The role of Polycomb group proteins in mouse pre-implantation development

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Julia Bammer

aus Österreich

Basel, 2018

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Prof. Dr. Antoine Peters
(Fakultätsverantwortlicher und Referent)

Prof. Dr. Petra Hajkova
(Koreferent)

Basel, den 16.Oktober 2018

Dekan

Prof. Dr. Martin Spiess

Table of contents

Summary **4**

List of Abbreviations **5**

Chapter 1. Introduction and scope of the thesis **7**

1.1 Epigenetics *7*

1.2 Chromatin *7*

1.3 Polycomb-mediated repression *10*

1.4 Pre-implantation development as a system to study chromatin dynamics *30*

1.5 Scope of the thesis *40*

Chapter 2. Manuscript in preparation: PRC2 regulates gene expression and dosage compensation during pre-implantation development **43**

2.1 Abstract *43*

2.2 Introduction *44*

2.3 Results *48*

2.4 Discussion *58*

2.5 Material and Methods *60*

2.6 Figures *68*

Chapter 3. Manuscript in preparation: A non-canonical PRC1 complex serves as a backup for DNA methylation to repress germ line genes in pre-implantation embryos **94**

3.1 Abstract *94*

3.2 Introduction *95*

3.3 Results *96*

3.4 Discussion *101*

3.5 Material and Methods *102*

3.6 Figures *107*

Chapter 4. General discussion and outlook **120**

4.1 Main findings *120*

4.2 Role of PRC2 in gene expression and dosage compensation in early embryos *121*

4.3 Role of PRC1 in the repression of germ line genes in pre-implantation embryos *125*

4.4 Single-cell and single-embryo heterogeneity *126*

References **129**

Acknowledgements **149**

Summary

At the onset of life, two highly specialized gametes fuse to give rise to a totipotent embryo, which has the potential to differentiate into all tissues of the adult organism. This process is governed by transcriptional networks that arise during pre-implantation development and is orchestrated by dynamic changes in DNA methylation, chromatin organization and histone modifications. In my thesis, I addressed the role of the evolutionarily conserved Polycomb group (PcG) proteins for embryonic development and gene regulation during mouse pre-implantation development.

We demonstrated that genetic ablation of a core component of Polycomb repressive complex 2 (PRC2) affects gene expression and dosage compensation in early embryos. We have shown that besides dynamically regulating genes involved in development and differentiation at different stages of pre-implantation development, PRC2 also represses a subset of genes specifically on the maternal allele in early embryos. This allele-specific gene repression depends on the function of PRC2 in the oocyte, indicating that at these genes the H3K27me3 mark is inherited from oocytes to early embryos. We found that among the genes regulated in this fashion is the X inactivation regulator *Xist*, suggesting that H3K27me3 is the maternal repressive imprint that prevents inactivation of the maternal X chromosome in pre-implantation embryos. I have shown that although embryos maternally and zygotically deficient for PRC2 are able to develop to the blastocyst stage, they exhibit differences in the expression of lineage markers. Moreover, I demonstrated that maternal deletion of PRC2 impairs post-implantation development.

We further discovered that Polycomb repressive complex 1 (PRC1) represses a set of genes with functions in germ line development in early embryos. In embryonic stem cells (ESCs), these genes are regulated by DNA methylation as well as the variant PRC1 complex PRC1.6, suggesting that in pre-implantation embryos, where DNA methylation levels are low, PRC1.6 functions as a backup mechanism for repression of germ line genes. We are currently validating our hypothesis in embryos deficient for components of the PRC1.6 complex.

In summary, our work demonstrates that Polycomb group proteins function as dynamic regulators of gene expression in pre-implantation embryos *in vivo*.

List of Abbreviations

5mC	5-methylcytosine
5hmC	5-hydroxymethylcytosine
AG	Androgenote
ATAC-seq	Assay for transposase accessible chromatin with high-throughput sequencing
ATP	Adenosine triphosphate
BSA	Bovine serum albumin
CGI	CpG island
ChIP	Chromatin immunoprecipitation
ChIP-seq	Chromatin immunoprecipitation with high-throughput sequencing
DKO	Double knock out
DMR	Differentially methylated region
DNA	Deoxyribonucleic acid
EPI	Epiblast
ESC	Embryonic stem cell
F0	Parental generation
F1	First filial generation
F2	Second filial generation
Fg	Fully grown (oocyte)
HMTase	Histone methyl transferase
GG	Gynogenote
GO	Gene ontology
GV	Germinal vesicle (oocyte)
hCG	Human chorion gonadotropin
HCP	High CpG
ICM	Inner cell mass
IF	Immunofluorescence
KO	Knock out
LCP	Low CpG
Log2(FC)	Log2 (Fold change)
MII	Metaphase II (oocyte)
MuERV-L	Murine endogenous retrovirus-like
ncRNA	Non-coding RNA
Ng	Non-growing (oocyte)
NSN	Non-surrounded nucleolus (GV oocyte)
ORF	Open reading frame
PBS	Phosphate buffered saline
PCA	Principal component analysis
PcG	Polycomb group
PE	Primitive endoderm

PFA	Paraformaldehyde
PhoRC	Pho repressive complex
PMSG	Pregnant mare serum gonadotropin
PRC1	Polycomb repressive complex 1
PRC2	Polycomb repressive complex 2
PRE	Polycomb response element
PTM	Post-translational modification
RNA	Ribonucleic acid
RNA Pol II	RNA Polymerase II
RNA-seq	RNA sequencing
RT	Room temperature
SAM	Sterile alpha motif
SN	Surrounded nucleolus (GV oocyte)
TAD	Topologically associated domain
TE	Trophectoderm
TF	Transcription factor
TrxG	Trithorax group
t-SNE	t-distributed stochastic neighbor embedding
TSS	Transcription start site
Wt	Wild type
XCI	X chromosome inactivation
XIC	X inactivation center
Xm	maternal X chromosome
Xp	paternal X chromosome
ZGA	Zygotic genome activation

Chapter 1. Introduction and scope of the thesis

1.1 Epigenetics

How does a single genotype, the same information that is encoded in the DNA of each cell, give rise to the plethora of different cell types and traits in the organism?
In order to describe this concept, the developmental biologist Conrad Hal Waddington in 1942 coined the term “epigenetics”, which he defined as “the branch of biology which studies the causal interactions between genes and their products, which bring the phenotype into being” [1]. Waddington used the model of the “epigenetic landscape” to illustrate the process of cellular decision-making during development. This visual metaphor pictures the cells as marbles rolling down a hill. At various points in this landscape, the marbles will sample the grooves on the hill and take specific permitted trajectories to reach the lowest points. These points represent the eventual cell fates that a cell can differentiate into [2].

To date, epigenetics is more specifically defined as the study of any stable and heritable changes in gene expression that occur without changes in the DNA sequence. However, there is still considerable disagreement about what is included in the definition. According to the strict definition, an epigenetic phenomenon has to be inherited over multiple rounds of mitotic division in absence of the initial trigger. Since it has been inherently difficult to differentiate between real self-propagation of epigenetic states that is independent of the underlying DNA sequence and re-establishment after mitosis via sequence-specific recruitment of enzymes, many chromatin modifications have been called epigenetic in absence of the formal proof that they are truly heritable [3].

1.2 Chromatin

Chromatin is the organizational form of DNA in the nucleus of eukaryotes. Its fundamental repeating subunit is the nucleosome core particle, which consists of two copies each of four histone proteins (H2A, H2B, H3 and H4), the so-called histone octamer, that packages 147bp of DNA in 1.65 turns [4]. A ‘linker’ region of DNA of variable length between the core particles is associated with histone H1. The interactions between H1 molecules facilitate the assembly of the nucleosomal array into a chromatin fiber, which is then further packaged into a higher order structure [5, 6]. Although chromatin certainly has a role in compaction, it has long been

speculated that its function is more than just packaging of DNA. The first evidence for the impact of chromatin on transcription came from the discoveries that nucleosomes interfere with transcription *in vitro* and that deletions of histones or histone tails affect gene expression *in vivo* [7-10]. More recently, it has become evident that the nucleosome is not a static packaging element, but instead has dynamic features that are tightly controlled by different protein complexes.

Both the N-terminal histone tails protruding out from the nucleosome as well as the globular histone domains are subject to post-translational modifications (PTMs). These include methylation of arginines (R), methylation, acetylation, ubiquitination, ribosylation and sumoylation of lysines (K) and phosphorylation of serines (S) and threonines (T) [reviewed in 11]. They are deposited by specific enzymes, so called "Writers". Some of these modifications, including acetylation of H3 and H4 or di- and tri-methylation of H3K4 are marks of active transcription and are thus termed euchromatic modifications. Other modifications, including methylation on H3K9 and H3K27, are associated with repressive functions and are termed heterochromatic modifications. Most of these modifications have characteristic localization patterns at the upstream region, promoter, 5' or 3' end of genes, which are important for their role in transcriptional regulation [reviewed in 12]. But how do histone modifications regulate transcription of genes? First, except for methylation, histone modifications change the net charge of nucleosomes, which can loosen the interaction between DNA and histones. Second, they can be read by other proteins, so-called "Readers". It has been suggested that individual histone modifications or combinations of different modifications can recruit different reader proteins that can influence chromatin dynamics and function, leading to the "histone code" hypothesis [13, 14].

Another way of transcriptional regulation on the level of chromatin is by action of specialized ATP-dependent chromatin-remodeling complexes. They can be grouped into different subfamilies with distinct functions that are all based on their ability to utilize ATP hydrolysis to translocate DNA from the nucleosome. One group of chromatin remodelers ensures the proper density and spacing of nucleosomes, thereby possibly contributing to gene repression. A second set of remodelers slides or ejects nucleosomes in order to facilitate binding of transcription factors (TFs) to DNA. A third set of remodelers is involved in the replication-independent removal of histones and their replacement with histone variants in order to generate specialized chromosomal regions [reviewed in 15].

Therefore, the third layer of transcriptional regulation via chromatin is the incorporation of histone variants. Initially it has been thought that the core histones, that are synthesized in S-phase, are the universal components of all nucleosomes [reviewed in 16]. However, it has later been revealed that next to the canonical histones there are variant forms of histones, so-called replacement histones, which are expressed outside S phase and incorporated into chromatin independent of replication [reviewed in 17]. Differences between canonical and variant histones range from only a few amino acids to more divergent changes in the histone tails or histone fold domains. Some histone variants differ in their biophysical characteristics and can thus influence the properties of nucleosomes, while others are incorporated at specific locations in the genome [reviewed in 12]. One of the best studied examples is the H3 variant H3.3, which is expressed throughout the cell cycle and is localized in transcriptionally active regions of the genome including promoters, enhancers and gene bodies [18]. H3.3 differs from the canonical histones H3.2 in 4 and from H3.1 in 5 positions. The majority of these changes are located in the histone fold domain and the affected residues are involved in the deposition of H3.3 outside of replication in *Drosophila* as well as mammals [18, 19]. This sequence-specific deposition is achieved by interactions with different assembly machineries. While the chaperone CAF-1 deposits canonical H3 during S phase, HIRA and DAXX selectively bind amino acids specific for H3.3 and mediate deposition in a replication-independent manner [20-24]. The relatively minor differences between canonical H3 and H3.3 make it unlikely that H3.3 incorporation causes major changes in chromatin structure or stability. Instead, it has been shown that H3.3 interferes with higher order chromatin folding by counteracting incorporation of the linker histone H1. The mechanism of how H3.3 restricts H1 binding remains unknown, but several possible mechanisms have been proposed. First, amino acids and posttranslational modifications that differ between H3 and H3.3 may regulate binding of H1 to the nucleosome. A second possibility is that H1 binding is influenced by reduced nucleosome stability upon H3.3 incorporation. Third, H3.3 may promote binding of other proteins like chromatin remodelers, which have been shown to affect localization of H1 [25]. Incorporation of H3.3 promotes a more open chromatin structure and is thus generally considered as a mark of transcriptional activation. H3.3 also carries an increased proportion of histone marks associated with active transcription, such as acetylation and H3K4 methylation [26, 27]. However, it remains unknown whether this is an inherent property of H3.3 or a secondary effect resulting from

the preferred localization of the histone variant at sites of active transcription. In addition, H3.3 is also incorporated at promoters of inactive genes that were previously active or that are poised for transcription, thereby forming a memory system of active transcriptional states [28, 29]. Interestingly, upon fertilization, after decondensation of the highly compacted sperm chromatin and removal of protamines, nucleosomes containing H3.3 are incorporated into paternal chromatin prior to the first round of DNA replication [30, 31]. In addition, H3.3 is specifically enriched in the paternal pericentromeric heterochromatin of the male pronucleus [32].

1.3 Polycomb-mediated repression

Polycomb group (PcG) proteins are an evolutionarily conserved group of chromatin regulators, which play a crucial role in gene regulation during differentiation and development of multicellular organisms. They regulate transcription by conferring a repressive chromatin state through their capacity to catalyse histone modifications and chromatin compaction [33]. PcG proteins were first characterized in *Drosophila melanogaster* as repressors of Hox genes based on the observation that PcG mutants exhibit posterior transformations of body segments [34]. During *Drosophila* embryonic development, the activity of maternally and zygotically provided TFs sets up a distinct pattern of Hox gene expression that defines the identity of each segment of the body. This segment-specific expression of Hox genes persists throughout fly development, even after disappearance of the early transcriptional regulators. It has been shown that this cellular memory system depends on the antagonistic functions of the PcG and Trithorax (TrxG) proteins. While PcG proteins are required for maintenance of Hox gene repression, TrxG proteins are involved in maintaining them active [reviewed in 35]. A study using TrxG and PcG double mutants revealed that the TrxG histone methyl transferases (HMTases) Trx and Ash1 are not 'coactivators' functioning in transcriptional activation of Hox genes, but rather function as 'anti-repressors', preventing PcG-mediated repression of Hox genes in their physiological expression domains [36]. It has further been shown that this anti-repressive function of Ash1 is due to its ability to prevent trimethylation by PRC2 [37].

The function of PcG and TrxG proteins is conserved in vertebrates, where mutations in PcG genes typically result in homeotic transformations of vertebrae [38-41]. More recently,

genome-wide studies in flies and mammals have revealed a more dynamic role of PcG proteins, regulating a plethora of genes that control key developmental processes. PcG proteins are generally found as components of two major classes of large multimeric complexes, the Polycomb repressive complexes 1 and 2 (PRC1 and PRC2) (Fig. 1.1).

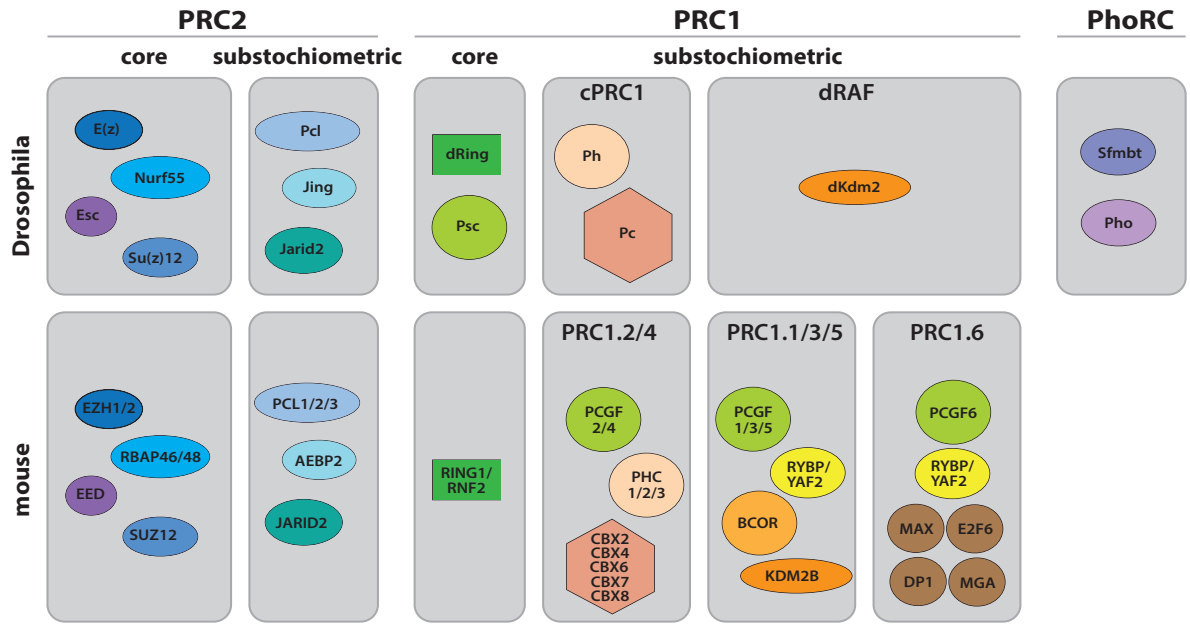


Figure 1.1. Polycomb complex composition in fly and mouse. PcG proteins are divided into two major groups based on their incorporation into distinct multimeric complexes, called Polycomb repressive complexes 1 and 2 (PRC1 and PRC2). These complexes can be subdivided into different sub-complexes, which share core components, but differ in additional components that have been identified in the complexes in substochiometric amounts. In *Drosophila*, the Pho repressive complex (PhoRC), which is not conserved in mammals, has DNA-binding activity and is important for recruitment of PRC complexes to Polycomb response elements (PREs). Homologues between *Drosophila* and mouse are depicted in the same colors.

1.3.1 PRC2 and H3K27me3

Drosophila PRC2 consists of the key components Enhancer of zeste (E(z)), Extra sex combs (Esc), Suppressor of zeste 12 (Su(z)12) and the nucleosome remodeling factor 55 (Nurf55). PRC2 deposits repressive methylation on Lysine 27 of histone H3 (H3K27me3) via the SET domain of E(z), which exerts methyltransferase activity [42, 43]. However, E(z) alone is inactive and needs to be in association with Esc and Su(z)12 to exert its catalytic activity [44-46]. Besides these key components, several other proteins have been identified in *Drosophila* PRC2 in substochiometric amounts. One of these components is Polycomb-like (Pcl), which has been shown to promote the addition of the last methyl group in order to generate H3K27me3 [47].

The mammalian PRC2 complex consists of the homologous enzymes Enhancer of zeste 1 (EZH1) and Enhancer of zeste 2 (EZH2), the scaffold protein Suppressor of zeste 12 (SUZ12), the chromatin-binding protein Embryonic ectoderm development (EED) and the histone binding proteins RBAP46/48. EZH1 and EZH2 contain a catalytic SET domain, which successively methylates lysine 27 on histone H3 (H3K27me1, me2 and me3) [42]. Furthermore, *in vitro* assays revealed the ability of PRC2 to methylate histone H1 on lysine 26, yet the biological significance of this remains elusive [48, 49]. EZH1 and EZH2 proteins alone are inactive and rely on the presence of SUZ12 and EED for their catalytic activity [50, 51]. Therefore, these three proteins constitute the key PRC2 components, which are both necessary and sufficient for methyltransferase function. EED contains a WD-repeat structure, consisting of seven copies of the WD-repeat motif forming a seven-bladed propeller structure that serves as a binding platform for EZH2 [52]. As a product of alternative transcriptional start sites, there are four isoforms of mammalian EED. They have been shown to be associated with different PRC2 functions and direct the HMT activity towards H1K26 or H3K27 [48, 49]. SUZ12, the third core component of PRC2, contains a VEFS and a C2H2 Zinc finger domain. The VEFS domain is responsible for binding to EZH1 or EZH2 and functions as an allosteric regulator of HMT activity [50].

In addition to the core components of PRC2, a number of other, substochiometric proteins, including AEBP2, PCL1/2/3 and JARID2, have been found as part of mammalian PRC2 complexes. AEBP2 is a zinc-finger protein that co-localizes with PRC2 at some of its targets and has been shown to interact with several PRC2 subunits to enhance enzymatic activity of the complex [44, 53]. PCL1, PCL2 and PCL3 (also referred to as PHF1, MTF2 and PHF19, respectively) are the mammalian orthologues of Pcl in the fly. They interact with PRC2 mainly through EZH2 and co-occupy target genes [47, 54]. Mammalian PCLs are expressed in a tissue-specific manner and their function in the PRC2 complex is not yet fully understood [54]. Yet, it has been suggested that they play a role in regulation of enzymatic activity and PRC2 recruitment in mouse and *Drosophila* [54-56]. PCL1 and PCL2 have been shown to contain extended homologous regions that bind to unmethylated CpG motifs and are required for efficient targeting of PRC2 to CpG island-containing promoters in embryonic stem cells (ESCs) [57]. JARID2 is part of the Jumonji family of histone demethylases, yet is lacking the amino acids required for enzymatic activity [58]. It has been shown to interact with EZH2 and its binding overlaps largely with PRC2 binding. JARID2 stimulates enzymatic

activity of PRC2 *in vitro* and *in vivo* and is required for PRC2 recruitment, although its absence has limited influence on H3K27me3 deposition [59-66]. A recent study has revealed that both JARID2 and PCL2 are involved in *de novo* establishment of PRC2 domains by increasing the stability of PRC2 at its nucleation sites, thereby allowing the formation of H3K27me3-marked Polycomb foci within the nucleus from which H3K27me2 and H3K27me3 spread across the genome [67].

The molecular function of the H3K27me3 modification is still not fully understood. In flies, it has been shown that a point mutation in lysine 27 of H3 fails to repress transcription of PcG target genes and causes homeotic transformations mimicking PRC2 mutant phenotypes. These results indicate that in flies, H3K27 is the crucial PRC2 substrate that is responsible for gene repression [68]. However, another study in *Drosophila* has shown that H3K27me3 is not the initial cue at the base of the PcG recruitment hierarchy, but plays a role in stabilizing the interaction of PcG complexes at Polycomb response elements (PREs) with surrounding gene regulatory elements [69]. In addition, H3K27me3 has been shown to be recognized by the aromatic cage at the center of the WD-propeller structure of EED *in vitro*, possibly allowing maintenance and propagation of H3K27me3 at sites that already carry the PRC2 mark. Interestingly, binding of Esc, the *Drosophila* homologue of EED, to H3K27me3 promotes the lysine methyltransferase activity of E(z). This self-reinforcing, positive feedback loop might contribute to the maintenance of transcriptionally silent chromatin domains and the transmission of the H3K27me3 mark from mother to daughter cells [70]. It has recently been demonstrated that the aromatic cage of EED is required for the efficient catalysis of di- to trimethylated H3K27 at PRC2 nucleation sites via its ability to bind to H3K27me3, thereby promoting the efficient spreading of PRC2 domains [67].

1.3.2 PRC1 and H2AK119ub

The *Drosophila* PRC1 core complex consists of Polycomb (Pc), Polyhomeotic (Ph), Posterior sex combs (Psc) and Sex combs extra (Sce/dRing) [71]. *In vitro* assays have shown that the reconstituted PRC1 complex inhibits chromatin remodeling by SWI/SNF and restricts access by RNA Polymerase II (RNA Pol II) [71, 72].

The mammalian PRC1 core complex contains the same subunits, but the genes encoding them have been subject to duplications [73]. The mammalian homologues of dRing,

RING1 (RING1A) and RNF2 (RING1B) contain a RING domain, which confers E3 ubiquitin ligase activity, resulting in mono-ubiquitination of histone H2A at lysine 119 (H2AK119ub1) [74-76]. RING1/RNF2 associates with one of the six alternative PcG RING finger proteins (PCGF1-6), which are homologues of fly Psc. The RING1/RNF2-PCGF heterodimers constitute a minimal core of the PRC1 complex that can deposit H2AK119ub1. The identity of the PCGF protein determines further composition of the PRC1 complex, leading to different molecular functions of the PRC1 variants.

PCGF2 (MEL18) and PCGF4 (BMI1) contain RING domains and are required to enhance the catalytic activity of RING1/RNF2 [77-79]. They are incorporated in so-called canonical PRC1 complexes (also known as PRC1.2 and PRC1.4). These complexes further contain polyhomeotic (PHC) and chromobox-containing protein (CBX) subunits. The CBX (CBX2, CBX4, CBX6, CBX7 and CBX8) subunits are homologues of *Drosophila* Pc and contain a well-conserved chromodomain. They have been shown to provide binding affinity of PRC1 through binding of DNA and methylated histone residues [80-82]. The distinct CBX proteins differ in their specificity towards methylation of H3K9 or H3K27, with CBX2 and CBX7 binding to both H3K9me3 and H3K27me3, CBX4 preferably binding to H3K9me3 and CBX6 and CBX8 not binding significantly to either modification [80]. PHC proteins (PHC1, PHC2 and PHC3) mediate chromatin compaction via polymerization of the SAM (sterile alpha motif) domain, thereby facilitating transcriptional silencing [83].

Heterodimers containing PCGF1, 3, 5 or 6 constitute so-called variant or non-canonical PRC1 complexes. They all contain Ring and YY1 binding protein (RYBP) or YY1-associated factor 2 (YAF2), but differ greatly in additional subunits [84, 85]. PCGF1-containing PRC1 (PRC1.1) contains KDM2B, a histone H3K36 demethylase. KDM2B is able to bind unmethylated CpG sequences via its CXXC zinc-finger domain leading to the recruitment of variant PRC1 to CpG islands (CGIs) [86, 87]. PCGF6-containing PRC1 (PRC1.6) contains the transcriptional repressor E2F6, the oncoprotein L3MBTL2, the H3K9me3-binding protein CBX3, the histone deacetylases HDAC1 and HDAC2, the H3K9 methytransferase G9A, the H3K4 demethylase JARID1A, the WD40 repeat protein WDR5 and the TFs MAX, MGA and DP-1 [84, 88-92]. It has been shown that this PRC1 variant binds DNA via a heterodimer of MGA and MAX to directly recognize and repress germ cell-related genes in ESCs [93].

The H2AK119ub1 mark was initially believed to have a key function in mediating PcG repression by interfering with different steps of transcription. It has been suggested that at

bivalent promoters, H2AK119ub1 restrains the activity of poised RNA Pol II [94]. Another study has shown that it prevents recruitment of the FACT complex, hindering the release of RNA Pol II at the start of transcriptional elongation [95]. In addition, *in vitro* experiments have suggested that the presence of H2AK119ub1 prevents di- and trimethylation of H3K4, leading to a block in transcriptional initiation [96]. On the contrary, two recent studies in flies and mouse came to the conclusion that H2A ubiquitination is not necessary for PcG repression [97, 98]. However, the study in mammals was performed in ESCs and mice carrying a point mutation in RNF2 that ablates its catalytic activity, but a wildtype allele of RING1, which can also catalyse H2AK119ub and could potentially compensate for the catalytic inactive RNF2 [98]. A study in which mutant RNF2 defective in the interaction with the E2 component was expressed in RING1/RNF2 double knock out (DKO) ESCs showed that the enzymatic activity of PRC1 is dispensable for target binding and compaction of Hox loci, but required for maintenance of ES cell identity via proper repression of target genes [99]. In line with these results, it has been shown that PRC1 produces discrete compacted chromatin domains, which change during ESC differentiation and are independent of ubiquitination. Instead, their formation depends upon the PHC component of canonical PRC1 [100]. These results suggest that PRC1-mediated repression depends on the concerted action of multiple effector mechanism including its enzymatic activity as well as its function in chromatin compaction.

1.3.3 Targeting of Polycomb complexes

A crucial question in the field of Polycomb is how PcG complexes are recruited to chromatin and how the different PRCs interact with each other. In *Drosophila,* both PRC1 and PRC2 are recruited to their target genes by specific DNA elements with a length of a few hundred base pairs, so-called Polycomb response elements (PREs) [101-105]. PREs can be located several kilobases up- or downstream of the gene promoter, within introns or in close proximity to the transcription start site (TSS) [106]. They frequently contain binding sites for Pleiohomeotic (Pho), PHO-like (Phol), Trithorax-like (Trl), Dorsal switch protein 1 (Dsp1) and other DNA-binding factors [105, 107, 108]. Pho and the Scm-related gene containing four MBT domains (Sfmbt) form a heterodimeric complex, called Pho repressive complex (PhoRC) [109]. Although this complex has a central role in recruitment of PcG proteins, there

are many loci that bind PcG proteins in the absence of PhoRC, suggesting that multiple factors may cooperate to recruit PcG complexes to PREs [110].

In contrast to *Drosophila,* attempts to find PREs in mammals have mostly failed and only very few PRE-like elements have been identified. These exceptions are a 1.5kb DNA element of the mouse *MafB/Kreisler* locus that can recruit PRC1, but not PRC2 and a 1.8kb fragment of the human *HOXD* cluster, which recruits both PRC1 and PRC2 [111, 112]. Instead, emerging evidence suggests that mammalian PcG complexes are recruited to their DNA targets by a combination of locus-specific and more generic targeting mechanisms (Fig. 1.2A).

Inspired by the function of TFs in PcG recruitment in *Drosophila,* it has been proposed that mammalian TFs might also play a role in recruiting PcG complexes to chromatin. However, only few site-specific TFs have been identified as PcG binding proteins using unbiased biochemical approaches. These include E2F, MAX and MGA, which have been found to interact with PRC1 [84, 88]. Further, REST, SNAIL and RUNX1 have also been identified to interact with Polycomb complexes in candidate-based screens. However, it has also been demonstrated that these TFs contribute to Polycomb targeting only in specific cases and at a subset of their target loci [113-117]. In recent years, several studies have demonstrated a role of non-coding RNAs (ncRNAs) in recruitment of PcG complexes, serving either as a scaffold for complex assembly or as a targeting factor. The human 2.2kb ncRNA HOX transcript antisense RNA (*HOTAIR*) binds to PRC2 and the Co-REST complex containing the H3K4 demethylase KDM1A (LSD1) and targets both complexes in *cis* to *HOXC* genes and in *trans* to *HOXD* genes [118, 119]. The ncRNA-based recruitment of PRC2 also has an important function for the establishment of mammalian X chromosome inactivation (XCI). The 17kb ncRNA *X-inactive specific transcript* (*Xist*) contains a 28bp repeat element (*RepA*) that interacts with EZH2 and initiates the process that leads to spreading of PRC2 binding in *cis* along the inactive X chromosome [120].

Even though it has clearly been demonstrated that TFs and ncRNAs can contribute to locus-specific targeting of PcG complexes in specific instances, it is unlikely that these mechanisms are sufficient to recruit PcG complexes to all of their targets genome-wide. It is highly likely that more generic targeting mechanisms are involved in setting up the extensive and in many cases tissue-specific binding patterns found in different cell types *in*

vivo. One of the possible generic targeting mechanisms also involves RNAs and comes from the observation that PRC2 binds RNA promiscuously [121, 122]. This has led to different hypotheses of how the interaction between PRC2 and RNA could play a role in PcG recruitment or function. One idea is that binding of PRC2 to short transcripts that arise from aberrant transcription events may help to retain PcG complexes and ensure stable repression [123, 124]. Alternatively, PRC2 might interact with nascent transcripts arising from active genes, which serve as a decoy to prevent stable binding of PRC2 to chromatin, thereby preventing repression of active genes [125, 126]. A third possibility is that PRC2 binds to nascent transcripts to increase its dwell time around promoters genome-wide and that counteracting active signals are required to prevent the formation of stable PcG complexes at transcribed genes [51]. Therefore, the exact role of RNAs in the targeting of PcG complexes remains controversial and further studies will be needed to reveal the mechanism and relevance of RNA-mediated PcG recruitment.

In general, most mammalian PcG complexes bind target genes over a broad region in proximity to the TSS, suggesting that there is no specific recruiting sequence. It has been shown that in ESCs, domains that are bound by both PRC1 and PRC2 correspond well with unmethylated and transcriptionally inactive CGIs [127, 128]. Comparison across different species further supports the hypothesis that clusters of CpGs lacking DNA methylation and active transcription can recruit PcG complexes [129]. As mentioned above, KDM2B, a component of the variant PRC1.1 complex, is able to bind to unmethylated CpG dinucleotides via its CXXC domain and therefore recruits PRC1.1 to CGIs genome-wide [86, 87, 130]. However, loss of KDM2B does not prevent binding of PRC1 at all CGI targets, suggesting that there are different ways how PcG complexes can be recruited to CGIs [86, 87, 131]. Moreover, although KDM2B binds to virtually all CGIs, only around 30% of them show recruitment of PcG complexes in ESCs [132, 133]. This fraction of bound CGIs represents promoters of repressed genes and inhibition of transcription in ESCs induces genome-wide ectopic recruitment of PcG complexes to newly silenced CGIs [128]. These results suggest a “chromatin sampling” model, which proposes weak interactions of PcG proteins with all potential binding sites, but at sites of active transcription stable binding is prevented by transcription itself or by the binding of activating TFs [134]. Recently it has been shown that also PRC2 can be recruited to CGIs independently of KDM2B. Using a

system where PRC2-mediated domains were disrupted and restored in an inducible manner in ESCs, Oksuz *et al.* revealed that PRC2 is stably recruited to a subset of CGIs, so-called PRC2 "nucleation sites", by JARID2 and PCL2 [67].

A third factor involved in generic PcG recruitment is the recognition of pre-existing histone modifications either catalysed by other histone-modifying complexes or PcG complexes themselves. It has been demonstrated that the H3K36me2/3 marks, which are usually associated with transcribed regions or sites of DNA damage, can contribute to PcG targeting. The PRC2 subunit PCL3 (PHF19) binds to H3K36me2 and me3 via its Tudor domain, facilitating the re-silencing of transcribed genes in ESCs [135]. On the other hand, it has been shown that H3K36me2/3 inhibits the catalytic function of PRC2 *in vitro* [51]. Polycomb recruitment has also been linked with H3K9 methylation. The PRC2 component EED has been shown to bind the H3K9me3 mark *in vitro* [70]. Further, PRC2 directly interacts with the H3K9 methyltransferases GLP and G9A and it has been shown that their deletion in ESCs leads to a reduction in PRC2 binding to chromatin [136].

Taken together, a complex combination of factors is involved in the recruitment of Polycomb complexes in mammals. Although TFs and ncRNAs can target PcG complexes to specific genomic regions, recent evidence highlights the importance of more generic targeting mechanisms including the interaction of PcG complexes with RNAs, unmethylated CGIs and chromatin modifications, either deposited by other complexes or PcG complexes themselves.

1.3.4 Hierarchy and interaction of Polycomb complexes

It has been shown that H3K27me3 increases the affinity for canonical PRC1 complexes to chromatin [137]. Therefore, over the last decade, the predominant view for Polycomb-mediated transcriptional repression has been a strict hierarchical model for recruitment of PRCs to chromatin, where binding of PRC2 catalyses H3K27me3, which then leads to recruitment of canonical PRC1 via the chromodomain of CBX and H2A monoubiquitination [138] (Fig. 1.2B). In this classical model, transcriptional repression is achieved through H2A monoubiquitination and chromatin compaction, whereas H3K27me3 ensures epigenetic maintenance of the repressed state.

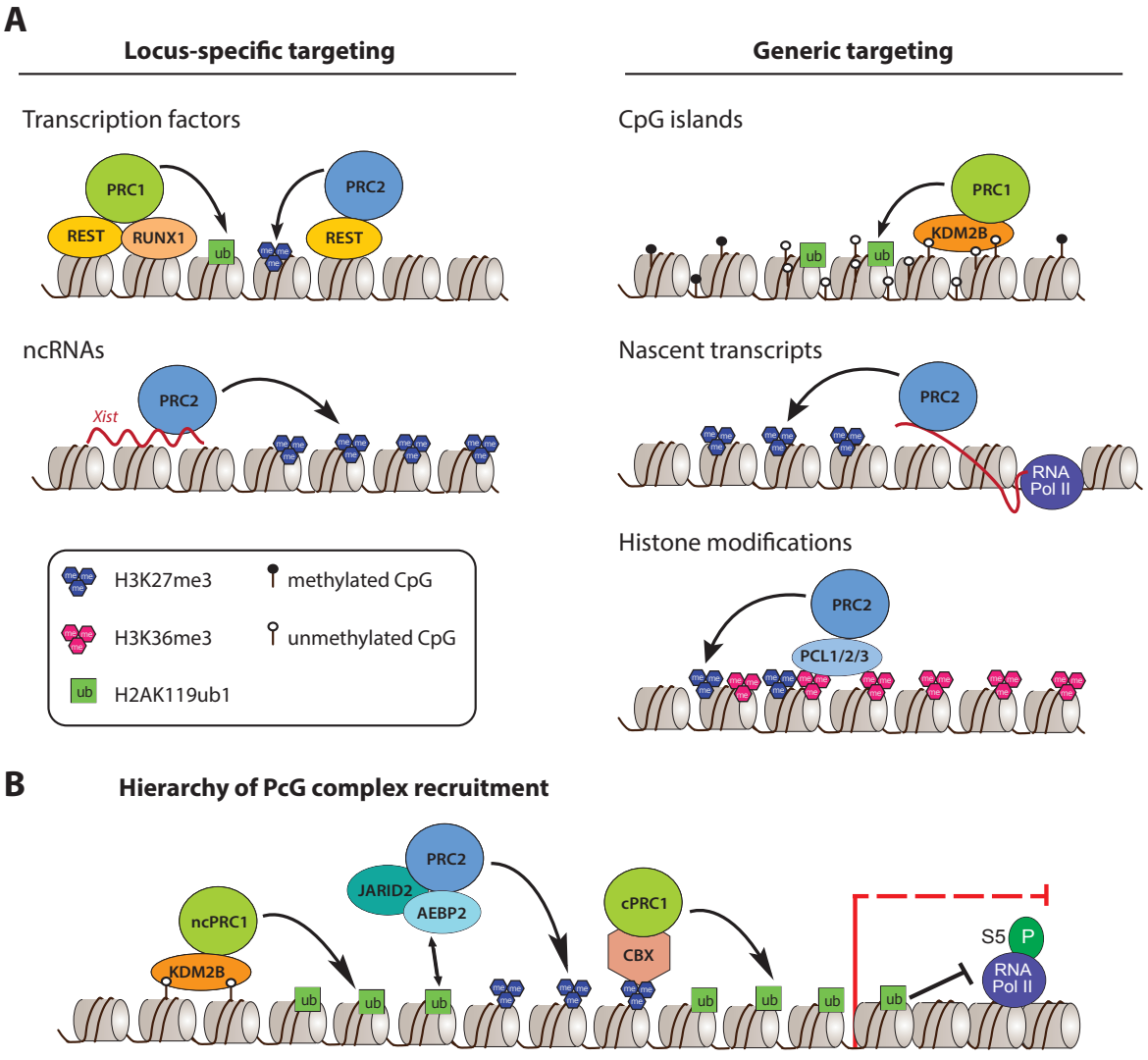


Figure 1.2. Mechanisms of Polycomb recruitment and interaction. A) PcG complexes can be targeted to chromatin through both locus-specific and generic targeting mechanisms. Locus-specific recruitment can be achieved through association of PRCs with transcription factors such as REST or RUNX1 or non-coding RNAs such as *Xist*. Generic targeting mechanisms include binding of a non-canonical PRC1 complex to unmethylated CpGs via KDM2B. Further, PRC2 has been shown to bind RNA promiscuously. One hypothesis of how this interaction contributes to PcG targeting is that the binding of PRC2 to nascent RNA Pol II transcripts at 5' ends of genes may help to retain PcG complexes and stable repression of silent genes following stochastic transcription initiation events. The third generic targeting mechanism is the interaction of PRC complexes with pre-existing histone modifications. A subset of PRC2 complexes contains PCL proteins which bind to H3K36me3, a mark for active transcription, possibly allowing PcG domains to spread into previously transcribed regions. **B) Model for the hierarchical recruitment of PcG complexes.** Non-canonical PRC1 containing KDM2B is recruited to unmethylated CpG dinucleotides and deposits H2AK119ub1. This mark can be recognized by PRC2 complexes containing AEBP2 and JARID2. PRC2 then deposits H3K27me3, which in turn is recognized by canonical PRC1 complexes via CBX subunits. Canonical PRC1 then deposits H2AK119ub1 and mediates chromatin compaction. Modified from [139].

However, a number of recent publications have challenged the generality of this model suggesting alternative mechanisms of Polycomb function. Blackledge *et al.* used a

targeting system in ESCs to show that ectopic recruitment of the variant PRC1 subunits PCGF1, 3 or 5 to chromatin leads to robust PRC2 recruitment and H3K27me3. Moreover, they demonstrated that H2AK119ub1 is sufficient to recruit catalytically active PRC2 and that canonical PRC1 complexes fail to catalyse significant levels of H2AK119ub1 in ESCs. Finally, they showed that recognition of unmethylated DNA by KBM2B is important for H2A monoubiquitination and recruitment of PRC2 to a subset of CGIs and that this mechanism is crucial for mouse development [140]. Cooper *et al.* demonstrated that DNA methylation does not affect the catalytic activity of PRC2 and variant PRC1, indicating that DNA methylation only affects PcG recruitment [141]. A complementary study by Kalb *et al.* showed that PRC2 preferentially binds polynucleosomes bearing H2AK119ub1 and that the ubiquitin mark can potentiate the catalytic activity of PRC2, suggesting a positive feedback loop between H2AK119ub1 and H3K27me3 at Polycomb target genes [142]. Recruitment of PRC2 to ubiquitinated H2A might depend on JARID2 and AEBP2, which have been shown to bind to H2AK119ub1 *in vitro* [66, 142].

Taken together, these studies suggest that the classical Polycomb recruitment model cannot explain all Polycomb targeting and that the order of recruitment might even be reversed. It is tempting to speculate that the variant PRC1.1 complex binds to unmethylated CpGs, recruits PRC2 via H2AK119ub and PRC2 in turn recruits canonical PRC1 via H3K27me3 to ensure robust transcriptional repression of developmental genes that harbour unmethylated CGIs (Fig. 2B). However, as mentioned above, PRC2 can also be recruited to PRC2 nucleation sites independently, via JARID2 and PCL2 [67]. Therefore, it seems that there is not a single strict hierarchy of PcG recruitment, but rather a number of targeting mechanisms and a range of interactions between different PcG subunits and histone modifications that ensure proper PcG-mediated gene repression at target genes.

1.3.5 Mechanisms of Polycomb repression

Another key question in the Polycomb field is how PcG complexes achieve gene repression and what determines the transcriptional activity at genes which can be activated by TrxG or repressed by PcG proteins, depending on the cell type. In addition to the partially controversial functions of the PcG-dependent histone modifications discussed above, an important mechanism of PcG repression is chromatin compaction, resulting in chromatin

inaccessibility and inhibition of SWI/SNF-mediated chromatin remodeling [71]. CBX2 contains a positively charged 'compaction region' that is required for bridging adjacent nucleosomes in order to compact chromatin [143]. Moreover, also EZH1-containing PRC2 is able to compact chromatin, indicating that PRC1 and PRC2 might cooperate in order to achieve chromatin compaction and inhibition of transcription [144]. It has recently been shown that there is a constant, dynamic competition between the SWI/SNF complex and PRC1, suggesting that the balance of these two complexes might be key to switch between repressed and active chromatin states [145].

In recent years, it has become evident that in metazoans, chromatin is organized in a hierarchical system of 3D architecture. Nucleosome fibres form chromatin loops, which then form topologically associated domains (TADs) [146-148]. The size of these TADs ranges from thousands to millions of bases and they have been shown to be stable over many cell divisions, invariant across different cell types and evolutionary conserved. TADs interact with each other and form active and repressive chromosome compartments, which again are organized into chromosome territories [reviewed in 149].

PcG proteins have been shown to appear clustered in the cell nucleus, forming so-called PcG foci [150-153]. Recently, with new technologies available, the molecular basis of PcG-mediated genome organization has become clearer. On a lower level, PcG proteins are involved in the formation of regulatory chromatin loops. This local looping, which may support PcG spreading in *cis,* is probably achieved by the combined function of PRC1 and PRC2 subunits [100, 154]. It has been suggested that PcG foci are involved in *de novo* establishment of PRC2 domains by functioning as nucleation hubs from which H3K27me2 and H3K27me3 spreads to proximal as well as distal regions in the genome via long-range 3D contacts [67]. On a higher level, PcG proteins organize the 3D architecture of PcG-containing TADs. It has been shown that repressed Hox clusters in mammals form a spatially separated, repressive domain [155]. Upon activation of Hox clusters, the CTCF insulator protein is responsible for the separation of the transcribed region of the cluster from the PcG-repressed domain [156, 157]. On the next level, PcG domains contact each other in the nucleus and form clusters of Polycomb TADs [158, 159]. These PcG-mediated long-range interactions are dynamic. It has been shown that when mouse ESCs change from the naïve to the primed state, new PcG-dependent interactions are established [160].

Both PRC1 and PRC2 have been shown to play a role in establishing long-range interactions between PcG-repressed domains in ESCs [154, 158, 161, 162]. However, PRC2 does not have a critical role in the establishment of TADs and its depletion does not have major effects on long-range interactions between PcG-repressed regions [147, 158, 162]. PRC1, on the other hand, has important functions in local condensation of PcG-repressed regions as well as in establishment of long-range interactions between domains [154, 158]. In particular, the SAM domain of the Ph/PHC1 subunit of canonical PRC1 mediates oligomerization, leading to local chromatin condensation of PcG domains as well as long-range domain interactions [83, 100, 154].

Drosophila Pc and mammalian CBX proteins can also repress gene activation by inhibiting the acetyltransferase activity of dCBP, thereby blocking H3K27 acetylation and favouring H3K27 methylation [163]. Further, the widespread H3K27 dimethylation mark has been shown to prevent pervasive chromatin opening and transcription by blocking H3K27 acetylation [164, 165].

1.3.6 Role of Polycomb-mediated repression in stem cells and development

1.3.6.1 Role of PcG proteins in embryonic development

Analysis of knock out (KO) mice for different PcG subunits has demonstrated important roles of these proteins in embryogenesis, with mutant embryos typically showing lethality at the gastrulation stage (Tab. 1.1). These results point towards a role of PcG proteins in the specification and maintenance of cell lineages. Specifically, constitutive KO of the PRC2 components EED, EZH2 and SUZ12 causes lethality at early post-implantation stages, but does not affect pre-implantation development [45, 166, 167]. In contrast to the major developmental defects observed upon deletion of the core PRC2 subunits, embryos deficient for JARID2 display defects in neural tube formation at later post-implantation stages [168]. PCL2, which regulates left-right asymmetry in chicken, is dispensable for mouse embryonic development upon zygotic deletion, possibly resulting from functional redundancy with other PCL homologues [169, 170]. These results suggest that JARID2 and PCL2 are not core components of PRC2, but are required for PRC2 function only in specific circumstances.

Zygotic deficiency of the core PRC1 component RNF2 results in embryonic lethality during gastrulation, whereas its paralog RING1 is not essential for development [40, 171]. Mice carrying a hypomorphic *Rnf2* allele exhibit posterior homeotic transformations of the axial skeleton [172]. It has further been shown that transcriptional regulation by PRC1 during oogenesis is required for proper zygotic genome activation (ZGA), replication and cell cycle progression during pre-implantation development and that maternal DKO of RING1 and RNF2 leads to a developmental arrest at the 2-cell stage [173]. Nuclear transfer experiments revealed cytoplasmic and chromatin-based contributions of PRC1 to pre-implantation development [173].

Moreover, several non-canonical PRC1 subunits have been shown to be required for embryogenesis. Embryos deficient for RYBP die during early post-implantation stages [174]. KDM2B exists in two isoforms, a long form that carries the histone demethylase JmjC domain and a short form, which lacks the JmjC domain. The short form is the predominant isoform in the embryo and in most adult tissues [175]. Deletion of the JmjC domain of the long form, that left the short form intact, has been shown to perturb neural development in around 50% of embryos [175]. In contrast, deletion of the CxxC domain from both KDM2B isoforms leads to early post-implantation lethality [140]. Another study demonstrated that homozygous embryos carrying a retroviral gene-trap allele that truncates both isoforms and functions as a null allele, exhibit pervasive and completely penetrant developmental defects and post-implantation lethality [176]. Taken together, these results suggest that the histone demethylase activity of KDM2B is largely dispensable for its function in embryogenesis, while the C-terminal domains are essential for post-implantation development.

In contrast, deletion of single CBX proteins does not impair embryonic development. Homozygous CBX2 KO mice exhibit growth retardation, homeotic transformations and sternal and limb malformations [38]. Knock out of CBX4 leads to severe hypoplasia of the fetal thymus as a result of reduced thymocyte proliferation, leading to lethality of homozygous KO mice within 1 hour of birth [177]. CBX7 KO mice develop to adulthood, but show an increase in liver and lung adenomas and carcinomas [178]. Knock out of PCGF2 or PCGF4 causes defects in anterior-posterior patterning of the axial skeleton and PCGF2/4 DKO leads to more severe defects and post-implantation lethality, suggesting that they have partially redundant functions [39, 41, 179].

1.3.6.2 Role of PcG proteins in pluripotency and differentiation

The observation that KO of most PcG components in mice causes embryonic lethality at gastrulation suggests an important function of PcG proteins in differentiation and cell fate specification. The function of PcG-mediated repression in self-renewal and differentiation has been extensively studied in ESCs (Tab. 1.1). In both embryonic and adult stem cells, many gene promoters are marked by both active H3K4me3 and repressive H3K27me3. These 'bivalent' promoters are found at genes with functions in cell-fate decisions that are transcribed at very low levels. Hence it has been proposed that the function of bivalency is to keep crucial genes off, but in a poised state, until their expression is required [133, 180]. However, it has also been shown that while bivalent promoters are only bound by the Trithorax H3K4 methyltransferase MLL2, active promoters are bound by MLL2 as well as CXXC1, a subunit of the SET1C complex. Upon removal of MLL2, which ablates H3K4me3 only at bivalent promoters, most bivalent promotes respond normally to retinoic acid, arguing against a priming function of bivalency. Removal of MLL1 had almost no effect on H3K4me3 at most promoters, unless MLL2 was removed as well, arguing that there is functional redundancy between the two enzymes. In light of their results, the authors proposed a different explanation for bivalency, in which MLL2 acts as a pioneer in the definition of CGIs and H2K27me3 at bivalent promoters serves to prevent the maturation of CGIs into SET1C-bound active promoters [181].

In mouse ESCs, several PcG proteins have been shown to repress genes involved in differentiation. KO of EED or RNF2 leads to an increased expression of genes involved in differentiation, but does not affect expression of pluripotency markers and the self-renewing capacity of cells [182-185]. Depletion of both RING1 and RNF2 however impairs self-renewal, indicating that the two RING1 subunits have redundant functions and are required for maintaining ESC identity [185]. Moreover, the PRC2 components PCL2 and PCL3 are necessary for the expression of key pluripotency genes [54, 135, 186]. Since deletion of EED does not affect pluripotency gene expression, PCL2/3 likely have PRC2-independent functions in ESC self-renewal.

