# The role of Polycomb Repressive Complex 2 in mouse preimplantation development

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch–Naturwissenschaftlichen Fakultät

der Universität Basel

von

Peter Dobrinov Nestorov

aus Bulgarien

**Basel, 2015**

Originaldokument gespeichert auf dem Dokumentenserver der Universität Basel
edoc.unibas.ch

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät der Universität Basel

auf Antrag von

Prof. Antoine HFM Peters

(Fakultätsverantwortlicher und Referent)

Prof. Rolf Zeller

(Korreferent)

Basel, den 9. Dezember 2014

Prof. Jörg Schibler

(Dekan)

**TABLE OF CONTENTS**

# I. Summary

Pluripotency is the ability of a cell to differentiate into any of the embryonic germ layers and is therefore referred to as the ground state of development. The totipotent/pluripotent state represents a bridge between generations – on one hand, it is initiated by the fusion of the two gametes that represent the previous generation, and on the other hand it gives rise to the germline of the next generation. Studies in model systems suggest that pluripotency is governed by a core transcriptional network that arises during mammalian preimplantation development, in a process accompanied by dynamic changes in chromatin organization, histone modifications and DNA methylation. In my thesis, I addressed the developmental and regulatory role of the evolutionary conserved Polycomb Repressive Complex 2 (PRC2) at the interface between two generations – in the oocyte and the preimplantation embryo.

I demonstrated that genetic ablation of core members of PRC2 has an effect on H3K27me3 *in vivo,* and leads to a developmental and transcriptional response in late oocytes and early embryos. Furthermore, the observed mutant phenotypes revealed a dosage-dependent requirement for PRC2/H3K27me3 in the preimplantation embryo. I also found genetic evidence for an interplay between the two major Polycomb complexes, PRC1 and PRC2, in preimplantation embryos. I further described the transcriptional dynamics during early embryonic development of genes encoding chromatin modifiers. This single-cell profiling study highlighted the existence of maternal and embryonic variants of the major chromatin modifying complexes.

In summary, my work reveals an important role of chromatin-based regulation in the preparation and acquisition of totipotency *in vivo,* manifested by a dosage-dependent PRC2/H3K27me3 requirement during the maternal-to-zygotic transition.

# II. List of tables

# III. List of figures

# IV. List of abbreviations

| | |
|---|---|
| ChIP | Chromatin immuno-precipitation |
| DNA | desoxyribonucleic acid |
| EPI | epiblast |
| ESC | embryonic stem cell |
| GRN | gene regulatory network |
| HSC | Hematopoietic stem cell |
| ICM | inner cell mass |
| iPS | induced pluripotent stem cell |
| mRNA | messenger ribonucleic acid |
| ncRNA | non-coding RNA |
| NLB | Nucleolar-Like Bodies |
| PcG | Polycomb group |
| PGC | Progenitor germ cell |
| PRC | Polycomb repressive complex |
| PTM | post-translational modification |
| RNA | ribonucleic acid |
| TE | trophectoderm |
| TF | transcription factor |
| ZGA | Zygotic genome activation |

# Chapter 1. Introduction and scope of thesis

## 1.1 The epigenetic landscape

How does a single cell give rise to a multicellular organism, consisting of different cell types, organized in tissues and organs and forming a complex three-dimensional body plan? Developmental biologists have addressed this question from many different perspectives in a variety of model organisms, and it has become evident that the processes taking place during development are pre-programmed in the hereditary molecule of nature – DNA, which can be pictured as a blueprint of the whole organism. However, in order to achieve cellular diversity and adequate response to signals from the environment, different cells read and implement only selected parts of the blueprint by regulating the flow of genetic information. The central dogma of molecular biology postulates that genes encoded in the DNA are transcribed into messenger RNA and then the mRNA is translated into protein, which is the functional product of gene expression (Crick, 1970). From a biochemical standpoint, the flow of genetic information is a combination of enzymatic reactions and binding events, which are controlled on multiple levels (Ptashne and Gann, 2002). For instance, the transcription of DNA into mRNA is performed by an enzymatic complex, which is a DNA-dependent RNA polymerase, i.e. it requires DNA as a template to catalyse the polymerisation of single ribose nucleotides to form mRNA. Both the enzymatic activity and the DNA-binding affinity of the RNA Polymerase can be positively or negatively regulated, which can ultimately explain 80% of the variation of mRNA levels in a given cell (Tippmann et al., 2012). However, the RNA Polymerase lacks gene specificity and requires the aid of other players, which orchestrate the execution of the genetic program.

Now to rephrase the question from the beginning: how does a single genotype (the information encoded in the DNA) give rise to a myriad of phenotypes (different sets of traits, characterizing different cell types in an organism)? In 1957, in order to illustrate this, Conrad Hal Waddington put forward the model of the epigenetic landscape, which pictures the cells as marbles rolling down a rugged area towards a wall (Waddington, 1957). The grooves in the hill are the

permitted states of the system, which Waddington called "chreodes". The wall, towards which the cells are heading is their terminal differentiation state, which they will eventually reach, regardless of minor intrinsic or extrinsic changes. This latter statement is what Waddington coined as the "canalization of development" (Waddington, 1942), integrating together two principles of development: the compulsory link between genotype and phenotype on the one hand, and the robustness of the system on the other. Waddington's models, which arose before the era of modern developmental genetics and molecular biology, are still valid and serve as a reference and basis of other theories. One such theory by Siu Huang incorporates Waddington's epigenetic landscape, the existence of gene regulatory networks (GRN) and the transcriptional noise (shown in Figure 1.1 (Huang, 2012a)). A GRN has been defined as the basic invariant unit of development, which is encoded by the genome and by this serves also as the platform of evolution (Davidson and Erwin, 2006). GRNs include all the regulatory links in a given genome (therefore a genome encodes a single GRN only) and have a modular and hierarchical structure. To cite Siu Huang: "GRNs are akin to law: carved in stone, but applied to situations only when relevant" (Huang, 2012a). Therefore, one GRN allows the existence of many gene expression patterns (GRN states), as long as they follow the general rules. The transition between these different states results in the generation of different phenotypes from the same genotype. During developmental progression the cells move through the rugged epigenetic landscape and reach "attractors" – different stable states of the GRN. Attractors are characterized by a local decrease of the quasi-potential energy that specifies a given GRN state. Thus, it is more likely that a cell remains captured in this state for a longer time (or for many cell generations). Again, the surrounding high potential and "prohibited" regions in the epigenetic landscape are determined by the invariable links in the GRN. Any "trespassing" is either not tolerated (energetically unfavourable) or is not possible because of non-existent relationships between genes in the GRN (unless a spontaneous genetic mutation allows for a previously missing interaction). Furthermore, transitions between two GRN states with different quasi-potential energies will generally occur in one direction only, namely from a higher to a lower energetic state. Terminally differentiated cells are in the lowest possible energetic state in a given Waddington's chreode (i.e. developmental path; blue arrows in Figure 1.1D), and therefore would remain in this state under normal developmental conditions. However, research in

recent years has shown that under certain *in vitro* conditions it is possible to dedifferentiate a cell (a change towards a higher energetic state), or even to transdifferentiate cells (i.e. a direct switch from one cell lineage to another) (Ladewig et al., 2013). These developmental transitions occur by artificially introducing genes that encode for master transcription factors (master TFs), or pioneer factors, which upon expression are able to control the GRN at the highest hierarchy level (Magnani et al., 2011; Zaret et al., 2008). In 2006 Kazutoshi Takahashi and Shinya Yamanaka reported the successful reprograming and dedifferentiation of mouse fibroblasts (Takahashi and Yamanaka, 2006). They used a combination of four TFs - Oct3/4, Sox2, c-Myc, and Klf4 to switch the GRN state of the fibroblasts to the GRN state of an induced pluripotent stem cells (iPSC). iPSC resemble the embryonic stem cells (ESC), which are derived from the early embryo and have the potential to differentiate into any embryonic cell type. Recent research by Richard Young and colleagues gives a hint on how a few transcription factors (TFs) can induce major changes in gene expression and facilitate a transition in cell fate. His team showed that the master regulators of ESC identity Oct3/4, Sox2, Nanog, Klf4 and Esrrb form large protein complexes and bind specific DNA loci called super-enhancer elements (Whyte et al., 2013). When activated, the super enhancers trigger the expression of another set of cell-specific genes (a GRN module) and thus lead to a shift of the GRN state. Furthermore, the master TFs positively regulate their own expression through feedback loops, which leads to a stronger activation of the downstream genes and ultimately drives the cell into the respective stable GRN state. During normal development, cell differentiation is regulated and coordinated in the context of the whole organism and goes towards a more differentiated state (with the notable exception of germ cells, which will be discussed below). On the other hand, “enforced” reprogramming by TFs *in vitro* leads to a seemingly normal cell that resembles a certain *in vivo* cell type. However, the efficiency is usually very low and there are certain differences on the molecular level, particularly the methylation state of DNA, as well as the post-translational chromatin modifications (Hanna et al., 2010; Hasegawa et al., 2010; Kim et al., 2010). This suggests that the TFs may be the master regulators of the GRN, but as we will see below, there are multiple layers of gene regulation, which act along with the TFs to support the smooth and “canalized” transition between GRN states, as well as to maintain the stability of a GRN over multiple cell cycles.

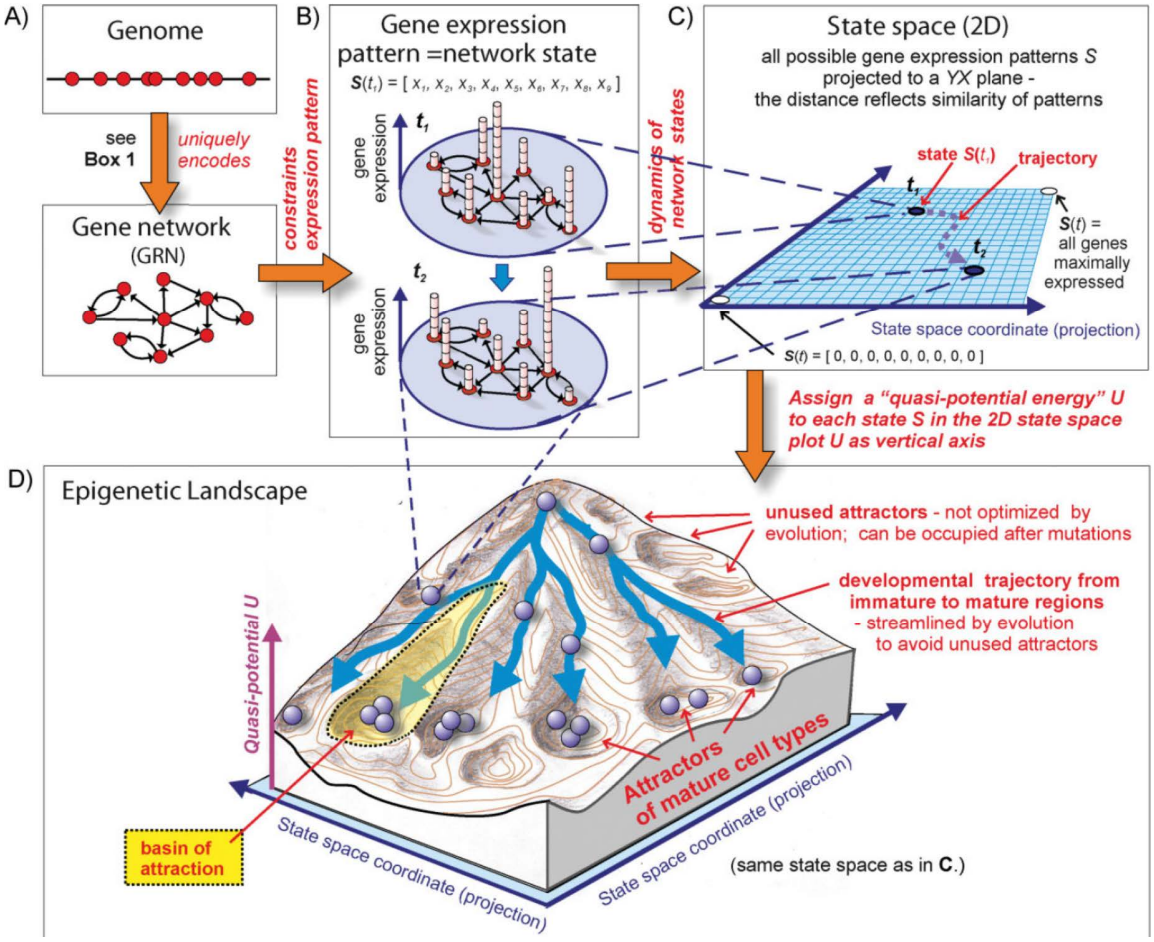



***Figure 1.1. Gene regulatory networks and the epigenetic landscape***

*The unequivocal correspondence (unique mapping) between genome (A) and associated network architecture and the epigenetic landscape (D) via the dynamics of the expression patterns (B) in state space (C) controlled by the GRN. The schematic representation is for a 9-gene GRN. The central concept to understanding the landscape is that each network state S (gene expression pattern, hence cell state = blue discs in B and C) maps into a point (= blue balls in D) on the landscape. The position of the point (network state) S, is determined as follows: The N gene expression values defining a given state S act as the coordinates in defining its position in that N-dimensional space, where each dimension (axis) represents the expression level of a gene. Each step (orange arrow) in entering a new (more abstract) conceptual domain (boxes A, B, C, D) can be formalized in terms of mathematical principles. In B, the two time points t1 and t2 represent the dynamics and the constrained change of gene expression pattern. Note that the quasi-potential is not a true potential energy since the gene network dynamics is a non-equilibrium, typically non-integrable system. The value of U can be intuitively (but formally not correctly) approximated by the negative logarithm of the steady-state probability P(S) to find the network in state S, i.e. U r v -ln [P(S)], or by decomposing the vector field that contains the forces F(S) that drive S into two perpendicular components, one of which is a gradient of some quasi-potential function U. Red circles = genes; blue axes = state space coordinates after hypothetical dimension-reduction to two dimensions, permitting the projection of the state space into an XY-plane (light blue in C, D), so that it can be used to display U as a third dimension.*

*Figure and legend reprinted from (Huang, 2012a) with permission provided by John Wiley and Sons (RightsLink licence agreement 3304950963715)*

## 1.2 Chromatin and Polycomb

### 1.2.1 Chromatin organization

In eukaryotic cells, the negatively charged DNA is wrapped around histones, which are small positively charged proteins that contribute to the compaction of the long DNA fibre in the cell nucleus. The complex between DNA and histones is called chromatin and it plays an important role in the process of gene regulation. The basic unit of chromatin is the nucleosome, which is formed by an octamer of the core histone proteins H2A, H2B, H3 and H4 (each one present in two copies) and 146 base pairs of DNA (Luger et al., 1997). The nucleosomes are separated by shorter stretches of DNA and the linker histone H1. However, the exact length of the DNA wrapped around a histone octamer, as well as the spacing between nucleosomes depends on the DNA sequence, the surrounding chromatin context and the transcriptional activity at a given DNA site. Furthermore, in the context of a living cell, chromatin is not just DNA and histones, there are thousands of protein and RNA molecules interacting with the chromatin in the cell nucleus. One of the leading scientists in the field of chromatin organization, Bas van Steensel, has summarized recent achievements in the field and identified two main principles that are responsible for the organization of the DNA-RNA-protein complex in the nucleus (van Steensel, 2011). The first one considers the three-dimensional architecture, which is driven by the physical and biochemical characteristics of the chromatin polymer. In particular, some of the elements of the 3D architecture are folding and compaction, as well as the local and long-range contacts between different parts of the polymer. The second principle is the chromatin composition, which defines the combination of proteins and RNA interacting with chromatin on the one hand, and the post-translational modifications of the histones on the other (referred to as “histone marks” when found on amino acids of the N-terminal histone tails).The existence of dozens of modifications on histone tails has even led some scientists to propose the idea of a “histone code” – a combination of PTMs that encodes an additional layer of information on top of DNA (Turner, 2002), however, this still remains a highly debated concept (Kouzarides, 2007a; Rando, 2012). But how does this really work in the living cell and what are the real dimensions of the chromatin complexity? The human genome is made up of three billion DNA

base pairs (3 Gb) and is present in two copies in somatic cells (diploid), resulting in 6 Gb of DNA. The DNA is wrapped around 30 million nucleosomes and in addition, there are an estimated one billion protein and several million RNA molecules present in the nucleus (van Steensel, 2011). And yet, the biochemical and biophysical forces, as well as the "hard wiring" encoded by the GRN result in a well-organized nucleus with clearly distinguishable types of chromatin. A systematic, genome-wide study in the fruit fly *Drosophila melanogaster* has classified chromatin into five categories, which are characterized by different chromatin composition and transcriptional activity (Filion et al., 2010). In the broadest terms, there are just two types of chromatin – active and repressed (in respect to gene expression), which correlate with the two classical forms of chromatin – euchromatin (active and open) and heterochromatin (silenced and condensed). Filion and colleagues have identified two subtypes of active chromatin – YELLOW and RED (Figure 1.2). YELLOW chromatin comprises mainly ubiquitously expressed genes and is associated with the active transcriptional machinery, as well as three histone modifications, indicative of actively transcribed genes: H3K4me2 (dimethylation of lysine 4 on histone H3), H3K36me3 (trimethylation of lysine 36 on histone H3) and H3K79me3 (trimethylation of lysine-36 on histone H3). RED chromatin is associated with a higher diversity of proteins, including many DNA-binding factors and chromatin remodelling proteins. Furthermore, RED chromatin displays not only a diverse protein composition but also a high concentration of molecules associated with a certain DNA locus. These genomic hotspots have been identified in the three organisms that were so far subject to large scale, systematic studies of chromatin interactions (Consortium, 2012; Gerstein et al., 2010; Roy et al., 2010). Furthermore, a map of the contacts between chromatin regions in mouse and human ESC, has revealed that there are multiple topological domains on each chromosome, which are the basic units of the 3D architecture of chromatin (Dixon et al., 2012). Certain topological domains can be organized together in a condensed chromatin state, facilitated by the Polycomb group proteins (PcG) (Cheutin and Cavalli, 2014; Denholtz et al., 2013; Isono et al., 2013). These 3D structures, referred to as Polycomb bodies (Pirrotta and Li, 2012), largely correspond to the BLUE chromatin from Filion's model (Filion et al., 2010) and comprise many genes involved in differentiation.

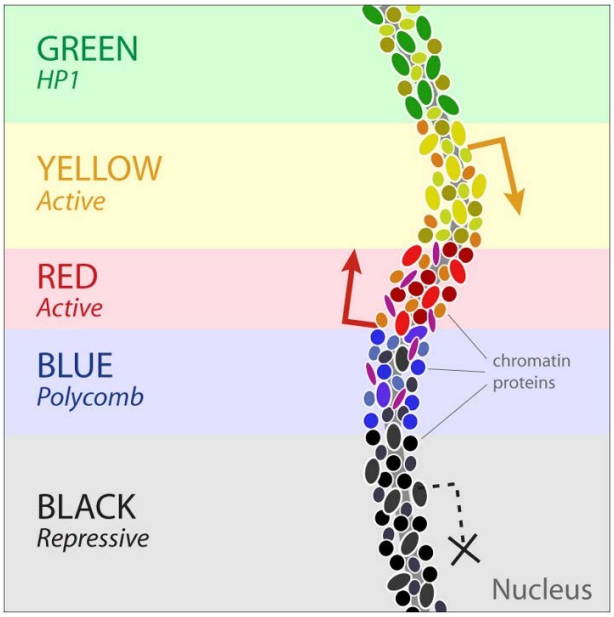



***Figure 1.2. Chromatin types in Drosophila***

*In a systematic genome-wide ChIP study, analysing the chromatin localization of 53 proteins in Kc167 embryonic cells, Filion and colleagues identified five distinct types of chromatin (Filion et al., 2010), which they termed as RED, YELLOW, BLUE, GREEN and BLACK (to fit with the etymology of "chromatin", a term coined in 1879 by Walther Flemming after using basophilic dyes, which strongly stained chromatin). According to this study, BLUE chromatin comprises the Polycomb targets bound by multiple PcG proteins and marked by H3K27me3. In the same time, H3K27me3 was found also in one of the two actively transcribed chromatin types (RED), arguing that H3K27me3 is not strictly associated with Polycomb-repressed targets.*

*Figure reprinted from (Filion et al., 2010) with permission provided by Elsevier (RightsLink licence agreement 3310430008253)*

### 1.2.2 The molecular diversity of mammalian PcG complexes

Polycomb group proteins (PcG) are chromatin-associated factors that are involved in the transcriptional regulation of many developmentally important genes. Historically, the first mutations related to Polycomb function were described in the 1940s by *Drosophila* developmental geneticists P. Lewis and E. Slifer (Lewis, 1947; Slifer, 1942). Already in 1958, A. Hannah-Alava suggested that the Polycomb mutations are changing the patterns set in the embryo: *"...it seems likely that the extra sex comb factors act by changing the pre-pattern of the embryonic legs, and to this changed pre-pattern the cells in the male genotype respond by formation of sex comb teeth.*" (Hannah-Alava, 1958). In 1978, Polycomb proteins have been identified as suppressors of the homeobox genes (Hox genes) in *Drosophila* and an anterior-posterior

gradient of repression was proposed as explanation of the embryonic posteriorization in Polycomb mutants (Lewis, 1978). In other words, the pioneer research on Polycomb shows clear examples of cells that end up in a different GRN state upon removal of certain Polycomb proteins. In order to understand how Polycomb proteins interact with chromatin and influence the GRN, we need to first focus on the biochemical function and molecular composition of PcG proteins. Studies in different organisms and cell types have identified a huge diversity in the composition of PcG complexes (Schwartz and Pirrotta, 2013). From an evolutionary perspective, this could be explained with the very early origin of Polycomb proteins, which are found in almost all eukaryotes (plants, animals and fungi), including unicellular organisms (Shaver et al., 2010). The latter indicates that Polycomb proteins have existed in the last common unicellular ancestor and have had 1.6 billion years to evolve. There are two distinct complexes formed by PcG proteins – Polycomb repressive complex 1 (PRC1) and PRC2, which are both catalytically active (Figure 1.3).

#### 1.2.2.1 PRC1

Core components of PRC1 are Rnf2 and its homologue Ring1, which are E3 ubiquitin ligases and can mono-ubiquitinate lysine 119 on histone H2A (H2AK119ub) (Buchwald et al., 2006; Wang et al., 2004). Other core members of the PRC1 complex are Cbx proteins (Cbx2, Cbx4, Cbx6, Cbx7 and Cbx8), which give the complex its binding affinity primarily through DNA-binding and recognition of methylated histone residues, including binding to the PRC2-mediated H3K27me3 (Tardat et al., unpublished work). Also an essential component of PRC1 are Pcgf proteins (Pcgf1, Pcgf2, Pcgf3 and Pcgf4), of which Pcgf2 (also known as Mel18) and Pcgf4 (also known as Bmi1) have been thoroughly studied in respect to their role as a PRC1 constituent (reviewed in (Nestorov et al., 2013a)). Other factors that have been associated with PRC1 function are the DNA-binding protein Rybp (Tavares et al., 2012) and the H3K4- and H3K36-specific histone demethylase Kdm2b (Farcas et al., 2012). This myriad of PRC1 components leads to the formation of multiple PRC1 entities, which may represent cell-type specific complexes with distinct molecular and developmental functions (Gao et al., 2012; Nestorov et al., 2013a; Schwartz and Pirrotta, 2013).

#### 1.2.2.2 PRC2 and the H3K27me3 mark

PRC2 consists of the homologous enzymes Ezh1 and Ezh2, the chromatin-binding protein Eed and the scaffold protein Suz12. The catalytical SET domain of Ezh1 and Ezh2 can successively methylate lysine 27 on histone H3 (H3K27me, H3K27me2 and H3K27me3, respectively) (Cao et al., 2002; Müller et al., 2002). Furthermore, *in vitro* biochemical assays have shown that PRC2 can methylate lysine 26 on histone H1 (H1K26) (Kuzmichev et al., 2004, 2005) and there also have been reports for some non-histone targets of PRC2 (Huang and Berger, 2008). These findings raise the question which of the molecular functions of PRC2 are indeed relevant for development. Recently, the importance of H3K27 methylation for *Drosophila* development has been tested by replacing the substrate residue on H3 with an "inert" one (lysine to arginine mutation at position 27), leading to homeotic defect phenotypes similar to what has been observed in Polycomb mutants (Pengelly et al., 2013). Methylation of H3K27 by Ezh1/Ezh2 depends on the presence of both Eed and Suz12 (Ketel et al., 2005; Schmitges et al., 2011), which makes these three components essential and sufficient for the molecular function of PRC2. The Eed protein contains WD-repeats, which fold into a seven-bladed beta-propeller domain that provides binding surface to PRC2 (Han et al., 2007). Mammalian Eed has four different isoforms, which are the product of alternative translation start sites and have been associated with distinct PRC2 functions, particularly in respect to the H1K26 recognition and regulation of enzymatic activity (Kuzmichev et al., 2004, 2005; Schmitges et al., 2011). The third core member of PRC2, Suz12, has a VEFS domain and a C2H2 Zinc-finger domain. The VEFS domain facilitates binding to Ezh1/Ezh2 and also serves as an allosteric regulator of the methyltransferase activity (Ketel et al., 2005). In addition to the essential components of PRC2, a number of other proteins have been associated with PRC2 function and termed as PRC2 cofactors. One of them is Rbbp4 (also known as NURF55), which has a WD-repeat domain similar to Eed and binds Suz12 and the tail of histone H3 (Nowak et al., 2011; Schmitges et al., 2011; Song et al., 2008).

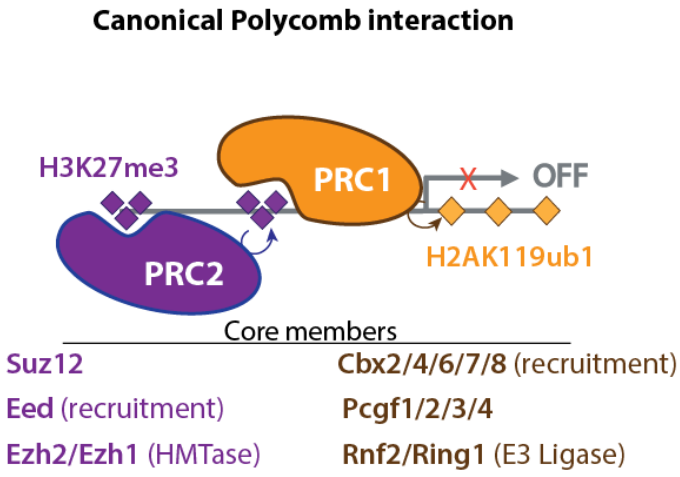



*Figure 1.3. Polycomb repressive complexes*

### 1.2.3 Polycomb targeting to chromatin in the mouse

Perhaps the most controversial question in the field of Polycomb is the mechanism by which PcG complexes are targeted to chromatin and how they repress transcription of a certain set of genes depending on the cell type. While in flies, PRC2 is targeted to chromatin by specific DNA-sequences (Polycomb responsive elements, PREs), in mammals it is still unclear whether Polycomb is targeted directly to DNA, or whether there are intermediate factors involved. The classical model of PcG silencing, depicted in Figure 1.3, postulates that PRC2 acts upstream of PRC1 to methylate H3K27 at target genes, which are subsequently silenced by the recruited PRC1 (Cao et al., 2002; Margueron and Reinberg, 2011; Simon and Kingston, 2009). However, there is a variety of other recruitment mechanisms, which also include independent targeting of the two PcG complexes (reviewed in (Nestorov et al., 2013a)). Recently, there have been reports of the reverse hierarchical model, where PRC1 is targeted to gene promoters through Kdm2b and subsequently PRC2 is recruited to the target genes by recognizing the PRC1-mediated H2AK119ub (Blackledge et al., 2014). The one common feature that these recruiting mechanisms have, is the preference for gene promoters containing unmethylated CpG islands (Jermann et al., 2014; Schwartz and Pirrotta, 2013).

Finally, there is also a self-perpetuating component to PRC2 recruitment, as H3K27me3 is recognized by the aromatic cage at the centre of the Eed WD-propeller structure (Margueron et al., 2009; Schmitges et al., 2011; Xu et al., 2010), therefore allowing enrichment and encroachment of PRC2 at sites that already have the PRC2 mark. This mechanism is potentially important for inheritance of the chromatin state.

## 1.3 The role of Polycomb during mouse development

### 1.3.1 Polycomb function and differentiation

Polycomb repressive mechanisms during differentiation might be paving the road to a given attractor (enforcing the developmental constraints), and once the cell has reached the basin, making it more difficult to switch states. This idea can be imposed on the observed homeotic phenotype in Polycomb mutants, where the constraints were not respected (or have been lowered) and a cell lineage that was pre-destined to become a posterior body part took the wrong way and became an anterior part. This transformation still respects the global epigenetic landscape, as it leads to an attractor state that is normally used. However, it could also lead to attractors that are not employed in normal development and form a new cell population (in a stable GRN state). In some cases the occupancy of unused attractors could be harmful for the organism, as is the case with cancer cells (Huang, 2012b). Indeed, changes in Polycomb function are often related to cancer (Albert and Helin, 2010; Richly et al., 2011).

Constitutive mutants for the core PcG genes obtained by crossing heterozygous parents display developmental failure around gastrulation. This has been shown for each of the PRC2 members Ezh2, Suz12 and Eed (Faust et al., 1998; O'Carroll et al., 2001; Pasini et al., 2004), as well as for the PRC1 gene Rnf2 (Voncken et al., 2003). Gastrulation is an early embryonic process, during which the three germ layers, ectoderm, endoderm and mesoderm, are set. The fact that PcG proteins are essential for this developmental event corroborate the role of Polycomb complexes in safeguarding cell identity during differentiation. Indeed, by using conditional knock-out models for Polycomb gene function that overcome the embryonic lethality, it has been shown that PRC1 and PRC2 take part in a variety of differentiation events.

For instance, a double knock-out for *Ezh1* and *Ezh2* in hair follicle stem cells revealed an essential function for PRC2 and H3K27me3 in the maintenance and differentiation potential of these stem cells (Ezhkova et al., 2011). A similar role for PRC2 was described also in hematopoietic stem cells upon removal of *Eed* (Xie et al., 2014). Transcriptional profiling of *Eed* knock-out HSCs highlighted multiple pathways that were regulated by PRC2, including differentiation, cell-cycle and apoptosis. In another study, Di Meglio and colleagues showed that *Ezh2* is responsible for the spatial organization of the precerebellar neurons in the mouse brain by regulating both intrinsic and extrinsic signals that guide the neurons (Meglio et al., 2013). These are just a few notable examples underscoring the complex role of PRC2 in development, which is determined by the broad spectrum of the regulated target genes. It also becomes evident that different cell types respond differently to loss of PRC2, which again brings up the GRN model discussed above (Huang, 2012a).

The role of PRC1 is more difficult to address due to the increased number of redundant members, as well as the existence of multiple variants of the complex. The latter fact makes it difficult to distinguish the developmental role of the different PRC1 variants or the cell type-specific roles of the redundant homologs. A study in hematopoietic stem cells tackled this issue from the perspective of the Cbx family proteins and revealed that *Cbx7* is responsible for maintaining the self-renewal capacity of the stem cells, while *Cbx2*, *Cbx4* and *Cbx8* are involved in the differentiation process (Klauke et al., 2013).

### 1.3.2 Polycomb function in the germline

The germline arises in the early mouse embryo around embryonic day 6.5 (E6.5), when a group of epiblast cells become induced as primordial germ cells (PGCs). Subsequently, the PGCs migrate and colonize the genital ridges at E11, followed by sex specification at E12.5 (PGC development is reviewed in (Saitou and Yamaji, 2012)). Female PGCs continue to proliferate until E13.5, followed by initiation of the first meiotic division and an arrest at the diplotene stage of prophase I. Arrested primordial oocytes reside in primordial follicles of the embryonic gonad until after birth when they are gradually triggered by hormonal waves to complete oogenesis. In contrast, male PGCs remain arrested at the G0/G1 cell cycle phase and resume proliferation

after birth as spermatogonial stem cells. The latter maintain a continuous pool of cells that enter meiosis and give rise to mature spermatozoa, thus displaying a significant difference between the male and female germlines- while the number of oocytes is limited to several thousand at birth (Sonne-Hansen et al., 2003), production of spermatozoa by the spermatogonial stem cells is virtually unlimited.

Germ cells are not only the vehicles of genetic information to the next generation, but they also transmit RNA, proteins, as well as DNA- and chromatin-borne information that is required to set the proper GRN state of the totipotent embryo. In particular, PGCs undergo erasure of genetic imprints prior to sex specific specification, followed by sex-specific reestablishment of imprinted DNA methylation loci (Saitou and Yamaji, 2012). Furthermore, there are also dynamic changes in histone modifications prior to sex specification. Between E7.5 and E9.5, during PGC migration, there is a strong reduction of H3K9me2 and a subsequent increase of H3K27me3 (Seki et al., 2007). Later, at E11.5, H3K27me3 and H3K9me3 are markedly reduced, which coincides with the genome-wide DNA demethylation (Hajkova et al., 2008). High levels of global H3K27me3 are re-established at E12.5. The rapid decrease of H3K27me3 at E11.5 has been suggested to depend on active demethylation by Utx (Kdm6a) and to play a crucial role in PGC development (Mansour et al., 2012). Finally, H3K27me3 is also associated with the inactive X–chromosome in females, which gradually loses H3K27me3 and gets reactivated by the end of PGC development at E13.5 (Chuva de Sousa Lopes et al., 2008; Sugimoto and Abe, 2007).

As evident, H3K27me3 is changing dynamically during the short time window of PGC development and disrupting this modulations leads to loss of the PGCs. The other Polycomb complex, PRC1, also seems to play an important role in PGCs. Ablation of PRC1 by removal of both *Ring1* and *Rnf2* resulted in loss of PGCs at E11.5 (Yokobayashi et al., 2013). Interestingly, reduced PRC1 activity, resulting from the deletion only of *Rnf2* but not *Ring1* (*Rnf2*$^{del/del}$ Ring1$^{+/-}$), revealed a sex-specific role for PRC1. At E13.5, *Rnf2*-deficient female gonads are depleted from germ cells, while male gonads are not affected. This is caused by the different wiring of retinoic acid signalling in the male versus female gonad and a requirement for PRC1 to counteract precautious activation of the pathway and subsequent entry into meiosis. Another study performed in the group of Antoine Peters demonstrated that PRC1 is required not only during the specification of germ cells, but also during oocyte growth and maturation – a process

associated with massive production of RNA. Deletion of both *Ring1* and *Rnf2* in early oocytes leads to the aberrant expression and accumulation of several thousand genes, which ultimately blocks embryonic development at the 2-cell stage (Posfai et al., 2012).

### 1.3.3 Potential role of Polycomb in epigenetic inheritance

The term "epigenetic inheritance" refers to the stable transmission of information that is not encoded by DNA across one or more generations. It is important to distinguish between intergenerational and transgenerational inheritance. In the case of traits induced by external signals it has to be considered that not only the parent animal (F0) is exposed to the signal, but also the germ cells that give rise to F1 could be affected, so transmission of the trait to F1 could be explained by the direct effect of the signal on the sperm and oocyte. Furthermore, in the case of inheritance via the female germline in mammals, the future F2 generation could be affected via *in utero* exposure of the F1 PGCs. Therefore, epigenetic inheritance is considered to be intergenerational and potentially influenced by the initial signal between F0 and F1 via the male germline and between F0 and F2 via the female germline. Occurrence of the acquired trait in later generations and in absence of the external signal is a case of transgenerational inheritance.

So far, most of the evidence for the propagation of acquired traits over multiple generations comes from plants and worms. The underlying molecular mechanisms are diverse and vary between different organisms. In plants, a complex mechanism is at place, involving transcriptional regulation of transposable elements and neighbouring loci via DNA methylation. Hypomethylated loci can be transmitted in plants through mitosis and meiosis for more than eight generations, bearing information for traits such as flowering time and root length (Heard and Martienssen, 2014). Perhaps one of the reasons for the proneness of plants to epigenetic inheritance is the somatic origin of the germline, along with the partial epigenetic reprogramming (only some of the DNA methylation is actively erased), which leaves room for the inheritance of acquired epigenetic states. In contrast to plants, worms and flies do not have DNA methylation and instead utilize a different molecular mechanism for epigenetic inheritance that involves small RNAs. Interestingly, even though the molecular pathways differ

between flies, worms and plants, they share a common feature – requirement and involvement of histone methylation of H3K9 (Heard and Martienssen, 2014).

In the mouse, only a few examples of epigenetic inheritance affecting endogenous genes have been documented, most notably the agouti viable yellow ($A^{vy}$) and the axin fused ($Axin^{Fu}$) alleles (Daxinger and Whitelaw, 2012). In all instances, the described epialleles were associated with transposable elements and presumably DNA methylation as an underlying mechanism. There are also a number of studies that demonstrate epigenetic effects on the expression of transgenes, as well as several reports that document heritable epigenetic changes caused by external signals. However, the observed effects rarely pass beyond F2 (and are thus intergenerational) and some of the studies have led to controversial results, which altogether shows that the question regarding the existence of transgenerational epigenetic inheritance in mice is still open (Daxinger and Whitelaw, 2012; Heard and Martienssen, 2014).

So what could be the role of histone modifications and Polycomb in this process? As discussed above, the epigenetic reprogramming in the germline affects both DNA methylation and histone modifications. Epigenetic inheritance occurs due to an escape of the epialleles from germline reprogramming, hence histone modification could also be prone to become epialleles. The importance of resetting the histone modification state in the germline, has been suggested by functional studies in mice lacking the H3K27 demethylase *Utx/Kdm6a* (Mansour et al., 2012), in worms deficient for the H3K4 demethylase LSD1 (Katz et al., 2009), as well as in plants lacking the H3K27 demethylase ELF6 (Crevillén et al., 2014). Genome-wide ChIP experiments in mouse and human sperm have revealed that the small amount of retained histones (10% in human and 1% in mouse) are specifically localized at the promoters of key developmental genes and also carry histone modifications, including H3K27me3 (Brykczynska et al., 2010; Erkek et al., 2013; Hammoud et al., 2009). Furthermore, H3K27me3 can be retained in the absence of the enzymatic PRC2 complex over several rounds of cell division (Gaydos et al., 2014; Puschendorf et al., 2008). However, these observations were made only on the global chromatin level and it remains to be tested whether and how H3K27me3 is retained at gene loci.

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

The majority of eukaryotic cell types share a common feature – the cell cycle. This is a highly regulated process that involves phases of growth, followed by cell division when the cellular components and the genetic material are distributed to the two daughter cells. In regard of these processes, there are two cell types that do not follow the common model. Oocytes are unusually large cells that grow without dividing and complete the asymmetric meiotic division only after fertilization, giving rise to the zygote. In contrast, the zygote starts off as a very large cell, which divides multiple times without growing in a process known as cell cleavage. These distinct features of the oocyte and preimplantation blastomeres are accompanied by a unique chromatin conformation in the zygote, major transcriptional shift, as well as dynamic changes in DNA and histone modifications during the cleavage stage (Figure 1.4). From a developmental perspective, preimplantation development is associated with the establishment and maintenance of pluripotency – the ability to form all embryonic germ layers. If we consider also the germline development, which starts shortly after implantation from a pool of pluripotent cells, we could envision both developmental events as a pluripotency life cycle, during which the involved cells prepare, establish and maintain the ability to form an embryo (Leitch and Smith, 2013).

