

**Epigenetic programming  
of mouse pre-implantation development**

**Inauguraldissertation**

zur

Erlangung der Würde eines Doktors der Philosophie  
vorgelegt der  
Philosophisch–Naturwissenschaftlichen Fakultät  
der Universität Basel

von

**Mareike Puschendorf**

aus Deutschland

Basel, 2008

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät  
auf Antrag von

Prof. Frederick Meins  
(Fakultätsverantwortlicher)

Dr. Antoine Peters  
(Referent)

Prof. Renato Paro  
(Koreferent)

Basel, den 22. April 2008

Prof. Hans-Peter Hauri  
(Dekan)

## Summary

In mammals, fertilization triggers a cascade of events leading to the formation of a totipotent embryo from two highly specialized gametes. During this process both parental genomes undergo major epigenetic programming, suggesting a potential causal relationship between the two events. Several studies in mouse indicate that chromatin states of maternal and paternal genomes are initially highly asymmetric. At gamete fusion, maternal chromatin exists in a nucleosomal configuration marked by many distinct types of histone lysine methylation. Paternally, inheritance of nucleosomal chromatin and modifications is limited because of an extensive exchange of histones by protamines during spermiogenesis. Consequently, the paternal genome becomes *de novo* methylated at different lysine residues in a highly spatially and temporally coordinated manner after the protamine to histone exchange. The aim of this PhD project was to study the dynamics and biological function of histone lysine methylation in the maternal germ line and during pre-implantation development. In particular, I addressed the parental influence on the establishment of constitutive heterochromatin after fertilization.

In eukaryotes, *Suv39h* H3K9 tri-methyltransferases are required for pericentric heterochromatin formation and function. In early mouse pre-implantation embryos, however, paternal pericentric heterochromatin lacks *Suv39h*-mediated H3K9me3 and downstream marks. Here we demonstrate *Ezh2*-independent targeting of a maternally provided Polycomb repressive complex 1 (PRC1) to paternal heterochromatin. In *Suv39h2* deficient zygotes, PRC1 also associates with maternal heterochromatin lacking H3K9me3, thereby revealing hierarchy between repressive pathways. In *Rnf2* maternally deficient zygotes, the PRC1 complex is disrupted and levels of pericentric major satellite transcripts are increased at the paternal but not maternal genome. We conclude that in early embryos PRC1 functions as the default repressive back-up mechanism in absence of H3K9me3. Parental epigenetic asymmetry, also observed along cleavage chromosomes, is resolved by the end of the 8-cell stage, concurrent with blastomere polarization, marking the end of the maternal to embryonic transition.

In mammals, Polycomb group targeting mechanisms remain poorly understood despite genome-wide studies identifying Polycomb target genes in different cell types. Using a heterologous ES cell system, we show that PRC1 recruitment to pericentric heterochromatin is mediated by the Cbx2 chromodomain and AT hook motif. We provide first evidence that the Cbx2 AT hook is also involved in heterochromatin and possibly euchromatin targeting in early embryos. Moreover, these studies are complemented by a bioinformatics approach to explore a potential role of the Cbx2 AT hook in the recognition of PRC1 target genes.

Taken together, our studies revealed an unanticipated function of PRC1 proteins in the establishment of paternal heterochromatin in early mouse embryos. We discuss the potential role of Polycomb proteins in guiding other important events during pre-implantation development. Finally, we characterize the PRC1 binding mode to heterochromatin which may however also be involved in PRC1 recruitment to target genes.

# Table of contents

<b>Summary .....</b>	<b>1</b>
<b>Abbreviations .....</b>	<b>4</b>
<b>1. Introduction .....</b>	<b>6</b>
1.1. Constitutive heterochromatin .....	7
1.1.1. Suv39h pathway.....	7
1.1.2. Heterochromatin establishment .....	9
1.2. Polycomb group complexes .....	10
1.2.1. PRC2 and H3K27 methylation .....	10
1.2.2. PRC1 and H2A ubiquitination .....	11
1.2.3. Polycomb mediated chromatin marks are reversible .....	12
1.2.4. Targeting of PcG complexes.....	14
1.2.5. Role of PcG proteins in development, pluripotency and disease .....	16
1.2.6. Recent advances in our understanding of PcG function.....	18
1.3. Epigenetic (re)programming during early embryonic development .....	19
1.3.1. Inheritance of maternal RNAs, proteins and histone modifications .....	20
1.3.2. Paternal <i>de novo</i> establishment of a nucleosomal chromatin structure .....	24
1.3.3. Parental asymmetry in histone modifications.....	27
1.3.4. Asymmetric DNA demethylation .....	29
1.3.5. Maternal to zygotic transition in gene expression .....	31
1.3.6. Insights from nuclear transfer and cloning .....	34
1.3.7. Trans-generational epigenetic inheritance .....	37
1.3.8. Constitutive heterochromatin reprogramming.....	39
1.3.9. Pluripotency and first lineage commitment .....	41
1.4. Scope of the thesis.....	44
<b>2. Results .....</b>	<b>46</b>
2.1. Abundant transcripts from retrotransposons are unstable in fully grown mouse oocytes .....	46
2.1.1. Published manuscript.....	47
2.2. PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos.....	55
2.2.1. Published manuscript.....	56
2.2.2. Supplementary information .....	66
2.3. The Cbx2 chromodomain and AT hook are required for heterochromatin targeting of PRC1.....	79
2.3.1. Introduction.....	80
2.3.2. Methods.....	81



2.3.3. Results .....	83
2.3.4. Discussion.....	97
<b>3. General conclusions.....</b>	<b>101</b>
<b>References .....</b>	<b>104</b>
<b>Appendix .....</b>	<b>120</b>
<b>Acknowledgements.....</b>	<b>123</b>
<b>Curriculum vitae .....</b>	<b>125</b>

## Abbreviations

aa	amino acid
AT1	AT hook motif 1
bp	base pair
BrdU	5-bromo-2-deoxyuridine
BrUTP	5-bromouridine-5-triphosphate
CD	chromodomain
cDNA	complementary DNA
ChIP	chromatin immuno-precipitation
CTD	carboxy terminal domain
DAPI	4,6-diamidino-2-phenylindole
Dnmt	DNA methyl transferase
dsRNA	double-stranded RNA
EM	equatorial division followed by meridional division
EPI	epiblast
ES	embryonic stem
FCS	fetal calf serum
FISH	fluorescence <i>in situ</i> hybridization
FRAP	fluorescence recovery after photobleaching
GFP	green fluorescent protein
GO	gene ontology
GV	germinal vesicle oocyte
H2AK119ub1	histone H2A lysine 119 mono-ubiquitination
H3K27me3	histone H3 lysine 27 tri-methylation
HAT	histone acetyl transferase
HCNE	highly conserved non-coding elements
HDAC	histone deacetylase
HDM	histone demethylase
HMG	high mobility group
HMT	histone methyl transferase
HP1	heterochromatin protein 1
ICM	inner cell mass
ICSI	intra-cytoplasmic sperm injection
IVF	<i>in vitro</i> fertilization
JmjC	Jumonji C
kb	kilo base
LIF	leukemia inhibitory factor
MII	metaphase II oocyte
ME	meridional division followed by equatorial division
5meC	5-methyl cytosine

MEF	mouse embryonic fibroblast
miRNA	micro RNA
NPB	nucleolar precursor body
NPC	neural progenitor cell
NSN	non-surrounding nucleolus oocyte
NT	nuclear transfer
PcG	Polycomb group
PCH	pericentric heterochromatin
PCR	polymerase chain reaction
PE	primitive endoderm
PN	pronucleus
PRC	Polycomb repressive complex
PRE	Polycomb response element
RNAi	RNA interference
RNAP	RNA polymerase
ROSI	round spermatid injection
RT-PCR	reverse transcriptase PCR
SCNT	somatic cell nuclear transfer
SELEX	systematic evolution of ligands by exponential enrichment
SET	Suppressor of variegation, Enhancer of Zeste, Trithorax
SN	surrounding nucleolus oocyte
ssRNA	single-stranded RNA
SUMO	small ubiquitin-related modifier
TCR	transcription requiring complex
TE	trophectoderm
TE buffer	Tris-EDTA buffer
TF	transcription factor
TrxG	Trithorax group
TS	trophoblast stem
TSA	trichostatin A
TSS	transcription start site
ZGA	zygotic genome activation

# 1. Introduction

Mammalian development begins at fertilization of the oocyte by the sperm leading to the formation of a totipotent zygote that has the ability to differentiate into an array of specific cells. A coordinated program of proliferation and differentiation will subsequently result in the formation of all the distinct cell types and tissues found in the adult organism. As embryonic cells differentiate, certain genes are activated while others are repressed resulting in a unique pattern of gene expression in each cell type. This process is directed by transcription factors binding to specific promoter and enhancer sequences. In addition chromatin plays a significant role as substrate for many genetic processes inside the nuclei of eukaryotic cells. Initially, chromatin was viewed as a static entity in which DNA was packaged to condense and maintain the integrity of such a large macromolecule. However, the chromatin template undergoes dynamic changes during many genetic processes such as DNA replication and repair, recombination, cell cycle progression and during gene expression. The basic unit of chromatin is the nucleosome consisting of 147 bp of DNA wrapped around a histone octamer that is composed of two copies each of H2A, H2B, H3 and H4. It is now well established that covalent modifications at the N-termini of histones and at histone cores, like histone phosphorylation, acetylation, methylation and ubiquitination, play a major role in maintaining differential patterns of gene expression. In addition to regulating chromatin compaction, modified nucleosomes affect gene expression by providing binding sites for regulatory proteins. The information that mediates inheritance of a given expression state without inducing alteration in the DNA sequence is termed epigenetic information.

DNA methylation and histone modifications have been shown to regulate specific gene expression but they also affect the organization of chromatin to maintain the basic structure of chromosomes, and thus are important as regulators at the chromosome level. DNA methylation results from the addition of a methyl group to the 5 position of cytosine (5mC) and in mammals occurs almost exclusively at CpG dinucleotides. DNA methylation is in general associated with gene silencing and chromosome stability. The DNA methylation pattern in somatic cells is relatively stable and is maintained during each replication cycle. In contrast, histone acetylation, found on N-terminal lysine residues of histones H3 and H4, is mainly present in regions of transcriptional activation. Histone acetylation is viewed as a relatively dynamic mark that depends on a balance of the opposing activities of histone acetyltransferases (HATs) and histone deacetylases (HDACs). Another well-studied histone modification is histone methylation, which occurs on lysine and arginine residues, and exists in a mono-, di- or tri-methylated state. Histone methylation can either be activating or repressive depending on the residue of the histone tail to which the mark is added. Histone methyltransferase (HMT) activity is found in a family of enzymes with a conserved SET (Suppressor of variegation, Enhancer of Zeste, Trithorax) domain.

In all, a complex epigenetic system has evolved to coordinate and maintain tissue-specific patterns of gene expression, and thus contributes to cellular differentiation and lineage commitment during embryonic development.

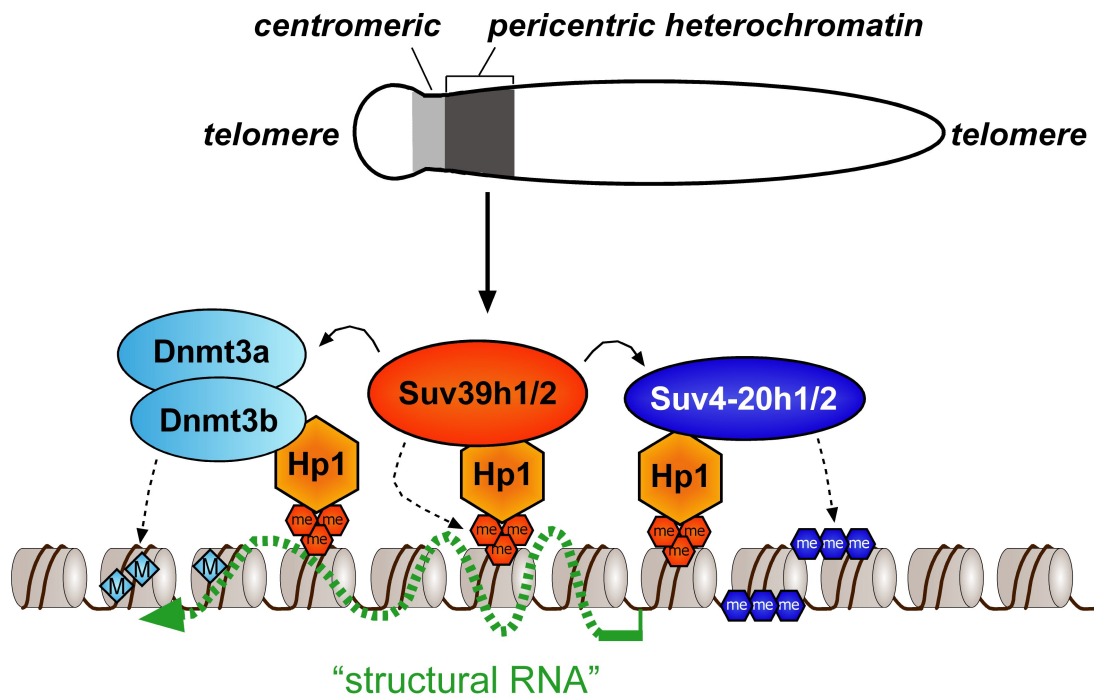
## 1.1. Constitutive heterochromatin

At the beginning of the twentieth century, cytological observations by Emil Heitz recognized regions of moss chromosomes that remained condensed and densely stained throughout the cell cycle <sup>4</sup>. These domains were named heterochromatin, while regions that de-condensed during interphase were called euchromatin. Later, the term constitutive heterochromatin was coined to describe domains that remain condensed during the cell cycle and during development, involving chromosomal regions with a high density of repetitive DNA elements like transposable elements and satellite sequences, present at telomeres and centromeres. In contrast, facultative heterochromatin, present at developmentally regulated genes or at the inactive X chromosome in female mammals, can change its chromatin state in response to developmental signals.

Centromeres are key chromosomal regions essential for chromosome segregation during mitosis and meiosis. The inner centromere, which is directly involved in kinetochore function, in mouse is composed of 120 bp minor satellite repeats <sup>5</sup>, giving rise to centromeric heterochromatin. The surrounding regions are made up of major satellite repeats consisting of 234 bp monomers that are tandemly arranged into arrays extending over a length ranging from 140 kb to more than 2,000 kb in mammalian cells <sup>6</sup>. These regions form large blocks of pericentric heterochromatin (PCH), a typical example of constitutive heterochromatin. Major satellite repeats are highly AT-rich in mouse, and therefore can be easily recognized by staining with dyes like 4,6-diamidino-2-phenylindole (DAPI), preferentially binding to AT base pairs. During interphase, PCH from several chromosomes cluster together as chromocenters, while minor satellites remain as separate entities, in close association with the chromocenters <sup>7</sup>. Centromeric sequences are not found to be conserved between various species, indicating that the repetitive nature rather than any particular DNA sequence triggers heterochromatin formation <sup>8</sup>. Once the chromatin organization is established, it has to be stably maintained through numerous cell divisions.

### 1.1.1. Suv39h pathway

Pericentric heterochromatin is characterized by a number of specific chromatin modifications, including hypoacetylation and tri-methylation at histone H3 lysine 9 (H3K9me3). In mammals, H3K9me3 is mediated by the Suv39h1 and Suv39h2 HMTs <sup>9</sup>, creating a binding site for the heterochromatin protein 1 (HP1) <sup>10,11</sup>. In mammals, three isoforms exist, HP1 $\alpha$ , HP1 $\beta$  and HP1 $\gamma$ , which are primarily associated with PCH, but HP1 $\beta$  and particularly HP1 $\gamma$  are also found at euchromatic sites <sup>12</sup>. HP1 contains an N-terminal chromodomain (CD) that recognizes the H3K9 methylation mark and a C-terminal chromo shadow domain (CSD) which functions as a dimerization module to promote self-association and compaction <sup>13</sup>. The CD and CSD are linked by a flexible hinge region. The HP1 chromodomain specifically recognizes H3K9me histone tails, which insert into a pocket of aromatic caging residues within the CD <sup>14,15</sup>. Although H3K9me3 is required for PCH targeting of HP1 <sup>11</sup>, it does not seem to be sufficient. In mammalian cells, HP1 detection at chromocenters requires a structural RNA of unknown identity <sup>12</sup>. The HP1 RNA binding domains has been shown to lie within the hinge region <sup>16</sup>.



**Figure 1: Pericentric heterochromatin.** *Suv39h*-mediated H3K9me3 directs chromatin binding of the heterochromatin protein HP1, which targets the two H4K20me2/3-specific *Suv4-20h* HMTs and the *Dnmt3a/3b* DNA methyltransferases, to establish a transcriptionally repressed state. It has been proposed that the methylated histone tails are arranged in a specific configuration dependent on a structural RNA which is required for HP1 accumulation at chromocenters, as detected by immunofluorescence.

Moreover, HP1 isoforms can be modified by various post-translational modifications, and phosphorylation of HP1 profoundly affects its chromatin and protein binding properties<sup>17</sup>. Similarly, the acetylation status of SUV39H1, which is regulated by the NAD<sup>+</sup>-dependent histone deacetylase SIRT1, has been recently shown to affect its HMT activity<sup>18</sup>. *Suv39h* HMTs are able to both methylate H3K9 and to bind to HP1, suggesting that once heterochromatin formation has been initiated, it can be propagated in a self-sustaining loop<sup>12</sup>.

Even so heterochromatin domains are perceived to be highly stable, HP1 binding to PCH is highly dynamic<sup>19,20</sup>. HP1 association with methylated H3K9 can be regulated by phosphorylation of the adjacent S10 residue, leading to dissociation of HP1 from chromatin during mitosis<sup>21,22</sup>. In mammals, HP1 interacts with the DNA methyl transferases *Dnmt1* and *Dnmt3a/b*<sup>23</sup>, directing DNA methylation to major satellite repeats<sup>1</sup>. Moreover, additional repressive histone methylation at H4K20 is mediated by the recruitment of the *Suv4-20h* HMTs<sup>24</sup>. Taken together, mammalian PCH is marked by a number of characteristic chromatin modifications (Fig. 1), establishing a transcriptionally repressed state. Loss of *Suv39h*-mediated chromatin modification leads to an increase in major satellite transcription<sup>1</sup>. Importantly, in *Suv39h*-deficient mice chromosome segregation is perturbed, indicating that proper marking of PCH is essential to ensure correct chromosome segregation and genome stability<sup>9</sup>.

### 1.1.2. Heterochromatin establishment

Most of what we know about the mechanisms of heterochromatin establishment comes from studies in the fission yeast *S. pombe*. In this organism H3K9 methylation at the outer centromeric *dg* and *dh* repeats is mediated by the Clr4 HMT, providing a binding site for the CD proteins Swi6, Chp1 and Chp2. *Clr4*-dependent H3K9me is initiated independent of Swi6, but subsequent spreading of H3K9me strictly requires Swi6<sup>25</sup>, suggesting that Swi6 when bound to H3K9me serves as assembly platform for chromatin modifying factors that are involved in spreading of PCH<sup>26</sup>. In *S. pombe*, deletion of key components of the RNAi machinery (including Dcr1, Ago1 or Rpd1) results in loss of H3K9me and Swi6 localization, and derepression of centromeric DNA repeats<sup>27,28</sup>. Chp1 is part of an RNA-induced transcription silencing (RITS) complex that also contains Ago1 and Tas3, a protein of unknown function. In addition, RITS contains siRNAs that originate from centromeric repeat regions. These siRNAs are required for efficient binding of RITS throughout centromeric repeats<sup>29</sup>. Although in conflict with heterochromatin being highly condensed and transcriptionally repressed, centromeric repeats are transcribed by RNA polymerase II to produce non-coding RNAs that could act as templates for RITS<sup>30</sup>. *S. pombe* has an RNA-dependent RNA polymerase (RdRP) which can use ssRNA templates to generate dsRNA. The RNA-dependent RNA polymerase is part of the RDRC complex which interacts with the RITS complex.

The emerging view is that siRNAs mediate the initial targeting of heterochromatin associated factors, resulting in the establishment of H3K9me. This in turn promotes the recruitment of the RITS complex to heterochromatin regions, functioning as a platform for the binding of other RNAi associated factors including RDRC, essential for the processing of nascent centromeric repeat transcripts into siRNAs. These siRNAs are part of positive feedback loop, triggering further the recruitment of the heterochromatin machinery<sup>27,28</sup>. Altogether, heterochromatin formation seems to require a delicate balance between the need for expression and the need for silencing<sup>28</sup>. Moreover, recent studies have shown that centromeric heterochromatin is not only required for the recruitment of cohesin but also to establish CENP-A on the central kinetochore domain<sup>31</sup>.

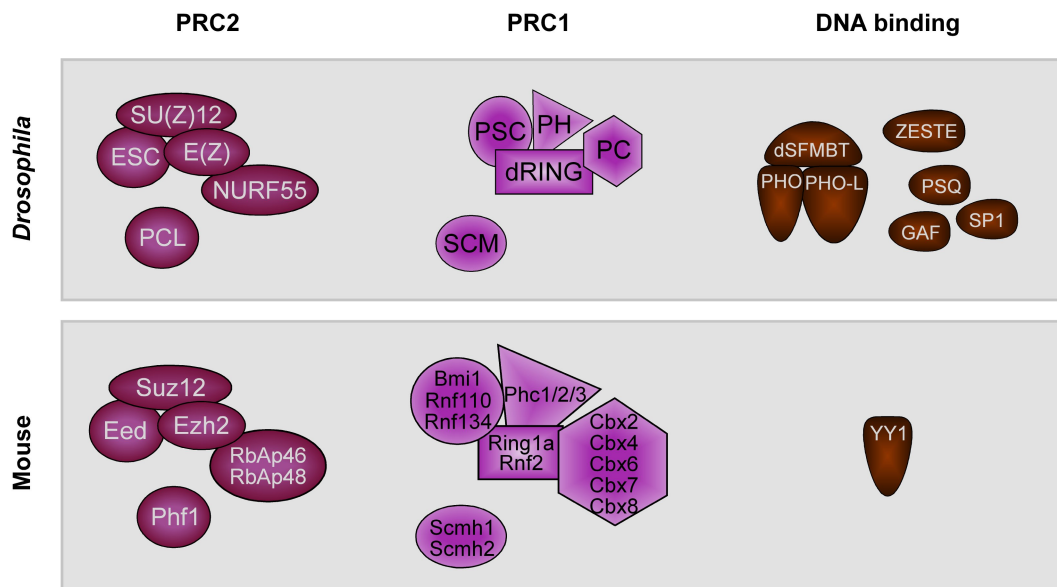
In contrast to fission yeast, *Drosophila* and mammals lack the canonical RNA-dependent RNA polymerase, suggesting that a different mechanism might be involved in heterochromatin propagation in these organisms<sup>27,28</sup>, although recent evidence indicates that the RNA polymerase II might also be able to provide RdRP activity<sup>32</sup>. Major satellite repeats are transcribed by RNA polymerase II in mammals, which is regulated in a cell cycle dependent manner<sup>1,33</sup>. Moreover, loss of *Dicer* results in the accumulation of transcripts from heterochromatin repeats, suggesting that epigenetic silencing is impaired<sup>34,35</sup>. The RNAi machinery has also been implicated in silencing and heterochromatin formation in *Drosophila*<sup>36</sup>. Clearly, further work is needed to establish the mechanisms and order of events of pericentric heterochromatin formation in mammals.

## 1.2. Polycomb group complexes

Polycomb proteins were originally identified in *Drosophila* as repressors of *Hox* genes based on their mutant phenotypes involving posterior transformations of body segments. In *Drosophila*, segmentation patterns along the Anterior-Posterior axis are established by the 'gap and pair rule' genes during the first hours of development. While the expression of these sequence specific activators and repressors is transient, the resulting transcriptional patterns persist. This memory system depends on the antagonistic function of the Trithorax (TrxG) and Polycomb (PcG) group of proteins. TrxG proteins maintain *Hox* gene expression pattern in the appropriate spatial domains, whereas PcG proteins restrict their expression. This function is conserved in vertebrates with several PcG mutants exhibiting skeletal transformations<sup>37-40</sup>. At the molecular level, PcG proteins act as repressors of gene activation through the concerted function of at least two distinct classes of multimeric complexes, termed Polycomb repressive complexes (PRCs) (Fig. 2).

### 1.2.1. PRC2 and H3K27 methylation

Key components of PRC2 in *Drosophila* include Enhancer of zeste (E(Z)), Suppressor of zeste 12 (SU(Z)12), Extra sex combs (ESC) and the nucleosome remodeling factor 55 (NURF55). PRC2 mediates repressive methylation at H3K27 through the histone methyltransferase E(Z) which does not show HMT activity alone but requires other complex members to be present<sup>41-43</sup>. E(Z) is responsible for global H3K27me1/2 and for H3K27me3 at PcG target genes<sup>44,45</sup>. High levels of H3K27me3 at target genes, but not global H3K27me1/2, require the PcG protein Polycomb-like (PCL)<sup>46</sup>, which has been proposed to act downstream of PHO/PHOL but



**Figure 2: Mouse and *Drosophila* Polycomb group complexes.** At the molecular level, PcG proteins are classified into two groups based on their association with distinct multimeric complexes, termed Polycomb repressive complexes (PRCs). In addition, several DNA binding proteins associate with Polycomb proteins in *Drosophila*, most of which are not conserved in mammals.



upstream of PRC2<sup>47</sup>.

The mammalian PRC2 homologs comprise Enhancer of zeste 2 (Ezh2), Suppressor of zeste 12 (Suz12), Embryonic ectoderm development (Eed) and the histone binding proteins RbAp46/RbAp48. While Ezh2 confers HMT activity, Suz12 is required to stabilize the complex including the recruitment of RbAp48<sup>48</sup>. In mammals, different isoforms of Eed direct the Ezh2 HMT activity towards H1K26 (PRC2), H3K27me2/3 (PRC3) and SirT1 (PRC4)<sup>49,50</sup>, though the *in vivo* functional significance of these specificities remains to be determined. Recently, the crystal structure of EZH2-EED interaction has been solved, revealing that EZH2 binds to the WD domain of EED<sup>51</sup>. So far no mammalian H3K27me1 HMT has been identified; however, one report suggests that Eed is required for all H3K27me as part of an additional complex specific for H3K27me1<sup>52</sup>. In contrast to H3K27me3 which is enriched at silent loci, H3K27me1 has been detected downstream of the transcription start of active genes<sup>53</sup>. In analogy to flies, the mammalian Polycomb-like protein PHF1 modulates EZH2 activity towards the repressive H3K27me3 mark<sup>54</sup>. Interestingly, PcG regulation is linked to DNA methylation in mammals, suggesting that different repressive modifications cooperate to establish a silent state. EZH2 has been shown to interact with DNMT1 and DNMT3A/B *in vivo* and is required for their binding and CpG methylation at PcG target promoters<sup>55</sup>.

### 1.2.2. PRC1 and H2A ubiquitination

PRC2 mediated H3K27me3 provides a docking site for the PRC1 complex which recognizes methyl-lysine residues via the N-terminal chromodomain of Polycomb (PC)<sup>15,56</sup>. PRC1 complexes purified from flies contain stoichiometric amounts of PC, Posterior sex combs (PSC), Polyhomeotic (PH) and dRING, lower amounts of Sex combs midlegs (SCM) and additional proteins including ZESTE and numerous TBP-associated factors (TAFs)<sup>57,58</sup>. In mammals, duplication of many PcG genes allows the assembly of PRC1 core complexes into various, functionally distinct sub-complexes depending on cell type and developmental stage<sup>59-61</sup>. The PC chromodomain is well conserved in mammalian Cbx homologs (Cbx2, Cbx4, Cbx6, Cbx7 and Cbx8)<sup>61</sup>, displaying distinct *in vitro* binding specificities towards H3K27me3 and H3K9me3<sup>62</sup>. In addition, the Cbx proteins contain a conserved Pc Box at the C-terminus, a 15 amino acid motif necessary for interaction with Ring1a and Rnf2<sup>61,63</sup>. The fly dRING and mammalian RING1A and RNF2 (RING1B) proteins contain a RING domain that mediates E3 ubiquitin ligase activity, resulting in mono-ubiquitination of histone H2A at K119 (H2AK119ub1)<sup>64-66</sup>. Additional RING domains are present in the fly protein PSC and in mammalian homologs Bmi1 and Mel18 (Rnf110), necessary to enhance the catalytic activity of Ring1a/Rnf2<sup>67-69</sup>. Specifically, Mel18 directs the E3 activity of Rnf2 towards H2AK119, requiring the prior phosphorylation of Mel18 for its substrate-targeting function<sup>68</sup>. Bmi1 directs self-ubiquitination of Rnf2 to generate atypical mixed chains necessary for H2A ubiquitination<sup>70</sup>.

H3K27me3 and H2AK119ub1 are thought to cooperate to mediate *Hox* gene silencing and the ubiquitination mark has been proposed to function downstream of H3K27me3<sup>64,71</sup>. Purified and reconstituted PRC1 complexes from fly and human inhibit chromatin remodeling *in vitro*<sup>58,59,72,73</sup>. PRC1 proteins induce compaction of nucleosome arrays which requires the presence

of nucleosomes but not histone tails<sup>74</sup>, suggesting a mechanism to account for the repressive activity of PcG proteins. However, there is no strong evidence to support PcG induced chromatin compaction *in vivo*<sup>75</sup>. Instead, *in vivo* binding of PRC1 proteins is dynamic, exchanging rapidly within minutes<sup>76</sup>. Recent live imaging studies suggest that mobility of PRC1 proteins increases upon induction of ES cell differentiation but decreases again as differentiation progresses<sup>77</sup>. Our understanding of how PcG proteins mediate their repressive function *in vivo* is still very limited. Access of the transcription machinery itself does not seem to be blocked; rather the activity of the transcriptional machinery at the promoter is affected by PcG proteins, preventing transcription initiation by the RNA polymerase (RNAP)<sup>78</sup>. A recent study analyzing bivalent genes in ES cells supports this idea by showing that RNAP assembles at the promoters of these genes but is held in check by PRC1-mediated H2A ubiquitination<sup>79</sup>.

Clearly, further studies are needed to elucidate how PcG proteins repress transcription at the molecular level. The classical model of PcG function suggests that the H3K27me3 mark placed by PRC2 leads to recruitment of PRC1 and ultimately H2A ubiquitination, resulting in a heritable repressed state. However, a number of reports studying different genomic contexts, including X chromosome inactivation<sup>80</sup>, gene promoters<sup>81</sup> and constitutive heterochromatin (Chapter 2.2), indicate that PRC1 targeting does not always rely on the presence of the H3K27 methylation mark.

In our way of thinking of PcG proteins as part of an epigenetic memory system, the H3K27me3 chromatin mark may provide a means to re-target the PRC complexes after mitotic divisions. Although modified histones are randomly distributed to daughter strands<sup>82</sup>, thereby diluting the chromatin marks, they could be reestablished by PRC complexes recruited through the remaining modifications. We tend to assume that PcG-mediated chromatin modifications fully account for repression, but it is possible that they are just by-products of PcG complexes delivering their function to the right place<sup>75</sup>. In light of this it is interesting to note that HMTs can modify other targets than histones<sup>83-85</sup> (reviewed in<sup>86</sup>). Moreover, the role of chromatin modifications versus histone-independent functions of PRC complexes can only be addressed by expressing catalytically dead HMTs and E3 ligases at endogenous levels. So far, there are few examples of such mutants. One allele has been described in flies with a point mutation in the catalytic domain of the RING ubiquitin E3-ligase (*Sce*<sup>33M2</sup>)<sup>87</sup>. *Sce*<sup>33M2</sup> mutant flies display homeotic transformations but to a lesser extent than *Sce*<sup>1</sup> mutants that express a truncated RING protein<sup>87</sup>, suggesting that ubiquitination of H2A contributes to PcG-mediated repression but cannot fully account for PRC1 function. Similar approaches should be applied in mammalian systems and would certainly shed further light on the role of chromatin modifications in maintaining cell identities during development as part of an epigenetic memory system.

### **1.2.3. Polycomb mediated chromatin marks are reversible**

With the recent identification of histone demethylases (HDMs) of the Jumonji (Jmj) family, our view of histone methylation as a permanent, heritable mark has changed drastically. The human genome encodes 27 proteins with JmjC domains, of which 15 have been shown to demethylate

histone lysines<sup>88</sup>. Two of these demethylases, UTX and JMJD3, are specific for H3K27me<sup>89-92</sup>. *In vitro*, both enzymes catalyze the transition from H3K27me3 and H3K27me2 to H3K27me1 on bulk histones, requiring Fe(II) and  $\alpha$ -ketoglutarate as cofactors for the oxidative demethylation reaction.

UTX is localized on the X chromosome, but escapes X inactivation in females. UTY is the male counterpart, localized on the Y chromosome, but is *in vitro* enzymatically inactive under the same assay conditions, possibly requiring different cofactors<sup>91</sup>. Inhibition of UTX in HeLa cells results in a global increase in H3K27me3, suggesting that histone demethylation is continuously required to maintain a precise level of methylation<sup>90</sup>. Consistent with the loss of H3K27me3 observed at *HOX* genes during differentiation, UTX is recruited to the promoters of several *HOX* genes, coinciding with the disappearance of H3K27me3 and decreased PRC2 occupancy<sup>90,91</sup>. Knockdown of UTX leads to increased H3K27me3, enhanced binding of PRC1 proteins and increased H2Aub at *HOX* genes, resulting in their downregulation<sup>91,92</sup>. In zebrafish, morpholino oligonucleotide inhibition of *Utx1*, one of the two zebrafish UTX homologs, leads to decreased expression of *Hox* genes and improper development of the posterior trunk<sup>91</sup>. These defects could be partially rescued by wild-type but not catalytically inactive human UTX, suggesting that demethylase activity is required for proper posterior patterning. Overexpression of *JMJD3*, but not UTX, results in global H3K27 demethylation *in vivo*, resulting in delocalization of CBX8<sup>90</sup>. *Jmjd3* is expressed in macrophages upon induction by inflammatory stimuli, suggesting that it might contribute to macrophage plasticity<sup>89</sup>. Moreover, mutation of one of the three *JMJD3* homologs in *C. elegans* results in impaired gonad development<sup>90</sup>.

In addition to HDMs, recently a number of mammalian deubiquitinating (DUB) enzymes have been characterized including the ubiquitin-specific proteases USP3, USP16 (Ubp-M), USP21 and 2A-DUB<sup>93-96</sup>. USP3 deubiquitinates both H2A and H2B, is required for S phase progression and is involved in the response to DNA double strand breaks<sup>95</sup>. In contrast, USP21 is specific for H2A. *In vitro* H2Aub specifically represses transcription initiation by inhibiting H3K4me, which is relieved by USP21-mediated deubiquitination<sup>95</sup>. The deubiquitinating enzyme USP16 might be linked to TrxG/PcG-mediated regulation, as blocking its function leads to defective posterior development in *Xenopus laevis*<sup>93</sup>.

How PcG complexes impose cellular memory when their modifications can be dynamically altered remains to be determined. In any case, a strict regulation of both the expression and recruitment of HDMs and deubiquitinating enzymes is required. Interestingly, a number of demethylases are found in complexes together with the HMTs specific for the lysine residues antagonistic to the marks removed by the HDM<sup>88,89,92,97-99</sup>. UTX has been found in a complex together with MLL2, a H3K4-specific HMT of the trithorax group of proteins<sup>92,97,98</sup>. Similarly, JMJD3 co-immunoprecipitates with RbAp5, which is an essential component of different multimeric MLL complexes<sup>89</sup>. Vice versa, the H3K4-specific demethylase RBP2 (JARID1A) co-purifies with the PRC2 complex<sup>88</sup>. The presence of both histone methyltransferases and demethylases in one complex allows switching of transcriptional states by erasing pre-existing modifications followed by their replacement with new chromatin marks – a mechanism that

might be important during developmental transitions allowing cellular identities to switch fate<sup>100-102</sup>.

#### 1.2.4. Targeting of PcG complexes

How PcG proteins are recruited to their target sequences is still poorly understood, especially in mammalian systems. In *Drosophila*, a number of DNA binding proteins have been found to associate with PcG complexes. *Drosophila* pleiohomeotic (PHO) and pleiohomeotic-like (PHOL) encode DNA binding proteins with homology to the mammalian transcription factor YY1<sup>63,103-106</sup>. PHO interacts with PRC2 and PRC1<sup>59,71,107</sup>, and complex purification identified a PHO repressive complex (PhoRC) containing the PcG protein dSFMBT<sup>106</sup>. In *Drosophila* a number of specific DNA modules, called Polycomb response elements (PRE), have been identified that control the transcriptional status of their associated promoters at a distance. Much information about the structure and function of these PREs comes from a few well characterized examples including the *engrailed*, *Fab-7* and *bxd* elements<sup>108</sup>. PREs are typically a few hundred base pairs long and contain many conserved short motifs that are recognized by DNA binding proteins. These include clusters of GAGAG bound by the GAGA factor (GAF) and pipsqueak (PSQ)<sup>109,110</sup>, binding sites for PHO/PHOL and some more PRE specific binders like SP1 (KLF)<sup>111</sup> and the HMG-like protein DSP1<sup>112</sup>, suggesting that a large number of factors might contribute to PcG targeting to PREs.

##### *Polycomb response elements (PREs) in Drosophila*

The number and topological order of consensus motifs for these proteins are not conserved between PREs<sup>75,108</sup>. Ringrose and colleagues have used an algorithm to search for putative PREs using the consensus binding sites for GAF, PHO and ZESTE as basis<sup>113</sup>. In addition, they took into account that clustered motifs may allow more effective binding. This approach identified more than hundred presumptive PREs in *Drosophila*, of which several candidates could be confirmed by ChIP and in transgenic assays<sup>113</sup>. Since then a number of large-scale PcG binding studies became available, comprising different cell types and developmental stages<sup>45,114,115</sup>. Taken together, only approximately 20% of the identified PcG binding sites were predicted by the algorithm<sup>108</sup>, suggesting that *in vivo* other proteins than GAF, PHO and ZESTE may recruit PcG. This is supported by a study indicating that a synthetic PRE requires DSP1 binding sites in addition to GAF, PHO and ZESTE to support PRE function<sup>112</sup>. Conversely, many predicted PREs were not bound by PcG proteins in these assays, including some that were confirmed in transgenic approaches, possibly reflecting tissue and stage specific usage of PREs during development, or technical limitations in the detection of weak PcG binding.

One classical hypothesis suggests that PRE binding factors recruit PRC2 proteins and the ensuing H3K27me3 would in turn recruit PRC1<sup>75</sup>. Evidence for sequential binding comes from studies of the *ubx* PRE where PHO/PHOL recruit E(Z) leading to H3K7me and subsequent PRC1 targeting<sup>71</sup>. Moreover, inactivation of E(Z) results in the loss of PRC1 binding from polytene chromosomes<sup>42</sup> and likewise PC binding can be competed with H3K27me3 peptides<sup>116</sup>. In contrast, ChIP data in flies argues against such a model<sup>45,117,118</sup>. While PRC2 and PRC1

complexes peak sharply at known PREs, H3K27me3 is distributed over the entire transcription unit and upstream regulatory regions, often spanning several 10 kb. Although the distribution of PC is broader than that of other PRC1 members, PC declines gradually from PREs and does not parallel the H3K27me3 mark. In light of several ChIP and nuclease mapping studies suggesting that PREs are depleted of nucleosomes *in vivo*, it is interesting that PHO has been suggested to wrap the PRE DNA around its own surface, rather than the histone core <sup>119</sup>.

Although a number of PREs are known in flies and more PcG binding sites have been identified by genome-wide mapping approaches, the molecular mechanisms of PcG targeting are not well characterized. It remains open how PcG present at PREs, which are often far away from their target promoters, can induce silencing. It has been suggested that PRE-bound complexes loop out to contact their target promoter. In *Drosophila*, silent endogenous and transgenic PREs have been found to associate, with long distance nuclear interactions enforcing PcG mediated silencing <sup>120-123</sup>. Moreover, PREs and PcG target promoters cluster and localize to PcG bodies <sup>121</sup>. But then, how do PcG proteins interfere with transcriptional activation? At PREs a strong correlation can be established between H3K27me3 levels surrounding the PRE and PC binding to the PRE, but there is not always a clear correlation found to their transcriptional activity <sup>116</sup>. Likely other chromatin determinants contribute to silencing of PcG target genes. The RNAi machinery is involved in the maintenance of long-range contacts of transgenic *Fab-7* copies and the association of endogenous homeotic genes, suggesting that RNAi might increase PcG-mediated silencing through an effect on nuclear organization <sup>124</sup>. The RNAi machinery does, however, not seem to be involved in the recruitment of PcG complexes to their target sequences.

#### *PcG targeting mechanisms in mammals*

So far, no PREs have been identified in mammals, despite the availability of large-scale ChIP data sets for chromatin modifications and different PRC2 and PRC1 members <sup>2,53,125-127</sup>. Likely, PcG target regions do not contain a simple arrangement of consensus motifs. The *Drosophila* DNA binding proteins GAF, PSQ and ZESTE are not conserved in mammals. Instead, DNA binding transcription factors might hold the key for PcG recruitment to their target genes in different cell types and at different developmental stages. In ES cells, there is a significant overlap between genes bound by the pluripotency transcription factors OCT4, SOX2 and NANOG and PcG target genes <sup>127</sup>.

Another valid approach that has been taken recently to identify regulatory regions in mammalian genomes is the search for highly conserved non-coding elements (HCNE). Comparison of human and chimpanzee genomes identified approximately 200 genomic regions where the CpG distribution is far more conserved than predicted <sup>128</sup>. Interestingly, many of these HCNEs coincide with domains bound by PRC2. Moreover, most of the HCNEs identified by comparison of human and the pufferfish *Fugu rubripes* are located around genes that act as developmental regulators <sup>129</sup>. However, at the global scale, only 8% of HCNE regions are bound by Suz12 and the highest peaks of PcG binding do not appear to correlate strongly with regions of highest conservation <sup>108</sup>. The small percentage of HCNEs bound by PcG does not necessary

argue against their role in PcG targeting, as likely other not PcG related regulatory sequences will be conserved. Functional tests of such elements will be required to elucidate their contribution to PcG recruitment.

Yet another interesting possibility for PcG targeting comes from the analysis of transcription across human *HOX* gene clusters that identified a large number of non-coding RNAs<sup>130,131</sup>. Sessa and colleagues suggest that intergenic antisense transcription might be important for the opening and maintenance of the active state at adjacent *HOXA* genes<sup>131</sup>. In analogy, in *Drosophila* it has been shown that non-coding transcription through a PRE prevents the establishment of PcG silencing<sup>132</sup>. Transcription through endogenous PREs can be detected at different stages of fly development, suggesting that it is required continuously as anti-silencing mechanism. In contrast, Rinn and colleagues identified a 2.2 kb non-coding RNA in the human *HOXC* locus, termed HOTAIR, which represses transcription in *trans*<sup>130</sup>. Most importantly, HOTAIR interacts with PRC2 and is involved in PRC2 recruitment and H3K27me3 of the *HOXD* locus. Clearly, a lot of work is still needed to fully understand PcG targeting in *Drosophila* and to get to targeting mechanisms in mammals. Probably a number of different mechanisms will collaborate to account for cell type specific regulation.

#### **1.2.5. Role of PcG proteins in development, pluripotency and disease**

##### *Polycomb functions during development*

Mouse mutants for any PRC2 members result in embryonic lethality<sup>48,133-135</sup>. They either cease development after implantation or initiate but fail to complete gastrulation, demonstrating that Polycomb function is essential for proper embryonic development. In addition, mice generated from oocytes lacking maternal *Ezh2* but expressing *Ezh2* from the paternal allele (resulting in *Ezh2* protein from the 4-cell stage onwards) are post-natally growth retarded<sup>136</sup>. This study by Erhardt and colleagues has often been held as an example in the literature for the very early requirement of PcG proteins during development, even before implantation. The idea is very tempting that the transient lack of a chromatin modifier early in development can have such long-term effects on the developmental outcome. Unfortunately, we were not able to reproduce these data with our maternal *Ezh2* mutant embryos (unpublished results) that were generated with a different conditional targeting strategy (Chapter 2.2), arguing that the different alleles might affect molecular mechanisms to a different degree.

Deletion of several PRC1 members leads to homeotic transformations<sup>37-40</sup>. The most severe phenotype, resulting in embryonic lethality, is observed for ablation of *Rnf2*<sup>137</sup>, indicating that PcG silencing requires both PRC complexes during development. Deletion of *Rnf2* results in gastrulation arrest and cell cycle inhibition. With respect to the differences in the severity of *Rnf2* and *Ring1a* phenotypes<sup>38,137</sup>, it is interesting that global H2A ubiquitination levels are drastically reduced in *Rnf2* but not *Ring1a* mutant ES cells<sup>138</sup>. The less severe phenotypes of the other PRC1 members might be to some extent explained by functional redundancy of the different mouse PRC1 homologs.

### *Polycomb proteins and the regulation of pluripotency*

In addition to the embryonic defects of PcG mutants, PRC1 members play a role in the maintenance and proliferation of adult stem cells<sup>139,140</sup>. Deletion of *Bmi1* results in loss of hematopoietic stem cells and also affects neuronal and mammary stem cells<sup>140-143</sup>. Moreover, PRC2 proteins seem to be involved in maintaining the pluripotent state of embryonic stem cells. ES cells deficient for *Eed* tend to differentiate spontaneously, whereas *Suz12* mutant ES cells can be derived and expanded in culture but have impaired differentiation potential<sup>81</sup>. O'Carroll and colleagues could not derive ES cells from *Ezh2* mutant blastocysts, suggesting that *Ezh2* is required to maintain the pluripotent state<sup>135</sup>. We used our conditional mice (Chapter 2.2) to establish *Ezh2* F/F ES cells. Transfection of these cells with Cre expression plasmids results in *Ezh2* mutant ES cells that can be normally maintained (M.P. and E. Posfai, unpublished). Maybe *Ezh2* is required for the generation of ES cells from blastocysts but dispensable thereafter? If so, then *Ezh2* is likely to have H3K27me3-independent functions during ES cell derivation as *Suz12* mutant blastocysts devoid of global H3K27me3 give rise to ES cells.

Large scale ChIP studies during recent years confirm that PcG target genes do not only include *Hox* genes but also other key regulators of development<sup>2,125,144</sup>. It seems that in any given cell, most alternative transcription programs, not required for this cell type, are repressed by PcG silencing mechanisms<sup>75</sup>. However, not all of the genes bound by PcG proteins are destined to be shut down permanently<sup>145</sup>. Many of the developmental transcription factors repressed in ES cells become expressed upon entry into specific differentiation programs. Given this flexibility, the discovery of so called 'bivalent domains' in ES cells<sup>144</sup>, that are marked by the presence of both active H3K4me and repressive H3K27me chromatin marks, was taken up with much excitement. It was suggested that bivalent domains silence developmental regulators in ES cells while keeping them 'poised' for later activation<sup>144,146</sup>. Subsequent studies indicate that most of the genes in ES cells fall into three groups<sup>3,147,148</sup>. Genes that are only marked by H3K4me3 are associated with high levels of expression and mainly comprise genes involved in the general metabolism. In contrast, the second group of genes, marked by both H3K4me3 and H3K27me3, is expressed at low levels<sup>79,126</sup> and is characterized by developmental regulators. Genes lacking either mark are generally not expressed and related to more tissue-specific functions. Upon differentiation of ES cells into neural progenitor cells (NPCs), approximately half of the bivalent domains are resolved to H3K4me3 alone and become highly expressed in NPC<sup>3</sup>. The remaining genes retain their low expression and either resolve into H3K27me3 alone or loose both marks, whereas only a small percentage remains bivalent<sup>3</sup>. Thus, bivalency does not only predispose genes for future activation but can also result in repression. Moreover, bivalent domains are not unique to ES cells as initially thought but are also found in differentiated cells, like NPCs, MEFs and T cells<sup>3,149</sup>. However, the number of genes marked by bivalency in differentiated cells is much lower compared to ES cells, suggesting that the resolution of these domains is closely related to the commitment of cells. With progressive commitment, lineage choices become increasingly limited by the resolution of 'poised' promoter states into either ON or OFF states<sup>145</sup>.

### *Polycomb proteins in cancer*

The PRC1 protein Bmi1 has been identified as a proto-oncogene that cooperates with Myc to promote the generation of B- and T-cell lymphomas<sup>150,151</sup>. Bmi1 inhibits Myc induced apoptosis through the repression of the Ink1a-Arf (Cdkn2a) locus<sup>152,153</sup>. INK4A encodes a cyclin-dependent kinase inhibitor that activates the retinoblastoma (RB) pathway, while ARF induces the transcription factor p53<sup>154</sup>. These proteins need to be activated upon stress, oncogene activation or senescence. The INK4A-ARF locus is also targeted by number of other PcG proteins, including CBX4, CBX7, CBX8, MEL18, RNF2, EZH2 and SUZ12<sup>137,152,155-157</sup>. CBX7 was shown to extend the life span of human and mouse cells by bypassing replicative senescence through downregulation of the INK4A-ARF locus<sup>157</sup>. Similarly, ectopic expression of *CBX8* leads to repression of the Ink4a locus and immortalization of mouse embryonic fibroblasts<sup>155</sup>. However, PcG proteins might also function through alternative pathways, as INK4A-ARF independent effects on cellular proliferation were reported for EED and PHC1<sup>140,158</sup>. There is evidence that PcG proteins are also involved in cell cycle regulation in *Drosophila*, suggesting that the *Cyclin A* gene is directly controlled by PcG proteins<sup>159</sup>.

The importance of proper levels of PcG expression is highlighted by the fact that several PcG proteins are misregulated in different types of human cancers<sup>160</sup>. For example, SUZ12 overexpression has been found in colon and breast cancers<sup>161</sup>. Moreover, EZH2 expression is associated with poor prognosis of cancer<sup>162,163</sup>. Strikingly, a significant proportion of PcG targets genes is silenced in human tumors by DNA methylation, suggesting that PRC2 could promote tumorigenesis by specifically recruiting DNA methyltransferases to tumor-suppressor genes<sup>160</sup>. Another interesting hypothesis suggests that PcG proteins may influence tumor development through 'misspecification' of cells towards the stem cell fate, particularly attractive in light of the 'cancer stem cell' hypothesis<sup>160</sup>.

#### **1.2.6. Recent advances in our understanding of PcG function**

One of the fundamental mechanistic challenges in development represents the need to maintain the transcription pattern of key regulatory genes throughout the life time of an organism<sup>164</sup>. The Polycomb proteins were identified as part of a cellular memory system that perpetuates *Hox* gene expression pattern after they have been established during embryogenesis. Recent studies in mammalian systems have, however, implicated Polycomb proteins in much more widespread functions in many different cell types. Especially the identification of key regulators of different developmental programs among PcG target genes suggests that PcG proteins may function already much earlier in development, before *Hox* gene expression patterns are established (~E6.5 in mouse). Along these lines, it is interesting to note that H3K27me3 is asymmetrically localized in inner and outer cells of morula and blastocyst embryos<sup>136</sup>. It will be exciting to dissect the role of PcG proteins during pre-implantation development when the first lineage choices are made.

Since the discovery of histone demethylating enzymes, our way of thinking of PcG mediated chromatin modifications as stable, heritable marks has been challenged. Moreover, active and repressive histone methylation can co-exist on the same promoter, indication that the



transcriptional outcome is mediated by a fine balance of two opposing systems<sup>108</sup>. Interestingly, a theoretical model predicts that histone modifications can be highly dynamic without compromising stability<sup>165</sup>. It suggests that bistable systems are resistant both to high noise and to random partitioning upon DNA replication.

In addition, the concept of epigenetic stable silencing, that is established during early development and maintained throughout life, does not seem to be a general rule<sup>75</sup>. Several examples<sup>100-102</sup> show that such a regulatory system can guide commitment and differentiation into different lineages, but at the same time remains flexible to respond to certain environmental and developmental cues<sup>166</sup>. Bivalent domains have been suggested to provide means to postpone the decision of gene activation or repression to later stages of development<sup>145</sup>. Accordingly, commitment of cells to a permanent cell fate may require recruitment of additional repressive mechanisms to manifest silencing.

Some of the questions that remain to be tackled include: How do PcG proteins induce silencing at the molecular level? How are they recruited to their target promoters in mammalian systems? What allows PcG proteins to promote heritable silencing on the one hand, but simultaneously permits, at least to some extent, that cell fates can be switched? Presumably there must be another layer of regulation, possibly through association with specific transcription factors or modulators, allowing PcG proteins to do justice to these diverse needs.

### **1.3. Epigenetic (re)programming during early embryonic development**

In mammals, development begins with the fertilization of the oocyte by the sperm and thus, brings together two of the most highly specialized cells found in an organism. The oocyte is arrested at meiotic metaphase II and is loaded with a pool of maternally stored transcripts and proteins, required to support early embryonic events up to zygotic genome activation. In contrast, sperm with its highly condensed chromatin, organized in a protamine-based structure, carries only very few transcripts and proteins. Despite the similar genetic content of the gametes, differences in epigenetic modifications reinforce their distinctive nature. Following fertilization, these cells undergo a series of reorganization events and epigenetic changes. In mouse several studies show that parental genomes are initially highly asymmetric with respect to their DNA and histone methylation status. But what is the significance of this epigenetic asymmetry for proper pre-implantation development? Does it influence zygotic genome activation and therefore gene expression from maternal and paternal genomes? Do chromatin modifications contribute to the establishment of the pluripotent state in early embryos? And what is their role in guiding first lineage decisions which are accompanied by differences in the expression of key transcriptional regulators?

The following sections aim to summarize our current knowledge of chromatin dynamics in the maternal germ line and following fertilization at maternal and paternal genomes. So far, most data available remains descriptive but functional links are emerging from studies using conventional and conditional knock out and knock down approaches. The potential roles of chromatin modifications in guiding major events occurring during early development are discussed.

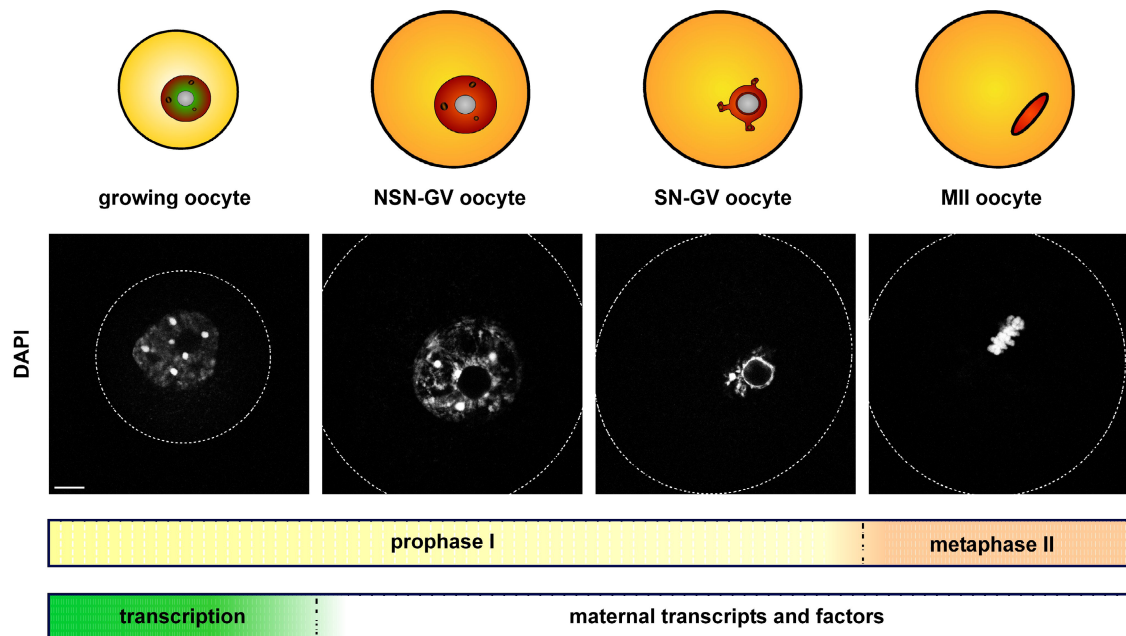
### 1.3.1. Inheritance of maternal RNAs, proteins and histone modifications

The oocyte is a unique cell whose development is characterized by a period of active meiotic programming and a long period of meiotic arrest. In addition, a period of high transcriptional and translational activity is followed by a phase of relatively low metabolic activity <sup>167</sup>.

In mammals, the initial stages of oogenesis take place during fetal development when the primordial germ cells enter the genital ridges. They begin to divide mitotically to create a large supply of oogonia, which then enter into meiosis and start forming primary follicles. Of the oocytes initially present, at least 50% die but some 10,000 remain to provide the female with functional gametes for her entire reproductive life. Follicular growth and oocyte maturation are induced by follicle stimulating hormone (FSH). Growing oocytes actively transcribe many genes and accumulate proteins, thereby creating in its cytoplasm a store of maternal factors necessary to bridge the period of transcriptional silence that begins with the completion of oocyte growth and lasts until zygotic genome activation. The fully grown oocyte, also referred to as germinal vesicle (GV) oocyte, is induced to mature by the lutenizing hormone (LH). It completes the first meiosis and then arrests in metaphase of the second meiotic division, waiting to be fertilized by the sperm.

#### *Chromatin remodeling during oogenesis*

During its phase of growth and maturation the oocyte undergoes a number of changes in nuclear architecture and chromatin state (Fig. 3). At the GV stage, two types of oocytes can be



**Figure 3: Oocyte maturation.** Growing oocytes are transcriptionally active but enter transcriptional quiescence with completion of oocyte growth. At the germinal vesicle (GV) stage, two types of oocytes are distinguished based on their chromatin organization either in a non-surrounding nucleolus (NSN) or surrounding nucleolus (SN) configuration. Upon hormonal stimuli, GV oocytes resume meiosis, complete the first meiotic division and arrest at the metaphase of the second meiotic division (MII), ready to be fertilized by sperm. Scale bar, 10  $\mu$ m.

distinguished on the basis of their chromatin organization<sup>168</sup>. The first one is referred to as non-surrounded nucleolus (NSN) oocyte and is characterized by a diffuse chromatin configuration, predominantly found in growing diplotene stage oocytes. Upon completion of oocyte growth, chromatin becomes progressively condensed around the nucleolus which is therefore called surrounded nucleolus (SN) configuration. NSN-oocytes represent an immature form of oocytes, not yet ready to be fertilized, whereas SN-oocytes are capable to support full development. The transition from NSN- to SN-oocytes is accompanied by the cessation of transcriptional activity. Centromeres, which are initially spread within the nucleus, progressively cluster around the periphery of the nucleolus in SN-oocytes, resulting in a ring of pericentric heterochromatin (PCH). PCH shows reduced H4K5 acetylation<sup>169</sup>, enriched incorporation of macroH2A<sup>170</sup> and is bound by the ATP-dependent SWI/SNF chromatin remodeler Atrx that is required for proper chromosome segregation during meiosis<sup>169</sup>. Chromatin remodeling does not seem to be required for transcriptional repression as in oocytes that are mutant for the nuclear chaperone Nucleoplasmin 2 (Npm2), transcription is shut down although nucleolar heterochromatin organization and chromatin compaction are impaired<sup>168,171,172</sup>. In addition, chromatin decompaction within euchromatin, induced by treatment with the histone deacetylase inhibitor trichostatin A (TSA), does not restore transcriptional activity in fully grown GV-oocytes<sup>168</sup>.