Table 1.1. Role of PRC1/2 components in development and embryonic stem cell self-renewal and differentiation. Entries indicate whether the component is essential for embryogenesis, ESC self-renewal and ESC differentiation. ND, not determined; ?, controversial. The last column includes corresponding references.

PcG complex	component	Embryonic development	ESC self-renewal	ESC differentiation	Ref.
PRC1	RING1	No	No	No	[40]
	RNF2	post-implantation lethality (gastrulation)	No	Yes	[171, 182, 183]
	RING1/RNF2	maternal KO causes 2-cell arrest	Yes	Yes	[173, 185]
	PCGF2	No	Yes	Yes	[39, 187]
	PCGF4	No	ND	ND	[41]
	PCGF2/4	Post-implantation lethality around E9.5	ND	ND	[179]
	PCGF6	partial peri-implantation lethality	Yes	No	[93, 188]
	CBX2	No	No	Yes	[38, 189]
	CBX4	No	No	Yes	[177, 189]
	CBX7	No	Yes	Yes	[178, 189]
	RYBP	Early post-implantation lethality around E6.0	No	Yes	[174, 190]
	KDM2B	defects in neural tube formation/early post-implantation lethality	?	Yes	[87, 130, 140, 175, 176]
PRC2	EED	post-implantation lethality (gastrulation)	No	Yes	[166, 182, 184, 191]
	EZH2	post-implantation lethality (gastrulation)	?	Yes	[167, 192]
	EZH1	No	ND	ND	[193]
	SUZ12	post-implantation lethality (gastrulation)	No	Yes	[45, 194]
	JARID2	lethality at E15.5, defects in neural tube formation	No	Yes	[63, 64, 168]
	PCL2	No	Yes	Yes	[54, 170]
	PCL3	ND	Yes	Yes	[135, 186]

Upon differentiation of ESCs, repression of lineage-specific genes is relieved and pluripotency genes become repressed. PcG proteins also still repress alternative lineage genes in order to prevent de- or trans-differentiation [reviewed in 195]. The PRC2 component EZH2 is required for differentiation of the mesendodermal lineages and ESCs deficient for SUZ12 show deficiencies in the generation of endodermal lineages [192, 194]. On the contrary, EED deficient ESCs are able to contribute to all three germ layers, although

it has been shown that their ability to form teratomas is compromised [182, 191]. The difference in differentiation phenotypes between EZH2/SUZ12 and EED deficient ESCs might be due to technical differences or it may suggest that EZH2 and SUZ12 have additional functions, independent of PRC2. In general, it seems that PRC2 plays a key role in differentiation, since also other PRC2 members, such as JARID2 and PCL subunits, have been implicated in the differentiation process [54, 63, 64, 135]. In addition, also the PRC1 complex has been shown to regulate differentiation of ESCs. It has been demonstrated that ESCs lacking RNF2 exhibit improper expression of differentiation markers upon embryoid body formation [183]. Additionally, other PRC1 components, like RYBP, KDM2B and CBX proteins are required for ESC differentiation [85, 87, 189, 190]. Interestingly, DKO ESCs lacking both EED and RNF2 showed a more severe phenotype than either of the two single KOs, indicating that PRC1 and PRC2 have at least partially independent functions in differentiation [182].

Since PcG proteins seem to function in both self-renewal and differentiation, one question that arises is how the balance between pluripotency and lineage commitment is achieved. It seems that these different functions are mediated by PcG complexes containing different subunits, which are dynamically assembled and disassembled. It has been shown that upon differentiation of ESCs, the CBX subunit composition of the canonical PRC1 complex changes. In self-renewing ESCs, CBX7 is the predominant CBX protein, while in differentiating ESCs it gets replaced by CBX2 and CBX4 [189, 196]. It has been suggested that the switch in CBX subunits is involved in the regulation of ESC differentiation, since CBX7-containing PRC1 is targeted to a different set of genes than CBX2/4-containing PRC1. Interestingly, it seems that the CBX composition of PRC1 is regulated by an autoregulatory loop. CBX7-containing PRC1 represses *Cbx2* and *Cbx4* in pluripotent ESCs, while CBX2/4-PRC1 represses *Cbx7* in differentiated ESCs [189]. In addition, also the different PCGF family members of PRC1 have been shown to have specific functions in ESC self-renewal, mesodermal differentiation and reprogramming [93, 187, 188].

The composition of the PRC2 complex also changes upon differentiation, with an exchange of EZH2 with EZH1 when cells exit pluripotency and start to differentiate. It has been shown that EZH2-PRC2 has higher enzymatic activity, while EZH1-PRC2 compacts chromatin more efficiently [144]. In addition, the interactions of PRC2 with accessory proteins are highly dynamic. In undifferentiated cells, the PRC2 core components

preferentially associate with EPOP, PCL2, PCL3 and JARID2, whereas in differentiated cells PCL1 and AEBP2 are more commonly present [90, 197]. These subunits can contribute to PRC2 targeting during differentiation, since PCL3 (PHF19) has been shown to recognize H3K36me3 and JARID2 shows affinity for DNA sequences [63, 135]. Depletion of JARID2 impairs the binding of PRC2 to target genes and ESC differentiation [63]. Similarly, knock down of PCL3 in ESCs causes a global reduction of H3K27me3 and PRC2 occupancy genome-wide and leads to loss of pluripotency and differentiation defects [135].

1.3.6.3 Role of PcG proteins in epigenetic inheritance

The term "epigenetic inheritance" describes the stable transmission of gene expression and function across cell divisions without changes in the DNA sequence. When looking at epigenetic inheritance across generations, it is important to distinguish between intergenerational and transgenerational epigenetic inheritance. In the case of traits induced by an external signal, not only the parental animal (F0) is exposed to the signal, but also its germ cells, which will give rise to the F1 generation. If a pregnant female is exposed to the external signal, not only the fetus can be affected *in utero* (F1), but also the germline of the fetus (F2). Therefore, epigenetic inheritance is considered to be intergenerational and potentially influenced by direct exposure to the external signal between F0 and F1 via the male germ line and between F0 and F2 via the female germline. Inheritance of the acquired trait in later generations and in absence of the initial signal is considered as transgenerational epigenetic inheritance.

Much of the evidence for transgenerational epigenetic inheritance comes from plants, where the germline arises from somatic cells stimulated by developmental and environmental cues. In plants, epigenetic inheritance involves transcriptional regulation of transposable elements and neighbouring loci via DNA methylation. Hypomethylated chromosomal segments can be propagated for eight or more generations through mitosis and meiosis. A significant proportion of the differentially methylated regions (DMRs) exist in natural populations, controlling complex traits such as flowering time and primary root length [reviewed in 198]. Further evidence for transgenerational epigenetic inheritance comes from worms and flies, both of which lack DNA methylation. Instead, they utilize a

different molecular mechanism for epigenetic inheritance, involving small RNAs and histone methylation of H3K9 [reviewed in 198].

Although intergenerational effects certainly occur in mammals, only a few examples of transgenerational epigenetic inheritance affecting endogenous genes have been described. The most prominent examples are the agouti viable yellow (A^{vy}) and the axin fused ($Axin^{Fu}$) alleles, both of which are associated with transposable elements [199, 200]. Transcription of the agouti (*A*) gene is initiated by an intracisternal A particle (IAP) retrotransposon inserted 100kb upstream of the gene. The resulting ectopic expression of the agouti protein leads to yellow fur colour, obesity, diabetes and increased risk of cancer [201]. Inbred mice carrying the allele show variable fur colours, from yellow to mottled to pseudoagouti (brown) [202]. The differences in coat colour are caused by the DNA methylation level of the IAP promoter, that is inversely correlated with transcriptional activity [199]. In addition, the A^{vy} locus in yellow mice is associated with active histone marks, while in pseudoagouti mice, repressive histone marks are found at the locus, suggesting an interplay between DNA methylation and histone modifications [203]. The distribution of coat colours in the offspring is related to the phenotype of the dam, meaning that pseudoagouti dams give rise to a higher percentage of pseudoagouti offspring than yellow dams [199]. This suggests that the erasure of epigenetic modifications at the A^{vy} locus is incomplete when it is passed through the female germline. In contrast, the coat colour of the offspring is not affected by the phenotype of the sire, suggesting that the epigenetic marks are cleared when they are passed on through the male germline [199]. It has been shown that the paternal A^{vy} allele is de-methylated rapidly after fertilization, whereas the maternal allele undergoes slow, presumably passive, de-methylation, leading to complete absence of DNA methylation on the maternal allele at the blastocyst stage. These results argue that DNA methylation is not the primary inherited epigenetic mark [204]. Further evidence for this conclusion comes from the fact that following paternal A^{vy} transmission and maternal transmission of a *Pcgf2* (*Mel18*) KO allele, haploinsufficient mice for PCGF2 display epigenetic inheritance, whereas their wild type littermates do not. Since the cytoplasmic environment of oocytes is set up before segregation of the diploid set of chromosomes, this argues that the effect is dependent on ZGA occurring at the 2-cell stage. The paternal genome is actively de-methylated within 6 hours after fertilization, thus the process is completed by the time that wild type and PCGF2 haploinsufficient embryos differ,

further arguing against a role of DNA methylation as the heritable mark [204]. Since DNA methylation does not seem to be the primary inherited mark, it is likely that histone modifications or non-coding RNAs are involved in the process. The role of PCGF2 in the process is not completely understood, but presumably it is involved in clearing of the underlying epigenetic mark in wild type embryos. Haploinsufficiency for PCGF2 might therefore lead to incomplete clearing of this mark due to altered composition of PRC1 complexes.

Further, several studies have reported epigenetic inheritance of expression of transgenes as well as environmentally induced epigenetic inheritance. Recently, a couple of studies in mice have focused on transmission of such effects through the paternal germline. Males fed a low-protein diet have been shown to pass on a high-cholesterol phenotype to their offspring, that was associated with changes in gene expression and modest differences in DNA methylation, including methylation changes at a likely enhancer of the key lipid regulator *Ppara* [205]. It has further been shown that *in utero* exposure to undernourishment results in DNA hypomethylation of discrete loci that are enriched for nucleosomes in sperm of male offspring. Although differential methylation does not persist in the next generation, the F2 generation still shows misexpression of genes neighbouring differentially methylated regions and a metabolic phenotype [206]. Siklenka *et al.* have shown that overexpression of the H3K4 demethylase KDM1A that leads to a reduction of H3K4me2 in sperm, severely impairs development and survival of offspring. These effects were observed for two subsequent generations, in the absence of transgene expression and were associated with changes in RNA content in sperm and offspring, but unchanged DNA methylation [207]. However, many of the effects that have been described do not pass beyond F2 and are thus not transgenerational and rarely exclude changes in DNA sequence as the cause of heritability. Thus, the question regarding the existence of transgenerational epigenetic inheritance in mammals remains controversial [reviewed in 198, 208].

What could be the role of PcG proteins in this process? The epigenetic reprogramming in the germline involves both DNA methylation and histone modifications. Since epigenetic inheritance is the result of incomplete reprogramming of certain loci called “epialleles”, histone modifications could potentially function as epialleles. A recent study suggests that in *Drosophila* intergenerational inheritance of maternally provided H3K27me3

prevents inappropriate activation of lineage-specific enhancers during embryonic development [209]. Further, PcG proteins were shown to be responsible for transgenerational inheritance of transgene expression in flies [210, 211]. A recent study, using a system of transiently enhancing the 3D chromatin interactions of a transgene in *Drosophila*, generated alternative epialleles that differed in the levels of H3K27me3. These epialleles were inherited transgenerationally, through both male and female germlines [212]. In worms, it has been shown that PRC2 and H3K27me3 are responsible for inter- and transgenerational inheritance of the repressed state of the X chromosome, via both male and female germlines [213].

In mammals, genome-wide chromatin immunoprecipitation (ChIP) studies have revealed that the small percentage of nucleosomes retained in human and mouse spermatozoa are localized specifically at promoters of key developmental genes and carry histone modifications, including H3K27me3 [214-216]. Recently, with the advances in ChIP-seq technologies for small samples, histone marks have also been mapped in oocytes and different stages of pre-implantation embryos. These studies show that both sperm and oocytes contain regions of H3K27me3, most of which are reprogrammed during early embryogenesis [217, 218]. However, a significant number of these regions, particularly those coming from the oocyte maintain the H3K27me3 mark during pre-implantation development [218].

1.4 Pre-implantation development as a system to study chromatin dynamics

1.4.1 Chromatin reorganization in early embryos

The onset of mammalian embryonic development is initiated by the fusion of two highly differentiated and transcriptionally silent germ cells, the oocyte and the spermatozoon. These two cell types differ substantially, not only in their morphology and cytoplasmic content, but also in chromatin organization. It has recently been shown that histone modifications in oocytes display largely non-canonical patterns. H3K4me3 appears as broad domains at both promoters and distal sites and has been linked to genome silencing, possibly by acting as a 'sponge' to attract TFs to distal sites, diluting them away from promoters [219-221]. H3K27me3 in oocytes displays relatively little enrichment at

promoters of developmental genes. In contrast, it forms large domains in intergenic regions and gene deserts [217, 218]. The specific pattern of H3K27me3 in oocytes could be related to the meiotic cell cycle arrest, which might allow accumulation of distal H3K27me3 in the absence of cell division. Therefore, promoter H3K27me3 might seem relatively reduced. Alternatively, the possible disruption of Polycomb TADs in meiosis might alter the efficiency of promoter H3K27me3 deposition. On the contrary, the ubiquitous acquisition of H3K27me3 at distal sites seems to occur promiscuously in regions devoid of active transcription or DNA methylation. In accordance with the ChIP-seq results, a study using superresolution imaging has revealed periodic H3K27me3 staining in roughly 500 nm intervals along chromosomes in mouse oocytes [222]. The function of this pervasive H3K27me3 in oocytes remains elusive, but possibly it serves to compensate for the absence of other repressive mechanisms like DNA methylation or H3K9me3. In order to fit the genome into the very small sperm head, the majority of histones in sperm (99% in mouse, 90% in humans) get replaced by protamines. Notably, the remaining fraction of histones carries specific histone modifications and is enriched at promoters of developmental genes, in both mouse and human [214-216]. Compared to oocytes, histone modifications in sperm display mostly canonical patterns [223, 224].

Upon fertilization, the male and female genomes remain spatially separated, forming the paternal and maternal pronuclei [225]. The initial chromatin state of the maternal genome is inherited from the oocyte and is similar to that of somatic cells, with high abundance of H3K4me3, H3K9me2/3, H3K27me2/3, H3K36me3, H3K64me3 and H4K20me3 [30, 226-233]. On the other hand, the paternal genome undergoes rapid *de novo* chromatin formation incorporating maternally provided histones. This histone incorporation takes place before DNA replication and involves the histone variant H3.3 and its chaperone Hira [30] (Fig. 3). The newly incorporated histones then become *de novo* methylated at different lysine residues in a highly spatially and temporally coordinated manner. While H3K4me1, H3K9me1 and H3K27me1 are detected on the paternal pronucleus at fertilization, H3K4me2/3, H3K9me2/3 and H3K27me2/3 only appear after DNA replication [32, 226, 228, 232].

Recently, allele-specific ChIP-seq data have shed more light on the chromatin reorganization in early embryos. It has been suggested that most of the canonical H3K4me3

and H3K27me3 peaks from sperm disappear after fertilization. However, the paternal genome acquires broad, but weak domains of H3K4me3 in gene-rich regions and H3K27me3 in gene deserts [218, 220]. In contrast, maternal non-canonical H3K4me3 is inherited to embryos and maintained before ZGA. In late 2-cell embryos, after ZGA, both maternal and paternal broad H3K4me3 domains are dramatically reduced, while canonical H3K4me3 starts to appear at promoters [220]. Reprogramming of H3K27me3 on the maternal allele has a different dynamic, with removal of the mark from promoters of developmental genes after fertilization. The reduction of H3K27me3 at promoters persists throughout pre-implantation development until promoter H3K27me3 re-appears at canonical PcG targets in the inner cell mass (ICM) of the blastocyst and becomes stronger in the epiblast of post-implantation embryos [218]. Despite the reduction of H3K27me3 at promoters in pre-implantation embryos, most developmental genes remain silent, suggesting either the presence of other repressive mechanisms or a lack of activating signals. H3K27me3 in distal regions in contrast is maintained throughout pre-implantation development, but lost in post-implantation embryos [218].

A recent study demonstrated that maternal H3K27me3 inherited from oocytes can regulate maternal allele-specific gene repression independently of DNA methylation. While most of the identified genes show maternal-specific repression during early pre-implantation development, a few of them maintain their allelic expression pattern in the extraembryonic lineage of post-implantation embryos [234]. Similarly, in *Drosophila*, it has been shown that H3K27me3 inherited from oocytes and propagated after fertilization, regulates the function of enhancers in embryos [209].

In addition to reprogramming of chromatin marks, extensive and rapid genome-wide DNA de-methylation of the paternal genome, which has been suggested to be mediated by TET3-driven oxidation of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC), occurs prior to the first round of replication [235-240]. However, a recent study has shown that the initial reduction of paternal 5mC does not require the formation of 5hmC. In contrast, Amouroux *et al.* demonstrated that the maternally inherited DNA methyltransferases DNMT1 and DNMT3A are required for 5hmC accumulation, suggesting that the function of TET3 is to protect the hypomethylated state in the early embryo from accumulating *de novo* methylation [241]. In addition to active DNA methylation of the

paternal genome, both maternal and paternal genomes undergo passive DNA de-methylation happening along with DNA replication throughout pre-implantation development [242, 243] (Fig. 1.3).

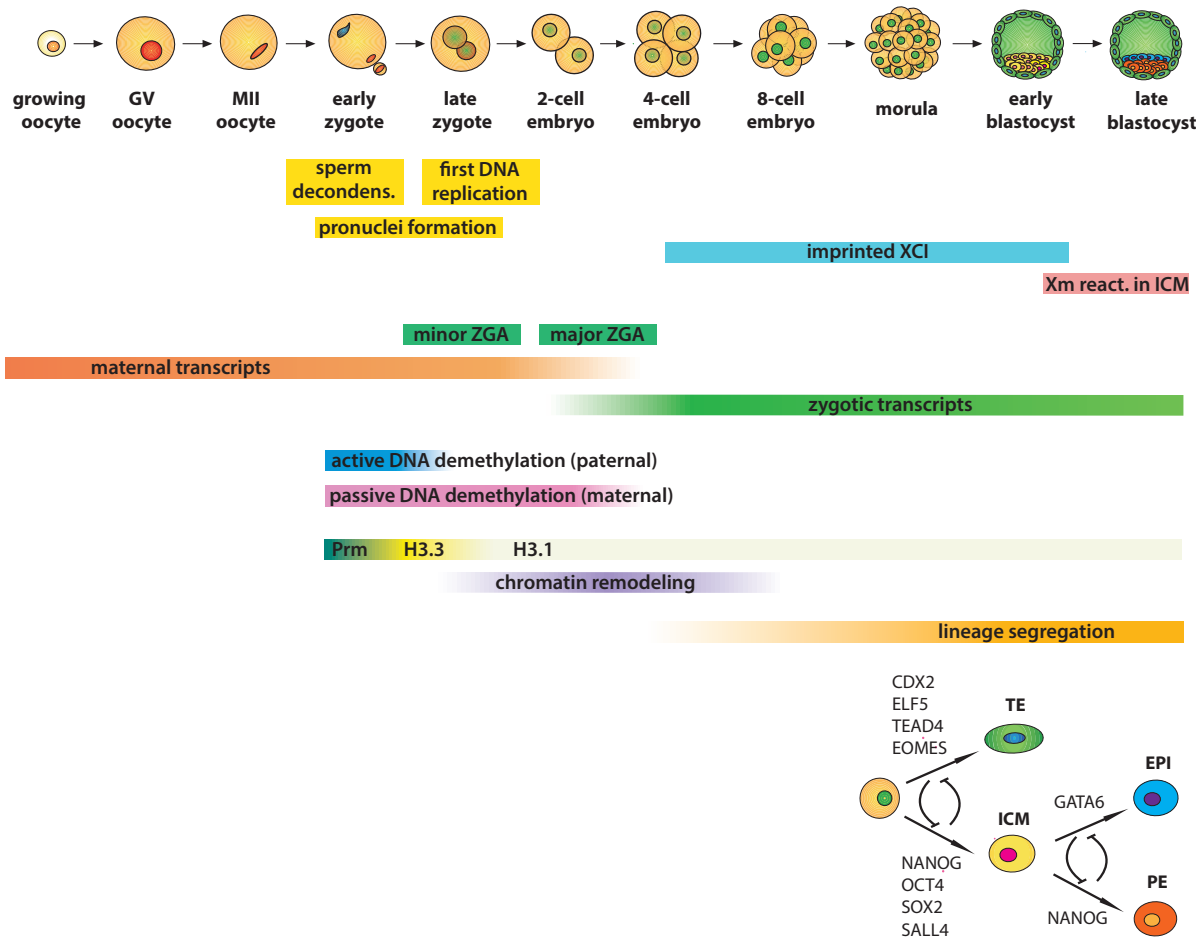


Figure 1.3. Major events during pre-implantation development. Schematic overview of mouse early embryogenesis in respect to X chromosome inactivation, global gene expression events, epigenetic remodeling and lineage segregation. Lineage specification occurs in two steps, which lead to the formation of three distinct cell lineages in the late blastocyst: the epiblast (EPI), primitive endoderm (PE) and trophectoderm (TE). The key transcription factors (TFs) regulating the cell fate decisions are indicated.

Although the epigenetic states of the two genomes become less distinct, some major differences are retained during DNA replication and subsequent cell divisions up to the 8-cell stage [232, 244]. It has been shown that in zygotes maternal constitutive heterochromatin is marked by Suv39h-mediated H3K9me3, which is absent from the paternal genome, where H3K27me3 and PRC1 accumulate and function as a back-up mechanism for repression in the absence of H3K9me3. This parental epigenetic asymmetry

persists until the end of the 8-cell stage and its resolution coincides with the end of the maternal-to-embryonic transition [244].

1.4.2 Zygotic genome activation

The rapid and dynamic changes in the early zygote occur in the absence of zygotic transcription and are fully driven by maternal factors, which were accumulated in the oocyte. Zygotic genome activation (ZGA) is the first major developmental transition in the embryo in which the initial developmental program orchestrated by maternally inherited factors gets replaced by a new program that is driven by newly expressed genes in the embryo. It commences at the end of 1-cell stage with a minor activation of transcription (minor ZGA) and continues at 2-cell stage with major ZGA [245, 246]. The main functions of ZGA are the degradation of oocyte-specific mRNAs, the replacement of maternal transcripts that are common to oocyte and early embryo with zygotic ones and most importantly the generation of novel transcripts not expressed in the oocyte, thereby promoting a dramatic shift in gene expression patterns [reviewed in 247, 248] (Fig. 1.3).

Notably, ZGA is characterized by activation of a large number of transposons and repetitive elements including endogenous retroviruses that are silenced in most other cell types [249-251]. Many of these elements are expressed early and become repressed after the 2-cell stage to prevent mutations and maintain genome integrity in the early embryo [252, 253]. The most highly activated repeat in 2-cell embryos is the Murine endogenous retrovirus-like (MuERV-L) family of retroviruses [254]. The promiscuous transcription of repeats is likely due to the permissive chromatin state before ZGA, which has been shown in different studies, including ATAC-seq (Assay for Transposase Accessible Chromatin with high-throughput sequencing) and Hi-C analyses [255-257]. Although the exact mechanism underlying the high chromatin accessibility in early embryos is not yet fully understood, several factors including chromatin assembly factor CAF-1 and the TF DUX4 have been shown to be involved [258-261]. Many MuERV-L elements expressed at the 2-cell stage serve as alternative promoters for the activation of protein-coding genes by generating chimaeric transcripts [254].

Further, the *Zscan4* gene family shows temporally and spatially restricted expression during ZGA at the 2-cell stage and is essential for early mouse development [262]. In recent

years, a rare, transient proportion of cells in ESC populations has been identified, whose transcriptome is highly similar to that of 2-cell embryos, including expression of genes like the *Zscan4* family and MuERV-L retrotransposons [254, 263]. These so-called 2C-like ESCs that have the potential to contribute to both embryonic and extraembryonic lineages have great potential to identify the molecular mechanisms controlling ZGA *in vivo.*

1.4.3 First lineage specification events in the early embryo

In many organisms, the first cell fate decisions in the embryo are dictated by maternally provided determinants like morphogens that are asymmetrically distributed in the zygote. In contrast, mammalian embryos lack clearly visible pre-patterning determinants, raising the question how three days after fertilization two different cell lineages arise in the mouse embryo. Therefore, two very different viewpoints of development in early mammalian embryos have been proposed. The first one, which represents the traditional view on mammalian development, argues that cell fate arises through stochastic events because all blastomeres of the early embryo are identical in their cell fate potential. The second one argues that cell fate is predictable in early embryos, because they are not perfectly homogenous and not all blastomeres are identical due to differential localization or expression of molecules [reviewed in 264]. The later viewpoint has been gaining more support recently through several studies that showed inequality in the potency of blastomeres in 2-cell and 4-cell mouse embryos. The first evidence supporting this idea comes from the notion that upon separation of 2-cell embryos, only one blastomere has the full potential to develop into a mouse, while the other will give rise to fewer epiblast cells leading to developmental failure [265-269]. More recent studies, involving single-blastomere RNA sequencing (RNA-seq) have revealed small, but significant differences between the blastomeres at the 2-cell stage in both mouse and human embryos [270, 271].

Moreover, it has been shown that heterogeneity in the nuclear localization of SOX2 and OCT4 among blastomeres of 4-cell embryos drives lineage segregation. Blastomeres with long-lived SOX2 or OCT4 binding contribute more likely to the inner cell mass (ICM), while blastomeres with shorter binding of these TFs more likely contribute to the trophectoderm (TE) [272, 273]. The ability of SOX2 to bind DNA is regulated by histone H3R26me2 that is deposited by the methyltransferase CARM1 together with PRDM14, both

of which are heterogeneously expressed at the 4-cell stage [272, 274, 275]. Further, SOX21, a target of OCT4 and SOX2, is also heterogeneously expressed in 4-cell embryos. This heterogeneity has been shown to drive cell fate because higher levels of SOX21 lead to lower expression of CDX2, which induces TE differentiation [276]. Together, these results indicate that differences in CARM1 activity at the 4-cell stage lead to differential expression of OCT4 and SOX2 target genes such as SOX21, thereby biasing cells towards ICM vs. TE fate.

Subsequently, the first asymmetric cell divisions take place at the 8-cell stage, where the cell-cell contacts established during compaction are of great importance for polarization and cell fate specification [277, 278]. During compaction, CDX2, which is a key determinant of the TE lineage, localizes to the apical half of blastomeres, leading to higher CDX2 expression in outer cells of 16-cell embryos that thereby become biased towards TE fate [279]. Proteins involved in the Hippo pathway play a key role in setting up the differential lineage bias of the two populations by sensing the compaction state of the cells and activating the signalling cascade in the more densely packed inner cells [280]. Therefore, in the outer cells, the kinases LATS1/2 are not active and do not interfere with the cytoplasmic transcriptional co-activator YAP1. Thus, YAP1 is relocated to the nucleus and activates TEAD4, which in turn activates CDX2 [280, 281]. Expression of CDX2 induces the activation of further TE differentiation factors like ELF5 and EOMES and the suppression of pluripotency factors, thereby promoting the irreversible TE lineage commitment of the outer cells [282, 283]. In contrast, Hippo signalling is active in inner cells, leading to phosphorylation of YAP1 and its retention in the cytoplasm, thereby preventing nuclear activity of YAP1 and keeping the inner cells in the pluripotent state [280]. The differences between inner and outer cells become more pronounced upon the first cell fate specification event in the early blastocyst, where the inside cells form the pluripotent inner cell mass (ICM) and the outer cells the extraembryonic trophectoderm (TE). Once these two populations are separated, they establish stable gene regulatory networks, including SALL4, OCT4, SOX2 and NANOG in inner cells and TEAD4, CDX2 and EOMES in outer cells [reviewed in 284] (Fig. 1.3).

The second lineage specification event differentiates the ICM into two separate lineages, the epiblast (EPI), which will give rise to all the cells of the future body and the primitive endoderm (PE), the second extraembryonic tissue. This developmental choice is

influenced by differential levels of the TFs GATA6 in PE cells and NANOG in EPI cells [285] (Fig. 1.3). However, this process begins much earlier, since the timing of the internalization of blastomeres affects the probability of their later affiliation with either EPI or PE lineages. It has been shown that blastomeres internalized during the first asymmetric division at the 8-16 cell stage transition are more likely to contribute to the EPI lineage, while blastomeres internalized in the second or third asymmetric division, between 16 and 64-cell stage, are more likely to become PE [286, 287]. This bias is due to differential levels of FGFR2, since blastomeres internalized later upregulate FGFR2, which makes them more sensitive to FGF signalling that is required for formation of the primitive endoderm [288]. Recently it has also been suggested that OCT4 is involved in the second cell fate decision event, by promoting FGF expression, maintenance of GATA6 expression and expression of SOX17 and PDGFRA, which are late makers for PE differentiation [289, 290].

The segregation of the first cell lineages in mouse embryos is accompanied by changes in chromatin. The asymmetries in chromatin organization between lineages is visible already at the global level. TE cells exhibit a more compact chromatin organization with heterochromatic foci at the nuclear periphery, while ICM cells show a looser chromatin organization of mostly euchromatin [291]. Histone marks such as H3K27me3 are enriched in the ICM [228]. It has also been suggested that epigenetic marks play a role in lineage specification. The *Cdx2* gene is marked by repressive H3K9me2 marks in the ICM, but enriched for activating H3K4me3 and H4K16ac in the TE. In contrast, the pluripotency genes *Nanog* and *Oct4* show the inverse marking [292]. Moreover, the asymmetry in H3R26me2 at the 4-cell stage is involved in very early cell fate specification, preceding the establishment of differential gene regulatory networks [274]. In addition, also the global DNA methylation status differs between the ICM and TE lineages. ICM cells show more extensive re-establishment of DNA methylation at the blastocyst stage compared to cells of the TE lineage [293]. Finally, differential methylation of the *Elf5* gene might play a role in maintaining the TE fate through a feedback loop to *Cdx2* and *Eomes*, but prevents expression of these genes in the ICM [282].

1.4.4 Imprinted X inactivation

In mammals, asymmetry in the constitution of sex chromosomes in XY males and XX females caused the evolution of dosage compensation, which is achieved by silencing of one X chromosome in females during early embryogenesis [294] (Fig. 1.4a). In mouse, X chromosome inactivation (XCI) initiates in pre-implantation embryos around the 4-cell stage in an imprinted fashion, with inactivation of the paternal X chromosome (Xp) [295]. This imprinted XCI persists in extraembryonic tissues, while in the ICM of the late blastocyst the Xp is reactivated and becomes subject to random XCI of either the maternal (Xm) or paternal X chromosome in the epiblast [296-298] (Fig. 1.3).

Both imprinted and random XCI are governed by a master regulatory locus, the X-inactivation center (XIC), which, when present in two copies, is both necessary and sufficient to trigger XCI [299, 300]. The XIC contains the noncoding locus *Xist* and some of its regulators (Fig. 1.4c). *Xist* is expressed monoallelically from the future inactive X chromosome during imprinted as well as random XCI and the resulting long noncoding RNA coats the chromosome in *cis* [301-305] (Fig. 1.4b). This further induces changes in chromosome conformation and creates a silenced nuclear compartment devoid of RNA Pol II and euchromatic histone modifications [297, 306, 307].

It has been shown that the protein SHARP (SPEN), which is a member of a complex that activates HDAC3, interacts with *Xist* RNA and mediates the early steps of *Xist*-induced gene repression [308-310]. Knock down of SHARP or HDAC3 has been shown to abolish PRC2 targeting to the *Xist* domain in ESCs. These results argue that SHARP acts upstream of PRC2 and that *Xist*-dependent recruitment of PRC2 at the X chromosome is mediated by SHARP and HDAC3 [309]. McHugh *et al.* proposed a model in which *Xist* directly interacts with SHARP to recruit SMRT to the inactive X chromosome [311]. The *Xist*-SHARP-SMRT complex then recruits HDAC3 directly or may induce its enzymatic activity [312]. HDAC3 removes activating histone acetylation marks, leading to chromatin compaction and transcriptional silencing. Upon initiation of the repressed state, PRC2 is recruited in an HDAC3-dependent manner, either by direct interaction or via HDAC3-mediated transcriptional repression or chromatin compaction [309]. Subsequently, numerous repressive complexes are recruited and help to lock in the silenced state [reviewed in 313].

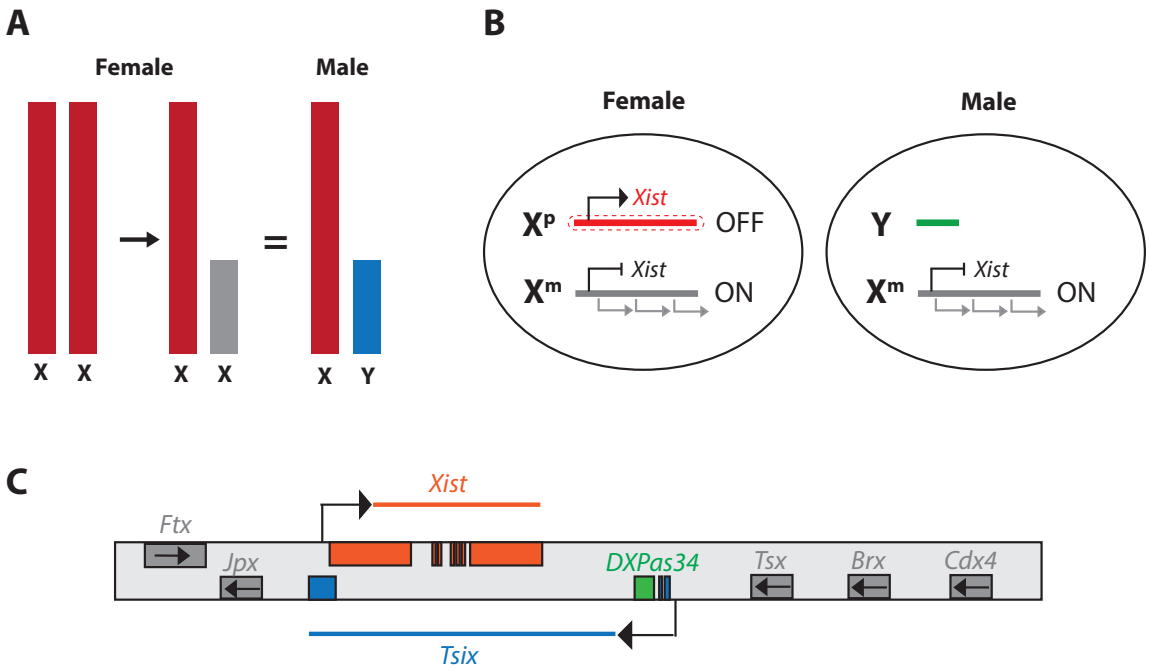


Figure 1.4. Imprinted X inactivation and the X inactivation center. A) Schematic diagram to illustrate the mechanism of X inactivation (XCI) in mammals. In order to compensate for the differences in X chromosome dosage between males and females, one of the female X chromosomes undergoes inactivation. **B) Scheme of imprinted X inactivation**. In females, during pre-implantation development, the paternal X (Xp) expresses the non-coding RNA *Xist*, which coats the chromosome in cis and induces silencing. The maternal X chromosome (Xm) in both males and females does not express *Xist* and thus remains active. **C) Schematic overview of the X inactivation center (XIC) in mice.** *Xist* and *Tsix* are non-coding RNAs which are expressed from opposite strands. While *Xist* is required to initiate XCI on the future inactive X, *Tsix* expression represses *Xist* on the active X in order to keep it active. However, it has been shown that in pre-implantation embryos, *Tsix* expression is dispensable for inhibiting *Xist* expression on the Xm.

However, it has recently been shown that the non-canonical PRC1 complex containing PCGF3/5 is recruited to the *Xist* RNA B-repeat via hnRNPK [314]. Hence, ncPRC1 acts upstream of PRC2 at the inactive X chromosome and initiates recruitment of Polycomb complexes via deposition of H2AK119ub1, leading to recruitment of PRC2 and other PRC1 complexes [315]. However, these results come from an ESC model where *Xist* transgenes were located on autosomes. Therefore, it cannot be excluded that additional factors contribute to X inactivation *in vivo.*

In addition to PcG proteins, other factors associated with the inactive X chromosome have been identified, including G9a, SETDB1 and CDYL, which may read the combination of H3K9me2 and H3K27me3 marks present on the inactive X [316, 317]. Taken together, a number of repressive mechanisms act in concert in order to establish and maintain the repressed state of the inactive X chromosome.

A long-standing question in the field concerns the nature of the parent-specific imprint underlying silencing of the Xp during imprinted XCI. It has been a matter of debate whether it is established in the maternal, the paternal or both germ lines [295, 318, 319]. The existence of a maternal imprint, required to prevent *Xist* expression on the Xm, so that it remains active, is suggested by experiments using mouse embryos generated by parthenogenesis. In these embryos both X chromosomes are inherited maternally and stay active throughout pre-implantation development until the morula stage, where the imprint seems to diminish, resulting in monoallelic *Xist* expression [320]. Nuclear transfer experiments have revealed that the imprint on the Xm is established during oocyte maturation in the prophase of meiosis I, since embryos that are generated by combining two maternal genomes from a fully grown (fg) oocyte and an early non growing (ng) oocyte preferentially inactivate the ng X chromosome [321].

Transcription of *Tsix,* which covers the entire *Xist* locus in antisense orientation, is required for *Xist* repression on the Xm in extraembryonic tissues of post-implantation embryos [322]. However, it has been shown that in pre-implantation embryos, *Tsix* is dispensable for inhibiting *Xist* expression and Xm silencing [323]. Further, analysis of embryos maternally deficient for *de novo* DNA methyltransferases has revealed that oocyte DNA methylation is dispensable for *Xist* imprinting [324]. It has been shown that overexpression of *Kdm4b*, a H3K9me3 demethylase, leads to partial de-repression of *Xist* in parthenogenetic embryos [325]. However, H3K9me3 does not become enriched at the *Xist* promoter during oocyte maturation, when the imprint is established, and thus does not seem to be the primary imprinting mark [326]. Recently, Inoue *et al.* showed that injection of the histone H3K27me3 demethylase *Kdm6b* into zygotes leads to maternal *Xist* expression and maternal XCI in pre-implantation embryos. This result together with the fact that H3K27me3 is enriched at the maternal allele of *Xist* in oocytes and pre-implantation embryos, suggests that maternal H3K27me3 serves as the imprinting mark for *Xist* [327].

1.5 Scope of the thesis

The role of PcG proteins has been studied extensively in ESCs, where they have been shown to have crucial functions for both maintenance of pluripotency and differentiation. However, much less is known about the role of PcG-mediated repression *in vivo* during mammalian

early embryonic development. During pre-implantation development, the fusion of two highly specialized games gives rise to a totipotent zygote, which then undergoes subsequent cell divisions to form all the lineages of the entire organism. This system provides a perfect *in vivo* model to address the role of PcG repressive mechanisms in the establishment of totipotency and in cellular differentiation.

Although the epigenetic landscape is heavily reprogrammed upon fertilization, there is some evidence that chromatin marks can be passed on to the embryo and exert specific functions during embryonic development. It has been shown that the small percentage of histones retained in human and mouse sperm are enriched for H3K27me3 and localize at promoters of developmental genes, suggesting a potential function during embryonic development [214-216]. However, recent ChIP-seq studies have suggested that the majority of H3K27me3 sites in sperm are reprogrammed upon fertilization. On the contrary, these studies showed maintenance of broad H3K27me3 domains from the oocyte in early embryos, pointing towards a maternal function of Polycomb-mediated repression for pre-implantation development [217, 218].

It has further been shown that PRC1 acts as a major transcriptional regulator in oocytes controlling expression of a large number of genes. These transcriptional functions of PRC1 in oocytes are required for ZGA, replication and cell cycle progression in pre-implantation embryos and maternal DKO of the two core PRC1 enzymatic components RING1 and RNF2 leads to a developmental arrest at the 2-cell stage [173]. Zygotic deficiency of the core PRC1 component RNF2, as well as of the PRC2 components EED, EZH2 and SUZ12 results in embryonic lethality during gastrulation, but does not affect pre-implantation development [45, 166, 167, 171].

Since at the time when I started my thesis, much less was known about the role of PRC2 during pre-implantation development, I set out to study the role of PRC2 and the H3K27me3 mark at the onset of life. I used a cre-lox based approach in order to abolish the function of PRC2 in the female and male germlines and to generate embryos that are maternally and zygotically deficient for PRC2. KO mouse lines carrying a conditional allele for *Eed* (*Eed*$^{F/F}$) and expressing Cre recombinase under oocyte- and spermatozoon-specific promoters (*Gdf9*-iCre and *Prm1*-Cre, respectively) had already been established in my laboratory. From a developmental perspective, I wanted to find out whether embryos maternally and zygotically deficient for EED can reach the blastocyst stage and if not, at what

stage they would die. Further, I was interested in a potential function of PRC2 in the first lineage specification events in early embryos.

In addition to studying effects of *Eed* deletion on embryonic development, I was interested in the function of PRC2 in transcriptional regulation in pre-implantation embryos. The following questions were key to the transcriptional analysis: What is the role of maternal PRC2 via inheritance of H3K27me3 from the oocyte and what are *de novo* functions of PRC2 in embryos? Does PRC2 regulate the maternal and paternal alleles differently during pre-implantation development? Does maternal deficiency for PRC2 affect gene expression in post-implantation embryos, indicating a role of PcG proteins in chromatin-based intergenerational inheritance?

In order to answer these questions, I performed RNA sequencing (RNA-seq) at different stages of pre-implantation embryos as well as of post-implantation embryos. In order to capture possible heterogeneity in gene expression profiles between embryos, I set up a protocol in the lab which allows RNA-seq of single pre-implantation embryos. I performed RNA-seq on both maternally and zygotically deficient as well as solely maternally deficient embryos. In addition, I sequenced wildtype (wt) embryos treated with PRC2 inhibitors after fertilization, in order to study the zygotic role of PRC2. Part of the experiments were performed using hybrid embryos, allowing to distinguish between reads from the maternal and paternal alleles. This was particularly interesting since during the course of my PhD, it has been shown that H3K27me3 in pre-implantation embryos differs greatly between the alleles and that H3K27me3 can repress genes maternally which are expressed paternally [218, 234]. Finally, I performed breeding tests as well as RNA-seq of post-implantation embryos maternally deficient for PRC2 in order to further study a potential role of PcG proteins in chromatin-based intergenerational epigenetic inheritance.

For the second part of my PhD project, I was interested in the contributions of PRC1 and PRC2 to Polycomb-mediated gene repression during pre-implantation development. Therefore, I generated compound mutants for *Eed* and *Rnf2* as well as for *Ezh2* and *Rnf2*. In these mouse models, I was interested in the developmental potential as well as changes on the transcriptional level. I set out to perform single-embryo RNA-seq on the different mutants for PRC1, PRC2 or both and to compare their transcriptional profiles in order to find specific or common targets of the two Polycomb repressive complexes.

Chapter 2. Manuscript in preparation: PRC2 regulates gene expression and dosage compensation during pre-implantation development

Julia Bammer[1,2], Evgeniy Ozonov[1], Helene Royo[1,4], Peter Nestorov[1,3], Markus Rempfler[1], Sebastien Smallwood[1] and Antoine Peters[1,5]

[1] Friedrich Miescher Institute for Biomedical Research (FMI), Maulbeerstrasse 66, 4058 Basel, Switzerland
[2] Faculty of Sciences, University of Basel, 4056 Basel, Switzerland
[3] Current address: Witec AG, Industriestrasse 12, 6210 Sursee, Switzerland
[4] Current address: Roche Pharmaceutical Research and Early Development (pRED), F. Hoffmann-La Roche AG, Grenzacherstrasse 124, 4070 Basel, Switzerland
[5] Corresponding author: Email: antoine.peters@fmi.ch, Phone: +41616978761, Fax: +41616973976

Keywords
Polycomb repressive complex 2, intergenerational inheritance, epigenetic memory, X chromosome inactivation

2.1 Abstract

Polycomb group proteins have been shown to be crucial for differentiation of embryonic stem cells as well as during post-implantation development of embryos *in vivo.* However, much less is known about their function during pre-implantation development. In order to investigate the role of PRC2 during early mouse embryogenesis, we generated embryos maternally and/or zygotically deficient for the core PRC2 component EED. We show that these embryos develop to the blastocyst stage, but are delayed compared to wild type controls. By single embryo RNA-seq, we demonstrate that PRC2 establishes H3K27me3 in the oocyte, which is transmitted to embryos where it mediates maternal allele-specific repression at a set of genes. In line with this allele-specific mode of gene regulation, we show that deletion of *Eed* in oocytes leads to ectopic expression of maternal *Xist* and inactivation of the maternal X chromosome in embryos, indicating that H3K27me3 serves as the maternal repressive imprint for *Xist.* We further show that PRC2 dynamically regulates genes involved in development and differentiation at different stages of pre-implantation development. Finally, we demonstrate that loss of maternal PRC2 affects post implantation development, suggesting that the maternal function of PRC2 during early pre-implantation development is important for proper embryonic development after implantation.

2.2 Introduction

Polycomb group (PcG) proteins are evolutionarily conserved transcriptional repressors that were first identified in *Drosophila* as factors required to maintain repression of homeotic genes during embryonic development [34]. More recently, it has been shown that PcG proteins are involved in more dynamic modes of gene silencing and regulate a plethora of cellular processes during development and tumorigenesis [reviewed in 195]. PcG proteins are generally found as components of two major classes of complexes, the Polycomb repressive complexes 1 and 2 (PRC1 and PRC2), that catalyze monoubiquitination of H2AK119 (H2AK119ub1) and trimethylation of H3K27 (H3K27me3), respectively [reviewed in 138]. The molecular function of H3K27me3 is still not fully understood. In flies, it has been shown that a point mutation in lysine 27 of H3 fails to repress transcription of PcG targets and causes homeotic transformations, suggesting that H3K27 is a crucial PRC2 substrate responsible for gene repression [68]. In addition, H3K27me3 has been shown to be recognized by the aromatic cage at the center of the WD-propeller structure of EED *in vitro*, possibly allowing maintenance and propagation of H3K27me3 at sites that already carry the PRC2 mark. Interestingly, binding of Esc, the Drosophila homologue of EED, to H3K27me3 promotes the lysine methyltransferase activity of E(z). This self-reinforcing, positive feedback loop might contribute to the maintenance of transcriptionally silent chromatin domains and the transmission of the H3K27me3 mark from mother to daughter cells [70].

PcG proteins have been extensively studied in mouse embryonic stem cells (ESCs), where they occupy genes that are required later during development and are thus important regulators of pluripotency and differentiation [reviewed in 328]. However, much less is known about their role *in vivo* during mammalian early embryonic development. At the onset of life, fusion of two dimorphic gametes gives rise to a totipotent embryo with the potential to form all embryonic and extraembryonic lineages of the organism. Initially, both parental genomes are transcriptionally silent and the embryo relies on transcripts and proteins inherited from the oocyte. At the late 1-cell and 2-cell stages, two waves of zygotic genome activation (ZGA) occur, in which the developmental program that is initially directed by maternally inherited factors is replaced by a new program as a consequence of zygotic gene expression. Notably, ZGA is characterized by activation of a large number of transposons and repetitive elements including endogenous retroviruses that are silenced in

most other cell types [249-251]. Many of these elements are expressed early and become repressed after the 2-cell stage to prevent mutations and maintain genome integrity in the early embryo [252, 253]. The most highly activated family of repeats in 2-cell embryos are MuERV-L retroviruses, of which many serve as alternative promoters for the activation of protein-coding genes by generating chimaeric transcripts [254].

As development progresses, the embryo undergoes subsequent cell divisions and cells begin to acquire different cell fates. The first lineage specification event gives rise to two distinct cell populations at the early blastocyst stage, with outer cells forming the extraembryonic trophectoderm (TE) and inner cells the embryonic inner cell mass (ICM). This first cell fate choice is initiated by the first asymmetric cell divisions in the embryo prior to blastocyst formation and results in outer cells expressing transcription factors (TFs) like CDX2 and EOMES and inner cells expressing pluripotency factors like OCT4, SOX2 and NANOG [reviewed in 284]. It has been shown that CDX2 is crucial for setting up these gene regulatory networks by repression of OCT4 and NANOG in the TE lineage [283, 329, 330]. Interestingly, the first lineage choice is already initiated at the 4-cell stage, where heterogeneity in the nuclear localization of SOX2 and OCT4 among blastomeres drives lineage segregation [272, 273]. The ability of SOX2 to bind DNA is regulated by histone H3R26me2 that is deposited by the methyltransferase CARM1 together with PRDM14, both of which are heterogeneously expressed at the 4-cell stage [272, 274, 275]. Recently it has been shown that also SOX21, a target of OCT4 and SOX2, is heterogeneously expressed in 4-cell embryos and drives cell fate by repressing expression of CDX2 [276]. The second lineage specification event takes place at the late blastocyst stage and further differentiates the ICM into the epiblast (EPI), which will give rise to all the cells of the future body and the primitive endoderm (PE), the second extraembryonic tissue. This developmental choice is orchestrated by expression of the TFs GATA6 in PE cells and NANOG in EPI cells [285].