### 1.4.1 Chromatin rearrangements in the zygote

In sperm, the DNA is very densely packaged, which is due to exchange of more than 95% of the nucleosomes for protamines. Notably, the small fraction of remaining histones carries specific epigenetic modifications and occupies the promoters of developmental genes both in mouse and human (Brykczynska et al., 2010; Hammoud et al., 2009). Upon fertilization, the male genome remains spatially separated from the female and undergoes rapid *de novo* chromatin formation (Mayer et al., 2000). The male pronucleus gets loaded with maternally provided histones, including the non-canonical histone variant H3.3 (van der Heijden et al., 2005).

Another unique feature of the mouse zygote is the organization constitutive heterochromatin, which consists mainly of AT-rich, repetitive DNA sequences. It has been shown that the pericentromeric satellite repeats are arranged around nucleolar-like bodies (NLBs) in the two pronuclei (Probst et al., 2007). In contrast, in somatic cells, constitutive heterochromatin is organized in a number of small densely-packed regions called chromocenteres. The histone modifications that are usually associated with constitutive heterochromatin, i.e. H3K9me2/3, H4K20me2/3, H3K64me3 as well as H3K79me2/3, are present only in the female pronucleus, while the male pronucleus acquires modifications of facultative heterochromatin instead (Arney et al., 2002; Daujat et al., 2009; van der Heijden et al., 2005; Liu et al., 2004; Ooga et al., 2008; Santos et al., 2005). In fact, the male pronucleus histones are largely hypomethylated until after DNA replication at the late zygote stage. The maternal histones that form the paternal chromatin are initially hyperacetylated and gradually acquire monomethylation modifications, including PRC2-mediated H3K27me1 (Puschendorf et al., 2008; Santos et al., 2005). It is suggested that the hypomethylation status of the male chromatin and especially the lack of H3K9me3 could be linked to another parent-specific process – active DNA demethylation (Burton and Torres-Padilla, 2010). Both genomes are subject to DNA demethylation up to the blastocyst stage, which happens passively along with DNA replication. However, the male pronucleus appears to undergo active hydroxymethylation in the zygote (5mC to 5hmC conversion), which is mediated by Tet proteins (Oswald et al., 2000; Wossidlo et al., 2011). This process adds to the asymmetry between the male and the female pronuclei in the zygote.

Overall, the male and female genomes not only come in a different conformation, but are also subject to different chromatin regulation mechanisms in the zygote and subsequent preimplantation stages. Specifically, the male genome becomes organized in a more open configuration, which may be a prerequisite for proper epigenetic programing and also a preparation for the zygotic genome activation.

### 1.4.2 Zygotic genome activation

The dynamic events that take place in the early zygote are fully driven and carried out by maternal factors, which were accumulated and stored in the oocyte. During this time the two

parental genomes remain transcriptionally silent and the cell is under maternal control. The transition from maternal to zygotic gene expression begins at the end of the zygote stage with a minor activation of transcription (minor ZGA) and goes on at the 2-cell stage with the major ZGA, while at the same time maternal mRNAs are actively degraded (Aoki et al., 1997; Schultz, 2002; Wang and Dey, 2006). There are two main components that underlie the maternal-zygotic transition: degradation of maternal transcript and transcriptional activation (Tadros and Lipshitz, 2009). The maternally provided RNA is rapidly degraded until the two cell stage (Pikó and Clegg, 1982), with some specific transcripts being selectively degraded already in the zygote (Alizadeh et al., 2005). The latter group includes the oocyte-specific genes Gdf9, H1foo, Mos and Hprt.

In parallel to the maternal RNA degradation, the transcriptional machinery is set at place and gradually activates the zygotic genome. A recent whole-transcriptome study in preimplantation embryos gave a detailed view on the transcriptional changes taking place around ZGA (Park et al., 2013a). The RNA-seq profiling allowed for the distinction between intronic and exonic expression, which could be used to assess the level of *de novo* transcription of a given gene. The findings by Park et al. confirmed the two waves of transcriptional activation, as described in earlier studies (Aoki et al., 1997; Schultz, 2002). It also revealed a large number of non-coding RNAs that are affected by the maternal-zygotic transition. Furthermore, Park et al. identified a number of transcription factors that regulate emerging gene networks, including Myod1, Sox9, Sox18, Mafb, Egr3, Runx1, Nkx2-5, Foxd1, Hnf1a and Nfatc2. All of these genes show *de novo* expression in the zygote and were not detected in the oocyte, hence they are either activated early in the zygote or are supplied by the sperm. The functional relevance of these findings remains to be tested experimentally.

Finally, there is evidence that ZGA is accompanied and requires chromatin remodelling. It has been shown that maternal Brg1 (also known as. Smarca4), a component of the SWI/SNF chromatin remodelling complexes, is essential for the genome activation (Bultman et al., 2006). Other chromatin remodelling factors, Tif1alpha and Snf2h, are also needed for initiation of the ZGA (Torres-Padilla and Zernicka-Goetz, 2006).

### 1.4.3 Embryo patterning and first cell fate decisions

The first cell fate decisions in embryos are often dictated by the mother and involve asymmetric cell divisions, during which specific factors segregate into a subset of the embryonic cells or form morphogen gradients. For instance, the oocytes of the fruit fly *D. melanogaster* show localization of the *bicoid* mRNA and subsequent formation of a Bicoid protein gradient, which is responsible for the specification of anterior cell fate in the early fly embryo (Frohnhöfer and Nüsslein-Volhard, 1986). Following this ground-breaking discovery, it has been shown in various other organisms that localized maternal mRNAs play a role in embryo pattern formation and cell fate specification (Palacios and Johnston, 2001). Perhaps the most striking case of maternally-driven cell fate specification in early embryos comes from the crustacean *Parhyale hawaiensis*, where all embryonic layers are invariantly and irreversibly set during the maternal control of the embryo (Gerberding et al., 2002; Nestorov et al., 2013b). Hence, in many eukaryotic organisms the oocyte and early embryos are pre-patterned and the first cell lineage specification is pre-determined to some extent. In contrast, mammalian organisms display a more flexible type of development, where the early blastomeres are totipotent and do not differ in their developmental potential. The difference between the pre-patterned embryogenesis and the plastic early development has led to a general classification of embryos as mosaic, i.e. pre-patterned and relying on asymmetric distribution of transcription factors, or regulative, which are more dependent on signalling cues and have the potential to modulate the transcriptional program if needed. Of course, there are rarely black-and-white situations when it comes to biological processes and a more closer look at early embryos reveals that mosaic embryos display some features of regulative embryos and *vice versa* (Lawrence and Levine, 2006).

A recent lineage tracing study in mouse embryos indicates that the blastomeres have a preference towards one of the lineages already at the 4-cell stage, well before the late blastocyst when the three distinct lineages are irreversibly defined (Tabansky et al., 2013). This finding corroborates earlier results suggesting a link between the spatial orientation of 4-cell blastomeres and their cell fate (Piotrowska-Nitsche and Zernicka-Goetz, 2005). Subsequently, the first asymmetric cell divisions happen at the 8-cell stage and cell polarity depends on a sophisticated and yet not well understood network of cell skeleton components, signalling

molecules and transcription factors (Ajduk et al., 2014). At the 16-cell stage there are already two distinct cell populations - smaller inner cells, surrounded by larger outer cells. The differences between the inner and outer cells become more pronounced and stable upon the first cell fate specification event in the early blastocyst, with the formation of the pluripotent inner cell mass (ICM) and the trophectoderm (TE). The master transcription factors that are associated with the first two lineages are Oct4, Sox2, Nanog and Cdx2, which are still co-expressed in all cells of the 8-cell embryo and segregate to the respective lineages only in the 16- to 32-cell embryo (Guo et al., 2010). While the pluripotency transcriptional network has been described *in vitro* in embryonic stem cells, the hierarchy among the master regulators *in vivo* is not well understood. For instance, genetic studies have indicated that Oct4 is not required for the establishment but only for the maintenance of pluripotency in preimplantation embryos, suggesting that another factor is responsible for the establishment of the totipotency-pluripotency state (Nichols et al., 1998; Wu and Schöler, 2014).

Upon compaction, the gene regulatory networks of the inner and outer cells start to change. This process involves the Hippo signalling pathway, which is suggested to sense the compaction state of cells by activating the signalling cascade in the more densely-packed inner cells (Bergsmedh et al., 2011). In the Hippo-off outer cells, the kinases Lats1/2 are not active and do not interfere with the cytoplasmic Yap1 transcriptional co-activator (Nishioka et al., 2009). This leads to relocation of Yap1 to the nucleus and activation of Tead4, which is in turn an activator of Cdx2 (Vassilev et al., 2001; Yagi et al., 2007; Zhao et al., 2008). The stabilized and amplified expression of Cdx2 induces the activation of further differentiation factors like Elf5 and Eomes, and at the same time suppresses the expression of the pluripotency transcription factors Oct4, Sox2 and Nanog (Ng et al., 2008; Strumpf et al., 2005). This ultimately leads to the irreversible commitment of the outer cells to the TE lineage. On the other hand, the inner cells display active Hippo signalling, which causes the phosphorylation of Yap1 and prevents its nuclear activity. Thus the ICM cells remain in their naïve pluripotent state.

In the second cell fate specification event during preimplantation, the inner cells give rise to the epiblast (EPI) and the primitive endoderm (PE). The main determinant of this developmental choice is the differential expression of Nanog and Gata6, which is regulated by the Erk/MAPK signalling pathway (Chazaud et al., 2006). The Gata6-positive cells will start expressing Sox17,

which maintains the cell fate and induces expression of two more transcription factors – Gata4 and Sox7 (Artus et al., 2011). During the lineage commitment process, the prospective EPI and PE cells are initially intermingled in the ICM. Subsequently, the Gata6/Gata4/Sox17-positive cells migrate to the surface of the ICM towards the blastocyst cavity and form a monolayer, which is the initiation of epithelium formation (Plusa et al., 2008). Sox7 gets activated in polarized monolayer cells, while at the same time all Gata6-positive cells that are still on the inside undergo selective apoptosis (Meilhac et al., 2009). After the two extra-embryonic lineages PE and TE are formed, the embryo is ready for implantation and subsequent differentiation of the epiblast during gastrulation.

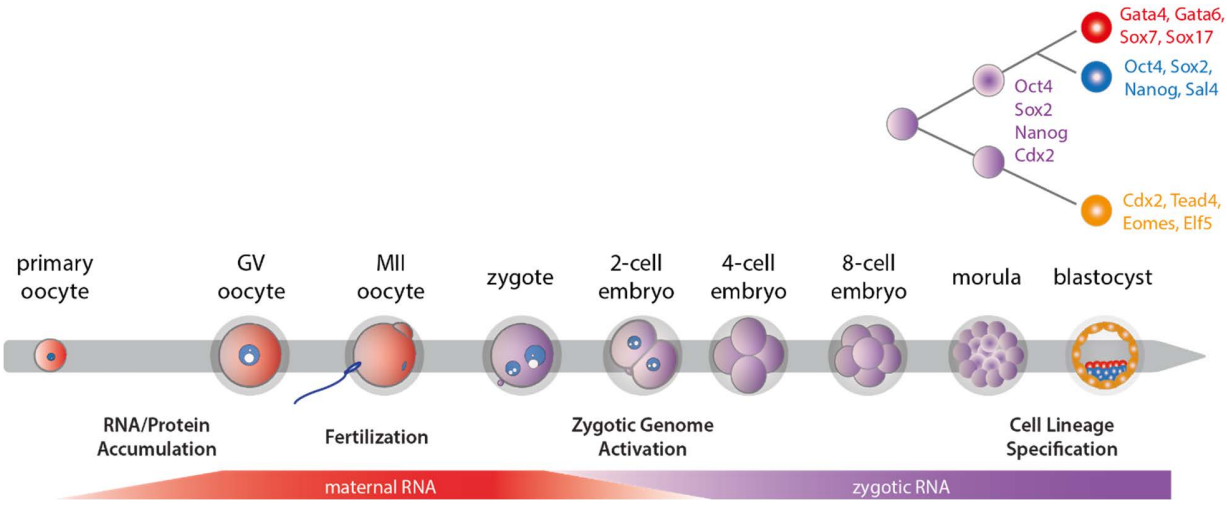



***Figure 1.4. Major events during pre-implantation development***

*Schematic overview of mouse early embryogenesis in respect to global gene expression events and lineage specification. Fusion of the two gametes forms the zygote (red: maternally provided messages; violet: zygotic transcripts). Lineage specification occurs in two steps, ultimately leading to the formation of three distinct lineages in the blastocyst: the epiblast (blue), the primitive endoderm (red) and the trophectoderm (yellow). The key transcription factors regulating the cell fate decisions are shown next to the respective cell lineage.*

The establishment of the first cell lineages in mouse embryos is accompanied by changes in chromatin, which is visible already at the global organization of chromatin. The differentiated outer cells have a more compact nucleus with heterochromatic foci at the periphery of the nucleus, while the pluripotent cells show a looser chromatin organization, dominated by

euchromatin (Ahmed et al., 2010). Furthermore, some post-translational histone modifications also differ between the early lineages. In the blastocyst, H3K27me3 is globally enriched in the ICM cells and in female embryos it marks the inactive X chromosome (Erhardt et al., 2003; Puschendorf et al., 2008). An example for an early asymmetry, which may be setting the state for the upcoming GRN change is the differential methylation of H3R26 (H3R26me2) in 4-cell embryos (Torres-Padilla et al., 2007). It has been shown that this modification depends on the activity of Carm1 and that aberrant enrichment of H3R26me2 promotes the pluripotent cell fate.

In addition to the chromatin changes in the blastocyst, there is also a difference in the global DNA methylation status. As mentioned before, the levels of 5mC-DNA decrease gradually from the 2-cell through the morula stage. Subsequently, DNA methylation is re-established in the ICM cells first (Santos et al., 2002). There is also one example for a direct link between gene expression and DNA methylation that affects lineage specification - the Elf5 promoter gets methylated in the ICM cells, which represses the TE differentiation program (Ng et al., 2008).

## 1.5 Scope of the thesis

The mechanism of Polycomb-mediated gene repression has been studied in detail in pluripotent ESCs and a link has been suggested between the transcription factor network and PcG proteins (Boyer et al., 2006; Endoh et al., 2008; Mohn et al., 2008). But what is the role of PcG proteins in establishing pluripotency *in vivo*? In order to determine the links in the regulatory network and the position of the players, one needs to study the molecular network as it emerges.

The pluripotent GRN state is established in a dynamic developmental time window during preimplantation development, which offers an *in vivo* model to address the role of Polycomb repressive mechanisms in establishing and maintaining pluripotency. The asymmetries between the blastomeres of the preimplantation embryos become evident as early as the 4-cell stage with a global change in H3R26 methylation. The PRC1- and PRC2-mediated H2AK119ub

and H3K27me3 respectively, are rapidly established on the paternal genome in the zygote and later become enriched in the ICM cells. Other chromatin modifications, as well as DNA methylation also change dynamically during preimplantation. However, it is not yet clear whether the chromatin changes are causative for the lineage specification, or whether they act downstream of the transcriptional and signalling machinery. A potential causal role is suggested by the specific retention of nucleosomes bearing H3K27me3 at the promoters of developmentally regulated genes in sperm (Brykczynska et al., 2010; Erkek et al., 2013; Hammoud et al., 2009). There is a link between Polycomb and preimplantation also coming from the oocyte side, where PRC1 regulates the accumulation of maternal RNA and thus plays a crucial role for the initiation of embryonic development (Posfai et al., 2012). Furthermore, several loss-of-function studies indicate the importance of Ezh2 around gastrulation (Erhardt et al., 2003; O'Carroll et al., 2001). Finally, gain of PRC2 function in preimplantation embryos has been reported to suppress the TE lineage through specific transcriptional repression of Cdx2 and Gata3 (Saha et al., 2013).

Considering the importance of Polycomb in development and the potential transgenerational role it may play, I set out to reveal the function of PRC2 in preimplantation embryos. Since Polycomb proteins are mainly associated with the regulation of gene expression in the various organisms and systems studied to date, it is logical to predict that PRC2 may have an effect on the main developmental events that are accompanied by transcriptional changes. There are three such events in oogenesis and preimplantation development: the accumulation of maternal RNA during oocyte growth, the shift from maternal to zygotic expression during ZGA around the 2-cell stage, and the first lineage specification event around the 16-cell stage (Figure 1.4). Based on the existing experimental evidence about PRC2 and the developmental characteristics of the system that I address, I formulated three hypotheses to test in my thesis:

***PRC2-mediated repressive mechanisms:***

***(1) Regulate the accumulation of maternal RNA in the oocyte;***

***(2) Regulate zygotic genome activation in mouse embryos;***

***(3) Are involved in cell fate specification during mouse pre-implantation development***.

I took a cre-lox based approach to abolish the function of PRC2 in the female and male germlines and thus generated maternally and zygotically deficient embryos. I used two independent models, an *Eed* knock-out and an *Ezh1/Ezh2* double knock-out respectively, in combination with two different cre drivers in the female germline and one in the male germline. The *Ezh1/Ezh2* double knock-out was aimed at overcoming the reported redundancy between the homologous genes *Ezh1* and *Ezh2* (Ezhkova et al., 2011; Margueron et al., 2008), while at the same time it gave the possibility to address possible dosage dependent effects. The use of different cre divers combined with the *Eed* conditional allele allowed for addressing the temporal component of PRC2 activity during oogenesis and its effect on preimplantation development. From the developmental perspective, I wanted to know whether PRC2-deficient embryos can reach the blastocyst stage, and if not, at which stage during preimplantation the problems occur. From the molecular perspective, I was mainly interested in the transcriptional output of the oocytes and early embryos in the absence of PRC2, as this should be directly influenced by the loss of PRC2. Another important question that I asked was about the dynamics and role of the PRC2-mediated mark H3K27me3 during oogenesis and in preimplantation embryos.

Finally, at the time when I started my thesis, the generally accepted model of Polycomb repression implied that PRC2 is upstream of PRC1 (Figure 1.3). This motivated my interest to describe the PRC1 activity in PRC2 mutant embryos, as well as to study a further mouse model, which lacks the major components of each PRC1 and PRC2 (*Rnf2* and *Ezh2* respectively) but still retains the Polycomb activity through the expression of the respective homologs *Ring1* and *Ezh1*.

# Chapter 2. Published review: H3K9/HP1 and Polycomb: two key epigenetic silencing pathways for gene regulation and embryo development

Peter Nestorov*,†,1, Mathieu Tardat*,1, Antoine H.F.M. Peters*,†,2

*Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland †Faculty of Sciences, University of Basel, Basel, Switzerland 1Equal contribution.

2Corresponding author: e-mail address: antoine.peters@fmi.ch

Current Topics in Developmental Biology

*ISSN 0070-2153* *http://dx.doi.org/10.1016/B978-0-12-416027-9.00008-5*

Contents

**Abstract**

Proper development of an embryo requires tightly controlled expression of specific sets of genes. In order to generate all the lineages of the adult, populations of pluripotent embryonic stem cells differentiate and activate specific transcriptional programs whereas others are shutdown. The role of transcription factors is obvious in promoting expression of such developmental genes; however maintenance of specific states throughout cell division needs additional mechanisms. Indeed, the nucleoprotein complex of DNA and histones, the chromatin, can act as a facilitator or barrier to transcription depending on its configuration. Chromatin-modifying enzymes regulate accessibility of DNA by establishing specific sets of chromatin, which will be either permissive or repressive to transcription. In this review, we will describe the H3K9/HP1 and Polycomb path- ways, which mediate transcriptional repression by modifying chromatin. We discuss how these two major epigenetic silencing modes are dynamically regulated and how they contribute to the early steps of embryo development.

## 2.1 Introduction and evolutionary perspective

The eukaryotic genome is organized in the nucleus of a cell as chromatin - a dynamic and highly organized DNA-histone complex. In the nucleus of an interphase cell, chromatin appears as two distinct subtypes referred to as euchromatin and heterochromatin. Euchromatin is decondensed, gene-rich, replicates early and is enriched for histone post-translational modifications (PTM) associated with active transcription (i.e. H3K4me2/3, H3K36me2/3). In contrast, heterochromatin remains compacted throughout the cell cycle, contains repetitive sequences, replicates late in S phase and is enriched for DNA methylation and histone PTMs related to transcriptional silencing (i.e. H3K9me3, H4K20me3) (Probst et al., 2009). These different chromatin marks have been intensively mapped to the genome of various model organisms and correlated to key biological processes (Campos and Reinberg, 2009; Consortium,

2012; Gerstein et al., 2010; Kouzarides, 2007b; Roy et al., 2010). Multicellular organisms originate from a single totipotent cell, the zygote that gives rise to a variety of cell types, which share the same genome but differ greatly in their morphology, function and chromatin structure. This developmental diversity is achieved by complex genome regulation, involving transcription factors and chromatin modifiers.

In this review, we will specifically focus on two major chromatin repressive pathways. The first one is the H3K9/HP1 pathway, which involves lysine-specific methyltransferases (KMTs) that methylate H3K9, a mark recognized by the chromodomain (CD) containing family of HP1 proteins. The second one is the Polycomb repressive pathway, which in mammals is classically related to gene silencing by two distinct complexes, Polycomb repressive complexes 1 and 2 (PRC1 and PRC2) and the associated PTMs H2AK119ub1 and H3K27me3 respectively.

Both pathways are conserved throughout eukaryotic evolution, since homologues for the core components are found in animals, plants and fungi (Garcia et al., 2007; Krauss, 2008; Schuettengruber et al., 2007; Shaver et al., 2010). In fact, both H3K9 and H3K27 methylation have been found in unicellular algae and protozoa, which suggests that this type of chromatin regulation has a very early origin (Krauss, 2008; Liu et al., 2007; Shaver et al., 2010). Furthermore, it seems that the two pathways not only have a common origin but possibly shared the same biological function. In the protozoa *Tetrahymena thermophila*, Ezl1 (the homologue of the PRC2 KMT Ezh2) catalyses both H3K9 and H3K27 methylation and is required for the formation of constitutive heterochromatin (Liu et al., 2007). Another example of PRC2-dependent heterochromatin formation comes from the green alga *Chlamydomonas reinhardtii,* where the H3K27 methylation activity is needed for the silencing of repetitive sequences (Shaver et al., 2010).

During evolution and as a result of multiple genome duplication events, the two pathways diverged and were attributed with specific functions in multicellular organisms. Below, we will describe the most important developmental roles of these two pathways and how they contribute to gene regulation.

## 2.2 The H3K9/HP1 pathway and its role in development

The seminal discovery of enzymes mediating H3K9 methylation on pericentric heterochromatin (Rea et al., 2000) and thereby generating a binding site for HP1 (Lachner et al., 2001), emphasizes the role of this pathway in regulation of heterochromatin structure (Cheutin et al., 2003). The discovery of the Suv39h enzymes was soon followed by the identification of other H3K9-methylating enzymes (H3K9 KMT) (see Table I) catalysing different degrees of methylation and regulating repression of different classes of sequences (Krauss, 2008). The concept of division of labour applies also to HP1, as most eukaryotes express different isoforms of HP1 that differ in their nuclear localization (Zeng et al., 2010) and serve partially redundant as well as isoform specific functions (Rosnoblet et al., 2011). For example, these proteins have been implicated in maintenance of heterochromatin, chromosome segregation, transcriptional silencing/activation (Kwon and Workman, 2011; Schotta et al., 2003), DNA replication (Hayashi et al., 2009; Schwaiger et al., 2010), and the DNA damage response (Dinant and Luijsterburg, 2009).

### 2.2.1 Developmental role through regulation of gene expression.

#### 2.2.1.1 Su(var)

With their pleiotropic roles in diverse biological pathways, H3K9 modifying enzymes and HP1 proteins are of particular interest for embryonic development. Functional studies in several organisms revealed that with the exception of Suv39h1/2, removal of H3K9/HP1 components results in developmental defects in most organisms. *Suv39h1/2* double knock-out (DKO) mice are viable although smaller in size and exhibit chromosomal defects during meiosis in the germline (Peters et al., 2001), partial loss of DNA methylation on pericentric heterochromatin (PCH) (Lehnertz et al., 2003) and increased tumour risks (Braig et al., 2005; Peters et al., 2001). Embryonic stem cells (ESCs) deficient for *Suv39h1/2* can maintain stemness, proliferate and show enrichment of H3K27me3 on PCH, suggesting plasticity between Suv39h1/2 and PcG repressive pathways (Peters et al., 2003). Inactivation of *Suv39h* homologues in other organisms

does not strongly impair their viability as it was shown in *Drosophila* (Schotta et al., 2002; Tschiersch et al., 1994), plants (Jackson et al., 2002; Naumann et al., 2005) or even yeast (Allshire et al., 1995), with the exception of the fungi *Neurospora crassa*, where the mutation of *Dim-5* results in growth defects and sterility (Tamaru and Selker, 2001). Interestingly, in *Dim-5* mutant, both H3K9 and DNA methylation are affected, suggesting an interdependency of these two chromatin marks in this organism (for more detail about the link between DNA and H3K9 methylation, see Box1). This suggest that the Suv39h enzymes act principally as gatekeeper of genome integrity during development by regulating constitutive heterochromatin more than affecting gene transcription. Interestingly, a recent report demonstrated the requirement of transcription factors Pax3 and Pax9 for repression of pericentric transcripts and maintenance of pericentromeric heterochromatin (PCH) in mouse embryonic fibroblast (Bulut-Karslioglu et al., 2012). However, whether Pax transcription factors directly interact with Suv39h enzymes and how they could promote their targeting in a developmental manner remains an open question. Transcriptional regulation by Suv39h1 was recently pointed out in a report from Alder *et al*, suggesting that developmental genes targeted by the PcG proteins in mouse ESCs, are repressed in a Suv39h1-dependent manner in the trophoblast lineage (Alder et al., 2010). Knockdown of *Suv39h1* in cultured trophoblast stem cells (TSCs) resulted in decreased expression of the TSC key transcription factor Cdx2, and enhanced TSC differentiation. This suggests that, at least during the blastocyst stage of pre-implantation development, Suv39h enzymes could be required for proper trophoblast formation through specific gene regulation. In plants, SUPERMAN (a C2H2 type zinc finger protein), which is required for maintaining boundaries between floral organs in *Arabidopsis*, is controlled by the Su(var)3-9 homologue SUVH4/KRYPTONITE (KYP). KYP indirectly represses the SUPERMAN locus through the recruitment of the DNA methyltransferase CHROMETHYLASE3 (CMT3), allowing development of flowering organs (Jackson et al., 2002; Yun et al., 2002). However, the principal role of plant Su(var)3-9 homologues SUVH and SUVR seems to be the control of transposons in heterochromatin during plant development (Kuhlmann and Mette, 2012; Naumann et al., 2005; Thorstensen et al., 2011).

#### 2.2.1.2 SETDB1

During mouse pre-implantation development, transcriptional regulation seems to rely more on other classes of the H3K9 KMT including G9a, and ESET/SETDB1 and on the H3K9 KDMs of the JMJD2 family. Indeed, murine ESCs inactivated for SETDB1 or the two KDM JMJD2A and JMJD2C are unable to maintain self-renewal and differentiate (Loh et al., 2007). SETDB1 associates with the core pluripotency transcription factor Oct3/4 and regulates a specific set of developmental genes, most of them related to the trophectoderm lineage (Bilodeau et al., 2009; Cho et al., 2012; Lohmann et al., 2010; Yeap et al., 2009; Yuan et al., 2009). This set of genes is also targeted by PcG proteins (discussed again in the Polycomb section). This suggests cooperation between different epigenetic repressive pathways for maintaining the stemness state of ESC (Azuara et al., 2006; Boyer et al., 2006). Furthermore, SETDB1 association with the zinc finger transcription factor ZNF274, allow repression of ZNF genes (Frietze et al., 2010) whereas its association with the serine/threonine kinase Akt/PKB mediates repression of certain transcription factors such as Forkhead family member (Gao et al., 2007). This suggests that SETDB1 mediates transcriptional silencing of specific sets of genes depending on its binding partners. Zygotic expression of SETDB1 begins at the blastocyst stage, while the maternal transcript is present in the oocyte and persists throughout preimplantation development (Dodge et al., 2004). It would be of great interest to identify binding partners of SETDB1 that could modulate its targets and thereby its biological output during the earliest stages of embryo development. Despite this role in euchromatic regions, one striking feature of SETDB1 is the repression of transposable elements and repeats in mESC (Karimi et al., 2011; Matsui et al., 2010), a function achieved in cooperation with the co-repressor KAP1/Trim28 and HP1 proteins (Schultz et al., 2002; Sripathy et al., 2006), although the later appear dispensable for this function (Maksakova et al., 2011). Indeed, even if heterochromatin is more compacted and silent than euchromatin, there is evidence that transcription of repeats is tightly controlled (Eymery et al., 2009). SETDB1 is also implicated in the regulation of the structure of promyelocytic leukaemia-nuclear body (PML-NBs) and the transcription of its associated genes (Cho et al., 2011). PML-NBs that have been linked to many cellular processes such as apoptosis, DNA damage responses, and transcriptional regulation (Torok et al., 2009), expanding the many potential biological roles of SETDB1 for proper embryo development. Altogether, these multiple functions achieved by

SETDB1 could therefore explain in part the early phenotype seen in SETDB1 knock-out mice, which die around 3.5-5.5dpc (Dodge et al., 2004).

#### 2.2.1.3 G9a

G9a takes part in the control of genes like Mage-a (Tachibana et al., 2002), p21/waf1 (Nishio and Walsh, 2004), some imprinted genes in the trophoblast (Wagschal et al., 2008), interferon beta through its association with PRDM1 (human homologue of mouse Blimp1) (Gyory et al., 2004) and the key developmental regulators Oct3/4 and Nanog (Epsztejn-Litman et al., 2008; Feldman et al., 2006; Yamamizu et al., 2012). G9a is targeted to the promoter of the transcription factors Oct3/4 and Nanog, where it deposits H3K9me2 and allows the recruitment of HP1 and Dnmt3a/b. Gene silencing of *Oct3/4* and *Nanog* is tightly controlled by the joint action of APC/Cdh1-mediated degradation of G9a by the proteasome, and removal of the H3K9me2 mark from their promoters by the JMJD2A and JMJD2C KDM (Loh et al., 2007; Wang et al., 2010; Whetstine et al., 2006). Indeed, JMJD2C KO in mice results in down-regulation of *Oct3/4*, *Nanog* and *Sox2* mRNA (Wang et al., 2010), whereas G9a KO aberrantly prolongs expression of these genes up to embryonic day 7.5 (Yamamizu et al., 2012). Both mutant mouse models show developmental defects, suggesting that proper control of G9a mediated gene silencing is crucial for the embryo. In fly, dG9a was shown to be a suppressor of position effect variegation (PEV) (Mis et al., 2006), and although *dG9a* mutant flies show minor developmental defects and are viable (Seum et al., 2007a; Stabell et al., 2006), such role in gene regulation in the embryo could apply. In fact, dG9a overexpression affects transcription of genes involved in the pupal eye formation (Kato et al., 2008). In plants, the SUVR proteins are the most closely related to G9a (Baumbusch et al., 2001; Thorstensen et al., 2006), but so far, they have been described to function essentially in the repression of transposons and ribosomal DNA (rDNA) (Thorstensen et al., 2006; Veiseth et al., 2011).

#### 2.2.1.4 HP1

Given its affinity for H3K9 methylated residues, involvement of HP1 for gene silencing is anticipated (Bannister et al., 2001; Lachner et al., 2001). HP1 can induce compaction of a targeted loci (Verschure et al., 2005) with heritability of the repressed state over mitotic division (Ayyanathan et al., 2003). Indeed, artificial targeting of exogenous HP1α to a specific locus induces increased H3K9me3 overtime which can be maintained through cell division even in absence of further exogenous HP1α expression (Hathaway et al., 2012). However, it is now clear that HP1 isoforms, which show different localizations within the nucleus (Minc et al., 1999), do not completely share redundant function (Cammas et al., 2007). Some HP1 isoforms are required for either promoting or repressing transcription, or both (Vermaak and Malik, 2009; de Wit et al., 2007). Indeed, there is now substantial evidence arguing that HP1 can promote transcription (Kwon and Workman, 2011). The mammalian HP1β has been found to play a role in the control of the expression of rDNA genes by RNA polymerase I (RNAPI) in a Suv39h-dependent manner (Horáková et al., 2010), whereas HP1γ regulates transcriptional elongation through its association with RNA polymerase II (RNAPII) (Vakoc et al., 2005) and RNA processing (Smallwood et al., 2012). Moreover, HP1γ is part of a complex with the Polycomb group protein L3mbtl2, forming a PRC1-like complex involved in gene repression in mouse ESC and early embryogenesis (Qin et al., 2012; Trojer et al., 2011).

Interestingly, *Drosophila* HP1a (Su(var)2-5), which is required for fly development (Eissenberg et al., 1992; Kellum and Alberts, 1995; Kellum et al., 1995), may share some function in RNA processing as it was found to positively regulates euchromatic gene expression by interacting with the heterogeneous ribonucleoprotein (hnRNP) DDP1, HRB87F, and PEP (Piacentini et al., 2009). In *Drosophila,* HP1 isoforms are also involved in gene regulation. Indeed, HP1a (Su(var)2-5), positively regulates (indirectly or not) heterochromatic as well as euchromatic genes, like heat-shock (Hsp70) or cell-cycle related genes (Mcms, ORC4, CAF-1, Cdc45 and Aurora B). More generally, HP1a is a positive regulator of transcription by facilitating H3K36 demethylation via chromoshadow domain (CSD)-mediated recruitment of dKDM4A at active and/or heterochromatic regions (Lin et al., 2008, 2012). Either knockdown or over-expression of HP1b is lethal for fly development, where it seems to play an important role in transcription. HP1c associates with the transcription factors WOC (*without children*) and ROW

(*related of WOC*) via their PxVxL motif to regulate a common set of genes involved in nervous system development (Abel et al., 2009; Font-Burgada et al., 2008). Moreover, HP1c appears also to be required for proper recruitment of FACT (facilitates chromatin transcription) to chromatin (Kwon et al., 2010).

In addition to fly and mammals, both isoforms of HP1 in *C. elegans* (HPL-1, HPL-2) show preferential euchromatic localization, HPL-2 mutants showed that it is required for the regulation of germline genes, as well as sets of genes involved in lipid metabolism or innate immunity (Couteau et al., 2002; Meister et al., 2011; Studencka et al., 2012). The only HP1 homologue identified in *Arabidopsis thaliana* (LHP1/TLF2) also localizes to euchromatic domains, where it associates with PcG proteins to represses genes involved in plant development (Gaudin et al., 2001; Kotake et al., 2003; Nakahigashi et al., 2005; Turck et al., 2007).

The role of HP1 isoforms for gene regulation during mouse pre-implantation development is, however, still poorly understood. HP1β is predominantly expressed in oocytes and zygotes whereas HP1α appears at the 2-cell stage (Puschendorf et al., 2008). HP1γ is expressed later and throughout development (Meglicki et al., 2012). Investigation of germline knock-out mutant mice for either HP1 isoform would help us to understand better their role, if any, for proper transcriptional regulation during early embryo development.

### 2.2.2 Function in the germline.

The germline can be viewed as the immortal lineage of cells that gives rise to haploid gametes in sexually reproducing organisms. Germ cells undergo numerous DNA-directed events that must be tightly coordinated and controlled while these cells progress through their development. Several recent reports indicate that members of the H3K9/HP1 pathway have important biological functions in germline maintenance, differentiation and possibly in the process of meiotic silencing of unpaired chromosomes (MSUC) and meiotic sex chromosome inactivation (MSCI). These last two phenomena, collectively called "meiotic silencing," target sex chromosomes in the heterogametic sex (the X chromosome in male nematodes and the XY-body in male mice) and also any other chromosomes that fail to synapse due to mutation or

chromosomal rearrangement. Meiotic silencing is of crucial importance as it is hypothesized to maintain genome integrity (Turner, 2007; Zamudio et al., 2008). Many members of the H3K9/HP1 pathway are highly expressed in germ cells (van der Heijden et al., 2007; Khalil et al., 2004; Peters et al., 2001). Mouse *Suv39h2* is abundant in testes and in oocytes compared to *Suv39h1* which is more ubiquitously expressed in somatic tissues (O'Carroll et al., 2000; Puschendorf et al., 2008). *Suv39h1/2* double knockout (DKO) mice are viable, but display impaired spermatogenesis. Spermatocytes undergo apoptosis at the pachytene stage as a consequence of incomplete homologue pairing and synapsis defects (Peters et al., 2001). Interestingly, *G9a* germline conditional knock-out males are sterile while fertility is severely impaired in females (Tachibana et al., 2007). Mutant spermatocytes fail to progress through the pachytene stage, show defects in double-strand break (DSB) repair and undergo apoptosis. It can, however not be excluded that these defects are a consequence of misregulation of G9a target genes.

In mammals, knock-out of HP1γ induces a dramatic loss of the number of PGCs due to cell cycle defects (Abe et al., 2011). These animals are sterile and exhibit defects in centromere clustering and synapsis in spermatocytes (Brown et al., 2010; Naruse et al., 2007; Takada et al., 2011). Takada and colleagues further show that *Suv39h1/2* H3K9-dependent methylation at pericentric heterochromatin serves as a platform to recruit HP1γ, which then recruits G9a, highlighting a putative cooperative role between H3K9 KMTs and HP1 for meiotic progression. Strikingly, HP1γ and HP1β, but not HP1α associate with the transcriptionally silent XY body during male meiosis, suggesting possibly isoform specific functions (Metzler-Guillemain et al., 2003). Recently, it was shown that the SETDB1 homologue in worms, MET-2, is involved in MSCI and protects the germline from undergoing apoptosis (Checchi and Engebrecht, 2011). Such function has not yet been described for mammalian SETDB1 and SETDB2 proteins. In murine embryonic ovary and postnatal testis, *Setdb2* expression correlates with that of *Stra8*, a gene involved in the onset of meiosis in germ cells (Hogarth et al., 2011), suggesting a potential role for *Setdb2* in mitotic to meiotic transition in germ cells.

In the mouse female germline, HP1β is the predominantly expressed isoform, and so far, no report indicates that its inactivation induces defects in this lineage (Aucott et al., 2008). In *Drosophila*, among the five isoforms of HP1, *HP1d* and *HP1e* show a germline specific expression

in female ovary and male testis respectively (Vermaak et al., 2005; Volpe et al., 2001). *HP1d*, also named *Rhino*, was found in a screen for female sterile mutant flies. *HP1d/Rhino* mutant flies are characterized by defects in polytene chromosome structure of nurse cells and egg polarity defects. HP1d/Rhino is thought to act as a safeguard of the female germline against mobile elements through modulation of the piRNA pathway (Klattenhoff et al., 2009). Even though there is less information about HP1e, it is hypothesized to have a function similar to HP1d but in the male germline (Vermaak et al., 2005).