Even so GV-oocytes are not transcriptionally active, their histones are acetylated at K9 and K14 of histone H3 and at lysines K5, K8, K12 and K16 of H4<sup>173-175</sup> – modifications that are associated with an active chromatin state. All histone acetylation marks disappear within hours after resumptions of meiosis and acetylation remains low until after fertilization, except for H4K12 acetylation which transiently reappears after first polar body extrusion on condensed metaphase-I chromosomes<sup>173</sup>. Histone deacetylation is dependent on Hdac activity and incubation of oocytes with TSA results in a hyperacetylated state which delays germinal vesicle breakdown<sup>176</sup>. In addition, hyperacetylation induced chromatin decondensation disturbs chromosome interactions with the microtubular network and proper alignment of chromosomes during meiotic metaphase resulting in frequent aneuploidy in zygotes<sup>168,176,177</sup>. Strikingly, histone hyperacetylation is also observed in oocytes of aged female mice<sup>177</sup> suggesting that the high incidence of aneuploidy in older pregnancies (also observed in humans) might be due to inadequate histone deacetylation. Histone deacetylation during meiosis is conserved in pig oocytes, where it is required for phosphorylation of H3S10 and H3S28 to occur<sup>178</sup>. Like in mouse, histones become reacylated after oocyte activation in pig. Immunofluorescence with a newly characterized antibody recognizing dual di-methylated and phosphorylated H3K9/S10 and H3K27/S28 showed that mouse metaphase chromosomes become also phosphorylated during meiosis I and II<sup>179</sup>. This is however not specific for meiosis, as the H3K9me2/S10ph and H3K27me2/S28ph sites are also involved in the acquisition of a specific chromatin conformation during mitosis<sup>179</sup>. Taken together, both histone acetylation and phosphorylation undergo dynamic changes during mouse and pig oocyte meiosis suggesting that the underlying mechanisms are conserved in different species and might therefore serve an important function during meiosis.

The kinetics of histone deacetylation suggest that it is probably not required to establish the transcriptional quiescent state in mature oocytes but might contribute to manifest silencing. Interestingly, by the end of the first meiotic cycle, a large number of transcription factors and chromatin associated proteins become dissociated from chromatin and are only retargeted after fertilization before first zygotic genome activation<sup>180</sup>. Together these mechanisms might facilitate the transition from a maternal to a zygotic transcription program by erasing previously set transcriptional marks. At the global level, both active as well as repressive histone lysine methylation marks are thought to remain stable during oocyte meiosis. However, this assumption is probably based on the fact that histone marks are detectable at maternal chromatin prior to fertilization<sup>169,181</sup> (our data, Chapter 2.2), as the dynamics of few histone methylation marks have been studied in detail during oogenesis<sup>182</sup>. Moreover, it is not known how much reprogramming of histone methylation is occurring at the level of individual promoters or regulatory sequences and whether changes in histone methylation contribute to the oocyte to embryo transition.

#### *Maternal transcripts are required for embryonic development*

At fertilization, the oocyte is loaded with a huge store of maternal transcripts. In both, human and mouse oocytes more than 5,000 different transcripts can be detected<sup>183,184</sup>. Microarray analysis has shown that of the transcripts accumulated during oocyte growth only ~25% are specifically degraded from the GV to the MII stage<sup>185</sup>. This degradation is a selective process and targets transcripts involved in meiotic arrest at the GV stage and progression of oocyte meiosis. Overrepresented processes include oxidative phosphorylation, energy production and protein metabolism which are among other things required to generate the high levels of cAMP from ATP necessary to keep GV oocytes arrested. Other transcripts that are not required during oocyte maturation are masked by deadenylation generating short poly A tails to maintain mRNA stability for long periods of time<sup>167,186,187</sup>. In *Xenopus* oocyte, three different *cis* regulatory elements in the 3' untranslated region of maternal mRNAs regulate their polyadenylation<sup>188</sup>. In mouse, the oocyte-specific RNA-binding protein Msy2 plays an important role in the translational repression and storage of maternal mRNAs<sup>189</sup>.

Maternal transcripts serve important roles during oogenesis and early embryonic development including embryonic genome activation. Gene knockout approaches during recent years have identified a limited number of murine "maternal effect genes" which refer to the dependence of early embryonic development on maternal products and their deletion results in female sterility<sup>167,190</sup>. Unlike in invertebrates and lower vertebrates, the molecular mechanisms of mammalian maternal effect genes are not well understood yet and few have been studied in detail. For example, *Zar1* was identified as a gene exclusively expressed in the growing oocyte and accordingly, *Zar1*<sup>-/-</sup> mice are viable and grossly normal but *Zar1*<sup>-/-</sup> females are sterile<sup>191</sup>. Oocytes from *Zar1*<sup>-/-</sup> mice progress normally through oogenesis, but early embryonic development arrests at the 1- or 2-cell stage though the exact mechanisms remain elusive. Similarly, oogenesis is unaffected in females mutant for *Maternal antigen that embryos require (Mater)*<sup>192</sup>, a factor that has been identified as an oocyte antigen in a mouse model of

autoimmune premature ovarian failure. Like in mice mutant for the ubiquitin-conjugating DNA repair enzyme *Hr6a*<sup>193</sup> or the heat shock factor *Hsf1*<sup>194</sup> early embryos from these females arrest at the 1- or 2-cell stage.

Another maternal effect gene, whose molecular role is still unclear in mouse, is the nucleoplasmin *Npm2*<sup>171</sup>. In *Npm2*<sup>-/-</sup> oocytes, DNA appears amorphous and diffuse and does not condense around the nucleolus, which does however not impair meiotic progression. In *Xenopus laevis*, NPM2 removes sperm protamines, facilitates nuclear assembly and replication of the paternal genome. In contrast, sperm decondensation occurs normally in *Npm2* maternal mutant mouse embryos; however, no nucleoli are visible in zygotes, first mitosis is delayed followed by fragmentation and death of most embryos. The fact that no nucleoli are formed is very interesting in light of a very recent study showing that maternal nucleoli are essential for embryonic development<sup>195</sup>.

Knockdown of *Basonuclin* in growing oocytes using a Zp3-hairpin approach also results in female sub-fertility<sup>196</sup>. This is not strictly a maternal effect gene as oogenesis is also affected. Basonuclin is a zinc-finger protein involved in the transcription of rRNA. In oocytes, Basonuclin co-localizes with RNAP I activity in the nucleus, however, Basonuclin is also abundant in the nucleoplasm and interacts with RNAP II promoters. In *Basonuclin* knockdown oocytes, RNA polymerase I and II mediated transcription and normal oocyte morphology are perturbed. However, some oocytes do mature and are capable to support fertilization. In the resulting embryos chromatin decondensation of the paternal pronucleus is decreased, chromatin is frequently observed to fragment and DNA is unequally distributed between daughter cells resulting in embryonic arrest at the 2-cell stage. Other maternal factors include *Brg1*, *Tif1α*, *Stella* (*PGC7*) and *Dnmt1* that will be discussed in detail later.

It is striking to note that most of these maternal affect genes lead to embryonic arrest at the late 1- or 2-cell stage. Similar observations were made in early *in vitro* culture studies, where embryos from the majority of inbred and outbred mouse strains underwent cleavage arrest at the 2-cell stage<sup>197</sup>. This stage represents a crucial and very sensitive period of early embryonic development in mouse as it coincides with first zygotic genome activation and therefore marks the maternal to embryo transition.

#### *Role of maternal micro RNAs*

In addition to protein coding RNAs, small regulatory RNAs are maternally inherited from the oocyte. MicroRNAs (miRNAs) are single-stranded RNA molecules of about 21-23 nucleotides in length with 3' two-nucleotide overhangs and are complementary to sites in the 3' UTR of their target messages<sup>198</sup>. miRNAs downregulate gene expression by inhibition of protein translation or by induced cleavage of the target mRNA. miRNAs are first transcribed as primary transcripts (pri-miRNA) and processed to 70 nucleotide stem-loop structures known as pre-miRNAs by the RNase III nuclease Drosha. These pre-miRNAs are further processed to mature miRNAs by the endonuclease Dicer which also initiates the formation of the RNA-induced silencing complex (RISC). The endonuclease Dicer is expressed in growing and mature oocyte as well as in early embryos<sup>199</sup>. Zp3-Cre mediated deletion of *Dicer* in growing oocytes results in loss of most

maternal miRNAs. The mutant oocytes mature and undergo GV breakdown but are defective in meiotic spindle organization<sup>199,200</sup>. Some *Dicer* mutant oocytes are capable to support fertilization but do not progress through the first cell cycle<sup>200</sup>. Microarray analysis showed that *Dicer* is essential for the turnover of a substantial subset of maternal messages that are normally degraded during oocyte maturation. Interestingly, the upregulated set includes a number of genes involved in microtubule associated processes which might explain the observed chromosome segregation defect. Furthermore, expression of certain transposon-derived sequences is elevated in *Dicer* mutant oocytes which is interesting in light of another recent study that identified a novel class of 20-24 nucleotides long small RNAs from oocytes that are derived from retroelements<sup>201</sup>. The functional significance of these siRNAs is not known but it has been proposed that they might act to suppress retrotransposition events in germ cells.

Taken together, at fertilization the oocyte is equipped with a pool of maternal messages, small RNAs and proteins, some of which are absolutely essential to support the earliest stages of mouse embryonic development. Maternal chromatin is inherited in a chromosomal configuration and is globally characterized by various histone acetylation and methylation marks<sup>202</sup>. Although several transcription factors and histone modifiers are not chromatin bound at fertilization, they are present in the oocyte cytoplasm, ready to be retargeted before zygotic genome activation is initiated.

### **1.3.2. Paternal *de novo* establishment of a nucleosomal chromatin structure**

The production of male gametes is a result of a complex multistep process of cellular differentiation. During spermatogenesis, the paternal genome undergoes major changes including meiotic recombination and chromosome segregation<sup>203</sup>. The aim of meiosis is to produce great genetic diversity within a population to allow a species to maintain stability under environmental fluctuations. Chromatin changes are characterized by the appearance of testis-specific histone variants<sup>204</sup>, the silencing of the XY-body by acquisition of heterochromatic histone modifications<sup>9,205-207</sup>, the epigenetic marking of imprinted genes<sup>208</sup> as well as the post-meiotic packaging of the genome in a highly condensed chromatin structure<sup>203,204</sup>. The final product, the mature spermatozoon, is designed for the safe delivery of a haploid copy of the paternal genetic information to the future zygote.

#### *Chromatin packaging in mature sperm*

High density packaging of the paternal genome is achieved by the removal of histones that are first replaced by transition proteins (TPs) which are later exchanged by sperm-specific nucleoproteins, the protamines<sup>203,204</sup>. Protamines are highly basic proteins of a low molecular mass that essentially allow complete neutralization of DNA negative charges. In mouse and human, two classes of protamines, called Prm1 and Prm2, are present in sperm nuclei. Protamines can be post-translationally modified, e.g. phosphorylated, and during final sperm maturation form di-sulfide bridges between neighboring protamines to further increase stability. However, in several mammalian species, including mouse and human, the protamine exchange is not complete resulting in the retention of varying amounts of histones in mature sperm<sup>203,204</sup>.

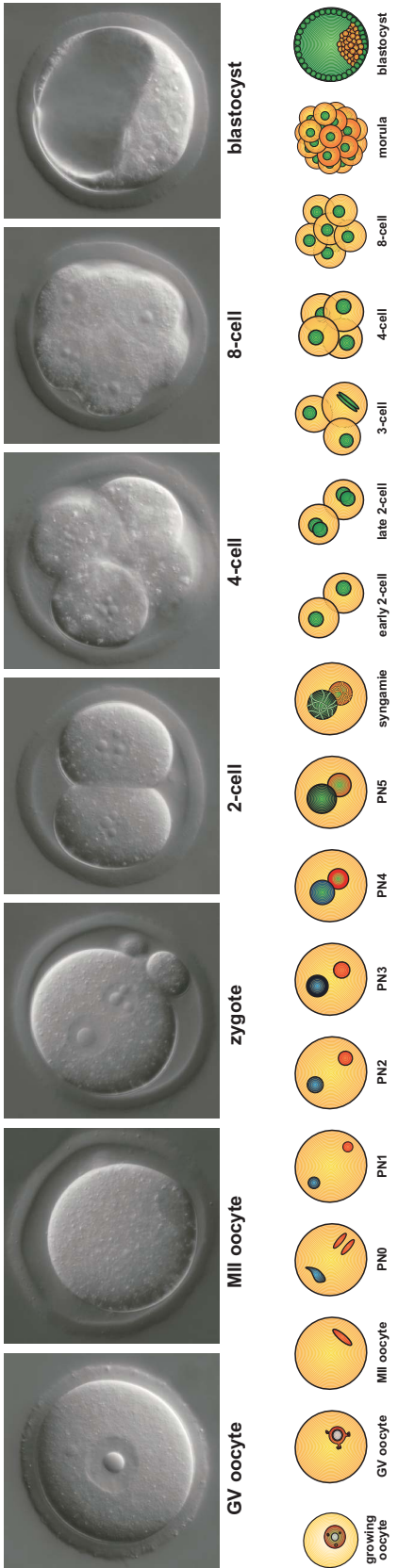
In human, a substantial amount of histones is kept corresponding to approximately 15% of the usual haploid histone complement. Biochemical purifications of human sperm have identified all four core histones<sup>204</sup>. The precise localization of these histones within the genome is not known, but it was noted that sperm telomeres are clustered at the nuclear periphery and retain a nucleosomal configuration<sup>209</sup>. In contrast, mouse sperm is more homogeneous containing up to 99% of protamines.

Recently it was shown that the retained histones are post-translationally modified<sup>210</sup>. In mature sperm derived from the caput epididymis, histone acetylation at H4K8 and H4K12 was detected in chromocenters, suggesting that sperm contributes more than just its genetic information to the embryo. Currently it is not clear which genomic regions in addition to centromeric and telomeric heterochromatin are associated with a nucleosomal configuration in sperm. Furthermore, it will be very interesting to determine the full post-translational modification status of retained histones, especially with respect to more stable modifications like histone methylation. It is intriguing to speculate that the retained histones are not merely leftover artifacts from an inefficient exchange process but might indeed serve a biological function. Modified histones might direct *de novo* histone deposition after fertilization and targeting of HMTs to reestablish the correct chromatin state, thereby guiding embryonic gene expression in the newly formed embryo.

#### *Acquisition of a nucleosomal chromatin configuration*

After fertilization, paternal protamines are re-exchanged to histones that are maternally provided (Fig. 3). Shedding of protamines from DNA is a rapid process that occurs within 30 minutes after sperm entry<sup>202</sup>. Upon start of sperm decondensation the nucleosomal density increases, coinciding with the presence of the histone chaperone Hira in paternal chromatin<sup>202</sup>. Hira is specifically involved in the replication-independent assembly of nucleosomes containing the histone variant H3.3<sup>211</sup>. In contrast, H3.1, which differs from H3.3 by only five amino acids, is deposited by the histone chaperone Caf-1 exclusively during DNA replication. Indeed, H3.3 is incorporated preferentially into the paternal genome following sperm decondensation<sup>212</sup>. In contrast, H3.1 is not detectable in early paternal chromatin and its appearance is delayed until the start of DNA replication in later zygotes<sup>202</sup>.

In *Drosophila* it was shown that the H3.3 histone chaperone Hira is essential for *de novo* assembly of paternal chromatin after fertilization<sup>213,214</sup>. *Hira<sup>ssm</sup>* (sesame) mutant flies which carry a point mutation in the Hira gene<sup>214</sup> as well as *Hira<sup>HR1</sup>* mutant flies generated by homologous recombination<sup>213</sup> show defects in paternal pronucleus formation, resulting in haploid embryos that develop with only maternal chromosomes and die before hatching. Maternally provided FLAG-tagged H3.3 accumulates specifically in paternal chromatin. It strongly labels the paternal genome during first mitosis but fades away after a few nuclear divisions. Interestingly, these mutations of Hira only affect paternal pronucleus formation but not H3.3 deposition in other tissues<sup>213,214</sup>. In addition to Hira, the *Drosophila* molecular motor protein CHD1 is required for deposition of H3.3 into paternal chromatin, but is dispensable for protamine removal<sup>215</sup>. Molecular motors are known to assemble nucleosomes *in vitro* and



Biological Processes



mRNAs and proteins

DNA demethylation



Acquisition of paternal chromatin marks



CHD1 belongs to the CHD (chromo, helicase and DNA-binding domain) subfamily of Snf2-like ATPases. In analogy to the *Hira* phenotype, in maternal *CHD1* mutant embryos, the paternal genome does not participate in mitosis resulting in haploid embryos. Although H3.3 is recruited to the paternal genome it remains at the nuclear periphery, suggesting that Hira is essential for histone delivery whereas CHD1 facilitates histone deposition.

Further research is needed to elucidate how the protamine-to-histone exchange is accomplished in mouse and whether the same players, most of which have homologs in mammals, are involved. Moreover, it will be very exciting to dissect the epigenetic contributions of the paternal complement to early embryonic development.

### 1.3.3. Parental asymmetry in histone modifications

Following fertilization, the highly specialized maternal and paternal genomes need to undergo a series of structural changes and epigenetic programming events to give rise to the pluripotent embryonic cells that ultimately have the power to regenerate a complete organism consisting of hundreds of distinct cell types. Upon sperm entry into the metaphase II oocyte, female meiosis resumes with the extrusion of the second polarbody and the sperm head decondenses followed by a transient recondensation after nucleosome deposition. The paternal pronucleus (PN) forms shortly after the maternal one and then the pronuclei swell and migrate toward the center of the zygote where they become apposed. Based on morphology, zygotic substages have been defined<sup>216,217</sup>. PN0 refers to the zygote immediately after fertilization characterized by maternal chromosome segregation and paternal sperm decondensation, PN1 pronuclei are small and reside at the periphery of the embryo, PN2 pronuclei have an increased size and have started to migrate toward the center of the embryo, PN3 pronuclei have migrated toward the center, large PN4 pronuclei are close to each other and PN5 refers to large central pronuclei.

#### *Paternal histones are de novo methylated in the zygote*

During the first embryonic cell cycle, reprogramming of maternal and paternal genomes takes place in the same environment of the zygotic cytoplasm but the epigenetic outcomes are markedly different. While maternal DNA is embedded in a chromatin structure characterized by the presence of histone methylation marks, *de novo* histone deposition occurs paternally. Immediately upon incorporation, these histones show strong signals for acetylation at various

---

**Figure 4: Epigenetic (re)programming during mouse pre-implantation development.** Fertilization triggers the resumption of the second meiotic division including the extrusion of the second polar body. Following sperm decondensation, paternal and maternal pronuclei (PN) are formed which undergo the first round of DNA replication while remaining spatially separated. First minor zygotic genome activation (ZGA) occurs at the late 1-cell stage, followed by degradation of maternal messages which are replaced by zygotic ones during major ZGA at the 2-cell stage. In pre-implantation embryos, paternal and maternal DNA methylation is erased by active and passive mechanisms, respectively, giving rise to asymmetric labeling of parental genomes. *De novo* methylation at the morula stage results in high levels of DNA methylation specifically in the cells of the inner cell mass (ICM). During sperm decondensation, paternal protamines (Prm) are exchanged to maternally provided histones, involving the specific incorporation of the replication independent histone variant H3.3. Subsequently, various different histone methylation marks are acquired at the paternal genome in a highly spatially and temporally coordinated manner.

lysine residues<sup>181,210,216,218</sup> but lack detectable amounts of known histone methylation marks. In contrast, acetylation of maternal chromatin is delayed until after pronucleus formation resulting in differential acetylation levels between parental genomes which are resolved during the mid zygote stage. At subsequent embryonic stages, histone acetylation remains high during interphase but is absent from cleavage chromosomes<sup>174,181</sup>.

In addition, asymmetry of histone lysine methylation has been observed for most residues studied. Di- and tri-methylation of H3K4, H3K9, H3K27 and H4K20 are initially only present in the maternal pronucleus. The earliest methylation mark that appears on paternal chromatin is H4K20me1 which is detected strikingly fast after gamete fusion, in the periphery of expanding sperm<sup>202</sup>. Following that mono-methylation of H3K4, H3K9 and H3K27 appear after paternal pronucleus formation (PN1/2), resulting in equal levels at both parental genomes. Generally, the acquisition of histone methylation seems to follow a strict spatially and temporally coordinated program with mono-methylation marks appearing first, followed by di- and tri-methylation (Fig. 3). Interestingly, the Polycomb proteins Ezh2 and Eed that confer the di- and tri-methylation of H3K27 are present very early in the paternal pronucleus (at PN1/2) but nonetheless H3K27me3 only appears around PN3/4 concurrent with DNA replication<sup>136,218</sup>. The H3K27 mono-methylase has not been identified in mammals, however, since H3K27me1 is already present at PN1 in paternal chromatin, lack of substrate for PRC2 cannot be the limiting factor. Absence of other essential PRC2 complex members or post-translational modifications on PRC2 itself<sup>219</sup> may explain the catalytic inactivity. Alternatively, H3K27me3 on the paternal genome might be directly linked to DNA replication. It is interesting to note that deletion of the maternally inherited Ezh2 protein in the oocyte results in growth retardation of resulting offspring apparent until weaning, although Ezh2 is expressed from the paternal allele starting from the 4-cell stage, suggesting that Ezh2 has an important function very early in embryonic development<sup>136</sup>.

#### *De novo methylation of H3K9 is delayed*

Whereas all three H3K4 and H3K27 methylation states are acquired on the paternal genome until the end of the first cell cycle<sup>181,218,220</sup>, acquisition of H3K9me2 and H3K9me3 is further delayed until the 4- to 8-cell stage<sup>218,220-223</sup>. H3K9me2 on the maternal genome, which is inherited from the oocyte, declines from the 1- to the 2-cell stage, indicating that no active H3K9me2 HMT is present in the zygotic cytoplasm to maintain H3K9me2 during DNA replication<sup>221,223</sup>. In contrast, nuclear transfer of the unmethylated paternal pronucleus into enucleated GV or MII oocytes allows *de novo* H3K9me2 to occur suggesting that the oocyte cytoplasm but not the embryonic cytoplasm contains H3K9me2 activity. Moreover, inhibition of protein synthesis with cyclohexamid in the zygote results in H3K9me2 at the paternal genome, suggesting that the H3K9me2 HMT is active before fertilization but is deactivated by a newly synthesized protein in the early embryo.

It remains to be resolved whether the stepwise appearance of the various histone lysine methylation marks simply reflects the consecutive mode of action of the mono-, di- and tri-HMTs or whether it is used as a means to distinguish the parental genomes. So far, it is not known which of the mammalian HMTs establish histone methylation marks at the paternal genome

following fertilization, except for Ezh2 which mediates *de novo* H3K27me2/3<sup>136</sup> (our results, Chapter 2.2). Notably, at the zygote stage, the paternal pronucleus is characterized by the presence of hyperacetylated histones associated with an active transcriptional state but lacks repressive methylation marks like H3K9me2/3. One possibility is that the open chromatin configuration of the paternal genome is required to allow efficient reprogramming of the paternal genome into a proper embryonic chromatin configuration.

#### 1.3.4. Asymmetric DNA demethylation

Mammalian development is accompanied by two-major waves of genome-wide DNA demethylation and remethylation: one during germ cell formation and the other during pre-implantation development. DNA demethylation following fertilization occurs asymmetrically. While the paternal genome loses its methylation very rapidly within hours after sperm decondensation through an active process, maternal demethylation is reduced passively through successive DNA replication cycles up to the morula stage (Fig. 3)<sup>217,224</sup>. At the 2-cell stage, half of the nucleus is stained by an antibody against 5-methyl cytosine (5meC)<sup>224</sup>, corresponding to the maternal complement which is still spatially separated from the paternal one at that stage<sup>225</sup>. Methylation is further reduced in 4-cell embryos. A few genes have been studied in detail and those that are highly methylated in sperm rapidly lose their methylation in zygotes<sup>226</sup>, except for paternally imprinted genes<sup>227</sup>. In addition, pericentric heterochromatin<sup>217,218,228</sup> and the intracisternal A particle (IAP) retrotransposon<sup>229</sup> are excluded from active DNA demethylation. Using polyspermic zygotes, it was shown that active paternal demethylation is a very powerful process as up to five paternal pronuclei lose their DNA methylation within a zygote<sup>217</sup>. *De novo* methylation occurs specifically in the inner cell mass (ICM) of the blastocyst whereas much lower levels of methylation are detected in the trophectoderm (TE)<sup>217</sup>.

#### *Maintenance DNA methyl transferases*

Maintenance of genomic DNA methylation patterns in somatic cells depends on the DNA methyltransferase 1 (Dnmt1). In growing oocytes, a specific variant of the maintenance methyltransferase, called Dnmt1o, is expressed from an oocyte specific promoter. Dnmt1o can be detected in the nucleus of growing oocytes, but is excluded from the nucleus upon oocyte maturation and remains cytoplasmic during pre-implantation development<sup>230</sup>. It only specifically translocates into the nucleus for one cell cycle at the 8-cell stage. Homozygous *Dnmt1o* mutant mice are normal, but embryos from *Dnmt1o*<sup>-/-</sup> females die during the last third of gestation. Deletion of *Dnmt1o* does not effect the establishment of genomic methylation patterns and embryos show normal levels of 5meC content as well as normal methylation at IAPs, major satellites and single copy genes. However, around 50% of methylation is lost post-zygotically at some imprinted genes resulting in defective allele specific expression, suggesting that transient nuclear localization of Dnmt1o during the 8-cell stage is required to maintain methylation at specific imprinted genes. In addition, during pre-implantation development, Dnmt1o is important to maintain methylation of IAP sequences inserted in the agouti locus linked to transcriptional silencing of the agouti gene<sup>231</sup>.

In contrast, the longer somatic Dnmt1 isoform, Dnmt1s, is required to maintain IAP methylation during post-implantation stages. This is in agreement with earlier studies that failed to detect Dnmt1s in pre-implantation embryos, whereas two recent reports found Dnmt1s to be expressed early<sup>232,233</sup>. Using newly characterized antibodies, they showed that Dnmt1s localizes to maternal chromosomes in MII oocytes and is recruited to the paternal pronucleus during PN3-4. It remains in pronuclei during first G1/S phase but is excluded during G2. Following that, Dnmt1s can be found in nuclei up to the blastocyst stage. Antibody or siRNA microinjections directed against Dnmt1s during the zygote stage result in reduced methylation at IAPs and the paternally imprinted *H19* gene<sup>233</sup>. These studies suggest that both variants of the maintenance methyltransferase Dnmt1 are required during early embryonic development to maintain specific methylation of imprinted genes. It remains to be resolved, however, how specific sequences are protected from active as well as passive DNA demethylation. Despite many speculations about possible mechanisms to remove DNA methylation<sup>234-236</sup> and targeted deletion of several candidate demethylases (including Mbd2 and Mbd4)<sup>217,234</sup>, the enzyme responsible for paternal specific active demethylation has not been identified.

#### *Importance of asymmetric DNA demethylation*

The lack of a specific enzyme also hampers further studies to address the role of asymmetric DNA demethylation. Like genomic imprinting, active demethylation seems to have evolved in mammals and does not take place in *Xenopus laevis* or zebrafish. Paternal specific loss of DNA methylation has been observed in mouse, rat, pig, bovine and humans but not in rabbit and sheep<sup>217,234,237-240</sup>. In sheep, no differences between maternal and paternal methylation is detected in the zygote and limited demethylation is observed from the 2- to the 8-cell stage. There is no evidence for remethylation of the sheep ICM; instead the TE lineage is further demethylated from the morula to the blastocyst stage<sup>238,239</sup>. However, sheep oocytes are capable to demethylate mouse sperm and sheep sperm becomes demethylated in mouse oocytes<sup>241</sup>. The fact that asymmetric DNA demethylation is conserved in most mammals studied so far suggests that it may be needed for proper embryonic development and might facilitate paternal epigenetic remodeling by erasing previously acquired spermatogenesis specific DNA methylation marks. Alternatively, paternal DNA demethylation may be required to allow for the generalized de-repression of paternal alleles to accommodate the minor transcriptional burst at the zygote stage. Paternal specific DNA demethylation is in line with the general more active paternal chromatin state defined by histone hyperacetylation and lack of repressive histone methylation marks.

Some further mechanistic insights come from studies of the recently identified maternal effect gene *Stella* (also known as *PGC7*)<sup>242</sup>. *Stella* mutant mice are viable, but females show reduced fertility due to a failure during pre-implantation development. In MII oocytes, *Stella* localizes to the cytoplasm but translocates to pronuclei in zygotes which is mediated by the nuclear transport shuttle Ran binding protein 5 (RanBP5)<sup>243</sup>. Although *Stella* is present in both maternal and paternal pronuclei, it is specifically required to protect the maternal genome from active demethylation after localizing to the nucleus. In addition, *Stella* is important to maintain

the methylation of some maternally and paternally imprinted genes. This is the first study to provide some evidence for the requirement of asymmetric DNA demethylation for proper pre-implantation development, but the ultimate proof will have to await the identification of the paternal specific DNA demethylase.

### **1.3.5. Maternal to zygotic transition in gene expression**

The maternal to zygotic transition that occurs following fertilization entails a dramatic reprogramming of gene expression that is required for continued development. In mouse, the first wave of zygotic transcription is observed at the late zygote stage, also known as minor zygotic genome activation (ZGA), followed by the major ZGA in the 2-cell embryo (Fig. 3). Genome activation is accompanied by a global degradation of maternally inherited transcripts at the 2-cell stage which restricts the period of time in which these genes can function and allows the embryo to replace oocyte-specific transcripts with those required for embryonic development<sup>187,244-246</sup>.

#### *First zygotic transcription takes place at the 1-cell stage*

Transcription at the 1-cell stage has been studied by BrUTP incorporation and microinjection of reporter plasmids<sup>187</sup>, revealing that the paternal pronucleus supports approximately five-fold higher levels of transcription<sup>246</sup>. Inhibition of the first round of DNA replication results in reduced BrUTP incorporation at the zygote stage, suggesting that first DNA replication and transcription are functionally linked. More recent microarray profiling studies detected a large number of transcripts that increase during the 1-cell stage<sup>247</sup>. Based on gene ontology (GO) analysis, a significant number of these genes are involved in the regulation of transcription and in chromatin assembly, two processes that play major roles during early pre-implantation development. At the 1-cell stage, RNA polymerases I, II and III are functional<sup>187</sup>. RNA polymerase II (RNAP II) gradually translocates into the nucleus throughout the minor ZGA phase and becomes phosphorylated at its C-terminal domain (CTD) at the onset of major ZGA<sup>248</sup>. Analysis of polyspermic zygotes suggests that the availability of the transcription machinery represents a rate limiting step<sup>246,249</sup>. Surprisingly, although the 1-cell embryo is transcriptionally active, no  $\alpha$ -amanitin sensitive transcripts were detected in the zygote, suggesting that the apparent increase in transcripts is due to the recruitment of maternal messages coupled with polyadenylation of their poly(A) tails<sup>250</sup>. This does, however, not explain what the BrUTP incorporation in the 1-cell embryo reflects. Moreover, from earlier studies using luciferase reporters, it is not clear whether the transcripts made at the 1-cell stage are efficiently translated<sup>187</sup>. But if that is the case, what is the purpose of 1-cell transcription?

Maybe it is simply a consequence of chromatin remodeling allowing easy access of the transcription machinery. If so, a first burst of transcription at the 1-cell stage, where paternal chromatin is reorganized into a chromosomal configuration, should be common to other animal species. Two transcriptional waves are indeed no exception and also occur in *Drosophila melanogaster*, *Caenorhabditis elegans* and *Xenopus laevis*, however, the timing is markedly different<sup>244,249</sup>. Following pronuclear fusion, *Drosophila* embryos go through 13 synchronous nuclear divisions without cellular divisions. The first weak wave of transcription occurs during

nuclear divisions 8 to 10, clearly after initial remodeling of the paternal genome, and for example involves pair rule and sex determination genes. In *C. elegans*, first zygotic transcription is observed at the 4-cell stage, which is however not required for embryonic development, followed by major ZGA at the mid-blastula transition (MBT) at the 100-cell stage. Early *Xenopus* development is characterized by 11 rapid mitotic cycles with a small transcriptional burst at the 6th mitotic division and ZGA at the MBT when the embryo is at the 4000-cell stage. Summing up, in all four organisms described here, two transcriptional waves occur during early embryogenesis, though first transcription only happens immediately following paternal chromatin remodeling in mouse zygotes, indicating that other mechanisms are likely to contribute to initiate first genome activation.

Notably, at the time when first zygotic transcription is initiated in the mouse, both parental genomes are marked by histone hyperacetylation and H3K4me3, RNAP II is present in the pronuclei, and the paternal genome is lacking repressive H3K9me2/3 and DNA methylation. Genome-wide studies in yeast and *Drosophila* have shown that specific histone modifications around transcription start sites of genes, including H3K4me3 and H3K9/14ac, are excellent predictors of active transcription<sup>251,252</sup>. The situation is more complex in human, where the majority of genes are marked by H3K4me3 and H3K9/14ac, although active genes show clearly higher enrichments<sup>126</sup>. In addition, the form of RNAP II that initiates transcription is present at low levels at 5' regions of many inactive genes that do not generate detectable levels of full-length transcripts<sup>53,126</sup>, suggesting that many genes initiate transcription but do not proceed to productive elongation<sup>126,253-255</sup>. Consistently, histone modifications characteristic of ongoing transcriptional elongation, like H3K36me3 and H3K79me2, are found exclusively downstream of promoters that do generate full-length transcripts. It was hypothesized that the unique capacity of human ES cells to initiate various differentiation programs might require inactive genes to be poised for rapid activation<sup>126</sup>, yet similar chromatin states and RNAP II patterns were detected in human hepatocytes and B lymphocytes. First zygotic transcription in the mouse embryo may follow a similar mechanism. Maternal and paternal genomes are being remodeled into a configuration compatible with major ZGA at the 2-cell stage and establishment of a pluripotent epigenetic state. One possible is that RNAP II<sup>248</sup> and other transcription factors and chromatin modifiers<sup>180</sup> are targeted in preparation of the following ZGA, establishing a state poised for transcriptional activation. Interestingly, H3K36me3 that is maternally inherited from the oocyte, declines during the first cell cycle and is almost undetectable at the time when first transcription occurs (our unpublished results). It only reappears at the late 2-cell stage, around the time of ZGA, and remains globally present at both parental genomes throughout further pre-implantation development. Based on these observations, a possible hypothesis is that BrUTP incorporation at the 1-cell stage reflects spurious transcription initiation which would not be detected by the microarrays used to analyze global transcription profiles. It could be analyzed, however, using quantitative RT-PCR along candidate genes in early embryos. Clearly, further studies are also needed to elucidate the localization of the involved histone modifications within the genome and with respect to transcription start sites in early embryos. New chromatin immunoprecipitation (ChIP) protocols, adjusted to permit the analysis of small samples, have

been described<sup>256</sup> and should make higher resolution analysis of histone modifications in pre-implantation embryos possible in the future.

#### *Major zygotic genome activation occurs at the 2-cell stage*

Super-imposed on genome activation at the 2-cell stage is the development of a chromatin-based transcriptionally repressive state<sup>246</sup> that leads to the requirement of enhancer elements for efficient transcription following ZGA<sup>187</sup>. In contrast to differentiated cells, transcription in early blastomeres and ES cells does not require a TATA box for enhancer-driven expression<sup>187</sup>. Of the transcripts detected in 2-cell embryos, 17% are sensitive to  $\alpha$ -amanitin treatment which represents a significant proportion, however, also indicates that genome activation is not as global as previously anticipated but rather specific<sup>250</sup>. Compatible with ZGA taking place, 20% of the genes that are transiently up-regulated at the 2-cell stage are related to transcription<sup>247</sup>.

Analysis of global gene expression changes during pre-implantation development revealed one wave of transcription (ZGA) that contributes mainly to the preparation of the basic cellular machinery and a second one, named mid-pre-implantation genome activation (MGA), which induces dramatic biological and morphological events<sup>257,258</sup>. GO term analysis identifies the switch from pyruvate as primary energy source in early embryos (GO: pyruvate metabolism, tricarboxylic acid cycle) to glucose-dependency in later pre-implantation development (GO: glycolysis, gluconeogenesis). Furthermore, GO terms for cell communication are overrepresented in unfertilized oocytes and 1-cell embryos, likely reflecting the communication between oocytes and follicle cells, and then again from the 8-cell stage when gap junctions are formed between the blastomeres<sup>247,258</sup>.

It has been proposed that the function of the repressive state established in 2-cell embryos is to dictate an appropriate profile of gene expression that is compatible with further development<sup>247</sup>. Yet it remains to be determined, how specific genes are activated in early embryos and what exactly is keeping the remaining ones repressed. A number of maternal effect genes affect the overall transcriptional competence of the 2-cell embryo, including *Npm2*<sup>171</sup> and *Mater*<sup>192</sup> (discussed in Chapter 1.3.1). In addition, maternal Brg1 has been implicated in the regulation of zygotic genome activation<sup>259</sup>. *Brg1* encodes the catalytic subunit of SWI/SNF chromatin remodeling complexes and exhibits DNA dependent ATPase activity. The energy derived from ATP hydrolysis alters the conformation and position of nucleosomes. Maternal deletion of *Brg1* results in embryonic arrest at the 2-cell (78%) and 3/4-cell (10%) stage. In such mutant embryos, the transcription requiring complex (TCR), an accepted marker of ZGA, is markedly reduced, BrUTP incorporation is decreased to 35% and approximately one third of the  $\alpha$ -amanitin sensitive genes are downregulated. Reduced transcriptional competence is accompanied by significantly lower levels of H3K4me2. Treatment of the mutant embryos with TSA induces increased levels of histone acetylation, H3K4me and BrUTP incorporation, however, cannot overcome the 2/4-cell arrest. These results suggest that maternal Brg1 is required for the activation of a significant number of genes, some of which seem to be important for further embryonic development. Another modulator of embryonic transcription is the Transcription intermediary factor 1  $\alpha$  (Tif1 $\alpha$ )<sup>260</sup>. Tif1 $\alpha$  contains a bromodomain and was first

identified as transcriptional regulator of nuclear receptors. At the onset of first genome activation in the zygote, Tif1 $\alpha$  translocates from the cytoplasm into the pronuclei to sites of active transcription, that are also enriched in Brg1 and Snf2h, the ATPase subunit of the mammalian ISWI complex. Ablation of Tif1 $\alpha$  by microinjection of dsRNAs or anti-Tif1 $\alpha$  antibodies results in embryonic arrest at the 2/4-cell stage. Moreover, ablation of Tif1 $\alpha$  induces mislocalization of active RNAP II and the chromatin remodelers Brg1 and Snf2h as well as a small albeit significant increase in BrUTP incorporation at the late 1-cell stage. Using a ChIP cloning approach, the authors show that Tif1 $\alpha$  has no general effect on transcription but affects specific sets of genes.

It is noteworthy that even at the 2-cell stage parental genomes are differentially marked at the global level, with paternal chromatin lacking repressive H3K9me2/3 and DNA methylation, although this might be compensated by the presence of Polycomb repressive complexes (our results, Chapter 2.2). Given the correlation of chromatin configurations and transcriptional states, several questions arise. Does the differential labeling of maternal and paternal genomes result in preferential expression of one parental allele over the other? Allele specific transcription due to distinct chromatin states has been described for imprinted genes<sup>208</sup>. It would be very interesting to analyze levels of maternal versus paternal alleles in early pre-implantation embryos. Along these lines, does deletion of specific chromatin modifiers affect parental genomes and their expression in the same extent? Currently, it is not completely clear whether the high levels of histone acetylation and H3K4me are causative factors of transcription or merely consequential to acquisition of transcriptional competence. Conditional deletion of chromatin modifying enzymes in maternal and paternal germlines in combination with local analysis of histone modifications by ChIP in early embryos should shed further light on these issues. Furthermore, when are maternal specific histone modifications reset to allow establishment of modifications compatible with embryonic development? In contrast to spermatogenic chromatin remodeling, global histone methylation appears stable during oogenesis<sup>182</sup>. Possibly, histone methylation marks are specifically erased at regulatory sequences through the action of recently identified histone demethylases<sup>88</sup>. Alternatively, passive dilution through the first and second round of replication might be sufficient to erase undesired modifications before ZGA. Recent studies in zebrafish have identified the miRNA miR-430 as a potential link between zygotic genome activation and the decay of maternal mRNAs<sup>261</sup>. In mouse, maternally inherited miRNAs are down-regulated between the 1- and 2-cell stage and replaced by *de novo* synthesized miRNAs in 2- to 4-cell embryos<sup>200</sup>. It remains to be revealed whether mouse miRNAs participate in the regulation of the maternal to zygotic transition.

### **1.3.6. Insights from nuclear transfer and cloning**

In the 1990s, the first reports from mammals born after nuclear transfer (NT) cloning appeared<sup>262</sup>. These initial experiments used donor cells derived from pre-implantation embryos. The first successful cloning using the nucleus of a somatic cell from an adult animal to create Dolly the sheep<sup>263</sup> gained much public recognition. Since then it has become possible to clone a number



of species by somatic cell nuclear transfer (SCNT) including pigs, goats, horses, cats and mice. Although many different donor cells have been used for NT, all experiments show that cloning is an inefficient process resulting in embryonic lethality of most cloned embryos and aberrant development of surviving embryos<sup>264</sup>. In many cloned species abnormal placental development has been described, which may lead to malnutrition of seemingly normal cloned embryos<sup>265</sup>. The precise mechanisms leading to these severe defects are unclear, however, accumulating evidence suggests that epigenetic reprogramming is deficient in cloned embryos. Epigenetic reprogramming involves two steps. First, the differentiated donor cell nucleus has to undergo dedifferentiation including the erasure of somatic epigenetic marks to reestablish a totipotent embryonic state. Second, the totipotent cells of the cloned embryo need to redifferentiate to give rise to specific somatic cell types during later development<sup>264</sup>.

#### *Abnormal DNA and histone methylation*

Cloning experiments in bovine reveal an inverse correlation between high levels of H3K9me3 and DNA methylation versus developmental potential of cloned embryos<sup>266</sup>. They demonstrate that H3K9me3 is reprogrammed in parallel with DNA methylation in normal embryos whereas the majority of cloned embryos are characterized by H3K9 hyperacetylation and -methylation associated with DNA hypermethylation. In such cloned embryos, DNA methylation is reduced at the zygote stage but no further demethylation takes place during cleavage divisions, *de novo* methylation occurs precociously and at the morula stage all cells are highly methylated including the TE that is normally characterized by hypomethylation<sup>240</sup>. Along these lines, it was shown that TSA treatment following oocyte activation results in more efficient development of SCNT embryos to the blastocyst stage and allowed the cloning of an outbred mouse strain which was not previously achieved<sup>267,268</sup>, suggesting that induction of histone hypoacetylation may facilitate reprogramming.

Intracytoplasmic sperm injection (ICSI) can be used to create progeny for males who have fertility problems due to immobile mature sperm. In addition, round spermatid injection (ROSI) has also been used for males who do not produce mature sperm, although the rate of successful embryogenesis is significantly lower compared to that for ICSI. The poor development of embryos from ROSI may be associated with epigenetic errors after fertilization<sup>269</sup>. Even though the paternal genome is normally demethylated, it becomes remethylated in the zygote which is dependent on Dnmts and DNA replication. This remethylation is also observed when elongated spermatids are injected but remethylation is less strong, suggesting that during spermiogenesis paternal germ cells gradually acquire the ability to maintain an undermethylated state in the zygote which seems to be required for proper embryonic development. Consistent with reports from SCNT cloning experiments<sup>267,268</sup>, aberrant DNA methylation in the spermatid-derived genome can be reduced by TSA treatment of zygotes<sup>269</sup>.

#### *Nuclear transfer experiments*

For NT experiments in mouse, MII oocytes are used as recipients as the use of activated oocytes even minutes after resumption of meiosis results in chromosome damage leading to 100% abnormal karyotypes and failure to develop to the blastocyst stage<sup>270</sup>. Oocyte sperm

fusion induces a series of  $\text{Ca}^{2+}$  oscillations within one to three minutes that are transduced by Calmodulin kinase II. As a consequence, within minutes after fertilization, the early embryo is either biochemically or functionally different to MII oocytes. The chromatin remodeling capacity of MII oocytes has been studied using demembrated sperm heads that have been heated to  $48^{\circ}\text{C}$ , preventing meiotic exit normally induced by sperm <sup>270</sup>. These studies show that the paternal genome undergoes DNA demethylation and acquires H4K12 acetylation in MII oocytes, delineating non-zygotic chromatin remodeling. One hypothesis to explain why zygotes are not suitable for NT suggests that factors required for reprogramming or embryonic development become sequestered in the pronuclei and are therefore removed by enucleation before NT <sup>271</sup>. This model does not explain the rapid loss of developmental potential following oocyte activation as pronuclei are formed only hours after fertilization. However, the authors show that developmental reprogramming is possible in mouse zygotes that are in mitosis suggesting that the breakdown of the pronuclear envelop at entry into mitosis releases the required factors. Reversible inhibition of the metaphase-to-anaphase transition with the proteasome inhibitor MG-132 allows the mitotic spindle to form which can be observed under the light microscope and therefore easily removed. Chromosome transfer into mitotic mouse zygotes allows the production of cloned mice as well as embryonic stem (ES) cell lines. These results open up new perspectives for the generation of human genetically tailored ES cell lines without the need of fresh oocytes. Instead "left over" polyspermic zygotes from assisted reproduction treatments could be used, which are aneuploid and therefore are excluded from clinical use.

Oocytes and pre-implantation embryos are characterized by the presence of a special linker histone subtype, called H1foo <sup>272</sup>. H1foo expression, which is regulated by DNA methylation <sup>273</sup>, is coupled to the initiation of oocyte growth <sup>274</sup> and is indispensable for meiotic maturation of mouse oocytes <sup>275</sup>. During the late 2- to 4-cell stage, H1foo is replaced by somatic linker histones <sup>272,276</sup>. Linker histone composition does not seem to be affected in SCNT, as H1foo rapidly populates the donor nucleus (within minutes) and subsequently somatic H1 subtypes are lost <sup>276-278</sup>.

To study the extent of nuclear reprogramming at the molecular level, several studies have analyzed gene expression in cloned pre-implantation embryos. Global gene expression patterns of bovine SCNT blastocysts closely resemble those of fertilized control embryos and are very different from somatic donor cells, suggesting that significant reprogramming does occur <sup>279</sup>. Instead, small changes in the expression of a few pluripotency related genes seem to compromise embryonic development in cloned embryos <sup>280,281</sup>. Incomplete re-activation of Oct4-related genes has been observed in mouse SCNT embryos <sup>280</sup> and increased Oct4 expression correlates with higher success in the blastocyst rate <sup>281</sup>. Taken together, studies of cloned embryos from several species indicate that epigenetic reprogramming following fertilization is crucial for proper embryogenesis, and that even tiny changes in the expression of key regulatory genes may have a tremendous impact on the developmental outcome.

### 1.3.7. Trans-generational epigenetic inheritance

Epigenetic information is inherited through mitotic rounds of cell divisions to maintain active and repressive transcriptional states. At certain times in development – during gametogenesis and early embryogenesis – the epigenetic state is reset. This clearing of epigenetic marks between generations is necessary to provide a 'clean state', correlating with totipotency, on which new heritable gene expression programs can be established to initiate various differentiation programs. However, evidence that epigenetic information is not always completely erased and can be inherited from parent to offspring comes from studies in several different eukaryotic organisms<sup>282-285</sup>.

The best studied example is paramutation, which is a widespread epigenetic phenomenon in plants (reviewed in<sup>284</sup>). The term paramutation was first used to describe non-Mendelian inheritance of pigmentation in maize, and subsequently further examples involving phenotypes that were easy to score, like pigment levels, morphological changes or drug resistance, were described in several other plants. Paramutation involves communication between homologous sequences that are present in *trans* to set up distinct epigenetic states that are heritable. Two models have been proposed for *trans* communications: the pairing model where epigenetic states are altered by direct interaction between chromatin complexes, and the *trans*-RNA model which involves the participation of RNA-mediated chromatin changes<sup>284,285</sup>. One well studied example of paramutation in the *b1* locus in maize, where communication between different alleles is mediated by RNA leading to the establishment of distinct chromatin states at downstream tandem repeats – the key sequences required for paramutation<sup>286</sup>.

Remarkably little is known about the contribution of histone modifications and associated proteins towards trans-generational epigenetic inheritance in animals. In flies and mammals, female germ cells maintain a nucleosomal chromatin configuration during oogenesis that is transmitted to the zygote whereas paternal inheritance of nucleosomal chromatin and modifications is limited. Two studies in flies support the concept of maternal epigenetic transmission. Firstly, transmission of a hyperactive SU(VAR)3-9 allele (*pitkin*<sup>Dominant</sup>) through the maternal but not paternal germline results in impaired development of wild-type progeny, likely due to a combined effect of aberrantly established H3K9 methylation patterns and inheritance of mutant maternal protein<sup>44,287</sup>. Secondly, trans-generational epigenetic inheritance through female meiosis of an activated state, mediated by release of Polycomb-dependent silencing, has been reported at the *Fab-7* chromosomal element<sup>288</sup>.

#### *Examples of trans-generational inheritance in mouse*

In mice, maternal and grand-maternal (but not paternal) inheritance of epigenetic information has been observed at the *agouti viable yellow* (*A<sup>vy</sup>*) locus, in which an IAP retrotransposon inserted upstream of the *agouti* gene (*A*)<sup>283,285</sup>. *A* encodes a paracrine signaling molecule that induces hair follicle melanocytes to switch from synthesis of eumelanin (black) to pheomelanin (yellow), resulting in yellow fur color. Mosaic expression of the IAP correlates with its DNA methylation status and is responsible for a range of fur color phenotypes. The maternal effect at this locus is not the result of the maternally contributed environment but is due to incomplete

erasure of epigenetic modifications in the germ line <sup>289</sup>. Recently it was shown, that DNA methylation is not the inherited mark at the  $A^{vy}$  locus as DNA methylation is cleared immediately post-fertilization and is stochastically re-established before mid-gestation <sup>290</sup>. Interestingly, maternal heterozygosity for the PRC1 component *Mel18* (*Rnf110*) introduces epigenetic inheritance at the paternal  $A^{vy}$  allele only in heterozygous offspring, arguing that this is dependent on events occurring after zygotic genome activation and therefore, despite rapid DNA demethylation at the locus <sup>290</sup>.

Moreover, recent screening experiments in mice have identified a limited number of paternal effect genes, in which mutation in the male parental germ line affects wild-type offspring <sup>291</sup>. One example is paternal heterozygosity of the ISWI chromatin remodeler *Snf2h* (*Smarca5*) that enhances variegation at the  $A^{vy}$  allele. *Snf2h* is highly expressed in testis and therefore could induce epigenetic modifications at chromosomes that would enter both wild-type and mutant sperm. Such differences may be retained and could affect maternal chromatin in *trans* following fertilization. Another example of a paternal effect gene is *Dnmt1* which is highly expressed in spermatocytes, providing an opportunity to modify paternally transmitted DNA. The chromosomes from heterozygous *Dnmt1* fathers may enter the zygotes in a hypomethylated state and may act as a sink for chromatin factors, diverting them from sensitive alleles like the  $A^{vy}$  locus. These examples of paternal effect genes show that there is potential for transmission of paternal epigenetic information across generations despite extensive chromatin remodeling during spermatogenesis.

Another recent study supports the idea that sperm can transmit more to the embryo than the information encoded in the primary DNA sequence <sup>282</sup>. A paramutation-like phenomenon was described at the mouse *Kit<sup>tm1Alf</sup>* locus in which a *lacZ-neo* cassette was inserted downstream of the *Kit* promoter <sup>292</sup>. The *Kit* gene encodes a receptor tyrosine kinase, crucial for development. Mice heterozygous for *Kit<sup>tm1Alf</sup>* have distinctive pigmentation patterns: a white tail tip and white feet. This phenotype can be transmitted from heterozygous parents to wild-type offspring through both the male and female germ line for several generations. Paramutated mice have reduced amounts of *Kit* mRNA and accumulate non-polyadenylated RNAs of abnormal size, particularly in testis. Strikingly, these RNAs are also present in *Kit<sup>tm1Alf</sup>* mature sperm and microinjection of total RNA from *Kit<sup>tm1Alf</sup>* heterozygotes or of *Kit*-specific miRNAs into wild-type zygotes induces the heritable white tail phenotype. In contrast to the *b1* locus in maize where gene silencing occurs at the transcriptional level, posttranscriptional silencing may be employed in the *Kit* system <sup>286</sup>. One possibility is that the transferred RNA leads to degradation of wild-type *Kit* transcripts. Alternative, it is also possible that the transmitted RNA establishes a chromatin state that results in reduced transcription. Taken together, transmission of parental RNAs can induce permanent and heritable epigenetic changes affecting gene expression.

#### *Evidence for trans-generational inheritance in human*

In human, there is emerging evidence of germ-line inheritance of epigenetic information at the promoters of two DNA mismatch repair genes, *MLH1* and *MSH2* <sup>293,294</sup>. However, due to great genetic diversity between individuals it is very difficult to rule out alternative explanations and to

argue for a case of trans-generational inheritance in humans. Two individuals with soma-wide, allele-specific and mosaic hypermethylation of *MLH1* were reported<sup>294</sup>. Both had multiple tumors that show mismatch repair deficiency but no mutation in any repair gene was detected. Instead, both carried germ-line epimutations in *MLH1* and one showed the epimutation in a low proportion of sperm cells. Although compelling, definite proof of inheritance of the germ-line mutation is still lacking. Germ-line allele-specific and mosaic somatic hypermethylation has also been described for the *MSH2* gene in a family with inheritance in three successive generations<sup>293</sup>. Three siblings developed cancer, all associated with microsatellite instability and MSH2 protein loss. Again, evidence is missing that the epimutation is indeed a germ-line event and it cannot be ruled out that a mutation in the disease haplotype is responsible<sup>295,296</sup>.

To sum up, evidence for incomplete clearance of epigenetic information in the germ-line resulting in trans-generational epigenetic inheritance has been presented in a number of species. In many of the reported cases, inheritance occurs at transgenes or genes under transcriptional control of retrotransposons or other repetitive elements. These repeat elements often have critical roles in chromosome function and the maintenance of their epigenetic state may be required for proper segregation and pairing of chromosomes during meiosis<sup>285</sup>. Moreover, resistance of IAP elements to DNA demethylation following fertilization<sup>229</sup> may be desirable to prevent IAP retrotransposition but in turn can lead to heritable epimutations of neighboring genes. In the future, it will be challenging to study trans-generational inheritance at endogenous loci, although the phenotypes will be more difficult to score. Such analysis will give us a better understanding of the significance of epigenetic inheritance across generations and may shine further light on the question of the evolutionary purpose of paramutation-like events<sup>282,283,285</sup>.

### **1.3.8. Constitutive heterochromatin reprogramming**

During early pre-implantation development, both maternal and paternal genomes undergo major structural and epigenetic changes which also involve constitutive heterochromatin of pericentric and centromeric regions<sup>297</sup>. Upon pronucleus formation, maternal PCH regions form a discontinuous ring surrounding the nucleolar precursor bodies (NPBs)<sup>298</sup>. PCH can be visualized by fluorochromes like 4,6-diamidino-2-phenylindole (DAPI), preferentially binding to the underlying AT-rich major satellite sequences or by DNA fluorescence *in situ* hybridization (FISH) using probes specific for major satellites. Centromeric minor satellite foci insert into pockets within major satellite rings<sup>298</sup>. In maternal pronuclei, at least one centromere resides at the nuclear periphery and is not associated with the NPBs<sup>298,299</sup>. Paternal PCH initially occupies a single cluster in the centre of the nucleus of decondensing sperm but is subsequently relocated to NPBs upon pronucleus formation<sup>298</sup>. In contrast to maternal PCH, paternal major satellites form a continuous ring with minor satellites positioned on either side and all centromeres associated to NPBs. Notably, in female pronuclei, all NPBs are associated with centromeres whereas this is only the case for 60% of paternal NPBs<sup>299</sup>.

### *Heterochromatin marking following fertilization*

In zygotes, only maternal PCH is labeled by H3K9me3 and H4K20me3<sup>202,218,298-300</sup>, characteristic marks of constitutive heterochromatin that have been inherited from the oocyte<sup>202</sup>. HP1 $\beta$  is loaded onto maternal heterochromatin upon gamete fusion and is subsequently present within euchromatin and is enriched at PCH rings. Both histone methylation marks as well as HP1 $\beta$  are excluded from minor satellite foci<sup>298</sup>. Despite the lack of H3K9me3 binding sites at paternal chromatin, HP1 $\beta$  is also present in paternal pronuclei, albeit less strongly bound<sup>202,218</sup>. Depending on the antibody used, HP1 $\beta$  shows either weak<sup>202,218,298</sup> or no<sup>299</sup> (our data, Chapter 2.2) enrichment at paternal PCH. It is unclear why paternal chromatin is initially refractory towards H3K9me3. One possibility is that the Suv39h enzymes, if present, are enzymatically inhibited as has been reported for a H3K9 di-methyltransferase<sup>223</sup> (Chapter 1.3.3). Alternatively, conditions that enable *de novo* H3K9me3 may first need to be established in analogy to heterochromatin maturation in fly embryos<sup>301</sup> or *S. pombe*<sup>26</sup>. In fly embryos, PCH becomes first visible at the apical pole of early blastoderm nuclei. Establishment of PCH involves removal of H3K4me2 by the histone demethylase SU(VAR)3-3, the *Drosophila* homolog of Lsd1, which facilitates subsequent H3K9me by SU(VAR)3-9<sup>301</sup>. SU(VAR)3-3 co-immunoprecipitates with SU(VAR)3-9, HP1 and the histone deacetylase RPD3 which together form a silencing complex that provides activities for erasing pre-existing chromatin modifications and replacing them with new ones. In line with this it is interesting to note that extensive reprogramming of PCH is occurring during mouse spermatogenesis<sup>302</sup>, resulting in transmission of acetylated core histones as well as specific histone variants at PCH with mature sperm<sup>210,302</sup>. Intriguingly, despite the lack of H3K9me3 at paternal centromeres, chromosomes are properly segregated during first mitosis suggesting that paternal chromatin can compensate for the lack of canonical heterochromatin marks. Future studies should be aimed to understand how parental asymmetry at constitutive heterochromatin could influence chromatid segregation during cleavage divisions, which in humans is remarkable irregular<sup>303</sup>. Maybe the differential epigenetic marking of maternal and paternal PCH is even required to "sort" chromosomes into separate nuclear compartments according to their parental origin<sup>225</sup>, which may facilitate further remodeling of paternal chromatin or selective transcriptional activation.

### *Further reprogramming during pre-implantation stages*

At the 2-cell stage, an increasing number of centromeres relocate from the periphery of the NPBs to the nucleoplasm where they form clusters with centromeres of other chromosomes<sup>222,298,299</sup>, resembling chromocenters of mouse somatic cells<sup>7</sup>. At the 4-cell stage, only 10% of centromeres remain associated with NPBs and relocation is complete at the 8-cell stage<sup>222</sup>. In parallel to chromocenter formation, HP1 $\beta$  which is initially only present at half of the PCH foci at the 2-cell stage, is progressively acquired, marking all chromocenters at the late 8-cell stage<sup>222</sup>. Moreover, fluorescence recovery after photobleaching (FRAP) analysis indicates that HP1 $\beta$  becomes more stably bound to chromatin in later pre-implantation embryos compared to zygotes where it is freely mobile<sup>304</sup>.