Another key event in early embryos is the inactivation of one X chromosome in females in order to achieve dosage compensation for X-linked genes. In pre-implantation development, X chromosome inactivation (XCI) commences around the 4-cell stage in an imprinted fashion, meaning that the paternal X chromosome (Xp) is silenced [295]. XCI is controlled by the X-inactivation centre (XIC) that contains the non-coding locus *Xist*, which is expressed monoallelically from the future inactive X chromosome and its long non-coding

RNA coats the chromosome in *cis* [301-305]. A long-standing question in the field is the nature of the parent-specific imprint underlying silencing of the Xp during imprinted XCI. The existence of a maternal imprint is suggested by experiments in parthenogenetic embryos, in which both X chromosomes are maternally inherited and remain active during pre-implantation development [320]. Nuclear transfer experiments have revealed that the imprint on the maternal X (Xm) is established during oocyte maturation [321]. Analysis of embryos maternally deficient for *de novo* DNA methyltransferases has shown that oocyte DNA methylation is dispensable for *Xist* imprinting [324]. Overexpression of *Kdm4b*, a H3K9me3 demethylase, leads to partial de-repression of *Xist* in parthenogenetic embryos [325]. However, H3K9me3 does not become enriched at the *Xist* promoter during oocyte maturation, when the imprint is set up, and thus does not seem to be the primary imprinting mark [326]. Recently, Inoue *et al.* showed that injection of the histone H3K27me3 demethylase *Kdm6b* into zygotes leads to maternal *Xist* expression and maternal XCI in pre-implantation embryos, suggesting that maternal H3K27me3 serves as the imprinting mark for *Xist* [327].

Typically, zygotic deletion of PcG proteins causes a developmental arrest during gastrulation, pointing towards a role in specification and maintenance of cell lineages [45, 166, 167, 171]. Maternal deletion of both catalytic PRC1 subunits, *Ring1* and *Rnf2*, leads to a developmental arrest at the 2-cell stage, caused by massive transcriptional misregulation in the oocyte [173]. However, much less is known about the maternal function of PRC2. Recent advances in ChIP-seq have shed light on the dynamics of H3K27me3 in early embryos [217, 218]. Most of the canonical H3K27me3 peaks from sperm are not detected on the paternal allele after fertilization. On the maternal allele, broad domains of H3K27me3 in intergenic regions and gene deserts are transmitted from oocytes to early embryos. In contrast, the mark is decreased at promoters of developmental genes where it starts to re-appear in the ICM of the blastocyst and increases in the epiblast of post-implantation embryos. Despite the reduction of H3K27me3 at promoters in pre-implantation embryos, most developmental genes remain silent, suggesting either the presence of additional repressive mechanisms or a lack of activators. H3K27me3 in distal regions in contrast is maintained throughout pre-implantation development, but lost in post-implantation embryos. The inheritance of H3K27me3 from oocytes to early embryos raises the possibility for a maternal function of

PRC2 in pre-implantation development. More evidence in this direction comes from a recent study that identified 76 genes with paternal allele-specific DNase I hypersensitivity sites that lack DNA methylation, but harbor maternal allele-specific H3K27me3. The authors showed that these genes are paternally expressed in wild type embryo and gain maternal expression upon removal of H3K27me3 by injection of *Kdm6b*. While most of the identified genes show maternal-specific repression only during early pre-implantation development, at least five of them (*Gab1*, *Phf17*, *Sfmbt2*, *Slc38a4* and *Smoc1*) maintain their allelic expression pattern in the extraembryonic lineage of post-implantation embryos, suggesting a possible function in placenta development [234]. Among the genes upregulated on the maternal allele upon demethylase injection is *Xist,* suggesting that H3K27me3 serves as the maternal repressive imprint that keeps the maternal X chromosome active in early embryos [327]. In Drosophila, it has been shown that maternal H3K27me3 is propagated in early embryos, where it prevents the inappropriate activation of lineage-specific enhancers [209].

Here we address the maternal and zygotic functions of PRC2 in early embryogenesis by deleting *Eed* in growing oocytes and during spermatogenesis. We study the transcriptional effects in embryos maternally and/or zygotically deficient for *Eed* by single embryo RNA-sequencing (RNA-seq). We show that loss of maternal and zygotic PRC2 does not affect development of embryos to the blastocyst stage, but causes a delay in developmental progression. Using a method to integrate the developmental timing of embryos into differential gene expression analysis, we identify that PRC2 regulates a set of genes that is repressed by H3K27me3 inherited from oocytes to embryos specifically on the maternal allele. We further show that genes repressed by PRC2 during pre-implantation development follow different dynamics and have functions in development and differentiation. We identify a set of genes that likely rely on maternally inherited H3K27me3 for repression during early stages of pre-implantation development and another set that depends on the *de novo* function of PRC2 in order to be repressed at later stages of pre-implantation development. We further provide more evidence that H3K27me3 serves as the maternal repressive imprint for *Xist* by showing that deletion of maternal *Eed* leads to de-repression of maternal *Xist* and ectopic inactivation of the maternal X chromosome in both female and male embryos. Finally, we demonstrate that the function of maternal PRC2 is required for proper post-implantation development.

2.3 Results

2.3.1 Deletion of *Eed* in oocytes leads to loss of H3K27me3 and minor transcriptional misregulation

We used a conditional knock out (KO) approach to delete exons 3-6 of *Eed* in primary oocytes (*Eed* $^{F/F}$ *Gdf9-icre*) as well as in late spermiogenesis (*Eed* $^{F/-}$ *Prm1-cre*) (Supp. Fig. 2.1A-D). Deletion of *Eed* in primary oocytes by *Gdf9-icre* leads to a global depletion of H3K27me3, which by the end of oogenesis is reduced to near-background levels in non-surrounded nucleolus (NSN) as well as surrounded nucleolus (SN) germinal vesicle (GV) oocytes (Fig. 2.1A, B). We did not observe a difference in the numbers of metaphase II (MII) oocytes upon super ovulation of *Eed* $^{F/F}$ *Gdf9-icre* and *Eed* $^{F/F}$ females (Supp. Fig. 2.1E). Moreover, also the number of misformed oocytes was comparably low with around 10% of isolated oocytes upon super ovulation in both genotypes (data not shown). In order to investigate transcriptional changes upon deletion of *Eed* during oogenesis, we performed RNA-seq of GV oocytes. Principal component analysis (PCA) separates *Eed* $^{F/F}$ from *Eed* $^{F/F}$ *Gdf9-icre* oocytes along principal component 1 that explains around 35% of the variance in the dataset (Supp. Fig. 2.1F). Differential gene expression analysis revealed upregulation of 277 genes and downregulation of 19 genes (maxFDR=0.05; minlog2(FC)=1) in *Eed* $^{F/F}$ *Gdf9-icre* compared to *Eed* $^{F/F}$ GV oocytes (Fig. 2.1C). In contrast, deletion of *Ring1* and *Rnf2* by *Zp3-cre* results in upregulation of 2448 genes and downregulation of 507 genes (maxFDR=0.05; minlog2(FC)=1), confirming the fundamental role of PRC1 in transcriptional regulation in oocytes (Fig. 2.1D, Supp. Fig. 2.1G) [173]. Thus, we conclude that PRC2 plays a role in gene regulation in oocytes, but to a much lower extent than PRC1 (Fig. 2.1E). Taken together, our analysis of *Eed* deficient oocytes shows that absence of PRC2 in oocytes leads to a strong reduction in H3K27me3, but causes only minor transcriptional misregulation and does not affect oocyte development.

2.3.2 Embryos deficient for PRC2 develop to the blastocyst stage, but are delayed and show defects in lineage specification

We next investigated the function of PRC2 in early embryogenesis during the acquisition of totipotency and lineage specification. *Eed* $^{F/F}$ *Gdf9-icre* females were super ovulated and mated naturally with *Eed* $^{F/-}$ *Prm1-cre* males to generate embryos maternally and zygotically deficient for *Eed* (referred to as *Eed* $^{m-z-}$ embryos). Control embryos (*Eed* $^{m+z+}$) were generated

by crosses of *Eed* $^{F/F}$ males and females (Fig. 2.2A). Upon 4 days in culture, at embryonic day 4.5 (E4.5), around 80% of both *Eed* $^{m+z+}$ and *Eed* $^{m-z-}$ embryos reached the blastocyst stage. However, *Eed* $^{m-z-}$ embryos developed at a lower developmental rate, leading to a developmental delay. This was most pronounced at embryonic day 3.5 (E3.5) where around 50% of *Eed* $^{m+z+}$ embryos had formed a blastocyst, compared to only 20% of *Eed* $^{m-z-}$ embryos (Fig. 2.2B). By immunofluorescence (IF) staining, we confirmed absence of H3K27me3 from *Eed* $^{m-z-}$ embryos at the morula/early blastocyst stage at E3.5 (Fig. 2.2C). H3K27me3 was still undetectable in *Eed* $^{m-z-}$ embryos at the late blastocyst stage at E4.5 (data not shown). In order to further characterize the developmental delay of *Eed* $^{m-z-}$ embryos and to investigate possible effects of *Eed* deletion on lineage specification, we performed IF staining for CDX2 and OCT4 in E4.5 *Eed* $^{m+z+}$ and *Eed* $^{m-z-}$ embryos (Fig. 2.2D). We developed a workflow for segmentation of nuclei in 3D-stacks of stained blastocysts, allowing us to automatically count nuclei and quantify staining intensities. Our analysis revealed that *Eed* $^{m-z-}$ embryos consist of significantly less cells than *Eed* $^{m+z+}$ embryos (average of 52 vs. 61 nuclei per embryo), suggesting that *Eed* $^{m-z-}$ embryos are still slightly delayed at E4.5 (Fig. 2.2E). Interestingly, we observed a significant increase in OCT4 staining and a significant decrease in CDX2 staining in nuclei of *Eed* $^{m-z-}$ embryos (Fig. 2.2D,F). Together, these results indicate that PRC2 and H3K27me3 are not absolutely required for the acquisition of totipotency and formation of blastocyst embryos. However, deletion of *Eed* leads to a substantial delay in pre-implantation development and significant changes in the levels of lineage markers, suggesting that PRC2 plays a role in orchestrating the processes involved in early embryogenesis including proper specification of the ICM and TE lineages.

2.3.3 Single-embryo RNA-seq to investigate the transcriptional role of PRC2 in pre-implantation development

To investigate the transcriptional function of PRC2 during pre-implantation development, we performed Smart-seq2 single embryo RNA-seq of *Eed* $^{m-z-}$ and *Eed* $^{m+z+}$ embryos at embryonic day 3.5 (E3.5) and 4.5 (E4.5). To study potential allele-specific effects of maternal *Eed* deletion on gene expression, we performed a second single embryo RNA-seq experiment using hybrid embryos that were generated by crossing *Eed* $^{F/F}$ *Gdf9-icre* (C57Bl/6 background) females with JF1/MsJ wild type males. The resulting embryos are maternally deficient, but zygotically heterozygous for *Eed (Eed* $^{m-z+}$). As a control, we crossed *Eed* $^{F/F}$ females with JF1/MsJ wild

type males to generate *Eed* $^{m+z+}$ hybrid embryos. In order to investigate the importance of the enzymatic function of PRC2 during embryonic development on gene expression, we treated part of the *Eed* $^{m+z+}$ embryos with a combination of two PRC2 inhibitors, the EED inhibitor A-395 and the EZH1/2 inhibitor UNC1999 [331, 332]. The second experiment included embryos at E2.5, E3.5 and E4.5 (Fig. 2.3A). Each embryo was imaged prior to library preparation and sub-staged according to morphological features in order to relate the heterogeneity in embryo morphology to gene expression profiles (Supp. Fig. 2.2A). We next assessed the level of H3K27me3 in embryos of these different genotypes/ treatments by IF at E3.5. H3K27me3 is highly reduced in *Eed* $^{m-z-}$ as well as *Eed* $^{m-z+}$ embryos, arguing that the zygotic function of *Eed* is not sufficient to rescue the global H3K27me3 levels in pre-implantation embryos (Fig. 2.3B,C). In *Eed* $^{m+z+}$ embryos treated with PRC2 inhibitors, we observed a strong reduction of H3K27me3 staining in late zygotes, particularly on the male pronucleus (data not shown). However, at E3.5, inhibitor treated embryos show substantial, yet significantly reduced levels of H3K27me3, compared to untreated embryos (Fig. 2.3B,C). These results suggest that the inhibitors are efficiently reducing H3K27me3 levels after fertilization, but not during late stages of pre-implantation development.

Since we knew from previous experiments that the deletion efficiency of Prm1-cre is only around 90%, we checked for efficient deletion of exons 3 to 6 in *Eed* $^{m-z-}$ embryos. In most of the embryos, we did not detect any reads from these exons. However, 2 embryos at E3.5 and E4.5, respectively, showed a significant number of reads coming from exons 3 to 6 of *Eed* (Supp. Fig. 2.2B). We therefore suspected that these embryos are heterozygous and inherited an intact allele from the father. In order to confirm this, we quantified reads at the junction between exons of *Eed*. In *Eed* $^{m+z+}$ embryos, reads at the junction between all exons around the deletion site are detected. In *Eed* $^{m-z-}$ embryos, such reads are mostly missing, and a new splicing event is detected, from the exon upstream to the exon downstream of the deletion site. However, 2 *Eed* $^{m-z-}$ embryos at E3.5 and E4.5, respectively exhibit junction reads that resemble the pattern in *Eed* $^{m+z+}$ embryos, confirming that these embryos carry an intact allele of *Eed*, most probably due to inefficient deletion by Prm1-cre (Supp. Fig. 2.2C). We therefore excluded these embryos from further analysis. The sex of each embryo was determined by the presence or absence of reads coming from the Y-chromosome and expression of the Y-linked gene *Eif2s3y* (Supp. Fig. 2.2D).

2.3.4 Pseudotiming of embryos confirms developmental delay and allows investigation of PRC2-dependent differential gene expression

In light of the developmental delay of *Eed* $^{m-z-}$ embryos, we were interested in the developmental timing of embryos in the RNA-seq experiments. Therefore, we compared relative expression in our samples to that of an external dataset including single embryo RNA-seq of embryos from 8-cell to E4.5 stage [333]. In general, our samples correlate well with the respective time points in the external dataset. However, *Eed* $^{m-z-}$ as well as *Eed* $^{m-z+}$ embryos are systematically behind the control embryos in the developmental path, confirming that deletion of *Eed* in early embryos causes a developmental delay (Fig. 2.3D). Since this could potentially bias the results of differential gene expression analysis towards effects that are solely due to the delay and not directly related to the loss of *Eed*, we developed a method to assign a pseudotime to each embryo, for which we can correct in differential gene expression analysis. Comparison of pseudotime to morphological staging of embryos showed a good correlation. The biggest difference in pseudotime was detected at E3.5, matching the transition between morula and early blastocyst stage (Fig. 2.3E). Comparison of pseudotime for the different genotypes in the two experiments confirmed the developmental delay observed before. *Eed* $^{m-z-}$ embryos are substantially delayed at all timepoints analyzed (E3.5 and E4.5). *Eed* $^{m-z+}$ embryos are delayed at E2.5 and E3.5, but seem to catch up at E4.5. Interestingly, embryos treated with the PRC2 inhibitors are substantially delayed at E3.5, but not at E2.5 and E4.5 (Supp. Fig. 2.2E). These results are in accordance with the data from morphological staging of embryos, where we observed a delay of *Eed* $^{m-z-}$, *Eed* $^{m-z+}$ and inhibitor treated embryos at all stages, with the biggest effect at E3.5 (Supp. Fig. 2.2F).

After correcting for time, strain and differences in library preparation, we performed differential gene expression analysis. We observe comparable numbers of up- and downregulation of genes at all stages analyzed. In general, more genes are misregulated in male embryos, particularly at E3.5 and E4.5. In the PRC2 inhibitor treated embryos, we observe considerable misregulation of genes at E2.5, but not at later stages. This result is in line with the relatively high global H3K27me3 levels at E3.5 detected by IF (Fig. 2.3B, C). X-linked genes show a particularly interesting pattern, with a few upregulated and a relatively high number of downregulated genes in both male and female embryos at all stages, in *Eed* $^{m-z-}$ and *Eed* $^{m-z+}$, but not in inhibitor treated embryos (Supp. Fig. 2.3A).

2.3.5 Identification of maternal allele-specific PRC2 targets using t-SNE analysis

In order to identify groups of differentially expressed genes with specific expression patterns, we performed t-distributed stochastic neighbor embedding (t-SNE) analysis on fitted log2 Fold-changes in each contrast. For the purpose of visualization and subsequent analysis, we performed k-means clustering of t-SNE coordinates, resulting in 12 clusters (Fig. 2.4A). We first analyzed expression changes in t-SNE space for autosomal genes. In general, we find that genes, which change upon deletion or inhibition of *Eed* are located on the outside of the t-SNE plot, whereas genes at the center do not change between the conditions analyzed (Fig. 2.4B). Overall, we observe similar changes in gene expression between the three conditions analyzed. In embryos, maternally and zygotically deficient for *Eed*, the observed transcriptional effects increase from E3.5 to E4.5. The effect of maternal deletion of *Eed* is strongest at E2.5 and then decreases gradually. These overall dynamics illustrate the transition from maternal to zygotic functions of PRC2 during pre-implantation development. Embryos treated with PRC2 inhibitors show similar expression patterns to *Eed*$^{m-z+}$ embryos at E2.5, but display only very minor gene misregulation at later stages. Even though the reason for the difference between stages remains elusive, the similarity between E2.5 embryos maternally deficient for *Eed* and embryos treated with PRC2 inhibitors highlights the importance of the enzymatic activity of the PRC2 complex for gene regulation in pre-implantation embryos. Clusters 1, 2 and 3 contain mostly genes that are upregulated upon *Eed* deletion/PRC2 inhibition, while clusters 5,6 and 8 contain genes that go down upon *Eed* deletion/PRC2 inhibition. Interestingly, we observe both stage- and sex-specific effects. At E2.5 most downregulated genes are found in cluster 8, while at E3.5 they are most enriched in cluster 6. Upregulated genes at E2.5 are mostly located in cluster 3, while at E3.5 and E4.5 they are enriched in cluster 1. These results suggest that PRC2 targets are dynamic during pre-implantation development. Sex-specific effects are present in clusters 2 and 3, where some genes show the inverse behavior in male and female embryos at E2.5. This difference between the sexes is present in both maternal deficient as well as inhibitor treated embryos (Fig. 2.4B).

On the level of allelic expression, we find maternally and paternally expressed genes mostly intermingled in the t-SNE space with a slight enrichment for paternally expressed genes in cluster 1 (Supp. Fig. 2.3B). However, when we analyzed the change in allelic

expression between embryos maternally deficient for *Eed* and control embryos, we observed enrichment for genes that gain maternal expression in mutant embryos in clusters 1,2 and 3. Interestingly, this effect was only observed in *Eed* $^{m-z+}$ and not in inhibitor treated embryos (Fig. 2.4C, Supp. Fig. 2.3C). These results suggest that a set of genes is repressed by PRC2 specifically on the maternal allele in pre-implantation embryos. Since these genes did not gain maternal expression in embryos treated with PRC2 inhibitors after fertilization, we speculate that they rely on the function of PRC2 in the oocyte.

To relate changes in gene expression upon *Eed* KO to the PRC2 mediated chromatin mark, we used a published dataset from Liu *et al*. including chromatin immunoprecipitation sequencing (ChIP-seq) for H3K27me3 in MII oocytes and different stages of pre-implantation embryos [217]. We observe that the levels of H3K27me3 are non-uniform across stages (Supp. Fig. 2.4A). Therefore, in order to compare between H3K27me3 at different stages, we applied gaussian mixture models with 3 components to classify promoters (±2.5kb around TSS) into 3 categories (Low, Medium and High) for H3K27me3 at each stage. We observe a general trend of de-methylation in 2-cell embryos and re-methylation in the ICM of the blastocyst (Supp. Fig. 2.4B). In t-SNE space, genes containing medium and high H3K27me3 promoters are mostly concentrated in clusters 1 and 3. However, also clusters 6 and 8 that contain genes downregulated upon *Eed* deletion, are enriched for medium and high H3K27me3 targets. (Fig. 2.4D, Supp. Fig. 2.4C). Genes with low H3K27me3 promoters are mostly concentrated in the center of the t-SNE space, coinciding with genes that do not change upon *Eed* deletion (Supp. Fig. 2.4C). We further investigated the allelic levels of H3K27me3 using a published dataset from Zheng *et al*. [218]. After confirming reproducibility of the data for 2-cell stage embryos (Supp. Fig. 2.4D,E), we analyzed allelic H3K27me3 signal at the 2-cell stage in t-SNE space (Supp. Fig. 2.4F). We found that maternal H3K27me3 target genes are enriched in cluster 1 (Fig. 2.4E). We therefore speculate that cluster 1 contains a set of genes that are repressed by H3K27me3 specifically on the maternal allele and become de-repressed maternally in early embryos upon deletion of *Eed* in the oocyte.

Further support for this hypothesis comes from analysis of biased expression in androgenetic (AG) and gynogenetic (GG) morula stage embryos from Inoue *et al.* (2017) in t-SNE space [234]. We confirmed reproducibility of the data and determined log2(FC) between AG and GG embryos (Supp. Fig. 2.4G). In t-SNE space, AG-specific genes are enriched in Cluster 1, coinciding with genes that carry maternal H3K27me3 and become de-repressed maternally

upon *Eed* deletion. In contrast, genes with GG-specific expression are most concentrated in clusters 6 and 8 (Fig. 2.4F, Supp. Fig. 2.4H). Finally, we show that a set of genes carries maternal H3K27me3 in pre-implantation embryos and becomes de-repressed maternally upon deletion of *Eed* in oocytes, but not upon PRC2 inhibitor treatment of embryos (Fig. 2.4G, Supp. Fig. 2.4I).

In order to exclude the possibility that the misregulation of genes in pre-implantation embryos is due to inheritance of aberrant transcripts from the oocyte, we analyzed expression changes in *Eed* deficient GV oocytes for genes in t-SNE space. We observe that only few genes are differentially expressed in GV oocytes, with a slight enrichment of upregulated genes in cluster 1. In general, we observe low correlation between differentially expressed genes in *Eed* deficient GV oocytes and embryos (Supp. Fig. 2.4K). These results suggest that the transcriptional effects in embryos arise from aberrant zygotic transcription and are independent of gene misregulation in the oocyte.

Taken together, our results demonstrate that maternal PRC2 is required to establish H3K27me3 in oocytes that is inherited to pre-implantation embryos, where it controls the maternal-specific repression of genes. However, this allele-specific mode of gene regulation concerns relatively few genes and the gain in maternal expression upon *Eed* deletion does not correlate with a major upregulation of total gene expression (Supp. Fig. 2.4J). Our results are in line with a recent study that reported H3K27me3-dependent imprinting for a small group of genes in early embryos [234].

2.3.6 PRC2 dynamically regulates genes involved in development and differentiation

Based on the t-SNE analysis, we identified a group of genes in clusters 1,2 and 3, that is upregulated upon deletion of *Eed* and enriched for H3K27me3 on their promoters in early embryos, suggesting that this group of genes constitutes a set of direct PRC2 targets in pre-implantation embryos. We therefore focused our further analysis on genes in these three clusters. In order to understand more about the developmental dynamics and functions of these genes, we selected autosomal genes from t-SNE clusters 1, 2 and 3 that are up- or downregulated upon *Eed* deletion/PRC2 inhibition (FDR ≤ 10% and abs(log2(FC)) ≥ 1) or show a strong change in allelic bias upon *Eed* deletion/PRC2 inhibition (FDR ≤ 5% and abs(log2(FC)) ≥ 2). According to these criteria, we identified 1767 genes. We further split them into high

and low CpG genes (HCP and LCP) and performed k-means clustering based on expression dynamics in wild type embryos and the effect of *Eed* deletion/PRC2 inhibition (Fig. 2.5A, Supp. Fig. 2.5A). In both HCP and LCP classes, we identified clusters of genes with an increase in expression from E2.5 to E4.5 (clusters HCP K1, K2, K3 and LCP K1, K2). A second group of gene clusters shows inverse dynamics with decreasing expression from E2.5 to E4.5 (HCP K6, K7 and LCP K4, K5). We also identified clusters of genes with increasing expression from E2.5 to E3.5 and decreasing expression at E4.5 (HCP K4 and LCP K3). Cluster HCP K5 shows relatively stable expression with little changes from E2.5 to E4.5. In general, we observe that genes, which show increasing expression during development are upregulated early upon *Eed* deletion/PRC2 inhibition, suggesting that these genes either rely on maternally inherited H3K27me3 for repression during early stages of pre-implantation development or become *de novo* repressed by PRC2 at early pre-implantation stages. Surprisingly, we identified 2 clusters of HCP genes (K2 and K3) that are upregulated specifically in either male or female embryos. In contrast, genes with increasing expression during development are mostly upregulated at E4.5 upon maternal and zygotic deletion of *Eed*, arguing that these genes rely on *de novo* function of PRC2 during embryonic development in order to repress them at later stages of pre-implantation development (Fig. 2.5A).

In order to investigate the function of PRC2 target genes in pre-implantation embryos, we performed slim Gene Ontology (GO) term analysis for t-SNE clusters 1,2 and 3. Genes in the respective clusters were grouped into HCP and LCP genes and GO term analysis was performed separately (Fig. 2.5B, Supp. Fig. 2.5B). We found that HCP genes in clusters 1,2 and 3 are enriched for GO terms relating to cell differentiation, development, cell signaling, cell motility and adhesion and metabolic, nervous system and circulatory system processes (Fig. 2.5B). Genes in these GO term categories include genes involved in cell fate choices in early embryos like *Otx2, Pou3f1 and Sox21*, several *Hox* family members and genes involved in WNT- and FGF-signaling pathways (Fig. 2.5C). These results indicate that PRC2 regulates key developmental processes and cellular differentiation in pre-implantation embryos.

2.3.7 PRC2 functions in the maternal allele-specific repression of X-linked genes

We next investigated the expression dynamics for X-linked genes in t-SNE space (Fig. 2.6A). We observe particularly interesting changes on the X chromosome upon maternal deletion of

Eed. In both male and female *Eed* $^{m-z+}$ and *Eed* $^{m-z-}$ embryos, many X-linked genes are downregulated (located mainly in t-SNE clusters 6 and 8), while few genes are upregulated, including genes from the *Bex*, *Rhox* and *Fthl17* families that appear in gene clusters as well as the X chromosome inactivation regulator *Xist* (located in t-SNE clusters 1 and 3) (Fig. 2.6B). Interestingly, many of these genes have been shown to escape X inactivation and to be expressed predominantly from the paternal allele in pre-implantation embryos, leading to expression exclusively in female embryos [301, 334, 335]. For *Rhox5* and *Xist*, it has been suggested that this imprinted expression pattern in pre-implantation embryos is independent of DNA methylation [336, 337]. Notably, the changes on the X chromosome are much stronger upon maternal deletion of *Eed* than upon inhibition of PRC2 after fertilization (Fig. 2.6B, Supp. Fig. 2.6A). We further demonstrate that the X-linked genes that become upregulated upon *Eed* deletion are paternally expressed in wild type embryos, while genes that are downregulated tend to be maternally expressed. Upon maternal deletion of *Eed,* the upregulated genes gain maternal expression, whereas the downregulated genes progressively gain paternal expression over time (Fig. 2.6C, Supp. Fig. 2.6B). In contrast, embryos treated with PRC2 inhibitors do not show any change in allelic bias of X-linked genes compared to controls (Supp. Fig. 2.6B,C). Comparison to allelic ChIP-seq data at 2-cell stage shows that maternal H3K27me3 targets are concentrated mostly in cluster 1 (Supp. Fig. 2.6D). Finally, none of the effects on the X chromosome that we observe in embryos is present in *Eed* deficient GV oocytes (Supp. Fig. 2.6E). Taken together, these results suggest that few genes on the X chromosome, including *Xist*, are repressed by H3K27me3 specifically on the maternal allele. We therefore speculated that deletion of *Eed* in the oocyte might lead to de-repression of the maternal *Xist* allele in early embryos, resulting in global silencing of X-linked genes on the maternal allele. In line with our hypothesis, it has recently been shown that injection of an H3K27me3 demethylase into zygotes leads to de-repression of maternal *Xist* and ectopic inactivation of the maternal X chromosome [327].

2.3.8 H3K27me3 serves as the maternal repressive imprint for *Xist*

We therefore investigated, whether H3K27me3 could be the maternal repressive imprint for *Xist.* Analysis of published ChIP-seq datasets revealed the presence of a broad H3K27me3 domain over the *Xist* locus. This ~450 kb domain spans the region from *Tsix* to *Slc16a2* and

strikingly, we observe upregulation of 4 out of 8 genes and non-coding RNAs within this domain in *Eed* $^{m-z-}$ embryos. Notably, genes outside of this domain are mostly downregulated in *Eed* $^{m-z-}$ embryos, starting with *Chic1* and *Rlim*, that are located directly adjacent to the H3K27me3 domain. The H3K27me3 domain is established in oocytes and transmitted to pre-implantation embryos specifically on the maternal allele (Fig. 2.6D). These data suggest that H3K27me3 might be involved in the imprinted expression of *Xist* during pre-implantation development. In order to confirm that deletion of *Eed* leads to de-repression of the maternal *Xist* allele, we performed RNA FISH using probes against *Xist* as well as the Y-linked gene *Eif2s3y* to distinguish between female and male embryos. Strikingly, we found that *Eed* $^{m-z+}$ females show 2 *Xist* clouds in most nuclei, compared to 1 cloud in control embryos. Male *Eed* $^{m-z+}$ embryos show one cloud in most nuclei, compared to no clouds in control embryos (Fig. 2.6E,F). We observe the presence of ectopic *Xist* clouds in 8-cell embryos at E2.5 as well as in morula embryos at E3.5 (data not shown). Further, the de-repression of maternal *Xist* is present in both *Eed* $^{m-z-}$ as well as in *Eed* $^{m-z+}$ embryos (data not shown). In conclusion, we demonstrate that upon deletion of *Eed* in the oocyte, the maternal repressive imprint for *Xist,* H3K27me3, is lost, leading to de-repression of the maternal *Xist* allele, coating and ectopic XCI of the maternal X chromosome. Interestingly, we do not observe this effect upon treatment with PRC2 inhibitors after fertilization, indicating that removal of the H3K27me3 mark after fertilization is not sufficient to de-repress the maternal *Xist* allele. We thus conclude that the H3K27me3 imprint on *Xist* is inherited from oocytes and required during early stages of pre-implantation development to establish the imprinted fashion of *Xist* expression and X chromosome inactivation.

2.3.9 Maternal deletion of PRC2 impairs post-implantation development

In order to investigate whether the maternal function of PRC2 is required for proper post-implantation development, we performed a breeding experiment in which we crossed *Eed* $^{F/F}$ *Gdf9-icre* or *Eed* $^{F/F}$ females to wild type males. We found that *Eed* $^{m-z+}$ litters were significantly smaller compared to *Eed* $^{m-z+}$ control litters, with a reduction from around 6 to 3 pups per litter (Fig. 2.7A). This result is in line with a recent report that reported reduced litter sizes upon maternal deletion of *Eed* by *Zp3-cre* [338].

2.4 Discussion

In this study, we present a comprehensive investigation of the role of PRC2 in oocytes and pre-implantation embryos. We demonstrate that despite the ability of embryos maternally and zygotically deficient for *Eed* to develop to the blastocyst stage, PRC2 acts as a transcriptional repressor in pre-implantation embryos. We show that PRC2 regulates the expression of genes involved in key functions in early embryos, including cell differentiation, signaling and developmental processes. Notably, we find that PRC2 regulates genes that have been shown to be involved in early lineage specification events including *Pou3f1, Otx2* and *Sox21* [276, 339-341]. *Sox21* has been shown to be asymmetrically expressed among blastomeres of 4-cell embryos and to contribute to the first lineage specification event by counteracting expression of the key TE determinant *Cdx2* [276]. Interestingly, by IF staining, we observe a reduction of CDX2 in nuclei of *Eed* deficient blastocysts. Therefore, it is possible that PRC2 contributes to the correct specification of ICM and TE lineages by regulating the expression of *Sox21* in early embryos.

We find that PRC2 represses different sets of genes in a stage-specific manner during pre-implantation development. On one hand, at early stages of pre-implantation development, it represses genes that need to be activated only at later stages. On the other hand, it represses genes at later stages of pre-implantation development, which are required early and need to be silenced afterwards. These results suggest that the role of PRC2 during pre-implantation development can be divided into two main functions: an early function, which depends on maternally inherited H3K27me3 that is propagated in early embryos and a late function, which depends on *de novo* activity of PRC2 in embryos. Interestingly, among the early targets of PRC2, we observe that certain genes are more affected in male and others in female embryos, suggesting that PRC2 has sex-specific functions in early embryos. It has been shown that male embryos develop more quickly than female embryos and expression of genes, including autosomal genes, is different between the sexes at the blastocyst stage in mouse and bovine embryos [335, 342-344]. Our results raise the possibility that PRC2 is involved in setting up this early asymmetry between the sexes.

It has recently been shown that H3K27me3 represses genes specifically on the maternal allele in pre-implantation embryos and for a few genes, this allelic expression pattern is maintained in extraembryonic lineages of post-implantation embryos [234]. While

the published study was based on injection of an H3K27 demethylase into zygotes, we confirm these results using a genetic approach to abolish the function of PRC2 in oocytes. By single-cell RNA-seq of hybrid embryos, we identified a group of genes that is paternally expressed in wild type embryos and gains maternal expression upon deletion of *Eed.* We show that these genes are enriched for maternal H3K27me3 in 2-cell embryos and preferentially expressed in androgenetic embryos. We further extend the list of genes maternally repressed by H3K27me3 in pre-implantation embryos by genes on the X-chromosome, which are mainly located in gene clusters and have in part been previously described as being imprinted [301, 334, 335, 344, 345]. Among these genes are genes from the *Rhox, Bex* and *Fthl17* families, but also the XCI regulators *Xist* and *Ftx.* It has recently been shown that injection of an H3K27me3 demethylase into zygotes leads to *Xist* expression from the maternal allele and ectopic inactivation of the maternal X chromosome [327]. Here we provide more evidence that H3K27me3 serves as the maternal repressive imprint for *Xist* by showing that genetic deletion of maternal *Eed* leads to de-repression of maternal *Xist* and ectopic inactivation of the maternal X chromosome in both female and male embryos.

Finally, we demonstrate that the function of maternal PRC2 is required for proper post-implantation development. Our findings are in line with a recent report that also showed a reduction in litter sizes, when *Eed* was deleted in oocytes by *Zp3-cre* [338]. Interestingly, the authors further report a postnatal overgrowth phenotype of the offspring that develops to term. We also observed higher body weights of *Eed* $^{m-z+}$ offspring, but attributed them to the smaller litter sizes compared to wild type. It has recently been shown that in embryos generated by somatic cell nuclear transfer (SCNT), H3K27me3-dependent imprinting is lost, likely contributing to post-implantation lethality and a placental overgrowth phenotype [346]. Interestingly, the dysregulated imprinted genes in SCNT embryos include *Slc38a4*, *Sfmbt2* and *Gab1*, which have been previously shown to be involved in the regulation of placental growth and *Runx1*, *Otx2* and *Etv6,* which play important roles in mouse early embryonic development [347-352]. Upon deletion of *Eed*, we observe upregulation of *Etv6*, *Gab1*, *Runx1* and *Sfmbt2* and a shift towards maternal expression of *Gab1, Runx1, Sfmbt2* and *Slc38a4* at different stages of pre-implantation development, suggesting that the post-implantation lethality observed in embryos maternally deficient for *Eed* might result from loss of H3K27me3-dependent imprinting of these genes.

2.5 Material and Methods

2.5.1 Mice

Ring1; Rnf2 double knock out mice were generated as described previously [173]. *Eed* conditional knock out mice were a gift from Prof. Stuart H. Orkin [353]. To obtain oocytes maternally deficient for *Eed,* we generated *Eed* $^{F/F}$ mice that carry the *Gdf9-iCre* recombinase transgene, mediating deletion in growing oocytes. In order to generate sperm deficient for *Eed,* we generated *Eed* $^{F/F}$ mice that carry the *Prm1-cre* recombinase transgene, mediating deletion during late spermatogenesis. Both mouse lines were maintained on C57BL/6 background by transmitting the *cre* though the non-affected germline as described previously [244]. Maternal and zygotic *Eed* deficient (*Eed* $^{m-z-}$) embryos were obtained from matings between *Eed* $^{F/F}$; *Gdf9-icre/+* females and *Eed* $^{F/-}$; *Prm1-cre/+* males. Control (*Eed* $^{m+z+}$) embryos were obtained from matings between *Eed* $^{F/F}$ females and *Eed* $^{F/F}$ males. Maternal *Eed* deficient and control hybrid embryos (*Eed* $^{m-z+}$ and *Eed* $^{m+z+}$) were generated by *in vitro* fertilization (IVF) of *Eed* $^{F/F}$; *Gdf9-icre/+* or *Eed* $^{F/F}$ oocytes, respectively, with sperm from JF1/MsJ males. *Eed* conditional mice were genotyped by PCR on ear biopsy DNA to identify wild-type, floxed and deleted alleles (Supp. Fig. 2.1C, D). The 3 alleles were distinguished using the PCR primers in Table 2.1 resulting in 178bp (wild-type), 233bp (floxed) or 453bp (excised) products, respectively. The mice were genotyped for *Gdf9-icre* or *Prm1-cre* with the primers listed in Table 2.1, yielding a 200bp (*Gdf9-icre*) or 700bp (*Prm1-cre*) product. All experiments were performed in accordance with the Swiss animal protection laws and institutional guidelines.

Table 2.1 PCR primers used for genotyping

Gene name	**Primer pairs**	
	Forward	**Reverse**
Eed flox	CTACGGGCAGGAGGAAGAG	CCACATAGGCTCATAGAATTG
Eed Δ	CTACGGGCAGGAGGAAGAG	GGGGGAGAGGGAGTTGTC
Gdf9-icre	AGATGCCAGGACATCAGGAACCTG	ATCAGCCACACCAGACACAGAGATC
Prm1-cre	GTTCCCTCAGCAGCATTCTC	AGGCAAATTTTGGTGTACGG

2.5.2 Collection of GV oocytes

12-week-old females were injected with pregnant mare serum gonadotropin (PMSG, 5U, MSD). Females were sacrificed 20 hours after PMSG injection and ovaries were transferred to M2 medium. The ovarian follicles were punctured with a 30-gauge needle and the cumulus cells were removed from the oocyte-cumulus complex by gentle pipetting through a narrow glass pipette.

2.5.3 Collection and culture of early mouse embryos

6-20 weeks old female mice were super ovulated by injection of pregnant mare serum gonadotropin (PMSG, 5U, MSD) and 48h later injection of human chorionic gonadotropin (hCG, 5U, MSD). *Eed* $^{m-z-}$ and *Eed* $^{m+z+}$ embryos were generated by natural fertilization. Fertilization occurred at 13-14 hours after hCG injection, which was used as a reference time point for zygote sub-staging and embryonic development. Females with a vaginal plug were sacrificed 20 hours post-hCG (hp-hCG) and zygote-cumulus complexes were isolated from the swollen ampullae. After removal of cumulus cells in Hyaluronidase (1mg/ml, Sigma) in M2 medium, zygotes were washed in M2 medium (Sigma) and cultured in 50µl drops of pre-equilibrated M16 medium (Sigma) covered by mineral oil (Sigma) until 2-cell stage (44hp-hCG). 2-cell embryos were transferred to 50µl drops of pre-equilibrated KSOM+AA with D-Glucose (Millipore) covered by mineral oil (Sigma) and cultured until blastocyst. Embryos were cultured in hypoxia chambers at 37°C in 5%CO_2, 5%O_2, 90%N_2. Hybrid embryos for single-embryo RNA-seq were generated by *in vitro* fertilization (IVF) of MII oocytes from super ovulated females with sperm from JF1/MsJ males. Sperm was isolated from epididymides by gentle squeezing and capacitated in 150µl drops of pre-equilibrated HTF medium covered by mineral oil (Sigma) at 37°C in 5%CO2 for one hour. Oocyte-cumulus complexes were isolated from swollen ampullae and IVF was performed in 50µl drops of pre-equilibrated HTF medium covered by mineral oil (Sigma) at 37°C in 5%CO_2, 5%O_2, 90%N_2 for 5 hours. Zygotes were washed twice in M2 medium and cultured as described above. Part of the control embryos were cultured in media supplemented with 3µM of the EED inhibitor A-395 (SGC) and 3µM of the EZH1/2 inhibitor UNC1999 (gift from M.Frederiksen, NIBR, Basel) [331, 332]. At E2.5, after embryo collection for RNA-seq, the remaining embryos were transferred to fresh drops of pre-equilibrated KSOM+AA with D-Glucose with/without inhibitors.

2.5.4 Immunofluorescence staining (IF) and microscopy

Oocytes or embryos were fixed in 4% paraformaldehyde (PFA, Electron Microscopy Sciences) in PBS for 20 min and permeabilized with 0.5% Triton X-100 in PBS for 30 min at room temperature (RT). After one wash in PBS with 0.1% Tween 20 for 10 minutes and one wash in PBS with 2%BSA and 0.1% Tween 20, for 10 minutes, embryos were blocked for 4 hours in PBS containing 2% bovine serum albumin (BSA), 0.1% Tween 20 and 5% horse serum at 4°C. Oocytes and embryos were incubated with primary antibodies in blocking solution at suitable concentration (Tab. 2.2) overnight at 4˚C. After three 10 minute rinses with PBS with 2% BSA and 0.1% Tween, samples were incubated with secondary antibodies diluted 1:500 in blocking solution for 1 hour at RT. After three rinses with PBS with 0.1% Tween 20, samples were directly transferred onto glass slides (Thermo Scientific), mounted with adhesive imaging spacers (Grace Bio) in Vectashield medium containing DAPI (H-1200, Vector Laboratories) and covered by coverslips. Image acquisition was performed using a Zeiss LSM 700 laser scanning confocal microscope and a Zeiss Axio Imager M2 microscope equipped with a Yokogawa CSU W1 Dual camera T2 spinning disk confocal scanning unit using 40x or 63x objectives. Z-stacks were acquired with a step size of 0.39µM.

Table 2.2. Antibodies used in this study

Antibody	Origin	Catalog number	Dilution
H3K27me3	Cell Signaling	9733	1:100
Oct4	Santa cruz	sc-8628	1:200
Cdx2	BioGenex	MU392A-UC	1:200

2.5.5 Automated segmentation and quantification of 3D IF images of blastocysts

Nuclei segmentation of the 3D stacks was generated with a workflow using fully convolutional neural networks (FCNs) and the watershed algorithm [354, 355]. It employs two FCNs: One is applied to xy-slices and the other to yz-slices of the DAPI channel of the anisotropic 3D stacks. They are designed to predict a dense probability map distinguishing foreground (nuclei) vs background (rest), as well as a separator probability map. This separator map is constructed to peak in between close-by nuclei, thereby acting as a barrier between individual nuclei when partitioning the foreground map into individual nuclei. The partitioning is performed by a

watershed algorithm that is applied to the averaged prediction maps of both FCNs [355]. Both FCNs have the same architecture, a hybrid of the Resnet-50 and the Unet, but are trained individually on slices sampled from three fully annotated blastocyst image volumes [356, 357]. Parameters of the Watershed were hand-tuned. All generated segmentations were manually reviewed and corrected if necessary. To this end, we devised three semi-automatic routines to selectively correct each of the following types of segmentation errors: 1) A nucleus is segmented into multiple pieces (False split), 2) Two or more nuclei are segmented as one (False merge) or 3) a nucleus is merged with parts of apoptotic cells. The CDX2 and OCT4 channels were aligned with the DAPI channel by a translation transformation such that it maximizes the mutual information between DAPI and the OCT4 channel. The final quantification is conducted on the reviewed and corrected segmentations and the aligned image channels. This is implemented in KNIME [358]. For each segmented nucleus, we calculate geometric 3D shape features (volume, surface, compactness, convexity, solidity and sphericity) using the ImageJ extension [359]. In addition to that, we collect its intensity statistics under each nucleus mask in the following channels: DAPI, CDX2 and OCT4.

2.5.6 RNA FISH

Zona pellucida of embryos was removed by incubation in Protease (Sigma, P5147-1 G, 5 mg/ml in M2 medium) for around 10 minutes at 37°C. Embryos were rinsed in M16 medium and fixed in 4% PFA in PBS for 20 minutes. Embryos were permeablized in 0.5% Triton X-100 in PBS for 30 minutes at RT. After 2 washes in PBS, embryos were dried on glass slides and the area around the embryos was marked with a DAKO pen (Agilent). After two 5 minute washes in Stellaris Wash Buffer A (freshly mixed 2ml Stellaris RNA FISH Wash Buffer A (LGC Biosearch Technologies Cat# SMF-WA1-60), 7ml nuclease-free water (Ambion AM9937) and 1 ml deionized formamide (Sigma)), embryos were incubated in hybridization buffer (10% formamide, 2× SSC (Sigma), 10% w/v dextran sulfate) containing 125nM of each probe at 37°C in a humidified chamber. The *Xist* probe was labeled with Quasar 570 fluorophore (LGC Biosearch Technologies SMF-3011-1) and the *Eif2s3y* probe was labeled with Quasar 670 fluorophore (LGC Biosearch Technologies, Custom Stellaris® RNA FISH probe). Hybridization buffer was aspirated and embryos were washed 3 times for 10 min each in wash buffer A at 37˚C. After two 5 minute washes with wash buffer B (LGC Biosearch Technologies Cat# SMF-

WB1-20) at room temperature, embryos were directly mounted in Vectashield medium containing DAPI (H-1200, Vector Laboratories). Image acquisition was performed using a Zeiss Axio Imager M2 microscope equipped with a Yokogawa CSU W1 Dual camera T2 spinning disk confocal scanning unit using 40x or 63x objectives. Z-stacks were acquired with a step size of 0.39µM.

2.5.7 RNA sequencing of GV oocytes

For RNA isolation, 15-25 GV oocytes from one mouse were pooled and RNA was extracted using the RNAEasy Micro Kit (Qiagen, 74004) according to the manufacturer's instructions. Total RNA was quantified and checked for quality on Agilent's BioAnalyzer using the RNA 6000 Pico Kit (Agilent, 5067-1513). RNA was amplified and converted to cDNA using the Ovation v2 Kit (NuGen, 7102). Sequencing libraries were prepared with the Truseq DNA LT kit (Illumina, FC-121-2001) and multiplexed, barcoded libraries were sequenced on an Illumina HiSeq instrument.

2.5.8 RNA sequencing of pre-implantation embryos

Eed $^{m+z+}$ and *Eed* $^{m-z-}$ embryos were generated by super ovulation and natural fertilization and collected at E3.5 and E4.5. Hybrid *Eed* $^{m+z+}$, *Eed* $^{m-z+}$ and PRC2 inhibitor treated embryos were generated by IVF and harvested at E2.5, E3.5 and E4.5. Single embryos were imaged using an Olympus IX71 microscope. After imaging, embryos were washed in PBS and transferred by mouth pipetting to 96-well plates with Smart-seq2 lysis buffer containing ERCC spike-in Mix 1 at a 1/100.000 dilution on ice. Plates were snap-frozen and stored at -80°C until further processing. Library preparation was done using the Smart-seq2 protocol [360]. 14 cycles of preamplification were performed for E3.5 and E4.5 embryos and 16 cycles were performed for E2.5 embryos. RNA-seq libraries were prepared from the single-embryo cDNA libraries (0.2ng/sample) using the Nextera XT kit (Illumina) following the manufacturer's instructions, but using one-fourth volumes. Multiplexed library pools were sequenced on an Illumina HiSeq2500 (around 20 single embryos per lane), generating 50-bp single-end reads.

2.5.9 Alignment of RNA-seq data and expression quantification

Spliced alignment of RNA-seq data to *M.musculus* genome assembly (GRCm38/mm10 Dec. 2011) was done using STAR [361] with parameters *-outFilterMultimapNmax 300 -outMultimapperOrder Random -outSAMmultNmax 1 -alignIntronMin 20 -alignIntronMax 1000000,* allowing multimappers with up to 300 matches in the genome and choosing positions for multimappers randomly.

Read quantification for genes was done using Bioconductor annotation package *TxDb.Mmusculus.UCSC.mm10.knownGene* (version 3.4.0) and qCount function from *QuasR* R package [362] selecting only uniquely mapped reads (parameter mapqMin=255). RPKM values were calculated as described previously [363] and log2 transformed using formula *log2(RPKM + psc) – log2(psc)* where pseudo-count *psc* was set to 0.1. Genotyping of *Eed* $^{m+z+}$ and *Eed* $^{m-z-}$ samples and detection of heterozygous embryos was done by quantification of expression of each exon and each exon-exon splice junctions of *Eed* gene.

2.5.10 Gender inference for embryos using single-embryo RNA-seq data

Chromosomes X and Y (chrX and chrY) were partitioned into 5kb non-overlapping tiles and only those tiles which have at least 80% mappability (80% or more of 50bp windows in a tile are unique in the genome) were selected for further analysis. Uniquely mapped RNA-seq reads overlapping selected 5kb tiles were aggregated for chrX and chrY and percentages were calculated by dividing number of reads in chrX and chrY by total number of mapped reads for each embryo. Genders for embryos were assigned using appropriate cutoff for percentage of reads mapping to chrY.

2.5.11 Pseudotime inference and analysis of differential expression of genes in single embryo RNA-seq data

Pseudotemporal ordering of single embryos was done using functions *reduceDimension* and *orderCells* from *monocle* package of version 2.8.0 [364] with default parameters. Analysis of differential expression was done using generalized linear model (GLM) functionality in *edgeR* package [365]. All single embryo samples and genes with at least 1 read per million in at least 3 samples were included in the analysis. Possible difference in transcriptomes between genotypes due to developmental delay was taken into account by fitting coefficients for basis

natural cubic splines with 3 components for each stage generated by *ns* function in *splines* R package (version 3.5.1), using pseudotime scaled to interval [0;1] for each stage as knots. Possible variation due different genetic background and library preparation was taken into account by adding two additional factors in generalized linear model. More explicitly, design matrix for GLM was generated using *model.matrix* function with model formula *~0 + Group:Gender:Stage + ns(ScaledPseudotime):Stage + Strain + LibraryPrep*, where *Group* includes genotypes or treatment (i.e. *m+/z+, m-/z+, m-/z-* and *PRC2 inh.*), *Strain* and *LibraryPrep* factors reflect differences in genetic background and library preparation. P-values for differential expression were calculated using log-likelihood tests and correction for multiple testing was done using Benjamini-Hochberg method. t-SNE plots were generated using *Rtsne* package [366] with perplexity 100 on matrix of expression changes (log2(FC)) for all contrasts of interest. Gene Ontology analysis was performed using Slim GO annotation provided by Gene Ontology Consortium and topGO R package. Heatmaps were generated using *ComplexHeatmap* package [367].