Among the H3K9 KMTs in *Drosophila*, only dSETDB1, has been shown to have a crucial role in the female germline (Clough et al., 2007; Koch et al., 2009; Wang et al., 2011; Yoon et al., 2008). *dSETDB1* is expressed most strongly at early stages of oogenesis, in germ cells in the germarium. *dSETDB1* mutant ovaries primarily exhibit germ cell differentiation defects and apoptosis in young females, and gradually the niche of germ stem cells (GSCs) is lost, indicating that dSETDB1 regulates both germ cell maintenance and differentiation. It has been proposed that dSETDB1 and Su(var)3-9 cooperate during germ cell maturation. As germ cells mature and differentiate, *dSETDB1* expression decreases and its function is gradually taken over by *Su(var)3-9* (Yoon et al., 2008). Some interplay between these two H3K9 KMTs is reinforced by the fact that the *dSETDB1* mutant phenotype is less severe in *Su(var)3-9*-deficient flies (Brower-Toland et al., 2009; Seum et al., 2007b). This puzzling phenotypic connection remains unresolved at the molecular level. In mammals, the exact function of SETDB1 in the germline is unknown and possible links between the different H3K9 KMT enzymes are not clear either. One study even suggested that Suv39h1, G9a, Glp and SETDB1 are part of the same complex (Fritsch et al., 2010), However, since these enzymes show different expression pattern during development, the molecular composition and function of such putative H3K9 methylation complex is still an open question.

Recently, PRDM3/16 were characterized as H3K9me1 KMTs that provide a template for Suv39-mediated H3K9me3 conversion at PCH (Pinheiro et al., 2012). Importantly, these proteins are essential for the clustering of pericentromeric regions into chromocenteres. Although using different enzymes, this pathway seems to be by largely conserved in *C. elegans*, where it is required for maintenance of the anchoring of heterochromatic regions to the nuclear periphery

(Towbin et al., 2012). The role of these enzymes, during germline and early embryonic development awaits further investigations.

### 2.2.3 New insights into the function of H3K9/HP1 pathway?

During the last decade, we gathered a better understanding of the function of H3K9 KMT and HP1 proteins. Although many points need further investigation, knockout and knockdown studies in various organisms allowed us to appreciate more their role for proper development. However, most of these experiments focus on the catalytic activity of KMT towards histones. Since its discovery, methylation of proteins involves many different substrates (Paik et al., 2007). For instance, G9a has substrates other than the canonical histones (Chin et al., 2007; Sampath et al., 2007). Therefore, identification of non-histone substrates for the other H3K9 KMT would be of great interest, as PTMs of proteins could lead to a variety of biological outputs.

Indeed, HP1 is dynamically post-translationally modified, which can affect its localization toward PCH (Maison et al., 2011) or following DNA damage (Ayoub et al., 2008; Dinant and Luijsterburg, 2009). Moreover, the different HP1 isoforms don't show similar localization, associate with different proteins and harbour non-completely redundant functions in various organisms. Thus, HP1 isoforms and their PTMs may constitute additional layer of information, increasing the complexity of genome control by chromatin modifying enzymes.

## 2.3 Polycomb repressive pathways

### 2.3.1 Composition and developmental role of PcG proteins

In recent years, the composition and variety of Polycomb complexes in different cell types and organisms have received substantial attention. Many of the core components of PRCs have been duplicated during evolution, but instead of acting as "spare parts" serving redundant functions, these paralogs acquired different developmental roles during their divergence. In addition, there are other components that associate with Polycomb complexes ("Polycomb cofactors"), which further increase the variety and have an effect on the functionality and the

targeting of the complexes in different cell types. Remarkably, the divergence of Polycomb complexes is not a rare event in evolution, since it occurred in plants, flies and mammals. Below we will focus on the composition and developmental roles of PRC1 and PRC2 in three model organisms.

#### 2.3.1.1 PRC2

##### 2.3.1.1.1 Drosophila

E(z), Su(z)12, Esc/Escl and Nurf55 are the core PRC2 components in *Drosophila* (see Table II), which together have been isolated as a 600-kDa complex (Müller et al., 2002; Tie et al., 2001). Pcl has been identified as part of a bigger Pcl-PRC2 complex (Tie et al., 2003), where Pcl functions as an enhancer of the H3K27 trimethylation *in vivo* (Nekrasov et al., 2007). In *Drosophila* larvae, Pcl facilitates the recruitment of PRC2 to chromosomes (Savla et al., 2008). Classical Polycomb phenotypes related to the misregulation of Hox genes have been described for all PRC2 members and with the exception of *esc* and *escl*, all homozygous null alleles show larval lethality (Anderson et al., 2011; Birve et al., 2001; Duncan, 1982; Phillips and Shearn, 1990; Struhl and Brower, 1982). According to the *Drosophila* developmental transcriptome project (Gelbart and Emmert, 2011; McQuilton et al., 2012), mRNA levels for all PRC2 genes peak in early embryos, then decline at larval stages and increase in female but not male adults. The only exception are the partially redundant paralogs *Esc* and *Escl* with *Esc* predominantly expressed in embryos and *Escl* in late larval/early pupal stages (Kurzhals et al., 2008). In addition, *Esc* and *Escl* mRNAs undergo splicing with different efficiencies and the two proteins control the enzymatic activity of PRC2 complexes differentially (Ohno et al., 2008). As a consequence, the maternal *Esc* contribution is required for development.

##### 2.3.1.1.2 Mammals

The mammalian PRC2 homologues are *Ezh1*/*Ezh2*, *Eed*, *Suz12*, *Rbbp4*/*Rbbp7* and *Pcl1*/*Pcl2*/*Pcl3* (see Table II). Homozygous mutations for *Ezh2*, *Eed* or *Suz12* are lethal in early post-implantation development (Faust et al., 1998; O'Carroll et al., 2001; Pasini et al., 2004),

while *Pcl2* mutant mice are viable but show posterior transformations as the result of *Hox* gene misregulation (Li et al., 2011). Gene expression data from mouse and human for the two KMT homologues *Ezh1* and *Ezh2* suggests that *Ezh2* is predominant in embryogenesis and in proliferating cells, while Ezh1 might be more important for postnatal development (Margueron et al., 2008; Shen et al., 2008). Biochemical characterization has identified differences in the activity of Ezh1-PRC2 and Ezh2-PRC2, implying a strong chromatin compaction activity and a comparably weaker H3K27me2/3 KMT activity for Ezh1-PRC2 (Margueron et al., 2008; Shen et al., 2008). We could speculate that the enzymatic activity of PRC2 is more needed early in development, when the chromatin is still "young" and changing, so keeping the right set of genes silent in a dynamic environment would require a more active repressor. On the other hand, a number of studies using *Ezh2* conditional mutants revealed that Ezh2 is required for the terminal somatic cell differentiation and in some cases also for the maintenance of the multipotent or progenitor cell state (reviewed in (Aldiri and Vetter, 2012)). In most of these cells, *Ezh2* is co-expressed together with *Ezh1*, so it is difficult to assign roles to one or the other. The developmental role of Ezh1-PRC2 is therefore still unclear, as it is either masked by the presence of Ezh2 or is not critical, since homozygous *Ezh1* mutant mice are healthy, fertile and do not show any transformations (Ezhkova et al., 2011). Another component of PRC2, *Eed,* does not have paralogs in mammals but instead has four isoforms, which are associated with three variants of the complex – PRC2, PRC3 and PRC4. These PRC2 variants show biochemical differences, as PRC2 and PRC4 are suggested to have the canonical H3K27- and an additional H1K26-KMT specificity, the biological significance of which remains unknown (Kuzmichev et al., 2004, 2005).

#### 2.3.1.1.3 Plants

Homologues of all four core PRC2 members are found in *Arabidopsis thaliana*: MEA, SWN, CLF (homologues of Ezh2); VRN2, FIS2, EMF2 (homologues of Suz12); the Rbbp4 homologue MSI1; and the Eed homologue FIE (see Table II). Furthermore, there is a group of three PHD-finger proteins – VRN5, VIN3 and VEL1, which are considered as functionally related to Pcl because they enhance the KMT activity of the complex (Lucia et al., 2008). As expected from the

existence of numerous paralogs, there are multiple PRC2 complexes, which are involved in different processes. The development of the female gametophyte and the initiation of embryogenesis are controlled by the MEA-FIS2-MSI1-FIE complex and mutation in any of the four genes is embryonically lethal (Grossniklaus et al., 1998; Köhler et al., 2003; Luo et al., 1999; Ohad et al., 1999). The PRC2 maternal effect phenotype is caused by imprinting defects in the endosperm (functionally analogous to the mammalian placenta) and derepression of MADS box TFs, which leads to hyperproliferation of the endosperm and eventually seed abortion. Furthermore, by expressing a FIE transgene in *fie* mutant plants, it was possible to overcome the maternal requirement for PRC2 and identify pleiotropic phenotypes as the result of derepression of KNOX homeobox genes (Katz et al., 2004). Another PRC2 variant consisting of CLF/SWN-EMF2/VRN2-MSI1-FIE is involved in the regulation of several key transcription factors (AG, FLC), which control the transition from vegetative-to-reproductive development and the cold-induced flowering response ((Coustham et al., 2012), reviewed in (Holec and Berger, 2012; Song et al., 2012)).

#### 2.3.1.2 PRC1

##### 2.3.1.2.1 Drosophila

The *Drosophila* PRC1 consists of Sce, Pc, Psc and Ph and similarly to PRC2 has been implicated in the *Hox* gene regulation (see Table III). The expression of PRC1 genes throughout fly development resembles the one of PRC2 and mutations lead to classical homeotic phenotypes and embryonic lethality (Breen and Duncan, 1986; Dura et al., 1985, 1987; Graveley et al., 2011; Jürgens, 1985; Lewis, 1978, 1947; McQuilton et al., 2012). More recently, Sce and Psc have been identified as members of a distinct complex, dRING-associated factors (dRAF), containing also the H3K36-demethylase dKDM2 (Lagarou et al., 2008). The *kdm2* mutant allele significantly enhances the *Pc* homeotic phenotype and in the same time rescues the *Trx* and *Ash1* mutations (ASH1 and TRX are H3K36- and H3K4-specific KMTs, respectively). Biochemical analysis revealed that dRAF but not PRC1 is the major complex involved in the ubiquitination of H2AK119 and this activity is directly linked to the removal of the H3K36me2 PTM. This is a striking example of how changing the components of the complex can dramatically alter the

enzymatic activity of Sce. In fact, the discovery of the Polycomb repressive deubiquitinase complex "PR-DUB" revealed that the ubiquitination activity of dRAF/PRC1 needs to be counterbalanced in order to prevent repression of unintended targets (Scheuermann et al., 2010). PR-DUB consists of the deubiquitinating enzyme Calypso and the PcG protein ASX that co-occupy Polycomb target genes. The functional combination of dRAF and PR-DUB allows precise regulation of the dosage of Polycomb repression (Scheuermann et al., 2010). Genetic, genome- wide expression and chromatin analyses further show that SCE, PSC and PR-DUB, regulating H2A mono-ubiquitination levels, are only required for repression of a subset of PRC1 target genes. Repression of other targets depends on the function of the PSC paralog Su(z)2 and the Ph protein, possibly by mediating chromatin compaction (Gutiérrez et al., 2012).

#### 2.3.1.2.2 Mammals

The PRC1 members in mammals have undergone multiple duplications during evolution and there are six homologues of *Drosophila Psc* (*Pcgf1/2/3/4/5/6*), five homologues of *Pc* (*Cbx2/4/6/7/8*), three homologues of *Ph* (*Phc1/2/3*) and two homologues of *Sce* (*Ring1/Rnf2*). The only gene that has been shown to be embryonic lethal at early post-implantation is *Rnf2* (see Table III), while *Pcgf2*, *Pcgf4*, *Cbx2* and *Phc1* show perinatal lethality and/or homeotic transformations (Akasaka et al., 1996; Coré et al., 1997; Katoh-Fukui et al., 1998; van der Lugt et al., 1994; Takihara et al., 1997). *Phc2* and *Ring1* mutant mice are healthy and fertile with minor homeotic transformations in the anterior-posterior axis (Isono et al., 2005; Lorente et al., 2000). Furthermore, two studies analysing the role of Cbx2 point to a function of PRC1 in sex determination and meiotic regulation (Baumann and De La Fuente, 2011; Katoh-Fukui et al., 1998). The fact that only *Rnf2* is indispensable for embryonic development could be explained with the presence of redundant paralogs. This is evident from several studies of double knockout mice, which show dramatically enhanced phenotypes compared to the single mutations. In the first study, Akasaka and colleagues showed that double deficiency for *Pcgf2* and *Pcgf4* results in post-implantation lethality during somite formation and organogenesis as a result of misregulated *Hox* gene expression (Akasaka et al., 2001). In another study, Posfai and colleagues removed both *Ring1* and *Rnf2* in the female germline and observed a strong

maternal effect leading to a 2-cell embryonic arrest, i.e. the stage before zygotic genome activation (Posfai et al., 2012). The authors profiled the transcriptome of *Ring1*/*Rnf2*-deficient fully-grown oocytes and revealed massive gene misregulation. Furthermore, they presented evidence that Ring1/Rnf2 are responsible for global H2AK119 ubiquitination in mammalian oocytes. With a similar genetic approach, Lapthanasupkul *et al.* examined the function of *Ring1* and *Rnf2* in mesenchymal stem cells and revealed that upon double knock-out the proliferation of the stem cells and the differentiation process are severely affected (Lapthanasupkul et al., 2012). A number of developmental regulator genes that show a restricted expression pattern in normal tissue were broadly expressed in the mutant tissue. Interestingly, these authors reported that the removal of Ring1/Rnf2 resulted in a massive depletion of Kdm2b protein. The latter has been identified as a member of the human BCOR complex, which is the homologue of dRAF (Gearhart et al., 2006). It remains to be shown whether mammalian BCOR has specialized as the major H2AK119 ubiquitin ligase and whether the H3K36 demethylase activity is coupled to Polycomb-mediated repression. A hint into this direction presents the recently published large-scale genomics data from the human ENCODE project, which shows that H3K27me3 and H3K36me3 are largely mutually exclusive (Consortium, 2012; Voigt et al., 2012). The study that identified the fly PR-DUB complex also showed that the human homologues of Calypso and ASX (BAP1 and Asxl1 respectively) form a complex and have deubiquitinase activity *in vitro*, suggesting a conserved molecular function for this Polycomb-related complex (Scheuermann et al., 2010). Indeed, homozygous *Asxl1*-deficient mice have homeotic transformations and die shortly after birth (Fisher et al., 2010). Finally, Gao and colleagues addressed the variety of mammalian PRC1-like complexes in a systematic way and identified four major PRC1 subtypes, which differ by the presence of different Pcgf and Cbx homologues and target sequences (Gao et al., 2012). Tavares and colleagues identified another PRC1 complex containing RYBP, Rnf2 and Mel18/Pcgf2 that is highly catalytically active and is targeted to chromatin in an unknown manner, independently of *Eed* function and H3K27me3 (Tavares et al., 2012). *Rybp* is essential for gastrulation (Pirity et al., 2005) and *in vitro* differentiation of ESCs. Rybp is not required for self-renewal of ESCs and mediates repression of certain endogenous retroviruses and pre-implantation and germ line genes (Hisada et al., 2012). The extensive variety in complexes potentially underlies different biological roles in specific

developmental contexts. To address this issue, future studies will have to target individual Pcgf, Cbx and Rybp/Yaf2 components rather than the common Ring1 and Rnf2.

#### 2.3.1.2.3 Plants

The composition of PRC1 complexes in *Arabidopsis* has not been studied in a systematic way and to the depth as in mammalian systems or in *Drosophila*, mostly because the PRC1 homologues have been identified only recently in *Arabidopsis* (Sanchez-Pulido et al., 2008). A plant PRC1-like complex has been characterized by Xu and Shen in 2008 (Xu and Shen, 2008), showing interaction between AtRING1a/AtRING1b and the H3K27me3-binding chromodomain protein LHP1 (Turck et al., 2007). While single AtRING1a or AtRING1b homozygous mutants did not show any abnormal phenotype, double mutant plants displayed homeotic transformations and meristem stem cell phenotypes similar to the ones observed in lhp1-/- and clf-/- plants (Gaudin et al., 2001; Goodrich et al., 1997; Larsson et al., 1998).The loss of AtRING1a/b caused derepression of KNOX genes promoting meristem proliferation, whereas H3K27me3 levels at the promoters of affected genes did not change, which shows that PRC2 function was still intact but transcriptional repression requires AtRING1a/b (Xu and Shen, 2008). The latter observation was confirmed in a study by Bratzel and colleagues who examined the role of AtBMI1a and AtBMI1b, two of the three identified Pcgf homologues in *Arabidopsis* (Bratzel et al., 2010; Sanchez-Pulido et al., 2008). Plants deficient for both AtBMI1a and AtBMI1b show a variety of phenotypes, ranging from complete developmental arrest of early seedlings to more mild effects related to cell differentiation and formation of callus-like structures. This phenotypes correspond to the observed upregulation of stem cell regulators such as WUS; STM, FUS3, LEC1 and WOX5. Furthermore, Bratzel and co-workers identified for the first time that Arabidopsis PRC1 proteins AtBMI1a/B and AtRING1a/b are involved in mono-ubiquitination of H2A.1 (the *Arabidopsis* H2A homologue that retained the lysine substrate at position 121) and are associated in a complex with LHP1 and a non-conserved plant-specific protein EMF1. Interestingly, EMF1 has been implicated with a dual role as being part of both PRC1 and PRC2 complexes in *Arabidopsis* and repressing two independent sets of genes (Kim et al., 2012). Plants

depleted of *emf1* show homeotic and flowering phenotypes (Aubert et al., 2001), as well as de-repression of PRC2 targets like AG, FLC and SEP1-3 but not the imprinted PHE genes.

## 2.3.2 Polycomb regulation

As evident from the many examples given above, PcG proteins are required throughout the life cycle of eukaryotes: during embryogenesis, for somatic cell differentiation, in germ line development as well as in disease (Albert and Helin, 2010; Richly et al., 2011). A common feature of all these biological processes is the transition from one developmental state to another, which is accompanied by major changes of gene expression. Original genetic studies in various species demonstrated major roles of Polycomb group proteins in the maintenance of stable repression during differentiation (e.g. Hox genes in flies and mammals and MADS box genes in plants). More recently, however, multiple chromatin immunoprecipitation (ChIP) experiments in different organisms have revealed thousands of putative Polycomb targets that are controlled in a cell-type specific manner (Bouyer et al., 2011; Consortium, 2012; Endoh et al., 2012; Gerstein et al., 2010; Gutiérrez et al., 2012; Kwong et al., 2008; Mohn et al., 2008; Roy et al., 2010). This data allows us to address the divergent developmental roles of Polycomb by classifying the Polycomb targets and considering the specificity of distinct PRC complexes. Finally, we will discuss the various mechanisms for targeting Polycomb to the chromatin.

### 2.3.2.1 Dynamics of core and specific Polycomb target sets

On the basis of chromatin profiling experiments, Polycomb targets in mammals can be classified into several classes: (I) core Polycomb target loci, co-occupied by PRC1 and PRC2 proteins and labelled by H3K27me3 and H2AK119ub1; (II) PRC2-only targets, marked by H3K27 methylation and sometimes co-occupied by core PRC2 members; (III) PRC1 only targets, bound by one or more PRC1 members and harbouring H2AK119ub1. There is a fourth class that includes targets bound by PRC1 or PRC2 members acting in a Polycomb-independent manner (for instance as part of other complexes), which we will not consider here. Although

classification may in part relate to differences in immuno-precipitation efficiencies and definitions of enrichment threshold values, genes in classes I and II do encode for different biological functions and respond differently in ESCs upon induction of differentiation (Ku et al., 2008). H2AK119ub1 mediated by Ring1 and Rnf2 is required for repression of class I genes as well as for the maintenance of ESC identity. PRC1 activity is, however, not required for binding of PRC1 proteins to target genes, neither for compaction of e.g. the HoxB cluster in ESCs (Endoh et al., 2012). Generally, these recent results are in accordance with the first genome-wide study that compared PRC1/PRC2 co-occupancy in ESCs and identified common 512 targets (Boyer et al., 2006), from which 291 overlap with the 510 class I targets found by Endoh and co-workers (Endoh et al., 2012). In summary, PRC1 and particularly the E3-ligase activity of its Ring finger proteins are required in ESCs and during mouse development for silencing of a rather small but important set of evolutionary conserved genes encoding for developmental regulators.

Interestingly, from the thousands of Polycomb targets described, only several hundred genes belong to class I marked by PRC1 and PRC2 loci. A similar observation was made in *Drosophila*. Several independent genome-wide experiments identified around 200-400 Polycomb targets co-occupied by several PcG members (Gutiérrez et al., 2012; Schuettengruber et al., 2009; Schwartz et al., 2006), which roughly corresponds to 20% of all H3K27me3 target sites. In a systematic genome-wide ChIP study, analysing the chromatin localization of 53 proteins in Kc167 embryonic cells, Filion and colleagues identified five distinct types of chromatin (Filion et al., 2010). Class I target genes were represented by one chromatin type. H3K27me3 only states were found in one of the two actively transcribed chromatin types, suggesting that class II genes might be partially associated with active chromatin. Furthermore, upon removal of Polycomb, the number of upregulated genes (derepressed targets) is significantly lower than the number of targets bound and most of the misexpressed genes are "classical" Polycomb targets, or class I targets (e.g. Hox genes, Wnt-, Fgf-, Tgf-signalling genes and other developmental regulators) (Bracken et al., 2006; Ezhkova et al., 2011; Posfai et al., 2012). Therefore, the role of class II and III target loci remains to be determined. Are they serving as a buffer of regulation, or perhaps as a structural component of the chromatin landscape?

#### 2.3.2.2 Mechanisms of Polycomb recruitment

Over the years, a lot of effort has been put into revealing the targeting mechanism of PRCs and propagation of the modified state. The classical model of epigenetic inheritance of the H3K9 methylated state by HP1 proteins recognizing the methylated histone as well as interacting with H3K9 HMTs fuelled the field (Jenuwein and Allis, 2001). Seminal work by Margueron et al. demonstrated that binding of the PRC2 component ESC/EED via its WD40 propeller to H3K27me3 stimulates E(Z)/EZH2 to catalyse tri-methylation of the unmodified substrate (Margueron et al., 2009). Together with work by Hansen et al. (Hansen et al., 2008), these data provided, in principle, a mechanism for epigenetic inheritance. More recent work indicates that re-establishment of H3K27me3 levels (and also for H3K9me3) does not occur during replication but gradually during subsequent cell cycle stages (Xu et al., 2012). Importantly, instead of newly incorporated histones, parental histones with intermediate methylation states are preferentially used as substrate. Together, these data suggest that number of modified nucleosomes within a region likely affects the efficiency of propagation (Brykczynska et al., 2010; Xu et al., 2012). The catalytic activity of the PRC2 complex is inhibited by H3K4 and H3K36 tri-methylation, when residing on the same histone tail in the nucleosome (Schmitges et al., 2011; Voigt et al., 2012; Yuan et al., 2011), potentially providing means for inhibition of spreading of the PRC2 repressed state by Trithorax group proteins. Furthermore, independent of the stimulatory effect of pre-existing H3K27me3, PRC2 activity is stimulated by high nucleosomal density that is sensed by the VEFS-box domain of the Su(z)12 protein interacting with amino acids 35-42 of H3 protruding from the nucleosome core. For PRC1, Psc interacts with nucleosomes and self-interacts in cell-fee replication systems thereby forming oligomeric structures. Since some Psc-chromatin contacts are dynamic while others are stable, Psc may enable inheritance of PRC1 on chromatin during replication (Lo et al., 2012).

*De novo* targeting and propagation may also be in part mediated by interactions of PRC proteins with the underlying DNA. Indeed, fly PcG proteins interact with specific DNA-binding factors such as PHO that associate with complex DNA elements termed Polycomb response elements (PREs) (Ringrose and Paro, 2007). Regions around such sites are marked by H3K27me3 and are co-occupied by Pc, likely due to its chromodomain that has a high binding affinity to H3K27me3 (Fischle et al., 2003; Schuettengruber et al., 2009). In *Drosophila* embryos, E(z) and

Pc proteins, but not H3K27me3, have been reported to be associated with PREs on recently replicated DNA, suggesting that these proteins may be directly involved in epigenetic heritability. In mammals, the role of transcription factors is less understood and only two PRE-like sequences have been identified up to date (Sing et al., 2009; Woo et al., 2010) (Sing et al., 2009; Woo et al., 2010). The role of PHO in targeting is probably not conserved in mammals since YY1, the mammalian orthologue, is not localized at PRC target genes in ESCs (Ku et al., 2008; Mendenhall et al., 2010). PRC proteins generally localize at CpG-rich sequences suggesting a possible function of transcription factors binding within such elements (Ku et al., 2008; Zheng et al., 2009). Arnold and colleagues revealed a role for Rest in H3K27me3 establishment at specific target sequences in neuronal progenitor cells during differentiation of ESCs (Arnold et al., 2012). Sequences containing Rest and Snail transcription factors are sufficient for the recruitment of H3K27me3 at targeted transgenic insertion sites suggesting that transcription factors can target PRC2 for gene repression which is consistent with reported biochemical interactions between REST and PcG proteins (Dietrich et al., 2012; Ren and Kerppola, 2011).

Finally, both PRC2 and PRC1 complexes are independently required for contraction of the *Kcnq1* imprint cluster and imprinted gene silencing during early mouse development (Terranova et al., 2008), as well for the formation of facultative heterochromatin at one of the two X-chromosomes in female mammalian cells (Plath et al., 2003). In these processes, non-coding RNAs (ncRNAs) may target PRC complexes to chromatin. Indeed, over the last few years an ever-increasing amount of data has been accumulating on the link between ncRNA and Polycomb repression. A pre-requisite for ncRNA-mediated targeting is RNA-binding affinity by PRC members and this has been (so far) reported for the PRC2 members Ezh2 and Suz12 (Guil et al., 2012; Kanhere et al., 2010; Ng et al., 2012), as well as for PRC1 member Cbx7 (Yap et al., 2010). Until recently, only a few of the known ncRNAs have been functionally analysed and prominent examples linked to PcG targeting are the HOTAIR ncRNA responsible for the silencing of the HoxD cluster in mammals (Rinn et al., 2007; Tsai et al., 2010), COLDAIR/COOLAIR required for the cold-induced silencing of the flowering repressor FLC in *Arabidopsis* (reviewed in (Song et al., 2012)), as well as Polycomb/Trithorax-related ncRNAs in flies (Hekimoglu and Ringrose, 2009). A recent study in mouse ESC identified the "Polycomb transcriptome"

consisting of almost 10,000 PRC2-bound RNAs (Zhao et al., 2010). Another report focused on a subclass of ncRNAs, and found 24 of the 226 lincRNAs (large intergenic ncRNAs) in ESC to be bound by PcG proteins (Guttman et al., 2011). Ng et al. described lincRNA-dependent PcG recruitment in human ESC (Ng et al., 2012). The molecular mechanisms underlying target selection *in cis* and *in trans* in relation to timing of ncRNA expression remain little understood. Interestingly, HOTAIR is able to bind to PRC2 and a REST complex, also containing LSD1 and CoREST (Tsai et al., 2010), suggesting scaffold functions for ncRNAs bridging DNA binding factors and PRC2. In all, future work is required to determine the relative contributions of transcription factors, ncRNAs and recognition of existing chromatin states in the de novo formation versus maintenance of Polycomb gene repression.

## 2.4 Conclusion

One role of chromatin silencing mechanisms is in the maintenance of repressed states through cell division. But this does not mean that this is some rigid regulatory mechanism, on the contrary, there is a need of plasticity to support the development of multicellular organisms. One way of achieving plasticity is through the interconnection of pathways. The two major chromatin-based silencing mechanisms that we reviewed above share many common ways of targeting based on protein interaction, ncRNA or recognition of PTM on histones (Figure 1). However, they are often considered as distinct and mutually exclusive in regard to their distribution in the nucleus and genome-wide (de Wit et al., 2007). As mentioned in the beginning, studies in *T. thermophila* would argue that these pathways have common ancestral functions (Liu et al., 2007). Could these two silencing mechanisms be independent of each other? HP1 is found in an atypical PRC1 complex (Gao et al., 2012; Qin et al., 2012) and SETDB1 represses some sets of developmental genes also targeted by the PcG pathway (Bilodeau et al, 2009). The PRC1 subunits Cbx2 and Cbx7 have affinity for H3K27me3 and H3K9me3 *in vitro* (Bernstein et al., 2006b; Kaustov et al., 2011) and were suggested to associate with Suv39h KMT (Li et al., 2010b; Sewalt et al., 2002). Suv39h KMT could then influence PRC1 targeting as suggested by Yang and colleagues (Yang et al., 2011), who showed that Cbx4/hPC2 localization in PcG is dependent on Suv39h1 in quiescent cells. Interestingly, in the zygote, PRC1 is

prevented from binding to maternal PCH in a Suv39h2 dependent manner (Puschendorf et al., 2008). Many of these enzymes and their relative marks are present in germ cells of both sex and could therefore influence the outcome of embryogenesis (Brykczynska et al., 2010; Hammoud et al., 2009; Posfai et al., 2012). Though the concept of transgenerational inheritance is still under debate (Gill et al., 2012), the interplay between these two major chromatin silencing pathway would be an interesting way to accommodate chromatin plasticity during developmental transition within germ cell maturation and the developing embryo.

**Box1. Interplay between H3K9 and DNA methylation.**

Generally, DNA methylation (Jones, 2012) and H3K9 methylation serve similar purposes in long-term silencing. However a direct (inter-)dependence between these two silencing mechanisms is not always clear. DNA demethylation for the generation of induced pluripotent stem (iPS) cell is critical, as it can be enhanced with inhibitors like azacytidine (Yamanaka and Blau, 2010). Interestingly, a screen for the identification of barriers against *in vitro* reprogramming identified *Suv39h1* (Onder et al., 2012), suggesting that losing DNA methylation and H3K9me3 is required for reprogramming. On the other hand, MEFs deficient for *Suv39h1/2* show a concomitant loss of H3K9me3 and DNAme at PCH (Lehnertz et al., 2003). This decrease could be a consequence of the loss of HP1α binding. Indeed, artificial targeting of HP1 to euchromatic genes allows recruitment of DNMT1 and transcriptional silencing (Smallwood et al., 2007). G9a appears to act synergistically with DNMTs to mediate *de novo* epigenetic silencing. Dnmt3a/b associate with G9a through its ankyrin domains to silence key pluripotency factors like Oct3/4 and suppresses proviruses (Epsztejn-Litman et al., 2008; Leung et al., 2011). G9a also interacts directly with DNMT1 to form a ternary complex with PCNA required for maintaining DNA and histone methylation on rDNA repeats throughout replication (Estève et al., 2006). However, DNA methylation and G9a-mediated H3K9 methylation seem to occur mostly as two parallel pathways (Tachibana et al., 2008), because most of the DNA methylation defects seen in *G9a*$^{-/-}$ ESCs can be rescued by a catalytically dead enzyme (Dong et al., 2008). DNMT1 associates with Np95 (also known as Uhrf1 or ICBP90) (Sharif et al., 2007), which has the ability to bind hemimethylated DNA through its SET and RING finger-associated (SRA) domain and H3K9me2/3 through its tandem tudor domain. Np95 can bind H3K9me2/3 regardless of the presence of H3S10P. Insensitive to the phospho-methyl switch (Fischle et al., 2005), Np95 provides a way of maintaining DNA methylation during mitosis (Rothbart et al., 2012). This bridging protein would then allow DNMT1 to prevent the loss of DNA methylation states during critical steps of the cell-cycle.

Such coordination between DNA replication and maintenance of histone and DNA methylation seems to apply for PCH through the association of the methyl-CpG binding protein MBD1 and SETDB1 to replication forks (Sarraf and Stancheva, 2004). Indeed, it has been suggested that SETDB1 mediated H3K9me1 would serve as a substrate for the Suv39h enzymes to restore H3K9me3 on PCH as replication is ongoing (Dambacher et al., 2010). SETDB1 associates with the co-repressor KAP1 (also known as Trim28/Tif1b) (Schultz et al., 2002). Biochemical studies showed that KAP1 is in a complex with remodelling enzymes (Mi2a), DNA methyltransferases (DNMT1, DNMT3a, DNMT3b), KMT (SETDB1), HP1 and KRAB-ZNF proteins (for review see (Iyengar and Farnham, 2011)). KAP1 repressive complex was shown to be implicated in the overall silencing of euchromatic genes, retrotransposons and imprinted control regions (ICRs) during development (Quenneville et al., 2011; Rowe et al., 2010; Schultz et al., 2002). Targeting to specific genomic regions is mediated by KRAB-ZNF proteins, like ZNF57, which can target this complex to ICRs when they are methylated, a mechanism suggested to protect these loci from demethylation during early embryo development (Quenneville et al., 2011, 2012; Zuo et al., 2012). KAP1 therefore acts as a scaffold linking DNA methylation, H3K9me3 and HP1 binding to specific loci which repressive state needs to be maintained for proper development. With the hundreds of KRAB-ZNF proteins encoded by the human genome (Huntley et al., 2006), it will be challenging to understand the contribution of these proteins for targeting KAP1-repressive complex (and then DNA and H3K9 methylation) to specific genomic sites during early embryo development.

HP1γ, which was recently mapped in HCT116 cells show a localization towards gene bodies and is thought to be involved in RNA processing (Smallwood et al., 2012). Strikingly, DNA methylation in gene bodies is often associated with transcriptional activity and it has even been suggested to participate in RNA splicing (Laurent et al., 2010), however it is not known if there is

a direct interplay between these two pathways for achieving this function. Recently, new evidence was brought for a direct link between H3K9 and DNA methylation, as PGC7 (also known as Dppa3/Stella) has been shown to bind H3K9me2. PGC7 is involved in protecting from active demethylation the maternal genome, marked by H3K9me2, and several paternally imprinted genes in the zygote (Nakamura et al., 2012). In plants and mammals, DNA methylation can occur on CG, CHG and CHH context (where H is either a C, T or A). In *A. thaliana*, methylation on CHG is maintained by a positive feedback loop between the Su(var)3-9 homologue SUVH4/KYP and DNA methyltransferase CMT3 in order to silence genes and retrotransposons (Jackson et al., 2002, 2004). Indeed, H3K9me1/2 catalysed by SUVH4/KYP (the major enzyme responsible for H3K9me1/2 in plant heterochromatin) allows recruitment of HP1, which then allows binding of CMT3 and DNA methylation. Recently, it was also shown that SUVH4/KYP is required for the chromatin remodeler DDM1 (decrease in DNA methylation 1) dependent *de novo* methylation (Sasaki et al., 2012). Furthermore, plant SUVH homologs contain a YDG/SRA domain in their N-terminus that can bind methylated DNA (Johnson et al., 2007), suggesting that methylated DNA can reinforce the silencing signal by enhancing the recruitment of SUVH enzymes (Ebbs and Bender, 2006).

However, the most striking evidence of a direct link between DNA and H3K9 methylation comes from studies in the fungi *N. crassa*. In this organism, mutation of *DIM-5*, a *Su(var)3-9* homologue, results in a global loss of both H3K9me3 and DNA methylation (Tamaru and Selker, 2001), similar to what is seen in the DNA methyltransferase *DIM-2* mutants (Kouzminova and Selker, 2001), although the latter does not affect H3K9 methylation. The link between these two marks is HP1, which binds to H3K9me3 through its CD and to DIM-2 through its CSD (Honda and Selker, 2008). Mutation of the *N. crassa* HP1 gene, *hpo,* results in severe defects of DNA methylation without altering H3K9me3, like *DIM-2* mutants (Freitag et al., 2004). Strikingly, this suggests that DIM-5 and HP1 are upstream of DNA methylation in *N. crassa*. Moreover, H3K9me3, HP1 and DNA methylation colocalize almost perfectly on 44 defined heterochromatic domains on linkage group VII (Lewis et al., 2009). Interestingly, HP1 also prevents the spreading of heterochromatic domains by association with the jmjC domain containing KDM DMM-1/2 (DNA methylation modulator 1/2). DMM-1/2 remove H3K9me3 and then prevent further accumulation of HP1 and DNA methylation (Honda et al., 2010). Similarly, such anti-silencing mechanisms also exist in *S. pombe* and *A. thaliana* (Miura et al., 2009; Saze et al., 2008; Zofall and Grewal, 2006). In the yeast *S. pombe*, there is clear evidence of the role played by the RNAi pathway in recruitment of the H3K9me3 KMT Clr4 and heterochromatin formation (Zhang et al., 2008) (although it lacks DNA methylation) and to a lesser extend in plants and *Drosophila* from the RNA-directed DNA methylation (RdDM) (Law and Jacobsen, 2010) and the piRNAs pathway respectively (Pal-Bhadra et al., 2004). In *N. crassa*, targeting of DIM-5 to chromatin relies on its interaction with another factor, DIM-7 (Lewis et al., 2010). Therefore, despite the differences between the model organisms cited so far, it seems that methylation of DNA and H3K9 cooperate for mediating chromatin silencing.

**Box2. Removal of H3K9 methylation by KDM during development.**

The balance between establishment and erasure of H3K9 methylation is also crucial for proper embryo development. The expression of the H3K9me1/2 JmjC-domain-containing histone demethylase 2a (Jhdm2a) enzyme partially overlaps with that of G9a. Whereas *G9a* is continuously expressed in the germ line until its down-regulation during meiotic prophase, *Jhdm2a* is found transcribed from late pachytene onwards, with high expression levels in round and elongating spermatids (Okada et al., 2010). Although *Jhdm2a* knock-out mice are viable, they display smaller testis, infertility and obesity (Okada et al., 2007, 2010). Disruption of this enzyme causes defects in post-meiotic chromatin condensation in elongating spermatids leading to impaired nuclear elongation. Interestingly, Jhdm2a was shown to bind to the promoter region of genes encoding Transition Protein 1 (*Tnp1*) and Protamine 1 (*Prm1*). Tnp1 and Prm1 are required for correct nuclear condensation during spermiogenesis (Cho et al., 2001; Zhao et al., 2004). Furthermore, deficiency for *Jhdm2a* impairs transcriptional activation of *Tnp1* and *Prm1* leading to infertility (Okada et al., 2010). Interestingly, the four members of the Jmjd2 family of KDMs in mammals show dual specificity for the removal of H3K9me3/2 and H3K36me3/2 *in vitro* (Fodor et al., 2006; Klose et al., 2006; Whetstine et al., 2006). RNAi depletion of the *C. elegans* homologue jmjd-2 induces increased levels of H3K9me3 (and H3K36me3 on one end of the X chromosome) in the germline. Germ cells depleted for jmjd-2 showed increased apoptosis and altered DSB repair although they don't harbour defects in pairing and synapsis (Whetstine et al., 2006).This dynamic interplay between the deposition and the erasure of H3K9 methylation is reinforced by the fact that the phenotype of *jmjd-2-/-* animals can be partially rescued by the deletion of HPL-2, the *C. elegans* homologue of HP1 (Black et al., 2010).