DNA replication is temporally regulated in mammals with PCH replicating during mid S-phase in somatic cells <sup>7</sup>. During pre-implantation development, replication timing of PCH switches from late replication at the 1-cell stage to replication during mid S-phase in blastocysts <sup>299</sup>. Heterochromatin organization in embryonic cells depends on the Chromatin assembly factor 1 (Caf-1) <sup>305</sup>. CAF-1 is known as a histone chaperone that promotes deposition of histones, specifically H3.1, onto newly synthesized DNA during S-phase <sup>211</sup>. Deletion of Caf-1 in embryos results in a perturbed 3D organization of PCH and leads to embryonic arrest at the 16-cell stage, that is one to two cell divisions after loss of maternal Caf-1 protein. After knockdown of Caf-1 in ES cells, PCH of individual chromosomes appears isolated or less densely clustered and shows reduced H3K9me3 and H4K20me3 levels. In contrast, such severe alterations are not observed after knockdown of Caf-1 in fibroblast cells, suggesting that Caf-1 is specifically required for the organization of PCH in pluripotent cells.

Notably, the most dramatic rearrangement of PCH from the zygotic ring shape into chromocenters takes place at the 2-cell stage, concurrent with zygotic genome activation. In bovine, where ZGA is initiated in 8-cell embryos, the formation of local HP1 $\beta$  accumulations at centromeres starts at the beginning of the 8-cell stage <sup>299</sup>. This striking coincidence between chromocenter formation and zygotic genome activation suggests that these two processes may be functionally linked.

### **1.3.9. Pluripotency and first lineage commitment**

The zygote and to some extent the blastomeres of early embryos are totipotent, as they can give rise to an entire new organism. After the third division, blastomeres of the 8-cell embryo undergo a morphological change known as compaction. Symmetric and asymmetric cell divisions at the 8-cell stage result in polarized blastomeres. Symmetric divisions give rise to two "outside" daughter cells whereas asymmetric divisions result in one polar "outside" and one apolar "inside" cell, leading to the formation of a 16-cell morula consisting of small inner cells enclosed with larger outer cells. Both compaction and polarization depend upon cell adhesion, involving the intercellular adhesion molecule E-cadherin <sup>306</sup>. The polar cells remain outside and form the trophectoderm (TE) which will give rise to the majority of the embryonic portion of the placenta. In contrast, the inner cells generate the inner cell mass (ICM) that gives rise to the pluripotent cells of the epiblast (EPI) which are covered by a monolayer of the primitive endoderm (PE) on the blastocoelic surface, fated to form extraembryonic tissues. Thus, by the time of implantation the blastocyst has developed three distinct cell lineages which are no longer interconvertible. The TE and PE extraembryonic lineages are required to support the growth of the fetus in the uterine environment and provide a source of signals to the epiblast to direct differentiation and initiation of gastrulation <sup>307-310</sup>.

#### *Differential expression of transcription factors during lineage commitment*

The zygote contains a number of transcription factors, some of which are essential for pluripotency, such as the homeodomain protein Oct4 (Pou5f1) and the SRY-box containing gene 2 (Sox2). Differentiation of the first cell lineages in mammalian embryos is accompanied by the restricted expression of specific sets of key transcription factors in the distinct cell

lineages (reviewed in <sup>311,312</sup>). This lineage restriction is mimicked by the different cell types that can be derived from the mouse blastocyst, providing valuable tools for studying pluripotency and differentiation in the Petri dish. Embryonic stem (ES) cells, derived from the ICM of the blastocyst, are dependent on the expression of the pluripotency factors Oct4, Sox2 and Nanog, and on signaling through the leukemia inhibitory factor (LIF) and bone morphogenetic protein 4 (Bmp4) <sup>311</sup>. In contrast, trophectodermal stem (TS) cells, derived from the TE, require expression of Cdx2, Eomes and other TS specific transcription factors and are dependent on fibroblast growth factor 4 (FGF4) <sup>313</sup>.

The POU transcription factor Oct4 is distinguished by its expression in all pluripotent cell types, including blastomeres of early embryos, ICM, early EPI and ES cells as well as germ cells, and is downregulated upon formation of extraembryonic and somatic lineages. *Oct4*-deficient embryos develop to the blastocyst stage but cells within the ICM lack pluripotency and are instead restricted to differentiate along the TE lineage <sup>314</sup>. Oct4 is, however, not sufficient to maintain pluripotency as forced Oct4 expression fails to render ES cells independent of gp130 signaling through LIF. In contrast, forced expression of Nanog is sufficient to maintain mouse ES cells in the absence of LIF <sup>315</sup>. Upon cell differentiation, Oct4 target genes are predominantly but not exclusively downregulated, consistent with Oct4 acting as a transcriptional activator or repressor, depending on different promoter contexts <sup>311</sup>. Transcription of Oct4 itself needs to be tightly regulated to sustain the ES cell state, as reduction of expression by approximately half induces loss of pluripotency and differentiation to TE whereas a less than two-fold increase of Oct4 causes differentiation into PE and mesoderm <sup>316</sup>. Oct4 expression is controlled through two enhancer regions: the distal enhancer is required for normal Oct4 levels in pre-implantation embryos and ES cells (but is inactive in the epiblast) and the proximal element directs epiblast-specific expression patterns <sup>317</sup>. Expression of Oct4 is positively regulated by the pluripotency transcription factors Nanog and Sox while in turn Oct4 is present at the promoter of these genes, resulting in a transcriptional network that establishes and maintains pluripotency <sup>311</sup>. Downregulation of *Oct4* during differentiation is accompanied by a loss of active chromatin marks, including deacetylation of H3K9/K14 and demethylation of H3K4 as well as an increase in G9a-mediated H3K9me and HP1 binding followed by *de novo* methylation of DNA <sup>318</sup>. While H3K9me is slowing down Oct4 reactivation, DNA methylation of the Oct4 promoter seems to be the locking mechanism that serves to stably inhibit Oct4 re-expression.

During pre-implantation development, the first differences in the distribution of key pluripotency factors have been described at the early morula stage. In 8-cell embryos, both Oct4 and Cdx2 are expressed in all nuclei of the embryo. Cdx2 expression starts to decrease in some inner cells of the morula while Oct4 is still detectable in all cells, suggesting that loss of Cdx2 in the inner cells of morula embryos might be the primary event in the segregation between ICM and TE <sup>319</sup>. How differential expression of Cdx2 is achieved in pre-implantation embryos is still unclear. ES cells can be induced to differentiate into TE by forced repression of Oct4, independent of Cdx2, but Cdx2 is essential for the self-renewal capacity of TS cells <sup>319</sup>.

Segregation of the ICM into PE and EPI is preceded by the mutual exclusive expression of Nanog and Gata6 in ICM cells in a random "salt and pepper" pattern at E3.5 <sup>320</sup>. Lineage tracing



experiments have shown that single ICM cells are restricted to either the PE or EPI lineage at E3.5 – one day before morphological differences become apparent between these tissues. In embryos deficient for the adaptor molecule Grb2, which has important roles in signal transduction downstream of several different receptor tyrosine kinases, Gata6 expression is lost and all ICM cells are positive for Nanog. In such mutant embryos, no PE is formed, suggesting that the ICM develops as a mosaic of PE and EPI progenitors at E3.5, dependent on Grb2 signaling, followed by later segregation of the progenitors into the appropriate cell layers.

#### *Differences in epigenetic modifications*

Epigenetic mechanisms are thought to be involved in maintaining pluripotency as well as directing cells towards distinct developmental programs<sup>312</sup>. Global differences for chromatin modifications become first apparent at the morula to blastocyst transition in early mouse embryos. Following global DNA demethylation during pre-implantation development, *de novo* DNA methylation takes place specifically in the cells of the ICM whereas TE cells keep their hypomethylated state, resulting in epigenetic asymmetry between the first two cell lineages<sup>217,240</sup>. Asymmetry is also observed for Polycomb-mediated H3K27me3 with strongly increased levels in the ICM compared to TE<sup>136</sup>. What is the role of this epigenetic asymmetry? Is it involved in directing the expression patterns of key transcriptional regulators? Very little is known towards this end, mainly due to technical difficulties in using high resolution ChIP analysis on small samples like early embryos. One pilot study has however succeeded in applying ChIP to as little as 100 cells, permitting the analysis of dissected ICM and TE cells of blastocyst embryos using *Drosophila* chromatin as carrier<sup>256</sup>. They show that the pluripotency factors Oct4 and Nanog are enriched in active chromatin marks in the ICM but lack these modifications in TE cells whereas Cdx2 shows inverse enrichments. More studies are required to specifically follow chromatin modifications during early development to monitor their changes at the promoters of key developmental regulators during specification of the first cell lineages. Such analysis will allow us to draw conclusions with respect to the order of events and will be especially informative in embryos deficient for certain chromatin modifiers, a number of which became available during recent years.

The long-standing view that blastomeres of the early mouse embryo are identical at least until the 8-cell stage, has been challenged by a number of recent studies. In several animal species, the polarity of the embryo is established from the very beginning, as a result of maternally inherited factors that are either asymmetrically localized in the egg or asymmetrically distributed following fertilization. Mammalian embryos have often been held as an exception from this rule and it is still controversial, when and how polarity is acquired (reviewed in<sup>321,322</sup>). Some studies suggest that formation of the embryonic-abembryonic axis happens at random, thus independent of earlier developmental events<sup>323-325</sup>. Another viewpoint is that the first cleavage in the mouse embryo could predict blastocyst polarity. Consequently, although each 2-cell blastomere can give rise to both ICM and TE, one of the 2-cell blastomeres tends to contribute more to embryonic parts of the blastocyst and the other blastomere more to extraembryonic parts<sup>326-328</sup>. In light of these findings it is interesting to note that while all 4-cell

blastomeres can have full developmental potential, they differ in their individual developmental properties according to their spatial origin <sup>329</sup>. Especially blastomeres from embryos that result from one meridional (M) and later equatorial (E) division ("ME" embryos) differ in their fate and their developmental properties. Chimeras made entirely from the later E blastomeres are developmentally compromised. Such E blastomeres can however contribute to all embryonic cell lineages when they are surrounded by blastomeres from random positions <sup>329</sup>. It has been suggested many times in the literature that epigenetic mechanisms may guide cell fate decisions during embryogenesis and a recent report by Torres-Padilla and colleagues <sup>330</sup> provides the first indication that this hypothesis might actually be true. They show that levels of arginine methylation at H3R2, H3R17 and especially H3R26, mediated by CARM1, differ in individual blastomeres of EM and ME embryos, i.e. in the embryos that have been previously shown to have a developmental bias. Blastomeres with high levels of H3R26me preferentially contribute to ICM and polar TE whereas progeny of blastomeres with low H3R26me contribute more to mural TE. Most strikingly, overexpression of CARM1 in individual blastomeres directs the fate of their progeny towards the ICM lineage, resulting in upregulation of the pluripotency factors Nanog and Sox2.

Taken together, histone arginine methylation has been identified as the earliest epigenetic mark that contributes to the segregation of the first cell lineages in the early embryo. It remains to be studied whether other epigenetic modifications play similar roles in guiding cell fate decisions in mouse embryos and other species.

#### **1.4. Scope of the thesis**

When I started my PhD, it was known that different types of histone methylation are initially asymmetrically distributed between maternal and paternal genomes. Except for Ezh2 which is required for H3K27me3, the histone methyltransferases promoting the *de novo* methylation reactions had not been identified. By RT-PCR expression profiling, we found that different HMTs specific for H3K4me that belong to the Trithorax group of proteins are expressed in oocytes and early embryos. Moreover, in addition to PRC2, several members of the PRC1 complex were found to be expressed maternally and zygotically. Since the TrxG and PcG proteins are known to be part of a cellular memory system, we wanted to analyze their role in transmitting epigenetic information through the maternal germ line and in guiding major events occurring in the early embryo.

Initially, it was thought that major zygotic genome activation at the 2-cell stage is global, but recent microarray profiling studies suggest that specific sets of genes are activated. The contribution of histone modifications to this process is largely unknown. Like observed later during development, it is possible that TrxG proteins are involved in promoting active gene transcription whereas PcG proteins might repress other genes during zygotic genome activation. Moreover, following induction of lineage commitment, the cells of the TE and ICM need to 'remember' their cellular fate, possibly involving the antagonistic action of TrxG and PcG proteins. The asymmetric distribution of H3K27me3 between TE and ICM <sup>136</sup> (our data) might point towards such a role.

Taken together, the players involved in setting up chromatin states in early embryos and their contribution to pre-implantation development was not well understood. In this PhD project, I initially aimed to analyze the antagonistic function of TrxG and PcG proteins. However, due to the high redundancy of H3K4 specific HMTs we decided to focus on PcG proteins. We aimed to dissect the contributions of PRC2 and PRC1 complexes to pre-implantation development using conditional knock-out approaches.

## **2. Results**

### **2.1. Abundant transcripts from retrotransposons are unstable in fully grown mouse oocytes**

Mareike Puschendorf, Paula Stein, Edward J. Oakeley, Richard M. Schultz, Antoine H.F.M. Peters and Petr Svoboda

## Abundant transcripts from retrotransposons are unstable in fully grown mouse oocytes

Mareike Puschendorf<sup>a</sup>, Paula Stein<sup>b</sup>, Edward J. Oakeley<sup>a</sup>, Richard M. Schultz<sup>b,\*</sup>, Antoine H.F.M. Peters<sup>a,\*</sup>, Petr Svoboda<sup>a</sup>

<sup>a</sup> *Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland*

<sup>b</sup> *Department of Biology, School of Arts and Sciences, University of Pennsylvania, PA, USA*

Received 17 May 2006

Available online 27 June 2006

### Abstract

One physiological function proposed for RNA interference (RNAi) is to constrain expression of repetitive elements and thereby reduce the incidence of retrotransposition. Consistent with this model is that inhibiting the RNAi pathway results in an increase in expression of repetitive elements in preimplantation mouse embryos. Mouse oocytes are essentially transcriptionally quiescent providing a unique opportunity to assess the stability of repetitive element-derived transcripts in these cells. We compared the transcriptome of freshly isolated fully grown germinal vesicle (GV)-intact oocytes to that of oocytes in which meiotic maturation *in vitro* was inhibited for 48 h by milrinone. Consistent with the aforementioned function for RNAi is that the abundance of only a relatively small number of transcripts decreased in the cultured oocytes, when compared to changes that occur during maturation or following fertilization, and of those, several belonged to mobile elements.

© 2006 Elsevier Inc. All rights reserved.

**Keywords:** GV-oocyte; Meiosis; mRNA stability; mRNA degradation; Milrinone; L1; IAP; VL30; ETn

Approximately 40% of the human and mouse genomes are composed of a diverse group of transposable elements (TEs) [1–3], which differ in many aspects such as structure, copy-number, expression pattern, and rate and mechanism of mobility. Most of TEs are retroelements that transpose through a “copy and paste” mechanism utilizing an RNA intermediate. Although transcribed, these sequences are typically not mobile because the vast majority bear mutations, truncations, and deletions. For example, of the ~400,000 human L1 insertions (which occupy 17% of the genome [2]), only 90 are intact and full-length, of which 40 are active in a cell-culture retrotransposition assay [4]. TEs have a capacity to cause deleterious mutations and they are often viewed as harmful parasites [5]. Consistent with this idea is that numerous mechanisms operate in ani-

mals to silence TEs that include transcriptional silencing mediated by DNA methylation [6–8] and chromatin changes [9,10].

In invertebrates, repression of mobile elements occurs post-transcriptionally by an RNA interference (RNAi)-like mechanism. RNAi refers to the selective degradation of mRNA induced by double-stranded RNA (dsRNA), and is one of the mechanistically related RNA silencing pathways (reviewed in [11]). It is viewed as a form of defense against viruses and other parasitic sequences and one of the proposed roles for RNAi in the germ-line in metazoa is to inhibit TEs [12–14]. Whether this role extends to mammals is not clear.

RNAi operates in mammalian cells and appears to be a major pathway responding to long dsRNA in germ cell lineage-competent cells such as oocytes, early embryos, and undifferentiated embryonic stem (ES) cells [15–17]; these cells lack an interferon response [15,18]. Because mammalian TEs apparently can generate dsRNA [19–21]

\* Corresponding authors.

E-mail addresses: [rschultz@sas.upenn.edu](mailto:rschultz@sas.upenn.edu) (R.M. Schultz), [antoine.peters@fmi.ch](mailto:antoine.peters@fmi.ch) (A.H.F.M. Peters).

and expression of several retrotransposons occurs in mammalian oocytes and early embryos [21–23], a likely function of RNAi in oocytes and early embryos would be to constrain expression of TEs, thereby limiting their activity in germ-line cells. Indeed, inhibiting Dicer in preimplantation mouse embryos or ES-cells results in an increased abundance of mRNAs of retrotransposons L1, IAP, and MuERV-L [20,24].

To address post-transcriptional silencing of TEs in mammals, we analyzed the stability of TE-derived RNAs in mouse oocytes using microarrays and RT-PCR. Oocyte growth is accompanied by the cessation of transcription in the fully grown oocyte [25]; transcription resumes by the late 1-cell stage/2-cell stage [26]. *In vitro*, this window of transcriptional quiescence can efficiently be extended by treatment with the phosphodiesterase (PDE) inhibitor milrinone, which blocks the resumption of meiosis. Thus, milrinone treatment provides an opportunity to assess mRNA stability without interfering with developmental processes such as resumption of meiosis. The underlying assumption is that, under these conditions, TEs would exhibit greater instability due to RNAi-mediated degradation. Results of the experiments described here indicate that the abundance of transcripts of several (but not all) mobile elements is markedly reduced in oocytes inhibited from undergoing maturation *in vitro*. These elements include mVL30 and IAP, retrotransposons exhibiting the highest transcript abundance in the oocyte.

## Materials and methods

**Oocyte and embryo collection and culture.** Oocytes for microarray and RT-PCR analyses were isolated from superovulated 6- to 8-week-old CF1 or CD1 females, respectively. Fully grown germinal vesicle (GV)-intact oocytes were collected 46 h after eCG injection (5 IU) from cumulus cell-oocyte complexes. Metaphase II-arrested eggs were collected from eCG- and hCG-primed mice. One-cell embryos were harvested from eCG- and hCG-primed females that were mated.

Oocytes were pooled and one-half was immediately frozen in lysis buffer for RNA isolation at a later time. The other portion was cultured in CZB containing 1 mM glutamine and 2.5  $\mu$ M milrinone for 48 h at

37 °C in a humidified atmosphere of 5% CO<sub>2</sub> in air. MII eggs and 1-cell embryos were collected in FHM/Hepes 20 and 21 h post-eCG injection, respectively. For all stages, oocytes and embryos were pooled from several mice and RNA was isolated from batches of 50 oocytes/embryos.

**RNA isolation and RT reaction.** RNA was isolated using the “Absolutely RNA Nanoprep Kit” (Stratagene). Briefly, cells were transferred to lysis buffer and stored at –80 °C. After thawing, 100 ng of *Escherichia coli* rRNA was added to each sample as carrier. RNA was purified according to the manufacturer’s instructions. RNA was eluted using two separate elution steps resulting in a total volume of 20  $\mu$ l (i.e., 2.5 oocyte-equiv/ $\mu$ l). Reverse transcription was performed from total RNA corresponding to 20 oocytes or embryos using random primers (200 ng) and SuperScript II RNase H Reverse Transcriptase (Invitrogen) according to the manufacturer’s protocol.

**PCR.** For PCRs, cDNA corresponding to 0.2 embryos or oocytes was used as a template. Primer sequences are listed in Table 1. Amplifications were carried out using *Taq* DNA polymerase (Qiagen). Thermocycling was performed in a Bio-Rad iCycler using the following PCR conditions: 1 cycle at 95 °C for 5 min; the indicated number of cycles at 95 °C for 45 s, 55 °C for 45 s, and 72 °C for 1 min; 72 °C for 10 min and a final hold at 4 °C. For MT 22 cycles were used, for ActB, Mos, Oct4, Plat, and Zp3 28 cycles, for Orr1 30 cycles, for IAP, L1, mVL30, RLTR1B, SINE B1, SINEB2 32 cycles, for Ezh2 35 cycles and for Etn 38 cycles, respectively. PCR products were resolved on a 2% agarose gel and afterwards the gels were stained with SYBR green I (Molecular Probes, 1:10,000). Fluorescence was detected on a Typhoon 9400 scanner (Amersham Biosciences).

**Affymetrix microarray hybridization and analysis.** Total RNA from four replicates of each treatment was used for linear, two-round amplification by *in vitro* transcription and target cRNA preparation according to the Affymetrix Small Sample Prep Technical Bulletin ([www.affymetrix.com](http://www.affymetrix.com)). Total RNA from each replicate was reverse-transcribed using the Affymetrix cDNA synthesis kit and cRNA was produced by *in vitro* transcription (IVT) by T7 RNA polymerase using the Affymetrix IVT kit as per the manufacturer’s instructions. Twenty micrograms of biotinylated cRNA was fragmented by heating with magnesium (as per Affymetrix’s instructions) and 15  $\mu$ g of fragmented cRNA was serially hybridized to MOE430 2.0 GeneChips and then processed according to the manufacturer’s instructions (GeneChip Analysis Technical Manual, [www.affymetrix.com](http://www.affymetrix.com)) at the FMI Microarray Facility. Microarray analysis was performed as described in [27] and the resulting data from this study are available at the Gene Expression Omnibus ([www.ncbi.nlm.nih.gov/geo](http://www.ncbi.nlm.nih.gov/geo)). Briefly, GC-RMA algorithm from Bioconductor was used to estimate probe set expression values. The expression values were then exported as a text file and imported into GeneSpring 7 (Silicon Genetics) with default per-chip normalization to the 50th percentile and per-gene normalization to the median. The per-chip and per-gene normalized data are referred to as

Table 1  
Primers used for RT-PCR

Name	Forward	Reverse	References
Actb	5'-TGGGAATGGGTCAGAAGGACT-3'	5'-GGGTCATCTTTTCACGGTTGGC-3'	
Ezh2	5'-AGCCTTGTGACAGTTCGTGC-3'	5'-TTTAGAGCCCCGCTGAATG-3'	
Pou5f1	5'-GGAGAAGTGGGTGGAGGAAG-3'	5'-GGGAAACCCTGTAGCCTCATAAC-3'	
Zp3	5'-AAGCTCAACAAAGCCTGTTTCG-3'	5'-TATTGCGGAAGGGATAC AAGG-3'	
Mos	5'-CCATCAAGCAAGTAAACA-3'	5'-AGGGTGATTCCAAAAGAGTA-3'	[16]
Plat	5'-CATGGGCAAGAGTTACACAG-3'	5'-CAGAGAAGAATGGAGACGAT-3'	[16]
IAP	5'-GCACCCTCAAAGCCTATCTTAT-3'	5'-TCCCTTGGTCAGTCTGGATT-3'	[20]
mVL30	5'-CCTTTGTTGCCAGGTAAGTC-3'	5'-CACTGTAGCCAGTTGTGACCAG-3'	
L1	5'-TTTGGGACACAATGAAAGCA-3'	5'-CTGCCGTCTACTCCTTTGG-3'	[9]
RLTR1b	5'-TCCTTCCCTTTGCCCTATT-3'	5'-GGCTGGAAGTGGTGAGATGT-3'	[21]
Etn	5'-CAGGCTTTGGAGACAATAGGG-3'	5'-TCTCTCAGGGAAGTCTCAGAAACG-3'	
ORR1	5'-CTTAGTTGATGGCCAGGA-3'	5'-CCAACTGCCCCTGTAGC-3'	[21]
MT	5'-ATGTCTGGGGAGGACTGTG-3'	5'-AGCCCCAGCTAACAGAACT-3'	[21]
SINE B1	5'-GTGGCGCACGCCTTAATC-3'	5'-GACAGGGTTTCTCTGTGTAG-3'	[9]
SINE B2	5'-GAGATGGCTCAGTGGTAAAG-3'	5'-CTGTCTCAGACACTCCAG-3'	[9]

“raw” and “normalized” expression values, respectively. Because we expected down-regulation of an unknown number of genes during the milrinone treatment, we generated a list of 23,989 probe set for further analysis from probe sets present in all four control replicates (GV-oocyte,  $t = 0$ ; raw signal value >50).

**Results and discussion**

*General characterization of the microarray analysis*

We first compared fully grown GV-intact oocytes freshly isolated from the ovary, with oocytes cultured in the presence of 2.5  $\mu$ M milrinone for 48 h (Fig. 1); additional information can be found in the Supplemental Material. The profile of raw signal from MOE430 2.0 arrays was uniform

(Fig. 1A). The average number of probe sets with present signal (raw signal value >50) per chip was 23,577 ( $\pm$ 314). It should be noted here that the number of probe sets on the microarray is higher than the number of expressed genes because the array is partially redundant.

We performed several clustering analyses of the microarray data using various gene lists. In each case replicates of each experimental condition clustered together (Fig. 1B). For most genes, raw signal and normalized values did not vary much between replicates. Of note is the presence of many genes whose relative abundance was decreased in oocytes following culture when compared to their freshly isolated counterparts, as evidenced by the appearance of blue bands.

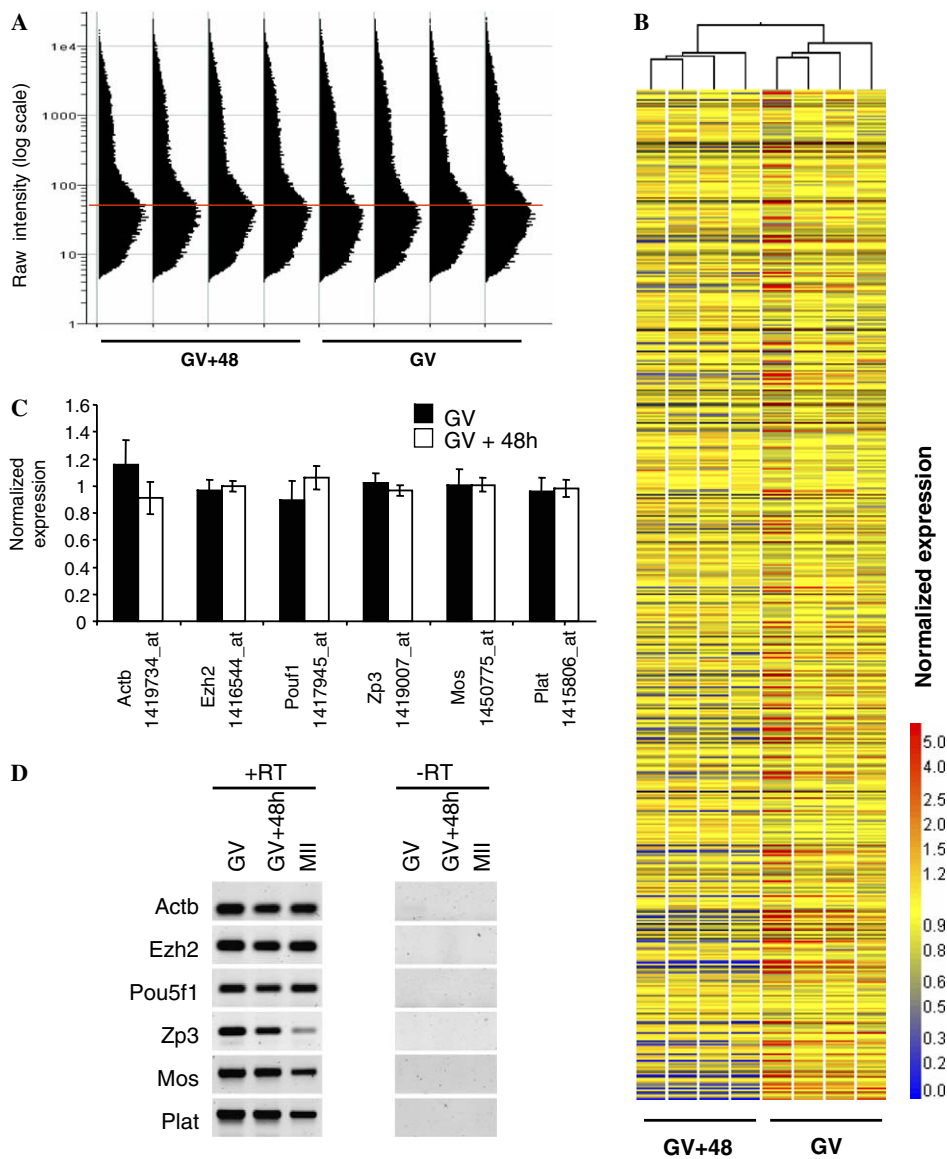


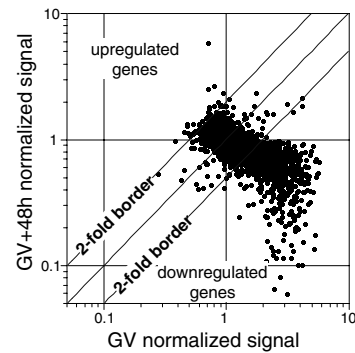
Fig. 1. Evaluation of microarray data. (A) Hybridization signal profile on all microarrays. X-axis shows relative abundance of probe sets with raw intensities depicted on the Y-axis. (B) Conditional clustering analysis. Displayed are all differentially expressed probe sets (ANOVA,  $p < 0.05$ ) with a raw signal >50 in the GV samples. (C) Mean normalized value for a set of six transcripts, which did not show significant down-regulation during milrinone treatment. Error bar = standard deviation. (D) Analysis of the same genes by RT-PCR. MII, metaphase II-arrested egg. +RT, samples were reverse-transcribed; –RT, samples were not reverse-transcribed.

To address further the reliability of the microarray data, we analyzed the relative abundance of six transcripts by RT-PCR, which showed little, if any, changes in their mean normalized expression values (Fig. 1C and D). These transcripts included: the dormant maternal mRNAs *Mos* and *Plat*, which are stored during oocyte growth and not translated until the resumption of meiosis and which are stable following extended culture of fully grown GV-intact oocytes [16]; the oocyte-specific gene *Zp3*; the ubiquitously expressed housekeeping gene *Actb*; and genes with regulatory functions *Pou5f1* (transcription factor Oct-4) and a polycomb gene *Ezh2*. This analysis yielded a good correlation between the microarray data and RT-PCR data, providing further support for the robustness of the microarray results (Fig. 1C and D).

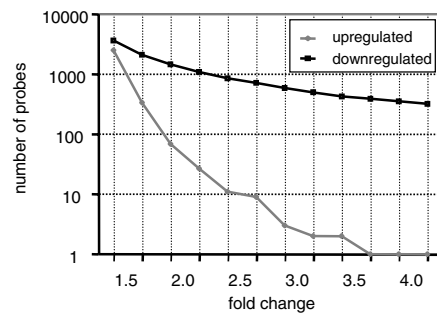
In the absence of transcription, the majority of probe sets that display a change should show a decrease in relative abundance, and such was the case (Fig. 2A and B). A scatter plot of 23,989 probe sets (raw value >50) where per-gene normalized values of milrinone-treated oocytes are plotted against control oocytes and using a 2-fold cut-off revealed a population of transcripts whose relative abundance decreased. There were 27 probe sets exhibiting up-regulation above the 2-fold cut-off in the milrinone treatment but the majority of probe sets showing more than 2-fold-change were down-regulated (1091). The 2-fold cut-off is a default for many microarray experiments. We have addressed the problem of fold-change cut-off in more detail because linear amplification likely introduced more variability into the analysis resulting in a higher rate of false positive changes. Fig. 2B shows relationship between fold-change cut-off and numbers of up-regulated and down-regulated transcripts. These results indicate that even at the 2-fold cut-off, the number of up-regulated transcripts is fairly small (27) and there are only three transcripts found up-regulated at the 3-fold cut-off. Only six of the 2-fold up-regulated probe sets (none of 3-fold) pass the *t*-test ( $p < 0.05$ ).

As mentioned above, 1091 probe sets were found to be down-regulated more than 2-fold, which represents 4.35% of all transcripts; Fig. 2C shows the distribution of the fold down-regulation across the 1091 probe sets down-regulated more than 2-fold. Most of the probe sets detected a smaller down-regulation but there were also transcripts showing a strong reduction in abundance (up to ~52-fold). The initial fold-change filter reduces the number of genes considered for statistical validation, which is an important step for minimizing the false discovery rate. For the statistical analysis we applied a one-way ANOVA ( $p < 0.05$ ) reducing the 2-fold list from 1091 to 424 probe sets, while most of the removed probe sets showed lower fold-changes (see Supplemental Material for gene lists). Although the number of actual transcripts is somewhat smaller because the probe sets are redundant, we estimate that the total number of affected transcripts does not exceed a few percent of all expressed genes. In particular, we rarely observed down-regulation of abundant transcripts, i.e., transcripts whose

### A Global comparison of expressed genes (raw >50)



### B Number of MOE 430 probes showing a given fold change



### C Fold change distribution of 1091 probes downregulated >2-fold

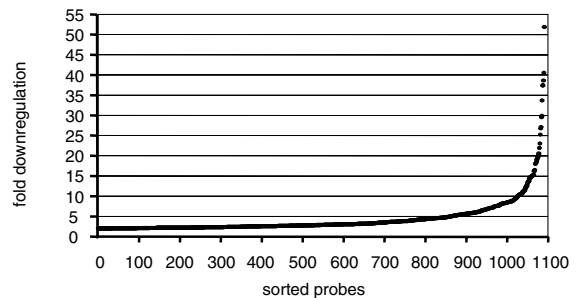


Fig. 2. Global view of differentially regulated transcripts. (A) Scatter plot showing comparison of per-gene normalized expression levels of 23989 transcripts detected present in the fully grown oocyte (GV). Each gene is represented by a dot. (B) Number of up-regulated and down-regulated transcripts as a function of fold-change cut-off. Y-axis (logarithmic scale) shows number of gene probes, which pass a fold-change cut-off depicted on the X-axis. (C) Distribution of fold-change for 1091 probe sets down-regulated more than 2-fold. Probe sets (X-axis) were sorted according to their increasing fold-down-regulation, which is displayed on the Y-axis.

raw expression value was higher than several thousands. Therefore, maternal mRNAs, which accumulate during oocyte growth, are highly stable in the fully grown GV-oocyte. This result is consistent with results of previous experiments which indicated that oocyte mRNA is stable during the growth phase [28,29].

The cause of the observed instability of down-regulated mRNAs is unknown. Down-regulated transcripts are not strongly enriched for AU-rich or other known mRNA-destabilizing sequences (data not shown). It should be noted that known mechanisms degrading maternal mRNAs in the oocytes are associated with resumption of meiosis or other developmental transitions [30], which do not occur



in our model system. It is possible that a fraction of the down-regulated mRNAs represents microRNA (miRNA) targets. It has been recently shown that post-transcriptional regulation by miRNAs may involve mRNA degradation even in the absence of extensive base pairing to their targets [31,32], possibly as a consequence of relocalization of repressed mRNAs to cytoplasmic domains known as P-bodies (reviewed in [33,34]). There are hundreds of known mammalian miRNAs [35]. Each mammalian cell studied so far expressed numerous miRNAs and oocytes are likely not an exception. Studies of miRNA targets estimate that up to 30% of mammalian genes could be targeted by miRNAs [36]. However, it is not known how many genes are regulated by miRNAs in specific cell types. Given the limited information about what makes a functional miRNA target site, prediction algorithms generally estimate that hundreds of mRNAs are targeted by individual miRNAs [37–41]. A recent study of *Drosophila* cells depleted of the miRNA pathway components indicates that miRNAs may down-regulate up to several percent of cellular transcripts [42]. If a similar fraction of transcripts would be targeted in the mouse oocytes, a majority of down-regulated mRNAs could be miRNA targets. However, this issue cannot be directly addressed until technical problems with cloning miRNAs from very small samples will be solved and miRNA pathway-deficient oocytes will be available.

How the down-regulation of hundreds of transcripts would affect developmental competence of mouse oocytes is not known. It is interesting to note, however, that culturing porcine oocytes for 22 h in the presence of IBMX, another phosphodiesterase inhibitor, which has a comparable effect to milrinone in preventing resumption of meiosis, does not negatively affect developmental competence as measured by culturing *in vitro* matured and fertilized eggs to the blastocyst stage [43]. Whether the developmental competence to term is affected is not known.

#### Stability of mRNA of repetitive elements

As mentioned above, several mobile elements in invertebrates were described as endogenous targets for RNAi [12,13,22,44] but there is presently little evidence supporting this role for RNAi in mammals [20,24]. If RNAi-mediated degradation of repetitive elements exists in the mouse oocyte, the targeted transcripts should be found among the mRNAs down-regulated after milrinone treatment.

To this end we conducted a systematic survey of the stability of transcripts derived from repetitive elements. We generated a list of murine repetitive elements highly similar to or perfectly matching MOE 430 2.0 probe sets (see Supplemental Material). We then extracted and analyzed raw and normalized data for these repetitive elements in GeneSpring 7. Although the majority of the Affymetrix probes for various repetitive elements appeared stable, several probes (perfectly matching intracisternal A particle (IAP) and mVL30 retrotransposons) showed a robust down-regulation (Fig. 3A). Instability of IAP and mVL30 transcripts

was further confirmed by RT-PCR with primers amplifying a specific region upstream of the 3' long terminal repeats (Fig. 3B). RT-PCR analysis of other repetitive elements revealed down-regulation of L1 transcripts (which show a low hybridization signal and a weak down-regulation on microarrays (Fig. 3A)) and of the RLTR1b element (absent on the microarrays).

No down-regulation was observed for ORR1, MT, and SINE B1 and B2 elements (Fig. 3B). It should be noted that a recent study reported that many gene transcripts expressed in oocytes are chimeric, containing LTR class III retrotransposon sequences fused to their 5' ends [21]. The function of these transcripts in preimplantation development remains unresolved. The MOE430 2.0 chip does not contain a perfectly matching probe set for MT-like retrotransposons and our data do not discriminate between these chimeric transcripts and their *bona fide* counterparts because the array probes used are biased towards the 3' end of the transcript. In any case, RT-PCR with MT primers showed no change in their mRNA levels after milrinone treatment (Fig. 3B).

Interestingly, RT-PCR analysis of repetitive transcripts down-regulated during milrinone treatment revealed variable behavior during normal development (Fig. 3B). The RT-PCR analyses of unfertilized and fertilized eggs (Fig. 3B) are consistent with previously reported expression data [21,23,27]. Fig. 3B illustrates differences between regulation of individual mobile elements. For example, IAP shows a strong down-regulation of mRNA during meiotic maturation while mVL30 and RLTR1B transcripts appear relatively stable. This latter observation does not necessarily contradict the milrinone result as milrinone treatment of GV-oocytes lasted for 48 h while the MII and zygote samples were isolated about 20 h after induction of meiotic maturation *in vivo*. Thus, mRNA degradation may not be apparent within this shorter time period and could explain why mVL30 and RLTR1B do not show a clear down-regulation. Alternatively, the mechanism degrading certain repetitive transcripts in the fully grown GV-oocytes may not efficiently work during meiotic maturation. The differential behavior of IAP illustrates however the diversity in post-transcriptional regulation of repetitive elements. The strong down-regulation of IAP mRNA levels in the MII eggs supports the idea that meiotic maturation enhances the degradation process of IAP mRNA possibly by incorporating one of the known mechanisms degrading maternal mRNAs upon resumption of meiosis [30].

The IAP retrotransposon is one of the most aggressive parasitic sequences known in the mouse genome and present in approximately 1000 copies per haploid [45,46]. IAP is highly expressed in the oocyte and early embryos [23]. It is estimated that there are ~13,000 transcripts in the oocyte, which are reduced ~10-fold in the ovulated egg and then increased up to ~150,000 transcripts between the one-cell and the blastocyst stages [23]. Microarray data also confirmed that IAP mRNA is an abundant transcript;

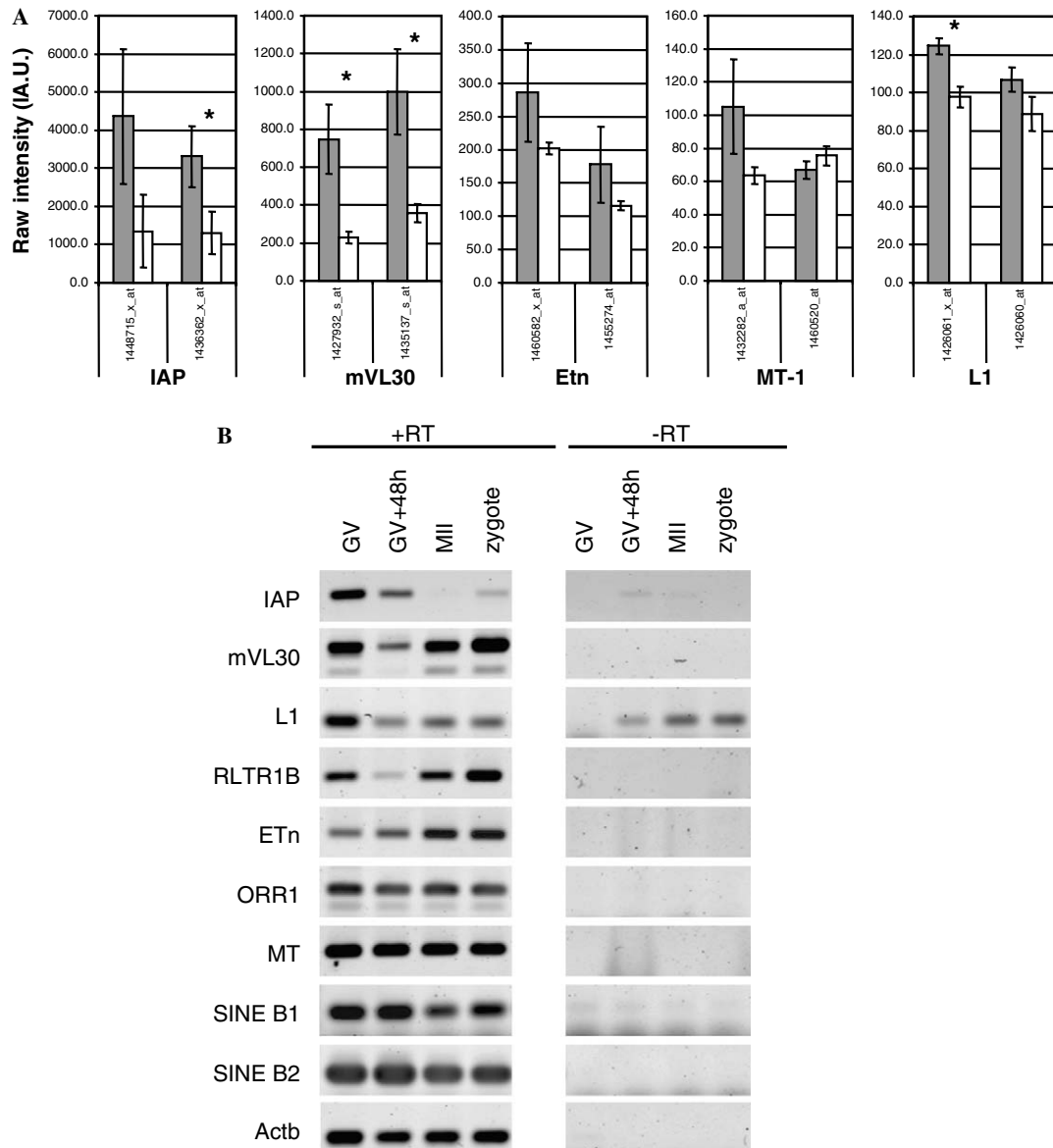


Fig. 3. Relative changes of transcript abundance for several different mobile elements. (A) Effect of milrinone treatment on expression of different repetitive elements. For IAP, mVL30, L1 and Etn, expression of two most specific Affymetrix MOE 430 2.0 probe sets with the highest hybridization signal (Supplemental Table 2) is displayed as a mean raw expression value (Y-axis) for control (GV) and milrinone-treated oocytes (GV + 48 h). \* indicates *t*-test *p*-value < 0.05. (B) RT-PCR analysis of expression of repetitive elements performed on the same material as shown in Fig. 1D.

IAP probe set raw signal reached 4133, which is more than eight times higher than the average raw signal for transcripts on the chip and places the IAP probe set among the 7% of probe sets with the highest signal. Our previous study of IAP expression in the early embryo found that the knockdown of Dicer in the embryo results in approximately 50% increase in the steady-state level of IAP (from ~60,000 to 90,000 transcripts [20]); this increase would likely be greater if maternal Dicer protein was depleted.

Our previous results indicate that ~1000 molecules of dsRNA are required to trigger RNAi-mediated mRNA degradation [16]. Thus, endogenous RNAi likely requires formation of sufficient amounts of dsRNA to elicit efficient transcript targeting. Production of such dsRNA likely correlates with the level of RNA expression. Retroelements

can produce dsRNA by a variety of mechanisms [47] that generate complementary strands. New insertions in the genome also increase the probability of generating antisense transcripts. Indeed, expression of antisense RNA has been demonstrated for MuERV-L [20,21] and IAP retrotransposons in the 2-cell mouse embryo [20]. Although there is no solid evidence for base-pairing of endogenous transcripts, it has been shown experimentally that simultaneous expression of longer sense and antisense transcripts can trigger sequence-specific mRNA degradation in mammalian cells [48]. Thus, high expression levels of a mobile element may increase both the probability of successful retrotransposition as well as the risk of pairing with antisense transcripts. Activation of RNAi by forming these dsRNAs would thereby constrain their expression. Repetitive

elements expressed at low levels would have a lower probability to form enough dsRNA and thereby may escape RNAi targeting. A consequence would be that RNAi-mediated targeting of mobile elements would be rather selective, i.e., it would constrain only those elements that produce enough dsRNA to trigger RNAi. IAP is likely the most abundant autonomous TE mRNA in the mouse oocyte and it is likely a target for RNAi because depletion of Dicer in early embryo cells lead to accumulation of IAP transcripts [20,24]. In addition, IAP mRNA instability in the GV-oocyte is also consistent with the hypothesis that IAP is a natural RNAi target in mouse.

## Acknowledgments

We thank Herbert Angliker from the FMI microarray facility. Research at the Friedrich Miescher Institute is supported by the Novartis Research Foundation. M.P. acknowledges the Boehringer Ingelheim Fonds for her PhD fellowship and The Company of Biologists for the Development Travelling Fellowship. P.S. is supported by an EMBO long-term fellowship ALTF 2003-199; research in the Schultz laboratory by NIH Grant HD22681 and research in the Peters laboratory by the NoE network “The Epigenome” (LSHG-CT-2004-503433).

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.bbrc.2006.06.106](https://doi.org/10.1016/j.bbrc.2006.06.106).

## References

- [1] R.H. Waterston, K. Lindblad-Toh, E. Birney, J. Rogers, J.F. Abril, P. Agarwal, R. Agarwala, R. Ainscough, M. Alexandersson, P. An, S.E. Antonarakis, J. Attwood, R. Baertsch, J. Bailey, K. Barlow, S. Beck, E. Berry, B. Birren, T. Bloom, P. Bork, M. Botcherby, N. Bray, M.R. Brent, D.G. Brown, S.D. Brown, C. Bult, J. Burton, J. Butler, R.D. Campbell, P. Carninci, S. Cawley, F. Chiaromonte, A.T. Chinwalla, D.M. Church, M. Clamp, C. Clee, F.S. Collins, L.L. Cook, R.R. Copley, A. Coulson, O. Couronne, J. Cuff, V. Curwen, T. Cutts, M. Daly, R. David, J. Davies, K.D. Delehaunty, J. Deri, E.T. Dermizakis, C. Dewey, N.J. Dickens, M. Diekhans, S. Dodge, I. Dubchak, D.M. Dunn, S.R. Eddy, L. Elnitski, R.D. Emes, P. Esvara, E. Eyra, A. Felsenfeld, G.A. Fewell, P. Flicek, K. Foley, W.N. Frankel, L.A. Fulton, R.S. Fulton, T.S. Furey, D. Gage, R.A. Gibbs, G. Glusman, S. Gnerre, N. Goldman, L. Goodstadt, D. Grafham, T.A. Graves, E.D. Green, S. Gregory, R. Guigo, M. Guyer, R.C. Hardison, D. Haussler, Y. Hayashizaki, L.W. Hillier, A. Hinrichs, W. Hlavina, T. Holzer, F. Hsu, A. Hua, T. Hubbard, A. Hunt, I. Jackson, D.B. Jaffe, L.S. Johnson, M. Jones, T.A. Jones, A. Joy, M. Kamal, E.K. Karlsson, et al., Initial sequencing and comparative analysis of the mouse genome, *Nature* 420 (2002) 520–562.
- [2] E.S. Lander, L.M. Linton, B. Birren, C. Nusbaum, M.C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczyk, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J.P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J.C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Showkeen, S. Sims, R.H. Waterston, R.K. Wilson, L.W. Hillier, J.D. McPherson, M.A. Marra, E.R. Mardis, L.A. Fulton, A.T. Chinwalla, K.H. Pepin, W.R. Gish, S.L. Chissoe, M.C. Wendl, K.D. Delehaunty, T.L. Miner, A. Delehaunty, J.B. Kramer, L.L. Cook, R.S. Fulton, D.L. Johnson, P.J. Minx, S.W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J.F. Cheng, A. Olsen, S. Lucas, C. Elkin, Initial sequencing and analysis of the human genome, *Nature* 409 (2001) 860–921.
- [3] J. Jurka, V.V. Kapitonov, A. Pavlicek, P. Klonowski, O. Kohany, J. Walichiewicz, Repbase Update, a database of eukaryotic repetitive elements, *Cytogenet. Genome Res.* 110 (2005) 462–467.
- [4] B. Brouha, J. Schustak, R.M. Badge, S. Lutz-Prigge, A.H. Farley, J.V. Moran, H.H. Kazazian Jr., Hot L1s account for the bulk of retrotransposition in the human population, *Proc. Natl. Acad. Sci. USA* 100 (2003) 5280–5285.
- [5] T.H. Bestor, Sex brings transposons and genomes into conflict, *Genetica* 107 (1999) 289–295.
- [6] D. Bourc’his, T.H. Bestor, Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L, *Nature* 431 (2004) 96–99.
- [7] K. Hata, Y. Sakaki, Identification of critical CpG sites for repression of L1 transcription by DNA methylation, *Gene* 189 (1997) 227–234.
- [8] C.P. Walsh, J.R. Chaillet, T.H. Bestor, Transcription of IAP endogenous retroviruses is constrained by cytosine methylation, *Nat. Genet.* 20 (1998) 116–117.
- [9] J.H. Martens, R.J. O’Sullivan, U. Braunschweig, S. Opravil, M. Radolf, P. Steinlein, T. Jenuwein, The profile of repeat-associated histone lysine methylation states in the mouse epigenome, *EMBO J.* 24 (2005) 800–812.
- [10] J. Huang, T. Fan, Q. Yan, H. Zhu, S. Fox, H.J. Issaq, L. Best, L. Gangi, D. Munroe, K. Muegge, Lsh, an epigenetic guardian of repetitive elements, *Nucleic Acids Res.* 32 (2004) 5019–5028.
- [11] G. Meister, T. Tuschl, Mechanisms of gene silencing by double-stranded RNA, *Nature* 431 (2004) 343–349.
- [12] T. Sijen, R.H. Plasterk, Transposon silencing in the *Caenorhabditis elegans* germ line by natural RNAi, *Nature* 426 (2003) 310–314.
- [13] A.I. Kalmykova, M.S. Klenov, V.A. Gvozdev, Argonaute protein PIWI controls mobilization of retrotransposons in the *Drosophila* male germline, *Nucleic Acids Res.* 33 (2005) 2052–2059.
- [14] V.J. Robert, N.L. Vastenhouw, R.H. Plasterk, RNA interference, transposon silencing, and cosuppression in the *Caenorhabditis elegans* germ line: similarities and differences, *Cold Spring Harb. Symp. Quant. Biol.* 69 (2004) 397–402.
- [15] S. Yang, S. Tutton, E. Pierce, K. Yoon, Specific double-stranded RNA interference in undifferentiated mouse embryonic stem cells, *Mol. Cell. Biol.* 21 (2001) 7807–7816.
- [16] P. Svoboda, P. Stein, H. Hayashi, R.M. Schultz, Selective reduction of dormant maternal mRNAs in mouse oocytes by RNA interference, *Development* 127 (2000) 4147–4156.
- [17] F. Wianny, M. Zernicka-Goetz, Specific interference with gene function by double-stranded RNA in early mouse development, *Nat. Cell Biol.* 2 (2000) 70–75.
- [18] P. Stein, F. Zeng, H. Pan, R.M. Schultz, Absence of non-specific effects of RNA interference triggered by long double-stranded RNA in mouse oocytes, *Dev. Biol.* 286 (2005) 464–471.
- [19] D.A. Kramerov, M.I. Bukrinsky, A.P. Ryskov, DNA sequences homologous to long double-stranded RNA. Transcription of intracisternal A-particle genes and major long repeat of the mouse genome, *Biochim. Biophys. Acta* 826 (1985) 20–29.
- [20] P. Svoboda, P. Stein, M. Anger, E. Bernstein, G.J. Hannon, R.M. Schultz, RNAi and expression of retrotransposons MuERV-L and IAP in preimplantation mouse embryos, *Dev. Biol.* 269 (2004) 276–285.

- [21] A.E. Peaston, A.V. Evsikov, J.H. Graber, W.N. de Vries, A.E. Holbrook, D. Solter, B.B. Knowles, Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos, *Dev. Cell* 7 (2004) 597–606.
- [22] C.E. Park, M.R. Shin, E.H. Jeon, S.H. Lee, K.Y. Cha, K. Kim, N.H. Kim, K.A. Lee, Oocyte-selective expression of MT transposon-like element, clone MTi7 and its role in oocyte maturation and embryo development, *Mol. Reprod. Dev.* 69 (2004) 365–374.
- [23] L. Piko, M.D. Hammons, K.D. Taylor, Amounts, synthesis, and some properties of intracisternal A particle-related RNA in early mouse embryos, *Proc. Natl. Acad. Sci. USA* 81 (1984) 488–492.
- [24] C. Kanellopoulou, S.A. Muljo, A.L. Kung, S. Ganesan, R. Drapkin, T. Jenuwein, D.M. Livingston, K. Rajewsky, Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing, *Genes Dev.* 19 (2005) 489–501.
- [25] C. Bouniol-Baly, L. Hamraoui, J. Guibert, N. Beaujean, M.S. Szollosi, P. Debey, Differential transcriptional activity associated with chromatin configuration in fully grown mouse germinal vesicle oocytes, *Biol. Reprod.* 60 (1999) 580–587.
- [26] F. Zeng, R.M. Schultz, RNA transcript profiling during zygotic gene activation in the preimplantation mouse embryo, *Dev. Biol.* 283 (2005) 40–57.
- [27] F. Zeng, D.A. Baldwin, R.M. Schultz, Transcript profiling during preimplantation mouse development, *Dev. Biol.* 272 (2004) 483–496.
- [28] R. Bachvarova, V. De Leon, Polyadenylated RNA of mouse ova and loss of maternal RNA in early development, *Dev. Biol.* 74 (1980) 1–8.
- [29] P.T. Brower, E. Gizang, S.M. Boreen, R.M. Schultz, Biochemical studies of mammalian oogenesis: synthesis and stability of various classes of RNA during growth of the mouse oocyte in vitro, *Dev. Biol.* 86 (1981) 373–383.
- [30] Z. Alizadeh, S. Kageyama, F. Aoki, Degradation of maternal mRNA in mouse embryos: selective degradation of specific mRNAs after fertilization, *Mol. Reprod. Dev.* 72 (2005) 281–290.
- [31] L.P. Lim, N.C. Lau, P. Garrett-Engle, A. Grimson, J.M. Schelter, J. Castle, D.P. Bartel, P.S. Linsley, J.M. Johnson, Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs, *Nature* 433 (2005) 769–773.
- [32] S. Bagga, J. Bracht, S. Hunter, K. Massirer, J. Holtz, R. Eachus, A.E. Pasquinelli, Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation, *Cell* 122 (2005) 553–563.
- [33] M.A. Valencia-Sanchez, J. Liu, G.J. Hannon, R. Parker, Control of translation and mRNA degradation by miRNAs and siRNAs, *Genes Dev.* 20 (2006) 515–524.
- [34] R.S. Pillai, MicroRNA function: multiple mechanisms for a tiny RNA? *RNA* 11 (2005) 1753–1761.
- [35] S. Griffiths-Jones, R.J. Grocock, S. van Dongen, A. Bateman, A.J. Enright, miRBase: microRNA sequences, targets and gene nomenclature, *Nucleic Acids Res.* 34 (2006) D140–D144.
- [36] B.P. Lewis, C.B. Burge, D.P. Bartel, Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets, *Cell* 120 (2005) 15–20.
- [37] A. Krek, D. Grun, M.N. Poy, R. Wolf, L. Rosenberg, E.J. Epstein, P. MacMenamin, I. da Piedade, K.C. Gunsalus, M. Stoffel, N. Rajewsky, Combinatorial microRNA target predictions, *Nat. Genet.* 37 (2005) 495–500.
- [38] M. Kiriakidou, P.T. Nelson, A. Kouranov, P. Fitziev, C. Bouyioukos, Z. Mourelatos, A. Hatzigeorgiou, A combined computational-experimental approach predicts human microRNA targets, *Genes Dev.* 18 (2004) 1165–1178.
- [39] B.P. Lewis, I.H. Shih, M.W. Jones-Rhoades, D.P. Bartel, C.B. Burge, Prediction of mammalian microRNA targets, *Cell* 115 (2003) 787–798.
- [40] B. John, A.J. Enright, A. Aravin, T. Tuschl, C. Sander, D.S. Marks, Human MicroRNA targets, *PLoS Biol.* 2 (2004) e363.
- [41] J. Krutzfeldt, N. Rajewsky, R. Braich, K.G. Rajeev, T. Tuschl, M. Manoharan, M. Stoffel, Silencing of microRNAs in vivo with ‘antagomirs’, *Nature* 438 (2005) 685–689.
- [42] J. Rehwinkel, P. Natalin, A. Stark, J. Brennecke, S.M. Cohen, E. Izaurralde, Genome-wide analysis of mRNAs regulated by Droscha and Argonaute proteins in *Drosophila melanogaster*, *Mol. Cell. Biol.* 26 (2006) 2965–2975.
- [43] T. Somfai, K. Kikuchi, A. Onishi, M. Iwamoto, D. Fuchimoto, A.B. Papp, E. Sato, T. Nagai, Meiotic arrest maintained by cAMP during the initiation of maturation enhances meiotic potential and developmental competence and reduces polyspermy of IVM/IVF porcine oocytes, *Zygote* 11 (2003) 199–206.
- [44] E. Sarot, G. Payen-Groschene, A. Bucheton, A. Pelisson, Evidence for a piwi-dependent RNA silencing of the gypsy endogenous retrovirus by the *Drosophila melanogaster* flamenco gene, *Genetics* 166 (2004) 1313–1321.
- [45] H.H. Kazazian Jr., J.V. Moran, The impact of L1 retrotransposons on the human genome, *Nat. Genet.* 19 (1998) 19–24.
- [46] E.L. Kuff, K.K. Lueders, The intracisternal A-particle gene family: structure and functional aspects, *Adv. Cancer Res.* 51 (1988) 183–276.
- [47] P.M. Waterhouse, M.B. Wang, T. Lough, Gene silencing as an adaptive defence against viruses, *Nature* 411 (2001) 834–842.
- [48] J. Wang, E. Tekle, H. Oubrahim, J.J. Mieyal, E.R. Stadtman, P.B. Chock, Stable and controllable RNA interference: investigating the physiological function of glutathionylated actin, *Proc. Natl. Acad. Sci. USA* 100 (2003) 5103–5106.