2.5.12 Maternal/paternal allele inference for RNA-seq data for Bl6xJF1 hybrid embryos and analysis of allelic bias in expression of genes

Annotation of single nucleotide polymorphisms (SNPs) between JF1 and Bl6 mouse strains from previous publication [368] was used for assigning allelic origin for RNA-seq reads. Locations of SNPs were mapped from mm9 to mm10 genome using UCSC *liftOver* tool. RNA-seq reads were separately aligned to Bl6 and JF1 genomes, which were created by injecting SNPs at corresponding positions in the reference genome. Allelic origin for each read was assigned by comparing number of mismatches in corresponding alignments to Bl6 and JF1 genomes. In vast majority of samples, percentage of maternal reads is around 50%. However, there are a few samples with abnormally high (higher than 60%, 5 samples) or low (lower than 40%, 2 samples) percentages of maternal reads which were removed from the analysis.

Number of maternal and paternal reads for genes was quantified using *qCount* function in *QuasR* package selecting only uniquely mapped reads. Only genes with at least 10 allelic reads in at least 6 samples were selected for the analysis. Statistical analysis of allelic bias in gene expression was performed using *edgeR* package. Design matrix for GLM was constructed using *model.matrix* function with formula *~0+Group:Gender:Stage:Allele*, where, as for differential expression analysis, *Group* includes genotypes or treatment (*m+/z+, m-/z+,*

m-/z- and *PRC2 inh.*) and library sizes for each allele were set up equal and chosen to be average between total maternal and paternal read counts. Statistical significance was calculated using log-likelihood tests and adjusted for multiple testing using Benjamini-Hochberg method.

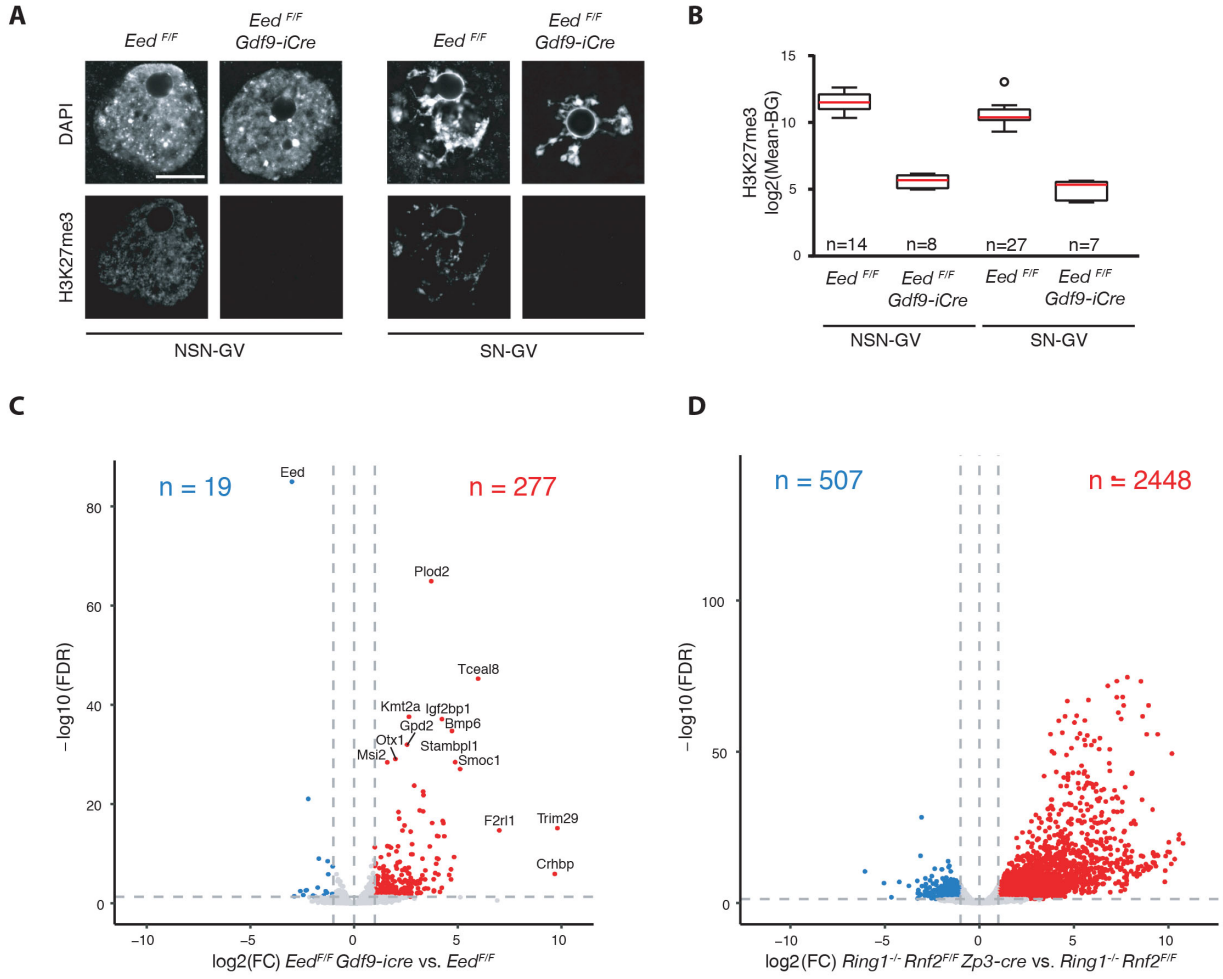
2.5.13 Analysis of ChIP-seq data

Published H3K27me3 ChIP-seq dataset [217] for pre-implantation embryo was downloaded from NCBI GEO repository (GSE73952). Illumina adapters were removed using *trim_galore* tool with parameters *--illumina --paired --stringency 5*. After removal adapters reads were aligned to *M.musculus* genome assembly (mm10) using STAR with parameters *-alignIntronMin 1 -alignIntronMax 1 -alignEndsType EndToEnd -alignMatesGapMax 1000 -outFilterMatchNminOverLread 0.85 -outFilterMultimapNmax 300 -outMultimapperOrder Random -outSAMmultNmax 1.* Reads for ChIP and Input samples were counted for promoters (±2.5kb around TSS) using QuasR package selecting only uniquely mapped reads and enrichments were calculated by subtracting corresponding input log2(RPKM) from ChIP log2(RPKM). Classification of promoters into "Low", "Medium" and "High" classes with respect to H3K27me3 enrichments for each stage was done by fitting Gaussian mixture model with 3 components using *mixtools* R package. Promoters were classified as "Low H3K27me3" if posterior probability for belonging to the lowest Gaussian component was above 0.1, and the rest were classified as "Medium H3K27me3" or "High H3K27me3" choosing class with the highest posterior probability. For analysis of maternal/paternal allelic bias in H3K27me3 enrichments published H3K27me3 dataset [218] for hybrid embryos (C57BL/6N females x PWK/PhJ males) was downloaded from NCBI GEO repository (GSE76687). SNPs for C57BL/6NJ and PWK/PhJ strains were obtained from The Mouse Genomes Project [369] and corresponding genomes were generated by injecting SNPs in reference genome. Alignment of reads to each genome was done using STAR as previously described and assignment of alleles was done identical to previously described analysis of allelic RNA-Seq data. Read counts for each allele was performed using QuasR package for promoters (±5kb around TSS). Promoters which have at least 10 allelic reads and log2(CPM) value higher than 5.5 in each replicate were selected for the analysis. Statistical analysis was performed using edgeR package, similarly to previously described analysis for allelic RNA-seq data.

2.5.14 Analysis of bulk RNA-seq datasets

Previously published RNA-seq dataset for AG/GG morula embryos was downloaded from NCBI GEO repository (GSE92605). Bulk RNA-seq datasets for GV oocytes and AG/GG morula embryos were aligned to mm10 genome using STAR with parameters *-outFilterMultimapNmax 300 -outMultimapperOrder Random -outSAMmultNmax 1 -alignIntronMin 20 -alignIntronMax 1000000* -outFilterMismatchNmax 5, and unique mappers were counted for each gene using QuasR package. Genes with at least 1 read per million in at least 2 samples were included in the analysis. Analysis of differential expression of genes was performed using edgeR package using log-likelihood tests and Benjamini-Hochberg correction for multiple testing.

2.6 Figures



E

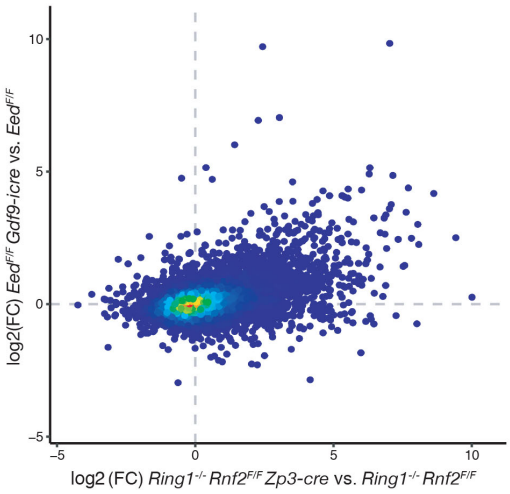


Figure 2.1. Deletion of *Eed* in oocytes leads to a loss of H3K27me3 and minor transcriptional misregulation. **(A)** Representative images of immunofluorescence (IF) staining for H3K27me3 in NSN-GV and SN-GV oocytes from *Eed* $^{F/F}$ and *Eed* $^{F/F}$ *Gdf9-icre* females. Upper panels show DNA stained with DAPI, lower panels show staining for H3K27me3. Scale bar 20µM. **(B)** Quantification of IF staining intensity for H3K27me3 in NSN-GV and SN-GV oocytes from *Eed* $^{F/F}$ and *Eed* $^{F/F}$ *Gdf9-icre* females. Quantification was done by measuring the mean intensity for H3K27me3 in the nucleus (nuclear area determined according to DAPI staining) and normalized by subtracting the background signal (measured in an area outside the nucleus). Normalized relative intensity values are represented as Tukey boxplots on a logarithmic scale for NSN-GV and SN-GV oocytes, separately. Red lines represent the median, black boxes include the 25th to 75th percentiles. Whiskers show the non-outlier range of the data. Circles represent outliers, which are defined as values that are more than 1.5 times the box height above or below the box. The number of analyzed oocytes per genotype is indicated in the graph. **(C)** Volcano plot showing differentially expressed genes between GV oocytes from *Eed* $^{F/F}$ *Gdf9-icre* and *Eed* $^{F/F}$ females. The x-axis corresponds to the level of misregulation (log2(FC)) and the y-axis displays significance (-log10(FDR)). Red dots represent significantly upregulated and blue dots significantly downregulated transcripts in GV oocytes from *Eed* $^{F/F}$ *Gdf9-icre* females (max. FDR=0.05, min. log2(FC)=1). The numbers of significantly up- and downregulated genes are indicated on the graph. **(D)** Volcano plot showing differentially expressed genes between GV oocytes from *Ring1* $^{-/-}$ *Rnf2* $^{F/F}$ *Zp3-cre* and *Ring1* $^{-/-}$ *Rnf2* $^{F/F}$ females. The x-axis corresponds to the level of misregulation (log2(FC)) and the y-axis displays significance (-log10(FDR)). Red dots represent significantly upregulated and blue dots significantly downregulated transcripts in GV oocytes from *Ring1* $^{-/-}$ *Rnf2* $^{F/F}$ *Zp3-cre* females (max. FDR=0.05, min. log2(FC)=1). The numbers of significantly up- and downregulated genes are indicated on the graph. **(E)** Scatter plot showing correlation between differentially expressed genes in *Eed* deficient GV oocytes and *Ring1/Rnf2* deficient GV oocytes. The x-axis corresponds to log2(FC) in GV oocytes from *Eed* $^{F/F}$ *Gdf9-icre* vs. *Eed* $^{F/F}$ females and y-axis displays log2(FC) in GV oocytes from *Ring1* $^{-/-}$ *Rnf2* $^{F/F}$ *Zp3-cre* vs. *Ring1* $^{-/-}$ *Rnf2* $^{F/F}$ females. Colors correspond to density of points.

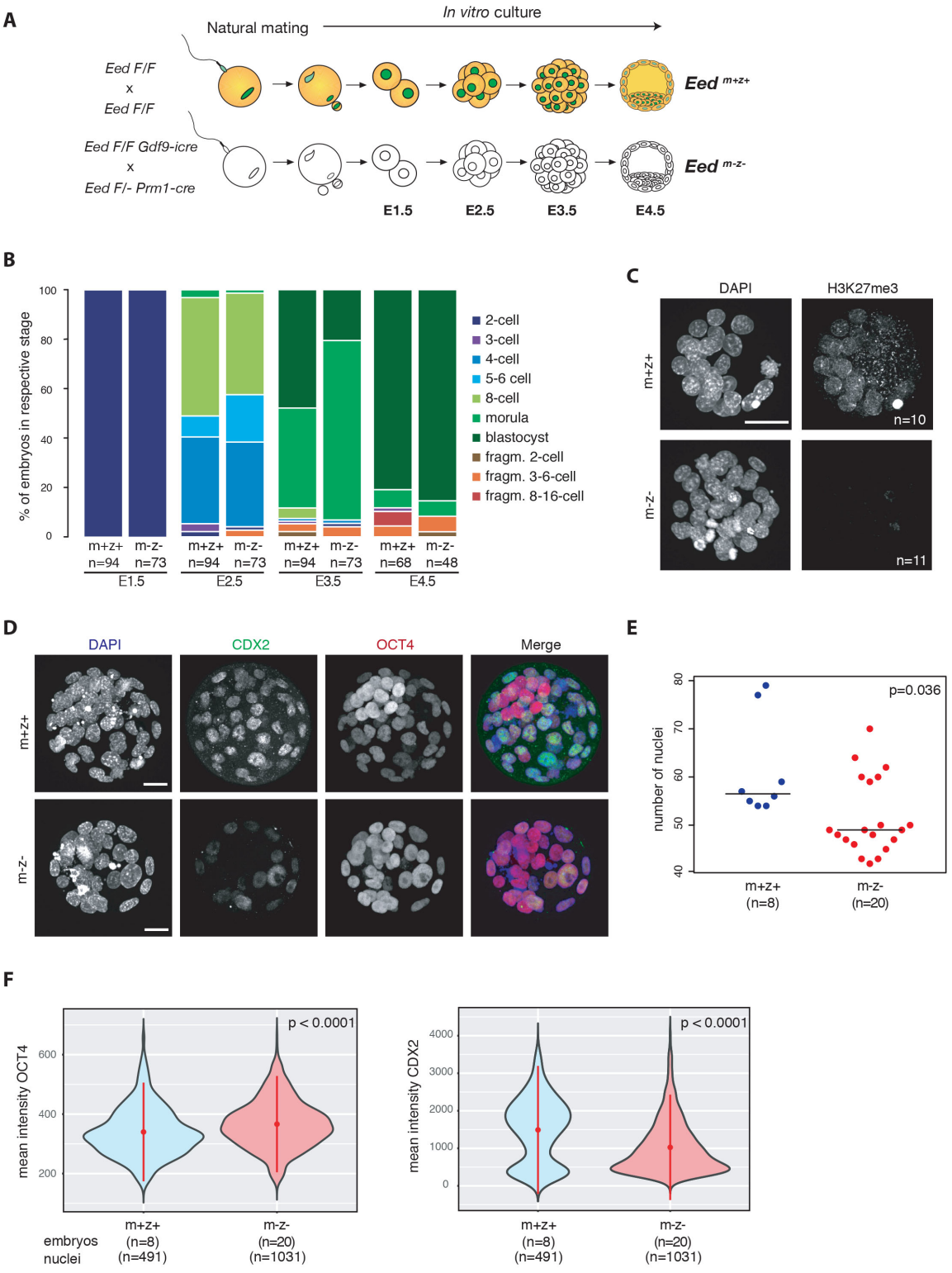


Figure 2.2. Embryos maternally and zygotically deficient for *Eed* are delayed in pre-implantation development. (A) Schematic overview of the generation and *in vitro* culture of *Eed* $^{m+z+}$ and *Eed* $^{m-z-}$ embryos. *Eed* $^{m+z+}$ embryos were generated by natural matings of *Eed* $^{F/F}$ females and males, *Eed* $^{m-z-}$ embryos by mating of *Eed* $^{F/F}$ *Gdf9-icre* females with *Eed* $^{F/F}$ *Prm1-cre* males. Embryos were cultured *in vitro* until the late blastocyst stage (E4.5). Yellow color in the cytoplasm indicates activity of EED protein, green color in the nucleus depicts

levels of H3K27me3. **(B)** Comparative time course developmental progression analysis of *Eed* $^{m+z+}$ and *Eed* $^{m-z-}$ embryos**.** Developmental stages were scored daily for 4 days from E1.5 to E4.5. The number of analyzed embryos is indicated on the graph. Data pooled from 3 independent experiments, including the Smart-seq2 experiment, resulting in a decrease in the number of embryos at E4.5. **(C)** Immunofluorescence (IF) staining for H3K27me3 in *Eed* $^{m+z+}$ and *Eed* $^{m-z-}$ E3.5 blastocyst embryos. Left panels show DNA stained with DAPI and right panels show staining for H3K27me3. Numbers of embryos analyzed per genotype are indicated on the respective panels. Scale bar 20µM. **(D)** Representative IF images of *Eed* $^{m+z+}$ and *Eed* $^{m-z-}$ E4.5 blastocyst embryos stained for DNA (DAPI), CDX2 and OCT4. Right panels show merged images of the 3 channels. Scale bars 20µM. **(E)** Beeswarm plot showing quantification of the number of nuclei in *Eed* $^{m+z+}$ and *Eed* $^{m-z-}$ E4.5 embryos resulting from automated image segmentation. Each dot represents one embryo and black lines represent the median for each group. The number of embryos analyzed per group is indicated on the plot. Significance was calculated by two-tailed t-test ($p=0.036$). **(F)** Violin plots showing mean signal intensity for OCT4 (left panel) and CDX2 (right panel) in automatically segmented nuclei of *Eed* $^{m+z+}$ and *Eed* $^{m-z-}$ E4.5 embryos. Red points represent the mean for each group and red lines standard deviation. The number of embryos and nuclei analyzed per group is indicated on the plot. Significance was calculated by two-tailed t-test ($p<0.0001$ for both panels).

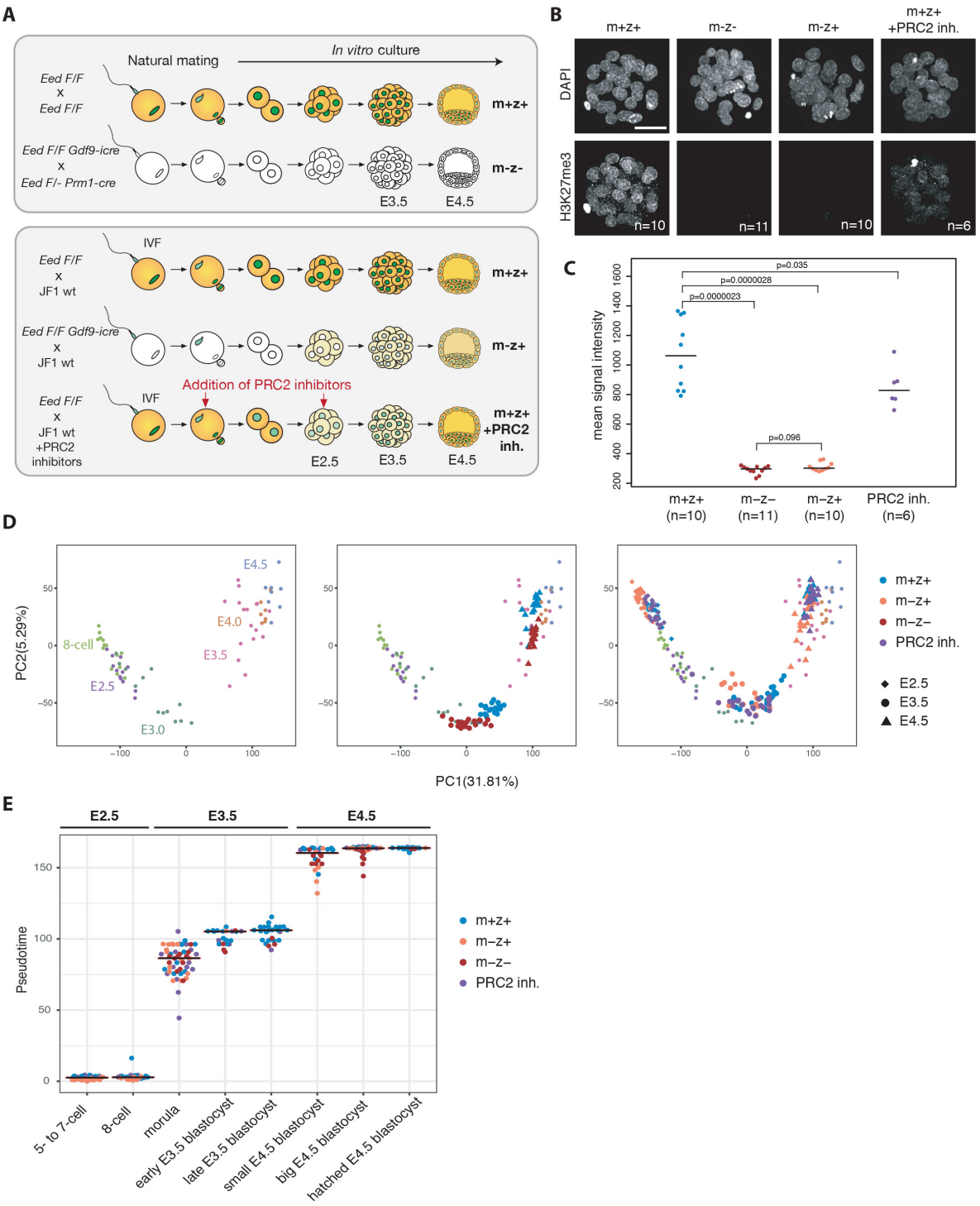


Figure 2.3. Single embryo RNA sequencing confirms delay of *Eed* $^{m-z-}$ and *Eed* $^{m-z+}$ embryos. (A) Schematic illustration depicting the conditions analyzed in the two Smart-seq2 experiments. Experiment 1 (upper panel) includes *Eed* $^{m+z+}$ and *Eed* $^{m-z-}$ embryos that were generated by natural mating of *Eed* $^{F/F}$ or *Eed* $^{F/F}$ *Gdf9-icre* females with *Eed* $^{F/F}$ or *Eed* $^{F/-}$ *Prm1-cre* males, respectively. Embryos were cultured *in vitro* and collected for RNA sequencing at E3.5 and E4.5. Experiment 2 (lower panel) includes *Eed* $^{m+z+}$ and *Eed* $^{m-z+}$ hybrid embryos generated by in vitro fertilization (IVF) of oocytes from *Eed* $^{F/F}$ or *Eed* $^{F/F}$ *Gdf9-icre* females with sperm from JF1 wild type males. Part of the *Eed* $^{m+z+}$ embryos were treated with a combination of 2 PRC2 inhibitors. Fresh inhibitors were added after fertilization and at E2.5. Embryos were harvested at E2.5, E3.5 and E4.5. Yellow color in the cytoplasm indicates activity of EED, green color in the nucleus depicts levels of H3K27me3. **(B)**

Immunofluorescence staining for H3K27me3 in *Eed* $^{m+z+}$, *Eed* $^{m-z-}$, *Eed* $^{m-z+}$ and PRC2 inhibitor treated *Eed* $^{m+z+}$ E3.5 blastocyst embryos. Upper panels show DNA stained with DAPI and lower panels show staining for H3K27me3. Numbers of embryos analyzed per genotype are indicated on the respective panels. Scale bar 20μM. **(C)** Beeswarm plot showing quantification of H3K27me3 signal in IF images of *Eed* $^{m+z+}$, *Eed* $^{m-z-}$, *Eed* $^{m-z+}$ and PRC2 inhibitor treated *Eed* $^{m+z+}$ E3.5 blastocyst embryos. Quantification was done by measuring the mean intensity for H3K27me3 in a Z-projection of the entire embryo. Numbers of embryos analyzed per genotype are indicated on the graph. Black lines represent the median for each group. Significance was calculated by two-tailed t-test (p-values are indicated on the graph). **(D)** Principal component analysis on relative expression including 8-cell to E4.5 wild type embryos from an external dataset [333] and embryos of both Smart-seq2 experiments. Analysis was performed including all samples, for visualization experiments were plotted separately. **(E)** Beeswarm plot displaying the estimated pseudotime for embryos of different morphological stages. Black lines represent the median for each morphology group.

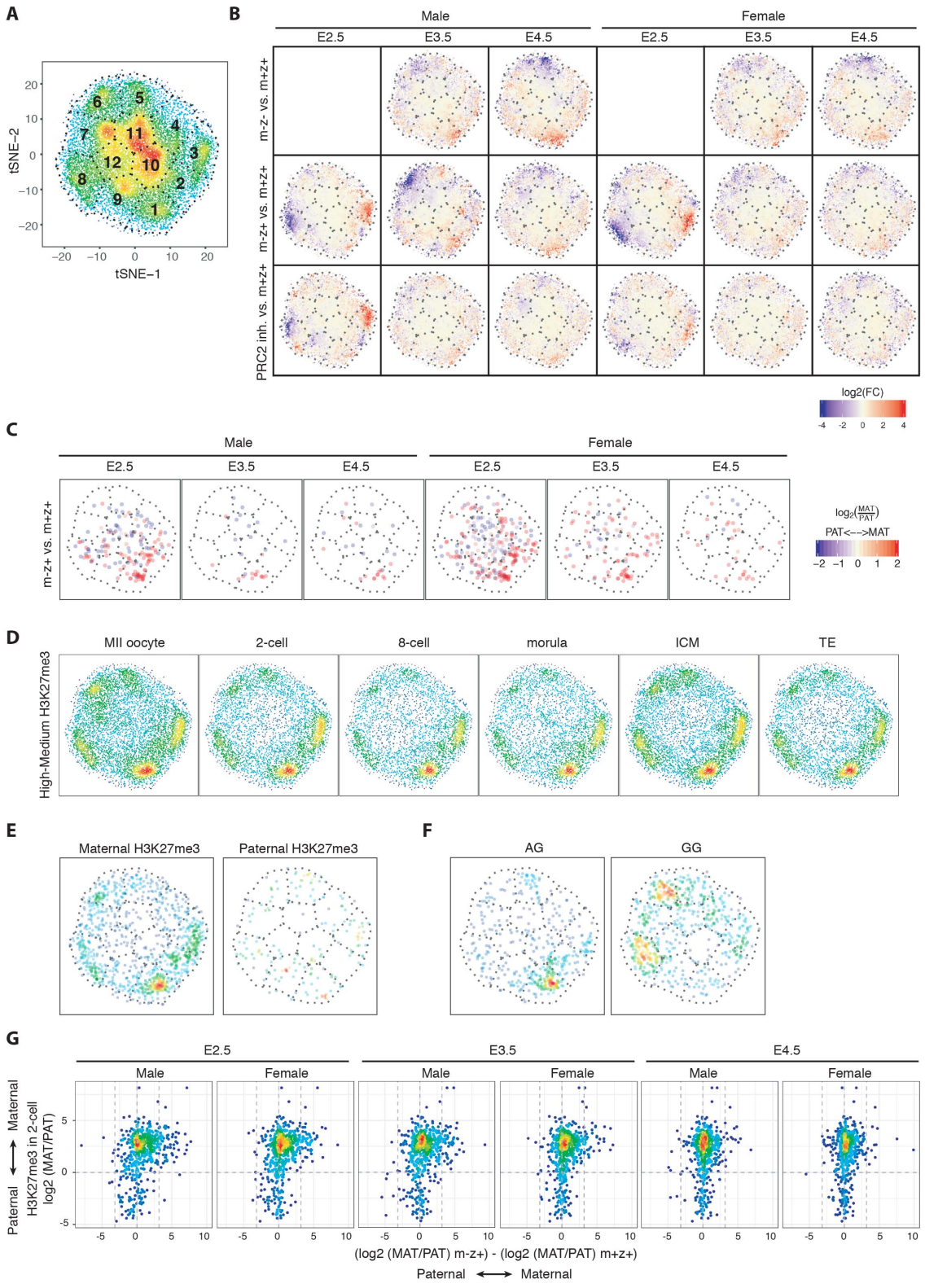
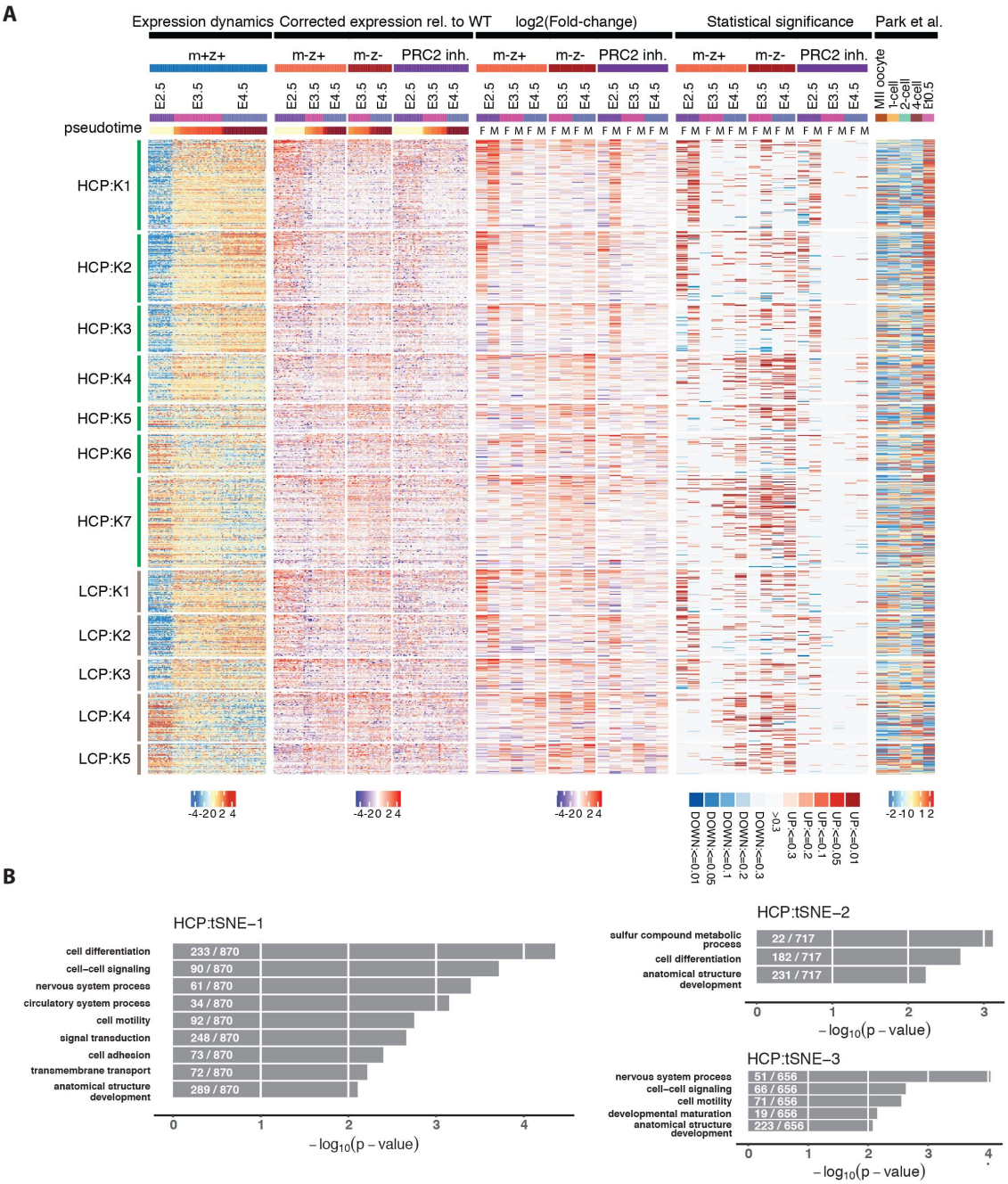


Figure 2.4. Identification of maternal PRC2 targets in pre-implantation embryos. (A) t-SNE plot for all expressed genes. Each point represents a gene and colors represent gene density in t-SNE space. Dotted lines represent partitioning of t-SNE space in k clusters, numbers correspond to cluster IDs. The outer border represents the area which encompasses 99% of all genes in the t-SNE space. **(B)** t-SNE representation of the effect of *Eed* $^{m-z-}$, *Eed* $^{m-z+}$ or PRC2 inhibition on autosomal gene expression in male and female embryos at E2.5, E3.5 and E4.5.

Each point represents a gene and colors represent log2(FC) in a corresponding comparison. **(C)** Change in allelic bias of expression upon maternal deletion of *Eed* for autosomal genes in male and female embryos at E2.5, E3.5 and E4.5. Each point represents a gene and colors represent log2(maternal/paternal expression). **(D)** t-SNE representation of densities for high and medium H3K27me3 promoters of autosomal genes as defined in Figure S4B [217]. Each point represents a gene and colors represent density in t-SNE space. **(E)** t-SNE representation of densities for autosomal genes with a significant allelic bias towards maternal or paternal H3K27me3 (FDR≤0.05 and abs(log2(FC)) ≥2) [218]. Each point represents a gene and colors represent density in t-SNE space. **(F)** t-SNE representation of densities of autosomal genes which are significantly differentially expressed in androgenetic (AG) or gynogenetic (GG) embryos (FDR≤0.05 and abs(log2(FC)) ≥2) [234]. Each point represents a gene and colors represent density in t-SNE space. **(G)** Scatter plots showing correlation between change in allelic expression upon maternal deletion of *Eed* in male and female embryos at E2.5, E3.5 and E4.5 (x-axis) and allelic H3K27me3 in 2-cell embryos (y-axis). Each point represents a gene and colors represent density.



	t-SNE 1	t-SNE 2	t-SNE 3
Cell differentiation	Bmp1, Notch2, Notch3, Pitx2, Pou3f1, Sall1, Sox1, Sox15, Sox6, Sox21, Stat6	Fgf20, Fzd2, Fzd8, Fzd9, Mapk12, Mov10l1, Nanog, Otx2, Pitx3, Pou3f2, Rnf17, Stat5a, Stat5b, Sycp1, Sycp3, Tfcp2l1, Wnt5b	
Cell-cell signaling	Fgf8, Fgf12, Fgfr3, Fzd1, Fzd10, Wnt3, Wnt8a		Bmp4, Lgr4, Rspo1, Wnt1, Wnt4, Wnt9a
Anatomical structure development	Bmi1, Dppa3, Foxl2, Hormad1, Hoxa1, Hoxa5, Hoxb6, Hoxb7, Hoxb13, Hoxd10, Hoxd13, Otx1, Pax6	Bmp7, Bmp8a, Hoxa11, Hoxa9, Hoxb1, Hoxc8, Hoxc12, Irak3, Pax9	Esrrg, Fos, Foxa3, Hoxa10, Hoxd3, Klf7, Pcgf2, Pdgfc, Pou4f2, Prdm14, Sall3, Socs5, Sox18

Figure 2.5. PRC2 regulates developmental genes with different dynamics in pre-implantation embryos. (A) Heatmap illustrating expression patterns and effect of each condition at different stages for autosomal genes from t-SNE clusters 1,2 and 3 that are up- or downregulated upon *Eed* deletion/PRC2 inhibition of (FDR ≤ 10% and abs(log2(FC)) ≥ 1) or show a strong change in allelic bias upon *Eed* deletion/PRC2 inhibition (FDR ≤ 5% and

abs(log2(FC)) ≥ 2). K-means clustering was performed based on expression dynamics in wild type embryos and effect of *Eed* deletion/PRC2 inhibition to identify 7 clusters of genes with high CpG promoters (HCP:K1-7) and 5 clusters of genes with low CpG promoters (LCP:K1-5). “Expression Dynamics” shows relative expression in *Eed* $^{m+z+}$ embryos. “Corrected expression relative to WT” displays values for each genotype/treatment, which are corrected for the time component and calculated by subtracting mean expression in *Eed* $^{m+z+}$ embryos at each stage. “log2(Fold-Change)” displays log2(fold changes) and “Statistical significance” displays FDRs for each contrast fitted by the model, separate for female (F) and male (M) embryos. “Park et al.” displays relative expression in an external dataset of wild type embryos [370]. Yellow-Red bars correspond to the pseudotime assigned to each embryo. Legends are plotted below each section of the heatmap. **(B)** Bar plots displaying slim GO terms for genes with high CpG promoters (HCP) of t-SNE clusters 1,2 and 3. The x-axis displays significance (-log10(p-value)) and the number of genes in the respective GO category is indicated on the bars. **(C)** Table showing selected genes in t-SNE clusters 1,2 and 3 belonging to GO term categories “Cell differentiation”, “Cell-cell signaling” and “Anatomical structure development”.

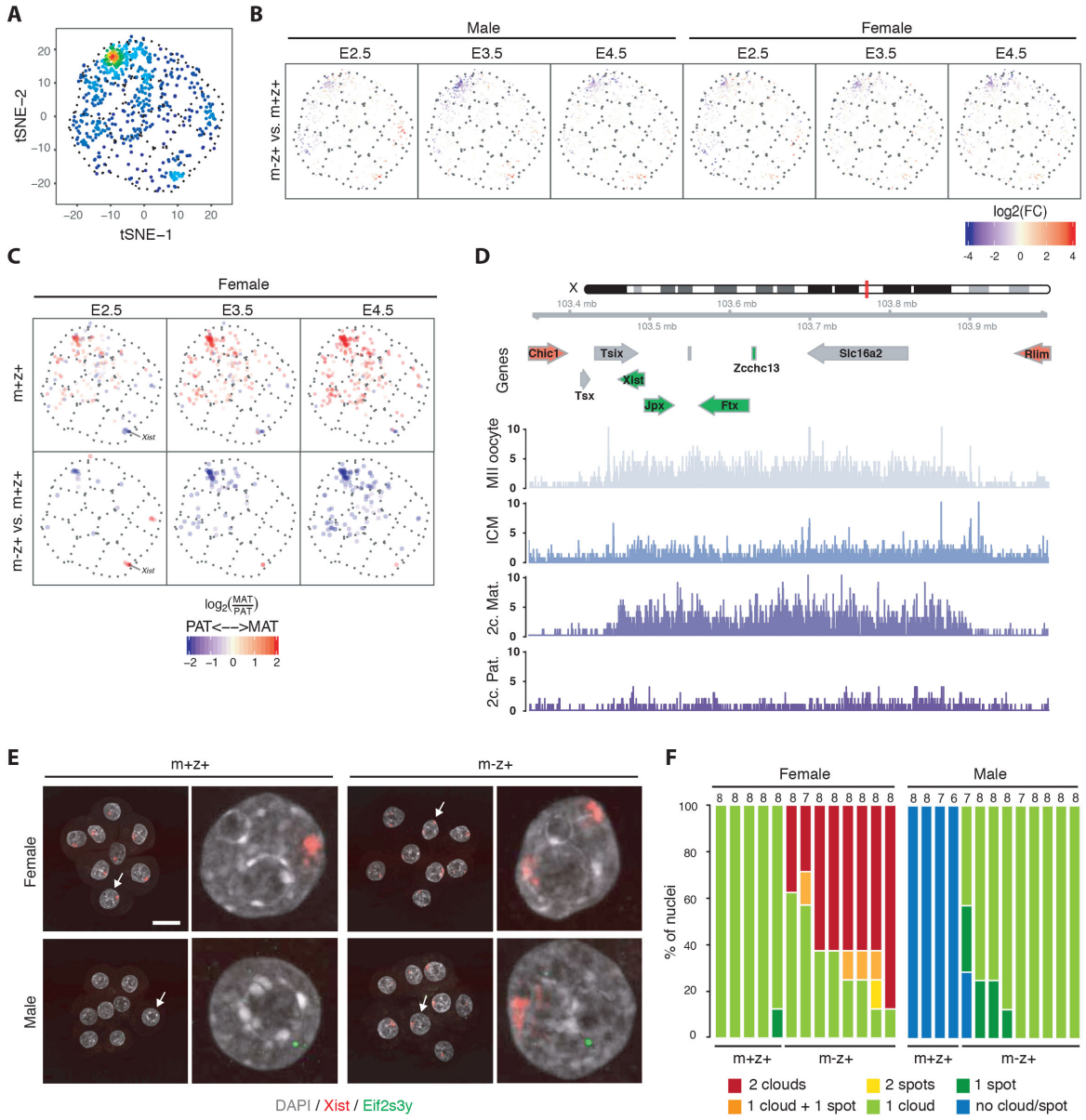


Figure 2.6. H3K27me3 functions as the repressive maternal imprint for *Xist*. (A) t-SNE plot for all X-linked genes. Each point represents a gene and colors represent gene density in t-SNE space. **(B)** t-SNE representation of the effect of maternal *Eed* deletion on X-linked gene expression in male and female embryos at E2.5, E3.5 and E4.5. Each point represents a gene and colors represent log2(FC). **(C)** Allelic bias of expression in *Eed* $^{m+z+}$ (upper panels) and change in allelic bias of expression upon *Eed* maternal deletion (lower panels) for X-linked genes in female embryos at E2.5, E3.5 and E4.5. Each point represents a gene and colors represent log2(maternal/paternal expression). **(D)** Genome browser view of H3K27me3 ChIP-seq signal at the *Xist* locus in MII oocytes, ICM of the blastocyst and 2-cell embryos (maternal and paternal alleles). Genes upregulated upon *Eed* deletion are indicated in green, while downregulated genes are marked in red. **(E)** Representative images of *Xist* RNA FISH in *Eed* $^{m+z+}$ and *Eed* $^{m-z+}$ male and female embryos at E2.5. The sex of each embryo was assessed by including an RNA FISH probe against the Y-linked gene *Eif2s3y* and DNA was stained with DAPI. White arrows indicate blastomeres that are shown as zoom-ups next to the full image. Scale bar 20µM. **(F)** Quantification of *Xist* RNA FISH in *Eed* $^{m-z+}$ and *Eed* $^{m+z+}$ male and female embryos at E2.5. Each bar represents one embryo, showing the percentage of blastomeres with the indicated number of *Xist* clouds and spots. The number of blastomeres per embryo is indicated above each bar.

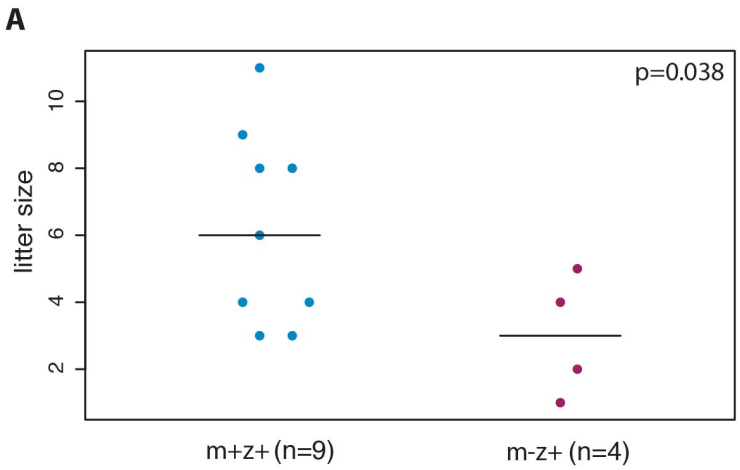
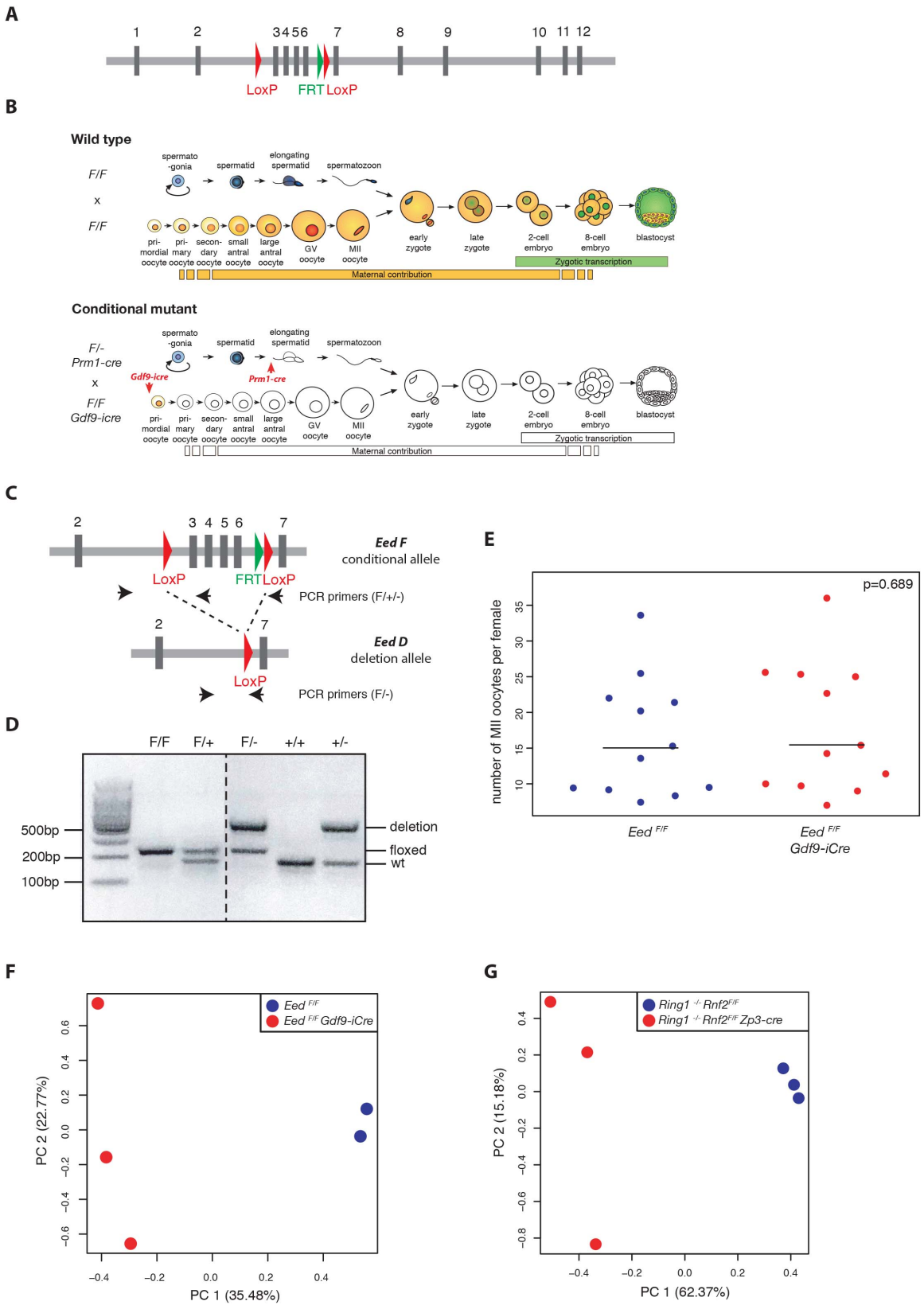
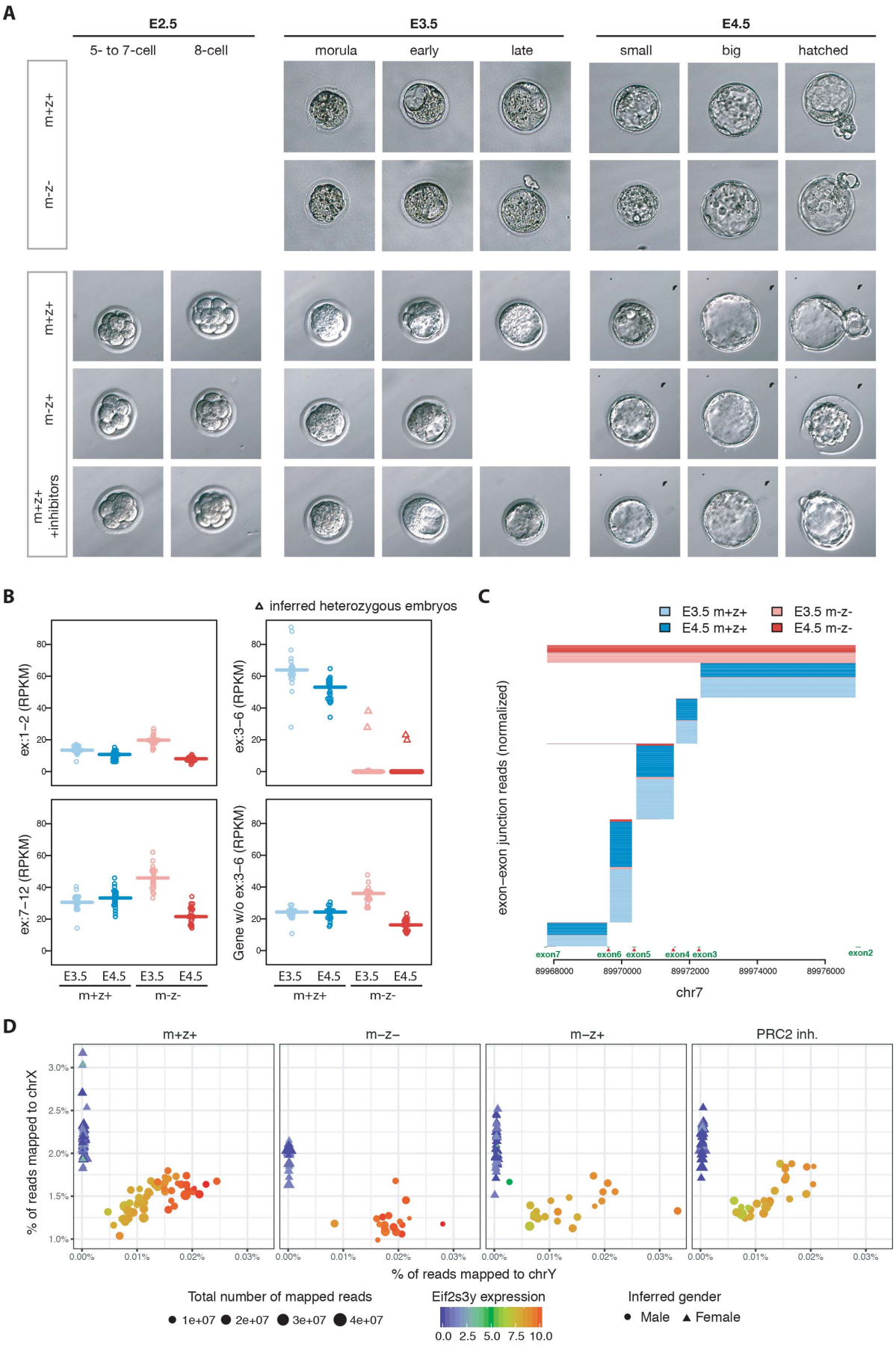


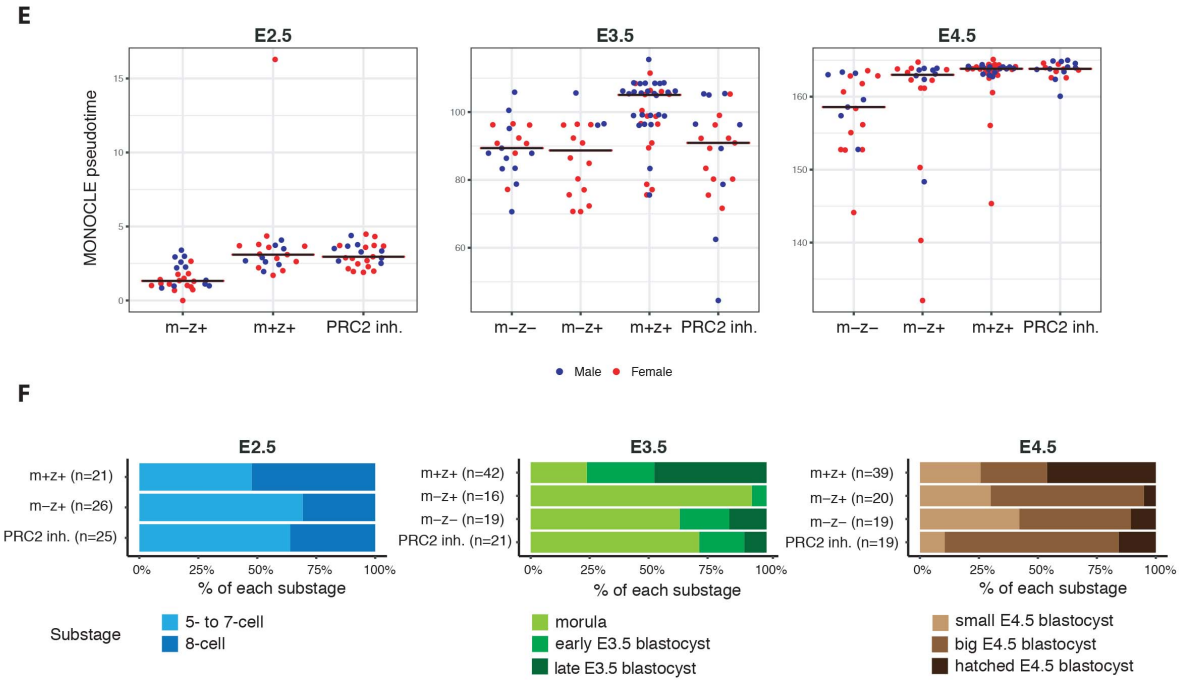
Figure 2.7. Maternal PRC2 is required for proper post-implantation development. (A) Beeswarm plot displaying litter sizes of *Eed* $^{m+z+}$ and *Eed* $^{m-z+}$ litters. Number of litters for each condition are indicated on the graph. Black lines represent the median for each group. Significance was calculated by two-tailed t-test (p=0.038).