## *Table 2.1 Components of the HP1/H3K9 pathway*

| *M. musculus* | | | *D. melanogaster* | | *A. thaliana* | |
|---|---|---|---|---|---|---|
| Suv39h1/ KMT1A<br>Suv39h2/ KMT1B | Embryo: DKO viable, reduced size, impaired gametogenesis, increased incidence of lymphoma *(Peters 2001, Braig 2005)*. | ESC: Normal proliferation, loss of H3K9me3 and HP1 binding on chromocenter, replaced by H3K27me3 *(Peters 2003, Lehnertz 2003)*. | Su(var)3-9 | Flies are viable, phenotype similar to mammals, HP1 binding is lost on chromocenters *(Tschiersch 1994, Schotta 2002)*. | SUVH2<br>SUVH4/ KYP | Associate with chromocenters. Mutants show loss of heterochromatic marks, loss of gene silencing. Heterochromatinisation if overexpressed *(Jackson 2002, Naumann 2005)*. |
| G9a/ Ehmt2/ KMT1C<br>Glp/ Ehmt1/ KMT1D | Embryo: Both G9a and Glp single KO are lethal around E9.5 due to severe growth defects. Conditional germ line KO impairs development of germ cells and leads to defects in meiosis *(Tachibana 2002, 2005, 2007)*. | ESC: Regulation of euchromatic genes, imprinted genes, rDNA repeats and transposons. Reduction of DNA methylation in G9a KO cells can be rescued by catalytically dead G9a *(Tachibana 2002, Estève 2006, Dong 2008, Wagschal 2008)*. | dG9a | Suppressor of PEV, also regulates euchromatic genes. Mutants show developmental defects but are viable, more severe effects when combined with Suv(3-9) mutation *(Mis 2006, Stabell 2006, Seum 2007b, Kato 2008)*. | SUVR4? | Contributes to the silencing of transposons and rDNA repeats. *(Thorstensen 2006, Veiseth 2011)*. |
| Setdb1/ ESET/ KMT1E<br>Setdb1/ KMT1F | Embryo: Setdb1 KO is lethal around E3.5-E5.5 and shows ICM growth defects. No KO mice reported for Setdb2 *(Dodge 2004)*. | ESC: Setdb1 KO leads to transcriptional repression of euchromatic genes and repeats; required for maintenance of pluripotency *(Dodge 2004, Bilodeau 2009, Lohmann 2010, Matsui 2010)*. | dSetdb | Homozygous lethal. Localized at euchromatin and on chromosome 4 *(Seum 2007a)*. | | |
| HP1α/ Cbx5<br>HP1β/ Cbx1<br>HP1γ/ Cbx3 | Embryo: HP1β KO is perinatal lethal and leads to genomic instability. HP1γ KO mice are infertile and show impaired spermatogenesis, cell cycle and meiosis defects. No HP1α KO mice have been reported. *(Aucott 2008, Naruse 2007, Brown 2010, Abe 2011, Takada 2011)*. | ESC: Modest upregulation of ERV *(Maksakova 2011)*. | Su(var)2-5/ HP1a | Flies have at least five HP1 homologues (HP1a/b/c/d/e), see text for detail. HP1a KO flies die at 3rd instar larva *(Eissenberg 1992, Kellum 1995)*. | LHP1/ TLF2 | Function with PRC1, see Table 3 and text for details *(Gaudin 2001, Turck 2007)*. |

## Table 2.2 Components of PRC2

| *M. musculus* | | | *D. melanogaster* | | *A. thaliana* | |
|---|---|---|---|---|---|---|
| Ezh1<br>Ezh2 | Embryo: *Ezh2* KO is lethal around E5.5-8.5 and displays abnormal gastrulation *(O'Carroll 2001)*. *Ezh1* KO does not have an embryonic phenotype. No embryonic/germ line DKO has been reported yet *(Ezhkova 2011)*. | ESC: Single or double KO do not have a significant effect on the maintenance of ESC, although the H3K27 methylation mark is lost. An effect of the KO is seen upon differentiation of ESC due to failure in extinguishing expression of pluripotency genes. *(Shen 2008)*. | E(z) | Loss of function of *e(z)* is 100% lethal around end of 3rd instar larva/puparium. Hypomorph mutations result in homeotic transformations and failure of germ cell development. *e(z)* is a maternal effect gene*(Phillips 1990)*. | CLF<br>MEA<br>SWN | Plants lacking CLF display pleiotropic effects on leafs and flowers and flower prematurely *(Goodrich 1997)*.<br>*MEA* is a maternal effect gene and loss of function leads to over-proliferation of the endosperm, resulting in embryo lethality later in seed development (*Grossniklaus 1998*). Mutations in the SWN gene do not show significant phenotype but enhance the clf phenotype *(Chanvivattana 2004)*. |
| Suz12 | Embryo: KO of *Suz12* is lethal around E7.5-E8.5 and gastrulation is not initiated *(Pasini 2004)*. | ESC: In *Suz12* deficient cells proliferation and self-renewal are not affected but some lineage-specific genes are upregulated. Differentiation of ESC is disturbed in KO cells *(Pasini 2007)*. | Su(z)12 | Loss of function is lethal around end of 1st/beginning of 2nd instar larva and results in misexpression of Hox genes and strong homeotic phenotypes. *Su(z)12* is required in the germ line and for suppression of PEV through heterochromatin-silencing *(Birve 2001)* | EMF2<br>FIS2<br>VRN2 | *FIS2* is a maternal effect gene, which phenocopies *MEA (Luo 1999)*. EMF2 mutation leads to premature flowering (*Yoshida 2001*). VRN2 is required for the maintenance of FLC repression in the process of vernalization (*Gendall 2001*). |
| Eed | Embryo: KO of *Suz12* is lethal around E8.5 and displays abnormal gastrulation *(Faust 1998)*. | ESC: In *Eed* deficient cells proliferation and self-renewal are mildly affected. Differentiation in the absence of Eed is biased towards endoderm *(Leeb 2010, Shen 2008, 2009)*. | Esc<br>Escl | *Esc* is needed for segment establishment in the early embryo, but not for maintenance at later stages *(Struhl 1982)*. *Escl* mutation does not show a phenotype but enhances the phenotype of *esc* and leads to homeotic transformations and lethality at 3rd instar larva (*Kurzhals* 2008, Ohno 2008). | FIE | A maternal effect gene phenocopying *MEA* and *EMF2* (*Ohad 1999*). *FIE* is required for embryo-to-seedling transition if maternal effect is rescued (*Bouyer 2011*). |

## Table 2.3 Components of PRC1

| *M. musculus* | | | *D. melanogaster* | | *A. thaliana* | |
|---|---|---|---|---|---|---|
| Ring1a/ Ring1<br>Ring1b/ Rnf2 | Embryo: Ring1b KO is lethal around E6.5-9.5, gastrulation defects. Ring1a KO mice are viable and fertile with abnormalities in the axial skeleton. Conditional DKO in oogenesis has a strong maternal effect and leads to 2-cell arrest *(Posfai 2012, Voncken 2003, del Mar Lorente 2000)*. | ESC: Single Ring1a or Ring1b KO ESC do not show significant phenotype, while DKO leads to derepression of lineage genes and loss of ESC identity. H2AK119ub is globally absent in DKO. *(Leeb 2007, 2010, von der Stoop 2008, Endoh 2008)*. | Sce/ dRING | Loss of function is lethal around 1st instar larva. *Hox* genes are derepressed in *Sce* mutants *(Breen 1986)*. | AtRING1a<br>AtRING1b | Single KO plants do not exhibit significant phenotypes, while DKO leads to derepression of KNOX genes, homeotic transformations and ectopic meristemic formation. AtRING1a/b interact with the plant HP1 homologue LHP1 *(Xu 2008)*. |
| Pcgf1<br>Pcgf2/ Mel18<br>Pcgf3<br>Pcgf4/ Bmi1<br>Pcgf5<br>Pcgf6 | Embryo: Only Pcgf2 and Pcgf4 have been functionally studied and single KO of either Pcgf2 or Pcgf4 results in posterior transformations of axial skeleton. DKO is lethal around E9.5 due to failure of somite formation and organogenesis *(Akasaka 2001)*. | | Psc | *Psc* KO is embryonic lethal, affecting the embryonic body pattern and most severely the head formation. Derepression of *Hox* genes in *Psc* mutants *(Jürgens 1985, Beh 2012)*. | AtBMI1a<br>AtBMI1b<br>AtBMI1c | Single KO mutants do not exhibit significant phenotypes. DKO for *AtBMI1a* and *AtBMI1b* leads to aberrant cell differentiation and early embryonic arrest, as well as upregulation of the stem-cell regulators WUS, STM, FUS3, LEC1 and WOX5 *(Bratzel 2010)*. |
| Cbx2/ M33<br>Cbx4/ Pc2<br>Cbx6<br>Cbx7<br>Cbx8/ Pc3 | Embryo: From the five homologues, only Cbx2 has been functionally studied and loss of function results in perinatal lethality, as well as male-to-female sex reversal, meiotic defects and chromosome instability *(Katoh-Fukui 1998, Baumann 2011)*. | | Pc | *Pc* loss of function is embryonic lethal and giving a strong homeotic phenotype, which is enhanced by overexpression of BX-C. Derepression of *Hox* genes in *Pc* mutants *(Lewis 1978)*. | LHP1/ TLF2 | A homologue of fly and mouse HP1, but localizing to euchromatin and interacting with the other PRC1 members in plants. Mutant *lhp1* plants show pleiotropic phenotypes with overall size reduction, changes in leaf morphology, early flowering. LHP1 is also required for the repression of FLC in the process of vernalization *(Gaudin 2001, Sung 2006)*. |

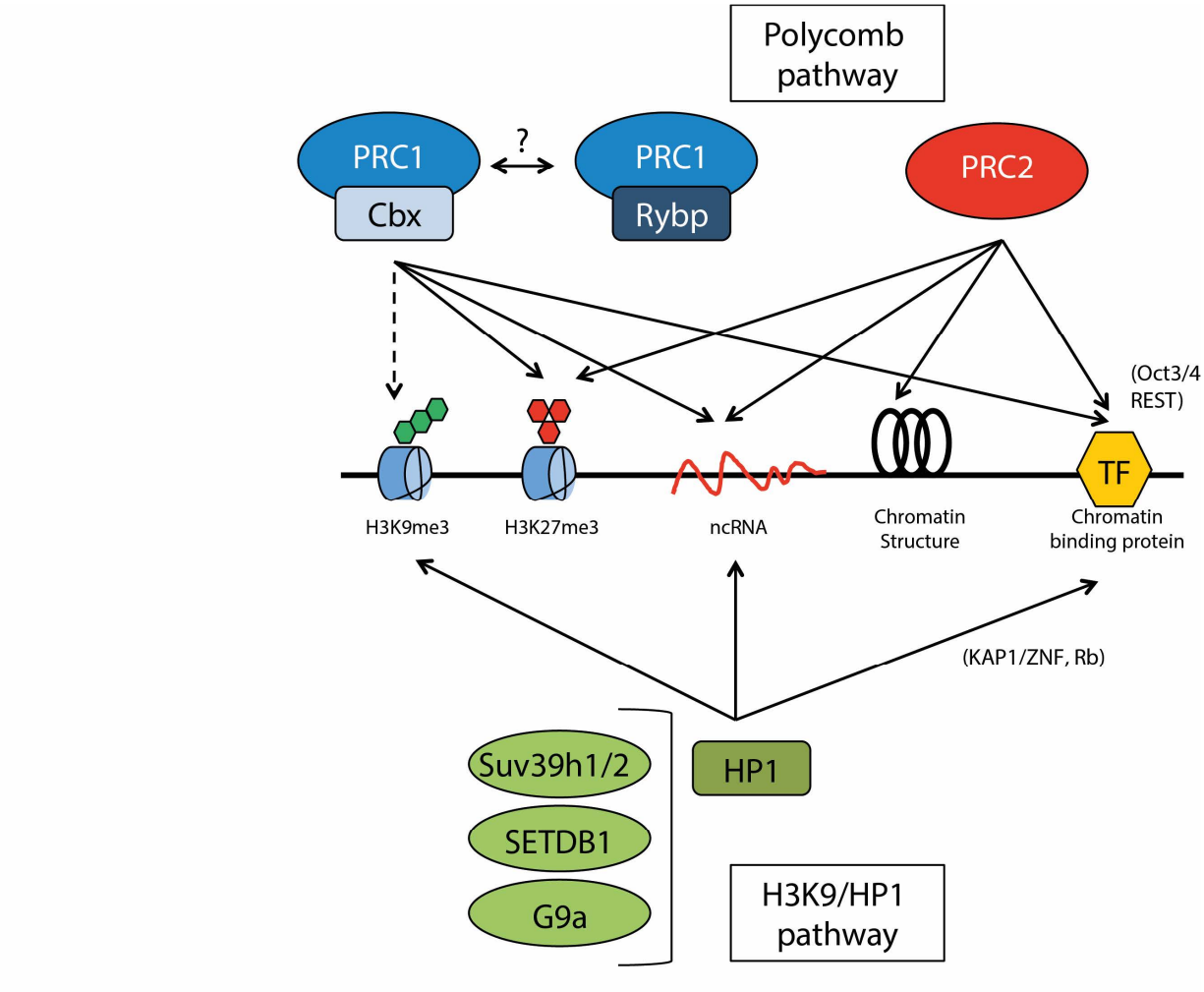



***Figure 2.1 Overview of the targeting of Polycomb complexes and H3K9 KMT/HP1 proteins***

Polycomb and H3K9 KMT/HP1 complexes can be targeted to chromatin through many diverse ways implicating direct interaction with modified histones, noncoding RNA (ncRNA), chromatin structure, and DNA-binding protein like transcription factors or zinc-finger-containing protein (ZNF). Solid lines indicate interactions. Dotted lines indicate potential interactions. Some examples are indicated between brackets. See text for details.

**Acknowledgements**

M. Tardat greatly acknowledges his EMBO long-term fellowship and P. Nestorov his Boehringer Ingelheim Fonds PhD fellowship. Research in the Peters group is supported by the Novartis Research Foundation, the Swiss National Science Foundation (31003A_125386 and NRP 63 – Stem Cells and Regenerative Medicine), SystemsX.ch (Cell plasticity), the Japanese Swiss Science and Technology Cooperation Program, and the EMBO Young Investigator Program.

# Chapter 3. Published manuscript: PRC1 coordinates timing of sexual differentiation of female primordial germ cells

LETTER

doi:10.1038/nature11918

# PRC1 coordinates timing of sexual differentiation of female primordial germ cells

Shihori Yokobayashi[1]†, Ching Yeu Liang[1,2], Hubertus Kohler[1], Peter Nestorov[1,2], Zichuan Liu[1], Miguel Vidal[3], Maarten van Lohuizen[4], Tim C. Roloff[1] & Antoine H. F. M. Peters[1,2]

**In mammals, sex differentiation of primordial germ cells (PGCs) is determined by extrinsic cues from the environment[1]. In mouse female PGCs, expression of stimulated by retinoic acid gene 8 (*Stra8*) and meiosis are induced in response to retinoic acid provided from the mesonephroi[2–5]. Given the widespread role of retinoic acid signalling during development[6,7], the molecular mechanisms that enable PGCs to express *Stra8* and enter meiosis in a timely manner are unknown[2,8]. Here we identify gene-dosage-dependent roles in PGC development for *Ring1* and *Rnf2*, two central components of the Polycomb repressive complex 1 (PRC1)[9,10]. Both paralogues are essential for PGC development between days 10.5 and 11.5 of gestation. *Rnf2* is subsequently required in female PGCs to maintain high levels of *Oct4* (also known as *Pou5f1*) and *Nanog* expression[11], and to prevent premature induction of meiotic gene expression and entry into meiotic prophase. Chemical inhibition of retinoic acid signalling partially suppresses precocious *Oct4* downregulation and *Stra8* activation in *Rnf2*-deficient female PGCs. Chromatin immunoprecipitation analyses show that *Stra8* is a direct target of PRC1 and PRC2 in PGCs. These data demonstrate the importance of PRC1 gene dosage in PGC development and in coordinating the timing of sex differentiation of female PGCs by antagonizing extrinsic retinoic acid signalling.**

In mammalian somatic cells, PRC1 and PRC2 proteins are transcriptional repressors that function in large multiprotein complexes and that modify chromatin by mono-ubiquitinating histone H2A at lysine 119 (H2AK119u1) and trimethylating histone H3 at lysine 27 (H3K27me3), respectively[9,12]. At day 12.5 of embryonic development (E12.5), in PGCs marked by Cdh1 (E-cadherin) staining[13], we observed nuclear localization of PRC1 components Rnf2 (also known as Ring1B), Mel18 (also known as Pcgf2) and Rybp (Fig. 1a and Supplementary Fig. 1) as well as a robust H2AK119u1 signal, suggesting the presence of catalytically active PRC1 complexes (Fig. 1a). To

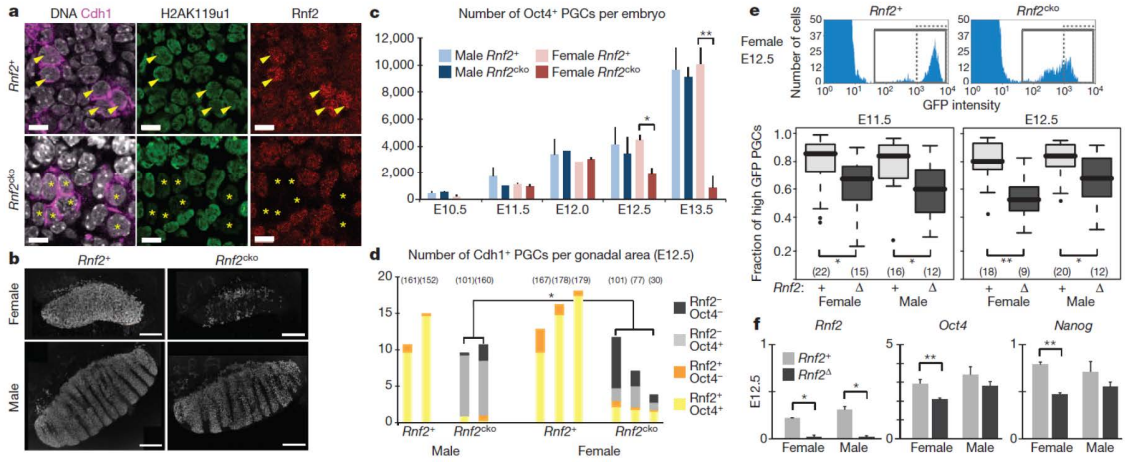



**Figure 1 | *Rnf2* regulates PGC development and *Oct4* and *Nanog* expression in *Rnf2*$^{cko}$ female gonads.** **a**, Immunofluorescence staining of H2AK119u1, Rnf2 and Cdh1 with DAPI (4′,6-diamidino-2-phenylindole) in *Rnf2*$^{+}$ and *Rnf2*$^{cko}$ gonadal sections from E12.5 female embryos. Arrowheads denote *Rnf2*$^{+}$ PGCs and asterisks denote *Rnf2*$^{\Delta}$ PGCs. Scale bars, 10 μm. **b**, Immunofluorescence staining of Oct4 in E13.5 *Rnf2*$^{+}$ and *Rnf2*$^{cko}$ whole gonads and mesonephroi. Scale bars, 300 μm. **c**, Average number of Oct4-positive cells in whole gonads at E10.5–E13.5. Error bars indicate +s.d. $n = 2$–12. $^{*}P < 0.005$; $^{**}P < 1.0 \times 10^{-5}$ (Student's *t*-test). **d**, Classification of Cdh1-positive PGCs according to Rnf2 and Oct4 protein levels in *Rnf2*$^{+}$ and *Rnf2*$^{cko}$ E12.5 gonads. *y* axis represents the number of PGCs that were normalized to areas analysed (10,000 μm$^2$). Numbers in brackets denote number of PGCs scored per embryo. $^{*}P < 1.0 \times 10^{-8}$ (chi-squared test). **e**, Representative histograms showing *Oct4*(Δ*PE*)–GFP signals in PGCs from female *Rnf2*$^{+}$ and *Rnf2*$^{cko}$ E12.5 gonads. Boxplots showing the ratios of PGCs with high GFP intensity (>10$^3$, enclosed by dashed line in histogram) over all GFP-positive cells (enclosed by solid line) in different embryos. Numbers in brackets denote number of embryos analysed. $^{*}P < 0.05$; $^{**}P < 0.005$ (Student's *t*-test). **f**, Representative qRT–PCR data of *Rnf2*, *Oct4* and *Nanog* in *Rnf2*$^{+}$ and *Rnf2*$^{\Delta}$ PGCs (normalized to *Tbp*). Error bars indicate +s.d. of 2–3 technical replicates. $^{*}P < 0.05$; $^{**}P < 0.01$ (Student's *t*-test).

[1]Friedrich Miescher Institute for Biomedical Research (FMI), Maulbeerstrasse 66, CH-4058 Basel, Switzerland. [2]Faculty of Sciences, University of Basel, CH-4056 Basel, Switzerland. [3]Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas (CSIC), 28040 Madrid, Spain. [4]Division of Molecular Genetics and Centre for Biomedical Genetics, the Netherlands Cancer Institute (NKI), 1066 CX Amsterdam, the Netherlands. †Present address: Department of Reprogramming Science, Center for iPS Cell Research and Application, Kyoto University, 53 Kawahara-cho, Shogoin Yoshida, Sakyo-ku, Kyoto 606-8507, Japan.

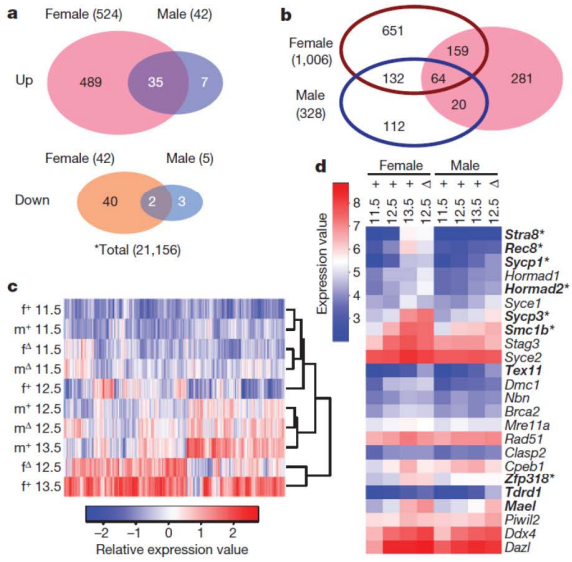



**Figure 2 | *Rnf2* deficiency induces extensive transcriptional misregulation in female PGCs.** **a**, Venn diagrams showing the numbers of genes misregulated in male and female $Rnf2^{\Delta}$ PGCs compared to $Rnf2^{+}$ PGCs at E12.5. **b**, Venn diagram showing the numbers of genes upregulated in female (red) and male (blue) control PGCs between E11.5 and E13.5 and in female $Rnf2^{\Delta}$ PGCs compared to $Rnf2^{+}$ PGCs at E12.5 (pink oval). **c**, Relative expression levels of 1138 probe sets upregulated in male (m) and female (f) $Rnf2^{+}$ PGCs between E11.5 and E13.5 (**b**) in various samples indicated. Unsupervised clustering analysis shows clustering of female E12.5 $Rnf2^{\Delta}$ PGCs with female E13.5 $Rnf2^{+}$ PGCs, whereas E12.5 male $Rnf2^{\Delta}$ and $Rnf2^{+}$ PGCs clustered together. **d**, Microarray expression values of early meiosis program genes. Genes significantly upregulated in $Rnf2^{\Delta}$ PGCs of both sexes are indicated in bold; those in female $Rnf2^{\Delta}$ PGCs only in bold with asterisk. $n = 3$ per condition; fold change >1.5; adjusted $P$ value <0.05.

address the function of PRC1 in PGC development (Supplementary Fig. 2), we conditionally deleted *Rnf2* in PGCs from E9.5 onwards by generating mice carrying a floxed and a mutant allele of *Rnf2* ($Rnf2^{fl/\Delta}$) and a Cre recombinase driven by the mouse *Tnap* promoter (*Tnap-cre*)[14,15]. To concomitantly assess possible functional redundancy with the *Rnf2* paralogue *Ring1* (refs 10, 16), we studied mice that were either heterozygous or homozygous deficient for *Ring1* (ref. 17) (referred to as *Rnf2* conditional knockout ($Rnf2^{cko}$) mice and $Ring1^{\Delta}$ $Rnf2^{cko}$ mice, respectively) (Supplementary Figs 3 and 6a). This strategy resulted in ~90% deletion efficiency at E11.5 (Supplementary Fig. 4). At E12.5, Rnf2, Mel18, Rybp and H2AK119u1 were lost in PGCs of $Rnf2^{cko}$ embryos but not of $Rnf2^{fl/\Delta}$ embryos, carrying one functional allele, indicating that complex stability and catalytic activity of PRC1 is regulated by Rnf2 in PGCs at E12.5 (Fig. 1a and Supplementary Figs 1 and 4). By contrast, Ezh2 and H3K27me3 levels were similar in $Rnf2^{cko}$ versus control PGCs, suggesting globally unaltered PRC2 function in $Rnf2^{cko}$ PGCs (Supplementary Fig. 5).

To study the fate of *Rnf2*-deficient PGCs, we analysed expression of the pluripotency and germ cell marker Oct4 in whole gonads of E10.5–E13.5 embryos (Fig. 1b and Supplementary Fig. 2). We observed a strong reduction of Oct4-positive PGCs, specifically in female $Rnf2^{cko}$ embryos but not in male $Rnf2^{cko}$ or control embryos, starting around E12.5 of gestation (Fig. 1b, c). By contrast, double deficiency of *Ring1* and *Rnf2* caused a strong reduction of $Oct4^{+}$ PGCs by E11.5 in both sexes (Supplementary Fig. 6b, c), indicating an essential role for PRC1 in PGCs after their migration into the embryonic gonad (Supplementary Fig. 2)[11].

To further dissect the role of *Rnf2* in regulating Oct4 expression versus PGC development, we assessed co-expression of Oct4 and Rnf2 at E12.5 in $Cdh1^{+}$ PGCs in $Rnf2^{cko}$ embryos (Fig. 1d). The number of PGCs lacking detectable Rnf2 protein was strongly reduced in female but not male gonads, despite the fact that gonads of both sexes harboured comparable numbers of *Rnf2*-deficient PGCs at E12.0 (Supplementary Fig. 4a, and data not shown). We further observed a pronounced downregulation of Oct4 protein in female and some male *Rnf2*-deficient PGCs (Fig. 1d). These data indicate that *Rnf2* contributes to maintaining Oct4 expression, particularly in female PGCs, beginning between E12.0 and E12.5.

We subsequently investigated the mechanism underlying the reduction in female *Rnf2*-deficient PGCs. We failed to observe increased levels of apoptosis or major changes in cell cycle progression (data not shown). To study changes in gene expression, we introduced a green fluorescent protein (GFP) transgene driven by the promoter of *Oct4* lacking the proximal enhancer (*Oct4*(ΔPE)–GFP))[18] into the $Rnf2^{cko}$ strain and isolated pure populations of PGCs by fluorescence-activated cell sorting (FACS) (Fig. 1e). By quantitative reverse transcriptase PCR (qRT–PCR), we barely detected any *Rnf2* transcripts in isolated $Rnf2^{cko}$ PGCs, confirming efficient deletion of the $Rnf2^{fl}$ allele by *Tnap-cre* (Fig. 1f; see also Supplementary Fig. 4). Hence, we subsequently refer to PGCs from $Rnf2^{cko}$ embryos as $Rnf2^{\Delta}$. We also noticed significant reductions in GFP intensities in $Rnf2^{\Delta}$ PGCs isolated from male and female E11.5 and E12.5 embryos compared to controls (Fig. 1e), suggesting decreased *Oct4*(ΔPE) promoter activity in these cells. qRT–PCR analysis showed significantly reduced *Oct4* and *Nanog* expression in female GFP-positive $Rnf2^{\Delta}$ PGCs at E11.5 and E12.5 (Fig. 1f and data not shown). Thus, *Rnf2* is required for maintaining the expression of pluripotency factors in PGCs.

We next analysed genome-wide expression in purified PGCs. The number of misregulated genes in female compared to male $Rnf2^{\Delta}$ PGCs at E12.5 was 12-fold higher (Fig. 2a and Supplementary Table 1). Consistent with the fact that PRC1 is a transcriptional repressor, ~90% of misexpressed genes were upregulated. At E11.5, we only observed a few misregulated genes in $Rnf2^{\Delta}$ PGCs (Supplementary Table 1), consistent with the timing of the mutant phenotype. According to Gene Ontology analysis, gene functions related to meiosis (synapsis, sister chromatid cohesion) were highly over-represented among genes upregulated in female $Rnf2^{\Delta}$ PGCs. By contrast, nucleosome functions were enriched in downregulated genes, as reported previously for *Ring1* and *Rnf2* double-deficient germinal vesicle oocytes[10] (Supplementary Fig. 7 and Supplementary Table 2). Notably, we barely found any over-representation of developmental gene functions in $Rnf2^{\Delta}$ PGCs, a classical feature of PRC1 deficiency in ESCs[16,19] and germinal vesicle oocytes[10]. To test whether this is due to functional redundancy of the *Ring1* paralogue, we profiled expression in $Ring1^{\Delta}$ $Rnf2^{\Delta}$ PGCs purified from E11.5 embryos. Consistently, many developmental gene functions were over-represented among upregulated genes (Supplementary Fig. 8 and Supplementary Table 2). These data indicate that although *Ring1* expression is sufficient to safeguard global repression of canonical Polycomb target genes in PGCs, *Rnf2* is required in female E12.5 PGCs for repression of genes driving meiosis.

At E11.5, gonads still possess the potential to develop into either ovaries or testes, but are committed to a sex-specific differentiation process 2 days later. To relate aberrant expression in $Rnf2^{\Delta}$ PGCs to changes occurring during normal PGC differentiation, we profiled expression in $Rnf2^{+}$ PGCs isolated between E11.5 and E13.5. In this developmental period, sixfold more genes were upregulated in female (810) than in male (132) $Rnf2^{+}$ PGCs, with an additional 196 genes being upregulated in both sexes, suggesting the activation of female- and male-specific 'PGC-differentiation programs' (PDPs) (Fig. 2b). Among genes upregulated in E12.5 female $Rnf2^{\Delta}$ PGCs, 223 (43%) were found to be part of the female PDP (Fig. 2b and Supplementary Fig. 7b) ($P$ value $2.67 \times 10^{-159}$; geometric test), being activated about 1 day ahead compared to those in $Rnf2^{+}$ PGCs (Fig. 2c).

We next identified that 24 out of 119 genes annotated with the Gene Ontology term 'meiosis' (GO:0007126) were part of the female PDP, probably reflecting activation of an 'early meiosis program' (Fig. 2d). Among these 24 genes, 10 were precociously upregulated in female $Rnf2^{\Delta}$ PGCs, including *Stra8*, required for meiotic initiation[5], and others such as *Rec8*, *Sycp3*, and *Hormad2*, which have key functions in cohesion, chromosome synapsis and recombination[20–22]. Using qRT–PCR, we measured high *Stra8*, *Rec8* and *Sycp3* messenger RNA levels at E13.5 in female $Rnf2^{+}$ PGCs (Fig. 3a). At E11.5 and E12.5, *Stra8* and *Rec8* were precociously activated, with up to 20-fold higher expression levels in female $Rnf2^{\Delta}$ versus $Rnf2^{+}$ PGCs. *Sycp3* was expressed in $Rnf2^{+}$ PGCs from both sexes with up to fourfold increased levels in female $Rnf2^{\Delta}$ PGCs (Fig. 3a). Immunofluorescence analyses revealed strong Stra8 nuclear localization by E12.5 in PGCs of $Rnf2^{cko}$ gonads, whereas the protein only started to accumulate in control PGCs at E13.5 (Fig. 3b and Supplementary Fig. 7c). For Sycp3, we observed focal nuclear staining at E13.5 and synaptonemal complex staining at E14.5 in $Rnf2^{\Delta}$ germ cells, whereas only axial elements of meiotic chromosomes were visible in $Rnf2^{+}$ germ cells at E14.5 (Fig. 3c, d, data not shown). These data indicate that precocious transcriptional activation of *Stra8* and other PGC differentiation and meiosis genes in female $Rnf2^{\Delta}$ PGCs induces these cells to prematurely stop proliferation and enter into meiotic prophase, hence accounting for the lower number of female $Rnf2^{\Delta}$ PGCs at E12.5. We also measured increased expression of genes functioning in retinoic acid metabolism, such as *Aldh1a2*, *Crabp1* and *Crabp2* in female $Rnf2^{\Delta}$ PGCs at E11.5 and 12.5 (Supplementary Fig. 7d and data not shown), probably enhancing retinoic acid signalling and meiotic entry in a feed-forward manner[23].

We next wanted to find out why *Stra8* transcription is abnormally activated only in female $Rnf2^{\Delta}$ PGCs. In male gonads, retinoic-acid-mediated induction of *Stra8* is counteracted by the somatically expressed retinoid-degrading enzyme Cyp26b1 and by the Fgf9 signalling pathway (Supplementary Fig. 2)[3,4,24]. PRC1 may therefore not be required to suppress *Stra8* expression in males. To study this possibility, we aimed to overcome the antagonizing activity of Cyp26b1 and cultured E11.5 genital ridges for 24 h in the presence of all-*trans* retinoic acid (ATRA)[3,4]. ATRA increased *Stra8* expression in isolated $Rnf2^{+}$ PGCs of both sexes to levels comparable with what we measured in female $Rnf2^{\Delta}$ PGCs treated with vehicle (Fig. 4a). Intriguingly, ATRA additionally enhanced *Stra8* expression in male as in female $Rnf2^{\Delta}$ PGCs to >fivefold higher levels compared to $Rnf2^{+}$ PGCs (Fig. 4a), indicating that PRC1 effectively suppresses retinoic-acid-induced *Stra8* activation in PGCs of both sexes. We observed comparable responses for *Rec8* and *Sycp3* expression (Fig. 4a). By contrast, *Stra8* activation was completely suppressed in all genotypes by treatment with Win-18446 (*N*,*N*′-octamethylenebis(dichloroacetamide)), an inhibitor of the retinoic acid biogenesis pathway[25] (Fig. 4a). Likewise, 2-day *in vivo* exposure of PGCs to Win-18446 in developing embryos suppressed premature Stra8 protein expression (Supplementary Fig. 9a, b). In addition, numbers of Oct4-expressing PGCs in $Rnf2^{cko}$ gonads were significantly increased upon treatment *in vitro* and *in vivo* with Win-18446 (Fig. 4b and Supplementary Fig. 9c, d). These results indicate that sensitization of PGCs to retinoic acid signalling by *Rnf2* deficiency contributes to precocious exit from the PGC state and activation of the meiotic program in female gonads.

Subsequently, most $Rnf2^{\Delta}$ PGCs were unable to complete meiosis and to develop into mature oocytes (Supplementary Fig. 10), possibly as a consequence of the precocious entry into meiosis that may eventually impair, perhaps in concert with other aberrant changes in gene expression (Fig. 2b), the execution of the natural female PDP and oogenesis. Alternatively, PRC1 may exert a separate essential function later during meiotic progression.

Finally, we performed a micro-chromatin immunoprecipitation (μChIP) assay on isolated $Rnf2^{+}$ PGCs. At E11.5, the *Stra8* promoter was strongly enriched with PRC2-mediated H3K27me3 and with

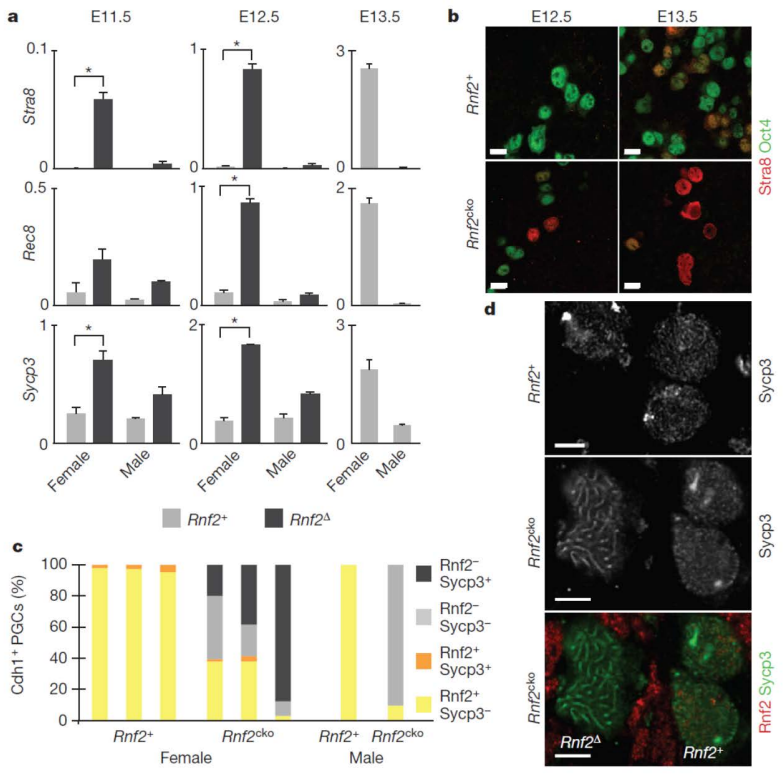



**Figure 3 | Female $Rnf2^{\Delta}$ PGCs enter precociously into meiotic prophase.**
**a**, Representative qRT–PCR data of *Stra8*, *Rec8 and Sycp3* transcripts (normalized to *Tbp*) in isolated PGCs. Error bars indicate +s.d. of two technical replicates. $*P < 0.05$ (Student's *t*-test).
**b**, Immunofluorescence staining of Oct4 and Stra8 in $Rnf2^{+}$ and $Rnf2^{cko}$ female gonads at E12.5 and E13.5. Scale bars, 10 μm. **c**, Classification of Cdh1-positive PGCs according to Rnf2 and Sycp3 protein levels in $Rnf2^{+}$ and $Rnf2^{cko}$ E13.5 gonads of individual embryos. Numbers in brackets indicate the number of PGCs scored per embryo.
**d**, Immunofluorescence staining of Sycp3 and Rnf2 in $Rnf2^{+}$ and $Rnf2^{cko}$ E14.5 ovaries. Scale bars, 10 μm.