## **2.2. PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos**

Mareike Puschendorf, Rémi Terranova, Erwin Boutsma, Xiaohong Mao, Kyo-ichi Isono, Urszula Brykczynska, Carolin Kolb, Arie P. Otte, Haruhiko Koseki, Stuart H. Orkin, Maarten van Lohuizen and Antoine H.F.M. Peters

# PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos

Mareike Puschendorf<sup>1</sup>, Rémi Terranova<sup>1</sup>, Erwin Boutsma<sup>2</sup>, Xiaohong Mao<sup>3,6</sup>, Kyo-ichi Isono<sup>4</sup>, Urszula Brykczynska<sup>1</sup>, Carolin Kolb<sup>1</sup>, Arie P Otte<sup>5</sup>, Haruhiko Koseki<sup>4</sup>, Stuart H Orkin<sup>3</sup>, Maarten van Lohuizen<sup>2</sup> & Antoine H F M Peters<sup>1</sup>

**In eukaryotes, *Suv39h* H3K9 trimethyltransferases are required for pericentric heterochromatin formation and function. In early mouse preimplantation embryos, however, paternal pericentric heterochromatin lacks *Suv39h*-mediated H3K9me3 and downstream marks. Here we demonstrate *Ezh2*-independent targeting of maternally provided polycomb repressive complex 1 (PRC1) components to paternal heterochromatin. In *Suv39h2* maternally deficient zygotes, PRC1 also associates with maternal heterochromatin lacking H3K9me3, thereby revealing hierarchy between repressive pathways. In *Rnf2* maternally deficient zygotes, the PRC1 complex is disrupted, and levels of pericentric major satellite transcripts are increased at the paternal but not the maternal genome. We conclude that in early embryos, *Suv39h*-mediated H3K9me3 constitutes the dominant maternal transgenerational signal for pericentric heterochromatin formation. In absence of this signal, PRC1 functions as the default repressive back-up mechanism. Parental epigenetic asymmetry, also observed along cleavage chromosomes, is resolved by the end of the 8-cell stage—concurrent with blastomere polarization—marking the end of the maternal-to-embryonic transition.**

In mammals, parental genomes are epigenetically distinct, despite their genetic resemblance<sup>1</sup>. During early mouse preimplantation development, parental genomes are highly asymmetric in epigenetic modifications of DNA and associated chromatin<sup>2–9</sup>. At gamete fusion, the maternal genome exists in a nucleosomal configuration marked by distinct types of histone lysine methylation inherited from the oocyte. In contrast, following the histone-to-protamine exchange occurring during spermiogenesis, the paternal genome incorporates maternally provided histones and becomes *de novo* methylated at different lysine residues in a highly spatially and temporally coordinated manner. The function of parental epigenetic asymmetry for gene expression and genome reorganization<sup>6,10</sup> is largely enigmatic, as are the mechanisms of establishment, maintenance and resolution. A key question is whether parentally inherited epigenetic states affect *de novo* targeting and function of (maternally provided) epigenetic modifiers in *cis* and/or in *trans* in the early embryo, thereby directing gene expression over shorter or longer developmental time windows. Notably, transmission of the paternal genome in a nucleosomal state impairs DNA methylation reprogramming in early embryos<sup>11</sup>. Here, we study the transgenerational contribution of two distinct evolutionarily conserved classes of epigenetic modifiers in defining parental

asymmetry at constitutive heterochromatin and euchromatin in preimplantation embryos.

The first class consists of the *Suv39h* histone methyltransferases (HMTs), which are essential for constitutive heterochromatin formation and function, gene repression and maintenance of genome integrity<sup>12–15</sup>. *Suv39h*-mediated H3K9me3 directs chromatin binding of the heterochromatic proteins HP1 $\alpha$ , HP $\beta$  and HP $\gamma$  (ref. 16), which target the two H4K20 di- and trimethylation-specific Suv4-20h HMTs and the Dnmt3a/3b DNA methyltransferases, to establish a transcriptionally repressed state<sup>17–19</sup>.

The second class consists of Polycomb group (PcG) proteins, which are repressive chromatin factors required for maintaining cell identity<sup>14,20</sup>. PcG proteins are classified into two groups of multimeric protein complexes termed polycomb repressive complexes (PRCs). In *Drosophila melanogaster*, PRC1 contains four core components for which multiple paralogs exist in mammals<sup>21</sup>. *In vitro*, PRC1 mediates repression by inhibiting chromatin remodeling, impairing the transcription machinery and by mediating chromatin compaction<sup>22</sup>. The mammalian and fly RING orthologs function as E3 ubiquitin ligases that monoubiquitinate H2A at lysine 119, a modification associated with gene repression<sup>23,24</sup>. PRC2 consists of *Ezh2*, *Suz12* and *Eed*, which

<sup>1</sup>Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, CH-4058 Basel, Switzerland. <sup>2</sup>Division of Molecular Genetics and Centre for Biomedical Genetics, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands. <sup>3</sup>Department of Pediatric Oncology, Dana Farber Cancer Institute, Harvard Stem Cell Institute and Howard Hughes Medical Institute, Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>4</sup>RIKEN Research Center for Allergy and Immunology, RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku Yokohama City, Kanagawa 230-0045, Japan. <sup>5</sup>Swammerdam Institute for Life Sciences, University of Amsterdam, Kruislaan 406, 1098 SM Amsterdam, The Netherlands. <sup>6</sup>Present address: Novartis Institutes for BioMedical Research, Inc., 250 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA. Correspondence should be addressed to A.H.F.M.P. (antoine.peters@fmi.ch).

Received 14 November 2007; accepted 22 January 2008; published online 2 March 2008; corrected online 16 March 2008 (details online); doi:10.1038/ng.99



together mediate H3K27me2 and H3K27me3 (H3K27me2/3)<sup>21,25</sup>. Genome-wide chromatin profiling has shown that components of PRC1 and PRC2 co-occupy promoters of genes enriched in H3K27me3 (ref. 26). At selected genes, PRC2 is required for binding of PRC1 (ref. 26), suggesting that PRC1 can be targeted via mammalian orthologs of *Drosophila* Polycomb that bind H3K27me3 (refs. 27,28). In *Eed*<sup>-/-</sup> embryonic stem cells, most joint targets of PRC1 and PRC2 are transcriptionally activated<sup>26</sup>.

In this study, we addressed the parental influence on constitutive heterochromatin formation in early mouse embryos. In zygotes, only maternal constitutive heterochromatin is labeled by HP1 $\beta$  and H3K9me2/3 marks inherited from the oocyte<sup>2,6,7,9</sup>. We investigated whether absence of the *Suv39h*-dependent chromatin signature at paternal constitutive heterochromatin is compensated by targeting of other repressive histone modifications and/or proteins. We show that H3K27me2/3 and PRC1 components accumulate at paternal heterochromatin. In contrast to promoters of certain developmental regulators<sup>26</sup>, PRC1 targeting to heterochromatin is independent of PRC2 function. PRC1 is required for transcriptional repression of heterochromatic major satellite repeats. We further demonstrate that parental-specific *Suv39h* and PRC1-defined states are inherited over the first three cleavage divisions and that asymmetry is not limited to constitutive heterochromatin only. Finally, we explain the basic principle underlying parental origin-specific definition of

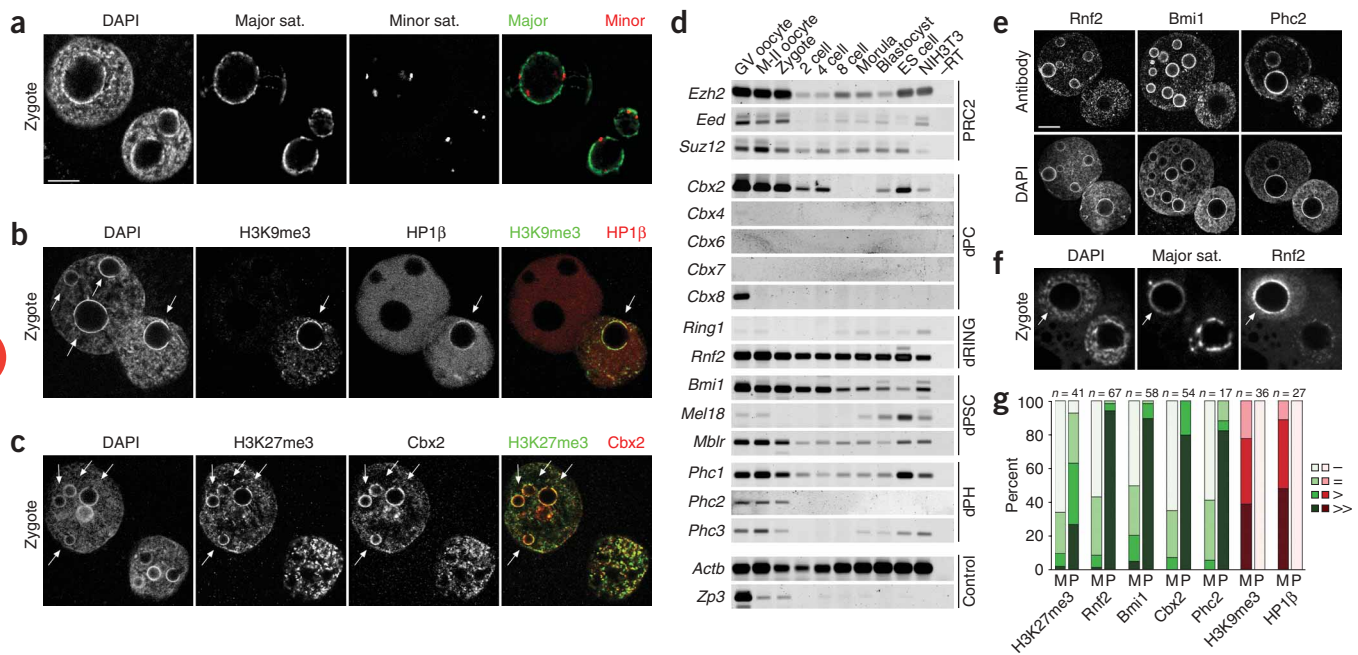
constitutive heterochromatin by *Suv39h* and Polycomb-based repression mechanisms.

## RESULTS

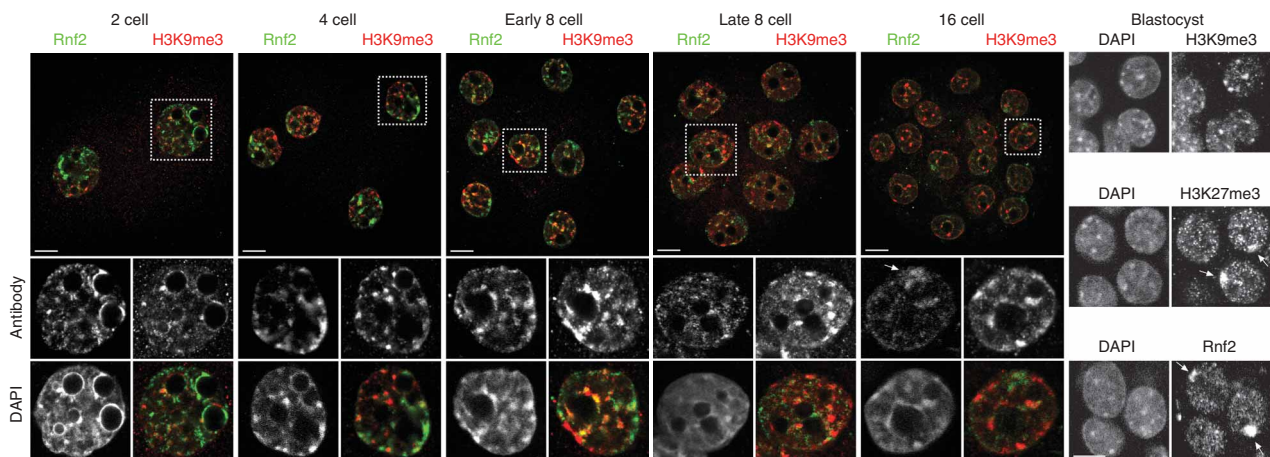
### PRC1 components define paternal heterochromatin

In mouse somatic cells, pericentric heterochromatin of different chromosomes cluster into chromocenters that can be visualized by fluorochromes such as 4,6-diamidino-2-phenylindole (DAPI) that preferentially bind to the underlying AT-rich major satellite sequences<sup>29</sup>. In zygotes, DAPI-intense chromatin is organized into ring structures around nucleolar-precursor bodies (NPBs)<sup>6</sup>. By DNA FISH for major satellites, we show that pericentric heterochromatin co-localizes with DAPI-intense rings around NPBs (Fig. 1a)<sup>30</sup>. Centromeric minor satellite foci are interspersed within major satellites (Fig. 1a)<sup>30</sup>. We confirm that in zygotes, only maternal constitutive heterochromatin is labeled by H3K9me3 and H4K20me3 (Fig. 1b and Supplementary Fig. 1 online)<sup>2,6,7,9</sup>, two modifications that are inherited from the oocyte. HP1 $\beta$  is loaded onto maternal heterochromatin upon gamete fusion (Fig. 1b and Supplementary Fig. 1).

In embryonic stem cells, pericentric heterochromatin acquires H3K27me3 in the absence of the *Suv39h*-mediated chromatin configuration<sup>31</sup>. To determine whether such a compensatory mechanism also operates in early embryos, we analyzed H3K27 methylation states in late zygotes. Indeed, H3K27me2/3 modifications were



**Figure 1** Maternal and paternal heterochromatin are marked by distinct repressive complexes in preimplantation embryos. **(a)** DNA-FISH analysis showing localization of major satellite and minor satellite sequences along DAPI-intense ring structures around NPBs in late mouse zygotes. **(b)** Immunofluorescence analysis of late zygotes showing euchromatic localization of HP1 $\beta$  in the large paternal pronucleus (left) and genome-wide H3K9me3 and HP1 $\beta$  staining in the maternal pronucleus (right). Constitutive heterochromatin (arrows) is only maternally labeled by H3K9me3 and HP1 $\beta$ . **(c)** Immunofluorescence analysis of late zygotes showing euchromatic staining of H3K27me3 and Cbx2 in both pronuclei. Constitutive heterochromatin is only paternally labeled by H3K27me3 and Cbx2 (arrows). **(d)** RT-PCR analysis of PRC2 and PRC1 components in oocytes and preimplantation embryos. RNA from embryonic stem cells and NIH3T3 cells was used as control. -RT refers to analysis of M-II oocytes without reverse transcriptase. Ubiquitously expressed  $\beta$ -actin (*Actb*) and oocyte-restricted zona pellucida 3 (*Zp3*) serve as controls. For mouse PRC1 genes, *Drosophila* orthologs are given on the right. **(e)** Immunofluorescence analysis of late zygotes showing heterochromatic enrichment of Rnf2, Bmi1 and Phc2 paternally only. All three proteins label euchromatin of both pronuclei. **(f)** Co-immuno-DNA-FISH analysis showing co-localization of major satellite sequences with Rnf2 in the paternal pronucleus (arrow). **(g)** Levels of enrichment of H3K9me3, HP1 $\beta$ , H3K27me3 and PRC1 components at maternal (M) versus paternal (P) constitutive heterochromatin in late zygotes were scored as follows: (-) no staining; (=) equal hetero- and euchromatic staining; (>) enhanced and (>>) strongly enhanced staining of heterochromatin versus euchromatin. Scale bars, 10  $\mu$ m.



**Figure 2** Differential heterochromatic states are maintained up to the 8-cell stage in a parental origin-dependent manner. Immunofluorescence analyses of Rnf2 and H3K9me3 in 2-cell, 4-cell, early 8-cell (before compaction), late 8-cell (after compaction), 16-cell and blastocyst embryos. In 2-cell embryos, exclusive enrichment of either H3K9me3 or Rnf2 is detected at individual DAPI-intense heterochromatic chromocenters. Following the gradual intermingling of parental genomes, chromocenters become progressively doubly labeled by H3K9me3 and Rnf2. From the late 8-cell stage onwards, Rnf2 is completely replaced by H3K9me3 at heterochromatin. From the 16-cell stage onwards, Rnf2 is enriched at the presumptive inactive X chromosome (arrow) in female embryos. Scale bars, 20  $\mu\text{m}$ . In blastocysts, H3K9me3 marks DAPI-intense chromocenters, whereas H3K27me3 and Rnf2 are enriched at the presumptive inactive X chromosome (arrows). Other matPRC1 components were not enriched (Phc2) or occasionally weakly enriched (Cbx2) at the inactive X (data not shown). Scale bar, 10  $\mu\text{m}$ .

enriched at pericentric heterochromatin of paternal origin only (Fig. 1c and Supplementary Fig. 1), whereas the remainder of both parental genomes (referred to as ‘euchromatin’) was labeled by both marks<sup>7</sup>. H3K27me1 labeled euchromatin and was enriched at heterochromatin of both pronuclei (Supplementary Fig. 1). Therefore, H3K27me states define constitutive heterochromatin depending on parental descent.

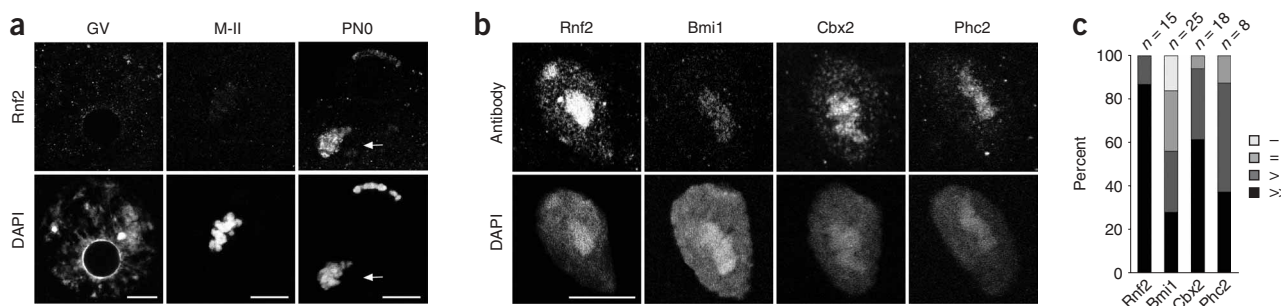
To identify PcG complexes functioning in early embryos<sup>21,25</sup>, we profiled RNA expression of PRC2 and PRC1 genes in oocytes and preimplantation embryos. For PRC2 components, we observed strong maternal and moderate zygotic expression (Fig. 1d). Similarly, at least one mammalian paralog of each *Drosophila* PRC1 core member was maternally and zygotically expressed (Fig. 1d). These data suggest that PRC2 and PRC1 have roles in preimplantation development.

We next asked whether H3K27me3 could, in principle, target PRC1 complexes to pericentric heterochromatin. We observed that,

analogous to the binding of HP1 $\beta$  to H3K9me3<sup>16</sup>, H3K27me3 co-localized with Cbx2 (Fig. 1c), a Polycomb protein known to bind to H3K27me3 via its chromodomain<sup>27</sup>. In addition, Rnf2 (Ring1b), Bmi1 and Phc2 also selectively accumulated at paternal constitutive heterochromatin (Fig. 1e), suggesting the presence of a functional ‘matPRC1’ complex (‘matPRC1’). Immuno-DNA-FISH analysis confirmed co-localization of the matPRC1 complex with major satellites in the paternal genome (Fig. 1f). Euchromatin of both genomes was labeled by all four matPRC1 proteins as well as by Phc1 that was absent from paternal heterochromatin (data not shown). These data convincingly support a parental origin-dependent definition of constitutive heterochromatin (Fig. 1g).

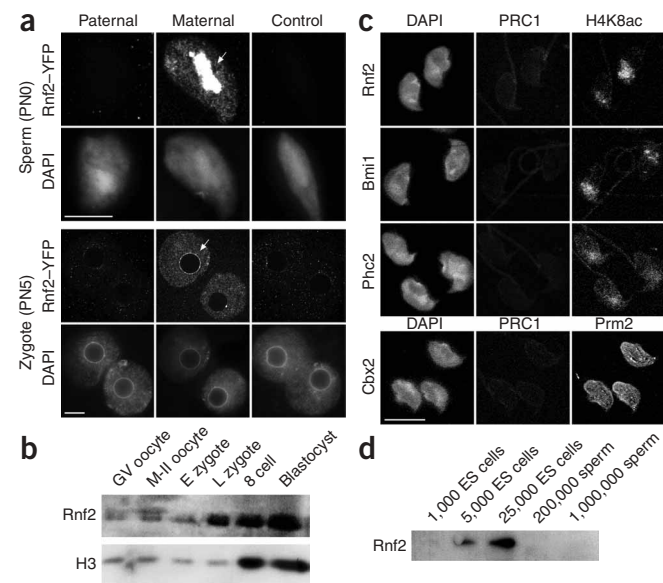
### Parental-specific heterochromatic states are heritable

To investigate whether parental-specific heterochromatic states are transmitted, we profiled successive cleavage-stage embryos (Fig. 2 and Supplementary Fig. 2 online;  $n = 15\text{--}25$  per stage) in which



**Figure 3** PRC1 components are targeted to chromatin upon gamete fusion. (a) Immunofluorescence analyses of germinal vesicle (GV) and M-II oocytes and pronuclear stage 0 (PNO) zygotes show that Rnf2 starts to bind to the maternal genome shortly after gamete fusion. At PNO, the maternal genome is at anaphase of the 2nd meiotic division. The strongly labeled genome complement will constitute the embryo proper (bottom; arrow). The top set will segregate into the second polar body. Similar data were obtained for matPRC1 components Cbx2, Bmi1 and Phc2 (Supplementary Fig. 4). (b) Immunofluorescence analyses of matPRC1 enrichment at constitutive heterochromatin in decondensed sperm nuclei shortly after gamete fusion (PNO stage). Constitutive heterochromatin is organized into one large chromocenter in most mature spermatozoa. (c) Numerical evaluation of data presented in b and c. Level of enrichment was scored as described in Fig. 1g. Scale bars, 10  $\mu\text{m}$ .





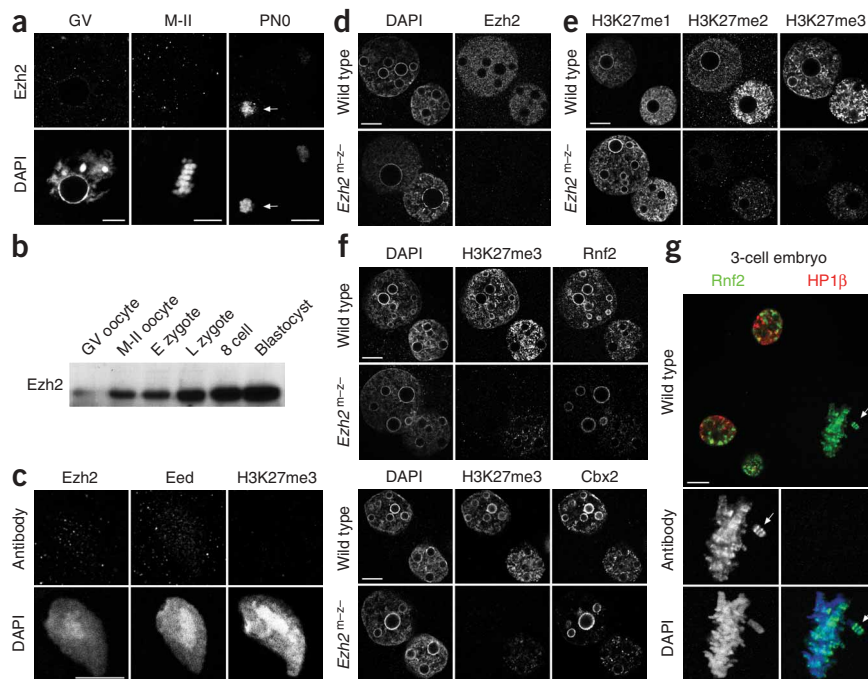
**Figure 4** Maternally provided Rnf2 is targeted to paternal constitutive heterochromatin and euchromatin. **(a)** Immunofluorescence analysis with antibodies to GFP shows binding of Rnf2-YFP fusion protein to paternal heterochromatin of decondensing spermatozoa and late zygotes only after maternal (arrows in middle panels) but not paternal (left panels) transmission of a functional Rnf2-YFP knock-in allele. Paternal and maternal euchromatin is also bound by maternal fusion protein (bottom middle panel). Scale bars, 10  $\mu$ m. **(b)** Protein blot analysis of Rnf2 and histone H3 in germinal vesicle and M-II oocytes, in early (E) and late (L) zygotes, and in 8 cell and blastocyst embryos. Rnf2 protein levels strongly increased during the development of 1-cell embryos. We loaded 200 oocytes or embryos per lane. **(c)** Immunofluorescence analysis of isolated caudal epididymal spermatozoa shows labeling with H4K8ac and Protamine2 but not with PRC1 components. **(d)** Protein blot analysis of Rnf2 in isolated caudal epididymal spermatozoa and mouse embryonic stem cells fails to detect Rnf2 protein in 1,000,000 spermatozoa that are estimated to contain about 1% histones (U.B. and A.H.F.M.P., unpublished data), which is equivalent to the amount of histones present in 5,000 diploid embryonic stem cells.

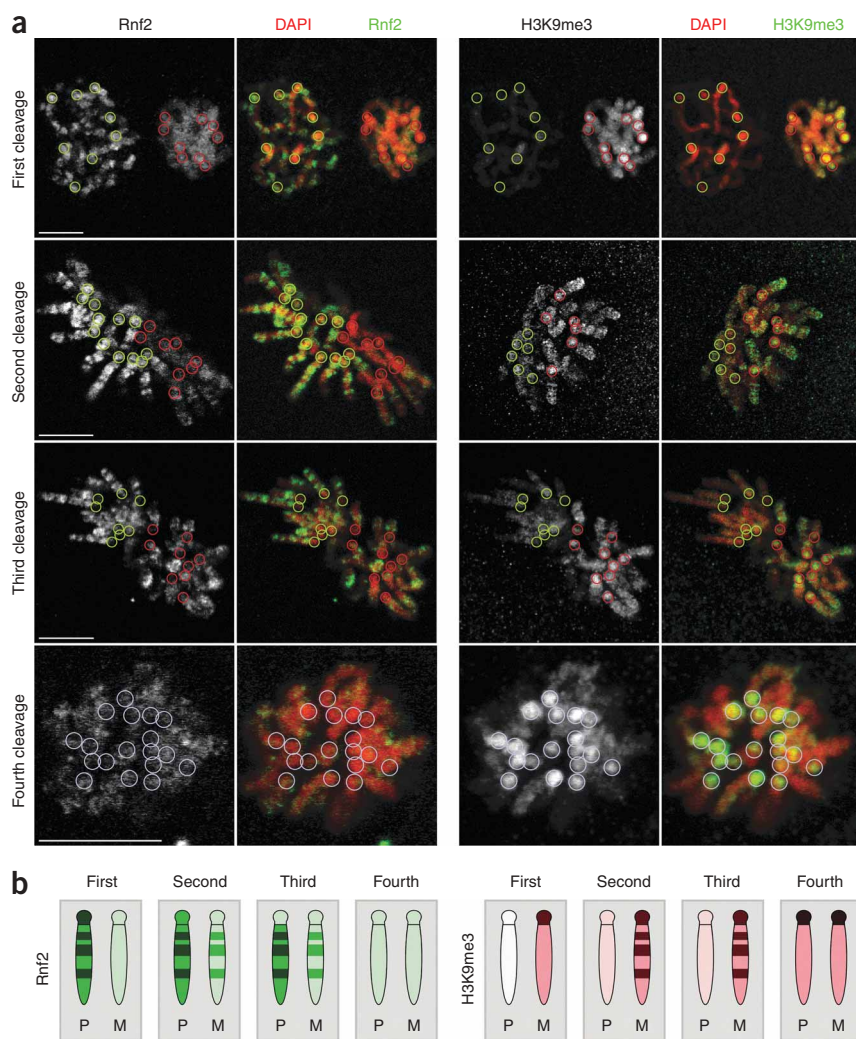
heterochromatin becomes progressively reorganized and clustered into chromocenters<sup>6</sup>. In 2- and 4-cell embryos, individual DAPI-intense areas were labeled with either H3K9me3 (and HP1 $\beta$ ) or matPRC1 components. Besides HP1 $\beta$ , HP1 $\alpha$  and HP1 $\gamma$  were also associated to maternal H3K9me3-positive heterochromatin from the 4-cell stage onwards (data not shown). Following gradual intermingling of parental genomes<sup>32</sup>, chromocenters became doubly labeled by H3K9me3 and Rnf2, particularly in early 8-cell embryos. Rnf2 association with

pericentric heterochromatin strongly declined during the 8-cell stage and was completely replaced by H3K9me3 from the 16-cell stage onwards. At that point, Rnf2 started to accumulate at the presumptive Xi chromosome. In blastocysts, constitutive heterochromatin was labeled by H3K9me3 and HP1 $\beta$  ( $n = 10$ ), whereas H3K27me3 and Rnf2 were enriched at facultative heterochromatin of the Xi ( $n = 20-30$ )<sup>23</sup>.

To relate labeling of heterochromatin by Rnf2 versus H3K9me3 and HP1 $\beta$  to parental origin, we studied 4-cell embryos that were hybrid for the C57BL/6J and JF1 genetic backgrounds, which differ in DNA sequence composition at pericentric heterochromatin. By carrying out immuno-FISH using a probe specific for maternal C57BL/6J major satellites, we connected differential labeling of pericentric heterochromatin to parental origin (**Supplementary Fig. 3** online). We conclude

**Figure 5** Heterochromatic but not euchromatic matPRC1 targeting is *Ezh2* independent. **(a)** Immunofluorescence analysis of germinal vesicle and M-II oocytes and PNO zygotes shows that after gamete fusion (PNO), Ezh2 preferentially accumulates at the maternal genome complement that will constitute the embryo (bottom; arrow). The top chromosome set will segregate into the second polar body. Similar data were obtained for Eed and Suz12 (**Supplementary Fig. 4**). **(b)** Protein blot analysis of Ezh2 in germinal vesicle and M-II oocytes, in early (E) and late (L) zygotes, and in 8 cell and blastocyst embryos show increasing Ezh2 protein levels from M-II stage onwards. Analyses were done on the same material as shown in **Figure 4b**. **(c)** Immunofluorescence analysis fails to show accumulation of Ezh2, Eed and H3K27me3 at DAPI-intense constitutive heterochromatin in decondensing sperm nuclei (PNO stage). **(d)** Immunofluorescence analysis of wild-type and *Ezh2*<sup>m-z-</sup> late zygotes showing absence of Ezh2 protein upon maternal *Zp3*-cre mediated and paternal *Prr1*-cre mediated deletion. **(e)** Accordingly, the establishment and maintenance of H3K27me2 and H3K27me3 is impaired at paternal and maternal genomes in mutant zygotes, whereas H3K27me1 is unaffected. **(f)** Nevertheless, Rnf2 and Cbx2 localization to paternal heterochromatin is unaltered in *Ezh2*<sup>m-z-</sup> zygotes, whereas levels of euchromatic binding correlate with the level of H3K27me3. **(g)** Immunofluorescence analysis of a wild-type 3-cell embryo showing Rnf2 bound to interphase and mitotic chromatin, whereas HP1 $\beta$  is only bound to interphase chromatin. Arrow marks Rnf2 binding to mitotic chromosomes in a banded pattern. Scale bars in **a** and **c-f**, 10  $\mu$ m. Scale bar in **g**, 20  $\mu$ m.





**Figure 6** Parent of origin-specific labeling of constitutive heterochromatin and chromosome arms by Rnf2 and H3K9me3. **(a)** Immunofluorescence analysis of pericentric heterochromatic regions of paternal prometaphase chromosomes (green circles) over successive cleavage divisions shows strong, moderate and weak Rnf2 enrichment at the first, second and third cleavage divisions, respectively. In contrast, maternal chromosome ends (red circles) are moderately enriched in H3K9me3 at all three divisions. At the subsequent fourth cleavage, all chromosome ends (white circles) are strongly enriched in H3K9me3. During the first three cleavages, the arms of paternal and maternal chromosomes are also more intensely labeled by Rnf2 and H3K9me3 in a banded pattern, respectively. This parental-specific labeling of chromosome arms is lost at the fourth cleavage division. Before syngamy at the first cleavage, paternal and maternal chromosome sets are detected as separate entities. At the second and third cleavage divisions, parental chromosome sets are often not yet intermingled and lie adjacent to each other at the metaphase plate. Parental identity was assigned on the basis of levels of Rnf2 and H3K9me3 enrichment and position on the metaphase plate. Scale bars, 10  $\mu$ m. **(b)** Graphical representation of the dynamics of asymmetric distribution of Rnf2 and H3K9me3 along paternal (P) and maternal (M) chromosomes at successive cleavage stages. Relative levels of enrichment are represented by color intensities. Stripes along chromosome arms represent banding patterns. Note that Rnf2 enrichment at paternal pericentric heterochromatin is progressively reduced at the second and third cleavages, whereas labeling along chromosome arms is not affected.

that the matPRC1-defined state at paternal constitutive heterochromatin is heritable over three consecutive cleavages and is subsequently replaced by the canonical *Suv39h*-mediated state.

### Paternally bound PRC1 is maternally provided

Next, we analyzed the onset of matPRC1 targeting. In fully grown germinal vesicle and metaphase-II oocytes, Rnf2, Bmi1, Cbx2 and Phc2 seemed not to be chromatin bound (**Fig. 3a** and **Supplementary Fig. 4** online;  $n = 25\text{--}35$  per stage) although the maternal genome was trimethylated at H3K27 (**Supplementary Fig. 1**;  $n = 10$  per stage). Shortly after gamete fusion, however, all four matPRC1 components rapidly accumulated at the maternal chromosomes that would constitute the embryo (arrow, **Fig. 3a**), whereas lower protein levels built up at chromosomes segregating into the second polar body. Concurrently, matPRC1 proteins were particularly enriched at DAPI-intense heterochromatin of paternal decondensing chromosomes (**Fig. 3b,c**).

We further asked whether matPRC1 proteins at paternal chromatin are transmitted through spermatozoa or maternally provided. We used a functional Rnf2 and yellow fluorescent protein (Rnf2-YFP) fusion knock-in allele (K.I. and H.K., unpublished data) to probe for YFP signal in heterozygous embryos. The Rnf2-YFP signal was only detectable after maternal transmission of the knock-in allele (**Fig. 4a**), indicating that most matPRC1 in zygotes is of maternal origin ( $n = 10\text{--}15$  per genetic condition). Accordingly, using protein blot analysis, we

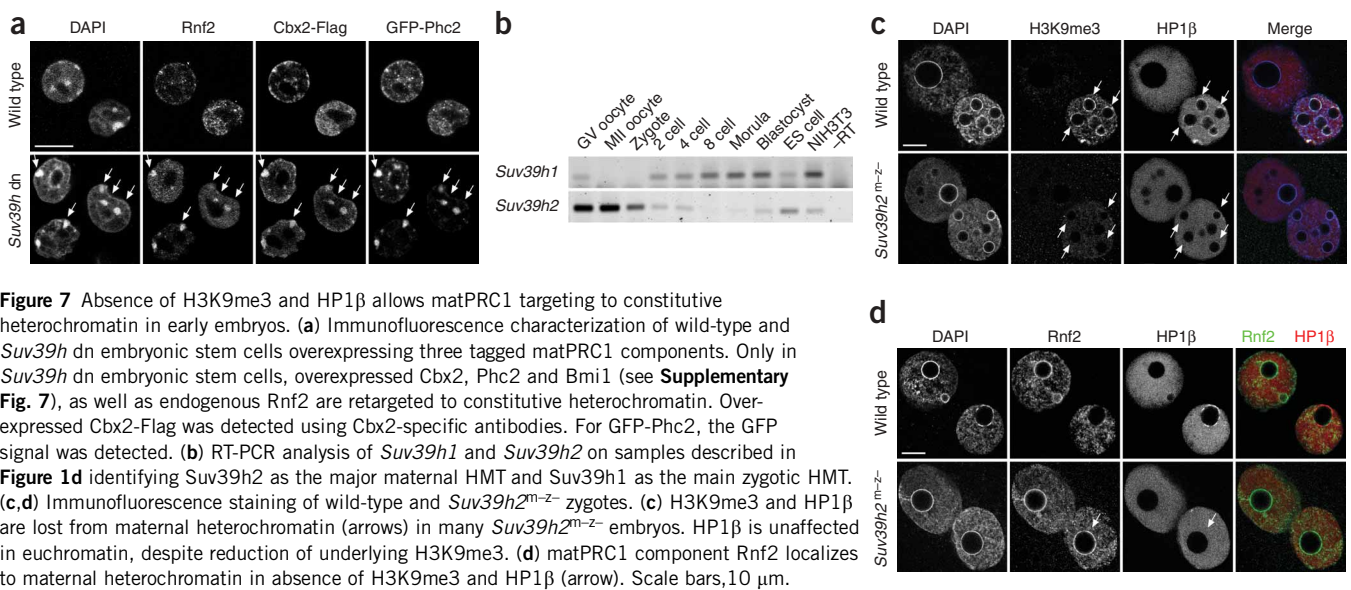
found that Rnf2 protein levels increased upon gamete fusion (**Fig. 4b**), likely as a result of translational activation of maternal message.

Finally, we determined Rnf2 protein levels in mature mouse spermatozoa isolated from the caudal epididymus. By immunofluorescence, we did not detect any PRC1 signal in sperm (**Fig. 4c**). This was not a result of antibody inaccessibility, as we were able to detect chromatin markers, known to be present in mature sperm<sup>8</sup>. By protein blot analysis, we also failed to detect Rnf2 protein in spermatozoa, although we could detect the protein in low numbers of embryonic stem cells (**Fig. 4d**). These data support our finding that PRC1 in zygotes is primarily of maternal origin, although we cannot exclude the possibility that paternal proteins may seed binding of maternally provided PRC1.

### matPRC1 targeting to heterochromatin is *Ezh2* independent

At certain genes, PRC1 binding to chromatin depends on PRC2 function<sup>27,28</sup>. Therefore, we evaluated the role of PRC2 components in oocytes and early embryos. Like matPRC1 components (**Fig. 3a**), PRC2 components become enriched after gamete fusion at maternal chromosomes that will be retained in the embryo (**Fig. 5a** and **Supplementary Fig. 4**;  $n = 25\text{--}30$  per stage). Protein blots showed increasing *Ezh2* protein levels during oocyte meiosis and early embryogenesis (**Fig. 5b**), suggesting translational activation of maternal message. Nevertheless, in contrast to matPRC1 (**Fig. 3b**), PRC2





**Figure 7** Absence of H3K9me3 and HP1 $\beta$  allows matPRC1 targeting to constitutive heterochromatin in early embryos. **(a)** Immunofluorescence characterization of wild-type and *Suv39h* dn embryonic stem cells overexpressing three tagged matPRC1 components. Only in *Suv39h* dn embryonic stem cells, overexpressed Cbx2, Phc2 and Bmi1 (see **Supplementary Fig. 7**), as well as endogenous Rnf2 are retargeted to constitutive heterochromatin. Overexpressed Cbx2-Flag was detected using Cbx2-specific antibodies. For GFP-Phc2, the GFP signal was detected. **(b)** RT-PCR analysis of *Suv39h1* and *Suv39h2* on samples described in **Figure 1d** identifying *Suv39h2* as the major maternal HMT and *Suv39h1* as the main zygotic HMT. **(c,d)** Immunofluorescence staining of wild-type and *Suv39h2<sup>m-m</sup>* zygotes. **(c)** H3K9me3 and HP1 $\beta$  are lost from maternal heterochromatin (arrows) in many *Suv39h2<sup>m-m</sup>* embryos. HP1 $\beta$  is unaffected in euchromatin, despite reduction of underlying H3K9me3. **(d)** matPRC1 component Rnf2 localizes to maternal heterochromatin in absence of H3K9me3 and HP1 $\beta$  (arrow). Scale bars, 10  $\mu$ m.

members ( $n = 27$ ) or H3K27me3 ( $n = 22$ ) did not label paternal chromatin shortly after gamete fusion (**Fig. 5c** and **Supplementary Fig. 1**), suggesting that initial targeting of matPRC1 to paternal heterochromatin is PRC2 independent.

To unequivocally determine the role of PRC2 in matPRC1 targeting to paternal heterochromatin, we used a conditional deficiency allele for *Ezh2* in which exons 16 and 17, containing parts of the catalytic SET domain, are flanked by loxP sites (**Supplementary Fig. 5** online; X.M. and S.H.O., unpublished data). Loss of *Ezh2* function using this allele leads to gastrulation defects and post-implantation embryonic lethality (**Supplementary Fig. 5**), as described for a conventional deletion allele<sup>33</sup>. We subsequently bred in *Zp3-cre* and *Prm1-cre* transgenic lines that mediate gene deletion in primary oocytes and elongating spermatids, respectively, in order to generate embryos that were deficient for maternal and zygotic *Ezh2* expression (*Ezh2<sup>m-m</sup>*) (**Supplementary Fig. 5**), and that lacked detectable *Ezh2* protein (**Fig. 5d**;  $n = 35$ ). We did not detect *de novo* H3K27me2/3 in late-stage paternal pronuclei (**Fig. 5e**;  $n = 75$ ). In maternal pre-replication pronuclei, H3K27me2/3 levels were similar in wild type and *Ezh2<sup>m-m</sup>* zygotes, indicating that after *Zp3-cre*-mediated *Ezh2* depletion, previously established H3K27me2/3 remains stably present in maturing oocytes. After replication, however, H3K27me2/3 levels were strongly reduced in maternal pronuclei (**Fig. 5e**). H3K27me3 was undetectable from late 2-cell to blastocyst stage (data not shown), indicating that *Ezh2* is required for the establishment and maintenance of global H3K27me3 at both parental genomes. H3K27me1 is, however, independent of *Ezh2* (**Fig. 5e**;  $n = 10$ ).

Through double-labeling experiments, we found that in zygotes targeting of all four matPRC1 components to euchromatin correlated with levels of H3K27me3 labeling. MatPRC1 staining was undetectable on paternal euchromatin, whereas it was severely reduced on maternal euchromatin (**Fig. 5f** and **Supplementary Fig. 6** online; wild type,  $n = 43$ ; *Ezh2<sup>m-m</sup>*,  $n = 42$ ). In contrast, matPRC1 remained enriched at paternal constitutive heterochromatin in *Ezh2<sup>m-m</sup>* embryos up to the 8-cell stage (**Fig. 5f** and **Supplementary Fig. 6**; data not shown), indicating PRC2-independent targeting of matPRC1 to paternal constitutive heterochromatin.

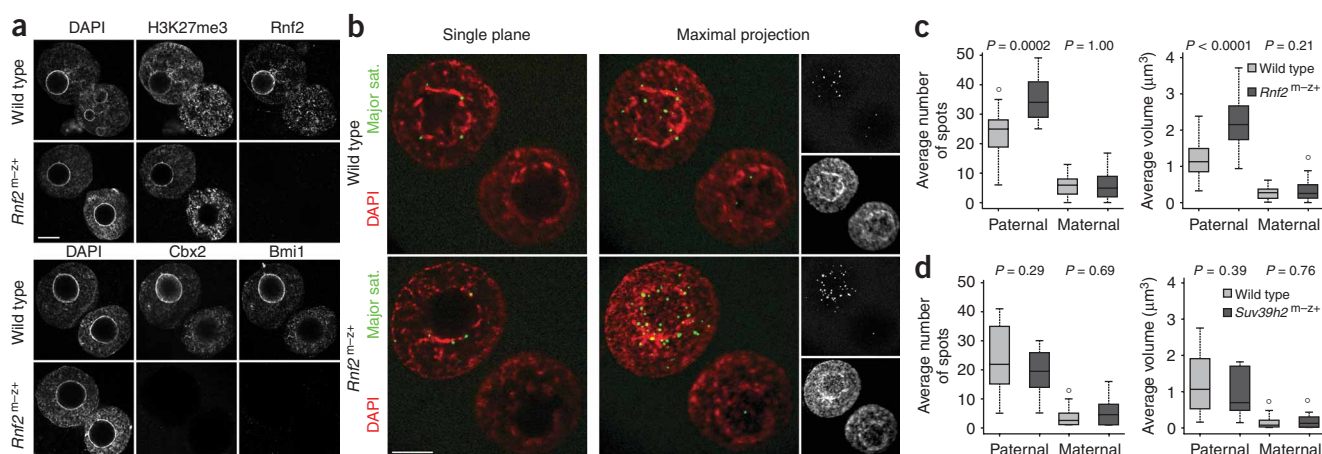
### Parental-specific marking of cleavage chromosomes

We subsequently studied the mechanism of mitotic transmission. Unlike HP1 $\beta$ , which dissociated from mitotic chromosomes (**Fig. 5g**)<sup>34</sup>, Rnf2 remained bound along metaphase chromosomes in a banded pattern ( $n = 10$ ). For higher-resolution analyses, we studied prometaphase chromosomes of embryos at successive cleavage stages (**Fig. 6a**). At the first cleavage division, Rnf2 was strongly enriched at the proximal DAPI-intense heterochromatic chromosome ends of paternal origin but not at maternal chromosome ends, which were enriched in H3K9me3. At the second and third divisions, Rnf2 labeling of paternal pericentric heterochromatin was moderately and strongly reduced, respectively (relative to euchromatin; see below). In contrast, pericentric heterochromatin of maternal chromosomes was moderately labeled by H3K9me3 during the first three cleavages. At the fourth division, all chromosomes ends were highly enriched in H3K9me3, indicating acquisition and consolidation of the canonical *Suv39h*-mediated identity at paternal and maternal heterochromatin, respectively, during the 8-cell stage. In summary, these analyses demonstrate mitotically stable but transient marking of paternal heterochromatin by matPRC1 components (**Fig. 6b**).

In addition to asymmetric heterochromatic labeling, we also observed differential marking along chromosome arms over the first three cleavage divisions (**Fig. 6**). Rnf2 showed clear banding patterns along paternal chromosomes that were much less pronounced along maternal chromosomes. Conversely, the maternal complement was more strongly marked by H3K9me3 banding patterns. This dual asymmetry was lost at the fourth cleavage.

### *Suv39h*-pathway impairs PRC1 targeting to heterochromatin

We next addressed the mechanism of parental-specific matPRC1 targeting. We reasoned that matPRC1 targeting to paternal constitutive heterochromatin could be due to a chromatin configuration inherited from sperm<sup>4,8</sup> and/or the absence of a functional *Suv39h* pathway. Intrigued by the mutually exclusive labeling by matPRC1 versus H3K9me3 (**Figs. 2** and **6**), we first analyzed *Suv39h1* and *Suv39h2* double null (*Suv39h* dn) embryonic stem cells that harbor H3K27me3 at constitutive heterochromatin in the absence of H3K9me3 (ref. 31; **Supplementary Fig. 7** online). Nonetheless, we



**Figure 8** MatPRC1 is required for transcriptional repression of major satellites on the paternal genome in early embryos.

(a) Immunofluorescence characterization of wild-type and *Rnf2*<sup>m-z+</sup> zygotes showing global loss of Rnf2, Cbx2 and Bmi1 upon *Zp3*-cre-mediated deletion of maternal *Rnf2*. The distribution of H3K27me3 is unchanged. (b) RNA-FISH analysis of nascent major satellite transcripts in wild-type and *Rnf2*<sup>m-z+</sup> zygotes showing close proximity of satellite transcripts to DAPI-intense heterochromatin (single planes). Maximal projections visualize total number of transcription sites per pronucleus. Scale bars, 10  $\mu$ m. (c) Average number and volume of major satellite transcription sites for wild-type ( $n = 20$ ) and *Rnf2*<sup>m-z+</sup> ( $n = 25$ ) pronuclei represented as boxplots (see Methods). In *Rnf2*<sup>m-z+</sup> zygotes, the number and volume of transcription sites are significantly increased in paternal but not maternal pronuclei compared to wild-type embryos.  $P$  values were calculated using a paired  $t$ -test. (d) Average number and volume of major satellite transcription sites for wild-type ( $n = 28$ ) and *Suv39h2*<sup>m-z+</sup> ( $n = 10$ ) pronuclei represented as boxplots. Major satellite levels remain unchanged in maternal pronuclei of *Suv39h2*<sup>m-z+</sup> zygotes. (e) Quantitative RT-PCR analysis of major satellite transcription in late 2-cell embryos showing a threefold increased transcript level in *Rnf2*<sup>m-z+</sup> versus wild-type embryos, whereas levels remain unaltered in *Suv39h2* maternal mutants compared to wild-type controls. Error bars represent s.d. of three PCR amplifications for each sample. Similar results were obtained in two independent experiments. (f) Model for the establishment of distinct chromatin states at constitutive heterochromatin in the early embryo by Suv39h and PRC1 pathways in a parent of origin-dependent manner.

failed to detect heterochromatic enrichment for Rnf2, Phc1 or Mel18 (Supplementary Fig. 7), expressed in embryonic stem cells (Fig. 1d)<sup>35</sup>. Likewise, we failed to detect PRC1 members at constitutive heterochromatin in quiescent B cells (data not shown), devoid of *Suv39h*-mediated repressive modifications<sup>36</sup>. To test whether lack of detection was a result of low expression levels of one or more matPRC1 components, we reconstituted matPRC1 by co-overexpressing Cbx2-Flag, GFP-Phc2 and 3Flag-Bmi1 in wild-type and *Suv39h* dn embryonic stem cells and analyzed their nuclear localization (Fig. 7a and Supplementary Fig. 7). All overexpressed PRC1 components as well as endogenously expressed Rnf2 were targeted to constitutive heterochromatin in *Suv39h* dn embryonic stem cells but not in wild-type cells. These data show that PRC1 enrichment at heterochromatin is not unique to early embryos, excluding the necessity of a 'paternal imprint' for targeting, and that the *Suv39h* pathway blocks matPRC1 loading to heterochromatin.

We subsequently investigated the interplay between *Suv39h* and matPRC1 *in vivo*. As most *Suv39h* dn mice die during late gestation or are postnatally growth retarded<sup>13</sup>, we studied whether maternal deficiency for Suv39h2, the enzyme predominantly expressed in oocytes (Fig. 7b), creates a hypomorphic condition in zygotic embryos. Indeed, in 71% ( $n = 34$ ) of maternally (and zygotically) deficient *Suv39h2* zygotes, maternal heterochromatic labeling of

H3K9me3 and HP1 $\beta$  was lost, although euchromatic HP1 $\beta$  labeling was unaffected (Fig. 7c). In such embryos, we observed strong enrichment for matPRC1 components at maternal heterochromatin and thus alleviation of parent of origin-specific labeling (Fig. 7d and Supplementary Fig. 7). Through analysis of wild-type and *Suv39h2*<sup>m-z+</sup> embryos, we found a highly significant inverse correlation between levels of H3K9me3 and HP1 $\beta$  versus those of matPRC1 components at constitutive heterochromatin ( $P = 8.47 \times 10^{-10}$ ;  $n = 103$ ; Fisher's exact test for dependency). These data unequivocally demonstrate that paternal heterochromatic enrichment of matPRC1 is due to the lack of local Suv39h activity instead of germline transmission of a paternal-specific factor or chromatin state.

#### MatPRC1 represses major satellites in early embryos

Finally, we generated embryos lacking matPRC1 to determine the functional significance of matPRC1 for transcriptional repression. We used a conditional deficiency allele for *Rnf2* in which major parts of the Ring finger domain were flanked by loxP sites (E.B. and M.v.L., unpublished data). The same targeting strategy was used to generate a conventional deletion allele that causes gastrulation defects in *Rnf2*<sup>-/-</sup> mice<sup>37</sup>. We crossed conditional *Rnf2* mice with a *Zp3*-cre transgenic line to generate embryos that were maternally deficient for *Rnf2* and that lacked detectable Rnf2 protein in zygotes (Fig. 8a). H3K27me3

distribution was unchanged in *Rnf2<sup>m-z+</sup>* zygotes (Fig. 8a). In contrast, Cbx2 and Bmi1 were undetectable (Fig. 8a), indicating that Rnf2 is required for the stability of matPRC1, as observed in *Rnf2<sup>-/-</sup>* embryonic stem cells<sup>35</sup>. To measure nascent transcript levels of major satellites<sup>38</sup>, we carried out RNA-FISH in zygotes (Fig. 8b). We observed foci adjacent to DAPI-intense heterochromatin. As detection of foci was sensitive to treatment with the transcription elongation inhibitor DRB (data not shown), our results show *de novo* transcription in one-cell embryos, before genome-wide activation at the 2-cell stage. In wild-type zygotes, major satellite transcript levels were fivefold higher in paternal versus maternal pronuclei (Fig. 8c), consistent with the higher transcriptional potential of the paternal pronucleus<sup>39</sup>. Major satellite transcription was unchanged in maternal pronuclei of wild-type versus *Rnf2<sup>m-z+</sup>* zygotes, correlating with the local presence of *Suv39h*-mediated repressive chromatin marks. In paternal pronuclei of *Rnf2<sup>m-z+</sup>* zygotes, however, the number of major satellite transcription sites and the total level of nascent transcripts were significantly increased compared to wild-type paternal pronuclei (Fig. 8c). Consistently, in *Rnf2<sup>m-z+</sup>* 2-cell embryos, major satellites were threefold upregulated (Fig. 8e). Thus, these data unambiguously demonstrate that matPRC1 is required for transcriptional repression of underlying major satellites.

As reported before<sup>18</sup>, we measured a fourfold upregulation of major satellite expression in *Suv39h* dn embryonic stem cells over wild-type controls (Supplementary Fig. 7). Through RNA-FISH analysis of *Suv39h2<sup>m-z+</sup>* zygotes, however, we found no significant increase in major satellite transcript levels in maternal pronuclei despite the lack of H3K9me3 and HP1 $\beta$  at pericentric heterochromatin (Fig. 8d). Similarly, we observed no change in transcript levels in *Suv39h2<sup>m-z+</sup>* 2-cell embryos (Fig. 8e). These results strongly suggest that matPRC1 recruited to maternal constitutive heterochromatin in *Suv39h2* mutant early embryos functions as a repressive back-up mechanism.

## DISCUSSION

Parent of origin-dependent differential marking by active and repressive epigenetic modifications is reminiscent of allelic specification underlying imprinted X inactivation and genomic imprinting. Here we show that two repressive pathways specify the 'allelic' states of maternal and paternal constitutive heterochromatin, suggesting functional compensation. Maternally, pericentric heterochromatin is marked by H3K9me3 and H4K20me3, two modifications established by the *Suv39h* and *Suv4-20h* HMTs in oocytes, and by HP1 $\beta$  loaded onto chromatin upon gamete fusion. In contrast, maternally provided PRC1 complexes that are required for transcriptional repression of underlying major satellites associate with paternal heterochromatin during sperm nuclear decondensation. Deficiency for *Suv39h2* results in targeting of matPRC1 to maternal heterochromatin only when it is devoid of detectable H3K9me3 and/or HP1 $\beta$ . Absence of paternal germ-line derived H3K9me3 therefore serves as the primary germ-line imprint for targeting matPRC1 to major satellites of the paternal genome (Fig. 8f). Consistently, major satellite expression was not upregulated in *Suv39h2<sup>m-z+</sup>* early embryos but was upregulated in *Suv39h* dn embryonic stem cells lacking heterochromatic PRC1. In absence of H3K9me3, PRC1 thus constitutes the default repressive pathway for constitutive heterochromatin formation in early embryos. The identity of maternal heterochromatin is inherited from the oocyte, assigning a crucial role to *Suv39h*-mediated H3K9me3 in *trans*-generational inheritance of maternal epigenetic states in mouse and likely other mammals.

The fact that a matPRC1-like complex, when overexpressed, is targeted to heterochromatin of *Suv39h* dn embryonic stem cells

underscores the absence of a 'selective paternal imprint' needed for heterochromatic association. This result confirms the dominant role of the *Suv39h* pathway in blocking matPRC1 binding to heterochromatin and emphasizes the dosage sensitivity of matPRC1 association to chromatin. Notably, absence of H3K9me3 correlates with H3K27me3 enrichment at pericentric heterochromatin in wild-type early embryos and *Suv39h* dn embryonic stem cells, indicating that *Suv39h* is also dominant over PRC2.

Given the dominance of *Suv39h* in defining heterochromatin, it is of note that the paternal genome is initially refractory toward *de novo* H3K9 trimethylation. It is possible that the maternal *Suv39h2* protein, if present, is enzymatically inhibited, as reported for a H3K9 dimethyltransferase<sup>5</sup>, or its activity is locally counteracted by a histone demethylase. In analogy to heterochromatin maturation in fly embryos<sup>40</sup> or *Schizosaccharomyces pombe*<sup>17</sup>, conditions that enable 'de novo' targeting of *Suv39h* enzymes may first need to be established in mammalian embryos.

In mature quiescent oocytes, PcG proteins are not detectable at chromatin, although they are expressed. Only after germ cell fusion do they associate to maternal anaphase II chromosomes. These dynamics are reminiscent of those of other transcription and chromatin factors in quiescent oocytes<sup>41</sup> and M-II oocytes<sup>42</sup>, and may facilitate the transition of the maternal-to-zygotic transcription program. At the paternal genome, matPRC1 is loaded onto heterochromatin during sperm decondensation, before repressive histone methylation marks are acquired. Consistently, Ezh2-mediated H3K27me3 is not required for matPRC1 targeting to paternal heterochromatin. This shows that despite the binding affinity of Cbx2 toward H3K27me3<sup>27</sup>, the interaction between Cbx2 and H3K27me3 is not the sole mechanism for chromatin targeting of matPRC1. In analogy to X inactivation<sup>23,43</sup>, genomic imprinting<sup>44</sup> and constitutive heterochromatin formation in somatic cells<sup>45</sup>, a noncoding RNA may be required for matPRC1 heterochromatic localization. Alternatively, matPRC1 recruitment may depend on the repetitive nature and/or AT-richness of the underlying sequences.

The kinetics of H3K9me3 and matPRC1 chromatin association suggest two phases of epigenetic programming during preimplantation development. The first phase, lasting until the 8-cell stage, is characterized by parental asymmetry in histone modifications<sup>7,9</sup>, DNA methylation<sup>3</sup> and PcG proteins. Besides pericentric regions, paternal and maternal mitotic chromosome arms are also differentially labeled by matPRC1 and H3K9me3, respectively. Given the requirement of matPRC1 for transcriptional repression of major satellites, matPRC1 likely represses other sequences in the paternal genome, whereas H3K9me3 could contribute to repression in the maternal genome. The *Suv39h2* loss-of-function study demonstrates that epigenetic programming is flexible in the early embryo, being adaptable to variable chromatin states established during preceding oogenesis (Fig. 7) and likely spermatogenesis<sup>46</sup>. This plasticity may facilitate the maternal-to-zygotic transition in gene expression at both genomes as well as the reacquisition of totipotency.