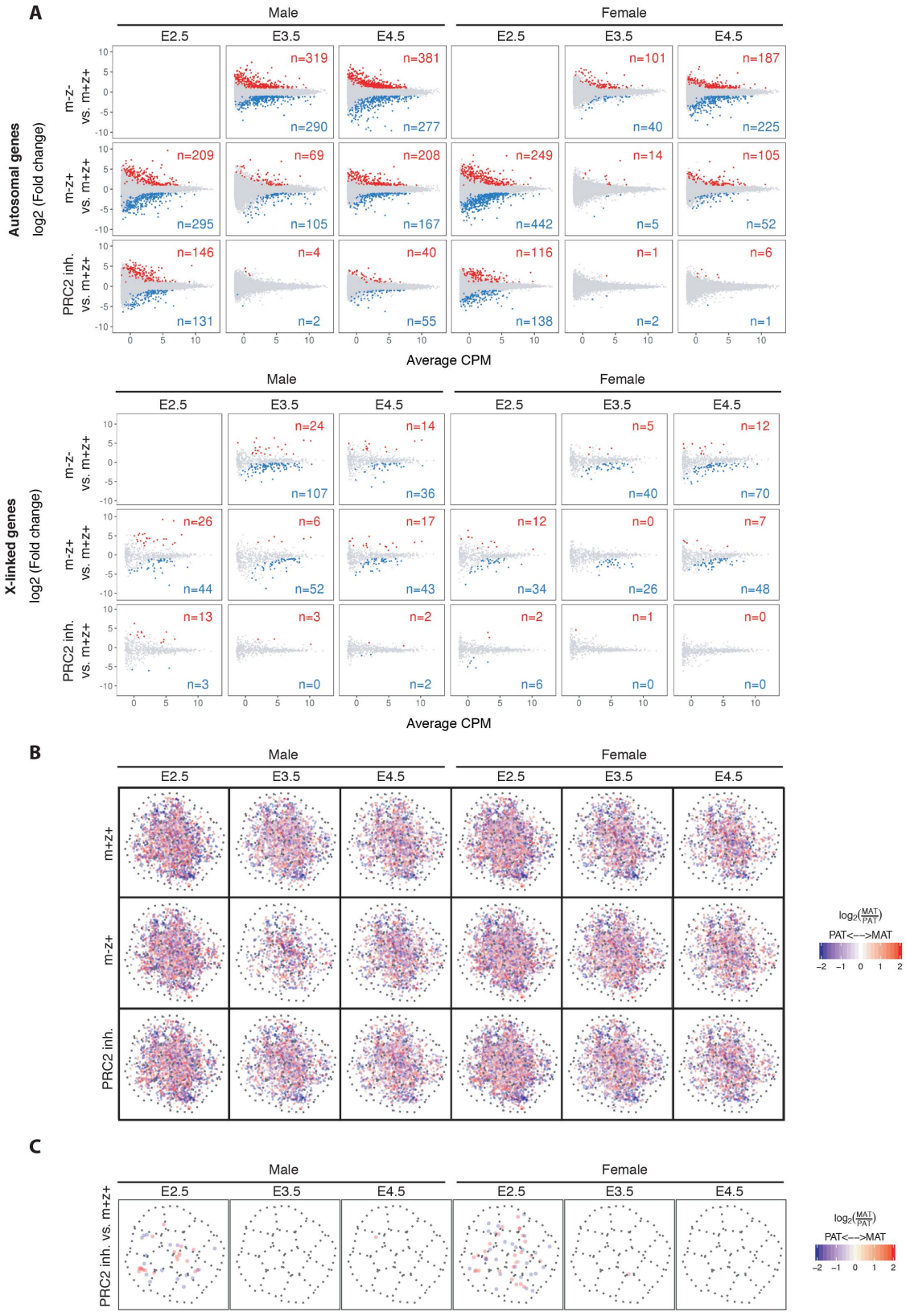
Supplementary Figure 2.1. Characterization of *Eed* deficient oocytes. (A) Schematic diagram of the *Eed* genomic locus. Grey boxes and numbers refer to exons. The green triangle represents an FRT site and red triangles represent LoxP sites flanking exons 3 to 6 in the conditional floxed *Eed* allele. **(B)** Schematic representation illustrating different stages of gametogenesis and early embryonic development and conditional genetic strategy

to generate embryos maternally and zygotically deficient for Eed function (*Eed* $^{m-z-}$). A conditional deletion in the female germline by Gdf9-icre results in *Eed* deficient oocytes, which were fertilized with *Eed* deficient sperm (deletion by Prm1-cre) to generate *Eed* $^{m-z-}$ embryos. Control (*Eed* $^{m+z+}$) embryos were generated by crossing females and males carrying floxed alleles of *Eed*, but no cre. **(C)** PCR genotyping strategy to identify wild type, floxed and deletion *Eed* alleles. **(D)** Agarose gel showing the results of PCR genotyping for *Eed* wild type, floxed and deletion alleles. **(E)** Beeswarm plot showing the number of MII oocytes isolated from super ovulated *Eed* $^{F/F}$ and *Eed* $^{F/F}$ *Gdf9-icre* females. Each dot represents the average number of oocytes per female from a group of 2-11 females. Black lines represent the median of each genotype. Significance was calculated by two-tailed t-test (p=0.689). **(F)** Principal component analysis including GV oocytes from *Eed* $^{F/F}$ and *Eed* $^{F/F}$ *Gdf9-icre* females. **(G)** Principal component analysis including GV oocytes from *Ring1* $^{-/-}$ *Rnf2* $^{F/F}$ and *Ring1* $^{-/-}$ *Rnf2* $^{F/F}$ *Zp3-cre* females.


A
E2.5
E3.5
E4.5
5- to 7-cell
8-cell
morula
early
late
small
big
hatched
m+z+
m-z-
m+z+
m-z+
m+z+ +inhibitors
B
inferred heterozygous embryos
ex:1–2 (RPKM)
ex:3–6 (RPKM)
ex:7–12 (RPKM)
Gene w/o ex:3–6 (RPKM)
E3.5 E4.5 E3.5 E4.5
m+z+ m-z-
C
E3.5 m+z+
E3.5 m-z-
E4.5 m+z+
E4.5 m-z-
exon–exon junction reads (normalized)
exon7 exon6 exon5 exon4 exon3 exon2
89968000 89970000 89972000 89974000 89976000
chr7
D
m+z+
m–z–
m–z+
PRC2 inh.
% of reads mapped to chrX
% of reads mapped to chrY
Total number of mapped reads
1e+07 2e+07 3e+07 4e+07
Eif2s3y expression
0.0 2.5 5.0 7.5 10.0
Inferred gender
Male Female

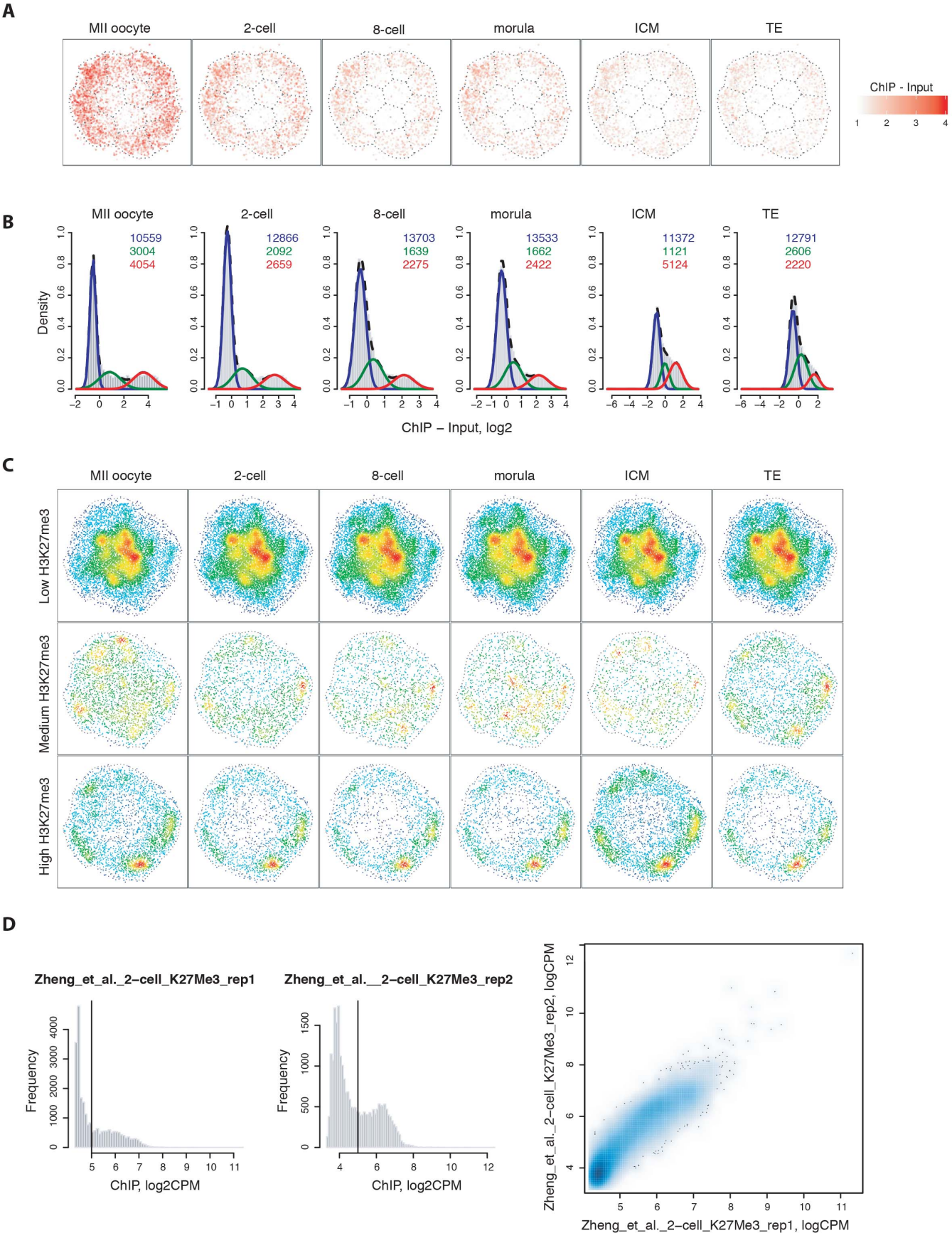


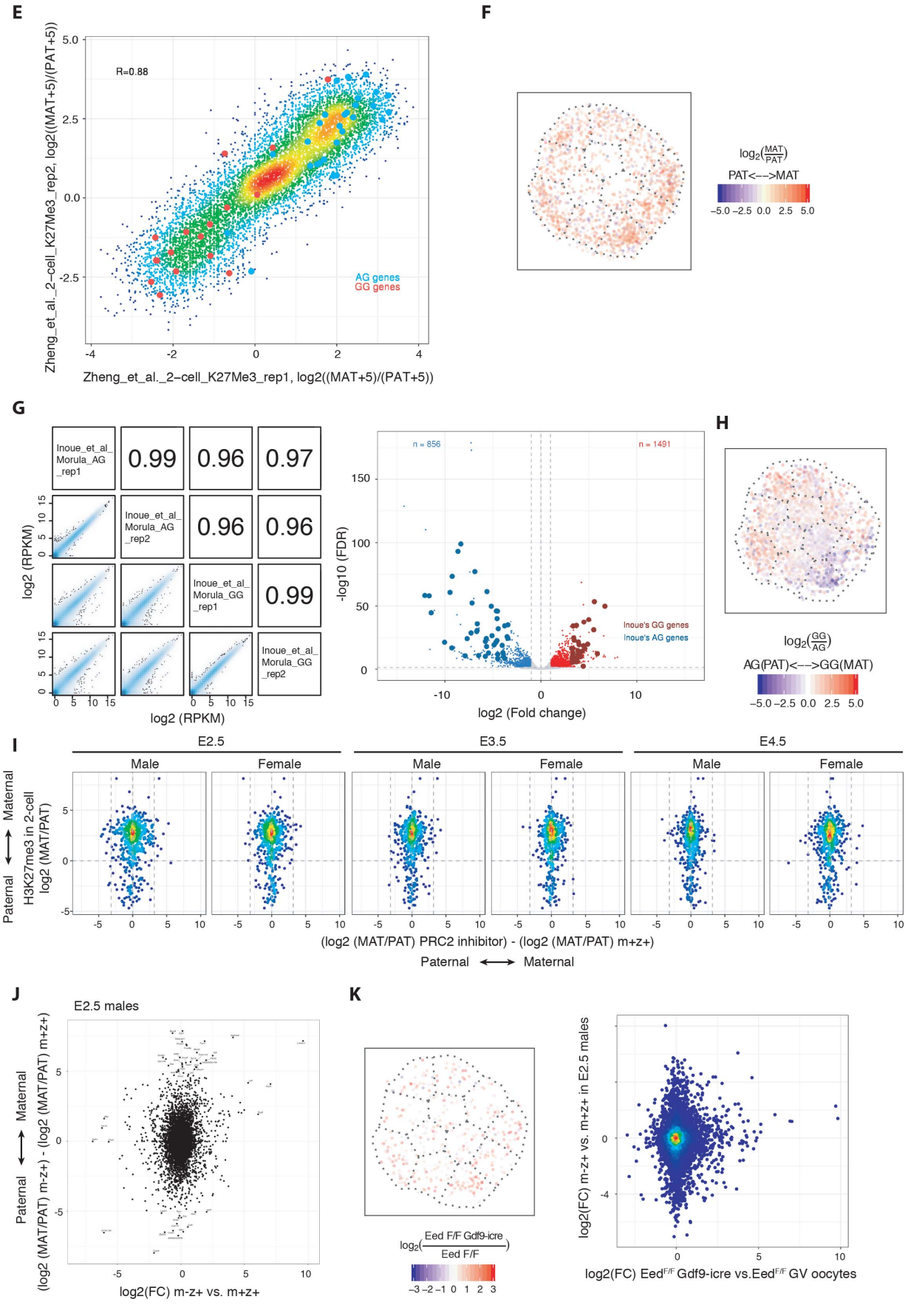
Supplementary Figure 2.2. Genotyping, sexing and pseudo-timing of embryos from RNA-seq data. (A) Representative images of *Eed* $^{m+z+}$, *Eed* $^{m-z-}$, *Eed* $^{m-z+}$ and PRC2 inhibitor treated *Eed* $^{m+z+}$ embryos analyzed in the two Smart-seq2 experiments. Embryos were sub-staged based on morphological criteria in 5- to 7-cell and 8-cell embryos (E2.5), morula, early E3.5 blastocyst and late E3.5 blastocyst (E3.5) and small, big and hatched blastocyst (E4.5). **(B)** Beeswarm plots showing number of reads (RPKM) for different exons of *Eed* in *Eed* $^{m+z+}$ and *Eed* $^{m-z-}$ embryos at E3.5 and E4.5. Each dot represents one embryo and lines represent the median for each group. Inferred heterozygous embryos are represented as triangles. **(C)** Exon junction reads for *Eed* in *Eed* $^{m+z+}$ and *Eed* $^{m-z-}$ embryos at E3.5 and E4.5. **(D)** Scatter plots showing the percentage of reads mapped to chromosome Y (X-axis) and chromosome X (Y-axis) in *Eed* $^{m+z+}$, *Eed* $^{m-z-}$, *Eed* $^{m-z+}$ and inhibitor treated *Eed* $^{m+z+}$ embryos. Each plot contains all embryos from the respective genotype pooled from all stages analyzed. Each point represents one embryo, point size corresponds to the total number of mapped reads and colors correspond to expression of *Eif2s3y*. Shapes correspond to the inferred gender of each embryo. **(E)** Beeswarm plots displaying the estimated pseudotime for different genotypes/treatments at E2.5, E3.5 and E4.5. Black lines represent the median for each group. **(F)** Bar plots displaying the fraction of embryos in different morphological states for different genotypes/treatments at E2.5, E3.5 and E4.5.



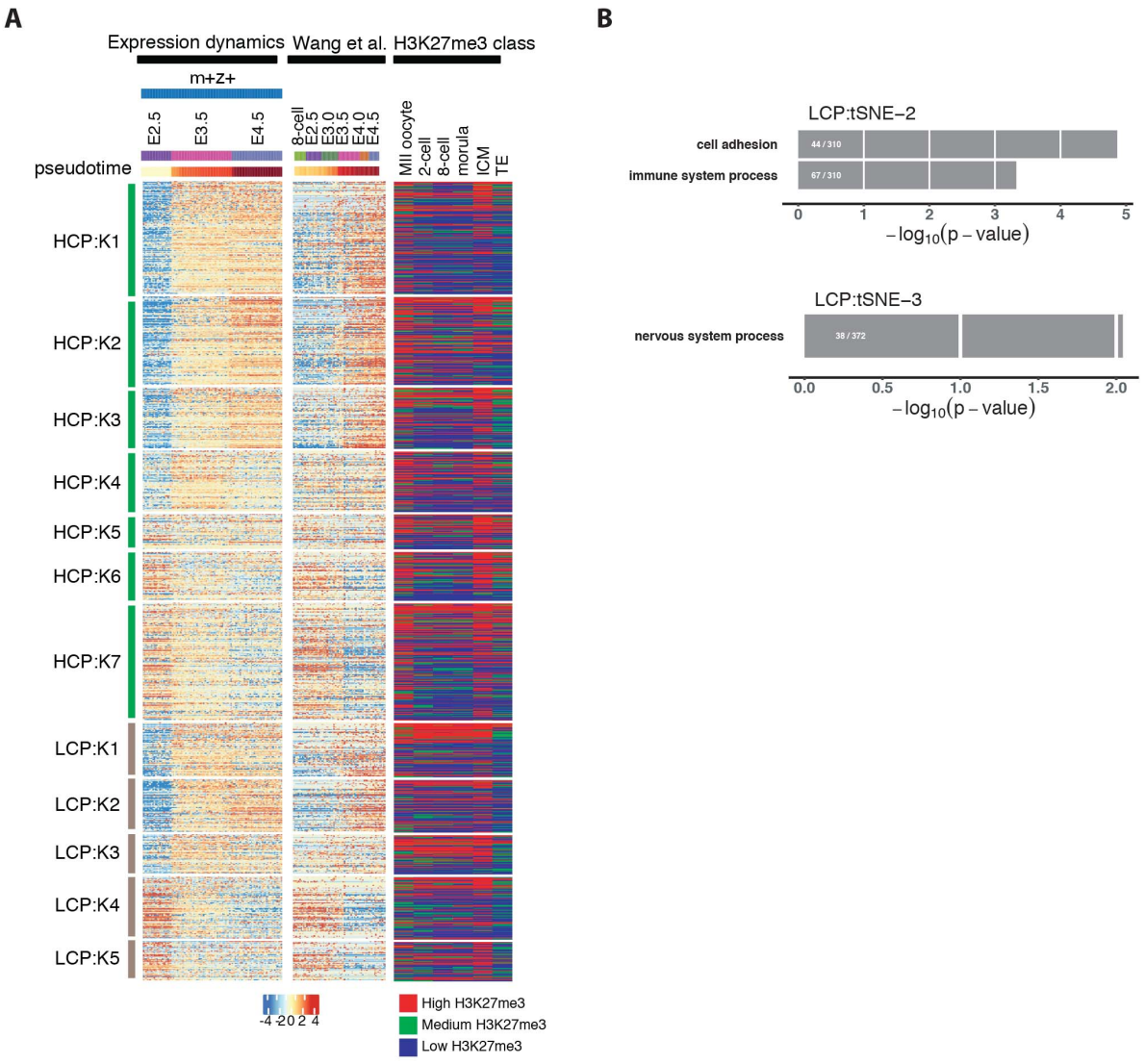
Supplementary Figure 2.3. Differential gene expression and allelic expression analysis (A) MA plots showing differentially expressed genes in *Eed* $^{m\text{-}z\text{-}}$, *Eed* $^{m\text{-}z+}$ and PRC2 inhibitor treated embryos at E2.5, E3.5 and E4.5, separated by embryo sex and split into autosomal (upper panel) and X-linked (lower panel) genes. The x-axis

corresponds to the level of expression (Average CPM) and the y-axis displays the level of misregulation (log2(FC)). Red dots represent significantly upregulated and blue dots significantly downregulated transcripts in the respective contrast (max. FDR=0.05, min. log2(FC)=1). The numbers of significantly up- and downregulated genes are indicated on the graphs. **(B)** t-SNE plot showing allelic bias of expression for autosomal genes in *Eed* $^{m+z+}$, *Eed* $^{m-z+}$ and PRC2 inhibitor treated male and female embryos at E2.5, E3.5 and E4.5. Each point represents a gene and colors represent log2(maternal/paternal expression). **(C)** t-SNE plots showing the change in allelic bias of expression upon PRC2 inhibitor treatment for autosomal genes in male and female embryos at E2.5, E3.5 and E4.5. Each point represents a gene and colors represent log2(maternal/paternal expression).

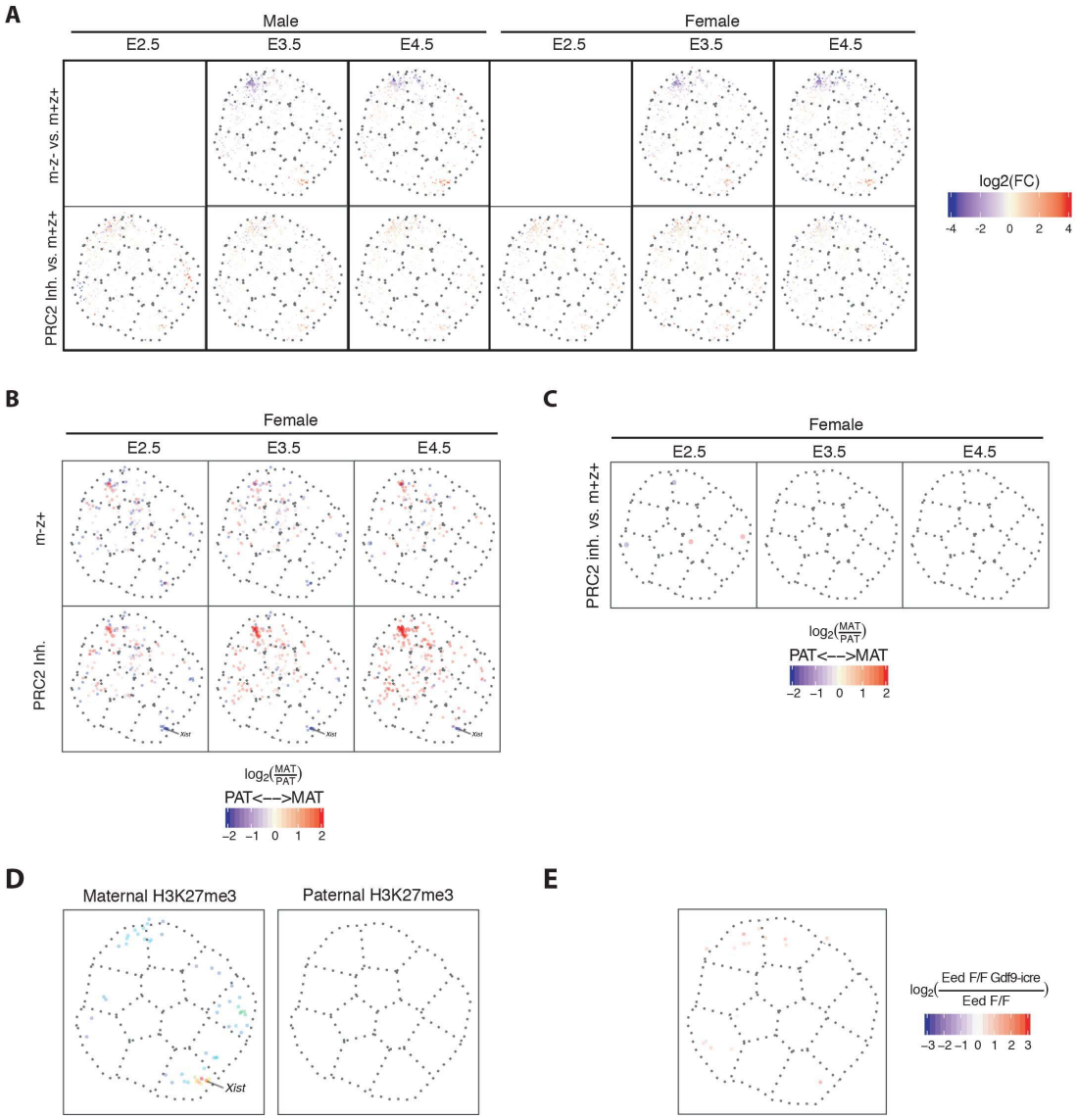

A
MII oocyte
2-cell
8-cell
morula
ICM
TE
ChIP - Input
1 2 3 4
B
MII oocyte
2-cell
8-cell
morula
ICM
TE
10559
3004
4054
12866
2092
2659
13703
1639
2275
13533
1662
2422
11372
1121
5124
12791
2606
2220
Density
ChIP – Input, log2
C
MII oocyte
2-cell
8-cell
morula
ICM
TE
Low H3K27me3
Medium H3K27me3
High H3K27me3
D
Zheng_et_al._2–cell_K27Me3_rep1
Zheng_et_al.__2–cell_K27Me3_rep2
Frequency
ChIP, log2CPM
Zheng_et_al._2–cell_K27Me3_rep2, logCPM
Zheng_et_al._2–cell_K27Me3_rep1, logCPM


E
R=0.88
AG genes
GG genes
Zheng_et_al._2–cell_K27Me3_rep2, log2((MAT+5)/(PAT+5))
Zheng_et_al._2–cell_K27Me3_rep1, log2((MAT+5)/(PAT+5))
F
log2(MAT/PAT)
PAT<--->MAT
G
Inoue_et_al_Morula_AG_rep1
Inoue_et_al_Morula_AG_rep2
Inoue_et_al_Morula_GG_rep1
Inoue_et_al_Morula_GG_rep2
log2 (RPKM)
n = 856
n = 1491
Inoue's GG genes
Inoue's AG genes
-log10 (FDR)
log2 (Fold change)
H
log2(GG/AG)
AG(PAT)<--->GG(MAT)
I
E2.5
E3.5
E4.5
Male
Female
Maternal
Paternal
H3K27me3 in 2-cell log2 (MAT/PAT)
(log2 (MAT/PAT) PRC2 inhibitor) - (log2 (MAT/PAT) m+z+)
Paternal ⟷ Maternal
J
E2.5 males
(log2 (MAT/PAT) m-z+) - (log2 (MAT/PAT) m+z+)
log2(FC) m-z+ vs. m+z+
K
log2(Eed F/F Gdf9-icre / Eed F/F)
log2(FC) m-z+ vs. m+z+ in E2.5 males
log2(FC) Eed F/F Gdf9-icre vs.Eed F/F GV oocytes

Supplementary Figure 2.4. Integration of H3K27me3 ChIP-seq data and expression in androgenetic and gynogenetic embryos (A) t-SNE representation showing ChIP-seq enrichment of H3K27me3 at promoters in MII oocytes, 2-cell embryos, 8-cell embryos, morula embryos, ICM and TE of blastocyst embryos (TSS +/- 2.5kb). **(B)** Histograms displaying the distribution of H3K27me3 enrichments at promoters and approximation produced by fitting Gaussian mixture model with 3 components for each stage (low, medium and high H3K27me3 promoters). The x-axis displays log2(ChIP-Input) and the y-axis corresponds to density. The number of promoters in each class is displayed on the graphs in red (high H3K27me3), green (medium H3K27me3) and blue (low H3K27me3). **(C)** t-SNE representation of densities for low, medium and high H3K27me3 promoters of autosomal genes for each stage as defined in Figure S4B. Each point represents a gene and colors represent density in t-SNE space. **(A)-(C)** Data from *Liu et al.* (2016) [217]. **(D)** Reproducibility of H3K27me3 ChIP-seq data at 2-cell stage from Zheng *et al.* (2016) [218]. Left and middle panels show histograms of total H3K27me3 enrichments in the two replicates. Cut-off of 5.5 is displayed as a vertical line. Right panel displays a scatter plot showing the correlation between the two replicates at promoters (TSS +/-5kb). **(E)** Scatter plot showing reproducibility of allelic biases of H3K27me3 levels at promoters (+/-5kb around TSS) between the two replicates at 2-cell stage. Colors of small points represent density and big colored points correspond to promoters of androgenote/gynogenote-specific genes from Inoue *et al.* (2017) [234]. **(F)** t-SNE plot representing allelic bias of H3K27me3 at autosomal genes in 2-cell stage embryos. Each point represents a gene and colors correspond to log2 (maternal/paternal H3K27me3). Data from Zheng *et al.* (2016) [218] **(G)** Left panel: Scatter plots showing correlation between replicates of RNA-seq samples in androgenetic (AG) and gynogenetic (GG) morula stage embryos (log2(RPKM)). Right panel: Volcano plot showing differentially expressed genes in AG and GG morula stage embryos. The x-axis corresponds to log2(FC) and the y-axis displays significance (-log10(FDR)). Big points correspond to AG- and GG-specific genes, as defined by Inoue *et al.* (2017) [234]. **(H)** t-SNE plot representing biased expression in AG and GG morula stage embryos. Each point represents a gene and colors correspond to log2(expression in GG/expression in AG). **(I)** Scatter plots showing correlation between change in allelic expression upon PRC2 inhibition in male and female embryos at E2.5, E3.5 and E4.5 (x-axis) and allelic H3K27me3 in 2-cell embryos (y-axis). Each point represents a gene and colors represent density. **(J)** Scatter plot showing correlation between total change in expression (log2(FC), x-axis) and change in allelic expression (log2(MAT/PAT), y-axis) between *Eed* $^{m-z+}$ and *Eed* $^{m+z+}$ male embryos at E2.5. **(K)** Left panel: t-SNE plot showing the effect of *Eed* deletion by *Gdf9-icre* on expression of autosomal genes in GV oocytes. Each point represents a gene and colors represent log2(FC). Only genes with FDR≤0.05 are shown. Right panel: Scatter plot showing the correlation between differentially expressed genes in *Eed* deficient GV oocytes and in male *Eed* $^{m-z+}$ embryos at E2.5. X-axis corresponds to log2(FC) in GV oocytes and y-axis displays log2(FC) in E2.5 male *Eed* $^{m-z+}$ embryos. Color corresponds to density of points.



Supplementary Figure 2.5. Expression dynamics and H3K27me3 classes of PRC2 target genes. (A) Heatmap showing expression dynamics in wild type embryos and H3K27me3 class at different stages for genes in clusters defined in Figure 5A. "Expression Dynamics" shows relative expression in *Eed* $^{m+z+}$ embryos. "Wang et al." displays relative expression in an external dataset of wild type embryos [333]. "H3K27me3 class" shows affiliation with low, medium or high H3K27me3 promoters as defined in Figure S4B. **(B)** Bar plots displaying slim GO terms for genes with low CpG promoters (LCP) of t-SNE clusters 2 and 3. The x-axis displays significance (-log10(p-value)) and the number of genes in the respective GO category is indicated in the bars.



Supplementary Figure 2.6. t-SNE analysis for X-linked genes (A) t-SNE representation of the effect of maternal and zygotic deletion of *Eed* and PRC2 inhibition on X-linked gene expression in male and female embryos at E2.5, E3.5 and E4.5. Each point represents a gene and colors represent log2(FC) in a corresponding comparison. **(B)** Allelic bias of expression for X-linked genes in *Eed* $^{m-z+}$ and PRC2 inhibitor treated female embryos at E2.5, E3.5 and E4.5. Each point represents a gene and colors represent log2(maternal/paternal expression). **(C)** Change in allelic bias of expression upon PRC2 inhibitor treatment for X-linked genes in female embryos at E2.5, E3.5 and E4.5. Each point represents a gene and colors represent log2(maternal/paternal expression). **(D)** t-SNE representation of densities for X-linked genes with a significant allelic bias towards paternal or maternal H3K27me3 (FDR≤0.05 and abs(log2(FC)) ≥2). Each point represents a gene and colors represent density in t-SNE space. **(E)** t-SNE plot showing the effect of *Eed* deletion by *Gdf9-icre* on expression of X-linked genes in GV oocytes. Each point represents a gene and colors represent log2(FC). Only genes with FDR≤0.05 are shown.

Author contributions

J.B and A.P. conceived and designed the experiments. J.B. performed most of the experiments, including IF, RNA FISH, developmental progression and breeding analysis and single-embryo RNA-seq. P.N. performed RNA-seq of GV oocytes and quantification of H3K27me3 IF signal in GV oocytes. E.O and H.R. performed bioinformatic analysis of RNA-seq data and external datasets. H.R. performed the initial steps of the single-embryo RNA-seq analysis including genotyping and sexing of embryos. E.O conceived the pseudotiming method, performed differential gene expression and t-SNE analysis and comparison to published datasets. M.R. developed the workflow for automated quantification in IF images of blastocysts. S.S helped in setting up single-embryo RNA-seq and provided technical support. J.B. wrote the manuscript with input from E.O, M.R. and A.P. A.P supervised the project.

Detailed author contributions

Figure 2.1
A and B: Oocyte collection, IF staining and imaging was done by JB. Quantification was done by PN.
C: Oocyte isolation and RNA-seq was done by PN, data analysis was performed by EO.
D: Oocyte isolation and RNA-seq was done by Nathalie Veron, data analysis was performed by EO.
E: Oocyte isolation and RNA-seq was done by PN (for Eed KO) and Nathalie Veron (for Ring1 Rnf2 DKO). Data analysis was performed by EO.

Figure 2.2
A: Cartoon prepared by JB (adapted from AP).
B: Data collected and analyzed by JB.
C: Embryo collection, IF staining and analysis done by JB.
D: Embryo collection, IF staining and analysis done by JB.
E and F: Embryo collection and IF staining done by JB, workflow for automated image quantification was developed by MR. Data analysis was done by JB with help from MR.

Figure 2.3
A: Cartoon prepared by JB (adapted from AP).
B and C: Embryo collection, IF staining and analysis done by JB.
D and E: Embryo collection and RNA-seq library preparation was done by JB with support from SS, analysis was performed by EO.

Figure 2.4
A – C: Embryo collection and RNA-seq library preparation was done by JB with support from SS, analysis was performed by EO.
D – F: Analysis was performed by EO.
G: Embryo collection and RNA-seq library preparation was done by JB with support from SS, analysis was performed by EO.

Figure 2.5
A and B: Embryo collection and RNA-seq library preparation was done by JB with support from SS, analysis was performed by EO.
C: GO term analysis was done by EO, selection of genes from GO term lists was done by JB and AP.

Figure 2.6
A – C: Embryo collection and RNA-seq library preparation was done by JB with support from SS, analysis was performed by EO.
D: Analysis was performed by EO and HR.
E and F: Embryo collection, RNA FISH, imaging and analysis was done by JB.

Figure 2.7
A: Breeding of mice, data collection and analysis done by JB.

Supp. Figure 2.1
A: Cartoon prepared by JB.
B: Cartoon prepared by JB (adapted from AP).
C: Cartoon prepared by JB.
D: Genotyping was performed by JB.
E: Isolation of oocytes and analysis done by JB.
F: Oocyte isolation and RNA-seq was done by PN, data analysis was performed by EO.
G: Oocyte isolation and RNA-seq was done by Nathalie Veron, data analysis was performed by EO.

Supp. Figure 2.2
A: Embryo collection, imaging and sub-staging was done by JB.
B and C: Embryo collection and RNA-seq library preparation was done by JB with support from SS, analysis was performed by HR.
D - F: Embryo collection and RNA-seq library preparation was done by JB with support from SS, analysis was performed by EO.

Supp. Figure 2.3
A - C: Embryo collection and RNA-seq library preparation was done by JB with support from SS, analysis was performed by EO.

Supp. Figure 2.4
A – H: Analysis was performed by EO.
I and J: Embryo collection and RNA-seq library preparation was done by JB with support from SS, analysis was performed by EO.

K: Oocyte isolation and RNA-seq was done by PN, Embryo collection and RNA-seq library preparation was done by JB with support from SS, analysis was performed by EO.

Supp. Figure 2.5
A and B: Embryo collection and RNA-seq library preparation was done by JB with support from SS, analysis was performed by EO.

Supp. Figure 2.6
A – C: Embryo collection and RNA-seq library preparation was done by JB with support from SS, analysis was performed by EO.
D: Analysis was performed by EO.
E: Oocyte isolation and RNA-seq was done by PN, data analysis was performed by EO.

Chapter 3. Manuscript in preparation: A non-canonical PRC1 complex serves as a backup for DNA methylation to repress germ line genes in pre-implantation embryos

Julia Bammer[1,2], Helene Royo[1,4], Evgeniy Ozonov[1], Peter Nestorov[1,3], Sebastien Smallwood[1] and Antoine Peters[1,5]

[1] Friedrich Miescher Institute for Biomedical Research (FMI), Maulbeerstrasse 66, 4058 Basel, Switzerland
[2] Faculty of Sciences, University of Basel, 4056 Basel, Switzerland
[3] Current address: Witec AG, Industriestrasse 12, 6210 Sursee, Switzerland
[4] Current address: Roche Pharmaceutical Research and Early Development (pRED), F. Hoffmann-La Roche AG, Grenzacherstrasse 124, 4070 Basel, Switzerland
[5] Corresponding author: Email: antoine.peters@fmi.ch, Phone: +41616978761, Fax: +41616973976

Keywords
Polycomb repressive complex 1, variant PRC1, PRC1.6, PCGF6-PRC1

3.1 Abstract

Polycomb group (PcG) proteins function in two main complexes, Polycomb repressive complex 1 and 2 (PRC1 and PRC2). While the canonical model of PcG repression suggests that PRC2 functions upstream of PRC1, it has become clear that there is a diversity of PRC1 complexes and some of them can be targeted to chromatin independently and act upstream of PRC2. It has been shown that zygotic deficiency for several PRC1 components leads to embryonic lethality at early post-implantation stages. Further, maternal deficiency for *Ring1* and *Rnf2* causes an embryonic arrest at the 2-cell stage due to massive transcriptional misregulation in the oocyte. However, much less is known about the function of PRC1 in pre-implantation embryos. We therefore generated embryos maternally deficient for *Ezh2/Eed* and *Rnf2*. We show that these embryos develop to the blastocyst stage, but exhibit gastrulation defects. By comparing the transcriptomic profiles of single PRC1/2 double knock out (DKO) embryos to those of PRC2 knock out (KO) embryos, we show that PRC1 represses a set of genes with functions in germ line development. We further present a comparative analysis to external datasets and provide evidence that germ line genes are repressed by the variant PRC1.6 complex that functions as a backup for DNA methylation in early embryos.

3.2 Introduction

The canonical model of PcG repression suggests that PcG complexes are linked in a hierarchical way, with PRC2 acting upstream of PRC1 and H3K27me3 being necessary for proper targeting of PRC1 to promoters and gene silencing [42, 43, 371]. This model is supported by the notion that H3K27me3 increases the affinity for CBX-containing PRC1 to chromatin [80, 137]. However, in recent years, a number of studies have challenged the generality of this model and suggested alternative modes of PcG targeting and interaction. It has been shown that many variants of PRC1 complexes that differ in their subunit composition exist and a number of them can be targeted to chromatin independently of PRC2 [reviewed in 139]. For example, PRC1.1 has been shown to be targeted to chromatin via binding of KDM2B to unmethylated CpG islands [140]. Moreover, it was proposed that H2AK119ub1, deposited by variant PRC1 complexes, can recruit PRC2 and that the ubiquitin mark can potentiate the catalytic activity of PRC2, suggesting a positive feedback loop between H2AK119ub1 and H3K27me3 [140, 142]. Recruitment of PRC2 to ubiquitinated H2A might depend on the PRC2 subunits AEBP2 and JARID2 that have been shown to bind to H2AK119ub1 *in vitro* [66, 142]. Taken together, these studies suggest that the classical Polycomb recruitment model is not sufficient to explain all Polycomb targeting and the interactions between PRC1 and PRC2 are more diverse than initially thought. However, most of the studies on the interplay between PcG complexes have been done biochemically in *in vitro* systems and much less is known about the relevance of the findings *in vivo.*

The role of the different PRC1 complexes in embryonic development has been studied to some extent. Zygotic deficiency for the core PRC1 component *Rnf2 (Ring1b)* causes embryonic lethality during gastrulation, whereas its paralog *Ring1 (Ring1a)* is not essential for development [40, 171]. Maternal DKO of *Rnf2* and *Ring1* leads to a developmental arrest at the 2-cell stage due to massive transcriptional misregulation in the oocyte [173]. Moreover, several components of variant PRC1 complexes are essential for embryonic development. Deficiency for *Rybp* causes lethality during early post-implantation stages [174]. The C-terminal domains of *Kdm2b* have also shown to be essential for post-implantation development [140, 176]. In contrast, deletion of single *Cbx* genes has no effect on embryonic development [38, 177, 178].

In the present study, we aim to investigate the role of PRC1 and PRC2 for mouse pre-implantation development by using conditional *Ezh2; Rnf2* DKO and *Eed; Rnf2* DKO mouse models. We show that embryos maternally and zygotically deficient for *Ezh2* and *Rnf2* or *Eed* and *Rnf2* progress relatively normal through pre-implantation development, but fail to complete gastrulation. We further study the transcriptional effects of PRC1/2 DKO in pre-implantation embryos and compare them to the effect in PRC2 deficient embryos in order to reveal possible PRC1-specific functions in gene regulation in early embryos. We show that PRC1 represses a set of genes with functions in germ line development. By comparing our data to external datasets, we provide evidence that germ line genes are repressed by the variant PRC1.6 complex that functions as a backup for DNA methylation in early embryos.

3.3 Results

3.3.1 Deletion of *Ezh2* and *Rnf2* leads to a reduction of H3K27me3 in GV oocytes

In order to study the role of PRC1 and PRC2 for pre-implantation development, we generated conditional *Ezh2; Rnf2* as well as *Eed; Rnf2* DKO mice. The presence of *Ring1* in these mice rescues the maternal effect that is observed in PRC1 DKO mice [173]. We used *Zp3-cre* to delete *Ezh2* and *Rnf2* in growing oocytes and *Gdf9-icre* to delete *Eed* and *Rnf2* in primary oocytes. Females were crossed to males harboring floxed alleles and *Prm1-cre* that deletes in late spermiogenesis, in order to generate embryos maternally and zygotically deficient for the respective genes. By Immunofluorescence (IF) staining, we confirm that deletion of *Ezh2* and *Rnf2* in growing oocytes by *Zp3-cre* leads to a reduction of H3K27me3 staining at the germinal vesicle (GV) oocyte stage, in both non-surrounded nucleolus (NSN) as well as surrounded nucleolus (SN) GV oocytes (Fig. 3.1A,B).

3.3.2 Embryos maternally and zygotically deficient for *Ezh2 and Rnf2* or *Eed* and *Rnf2* develop to the blastocyst stage, but exhibit gastrulation defects

We next investigated the developmental potential of embryos maternally and zygotically deficient for *Ezh2* and *Rnf2* (*Ezh2; Rnf2* DKO) or *Eed* and *Rnf2 (Eed; Rnf2* DKO). After 4 days of *in vitro* culture, at embryonic day 4.5 (E4.5), we observe that *Ezh2; Rnf2* DKO embryos develop to blastocyst at the same rate as control embryos (around 80%). *Eed; Rnf2* DKO embryos

develop to blastocyst stage at around 75%, compared to almost 100% of control embryos. Despite the high numbers of embryos in both DKO conditions that reach the blastocyst stage, we observe a developmental delay of both genotypes at embryonic days 2.5 (E2.5) and 3.5 (E3.5) (Fig. 3.2A). We used live imaging by light-sheet microscopy in order to track developmental progression of single embryos over time to see when the developmental delay in DKO embryos arises. Microinjection of embryos with *H2B-mCherry* mRNA allowed us to track cell divisions from 2-cell (E1.5) until blastocyst stage (E4.5). We observe that *Ezh2; Rnf2* DKO embryos are delayed already from the first time-point analyzed, where most control embryos have already reached the 4-cell stage, while DKO embryos still consist of 2 or 3 cells. Subsequently, both control and DKO embryos doubled their cell number at roughly the same rate, resulting in an average of 30 cells in control and 20 cells in DKO embryos at the last time point analyzed (Fig. 3.2B). These results suggest that the developmental delay of *Ezh2; Rnf2* DKO embryos arises already early in pre-implantation development, from the second or possibly even the first cell division after fertilization. IF analysis of *Ezh2; Rnf2* DKO embryos shows absence of EZH2 and RNF2 protein as well as H3K27me3 at the blastocyst stage (Fig. 3.2C, Supp. Fig. 3.2A). However, the level of H2AK119ub1 in *Ezh2; Rnf2* DKO embryos is comparable to control embryos, suggesting that the presence of *Ring1* rescues the ubiquitination function of PRC1 (Fig. 3.2D).

It has been reported that zygotic deletion of *Ezh2* or *Rnf2* causes embryonic lethality at early post-implantation stages due to gastrulation defects [167, 171]. In order to investigate when embryos maternally and zygotically deficient for both genes die, we first performed breeding experiments, which yielded no live offspring, indicating that DKO of *Ezh2* and *Rnf2* is embryonic lethal. Histological analysis of post-implantation embryos at different stages shows that *Ezh2; Rnf2* DKO embryos develop to E7.5, but are delayed to different degrees and do not progress through gastrulation normally (Supp. Fig. 3.2B). IF staining for lineage markers shows that *Ezh2; Rnf2* DKO E7.5 embryos express OCT4 in the epiblast and a subset even shows formation of the primitive streak, indicated by EOMES staining. However, none of the analyzed DKO embryos completed gastrulation as control embryos did (Fig. 3.2E).

3.3.3 Single embryo RNA sequencing of embryos maternally and zygotically deficient for *Eed, Ezh2* and *Rnf2*, and *Eed* and *Rnf2*

In order to investigate the transcriptional changes in PRC1 and PRC2 compound mutants, we performed single-embryo RNA-sequencing (RNA-seq) of *Ezh2; Rnf2* DKO and control as well as *Eed; Rnf2* DKO and control embryos at E3.5 and E4.5 using the Smart-seq2 protocol [360]. In order to analyze PRC1-specific effects on transcription in early embryos, we compared the transcriptional effects in DKO embryos to those in *Eed* KO embryos (Fig. 3.3A).

Since previous experiments had shown that the deletion efficiency of Prm1-cre is only around 90%, we checked for efficient deletion of floxed exons (exons 16 and 17 of *Ezh2* and exons 4 and 5 of *Rnf2*) in *Ezh2; Rnf2* DKO embryos. In most of the embryos, we do not detect any reads from these exons. However, 1 and 6 embryos show a significant number of reads coming from the deletion region of *Ezh2* and *Rnf2*, respectively (Supp. Fig. 3.3A). We inferred that these embryos are heterozygous and inherited an intact allele from the father. We performed the same analysis for *Eed; Rnf2* DKO embryos and identified 6 and 7 embryos presumably heterozygous for *Eed* and *Rnf2*, respectively (data not shown). All inferred heterozygous embryos were excluded from further analysis. The sex of each embryo was determined by the presence or absence of reads coming from the Y-chromosome (Supp. Fig. 3.3B). Multidimensional scaling (MDS) analysis of embryos from the 3 RNA-seq experiments shows that the first dimension corresponds to the developmental stage and separates E3.5 from E4.5 embryos. Dimension 2 clearly separates KO/DKO from control embryos. Control as well as KO/DKO embryos from the 3 experiments are largely intermingled with each other, indicating high technical reproducibility between the 3 experiments and that the 3 different knock out conditions have overall very similar transcriptional effects (Fig. 3.3B).