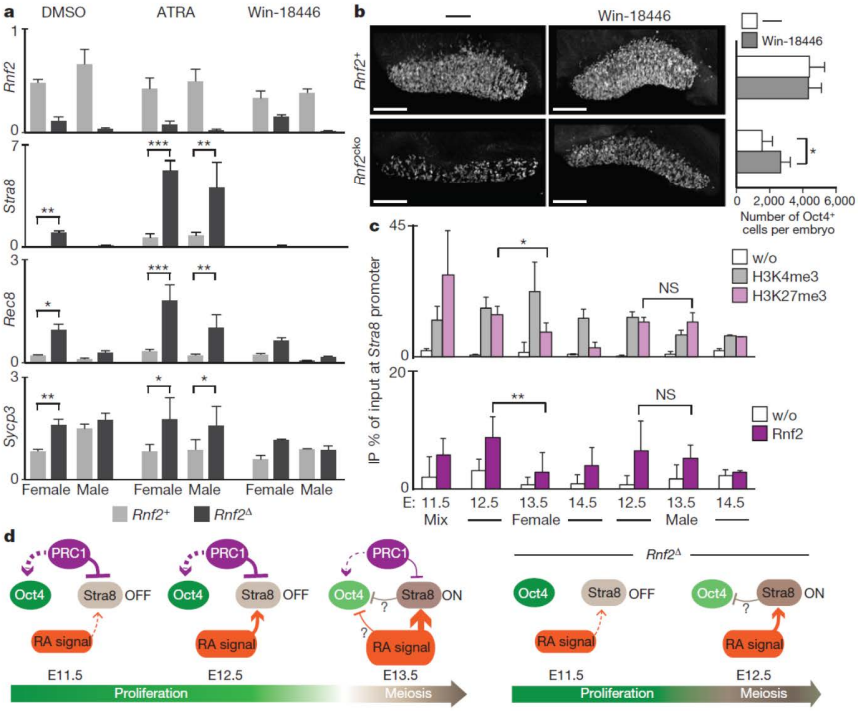



**Figure 4 | PRC1 antagonizes retinoic acid signalling and maintains *Stra8* in a repressive chromatin state.** **a**, Representative qRT–PCR data of *Rnf2*, *Stra8*, *Rec8* and *Sycp3* transcripts (normalized to *Tbp*) in isolated PGCs. Error bars indicate +s.d. of 2–4 technical replicates. $*P < 0.01$; $**P < 0.001$; $***P < 0.0001$ (Student's *t*-test). **b**, Immunofluorescence staining of Oct4 and number of Oct4-positive PGCs in $Rnf2^{+}$ and $Rnf2^{cko}$ whole gonads and mesonephroi from E12.5 female embryos, treated with Win-18446 or vehicle at E10.5. Graph shows means with +s.d. $n = 5–6$. Scale bars, 300 μm. **c**, μChIP analysis for H3K4me3, H2K27me3 and Rnf2 at the promoter region of *Stra8* in PGCs isolated from E11.5, 12.5, 13.5 and 14.5 $Rnf2^{+}$ gonads. Percentages of immunoprecipitated DNA (IP) compared to input are shown. w/o, without primary antibodies. Error bars indicate +s.d. of 1–3 biological and 2–4 technical replicates. $*P < 0.05$; $**P < 0.001$ (Student's *t*-test). **d**, Model for transcriptional regulation of meiotic genes (for example, *Stra8*) and pluripotency genes (for example, *Oct4*) during sex differentiation of female PGCs. RA, retinoic acid.

H3K4me3, indicative of a Polycomb repressed, yet potentially transcriptionally primed, chromatin state. During subsequent stages, we noticed a progressive decrease in H3K27me3 levels in female PGCs consistent with *Stra8* transcriptional activation, whereas the promoter remained bivalent in male PGCs (Fig. 4c). By contrast, the *Hoxa9* promoter was bivalent in all conditions (Supplementary Fig. 11). Likewise, we detected Rnf2 association with the *Stra8* promoter at E11.5 and E12.5 PGCs and a significant decrease in Rnf2 occupancy in female but not male PGCs at E13.5. Together, these experiments suggest that *Stra8* is directly regulated by PRC2 and PRC1 in PGCs of both sexes.

In summary, we identified an essential role for PRC1 in PGC development between E10.5 and E11.5. Later, between E11.5 and E13.5, *Rnf2* effectively modulates the sensitivity of PGCs to retinoic-acid-mediated induction of meiosis by directly controlling the competence of *Stra8* and probably of other meiotic genes for transcriptional activation. Like the proposed role of *Cbx2* in temporal co-linearity of *Hox* gene activation[26], we speculate that PRC1 maintains repression of *Stra8* and other genes of the PGC differentiation and early meiosis programs until retinoic acid signalling has reached a certain threshold (Fig. 4d). In addition, PRC1 regulates the expression of pluripotency genes in PGCs, possibly indirectly by suppressing transcription of negative regulators. Impairing PRC1 function in these parallel pathways probably leads to a synergistic effect, thereby promoting premature transition from proliferation into meiosis.

## METHODS SUMMARY

Embryos were obtained by timed matings, by scoring noon of the day after mating as E0.5. Genital ridges were cultured in drops of Dulbecco's minimal eagle medium supplemented with 10% fetal calf serum at 37 °C with 5% $CO_2$ in air. For expression profiling, we collected in triplicate 500 PGCs per embryo by fluorescence-activated cell sorting and processed, hybridized to Affymetrix Mouse Gene 1.0 arrays and analysed as described previously[10]. Gene Ontology terms were obtained using GOstat (http://gostat.wehi.edu.au). μChIP was performed as described previously[27].

**Full Methods** and any associated references are available in the online version of the paper.

**Received 27 April 2012; accepted 16 January 2013.**


1. Brennan, J. & Capel, B. One tissue, two fates: molecular genetic events that underlie testis versus ovary development. *Nature Rev. Genet.* **5**, 509–521 (2004).
2. Menke, D. B., Koubova, J. & Page, D. C. Sexual differentiation of germ cells in XX mouse gonads occurs in an anterior-to-posterior wave. *Dev. Biol.* **262**, 303–312 (2003).
3. Koubova, J. *et al.* Retinoic acid regulates sex-specific timing of meiotic initiation in mice. *Proc. Natl Acad. Sci. USA* **103**, 2474–2479 (2006).
4. Bowles, J. *et al.* Retinoid signaling determines germ cell fate in mice. *Science* **312**, 596–600 (2006).
5. Baltus, A. E. *et al.* In germ cells of mouse embryonic ovaries, the decision to enter meiosis precedes premeiotic DNA replication. *Nature Genet.* **38**, 1430–1434 (2006).
6. Deschamps, J. Ancestral and recently recruited global control of the *Hox* genes in development. *Curr. Opin. Genet. Dev.* **17**, 422–427 (2007).
7. Duester, G. Retinoic acid synthesis and signaling during early organogenesis. *Cell* **134**, 921–931 (2008).
8. Oulad-Abdelghani, M. *et al.* Characterization of a premeiotic germ cell-specific cytoplasmic protein encoded by Stra8, a novel retinoic acid-responsive gene. *J. Cell Biol.* **135**, 469–477 (1996).
9. Sparmann, A. & van Lohuizen, M. Polycomb silencers control cell fate, development and cancer. *Nature Rev. Cancer* **6**, 846–856 (2006).
10. Posfai, E. *et al.* Polycomb function during oogenesis is required for mouse embryonic development. *Genes Dev.* **26**, 920–932 (2012).
11. Saitou, M., Kagiwada, S. & Kurimoto, K. Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells. *Development* **139**, 15–31 (2012).
12. Simon, J. A. & Kingston, R. E. Mechanisms of polycomb gene silencing: knowns and unknowns. *Nature Rev. Mol. Cell Biol.* **10**, 697–708 (2009).
13. Di Carlo, A. & De Felici, M. A role for E-cadherin in mouse primordial germ cell development. *Dev. Biol.* **226**, 209–219 (2000).
14. Puschendorf, M. *et al.* PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos. *Nature Genet.* **40**, 411–420 (2008).
15. Lomelí, H., Ramos-Mejia, V., Gertsenstein, M., Lobe, C. G. & Nagy, A. Targeted insertion of Cre recombinase into the TNAP gene: excision in primordial germ cells. *Genesis* **26**, 116–117 (2000).
16. Endoh, M. *et al.* Polycomb group proteins Ring1A/B are functionally linked to the core transcriptional regulatory circuitry to maintain ES cell identity. *Development* **135**, 1513–1524 (2008).

17. del Mar Lorente, M. *et al.* Loss- and gain-of-function mutations show a polycomb group function for Ring1A in mice. *Development* **127,** 5093–5100 (2000).
18. Yoshimizu, T. *et al.* Germline-specific expression of the Oct-4/green fluorescent protein (GFP) transgene in mice. *Dev. Growth Differ.* **41,** 675–684 (1999).
19. Leeb, M. *et al.* Polycomb complexes act redundantly to repress genomic repeats and genes. *Genes Dev.* **24,** 265–276 (2010).
20. Xu, H., Beasley, M. D., Warren, W. D., van der Horst, G. T. & McKay, M. J. Absence of mouse REC8 cohesin promotes synapsis of sister chromatids in meiosis. *Dev. Cell* **8,** 949–961 (2005).
21. Yuan, L. *et al.* Female germ cell aneuploidy and embryo death in mice lacking the meiosis-specific protein SCP3. *Science* **296,** 1115–1118 (2002).
22. Wojtasz, L. *et al.* Meiotic DNA double-strand breaks and chromosome asynapsis in mice are monitored by distinct HORMAD2-independent and -dependent mechanisms. *Genes Dev.* **26,** 958–973 (2012).
23. Maden, M. Retinoic acid in the development, regeneration and maintenance of the nervous system. *Nature Rev. Neurosci.* **8,** 755–765 (2007).
24. Bowles, J. *et al.* FGF9 suppresses meiosis and promotes male germ cell fate in mice. *Dev. Cell* **19,** 440–449 (2010).
25. Hogarth, C. A. *et al.* Suppression of *Stra8* expression in the mouse gonad by WIN 18,446. *Biol. Reprod.* **84,** 957–965 (2011).
26. Bel-Vialar, S. *et al.* Altered retinoic acid sensitivity and temporal expression of *Hox* genes in *polycomb-M33*-deficient mice. *Dev. Biol.* **224,** 238–249 (2000).
27. Dahl, J. A. & Collas, P. A rapid micro chromatin immunoprecipitation assay (ChIP). *Nature Protocols* **3,** 1032–1045 (2008).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank A. Nagy and H. Schöler for the *Tnap-cre* and *Oct4(ΔPE)*–GFP mice, respectively, and M. Griswold for the Stra8 antibody. We are grateful to L. Gelman (microscopy and imaging), L. Burger, M. Stadler (bioinformatics), E. Cabuy, S. Thiry, K. Jacobeit (functional genomics) and the FMI animal facility for assistance. We thank members of the Peters laboratory, particularly M. Tardat, S. Erkek and M. Gill, for experimental support and discussions. S.Y. is a recipient of a Human Frontier Science Program long-term fellowship and a Japan Society for the Promotion of Science postdoctoral fellowship. Research in the Peters laboratory is supported by the Novartis Research Foundation, the Swiss National Science Foundation (31003A_125386 and NRP 63 - Stem Cells and Regenerative Medicine), SystemsX.ch (Cell plasticity), the Japanese Swiss Science and Technology Cooperation Program, the European Network of Excellence 'The Epigenome' and the EMBO Young Investigator program.

**Author Contributions** S.Y. and A.H.F.M.P. conceived and designed the experiments. S.Y. performed almost all experiments. C.-Y.L. performed μChIP experiments. H.K. performed FACS isolations. Z.L. isolated germinal vesicle oocytes. P.N. isolated RNA for RNA sequencing and provided advice on specific target amplification qRT-PCR. M.V. provided *Ring1*-deficient mice. M.v.L. provided *Rnf2* conditionally deficient mice. T.C.R. assisted in microarray and RNA-sequencing analysis. S.Y., C.-Y.L. and A.H.F.M.P. analysed the data. S.Y. and A.H.F.M.P. wrote the manuscript.

**Author Information** Microarray and RNA-sequencing data have been deposited in the Gene Expression Omnibus under accession numbers GSE42782 and GSE42852, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.H.F.M.P. (antoine.peters@fmi.ch).

## METHODS

**Mice and embryo collection.** *Rnf2*$^{cko}$ mice with *Rnf2*-deficient PGCs were generated by combining a floxed *Rnf2* (*Rnf2*$^{fl}$) allele[14] with the *Tnap-cre* transgene allele as illustrated in Supplementary Fig. 3. Introduction of the *Ring1* mutation[17] is illustrated in Supplementary Fig. 6a. For PGC isolation, embryos were sired by fathers that were homozygous for the *Oct4*(*ΔPE*)–GFP transgene. Mice were maintained on a mixed background of 129/Sv and C57BL/6J. Embryos were obtained by timed matings, by scoring noon of the day following mating as 0.5 embryonic day of development (E0.5). The genotype of embryos was determined by PCR as described previously[14]. Sex of gonads was determined by PCR for *Ubex1* using the following primers: forward, 5′-TGGTCTGGACCCAAACGCTG TCCACA-3′; reverse, 5′- GGCAGCAGCCATCACATAATCCAGATG-3′. All experiments were performed in accordance with the Swiss Animal Protection laws and institutional guidelines.

**Antibodies.** For immunofluorescence analyses, the following primary and secondary antibodies were used: polyclonal anti-Oct4 (sc-8628, 1:150), monoclonal anti-Rnf2 (gift from H. Koseki, 1:400)[28], monoclonal anti-E-cadherin (Invitrogen, 1:250), polyclonal anti-Stra8 (rabbit, gift from M. Griswold, 1:1,000)[29], polyclonal anti-Sycp3 (rabbit, gift from C. Heyting, 1:500)[30], polyclonal anti-Mel18 (sc-10774, 1:50), polyclonal anti-Rybp (1:400)[31], monoclonal anti-H2AK119u1 (Cell signaling, 1:500), polyclonal anti-H3K27me3 (gift from T. Jenuwein, 1:500)[32], monoclonal anti-Ezh2 (Novocastra, 1:200), anti-goat IgG-Alexa 488, anti-rabbit IgG-Alexa 488, anti-mouse IgG-Alexa 555 and anti-rat Cy5. For ChIP analysis, anti-H3K4me3 (Millipore, 17-614), anti-H3K27me3 (Millipore, 07-449) and anti-Rnf2 (Active motif, 39663) were used.

**Immunofluorescence.** For whole-mount stainings, dissected gonads with mesonephroi were fixed for 15 min in 3% paraformaldehyde in PBS (pH 7.4) and permeabilized with 0.5% Triton X-100 in PBS for 20 min on ice. Fixed embryos were blocked overnight at 4 °C in PBS containing 0.1% Triton X-100, 10% BSA and 5% normal donkey serum, and were then incubated with primary antibodies in blocking solution overnight at 4 °C. Gonads were washed three times for 1 h in PBS containing 0.1% Triton X-100 and 2% BSA before application of secondary antibodies. For detection, secondary antibodies were diluted 1:500 in blocking solution and gonads were incubated overnight at 4 °C followed by three washing steps for 1 h in PBS with 0.1% Triton X-100. Gonads were stained briefly with DAPI and mounted in Vectashield (Vector). For gonadal section stainings, the posterior part of embryos or gonads with mesonephroi were frozen in Tissue-Tek Optimal Cutting Temperature (OCT) compound (Sakura Finetek) on dry ice. Alternatively, the materials were fixed with 3% paraformaldehyde in PBS (pH 7.4) for 10 min, soaked in 30% sucrose solution overnight and embedded in OCT compound. Twelve-micron-thick cryo-sections were cut from frozen blocks with Microm HM355S. Cryo-sections were fixed with 3% paraformaldehyde for 10 min at room temperature (about 20 °C), permeabilized in 0.5% Triton X-100 in PBS for 4 min at 4 °C and blocked for 30 min in PBS containing 1% BSA at room temperature. Sections were incubated with primary antibodies in blocking solution overnight at 4 °C and subsequently washed three times for 10 min in PBS with 0.05% Tween-20. Incubation of secondary antibodies was done in the blocking solution for 1 h at room temperature.

**Microscopy and image analysis.** Immunofluorescence stainings of gonads were analysed using the Zeiss LSM 700 confocal microscope. For whole-mount gonads, images were acquired by using a tile function with a *z*-series of 1-µm slices in ZEN software and whole image was reconstructed using the XUV-tools software. We counted the number of Oct4-positive cells using a spot function in Imaris (Bitplane) software.

**Isolation of PGCs expressing the *Oct4*(*ΔPE*)–GFP transgene by FACS.** Dissected gonads were enzymatically disrupted using 0.025% trypsin at 37 °C for 8 min. Trypsin activity was inhibited by adding fetal calf serum in Hank's buffer salt solution without phenol red. Gonads were dispersed by pipetting and subjected to FACS. Embryos were processed individually for expression analysis. PGCs isolated from several embryos were pooled for ChIP analysis.

**qRT–PCR.** Total RNA was extracted from isolated PGCs or surrounding somatic cells from individual embryos using the PicoPure RNA Isolation Kit (KIT0202) according to the manufacturer's instructions (Stratagene) with the addition of 100 ng *Escherichia coli* ribosomal RNA as carrier. RT–PCR was performed using SuperScript III Reverse Transcriptase (Invitrogen) according to the manufacturer's protocol. qPCR reactions were performed with complementary DNA corresponding to 20 cells using the SYBR Green PCR Master Mix (Applied Biosystem) in an ABI Prism 7000 Real time PCR machine. All qPCR measurements were normalized to the endogenous expression level of *Tbp*. We performed qRT–PCR analyses on multiple pairs of *Rnf2*$^{cko}$ and control littermates. Data in figures present technical replicates for pairs of genotypes. Primers used were as follows: *Rnf2* (forward, 5′-TTAGAAGTGGCAACAAAGAGTG-3′; reverse, 5′-CGCTTCATACTCATCACGAC-3′), *Oct4* (forward, 5′-GATGCTGTGAGCCA AGGCAAG-3′; reverse, 5′-GGCTCCTGATCAACAGCATCAC-3′), *Nanog* (forward, 5′-CTTTCACCTATTAAGGTGCTTGC-3′; reverse, 5′-TGGCATCGG TTCATCATGGTAC-3′), *Stra8* (forward, 5′-CAAAAGCCTTGGCTGTGTTA-3′; reverse, 5′-AAAGGTCTCCAGGCACTTCA-3′), *Rec8* (forward, 5′-CCAA CAAGGAGCTGGACTTC-3′; reverse, 5′-GGACAGCACCAAGAGCAGAT-3′), *Sycp3* (forward, 5′-GTGTTGCAGCAGTGGGAAC-3′; reverse, 5′-GCTTT CATTCTCTGGCTCTGA-3′), *Aldh1a2* (forward, 5′-CCCTGACAGTGGCT TTGAGT-3′; reverse, 5′-CTGTGGGTTGAAGGGAGCTA-3′), *Crabp1* (forward, 5′-GCTTCGAGGAGGAGACAGTG-3′; reverse, 5′- CAGCTCTCGGGTCCAG TAAG-3′), *Crabp2* (forward, 5′-GCCGAGAACTGACCAATGAT-3′; reverse, 5′-GGAAGTCGTCTCAGGCAGTT-3′), *Tbp* (forward, 5′-TGCTGTTGGTGAT TGTTGGT-3′; reverse, 5′-AACTGGCTTGTGTGGGAAAG-3′).

**Expression profiling of PGCs and data analysis.** We performed expression profiling on PGCs isolated from three pairs of *Rnf2*$^{cko}$ and control littermates for each developmental time point. RNA was extracted from 500 PGCs isolated per embryo using the PicoPure RNA Isolation Kit (KIT0202) according to the manufacturer's instructions (Stratagene). The quality of the RNA was assessed using the Agilent 2100 Bioanalyzer and RNA 6000 Pico Chip. The extracted RNA was converted into OmniPlex Whole transcriptome amplification (WTA) cDNA libraries and amplified by WTA PCR using reagents supplied with the TransPlex WTA1 kit (Sigma) following the manufacturer's instructions with minor modifications. The obtained cDNA was purified using the GeneChip cDNA Sample Cleanup Module (Affymetrix). The labelling, fragmentation and hybridization of cDNA was performed according to Affymetrix instructions (GeneChip Whole Transcription Sense Target Labelling technical manual, Rev. 2) with minor modifications. Samples were hybridized to Affymetrx Mouse Gene 1.0 arrays. Microarray quality control and analysis was carried out in R 2.10.0 and Bioconductor 2.5. In brief, array quality was assessed using the 'arrayQualityMetrics' package. Raw data was read into R and normalized with RMA using the 'affy' package and differentially expressed genes were identified using the empirical Bayes method (*F* test) implemented in the LIMMA package. *P* values were adjusted for false discovery rate (FDR) using the Benjamini and Hochberg correction. Probe sets with a log2 average contrast signal of at least 3, an adjusted *P* value of $<0.05$, and an absolute linear fold-change of at least 1.5-fold were selected. The *P* values reported for enriched Gene Ontology terms were obtained using GOstat (http://gostat.wehi.edu.au) (Supplementary Figs 7 and 8 and Supplementary Table 2). A list of genes belonging to the GO term 'meiosis' (GO:0007126) were obtained using R annotation packages (library(org.Mm.eg.db) and library(GO.db)) (Fig. 2d and Supplementary Table 1). For RNA-sequencing experiments, RNA was extracted from 500 PGCs per embryo using RNeasy Micro Kit (Qiagen). The RNA amplification and cDNA generation were performed using NuGEN Ovation RNA-seq System V2 (Part no. 7102) and sequencing libraries were prepared using TruSeq DNA Sample Preparation Kit (low-throughput protocol) (Part no. 15005180 Rev. C). Barcoded libraries were sequenced in one lane of Illumina HiSeq instrument. The resulting sequencing reads were filtered, aligned to Refseq gene models and weighted as described previously[33].

**Treatment of gonads with agonist and antagonist of retinoic acid signalling.** Both gonads per embryo were dissected and were cultured together or separately in a drop of Dulbecco's minimal eagle medium supplemented with 10% fetal calf serum at 37 °C in a humidified atmosphere of 5% $CO_2$ in air. ATRA (Sigma) and Win-18446 (ABCR) were dissolved in dimethylsulphoxide (DMSO). These compounds were added to culture media with concentration of 0.5 µM for ATRA or 2 µM for Win-18446. Control cultures were treated with DMSO vehicle as appropriate. For *in vivo* administration of Win-18446, 100 mg per ml stock solution (in DMSO) was diluted with oil and 2.5 mg was injected intraperitoneally into pregnant female mice at E10.5. Embryos were collected at E12.5. DMSO mixed with oil was injected into pregnant female mice for control experiments.

**Specific target amplification (STA) qRT–PCR on single germinal vesicle oocytes.** Germinal vesicle oocytes were isolated from ovaries and carefully washed with removing cumulus cells. STA was performed using the CellsDirect One-Step qRT–PCR Kit (11753-100) according to the manufacturer's instructions (Invitrogen). In brief, the oocytes were individually added to the reaction mix containing RNase Inhibitor (Ambion) and primers of target genes of interest. After amplification, the samples were treated with Exonuclease I (New England Biolabs) and used as template for qPCR. Primers used were as follows: *Rnf2* (forward, 5′-CAGGCCCCATCCAACTCTTA-3′; reverse, 5′-CAACAGTGGCA TTGCCTGAA-3′), *Ssu72* (forward, 5′-GGTGTGCTCGAGTAACCAGAA-3′; reverse, 5′- CAAAGGAGCGGACACTGAAAC-3′).

**µChIP analysis.** Small-scale ChIP experiment was performed as previously described[27] with some modifications. In brief, 15,000 FACS-sorted PGCs were cross-linked with 0.5% paraformaldehyde in PBS for 10 min at room temperature and quenched with 125 mM glycine. PGCs were then lysed in 50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 1% SDS, and protease inhibitors. The cell lysate was sonicated in a

Diagenode Bioruptor to achieve a mean DNA fragment size of around 200–400 base pairs. After centrifugation, the supernatants were diluted with radio immunoprecipitation assay (RIPA) buffer to an equivalent of 2,500 PGCs and incubated with antibody-bound protein G-magnetic beads overnight at 4 °C. The beads were washed four times with the RIPA buffer and one time with 10 mM Tris HCl, pH 8.0, 10 mM EDTA buffer, and bead-bound complexes were incubated with complete elution buffer (20mM Tris-HCl, pH 7.5, 5 mM EDTA, 50 mM NaCl, 1% SDS, 50 mg per ml proteinase K) at 68 °C for DNA elution, crosslink reversal and protein digestion. Finally, immunoprecipitated DNA was purified by phenol-chloroform extraction and ethanol precipitation and dissolved in MilliQ water for qRT–PCR analysis. Primers used were as follows: *Stra8* (forward, 5′-GTATCGCC GTAACTCCCAGA-3′; reverse, 5′-GCAGATGACCCTCACACAAG-3′), *HoxA9* (forward, 5′- GGAGGGAGGGGAGTAACAAA-3′; reverse, 5′-TCACCTCGCCT AGTTTCTGG-3′).


28. Atsuta, T. *et al.* Production of monoclonal antibodies against mammalian Ring1B proteins. *Hybridoma* **20,** 43–46 (2001).
29. Zhou, Q. *et al.* Expression of stimulated by retinoic acid gene 8 (Stra8) and maturation of murine gonocytes and spermatogonia induced by retinoic acid in vitro. *Biol. Reprod.* http://dx.doi.org10.1095/biolreprod.107.064337 (2008).
30. Lammers, J. H. *et al.* The gene encoding a major component of the lateral elements of synaptonemal complexes of the rat is related to X-linked lymphocyte-regulated genes. *Mol. Cell. Biol.* **14,** 1137–1146 (1994).
31. García, E., Marcos-Gutierrez, C., del Mar Lorente, M., Moreno, J. C. & Vidal, M. RYBP, a new repressor protein that interacts with components of the mammalian Polycomb complex, and with the transcription factor YY1. *EMBO J.* **18,** 3404–3418 (1999).
32. Peters, A. H. *et al.* Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Mol. Cell* **12,** 1577–1589 (2003).
33. Tippmann, S. C. *et al.* Chromatin measurements reveal contributions of synthesis and decay to steady-state mRNA levels. *Mol. Syst. Biol.* **8,** 593 (2012).

# Chapter 4. Submitted manuscript: Dynamic expression of chromatin modifiers during developmental transitions in preimplantation embryos

Peter Nestorov*,†, Hans-Rudolf Hotz*, Antoine H.F.M. Peters*,†,2

*Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland †Faculty of Sciences, University of Basel, Basel, Switzerland [1]Equal contribution.

[2]Corresponding author: e-mail address: antoine.peters@fmi.ch

## Abstract

There are two major developmental transitions that take place during the first four days of mouse embryogenesis. The first one is the maternal-to-zygotic transition in the zygote and 2-cell embryo, the second one is the first cell lineage commitment at the 32-cell stage. A number of transcription factors govern zygotic expression and the specification of pluripotent versus trophectoderm cells, while at the same time there are dynamic changes in DNA methylation and chromatin organization. It is not fully understood which chromatin modifying complexes take part in these processes and whether there is a particular spatiotemporal distribution of certain chromatin players in preimplantation development. To address this question, we performed gene expression profiling of single cells from the oocyte to the blastocyst stage. We describe the patterns of expression of over 100 genes involved in histone methylation, DNA methylation and chromatin remodelling. For a number of these genes we observed differential maternal versus zygotic expression. This is particularly the case for homologous genes like Tet1 and Tet3 (zygotic and maternal respectively), or Suv39h1 and Suv39h2 (zygotic and maternal respectively). In contrast to the maternal-to-zygotic switch, most chromatin modifying genes are ubiquitously expressed in the subsequent lineage specification at the 32-cell stage. The few exceptions that showed a differential expression pattern are Kdm1b, Tet1 and Prdm14. Our data suggests that genes coding for histone modifying complexes are ubiquitously expressed in early embryos and that there are maternal and zygotic variants of most of these complexes. Furthermore, the method that we applied, allows the distinction of female versus male embryos based on the expression of Xist (X-linked, expressed in female embryos) and Kdm5d (Y-linked, expressed strongly in male embryo). With the exception of the X- and Y-linked genes, we did not observe sex-specific gene expression patterns of chromatin modifiers or transcription factors.

## 4.1 Introduction

The first steps of mouse embryogenesis involve chromatin remodelling and fusion of the two parental genomes, followed by activation of zygotic transcription and first cell fate specifications. These critical processes take place in a narrow time window of four days and are controlled in great part by the maternally provided components (Li et al., 2010a). The zygotic genome activation (ZGA) is a process occurring in two waves, starting in the zygote stage and continuing with a major burst at the 2-cell stage, while at the same time the maternally provided transcripts are actively degraded (Schultz 2002, Wang and Dey 2006). After the embryo has taken control over its genome at the 2-cell stage, the blastomeres undergo several divisions leading to the 16-cell morula stage. Upon compaction of the morula and progression to the 32-cell blastocyst stage, the outer cells are committed to the trophectoderm lineage (epithelium cells required for embryo implantation) and enclose the blastocyst cavity, in which the cells of the inner cell mass (ICM) reside. Thus the early blastocyst consists of two distinct cell types. Finally, before implantation, the ICM further differentiates into primitive endoderm (PE, contributing to extra-embryonic tissue) and epiblast (EPI, pluripotent cells that will form the embryo matter) (Rossant and Tam, 2009).

Lineage specification in the early embryo is regulated by transcriptional networks, driven by lineage-specific transcription factors, such as Cdx2, Sox2 and Oct4 (alias Pou5f1, referred to hereafter a Pou5f1). The first significant differences between inner and outer cells in terms of cell fate and gene expression become evident at the 16-cell stage (Guo et al., 2010), but a number of studies suggest that cell fate specification is initiated well before the 16-cell stage (Bruce and Zernicka-Goetz, 2010; Wennekamp et al., 2013). However, it is still not fully understood what triggers the process and how cell fate is maintained early on in a system with high cell plasticity and stochastic fluctuations.

Before the first cell lineages are specified, there are dynamic changes that happen to chromatin and DNA, which are potentially crucial for the subsequent developmental decisions. The most striking chromatin rearrangements and modifications happen in the zygote, where the two parental genomes are organized in two separate pronuclei and fuse only after DNA replication has taken place. The paternal genome comes in a highly compacted and histone-depleted

conformation, and rapidly undergoes decompression and incorporation of maternal histones and protein complexes (Gill et al., 2012). The histone and nucleosome assembly on the paternal genome are accompanied by post-translational modifications on the histones, as well as alteration of the DNA methylation state by Tet proteins (Albert and Peters, 2009; Gkountela and Clark, 2014; Saitou et al., 2012). In particular, the paternal pronucleus gains methylation on lysines 4 and 27 on histone H3 (H3K4me3 and H3K27me3 respectively), while these marks are pre-established and maintained in the female pronucleus. There is also an asymmetry between the female and male pronuclei in respect to the heterochromatin component - paternally it is marked by H3K27me3, while maternally there is methylation of lysine 9 on histone H3 (H3K9me3) and this difference persists until the 8-cell stage (Puschendorf et al., 2008). As for DNA methylation, the paternal pronucleus in the zygote undergoes active hydroxylation of the methylated cytosines (5-hmC), while the maternal genome largely retains its methylation status up to the blastocyst stage (Smith et al., 2012). Interestingly, the histone modification H3K9me2 and DNA methylation have been linked through the activity of Dppa3 (also known as PGC7/Stella), which has been shown to bind H3K9me2 and protect the maternal genome from hydroxymethylation by Tet proteins (Nakamura et al., 2012).

Another process occurring during preimplantation development and associated with chromatin changes is the imprinted inactivation of the paternal X (Xp) chromosome in female embryos (Takagi and Sasaki, 1975). It takes place after ZGA and is presumably triggered by the expression of the non-coding RNA Xist from the paternal X (Jeon et al., 2012). Thereafter, Xp remains silenced in the trophectoderm lineage, while in the ICM it becomes reactivated. The silencing of Xp is reinforced by Polycomb-mediated H3K27me3 and is also accompanied by hypomethylation of H3K4 and hypoacetylaion of H3K9 (Okamoto et al., 2004; Plath et al., 2003; Wang et al., 2001). Furthermore, imprinted X inactivation seems to be a reversible and stochastic process that ensues at different times in different blastomeres of the preimplantation embryo (Mak et al., 2004; Patrat et al., 2009).

The dynamic changes that happen at chromatin during preimplantation suggest an important role for histone modifying enzymes. It is also suggested that chromatin-borne parental information, recognized by chromatin-binding proteins, may influence ZGA and lineage specification in the mouse embryo (Brykczynska et al., 2010; Erkek et al., 2013). Genome-wide

expression profiling on pooled embryos gives insights into the global expression changes that happen and reveals the core transcriptional networks at work (Park et al., 2013b; Tan et al., 2013; Zeng et al., 2004). However, given the stochastic nature of the developmental processes in preimplantation embryos (Bruce and Zernicka-Goetz, 2010; Dietrich and Hiiragi, 2007; Ohnishi et al., 2014; Wennekamp et al., 2013), a single-cell profiling approach has the potential to shed more light on the possible gene expression patterns and the relation between chromatin modifiers and transcription factors (Burton et al., 2013).

The aim of the current study was to describe the expression dynamics of chromatin modifying complexes in preimplantation embryos, particularly emphasizing on histone methylation. The most prominent histone methylation pathways are related to the post-transcriptional modification of Lysine residues 4, 9, 27 and 36 on Histone H3 (H3K4, H3K9, H3K27 and H3K26 respectively). Generally, methylated H3K4 and H3K36 are associated with actively transcribed loci, while methylated H3K9 and H3K27 are linked to transcriptional repression. The histone methylation state can be altered by histone methyltransferases or by histone demethylases, which can be targeted to a given genomic locus by chromatin binding proteins or transcription factors.

## 4.2 Results

### 4.2.1 Selection of genes and single-cell profiling

In order to understand how chromatin modifying complexes take part in the dynamic transcriptional changes during mouse preimplantation development, we profiled the expression of 192 genes in a total of 168 oocytes, zygotes and single cells or groups of cells isolated from preimplantation embryos (Figure 4.1A). For the qPCR detection on the BioMark 96.96 chips, we used the EvaGreen chemistry, which gives the opportunity to check for unspecific products and primer dimers. After performing the QC analysis using the BioMark software, we had to discard the data for 36 genes due to non-specific or noisy signal. Data analysis was performed for the remaining 156 targets, which comprised genes coding for

histone methyltransferases, histone demethylases, chromatin-binding proteins, as well as key transcription factors and signalling molecules (Figure 4.1B, Table 4.1). As a reference set for stage- and lineage-specific expression, we included 20 of the 48 genes used in the first Fluidigm-based study of mouse embryos by Guo et al. (Guo et al., 2010). The complete dataset presented in the current study consists of 26,208 data points (156 genes by 168 samples), which were processed and normalized to the mean expression signal of the three endogenous control genes *Hnrnpr*, *Ssu72* and *Ube2e1* (see Methods for detailed description of the bioinformatics analysis). The obtained normalized expression values are in the logarithmic space and are centred around the mean value of the control genes (i.e. for every sample the mean value of the three endogenous controls is close to zero). Based on the frequency distribution of all normalized expression values, we determined the threshold of expression to be -10 (Figure 4.1C). Using this threshold, we identified the majority of the 156 analysed genes as maternal transcripts, with only 17 genes showing no expression in MII oocytes (the genes had a mean expression value for the MII oocyte stage below -10). The dynamic expression of the selected genes allowed us to confidently resolve the developmental transitions that occur during preimplantation development, which is visualized in the principle component plot in Figure 4.1E. The strongest difference was the transition from the transcriptionally silent MII and zygote stages to the transcriptionally active 2-cell and later preimplantation stages, i.e. the maternal-to-zygotic transition. We also observed lineage-specific expression in blastomeres of the morula (16-cell) and blastocyst (32-cell) stages. The contribution of single genes to the two principal components in Figure 4.1E is visualized in the loadings plot in Figure 4.1F and shows a predominantly maternal-zygotic difference. The genes that account for most of the differences between single cells (i.e. genes that have highest loading value for the principle component analysis in the first two projections) are the maternal genes *Dazl, Mecom* (alias *Prdm3*), *Prdm6, Scml2, Tet3* and *Zp3* on one side, and the zygotic genes *Cbx6, Elf5, Esrrb, Fgf4, Fgfr3, Hand1, Nanog, Suv39h1, Tet1, Xist* and *2410016O06Rik* (alias *NO66*, referred to below as *NO66*) on the other (see Table 4.1 for detailed gene information and for the normalized dataset). Furthermore, there is a group of genes that are expressed through all stages from oocyte to blastocyst but exhibit very strong changes between stages and therefore contribute significantly to the principle component analysis. This group of genes comprises the

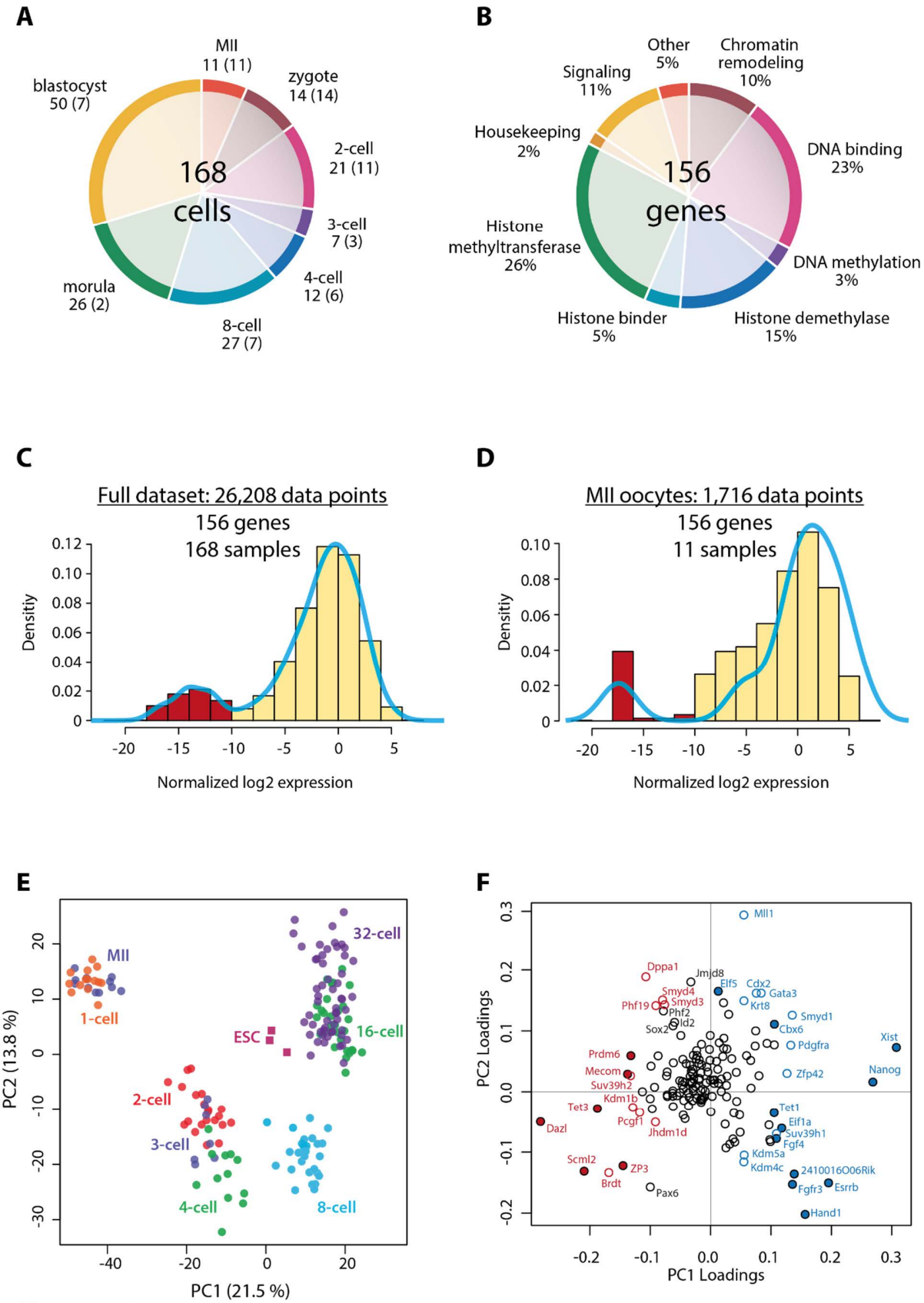

A
MII
11 (11)
zygote
14 (14)
2-cell
21 (11)
3-cell
7 (3)
4-cell
12 (6)
8-cell
27 (7)
morula
26 (2)
blastocyst
50 (7)
168
cells
B
Other
5%
Chromatin
remodeling
10%
Signaling
11%
Housekeeping
2%
DNA binding
23%
156
genes
Histone
methyltransferase
26%
DNA methylation
3%
Histone binder
5%
Histone demethylase
15%
C
Full dataset: 26,208 data points
156 genes
168 samples
Density
Normalized log2 expression
D
MII oocytes: 1,716 data points
156 genes
11 samples
Density
Normalized log2 expression
E
MII
1-cell
2-cell
3-cell
4-cell
ESC
32-cell
16-cell
8-cell
PC1 (21.5 %)
PC2 (13.8 %)
F
Mll1
Dppa1
Jmjd8
Elf5
Cdx2
Gata3
Smyd4
Smyd3
Krt8
Phf19
Phf2
Id2
Sox2
Smyd1
Cbx6
Xist
Pdgfra
Prdm6
Mecom
Zfp42
Nanog
Suv39h2
Kdm1b
Tet3
Tet1
Pcgf1
Eif1a
Dazl
Jhdm1d
Suv39h1
Fgf4
Kdm5a
Scml2
Kdm4c
ZP3
2410016O06Rik
Brdt
Fgfr3
Esrrb
Pax6
Hand1
PC1 Loadings
PC2 Loadings

***Figure 4-1.***

predominantly maternal *Brdt, Dppa1, Jhdm1d, Kdm1b, Pcgf1, Phf19, Smyd3, Smyd4* and *Suv39h2*, as well as the predominantly zygotic *Cbx6, Cdx2, Eif1a, Gata3, Kdm4c, Kdm5a, Krt8, Mll1, Pdgfra, Smyd1* and *Zfp42*.