The resolution of parental epigenetic asymmetry during the 8-cell stage marks the onset of the second phase of epigenetic programming. Besides a possible gain in *Suv39h1* function, the change in paternal repressive identity is probably supported by reduced expression of certain matPRC1 components. Accordingly, we did not detect Phc2, and we found that Cbx2 was strongly downregulated in 16-cell and later stage embryos (Supplementary Fig. 2). Of note, the resolution of epigenetic asymmetry coincides with a number of key developmental changes characteristic of the 8-cell stage that reflect preparation for the successive cell determination decisions. For example, blastomeres start



to compact and polarize, leading to the formation of precursors of the embryonic and extra-embryonic cell lineages<sup>47</sup>. Concurrently, active histone modifications start to be removed from the paternal X chromosome undergoing imprinted inactivation<sup>48</sup>. We observed Rnf2 accumulation at the presumptive Xi domain from the 16-cell stage onwards, thus directly after Rnf2 displacement from constitutive heterochromatin. On the basis of the overall kinetics, it is likely that resolution of global parental epigenetic asymmetry is needed to enable subsequent specification of cellular lineages.

In conclusion, our analyses establish the concept of epigenetic asymmetry between parental genomes during preimplantation development as a dynamic response to differential maturation of chromatin states during oogenesis versus spermatogenesis. We anticipate that PRC1 has a role in epigenetic gene regulation during preimplantation development<sup>49</sup>.

## METHODS

**Mice and cell lines.** To produce maternally deficient *Ezh2* oocytes, we generated *Ezh2*<sup>F/F</sup> mice that carried the *Zp3-cre* (zona pellucida3-cre) recombinase transgene, mediating efficient deletion in dictyate-stage growing primary oocytes. In addition, we used a *Prm1-cre* (protamine1-cre) line, which expresses Cre recombinase during late spermatogenesis, to generate sperm carrying the mutant *Ezh2* allele. Maternal and paternal *Ezh2*-deficient embryos were obtained from matings between *Ezh2*<sup>F/F</sup>; *Zp3-cre*/+ females and *Ezh2*<sup>F/F</sup>; *Prm1-cre*/+ males. Control embryos designated as wild-type were *Ezh2*<sup>F/F</sup>. Maternally deficient *Rnf2* zygotes were generated by crossing *Rnf2*<sup>F/F</sup>; *Zp3-cre*/+ females with *Rnf2*<sup>F/F</sup> males. Wild-type control embryos were generated by *Rnf2*<sup>F/F</sup>; *Rnf2*<sup>F/+</sup> or *Rnf2*<sup>F/-</sup> females and *Rnf2*<sup>F/F</sup> males. The generation of *Suv39h2* mutant mice has been described previously<sup>13</sup>. Wild-type control embryos were generated by *Suv39h2*<sup>+/-</sup> females and *Suv39h2*<sup>+/+</sup> males. Embryos polymorphic for pericentromeric heterochromatin were obtained from matings of C57BL6 females with JF1 males. For further information on mouse lines, embryo isolation and description of cell lines, see **Supplementary Methods** online. Housing and handling of mice conformed to the Swiss Animal Protection Ordinance, chapter 1.

**Immunofluorescence and FISH.** Immunofluorescence<sup>7</sup> and FISH<sup>48</sup> of embryos were carried out as previously described, with some modifications described in the **Supplementary Methods**.

**Microscopy and image analysis.** Immunofluorescence stainings of embryonic stem cells and embryos were analyzed using a laser scanning confocal microscope LSM510 META (Zeiss) and LSM510 software. For embryos, we recorded a z series of either 0.3 μm or 0.5-μm slices or scanned one confocal slice through the maximal radius of each (pro)nucleus; we then exported the slices as 8-bit TIFF files, later projected using Photoshop 9.0. For numerical evaluation, all images of embryos taken for a given modification were analyzed individually and scored as follows: (–) no staining; (=) equal staining at hetero- and euchromatin; (>) enhanced and (>>) strongly enhanced staining at heterochromatin versus euchromatin. Cleavage-stage embryos were scanned in three dimensions, recording a z series of 0.5-μm slices. Only long prometaphase chromosomes were evaluated for scoring heterochromatin labeling and banding patterns along chromosome arms. Embryos processed by Immunofluorescence were analyzed with a spinning disk confocal microscope (Yokogawa CSU-22). For RNA-FISH analysis of major satellite expression in zygotes, a z series of 0.3 μm was recorded for each zygote using the confocal microscope LSM510 META (Zeiss). The number of major satellite transcription sites and the volume of each spot were calculated for paternal and maternal pronuclei using 3D Spotfinder software in Imaris 5.9.0. alpha. Sites of major satellite transcription with a minimum volume of 0.01 μm<sup>3</sup> were identified, and the volume of each spot was summed up to represent the total volume of major satellite transcription per pronucleus.

**Statistical analysis.** We investigated a possible functional association between PRC1 and HP1β using wild-type (*n* = 69) and *Suv39h2* maternally and zygotically deficient embryos (*n* = 34). For each embryo, PRC1 and HP1

complexes were stained and classified according to staining intensity into four categories. We summarized the data in a four-by-four contingency table and tested the significance of the association between PRC1 and HP1 using Fisher's exact test.

Results of major satellite transcription levels measured by RNA-FISH are presented as boxplots with the thick line representing the median, the top and bottom boundaries of the boxes corresponding to the 75 and 25 percentiles, respectively, and the top and bottom whiskers presenting maximum and minimum values, respectively. Values that deviate more than the 1.5-fold distance between quartiles from the median are drawn as black circles and represent outliers. *P* values were calculated using a paired *t*-test.

**RT-PCR and protein blot.** Gene expression analysis of oocytes and embryos was carried out as described in the **Supplementary Methods**. Protein blot analysis was carried out using standard procedures. Preparation of oocytes, embryos and sperm samples for protein blot analysis is described in the **Supplementary Methods**.

*Note: Supplementary information is available on the Nature Genetics website.*

## ACKNOWLEDGMENTS

We thank M. Vidal (Centro de Investigaciones Biológicas, Spain) and T. Jenuwein (Research Institute of Molecular Pathology, Austria) for providing antisera. Moreover, we are grateful to T. Jenuwein and B. Knowles (The Jackson Laboratory, USA) for providing *Suv39h2* deficient and *Zp3-cre* transgenic mice, respectively. We acknowledge excellent assistance by Friedrich Miescher Institute (FMI) colleagues P. Schwarb and J. Rietdorf (microscopy and imaging facility), B. Heller-Stilb and J.-F. Spetz (animal facility), S. Bichet (histology) and M. Stadler (bioinformatics). We thank members of the Peters laboratory for fruitful discussions and P. de Boer, D. Schübeler, S. Gasser and P. Hublitz for valuable comments on the manuscript. Research at the Friedrich Miescher Institute is supported by the Novartis Research Foundation. M.P. acknowledges the Boehringer Ingelheim Fonds for her PhD fellowship. Research in the Peters laboratory is supported by the EU NoE network 'The Epigenome' (LSHG-CT-2004-503433).

## AUTHOR CONTRIBUTIONS

M.P. and A.H.F.M.P. conceived and designed the experiments. M.P., R.T., U.B. and C.K. performed the experiments. M.P., R.T., U.B. and A.H.F.M.P. analyzed the data. A.P.O. provided antibodies. E.B. and M.v.L. provided conditionally deficient *Rnf2* mice. X.M. and S.H.O. provided conditionally deficient *Ezh2* mice. K.I. and H.K. provided *Rnf2*-YFP knock-in mice. M.P. and A.H.F.M.P. wrote the manuscript.

Published online at <http://www.nature.com/naturegenetics>

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions>

1. Surani, M.A., Hayashi, K. & Hajkova, P. Genetic and epigenetic regulators of pluripotency. *Cell* **128**, 747–762 (2007).
2. Arney, K.L., Bao, S., Bannister, A.J., Kouzarides, T. & Surani, M.A. Histone methylation defines epigenetic asymmetry in the mouse zygote. *Int. J. Dev. Biol.* **46**, 317–320 (2002).
3. Dean, W. *et al.* Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos. *Proc. Natl. Acad. Sci. USA* **98**, 13734–13738 (2001).
4. Govin, J. *et al.* Pericentric heterochromatin reprogramming by new histone variants during mouse spermiogenesis. *J. Cell Biol.* **176**, 283–294 (2007).
5. Liu, H., Kim, J.M. & Aoki, F. Regulation of histone H3 lysine 9 methylation in oocytes and early pre-implantation embryos. *Development* **131**, 2269–2280 (2004).
6. Martin, C. *et al.* Genome restructuring in mouse embryos during reprogramming and early development. *Dev. Biol.* **292**, 317–332 (2006).
7. Santos, F., Peters, A.H., Otte, A.P., Reik, W. & Dean, W. Dynamic chromatin modifications characterise the first cell cycle in mouse embryos. *Dev. Biol.* **280**, 225–236 (2005).
8. van der Heijden, G.W. *et al.* Transmission of modified nucleosomes from the mouse male germline to the zygote and subsequent remodeling of paternal chromatin. *Dev. Biol.* **298**, 458–469 (2006).
9. van der Heijden, G.W. *et al.* Asymmetry in histone H3 variants and lysine methylation between paternal and maternal chromatin of the early mouse zygote. *Mech. Dev.* **122**, 1008–1022 (2005).
10. Merico, V. *et al.* Epigenomic differentiation in mouse preimplantation nuclei of biparental, parthenote and cloned embryos. *Chromosome Res.* **15**, 341–360 (2007).

11. Kishigami, S. *et al.* Epigenetic abnormalities of the mouse paternal zygotic genome associated with microinsemination of round spermatids. *Dev. Biol.* **289**, 195–205 (2006).
12. Ekwall, K. *et al.* Mutations in the fission yeast silencing factors *clr4+* and *rik1+* disrupt the localisation of the chromo domain protein Swi6p and impair centromere function. *J. Cell Sci.* **109**, 2637–2648 (1996).
13. Peters, A.H. *et al.* Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107**, 323–337 (2001).
14. Peters, A.H. & Schubeler, D. Methylation of histones: playing memory with DNA. *Curr. Opin. Cell Biol.* **17**, 230–238 (2005).
15. Wustmann, G., Szidonya, J., Taubert, H. & Reuter, G. The genetics of position-effect variegation modifying loci in *Drosophila melanogaster*. *Mol. Gen. Genet.* **217**, 520–527 (1989).
16. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. & Jenuwein, T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116–120 (2001).
17. Grewal, S.I. & Jia, S. Heterochromatin revisited. *Nat. Rev. Genet.* **8**, 35–46 (2007).
18. Lehnertz, B. *et al.* Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr. Biol.* **13**, 1192–1200 (2003).
19. Schotta, G. *et al.* A silencing pathway to induce H3–K9 and H4–K20 trimethylation at constitutive heterochromatin. *Genes Dev.* **18**, 1251–1262 (2004).
20. Ringrose, L. & Paro, R. Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu. Rev. Genet.* **38**, 413–443 (2004).
21. Schwartz, Y.B. & Pirrotta, V. Polycomb silencing mechanisms and the management of genomic programmes. *Nat. Rev. Genet.* **8**, 9–22 (2007).
22. Levine, S.S., King, I.F. & Kingston, R.E. Division of labor in polycomb group repression. *Trends Biochem. Sci.* **29**, 478–485 (2004).
23. de Napoles, M. *et al.* Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev. Cell* **7**, 663–676 (2004).
24. Wang, H. *et al.* Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431**, 873–878 (2004).
25. Sparmann, A. & van Lohuizen, M. Polycomb silencers control cell fate, development and cancer. *Nat. Rev. Cancer* **6**, 846–856 (2006).
26. Boyer, L.A. *et al.* Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441**, 349–353 (2006).
27. Bernstein, E. *et al.* Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. *Mol. Cell Biol.* **26**, 2560–2569 (2006).
28. Fujimura, Y. *et al.* Distinct roles of Polycomb group gene products in transcriptionally repressed and active domains of Hoxb8. *Development* **133**, 2371–2381 (2006).
29. Guenatri, M., Bailly, D., Maison, C. & Almouzni, G. Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. *J. Cell Biol.* **166**, 493–505 (2004).
30. Probst, A.V., Santos, F., Reik, W., Almouzni, G. & Dean, W. Structural differences in centromeric heterochromatin are spatially reconciled on fertilisation in the mouse zygote. *Chromosoma* **116**, 403–415 (2007).
31. Peters, A.H. *et al.* Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Mol. Cell* **12**, 1577–1589 (2003).
32. Mayer, W., Smith, A., Fundele, R. & Haaf, T. Spatial separation of parental genomes in preimplantation mouse embryos. *J. Cell Biol.* **148**, 629–634 (2000).
33. O'Carroll, D. *et al.* The polycomb-group gene *Ezh2* is required for early mouse development. *Mol. Cell Biol.* **21**, 4330–4336 (2001).
34. Fischle, W. *et al.* Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* **438**, 1116–1122 (2005).
35. Leeb, M. & Wutz, A. Ring1B is crucial for the regulation of developmental control genes and PRC1 proteins but not X inactivation in embryonic cells. *J. Cell Biol.* **178**, 219–229 (2007).
36. Baxter, J. *et al.* Histone hypomethylation is an indicator of epigenetic plasticity in quiescent lymphocytes. *EMBO J.* **23**, 4462–4472 (2004).
37. Voncken, J.W. *et al.* Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. *Proc. Natl. Acad. Sci. USA* **100**, 2468–2473 (2003).
38. Lu, J. & Gilbert, D.M. Proliferation-dependent and cell cycle regulated transcription of mouse pericentric heterochromatin. *J. Cell Biol.* **179**, 411–421 (2007).
39. Aoki, F., Worrall, D.M. & Schultz, R.M. Regulation of transcriptional activity during the first and second cell cycles in the preimplantation mouse embryo. *Dev. Biol.* **181**, 296–307 (1997).
40. Rudolph, T. *et al.* Heterochromatin formation in *Drosophila* is initiated through active removal of H3K4 methylation by the LSD1 homolog SU(VAR)3–3. *Mol. Cell* **26**, 103–115 (2007).
41. Sun, F. *et al.* Nuclear reprogramming: the zygotic transcription program is established through an “erase-and-rebuild” strategy. *Cell Res.* **17**, 117–134 (2007).
42. Yoshida, N., Brahmajosyula, M., Shoji, S., Amanai, M. & Perry, A.C. Epigenetic discrimination by mouse metaphase II oocytes mediates asymmetric chromatin remodeling independently of meiotic exit. *Dev. Biol.* **301**, 464–477 (2007).
43. Schoeftner, S. *et al.* Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. *EMBO J.* **25**, 3110–3122 (2006).
44. Umlauf, D. *et al.* Imprinting along the Kcnq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. *Nat. Genet.* **36**, 1296–1300 (2004).
45. Maison, C. *et al.* Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. *Nat. Genet.* **30**, 329–334 (2002).
46. Chong, S. *et al.* Modifiers of epigenetic reprogramming show paternal effects in the mouse. *Nat. Genet.* **39**, 614–622 (2007).
47. Johnson, M.H. Manipulation of early mammalian development: what does it tell us about cell lineages? *Dev Biol (N Y 1985)* **4**, 279–96 (1986).
48. Okamoto, I., Otte, A.P., Allis, C.D., Reinberg, D. & Heard, E. Epigenetic dynamics of imprinted X inactivation during early mouse development. *Science* **303**, 644–649 (2004).
49. Blewitt, M.E., Vickaryous, N.K., Paldi, A., Koseki, H. & Whitelaw, E. Dynamic reprogramming of DNA methylation at an epigenetically sensitive allele in mice. *PLoS Genet.* **2**, e49 (2006).

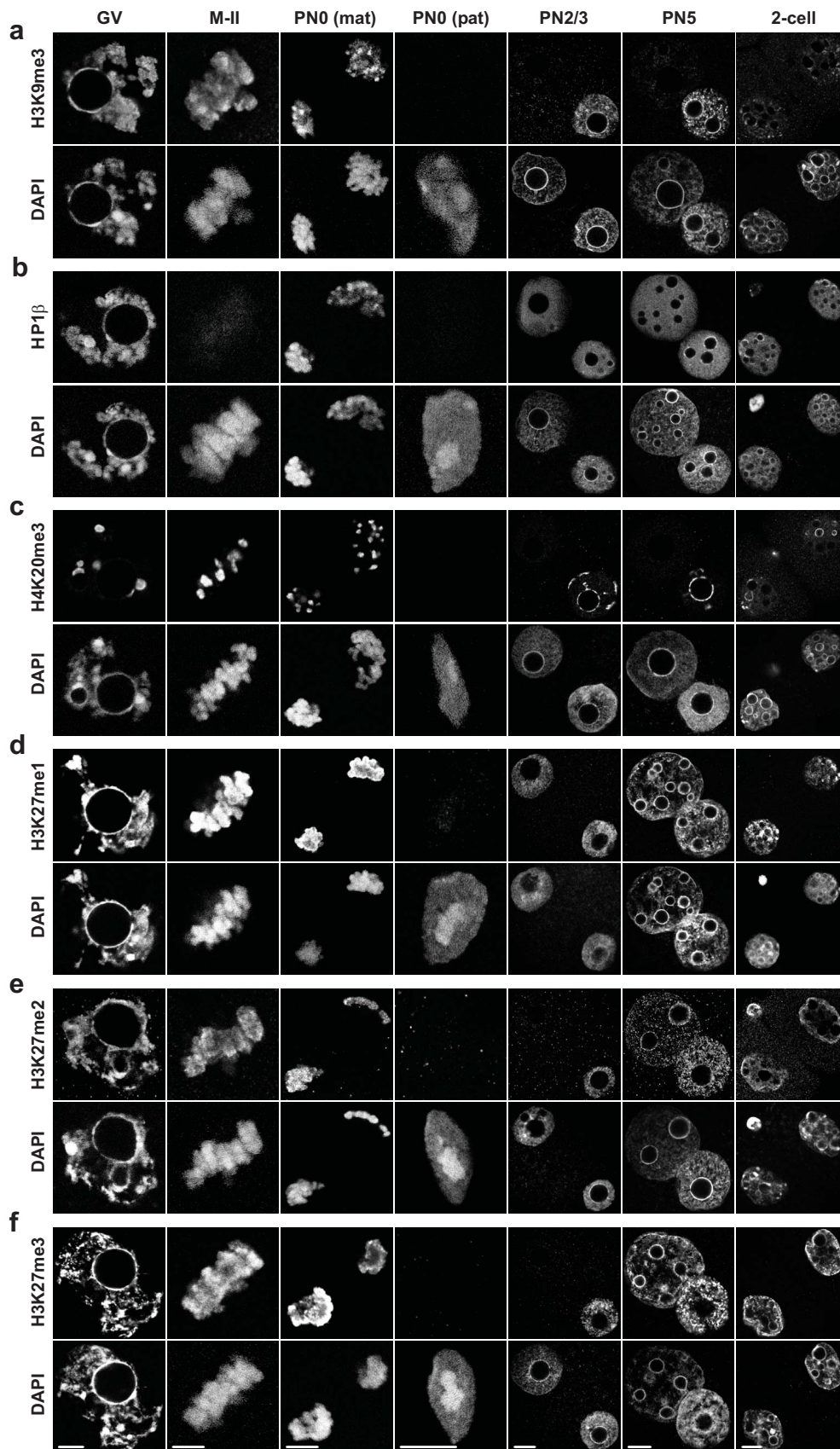
Supplementary information to the manuscript entitled:

## **PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos**

Mareike Puschendorf, Rémi Terranova, Erwin Boutsma, Xiaohong Mao, Kyo-ichi Isono, Urszula Brykczynska, Carolin Kolb, Arie P. Otte, Haruhiko Koseki, Stuart H. Orkin, Maarten van Lohuizen and Antoine H.F.M. Peters

Supplementary Figure 1	Immunofluorescence analyses of various histone methylation marks in oocytes and early pre-implantation embryos
Supplementary Figure 2	Differential heterochromatic states are maintained up to the 8-cell stage in a parental-origin-dependent manner
Supplementary Figure 3	HP1 $\beta$ is specifically enriched at maternal heterochromatin in 4-cell embryos
Supplementary Figure 4	Immunofluorescence analyses of PRC1 and PRC2 proteins in oocytes and zygotes
Supplementary Figure 5	Histological analysis of <i>Ezh2</i> deficient mice and breeding strategies to generate <i>Ezh2</i> maternally and zygotically deficient ( <i>Ezh2</i> <sup>m-z</sup> ) embryos
Supplementary Figure 6	<i>Ezh2</i> independent targeting of PRC1 proteins to paternal pericentric heterochromatin in the zygote
Supplementary Figure 7	<i>Suv39h</i> -pathway impairs PRC1 targeting to pericentric heterochromatin
Supplementary Methods	

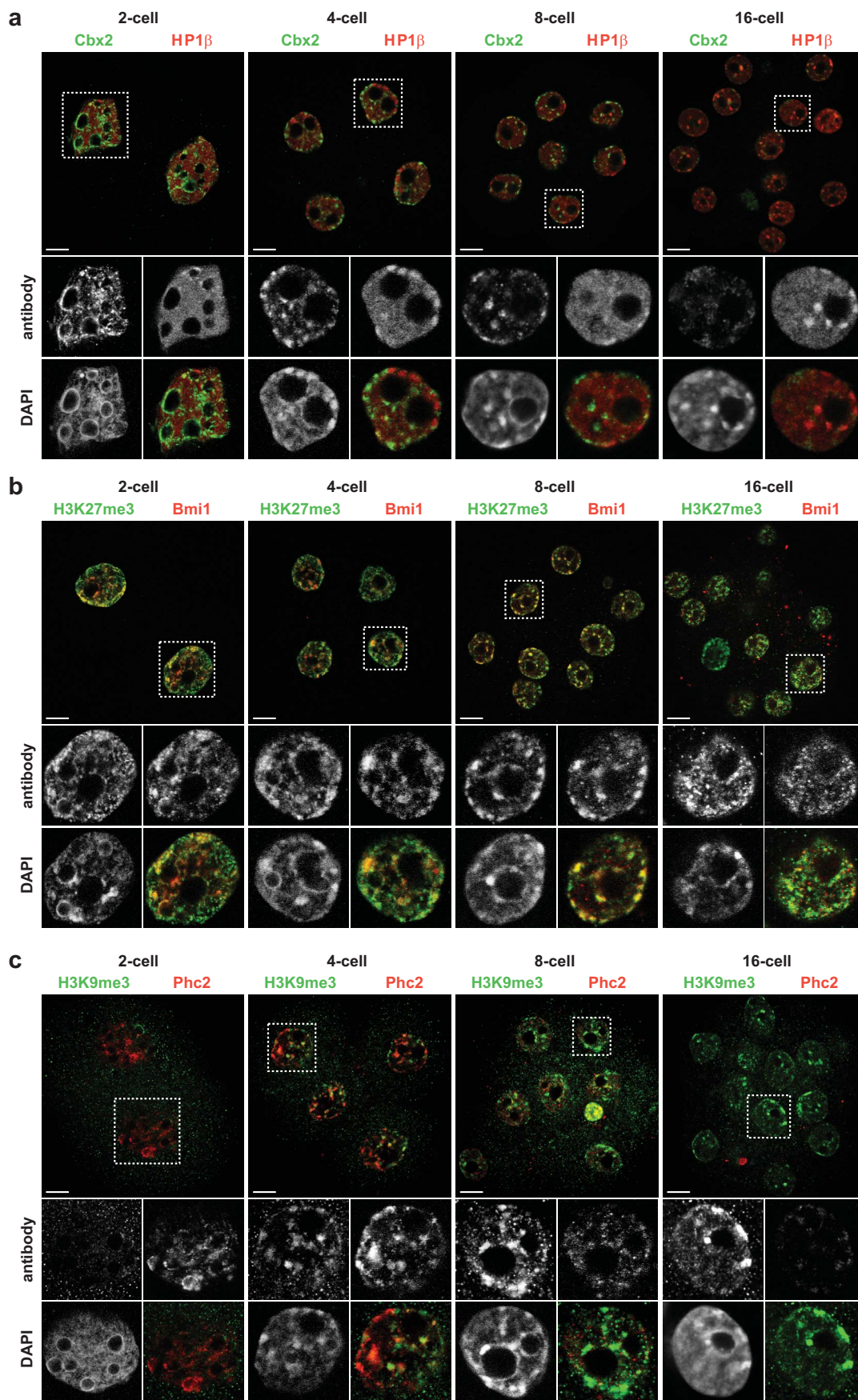




**Supplementary figure 1**

**Immunofluorescence analyses of various histone methylation marks in oocytes and early pre-implantation embryos.**

For all zygotes, the smaller maternal pronucleus is depicted on the right. Zygotes were obtained by IVF and 2-cell embryos after *in vivo* fertilization. (a) H3K9me3 is present in maternal euchromatin and is enriched at maternal DAPI-intense heterochromatin in oocytes and zygotic embryos. Paternal chromatin is lacking H3K9me3 in the zygote. In 2-cell embryos, H3K9me3 is globally down-regulated. (b) HP1β is present at maternal and paternal euchromatin but is only maternally enriched at DAPI-intense heterochromatin. Immunofluorescence of M-II oocytes shows that HP1β dissociates from metaphase chromosomes. (c) H4K20me3 is only present at maternal DAPI-intense heterochromatin in oocytes and zygotic embryos. Paternal chromatin is devoid of H4K20me3. (d) H3K27me1 stains euchromatin and heterochromatin of both parental genomes. Initially, chromatin of decondensing sperm is lacking H3K27me1. The modification appears at paternal chromatin around pronuclear stage 1/2 (PN1/2), before replication. (e, f) Initially, paternal chromatin is lacking H3K27me2 (e) and H3K27me3 (f). Both marks appear during the time of replication (PN4/5) and after the appearance of H3K27me1. H3K27me2 and me3 are present in euchromatin of both parental genomes but are only enriched at paternal heterochromatin. Scale bars: 10 micrometer.

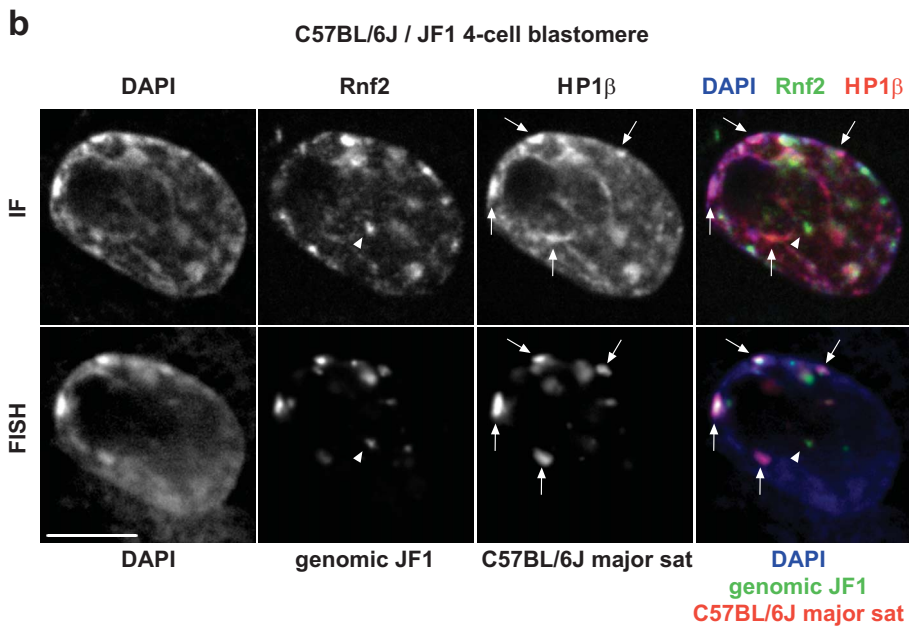
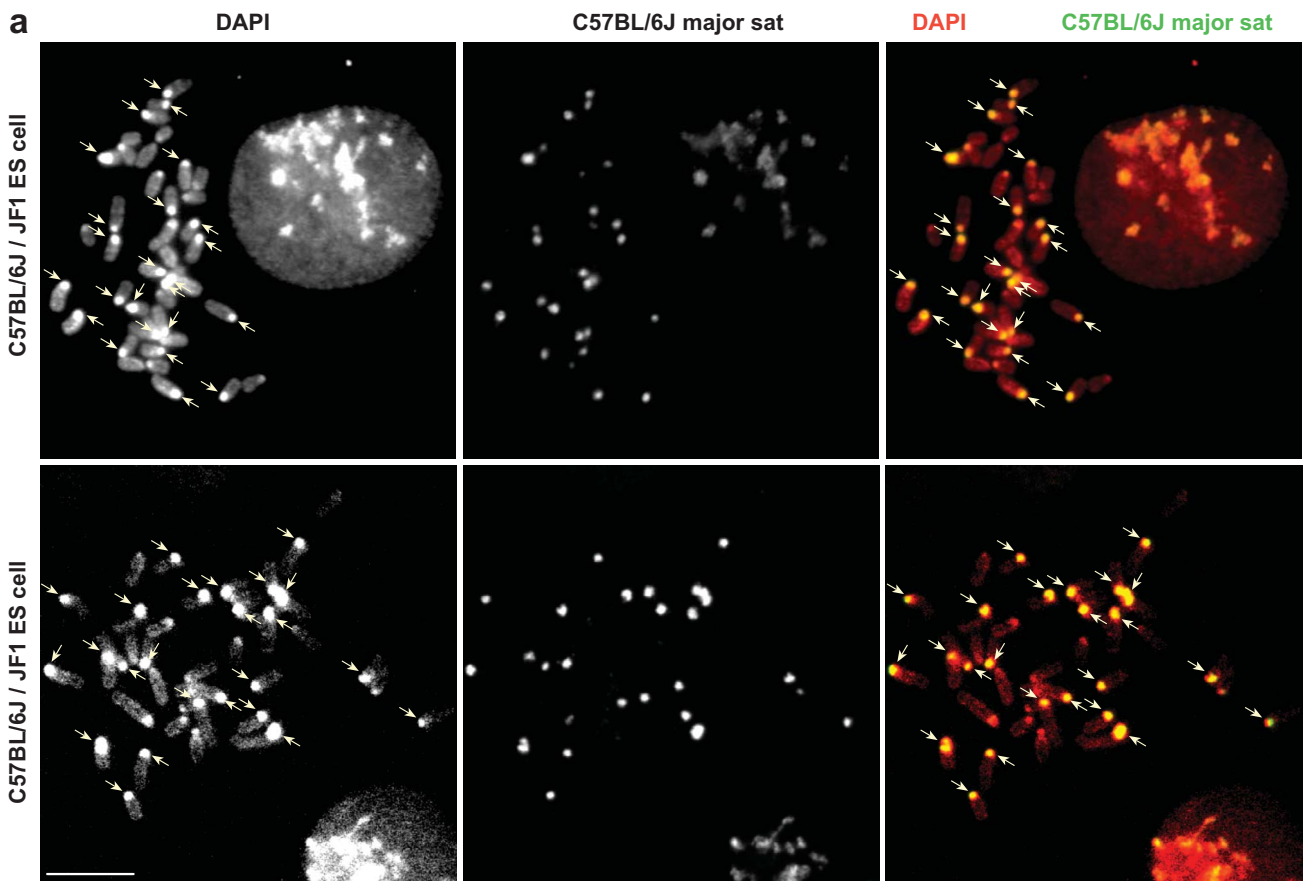


Supplementary figure 2

**Differential heterochromatic states are maintained up to the 8-cell stage in a parental-origin-dependent manner.**

(a) Immunofluorescence analyses of 2-, 4- and 8-cell embryos show almost exclusive enrichment for either Cbx2 or HP1 $\beta$  at DAPI-intense heterochromatic chromocenters. From the 16-cell stage onwards Cbx2 expression is reduced and instead heterochromatin is now solely labeled by HP1 $\beta$ . (b) Immunofluorescence analyses of 2-, 4-, and 8-cell embryos show substantial co-localization for H3K27me3 and Bmi1 in euchromatin and at approximately half of DAPI-intense heterochromatic chromocenters whereas remaining chromocenters are devoid of both H3K27me3 and Bmi1. At the 16-cell stage, H3K27me3 and Bmi1 are almost exclusively localized in euchromatic regions. (c) Immunofluorescence analysis of 2-cell embryos shows preferential enrichment of Phc2 at approximately half of DAPI-intense chromocenters. At the 4-cell stage, about half of chromocenters are labeled with H3K9me3 and half with Phc2. Maternal Phc2 protein is strongly reduced at the 8-cell stage and not detectable any more in consecutive embryonic stages where all pericentric heterochromatin domains are enriched in H3K9me3. Scale bars: 20 micrometer.

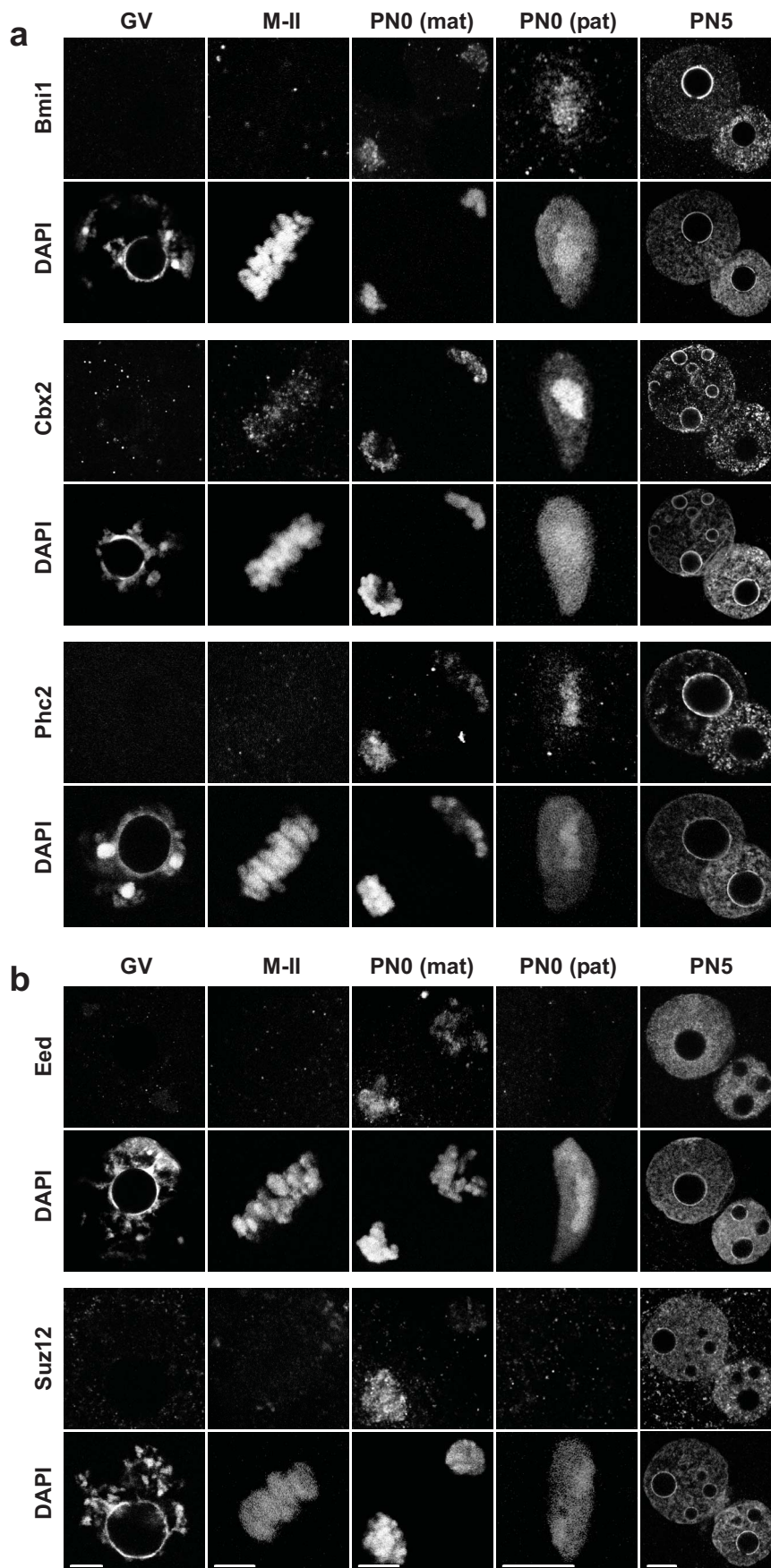




### Supplementary figure 3

#### HP1 $\beta$ is specifically enriched at maternal heterochromatin in 4-cell embryos.

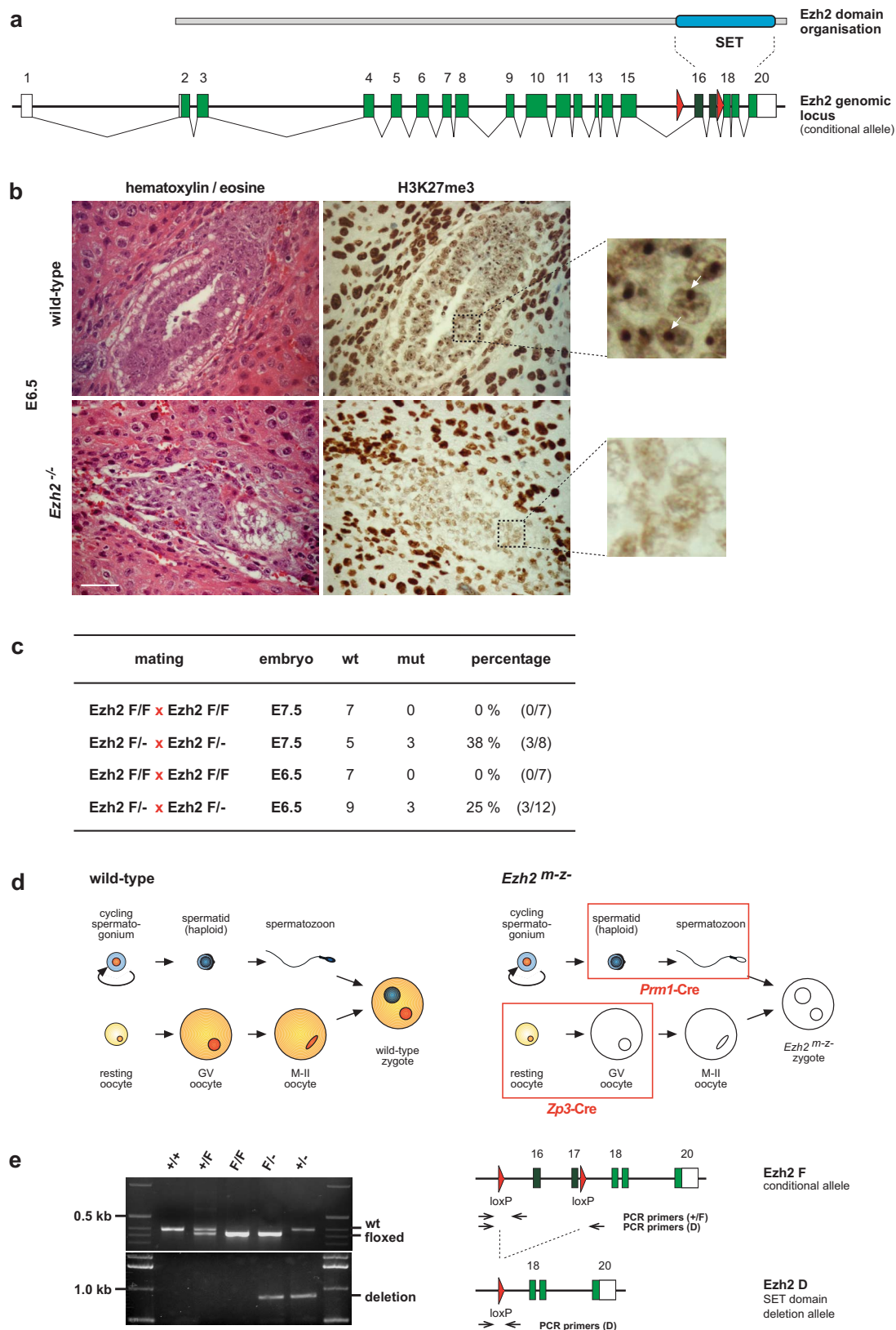
(a) DNA fluorescence *in situ* hybridization (FISH) analysis of ES cells, hybrid for *Mus musculus musculus* C57BL/6J and *Mus musculus molossinus* JF1 genomes, with a *Mus musculus musculus* derived major satellite probe. Mitotic metaphase analyses show that the FISH probe recognizes pericentromeric regions of approximately half of the chromosomes (24/40 in upper and 24/39 in lower image) indicating that the probe is largely specific for DNA sequences of C57BL/6J origin. Note that pericentromeric regions of JF1 chromosomes, not labeled by the major satellite probe, are less intensely stained with DAPI, suggesting that the underlying sequences are less AT-rich. (b) Sequential immuno-DNA-FISH of hybrid C57BL/6J<sup>mat</sup>/JF1<sup>pat</sup> 4-cell embryos shows that HP1 $\beta$  is enriched at DAPI-intense heterochromatin that is labeled by the C57BL/6J specific major satellite probe (arrows) indicating specific accumulation of HP1 $\beta$  at maternal heterochromatin. In contrast, several Rnf2 positive foci do not co-localize with the C57BL/6J major satellite probe but instead co-localize with a DNA probe generated from total genomic JF1 DNA (arrow head). Note that in addition to labeling pericentromeric regions of JF1 chromosomes, the JF1 genomic probe also detects major satellite regions of some C57BL/6J chromosomes (data not shown), explaining co-hybridisation of this probe with the C57BL/6J major satellite probe at DAPI-intense regions. Embryos were first labeled for Rnf2 and HP1 $\beta$  and processed by confocal microscopy (top row of images). Subsequently, FISH was performed and sampled by microscopy (bottom row). Confocal sections obtained by the two procedures were matched according to confocal planes of DAPI images. Scale bars: 10 micrometer.



**Supplementary figure 4**

**Immunofluorescence analyses of PRC1 and PRC2 proteins in oocytes and zygotes.**

(a) PRC1 proteins are not detectable at chromatin of GV and M-II oocytes but start to bind to the maternal genome shortly after fertilization. The more strongly labeled chromatin complement will constitute the embryo proper (left bottom). The top set will segregate into the polar body. In the zygote, PRC1 proteins are present in euchromatin of both parental genomes and are specifically enriched at paternal heterochromatin of decondensing sperm nuclei and of paternal pronuclei. (b) PRC2 proteins start to bind to maternal chromatin shortly after gamete fusion at meiotic anaphase-II. In contrast to PRC1 proteins, PRC2 proteins do not accumulate in decondensing sperm nuclei but in paternal chromatin from PN2 onwards. Scale bars: 10 micrometer.

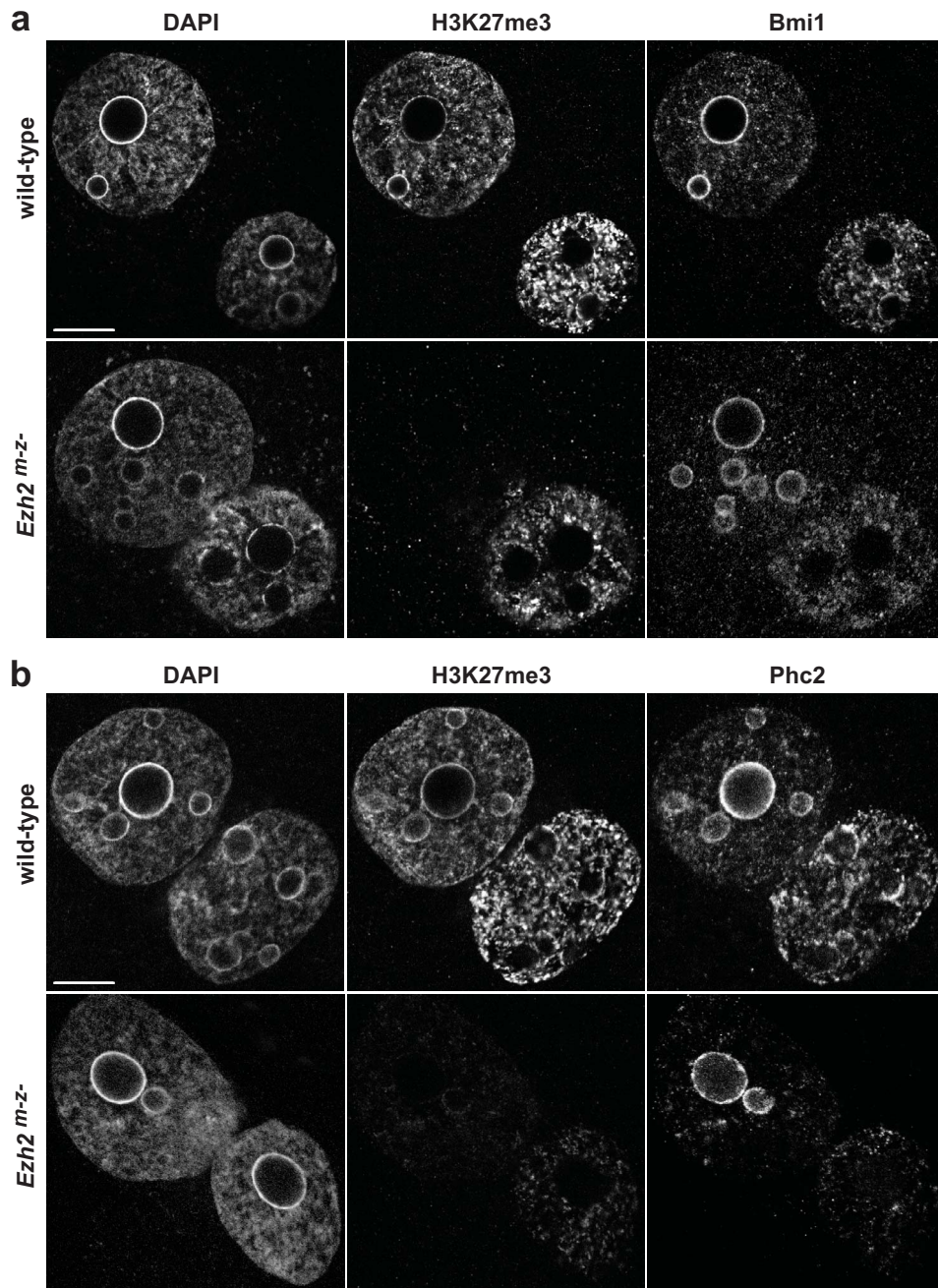


Supplementary figure 5

### Histological analysis of *Ezh2* deficient mice and breeding strategies to generate *Ezh2* maternally and zygotically deficient (*Ezh2<sup>m-z-</sup>*) embryos.

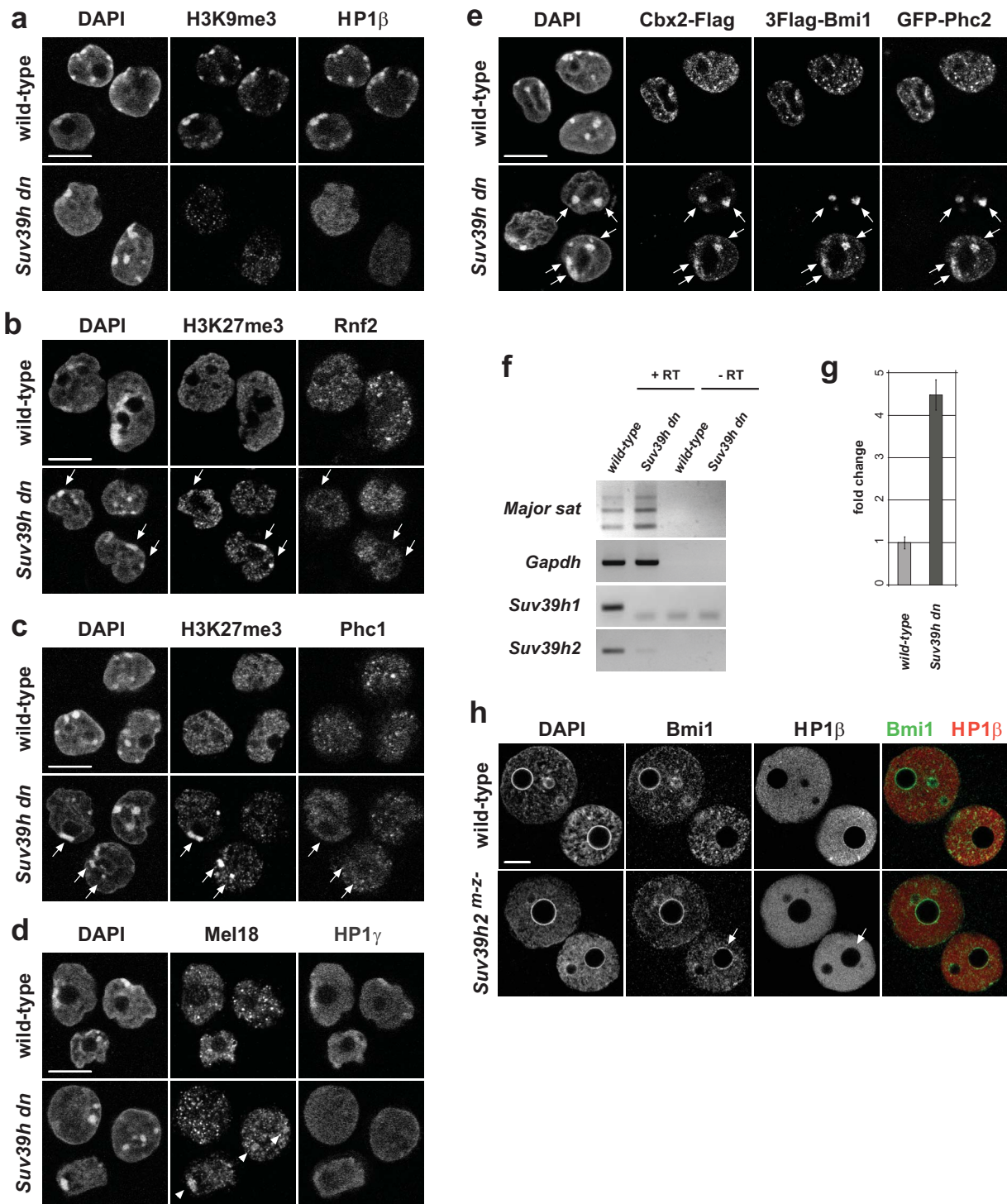
(a) Schematic diagram of *Ezh2* protein and genomic locus. Boxes and numbers refer to exons. Triangles represent LoxP sites in conditional floxed *Ezh2* allele. (b) Developmental arrest and gastrulation failure in *Ezh2* deficient embryos. Histological analyses of wild-type and *Ezh2* deficient embryos at day 6.5 of gestation, produced in crosses of heterozygotes. Five micron sections of deciduae were stained with hematoxylin and eosine (left) or with antibodies specific for H3K27me3 (right). Embryos were classified as normal (wild-type and heterozygous) and abnormal (mutant) according to size and morphology. In wild-type embryos, H3K27me3 staining was observed as numerous small foci in embryonic tissues. In female embryos, the presumptive inactivated X chromosome was strongly labeled (arrows in magnification). Lack of H3K27me3 specific staining in small and developmentally retarded embryos confirmed *Ezh2* deficiency. Scale bar: 50 micrometer. (c) Percentage of *Ezh2* deficient embryos at E6.5 and E7.5, as determined by histological analyses. (d) Breeding strategy to produce *Ezh2* maternally and zygotically deficient embryos. The maternal allele was deleted in dictyate stage growing primary oocytes by Cre-recombinase, expressed from the *Zona pellucida 3-cre* transgene, present in *Ezh2<sup>F/F</sup>* females. Likewise, the paternal allele was deleted in elongating spermatids by Cre-recombinase, expressed from the *Protamine 1-cre* transgene, present in *Ezh2<sup>F/F</sup>* males. Compound *Ezh2<sup>F/F</sup>; Zp3-cre* and *Ezh2<sup>F/F</sup>; Prr1-cre* mice are maintained by transmission of the cre allele via the sex in which the transgene is inactive. (e) PCR genotyping strategy to identify wild-type, floxed and deletion *Ezh2* alleles.





Supplementary figure 6

***Ezh2* independent targeting of PRC1 proteins to paternal pericentric heterochromatin in the zygote.** (a, b) Immunofluorescence analyses of wild-type and *Ezh2*<sup>m-z-</sup> zygotes show that Bmi1 (a) and Phc2 (b) localization to paternal heterochromatin is unaltered in *Ezh2*<sup>m-z-</sup> zygotes whereas levels of euchromatic binding correlate with the level of H3K27me3. Scale bars: 10 micrometer.



Supplementary figure 7

***Suv39h*-pathway impairs PRC1 targeting to pericentric heterochromatin.**

(a-c) Immunofluorescence characterization of wild-type and *Suv39h* dn ES cells. (a) *Suv39h* dn ES lack H3K9me3 and HP1 $\beta$  at constitutive heterochromatin. Although H3K27me3 (arrows) is enriched in absence of H3K9me3 and HP1 $\beta$ , Rnf2 (b, arrows) and Phc1 (c, arrows) fail to accumulate at heterochromatin in *Suv39h* dn ES cells. (d) Mel18 is enriched at a few heterochromatin foci (arrow heads) in wild-type (1.5%; n=266) as well as in *Suv39h* dn mutant (3.6%; n=398) cells. (e) Immunofluorescence characterization of wild-type and *Suv39h* dn ES cells overexpressing three tagged matPRC1 components. Only in *Suv39h* dn ES cells, overexpressed Cbx2, Phc2 and Bmi1 are retargeted to constitutive heterochromatin. Overexpressed Cbx2-Flag and 3Flag-Bmi1 were detected using antibodies specific for the respective protein. For GFP-Phc2, the GFP signal was detected. (f) Semi-quantitative and (g) quantitative RT-PCR analysis of major satellite transcripts revealing increased expression of pericentric sequences in *Suv39h* dn ES cells over wild-type controls. *Suv39h1* and *Suv39h2* expression analyses confirm genotypes of used material. Arrow bars represent standard deviation of three PCR amplifications for each sample. Similar results were obtained in two independent experiments. (h) Immunofluorescence staining of wild-type and *Suv39h2*<sup>m-z-</sup> deficient zygotes. matPRC1 component Bmi1 localizes to maternal heterochromatin in absence of H3K9me3 and HP1 $\beta$  (arrow). Scale bars: 10 micrometer.

## Supplementary information: Methods

### Collection, *in vitro* fertilization and culture of mouse oocytes and embryos

Mouse oocytes and embryos were derived from superovulated 5-10 week old females according to standard procedures<sup>50</sup>. Fully-grown germinal vesicle (GV)-intact oocytes were collected 46 h after PMSG injection (5 U, Intervet) in M2 medium (Sigma) containing 2.5  $\mu$ M milrinone (Sigma). Metaphase II-arrested eggs were collected from PMSG- and hCG-primed (5 U each, Intervet) mice 14 h after hCG injection. Embryos were harvested from superovulated females mated to appropriate males in FHM medium (Chemicon) at indicated time points after hCG injection: late zygotes (26 h), 1/2 cleavage (30 h), early 2-cell (36 h), late 2-cell (46 h), 2/4-cell cleavage (48-50 h), 4-cell (54 h), 4/8 cleavage (58-60 h), early 8-cell (62 h), late 8-cell (68 h), 8/16 cleavage (69-71 h), 16-cell (73 h) and blastocyst stage embryos (94 h). Where precise timing of progression of zygote stages was required, oocytes were fertilized *in vitro*. Sperm was obtained from 10-16 week old CD1 males, and M-II oocytes used for IVF were collected from CD1 females 14 h after hCG injection. Sperm capacitation was carried out in HTF containing 9 mg/ml BSA for 2h. IVF was performed in capacitation medium for 2 h and thereafter the embryos were cultured in KSOM medium plus amino acids (Chemicon) in a humidified atmosphere of 5% CO<sub>2</sub> in air until required. Zygotes were substaged according to morphology of pronuclei using criteria as defined previously<sup>51,52</sup>. In brief, PN0 refers to oocytes immediately after fertilization, PN1 pronuclei are small and reside at the periphery of the embryo, PN2 pronuclei have an increased size and have started to migrate towards the center of the embryo, PN3 pronuclei have migrated towards the center, large PN4 pronuclei were close to each other in the center of the embryo and PN5 refers to large central pronuclei.

### Mice and cell lines

*Ezh2* conditional mice were genotyped by PCR on genomic tail tip DNA to identify wild-type, conditionally floxed and deleted *Ezh2* alleles (Supplementary Fig. 5). Wild-type and floxed alleles were discriminated using 5'-TGACATGGGCCTCATAGTGAC forward and 5'-ACCATGTGCTGAAACCAACAG reverse primers resulting in 315 bp (wild-type) or 280 bp (floxed) products respectively. The mutated *Ezh2* allele was detected with 5'-CCCATGTTTAAGGGCATAGTG forward and 5'-TACTTCCTCAGGATTCGACTTAAGG reverse primers. *Ezh2*<sup>F/F</sup>; *Zp3-cre*<sup>53</sup> and *Ezh2*<sup>F/F</sup>; *Prm1-cre*<sup>54</sup> mice were maintained on a mixed background of 129/Sv and C57BL/6J. All experimental *Ezh2*<sup>F/F</sup> *Prm1-Cre*<sup>+</sup> males were tested for deletion efficiency by crossing them to C57BL/6J females and genotyping their offspring for the *Ezh2* deletion allele, revealing a deletion efficiency of >96% (n=25 males). For genotyping of *Rnf2* conditional mice, 5'-GTCTCATTTCCCAGTGTGTCTC forward and 5'-ACTGACCCATGGCTCTTGATG reverse primers were used to discriminate wild-type (418bp) and floxed (488bp) alleles. The deletion allele was detected with 5'-GTCTCATTTCCCAGTGTGTCTC forward and 5'-GATGCACTGTCCTGATGGC reverse primers. Mice were maintained on a mixed background of 129/Sv and C57BL/6J. *Suv39h2* mice were genotyped with 5'-TTTGAGGGGACGACGACAGTATCG, 5'-CTTATTGTAGCCTGGTGTGTGCC and 5'-GCAAACAGTCAAGAGTTGGATGC primers resulting in a 377 bp and 490 bp band for wild-type and targeted alleles, respectively. *Suv29h2* mice were maintained on a C57BL/6J and 129/Sv background. Wild-type and *Suv39h* dn ES cells<sup>13,18</sup>, C57BL/6J/JF1 hybrid ES cells and CCE ES cells<sup>55,56</sup> were cultured in DMEM medium with 4.5 g/l glucose (Gibco) containing 15% FCS



(foetal calf serum, Chemicon), penicillin, streptomycin, 2 mM L-glutamine, 0.1 mM 2 $\beta$ -mercaptoethanol, non-essential amino acids and 1 mM sodium pyruvate (Gibco). NIH3T3 fibroblasts were cultured in DMEM medium containing 10% FCS (BioConcept). Wild-type and *Suv39h* dn ES cells<sup>13,18</sup> were triple transfected with Cbx2-Flag (pCAG-Puro), 3Flag-Bmi1 (pCAG-Puro) and GFP-Phc2 (pEGFP-C2) using Lipofectamine 2000 (Invitrogen) and after 12h were seeded on poly-L-lysine coated coverslips, fixed with PFA and processed for immunofluorescence.

### Antibodies

For immunofluorescence analyses of Polycomb group proteins, the following antibodies were used: polyclonal anti-Rnf2 (van Lohuizen, 1:50), monoclonal anti-Rnf2<sup>57</sup> (1:400), polyclonal and monoclonal anti-Bmi1 (van Lohuizen, 1:400 and 1:50 respectively), anti-Rnf110 (Mel18, 10744, Santa Cruz, 1:50), polyclonal anti-Phc1 (van Lohuizen, 1:500), monoclonal anti-Phc1<sup>58</sup> (1:2), monoclonal anti-Phc2<sup>59</sup> (1:50), polyclonal anti-Cbx2 (Otte, 1:500), monoclonal anti-Ezh2 (M5 and M18<sup>60</sup>, undiluted), monoclonal anti-Eed (M26,<sup>60</sup> undiluted) and polyclonal anti-Suz12 (07-379, Upstate, 1:500).