3.3.4 Comparative analysis of differentially expressed genes reveals regulation of germ line genes by PRC1

Differential gene expression revealed that more genes are up- than downregulated in both *Ezh2; Rnf2* as well as *Eed; Rnf2* embryos at E3.5 and E4.5. We identified 130 to 287 upregulated and 24 to 124 downregulated genes in the different contrasts (max. FDR=0.01, min. log2(FC)=2) (Supp. Fig. 3.4A). In order to compare the transcriptional effects in the two DKO conditions as well as in *Eed* KO embryos, we performed a meta-analysis of all 3 genotypes. K-means clustering based on differential expression in all genotypes resulted in

the identification of 5 gene clusters with distinct expression patterns (Fig. 3.4A). Cluster 1 contains genes that are upregulated in all 3 KO conditions, however, the effect is much stronger in the DKO embryos than in *Eed* KO embryos. Cluster 2 contains genes, which are upregulated very strongly only in specific embryos. Since every embryo was imaged prior to RNA-seq, we could link upregulation of genes in cluster 2 to an unhealthy morphology of embryos. Clusters 3 and 4 are highly enriched for genes located on the X chromosome. Cluster 3 contains genes that are downregulated upon KO/DKO. We observe an increase in expression of these genes from E3.5 to E4.5, which is likely due to X hyper activation that takes place progressively throughout pre-implantation development in order to achieve X/AA dosage compensation [333]. Cluster 4 contains very few X-linked genes that are upregulated in KO/DKO embryos. These genes show an interesting expression pattern in wild type, where they are only expressed in female embryos. Cluster 5 contains genes that show relatively mild upregulation in all 3 conditions analyzed. We further compared our transcriptomic data to ChIP-seq signal for H3K27me3 (TSS+/- 20kb) in MII oocytes, 2-cell and morula stage embryos [217]. We find enrichment for H3K27me3 at genes in clusters 1,4 and 5, suggesting that these clusters contain direct Polycomb targets. In contrast, clusters 2 and 3 show low H3K27me3 enrichment, suggesting that the genes in these clusters are misregulated due to secondary effects.

In order to get a better understanding of the function of the direct and indirect PcG targets, we performed GO term analysis for genes in the 5 clusters. Interestingly, we find that cluster 1, which contains genes highly upregulated in the DKO conditions, is enriched for GO terms relating to germ line functions (Fig. 3.4B). Cluster 2 is enriched for functions relating to proteolysis and apoptosis (Supp. Fig. 3.4B). This is in line with the notion that genes in this cluster are highly upregulated only in specific embryos with unhealthy morphology. For cluster 3, which contains X-linked, downregulated genes, we did not find a significant group of GO terms enriched. Cluster 4, which contains upregulated X-linked genes is enriched for functions in dosage compensation and contains the X chromosome inactivation (XCI) regulators *Xist* and *Ftx* (Supp. Fig. 3.4B). These results suggest that upon deletion of PcG proteins, *Xist* is upregulated in male and female embryos, inducing a global downregulation of X-linked genes. We have demonstrated that this is due to loss of H3K27me3, which serves as the maternal imprint of *Xist,* and ectopic XCI of the maternal X chromosome (see chapter 2.3.8). Cluster 5 is enriched for developmental functions, indicating that this cluster contains

typical PcG targets (Fig. 3.4B). This is in line with the upregulation in all three genotypes analyzed and the enrichment for H3K27me3 in early embryos.

3.3.5 Regulation of germ line genes by PRC1.6 and DNA methylation in ESCs

We were particularly interested in the upregulation of germ line genes, that occurred predominantly in DKO embryos. We show by IF staining that one gene from cluster 1, *Mael*, is upregulated also on the protein level at E3.5 in *Ezh2; Rnf2* DKO embryos (Fig. 3.5A). It is well known that many genes involved in germ cell development are repressed by DNA methylation in embryonic stem cells (ESCs) as well as in somatic cells and premigratory germ cells [372-374]. We therefore speculated that in early embryos, where DNA methylation is low, PcG proteins could serve as a backup mechanism for the repression of germ line genes [293, 375]. In order to investigate whether the same set of germ line genes is repressed by DNA methylation in ESCs and by PcG in early embryos, we compared differentially expressed genes in *Ezh2; Rnf2* DKO embryos to differentially expressed genes in ESCs lacking DNA methylation. Strikingly, we observe upregulation of genes belonging to cluster 1 in ESCs grown in de-methylating conditions (2i+vitamin C) as well as in *Dnmt* triple KO (tKO) ESCs (Fig. 3.5B,C) [376, 377]. It has further been reported that in ESCs, KO of both *Rybp* and *Pcgf6* causes de-repression of germ line genes, indicating that they are regulated by the variant PRC1.6 complex [93, 190]. We therefore tested whether the genes upregulated in *Rybp* and *Pcgf6* KO ESCs are the same that are upregulated in *Ezh2 Rnf2* DKO embryos. Even though the datasets are in general not well correlated, we find that genes belonging to cluster 1 are upregulated in all conditions analyzed (Fig. 3.5D,E). The same pattern is observed, when we compare DKO embryos to ESCs deficient for other components of PRC1.6, *Ring1a/b, Max* and *Mga* (Supp. Fig. 3.5A-C). These results indicate that a group of genes with functions in germ cell development is regulated by DNA methylation as well as PRC1.6 in ESCs. In early embryos, where DNA methylation levels are low, PRC1.6 likely serves as a backup mechanism for DNA methylation in order to repress these genes.

3.4 Discussion

In this study, we present a comparative analysis of embryos maternally and zygotically deficient for both PRC1 and PRC2 or only PRC2. We show that embryos maternally and zygotically deficient for *Ezh2/Eed* and *Rnf2* are delayed in pre-implantation development, but develop to the blastocyst stage at levels comparable to wild type embryos. Typically for PcG mutants, *Ezh2 Rnf2* DKO embryos die during gastrulation. However, we also observe an upregulation of apoptosis-related genes in a number of PRC1/2 DKO pre-implantation embryos, which appear also morphologically unhealthy. This result indicates that although most DKO embryos are able to develop to the blastocyst stage, they are more likely to accumulate developmental defects during pre-implantation development. We further confirm that PRC2, likely together with PRC1, represses canonical PcG targets as well as maternally imprinted genes on the X chromosome.

Interestingly, we observe an upregulation of genes with germ line functions in all the mutants analyzed, but much more prominently in PRC1/2 DKO embryos. It is known that genes with functions in male germ cell differentiation are overrepresented on the X chromosome [378]. In line with this, we find an enrichment for X-linked genes in cluster 1, which contains genes with germ line functions that are mostly upregulated in DKO embryos.

It has been shown that genes with functions in postmigratory germ cell development are silenced by DNA methylation in premigratory germ cells and somatic tissues [374]. Similarly, DNA methylation has been found to repress genes related to germ cell development like the *Rhox* and *Mage* gene families, in somatic cells and ESCs [336, 379, 380]. More recently, genome-wide studies have revealed de-repression of germ line genes in ESCs deficient for DNA methyltransferases [372, 373]. Interestingly, it has been shown that the germ line specific genes de-repressed in DNA methylation deficient ESCs are also de-repressed in ESCs deficient for the variant PRC1 component *Rybp*, grown in standard serum/LIF conditions [190]. This gene upregulation in *Rybp* deficient ESCs is independent of DNA methylation, suggesting that variant PRC1 and DNA methylation act in parallel to repress germ line genes in ESCs. More recently, it has been demonstrated that germ line genes are bound by PRC1.6 in ESCs and become de-repressed upon deletion of *Pcgf6* in ESCs grown in standard serum/LIF conditions, affecting cell growth and viability [93].

In the present study, we demonstrate that the same set of germ line genes is re-repressed in PRC1/2 DKO embryos, ESCs lacking DNA methylation and ESCs deficient for PRC1.6 components. We therefore conclude that this set of genes with functions in germ cell development is repressed in parallel by DNA methylation and the PRC1.6 complex in ESCs. On the contrary, in early embryos, where DNA methylation levels are low, PRC1.6 takes over repression of these genes. We speculate that the presence of two repression mechanisms at these genes in ESCs might be a consequence of the necessity for a backup mechanism in absence of DNA methylation in early embryos.

Notably, *Pcgf6* deficient mice are viable and fertile, but born at sub-Mendelian ratios with lethality of homozygous embryos starting at the E3.5 blastocyst stage and continuing during post-implantation development [93]. We therefore speculate that the developmental delay and increased apoptosis in pre-implantation PRC1/2 DKO embryos might result from impaired PRC1.6 function. Analysis of embryos maternally and zygotically deficient for *Pcgf6* would be needed to confirm this hypothesis and to proof that PRC1.6 regulates germ line genes in early embryos. We are further planning to include RNA-seq of embryos maternally and zygotically deficient for *Rnf2* in the analysis, to investigate whether the de-regulation of germ line genes in DKO embryos is solely due to loss of *Rnf2* or the result of a genetic interaction between PRC1 and PRC2.

3.5 Material and Methods

3.5.1 Mice

Ezh2 and *Rnf2* conditional double knock out (DKO) mice were generated as described previously [244]. *Eed* conditional knock out (KO) mice were generated as described in chapter 2.5.1. *Eed; Rnf2* DKO mice were generated by crossing *Eed* conditional mice to *Rnf2* conditional mice. *Ezh2; Rnf2* DKO mice carry a *Zp3-cre* transgene in females and a *Prm1-cre* in males, while *Eed* KO and *Eed; Rnf2* DKO mice carry a *Gdf9-icre* transgene in females and a *Prm1-cre* transgene in males. Maintenance and crosses of mice were done as described in chapter 2.5.1. *Eed* conditional mice were genotyped as described in chapter 2.5.1. Genotyping for *Ezh2* was done as described previously [244]. *Rnf2* conditional mice were genotyped using PCR primers to identify wild-type, floxed and deleted alleles (Supp. Fig. 3.1A,B). The 3 alleles

were distinguished using the PCR primers in Table 3.1 resulting in 420bp (wild-type), 490bp (floxed) and 450bp (excised) products, respectively. The mice were genotyped for *Gdf9-icre*, *Zp3-cre* or *Prm1-cre* with the primers listed in Table 3.1, yielding a 200bp (*Gdf9-icre*), 700bp (*Zp3-cre*) or 700bp (*Prm1-cre*) product. All experiments were performed in accordance with the Swiss animal protection laws and institutional guidelines.

Table 3.1. PCR primers used for genotyping

Gene name	Primer pairs	
	Forward	**Reverse**
Rnf2 flox	GTCTCATTTCCCAGTGTGTCCTC	ACTGACCCATGGCTCTTGATG
Rnf2 Δ	GTCTCATTTCCCAGTGTGTCCTC	GATGCACTGTCCTGATGG
Zp3-cre	GGAATTCAGGTGGGAGGGT	ATCCTGGCAATTTCGGCT
Gdf9-icre	AGATGCCAGGACATCAGGAACCTG	ATCAGCCACACCAGACACAGAGATC
Prm1-cre	GTTCCCTCAGCAGCATTCTC	AGGCAAATTTTGGTGTACGG

3.5.2 Collection of GV oocytes

Collection of GV oocytes was done as described in chapter 2.5.2.

3.5.3 Collection and culture of early mouse embryos

Embryo collection and culture was performed as described in chapter 2.5.3.

3.5.4 Light sheet microscopy

Zygotes were isolated as described in chapter 2.5.3. and injected with 7ng/µl *H2B-mCherry* fused to a *3xKSH-ENE* sequence, which enhances mRNA stability [381, 382]. Embryos were cultured until 2-cell stage as described in chapter 2.5.3. At the 2-cell stage, embryos were transferred into preconditioned KSOMaa in specific imaging chambers and imaged in a custom built inverted light sheet microscope [383]. Every 10 minutes, a Z-stack of each embryo was acquired at 2µM intervals. Embryos were imaged for 3 days, until they reached the blastocyst stage.

3.5.5 Immunofluorescence staining (IF) and microscopy of oocytes and pre-implantation embryos

IF staining of oocytes and pre-implantation embryos was done as described in chapter 2.5.4.

3.5.6 Histology and IF staining of post-implantation embryos

Embryos were generated by natural matings. Females were sacrificed 7 days after a vaginal plug was detected and deciduae containing E7.5 embryos were isolated and fixed in 4% PFA overnight (o/n). Tissues were de-hydrated through a series of ethanol (2x70%, 80%, 2x96%, 3x100%), xylene and embedded in paraffin using an automated tissue processing center (TPC 15 Duo, Medite). Sectioning was done at 3µm thickness using an automatic microtome (HM355S, Thermo Fisher Scientific). Sections were mounted onto slides and dried at 37°C o/n. Staining was done with hematoxylin and eosin. Sections were heated to 65°C for 40 min., de-paraffinized in Neo-Clear (2x 3min.) (109843, Millipore) and re-hydrated in a series of decreasing concentrations of ethanol (2x 100%, 95%, 70%, 3 min. each) to deionized H2O. Re-hydrated tissues were stained in hematoxylin (3min.), rinsed in deionized water shortly and then for 5 minutes in tap water. After repeated fast dipping in acid ethanol (8x), slides were again rinsed in tap water (2x 1 min.) and deionized water (2 min.). Sections were counterstained with eosin (30 sec.), de-hydrated in ethanol (95%, 100%, 2 sec. each), cleared in Neo-Clear (2x 5 min.) and mounted with Neo-Mount (109016, Millipore) on slides using a CTM6 Coverslipper (970010-19, Thermo Fisher Scientific).

Immunofluorescence staining of paraffin sections was done following the same protocol as described above until re-hydration of sections. Antigen retrieval was performed by boiling of slides in sodium citrate buffer (10 mM sodium citrate, 0.05% Tween 20, pH 6.0) for 20 min. using a microMED T/T Mega (Hacker Instruments & Industries). Blocking was done by incubation in blocking solution (5% horse serum, 1% BSA (w/v), 0.1% TritonX in PBS) for 60 min. at RT. The area around the sections was marked using a Dako pen and the sections were incubated in primary antibodies (Tab. 3.2) diluted in antibody incubation solution (1% horse serum, 1% BSA (w/v), 0.1% TritonX in PBS) at 4°C o/n in a humidified chamber. Slides were washed 3 x 15 min. in washing solution (0.1% Tween 20 in PBS) before incubation in secondary antibodies (Invitrogen) diluted 1:500 in antibody incubation solution for 60 min. Slides were again washed 3x in washing solution, mounted in Vectashield Antifade Mounting Medium

with DAPI (H-1200, Vector Laboratories) and sealed with nailpolish.

Table 3.2 : Antibodies used in this study

Antibody	Origin	Catalog number	Dilution
H3K27me3	Cell Signaling	9733	1:100
EZH2	Leica	NCL-L-EZH2	1:200
H2AK119ub1	Cell Signaling	8240	1:500
RING1B	Cell Signaling	5694	1:500
OCT4	Santa cruz	sc-8628	1:200
CDX2	BioGenex	MU392A-UC	1:200
EOMES	Abcam	Ab23345	1:100

3.5.7 RNA sequencing of pre-implantation embryos

Single-embryo RNA-seq was done using the Smart-seq2 protocol, as described in chapter 2.5.8.

3.5.8 Alignment of RNA-seq data and expression quantification

Spliced alignment of RNA-seq data to *M.musculus* genome assembly (GRCm38/mm10 Dec. 2011) was done using STAR [361] with parameters *-outFilterMultimapNmax 300 -outMultimapperOrder Random -outSAMmultNmax 1 -alignIntronMin 20 -alignIntronMax 1000000*, allowing multimappers with up to 300 matches in the genome and choosing positions for multimappers randomly.

Read quantification for genes was done using Bioconductor annotation package *TxDb.Mmusculus.UCSC.mm10.knownGene* (version 3.4.0) and qCount function from *QuasR* R package [362] selecting only uniquely mapped reads (parameter mapqMin=255). Genotyping of samples and detection of heterozygous embryos was done by quantification of expression of each exon and each exon-exon splice junctions of *Ezh2, Eed* and *Rnf2* genes.

3.5.9 Gender inference for embryos using single-embryo RNA-seq data

Chromosomes X and Y were partitioned into 5kb non-overlapping tiles and only those tiles which have at least 80% mappability (80% or more of 50bp windows in a tile are unique in the genome) were selected for further analysis. Uniquely mapped RNA-seq reads overlapping selected 5kb tiles were aggregated for chrX and chrY and percentages were calculated by

dividing number of reads in chrX and chrY by total number of mapped reads for each embryo. Genders for embryos were assigned using appropriate cutoff for percentage of reads mapping to chrY.

3.5.10 Analysis of differential expression of genes in single embryo RNA-seq data

Differential expression analysis was performed using edgeR (version 3.14.0) by fitting a negative binomial generalized linear model to the read counts for each gene [365]. All genes with at least 1 read per million in at least 2 samples were included in the analysis. P-values for differential expression were calculated using log-likelihood tests and correction for multiple testing was done using Benjamini-Hochberg method. K-means clustering of genes was performed with the 'stats::kmeans' function, on log2(CPM) values centered by average expression. Heatmaps were generated using the *ComplexHeatmap* package [367]. Gene ontology term enrichment analysis was performed in R with the *GOstats* package [384], using a conditional hypergeometric test that uses the relationships among GO terms to decorrelate the results. Enrichments of genes from individual clusters were tested over all expressed genes. Biological Process (BP) GO-terms with a minimum gene size of 10 and with the most significant enrichments were displayed in barplots.

3.5.11 Analysis of ChIP-seq data

Published H3K27me3 ChIP-seq dataset [217] for pre-implantation embryo was downloaded from NCBI GEO repository (GSE73952). Reads were aligned to *M.musculus* genome assembly (mm10) using qAlign function from *QuasR* R package [362]. Reads for ChIP samples were counted for regions ±20kb around TSS using *QuasR* package selecting only uniquely mapped reads. ChIP signal was normalized for sex chromosome counts (counts on X and Y chromosomes were multiplied by 2) and by library size.

3.5.12 Analysis of bulk RNA-seq datasets

Previously published datasets for tKO ESCs, ESCs cultured in 2i+VitaminC, Rybp KO ESCs, Ring1a/b KO ESCs, Max and Mga siRNA knock down ESCs and Pcgf6 KO ESCs were downloaded

from NCBI GEO repository (GSE67867 [377], GSE71591 [376], GSE32294 [190], GSE84480 [93]).

Bulk RNA-seq datasets were aligned to mm10 genome using qAlign function from *QuasR* R package and unique mappers were counted for each gene using QuasR package. [362]. Genes with at least 1 read per million in at least 2 samples were included in the analysis. Analysis of differential expression of genes was performed using edgeR package using log-likelihood tests and Benjamini-Hochberg correction for multiple testing.

3.6 Figures

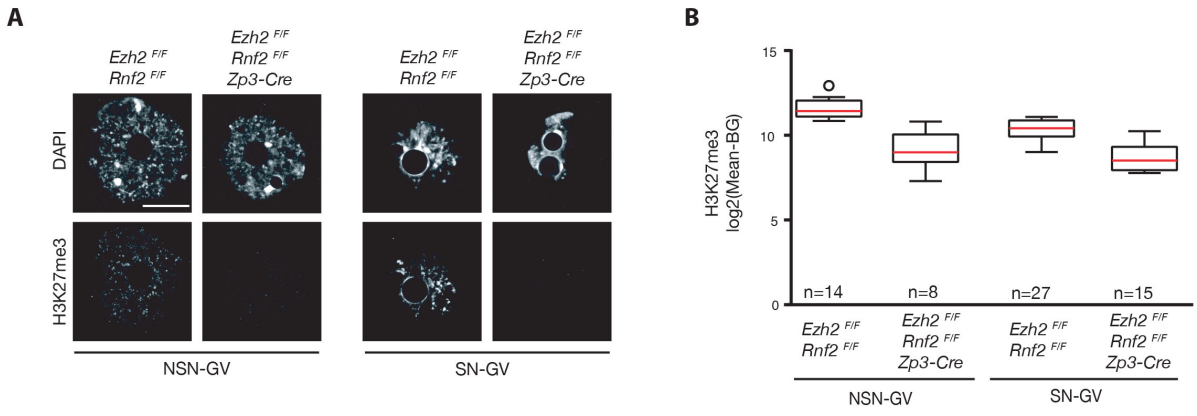


Figure 3.1. Deletion of *Ezh2* and *Rnf2* leads to a reduction of H3K27me3 in GV oocytes. (A) Immunofluorescence staining for H3K27me3 in NSN-GV and SN-GV oocytes from *Ezh2* $^{F/F}$;*Rnf2* $^{F/F}$ and *Ezh2* $^{F/F}$;*Rnf2* $^{F/F}$ *Zp3-cre* females. Upper panels show DNA stained with DAPI, lower panels show staining for H3K27me3. Scale bar 20µM. **(B)** Quantification of signal intensity for H3K27me3 in NSN-GV and SN-GV oocytes from *Ezh2* $^{F/F}$;*Rnf2* $^{F/F}$ and *Ezh2* $^{F/F}$;*Rnf2* $^{F/F}$ *Zp3-cre* females. Quantification was done by measuring the mean intensity for H3K27me3 in the nucleus (nuclear area determined according to DAPI staining) and normalized by subtracting the background signal (measured in an area outside the nucleus). Normalized relative intensity values are represented as Tukey boxplots on a logarithmic scale for NSN-GV and SN-GV oocytes, separately. Red lines represent the median, black boxes include the 25th to 75th percentiles. Whiskers show the non-outlier range of the data. Circles represent outliers, which are defined as values that are more than 1.5 times the box height above or below the box. The number of analyzed oocytes per genotype is indicated in the graph.

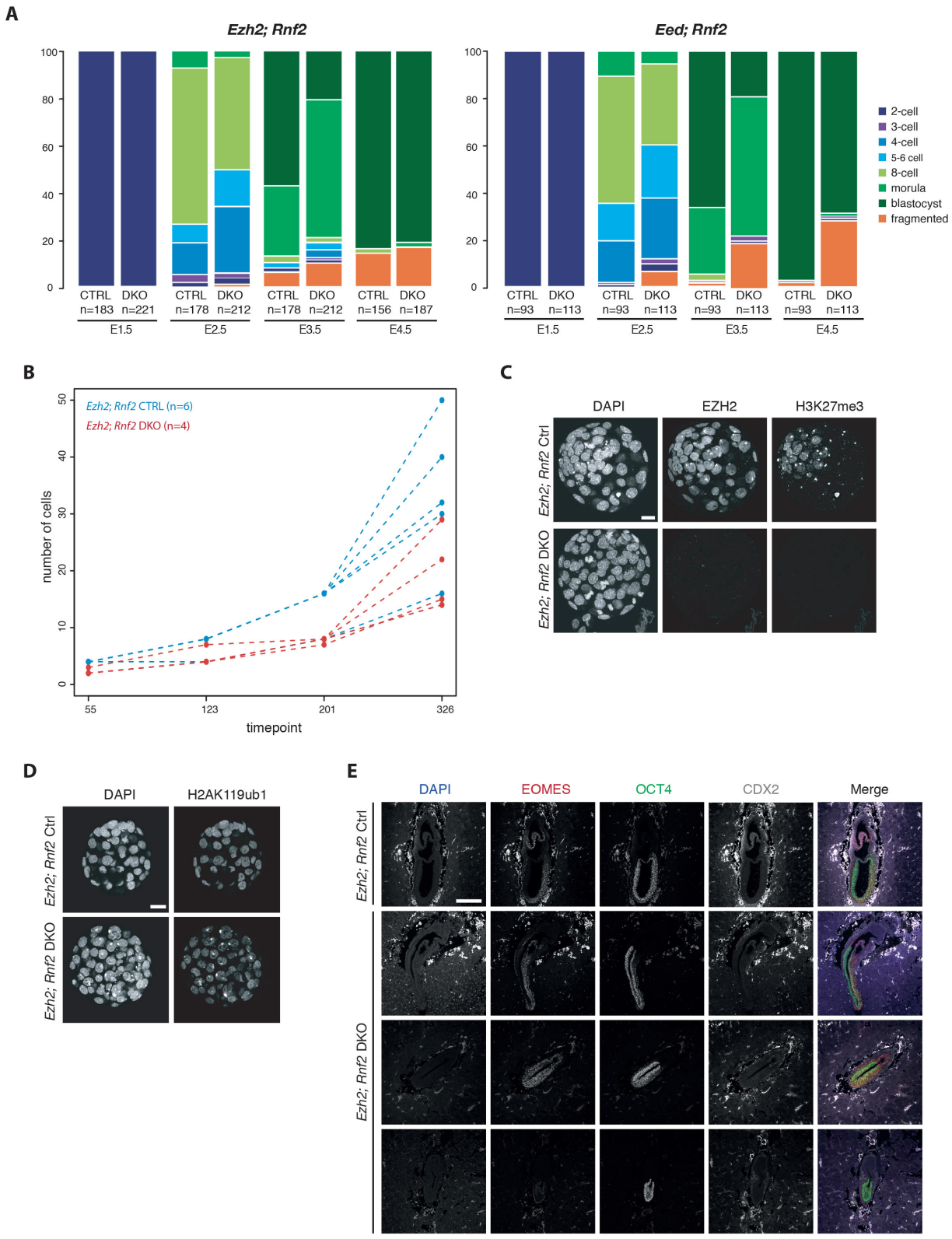


Figure 3.2. Embryos maternally and zygotically deficient for *Ezh2 and Rnf2* or *Eed* and *Rnf2* develop to the blastocyst stage. (A) Comparative time course developmental progression analysis of *Ezh2; Rnf2* Ctrl and *Ezh2; Rnf2* DKO embryos (left panel) and *Eed; Rnf2* Ctrl and *Eed; Rnf2* DKO embryos (right panel)**.** Developmental stages were scored daily for 4 days from E1.5 to E4.5. The number of analyzed embryos is indicated on the graphs. Data pooled from multiple independent experiments (5 experiments for *Ezh2;Rnf2*, 2 experiments for

Eed;Rnf2). **(B)** Plot showing the number of cells of individual *Ezh2; Rnf2* Ctrl and *Ezh2; Rnf2* DKO embryos at 4 time points of pre-implantation development measured by light sheet microscopy. Time point 55 corresponds to 9.2h, time point 123 to 20.5h, time point 201 to 33.5h and time point 326 to 54.3h after start of imaging at the mid 2-cell stage. The number of embryos analyzed per genotype is indicated on the graph. **(C)** Immunofluorescence (IF) staining for EZH2 and H3K27me3 in *Ezh2; Rnf2* Ctrl and DKO embryos at E4.5. Scale bar 20µM. **(D)** IF staining for H2AK119ub1 in *Ezh2; Rnf2* Ctrl and DKO embryos at E4.5. Scale bar 20µM. **(E)** IF staining for EOMES, OCT4 and CDX2 in histological sections of *Ezh2; Rnf2* Ctrl and DKO embryos at E7.5. Scale bar 200µM.

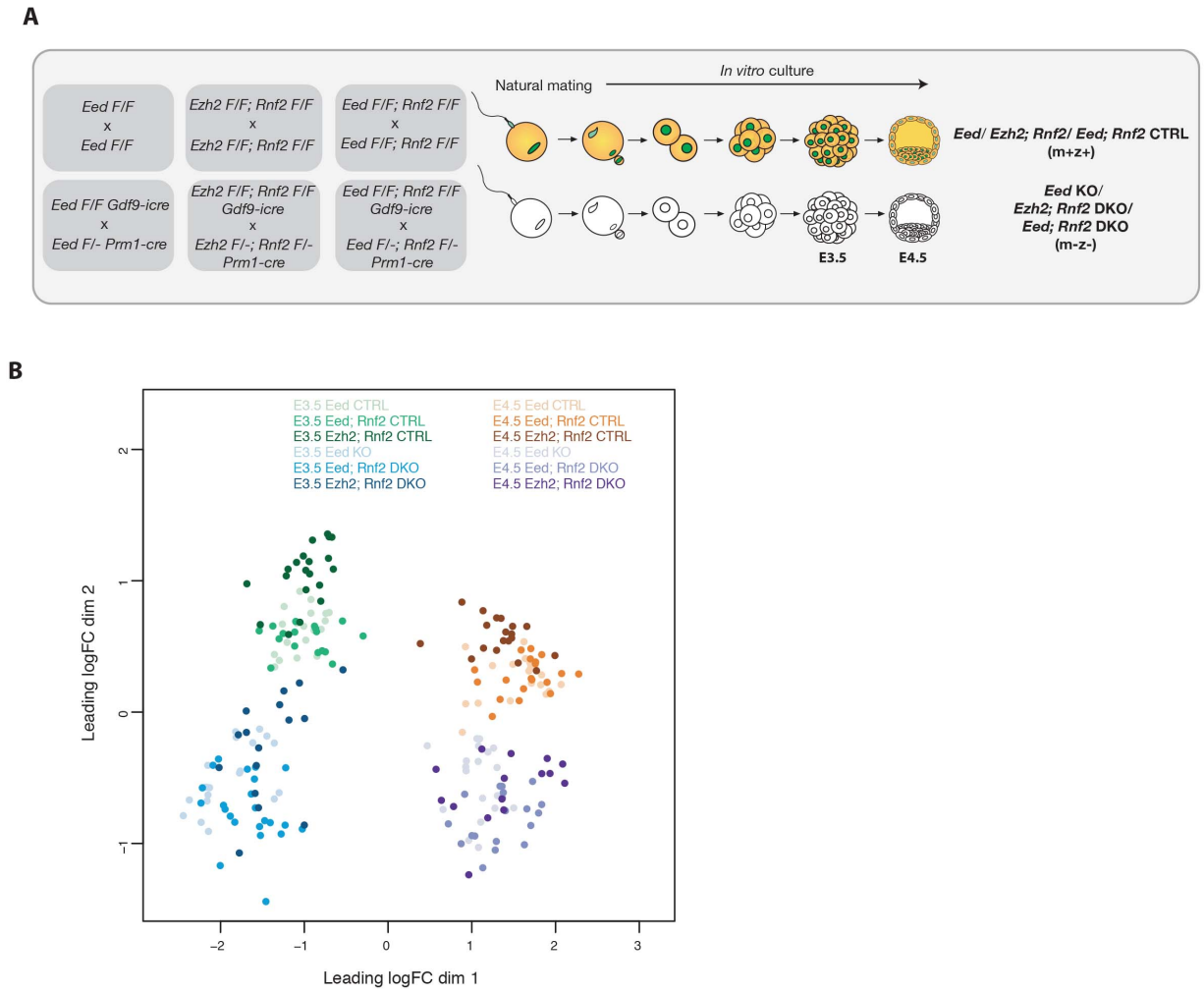


Figure 3.3. **Single embryo RNA sequencing of embryos maternally and zygotically deficient for *Eed, Ezh2* and *Rnf2*, and *Eed* and *Rnf2*. (A)** Schematic illustration showing the 3 Smart-seq2 experiments performed in this study. Experiment 1 includes *Eed* Ctrl and KO embryos that were generated by natural mating of *Eed* $^{F/F}$ or *Eed* $^{F/F}$ *Gdf9-icre* females with *Eed* $^{F/F}$ or *Eed* $^{F/-}$ *Prm1-cre* males, respectively. Experiment 2 includes *Ezh2; Rnf2* Ctrl and DKO embryos generated by natural mating of *Ezh2* $^{F/F}$ *Rnf2* $^{F/F}$ and *Ezh2* $^{F/F}$ *Rnf2* $^{F/F}$ *Zp3-cre* females with *Ezh2* $^{F/F}$ *Rnf2* $^{F/F}$ and *Ezh2* $^{F/-}$ *Rnf2* $^{F/-}$ *Prm1-cre* males, respectively. Experiment 3 includes *Eed; Rnf2* Ctrl and DKO embryos generated by natural mating of *Eed* $^{F/F}$ *Rnf2* $^{F/F}$ and *Eed* $^{F/F}$ *Rnf2* $^{F/F}$ *Gdf9-icre* females with *Eed* $^{F/F}$ *Rnf2* $^{F/F}$ and *Eed* $^{F/-}$ *Rnf2* $^{F/-}$ *Prm1-cre* males, respectively. Embryos were cultured *in vitro* and collected for RNA sequencing at E3.5 and E4.5. **(B)** Multidimensional scaling (MDS) analysis including samples from the 3 Smart-seq2 experiments.

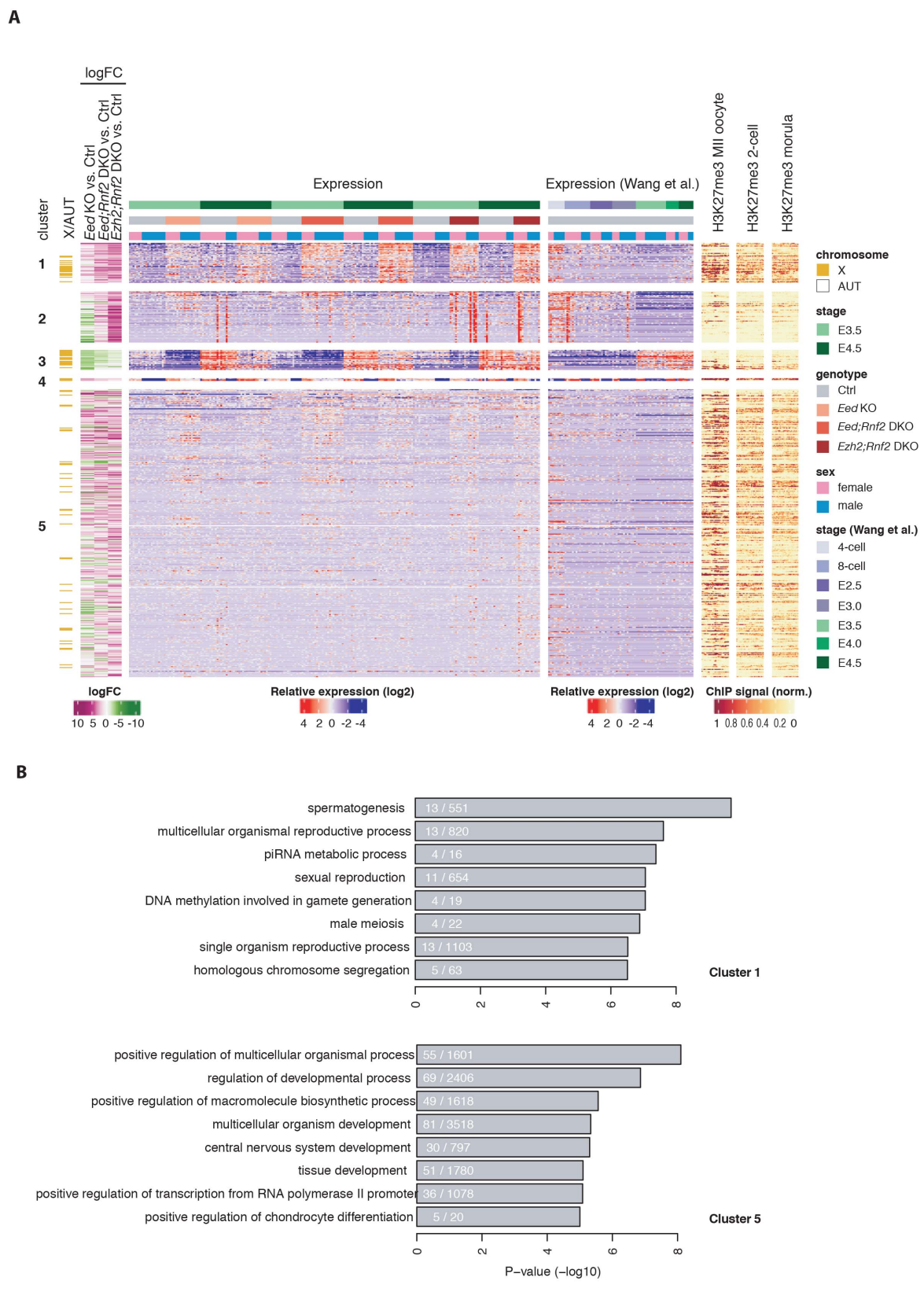


Figure 3.4. Differential gene expression analysis reveals regulation of germ line genes by PRC1. (A) Heatmap showing expression in embryos from the three RNA-seq experiments at E3.5 and E4.5 for genes with FDR ≤ 0.01 and log(FC) ≥ 2 in any of the analyzed contrasts. K-means clustering was performed to identify 5 clusters of differentially expressed genes. “X/AUT” indicates location of genes on autosomes or the X chromosome. “logFC” indicates mean log2(FC) for all samples of each genotype vs. the respective controls.

“Expression” shows relative expression in individual samples of the different genotypes at E3.5 and E4.5, separate for female and male embryos. “Expression (Wang et al.)” shows relative expression at different stages of pre-implantation development in single embryos from a published dataset, separated for female and male embryos [333]. Right columns show normalized ChIP signal for H3K27me3 in MII oocytes, 2-cell embryos and morula stage embryos, +/- 20kb around TSS, from a published dataset [217]. Legends are plotted below each section of the heatmap. **(B)** Bar plots showing GO terms for clusters 1 and 5. The x-axis displays significance (-log10(p-value)) and the number of genes in the respective GO category is indicated on the bars.

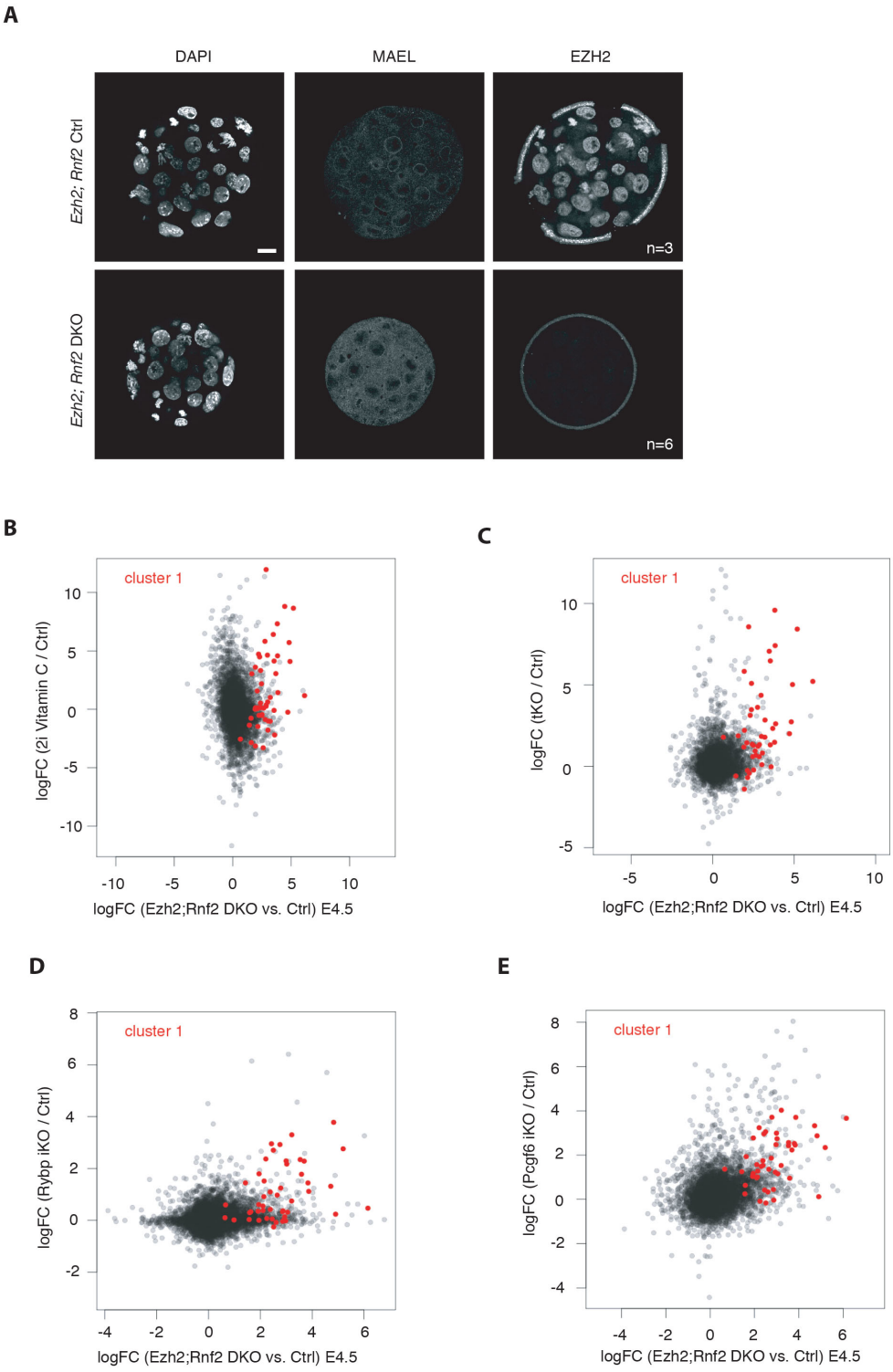
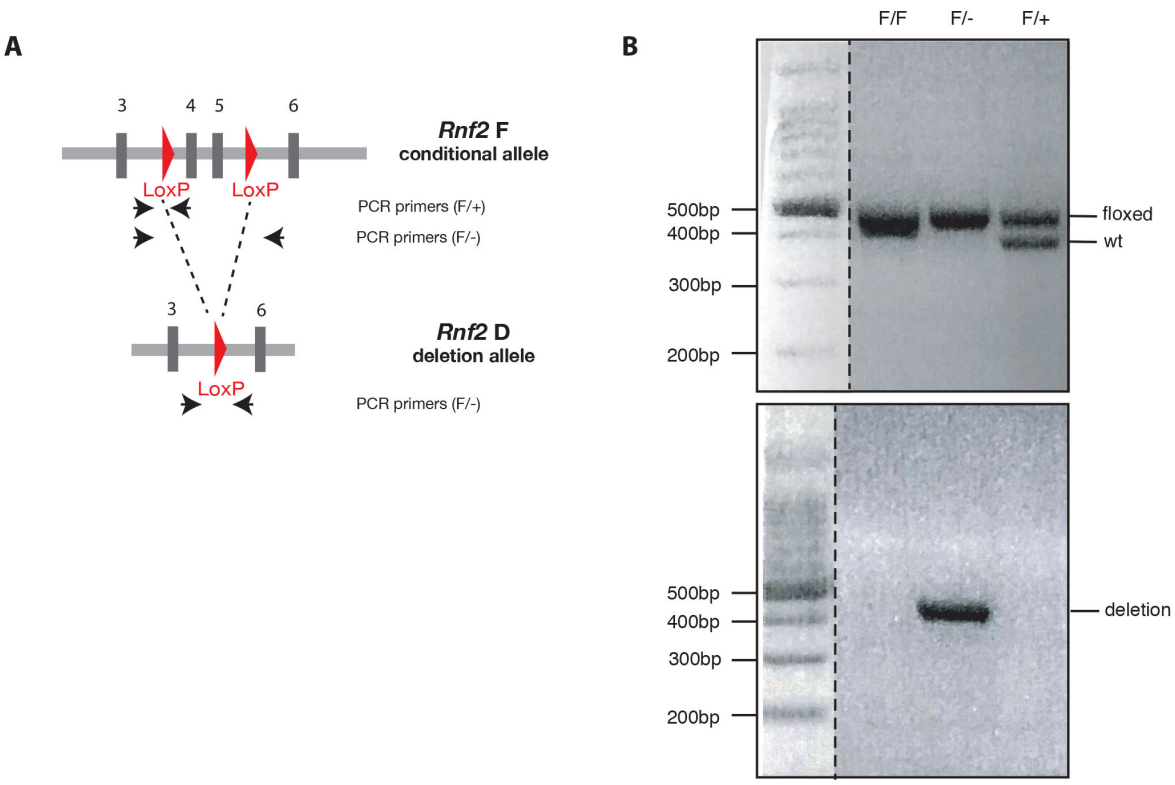
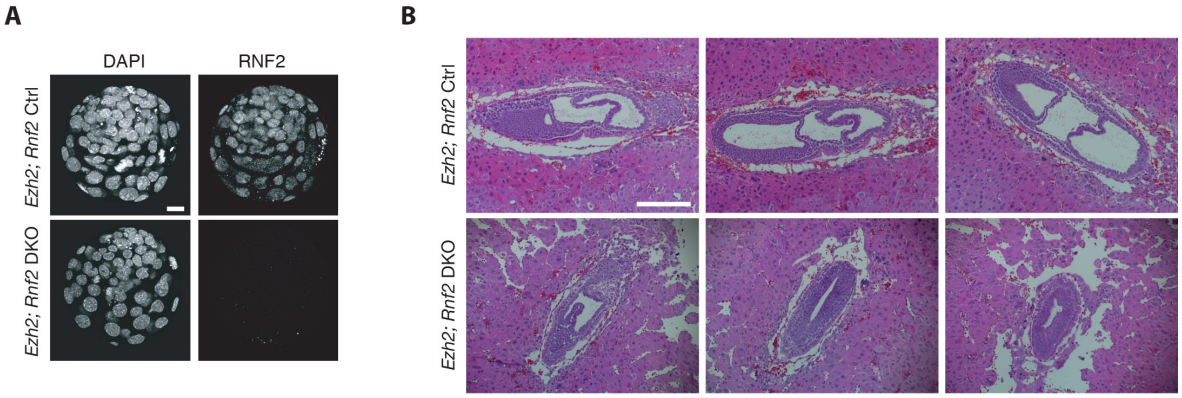


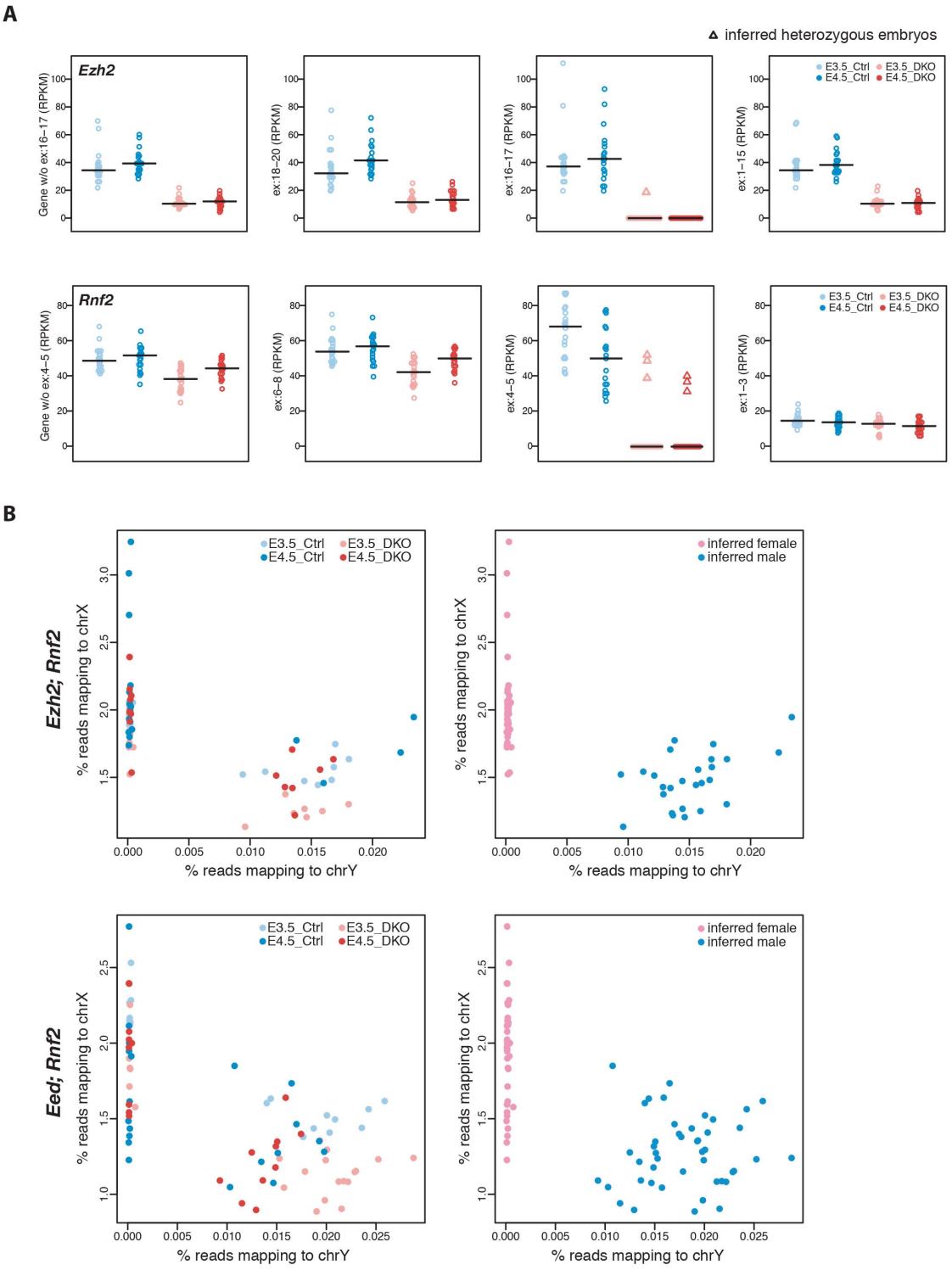
Figure 3.5. Regulation of germ line genes by PRC1.6 and DNA methylation in ESCs. (A) Immunofluorescence staining for MAEL and EZH2 in *Ezh2; Rnf2* Ctrl and DKO E3.5 embryos. Left panels show DNA stained with DAPI, middle panels show staining for MAEL and right panels staining for EZH2. Numbers of embryos analyzed per genotype are indicated on the respective panels. Scale bar 20µM. **(B-E)** Scatter plots showing correlation between differentially expressed genes in E4.5 embryos maternally and zygotically deficient for *Ezh2* and *Rnf2* (x-axis) and differentially expressed genes in **(B)** ESCs grown in 2i medium + Vitamin C for 6 days [376]. **(C)** ESCs deficient for *Dnmt* enzymes (tKO) [377] **(D)** ESCs deficient for *Rybp* [190]. **(E)** ESCs deficient for *Pcgf6* [93]. Red colored dots correspond to genes in cluster 1.