Next, we analysed the expression patterns of the 156 genes by centring the values for each gene on zero and applying hierarchical clustering to the centred dataset (Figure 4.2). By performing this manipulation, we could identify genes that show similar expression behaviour independent of the absolute expression intensity (for example two genes that show increasing signal from MII oocyte to blastocyst will be clustered together even if they differ in their absolute expression signal). The clustering yielded three main branches, which could be classified as "maternal", "ubiquitous" and "zygotic", respectively corresponding to a decreasing, constant, or increasing expression from MII oocyte to blastocyst. Most of the analysed genes (N=67) demonstrated stable expression levels across stages. A similar portion of the genes (N=58) were expressed stronger upon zygotic genome activation at the 2-cell stage. Finally, 31 genes exhibited constantly decreasing levels, indicating maternally provided transcripts subject to degradation. Lineage-specific genes, expressed only in a subset of the 16-cell and 32-cell stage blastomeres, were predominantly found in the "zygotic" class. Some "maternal" genes also showed differential expression in the morula and blastocyst embryos, though not necessarily correlated to the inner/outer lineage-specific cell identity.

***Figure 4.1 Single-cell expression of chromatin modifiers during preimplantation development***

***A*** *– Distribution of samples by stage. Embryo numbers are shown in brackets.*

***B*** *– Distribution of genes by function. Since some genes have multiple functions, the distributions are given as a percentage.*

***C*** *– Frequency distribution of normalized log2 expression signal for the complete dataset. Red bars indicate samples below the threshold of detection.*

***D*** *– Frequency distribution of normalized log2 expression values for the MII oocyte samples only (maternal expression).*

***E*** *– Principle component analysis (PCA) for complete dataset. Data points correspond to samples (single cells) and are coloured by embryonic stage. In addition, three samples from embryonic stem cells RNA (ESC) are included in the PCA. Percentages indicate the variability explained by the first and second principle component respectively.*

***F*** *– Loadings plot for the PCA shown in Figure 1E. The further away from the origin of the graph, the stronger it contributes to variability between the samples. Genes are colour-labelled based on their expression pattern. Full red data points correspond to exclusively maternal genes, full blue data points correspond to exclusively zygotic genes. Red circles correspond to predominantly maternal genes, blue circles correspond to predominantly zygotic transcript.*

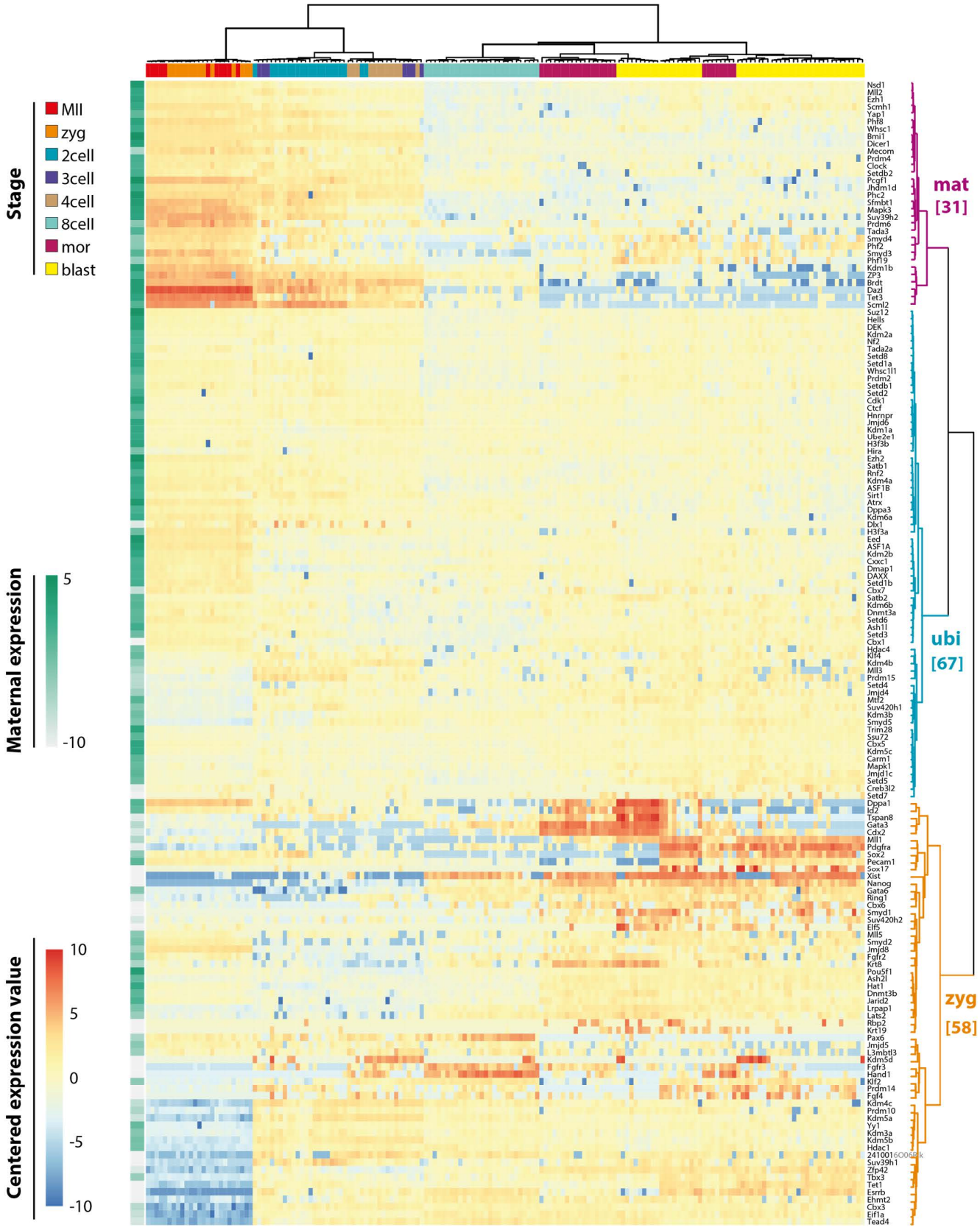

Stage
MII
zyg
2cell
3cell
4cell
8cell
mor
blast
Maternal expression
5
-10
Centered expression value
10
5
0
-5
-10
mat
[31]
ubi
[67]
zyg
[58]

***Figure 4-2.***

### 4.2.2 Expression of chromatin modifying complexes

We observed that most of the paralog genes, like *Tet1* and *Tet3*, display a maternal-zygotic differential expression pattern (Figure 4.3 and Figure 4.4). This suggests that major complexes, such as Trithorax or Polycomb, have maternal-specific and zygotic-specific variants comprised of different homologs of the core genes. Notable examples are the predominantly maternal H3K4 histone methyltransferase *Mll2* versus the zygotic homolog *Mll1* (Figure 4.3A), as well as the maternal H3K9 methyltransferase *Suv39h2* versus the zygotic *Suv39h1* (Figure 4.4A). The core members of Polycomb repressive complex 2 (PRC2), *Ezh2, Eed* and *Suz12*, are ubiquitously expressed throughout preimplantation development. The core members of PRC1, the other major Polycomb complex, show more diversity with *Bmi1* being maternal and Ring1 and *Cbx6* being strongly zygotic. Interestingly, most of the histone demethylases are predominantly zygotically expressed.

We also found this maternal-zygotic dichotomy for a family of poorly studied histone methyltransferases – the SMYD family proteins. Smyd1 is exclusively zygotic and gets expressed from the 8-cell stage on. The expression of *Smyd1* in the 16- and 32-cell stage embryo is non-uniform but does not follow the lineage-specific factors *Id2* and *Sox2*. *Smyd3* and *Smyd4* show the highest expression in MII oocytes and zygotes prior to ZGA. The maternal message for these two genes is rapidly degraded by the end of the 4-cell stage and expression is restarted in the 16-cell stage. A similar dual-expression pattern was observed also for *Smyd2*, however the level of maternal expression was equal to the level reached upon zygotic activation. The fifth paralog, *Smyd5*, displayed a ubiquitous expression pattern.

***Figure 4.2 Expression patterns during preimplantation development***

*The heatmap represents normalized data, centred on a by-gene basis. Thus, the levels of expression for each gene are shown irrespective of the absolute value (in order to visualize the expression trend relative to all genes). The level of absolute maternal expression is indicated with the green bar next to the heatmap. Hierarchical clustering highlighted three main branches, corresponding to "maternal", "ubiquitous" and "zygotic" expression patterns.*

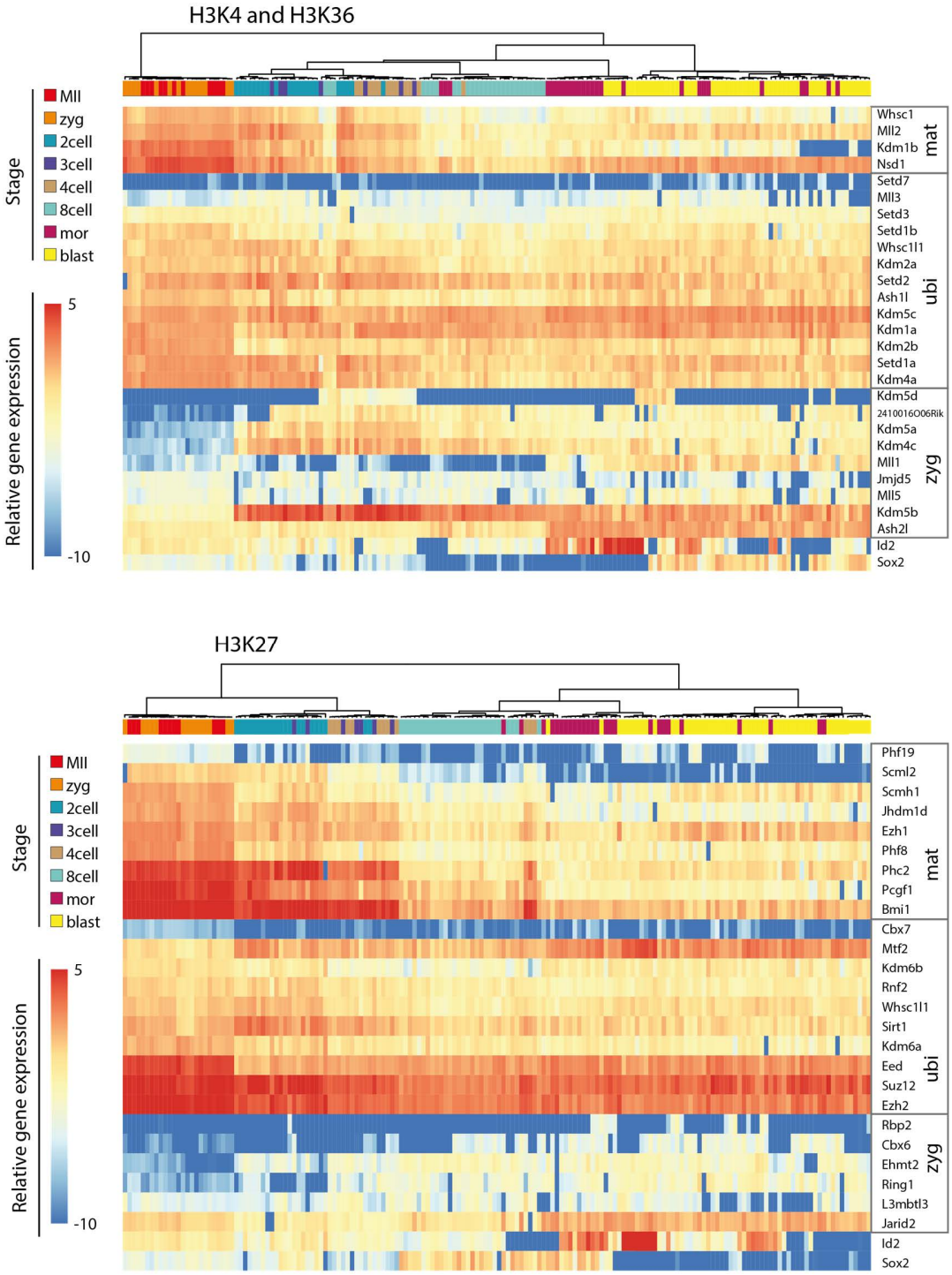



***Figure 4.3 H3K4, H3K36 and H3K27 methylation pathways***

*Heatmaps showing the expression dynamics of genes coding for proteins that are associated with (**A**) H3K4 and H3K36 methylation and (**B**) H3K27 methylation. The representation of the normalized log2 expression values allows comparison of expression intensities between genes. The annotation of the genes as "maternal, "ubiquitous" or "zygotic" corresponds to the clusters shown in Figure 4.2. As a reference, the expression dynamics of the lineage-specific genes Id2 and Sox2 is also depicted.*

Finally, we also looked in detail into the DNA methylation and chromatin remodelling players (Figure 4.4B). We saw the maternal-zygotic division of labour for the pair *Tet1/Tet3* (zygotic and maternal respectively), two homologous enzymes that oxidize methylated cytosine (5-mC) to hydroxymethyl-cytosine (5-hmC). Also *Dnmt3b*, required for *de novo* DNA methylation, was predominantly zygotic, while *Dnmt3a* was predominantly maternal. This observation corresponds to protein levels of the two enzymes detected by immunofluorescence in preimplantation embryos (Hirasawa et al., 2008). Remarkably, all of the screened histone chaperones and chromatin remodelers were ubiquitously expressed from oocyte to blastocyst.

### 4.2.3 Lineage-specific expression

After we identified a significant change from the maternal to the zygotic transcriptome, we asked whether there is also a difference in the expression of chromatin modifiers along with the first cell fate specification events. We looked for genes that show non-uniform distribution of the expression values in the 16-cell and 32-cell embryos (Figure 4.5A). Besides the already described lineage-specific transcription factors and signalling molecules (Guo et al., 2010), we found only a few other genes that show bimodal distribution. These were *Tet1, Kdm1b* and the previously studied *Prdm14* (Burton et al., 2013) We compared the expression of these genes to the strongest lineage markers, *Id2* and *Sox2*, which are expressed in the outer trophectoderm cells and the inner pluripotent cells respectively. *Prdm14* and *Tet1* follow the expression of Sox2 in the inner cell mass quite precisely. The H3K4-specific histone demethylase Kdm1b was expressed in most of the blastomeres, however, the Kdm1b-negative cells were almost exclusively coming from the inner cell mass. There were also a few other genes that showed non-uniform expression patterns in the morula and blastocyst stage, for example *Brdt, Cbx6, Jmjd5, Jmjd8, Mll2, Smyd1, Smyd2*, and *Suv420h2*, however the expression of these genes did not show any correlation with the lineage-specific markers.

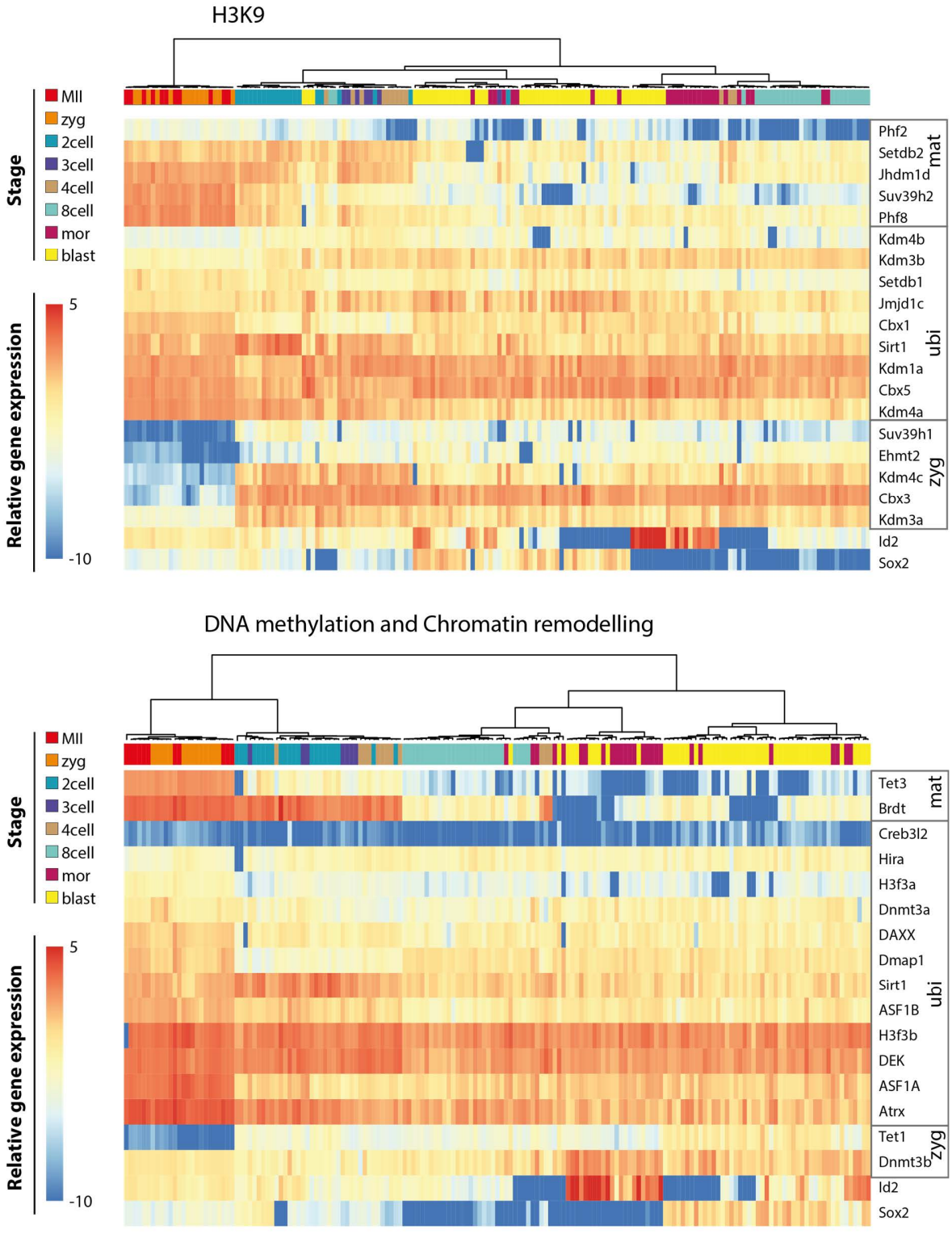



***Figure 4.4 H3K9 methylation, DNA methylation and chromatin remodeler genes***

*Heatmaps showing the expression dynamics of genes coding for proteins that are associated with (A) H3K9 methylation and (B) DNA methylation and chromatin remodelling*

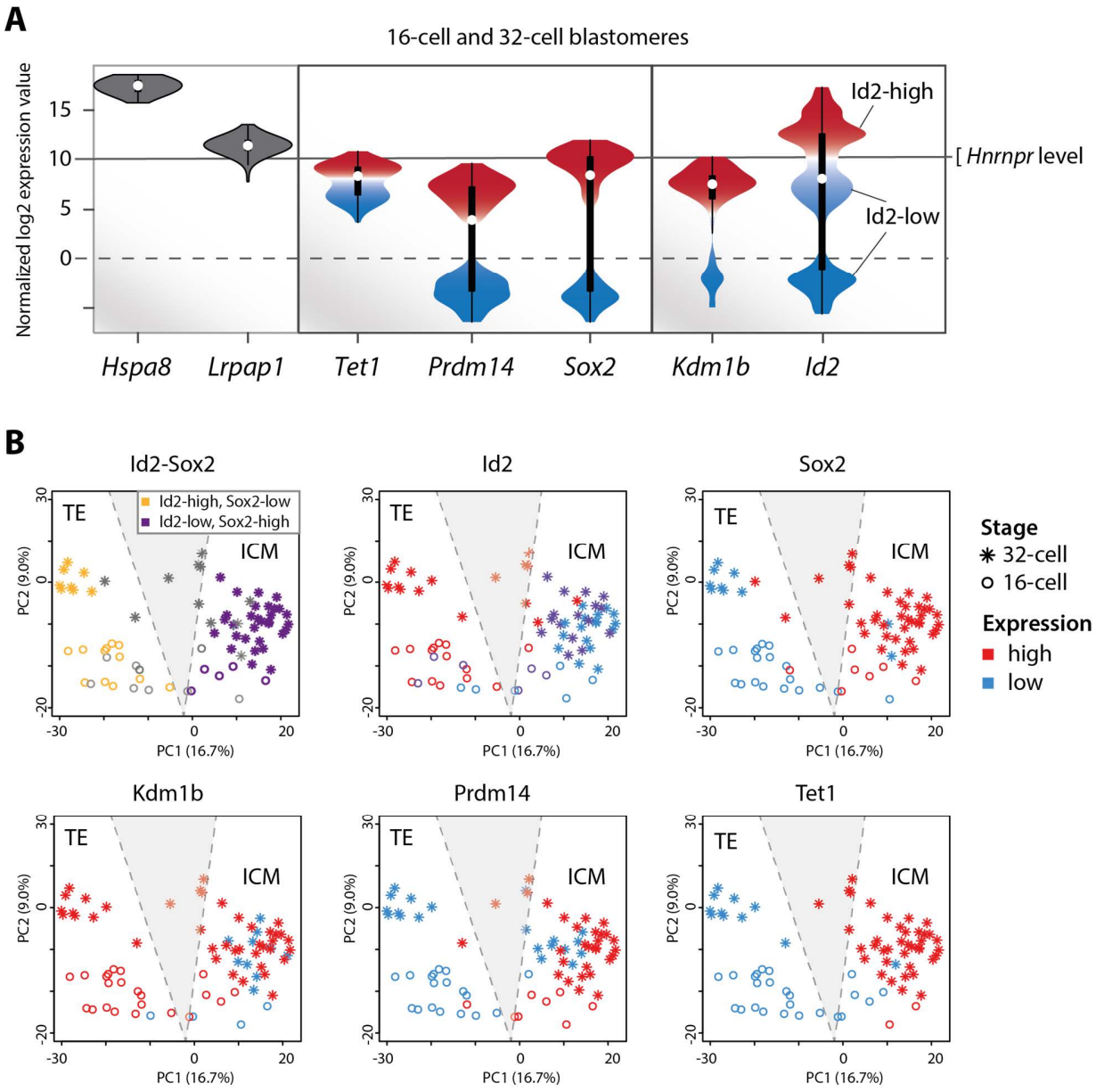



***Figure 4.5 Lineage-specific expression***

***A*** *– Violin plots showing the frequency distribution of the normalized log2 expression signal from 16- and 32-cell blastomeres for the uniformly expressed genes Hspa8 and Lrpap1, the ICM-enriched genes Tet1, Prdm14 and Sox2, as well as the TE-enriched genes Kdm1b and Id2. Red-to-blue gradients mark "high" and "low" expression for a given gene. The level of the endogenous control gene Hnrnpr is shown as a straight line and the detection level is represented as a dotted line.*

*B – PCA plot only for the cells of the 16-cell and 32-cell embryo, coloured based on the relative expression of the genes shown in (A). The shape of the points indicates the stage, while the colour corresponds to high or low expression of a given gene (gene-specific, not corresponding to the global levels or the expression of the other genes). The upper left panel shows the distribution of TE and ICM cells based on the mutually exclusive expression of the two lineage markers Id2 and Sox2. Grey-coloured data points did not display a mutually exclusive pattern of Id2/Sox2 expression.*

### 4.2.4 Sex-specific expression of X-linked genes

The analysed set of genes included some X- and Y-linked genes, in particular the non-coding RNA required for X inactivation in females *Xist* located on the X chromosome and the H3K4 demethylase *Kdm5d* (alias *Smcy*) located on the Y. Expression of *Xist* and *Kdm5d* was observed from the 2-cell stage onwards and allowed us to distinguish between female and male embryos (Figure 4.6A). In addition, there were six other genes on the X-chromosome: *Atrx, Kdm5c, Kdm6a, Phf8, Scml2* and *Suv39h1* (Figure 4.6B). We tested whether female and male embryos differ from each other by performing principle component analysis for female and male cells separately (Figure 4.6C-F). The distribution that we observed did not differ from the global analysis as shown in Figure 4.1. Also, excluding Xist and Kdm5d from the analysis did not change the result, hence they do not contribute critically to the global expression changes that drive the distinction between the different stages in our study.

We then took a closer look into the expression of the six X-linked genes (Figure 4.7). In females, the two X chromosomes are activated during the zygotic genome activation at the 2-cell stage and subsequently the paternal X is inactivated. We observed significant difference between female and male expression for *Atrx, Kdm5c, Kdm6a* and *Suv39h1* in the 2-/3-cell stage blastomeres. The higher expression in females versus males directly relates to the gene dosage, which is due to the presence of two X chromosomes in females versus only one in males. For the exclusively maternally expressed gene *Scml2* we did not detect differential expression between female and male blastomeres, suggesting that the gene is not expressed zygotically. *Phf8* displayed a trend for stronger expression in female cells, although the difference was not statistically significant based on the number of cells analysed. The six genes did not show pronounced sex-specific expression in the blastomeres of the 32-cell embryo.

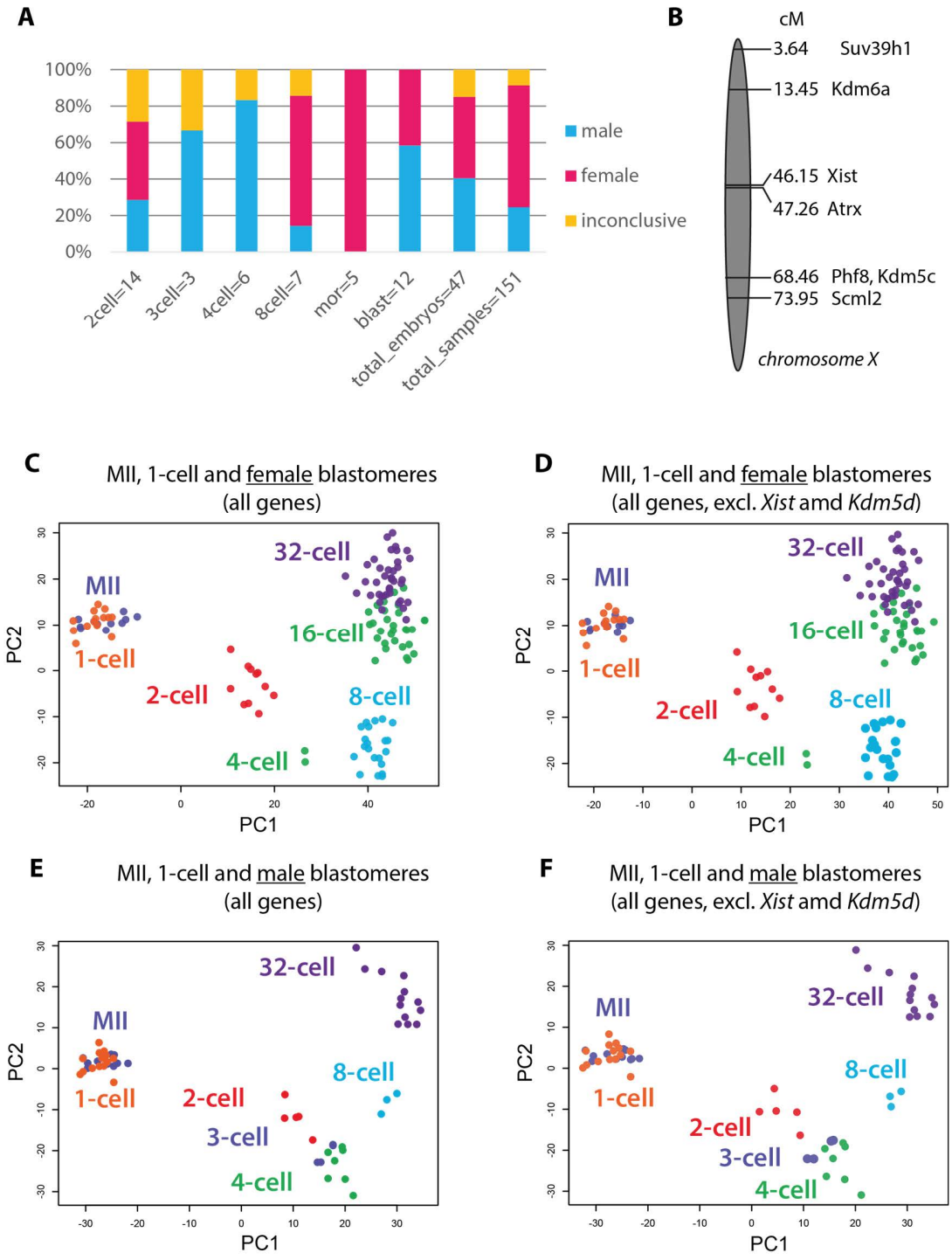



***Figure 4.6 Determination of sex based on Kdm5d and Xist***

***A*** *– Distribution of analysed embryos by sex, as determined by the expression of the Y-linked Kdm5d and the X-linked Xist. The total number of embryos is given below the bars.*

***B*** *– X-linked genes included in the study and their location on the X chromosome.*

***C, D*** *– PCA for female blastomeres only, including (C) and excluding (D) Kdm5d and Xist from the PCA.*

***E, F*** *– PCA for male blastomeres only, including (E) and excluding (F) Kdm5d and Xist from the PCA*

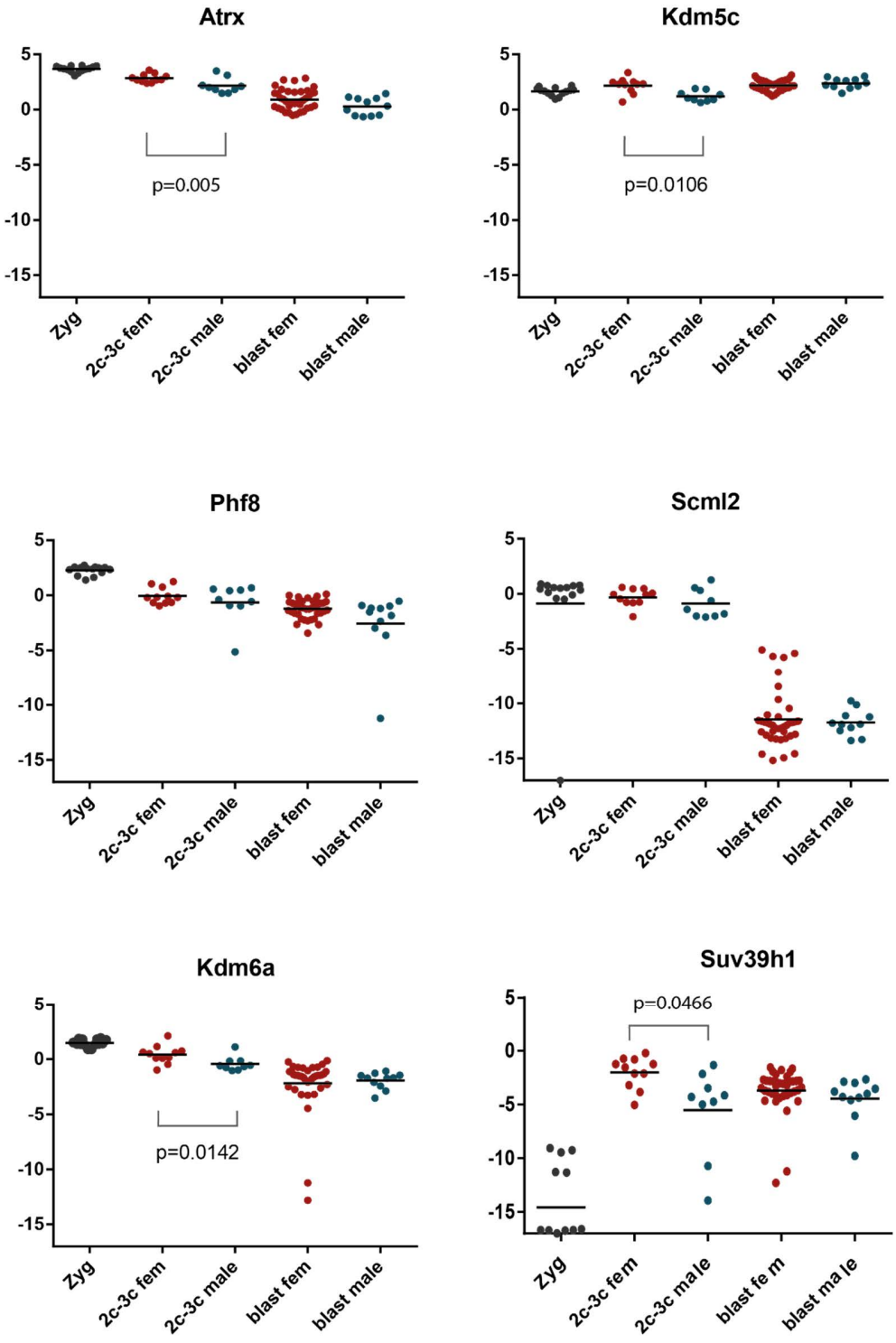



***Figure 4.7 Expression of X-linked genes***

*Scatter plots showing the normalized log2 expression values in zygotes, 2-/3-cell and blastocyst stage female and male blastomeres for Atrx, Kdm5c, Phf8, Scml2, Kdm6a and Suv39h1. P-values indicate statistically significant differences between female and male blastomeres, assessed by a Kolmogorov-Smirnov test.*

## 4.3 Discussion

The specific target amplification, followed by microfluidic qPCR, allowed us to profile the expression of 156 genes in single-cell resolution during the first stages of mouse embryogenesis. Our single-cell analysis reveals the dynamics of expression of a single gene during preimplantation development and in the same time allows comparison of expression intensities between genes. This method was already applied in a similar manner to study the expression of a smaller and partially overlapping set of genes in mouse embryos (Burton et al., 2013; Guo et al., 2010). Our results confirm the previously shown transcriptional dynamics during preimplantation development and at the same time significantly extend the depth of the single-cell profiling by including more than 100 additional genes coding for chromatin modifiers.

We identified many transcripts that showed a maternal- or zygotic-specific expression pattern. However, in every stage there were components of each of the major chromatin modifying complexes and the maternal-zygotic difference comes into place with the expression of maternal- or zygotic-specific gene paralogs. This has been previously shown by conventional qPCR experiments for the major H3K9 methyltransferases *Suv39h1* and *Suv39h2* (Puschendorf et al., 2008) and also for the DNA methyl-cytosine dioxygenases *Tet1* and *Tet3* (Iqbal et al., 2011). We also observed that the zygotic paralog *Tet1* is differentially expressed upon lineage specification and is enriched in the inner cell mass, which is in line with the described function of *Tet1* in maintaining the pluripotent state by preventing methylation of the Nanog promoter (Ito et al., 2010). We provide data for the poorly characterized genes of the SMYD family (associated with H3K4 and H3K36 methylation) and identified *Smyd2, Smyd3, Smyd4* and *Smyd5* as maternally provided, while *Smyd1* is expressed only after ZGA. SMYD genes have been shown to regulate transcription during heart and muscle development but their role in germ cells and early embryos has not been addressed yet (Leinhart and Brown, 2011). Our expression data provides a basis for selecting the appropriate Smyd paralogs in functional studies.

An interesting and not yet fully understood issue in chromatin-based gene regulation is the co-expression of histone methyltransferases and demethylases. The presence of both the "writer" and the "eraser" in the same nucleus suggests a highly dynamic methylation state, maintained by the balance between the two opposing enzymes (Agger et al., 2008). Our study included a total of 24 histone demethylases and 41 histone methyltransferases, the majority of which were co-expressed during preimplantation development. Among the demethylases, we should highlight the H3K4-specific histone demethylase Kdm1b, which is strongly maternal but remains expressed also after ZGA and eventually becomes excluded from many of the ICM cells. The non-ubiquitous expression pattern and correlation with Id2 expression, makes Kdm1b a good candidate for lineage-specification studies.