Antibodies against histone modifications and associated proteins: anti-HP1 $\beta$  (MCA1946, Serotec, 1:500), polyclonal anti-H3K9me3, anti-H3K27me1, anti-H3K27me2, anti-H3K27me3 and anti-H4K20me3<sup>31</sup> (all 1:500). For detection of Rnf2-YFP, a cross-reacting polyclonal anti-GFP antibody (Clontech, 1:200) was used.

### Immunofluorescence

Before fixation of oocytes and embryos, the zona pellucida was removed by incubation in acidic tyrode for 30 seconds. Embryos were washed twice in FHM, fixed for 15 min in cold 4% paraformaldehyde in PBS (pH 7.4) and permeabilized with 0.2% Triton-X 100 in PBS for 15 min at room temperature (RT). Fixed embryos were blocked overnight at 4°C in 0.1% Tween-20 in PBS containing 2% BSA and 5% normal donkey serum, and were then incubated with primary antibodies in blocking solution overnight at 4°C. Double antibody stainings were accomplished by mixing appropriate different primary and different secondary antibodies for simultaneous incubation. Embryos were washed three times for 20 min in 0.1% Tween-20 in PBS containing 2% BSA before application of secondary antibodies. For detection, anti-rabbit IgG-Alexa 488, anti-rat IgG-Cy3, anti-mouse IgG-Alexa 488 and anti-mouse IgG-Alexa 555 (Molecular Probes) secondary antibodies were diluted 1:500 in blocking solution and embryos were incubated for 1 h at RT followed by three washing steps a 20 min in 0.1% Tween-20 in PBS containing 2% BSA in the dark. Embryos were mounted in Vectashield containing DAPI (Vector)<sup>7</sup>.

ES cells were trypsinized and placed on poly-L-lysine coated coverslips for 10 min to attach. Cells were fixed with 2% paraformaldehyde in PBS (pH 7.4), permeabilized in 0.1% Triton-X100 in 0.1% sodium citrate and blocked for 30 min in 0.1% Tween-20 in PBS containing 2% BSA and 5% normal donkey serum at RT. Incubation with primary and secondary antibodies as well as mounting was performed as described above.

To open up the condensed chromatin structure of mature sperm, sperm from caudal epididymis was treated with a de-condensing solution containing 25 mM DTT, 0.2% Triton X-100 and 200 U heparin/ml<sup>8</sup>, and subsequently IF was performed as described above for ES cells.

### **Immuno-DNA-FISH and RNA-FISH**

Following removal of the zona pellucida with acid tyrode solution, zygotes and 4-cell embryos were rinsed twice in FHM, incubated for 5 minutes in PBS containing 6 mg/ml BSA and transferred onto Denhardt's solution-coated coverslips<sup>48</sup>. After removal of excess solution, embryos were dried down for 30 min at RT. Samples were fixed in 3% paraformaldehyde in PBS (pH 7.2) for 10 min at RT. After 3 washes in PBS, embryos were permeabilized in PBS containing 0.5% Triton X-100 for 3 min on ice. Fixed embryos were blocked for 20 min in PBS containing 1 % BSA, incubated with primary antibody (90 min) followed by secondary antibody (45 min). All immunostaining steps were performed at RT and excess antibodies were washed in PBS three times for 5 min. Preparations were post-fixed in 3% paraformaldehyde for 10 min at RT, rinsed in PBS and mounted in Vectashield with 1 µg/ml DAPI (Vector) prior to analyses of immunofluorescence stainings. For subsequent DNA fluorescence *in situ* hybridization (FISH), coverslips were washed extensively in PBS to remove mounting solution and incubated in a permeabilization solution (0.7% Triton X-100, 0.1 M HCl) for 10 min on ice. Preparations were washed in 2x SSC and DNA was denatured for 30 min at 80°C. Hybridization was performed overnight at 42°C in a humid chamber and excess probe was eliminated through two washes in 2x SSC (37°C for 30 min), followed by one wash in 1x SSC (RT for 30 min) and one wash in 0.5x SSC (RT for 30 min). Slides were mounted in Vectashield containing DAPI prior to analyses. The mouse major-satellite probe<sup>61</sup> was directly labeled with Spectrum-RED (Vysis).

For RNA-FISH, zygotes were prepared and fixed as described above. Embryos were permeabilized in PBS containing 0.5% Triton X-100 for 2 min on ice and washed in 2xSSC. Hybridization with the mouse major-satellite probe<sup>61</sup> was performed overnight at 37°C in a humid chamber. Subsequent washing and mounting was done as described for DNA-FISH.

### **RT-PCR**

For RNA isolation, oocytes and embryos were pooled from several mice and RNA was isolated from batches of 50 oocytes or embryos. Embryos were transferred into Trizol and 100 ng of E. coli rRNA was added to each sample as carrier. RNA was isolated according to the manufacturer's instructions. Reverse transcription and PCR reaction were performed as previously described<sup>62</sup>. Briefly, reverse transcription was performed from total RNA corresponding to 20 oocytes or embryos using random primers (200 ng) and SuperScript II RNase H Reverse Transcriptase (Invitrogen) according to the manufacturer's protocol. For PCR reactions, cDNA corresponding to 0.2 oocytes or embryos was used as a template. Amplifications were carried out using Taq DNA polymerase (Qiagen), PCR products were resolved on a 2% agarose gel and subsequently detected via SYBR green I staining (Molecular Probes, 1:10,000). Fluorescence was detected on a Typhoon 9400 scanner (Amersham Biosciences). The following primers and cycle numbers were used:

Ezh2 (F: 5'-AGCCTTGTGACAGTTCGTGC, R: 5'-TTTAGAGCCCCGCTGAATG, 33 cycles),

Eed (F: 5'-ACCAGCCATTGTTTGGAGTTC, R: 5'-ACCTCCGAATATTGCCACAAG, 32 cycles),

Suz12 (F: 5'-CTTCGATGGACAGGAGAAACC, R: 5'-AGGTCGTCTCTGGCTTCTGTGC, 32 cycles),

Cbx2 (F: 5'-GTAGTCCCAAAGCCCAGTCAG, R: 5'-CAAGTGCCTACATCAGCTTGC, 38 cycles),

Cbx4 (F: 5'-GTGGAGCCCTTGAGTGAGTTC, R: 5'-CCGGAGTAGAGTCAGCACTTG, 40 cycles),

Cbx6 (F: 5'-AGTGGAGAAGGAGCTGAATGC, R: 5'-CCCTTGTAGTCCAGGAGATGC, 40 cycles),

Cbx7 (F: 5'-GTCCTGGGCCCCACTCG, R: 5'-CATTGGTCAGGTCTGCTTCTG, 40 cycles),  
Cbx8 (F: 5'-CGGCGAGTCAACATGGAG, R: 5'-CCTGGAAGTAGACGCCAAATC, 40 cycles),  
Ring1 (F: 5'-GCCATCATGGATGGTACAGAG, R: 5'-TATTCCTCCCGGCTAGGGTAG, 38 cycles),  
Rnf2 (F: 5'-AGGCAATAACAGATGGCTTGG, R: 5'-GAGAGCCTGCTGATTGTTGTG, 33 cycles),  
Bmi1 (F: 5'-AAGACCGAGGAGAAGTTGCAG, R: 5'-CCCAGAGTCACTTTCCAGCTC, 32 cycles),  
Rnf110 (F: 5'-GGTGACACTTCCCAAATCTCC, R: 5'-ATCTGATGCTCAGCAGTGGTC, 33 cycles),  
Rnf134 (F: 5'-GAGCGCCTGATAAACCTTGTC, R: 5'-TCTCCTGAAACACGCACAAAC, 32 cycles),  
Phc1 (F: 5'-GCCTTCTTCAGGATTGACTGG, R: 5'-GATCACCCTTGCTTCTGCTG, 32 cycles),  
Phc2 (F: 5'-CAGTGCTCTACCACGCATGTC, R: 5'-GCTGGATGTTGGACTCTTG, 32 cycles),  
Phc3 (F: 5'-GTACCTGCAGCAGATGTACGC, R: 5'-CTGCAGACTGACAGGAAGGTG, 32 cycles),  
Suv39h1 (F: 5'-CGCATCGCATTCTTTGCC, R: 5'-AAGCCGTTGTCCCACATTTG, 35 cycles),  
Suv39h2 (F: 5'-AAATCCAACCAGGCACTCCC, R: 5'-CTCGTAGTCCAGGTCAAAGAGGTAG, 32 cycles),  
Actb (F: 5'-TCGCCATGGATGACGATA, R: 5'-AGGTGTGGTGCCAGATCTTC, 33 cycles),  
Gapdh (F: 5'-ACAACCCCTTCATTGACCTC, R: 5'-TTCTGAGTGGCAGTGATGGC), and  
Zp3 (F: 5'-AAGCTCAACAAAGCCTGTTTCG, R: 5'-TATTGCGGAAGGGATACAAGG, 32 cycles).

For analysis of major satellite transcription in late 2-cell embryos, RNA was isolated using the Absolutely RNA Nanoprep kit (Stratagene). cDNA was prepared using random primers as described above. For semi-quantitative RT-PCR in ES cells, 5'-GACGACTTGAAAAATGACGAAATC forward and 5'-CATATTCCAGGTCCTTCAGTGTGC reverse primers<sup>63</sup> were used (25 cycles). Quantitative RT-PCR of ES cells and 2-cell embryos was carried out using the same primers and qPCR MasterMix plus SYBR Green I w/o UNG (Eurogentec) on an ABI Prism 7000 light cycler. Error bars represent the standard deviation.

### **Western blot analyses of oocytes and embryos**

For western blot analysis, 200 oocytes or embryos per stage were pooled from several mice, washed through three drops of MEM containing 0.1% PVP (Polyvinyl pyrrolidone K90) and were then directly pipetted into 15 µl of 5x SDS loading buffer. Samples were denatured for 3 min at 95°C, proteins were separated on a 10% SDS gel and blotted onto Hybond-P membranes (Amersham). Membranes were probed for Ezh2 (M18<sup>60</sup>, undiluted), Rnf2 (polyclonal, van Lohuizen, 1:500) and histone H3 (1791, Abcam, 1:20,000). Primary antibodies were detected with anti-rabbit HRP or anti-mouse HRP followed by ECL detection (Amersham).

### **Western blot analyses of mature spermatozoa**

Mature sperm was obtained from caudal epididymi of adult CD1 males. Motile mature spermatozoa were recovered by a swim up method. Sample purity was verified by microscopy. Sperm proteins were isolated as described in<sup>64</sup> with minor modifications. Proteins were extracted under reducing conditions (6 M guanidine-HCl, 10 mM dithiothreitol), precipitated by 0.9 M HCl and 20 % trichloroacetic acid treatment and separated by SDS-PAGE. Western blot was performed as described in<sup>64</sup>. Protein extracts from CCE ES cells were used as a control and were prepared using the same procedure as described for the sperm extract.

## References

50. Hogan, B., Costantini, F. & Beddington, R. *Manipulating the mouse embryo*, (Cold Spring Harbor Laboratory Press, 1994).
51. Adenot, P.G., Mercier, Y., Renard, J.P. & Thompson, E.M. Differential H4 acetylation of paternal and maternal chromatin precedes DNA replication and differential transcriptional activity in pronuclei of 1-cell mouse embryos. *Development* **124**, 4615-25 (1997).
52. Santos, F., Hendrich, B., Reik, W. & Dean, W. Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev Biol* **241**, 172-82 (2002).
53. de Vries, W.N. et al. Expression of Cre recombinase in mouse oocytes: a means to study maternal effect genes. *Genesis* **26**, 110-2 (2000).
54. O'Gorman, S., Dagenais, N.A., Qian, M. & Marchuk, Y. Protamine-Cre recombinase transgenes efficiently recombine target sequences in the male germ line of mice, but not in embryonic stem cells. *Proc Natl Acad Sci U S A* **94**, 14602-7 (1997).
55. Keller, G., Kennedy, M., Papayannopoulou, T. & Wiles, M.V. Hematopoietic commitment during embryonic stem cell differentiation in culture. *Mol Cell Biol* **13**, 473-86 (1993).
56. Robertson, E., Bradley, A., Kuehn, M. & Evans, M. Germ-line transmission of genes introduced into cultured pluripotential cells by retroviral vector. *Nature* **323**, 445-8 (1986).
57. Atsuta, T. et al. Production of monoclonal antibodies against mammalian Ring1B proteins. *Hybridoma* **20**, 43-6 (2001).
58. Miyagishima, H. et al. Dissociation of mammalian Polycomb-group proteins, Ring1B and Rae28/Ph1, from the chromatin correlates with configuration changes of the chromatin in mitotic and meiotic prophase. *Histochem Cell Biol* **120**, 111-9 (2003).
59. Isono, K. et al. Mammalian polyhomeotic homologues Phc2 and Phc1 act in synergy to mediate polycomb repression of Hox genes. *Mol Cell Biol* **25**, 6694-706 (2005).
60. Hamer, K.M. et al. A panel of monoclonal antibodies against human polycomb group proteins. *Hybrid Hybridomics* **21**, 245-52 (2002).
61. Brown, K.E. et al. Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. *Cell* **91**, 845-54 (1997).
62. Puschendorf, M. et al. Abundant transcripts from retrotransposons are unstable in fully grown mouse oocytes. *Biochem Biophys Res Commun* **347**, 36-43 (2006).
63. Terranova, R., Sauer, S., Merckenschlager, M. & Fisher, A.G. The reorganisation of constitutive heterochromatin in differentiating muscle requires HDAC activity. *Exp Cell Res* **310**, 344-56 (2005).
64. Lee, K., Haugen, H.S., Clegg, C.H. & Braun, R.E. Premature translation of protamine 1 mRNA causes precocious nuclear condensation and arrests spermatid differentiation in mice. *Proc Natl Acad Sci U S A* **92**, 12451-5 (1995).

### **2.3. The Cbx2 chromodomain and AT hook are required for heterochromatin targeting of PRC1**

Mareike Puschendorf, Carolin Kolb, Jean-Francois Spetz, Michael Stadler and Antoine H.F.M. Peters

### 2.3.1. Introduction

The Polycomb group proteins are best known for their role in maintaining the silent state of *Hox* genes during development. More recently, PcG proteins have been found to target many other developmental regulators in a wide range of cell types<sup>75,108</sup>. Moreover, PcG proteins and H3K27me3 are enriched at facultative heterochromatin of the inactive X chromosome in female mammals<sup>331</sup>. We identified a novel role for PRC1 proteins at pericentric heterochromatin in early mouse pre-implantation embryos (Chapter 2.2). We showed that PRC1 proteins are required to repress pericentric major satellite transcripts.

It is not well understood, especially in mammals, how PcG proteins are recruited to their target genes. We showed that PRC1 binding to paternal pericentric heterochromatin in early embryos is independent of *Ezh2* function and H3K27me3. Moreover, PRC1 recruitment to maternal pericentric heterochromatin is blocked by *Suv39h*-mediated marks but does occur in *Suv39h2* mutant zygotes. We speculated that in analogy to X inactivation<sup>65,80</sup>, genomic imprinting<sup>332</sup> and constitutive heterochromatin formation in somatic cells<sup>333</sup>, a non-coding RNA may be required for matPRC1 heterochromatic localization. Alternatively, the repetitive nature and/or AT-richness of the underlying sequences might drive PRC1 recruitment. Here, we explore PRC1 targeting mechanisms to pericentric heterochromatin.

The Cbx2 AT hook has been suggested to serve as a nucleic acid binding module<sup>62</sup> and therefore might be involved in the targeting of PRC1 to AT-rich major satellite DNA, either by direct binding to the underlying DNA or to transcripts generated from satellite regions. AT hook motifs are also encoded by high mobility group (HMG) proteins<sup>334</sup>, which are, after histones, the most abundant chromatin proteins and are involved in a variety of nuclear processes<sup>335</sup>. HMG proteins are viewed as architectural factors that organize chromatin by appropriate bending of DNA, while their interactions with chromatin remain highly dynamic *in vivo*<sup>335,336</sup>. HMG proteins are divided into three families: HMGA, HMGB and HMGN.

The HMGA family consists of four members, each containing several AT hooks which are nine amino acids motifs that bind to AT-rich stretches of DNA in the minor groove<sup>335,337</sup>. HMGAs are abundant in undifferentiated and proliferating embryonic cells, but are expressed at low levels in most adult cells. *In vivo*, HMGA proteins preferentially, but not exclusively, localize to heterochromatin and are associated with condensed chromosomes during mitosis<sup>336</sup>. As key components of the enhanceosome, a complex of transcription factors and cofactors that binds on nucleosome-free control regions of genes, HMGA proteins are involved in the regulation of specific genes<sup>338,339</sup>. In mouse embryos, HMGA1 is involved in the activation of zygotic transcription and micro-injection of purified HMGA1 leads to precocious transcriptional activation at the 1-cell stage<sup>340</sup>. HMGA1a exerts its function by distorting DNA slightly which increases the affinity of other proteins for their binding sites<sup>338,339</sup>. Extensive post-translational modifications are involved in regulating the stability or disassembly of the enhanceosome (acetylation of K64 and K70, respectively) and in modulating the DNA binding affinity of HMGA1a (phosphorylation of T20, S43 and S63; methylation)<sup>335,338,341</sup>. Deletion of *Hmga1* or *Hmga2* in mouse results in a number of diverse phenotypes, indicating that the two proteins regulate distinct processes (reviewed in<sup>337</sup>). *Hmga1*<sup>-/-</sup> mice suffer from cardiac hypertrophy and

hematological disorders and develop type 2 diabetes. In contrast, *Hmga2*<sup>-/-</sup> mice are small and their spermatogenesis is deficient leading to male sterility.

HMGB proteins contain the 80 amino acid HMG Boxes as functional motifs that bind to the minor groove of DNA with limited sequence specificity and sharply bend DNA, which might be their main function<sup>335,337</sup>. While AT hook proteins cause modest distortions of DNA, HMG boxes can bend DNA to right angles or even more sharply. The third family of HMG proteins binds to nucleosomes, therefore called HMGN, through their positively charged domain<sup>335,337</sup>. Like the other HMG families, HMGNs contain a C-terminal acidic domain that is involved in chromatin unfolding. Taken together, the HMG proteins have been known for several decades and their chromatin binding properties have been well studied<sup>339,342</sup>. In contrast, nothing is known about the function and properties of the Cbx2 AT hooks, but some analogies might be drawn from our understanding of HMGA proteins.

In this study, we use a heterologous ES cell system to show that PRC1 binding to pericentric heterochromatin is dependent on Cbx2. Heterochromatin binding is mediated by a dual module consisting of the Cbx2 chromodomain and AT hook. Micro-injection experiments in mouse zygotes suggest that PRC1 heterochromatin binding in early embryos is RNA-independent but requires the Cbx2 AT hook. We discuss implications of our findings for PRC1 recruitment to target genes and provide first bioinformatics analysis.

### **2.3.2. Methods**

#### *Generation of constructs*

Constructs for C-terminal enhanced GFP (EGFP) fusion proteins of Cbx2, Cbx4, Cbx6, Cbx7 and Cbx8 in the EGFP-N1 backbone (BD Biosciences) were obtained from E. Bernstein<sup>62</sup>. We also received Cbx7-GFP with point mutations in the caging aromatic residues (F11A and W35A) and a swap of the Cbx4 CD into Cbx7-GFP. Point mutations of the Cbx2 CD and AT hooks were generated using a "Self-made Quickchange" protocol for site-directed mutagenesis. Constructs were amplified using Pfu Polymerase (Promega) and approximately 25 bp long primers containing the mutation in the centre. The PCR reaction was digested with the methylation sensitive restriction enzyme Dpn1 (New England BioLabs) to remove the original construct and subsequently used for transformation. Successful mutagenesis was confirmed by sequencing. The Cbx2 truncation constructs were generated by PCR from the Cbx2-EGFP construct and cloned with BglII/SalI into EGFP-N1. Hmga1 was cloned by PCR from NIH3T3 cDNA. All constructs were verified by sequencing throughout the entire coding region. The Cbx2 clone chart and a complete list of all constructs generated are attached in the Appendix.

#### *Transfection of ES cells and analysis of heterochromatin enrichment*

Wild-type and *Suv39h* dn ES cells<sup>1,9</sup> were cultured in DMEM medium with 4.5 g/l glucose (Gibco) containing 15% FCS (fetal calf serum, Chemicon), penicillin, streptomycin, 2 mM L-glutamine, 0.1 mM 2β-mercaptoethanol, non-essential amino acids and 1 mM sodium pyruvate (Gibco). ES cells were transfected with GFP fusion constructs using Lipofectamine 2000 (Invitrogen) and after approximately 16h were seeded on poly-L-lysine coated coverslips, fixed

with PFA and mounted in Vectashield containing DAPI (Vector). For more detailed information on the analysis of ES cells, see also Chapter 2.2.

#### *Microinjection of RNase into zygotes*

Mouse zygotes were isolated as previously described (Chapter 2.2.2). For micro-injection of RNase, zygotes were isolated 20h after hCG injection from CD1 mice and further cultured in M2 medium until the late pronuclear stage (PN4/5, 26h after hCG). RNase A (Roche, #109169)<sup>333</sup> was dissolved in TE buffer (pH 8.0) at 2 mg/ml. Approximately 5  $\mu$ l of RNase were micro-injected directly into either the maternal or paternal pronucleus. Control zygotes were either not injected or injected with TE buffer into the maternal or paternal pronucleus. Following micro-injection, zygotes were cultured for further 15 min. Per condition, 15 zygotes were transferred to lysis buffer for later RNA isolation (Chapter 2.1.1) and RT-PCR (Chapter 2.2.2) to monitor efficiency of RNase treatment. The remaining zygotes were fixed and processed for immunofluorescence (Chapter 2.2.2).

#### *Distamycin treatment during in vitro fertilization*

*In vitro* fertilization (IVF) was carried out as previously described using CD1 mice (Chapter 2.2.2). The fertilization reaction was performed in HTF (Chemicon) supplemented with 20  $\mu$ M or 50  $\mu$ M Distamycin (SIMGA, #D6135) for 2h. Control oocytes were fertilized in HTF without Distamycin. Following IVF, zygotes were fixed and processed for immunofluorescence as previously described (Chapter 2.2).

#### *Microinjection of polyadenylated messages into zygotes*

For *in vitro* transcription, full length and mutated Cbx2-GFP was cut from the EGFP-N1 backbone with BglII/NotI and cloned into BamHI/NotI sites of pBSSK<sup>+</sup>. Plasmids were linearized with NotI and gel purified. Cbx2-GFP fusions were *in vitro* transcribed from a T7 promoter using the mMMESSAGE mMACHINE High Yield Capped RNA Transcription Kit (Ambion, #1344) and messages were polyadenylated with the mMMESSAGE mMACHINE Poly(A) Tailing Kit (Ambion, #1350). mRNAs were diluted in nuclease-free water (not DEPC treated, Ambion) to yield a final concentration of 200 ng/ $\mu$ l. Maternal and paternal *Ezh2*-deficient embryos were obtained from matings between *Ezh2*<sup>F/F</sup>; *Zp3-cre*<sup>+</sup> females and *Ezh2*<sup>F/F</sup>; *Prm1-cre*<sup>+</sup> males. Control embryos designated as wild-type were *Ezh2*<sup>F/F</sup> (Chapter 2.2). Zygotes were isolated 14h after hCG injection and cultured in M2 medium until 19h after hCG. Approximately 5  $\mu$ l of polyadenylated messages (200 ng/ $\mu$ l) were micro-injected into the cytoplasm and zygotes were cultured for further 5h to allow production of Cbx2-GFP fusion proteins *in vivo*. Subsequently, zygotes were fixed for immunofluorescence at the late pronuclear stage (PN5). The direct GFP signal was detected by confocal microscopy (Chapter 2.2).

#### *Author contributions*

M.P. and A.H.F.M.P. conceived and designed the experiments. M.P. and C. Kolb designed the Cbx constructs. C. Kolb performed the cloning of Cbx constructs. M.P. performed and analyzed Cbx localization experiments in ES cells. C. Kolb prepared *in vitro* transcribed and



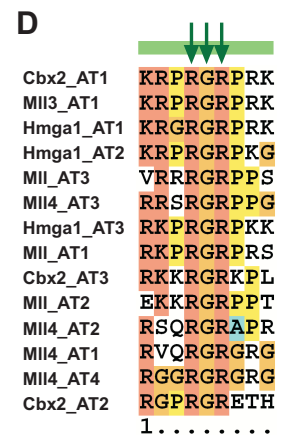
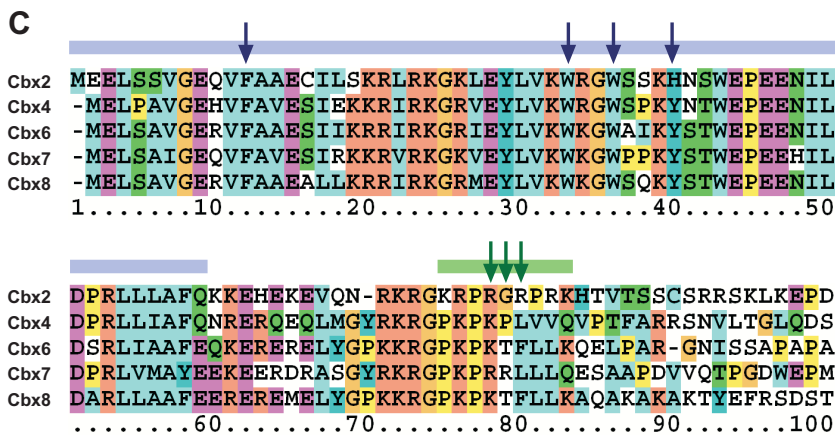
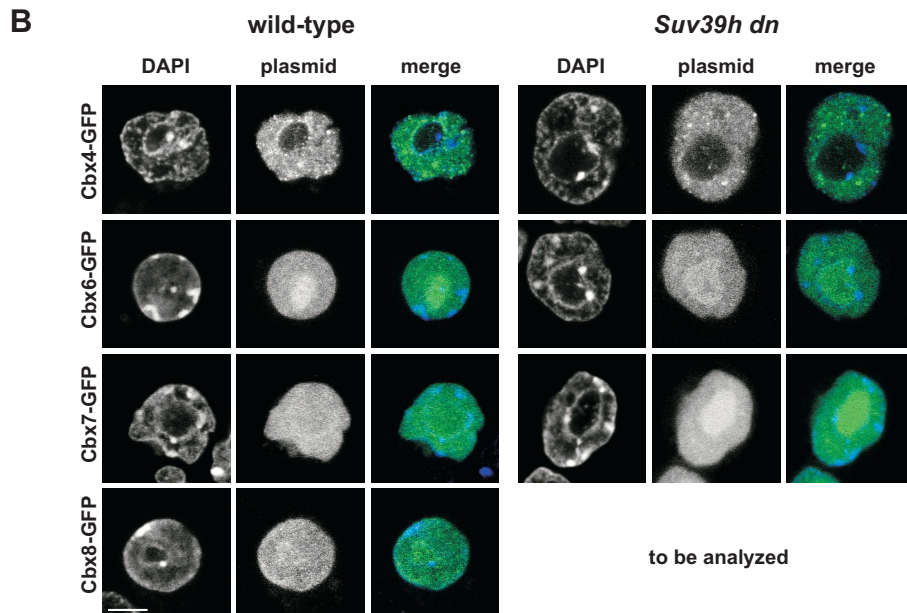
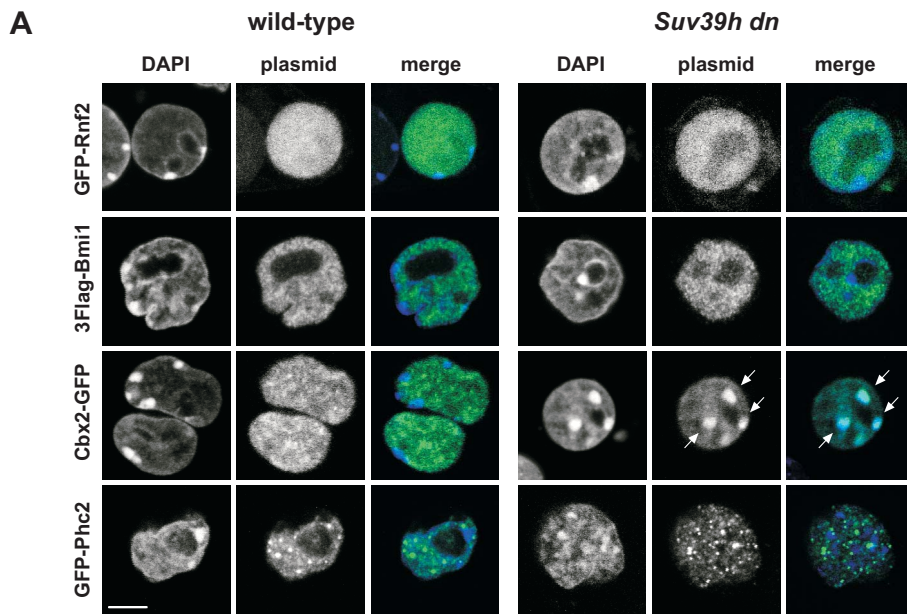
polyadenylated Cbx mRNAs. J.-F. Spetz performed micro-injections of RNase and Cbx mRNAs into zygotes. M.P. analyzed the micro-injected zygotes. M.P. performed the distamycin experiments. M. Stadler, M.P. and A.H.F.M.P. conceived the bioinformatics analysis. M. Stadler performed the bioinformatics tests.

### 2.3.3. Results

#### *Cbx2 mediates heterochromatin binding of PRC1 in *Suv39h* dn ES cells*

*Suv39h1/Suv39h2* double null (*Suv39h* dn) ES cells resemble paternal pronuclei as their pericentric heterochromatin is devoid of H3K9me3, H4K20me3 and HP1 $\beta$ <sup>343</sup>. Remarkably, *Suv39h* dn ES cells gain H3K27me3 at pericentric heterochromatin. In contrast to early embryos, however, PRC1 components that are expressed in ES cells, are not detectable at PCH in *Suv39h* dn ES cells (Chapter 2.2). PRC1 components expressed in ES cells include Rnf2, Mel18 (Rnf110) and Phc1 but we were unable to detect Cbx2, Bmi1 and Phc2, indicating that PRC1 complex constitution varies during development and between cell types<sup>60</sup>. We have previously shown that reconstitution of a matPRC1 complex in ES cells by overexpression of Cbx2-Flag, GFP-Phc2 and 3Flag-Bmi1 leads to targeting of these three PRC1 components as well as of endogenously expressed Rnf2 to PCH in *Suv39h* dn but not wild-type ES cells (Chapter 2.2). We concluded that enrichment at PCH is not unique to early embryos, but can in principle occur in other cell types, if a certain set of PRC1 components is expressed in the absence of *Suv39h*-mediated blocking activity.

To identify whether a single component of PRC1 is sufficient to mediate heterochromatin binding, we overexpressed matPRC1 proteins separately in wild-type and *Suv39h* dn ES cells. Of the four matPRC1 components, only overexpressed Cbx2 (C-terminally FLAG-tagged) is strongly enriched at PCH specifically in *Suv39h* dn ES cells (Fig. 5A). Cbx2-Flag is able to recruit endogenously expressed Rnf2, Mel18 and Phc1 to heterochromatin, likely through its C-terminal Polycomb box (Pc box)<sup>61,63</sup>, whereas recruitment of other PRC1 is impaired when a Cbx2 construct with a C-terminal GFP tag is used (data not shown), probably due to steric hindrance. *In vitro* binding assays have shown that the Cbx2 chromodomain (CD) binds with similar affinities to peptides tri-methylated at H3K9 and H3K27<sup>62</sup>. In contrast, overexpressed Cbx2 in ES cells and endogenous matPRC1 in early embryos do not bind to PCH in the presence of H3K9me3. In mouse, five different chromodomain-containing homologs of the fly Pc protein exist, of which Cbx4 and Cbx7 also bind to H3K9me3 *in vitro* and in addition Cbx7 binds to H3K27me3<sup>62</sup>. In our ES cell overexpression system, none of the other Cbx homologs accumulates at PCH, neither in wild-type nor in *Suv39h* dn ES cells (Fig. 5B). Thus, although Cbx2 and Cbx7 have comparable affinities towards methylated lysine residues *in vitro*, only Cbx2 is targeted to PCH in *Suv39h* dn ES cells.



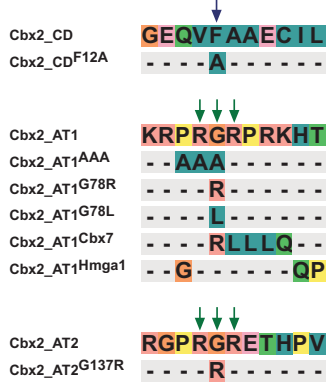
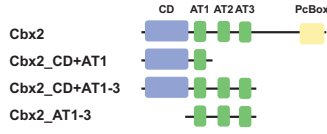
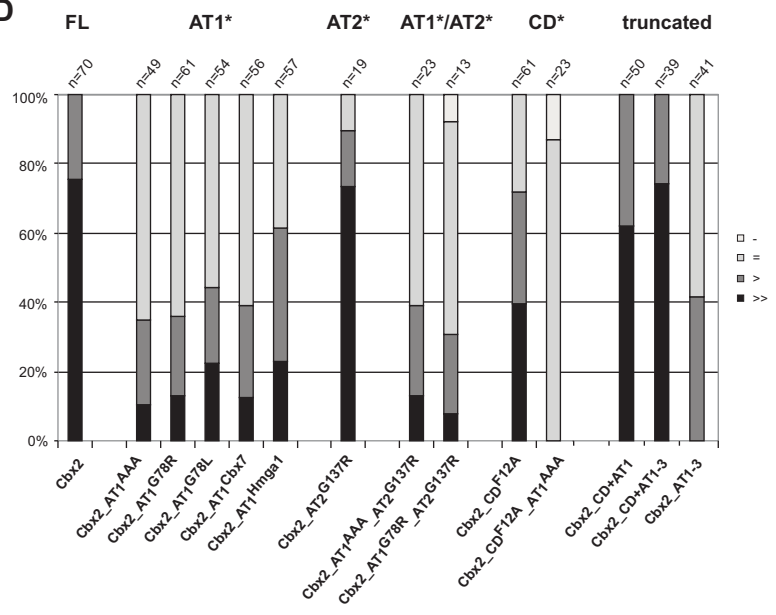
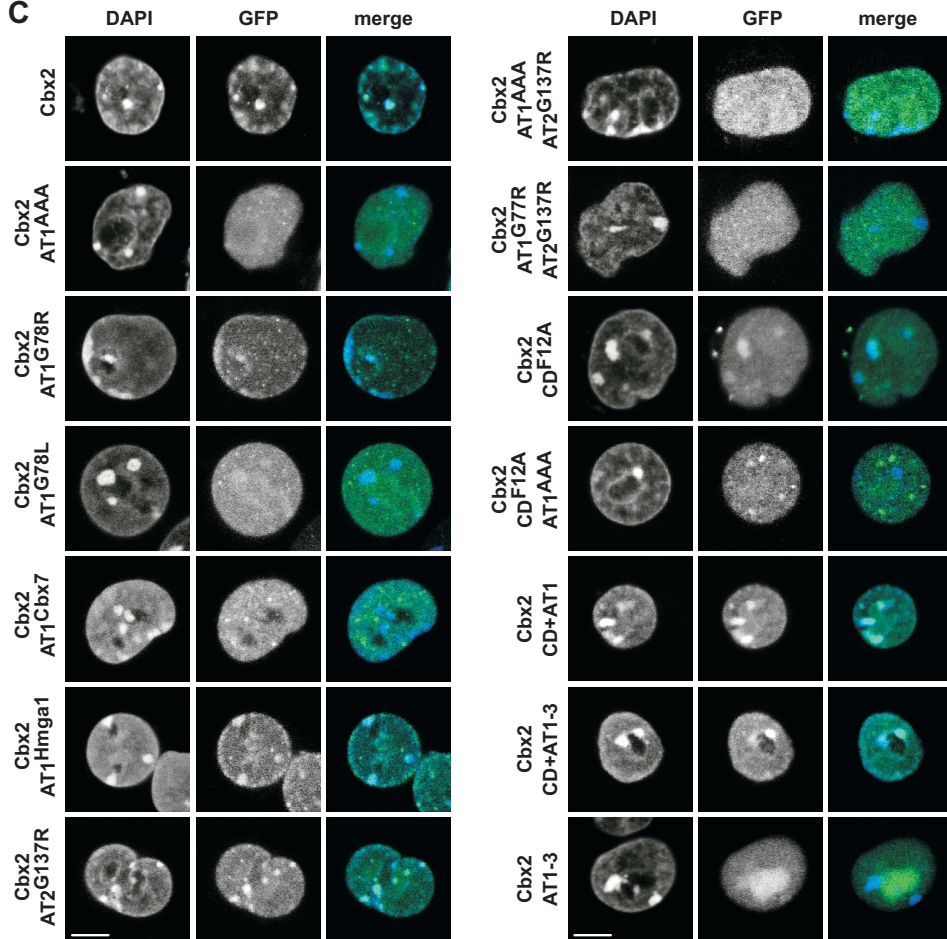
In Fig. 5C, an alignment of the first hundred amino acids of the five mouse Cbx homologs is shown, including the N-terminal chromodomain. The CD with the caging aromatic amino acids that mediate the histone methyl-lysine interaction is highly conserved<sup>61</sup>, suggesting that distinct binding properties are modulated by a few amino acid changes. Interestingly, in addition Cbx2 encodes three AT hook motifs immediately downstream of its CD that are not present in any of the other Cbx proteins but are conserved in several species including human, dog, chicken, zebrafish and *Xenopus* (excluding *Drosophila*)<sup>61</sup>. AT hooks are nine amino acid motifs that bind AT-rich stretches of DNA in the minor groove and have been extensively studied, especially in the high mobility group proteins (HMGs)<sup>337</sup>. HMGAs are architectural components that remodel the structure of chromatin and bind to AT-rich DNA without major preferences for the underlying DNA sequence. An alignment of the Cbx2 AT hooks with the ones from Hmga1 is shown in Fig. 5D. The solution structure of Hmga1 revealed that the central Arg-Gly-Arg (RGR) motif inserts into the minor groove of DNA, with neighboring residues at either side of the core mediating electrostatic and electrophobic interactions with the DNA backbone<sup>342</sup>. The second AT hook of Hmga1 is designated a "strong" AT hook as additional six amino acids C-terminal to the core form a more extensive interaction network with the DNA, resulting in significantly increased affinity<sup>342</sup>. The three Cbx2 AT hooks all have the central RGR core (Fig. 5D) but lack additional C-terminal stabilizing residues found in Hmga1 AT2. As mouse major satellites are strongly AT-rich, Cbx2 binding to PCH could be mediated by the Cbx2 AT hooks. It is interesting to note, that AT hook motifs are also found in the H3K4-specific histone methyltransferases MII, MII3 and MII4 (Fig. 5D)<sup>334,344</sup> that belong to the trithorax group (TrxG) of proteins and function antagonistically to PcG proteins.

#### *The Cbx2 chromodomain and AT hook confer heterochromatin binding*

To test the possibility that the Cbx2 AT hooks contribute to PCH targeting, we mutated three central amino acids of the Cbx2 AT hook 1 (AT1) to alanines (Cbx2\_AT1<sup>AAA</sup>, Fig. 6A) and overexpressed the GFP-tagged protein in *Suv39h* dn ES cells (Fig. 6C). In transfected cells, PCH enrichment was scored as follows: (-) GFP signal excluded from DAPI-dense heterochromatin domains; (=) equal GFP signal at hetero- and euchromatin; (>) enhanced and (>>) strongly enhanced GFP signal at heterochromatin versus euchromatin. Indeed, Cbx2\_AT1<sup>AAA</sup> shows strongly reduced enrichment at PCH compared to Cbx2 and in more than 60% of cells PCH accumulation is completely lost (Fig. 6D). To rule out the possibility that the

---

**Figure 5: The PRC1 component Cbx2 is recruited to constitutive heterochromatin in *Suv39h* dn ES cells.** (A) Separate overexpression of matPRC1 components in wild-type and *Suv39h* dn ES cells reveals that only Cbx2 is recruited to PCH specifically in *Suv39h* dn ES cells (arrows). (B) Overexpression of other Cbx homologs does not result in PCH enrichment in wild-type or *Suv39h* dn ES cells. Scale bars, 5  $\mu$ m. (C) ClustalX alignment of the first hundred amino acids of the five mouse Pc homologs. The blue line highlights the N-terminal chromodomain that is highly conserved between homologs. Blue arrows indicate the aromatic caging residues required for methyl-lysine binding. The Cbx2 AT hook 1 is marked by a green line, and green arrows indicate the central RGR core. (D) Alignment of the AT hook motifs found in the High mobility group protein Hmga1, the PRC1 protein Cbx2 and the Trithorax HMTs MII, MII3 and MII4.

**A****B****D****C**

Cbx2 structure is perturbed by the three amino acid changes, we also introduced single point mutations in the core of AT1. Mutation of the central small glycine residue to a large amino acid, either arginine (Cbx2\_AT1<sup>G78R</sup>) or leucine (Cbx2\_AT1<sup>G78L</sup>), results in a reduction in PCH enrichment to a similar extent as observed for Cbx2\_AT1<sup>AAA</sup> (Fig. 6C+D). PCH enrichment is also impaired by swapping the Cbx2 AT1 with amino acids found at the same position in Cbx7 (Cbx2\_AT1<sup>Cbx7</sup>). Moreover, exchanging one amino acid N-terminal of the core and two amino acids C-terminal of Cbx2 AT1 to match the Hmga1 AT1 (Fig. 6A) affects PCH enrichment (Fig. 6C+D), albeit less strongly than the other mutations, suggesting that the Cbx2 AT1 may confer some specificity that may not necessarily be mediated by all AT hooks.

Mutation of the first Cbx2 AT hook strongly impairs PCH binding but does not completely abrogate it. One possible explanation is that the second and third AT hook also contribute to heterochromatin targeting. Therefore, we introduced a point mutation in the central core of AT2 (Cbx2\_AT2<sup>G137R</sup>) in analogy to the glycine to arginine mutation of AT1 (Cbx2\_AT1<sup>G78R</sup>). In contrast to AT1, mutation of AT2 does not affect accumulation of Cbx2 at PCH (Fig. 6C+D). Likewise, mutation of both the first and second AT hook of Cbx2 does not further reduce PCH enrichment compared to a single AT1 mutation, suggesting that the second Cbx2 AT hook does not significantly contribute to targeting of Cbx2 to heterochromatin. Since H3K27me3 is enriched at PCH in approximately 30% of cells (data not shown), the chromodomain of Cbx2 might recognize this modification resulting in PCH targeting independent of a functional AT hook. Mutation of one of the caging residues known to be required for methyl-lysine binding (Cbx2\_CD<sup>F12A</sup>) results in reduced PCH enrichment (Fig. 6C+D), though less strongly than mutation of AT1. Strikingly, PCH is completely abolished when both the chromodomain and AT1 contain a point mutation.

In agreement with this, a truncated version of Cbx2 only containing the chromodomain and AT1 (Fig. 6B) strongly binds to PCH in *Suv39h* dn ES cells (Fig. 6C+D). Addition of the second and third AT hook does only slightly enhance heterochromatin enrichment. In contrast, a protein containing only the three AT hooks binds very weakly to PCH and instead accumulates within nucleoli. It is unclear whether this is a specific effect or presents an artifact caused by the large GFP-tag in relation to the short AT hook stretch. We will further test this using a construct in which the CD is deleted but which contains the C-terminus in addition to the AT hooks.

To sum up, both the mutation analyses as well as the truncation experiments suggest that heterochromatin binding of Cbx2 is mediated by a dual binding module consisting of the

---

**Figure 6: Cbx2 targeting to heterochromatin depends on its chromodomain and AT hook motif.** (A) Representation of the mutations introduced into the Cbx2 chromodomain and AT hooks. The blue arrow indicates the caging residue at position 12 of the chromodomain and green arrows highlight the central RGR core of AT hooks 1 and 2. (B) Schematic representation of Cbx2 truncation constructs. (C) Overexpression of Cbx2 mutation and truncation constructs in *Suv39h* dn ES cells. DAPI and direct GFP signals are shown for one representative cell per construct. Scale bar, 5  $\mu$ m. (D) Quantification of PCH enrichment for each construct, scored as follows: (-) GFP signal excluded from DAPI-dense heterochromatin domains; (=) equal GFP signal at hetero- and euchromatin; (>) enhanced and (>>) strongly enhanced GFP signal at heterochromatin versus euchromatin. CD, chromodomain; AT1, AT hook 1; PcBox, Polycomb box; FL, full length.

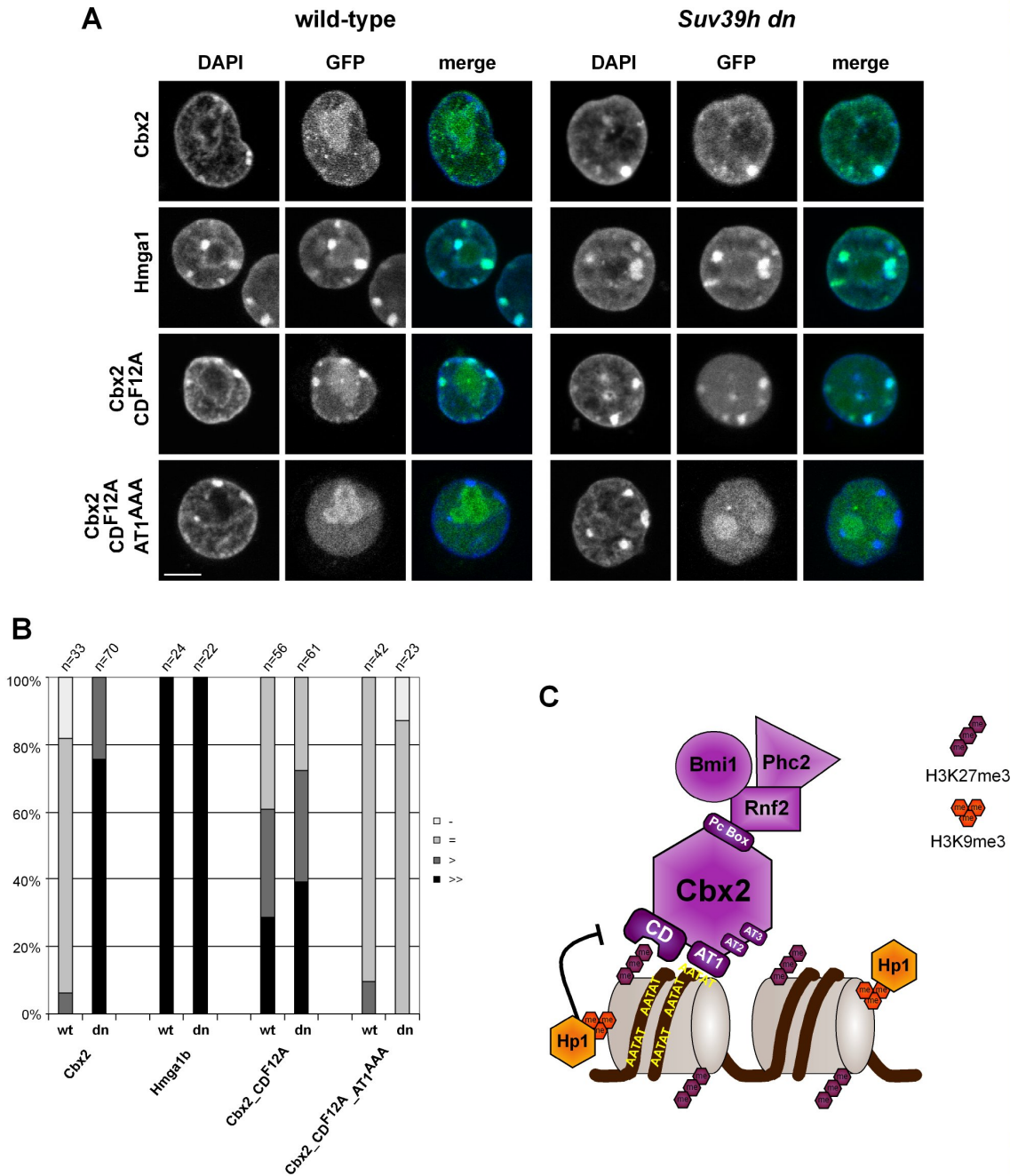
chromodomain and the first AT hook.

*Mutation of the Cbx2 chromodomain allows binding in the presence of H3K9me3*

Another question is, why does Cbx2 or matPRC1 only bind to heterochromatin in the absence of a functional *Suv39h* pathway? Possibly, access to the minor groove of the DNA is limited due to H3K9me3 and HP1 $\beta$  mediated compaction. We tested this hypothesis by overexpressing the AT hook protein Hmga1 in wild-type and *Suv39h* dn ES cells. Hmga1 is highly enriched at PCH, irrespective of the H3K9me3 status (Fig. 7A+B), arguing that inaccessibility alone is not prohibiting Cbx2 binding in wild-type cells. Alternatively, due to its strong second AT hook, Hmga1 may have a higher affinity towards pericentric major satellite sequences, allowing it to compete better with heterochromatin components.

Surprisingly, Cbx2 carrying a point mutation in the CD (Cbx2\_CD<sup>F12A</sup>) is able to bind PCH in wild-type ES cells with similar affinity as the same construct expressed in *Suv39h* dn ES cells (Fig. 7A+B). This is not caused by a change in the methyl-lysine affinity of the mutated CD toward H3K9me3 as Cbx7\_CD<sup>F11A</sup> does not accumulate at PCH in wild-type ES cells (data not shown). Instead, enrichment of Cbx2\_CD<sup>F12A</sup> to PCH in wild-type ES cells depends on the function of the first AT hook and is lost in a double CD and AT1 mutant construct (Cbx2\_CD<sup>F12A</sup>\_AT1<sup>AAA</sup>, Fig. 7A+B). Further experiments are needed to address the role of the Cbx2 chromodomain in "recognizing" the *Suv39h*-mediated blocking activity. For example, inserting an intact or mutated Cbx2 CD in front of the Hmga1 AT hook motifs would address whether the CD effect can be overruled by a high affinity AT hook. Further insights may also be gained by exchanging the Cbx2 CD with the CD of other Cbx proteins, thereby making use of the naturally provided variation between homologs. In case these constructs do not block PCH enrichment, we will make constructs of chimeric CDs to map the site required to recognize the blocking *Suv39h* activity.





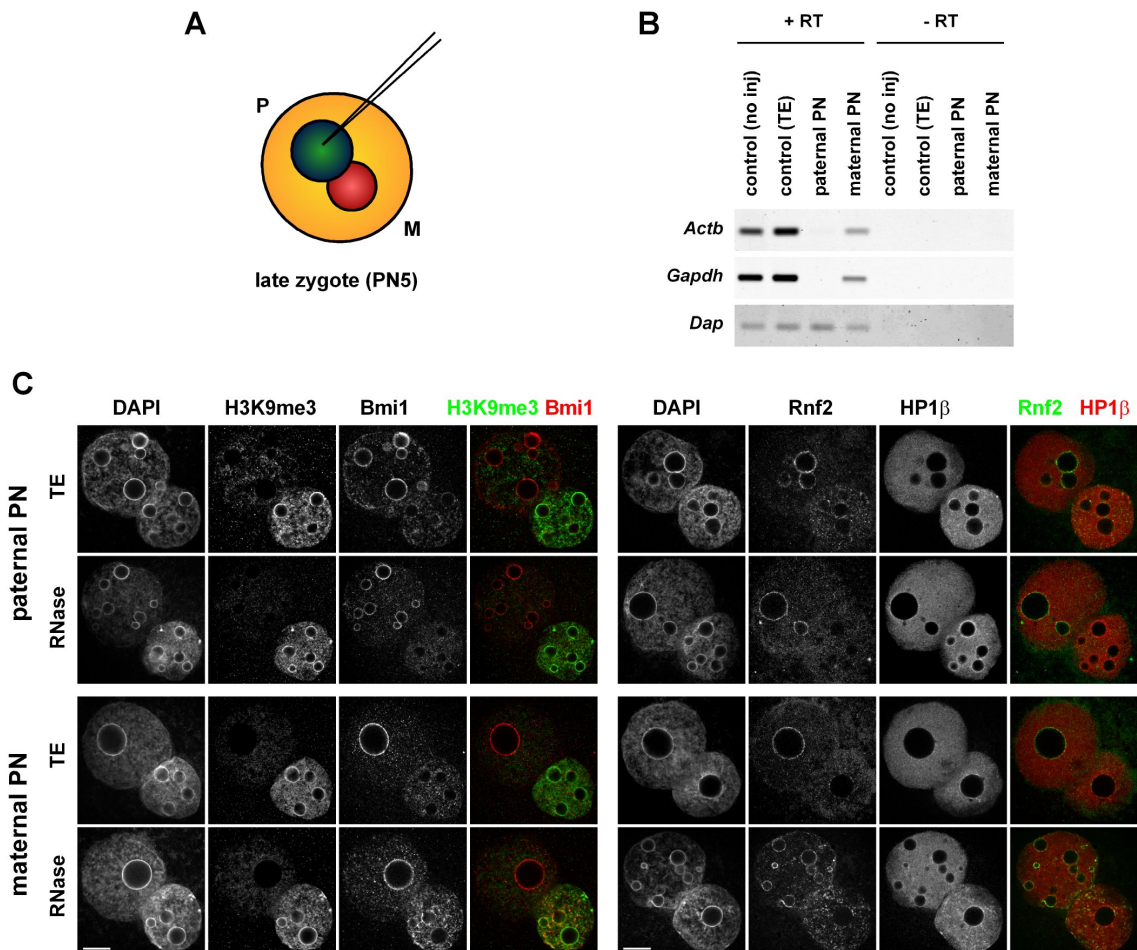
**Figure 7: Mutation of the Cbx2 chromodomain allows targeting to heterochromatin in wild-type ES cells.** (A) Overexpression of Cbx2 and Hmga1a constructs in wild-type and *Suv39h dn* ES cells. Whereas Cbx2 is only enriched at PCH in *Suv39h dn* ES, targeting of Hmga1 to PCH occurs irrespective of H3K9me3. Mutation of the Cbx2 CD allows binding to PCH also in wild-type ES cells, which is dependent on AT1. Scale bar, 5  $\mu$ m. (B) Quantification of PCH enrichments for each construct. (C) Model for the recruitment of PRC1 to constitutive heterochromatin via the Cbx2 chromodomain and AT hook.

In summary, our current data suggest that enrichment of Cbx2 at heterochromatin in *Suv39h dn* ES cells is mediated by the chromodomain, likely via binding to H3K27me3, and the first AT hook motif whereas AT hooks 2 and 3 seem to play minor roles (see model in Fig. 7C). Other PRC1 members are probably recruited to PCH through the C-terminal Pc Box of Cbx2.

How exactly the *Suv39h* pathway is blocking Cbx2 from PCH remains to be determined, but may involve the recognition of H3K9me3 by the Cbx2 chromodomain.

*PRC1 binding to heterochromatin is RNA-independent in zygotes*

In addition to marking of PCH by repressive histone methylation, it has been proposed that the methylated histone tails are arranged in a specific configuration dependent on a structural RNA which is required for the accumulation of HP1 proteins at chromocenters, as detected by immunofluorescence<sup>333</sup>. In analogy to this, we wondered whether targeting of matPRC1 proteins to paternal PCH in early embryos is also dependent on an RNA component. To test for a role of RNAs in PCH organization in early embryos, we adopted the RNase treatment previously used on cell lines<sup>333</sup> to early embryos. Unfortunately, the embryos did not survive the permeabilization procedure required before RNase digestion. Therefore, we decided to directly microinject RNase A either into the maternal or paternal pronucleus of the zygote (Fig. 8A).



**Figure 8: Binding of matPRC1 and HP1β to PCH is RNA-independent in zygotes.** (A) Schematic representation of the microinjection approach. (B) RT-PCR analysis of maternally provided housekeeping messages in control and RNase microinjected zygotes. *Actb* and *Gapdh* levels are decreased following RNase treatment, whereas a spike control (*Dap*) is unchanged. (C) Immunofluorescence analyses of H3K9me3/HP1β and Bmi1/Rnf2 reveals no difference between PCH localization in control versus RNase treated zygotes, suggesting that HP1β and PRC1 binding to PCH does not require RNA in zygotes. Scale bars, 10 μm.



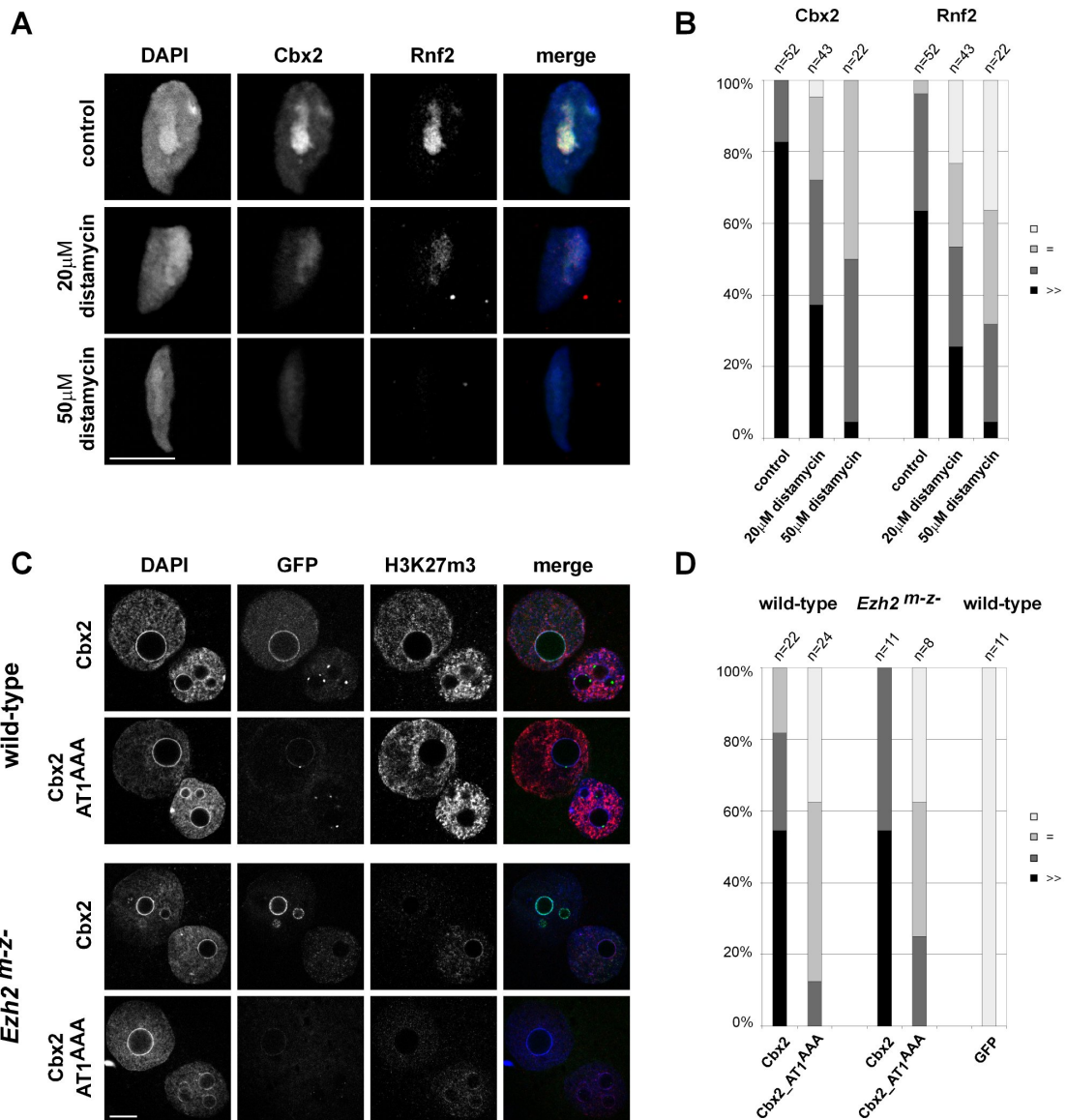
Efficiency of the treatment was confirmed by the isolation of total RNA from RNase and control injected zygotes. RT-PCR analysis shows that maternally provided mRNAs for house keeping genes are readily detected in control embryos, but are strongly reduced after RNase microinjection into both the maternal or paternal pronucleus (Fig. 8B). This was not due to inefficient RNA isolation or cDNA synthesis as an RNA spike control (*Dap*) added to the lysis buffer before RNA isolation is equally amplified in all samples. Following microinjection of RNase, zygotes were cultured for 15 minutes before fixation. IF for maternal PCH marks reveals no change in the distribution of H3K9me3 or HP1 $\beta$  at PCH rings (Fig. 8C). Similarly, PCH enrichment of matPRC1 components Bmi1 and Rnf2 is unaffected after RNase microinjection into the paternal pronucleus.