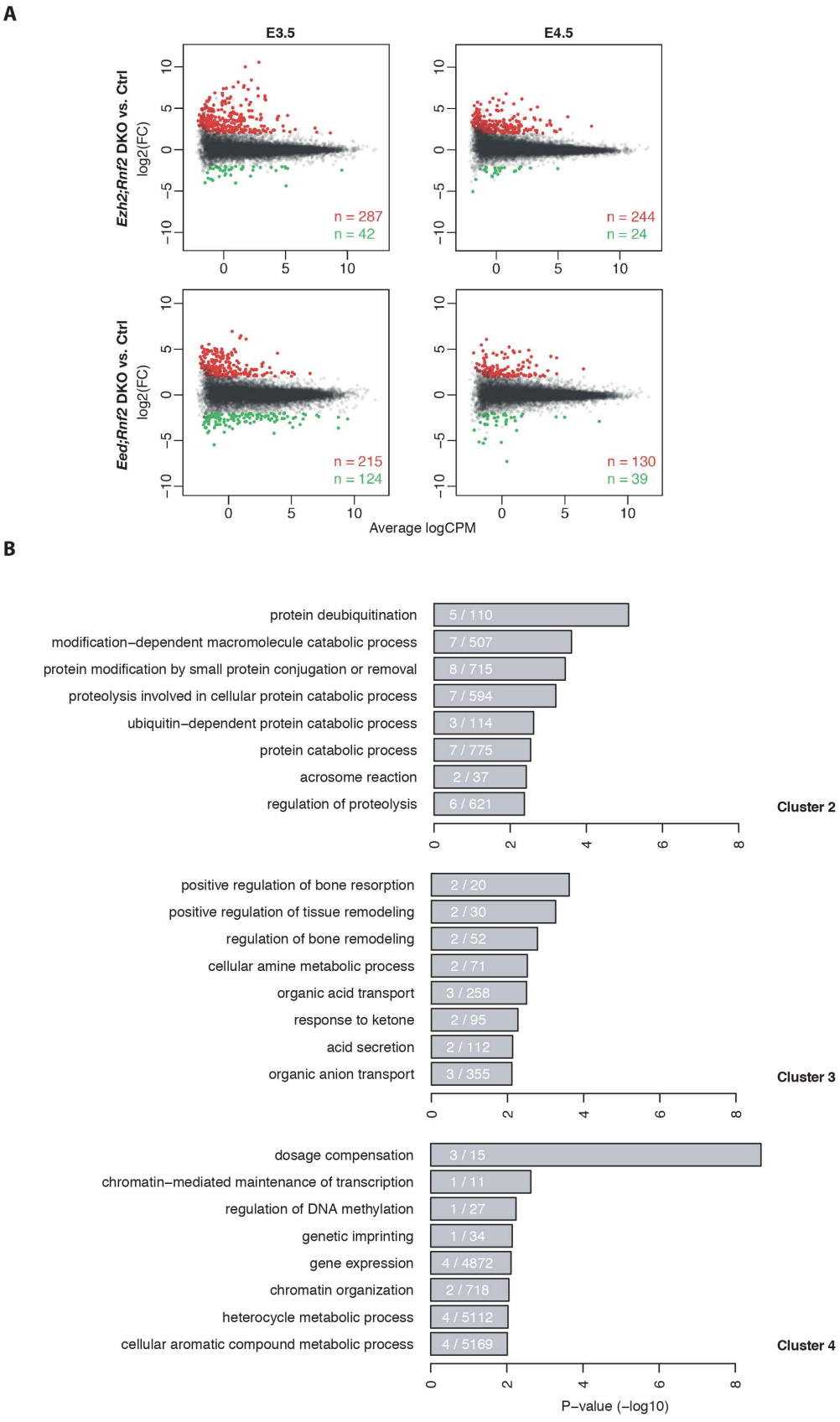
Supplementary Figure 3.1. Genotyping strategy for *Rnf2* conditional allele. (A) PCR genotyping strategy to identify wild type, floxed and deletion *Rnf2* alleles. **(B)** Agarose gels showing the results of PCR genotyping for *Rnf2* wild type, floxed (upper panel) and deletion (lower panel) alleles.



Supplementary Figure 3.2. Immunofluorescence and histological analysis. (A) IF staining for RNF2 in *Ezh2; Rnf2* Ctrl and DKO embryos at E4.5. Scale bar 20µM. **(B)** Hematoxylin-eosin (HE) staining on histological sections of *Ezh2; Rnf2* Ctrl and DKO embryos at E7.5.

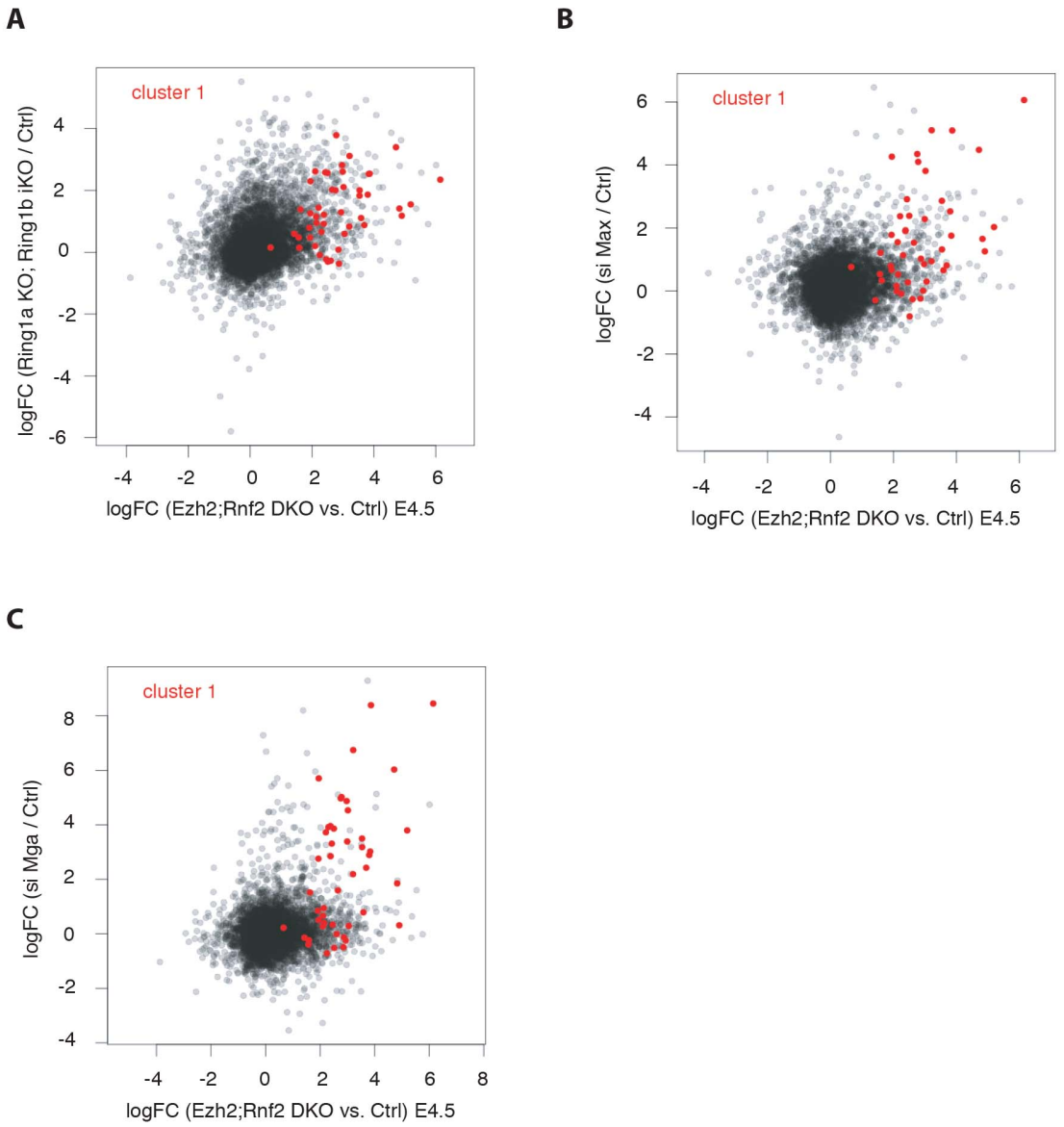


Supplementary Figure 3.3. Genotyping and sexing of embryos from RNA-seq data. (A) Beeswarm plots showing number of reads (RPKM) for different exons of *Ezh2* (upper panels) and *Rnf2* (lower panels) in *Ezh2; Rnf2* Ctrl and DKO embryos at E3.5 and E4.5. Each dot represents one embryo and black lines represent the median for each group. Inferred heterozygous embryos are represented as triangles. **(B)** Scatter plots showing the percentage of reads mapped to chromosome Y (X-axis) and chromosome X (Y-axis) in *Ezh2; Rnf2* Ctrl and DKO (upper panels) and *Eed; Rnf2* Ctrl and DKO (lower panels) embryos. Each point represents one embryo. Left panels show points colored according to embryo stage and genotype, right panels show points colored according to the inferred sex of the embryo.



Supplementary Figure 3.4. Differential gene expression and GO term analysis. (A) MA plots showing differentially expressed genes in *Ezh2; Rnf2* DKO (upper panels) and *Eed; Rnf2* DKO (lower panels) embryos, split for E3.5 and E4.5. The x-axis corresponds to the level of expression (Average CPM) and the y-axis displays

the level of misregulation (log2(FC)). Red dots represent significantly upregulated and green dots significantly downregulated transcripts in the respective contrast (max. FDR=0.01, min. log2(FC)=2). The numbers of significantly up- and downregulated genes are indicated on the graphs. **(B)** Bar plots showing GO terms for clusters 2,3 and 4. The x-axis displays significance (-log10(p-value)) and the number of genes in the respective GO category is indicated on the bars.



Supplementary Figure 3.5. Regulation of germ line genes by PRC1.6 in ESCs. (A-C) Scatter plots showing correlation between differentially expressed genes in E4.5 embryos maternally and zygotically deficient for *Ezh2* and *Rnf2* (x-axis) and differentially expressed genes in **(A)** ESCs deficient for *Ring1a/b* [93]. **(B)** ESCs treated with siRNA against *Max* [93]. **(C)** ESCs treated with siRNA against *Mga* [93]. Red colored dots correspond to genes in cluster 1.

Author contributions

J.B and A.P. conceived and designed the experiments. J.B. performed the experiments, including histology, IF, developmental progression analysis, light sheet microscopy and single-embryo RNA-seq. P.N. performed quantification of H3K27me3 IF signal in GV oocytes. H.R performed bioinformatic analysis of RNA-seq data and external datasets with support from E.O. S.S helped in setting up single-embryo RNA-seq and provided technical support. J.B. wrote the manuscript with input from E.O and A.P. A.P. supervised the project.

Detailed author contributions

Figure 3.1
A and B: Oocyte collection, IF staining and imaging was done by JB. Quantification was done by PN.

Figure 3.2
A: Data collected and analyzed by JB.
B: Embryo isolation, light sheet imaging and analysis performed by JB.
C and D: Embryo collection, IF staining and analysis done by JB.
E: Embryo collection, histology, IF staining and analysis done by JB.

Figure 3.3
A: Cartoon prepared by JB (adapted from AP).
B: Embryo collection and RNA-seq library preparation was done by JB with support from SS, analysis was performed by HR.

Figure 3.4
A and B: Embryo collection and RNA-seq library preparation was done by JB with support from SS, analysis was performed by HR.

Figure 3.5
A: Embryo collection, IF staining and analysis done by JB.
B: Embryo collection and RNA-seq library preparation was done by JB with support from SS, analysis was performed by HR and EO.

Supp. Figure 3.1
A: Cartoon prepared by JB.
B: Genotyping was done by JB.

Supp. Figure 3.2
A: Embryo collection, IF staining and analysis done by JB.
B: Embryo collection, histology and analysis done by JB.

Supp. Figure 3.3
A and B: Embryo collection and RNA-seq library preparation was done by JB with support from SS, analysis was performed by HR.

Supp. Figure 3.4
A and B: Embryo collection and RNA-seq library preparation was done by JB with support from SS, analysis was performed by HR.

Supp. Figure 3.5
A – C: Embryo collection and RNA-seq library preparation was done by JB with support from SS, analysis was performed by HR and EO.

Chapter 4. General discussion and outlook

4.1 Main findings

In my thesis, I have studied the role of Polycomb repressive mechanisms during mouse pre-implantation development. Most notably, we discovered that PRC2 regulates gene expression and dosage compensation in early embryos. By adapting a single-cell RNA-seq protocol, we were able to analyze the transcriptomes of single pre-implantation embryos deficient for the core PRC2 component *Eed.* Using this method, we have shown that a set of genes depends on the function of PRC2 in oocytes for maternal allele-specific repression in early embryos. We have demonstrated that the X inactivation regulator *Xist* is subject to this allele-specific repression, indicating that H3K27me3 serves as the maternal repressive imprint for *Xist.* We have further shown that PRC2 dynamically regulates genes involved in development and differentiation at different stages of pre-implantation development and that loss of maternal PRC2 impacts post-implantation development (Chapter 2). In a second project, we discovered that PRC1 represses germ line genes in early embryos. We performed single-embryo RNA-seq of PRC1/2 compound mutant embryos and found a strong upregulation of genes with functions in germ line development. Comparative analysis revealed that these genes are regulated by DNA methylation as well as the PRC1.6 complex in ESCs. We therefore hypothesize that in early embryos, where DNA methylation is low, PRC1.6 functions as a backup mechanism for repression of these genes (Chapter 3).

4.2 Role of PRC2 in gene expression and dosage compensation in early embryos

We have shown that PRC2 dynamically regulates expression of genes with functions in development and differentiation in pre-implantation embryos. We identified a group of genes that is repressed by PRC2 at early pre-implantation stages and a second group, which is active early and becomes repressed by PRC2 at later stages. While the second group clearly becomes silenced *de novo* at morula or blastocyst stages, the first group of genes might either become *de novo* repressed at the 8-cell stage or retain a repressed state by inheritance of H3K27me3 from the oocyte. In order to separate between the 2 possibilities, analysis of transcriptomic data at earlier stages of pre-implantation development and in oocytes will be informative. This would allow us to see whether the genes are already repressed in oocytes and early embryos or are active and become repressed *de novo* at early cleavage stages.

In line with a recent publication that reported DNA methylation independent imprinting by H3K27me3, we discovered that PRC2 represses a set of genes specifically on the maternal allele in early embryos [234]. We have shown that this allele-specific repression depends on the maternal function of PRC2, suggesting that at the identified genes, H3K27me3 is inherited from oocytes to early embryos. We discovered that the genes regulated in this fashion include a number of genes located on the X chromosome. These genes mostly appear in clusters, including *Rhox, Bex* and *Fthl17* gene families. We observed that all of these gene clusters are embedded in large H3K27me3 domains that are inherited from oocytes to early embryos. Many of the identified genes have been previously shown to be expressed exclusively in female embryos, indicating that they are maternally imprinted, but the nature of the imprint remained elusive [334, 335, 344, 345]. We have further shown that the master regulator of X inactivation, *Xist*, is repressed by H3K27me3 specifically on the maternal allele in early embryos and becomes de-repressed upon PRC2 deletion, leading to ectopic X inactivation of the maternal X chromosome in both male and female embryos. The same effect has recently been observed, when an H3K27me3 demethylase was injected into zygotes [327].

While maternal-specific repression at the *Xist* locus is required for keeping the maternal X chromosome active, the function of this mode of gene regulation at other X-linked and autosomal genes is less clear. Inoue *et al.* have shown that most genes retain maternal-

specific repression only during pre-implantation stages. In post-implantation embryonic tissues, maternal-specific repression by H3K27me3 is completely lost, while for few genes it is retained in placental tissues [234]. Therefore, the dynamics of this H3K27me3-dependent allele-specific repression is strikingly different from canonical methylation-dependent imprinting, which is largely maintained in both embryonic and extraembryonic lineages [reviewed in 385, 386]. It is well known that many methylation-dependent imprinted genes function to control pre-natal growth and the regulation of nutritional resources *in utero* through the placenta and that the placenta is a predominant site of tissue-specific imprinting [reviewed in 387, 388]. Therefore, it has been hypothesized that genomic imprinting evolved as a response to extrinsic or intrinsic signals in order to allow modulation of expression levels of these genes as required by the conditions [386]. Interestingly, the few H3K27me3-dependent imprinted genes that maintain imprinting in extraembryonic lineages of post-implantation embryos include *Gab1*, *Sfmbt2* and *Slc38a4*, all of which have been shown to have important functions for placental growth [234, 347-349]. However, it remains elusive why only few genes retain maternal repression by H3K27me3 and why it is lost at most other genes. How are the few genes selected and how is lineage-specific imprinting in extraembryonic tissues achieved? What is the function of this mode of regulation for the genes that do not retain it after implantation? Perhaps the asymmetric marking by H3K27me3 in early embryos is just a byproduct of the differential chromatin marking in oocytes and sperm and does not fulfill a particular function in pre-implantation embryos [217, 218]. Possibly, few genes with important functions in placental development have evolved to maintain this mode of regulation in extraembryonic tissues in order to regulate dosage compensation in response to changing conditions.

Inoue *et al.* have demonstrated that out of 76 genes that show maternal allele-specific repression in pre-implantation embryos, 5 maintain their imprinted expression in extraembryonic tissues of post-implantation embryos [234]. It will be interesting whether the same set of genes maintains maternal allele-specific de-repression after implantation in our system, where PRC2 is genetically deleted in oocytes. Therefore, I have recently performed RNA-seq of embryonic and extraembryonic lineages of single E7.25 embryos to investigate how much of the allele-specific repression is maintained in the different lineages of post-implantation embryos. While the results for post-implantation embryos by Inoue *et al.* were based on allele-specific transcriptomic analysis of wild type hybrid embryos and correlations

to H3K27me3-imprinted genes identified in pre-implantation embryos by injection of an H3K27me3 demethylase, we will be able to see a direct effect of maternal PRC2 deletion on allelic expression in post-implantation tissues [234].

Interestingly, I have observed decreased litter sizes in breedings maternally deficient for *Eed* compared to control breedings. These results are in line with a recent publication that further reported an overgrowth phenotype of the offspring that develops to term [338]. We also observed higher weights of pups maternally deficient for *Eed*, but attributed them to the smaller litter sizes. In my opinion, more careful analysis including embryo transfer experiments would be needed in order to rule out that this effect is due to differences in nutritional resources provided by the mother. However, it is highly interesting that maternal deficiency for PRC2 has an effect on development after implantation. In order to discover the reason for this lethality, we are currently analyzing RNA-seq of embryonic and extra-embryonic lineages of E7.25 embryos, as described above. This will provide information about the transcriptional effects in the different lineages of post-implantation embryos. In order to investigate the possibility that maternal PRC2 controls placental development, possibly by maternal-specific repression of genes involved in placental growth, we are planning to perform histological analysis of placental tissues from *Eed* maternal deficient and control embryos.

Another interesting question is why embryos maternally and zygotically deficient for *Eed* are able to complete pre-implantation development in absence of H3K27me3 and despite considerable transcriptional misregulation. Even though PRC2 deficient embryos develop to the blastocyst at normal levels, we observe a delay in pre-implantation development and an increase of OCT4 and a decrease of CDX2 by IF staining in *Eed* deficient embryos. Therefore, it might be possible that the PRC2-dependent regulation of genes involved in development and differentiation might not be absolutely required for the acquisition of totipotency, but rather prepare the embryo for the rapid cell divisions during specification of the 3 germ layers at gastrulation. Perhaps, H3K27me3 is not instructive for lineage segregation, but rather fine-tunes expression in the respective lineages and functions as a “memory” system at later stages of development, where cell fate decisions need to be remembered. Our results are in line with the fact that in ESCs, PRC2 components are dispensable for self-renewal and expression of pluripotency genes, but are required for differentiation into the three germ layers [182, 184, 194].

Taken together, our results argue that PRC2 and H3K27me3 play a substantial role for gene regulation in pre-implantation embryos. However, recent studies including ChIP-seq in germ cells and early embryos argue that on the paternal allele of early embryos, H3K27me3 is largely absent and on the maternal allele, it is specifically removed from promoters of developmental genes after fertilization and only retained as broad domains in distal regions of the genome [217, 218]. It has been suggested that H3K27me3 at promoters of canonical PcG targets only re-appears in the ICM of the blastocyst and increases in the post-implantation epiblast [218]. In order to relate the expression changes upon *Eed* KO to the chromatin mark, we re-analyzed the published H3K27me3 ChIP-seq data from Liu *et al.* for MII oocytes and different stages of pre-implantation development [217]. We classified promoters into 3 categories (low, medium and high H3K27me3) at each stage. Our analysis confirms a general trend of de-methylation after fertilization and re-methylation in the ICM of the blastocyst. However, we still identified around 2000 promoters with high levels of H3K27me3 in 2-cell, 8-cell and morula stage embryos (compared to 4000 in MII oocytes and 5000 in ICM), indicating that not all promoter H3K27me3 is lost upon fertilization. We observe a correlation between genes that have high and medium H3K27me3 promoters and genes that are de-repressed upon *Eed* deletion, further arguing that a substantial amount of H3K27me3 is retained at gene promoters in pre-implantation embryos and serves an important function for the repression of genes involved in development and differentiation. However, direct comparison of H3K27me3 levels and differential expression upon PRC2 deletion at the same stages of pre-implantation development is needed to prove our hypothesis.

As mentioned above, we identified a group of differentially expressed genes that are repressed by PRC2 at early pre-implantation stages, raising the possibility that they are regulated by H3K27me3 inherited from the oocyte. As discussed, for a subset of these genes we have shown that inheritance of H3K27me3 from the oocyte mediates maternal allele-specific repression in embryos. However, we also find many genes that show early de-repression upon *Eed* deletion that is not maternal-allele specific, but seems to be bi-allelic. We therefore wonder about the status of the paternal allele at these genes. One possibility is that many more genes are subject to maternal allele-specific repression, but the allele-specificity could not be detected in our analysis due to a lack of SNPs in the hybrids used. Another possibility is that there is substantially more H3K27me3 on the paternal allele than

it has been reported [217, 218]. This might be due to technical limitations of ChIP-seq studies in early embryos, possibly resulting from the limited amounts of input material. The third possibility is that PRC2 somehow contributes to silencing of the genes on the paternal allele in an H3K27me3-independent manner. In order to shed more light on the dynamics and mechanism of PcG-mediated repression on the parental alleles in early embryos, it would be interesting to perform allelic ChIP-seq for the PRC1-dependent H2AK119ub1 mark as well as for components of the PRC1 and PRC2 complexes themselves in germ cells and different stages of early embryos.

In summary, our results demonstrate that PRC2 is an important regulator of gene expression in early embryos. Our study therefore adds more evidence to the emerging view that Polycomb group proteins have much more dynamic functions than just maintaining repression of homeotic genes during embryonic development as it was initially suggested [34].

4.3 Role of PRC1 in the repression of germ line genes in pre-implantation embryos

We analyzed embryos maternally and zygotically deficient for *Ezh2*/*Eed* and *Rnf2* in order to investigate the functions of PRC1 and PRC2 for development and gene regulation in pre-implantation embryos. We discovered that PRC1 represses genes with functions in germ line development. Comparative analysis of our transcriptomic data and published datasets revealed that these genes are regulated by DNA methylation as well as the variant PRC1.6 complex in ESCs [93, 190, 376, 377]. It has been suggested that these two pathways act in parallel for repression of germ line genes in ESCs [190].

It is well known that in early embryos the genome is present in a hypomethylated state [375, 389]. This is due to extensive and rapid DNA de-methylation of the paternal genome as well as passive DNA de-methylation of both maternal and paternal genomes happening along with DNA replication [235, 236, 242, 243, 390]. The hypomethylated state of early embryos resulting from the epigenetic reprogramming events after fertilization is likely required for the acquisition of totipotency, initiation of zygotic gene expression and early lineage specification events. During implantation, global re-methylation, that is believed to contribute to lineage restriction and the loss of pluripotency, takes place mainly in epiblast

cells [391-393]. It has been shown that DNA methylation is primarily targeted to genes expressed in the male and female germ lines and hypomethylation in post-implantation embryos deficient for *Dnmt3b* is associated with massive upregulation of germ line genes. These results indicate that DNA methylation is important to repress the germ line expression program in post-implantation embryos *in vivo* [392]. Since germ line genes are silent at the time of implantation, when re-methylation takes place, it is assumed that DNA methylation does not initiate silencing, but rather acts as a system to maintain and lock these genes in a silent state to prevent deleterious effects caused by their ectopic activation in somatic cells [392, 394]. It remains unknown, what specifically recruits DNA methylation to germ line genes, many of which have CpG-rich promoters, a feature that is usually associated with hypomethylation in the genome [395].

We hypothesize that in the absence of DNA methylation in early embryos, PRC1.6 functions as a backup mechanism to repress genes with germ line functions that need to be kept in a silent state. In order to confirm this hypothesis, we will perform RNA-seq of embryos maternally and zygotically deficient for *Pcgf6*. By comparing the transcriptomic data from these mutant embryos to our existing datasets, we will unravel whether the same set of germ line genes that is upregulated in PRC1/2 DKO embryos, is also de-repressed in embryos lacking the PRC1.6 complex.

It is tempting to speculate that PRC1.6 recruits DNA methylation to germ line genes during the re-methylation that takes place around implantation. In line with this idea is the notion that in ESCs PCGF6-bound genes are enriched for short CGIs that are frequently marked by DNA methylation [93]. It has further been shown that the transcription factor E2F6, a component of the PRC1.6 complex, interacts with DNMT3B, which is the main DNA methyltransferase responsible for *de novo* methylation at implantation, to mediate silencing of germ line genes including *Mael*, *Syce1* and *Tex11* in somatic cells [88, 392, 396].

4.4 Single-cell and single-embryo heterogeneity

An important feature of cells is the non-genetic heterogeneity, which is defined as the property of a population that refers to the phenotypic variability between its genetically identical members [397]. It has been proposed that in early embryos, transcriptional heterogeneity between blastomeres contributes to lineage specification at different

timepoints of pre-implantation development. Single-blastomere RNA-seq of mouse and human embryos has revealed small, but significant differences in gene expression already at the 2-cell stage [270, 271]. Heterogeneity in the nuclear localization of SOX2 and OCT4 among blastomeres of 4-cell embryos has been shown to drive lineage segregation via heterogeneous expression of *Sox21* [272, 273, 276]. At later stages of pre-implantation development, stochastic expression of lineage markers has been shown to precede their lineage-resolved expression [398-400].

When I analysed the developmental progression of wild type and Polycomb mutant embryos, I observed substantial differences in the developmental timing of single embryos. This could result from asynchronous fertilization in natural matings or during *in vitro* fertilization, but also from stochastic transcription events. Since in my laboratory we have also observed heterogeneity of embryos generated by intracytoplasmic sperm injection (ICSI), which are fertilized at the same time, the second possibility certainly contributes to the observed differences in developmental timing. On top of the heterogeneity within wild type embryos, I have observed that embryos deficient for PcG proteins are delayed in their development to varying extent.

In light of these results, I have adapted a single-cell RNA-seq protocol for use in single pre-implantation embryos [360]. This approach has proven to be very successful for the transcriptomic analysis of early embryos and offers multiple advantages over bulk RNA-seq techniques. First of all, we were able to genotype the sequenced embryos from RNA-seq data and to exclude heterozygous escapers that did not delete the floxed alleles due to inefficient cre recombination. Further, we were able to distinguish between male and female embryos and discovered that some of the transcriptional effects upon PRC2 deletion are sex-specific. Notably, it has been shown that male embryos develop more quickly than females and a substantial number of genes, including autosomal genes, differs in their expression level between the sexes at the blastocyst stage in mouse and bovine embryos [335, 342-344]. Although clearly more analysis is needed to understand the sex-specific differences in PRC2 deficient embryos, our results raise the possibility that PRC2 is involved in setting up the early asymmetry between the sexes.

Importantly, we developed a method to assign a pseudotime to each embryo in order to assess the developmental timing of individual embryos and to be able to take the differences into account when performing differential gene expression analysis. Using the

pseudotiming approach, we were able to confirm the developmental delay of mutant embryos and correct for it in order to distinguish true PRC2-dependent transcriptional changes from those that arise from differential timing of embryos.

In PRC1/2 DKO embryos, we identified a group of genes specifically upregulated in a subset of embryos. The combination of single-embryo RNA-seq and imaging of each embryo prior to library preparation allowed us to identify these embryos as morphologically abnormal and correlate their morphological state to transcriptional upregulation of apoptosis-related genes.

As mentioned above, there is not only heterogeneity between embryos, but also between different blastomeres of the same embryo. Therefore, it would be interesting to perform single-cell RNA-seq of pre-implantation embryos at different stages. This would be particularly useful to further dissect the role of Polycomb proteins in the first lineage segregation events in early embryos. In PRC2 deficient embryos, we observe maternal allele-specific upregulation of *Sox21*, which has been shown to be heterogeneously expressed in 4-cell embryos and contribute to lineage specification by opposing the expression of the TE determinant *Cdx2* [276]. Interestingly, we observe a decrease in CDX2 by IF staining in PRC2 deficient blastocysts, raising the possibility that PRC2 controls the first lineage specification event via regulation of *Sox21*. Therefore, it would be highly interesting to perform RNA-seq of single blastomeres of 4-cell and later embryos, to compare the expression of *Sox21, Cdx2* and other genes involved in lineage choices between blastomeres of the same embryo. Further, this experiment would answer whether the transcriptional effects of *Eed* deletion in blastocysts are similar in cells belonging to ICM and TE lineages or whether different subsets of genes are misregulated in the two lineages. In order to further validate our hypothesis, we are planning to perform IF staining for SOX21 and CDX2 at different stages of pre-implantation development, in *Eed* deficient and control embryos. To test whether deficiency for PRC2 biases cells towards an embryonic fate, it would be interesting to generate chimaeric embryos by aggregation of 2-cell blastomeres, out of which one is deficient for PRC2 and one is fluorescently labelled. This experiment would allow to track whether PRC2 deficient blastomeres contribute more likely to the ICM than wildtype blastomeres.

References

1. Waddington, C.H., *The epigenotype.* Endeavour, 1942: p. 18-20
2. Waddington, C.H., *The Strategy of the Genes; a Discussion of Some Aspects of Theoretical Biology*. 1957, London: Allen & Unwin.
3. Bird, A., *Perceptions of epigenetics.* Nature, 2007. **447**(7143): p. 396-8.
4. Kornberg, R.D. and J.O. Thomas, *Chromatin structure; oligomers of the histones.* Science, 1974. **184**(4139): p. 865-8.
5. Thoma, F. and T. Koller, *Influence of histone H1 on chromatin structure.* Cell, 1977. **12**(1): p. 101-7.
6. Thoma, F., T. Koller, and A. Klug, *Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin.* J Cell Biol, 1979. **83**(2 Pt 1): p. 403-27.
7. Knezetic, J.A. and D.S. Luse, *The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II in vitro.* Cell, 1986. **45**(1): p. 95-104.
8. Lorch, Y., J.W. LaPointe, and R.D. Kornberg, *Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones.* Cell, 1987. **49**(2): p. 203-10.
9. Han, M. and M. Grunstein, *Nucleosome loss activates yeast downstream promoters in vivo.* Cell, 1988. **55**(6): p. 1137-45.
10. Kayne, P.S., et al., *Extremely conserved histone H4 N terminus is dispensable for growth but essential for repressing the silent mating loci in yeast.* Cell, 1988. **55**(1): p. 27-39.
11. Kouzarides, T., *Chromatin modifications and their function.* Cell, 2007. **128**(4): p. 693-705.
12. Li, B., M. Carey, and J.L. Workman, *The role of chromatin during transcription.* Cell, 2007. **128**(4): p. 707-19.
13. Strahl, B.D. and C.D. Allis, *The language of covalent histone modifications.* Nature, 2000. **403**(6765): p. 41-5.
14. Jenuwein, T. and C.D. Allis, *Translating the histone code.* Science, 2001. **293**(5532): p. 1074-80.
15. Clapier, C.R., et al., *Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes.* Nat Rev Mol Cell Biol, 2017. **18**(7): p. 407-422.
16. Kornberg, R.D. and Y. Lorch, *Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome.* Cell, 1999. **98**(3): p. 285-94.
17. Kamakaka, R.T. and S. Biggins, *Histone variants: deviants?* Genes Dev, 2005. **19**(3): p. 295-310.
18. Ahmad, K. and S. Henikoff, *The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly.* Mol Cell, 2002. **9**(6): p. 1191-200.
19. Goldberg, A.D., et al., *Distinct factors control histone variant H3.3 localization at specific genomic regions.* Cell, 2010. **140**(5): p. 678-91.
20. Ricketts, M.D., et al., *Ubinuclein-1 confers histone H3.3-specific-binding by the HIRA histone chaperone complex.* Nat Commun, 2015. **6**: p. 7711.
21. Tagami, H., et al., *Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis.* Cell, 2004. **116**(1): p. 51-61.
22. Ray-Gallet, D., et al., *HIRA is critical for a nucleosome assembly pathway independent of DNA synthesis.* Mol Cell, 2002. **9**(5): p. 1091-100.

23. Elsasser, S.J., et al., *DAXX envelops a histone H3.3-H4 dimer for H3.3-specific recognition.* Nature, 2012. **491**(7425): p. 560-5.
24. Liu, C.P., et al., *Structure of the variant histone H3.3-H4 heterodimer in complex with its chaperone DAXX.* Nat Struct Mol Biol, 2012. **19**(12): p. 1287-92.
25. Braunschweig, U., et al., *Histone H1 binding is inhibited by histone variant H3.3.* EMBO J, 2009. **28**(23): p. 3635-45.
26. McKittrick, E., et al., *Histone H3.3 is enriched in covalent modifications associated with active chromatin.* Proc Natl Acad Sci U S A, 2004. **101**(6): p. 1525-30.
27. Chow, C.M., et al., *Variant histone H3.3 marks promoters of transcriptionally active genes during mammalian cell division.* EMBO Rep, 2005. **6**(4): p. 354-60.
28. Mito, Y., J.G. Henikoff, and S. Henikoff, *Genome-scale profiling of histone H3.3 replacement patterns.* Nat Genet, 2005. **37**(10): p. 1090-7.
29. Tamura, T., et al., *Inducible deposition of the histone variant H3.3 in interferon-stimulated genes.* J Biol Chem, 2009. **284**(18): p. 12217-25.
30. 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, 2005. **122**(9): p. 1008-22.
31. Torres-Padilla, M.E., et al., *Dynamic distribution of the replacement histone variant H3.3 in the mouse oocyte and preimplantation embryos.* Int J Dev Biol, 2006. **50**(5): p. 455-61.
32. Santenard, A., et al., *Heterochromatin formation in the mouse embryo requires critical residues of the histone variant H3.3.* Nat Cell Biol, 2010. **12**(9): p. 853-62.
33. Schwartz, Y.B. and V. Pirrotta, *A new world of Polycombs: unexpected partnerships and emerging functions.* Nat Rev Genet, 2013. **14**(12): p. 853-64.
34. Lewis, E.B., *A gene complex controlling segmentation in Drosophila.* Nature, 1978. **276**(5688): p. 565-70.
35. Kennison, J.A., *The Polycomb and trithorax group proteins of Drosophila: trans-regulators of homeotic gene function.* Annu Rev Genet, 1995. **29**: p. 289-303.
36. Klymenko, T. and J. Muller, *The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins.* EMBO Rep, 2004. **5**(4): p. 373-7.
37. Papp, B. and J. Muller, *Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins.* Genes Dev, 2006. **20**(15): p. 2041-54.
38. Core, N., et al., *Altered cellular proliferation and mesoderm patterning in Polycomb-M33-deficient mice.* Development, 1997. **124**(3): p. 721-9.
39. Akasaka, T., et al., *A role for mel-18, a Polycomb group-related vertebrate gene, during theanteroposterior specification of the axial skeleton.* Development, 1996. **122**(5): p. 1513-22.
40. del Mar Lorente, M., et al., *Loss- and gain-of-function mutations show a polycomb group function for Ring1A in mice.* Development, 2000. **127**(23): p. 5093-100.
41. 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, 1994. **8**(7): p. 757-69.
42. Cao, R., et al., *Role of histone H3 lysine 27 methylation in Polycomb-group silencing.* Science, 2002. **298**(5595): p. 1039-43.

43. Czermin, B., et al., *Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites.* Cell, 2002. **111**(2): p. 185-96.
44. Cao, R. and Y. Zhang, *SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex.* Mol Cell, 2004. **15**(1): p. 57-67.
45. Pasini, D., et al., *Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity.* EMBO J, 2004. **23**(20): p. 4061-71.
46. Nekrasov, M., B. Wild, and J. Muller, *Nucleosome binding and histone methyltransferase activity of Drosophila PRC2.* EMBO Rep, 2005. **6**(4): p. 348-53.
47. Nekrasov, M., et al., *Pcl-PRC2 is needed to generate high levels of H3-K27 trimethylation at Polycomb target genes.* EMBO J, 2007. **26**(18): p. 4078-88.
48. Kuzmichev, A., et al., *Composition and histone substrates of polycomb repressive group complexes change during cellular differentiation.* Proc Natl Acad Sci U S A, 2005. **102**(6): p. 1859-64.
49. Kuzmichev, A., et al., *Different EZH2-containing complexes target methylation of histone H1 or nucleosomal histone H3.* Mol Cell, 2004. **14**(2): p. 183-93.
50. Ketel, C.S., et al., *Subunit contributions to histone methyltransferase activities of fly and worm polycomb group complexes.* Mol Cell Biol, 2005. **25**(16): p. 6857-68.
51. Schmitges, F.W., et al., *Histone methylation by PRC2 is inhibited by active chromatin marks.* Mol Cell, 2011. **42**(3): p. 330-41.
52. Han, Z., et al., *Structural basis of EZH2 recognition by EED.* Structure, 2007. **15**(10): p. 1306-15.
53. Kim, H., K. Kang, and J. Kim, *AEBP2 as a potential targeting protein for Polycomb Repression Complex PRC2.* Nucleic Acids Res, 2009. **37**(9): p. 2940-50.
54. Walker, E., et al., *Polycomb-like 2 associates with PRC2 and regulates transcriptional networks during mouse embryonic stem cell self-renewal and differentiation.* Cell Stem Cell, 2010. **6**(2): p. 153-66.
55. Sarma, K., et al., *Ezh2 requires PHF1 to efficiently catalyze H3 lysine 27 trimethylation in vivo.* Mol Cell Biol, 2008. **28**(8): p. 2718-31.
56. Savla, U., et al., *Recruitment of Drosophila Polycomb-group proteins by Polycomblike, a component of a novel protein complex in larvae.* Development, 2008. **135**(5): p. 813-7.
57. Li, H., et al., *Polycomb-like proteins link the PRC2 complex to CpG islands.* Nature, 2017. **549**(7671): p. 287-291.
58. Klose, R.J., E.M. Kallin, and Y. Zhang, *JmjC-domain-containing proteins and histone demethylation.* Nat Rev Genet, 2006. **7**(9): p. 715-27.
59. Li, G., et al., *Jarid2 and PRC2, partners in regulating gene expression.* Genes Dev, 2010. **24**(4): p. 368-80.
60. Son, J., et al., *Nucleosome-binding activities within JARID2 and EZH1 regulate the function of PRC2 on chromatin.* Genes Dev, 2013. **27**(24): p. 2663-77.
61. Sanulli, S., et al., *Jarid2 Methylation via the PRC2 Complex Regulates H3K27me3 Deposition during Cell Differentiation.* Mol Cell, 2015. **57**(5): p. 769-83.
62. Landeira, D., et al., *Jarid2 Coordinates Nanog Expression and PCP/Wnt Signaling Required for Efficient ESC Differentiation and Early Embryo Development.* Cell Rep, 2015. **12**(4): p. 573-86.

63. Pasini, D., et al., *JARID2 regulates binding of the Polycomb repressive complex 2 to target genes in ES cells.* Nature, 2010. **464**(7286): p. 306-10.
64. Peng, J.C., et al., *Jarid2/Jumonji coordinates control of PRC2 enzymatic activity and target gene occupancy in pluripotent cells.* Cell, 2009. **139**(7): p. 1290-302.
65. Shen, X., et al., *Jumonji modulates polycomb activity and self-renewal versus differentiation of stem cells.* Cell, 2009. **139**(7): p. 1303-14.
66. Cooper, S., et al., *Jarid2 binds mono-ubiquitylated H2A lysine 119 to mediate crosstalk between Polycomb complexes PRC1 and PRC2.* Nat Commun, 2016. **7**: p. 13661.
67. Oksuz, O., et al., *Capturing the Onset of PRC2-Mediated Repressive Domain Formation.* Mol Cell, 2018. **70**(6): p. 1149-1162 e5.
68. Pengelly, A.R., et al., *A histone mutant reproduces the phenotype caused by loss of histone-modifying factor Polycomb.* Science, 2013. **339**(6120): p. 698-9.
69. Kahn, T.G., et al., *Interdependence of PRC1 and PRC2 for recruitment to Polycomb Response Elements.* Nucleic Acids Res, 2016. **44**(21): p. 10132-10149.
70. Margueron, R., et al., *Role of the polycomb protein EED in the propagation of repressive histone marks.* Nature, 2009. **461**(7265): p. 762-7.
71. Shao, Z., et al., *Stabilization of chromatin structure by PRC1, a Polycomb complex.* Cell, 1999. **98**(1): p. 37-46.
72. King, I.F., N.J. Francis, and R.E. Kingston, *Native and recombinant polycomb group complexes establish a selective block to template accessibility to repress transcription in vitro.* Mol Cell Biol, 2002. **22**(22): p. 7919-28.
73. Levine, S.S., et al., *The core of the polycomb repressive complex is compositionally and functionally conserved in flies and humans.* Mol Cell Biol, 2002. **22**(17): p. 6070-8.
74. Cao, R., Y. Tsukada, and Y. Zhang, *Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing.* Mol Cell, 2005. **20**(6): p. 845-54.
75. de Napoles, M., et al., *Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation.* Dev Cell, 2004. **7**(5): p. 663-76.
76. Wang, H., et al., *Role of histone H2A ubiquitination in Polycomb silencing.* Nature, 2004. **431**(7010): p. 873-8.
77. Buchwald, G., et al., *Structure and E3-ligase activity of the Ring-Ring complex of polycomb proteins Bmi1 and Ring1b.* EMBO J, 2006. **25**(11): p. 2465-74.
78. Li, Z., et al., *Structure of a Bmi-1-Ring1B polycomb group ubiquitin ligase complex.* J Biol Chem, 2006. **281**(29): p. 20643-9.
79. Elderkin, S., et al., *A phosphorylated form of Mel-18 targets the Ring1B histone H2A ubiquitin ligase to chromatin.* Mol Cell, 2007. **28**(1): p. 107-20.
80. Bernstein, E., et al., *Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin.* Mol Cell Biol, 2006. **26**(7): p. 2560-9.
81. Kaustov, L., et al., *Recognition and specificity determinants of the human cbx chromodomains.* J Biol Chem, 2011. **286**(1): p. 521-9.
82. Tardat, M., et al., *Cbx2 targets PRC1 to constitutive heterochromatin in mouse zygotes in a parent-of-origin-dependent manner.* Mol Cell, 2015. **58**(1): p. 157-71.
83. Isono, K., et al., *SAM domain polymerization links subnuclear clustering of PRC1 to gene silencing.* Dev Cell, 2013. **26**(6): p. 565-77.

84. Gao, Z., et al., *PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes.* Mol Cell, 2012. **45**(3): p. 344-56.
85. Tavares, L., et al., *RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3.* Cell, 2012. **148**(4): p. 664-78.
86. Farcas, A.M., et al., *KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands.* Elife, 2012. **1**: p. e00205.
87. Wu, X., J.V. Johansen, and K. Helin, *Fbxl10/Kdm2b recruits polycomb repressive complex 1 to CpG islands and regulates H2A ubiquitylation.* Mol Cell, 2013. **49**(6): p. 1134-46.
88. Ogawa, H., et al., *A complex with chromatin modifiers that occupies E2F- and Myc-responsive genes in G0 cells.* Science, 2002. **296**(5570): p. 1132-6.
89. Hauri, S., et al., *A High-Density Map for Navigating the Human Polycomb Complexome.* Cell Rep, 2016. **17**(2): p. 583-595.
90. Kloet, S.L., et al., *The dynamic interactome and genomic targets of Polycomb complexes during stem-cell differentiation.* Nat Struct Mol Biol, 2016. **23**(7): p. 682-690.
91. Trojer, P., et al., *L3MBTL2 protein acts in concert with PcG protein-mediated monoubiquitination of H2A to establish a repressive chromatin structure.* Mol Cell, 2011. **42**(4): p. 438-50.
92. Qin, J., et al., *The polycomb group protein L3mbtl2 assembles an atypical PRC1-family complex that is essential in pluripotent stem cells and early development.* Cell Stem Cell, 2012. **11**(3): p. 319-32.
93. Endoh, M., et al., *PCGF6-PRC1 suppresses premature differentiation of mouse embryonic stem cells by regulating germ cell-related genes.* Elife, 2017. **6**.
94. 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, 2007. **9**(12): p. 1428-35.
95. Zhou, W., et al., *Histone H2A monoubiquitination represses transcription by inhibiting RNA polymerase II transcriptional elongation.* Mol Cell, 2008. **29**(1): p. 69-80.
96. Nakagawa, T., et al., *Deubiquitylation of histone H2A activates transcriptional initiation via trans-histone cross-talk with H3K4 di- and trimethylation.* Genes Dev, 2008. **22**(1): p. 37-49.
97. Pengelly, A.R., et al., *Transcriptional repression by PRC1 in the absence of H2A monoubiquitylation.* Genes Dev, 2015. **29**(14): p. 1487-92.
98. Illingworth, R.S., et al., *The E3 ubiquitin ligase activity of RING1B is not essential for early mouse development.* Genes Dev, 2015. **29**(18): p. 1897-902.
99. Endoh, M., et al., *Histone H2A mono-ubiquitination is a crucial step to mediate PRC1-dependent repression of developmental genes to maintain ES cell identity.* PLoS Genet, 2012. **8**(7): p. e1002774.
100. Kundu, S., et al., *Polycomb Repressive Complex 1 Generates Discrete Compacted Domains that Change during Differentiation.* Mol Cell, 2017. **65**(3): p. 432-446 e5.
101. Kahn, T.G., et al., *Polycomb complexes and the propagation of the methylation mark at the Drosophila ubx gene.* J Biol Chem, 2006. **281**(39): p. 29064-75.
102. Chan, C.S., L. Rastelli, and V. Pirrotta, *A Polycomb response element in the Ubx gene that determines an epigenetically inherited state of repression.* EMBO J, 1994. **13**(11): p. 2553-64.

103. Simon, J., et al., *Elements of the Drosophila bithorax complex that mediate repression by Polycomb group products.* Dev Biol, 1993. **158**(1): p. 131-44.
104. Horard, B., et al., *Structure of a polycomb response element and in vitro binding of polycomb group complexes containing GAGA factor.* Mol Cell Biol, 2000. **20**(9): p. 3187-97.
105. Muller, J. and J.A. Kassis, *Polycomb response elements and targeting of Polycomb group proteins in Drosophila.* Curr Opin Genet Dev, 2006. **16**(5): p. 476-84.
106. Kharchenko, P.V., et al., *Comprehensive analysis of the chromatin landscape in Drosophila melanogaster.* Nature, 2011. **471**(7339): p. 480-5.
107. Hodgson, J.W., B. Argiropoulos, and H.W. Brock, *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, 2001. **21**(14): p. 4528-43.
108. Dejardin, J., et al., *Recruitment of Drosophila Polycomb group proteins to chromatin by DSP1.* Nature, 2005. **434**(7032): p. 533-8.
109. Klymenko, T., et al., *A Polycomb group protein complex with sequence-specific DNA-binding and selective methyl-lysine-binding activities.* Genes Dev, 2006. **20**(9): p. 1110-22.
110. Schuettengruber, B., et al., *Functional anatomy of polycomb and trithorax chromatin landscapes in Drosophila embryos.* PLoS Biol, 2009. **7**(1): p. e13.
111. Sing, A., et al., *A vertebrate Polycomb response element governs segmentation of the posterior hindbrain.* Cell, 2009. **138**(5): p. 885-97.
112. Woo, C.J., et al., *A region of the human HOXD cluster that confers polycomb-group responsiveness.* Cell, 2010. **140**(1): p. 99-110.
113. Herranz, N., et al., *Polycomb complex 2 is required for E-cadherin repression by the Snail1 transcription factor.* Mol Cell Biol, 2008. **28**(15): p. 4772-81.
114. Dietrich, N., et al., *REST-mediated recruitment of polycomb repressor complexes in mammalian cells.* PLoS Genet, 2012. **8**(3): p. e1002494.
115. Ren, X. and T.K. Kerppola, *REST interacts with Cbx proteins and regulates polycomb repressive complex 1 occupancy at RE1 elements.* Mol Cell Biol, 2011. **31**(10): p. 2100-10.
116. Arnold, P., et al., *Modeling of epigenome dynamics identifies transcription factors that mediate Polycomb targeting.* Genome Res, 2013. **23**(1): p. 60-73.
117. Yu, M., et al., *Direct recruitment of polycomb repressive complex 1 to chromatin by core binding transcription factors.* Mol Cell, 2012. **45**(3): p. 330-43.
118. Rinn, J.L., et al., *Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs.* Cell, 2007. **129**(7): p. 1311-23.
119. Tsai, M.C., et al., *Long noncoding RNA as modular scaffold of histone modification complexes.* Science, 2010. **329**(5992): p. 689-93.
120. Zhao, J., et al., *Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome.* Science, 2008. **322**(5902): p. 750-6.
121. Davidovich, C., et al., *Promiscuous RNA binding by Polycomb repressive complex 2.* Nat Struct Mol Biol, 2013. **20**(11): p. 1250-7.
122. Davidovich, C., et al., *Toward a consensus on the binding specificity and promiscuity of PRC2 for RNA.* Mol Cell, 2015. **57**(3): p. 552-8.
123. Kanhere, A., et al., *Short RNAs are transcribed from repressed polycomb target genes and interact with polycomb repressive complex-2.* Mol Cell, 2010. **38**(5): p. 675-88.

