chromatin signatures mark the transition from pluripotency to cell specification during embryogenesis*. Genome Res. 2012 Oct;22(10):2043-53

* These authors contributed equally to this work

Dankwoord

Ik bedank iedereen van de afdelingen Moleculaire Ontwikkelingsbiologie en Moleculaire Biologie voor de fijne tijd die ik heb gehad tijdens mijn promotieonderzoek. Echt!

Er zijn een aantal mensen die ik in het bijzonder wil bedanken voor hun bijdrage aan het werk gebundeld in dit boekje en dat zijn de mensen van de kikkergroep.

Gert Jan, hartstikke bedankt dat ik in je groep heb mogen werken. Ik kan je geduld en kalmte waarderen. Ik heb veel van je geleerd. Niet alleen bedankt voor je directe bijdrage aan mijn ontwikkeling en die van het onderzoek, maar ook voor dat je zo een prettige groep hebt samengesteld.

Van dat team bedank ik vooral: Saartje! Ik ben erg blij dat ik samen met jou aan deze projecten mocht werken. Saartje, jij bent een perfect voorbeeld voor mij geweest voor hoe je moet doorknallen met embryo's en voor hoe je je niet gek moet laten maken door computers. Ik heb zo veel van je geleerd! En alsof dat nog niet genoeg was, heb je ook nog eens een fijne, gezellige persoonlijkheid.

Over fijne persoonlijkheden gesproken, Siebe bedankt. Goed dat jij het U-tje van Saartje en mij kwam versterken als we kraanwagens en kiepwagens aan het spotten waren. Altijd gezellig. Je was niet alleen onmisbaar voor de sfeer, maar ook voor de algehele intactheid van het lab; één dag op vakantie en alles loopt al stroef.

Het U-tje met Saartje is niet waar ik begon. I started in between Sarita and Emese. Sarita, you know all about the plenty of possible colour schemes. Even after you left, your bright pink coat hangers brightened up the start of my days in the lab. Emese, you would always end my day in the lab with some in depth philosophical thoughts. During the time with Sarita and Emese also Asif was there, who brought some calmness, which was nice. Thanks.

When George, Matteo, Georgina and Ann started in our group the quite times were over. I can appreciate that: thanks. George and Matteo always knew how to spice things up. Georgina and Ann were always there for nice chatters.

Saartje, Matteo, Siebe and Ann, I liked it a lot to have you around to team up with you to do frog stuff. It was super nice to share the injecting, squeezing and sorting tasks, because people need to sleep sometimes. Sharing embryo work was not only nice because of practical reasons, but also to stay positive in times of diseased frogs. Ron natuurlijk ook super bedankt voor al je hulp met de kikkers. Als je in de kelder komt en Ron is daar, heb je bovendien kans op een goeie mop of raadsel: mooi meegenomen.

Behalve alle praktische hulp met de kikkers was alle praktische hulp met computers ook onmisbaar. Simon and George, super awesome that you are so patient with people that do not understand programming.

Content wise I thank Simon, Saartje, George and Matteo for their discussions when we worked in the same projects. From outside our frog group I also thank Dei, Martijn, Ken, Jin and Ozren for their collaborations.

Wie ik vooral ook niet moet vergeten te bedanken, Juliette, Thomas en Madeleine, bedankt! Erg fijn dat jullie onze groep kwamen versterken. Niet alleen omdat het leerzaam voor me was, maar ook omdat het de sfeer ten goede kwam. Leuk!

Tijdens mijn promotieonderzoek zijn niet alleen de mensen uit de kikkergroep dierbaar voor me geweest. Ik heb het geluk gehad omringd te zijn geweest met heel heel veel sprankelende persoonlijkheden: iedereen van moleculaire (ontwikkelings-) biologie. Ik bedank oprecht iedereen van jullie voor goede discussies, lunches, koffietjes en vrijdagmiddagborrels. Echt, hoor! Jullie zijn allemaal top! I thank every single colleague from molecular (developmental) biology for the high quality discussions, lunches, coffees and Friday borrels. Ook mijn nieuwe collega's in het NKI bedankt voor jullie discussies en gezelligheid tijdens het laatste stukje van het schrijven van dit boekje.

Van buiten het lab bedank ik natuurlijk ook (schoon-)familie en vrienden, omdat ze er altijd voor me zijn. Melvin, Rob en Gitte en Leon voorop: Jullie zijn als eekhoorns en eikels in het geval van een sabeltandgnoe. Ik kan je één ding vertellen: dat is erg fijn, hoor. Gitte, voor jou is misschien de vergelijking met een engeltje passender. Dat ben je niet alleen op school, dat ben je 'gewoon' altijd!

Sjors, als ik nu zou afsluiten met hoeveel ik wel niet van je houd, enzo, dat zou maar ouderwets zijn. Toch?