These results suggest that maternal PCH rings may be organized differently from PCH clusters found in somatic cells with HP1 $\beta$  accumulating independent of an RNA-mediated structure. Moreover, matPRC1 enrichment at paternal PCH seems to be RNA-independent although we cannot exclude that an RNA component may be involved in initial matPRC1 targeting.

#### *The Cbx2 AT hook contributes to chromatin binding in zygotes*

Alternatively, matPRC1 proteins may be recruited to paternal PCH by direct binding to the underlying AT-rich DNA via the AT hook motifs present in Cbx2. This hypothesis can be approached by using natural or synthetic molecules that bind specifically to the minor groove of AT-rich DNA and therefore will compete with endogenous proteins for major satellite binding. One such molecule is the naturally occurring antibiotic distamycin that has been isolated from *Streptomyces distallicus*<sup>345</sup>. This oligo-peptide interacts reversibly with the DNA minor groove, recognizing sequences containing at least four AT base pairs, and shows limited cytotoxicity. We incubated oocytes with distamycin during *in vitro* fertilization (IVF) to see whether matPRC1 proteins are initially targeted to PCH via satellite DNA. Indeed, treatment with distamycin results in a dose-dependent decrease in matPRC1 enrichment (Fig. 9A). In control zygotes, Cbx2 is highly enriched at the central DAPI-dense heterochromatin domain in more than 80% of decondensing sperm heads (Fig. 9B). After IVF in 20 $\mu$ M distamycin, less than 40% of decondensing sperm show high enrichment for Cbx2 at PCH and incubation with 50 $\mu$ M distamycin almost completely abolishes high affinity binding to PCH. Enrichment of Rnf2 at paternal PCH is similarly compromised. Note that in addition to paternal heterochromatin, matPRC1 levels are also decreased in the surrounding euchromatic regions, suggesting that the AT dependent binding mode may not be exclusively required for heterochromatin only.

Drug treatment may have unspecific side effects and therefore we wished to address PCH targeting mechanisms more directly. Making use of the Cbx2 constructs tested for PCH enrichment in ES cells, we *in vitro* transcribed Cbx2-GFP and microinjected the polyadenylated message into the cytoplasm of early zygotes. Following microinjection, the zygotes were further cultured for five hours to allow translation and targeting of the overexpressed protein to occur. The zygotes were then fixed and processed for IF. Overexpressed Cbx2-GFP accumulates specifically at paternal PCH and is excluded from maternal PCH (Fig. 9C). Like endogenous



**Figure 9: matPRC1 binding to paternal constitutive heterochromatin involves the Cbx2 AT hook in zygotes.** (A) Immunofluorescence analyses of PN0 zygotes reveals enrichment of Cbx2 and Rnf2 at PCH of decondensing sperm. PCH enrichment is strongly reduced following treatment with the DNA minor groove binder distamycin. (B) Quantification of PCH binding in decondensing sperm, using the follow criteria: (-) no staining; (=) weak staining at heterochromatin; (>) strong and (>>) very strong enrichment at heterochromatin. (C) Immunofluorescence analyses of late zygotes microinjected with mRNA of GFP-tagged Cbx2 or Cbx2\_AT1<sup>AAA</sup>. Cbx2-GFP is enriched at paternal PCH in wild-type and *Ezh2*<sup>m-z-</sup> zygotes, like endogenous Cbx2. Mutation of Cbx2 AT1 leads to reduced binding to PCH and euchromatin. (D) Quantification of binding to paternal PCH in late zygotes. Scale bars, 10 μm.

Cbx2, it is also present in both maternal and paternal euchromatin. Additional highly fluorescent Cbx2-GFP foci are detected around the NPBs of both maternal and paternal pronuclei that we have not noticed during IF of endogenous Cbx2. They may represent minor satellites or rDNA, however, this is purely speculative and the nature of these foci remains to be determined. Microinjection of Cbx2-GFP into *Ezh2<sup>m-z-</sup>* zygotes results in paternal PCH enrichment independent of H3K27me3. In contrast, like endogenous matPRC1, targeting of Cbx2-GFP to euchromatin regions correlates with H3K27me3 levels. Thus, overexpressed Cbx2-GFP mimics the localization of endogenous Cbx2 at PCH and within euchromatin in both wild-type and *Ezh2<sup>m-z-</sup>* zygotes, although interaction with other PRC1 members is probably inhibited through the C-terminal GFP-tag (see experiments in ES cells), suggesting independent recruitment.

To address the role of the Cbx2 AT hook in H3K27me3-independent targeting to paternal PCH, we microinjected Cbx2 message with a mutation in the first AT hook (Cbx2\_AT1<sup>AAA</sup>) into zygotes. Like in *Suv39h* dn ES cells, PCH enrichment of Cbx2\_AT1<sup>AAA</sup> is reduced, both in wild-type and *Ezh2<sup>m-z-</sup>* zygotes (Fig. 9B). Unexpectedly, Cbx2\_AT1<sup>AAA</sup> accumulation was also lost within euchromatin. As the bright foci around NPBs were still detected, loss of PCH and euchromatin accumulation is probably not due to protein degradation, although should be confirmed using an independent method. Further microinjection experiment using additional mutation and truncation constructs are required to dissect the functional contributions of the Cbx2 CD and AT hook to chromatin binding.

Together our preliminary distamycin and microinjection experiments suggest that the Cbx2 AT hook may also be involved in PCH targeting of the matPRC1 complex in early embryos. The observation that binding of Cbx2 is also affected within euchromatin upon distamycin treatment or mutation of the AT hook suggested to us that the dual CD-AT-module may play a role also in targeting of PRC1 proteins to regions outside of constitutive heterochromatin.

#### *Does the Cbx2 AT hook contribute to the recognition of PRC1 target genes?*

So far no Polycomb response elements (PREs) have been identified in mammals, suggesting that they are possibly not characterized by specific, well defined sequence motifs like in *Drosophila*<sup>108</sup>. If our hypothesis is true, PRC1 proteins may be recruited to their target genes specifically by H3K27me3 via the CD of Cbx2 and additional affinity could be contributed by AT-rich sequence stretches recognized by the Cbx2 AT hooks. To further test this idea, we used a bioinformatics approach based on our results and the published data on AT hook proteins to analyze sequence characteristics of Polycomb target genes.

A number of biochemical and structural studies have analyzed AT hook binding properties and addressed the question whether AT hooks bind sequence specific or recognize any random stretch of AT-rich DNA. Early qualitative HMGA footprinting studies observed AT hook binding to any run of 5-6 AT bp with similar affinity, suggesting that HMGAs only recognize the minor groove of AT-rich DNA rather than a specific sequence<sup>346,347</sup>. More recent studies suggest that high affinity binding of TAF1 or HMGA requires two to three appropriately spaced AT-rich sequences to allow multivalent binding by neighboring AT hooks<sup>342,347,348</sup>. Examples of well characterized AT hook binding sites include the *interferon-β* (*IFN-β*) enhancer, the promoter of

*interleukin-2* and *interleukin-15*, and the *heat shock protein 70 (Hsp70)* transcription start site<sup>347-350</sup>. Each of these regulatory sequences contains at least one 5 bp AT stretch, required for AT hook binding. One study used a SELEX (systematic evolution of ligands by exponential enrichment) approach to identify sequences bound by HMGA2. Common features included two 5-6 bp AT-stretches separated by a 4-5 bp GC-rich linker, and surprisingly an AT consensus sequence (ATATT) was identified<sup>347</sup>.

Taken together, the consensus from the literature seems to be that high affinity binding by AT hook proteins requires appropriately spaced AT stretches to allow multivalent binding. Although there are a few examples hinting towards some sequence specificity, in general any AT-stretch seems to be sufficient. If Cbx2 is recruited to heterochromatin directly by DNA, these features should also be found within major satellites. The sequence of the 234 bp repeat unit of mouse major satellites, divided into four sub-repeats<sup>1</sup>, is shown in Fig. 10A. The sequence contains a number of different 5-7 bp AT motifs with the spacers varying in length from 3 to 14 bp, providing a range of options for AT hook binding. The required spacing may vary between AT hook proteins depending on their amino acid (aa) sequence and 3D structure. In Hmga1, AT1 and AT2 are separated by 22 aa, and a 10 aa linker is present between AT2 and AT3. The linkers are much longer in Cbx2 with 50 aa separating AT1 and AT2 (including long serine stretches), and a 20 aa linker between AT2 and AT3. However, our mutagenesis study in *Suv39h* dn ES cells does not support a major functional contribution of Cbx2 AT hooks 2 and 3. Therefore, our current knowledge does not allow us to define a specific sequence motif recognized by Cbx2, but rather suggests including all possible AT motifs with a minimum lengths of 5 bp into a bioinformatics search.

Next, we wanted to analyze whether AT motifs would be enriched in PcG target genes. Using the PRC1 and PRC2 ChIP on Chip data generated by Boyer and colleagues in mouse ES cells<sup>2</sup>, we defined a number of positive data sets containing Polycomb target genes (Table in Fig. 10B). The data set Pos 6 contains all sequences enriched in PRC1 (Phc1 and/or Rnf2) and PRC2 (Suz12 and/or Eed) within 10 kb (peaks only). The set Pos 8 contains sequences enriched in PRC2 that do not have any PRC1 peak within 10 kb, whereas Pos 4 contains sequences enriched in PRC1 without a PRC2 peak within 10 kb. For statistical analysis especially sets Pos 8 (PRC2 only) and Pos 4 (PRC1 only) are interesting as they contain a similar number of regions (580 and 516, respectively) spanning approximately 400,000 bp each. In addition, we defined a set Pos 2 that includes 2 kb around the transcription start sites (TSS) of all genes enriched in PRC1 (Phc1 and/or Rnf2) that lack PRC2 (Suz12, Eed and H3K27me3). Finally, we generated a negative data set (Neg 1) that includes 2 kb around the TSS of genes that are not targeted by Polycomb in ES cells (negative for PRC1, PRC2 and H3K27me3 based on the ChIP data by Boyer<sup>2</sup> and Mikkelsen<sup>3</sup>). It is important to note that the genes included in the negative set, may however be targeted by Polycomb proteins in other cell types and therefore may have the 'sequence characteristics' of a Polycomb gene. Moreover, AT hook motifs are present in a variety of DNA binding proteins with a wide range of functions<sup>334</sup>, suggesting that genes other than those targeted by Polycomb are expected to contain AT-rich sequences.

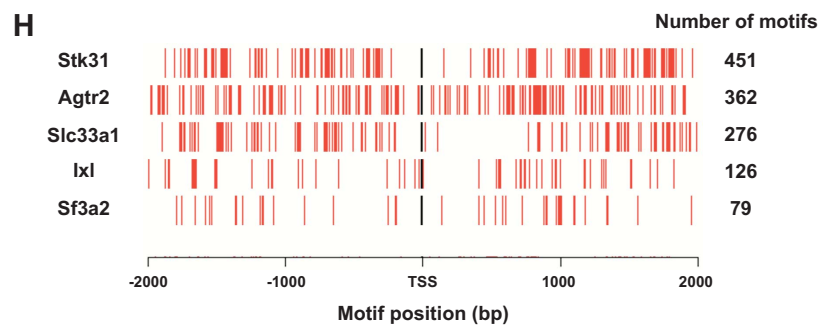
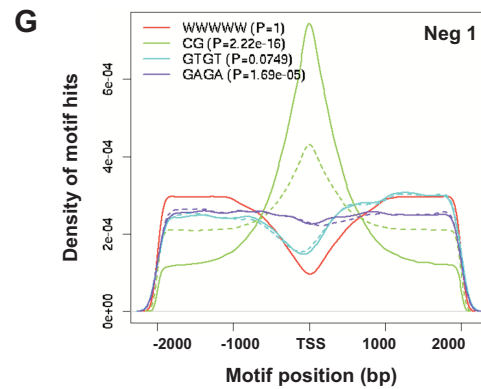
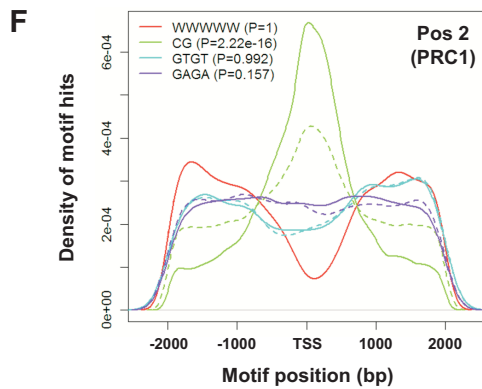
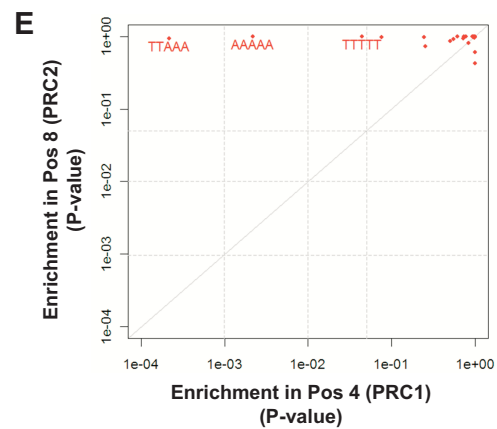
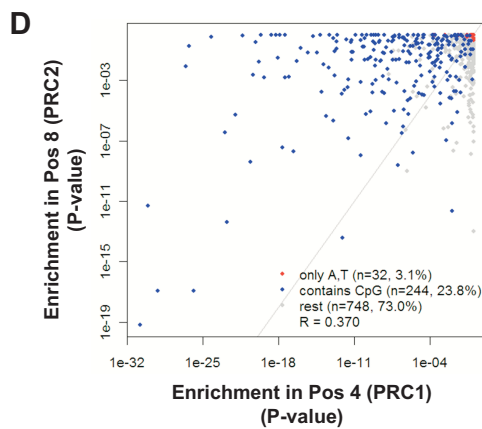
**A** Sub-repeat I GGACCTGG **AAATAT**GGCGAGAAAAGT**GGAAAAT**CACGG**AAAAT**GAG**AAATA**CACACTTTA  
 Sub-repeat II GGACGTG **AAATAT**GGCGAGGAAAAGT**GGAAAAT**TTAGAAATGTCCACTGTA  
 Sub-repeat III GGACGTGG **AAATAT**GGCAAGAAAAGT**GGAAAAT**CATGG**AAAAT**GAGAAACATCCACTTGA  
 Sub-repeat IV CGACTTG **AAAAT**GACGAAATCACT**AAAAAA**CGTG**AAAAT**GAGAAATGCACACTGAA

**B**

Name	PRC1	PRC2	Region	Number	Size
Pos 6	+	+	Peaks	1,242	2,159,571 bp
Pos 8	-	+	Peaks	580	422,799 bp
Pos 4	+	-	Peaks	516	357,882 bp
Pos 2	+	-	TSS +/- 2 kb	291	1,164,291 bp
Neg 1	-	-	TSS +/- 2 kb	10,389	41,566,389 bp

**C**

Name	Shuffled	Neg 1
Pos 6	1.59	1.03
Pos 8	1.37	0.85
Pos 4	1.35	0.92
Pos 2	1.52	1.09
Neg 1	1.46	1.00



We used an unbiased bioinformatics approach to test whether any 5mer sequence motif is enriched in the defined positive data sets either relative to a randomly 'shuffled' sequence (taking GC-richness into account) or relative to the negative set. This approach yielded a number of motifs that were significantly enriched in our data sets, including AT 5mers (data not shown). To test whether all possible AT 5mers are specifically enriched in any of the positive sets, we calculate the ratios between observed occurrences of AT 5mers versus occurrence of AT 5mers expected from a shuffled background or from the negative set. The observed/expected ratios for all AT 5mer motifs (n=32) are presented in Fig. 10C. Comparison of Pos 2 (PRC1) with Neg 1, which both include 2 kb around the TSS, indicates enrichment of AT 5mers in both sets relative to a shuffled background (no enrichment = 1.0), suggesting that AT-richness may be a general feature of TSS regions. Direct comparison of Pos 2 (PRC1) with Neg 1 shows that AT 5mers, although enriched in both sets, are more abundant within PRC1 target genes. As comparison of PRC1 peaks (Pos 4) and PRC2 peaks (Pos 8) yielded similar observed/expected ratios, we wanted to know whether the same or different 5mer motifs were enriched in these sets. Therefore, the P-values of all possible 5mers were visualized in a scatter blot between Pos 4 (PRC1) with Pos 8 (PRC2) (Fig. 10D), showing low correlation between both sets (R=0.37). Blotting the enrichments for all AT 5mers in Pos 4 (PRC1) versus Pos 8 (PRC2) (Fig. 10E) reveals that some AT motifs are significantly enriched in PRC1 target regions but none in PRC2 regions. The motifs enriched in PRC1 include TTAAA, AAAAA and TTTTT.

Furthermore, we were interested in the distribution of the AT 5mers in relation the TSS. For the PRC1 positive set (Pos 2), AT 5mers (WWWWW, according to IUPAC code for A or T) are present upstream and downstream of the TSS, but seem to be excluded from the TSS itself (Fig. 10F). In contrast, CpG containing 5mers are specifically enriched around the TSS. These motif distributions do however not seem to be specific for PRC1 genes as similar distributions are also found for the negative set (Fig. 10G), suggesting that they may instead represent general features of promoter regions. A few examples of AT motifs around the TSS of randomly selected PRC1 genes (Pos 2) are presented in Fig. 10H, with highly AT-rich genes at the top.

To sum up, our preliminary bioinformatics analysis indicates that AT motifs may indeed be enriched in a number of PRC1 target genes. Our current analysis is hampered by the fact that no specific ChIP data for Cbx2 itself is available yet. We are currently generating antibodies that will hopefully allow us to do such analysis. Meanwhile, we are planning to use the generated bioinformatics tools and data to select a number of candidate genes to test our hypothesis.

---

**Figure 10: Bioinformatics analysis of AT binding motifs within PRC1 and PRC2 target genes.** (A) Sequence of mouse major satellite repeats, divided into sub-repeats<sup>1</sup>. AT stretches of more than 5 bp are highlighted in green. (B) List of data sets generated from the ChIP on Chip analysis by Boyer<sup>2</sup> and Mikkelsen<sup>3</sup>. (C) Observed/expected ratios of AT 5mers motifs for each data set, either relative to a 'shuffled' background or relative to the negative set (Neg 1). (D) Scatter blot comparing enrichments (P-values) of all possible 5mer motifs between sets Pos 4 (PRC1) and Pos 8 (PRC2), showing a low overall correlation. (E) Scatter blot comparing enrichments (P-values) of all AT 5mer motifs between sets Pos 4 (PRC1) and Pos 8 (PRC2), revealing significant enrichment of three AT motifs in PRC1 target genes. (F+G) Distribution of motifs around the transcription start site (TSS) in Pos 2 (F) and Neg 1 (G). (H) Distribution of AT motifs around the TSS in five randomly selected PRC1 target genes (Pos 2). The number of AT 5mer motifs present within 2 kb around the TSS is indicated on the right.

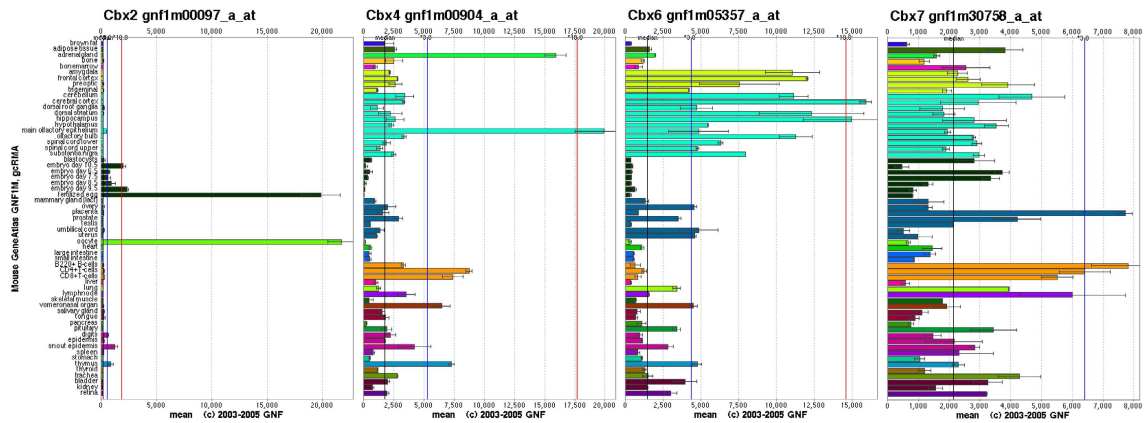
These genes will be selected from each of the three positive sets (PRC1 pos, PRC2 pos, and PRC1+PRC2 pos) based on their H3K27me3 and H3K9me3 status and the frequency of AT 5mers around the TSS. In addition, we will take into account possible conservation of AT-richness between species in which the Cbx2 AT hook is conserved (e.g. human, dog, chicken, zebrafish, *Xenopus* and *C. elegans*)<sup>61</sup>. This is based on the assumption that if the dual targeting mechanism is conserved in other species we would expect homologous genes to be enriched in AT motifs. Depending on the outcome of our candidate gene study, we may expand our analysis to look genome-wide for PRC1 target genes that rely on either the Cbx2 CD or AT hook or both.

#### 2.3.4. Discussion

Using a heterologous ES cell system, we could show that matPRC1 binding to constitutive heterochromatin is not limited to early embryos. We used this *in vitro* system to dissect which PRC1 members and domains contribute to heterochromatin recruitment. Separate overexpression of PRC1 members revealed that Cbx2 is sufficient for PCH targeting. Interestingly, in early embryos Cbx2 is maternally contributed but barely expressed following zygotic genome activation (Chapter 2.2, Fig. 1d). Consequently, Cbx2 levels gradually decline during pre-implantation development and at the 8-cell stage, once *Suv39h*-mediated marks have been acquired at paternal PCH, Cbx2 is hardly detectable any more (Chapter 2.2, Suppl. Fig. 2). This is in agreement with expression data available from the Genomics Institute of the Novartis Research Foundation (GNF), San Diego, detecting high levels of Cbx2 only in oocytes and 1-cell embryos and low expression in post-implantation embryos (E6.5 to E10.5) (Fig. 11). Thus, it seems that resolution of parental epigenetic asymmetry at heterochromatin and euchromatin at the 8-cell stage coincides with downregulation of Cbx2, the major determinant for matPRC1 binding to heterochromatin. During early pre-implantation development none of the other Cbx members are detectable, suggesting that Cbx2 could be involved in the recruitment of PRC1 to genes that are regulated by PcG during this developmental stage. Other Cbx members may take over during later development (Fig. 11), possibly targeting different sets of genes.

Site directed mutagenesis and truncation experiments of Cbx2 in ES cells identified the Cbx2 chromodomain and the AT hook 1 as important modules required for PCH targeting. The CD likely recognizes H3K27me3 present at PCH in a percentage of *Suv39h* dn ES cells and at paternal PCH in early embryos, although H3K27me3 is not essential for PCH recruitment in early embryos (Chapter 2.2). H3K27me3 and PRC2 independent targeting has also been observed for Cbx8 at some genes in embryonic bodies<sup>81</sup>. Similarly, Rnf2 binds independent of H3K27me3 to the inactive X chromosome during ES cell differentiation in *Eed* mutant cells<sup>80</sup>. In both cases, the mechanism of recruitment remains to be determined. Noticeably, Phc1 and Phc2 are not recruited to the inactive X in *Eed* mutant cells, suggesting that they may be part of a different complex. The general consensus from the literature seems to be that H3K27me3 can promote binding of PC/Cbx but other mechanisms contribute to the targeting of PRC1 complexes *in vivo* (Chapter 1.2). In *Drosophila*, it has been proposed that PC binding to





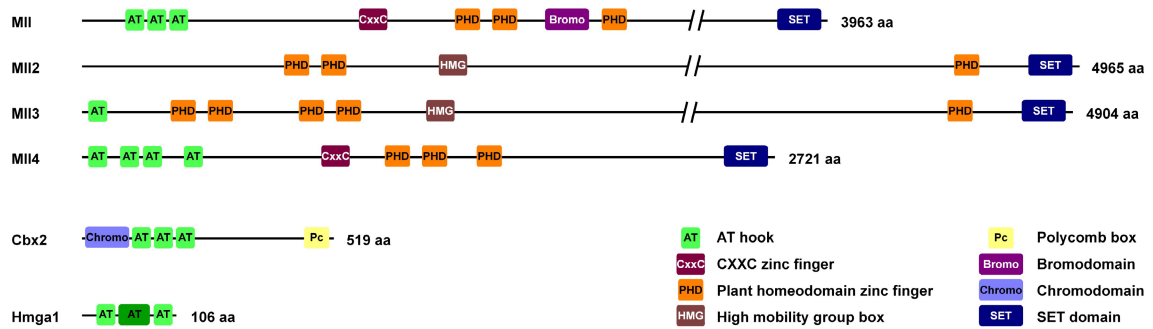
**Figure 11: Expression patterns of the mouse Cbx genes.** Expression data available from the Genomics Institute of the Novartis Research Foundation (GNF), San Diego, shows that high Cbx2 expression is restricted to oocytes and early embryos. In contrast, Cbx4, Cbx6 and Cbx7 are not expressed during that time but are present in various different adult cell types. The expression profile of Cbx8 was not available.

H3K27me3 peaks at putative PREs involves the recruitment by sequence specific binding proteins <sup>45</sup>. In contrast, weaker PC binding in regions surrounding PREs may be mediated by the CD alone, recognizing the H3K27me3 mark. We have identified one additional module in mammals, namely the Cbx2 AT hook 1, which is essential for heterochromatin recruitment of PRC1 in ES cells and likely early embryos.

In contrast to HP1 detection at chromocenters in somatic cells which depends on a structural RNA <sup>333</sup>, PRC1 binding to heterochromatin in early embryos seems to be RNA independent. This is agreement with *in vitro* studies showing that the Cbx4, Cbx6, Cbx7 and Cbx8 chromodomains all bind to RNA but not the one of Cbx2 <sup>62</sup>. Instead the Cbx2 AT hook might mediate direct binding to the underlying AT-rich major satellite DNA. The involvement of a structural RNA has also been assessed for higher order chromosome organization of the *Drosophila* bithorax complex <sup>121</sup>. Similar to our studies of PRC1 binding to PCH, RNase treatment did not affect PRE–PRE or PRE–promoter interactions.

In both early embryos and ES cells, recruitment of PRC1 proteins to PCH is blocked by *Suv39h*-mediated marks. Mutation of the Cbx2 CD, however, allows binding to PCH also in wild-type ES cells. One possibility is that Cbx2 binding is blocked directly by H3K9 methylation, supported by fact that the mutation of one of the caging residues in the Cbx2 CD, which is involved in methyl-lysine binding, abrogates *Suv39h*-mediated blocking activity. Alternatively, inhibition of binding might be mediated by HP1 proteins bound to H3K9me3 or by downstream chromatin marks like H4K20me2/3 or DNA methylation. However, DNA demethylation of the paternal genome in early embryos takes place after the initial targeting of PRC1 proteins to PCH, suggesting that DNA methylation is probably not blocking binding of PRC1. Further studies will be aimed at dissecting *Suv39h*-mediated inhibition of PCH recruitment.

The current literature on other AT hook containing proteins indicates that this motif recognizes AT-rich DNA with relatively little sequence specificity. In addition to heterochromatin recruitment, the Cbx2 AT hook could also be envisioned as a motif contributing additional



**Figure 12: Domain structure of chromatin associated AT hook proteins.** The mammalian Trithorax group HMTs MII, MII2 and MII4 contain one to four AT hook motifs at their N-terminus. The three AT hook motifs of Cbx2 are located immediately C-terminal of the chromodomain. Hmga1 encodes three AT hook motifs, of which the second one (dark green) is designated a strong AT hook due to additional stabilizing residues. Not drawn to scale.

affinity for PRC1 binding at target genes. The close vicinity of the Cbx2 CD and AT hook may allow Cbx2 to simultaneously scan the DNA sequence and the methylation marks on the nucleosome around which the DNA is wrapped (Model in Fig. 7C). Our preliminary bioinformatics study suggests that AT 5mer motifs may indeed be enriched in regions bound by PRC1 proteins in ES cells. AT-rich stretches per se might be sufficient, instead of a well-defined consensus sequence, which could explain why no such consensus sequences have been found so far in mammals. Interestingly, the H3K4me specific demethylase RBP2, known to interact with PRC2<sup>88</sup>, also contains a DNA binding domain, the AT-rich interaction domain (ARID), which however prefers binding to GC-rich sequences<sup>351</sup>.

Such DNA binding modules could be more commonly used for targeting of chromatin proteins than so far appreciated, which is supported by the presence of AT hook motifs in the H3K4me specific HMTs of the Trithorax group (Fig. 12). MII and MII4 encode three to four AT hook motifs at their N-terminus which in contrast to Cbx2 are distant from any other functional domain. Both MII and MII4 are huge proteins that contain a number of different functional modules in addition to the C-terminal SET domain. The CXXC motif is involved in the recognition of non-methylated CpG dinucleotides and PHD fingers promote protein-protein interactions but also have been recently found to be capable of binding to methylated H3K4 residues<sup>352,353</sup>. In addition, a bromodomain, specific for acetylated histones, is present in MII. The combination of these different motifs might allow the Trithorax HMTs to select their target genes and to control for their transcriptional status. In the case of the Trithorax HMTs, the AT hook motifs might be an evolutionary old invention as they are also found in the *Drosophila* SET domain protein ASH1<sup>334</sup>. Interestingly, MII2 and MII3, which have no or just one AT hook, instead encode a HMG box, also found in HMGB proteins where this domain penetrates into the minor groove of DNA and sharply bends it<sup>335,337</sup>. While the TrxG HMTs and the PcG protein Cbx2 combine several functional motifs in one protein, HMGA1 only contains the three AT hooks, suggesting that HMGA1 mediates its various functions through interactions with other proteins. Interestingly, HMGA1 is involved in the activation of zygotic transcription in the early embryo<sup>340</sup>, suggesting that AT hook proteins may play a role during that developmental stage.

It is worth noticing that the AT hook is not conserved in the *Drosophila* Cbx homolog (PC)<sup>61</sup>, suggesting that the Cbx2 AT hook might be a recent addition during evolution. Unlike in mammals, a number of DNA binding proteins have been identified to interact with PcG complexes in *Drosophila* and are involved in the recruitment of PcG to their target genes<sup>108</sup>. One of them, DSP1, is an ortholog of the mammalian HMGB proteins and contains two HMG boxes. Taken together, both fly and mammalian TrxG and PcG proteins encode DNA binding domains like the AT hook or HMG box, which might be involved in the targeting of TrxG/PcG proteins and in the case of the AT hook might provide additional specificity towards AT-rich sequence stretches.

So far, we have focused on PRC1 targeting mechanisms. But it is also unclear how PRC2 mediated H3K27me3 becomes enriched at heterochromatin in early embryos. Strikingly, PRC1 proteins are present at heterochromatin before H3K27me3 becomes detectable. Is it possible that PRC1 proteins recruit the PRC2 complex? Transient interactions between PRC2 (ESC) and PRC1 (PC) are required for the establishment of PcG silencing in early *Drosophila* embryos<sup>107</sup>. Moreover, it has been speculated in the literature that a dimer of PRC1 may function during replication to tether the PRC2 complex to the daughter strand in order to propagate the H3K27me3 mark<sup>354</sup>. Does a similar mechanism lead to H3K27me3 enrichment at paternal heterochromatin in early embryos? Analysis of maternal *Rnf2* deficient zygotes argues against such a strategy, as H3K27me3 is enriched at paternal PCH in late zygotes lacking the matPRC1 complex (Chapter 2.2). Thus, it seems that at heterochromatin PRC2 and PRC1 are targeted independently, which however ultimately leads to the co-localization of both complexes, possibly reinforcing efficient binding.

To sum up, we have identified the Cbx2 AT hook as novel motif contributing to PRC1 targeting to constitutive heterochromatin. We speculate that the Cbx2 chromodomain and AT hook could function as a dual module for PRC1 recruitment also at target genes. Further studies will be aimed at the identification of Cbx2 targets *in vivo*, followed by the dissection of mechanisms contributing to target gene recognition.

### 3. General conclusions

In this study, we aimed at dissecting role of Polycomb group complexes during pre-implantation mouse development, in particular during zygotic genome activation and first lineage commitment. We uncovered an unanticipated role of PRC1 proteins in pericentric heterochromatin formation of the paternal genome following fertilization. While the maternal genome is characterized by canonical heterochromatin marks of the *Suv39h*-pathway, paternal chromatin is initially devoid of these modifications. Studies in *S. pombe*, *Drosophila* and mammals have shown that *Suv39h*-mediated pericentric heterochromatin labeling is essential for repression of centromeric repeats, for proper chromosome segregation and genome stability<sup>9,12,26,355-357</sup>. How then is it possible that the paternal genome can do apparently fine for several cleavage divisions in the absence of these essential modifications?

Our studies suggest that maternally provided PRC1 functions as the default repressive back-up mechanism in the absence of *Suv39h*-mediated marks up to the 8-cell stage. Zygotes deficient for the *Suv39h2* HMT lack repressive H3K9me3 and HP1 $\beta$  at pericentric heterochromatin, and consequently PRC1 is also targeted to maternal heterochromatin in these zygotes. In zygotes maternally deficient for *Rnf2*, an essential component of PRC1, major satellite transcription is upregulated in the paternal but not maternal pronucleus, suggesting that PRC1 functions to repress centromeric repeat sequences. It remains to be determined whether PRC1 is also required for proper segregation of chromosomes during the first embryonic divisions. Interestingly, severe defects in sister chromatid segregation have been observed in *Drosophila* embryos from *Polycomb*, *Posterior sex combs* and *Additional sex combs* heterozygous mutant females<sup>358</sup>.

PcG proteins have been identified as part of a cellular memory system that maintains the expression pattern of key developmental regulators. In *Drosophila*, Polycomb proteins bind to Polycomb response elements (PREs) which contain consensus sequences for several known DNA binding proteins. In mammals, no such sequences have been identified to date and PcG targeting mechanisms remain poorly understood. Our analysis of early embryos revealed that targeting of PRC1 to heterochromatin can occur in the absence of *Ezh2* and H3K27me3, and is blocked by the presence of *Suv39h*-mediated marks. We used a heterologous ES cell system to dissect recruitment mechanisms of PRC1 to pericentric heterochromatin. We show that the PRC1 protein Cbx2 is essential for PCH recruitment, which is mediated by a dual module consisting of the Cbx2 chromodomain specific for H3K27me3 and an AT hook motif, recognizing AT-rich sequence stretches. Competition and mutation experiments in zygotes suggest that the Cbx2 AT hook is also functioning in heterochromatin targeting of Cbx2 in early embryos. It is tempting to speculate that the dual module of Cbx2 chromodomain and AT hook could, in addition, be involved in the recruitment of PRC1 to target genes.

Interestingly, asymmetric labeling of maternal chromatin by H3K9me3 and HP1 $\beta$  versus labeling of paternal chromatin by PRC1 is not restricted to pericentric regions. Analysis of cleavage embryos revealed that paternal chromosome arms are strongly labeled by Rnf2 in a banded fashion, whereas maternal chromosomes are enriched in H3K9me3 along the arms.

Similar to asymmetric labeling observed at heterochromatin, differential marking of the chromosome arms is resolved after the 8-cell stage. What are the implications of the transient epigenetic asymmetry for gene regulation during pre-implantation development? Does it affect the expression of genes during zygotic genome activation in the early embryo? It is interesting to note that asymmetry is resolved at the 8-cell stage, concomitant with a number of developmental changes including blastomere compaction and immediately prior to initiation of first asymmetric cell divisions, giving rise to the precursors of the embryonic and extra-embryonic lineages. The dynamics might imply that resolution of epigenetic asymmetry between parental genomes is a prerequisite to allow first lineage choices to occur.

Embryos maternally and zygotically deficient for *Ezh2* are morphologically normal during pre-implantation stages. *Ezh2*-deficient blastocyst embryos contain both cell lineages, the inner cell mass (ICM) and the trophectoderm (TE), which are correctly labeled by the lineage markers Oct4 and Cdx2, respectively (data not shown). However, expression profiling of *Ezh2* mutant blastocysts identified approximately 400 genes that are more than 1.5-fold changed (data not shown), of which some might have important roles during further development leading to the arrest at E6.5. These results suggest that asymmetric labeling of H3K27me3 in the ICM versus TE<sup>136</sup> might not be instructive for lineage segregation, but rather fine tunes gene expression in the respective lineages. This is in agreement with several studies of PcG mutant cells that often result in small expression changes<sup>2,125</sup>, even so such small changes in the expression of key developmental regulators may significantly affect developmental outcomes<sup>280,281</sup>. Possibly, the role of H3K27me3 only becomes apparent during later stages of development, when cell fate decisions need to be “remembered”. As part of an epigenetic memory system, H3K27me3 may not be required for the initial choices themselves.

Similarly, maternally and zygotically *Rnf2* mutant embryos appear morphologically normal during pre-implantation development (data not shown). It remains to be analyzed whether such mutants display defects in chromatid segregation during cleavage divisions, which in humans is remarkable irregular<sup>303</sup>, and which could lead to aneuploidy. *Rnf2* function during pre-implantation development might be compensated by its homolog *Ring1a*. Strikingly, maternal double deficiency for *Ring1a* and *Rnf2* leads to developmental arrest at the 2- to 4-cell stage (M.P. and E. Posfai, unpublished). The causes of this developmental failure remain to be dissected, but could be due to a maternal effect, possibly resulting from misregulation of gene expression during oocyte growth. Further studies are needed to dissect the roles of PcG proteins in the maternal germ line and during pre-implantation development.

Recent research showing that somatic cells can be reprogrammed into embryonic stem cell-like cells, so called induced pluripotent stem (iPS) cells, has challenged our view on the role of the epigenetic system. A number of groups have now demonstrated that adding just a handful of transcription factors can turn both mouse and human somatic cells into iPS cells<sup>359-364</sup>. Does this mean that chromatin states characteristic of a differentiated cell type can be easily overruled, or are not important at all? Yet, similar to nuclear transfer and cloning experiments, reprogramming of somatic cells into iPS cells is not a very efficient process<sup>365</sup>. While transcription factors might be able to induce reprogramming, the complete resetting of the

epigenetic state might hold the key to successful reprogramming. Interestingly, several of the genes that are targeted by the transcription factors used to induce pluripotency (including Oct4, Sox2 and Nanog) are also bound by PcG proteins<sup>127</sup>.

Taken together the emerging view seems that the epigenetic system has evolved to fine tune gene expression rather than being instructive for lineage choices. Our studies have added to our understanding of epigenetic 'programming' of mouse pre-implantation embryos, particularly with respect to the *de novo* establishment of chromatin states on the paternal genome following fertilization. Clearly, a lot remains to be learned and future research dissecting the role of different flavors of chromatin at paternal and maternal genomes will be very exciting.

## References

1. Lehnertz, B. et al. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol* **13**, 1192-200 (2003).
2. Boyer, L.A. et al. Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441**, 349-53 (2006).
3. Mikkelsen, T.S. et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553-60 (2007).
4. Heitz, E. Das Heterochromatin der Moose. *Jahrb. Wiss. Botanik* **69**, 762-818 (1928).
5. Wong, A.K. & Rattner, J.B. Sequence organization and cytological localization of the minor satellite of mouse. *Nucleic Acids Res* **16**, 11645-61 (1988).
6. Vissel, B. & Choo, K.H. Mouse major (gamma) satellite DNA is highly conserved and organized into extremely long tandem arrays: implications for recombination between nonhomologous chromosomes. *Genomics* **5**, 407-14 (1989).
7. Guenatri, M., Bailly, D., Maison, C. & Almouzni, G. Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. *J Cell Biol* **166**, 493-505 (2004).
8. Bloom, K. Centromere dynamics. *Curr Opin Genet Dev* **17**, 151-6 (2007).
9. Peters, A.H. et al. Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107**, 323-37 (2001).
10. Bannister, A.J. et al. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120-4 (2001).
11. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. & Jenuwein, T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116-20 (2001).
12. Maison, C. & Almouzni, G. HP1 and the dynamics of heterochromatin maintenance. *Nat Rev Mol Cell Biol* **5**, 296-304 (2004).
13. Fanti, L. & Pimpinelli, S. HP1: a functionally multifaceted protein. *Curr Opin Genet Dev* (2008).
14. Jacobs, S.A. & Khorasanizadeh, S. Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science* **295**, 2080-3 (2002).
15. Fischle, W. et al. Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes Dev* **17**, 1870-81 (2003).
16. Muchardt, C. et al. Coordinated methyl and RNA binding is required for heterochromatin localization of mammalian HP1alpha. *EMBO Rep* **3**, 975-81 (2002).
17. Lomberk, G., Bensi, D., Fernandez-Zapico, M.E. & Urrutia, R. Evidence for the existence of an HP1-mediated subcode within the histone code. *Nat Cell Biol* **8**, 407-15 (2006).
18. Vaquero, A. et al. SIRT1 regulates the histone methyl-transferase SUV39H1 during heterochromatin formation. *Nature* **450**, 440-4 (2007).
19. Cheutin, T. et al. Maintenance of stable heterochromatin domains by dynamic HP1 binding. *Science* **299**, 721-5 (2003).
20. Festenstein, R. et al. Modulation of heterochromatin protein 1 dynamics in primary Mammalian cells. *Science* **299**, 719-21 (2003).
21. Fischle, W. et al. Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* **438**, 1116-22 (2005).
22. Hirota, T., Lipp, J.J., Toh, B.H. & Peters, J.M. Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. *Nature* **438**, 1176-80 (2005).
23. Fuks, F., Hurd, P.J., Deplus, R. & Kouzarides, T. The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Res* **31**, 2305-12 (2003).



24. Schotta, G. et al. A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev* **18**, 1251-62 (2004).
25. Hall, I.M. et al. Establishment and maintenance of a heterochromatin domain. *Science* **297**, 2232-7 (2002).
26. Grewal, S.I. & Jia, S. Heterochromatin revisited. *Nat Rev Genet* **8**, 35-46 (2007).
27. Buhler, M. & Moazed, D. Transcription and RNAi in heterochromatic gene silencing. *Nat Struct Mol Biol* **14**, 1041-1048 (2007).
28. Grewal, S.I. & Elgin, S.C. Transcription and RNA interference in the formation of heterochromatin. *Nature* **447**, 399-406 (2007).
29. Verdel, A. et al. RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* **303**, 672-6 (2004).
30. Volpe, T.A. et al. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297**, 1833-7 (2002).
31. Folco, H.D., Pidoux, A.L., Urano, T. & Allshire, R.C. Heterochromatin and RNAi are required to establish CENP-A chromatin at centromeres. *Science* **319**, 94-7 (2008).
32. Lehmann, E., Brueckner, F. & Cramer, P. Molecular basis of RNA-dependent RNA polymerase II activity. *Nature* **450**, 445-9 (2007).
33. Chen, E.S. et al. Cell cycle control of centromeric repeat transcription and heterochromatin assembly. *Nature* **451**, 734-7 (2008).
34. Fukagawa, T. et al. Dicer is essential for formation of the heterochromatin structure in vertebrate cells. *Nat Cell Biol* **6**, 784-91 (2004).
35. Kanellopoulou, C. et al. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev* **19**, 489-501 (2005).
36. Pal-Bhadra, M. et al. Heterochromatic silencing and HP1 localization in *Drosophila* are dependent on the RNAi machinery. *Science* **303**, 669-72 (2004).
37. van der Lugt, N.M. et al. Posterior transformation, neurological abnormalities, and severe hematopoietic defects in mice with a targeted deletion of the *bmi-1* proto-oncogene. *Genes Dev* **8**, 757-69 (1994).
38. del Mar Lorente, M. et al. Loss- and gain-of-function mutations show a polycomb group function for Ring1A in mice. *Development* **127**, 5093-100 (2000).
39. Core, N. et al. Altered cellular proliferation and mesoderm patterning in Polycomb-M33-deficient mice. *Development* **124**, 721-9 (1997).
40. Akasaka, T. et al. A role for *mel-18*, a Polycomb group-related vertebrate gene, during theanteroposterior specification of the axial skeleton. *Development* **122**, 1513-22 (1996).
41. Cao, R. et al. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298**, 1039-43 (2002).
42. Czermin, B. et al. *Drosophila* enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell* **111**, 185-96 (2002).
43. Muller, J. et al. Histone methyltransferase activity of a *Drosophila* Polycomb group repressor complex. *Cell* **111**, 197-208 (2002).
44. Ebert, A. et al. *Su(var)* genes regulate the balance between euchromatin and heterochromatin in *Drosophila*. *Genes Dev* **18**, 2973-83 (2004).
45. Schwartz, Y.B. et al. Genome-wide analysis of Polycomb targets in *Drosophila melanogaster*. *Nat Genet* **38**, 700-5 (2006).
46. Nekrasov, M. et al. Pcl-PRC2 is needed to generate high levels of H3-K27 trimethylation at Polycomb target genes. *EMBO J* **26**, 4078-88 (2007).
47. Savla, U., Benes, J., Zhang, J. & Jones, R.S. Recruitment of *Drosophila* Polycomb-group proteins by Polycomblike, a component of a novel protein complex in larvae. *Development* **135**, 813-7 (2008).

48. Pasini, D., Bracken, A.P., Jensen, M.R., Lazzerini Denchi, E. & Helin, K. Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *Embo J* **23**, 4061-71 (2004).
49. Kuzmichev, A., Jenuwein, T., Tempst, P. & Reinberg, D. Different EZH2-containing complexes target methylation of histone H1 or nucleosomal histone H3. *Mol Cell* **14**, 183-93 (2004).
50. Kuzmichev, A. et al. Composition and histone substrates of polycomb repressive group complexes change during cellular differentiation. *Proc Natl Acad Sci U S A* **102**, 1859-64 (2005).
51. Han, Z. et al. Structural basis of EZH2 recognition by EED. *Structure* **15**, 1306-15 (2007).
52. Montgomery, N.D. et al. The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation. *Curr Biol* **15**, 942-7 (2005).
53. Barski, A. et al. High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823-37 (2007).
54. Sarma, K., Margueron, R., Ivanov, A., Pirrotta, V. & Reinberg, D. Ezh2 requires PHF1 to efficiently catalyze H3 lysine 27 tri-methylation in vivo. *Mol Cell Biol* (2008).
55. Vire, E. et al. The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **439**, 871-4 (2006).
56. Min, J., Zhang, Y. & Xu, R.M. Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. *Genes Dev* **17**, 1823-8 (2003).
57. Saurin, A.J., Shao, Z., Erdjument-Bromage, H., Tempst, P. & Kingston, R.E. A Drosophila Polycomb group complex includes Zeste and dTAFII proteins. *Nature* **412**, 655-60 (2001).
58. Shao, Z. et al. Stabilization of chromatin structure by PRC1, a Polycomb complex. *Cell* **98**, 37-46 (1999).
59. Levine, S.S. et al. The core of the polycomb repressive complex is compositionally and functionally conserved in flies and humans. *Mol Cell Biol* **22**, 6070-8 (2002).
60. Otte, A.P. & Kwaks, T.H. Gene repression by Polycomb group protein complexes: a distinct complex for every occasion? *Curr Opin Genet Dev* **13**, 448-54 (2003).
61. Whitcomb, S.J., Basu, A., Allis, C.D. & Bernstein, E. Polycomb Group proteins: an evolutionary perspective. *Trends Genet* **23**, 494-502 (2007).
62. Bernstein, E. et al. Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. *Mol Cell Biol* **26**, 2560-9 (2006).
63. Garcia, E., Marcos-Gutierrez, C., del Mar Lorente, M., Moreno, J.C. & Vidal, M. RYBP, a new repressor protein that interacts with components of the mammalian Polycomb complex, and with the transcription factor YY1. *EMBO J* **18**, 3404-18 (1999).
64. Cao, R., Tsukada, Y. & Zhang, Y. Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. *Mol Cell* **20**, 845-54 (2005).
65. de Napoles, M. et al. Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev Cell* **7**, 663-76 (2004).
66. Wang, H. et al. Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431**, 873-8 (2004).
67. Buchwald, G. et al. Structure and E3-ligase activity of the Ring-Ring complex of polycomb proteins Bmi1 and Ring1b. *EMBO J* **25**, 2465-74 (2006).
68. Elderkin, S. et al. A phosphorylated form of Mel-18 targets the Ring1B histone H2A ubiquitin ligase to chromatin. *Mol Cell* **28**, 107-20 (2007).
69. Li, Z. et al. Structure of a Bmi-1-Ring1B polycomb group ubiquitin ligase complex. *J Biol Chem* **281**, 20643-9 (2006).
70. Ben-Saadon, R., Zaaroor, D., Ziv, T. & Ciechanover, A. The polycomb protein Ring1B generates self atypical mixed ubiquitin chains required for its in vitro histone H2A ligase activity. *Mol Cell* **24**, 701-11 (2006).
71. Wang, L. et al. Hierarchical recruitment of polycomb group silencing complexes. *Mol Cell* **14**, 637-46 (2004).
72. Francis, N.J., Saurin, A.J., Shao, Z. & Kingston, R.E. Reconstitution of a functional core polycomb repressive complex. *Mol Cell* **8**, 545-56 (2001).

73. King, I.F., Francis, N.J. & Kingston, R.E. Native and recombinant polycomb group complexes establish a selective block to template accessibility to repress transcription in vitro. *Mol Cell Biol* **22**, 7919-28 (2002).
74. Francis, N.J., Kingston, R.E. & Woodcock, C.L. Chromatin compaction by a polycomb group protein complex. *Science* **306**, 1574-7 (2004).
75. Schwartz, Y.B. & Pirrotta, V. Polycomb silencing mechanisms and the management of genomic programmes. *Nat Rev Genet* **8**, 9-22 (2007).
76. Ficiz, G., Heintzmann, R. & Arndt-Jovin, D.J. Polycomb group protein complexes exchange rapidly in living *Drosophila*. *Development* **132**, 3963-76 (2005).
77. Ren, X., Vincenz, C. & Kerppola, T.K. Changes in the Distributions and Dynamics of Polycomb Repressive Complexes during Embryonic Stem Cell Differentiation. *Mol Cell Biol* (2008).
78. Dellino, G.I. et al. Polycomb silencing blocks transcription initiation. *Mol Cell* **13**, 887-93 (2004).
79. Stock, J.K. et al. Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. *Nat Cell Biol* **9**, 1428-35 (2007).
80. Schoeftner, S. et al. Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. *Embo J* **25**, 3110-22 (2006).
81. Pasini, D., Bracken, A.P., Hansen, J.B., Capillo, M. & Helin, K. The polycomb group protein Suz12 is required for embryonic stem cell differentiation. *Mol Cell Biol* **27**, 3769-79 (2007).
82. Ehrenhofer-Murray, A.E. Chromatin dynamics at DNA replication, transcription and repair. *Eur J Biochem* **271**, 2335-49 (2004).
83. Huang, J. et al. Repression of p53 activity by Smyd2-mediated methylation. *Nature* **444**, 629-32 (2006).
84. Chuikov, S. et al. Regulation of p53 activity through lysine methylation. *Nature* **432**, 353-60 (2004).
85. Shi, X. et al. Modulation of p53 function by SET8-mediated methylation at lysine 382. *Mol Cell* **27**, 636-46 (2007).
86. Huang, J. & Berger, S.L. The emerging field of dynamic lysine methylation of non-histone proteins. *Curr Opin Genet Dev* (2008).
87. Fritsch, C., Beuchle, D. & Muller, J. Molecular and genetic analysis of the Polycomb group gene *Sex combs extra/Ring* in *Drosophila*. *Mech Dev* **120**, 949-54 (2003).
88. Agger, K., Christensen, J., Cloos, P.A. & Helin, K. The emerging functions of histone demethylases. *Curr Opin Genet Dev* (2008).
89. De Santa, F. et al. The histone H3 lysine-27 demethylase Jmjd3 links inflammation to inhibition of polycomb-mediated gene silencing. *Cell* **130**, 1083-94 (2007).
90. Agger, K. et al. UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* **449**, 731-4 (2007).
91. Lan, F. et al. A histone H3 lysine 27 demethylase regulates animal posterior development. *Nature* **449**, 689-94 (2007).
92. Lee, M.G. et al. Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. *Science* **318**, 447-50 (2007).
93. Joo, H.Y. et al. Regulation of cell cycle progression and gene expression by H2A deubiquitination. *Nature* **449**, 1068-72 (2007).
94. Nakagawa, T. et al. Deubiquitylation of histone H2A activates transcriptional initiation via trans-histone cross-talk with H3K4 di- and trimethylation. *Genes Dev* **22**, 37-49 (2008).
95. Nicassio, F. et al. Human USP3 is a chromatin modifier required for S phase progression and genome stability. *Curr Biol* **17**, 1972-7 (2007).
96. Zhu, P. et al. A histone H2A deubiquitinase complex coordinating histone acetylation and H1 dissociation in transcriptional regulation. *Mol Cell* **27**, 609-21 (2007).

97. Cho, Y.W. et al. PTIP associates with MLL3- and MLL4-containing histone H3 lysine 4 methyltransferase complex. *J Biol Chem* **282**, 20395-406 (2007).
98. Issaeva, I. et al. Knockdown of ALR (MLL2) reveals ALR target genes and leads to alterations in cell adhesion and growth. *Mol Cell Biol* **27**, 1889-903 (2007).
99. Lee, M.G., Norman, J., Shilatifard, A. & Shiekhattar, R. Physical and functional association of a trimethyl H3K4 demethylase and Ring6a/MBLR, a polycomb-like protein. *Cell* **128**, 877-87 (2007).
100. Klebes, A. et al. Regulation of cellular plasticity in Drosophila imaginal disc cells by the Polycomb group, trithorax group and lama genes. *Development* **132**, 3753-65 (2005).
101. Chen, X., Hiller, M., Sancak, Y. & Fuller, M.T. Tissue-specific TAFs counteract Polycomb to turn on terminal differentiation. *Science* **310**, 869-72 (2005).
102. Lee, N., Maurange, C., Ringrose, L. & Paro, R. Suppression of Polycomb group proteins by JNK signalling induces transdetermination in Drosophila imaginal discs. *Nature* **438**, 234-7 (2005).
103. Brown, J.L., Fritsch, C., Mueller, J. & Kassis, J.A. The Drosophila pho-like gene encodes a YY1-related DNA binding protein that is redundant with pleiohomeotic in homeotic gene silencing. *Development* **130**, 285-94 (2003).
104. Brown, J.L., Mucci, D., Whiteley, M., Dirksen, M.L. & Kassis, J.A. The Drosophila Polycomb group gene pleiohomeotic encodes a DNA binding protein with homology to the transcription factor YY1. *Mol Cell* **1**, 1057-64 (1998).
105. Fritsch, C., Brown, J.L., Kassis, J.A. & Muller, J. The DNA-binding polycomb group protein pleiohomeotic mediates silencing of a Drosophila homeotic gene. *Development* **126**, 3905-13 (1999).
106. Klymenko, T. et al. A Polycomb group protein complex with sequence-specific DNA-binding and selective methyl-lysine-binding activities. *Genes Dev* **20**, 1110-22 (2006).
107. Poux, S., Melfi, R. & Pirrotta, V. Establishment of Polycomb silencing requires a transient interaction between PC and ESC. *Genes Dev* **15**, 2509-14 (2001).
108. Ringrose, L. & Paro, R. Polycomb/Trithorax response elements and epigenetic memory of cell identity. *Development* **134**, 223-32 (2007).
109. Hodgson, J.W., Argiropoulos, B. & Brock, H.W. Site-specific recognition of a 70-base-pair element containing d(GA)(n) repeats mediates bithoraxoid polycomb group response element-dependent silencing. *Mol Cell Biol* **21**, 4528-43 (2001).
110. Horard, B., Tatout, C., Poux, S. & Pirrotta, V. Structure of a polycomb response element and in vitro binding of polycomb group complexes containing GAGA factor. *Mol Cell Biol* **20**, 3187-97 (2000).
111. Brown, J.L., Grau, D.J., DeVido, S.K. & Kassis, J.A. An Sp1/KLF binding site is important for the activity of a Polycomb group response element from the Drosophila engrailed gene. *Nucleic Acids Res* **33**, 5181-9 (2005).
112. Dejardin, J. et al. Recruitment of Drosophila Polycomb group proteins to chromatin by DSP1. *Nature* **434**, 533-8 (2005).
113. Ringrose, L., Rehmsmeier, M., Dura, J.M. & Paro, R. Genome-wide prediction of Polycomb/Trithorax response elements in Drosophila melanogaster. *Dev Cell* **5**, 759-71 (2003).
114. Negre, N. et al. Chromosomal distribution of PcG proteins during Drosophila development. *PLoS Biol* **4**, e170 (2006).
115. Tolhuis, B. et al. Genome-wide profiling of PRC1 and PRC2 Polycomb chromatin binding in Drosophila melanogaster. *Nat Genet* **38**, 694-9 (2006).
116. Ringrose, L., Ehret, H. & Paro, R. Distinct contributions of histone H3 lysine 9 and 27 methylation to locus-specific stability of polycomb complexes. *Mol Cell* **16**, 641-53 (2004).
117. Kahn, T.G., Schwartz, Y.B., Dellino, G.I. & Pirrotta, V. Polycomb complexes and the propagation of the methylation mark at the Drosophila *ubx* gene. *J Biol Chem* **281**, 29064-75 (2006).
118. Papp, B. & Muller, J. Histone trimethylation and the maintenance of transcriptional ON and OFF states by *trxG* and PcG proteins. *Genes Dev* **20**, 2041-54 (2006).
119. Mohd-Sarip, A. et al. Architecture of a polycomb nucleoprotein complex. *Mol Cell* **24**, 91-100 (2006).