124. Kaneko, S., et al., *PRC2 binds active promoters and contacts nascent RNAs in embryonic stem cells.* Nat Struct Mol Biol, 2013. **20**(11): p. 1258-64.
125. Cifuentes-Rojas, C., et al., *Regulatory interactions between RNA and polycomb repressive complex 2.* Mol Cell, 2014. **55**(2): p. 171-85.
126. Kaneko, S., et al., *Nascent RNA interaction keeps PRC2 activity poised and in check.* Genes Dev, 2014. **28**(18): p. 1983-8.
127. Mendenhall, E.M., et al., *GC-rich sequence elements recruit PRC2 in mammalian ES cells.* PLoS Genet, 2010. **6**(12): p. e1001244.
128. Riising, E.M., et al., *Gene silencing triggers polycomb repressive complex 2 recruitment to CpG islands genome wide.* Mol Cell, 2014. **55**(3): p. 347-60.
129. Lynch, M.D., et al., *An interspecies analysis reveals a key role for unmethylated CpG dinucleotides in vertebrate Polycomb complex recruitment.* EMBO J, 2012. **31**(2): p. 317-29.
130. He, J., et al., *Kdm2b maintains murine embryonic stem cell status by recruiting PRC1 complex to CpG islands of developmental genes.* Nat Cell Biol, 2013. **15**(4): p. 373-84.
131. Boulard, M., J.R. Edwards, and T.H. Bestor, *Abnormal X chromosome inactivation and sex-specific gene dysregulation after ablation of FBXL10.* Epigenetics Chromatin, 2016. **9**: p. 22.
132. Ku, M., et al., *Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains.* PLoS Genet, 2008. **4**(10): p. e1000242.
133. Bernstein, B.E., et al., *A bivalent chromatin structure marks key developmental genes in embryonic stem cells.* Cell, 2006. **125**(2): p. 315-26.
134. Klose, R.J., et al., *Chromatin sampling--an emerging perspective on targeting polycomb repressor proteins.* PLoS Genet, 2013. **9**(8): p. e1003717.
135. Ballare, C., et al., *Phf19 links methylated Lys36 of histone H3 to regulation of Polycomb activity.* Nat Struct Mol Biol, 2012. **19**(12): p. 1257-65.
136. Mozzetta, C., J. Pontis, and S. Ait-Si-Ali, *Functional Crosstalk Between Lysine Methyltransferases on Histone Substrates: The Case of G9A/GLP and Polycomb Repressive Complex 2.* Antioxid Redox Signal, 2015. **22**(16): p. 1365-81.
137. Fischle, W., et al., *Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains.* Genes Dev, 2003. **17**(15): p. 1870-81.
138. Simon, J.A. and R.E. Kingston, *Occupying chromatin: Polycomb mechanisms for getting to genomic targets, stopping transcriptional traffic, and staying put.* Mol Cell, 2013. **49**(5): p. 808-24.
139. Blackledge, N.P., N.R. Rose, and R.J. Klose, *Targeting Polycomb systems to regulate gene expression: modifications to a complex story.* Nat Rev Mol Cell Biol, 2015. **16**(11): p. 643-9.
140. Blackledge, N.P., et al., *Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation.* Cell, 2014. **157**(6): p. 1445-59.
141. Cooper, S., et al., *Targeting polycomb to pericentric heterochromatin in embryonic stem cells reveals a role for H2AK119u1 in PRC2 recruitment.* Cell Rep, 2014. **7**(5): p. 1456-70.
142. Kalb, R., et al., *Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression.* Nat Struct Mol Biol, 2014. **21**(6): p. 569-71.
143. Lau, M.S., et al., *Mutation of a nucleosome compaction region disrupts Polycomb-mediated axial patterning.* Science, 2017. **355**(6329): p. 1081-1084.

144. Margueron, R., et al., *Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms.* Mol Cell, 2008. **32**(4): p. 503-18.
145. Kadoch, C., et al., *Dynamics of BAF-Polycomb complex opposition on heterochromatin in normal and oncogenic states.* Nat Genet, 2017. **49**(2): p. 213-222.
146. Dixon, J.R., et al., *Topological domains in mammalian genomes identified by analysis of chromatin interactions.* Nature, 2012. **485**(7398): p. 376-80.
147. Nora, E.P., et al., *Spatial partitioning of the regulatory landscape of the X-inactivation centre.* Nature, 2012. **485**(7398): p. 381-5.
148. Sexton, T., et al., *Three-dimensional folding and functional organization principles of the Drosophila genome.* Cell, 2012. **148**(3): p. 458-72.
149. Bonev, B. and G. Cavalli, *Organization and function of the 3D genome.* Nat Rev Genet, 2016. **17**(12): p. 772.
150. Buchenau, P., et al., *The distribution of polycomb-group proteins during cell division and development in Drosophila embryos: impact on models for silencing.* J Cell Biol, 1998. **141**(2): p. 469-81.
151. Saurin, A.J., et al., *The human polycomb group complex associates with pericentromeric heterochromatin to form a novel nuclear domain.* J Cell Biol, 1998. **142**(4): p. 887-98.
152. Grimaud, C., et al., *RNAi components are required for nuclear clustering of Polycomb group response elements.* Cell, 2006. **124**(5): p. 957-71.
153. Terranova, R., et al., *Polycomb group proteins Ezh2 and Rnf2 direct genomic contraction and imprinted repression in early mouse embryos.* Dev Cell, 2008. **15**(5): p. 668-79.
154. Wani, A.H., et al., *Chromatin topology is coupled to Polycomb group protein subnuclear organization.* Nat Commun, 2016. **7**: p. 10291.
155. Noordermeer, D., et al., *The dynamic architecture of Hox gene clusters.* Science, 2011. **334**(6053): p. 222-5.
156. Narendra, V., et al., *CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation.* Science, 2015. **347**(6225): p. 1017-21.
157. Narendra, V., et al., *CTCF-mediated topological boundaries during development foster appropriate gene regulation.* Genes Dev, 2016. **30**(24): p. 2657-2662.
158. Schoenfelder, S., et al., *Polycomb repressive complex PRC1 spatially constrains the mouse embryonic stem cell genome.* Nat Genet, 2015. **47**(10): p. 1179-1186.
159. Vieux-Rochas, M., et al., *Clustering of mammalian Hox genes with other H3K27me3 targets within an active nuclear domain.* Proc Natl Acad Sci U S A, 2015. **112**(15): p. 4672-7.
160. Joshi, O., et al., *Dynamic Reorganization of Extremely Long-Range Promoter-Promoter Interactions between Two States of Pluripotency.* Cell Stem Cell, 2015. **17**(6): p. 748-757.
161. Boettiger, A.N., et al., *Super-resolution imaging reveals distinct chromatin folding for different epigenetic states.* Nature, 2016. **529**(7586): p. 418-22.
162. Denholtz, M., et al., *Long-range chromatin contacts in embryonic stem cells reveal a role for pluripotency factors and polycomb proteins in genome organization.* Cell Stem Cell, 2013. **13**(5): p. 602-16.
163. Tie, F., et al., *Polycomb inhibits histone acetylation by CBP by binding directly to its catalytic domain.* Proc Natl Acad Sci U S A, 2016. **113**(6): p. E744-53.

164. Ferrari, K.J., et al., *Polycomb-dependent H3K27me1 and H3K27me2 regulate active transcription and enhancer fidelity.* Mol Cell, 2014. **53**(1): p. 49-62.
165. Lee, H.G., et al., *Genome-wide activities of Polycomb complexes control pervasive transcription.* Genome Res, 2015. **25**(8): p. 1170-81.
166. Faust, C., et al., *The eed mutation disrupts anterior mesoderm production in mice.* Development, 1995. **121**(2): p. 273-85.
167. O'Carroll, D., et al., *The polycomb-group gene Ezh2 is required for early mouse development.* Mol Cell Biol, 2001. **21**(13): p. 4330-6.
168. Takeuchi, T., et al., *Gene trap capture of a novel mouse gene, jumonji, required for neural tube formation.* Genes Dev, 1995. **9**(10): p. 1211-22.
169. Wang, S., et al., *Chick Pcl2 regulates the left-right asymmetry by repressing Shh expression in Hensen's node.* Development, 2004. **131**(17): p. 4381-91.
170. Wang, S., et al., *Polycomblike-2-deficient mice exhibit normal left-right asymmetry.* Dev Dyn, 2007. **236**(3): p. 853-61.
171. Voncken, J.W., et al., *Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition.* Proc Natl Acad Sci U S A, 2003. **100**(5): p. 2468-73.
172. Suzuki, M., et al., *Involvement of the Polycomb-group gene Ring1B in the specification of the anterior-posterior axis in mice.* Development, 2002. **129**(18): p. 4171-83.
173. Posfai, E., et al., *Polycomb function during oogenesis is required for mouse embryonic development.* Genes Dev, 2012. **26**(9): p. 920-32.
174. Pirity, M.K., J. Locker, and N. Schreiber-Agus, *Rybp/DEDAF is required for early postimplantation and for central nervous system development.* Mol Cell Biol, 2005. **25**(16): p. 7193-202.
175. Fukuda, T., et al., *Fbxl10/Kdm2b deficiency accelerates neural progenitor cell death and leads to exencephaly.* Mol Cell Neurosci, 2011. **46**(3): p. 614-24.
176. Boulard, M., J.R. Edwards, and T.H. Bestor, *FBXL10 protects Polycomb-bound genes from hypermethylation.* Nat Genet, 2015. **47**(5): p. 479-85.
177. Liu, B., et al., *Cbx4 regulates the proliferation of thymic epithelial cells and thymus function.* Development, 2013. **140**(4): p. 780-8.
178. Forzati, F., et al., *CBX7 is a tumor suppressor in mice and humans.* J Clin Invest, 2012. **122**(2): p. 612-23.
179. Akasaka, T., et al., *Mice doubly deficient for the Polycomb Group genes Mel18 and Bmi1 reveal synergy and requirement for maintenance but not initiation of Hox gene expression.* Development, 2001. **128**(9): p. 1587-97.
180. Azuara, V., et al., *Chromatin signatures of pluripotent cell lines.* Nat Cell Biol, 2006. **8**(5): p. 532-8.
181. Denissov, S., et al., *Mll2 is required for H3K4 trimethylation on bivalent promoters in embryonic stem cells, whereas Mll1 is redundant.* Development, 2014. **141**(3): p. 526-37.
182. Leeb, M., et al., *Polycomb complexes act redundantly to repress genomic repeats and genes.* Genes Dev, 2010. **24**(3): p. 265-76.
183. Leeb, M. and A. Wutz, *Ring1B is crucial for the regulation of developmental control genes and PRC1 proteins but not X inactivation in embryonic cells.* J Cell Biol, 2007. **178**(2): p. 219-29.

184. Chamberlain, S.J., D. Yee, and T. Magnuson, *Polycomb repressive complex 2 is dispensable for maintenance of embryonic stem cell pluripotency.* Stem Cells, 2008. **26**(6): p. 1496-505.
185. Endoh, M., et al., *Polycomb group proteins Ring1A/B are functionally linked to the core transcriptional regulatory circuitry to maintain ES cell identity.* Development, 2008. **135**(8): p. 1513-24.
186. Hunkapiller, J., et al., *Polycomb-like 3 promotes polycomb repressive complex 2 binding to CpG islands and embryonic stem cell self-renewal.* PLoS Genet, 2012. **8**(3): p. e1002576.
187. Morey, L., et al., *Polycomb Regulates Mesoderm Cell Fate-Specification in Embryonic Stem Cells through Activation and Repression Mechanisms.* Cell Stem Cell, 2015. **17**(3): p. 300-15.
188. Zdzieblo, D., et al., *Pcgf6, a polycomb group protein, regulates mesodermal lineage differentiation in murine ESCs and functions in iPS reprogramming.* Stem Cells, 2014. **32**(12): p. 3112-25.
189. Morey, L., et al., *Nonoverlapping functions of the Polycomb group Cbx family of proteins in embryonic stem cells.* Cell Stem Cell, 2012. **10**(1): p. 47-62.
190. Hisada, K., et al., *RYBP represses endogenous retroviruses and preimplantation- and germ line-specific genes in mouse embryonic stem cells.* Mol Cell Biol, 2012. **32**(6): p. 1139-49.
191. Morin-Kensicki, E.M., et al., *Cell and tissue requirements for the gene eed during mouse gastrulation and organogenesis.* Genesis, 2001. **31**(4): p. 142-6.
192. Shen, X., et al., *EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency.* Mol Cell, 2008. **32**(4): p. 491-502.
193. Ezhkova, E., et al., *EZH1 and EZH2 cogovern histone H3K27 trimethylation and are essential for hair follicle homeostasis and wound repair.* Genes Dev, 2011. **25**(5): p. 485-98.
194. Pasini, D., et al., *The polycomb group protein Suz12 is required for embryonic stem cell differentiation.* Mol Cell Biol, 2007. **27**(10): p. 3769-79.
195. Schuettengruber, B., et al., *Genome Regulation by Polycomb and Trithorax: 70 Years and Counting.* Cell, 2017. **171**(1): p. 34-57.
196. O'Loghlen, A., et al., *MicroRNA regulation of Cbx7 mediates a switch of Polycomb orthologs during ESC differentiation.* Cell Stem Cell, 2012. **10**(1): p. 33-46.
197. Oliviero, G., et al., *Dynamic Protein Interactions of the Polycomb Repressive Complex 2 during Differentiation of Pluripotent Cells.* Mol Cell Proteomics, 2016. **15**(11): p. 3450-3460.
198. Heard, E. and R.A. Martienssen, *Transgenerational epigenetic inheritance: myths and mechanisms.* Cell, 2014. **157**(1): p. 95-109.
199. Morgan, H.D., et al., *Epigenetic inheritance at the agouti locus in the mouse.* Nat Genet, 1999. **23**(3): p. 314-8.
200. Rakyan, V.K., et al., *Transgenerational inheritance of epigenetic states at the murine Axin(Fu) allele occurs after maternal and paternal transmission.* Proc Natl Acad Sci U S A, 2003. **100**(5): p. 2538-43.
201. Duhl, D.M., et al., *Neomorphic agouti mutations in obese yellow mice.* Nat Genet, 1994. **8**(1): p. 59-65.

202. Wolff, G.L., *Influence of maternal phenotype on metabolic differentiation of agouti locus mutants in the mouse.* Genetics, 1978. **88**(3): p. 529-39.
203. Dolinoy, D.C., et al., *Variable histone modifications at the A(vy) metastable epiallele.* Epigenetics, 2010. **5**(7): p. 637-44.
204. Blewitt, M.E., et al., *Dynamic reprogramming of DNA methylation at an epigenetically sensitive allele in mice.* PLoS Genet, 2006. **2**(4): p. e49.
205. Carone, B.R., et al., *Paternally induced transgenerational environmental reprogramming of metabolic gene expression in mammals.* Cell, 2010. **143**(7): p. 1084-96.
206. Radford, E.J., et al., *In utero effects. In utero undernourishment perturbs the adult sperm methylome and intergenerational metabolism.* Science, 2014. **345**(6198): p. 1255903.
207. Siklenka, K., et al., *Disruption of histone methylation in developing sperm impairs offspring health transgenerationally.* Science, 2015. **350**(6261): p. aab2006.
208. Daxinger, L. and E. Whitelaw, *Understanding transgenerational epigenetic inheritance via the gametes in mammals.* Nat Rev Genet, 2012. **13**(3): p. 153-62.
209. Zenk, F., et al., *Germ line-inherited H3K27me3 restricts enhancer function during maternal-to-zygotic transition.* Science, 2017. **357**(6347): p. 212-216.
210. Cavalli, G. and R. Paro, *The Drosophila Fab-7 chromosomal element conveys epigenetic inheritance during mitosis and meiosis.* Cell, 1998. **93**(4): p. 505-18.
211. Bantignies, F., et al., *Inheritance of Polycomb-dependent chromosomal interactions in Drosophila.* Genes Dev, 2003. **17**(19): p. 2406-20.
212. Ciabrelli, F., et al., *Stable Polycomb-dependent transgenerational inheritance of chromatin states in Drosophila.* Nat Genet, 2017. **49**(6): p. 876-886.
213. Gaydos, L.J., W. Wang, and S. Strome, *Gene repression. H3K27me and PRC2 transmit a memory of repression across generations and during development.* Science, 2014. **345**(6203): p. 1515-8.
214. Hammoud, S.S., et al., *Distinctive chromatin in human sperm packages genes for embryo development.* Nature, 2009. **460**(7254): p. 473-8.
215. Brykczynska, U., et al., *Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa.* Nat Struct Mol Biol, 2010. **17**(6): p. 679-87.
216. Erkek, S., et al., *Molecular determinants of nucleosome retention at CpG-rich sequences in mouse spermatozoa.* Nat Struct Mol Biol, 2013. **20**(7): p. 868-75.
217. Liu, X., et al., *Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos.* Nature, 2016. **537**(7621): p. 558-562.
218. Zheng, H., et al., *Resetting Epigenetic Memory by Reprogramming of Histone Modifications in Mammals.* Mol Cell, 2016. **63**(6): p. 1066-79.
219. Dahl, J.A., et al., *Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition.* Nature, 2016. **537**(7621): p. 548-552.
220. Zhang, B., et al., *Allelic reprogramming of the histone modification H3K4me3 in early mammalian development.* Nature, 2016. **537**(7621): p. 553-557.
221. Zuccotti, M., et al., *Chromatin organization during mouse oocyte growth.* Mol Reprod Dev, 1995. **41**(4): p. 479-85.
222. Prakash, K., et al., *Superresolution imaging reveals structurally distinct periodic patterns of chromatin along pachytene chromosomes.* Proc Natl Acad Sci U S A, 2015. **112**(47): p. 14635-40.

223. Hammoud, S.S., et al., *Chromatin and transcription transitions of mammalian adult germline stem cells and spermatogenesis.* Cell Stem Cell, 2014. **15**(2): p. 239-53.
224. Lesch, B.J., et al., *Parallel evolution of male germline epigenetic poising and somatic development in animals.* Nat Genet, 2016. **48**(8): p. 888-94.
225. Mayer, W., et al., *Spatial separation of parental genomes in preimplantation mouse embryos.* J Cell Biol, 2000. **148**(4): p. 629-34.
226. Lepikhov, K. and J. Walter, *Differential dynamics of histone H3 methylation at positions K4 and K9 in the mouse zygote.* BMC Dev Biol, 2004. **4**: p. 12.
227. Liu, H., J.M. Kim, and F. Aoki, *Regulation of histone H3 lysine 9 methylation in oocytes and early pre-implantation embryos.* Development, 2004. **131**(10): p. 2269-80.
228. Erhardt, S., et al., *Consequences of the depletion of zygotic and embryonic enhancer of zeste 2 during preimplantation mouse development.* Development, 2003. **130**(18): p. 4235-48.
229. Boskovic, A., et al., *Analysis of active chromatin modifications in early mammalian embryos reveals uncoupling of H2A.Z acetylation and H3K36 trimethylation from embryonic genome activation.* Epigenetics, 2012. **7**(7): p. 747-57.
230. Daujat, S., et al., *H3K64 trimethylation marks heterochromatin and is dynamically remodeled during developmental reprogramming.* Nat Struct Mol Biol, 2009. **16**(7): p. 777-81.
231. Kourmouli, N., et al., *Heterochromatin and tri-methylated lysine 20 of histone H4 in animals.* J Cell Sci, 2004. **117**(Pt 12): p. 2491-501.
232. Santos, F., et al., *Dynamic chromatin modifications characterise the first cell cycle in mouse embryos.* Dev Biol, 2005. **280**(1): p. 225-36.
233. Sarmento, O.F., et al., *Dynamic alterations of specific histone modifications during early murine development.* J Cell Sci, 2004. **117**(Pt 19): p. 4449-59.
234. Inoue, A., et al., *Maternal H3K27me3 controls DNA methylation-independent imprinting.* Nature, 2017.
235. Mayer, W., et al., *Demethylation of the zygotic paternal genome.* Nature, 2000. **403**(6769): p. 501-2.
236. Oswald, J., et al., *Active demethylation of the paternal genome in the mouse zygote.* Curr Biol, 2000. **10**(8): p. 475-8.
237. Wossidlo, M., et al., *5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming.* Nat Commun, 2011. **2**: p. 241.
238. Zhang, P., et al., *The involvement of 5-hydroxymethylcytosine in active DNA demethylation in mice.* Biol Reprod, 2012. **86**(4): p. 104.
239. Iqbal, K., et al., *Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine.* Proc Natl Acad Sci U S A, 2011. **108**(9): p. 3642-7.
240. Gu, T.P., et al., *The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes.* Nature, 2011. **477**(7366): p. 606-10.
241. Amouroux, R., et al., *De novo DNA methylation drives 5hmC accumulation in mouse zygotes.* Nat Cell Biol, 2016. **18**(2): p. 225-233.
242. Howlett, S.K. and W. Reik, *Methylation levels of maternal and paternal genomes during preimplantation development.* Development, 1991. **113**(1): p. 119-27.
243. Rougier, N., et al., *Chromosome methylation patterns during mammalian preimplantation development.* Genes Dev, 1998. **12**(14): p. 2108-13.

244. Puschendorf, M., et al., *PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos.* Nat Genet, 2008. **40**(4): p. 411-20.
245. Aoki, F., D.M. Worrad, and R.M. Schultz, *Regulation of transcriptional activity during the first and second cell cycles in the preimplantation mouse embryo.* Dev Biol, 1997. **181**(2): p. 296-307.
246. Hamatani, T., et al., *Dynamics of global gene expression changes during mouse preimplantation development.* Dev Cell, 2004. **6**(1): p. 117-31.
247. Schultz, R.M., *The molecular foundations of the maternal to zygotic transition in the preimplantation embryo.* Hum Reprod Update, 2002. **8**(4): p. 323-31.
248. Wang, H. and S.K. Dey, *Roadmap to embryo implantation: clues from mouse models.* Nat Rev Genet, 2006. **7**(3): p. 185-99.
249. Peaston, A.E., et al., *Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos.* Dev Cell, 2004. **7**(4): p. 597-606.
250. Evsikov, A.V., et al., *Systems biology of the 2-cell mouse embryo.* Cytogenet Genome Res, 2004. **105**(2-4): p. 240-50.
251. Kigami, D., et al., *MuERV-L is one of the earliest transcribed genes in mouse one-cell embryos.* Biol Reprod, 2003. **68**(2): p. 651-4.
252. Svoboda, P., et al., *RNAi and expression of retrotransposons MuERV-L and IAP in preimplantation mouse embryos.* Dev Biol, 2004. **269**(1): p. 276-85.
253. Ribet, D., et al., *Murine endogenous retrovirus MuERV-L is the progenitor of the "orphan" epsilon viruslike particles of the early mouse embryo.* J Virol, 2008. **82**(3): p. 1622-5.
254. Macfarlan, T.S., et al., *Embryonic stem cell potency fluctuates with endogenous retrovirus activity.* Nature, 2012. **487**(7405): p. 57-63.
255. Wu, J., et al., *The landscape of accessible chromatin in mammalian preimplantation embryos.* Nature, 2016. **534**(7609): p. 652-7.
256. Du, Z., et al., *Allelic reprogramming of 3D chromatin architecture during early mammalian development.* Nature, 2017. **547**(7662): p. 232-235.
257. Ke, Y., et al., *3D Chromatin Structures of Mature Gametes and Structural Reprogramming during Mammalian Embryogenesis.* Cell, 2017. **170**(2): p. 367-381 e20.
258. Ishiuchi, T., et al., *Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly.* Nat Struct Mol Biol, 2015. **22**(9): p. 662-71.
259. Whiddon, J.L., et al., *Conservation and innovation in the DUX4-family gene network.* Nat Genet, 2017. **49**(6): p. 935-940.
260. Hendrickson, P.G., et al., *Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons.* Nat Genet, 2017. **49**(6): p. 925-934.
261. De Iaco, A., et al., *DUX-family transcription factors regulate zygotic genome activation in placental mammals.* Nat Genet, 2017. **49**(6): p. 941-945.
262. Falco, G., et al., *Zscan4: a novel gene expressed exclusively in late 2-cell embryos and embryonic stem cells.* Dev Biol, 2007. **307**(2): p. 539-50.
263. Rodriguez-Terrones, D., et al., *A molecular roadmap for the emergence of early-embryonic-like cells in culture.* Nat Genet, 2018. **50**(1): p. 106-119.
264. Chen, Q., et al., *Tracing the origin of heterogeneity and symmetry breaking in the early mammalian embryo.* Nat Commun, 2018. **9**(1): p. 1819.

265. Tsunoda, Y. and A. McLaren, *Effect of various procedures on the viability of mouse embryos containing half the normal number of blastomeres.* J Reprod Fertil, 1983. **69**(1): p. 315-22.
266. Papaioannou, V.E. and K.M. Ebert, *Mouse half embryos: viability and allocation of cells in the blastocyst.* Dev Dyn, 1995. **203**(4): p. 393-8.
267. Katayama, M., M.R. Ellersieck, and R.M. Roberts, *Development of monozygotic twin mouse embryos from the time of blastomere separation at the two-cell stage to blastocyst.* Biol Reprod, 2010. **82**(6): p. 1237-47.
268. Morris, S.A., Y. Guo, and M. Zernicka-Goetz, *Developmental plasticity is bound by pluripotency and the Fgf and Wnt signaling pathways.* Cell Rep, 2012. **2**(4): p. 756-65.
269. Casser, E., et al., *Totipotency segregates between the sister blastomeres of two-cell stage mouse embryos.* Sci Rep, 2017. **7**(1): p. 8299.
270. Biase, F.H., X. Cao, and S. Zhong, *Cell fate inclination within 2-cell and 4-cell mouse embryos revealed by single-cell RNA sequencing.* Genome Res, 2014. **24**(11): p. 1787-96.
271. Shi, J., et al., *Dynamic transcriptional symmetry-breaking in pre-implantation mammalian embryo development revealed by single-cell RNA-seq.* Development, 2015. **142**(20): p. 3468-77.
272. White, M.D., et al., *Long-Lived Binding of Sox2 to DNA Predicts Cell Fate in the Four-Cell Mouse Embryo.* Cell, 2016. **165**(1): p. 75-87.
273. Plachta, N., et al., *Oct4 kinetics predict cell lineage patterning in the early mammalian embryo.* Nat Cell Biol, 2011. **13**(2): p. 117-23.
274. Torres-Padilla, M.E., et al., *Histone arginine methylation regulates pluripotency in the early mouse embryo.* Nature, 2007. **445**(7124): p. 214-8.
275. Burton, A., et al., *Single-cell profiling of epigenetic modifiers identifies PRDM14 as an inducer of cell fate in the mammalian embryo.* Cell Rep, 2013. **5**(3): p. 687-701.
276. Goolam, M., et al., *Heterogeneity in Oct4 and Sox2 Targets Biases Cell Fate in 4-Cell Mouse Embryos.* Cell, 2016. **165**(1): p. 61-74.
277. Lorthongpanich, C., et al., *Developmental fate and lineage commitment of singled mouse blastomeres.* Development, 2012. **139**(20): p. 3722-31.
278. Stephenson, R.O., Y. Yamanaka, and J. Rossant, *Disorganized epithelial polarity and excess trophectoderm cell fate in preimplantation embryos lacking E-cadherin.* Development, 2010. **137**(20): p. 3383-91.
279. Skamagki, M., et al., *Asymmetric localization of Cdx2 mRNA during the first cell-fate decision in early mouse development.* Cell Rep, 2013. **3**(2): p. 442-57.
280. Nishioka, N., et al., *The Hippo signaling pathway components Lats and Yap pattern Tead4 activity to distinguish mouse trophectoderm from inner cell mass.* Dev Cell, 2009. **16**(3): p. 398-410.
281. Yagi, R., et al., *Transcription factor TEAD4 specifies the trophectoderm lineage at the beginning of mammalian development.* Development, 2007. **134**(21): p. 3827-36.
282. Ng, R.K., et al., *Epigenetic restriction of embryonic cell lineage fate by methylation of Elf5.* Nat Cell Biol, 2008. **10**(11): p. 1280-90.
283. Strumpf, D., et al., *Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst.* Development, 2005. **132**(9): p. 2093-102.

284. Zernicka-Goetz, M., S.A. Morris, and A.W. Bruce, *Making a firm decision: multifaceted regulation of cell fate in the early mouse embryo.* Nat Rev Genet, 2009. **10**(7): p. 467-77.
285. Chazaud, C., et al., *Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway.* Dev Cell, 2006. **10**(5): p. 615-24.
286. Morris, S.A., et al., *Origin and formation of the first two distinct cell types of the inner cell mass in the mouse embryo.* Proc Natl Acad Sci U S A, 2010. **107**(14): p. 6364-9.
287. Krupa, M., et al., *Allocation of inner cells to epiblast vs primitive endoderm in the mouse embryo is biased but not determined by the round of asymmetric divisions (8-->16- and 16-->32-cells).* Dev Biol, 2014. **385**(1): p. 136-48.
288. Morris, S.A., et al., *The differential response to Fgf signalling in cells internalized at different times influences lineage segregation in preimplantation mouse embryos.* Open Biol, 2013. **3**(11): p. 130104.
289. Frum, T., et al., *Oct4 cell-autonomously promotes primitive endoderm development in the mouse blastocyst.* Dev Cell, 2013. **25**(6): p. 610-22.
290. Le Bin, G.C., et al., *Oct4 is required for lineage priming in the developing inner cell mass of the mouse blastocyst.* Development, 2014. **141**(5): p. 1001-10.
291. Ahmed, K., et al., *Global chromatin architecture reflects pluripotency and lineage commitment in the early mouse embryo.* PLoS One, 2010. **5**(5): p. e10531.
292. O'Neill, L.P., M.D. VerMilyea, and B.M. Turner, *Epigenetic characterization of the early embryo with a chromatin immunoprecipitation protocol applicable to small cell populations.* Nat Genet, 2006. **38**(7): p. 835-41.
293. Santos, F., et al., *Dynamic reprogramming of DNA methylation in the early mouse embryo.* Dev Biol, 2002. **241**(1): p. 172-82.
294. Lyon, M.F., *Gene action in the X-chromosome of the mouse (Mus musculus L.).* Nature, 1961. **190**: p. 372-3.
295. Okamoto, I., et al., *Evidence for de novo imprinted X-chromosome inactivation independent of meiotic inactivation in mice.* Nature, 2005. **438**(7066): p. 369-73.
296. Takagi, N. and M. Sasaki, *Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse.* Nature, 1975. **256**(5519): p. 640-2.
297. Okamoto, I., et al., *Epigenetic dynamics of imprinted X inactivation during early mouse development.* Science, 2004. **303**(5658): p. 644-9.
298. Mak, W., et al., *Reactivation of the paternal X chromosome in early mouse embryos.* Science, 2004. **303**(5658): p. 666-9.
299. Rastan, S., *Non-random X-chromosome inactivation in mouse X-autosome translocation embryos--location of the inactivation centre.* J Embryol Exp Morphol, 1983. **78**: p. 1-22.
300. Rastan, S. and E.J. Robertson, *X-chromosome deletions in embryo-derived (EK) cell lines associated with lack of X-chromosome inactivation.* J Embryol Exp Morphol, 1985. **90**: p. 379-88.
301. Kay, G.F., et al., *Imprinting and X chromosome counting mechanisms determine Xist expression in early mouse development.* Cell, 1994. **77**(5): p. 639-50.
302. Borsani, G., et al., *Characterization of a murine gene expressed from the inactive X chromosome.* Nature, 1991. **351**(6324): p. 325-9.

303. Penny, G.D., et al., *Requirement for Xist in X chromosome inactivation.* Nature, 1996. **379**(6561): p. 131-7.
304. Marahrens, Y., et al., *Xist-deficient mice are defective in dosage compensation but not spermatogenesis.* Genes Dev, 1997. **11**(2): p. 156-66.
305. Brockdorff, N., et al., *The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus.* Cell, 1992. **71**(3): p. 515-26.
306. Chaumeil, J., et al., *Integrated kinetics of X chromosome inactivation in differentiating embryonic stem cells.* Cytogenet Genome Res, 2002. **99**(1-4): p. 75-84.
307. Chaumeil, J., et al., *A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced.* Genes Dev, 2006. **20**(16): p. 2223-37.
308. Chu, C., et al., *Systematic discovery of Xist RNA binding proteins.* Cell, 2015. **161**(2): p. 404-16.
309. McHugh, C.A., et al., *The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3.* Nature, 2015. **521**(7551): p. 232-6.
310. Monfort, A., et al., *Identification of Spen as a Crucial Factor for Xist Function through Forward Genetic Screening in Haploid Embryonic Stem Cells.* Cell Rep, 2015. **12**(4): p. 554-61.
311. Shi, Y., et al., *Sharp, an inducible cofactor that integrates nuclear receptor repression and activation.* Genes Dev, 2001. **15**(9): p. 1140-51.
312. You, S.H., et al., *Nuclear receptor co-repressors are required for the histone-deacetylase activity of HDAC3 in vivo.* Nat Struct Mol Biol, 2013. **20**(2): p. 182-7.
313. Escamilla-Del-Arenal, M., S.T. da Rocha, and E. Heard, *Evolutionary diversity and developmental regulation of X-chromosome inactivation.* Hum Genet, 2011. **130**(2): p. 307-27.
314. Pintacuda, G., et al., *hnRNPK Recruits PCGF3/5-PRC1 to the Xist RNA B-Repeat to Establish Polycomb-Mediated Chromosomal Silencing.* Mol Cell, 2017. **68**(5): p. 955-969 e10.
315. Almeida, M., et al., *PCGF3/5-PRC1 initiates Polycomb recruitment in X chromosome inactivation.* Science, 2017. **356**(6342): p. 1081-1084.
316. Minkovsky, A., et al., *The Mbd1-Atf7ip-Setdb1 pathway contributes to the maintenance of X chromosome inactivation.* Epigenetics Chromatin, 2014. **7**: p. 12.
317. Escamilla-Del-Arenal, M., et al., *Cdyl, a new partner of the inactive X chromosome and potential reader of H3K27me3 and H3K9me2.* Mol Cell Biol, 2013. **33**(24): p. 5005-20.
318. Huynh, K.D. and J.T. Lee, *Inheritance of a pre-inactivated paternal X chromosome in early mouse embryos.* Nature, 2003. **426**(6968): p. 857-62.
319. Namekawa, S.H., et al., *Two-step imprinted X inactivation: repeat versus genic silencing in the mouse.* Mol Cell Biol, 2010. **30**(13): p. 3187-205.
320. Nesterova, T.B., et al., *Loss of Xist imprinting in diploid parthenogenetic preimplantation embryos.* Dev Biol, 2001. **235**(2): p. 343-50.
321. Tada, T., et al., *Imprint switching for non-random X-chromosome inactivation during mouse oocyte growth.* Development, 2000. **127**(14): p. 3101-5.
322. Sado, T., et al., *Regulation of imprinted X-chromosome inactivation in mice by Tsix.* Development, 2001. **128**(8): p. 1275-86.

323. Maclary, E., et al., *Differentiation-dependent requirement of Tsix long non-coding RNA in imprinted X-chromosome inactivation.* Nat Commun, 2014. **5**: p. 4209.
324. Chiba, H., et al., *De novo DNA methylation independent establishment of maternal imprint on X chromosome in mouse oocytes.* Genesis, 2008. **46**(12): p. spc one.
325. Fukuda, A., et al., *The role of maternal-specific H3K9me3 modification in establishing imprinted X-chromosome inactivation and embryogenesis in mice.* Nat Commun, 2014. **5**: p. 5464.
326. Fukuda, A., et al., *Chromatin condensation of Xist genomic loci during oogenesis in mice.* Development, 2015. **142**(23): p. 4049-55.
327. Inoue, A., et al., *Genomic imprinting of Xist by maternal H3K27me3.* Genes Dev, 2017. **31**(19): p. 1927-1932.
328. Aloia, L., B. Di Stefano, and L. Di Croce, *Polycomb complexes in stem cells and embryonic development.* Development, 2013. **140**(12): p. 2525-34.
329. Ralston, A. and J. Rossant, *Cdx2 acts downstream of cell polarization to cell-autonomously promote trophectoderm fate in the early mouse embryo.* Dev Biol, 2008. **313**(2): p. 614-29.
330. Jedrusik, A., et al., *Maternal-zygotic knockout reveals a critical role of Cdx2 in the morula to blastocyst transition.* Dev Biol, 2015. **398**(2): p. 147-52.
331. He, Y., et al., *The EED protein-protein interaction inhibitor A-395 inactivates the PRC2 complex.* Nat Chem Biol, 2017. **13**(4): p. 389-395.
332. Xu, B., et al., *Selective inhibition of EZH2 and EZH1 enzymatic activity by a small molecule suppresses MLL-rearranged leukemia.* Blood, 2015. **125**(2): p. 346-57.
333. Wang, F., et al., *Regulation of X-linked gene expression during early mouse development by Rlim.* Elife, 2016. **5**.
334. Kobayashi, S., et al., *The X-linked imprinted gene family Fthl17 shows predominantly female expression following the two-cell stage in mouse embryos.* Nucleic Acids Res, 2010. **38**(11): p. 3672-81.
335. Kobayashi, S., et al., *Comparison of gene expression in male and female mouse blastocysts revealed imprinting of the X-linked gene, Rhox5/Pem, at preimplantation stages.* Curr Biol, 2006. **16**(2): p. 166-72.
336. Oda, M., et al., *DNA methylation regulates long-range gene silencing of an X-linked homeobox gene cluster in a lineage-specific manner.* Genes Dev, 2006. **20**(24): p. 3382-94.
337. Chiba, H., et al., *De novo DNA methylation independent establishment of maternal imprint on X chromosome in mouse oocytes.* Genesis, 2008. **46**(12): p. 768-74.
338. Prokopuk, L., et al., *Loss of maternal EED results in postnatal overgrowth.* Clin Epigenetics, 2018. **10**(1): p. 95.
339. Buecker, C., et al., *Reorganization of enhancer patterns in transition from naive to primed pluripotency.* Cell Stem Cell, 2014. **14**(6): p. 838-53.
340. Boroviak, T., et al., *Lineage-Specific Profiling Delineates the Emergence and Progression of Naive Pluripotency in Mammalian Embryogenesis.* Dev Cell, 2015. **35**(3): p. 366-82.
341. Acampora, D., et al., *Loss of the Otx2-Binding Site in the Nanog Promoter Affects the Integrity of Embryonic Stem Cell Subtypes and Specification of Inner Cell Mass-Derived Epiblast.* Cell Rep, 2016. **15**(12): p. 2651-64.
342. Tsunoda, Y., T. Tokunaga, and T. Sugie, *Altered sex ratio of live young after transfer of fast- and slow-developing mouse embryos.* Gamete Reserch, 1985(12): p. 301-304.

343. Valdivia, R.P., et al., *PCR sexing and developmental rate differences in preimplantation mouse embryos fertilized and cultured in vitro.* Mol Reprod Dev, 1993. **35**(2): p. 121-6.
344. Bermejo-Alvarez, P., et al., *Sex determines the expression level of one third of the actively expressed genes in bovine blastocysts.* Proc Natl Acad Sci U S A, 2010. **107**(8): p. 3394-9.
345. Berletch, J.B., et al., *Female bias in Rhox6 and 9 regulation by the histone demethylase KDM6A.* PLoS Genet, 2013. **9**(5): p. e1003489.
346. Matoba, S., et al., *Loss of H3K27me3 Imprinting in Somatic Cell Nuclear Transfer Embryos Disrupts Post-Implantation Development.* Cell Stem Cell, 2018.
347. Miri, K., et al., *The imprinted polycomb group gene Sfmbt2 is required for trophoblast maintenance and placenta development.* Development, 2013. **140**(22): p. 4480-9.
348. Inoue, K., et al., *The Rodent-Specific MicroRNA Cluster within the Sfmbt2 Gene Is Imprinted and Essential for Placental Development.* Cell Rep, 2017. **19**(5): p. 949-956.
349. Itoh, M., et al., *Role of Gab1 in heart, placenta, and skin development and growth factor- and cytokine-induced extracellular signal-regulated kinase mitogen-activated protein kinase activation.* Mol Cell Biol, 2000. **20**(10): p. 3695-704.
350. Acampora, D., et al., *Forebrain and midbrain regions are deleted in Otx2-/- mutants due to a defective anterior neuroectoderm specification during gastrulation.* Development, 1995. **121**(10): p. 3279-90.
351. Okuda, T., et al., *AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis.* Cell, 1996. **84**(2): p. 321-30.
352. Wang, L.C., et al., *Yolk sac angiogenic defect and intra-embryonic apoptosis in mice lacking the Ets-related factor TEL.* EMBO J, 1997. **16**(14): p. 4374-83.
353. Xie, H., et al., *Polycomb repressive complex 2 regulates normal hematopoietic stem cell function in a developmental-stage-specific manner.* Cell Stem Cell, 2014. **14**(1): p. 68-80.
354. Long, J., E. Shelhamer, and T. Darrell. *Fully convolutional networks for semantic segmentation*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.
355. Vincent, L. and P. Soille, *Watersheds in Digital Spaces - an Efficient Algorithm Based on Immersion Simulations.* Ieee Transactions on Pattern Analysis and Machine Intelligence, 1991. **13**(6): p. 583-598.
356. He, K., et al. *Deep residual learning for image recognition*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
357. Ronneberger, O., P. Fischer, and T. Brox. *U-net: Convolutional networks for biomedical image segmentation*. in *International Conference on Medical image computing and computer-assisted intervention*. 2015. Springer.
358. Berthold, M.R., et al., *KNIME-the Konstanz information miner: version 2.0 and beyond.* AcM SIGKDD explorations Newsletter, 2009. **11**(1): p. 26-31.
359. Schneider, C.A., W.S. Rasband, and K.W. Eliceiri, *NIH Image to ImageJ: 25 years of image analysis.* Nat Methods, 2012. **9**(7): p. 671-5.
360. Picelli, S., et al., *Full-length RNA-seq from single cells using Smart-seq2.* Nat Protoc, 2014. **9**(1): p. 171-81.
361. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2013. **29**(1): p. 15-21.

362. Gaidatzis, D., et al., *QuasR: quantification and annotation of short reads in R.* Bioinformatics, 2015. **31**(7): p. 1130-2.
363. Mortazavi, A., et al., *Mapping and quantifying mammalian transcriptomes by RNA-Seq.* Nat Methods, 2008. **5**(7): p. 621-8.
364. Trapnell, C., et al., *The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells.* Nat Biotechnol, 2014. **32**(4): p. 381-386.
365. McCarthy, D.J., Y. Chen, and G.K. Smyth, *Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation.* Nucleic Acids Res, 2012. **40**(10): p. 4288-97.
366. Krijthe, J.H., *Rtsne: T-Distributed Stochastic Neighbor Embedding using Barnes-Hut Implementation.* 2015.
367. Gu, Z., R. Eils, and M. Schlesner, *Complex heatmaps reveal patterns and correlations in multidimensional genomic data.* Bioinformatics, 2016. **32**(18): p. 2847-9.
368. Takada, T., et al., *The ancestor of extant Japanese fancy mice contributed to the mosaic genomes of classical inbred strains.* Genome Res, 2013. **23**(8): p. 1329-38.
369. Keane, T.M., et al., *Mouse genomic variation and its effect on phenotypes and gene regulation.* Nature, 2011. **477**(7364): p. 289-94.
370. Park, S.J., et al., *Inferring the choreography of parental genomes during fertilization from ultralarge-scale whole-transcriptome analysis.* Genes Dev, 2013. **27**(24): p. 2736-48.
371. Kuzmichev, A., et al., *Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein.* Genes Dev, 2002. **16**(22): p. 2893-905.
372. Karimi, M.M., et al., *DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs.* Cell Stem Cell, 2011. **8**(6): p. 676-87.
373. Fouse, S.D., et al., *Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation.* Cell Stem Cell, 2008. **2**(2): p. 160-9.
374. Maatouk, D.M., et al., *DNA methylation is a primary mechanism for silencing postmigratory primordial germ cell genes in both germ cell and somatic cell lineages.* Development, 2006. **133**(17): p. 3411-8.
375. Smallwood, S.A., et al., *Dynamic CpG island methylation landscape in oocytes and preimplantation embryos.* Nat Genet, 2011. **43**(8): p. 811-4.
376. Walter, M., et al., *An epigenetic switch ensures transposon repression upon dynamic loss of DNA methylation in embryonic stem cells.* Elife, 2016. **5**.
377. Domcke, S., et al., *Competition between DNA methylation and transcription factors determines binding of NRF1.* Nature, 2015. **528**(7583): p. 575-9.
378. Wang, P.J., et al., *An abundance of X-linked genes expressed in spermatogonia.* Nat Genet, 2001. **27**(4): p. 422-6.
379. Chuang, J.C., et al., *Comparison of biological effects of non-nucleoside DNA methylation inhibitors versus 5-aza-2'-deoxycytidine.* Mol Cancer Ther, 2005. **4**(10): p. 1515-20.
380. De Smet, C., et al., *DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter.* Mol Cell Biol, 1999. **19**(11): p. 7327-35.

381. Conrad, N.K. and J.A. Steitz, *A Kaposi's sarcoma virus RNA element that increases the nuclear abundance of intronless transcripts.* EMBO J, 2005. **24**(10): p. 1831-41.
382. Conrad, N.K., et al., *Mutational analysis of a viral RNA element that counteracts rapid RNA decay by interaction with the polyadenylate tail.* Proc Natl Acad Sci U S A, 2007. **104**(25): p. 10412-7.
383. Strnad, P., et al., *Inverted light-sheet microscope for imaging mouse pre-implantation development.* Nat Methods, 2016. **13**(2): p. 139-42.
384. Falcon, S. and R. Gentleman, *Using GOstats to test gene lists for GO term association.* Bioinformatics, 2007. **23**(2): p. 257-8.
385. Saitou, M., S. Kagiwada, and K. Kurimoto, *Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells.* Development, 2012. **139**(1): p. 15-31.
386. Ferguson-Smith, A.C., *Genomic imprinting: the emergence of an epigenetic paradigm.* Nat Rev Genet, 2011. **12**(8): p. 565-75.
387. Frost, J.M. and G.E. Moore, *The importance of imprinting in the human placenta.* PLoS Genet, 2010. **6**(7): p. e1001015.
388. Coan, P.M., G.J. Burton, and A.C. Ferguson-Smith, *Imprinted genes in the placenta--a review.* Placenta, 2005. **26 Suppl A**: p. S10-20.
389. Smith, Z.D., et al., *A unique regulatory phase of DNA methylation in the early mammalian embryo.* Nature, 2012. **484**(7394): p. 339-44.
390. Inoue, A. and Y. Zhang, *Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos.* Science, 2011. **334**(6053): p. 194.
391. Kafri, T., et al., *Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line.* Genes Dev, 1992. **6**(5): p. 705-14.
392. Borgel, J., et al., *Targets and dynamics of promoter DNA methylation during early mouse development.* Nat Genet, 2010. **42**(12): p. 1093-100.
393. Dean, W., et al., *Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos.* Proc Natl Acad Sci U S A, 2001. **98**(24): p. 13734-8.
394. Simpson, A.J., et al., *Cancer/testis antigens, gametogenesis and cancer.* Nat Rev Cancer, 2005. **5**(8): p. 615-25.
395. Straussman, R., et al., *Developmental programming of CpG island methylation profiles in the human genome.* Nat Struct Mol Biol, 2009. **16**(5): p. 564-71.
396. Velasco, G., et al., *Dnmt3b recruitment through E2F6 transcriptional repressor mediates germ-line gene silencing in murine somatic tissues.* Proc Natl Acad Sci U S A, 2010. **107**(20): p. 9281-6.
397. Huang, S., *Non-genetic heterogeneity of cells in development: more than just noise.* Development, 2009. **136**(23): p. 3853-62.
398. Ohnishi, Y., et al., *Cell-to-cell expression variability followed by signal reinforcement progressively segregates early mouse lineages.* Nat Cell Biol, 2014. **16**(1): p. 27-37.
399. Dietrich, J.E. and T. Hiiragi, *Stochastic patterning in the mouse pre-implantation embryo.* Development, 2007. **134**(23): p. 4219-31.
400. Mohammed, H., et al., *Single-Cell Landscape of Transcriptional Heterogeneity and Cell Fate Decisions during Mouse Early Gastrulation.* Cell Rep, 2017. **20**(5): p. 1215-1228.

Acknowledgements

First, I would like to thank Antoine Peters for giving me the opportunity to pursue my PhD research in his laboratory. Thank you for giving me the freedom to follow my own ideas, while also providing guidance and support when needed.

I would like to express my gratitude to the members of my PhD committee, Petra Hajkova and Jörg Betschinger for their valuable input and support throughout the course of my PhD. My committee meetings have been very helpful and encouraging.

I would like to thank all past and present members of the Peters lab for their scientific input and for creating a stimulating and joyful working environment. Especially, I thank Peter Nestorov for helping me to get started in the lab at the beginning of my PhD, Zichuan Liu for help with microinjection and Helene Royo and Evgeniy Ozonov for their extremely valuable support with bioinformatic analysis.

I would like to thank the Peters and Schübeler lab members that I have been sharing my office with: Peter Nestorov, Mathieu Tardat, Philipp Bammer, Yumiko Kitazawa, Nhuong Nguyen, Rabih Murr, Sylvia Domcke, Darko Barisic, Arnaud Krebs, Dominik Hartl, Anna-Karina Felder and Marlene Lübke. Thank you for the fun times and stimulating scientific discussions we had in the past years.

I thank the members of the Betschinger, Galli and Schübeler labs for their input during our weekly joint lab meetings.

I acknowledge all the excellent facilities at FMI that I have been collaborating with in the past years. Especially, I thank Birgitte Lucas and Sebastien Smallwood from the Functional Genomics Facility for their help in setting up single embryo RNA sequencing and for the many helpful discussions we had. I would like to thank the FAIM facility for their support with imaging and especially Raphael Thierry and Markus Rempfler for help with image analysis. Special thanks also to Jean-Francois Spetz for help with embryo experiments. I

would also like to thank the FMI animal facility and especially all the animal caretakers who took great care of my mice.

I acknowledge the Boehringer Ingelheim Fonds for their financial support and the great meetings and network of BIF fellows.

I would like to thank Elida Keller, who always had an open ear for my questions and was very helpful to manage the administrative aspects of the PhD.

I would like to thank Haruhiko Koseki and Jafar Sharif for their valuable scientific input on my project and for providing *Pcgf6* KO embryos for RNA sequencing.

Finally, I would like to thank my husband Philipp, my parents, my brother and my friends for giving me constant motivation and support throughout my studies. Without them I would not have been able to complete this work.