There were a number of genes that exhibited variable expression in morula and blastocyst stages, which did not correspond to the cell lineage. This variation could be technical, but it could also be due to rapid fluctuations of the system and the fact that gene expression profiling experiments take snapshots of development, therefore making it difficult to describe highly fluctuating systems with a limited number of observations. Indeed, other single-cell studies have described high stochastic variability between the blastomeres of preimplantation embryos (Ohnishi et al., 2014). Another possibility might be post-transcriptional and post-translational regulation, which could ensure lineage-specific activity for a given protein even if the mRNA is ubiquitously expressed.

In conclusion, we performed a comprehensive analysis of gene expression of chromatin modifying genes during preimplantation and observed significant changes along the maternal-zygotic axis but not during lineage specification for a number of paralogs from different complexes.

## 4.4 Materials and methods

### 4.4.1 Oocyte and embryo collection, *in vitro* culture and single cell isolation

F1 female mice (C57BL/6 x DBA/2) were superovulated and in the case of embryo collection, mated to C57BL6 males. MII oocytes were collected 14 hours post hCG injection in M2 medium (Sigma, M7167). Zygotes were collected 14-16 hours post hCG collection in M2 medium. Cumulus cells were removed by incubation for 5 min in 1 mg/ml Hyaluronidase-containing M2. Embryos were cultured in KSOMaa (Milipore, MR-106-D) in a low oxygen air chamber. Single blastomeres were isolated by first removing the *Zona pellucida* with Proteinase (Sigma, P5147-1G, 5 mg/ml in M2 medium), followed by separating the cells with Trypsine (Trypsine-EDTA, T4049 Sigma) Single cells were picked with the help of a mouth pipette and a finely pulled glass capillary.

### 4.4.2 Preparation of pre-amplified single-cell cDNA

Specific target amplification was performed by pipetting single cells directly into 0,1 ml nuclease-free PCR tubes containing the CellsDirect™ One-Step qRT-PCR Kit (Invitrogen, 11753-100) reaction mix. Each reaction amounted 4.5 µl and contained 1.25 µl primers mix (192 forward and 192 reverse primers mixed at a final concentration of 500 nM each), 2.5 µl 2x CellsDirect master mix, 0.1 µl CellsDirect Enzyme mix, 0.1 µl Superase-In (Ambion, AM2696) and 0.55 µl DNA Resuspension Buffer (TEKnova, T0221). Reverse transcription was performed by incubation in a standard thermal cycler for 15 minutes at 50°C, followed by 2 minutes at 95°C and subsequently 14 cycles of PCR amplification (15 seconds at 95°C and 4 minutes at 60°C).

Residual primers were removed by adding 2 µl Exonuclease I master mix to the pre-amplified reactions. The Exonuclease I mastermix consisted of 0.4 µl Exonuclease I (New England BioLabs, M0293S), 0.2 µl 10x Reaction buffer (provided with enzyme) and 1.4 µl nuclease-free water. The samples were incubated for 30 min at 37°C, followed by 15 min at 80°C. The pre-amplified reaction products were diluted 5-fold and stored at -20°C.

### 4.4.3 Single-cell qPCR with BioMark 96.96 Dynamic Arrays

Sample Pre-Mix solutions were prepared in two 96-well plates for the 192 samples, using 2.7 µl of the pre-amplified cDNA, mixed with 3.0 µl 2X Sso Fast EvaGreen Supermix WithLow ROX (Bio-Rad Laboratories, PN 172-5211) and with 0.3 µl 20X DNA Binding Dye Sample Loading Reagent (Fluidigm, PN 100-0388). The assay mix for the 192 assays was prepared by mixing 3.0 µl 2X Assay Loading Reagent (Fluidigm, PN 85000736) with 2.7 µl 1X DNA Suspension Buffer (Teknova, PN T0221) and 0.3 µl 100 µM each of Forward and Reverse Primer Mix. Four 96.96 Dynamic Array chips were loaded and run on a BioMark system as described by the manufacturer (Protocol ADP 33, Fluidigm).

### 4.4.4 Data analysis and visualization

Data from the BioMark qPCR system was initially processed with Fluidigm Real-Time PCR Analysis Software, including quality control of the experiment and identification of unspecific products based on the product melting temperature. The Ct values obtained from the BioMark System were converted into relative expression levels by subtracting the values from the assumed baseline value of 30 (inverted Ct values). Cells with low or absent endogenous control gene expression levels were removed from analysis. Also genes with no specific signal were removed from the analysis. The resulting values (156 genes and 189 samples) were normalized to the average signal for the endogenous reference genes *Ssu72, Nnrnpr* and *Ube2e1*, by subtracting the average inverted Ct value for the three reference genes from the respective expression value.

The normalized expression values were analysed and visualized in R. The R scripts are provided in Appendix A: R scripts.

## 4.5 Additional information

### 4.5.1 Data availability

Upon acceptance of the manuscript for publication, the raw and normalized data, as well as the assay information will be published in the online database GEO at www.ncbi.nlm.nih.gov/geo/

Until then, the raw data is available upon request: peter.nestorov@fmi.ch

### 4.5.2 Author contributions

P. N. and A.H.F.M.P. conceived and designed the experiments. P. N. performed all experiments. P. N. analysed the data with the support of H.R.H. P. N. and A.H.F.M.P. wrote the manuscript.

### 4.5.3 Acknowledgements

P. N. greatly acknowledges his Boehringer Ingelheim Fonds PhD fellowship. Research in the Peters group is supported by the Novartis Research Foundation, the Swiss National Science Foundation (NRP 63—Stem Cells and Regenerative Medicine), SystemsX.ch (Cell plasticity), the Japanese Swiss Science and Technology Cooperation Program, and the EMBO Young Investigator Program.

### *Table 4.1 Gene expression of 156 genes in mouse pre-implantation embryos*

*The average gene expression value for each of the 156 assayed genes is calculated from the single-cell data for four major stages of preimplantation development: "mat" – MII oocytes; "zyg 2-4" – 2-cell to 4-cell embryos; "zyg 8-mor" – 8-cell to 16-cell stage embryos; "zyg blast" – blastocyst embryos. The values are normalized as described in the Materials and methods section. For easier visualization of expression trends, the table includes heatmap-colours, ranging from red-yellow-green (low-mid-high expression). The extended gene name gives information on the molecular function of a given gene. It first shows the official gene name, followed by an alternative gene name (where applicable), the main molecular activity (CR – chromatin remodelling; DNA – DNA binding; DNAme – DNA methylation; HDM – histone demethylation; HR – histone binding; HMT – histone methylation; HK – housekeeping; SIG – signalling), important functional domains (SET – SET histone methyltransferase domain; ZnF – Zn-finger domain; ATPase; Helicase; Kinase; CD – chromodomain; JmjC/JmjD - Jumonji C/D domains) and finally the associated histone modification.*

| Gene name | Extended Gene Name | mat | zyg 2-4 | zyg 8-mor | zyg blast |
|---|---|---|---|---|---|
| *2410016O06Rik* | 2410016O06Rik_NO66_HDM_JmjC_H3K4_H3K36 | -13.4094 | -2.50217 | -2.45851 | -2.7295 |
| *Asf1a* | ASF1A_CR | 3.184723 | -0.10574 | 0.654162 | 0.325795 |
| *Asf1b* | ASF1B_CR | 1.327016 | 0.246192 | -0.99898 | -1.15142 |
| *Ash1l* | Ash1l_HMT_SET_ZnF_H3K4_H3K36 | 0.842266 | -0.40972 | -0.79388 | 0.305304 |
| *Ash2l* | Ash2l_HMT_H3K4_H3K36 | -0.96062 | -0.93675 | 0.765861 | 1.617523 |
| *Atrx* | Atrx_CR_ATPase_Helicase_ZnF | 3.823325 | 2.148547 | 1.489486 | 0.695309 |
| *Bmi1* | Bmi1_HR_ZnF_PRC1_H3K27_H2AK119ub | 5.731631 | 5.016583 | 1.351498 | 0.691279 |
| *Brdt* | Brdt_CR | 3.42211 | 2.875662 | -3.69562 | -5.85643 |
| *Carm1* | Carm1_HMT_H3R17 | -0.65447 | 0.095537 | 0.375259 | 0.807442 |
| *Cbx1* | Cbx1_HR_CD_HP1b_H3K9 | 0.82327 | -1.92036 | -1.0629 | -0.00861 |
| *Cbx3* | Cbx3_HR_CD_HP1g_H3K9 | -6.3374 | 2.118737 | 1.951426 | 1.685447 |
| *Cbx5* | Cbx5_HR_CD_HP1a_H3K9 | 2.003803 | 1.136973 | 1.559681 | 2.103993 |
| *Cbx6* | Cbx6_HR_CD_PRC1_H3K27_H2AK119ub | -12.2458 | -11.9452 | -7.15761 | -6.37023 |
| *Cbx7* | Cbx7_HR_CD_PRC1_H3K27_H2AK119ub | -7.13002 | -12.3292 | -10.9724 | -10.6607 |
| *Cdk1* | Cdk1_SIG_Kinase | 3.196533 | 3.696929 | 3.511209 | 2.751352 |
| *Cdx2* | Cdx2_DNA | -8.44561 | -9.95532 | -3.61809 | -4.80225 |
| *Clock* | Clock_SIG | 0.322252 | -0.36674 | -3.13904 | -2.26761 |
| *Creb3l2* | Creb3l2_DNA | -9.16731 | -10.9473 | -11.9606 | -9.31158 |
| *Ctcf* | Ctcf_DNA_ZnF | 0.578697 | 1.132873 | 0.153328 | -0.17711 |
| *Cxxc1* | Cxxc1_DNA_ZnF | 0.20479 | -1.75712 | -1.84213 | -1.44336 |
| *Daxx* | Daxx_CR | 0.478316 | -1.4017 | -1.58786 | -1.25488 |
| *Dazl* | Dazl | 3.390408 | -1.26909 | -11.6988 | -10.6592 |
| *Dek* | Dek_CR | 2.93895 | 2.791206 | 1.667013 | 1.898964 |
| *Dicer1* | Dicer1 | 3.276403 | 1.762607 | -0.20912 | -0.74791 |
| *Dlx1* | Dlx1_DNA | -8.65415 | -11.4541 | -14.5756 | -12.811 |
| *Dmap1* | Dmap1_DNAme | 0.611418 | -2.60627 | -0.12804 | -0.72885 |
| *Dnmt3a* | Dnmt3a_DNAme_ZnF | -0.80076 | -2.72109 | -1.96251 | -1.7496 |
| *Dnmt3b* | Dnmt3b_DNAme_ZnF | -0.57428 | -1.50883 | 0.167656 | 0.797252 |
| *Dppa1* | Dppa1_DNA | -0.17321 | -4.76357 | -7.52598 | -4.66574 |
| *Dppa3* | Dppa3_DNA | 0.410565 | -0.64585 | -0.67596 | -2.01308 |
| *Eed* | Eed_HR_WD_PRC2_H3K27 | 4.25732 | 1.487241 | 2.260945 | 2.065337 |

| Gene name | Extended Gene Name | mat | zyg 2-4 | zyg 8-mor | zyg blast |
|---|---|---|---|---|---|
| *Ehmt2* | Ehmt2_G9a_HMT_SET_H3K9_H3K27 | -10.1156 | -4.13232 | -2.53649 | -3.12362 |
| *Eif1a* | Eif1a_DNA | -7.29717 | 0.349826 | 1.67899 | 1.485142 |
| *Elf5* | Elf5_DNA | -11.1227 | -14.1821 | -13.7214 | -8.12137 |
| *Esrrb* | Esrrb_SIG | -14.6121 | -1.12832 | 0.151008 | -0.52904 |
| *Ezh1* | Ezh1_HMT_SET_PRC2_H3K27 | 2.321562 | 1.210443 | -0.45705 | 0.494924 |
| *Ezh2* | Ezh2_HMT_SET_PRC2_H3K27 | 5.04822 | 3.752006 | 2.817665 | 2.536575 |
| *Fgf4* | Fgf4_SIG | -16.7853 | -9.15454 | -9.1962 | -7.64233 |
| *Fgfr2* | Fgfr2_SIG | -2.11295 | -5.86655 | -1.82019 | -2.32805 |
| *Fgfr3* | Fgfr3_SIG | -15.2996 | -7.36138 | -2.34303 | -8.36813 |
| *Gata3* | Gata3_DNA | -6.32575 | -12.3488 | -1.07227 | -5.11534 |
| *Gata6* | Gata6_DNA | -3.11395 | -5.18659 | 1.993592 | 1.625071 |
| *H3f3a* | H3f3a_CR | -1.64697 | -3.72997 | -4.96681 | -5.02955 |
| *H3f3b* | H3f3b_CR | 2.414977 | 2.651911 | 2.261671 | 2.37512 |
| *Hand1* | Hand1_DNA | -17.0705 | -10.6884 | -3.63798 | -9.61792 |
| *Hat1* | Hat1_H4K5ac_H4K12ac | 1.662405 | -0.20625 | 1.901374 | 2.758387 |
| *Hdac1* | Hdac1 | -2.30682 | 2.740473 | 2.710568 | 1.632739 |
| *Hdac4* | Hdac4 | -2.34435 | -3.34272 | -4.16456 | -2.87521 |
| *Hells* | Hells_Helicase | 2.851723 | 2.344366 | 1.757719 | 2.232458 |
| *Hira* | Hira_CR_WD | -2.50004 | -2.06257 | -1.78945 | -1.72539 |
| *Hnrnpr* | Hnrnpr_HK | -1.01931 | -0.2772 | -1.03585 | -1.36646 |
| *Id2* | Id2_DNA | -0.88037 | -2.99215 | -3.41146 | -3.05887 |
| *Jarid2* | Jarid2_HMT_JmjC_JmjN_PRC2_H3K27 | 0.25384 | -2.36009 | 0.522055 | 1.142791 |
| *Jhdm1d* | Jhdm1d_HDM_JmjC_ZnF_H3K27me2 _H3K9me2_H4K20me1 | 1.694712 | 0.943542 | -1.99713 | -3.53283 |
| *Jmjd1c* | Jmjd1c_HDM_JmjC_H3K9 | -0.41384 | 0.108787 | -0.46503 | 0.898349 |
| *Jmjd4* | Jmjd4_HDM_JmjC | -4.84636 | -2.49976 | -3.37208 | -2.26003 |
| *Jmjd5* | Jmjd5_HDM_JmjC_H3K36 | -3.54162 | -4.22156 | -4.69392 | -6.34773 |
| *Jmjd6* | Jmjd6_HDM_JmjC_H3R2_H4R3 | 2.085991 | 1.936403 | 1.132498 | 0.658881 |
| *Jmjd8* | Jmjd8_HDM_JmjC | -4.59889 | -12.5916 | -9.19656 | -7.69892 |
| *Kdm1a* | Kdm1a_Lsd1_HDM_H3K4_H3K9 | 1.798256 | 1.487365 | 1.616224 | 1.587258 |
| *Kdm1b* | Kdm1b_Lsd2_HDM_H3K4 | 3.084494 | 0.807089 | -2.49918 | -4.19756 |
| *Kdm2a* | Kdm2a_Jhdm1a_HDM_JmjC_ZnF_H3K36 | 0.618794 | 0.723149 | -0.62057 | 0.163353 |
| *Kdm2b* | Kdm2b_Jhdm1b_HDM_JmjC_ZnF_H3K4_H3K36 | 1.844747 | -0.38424 | 0.376727 | -0.40302 |
| *Kdm3a* | Kdm3a_ Jmjd1a_HDM_JmjC_H3K9 | -3.14165 | 0.892714 | 0.065521 | -0.49321 |
| *Kdm3b* | Kdm3b_Jmjd1b_HDM_H3K9 | -1.87259 | -0.58156 | 0.243299 | 0.228854 |
| *Kdm4a* | Kdm4a_Jmjd2a_HDM_JmjC_JmjN_ZnF_H3K9_H3K36 | 2.118296 | 1.404914 | -0.31082 | 0.017911 |
| *Kdm4b* | Kdm4b_Jmjd2b_HDM_JmjC_JmjN_ZnF_H3K9 | -3.96466 | -1.67929 | -3.77341 | -3.44015 |
| *Kdm4c* | Kdm4c_Jmjd2c_HDM_JmjC_JmjN_ZnF_H3K9_H3K36 | -6.93659 | 1.364509 | -0.27023 | -1.7332 |
| *Kdm5a* | Kdm5a_Rbp2_HDM_JmjC_JmjN_H3K4 | -7.64001 | 0.265061 | -1.56304 | -2.18492 |
| *Kdm5b* | Kdm5b_Jarid1b_HDM_JmjC_JmjN_ZnF_H3K4 | -2.13679 | 3.941703 | 2.265193 | 1.34355 |
| *Kdm5c* | Kdm5c_Jarid1c_HDM_JmjC_JmjN_ZnF_H3K4 | 1.771399 | 1.604586 | 1.89577 | 2.089474 |
| *Kdm5d* | Kdm5d_Jarid1d_HDM_JmjC_JmjN_ZnF_H3K4 | -16.8827 | -8.2905 | -14.2066 | -9.3683 |
| *Kdm6a* | Kdm6a_Utx_HDM_JmjC_H3K27 | 1.329679 | -0.16403 | -0.92473 | -2.12543 |

| Gene name | Extended Gene Name | mat | zyg 2-4 | zyg 8-mor | zyg blast |
|---|---|---|---|---|---|
| *Kdm6b* | Kdm6b_Jmjd3_HDM_JmjC_H3K27 | -0.19485 | -1.98463 | -2.04582 | -1.13218 |
| *Klf2* | Klf2_DNA_ZnF | -3.70931 | -0.90034 | -0.6673 | -0.39557 |
| *Klf4* | Klf4_DNA_ZnF | -1.21477 | -0.236 | -1.6295 | -0.68501 |
| *Krt19* | Krt19_SIG | -17.14 | -13.6065 | -12.4539 | -9.2062 |
| *Krt8* | Krt8_SIG | -5.38901 | -8.47909 | -3.08607 | -2.69988 |
| *L3mbtl3* | L3mbtl2_DNA_PRC1_H3K27_H2AK119ub | -4.4726 | -4.04921 | -4.92704 | -6.77519 |
| *Lats2* | Lats2_SIG_Kinase | -4.38549 | -5.42993 | -0.99976 | -1.09103 |
| *Lrpap1* | Lrpap1 | -2.56817 | -2.53127 | 0.940501 | 1.535071 |
| *Mapk1* | Mapk1_SIG_Kinase | -1.27655 | -1.24642 | -0.83888 | 0.290691 |
| *Mapk3* | Mapk3_SIG_Kinase | 2.707044 | -1.05785 | -3.44206 | -2.86114 |
| *Mecom* | Mecom_Prdm3_DNA | -5.60262 | -8.77918 | -13.9366 | -12.6496 |
| *Mll1* | Mll1_HMT_SET_H3K4 | -5.86074 | -9.23017 | -8.42683 | 0.191351 |
| *Mll2* | Mll2_HMT_SET_ZnF_H3K4 | 1.792448 | 1.555866 | -1.32055 | 0.299433 |
| *Mll3* | Mll3_HMT_SET_H3K4 | -5.36499 | -3.65861 | -5.20189 | -5.39128 |
| *Mll5* | Mll5_HMT_SET_ZnF_H3K4 | -3.42115 | -5.28908 | -4.53614 | -3.11187 |
| *Mtf2* | Mtf2_Pcl2_DNA_PRC2_ZnF_H3K27 | -0.32234 | 1.440708 | 1.425863 | 2.678584 |
| *Nanog* | Nanog_DNA | -16.683 | -7.10916 | -0.49084 | 0.696374 |
| *Nf2* | Nf2_SIG | 0.773878 | 2.034588 | 0.359585 | 1.116333 |
| *Nsd1* | Nsd1_HMT_H3K36_H4K20 | 3.878501 | 1.596254 | 0.373945 | 1.876201 |
| *Pax6* | Pax6_DNA | -5.58061 | -6.27963 | -7.9443 | -11.6167 |
| *Pcgf1* | Pcgf1_DNA_PRC1_ZnF_H3k27_H2AK119ub | 4.935701 | 2.414305 | -0.48226 | -2.24645 |
| *Pdgfra* | Pdgfra_SIG_Kinase | -9.36993 | -9.22028 | -4.72609 | -0.60495 |
| *Pecam1* | Pecam1_SIG | -0.42276 | -2.03491 | -3.7121 | -2.18778 |
| *Phc2* | Phc2_DNA_PRC1_ZnF_H3k27_H2AK119ub | 4.387853 | 3.370416 | -0.82601 | -0.26545 |
| *Phf19* | Phf19_Pcl3_DNA_PRC2_ZnF_H3K27 | -4.21629 | -9.54868 | -11.8185 | -8.16537 |
| *Phf2* | Phf2_HDM_JmjC_ZnF_H3K9 | -3.62019 | -6.85804 | -10.8258 | -5.22927 |
| *Phf8* | Phf8_HDM_JmjC_ZnF_H3K9_H3K27_H4K20 | 2.337232 | -0.22059 | -2.00575 | -1.5518 |
| *Pou5f1* | Pou5f1_Oct4_DNA | 3.445579 | 1.828929 | 4.042925 | 5.252264 |
| *Prdm10* | Prdm10_HMT | -5.76608 | -1.50556 | -1.92669 | -2.13859 |
| *Prdm14* | Prdm14_HMT_SET | -11.9425 | -6.73696 | -9.10593 | -6.47792 |
| *Prdm15* | Prdm15_HMT_SET | -6.47177 | -2.65539 | -5.16271 | -4.69145 |
| *Prdm2* | Prdm2_HMT_SET_ZnF | -0.94132 | -1.57469 | -1.97591 | -1.15349 |
| *Prdm4* | Prdm4_HMT_SET_ZnF | 0.609556 | -0.20207 | -2.55418 | -2.14799 |
| *Prdm6* | Prdm6_HMT_SET_ZnF | -3.33307 | -10.131 | -12.3459 | -11.323 |
| *Rbp2* | Rbp2_CR_PRC2_H3K27 | -16.7845 | -13.95 | -13.0447 | -10.2137 |
| *Ring1* | Ring1_CR_PRC1_ZnF_H3K27_H2AK119ub | -7.75459 | -7.03156 | -3.70647 | -2.4736 |
| *Rnf2* | Rnf2_CR_PRC1_ZnF_H3K27_H2AK119ub | 0.248717 | -0.18035 | -1.40552 | -1.20396 |
| *Satb1* | Satb1_DNA | 2.632253 | 1.833063 | 1.122044 | 1.451481 |
| *Satb2* | Satb2_DNA | -0.60722 | -1.65927 | -0.8039 | -0.07781 |
| *Scmh1* | Scmh1_CR_PRC1_ZnF_H3K27_H2AK119ub | 1.666023 | -0.79143 | -3.0938 | -0.73734 |
| *Scml2* | Scml2_CR_PRC1_ZnF_H3K27_H2AK119ub | -0.23086 | -1.22323 | -9.25151 | -11.5961 |
| *Setd1a* | Setd1a_HMT_SET_H3K4 | 1.934751 | 1.491483 | 0.144627 | 0.944842 |

| Gene name | Extended Gene Name | mat | zyg 2-4 | zyg 8-mor | zyg blast |
|---|---|---|---|---|---|
| *Setd1b* | Setd1b_HMT_SET_H3K4 | 0.543327 | -1.70634 | -1.72742 | -1.04772 |
| *Setd2* | Setd2_HMT_SET_H3K36 | 0.748591 | 1.893819 | 0.656534 | 0.839477 |
| *Setd3* | Setd3_HMT_SET_H3K36 | -1.56349 | -2.94212 | -3.28987 | -1.916 |
| *Setd4* | Setd4_HMT_SET | -6.28309 | -6.06375 | -4.62419 | -4.54503 |
| *Setd5* | Setd5_HMT_SET | -1.59689 | -1.41303 | -1.90621 | 0.546304 |
| *Setd6* | Setd6_HMT_SET | -0.50649 | -2.69964 | -2.99788 | -1.62489 |
| *Setd7* | Setd7_HMT_SET_H3K4 | -14.7902 | -12.4576 | -12.7372 | -10.267 |
| *Setd8* | Setd7_HMT_SET_H4K20 | 1.972844 | 1.916972 | 1.671982 | 2.258002 |
| *Setdb1* | Setdb1_Eset_HMT_SET_H3K9 | -0.46232 | -1.51639 | -2.97801 | -1.16945 |
| *Setdb2* | Setdb2_HMT_SET_H3K9 | 0.555912 | 0.76115 | -2.42807 | -1.80959 |
| *Sfmbt1* | Sfmbt1_HR_PRC_PhoRC | 3.536804 | 1.396842 | -2.93289 | -2.17584 |
| *Sirt1* | Sirt1_CR_PRC2_H3K27_H3K9 | 1.156116 | 2.021471 | 0.290388 | -0.30626 |
| *Smyd1* | Smyd1_HMT_SET_ZnF | -17.0705 | -13.9822 | -11.3972 | -6.36664 |
| *Smyd2* | Smyd2_HMT_SET_ZnF | -2.31191 | -9.00001 | -3.4256 | -2.49683 |
| *Smyd3* | Smyd3_HMT_SET_ZnF | -0.31127 | -5.47556 | -7.2315 | -4.11376 |
| *Smyd4* | Smyd4_HMT_SET_ZnF | -3.76172 | -9.17852 | -10.7475 | -6.73306 |
| *Smyd5* | Smyd5_HMT_SET_ZnF | -2.88044 | 0.267015 | 0.17901 | 1.097167 |
| *Sox17* | Sox17_DNA | -15.4521 | -13.5561 | -14.2688 | -7.04212 |
| *Sox2* | Sox2_DNA | -3.78329 | -5.00264 | -11.3045 | -2.60532 |
| *Ssu72* | Ssu72_HK | -0.59003 | -0.71286 | -0.20162 | 0.51319 |
| *Suv39h1* | Suv39h1_HMT_SET_H3K9 | -13.7387 | -3.47212 | -5.33536 | -3.88232 |
| *Suv39h2* | Suv39h2_HMT_SET_H3K9 | 2.136753 | -1.03465 | -5.27058 | -4.7308 |
| *Suv420h1* | Suv420h1_HMT_SET_H4K20 | -3.12268 | -0.70604 | -1.697 | -0.46718 |
| *Suv420h2* | Suv420h2_HMT_SET_H4K20 | -8.20156 | -9.19213 | -8.0665 | -5.60269 |
| *Suz12* | Suz12_HMT_PRC2_ZnF_H3K27 | 4.777871 | 4.257572 | 2.93068 | 3.523425 |
| *Tada2a* | Tada2a_DNA | 0.273034 | -0.68277 | -0.9685 | -0.34487 |
| *Tada3* | Tada3_DNA | -0.07551 | -2.81372 | -4.13196 | -5.79999 |
| *Tbx3* | Tbx3_DNA | -4.39342 | -0.91646 | 0.075945 | 0.937077 |
| *Tead4* | Tead4_DNA | -8.18299 | -0.54885 | 0.21673 | -0.59029 |
| *Tet1* | Tet1_DNAme | -10.5192 | -3.30493 | -3.59672 | -1.47902 |
| *Tet3* | Tet3_DNAme | 2.149591 | -2.40365 | -7.56028 | -8.18733 |
| *Trim28* | Trim28_Tif1b_DNA | 2.508978 | 2.901401 | 2.81115 | 3.756114 |
| *Tspan8* | Tspan8_SIG | -8.81299 | -7.29318 | -7.72968 | -6.71351 |
| *Ube2e1* | Ube2e1_HK | 1.609339 | 0.990055 | 1.237467 | 0.853271 |
| *Whsc1* | Whsc1_Nsd2_HMT_SET_ZnF_H3K4_H3K36_H4K20 | 1.23431 | 0.177135 | -2.96876 | -1.83563 |
| *Whsc1l1* | Whsc1l1_Nsd3_HMT_SET_ZnF_H3K4_H3K36_H3K27 | 0.552129 | 0.349473 | -0.98251 | -0.16107 |
| *Xist* | Xist | -15.164 | -10.091 | 1.735378 | 0.102833 |
| *Yap1* | Yap1_SIG | -0.708 | -0.96371 | -4.66745 | -3.15547 |
| *Yy1* | Yy1_DNA_PRC_ZnF_PhoRC | -1.27835 | 3.199492 | 2.894475 | 3.01209 |
| *Zfp42* | Zfp42_DNA_ZnF | -8.53607 | -3.8965 | -0.89094 | 0.35402 |
| *ZP3* | Zp3_SIG | 0.694934 | -0.9937 | -3.99109 | -9.01472 |

# Chapter 5. Manuscript in preparation: PRC2 is required for maintaining a repressive chromatin state at the onset of life

Peter Nestorov*,†, Zichuan Liu*, Julia Hacker*,†, Lukas Burger*, Antoine H.F.M. Peters*,†,2

*Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland †Faculty of Sciences, University of Basel.

2Corresponding author: e-mail address: antoine.peters@fmi.ch

## 5.1 Introduction

Multicellular organisms begin their life as a single cell that arises from the fusion of two gametes, which carry all the material and information required for the onset of life. The first events of embryonic development include the reorganization of the paternal genome by maternal factors and the switch between the maternal and the zygotic transcriptional programmes, which in the mouse happens in the first 48 hours. This short developmental window is associated with major changes in nuclear organization, chromatin modifications and DNA methylation. In the zygote stage, the paternal and maternal genomes are physically separated and display strong differences in DNA and histone methylation that partially hold up to the 8-cell stage (Albert and Peters, 2009). The epigenetic asymmetry in early embryos together with the observation of gene-specific histone retention in sperm, have been hypothesized to be associated with an inheritance mechanism acting in early embryos through intergenerational perpetuation of chromatin states (Brykczynska et al., 2010; Hammoud et al., 2009). The potential transmission of epigenetic information from the gametes to the embryo via H3K27me3 has been suggested both in worms and mouse (Gaydos et al., 2014; Puschendorf et al., 2008).Yet, it is still unclear how chromatin-based, epigenetic inheritance would function in mammalian development and what the relevance of the dynamic chromatin changes in the zygote might be. The evolutionary conserved Polycomb repressive complex 2 (PRC2) is a key

player involved in cellular memory that confers the repressive H3K27me3 histone mark and is essential for development in various organisms from plants to human (Nestorov et al., 2013a). PRC2 is required during gastrulation and for the establishment of the extra-embryonic tissue in mouse embryos (Erhardt et al., 2003; Faust et al., 1998; O'Carroll et al., 2001; Saha et al., 2013). However, the experimental approach in these studies does not properly deal with the maternal component of PRC2, which may be particularly important for the observed H3K27me3 dynamics in the zygote and the major transcriptional changes during ZGA and the subsequent cell fate specification. Moreover, PRC1, the other major Polycomb complex, is essential for regulating the accumulation of maternal RNA and this opens the question of how PRC2 is contributing to this process (Posfai et al., 2012). Here, we addressed the role of PRC2 in early embryos by removing both the maternal and the zygotic component of PRC2.

## 5.2 Results

PRC2 is highly abundant in mouse oocytes and the core genes *Ezh2*, *Suz12* and *Eed* show strong and increasing expression levels during oogenesis (Figure S 5-2). We used a conditional knock-out approach to delete *Eed* in primary and growing oocytes (*Eed* $^{fl/fl}$ *Gdf9-iCre* and *Eed* $^{fl/fl}$ *Zp3-cre* respectively), as well as in late spermatogenesis (*Eed* $^{fl/fl}$ *Prm1-cre*) Figure S 5-2C, D). Oocytes from *Eed* $^{fl/fl}$ *Gdf9-iCre* females were fertilized by intracytoplasmic sperm injection (ICSI) with *Eed* knock-out sperm (referred to as *Eed* m-z- embryos). After four days in culture, we observed abnormal development in 60% of the *Eed* m-z- embryos versus only 20% in the control *Eed* m+z+ embryos (Figure 5.1A, B). The incomplete penetrance of the phenotype could be explained by the reduced deletion efficiency of *Prm1-cre,* which resulted in *Eed* expression in some of the *Eed* m-z- embryos that reached the morula stage (Figure S 5-2E, F). We then compared the developmental effect upon removal of *Eed* in primary versus growing oocytes and observed that earlier maternal deletion has a more detrimental effect on preimplantation development (Figure 5.1D). Notably, the earlier *Gdf9-iCre* deletion causes a partial 2-cell arrest, while *Eed* $^{fl/fl}$ *Zp3-cre* m-z- embryos develop normally to the morula and early blastocyst stage (84 hpf) but undergo fragmentation before progressing to the expanded blastocyst stage (109 hpf).

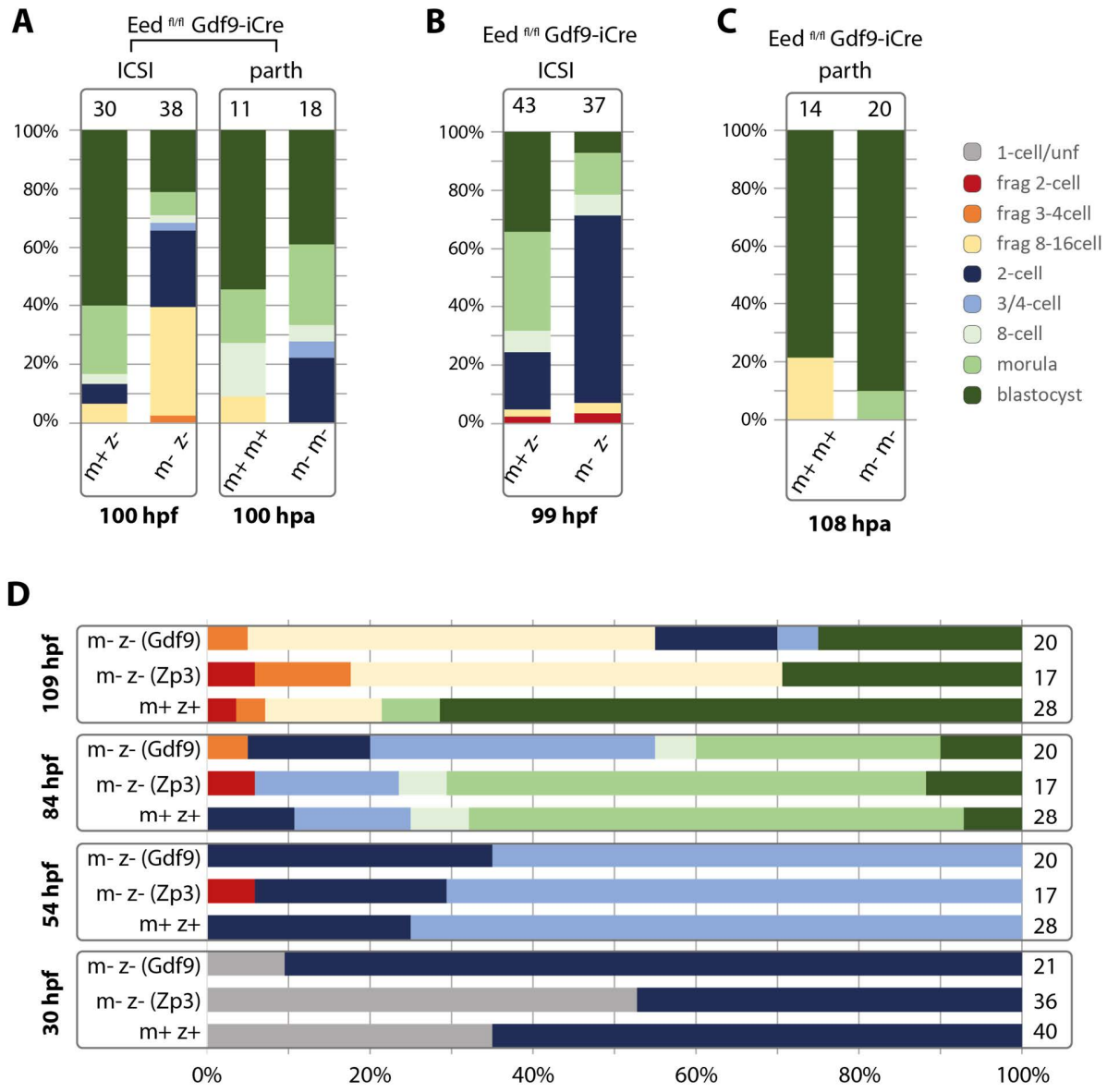



***Figure 5.1. Developmental progression of* Eed *deficient embryos***

***A, B, C*** *– Development of ICSI versus parthenote preimplantation embryos, scored at embryonic day E4 (ca. 96 hours post fertilization/activation). Eed knock-out embryos were generated by Gdf9-iCre deletion of Eed $^{fl/fl}$ alleles in early oocytes. ICSI embryos were fertilized with sperm form an Eed $^{fl/fl}$ Prm1-cre male, thus generating maternally deficient, homozygous Eed knock-out embryos (m-z-), or maternal wild type, heterozygous Eed embryos (m+z-) as a control group. Knock-out MII oocytes from Eed $^{fl/fl}$ Gdf9-iCre mothers were activated with $SrCl_2$ to generate homozygous Eed deficient embryos (m-m-). Control m+m+ parthenote embryos were generated by activating MII oocytes from cre-negative mothers. Panels A, B and C represent independent experiments. Numbers on top indicate the total number of embryos analysed.*

***D*** *– Comparative time course developmental progression of Eed knock-out embryos derived from Eed $^{fl/fl}$ Gdf9-iCre, Eed $^{fl/fl}$ Zp3-Cre or cre-negative mothers, indicated as m-z- (Gdf9), m-z- (Zp3) and m+z+ respectively. Fertilization by natural mating with Eed $^{fl/fl}$ Prm1-Cre males in the case of m-z- (Gdf9) and m-z- (Zp3) and cre-negative males in the case of m+z+. Unfertilized MII oocytes are included in the graph only for the first time point (30 hpf) and omitted for the later time points.*

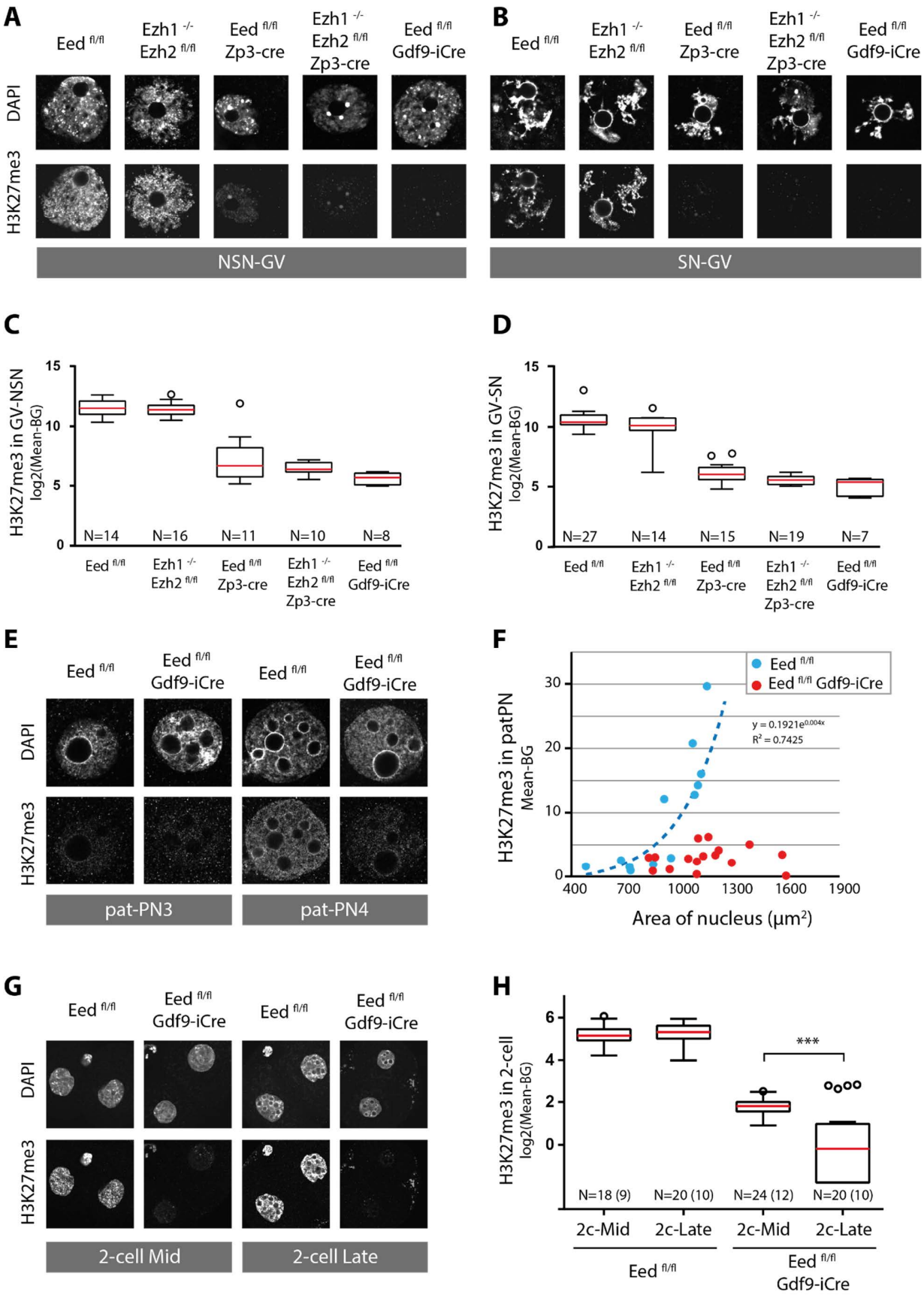
These results highlight two distinct phenotypes caused by the loss of maternal and zygotic Eed – an early block around ZGA and a developmental failure upon lineage specification. The latter defect was also observed in an alternative model of PRC2 m-z- embryos that are deficient for *Ezh1* and *Ezh2* (Figure S 5-3A). A role for PRC2 and H3K27me3 in lineage specification has been suggested by a recent study showing that PRC2 regulates the expression of trophectoderm-specific genes in the emerging blastocyst (Saha et al., 2013). In line with this, our preliminary results indicate aberrant *Eomes* expression in some of the *Ezh1/Ezh2* double knock-out embryos (Figure S 5-3B).