120. Cleard, F., Moshkin, Y., Karch, F. & Maeda, R.K. Probing long-distance regulatory interactions in the *Drosophila melanogaster* bithorax complex using Dam identification. *Nat Genet* **38**, 931-5 (2006).
121. Lanzaolo, C., Roure, V., Dekker, J., Bantignies, F. & Orlando, V. Polycomb response elements mediate the formation of chromosome higher-order structures in the bithorax complex. *Nat Cell Biol* **9**, 1167-74 (2007).
122. Bantignies, F., Grimaud, C., Lavrov, S., Gabut, M. & Cavalli, G. Inheritance of Polycomb-dependent chromosomal interactions in *Drosophila*. *Genes Dev* **17**, 2406-20 (2003).
123. Vazquez, J., Muller, M., Pirrotta, V. & Sedat, J.W. The Mcp element mediates stable long-range chromosome-chromosome interactions in *Drosophila*. *Mol Biol Cell* **17**, 2158-65 (2006).
124. Grimaud, C. et al. RNAi components are required for nuclear clustering of Polycomb group response elements. *Cell* **124**, 957-71 (2006).
125. Bracken, A.P., Dietrich, N., Pasini, D., Hansen, K.H. & Helin, K. Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes Dev* **20**, 1123-36 (2006).
126. Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R. & Young, R.A. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130**, 77-88 (2007).
127. Lee, T.I. et al. Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* **125**, 301-13 (2006).
128. Tanay, A., O'Donnell, A.H., Damelin, M. & Bestor, T.H. Hyperconserved CpG domains underlie Polycomb-binding sites. *Proc Natl Acad Sci U S A* **104**, 5521-6 (2007).
129. Woolfe, A. et al. Highly conserved non-coding sequences are associated with vertebrate development. *PLoS Biol* **3**, e7 (2005).
130. Rinn, J.L. et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129**, 1311-23 (2007).
131. Sessa, L. et al. Noncoding RNA synthesis and loss of Polycomb group repression accompanies the colinear activation of the human HOXA cluster. *RNA* **13**, 223-39 (2007).
132. Schmitt, S., Prestel, M. & Paro, R. Intergenic transcription through a polycomb group response element counteracts silencing. *Genes Dev* **19**, 697-708 (2005).
133. Faust, C., Schumacher, A., Holdener, B. & Magnuson, T. The eed mutation disrupts anterior mesoderm production in mice. *Development* **121**, 273-85 (1995).
134. Niswander, L., Yee, D., Rinchik, E.M., Russell, L.B. & Magnuson, T. The albino deletion complex and early postimplantation survival in the mouse. *Development* **102**, 45-53 (1988).
135. O'Carroll, D. et al. The polycomb-group gene Ezh2 is required for early mouse development. *Mol Cell Biol* **21**, 4330-6 (2001).
136. Erhardt, S. et al. Consequences of the depletion of zygotic and embryonic enhancer of zeste 2 during preimplantation mouse development. *Development* **130**, 4235-48 (2003).
137. Voncken, J.W. et al. Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. *Proc Natl Acad Sci U S A* **100**, 2468-73 (2003).
138. Fujimura, Y. et al. Distinct roles of Polycomb group gene products in transcriptionally repressed and active domains of Hoxb8. *Development* **133**, 2371-81 (2006).
139. Lessard, J. & Sauvageau, G. Bmi-1 determines the proliferative capacity of normal and leukaemic stem cells. *Nature* **423**, 255-60 (2003).
140. Lessard, J. et al. Functional antagonism of the Polycomb-Group genes eed and Bmi1 in hemopoietic cell proliferation. *Genes Dev* **13**, 2691-703 (1999).
141. Liu, S. et al. Hedgehog signaling and Bmi-1 regulate self-renewal of normal and malignant human mammary stem cells. *Cancer Res* **66**, 6063-71 (2006).
142. Molofsky, A.V. et al. Bmi-1 dependence distinguishes neural stem cell self-renewal from progenitor proliferation. *Nature* **425**, 962-7 (2003).

143. Park, I.K. et al. Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells. *Nature* **423**, 302-5 (2003).
144. Bernstein, B.E. et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315-26 (2006).
145. Pietersen, A.M. & van Lohuizen, M. Stem cell regulation by polycomb repressors: postponing commitment. *Curr Opin Cell Biol* (2008).
146. Azuara, V. et al. Chromatin signatures of pluripotent cell lines. *Nat Cell Biol* **8**, 532-8 (2006).
147. Pan, G. et al. Whole-Genome Analysis of Histone H3 Lysine 4 and Lysine 27 Methylation in Human Embryonic Stem Cells. *Cell Stem Cell* **1**, 299-312 (2007).
148. Zhao, X.D. et al. Whole-Genome Mapping of Histone H3 Lys4 and 27 Trimethylations Reveals Distinct Genomic Compartments in Human Embryonic Stem Cells. *Cell Stem Cell* **1**, 286-298 (2007).
149. Roh, T.Y., Cuddapah, S., Cui, K. & Zhao, K. The genomic landscape of histone modifications in human T cells. *Proc Natl Acad Sci U S A* **103**, 15782-7 (2006).
150. Haupt, Y., Alexander, W.S., Barri, G., Klinken, S.P. & Adams, J.M. Novel zinc finger gene implicated as myc collaborator by retrovirally accelerated lymphomagenesis in E mu-myc transgenic mice. *Cell* **65**, 753-63 (1991).
151. van Lohuizen, M. et al. Identification of cooperating oncogenes in E mu-myc transgenic mice by provirus tagging. *Cell* **65**, 737-52 (1991).
152. Jacobs, J.J., Kieboom, K., Marino, S., DePinho, R.A. & van Lohuizen, M. The oncogene and Polycomb-group gene bmi-1 regulates cell proliferation and senescence through the ink4a locus. *Nature* **397**, 164-8 (1999).
153. Jacobs, J.J. et al. Bmi-1 collaborates with c-Myc in tumorigenesis by inhibiting c-Myc-induced apoptosis via INK4a/ARF. *Genes Dev* **13**, 2678-90 (1999).
154. Sherr, C.J. The INK4a/ARF network in tumour suppression. *Nat Rev Mol Cell Biol* **2**, 731-7 (2001).
155. Dietrich, N. et al. Bypass of senescence by the polycomb group protein CBX8 through direct binding to the INK4A-ARF locus. *EMBO J* **26**, 1637-48 (2007).
156. Bracken, A.P. et al. The Polycomb group proteins bind throughout the INK4A-ARF locus and are disassociated in senescent cells. *Genes Dev* **21**, 525-30 (2007).
157. Gil, J., Bernard, D., Martinez, D. & Beach, D. Polycomb CBX7 has a unifying role in cellular lifespan. *Nat Cell Biol* **6**, 67-72 (2004).
158. Ohta, H. et al. Polycomb group gene rae28 is required for sustaining activity of hematopoietic stem cells. *J Exp Med* **195**, 759-70 (2002).
159. Martinez, A.M., Colomb, S., Dejardin, J., Bantignies, F. & Cavalli, G. Polycomb group-dependent Cyclin A repression in Drosophila. *Genes Dev* **20**, 501-13 (2006).
160. Sparmann, A. & van Lohuizen, M. Polycomb silencers control cell fate, development and cancer. *Nat Rev Cancer* **6**, 846-56 (2006).
161. Kirmizis, A., Bartley, S.M. & Farnham, P.J. Identification of the polycomb group protein SU(Z)12 as a potential molecular target for human cancer therapy. *Mol Cancer Ther* **2**, 113-21 (2003).
162. Kleer, C.G. et al. EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. *Proc Natl Acad Sci U S A* **100**, 11606-11 (2003).
163. Varambally, S. et al. The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature* **419**, 624-9 (2002).
164. Levine, S.S., King, I.F. & Kingston, R.E. Division of labor in polycomb group repression. *Trends Biochem Sci* **29**, 478-85 (2004).
165. Dodd, I.B., Micheelsen, M.A., Sneppen, K. & Thon, G. Theoretical analysis of epigenetic cell memory by nucleosome modification. *Cell* **129**, 813-22 (2007).
166. Swigut, T. & Wysocka, J. H3K27 demethylases, at long last. *Cell* **131**, 29-32 (2007).

167. Acevedo, N. & Smith, G.D. Oocyte-specific gene signaling and its regulation of mammalian reproductive potential. *Front Biosci* **10**, 2335-45 (2005).
168. De La Fuente, R. et al. Major chromatin remodeling in the germinal vesicle (GV) of mammalian oocytes is dispensable for global transcriptional silencing but required for centromeric heterochromatin function. *Dev Biol* **275**, 447-58 (2004).
169. De La Fuente, R., Viveiros, M.M., Wigglesworth, K. & Eppig, J.J. ATRX, a member of the SNF2 family of helicase/ATPases, is required for chromosome alignment and meiotic spindle organization in metaphase II stage mouse oocytes. *Dev Biol* **272**, 1-14 (2004).
170. Chang, C.C. et al. A maternal store of macroH2A is removed from pronuclei prior to onset of somatic macroH2A expression in preimplantation embryos. *Dev Biol* **278**, 367-80 (2005).
171. Burns, K.H. et al. Roles of NPM2 in chromatin and nucleolar organization in oocytes and embryos. *Science* **300**, 633-6 (2003).
172. Frehlick, L.J., Eirin-Lopez, J.M. & Ausio, J. New insights into the nucleophosmin/nucleoplasmin family of nuclear chaperones. *Bioessays* **29**, 49-59 (2007).
173. Akiyama, T., Kim, J.M., Nagata, M. & Aoki, F. Regulation of histone acetylation during meiotic maturation in mouse oocytes. *Mol Reprod Dev* **69**, 222-7 (2004).
174. Kim, J.M., Liu, H., Tazaki, M., Nagata, M. & Aoki, F. Changes in histone acetylation during mouse oocyte meiosis. *J Cell Biol* **162**, 37-46 (2003).
175. Spinaci, M., Seren, E. & Mattioli, M. Maternal chromatin remodeling during maturation and after fertilization in mouse oocytes. *Mol Reprod Dev* **69**, 215-21 (2004).
176. Wang, Q. et al. Histone deacetylation is required for orderly meiosis. *Cell Cycle* **5**, 766-74 (2006).
177. Akiyama, T., Nagata, M. & Aoki, F. Inadequate histone deacetylation during oocyte meiosis causes aneuploidy and embryo death in mice. *Proc Natl Acad Sci U S A* **103**, 7339-44 (2006).
178. Bui, H.T. et al. Regulation of chromatin and chromosome morphology by histone H3 modifications in pig oocytes. *Reproduction* **133**, 371-82 (2007).
179. Eberlin, A. et al. Histone H3 tails containing dimethylated lysine and adjacent phosphorylated serine modifications adopt a specific conformation during mitosis and meiosis. *Mol Cell Biol* (2008).
180. Sun, F. et al. Nuclear reprogramming: the zygotic transcription program is established through an "erase-and-rebuild" strategy. *Cell Res* **17**, 117-34 (2007).
181. Sarmento, O.F. et al. Dynamic alterations of specific histone modifications during early murine development. *J Cell Sci* **117**, 4449-59 (2004).
182. Kageyama, S. et al. Alterations in epigenetic modifications during oocyte growth in mice. *Reproduction* **133**, 85-94 (2007).
183. Evsikov, A.V. et al. Cracking the egg: molecular dynamics and evolutionary aspects of the transition from the fully grown oocyte to embryo. *Genes Dev* **20**, 2713-27 (2006).
184. Kocabas, A.M. et al. The transcriptome of human oocytes. *Proc Natl Acad Sci U S A* **103**, 14027-32 (2006).
185. Su, Y.Q. et al. Selective degradation of transcripts during meiotic maturation of mouse oocytes. *Dev Biol* **302**, 104-17 (2007).
186. Richter, J.D. Breaking the code of polyadenylation-induced translation. *Cell* **132**, 335-7 (2008).
187. Schultz, R.M. The molecular foundations of the maternal to zygotic transition in the preimplantation embryo. *Hum Reprod Update* **8**, 323-31 (2002).
188. Pique, M., Lopez, J.M., Foissac, S., Guigo, R. & Mendez, R. A combinatorial code for CPE-mediated translational control. *Cell* **132**, 434-48 (2008).
189. Yang, J., Medvedev, S., Yu, J., Schultz, R.M. & Hecht, N.B. Deletion of the DNA/RNA-binding protein MSY2 leads to post-meiotic arrest. *Mol Cell Endocrinol* **250**, 20-4 (2006).
190. Bettgowda, A., Lee, K.B. & Smith, G.W. Cytoplasmic and nuclear determinants of the maternal-to-embryonic transition. *Reprod Fertil Dev* **20**, 45-53 (2008).



191. Wu, X. et al. Zygote arrest 1 (Zar1) is a novel maternal-effect gene critical for the oocyte-to-embryo transition. *Nat Genet* **33**, 187-91 (2003).
192. Tong, Z.B. et al. Mater, a maternal effect gene required for early embryonic development in mice. *Nat Genet* **26**, 267-8 (2000).
193. Roest, H.P. et al. The ubiquitin-conjugating DNA repair enzyme HR6A is a maternal factor essential for early embryonic development in mice. *Mol Cell Biol* **24**, 5485-95 (2004).
194. Christians, E., Davis, A.A., Thomas, S.D. & Benjamin, I.J. Maternal effect of Hsf1 on reproductive success. *Nature* **407**, 693-4 (2000).
195. Ogushi, S. et al. The maternal nucleolus is essential for early embryonic development in mammals. *Science* **319**, 613-6 (2008).
196. Ma, J., Zeng, F., Schultz, R.M. & Tseng, H. Basonuclin: a novel mammalian maternal-effect gene. *Development* **133**, 2053-62 (2006).
197. Nagy, A., Gertsenstein, M., Vintersten, K. & Behringer, R. *Manipulating the mouse embryo: a laboratory manual*, (Cold Spring Harbor Laboratory Press, 2003).
198. Murchison, E.P. & Hannon, G.J. miRNAs on the move: miRNA biogenesis and the RNAi machinery. *Curr Opin Cell Biol* **16**, 223-9 (2004).
199. Murchison, E.P. et al. Critical roles for Dicer in the female germline. *Genes Dev* **21**, 682-93 (2007).
200. Tang, F. et al. Maternal microRNAs are essential for mouse zygotic development. *Genes Dev* **21**, 644-8 (2007).
201. Watanabe, T. et al. Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. *Genes Dev* **20**, 1732-43 (2006).
202. van der Heijden, G.W. et al. Asymmetry in histone H3 variants and lysine methylation between paternal and maternal chromatin of the early mouse zygote. *Mech Dev* **122**, 1008-22 (2005).
203. Rousseaux, S. et al. Establishment of male-specific epigenetic information. *Gene* **345**, 139-53 (2005).
204. Churikov, D., Zalenskaya, I.A. & Zalensky, A.O. Male germline-specific histones in mouse and man. *Cytogenet Genome Res* **105**, 203-14 (2004).
205. O'Carroll, D. et al. Isolation and characterization of Suv39h2, a second histone H3 methyltransferase gene that displays testis-specific expression. *Mol Cell Biol* **20**, 9423-33 (2000).
206. Takada, Y. et al. Mammalian Polycomb Scmh1 mediates exclusion of Polycomb complexes from the XY body in the pachytene spermatocytes. *Development* **134**, 579-90 (2007).
207. van der Heijden, G.W. et al. Chromosome-wide nucleosome replacement and H3.3 incorporation during mammalian meiotic sex chromosome inactivation. *Nat Genet* **39**, 251-8 (2007).
208. Reik, W. & Walter, J. Genomic imprinting: parental influence on the genome. *Nat Rev Genet* **2**, 21-32 (2001).
209. Zalenskaya, I.A., Bradbury, E.M. & Zalensky, A.O. Chromatin structure of telomere domain in human sperm. *Biochem Biophys Res Commun* **279**, 213-8 (2000).
210. van der Heijden, G.W. et al. Transmission of modified nucleosomes from the mouse male germline to the zygote and subsequent remodeling of paternal chromatin. *Dev Biol* **298**, 458-69 (2006).
211. Henikoff, S. & Ahmad, K. Assembly of variant histones into chromatin. *Annu Rev Cell Dev Biol* **21**, 133-53 (2005).
212. Torres-Padilla, M.E., Bannister, A.J., Hurd, P.J., Kouzarides, T. & Zernicka-Goetz, M. Dynamic distribution of the replacement histone variant H3.3 in the mouse oocyte and preimplantation embryos. *Int J Dev Biol* **50**, 455-61 (2006).
213. Bonnefoy, E., Orsi, G.A., Couble, P. & Loppin, B. The essential role of Drosophila HIRA for de novo assembly of paternal chromatin at fertilization. *PLoS Genet* **3**, 1991-2006 (2007).
214. Loppin, B. et al. The histone H3.3 chaperone HIRA is essential for chromatin assembly in the male pronucleus. *Nature* **437**, 1386-90 (2005).

215. Konev, A.Y. et al. CHD1 motor protein is required for deposition of histone variant H3.3 into chromatin in vivo. *Science* **317**, 1087-90 (2007).
216. Adenot, P.G., Mercier, Y., Renard, J.P. & Thompson, E.M. Differential H4 acetylation of paternal and maternal chromatin precedes DNA replication and differential transcriptional activity in pronuclei of 1-cell mouse embryos. *Development* **124**, 4615-25 (1997).
217. Santos, F., Hendrich, B., Reik, W. & Dean, W. Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev Biol* **241**, 172-82 (2002).
218. Santos, F., Peters, A.H., Otte, A.P., Reik, W. & Dean, W. Dynamic chromatin modifications characterise the first cell cycle in mouse embryos. *Dev Biol* **280**, 225-36 (2005).
219. Cha, T.L. et al. Akt-mediated phosphorylation of EZH2 suppresses methylation of lysine 27 in histone H3. *Science* **310**, 306-10 (2005).
220. Lepikhov, K. & Walter, J. Differential dynamics of histone H3 methylation at positions K4 and K9 in the mouse zygote. *BMC Dev Biol* **4**, 12 (2004).
221. Yeo, S., Lee, K.K., Han, Y.M. & Kang, Y.K. Methylation changes of lysine 9 of histone H3 during preimplantation mouse development. *Mol Cells* **20**, 423-8 (2005).
222. Merico, V. et al. Epigenomic differentiation in mouse preimplantation nuclei of biparental, parthenote and cloned embryos. *Chromosome Res* **15**, 341-60 (2007).
223. Liu, H., Kim, J.M. & Aoki, F. Regulation of histone H3 lysine 9 methylation in oocytes and early pre-implantation embryos. *Development* **131**, 2269-80 (2004).
224. Mayer, W., Niveleau, A., Walter, J., Fundele, R. & Haaf, T. Demethylation of the zygotic paternal genome. *Nature* **403**, 501-2 (2000).
225. Mayer, W., Smith, A., Fundele, R. & Haaf, T. Spatial separation of parental genomes in preimplantation mouse embryos. *J Cell Biol* **148**, 629-34 (2000).
226. Oswald, J. et al. Active demethylation of the paternal genome in the mouse zygote. *Curr Biol* **10**, 475-8 (2000).
227. Olek, A. & Walter, J. The pre-implantation ontogeny of the H19 methylation imprint. *Nat Genet* **17**, 275-6 (1997).
228. Rougier, N. et al. Chromosome methylation patterns during mammalian preimplantation development. *Genes Dev* **12**, 2108-13 (1998).
229. Lane, N. et al. Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. *Genesis* **35**, 88-93 (2003).
230. Howell, C.Y. et al. Genomic imprinting disrupted by a maternal effect mutation in the Dnmt1 gene. *Cell* **104**, 829-38 (2001).
231. Gaudet, F. et al. Dnmt1 expression in pre- and postimplantation embryogenesis and the maintenance of IAP silencing. *Mol Cell Biol* **24**, 1640-8 (2004).
232. Cirio, M.C. et al. Preimplantation expression of the somatic form of Dnmt1 suggests a role in the inheritance of genomic imprints. *BMC Dev Biol* **8**, 9 (2008).
233. Kurihara, Y. et al. Maintenance of genomic methylation patterns during preimplantation development requires the somatic form of DNA methyltransferase 1. *Dev Biol* **313**, 335-46 (2008).
234. Santos, F. & Dean, W. Epigenetic reprogramming during early development in mammals. *Reproduction* **127**, 643-51 (2004).
235. Morgan, H.D., Dean, W., Coker, H.A., Reik, W. & Petersen-Mahrt, S.K. Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *J Biol Chem* **279**, 52353-60 (2004).
236. Reik, W., Dean, W. & Walter, J. Epigenetic reprogramming in mammalian development. *Science* **293**, 1089-93 (2001).
237. Fulka, H., Mrazek, M., Tepla, O. & Fulka, J., Jr. DNA methylation pattern in human zygotes and developing embryos. *Reproduction* **128**, 703-8 (2004).

238. Beaujean, N. et al. Non-conservation of mammalian preimplantation methylation dynamics. *Curr Biol* **14**, R266-7 (2004).
239. Young, L.E. & Beaujean, N. DNA methylation in the preimplantation embryo: the differing stories of the mouse and sheep. *Anim Reprod Sci* **82-83**, 61-78 (2004).
240. Dean, W. et al. Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos. *Proc Natl Acad Sci U S A* **98**, 13734-8 (2001).
241. Beaujean, N. et al. The effect of interspecific oocytes on demethylation of sperm DNA. *Proc Natl Acad Sci U S A* **101**, 7636-40 (2004).
242. Payer, B. et al. Stella is a maternal effect gene required for normal early development in mice. *Curr Biol* **13**, 2110-7 (2003).
243. Nakamura, T. et al. PGC7/Stella protects against DNA demethylation in early embryogenesis. *Nat Cell Biol* **9**, 64-71 (2007).
244. Stitzel, M.L. & Seydoux, G. Regulation of the oocyte-to-zygote transition. *Science* **316**, 407-8 (2007).
245. Knowles, B.B., Evsikov, A.V., de Vries, W.N., Peaston, A.E. & Solter, D. Molecular control of the oocyte to embryo transition. *Philos Trans R Soc Lond B Biol Sci* **358**, 1381-7 (2003).
246. Aoki, F., Worrad, D.M. & Schultz, R.M. Regulation of transcriptional activity during the first and second cell cycles in the preimplantation mouse embryo. *Dev Biol* **181**, 296-307 (1997).
247. Zeng, F., Baldwin, D.A. & Schultz, R.M. Transcript profiling during preimplantation mouse development. *Dev Biol* **272**, 483-96 (2004).
248. Bellier, S. et al. Nuclear translocation and carboxyl-terminal domain phosphorylation of RNA polymerase II delineate the two phases of zygotic gene activation in mammalian embryos. *EMBO J* **16**, 6250-62 (1997).
249. Zurita, M., Reynaud, E. & Aguilar-Fuentes, J. From the beginning: the basal transcription machinery and onset of transcription in the early animal embryo. *Cell Mol Life Sci* (2007).
250. Zeng, F. & Schultz, R.M. RNA transcript profiling during zygotic gene activation in the preimplantation mouse embryo. *Dev Biol* **283**, 40-57 (2005).
251. Pokholok, D.K. et al. Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* **122**, 517-27 (2005).
252. Schubeler, D. et al. The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev* **18**, 1263-71 (2004).
253. Muse, G.W. et al. RNA polymerase is poised for activation across the genome. *Nat Genet* **39**, 1507-11 (2007).
254. Zeitlinger, J. et al. RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo. *Nat Genet* **39**, 1512-6 (2007).
255. Lorincz, M.C. & Schubeler, D. RNA polymerase II: just stopping by. *Cell* **130**, 16-8 (2007).
256. O'Neill, L.P., VerMilyea, M.D. & Turner, B.M. Epigenetic characterization of the early embryo with a chromatin immunoprecipitation protocol applicable to small cell populations. *Nat Genet* **38**, 835-41 (2006).
257. Wang, Q.T. et al. A genome-wide study of gene activity reveals developmental signaling pathways in the preimplantation mouse embryo. *Dev Cell* **6**, 133-44 (2004).
258. Hamatani, T., Carter, M.G., Sharov, A.A. & Ko, M.S. Dynamics of global gene expression changes during mouse preimplantation development. *Dev Cell* **6**, 117-31 (2004).
259. Bultman, S.J. et al. Maternal BRG1 regulates zygotic genome activation in the mouse. *Genes Dev* **20**, 1744-54 (2006).
260. Torres-Padilla, M.E. & Zernicka-Goetz, M. Role of TIF1alpha as a modulator of embryonic transcription in the mouse zygote. *J Cell Biol* **174**, 329-38 (2006).
261. Giraldez, A.J. et al. Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* **312**, 75-9 (2006).

262. Willadsen, S.M. Cloning of sheep and cow embryos. *Genome* **31**, 956-62 (1989).
263. Wilmut, I., Schnieke, A.E., McWhir, J., Kind, A.J. & Campbell, K.H. Viable offspring derived from fetal and adult mammalian cells. *Nature* **385**, 810-3 (1997).
264. Yang, X. et al. Nuclear reprogramming of cloned embryos and its implications for therapeutic cloning. *Nat Genet* **39**, 295-302 (2007).
265. Cibelli, J.B., Campbell, K.H., Seidel, G.E., West, M.D. & Lanza, R.P. The health profile of cloned animals. *Nat Biotechnol* **20**, 13-4 (2002).
266. Santos, F. et al. Epigenetic marking correlates with developmental potential in cloned bovine preimplantation embryos. *Curr Biol* **13**, 1116-21 (2003).
267. Kishigami, S. et al. Successful mouse cloning of an outbred strain by trichostatin A treatment after somatic nuclear transfer. *J Reprod Dev* **53**, 165-70 (2007).
268. Kishigami, S. et al. Significant improvement of mouse cloning technique by treatment with trichostatin A after somatic nuclear transfer. *Biochem Biophys Res Commun* **340**, 183-9 (2006).
269. Kishigami, S. et al. Epigenetic abnormalities of the mouse paternal zygotic genome associated with microinsemination of round spermatids. *Dev Biol* **289**, 195-205 (2006).
270. Yoshida, N., Brahmajosyula, M., Shoji, S., Amanai, M. & Perry, A.C. Epigenetic discrimination by mouse metaphase II oocytes mediates asymmetric chromatin remodeling independently of meiotic exit. *Dev Biol* **301**, 464-77 (2007).
271. Egli, D., Rosains, J., Birkhoff, G. & Eggan, K. Developmental reprogramming after chromosome transfer into mitotic mouse zygotes. *Nature* **447**, 679-85 (2007).
272. Fu, G. et al. Mouse oocytes and early embryos express multiple histone H1 subtypes. *Biol Reprod* **68**, 1569-76 (2003).
273. Maeda, C. et al. DNA Hypomethylation Circuit of the Mouse Oocyte-Specific Histone H1foo Gene in Female Germ Cell Lineage. *Biol Reprod* (2008).
274. Tanaka, M. et al. H1FOO is coupled to the initiation of oocytic growth. *Biol Reprod* **72**, 135-42 (2005).
275. Furuya, M. et al. H1foo is indispensable for meiotic maturation of the mouse oocyte. *J Reprod Dev* **53**, 895-902 (2007).
276. Gao, S. et al. Rapid H1 linker histone transitions following fertilization or somatic cell nuclear transfer: evidence for a uniform developmental program in mice. *Dev Biol* **266**, 62-75 (2004).
277. Becker, M. et al. Differential in vivo binding dynamics of somatic and oocyte-specific linker histones in oocytes and during ES cell nuclear transfer. *Mol Biol Cell* **16**, 3887-95 (2005).
278. Teranishi, T. et al. Rapid replacement of somatic linker histones with the oocyte-specific linker histone H1foo in nuclear transfer. *Dev Biol* **266**, 76-86 (2004).
279. Smith, S.L. et al. Global gene expression profiles reveal significant nuclear reprogramming by the blastocyst stage after cloning. *Proc Natl Acad Sci U S A* **102**, 17582-7 (2005).
280. Bortvin, A. et al. Incomplete reactivation of Oct4-related genes in mouse embryos cloned from somatic nuclei. *Development* **130**, 1673-80 (2003).
281. Cavaleri, F.M. et al. Subsets of cloned mouse embryos and their non-random relationship to development and nuclear reprogramming. *Mech Dev* **125**, 153-166 (2008).
282. Ashe, A. & Whitelaw, E. Another role for RNA: a messenger across generations. *Trends Genet* **23**, 8-10 (2007).
283. Blewitt, M.E., Chong, S. & Whitelaw, E. How the mouse got its spots. *Trends Genet* **20**, 550-4 (2004).
284. Chandler, V.L. & Stam, M. Chromatin conversations: mechanisms and implications of paramutation. *Nat Rev Genet* **5**, 532-44 (2004).
285. Chong, S. & Whitelaw, E. Epigenetic germline inheritance. *Curr Opin Genet Dev* **14**, 692-6 (2004).
286. Chandler, V.L. Paramutation: from maize to mice. *Cell* **128**, 641-5 (2007).

287. Kuhfittig, S. et al. pitkin(D), a novel gain-of-function enhancer of position-effect variegation, affects chromatin regulation during oogenesis and early embryogenesis in *Drosophila*. *Genetics* **157**, 1227-44 (2001).
288. Cavalli, G. & Paro, R. The *Drosophila* Fab-7 chromosomal element conveys epigenetic inheritance during mitosis and meiosis. *Cell* **93**, 505-18 (1998).
289. Morgan, H.D., Sutherland, H.G., Martin, D.I. & Whitelaw, E. Epigenetic inheritance at the agouti locus in the mouse. *Nat Genet* **23**, 314-8 (1999).
290. Blewitt, M.E., Vickaryous, N.K., Paldi, A., Koseki, H. & Whitelaw, E. Dynamic reprogramming of DNA methylation at an epigenetically sensitive allele in mice. *PLoS Genet* **2**, e49 (2006).
291. Chong, S. et al. Modifiers of epigenetic reprogramming show paternal effects in the mouse. *Nat Genet* **39**, 614-22 (2007).
292. Rassoulzadegan, M. et al. RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. *Nature* **441**, 469-74 (2006).
293. Chan, T.L. et al. Heritable germline epimutation of MSH2 in a family with hereditary nonpolyposis colorectal cancer. *Nat Genet* **38**, 1178-83 (2006).
294. Suter, C.M., Martin, D.I. & Ward, R.L. Germline epimutation of MLH1 in individuals with multiple cancers. *Nat Genet* **36**, 497-501 (2004).
295. Chong, S., Youngson, N.A. & Whitelaw, E. Heritable germline epimutation is not the same as transgenerational epigenetic inheritance. *Nat Genet* **39**, 574-5; author reply 575-6 (2007).
296. Suter, C.M. & Martin, D.I. Inherited epimutation or a haplotypic basis for the propensity to silence? *Nat Genet* **39**, 573; author reply 576 (2007).
297. Probst, A.V. & Almouzni, G. Pericentric heterochromatin: dynamic organization during early development in mammals. *Differentiation* (2007).
298. Probst, A.V., Santos, F., Reik, W., Almouzni, G. & Dean, W. Structural differences in centromeric heterochromatin are spatially reconciled on fertilisation in the mouse zygote. *Chromosoma* **116**, 403-15 (2007).
299. Martin, C. et al. Genome restructuring in mouse embryos during reprogramming and early development. *Dev Biol* **292**, 317-32 (2006).
300. Arney, K.L., Bao, S., Bannister, A.J., Kouzarides, T. & Surani, M.A. Histone methylation defines epigenetic asymmetry in the mouse zygote. *Int J Dev Biol* **46**, 317-20 (2002).
301. Rudolph, T. et al. Heterochromatin formation in *Drosophila* is initiated through active removal of H3K4 methylation by the LSD1 homolog SU(VAR)3-3. *Mol Cell* **26**, 103-15 (2007).
302. Govin, J. et al. Pericentric heterochromatin reprogramming by new histone variants during mouse spermiogenesis. *J Cell Biol* **176**, 283-94 (2007).
303. Baart, E.B. et al. Preimplantation genetic screening reveals a high incidence of aneuploidy and mosaicism in embryos from young women undergoing IVF. *Hum Reprod* **21**, 223-33 (2006).
304. Yamazaki, T., Kobayakawa, S., Yamagata, K., Abe, K. & Baba, T. Molecular dynamics of heterochromatin protein 1beta, HP1beta, during mouse preimplantation development. *J Reprod Dev* **53**, 1035-41 (2007).
305. Houliard, M. et al. CAF-1 is essential for heterochromatin organization in pluripotent embryonic cells. *PLoS Genet* **2**, e181 (2006).
306. De Vries, W.N. et al. Maternal beta-catenin and E-cadherin in mouse development. *Development* **131**, 4435-45 (2004).
307. Johnson, M.H. & McConnell, J.M. Lineage allocation and cell polarity during mouse embryogenesis. *Semin Cell Dev Biol* **15**, 583-97 (2004).
308. Loebel, D.A., Watson, C.M., De Young, R.A. & Tam, P.P. Lineage choice and differentiation in mouse embryos and embryonic stem cells. *Dev Biol* **264**, 1-14 (2003).
309. Rossant, J. Stem Cells and Early Lineage Development. *Cell* **132**, 527-531 (2008).

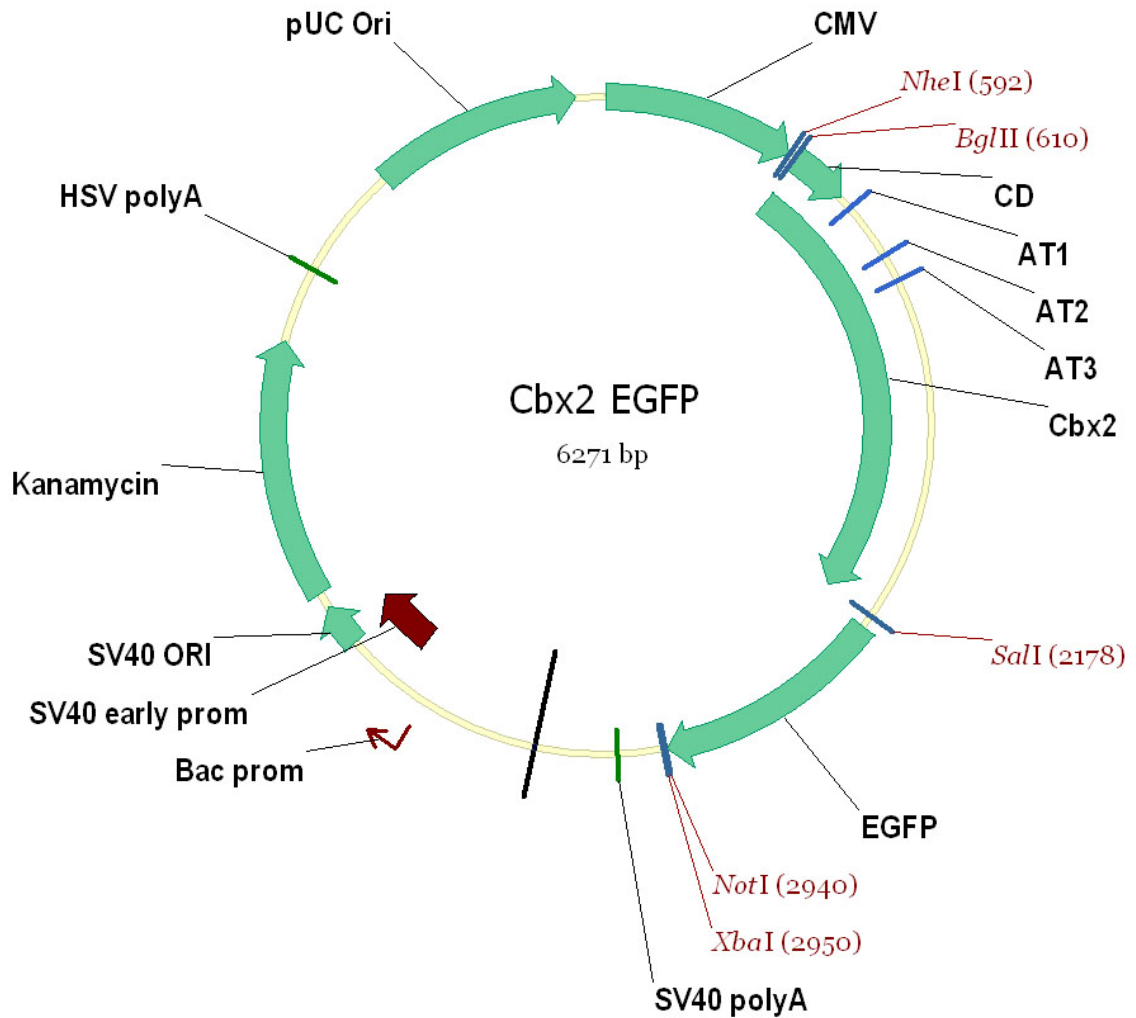
310. Rossant, J., Chazaud, C. & Yamanaka, Y. Lineage allocation and asymmetries in the early mouse embryo. *Philos Trans R Soc Lond B Biol Sci* **358**, 1341-8; discussion 1349 (2003).
311. Johnson, B.V., Rathjen, J. & Rathjen, P.D. Transcriptional control of pluripotency: decisions in early development. *Curr Opin Genet Dev* **16**, 447-54 (2006).
312. Surani, M.A., Hayashi, K. & Hajkova, P. Genetic and epigenetic regulators of pluripotency. *Cell* **128**, 747-62 (2007).
313. Roberts, R.M., Ezashi, T. & Das, P. Trophoblast gene expression: transcription factors in the specification of early trophoblast. *Reprod Biol Endocrinol* **2**, 47 (2004).
314. Nichols, J. et al. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95**, 379-91 (1998).
315. Mitsui, K. et al. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**, 631-42 (2003).
316. Niwa, H., Miyazaki, J. & Smith, A.G. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* **24**, 372-6 (2000).
317. Yeom, Y.I. et al. Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. *Development* **122**, 881-94 (1996).
318. Feldman, N. et al. G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nat Cell Biol* **8**, 188-94 (2006).
319. Niwa, H. et al. Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell* **123**, 917-29 (2005).
320. Chazaud, C., Yamanaka, Y., Pawson, T. & Rossant, J. Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. *Dev Cell* **10**, 615-24 (2006).
321. Henckel, A., Toth, S. & Arnaud, P. Early mouse embryo development: could epigenetics influence cell fate determination? *Bioessays* **29**, 520-4 (2007).
322. Zernicka-Goetz, M. The first cell-fate decisions in the mouse embryo: destiny is a matter of both chance and choice. *Curr Opin Genet Dev* **16**, 406-12 (2006).
323. Chroscicka, A., Komorowski, S. & Maleszewski, M. Both blastomeres of the mouse 2-cell embryo contribute to the embryonic portion of the blastocyst. *Mol Reprod Dev* **68**, 308-12 (2004).
324. Hiiragi, T. & Solter, D. First cleavage plane of the mouse egg is not predetermined but defined by the topology of the two apposing pronuclei. *Nature* **430**, 360-4 (2004).
325. Motosugi, N., Bauer, T., Polanski, Z., Solter, D. & Hiiragi, T. Polarity of the mouse embryo is established at blastocyst and is not prepatterned. *Genes Dev* **19**, 1081-92 (2005).
326. Plusa, B. et al. The first cleavage of the mouse zygote predicts the blastocyst axis. *Nature* **434**, 391-5 (2005).
327. Piotrowska, K., Wianny, F., Pedersen, R.A. & Zernicka-Goetz, M. Blastomeres arising from the first cleavage division have distinguishable fates in normal mouse development. *Development* **128**, 3739-48 (2001).
328. Gardner, R.L. Specification of embryonic axes begins before cleavage in normal mouse development. *Development* **128**, 839-47 (2001).
329. Piotrowska-Nitsche, K., Perea-Gomez, A., Haraguchi, S. & Zernicka-Goetz, M. Four-cell stage mouse blastomeres have different developmental properties. *Development* **132**, 479-90 (2005).
330. Torres-Padilla, M.E., Parfitt, D.E., Kouzarides, T. & Zernicka-Goetz, M. Histone arginine methylation regulates pluripotency in the early mouse embryo. *Nature* **445**, 214-8 (2007).
331. Wutz, A. & Gribnau, J. X inactivation Xplained. *Curr Opin Genet Dev* **17**, 387-93 (2007).
332. Umlauf, D. et al. Imprinting along the Kcnq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. *Nat Genet* **36**, 1296-300 (2004).
333. Maison, C. et al. Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. *Nat Genet* **30**, 329-34 (2002).

334. Aravind, L. & Landsman, D. AT-hook motifs identified in a wide variety of DNA-binding proteins. *Nucleic Acids Res* **26**, 4413-21 (1998).
335. Bianchi, M.E. & Agresti, A. HMG proteins: dynamic players in gene regulation and differentiation. *Curr Opin Genet Dev* **15**, 496-506 (2005).
336. Harrer, M., Luhrs, H., Bustin, M., Scheer, U. & Hock, R. Dynamic interaction of HMGA1a proteins with chromatin. *J Cell Sci* **117**, 3459-71 (2004).
337. Hock, R., Furusawa, T., Ueda, T. & Bustin, M. HMG chromosomal proteins in development and disease. *Trends Cell Biol* **17**, 72-9 (2007).
338. Thanos, D. & Maniatis, T. The high mobility group protein HMG I(Y) is required for NF-kappa B-dependent virus induction of the human IFN-beta gene. *Cell* **71**, 777-89 (1992).
339. Reeves, R. & Beckerbauer, L. HMGI/Y proteins: flexible regulators of transcription and chromatin structure. *Biochim Biophys Acta* **1519**, 13-29 (2001).
340. Beaujean, N. et al. Induction of early transcription in one-cell mouse embryos by microinjection of the nonhistone chromosomal protein HMG-I. *Dev Biol* **221**, 337-54 (2000).
341. Zou, Y. et al. A mass spectrometric study on the in vitro methylation of HMGA1a and HMGA1b proteins by PRMTs: methylation specificity, the effect of binding to AT-rich duplex DNA, and the effect of C-terminal phosphorylation. *Biochemistry* **46**, 7896-906 (2007).
342. Huth, J.R. et al. The solution structure of an HMG-I(Y)-DNA complex defines a new architectural minor groove binding motif. *Nat Struct Biol* **4**, 657-65 (1997).
343. Peters, A.H. et al. Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Mol Cell* **12**, 1577-89 (2003).
344. Ayton, P.M., Chen, E.H. & Cleary, M.L. Binding to nonmethylated CpG DNA is essential for target recognition, transactivation, and myeloid transformation by an MLL oncoprotein. *Mol Cell Biol* **24**, 10470-8 (2004).
345. Susbielle, G., Blattes, R., Brevet, V., Monod, C. & Kas, E. Target practice: aiming at satellite repeats with DNA minor groove binders. *Curr Med Chem Anticancer Agents* **5**, 409-20 (2005).
346. Henderson, A., Bunce, M., Siddon, N., Reeves, R. & Tremethick, D.J. High-mobility-group protein I can modulate binding of transcription factors to the U5 region of the human immunodeficiency virus type 1 proviral promoter. *J Virol* **74**, 10523-34 (2000).
347. Cui, T. & Leng, F. Specific recognition of AT-rich DNA sequences by the mammalian high mobility group protein AT-hook 2: a SELEX study. *Biochemistry* **46**, 13059-66 (2007).
348. Metcalf, C.E. & Wassarman, D.A. DNA binding properties of TAF1 isoforms with two AT-hooks. *J Biol Chem* **281**, 30015-23 (2006).
349. Cui, T., Wei, S., Brew, K. & Leng, F. Energetics of binding the mammalian high mobility group protein HMGA2 to poly(dA-dT)<sub>2</sub> and poly(dA)-poly(dT). *J Mol Biol* **352**, 629-45 (2005).
350. Singh, M., D'Silva, L. & Holak, T.A. DNA-binding properties of the recombinant high-mobility-group-like AT-hook-containing region from human BRG1 protein. *Biol Chem* **387**, 1469-78 (2006).
351. Tu, S. et al. The ARID domain of the H3K4 demethylase RBP2 binds to a DNA CCGCCC motif. *Nat Struct Mol Biol* (2008).
352. Li, H. et al. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature* **442**, 91-5 (2006).
353. Wysocka, J. et al. A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **442**, 86-90 (2006).
354. Trojer, P. & Reinberg, D. Histone lysine demethylases and their impact on epigenetics. *Cell* **125**, 213-7 (2006).
355. Ekwall, K., Olsson, T., Turner, B.M., Cranston, G. & Allshire, R.C. Transient inhibition of histone deacetylation alters the structural and functional imprint at fission yeast centromeres. *Cell* **91**, 1021-32 (1997).
356. Taddei, A., Maison, C., Roche, D. & Almouzni, G. Reversible disruption of pericentric heterochromatin and centromere function by inhibiting deacetylases. *Nat Cell Biol* **3**, 114-20 (2001).



357. Peng, J.C. & Karpen, G.H. H3K9 methylation and RNA interference regulate nucleolar organization and repeated DNA stability. *Nat Cell Biol* **9**, 25-35 (2007).
358. O'Dor, E., Beck, S.A. & Brock, H.W. Polycomb group mutants exhibit mitotic defects in syncytial cell cycles of *Drosophila* embryos. *Dev Biol* **290**, 312-22 (2006).
359. Yu, J. et al. Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**, 1917-20 (2007).
360. Maherali, N. et al. Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell* **1**, 55-70 (2007).
361. Wernig, M. et al. In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* **448**, 318-24 (2007).
362. Takahashi, K. et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861-72 (2007).
363. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-76 (2006).
364. Park, I.H. et al. Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* **451**, 141-6 (2008).
365. Cyranoski, D. Stem cells: 5 things to know before jumping on the iPS bandwagon. *Nature* **452**, 406-8 (2008).

## Appendix



Features	bp	aa
promoter (CMV)	1..589	
mRNA (Cbx2)	618..2174	1..519
CD	618..791	1..58
AT1	837..863	74..82
AT2	1014..1040	133..141
AT3	1101..1127	162..170
mRNA (EGFP)	2217..2936	534..773
polyA_signal (SV40)	3090..3140	
promoter (Bac)	3704..3732	
promoter (SV40 early)	3816..4045	
SV40\ORI	3983..4118	
mRNA (Kanamycin)	4167..4961	
polyA_signal (HSV)	5197..5215	
pUC\Ori	5546..6189	



List of Cbx, PRC1 and Hmga contracts

No.	Cloned by	Insert (systematic name)	Construct (tube label)	Backbone	Conc
1	E. Bernstein	Cbx2	Cbx2_GFP	EGFP-N1	1.23
2	E. Bernstein	Cbx4	Cbx4_GFP	EGFP-N1	1.34
3	E. Bernstein	Cbx6	Cbx6_GFP	EGFP-N1	1.10
4	E. Bernstein	Cbx7	Cbx7_GFP	EGFP-N1	1.37
5	E. Bernstein	Cbx7_CD <sup>F11A</sup>	Cbx7_F11A	EGFP-N1	0.98
6	E. Bernstein	Cbx7_CD <sup>W35A</sup>	Cbx7_W35A	EGFP-N1	0.72
7	E. Bernstein	Cbx7_CD4 <sup>SWAP</sup>	Cbx7_CD4_SWAP	EGFP-N1	1.17
8	E. Bernstein	Cbx8	Cbx8_GFP	EGFP-N1	1.09
9	C. Kolb	Cbx2	Cbx2_GFP	pBSSK+ (T7)	0.54
10	C. Kolb	Cbx2_AT1 <sup>AAA</sup>	Cbx2_mAT_GFP	pBSSK+ (T7)	0.68
11	C. Kolb	Cbx4	Cbx4_GFP	pBSSK+ (T7)	0.05
12	C. Kolb	Cbx6	Cbx6_GFP	pBSSK+ (T7)	0.44
13	C. Kolb	Cbx7	Cbx7_GFP	pBSSK+ (T7)	0.53
14	C. Kolb	Cbx7_CD <sup>F11A</sup>	Cbx7_F11A_GFP	pBSSK+ (T7)	0.06
15	C. Kolb	Cbx7_CD <sup>W35A</sup>	Cbx7_W35A_GFP	pBSSK+ (T7)	0.03
16	C. Kolb	Cbx7_CD4 <sup>SWAP</sup>	Cbx7_CD4_SWAP	pBSSK+ (T7)	
17	C. Kolb	Cbx8	Cbx8_GFP	pBSSK+ (T7)	0.11
35	M. Endoh (Koseki)	Ring1B	Myc-Ring1B (WT)	pCAGIPuro	1.30
41	M. Endoh (Koseki)	Cbx2	Cbx2-Flag	pCAGIPuro	1.20
44	M. Endoh (Koseki)	Bmi1	3 Flag-Bmi1	pCAGIPuro	1.04
45	K. Isono (Koseki)	Ring1B	GFP-Ring1B	pEGFP-C2	0.89
46	K. Isono (Koseki)	Phc2	GFP-Phc2	pEGFP-C2	1.01
47	K. Isono (Koseki)	Cbx2	GFP-Cbx2 (M33)	pEGFP-C2	1.12
48	K. Isono (Koseki)	Ring1A	GFP-Ring1A	pEGFP-C2	0.98
51	C. Kolb	Cbx2_AT1 <sup>AAA</sup>	Cbx2_mAT1_GFP (PRG-->AAA)	pEGFP-N1	2.07
52	C. Kolb	Cbx2_AT1 <sup>G78R</sup>	Cbx2_mAT2_GFP (G77R)	pEGFP_N1	1.05
53	C. Kolb	Cbx2_AT1 <sup>G78L</sup>	Cbx2_mAT3_GFP (G77L)	pEGFP_N1	1.28
54	C. Kolb	Cbx2_AT1 <sup>Cbx7</sup>	Cbx2_mAT4_GFP (GRPRK-->RLLLQ)	pEGFP_N1	0.71
57	C. Kolb	Cbx2_AT1 <sup>Hmga1</sup>	Cbx2_mAT6_GFP (AT_Hmga1)	pEGFP_N1	0.90
121	C. Kolb	Cbx2_AT2 <sup>G137R</sup>	Cbx2_m2AT2	pEGFP_N1	1.02
122	C. Kolb	Cbx2_AT1 <sup>AAA</sup> _AT2 <sup>G137R</sup>	Cbx2_mAT1_m2AT2	pEGFP-N1	0.88
123	C. Kolb	Cbx2_AT1 <sup>G78R</sup> _AT2 <sup>G137R</sup>	Cbx2_mAT2_m2AT2	pEGFP_N1	1.13
131	C. Kolb	Hmga1a	Hmga1a (full length, + 11aa linker)	pEGFP_N1	1.04
132	C. Kolb	Hmga1b	Hmga1b (- 11aa linker AT1-AT2)	pEGFP_N1	1.13
59	C. Kolb	Cbx2_CD <sup>F12A</sup>	Cbx2_mCD (F12A)	pEGFP_N1	0.86
119	C. Kolb	Cbx2_CD <sup>F12A</sup> _AT1 <sup>AAA</sup>	Cbx2_mCD_mAT1	pEGFP_N1	1.07
120	C. Kolb	Cbx2_CD <sup>F12A</sup> _AT1 <sup>G78R</sup>	Cbx2_mCD_mAT2	pEGFP_N1	mini
133	C. Kolb	Cbx2_CD	Cbx2_CD (1-58aa)	pEGFP_N1	1.11
127	C. Kolb	Cbx2_CD+AT1	Cbx2_CD_AT1 (1-88aa)	pEGFP_N1	1.0/1.10
128	C. Kolb	Cbx2_CD+AT1-3	Cbx2_CD_AT1-3 (1-184aa)	pEGFP_N1	1.07
129	C. Kolb	Cbx2_dCD	Cbx2_dCD (67-519aa)	pEGFP_N1	1.09
130	C. Kolb	Cbx2_AT1-3	Cbx2_AT 1-3 (67-184aa)	pEGFP_N1	1.07

## Acknowledgements

First of all, I would like to thank Antoine Peters for being a great PhD supervisor. His continuous support and guidance throughout my PhD studies, and especially his endless ideas and enthusiasm, have made the time truly inspiring. Starting a new lab also means that many things need to be established from scratch, but Antoine initiated various collaborations that enabled us to collect the tools and gain all the necessary expertise required for my project.

I would like to thank the members of my PhD committee, Renato Paro, Frederick Meins and Patrick Matthias, for their helpful discussions and suggestions over the years. Thanks to Susan Gasser for acting as chair person on my final defense.

Special thanks to Carolin Kolb for helping me with various different projects and for coping with these vast amounts of mouse genotypings. Her great motivation, high efficiency and joyful attitude have made working together very pleasant and effective.

I would also like to thank Frederic Zilbermann for supporting me with the mouse work and for sharing the efforts in getting things going during the initial months in the lab.

The entire Peters' group is acknowledged for creating such a great and friendly atmosphere in the lab which made coming to work every day very enjoyable. I would also like to thank everyone for suggestions, comments, support and advice on many occasions.

Many thanks to Dirk Schübeler and his lab members for various helpful suggestions and stimulating discussions, especially during our joint lab meetings.

I wish to thank the embryo experts who shared their knowledge with me to help setting up the pre-implantation work in our lab. I am grateful to Pawel Pelczar and Petr Svoboda for getting me started. Special thanks to the Schultz lab in Philadelphia, U.S., and in particular Paula Stein for giving much needed advise. Many thanks also to Godfried van der Heijden, Alwyn Derijck and Peter de Boer in Nijmegen, The Netherlands.

My visit to the Koseki lab in Yokohama, Japan, will remain unforgotten. I would like to thank Kyo-Ichi Isono and Haruhiko Koseki for the fruitful collaboration, and the other lab members for turning my stay in Japan into such a special cultural and personal experience.

I thank Nicolas Thomä for many helpful discussions and for sharing his structural views on AT hooks.

I acknowledge excellent help from people at the FMI facilities. In particular, I would like to thank former FMI member Thierry Laroche for making the functional principles of top-end microscopes

accessible to each user. Thereafter, microscopy expertise was provided by Jens Rietdorf and Laurent Gelman, and Patrick Schwarb was very supportive in questions related to imaging analysis. Jean-Francois Spetz is acknowledged for his help in embryo micro-injection experiments, which could not have been performed without him. I would like to thank Michael Rebhan for bioinformatics advice. Moreover, Michael Stadler has greatly supported us in our bioinformatics analysis. Finally, I would like to thank Birgit Heller-Stilb and all the friendly animal keepers for taking great care of our mice.

The Boehringer Ingelheim Foundation (BIF) is acknowledged, not only for providing funding for my PhD studies, but also for encouraging and enabling the attendance of international scientific meetings and for creating such a great personal atmosphere at their annual seminars.

My warmest thanks go to Peter, to my family and to my friends who are always there to support me in every way possible.

# MAREIKE PUSCHENDORF



Home address: Rainallee 49, CH-4125 Riehen, Switzerland  
+41 61 6014975

Work address: Friedrich Miescher Institute  
Maulbeerstrasse 66, 4058 Basel, Switzerland  
+41 61 6976626

Email: [mareike.puschendorf@fmi.ch](mailto:mareike.puschendorf@fmi.ch)

Date of birth: 24.06.1980

Nationality: German

## EDUCATION

---

- since 05/2004 **Friedrich Miescher Institute (FMI), Basel, Switzerland**  
PhD studies in the laboratory of Dr. Antoine H.F.M. Peters  
"Epigenetic programming of mouse pre-implantation development"
- 07/2002 – 06/2003 **Curtin University of Technology, Perth, Australia**  
MSc (Biomedical Science), Dr. David Sayer and Prof. John Wetherall  
Royal Perth Hospital, Dept of Clinical Immunology & Biochem Genetics  
"Genetic study of the molecular heterogeneity of the complement C4 genes"
- 10/1999 – 06/2002 **Friedrich Schiller University (FSU), Jena, Germany**  
Biochemistry / Molecular Biology  
Degree: Vordiplom (1.7 – highest grade: 1.0; lowest pass: 4.0)
- 09/1995 – 07/1999 **Carl Zeiss Gymnasium, Jena, Germany**  
specialized in natural, mathematical and technical sciences  
Degree: Abitur (1.0 – highest grade: 1.0; lowest pass: 4.0)

## RESEARCH AND WORK EXPERIENCE

---

- 10/2003 – 02/2004 **Roche Diagnostics GmbH, R&D, Penzberg, Germany**  
Laboratory of Dr. Michael Tacke  
"Development of immune diagnostic methods for the early detection of colorectal cancer"
- 11/2001 – 06/2002 **Institute of Molecular Biology, FSU, Jena, Germany**  
Laboratory of Prof. Anne S. Ulrich  
"Characterization of lipopolysaccharides using solid state NMR methods"
- 09/2001 – 10/2001 **Institute of Genetics, University of Leipzig, Germany**  
Laboratory of Prof. Heinz Sass  
"Sequencing of the *klett* gene in *Drosophila virilis*"
- 10/2000 – 06/2001 **Clinical Centre, FSU, Jena, Germany**  
Laboratory of Prof. Stefan H. Heinemann  
"Electrophysical characterization of ion channel proteins upon heterologous expression in *Xenopus* oocytes"



## PUBLICATIONS

---

**Puschendorf M**, Terranova R, Boutsma E, Mao X, Isono K, Brykczynska U, Kolb C, Otte AP, Koseki H, Orkin SH, van Lohuizen M, and AHFM Peters (2008). PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos. *Nature Genetics*. doi:10.1038/ng.99.

**Puschendorf M**, Stein P, Oakeley EJ, Schultz RM, Peters AHFM, and P Svoboda (2006). Abundant transcripts from retrotransposons are unstable in fully grown mouse oocytes. *Biochem Biophys Res Commun*. 347(1): 36-43.

Windsor L, **Puschendorf M**, Allcock R, Scott A, Sayer D, Kucharzak R, Gut I, McCann V, Davis E, Witt C, Christiansen F, and P Price (2005). Does a central MHC gene in linkage disequilibrium with HLA-DRB1\*0401 affect susceptibility to type 1 diabetes? *Genes Immun*. 6(4): 298-304.

## PRESENTATIONS

---

- 04/2008            **Talk - Keystone Symposium on Chromatin**, Snowmass, U.S.
- 09/2007            **Talk - EPITRON Brainstorming Meeting**, Crete, Greece
- 05/2007            Poster - EMBO Chromatin and Epigenetics, Heidelberg, Germany
- 01/2007            **Talk - TriRhena Chromatin & Transcription Meeting**, Strasbourg, France
- 12/2006            **Talk - Abcam Chromatin Structure & Function Meeting**, Punta Cana, Dominican Republic
- 03/2006            Poster - Novartis Epigenetics Symposium, Switzerland
- 11/2005            Poster - Germ Cells, Epigenetics & Reprogramming, Kyoto, Japan
- 09/2005            **Talk - FMI Annual Meeting**, Pontresina, Switzerland
- 10/2004            Poster - Novartis Corporate Research Conference, Boston, USA

## SCHOLARSHIPS

---

- 04/2008            **EPIGENOME Network of Excellence Travel Award**
- 09/2007            **EPITRON External Travel Grant**
- 11/2006            **EPIGENOME Network of Excellence Travel Award**  
Research visit to the laboratory of Prof. Haruhiko Koseki at the RIKEN Research Center for Allergy and Immunology, Yokohama, Japan
- 05/2005            **Boehringer Ingelheim Fonds (BIF) Travel Grant**  
Research visit to the laboratory of Dr. Peter de Boer at the Nijmegen Centre for Molecular Life Sciences, Nijmegen, The Netherlands
- 04/2005 – 03/2007    **Boehringer Ingelheim Fonds (BIF) PhD Scholarship**
- 02/2005            **Development Traveling Fellowship**  
Research visit to the laboratory of Prof. Richard M. Schulz at the University of Pennsylvania, Philadelphia, U.S.
- 05/2004            **FMI International PhD Program Scholarship**