Next, we tested whether PRC2 could be an essential maternal effect factor like PRC1 by performing breeding experiments with *Eed* $^{fl/fl}$ *Gdf9-iCre* and *Eed* $^{fl/fl}$ *Zp3-cre* females mated with wild type males. We observed live birth of healthy pups from *Eed*-deficient oocytes, as well as from *Ezh1*/Ezh2-deficient oocytes (Figure S 5-2F). However, the litter sizes upon maternal deletion of *Eed* were reduced in comparison to breedings with paternal deletion or no deletion of *Eed* (Table 5.1). Thus, maternal PRC2 has an effect on fertility but is not absolutely essential for development, since *Eed* m-z+ embryos can develop to term. To further understand the cause of the observed preimplantation defects in *Eed* m-z- embryos, we addressed the issue from the perspective of the dynamic changes that happen on the paternal genome in the zygote and compared the development of embryos with a paternal and a maternal genome (m-z-) versus embryos with a maternal genome only (parthenotes, m-m-). Surprisingly and in contrast to m-z-, the *Eed*-deficient parthenotes displayed rather normal development and blastocyst formation (Figure 5.1A, C, Figure S 5-1).

***Table 5.1. Effect on fertility of Eed deletion in the germline***

*The effect of Eed-deficiency in the germline was tested in breeding experiments with females deleting the conditional Eed allele during oogenesis by either Zp3-cre or Gdf9-iCre. The cre-positive females were mated with wild type males to generate heterozygous embryos. In the corresponding cross for the male germline, Eed* $^{fl/fl}$ *Prm1-cre males were mated with wild type females, resulting in m+z- pups. A breeding with cre-negative parents was used as a reference. The table gives the number of analysed breedings, the total number of litters and pups in the analysed breedings, as well as the average litter size for each cross.*

| Genotype | Breedings | Litters | Pups | Average litter size |
|---|---|---|---|---|
| *Eed* $^{m-z+}$ *(Gdf9-iCre)* | 2 | 6 | 16 | 2.7 |
| *Eed* $^{m-z+}$ *(Zp3-cre)* | 4 | 15 | 49 | 3.3 |
| *Eed* $^{m+z-}$ *(Prm1-cre)* | 2 | 5 | 32 | 6.4 |
| *Eed* $^{fl/fl}$ | 4 | 38 | 239 | 6.3 |


A
Eed fl/fl
Ezh1 -/- Ezh2 fl/fl
Eed fl/fl Zp3-cre
Ezh1 -/- Ezh2 fl/fl Zp3-cre
Eed fl/fl Gdf9-iCre
DAPI
H3K27me3
NSN-GV
B
Eed fl/fl
Ezh1 -/- Ezh2 fl/fl
Eed fl/fl Zp3-cre
Ezh1 -/- Ezh2 fl/fl Zp3-cre
Eed fl/fl Gdf9-iCre
SN-GV
C
H3K27me3 in GV-NSN log2(Mean-BG)
15
10
5
0
N=14
N=16
N=11
N=10
N=8
Eed fl/fl
Ezh1 -/- Ezh2 fl/fl
Eed fl/fl Zp3-cre
Ezh1 -/- Ezh2 fl/fl Zp3-cre
Eed fl/fl Gdf9-iCre
D
H3K27me3 in GV-SN log2(Mean-BG)
15
10
5
0
N=27
N=14
N=15
N=19
N=7
Eed fl/fl
Ezh1 -/- Ezh2 fl/fl
Eed fl/fl Zp3-cre
Ezh1 -/- Ezh2 fl/fl Zp3-cre
Eed fl/fl Gdf9-iCre
E
Eed fl/fl
Eed fl/fl Gdf9-iCre
Eed fl/fl
Eed fl/fl Gdf9-iCre
DAPI
H3K27me3
pat-PN3
pat-PN4
F
Eed fl/fl
Eed fl/fl Gdf9-iCre
y = 0.1921e0.004x
R2 = 0.7425
H3K27me3 in patPN Mean-BG
30
20
10
0
400
700
1000
1300
1600
1900
Area of nucleus (μm2)
G
Eed fl/fl
Eed fl/fl Gdf9-iCre
Eed fl/fl
Eed fl/fl Gdf9-iCre
DAPI
H3K27me3
2-cell Mid
2-cell Late
H
H3K27me3 in 2-cell log2(Mean-BG)
6
4
2
0
***
N=18 (9)
N=20 (10)
N=24 (12)
N=20 (10)
2c-Mid
2c-Late
2c-Mid
2c-Late
Eed fl/fl
Eed fl/fl Gdf9-iCre

In order to relate the observed phenotypic outcome to the core molecular function of PRC2, we assessed the levels of H3K27me3 in PRC2-deficient oocytes and embryos (Figure 5.2). Deletion of *Eed* in primary oocytes leads to a gradual and global depletion of H3K27me3, which by the end of oogenesis is reduced to near-background levels (Figure 5.2A, B, C, D, Figure S 5-4). A later deletion of *Eed* or *Ezh2* by *Zp3-cre* results in higher residual H3K27me3 levels, including some euchromatic signal. Upon fertilization and histone incorporation, the paternal pronucleus gradually gains methylation on H3K27 (Puschendorf et al., 2008) and therefore we looked for residual PRC2 activity in the developing zygote. We did not detect any gain of *de novo* H3K27me3 in *Eed* $^{fl/fl}$ *Gdf9-iCre* zygotes, which lead us to the conclusion that PRC2 activity was fully abolished following *Eed* deletion in primary oocytes. We further analysed 2-cell embryos before and after DNA replication and found a significant decrease in the residual H3K27me3 levels between mid-2-cell and late 2-cell m-z- embryos (Figure 5.2H), which could be due to passive loss of histone methylation in the absence of the active methyltransferase.

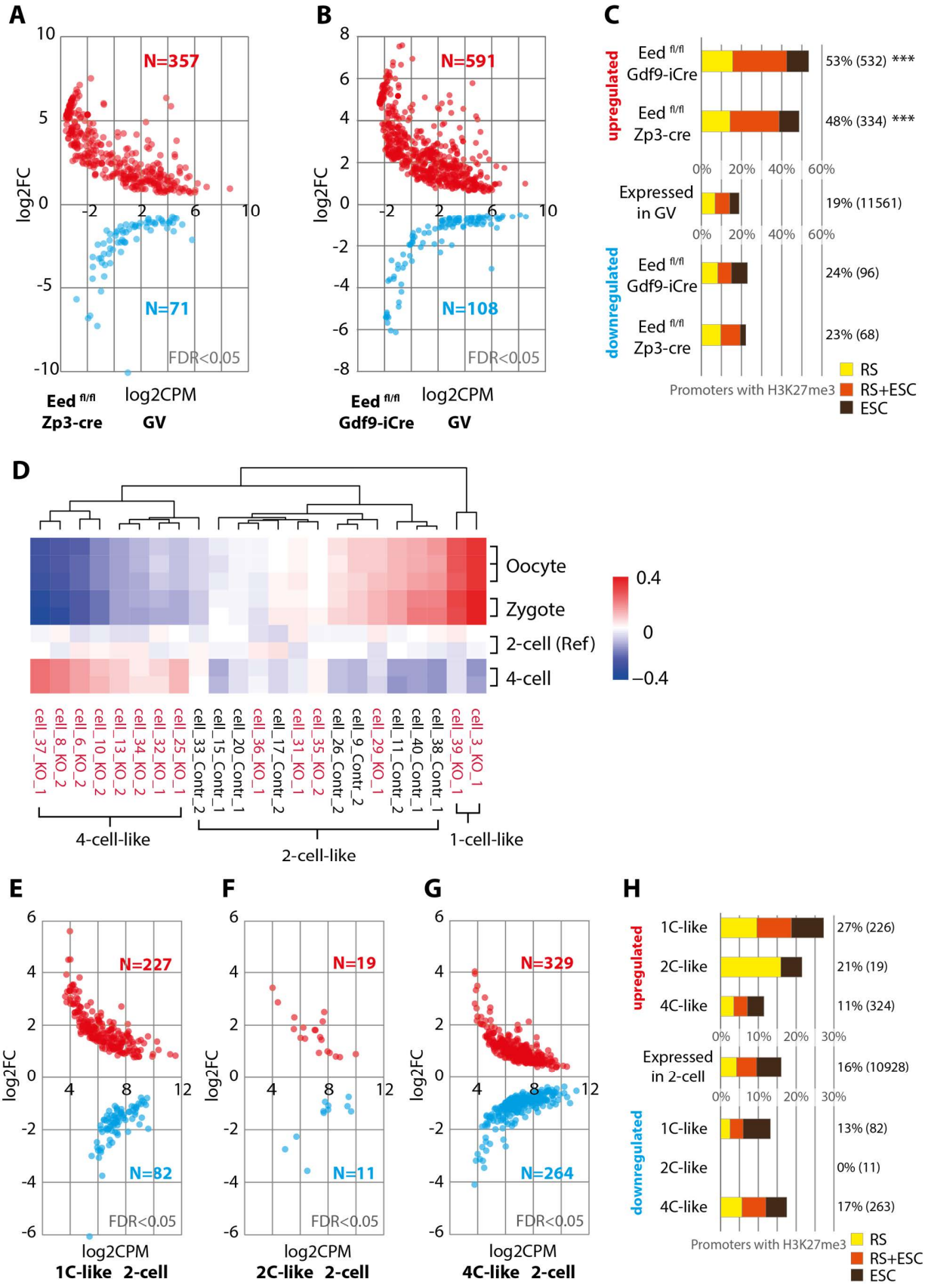
---

***Figure 5.2. Levels of H3K27me3 in PRC2-deficient oocytes and embryos***

***A, B, C, D*** *– Immunofluorescent staining and quantification of signal intensity for H3K27me3 in NSN-GV oocytes (A, C) and SN-GV oocytes (B, D). Upper panels in A and B show the DNA stained with DAPI, lower panels show the H3K27me3 signal under constant imaging conditions. Quantification was done by measuring the mean intensity for H3K27me3 in the nucleus (nuclear area determined according to the DAPI staining), normalization was done 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 the NSN-GV (C) and SN-GV (D). The number of analysed oocytes per genotype and stage is given on the graphs.*

***E, F*** *– Immunofluorescent staining and quantification of de novo H3K27me3 on the paternal pronucleus (PN) in zygotes. Upper panels in (E) show the DAPI signal, lower panels show the H3K27me3 signal at constant imaging conditions. Note that in the PN3 zygote stage, there is still no H3K27me3 both in the Eed $^{fl/fl}$ control zygotes and in the Eed $^{fl/fl}$ Gdf9-iCre knock-out. H3K27me3 comes up in the PN4 zygote stage in control embryos, but not in Eed $^{fl/fl}$ Gdf9-iCre knock-out zygotes. The H3K27me3 was quantified as described above and represented as function of the area of the pronucleus (F), which correlates with the zygote stage (i.e. the male PN expands during zygote progression). Blue dots represent single Eed $^{fl/fl}$ PN (control), red dots correspond to knock-out Eed $^{fl/fl}$ Gdf9-iCre male PN. Control embryos gain signal on the male PN as an exponential function of the pronuclear size, while Eed $^{fl/fl}$ Gdf9-iCre male PN do not show any de novo H3K27me3.*

***G, H*** *– Immunofluorescent staining and quantification of H3K27me3 in 2-cell embryos. Upper panels in (G) show the DNA by DAPI staining, lower panels show the H3K27me3 signal in the two blastomeres and the polar body for the representative embryos of the mid and late 2-cell stages in Eed $^{fl/fl}$ versus Eed $^{fl/fl}$ Gdf9-iCre. Control Eed $^{fl/fl}$ embryos were fertilized with wild type sperm, knock-out Eed $^{fl/fl}$ Gdf9-iCre embryos were fertilized with Eed $^{fl/fl}$ Prm1-Cre knock-out sperm. H3K27me3 was quantified as described above and represented as a Tukey boxplot on a logarithmic scale for the two substages and the two genotypes (H). Numbers below the boxes indicate the number of cells quantified (number of embryos is given in brackets). There are four outlier cells (coming from two embryos) in the late 2-cell Eed $^{fl/fl}$ Gdf9-iCre embryos, which presumably correspond to escaper knock-out embryos, expressing Eed from a non-deleted paternal allele. A Mann-Whitney test identified a significant difference ($P<0.0005$) between the two groups of Eed $^{fl/fl}$ Gdf9-iCre 2-cell embryos.*


A
N=357
N=71
FDR<0.05
log2FC
log2CPM
Eed fl/fl Zp3-cre GV
B
N=591
N=108
FDR<0.05
log2FC
log2CPM
Eed fl/fl Gdf9-iCre GV
C
upregulated
Eed fl/fl Gdf9-iCre
53% (532) ***
Eed fl/fl Zp3-cre
48% (334) ***
Expressed in GV
19% (11561)
downregulated
Eed fl/fl Gdf9-iCre
24% (96)
Eed fl/fl Zp3-cre
23% (68)
Promoters with H3K27me3
RS
RS+ESC
ESC
D
Oocyte
Zygote
2-cell (Ref)
4-cell
0.4
0
−0.4
cell_37_KO_1
cell_8_KO_2
cell_6_KO_2
cell_10_KO_2
cell_13_KO_2
cell_34_KO_2
cell_32_KO_1
cell_25_KO_1
cell_33_Contr_2
cell_15_Contr_1
cell_20_Contr_1
cell_36_KO_1
cell_17_Contr_2
cell_31_KO_1
cell_35_KO_2
cell_26_Contr_2
cell_9_Contr_2
cell_29_KO_1
cell_11_Contr_2
cell_40_Contr_1
cell_38_Contr_1
cell_39_KO_1
cell_3_KO_1
4-cell-like
2-cell-like
1-cell-like
E
N=227
N=82
FDR<0.05
log2FC
log2CPM
1C-like 2-cell
F
N=19
N=11
FDR<0.05
log2FC
log2CPM
2C-like 2-cell
G
N=329
N=264
FDR<0.05
log2FC
log2CPM
4C-like 2-cell
H
upregulated
1C-like
27% (226)
2C-like
21% (19)
4C-like
11% (324)
Expressed in 2-cell
16% (10928)
downregulated
1C-like
13% (82)
2C-like
0% (11)
4C-like
17% (263)
Promoters with H3K27me3
RS
RS+ESC
ESC

### *Figure 5.3. Transcriptome analysis of Eed deficient oocytes and embryos*

***A, B*** *– RNA-seq analysis comparing the transcriptomes of GV oocytes from Eed $^{fl/fl}$ females versus Eed $^{fl/fl}$ Zp3-cre (A) and Eed $^{fl/fl}$ Gdf9-iCre (B) oocytes respectively. Differentially expressed genes with FDR<0.05 are represented as MA plots with the logarithmic expression value (x-axis) against the logarithmic fold-change value (y-axis). Upregulated genes in the Eed knock-out oocytes are highlighted in red, downregulated genes are highlighted in blue. The number of up-/downregulated genes is shown on the graphs.*

***C*** *– Proportion of misregulated genes in Eed knock-out oocytes with H3K27me3-positive promoters in ESC and round spermatids (RS). The upper two bars correspond to the H3K27me3 profiles of the upregulated genes in Eed $^{fl/fl}$ Gdf9-iCre and Eed $^{fl/fl}$ Zp3-cre respectively. Middle bar shows the percentage of H3K27me3-positive promoters in the genes normally expressed at the GV stage. Lower two bars represent the percentage of H3K27me3-positive promoters among the downregulated genes in Eed $^{fl/fl}$ Gdf9-iCre and Eed $^{fl/fl}$ Zp3-cre respectively. The relative proportion of RS-specific, ESC-specific and common RS+ESC H3K27me3-positive promoters within each of the gene sets is shown in colour (yellow for RS-specific, orange for common and brown for ESC-specific, in this order). The percentage of H3K27me3-positive promoters is given for each gene set on the right and the number of all genes in a gene set is given in brackets (note that not all genes could be included in the analysis, because there is no information for some of the gene promoters). *** Chi-squared P<0.0005 (upregulated in knock-out versus expressed).*

***D****– Clustered correlation heatmap showing the contrast of each single 2-cell embryo transcription profile to the external 2-cell reference horizontally (expression data for wild type oocyte, zygote, 2-cell and 4-cell taken from (Park et al., 2013a)) versus the contrast of the wild type oocyte, zygote and 4-cell stage expression profiles to the 2-cell reference vertically. Stronger positive correlation between the contrasts is illustrated by a more intense red colour, while stronger negative correlation between contrasts results in a more intense blue colour. Similar contrasts, arising from similar expression profiles, show poor correlation and indicated by a light colours. The Eed knock-out embryos are highlighted in red below the heatmap. The last number in the sample name indicates the biological replicate (e.g. cell_37_KO_1 is an Eed knock-out and comes from knock-out mouse 1). The three main clusters are outlined below the sample names and labelled as "4-cell-like", "2-cell-like" and "1-cell-like" in respect to the similarity of the single embryo transcription profiles to the respective wild type expression profile.*

***E, F, G****– Differential expression analysis based on the clustering of the single 2-cell embryos relative to the external wild type transcriptomes for oocyte, 1-cell, 2-cell and 4-cell stages. Single Eed knock-out embryos from the same cluster were considered as biological replicates and compared to the control group (same for the three contrasts). Differentially expressed genes at FDR<0.05 are shown for the three contrasts "1-cell-like" Eed knock-out embryos (1C-like) versus Control 2-cell embryos (F), "2-cell-like" Eed knock-out embryos (2C-like) versus Control embryos (G) and "4-cell-like" Eed knock-out embryos (4C-like) versus Control (H).*

***H*** *– Proportion of misregulated genes in Eed knock-out 2-cell embryos having H3K27me3-positive promoters in RS and ESC (see above for explanation).*

We then focused on the transcriptional effect of the PRC2 loss in GV oocytes and 2-cell embryos. In *Eed* $^{fl/fl}$ *Zp3-cre* and *Eed* $^{fl/fl}$ *Gdf9-iCre* oocytes, we found 357 and 591 significantly upregulated genes, respectively, and just about a hundred downregulated transcripts in both contrasts (Figure 5.3A, B). A single-embryo RNA-seq on *Eed* $^{fl/fl}$ *Gdf9-iCre* m-z- and control 2-cell embryos revealed a heterogeneous response to the loss of PRC2 around ZGA. The *Eed* m-z- embryos clustered in three distinct branches based on their similarity to wild type 1-cell, 2-cell and 4-cell expression profiles (referred to as 1C-like, 2C-like and 4C-like 2-cell embryos) (Figure 5.3D). The difference between the three groups of PRC2-deficient 2-cell embryos was also reflected in the number and distribution of differentially expressed genes (Figure 5.3E, F, G).

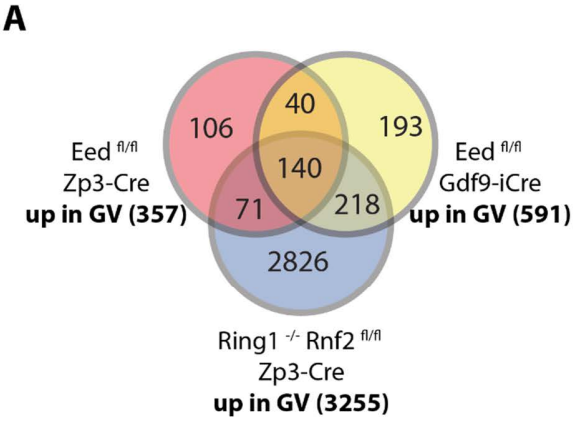
In order to explain the meaning of the observed transcriptional changes, we first looked at how much of these are a potential direct effect of the PRC2 loss. Since PRC2 is a transcriptional repressor, its primary targets would be upregulated in a loss-of-function condition. Indeed, based on the H3K27me3 profiles from the developmentally closely related round spermatids and ESC, we found a substantial enrichment for H3K27me3-positive gene promoters among the upregulated genes in oocytes (Figure 5.3C). Moreover, there was a substantial portion of the upregulated RNAs that were also aberrantly expressed in PRC1-deficient oocytes (Figure 5.4A). The 140 overlapping genes could be classified as core Polycomb targets, sensitive to the loss of either one of the two major Polycomb complexes. Among these genes, we found major transcription factors involved in differentiation such as Hox-, Klf- and Gata-family genes (Figure S 5-5). Based on the normal expression of these genes from the oocyte to the 4-cell stage, we could distinguish two components of the PRC2 loss-of-function effect on the maternal RNA pool – higher than normal upregulation of genuine maternal transcripts, and aberrant activation of genes that would normally be required in later stages of development.

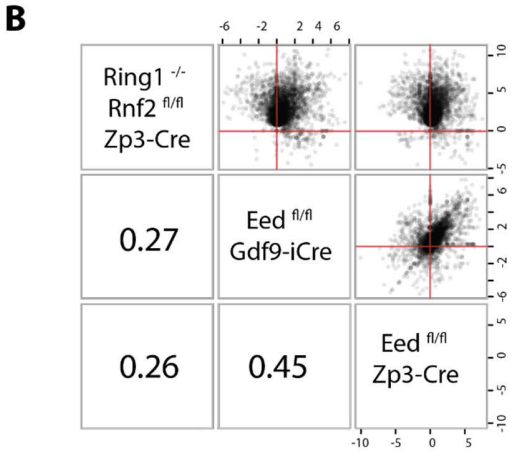
However, the transcriptional changes in GV oocytes could not be solely responsible for the observed phenotype in *Eed* m-z- embryos, particularly because *Eed* m-m- and m-z+ embryos develop to blastocyst and full term respectively. Consistent with this, we observed completely different sets of misregulated genes among the three groups of knock-out 2-cell embryos and versus the *Eed*-deficient oocytes (Figure 5.4C, D). Surprisingly, the upregulated genes in 2-cell m-z- embryos were not strongly enriched for H3K27me3 targets (Figure 5.3H), which argues for

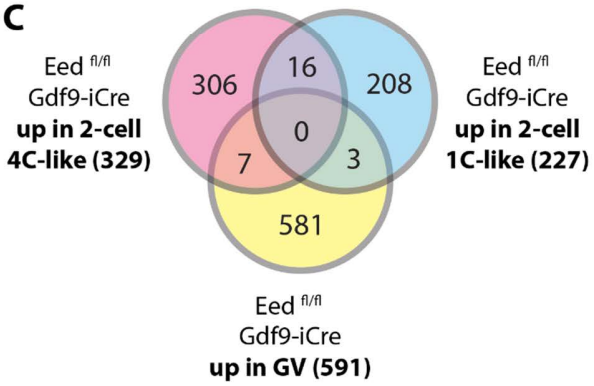
secondary effects or abnormal ZGA. Indeed, after relating the differentially expressed genes to the changes that normally occur around ZGA, we found that 1C-like, m-z- 2-cell embryos failed to degrade the maternally provided RNAs (Figure 5.4E). In the same time, 4C-like m-z- 2-cell embryos have developed to a more advanced ZGA state, characterized mainly by stronger depletion of maternal transcripts and an upregulation of the 2-cell zygotic transcriptome (Figure 5.4F). Gene ontology analysis highlighted the underlying functional difference between the sets of upregulated genes in oocytes and 2-cell embryos (Figure S 5-6, Table 5.2).

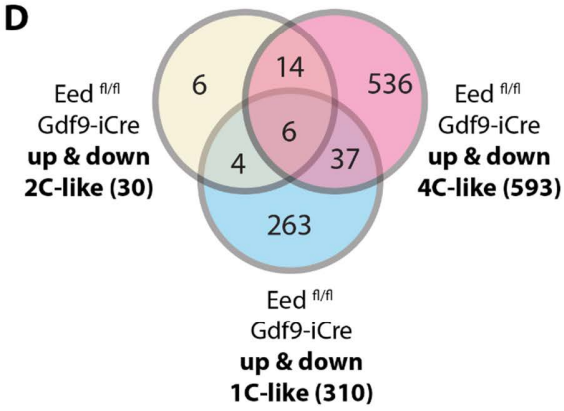
## 5.3 Discussion

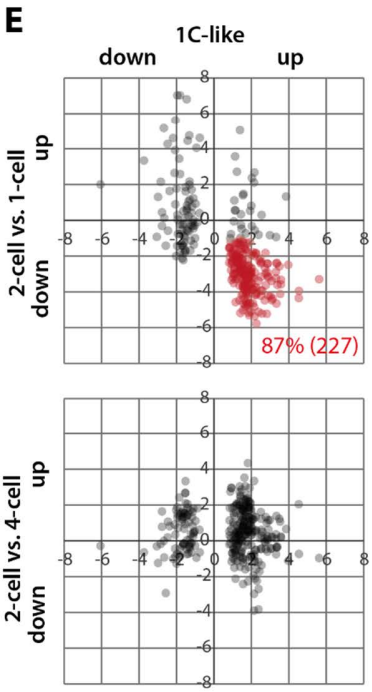
We demonstrated that genetic ablation of *Eed* has an effect on H3K27me3 *in vivo,* and leads to a developmental and transcriptional response in late oocytes and early embryos. These findings suggest a PRC2/H3K27me3-dosage dependent response to the loss of PRC2, which is first manifested during ZGA. The ZGA requirement seems to depend on the maternal PRC2/H3K27me3 dosage, as well as on the presence of the paternal genome. In the mouse zygote, H3K27me3 has three origins – inherited maternal, inherited paternal and *de novo* zygotic. In our study we do not alter the paternal contribution *per se,* but by removing the enzymatic complex PRC2, we prevent the perpetuation of any transmitted information via the male germline (Brykczynska et al., 2010; Hammoud et al., 2009). Nevertheless, *Eed* $^{fl/fl}$ *Zp3-cre* m-z- embryos go pass the 2-cell stage, which suggests that ZGA can occur also in the absence of *de novo* H3K27me3. This appears to be the case also in parthenote m-m- embryos. However, *Eed* $^{fl/fl}$ *Gdf9-iCre* m-z- arrest at the 2-cell stage and display a range of ZGA abnormalities. So what is the key determinant for this phenotype? We believe that there is a threshold for the H3K27me3 levels, below which the zygote cannot properly execute the ZGA program. According to this hypothesis, potential residual levels of H3K27me3 in *Eed* $^{fl/fl}$ *Zp3-cre* oocytes appear to be sufficient for maintaining the required repressed state in the zygote.

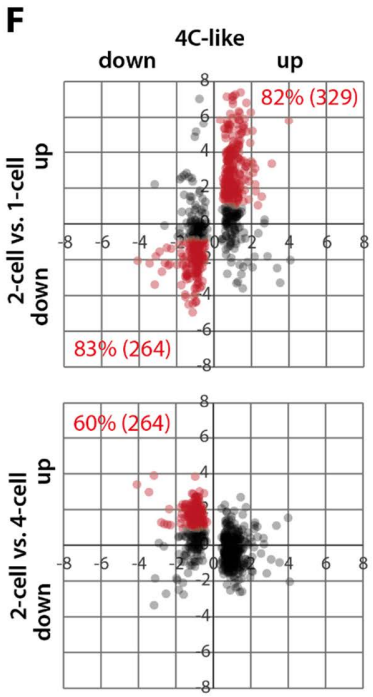

A
106
40
193
140
71
218
2826
Eed fl/fl
Zp3-Cre
up in GV (357)
Eed fl/fl
Gdf9-iCre
up in GV (591)
Ring1 -/- Rnf2 fl/fl
Zp3-Cre
up in GV (3255)


B
Ring1 -/-
Rnf2 fl/fl
Zp3-Cre
0.27
Eed fl/fl
Gdf9-iCre
0.26
0.45
Eed fl/fl
Zp3-Cre


C
306
16
208
0
7
3
581
Eed fl/fl
Gdf9-iCre
up in 2-cell
4C-like (329)
Eed fl/fl
Gdf9-iCre
up in 2-cell
1C-like (227)
Eed fl/fl
Gdf9-iCre
up in GV (591)


D
6
14
536
6
4
37
263
Eed fl/fl
Gdf9-iCre
up & down
2C-like (30)
Eed fl/fl
Gdf9-iCre
up & down
4C-like (593)
Eed fl/fl
Gdf9-iCre
up & down
1C-like (310)


E
1C-like
down
up
2-cell vs. 1-cell
up
down
87% (227)
2-cell vs. 4-cell
up
down


F
4C-like
down
up
2-cell vs. 1-cell
up
down
82% (329)
83% (264)
2-cell vs. 4-cell
up
down
60% (264)

We further speculate that reduction of global H3K27me may involve gene-specific retention of H3K27me3, which ensures that key developmental genes remain silent even upon reduction of PRC2 activity. The latter phenomenon has been already observed in ESCs (Shen et al., 2008). However, as the proposed mechanism is dosage-dependent, adding a genome that needs to be chromatinized presumably raises the total requirement for PRC2-mediated repression in the zygote and the residual maternal H3K27me3 is not sufficient to suppress the aberrant transcription, which ultimately leads to a 2-cell arrest. We therefore propose that PRC2-mediated H3K27me3 acts as a silencing buffer system to protect the emerging totipotent state in the mouse embryo.

This complex picture is in strong contrast to the function of the other major Polycomb complex, PRC1, which is required for the regulation of maternal RNA in the oocyte and PRC1 ablation in oogenesis leads to a 2-cell arrest (Posfai et al., 2012). The classical model of Polycomb

***Figure 5.4. Comparative analysis of genes misregulated in Eed knock-out oocytes and embryos***

***A*** *– Intersecting the significantly upregulated genes (FDR<0.05) in GV oocytes deficient for Eed (differential expression between Eed fl/fl and Eed fl/fl Gdf9-iCre or Eed fl/fl Zp3-cre respectively) with the genes upregulated in PRC1 knock-out GV oocytes (differentially expressed between Ring1 -/- Rnf2 fl/fl and Ring1 -/- Rnf2 fl/fl Zp3-cre). The 140 common genes are further analysed in Figure S 5-5.*

***B*** *– Correlation matrix of contrasts for the 3593 unique genes upregulated in one of the knock-out conditions shown in (A). The logarithmic fold-change values for the respective contrast are given on the side. The red lines in the scatter plots cross at (0, 0). The Spearman correlation coefficient is shown.*

***C*** *– Intersecting the significantly upregulated genes in the "1-cell-like" Eed knock-out embryos (1C-like), "4-cell-like" Eed knock-out embryos (4C-like) and the genes upregulated in Eed fl/fl Gdf9-iCre GV oocytes versus their respective control (see above).*

***D*** *– Intersecting the significantly differentially expressed genes (up- and downregulated) in 1C-like, 2C-like and 4C-like 2-cell Eed knock-out embryos versus 2-cell control embryos.*

***E, F*** *–Scatter plots relating the transcriptional changes in Eed knock-out 2-cell embryos to the transcriptional changes that normally occur in the first three stages of development (wild type data from (Park et al., 2013a)). The logarithmic fold change of significantly misregulated genes in the 1C-like Eed knock-out (E) or 4C-like Eed knock-out (F) vs control 2-cell embryos is plotted against the logarithmic fold change in expression between the 1-cell and 2-cell stages (upper panels) or against the logarithmic fold change between the 2-cell and 4-cell transcriptome profiles (lower panels). The highlighted data points in (E), upper panel, indicate the genes that were found upregulated in the 1C-like 2-cell knock-out embryos and that would normally go down in expression from the 1-cell to the 2-cell stage. Out of the 227 upregulated genes in 1C-like 2-cell embryos, 87% are more than 2-fold downregulated in the wild type 2-cell embryos vs the zygote stage. In the upper panel in (F), the highlighted points indicate the genes that are significantly up- or downregulated in 4C-like 2-cell Eed knock-out embryos and more than 2-fold up- or downregulated in wild type 2-cell versus 1-cell embryos. In particular, 82% of the 329 upregulated genes in 4C-like embryos are also going up in wild type 2-cell embryos. At the same time, 83% of the downregulated genes in 4C-like 2-cell embryos are also going down in wild type 2-cell embryos. In the lower panel in (F) the highlighted points correspond to the set of genes that are downregulated in 4C-like knock-out embryos and at the same time more than 2-fold downregulated in wild type 4-cell versus 2-cell embryos.*

interaction suggests that PRC1 is recruited to H3K27me3-positive genes, but recent studies have indicated that there may also be alternative ways of targeting Polycomb complexes to chromatin and DNA (Schwartz and Pirrotta, 2013). Interestingly, global H2AK119ub levels remained unchanged in the *Eed*-deficient oocytes ad embryos (Figure S 5-7), which suggests a potential PRC2-independent recruitment of PRC1 in preimplantation development.

In conclusion, we identified a core set of Polycomb target genes in oocytes and uncovered a new role for PRC2 and H3K27me3 in preimplantation development and particularly during ZGA – a dosage-dependent requirement for maternally provided H3K27me3 and *de novo* PRC2 activity on the paternal pronucleus are the prerequisite for proper ZGA and progression to blastocyst. Thus, PRC2 is essential for the maintenance and transmission of chromatin-borne information from the parents to the zygote.

There are some open questions that require further investigation. Most importantly, the gene-specific effect of PRC2 in early embryos needs to be elucidated in greater detail, particularly in light of the dosage effect and the observed developmental heterogeneity. A more sensitive genome-wide profiling method in combination with single-cell RT-qPCR could potentially uncover primary Polycomb target genes that would be responsible for the 2-cell arrest or the later developmental abnormalities. Next, the dosage-dependent requirement for PRC2/H3K27me3 could be validated and dissected further by actively removing H3K27me3 using ectopic expression of H3K27-specific demethylase at different stages of development. Finally, an interesting and still not fully answered question is whether *de novo* PRC2 activity on the paternal genome is essential for development and how it may affect parent-specific gene expression. Even though the results presented here suggest that this is potentially important, it is difficult to fully decompose the relative contribution of the maternal and paternal genomes in the zygote without manipulating the embryo. Nuclear transfer experiments could be used to generate embryos having only paternal genomes in a PRC2-deficient background, which should lead to a more severe developmental defect as compared to m-z- or m-m-. This kind of experiments could also address the contribution of the maternal chromatin versus the maternal cytoplasm (RNA and protein), as was done in the study describing the maternal role of PRC1 (Posfai et al., 2012).

## 5.4 Materials and methods

### 5.4.1 Mice

Conditional *Eed* mice were generated as described elsewhere (Xie et al., 2014). *Eed* $^{fl/fl}$ mice were crossed to mice carrying *Gdf9-iCre, Zp3-cre* or *Prm1-cre* and maintained by transmitting the cre though the non-affected germline as described previously (Lan et al., 2004; Puschendorf et al., 2008). In a similar manner, *Ezh1* $^{-/-}$ *Ezh2* $^{fl/fl}$ conditional knock out mice were generated, whereas the conditional and knock-out alleles for *Ezh1* and *Ezh2* were reported in the literature (Ezhkova et al., 2011; Puschendorf et al., 2008). Experiments were performed in accordance with the Swiss animal protection laws and institutional guidelines.

### 5.4.2 Oocyte and embryo collection, culture and manipulation

Oocytes and embryos were harvested and cultured upon superovulation as described above (4.4.1.) Parthenogenic activation of MII oocytes was done by 5 hours incubation in Calcium-free CZB medium containing 10 mM SrCl2 and 5 µg/ml cytochalasin B. After extensive washing in M2 medium (Sigma, M7167), activated oocytes were cultered in M16 medium (Sigma, M7292) until E2.0 and subsequently in KSOMaa (Milipore, MR-106-D). For ICSI, MII oocytes were collected from superovulated females. The microinjection procedure was performed on a Piezo (Eppendorf) drill micromanipulator (Olympus). The sperm head was aspirated in the injection pipette (10 µm internal diameter) and injected into 1 oocyte. ICSI-fertilized embryos were cultured as described above (4.4.1.).

### 5.4.3 Immunufluoresent analysis

Prior to fixation, embryos from *in vitro* culture were washed twice in M2, and freshly isolated GV oocytes were treated with Hyaluronidase to remove cumulus cells. Subsequently, embryos or oocytes were fixed for 15 min in 4% formaldehyde (Electron microscopy sciences,15713-S) in PBS (pH 7.4) and permeabilized with 0.2% Triton-X 100 in PBS for 15 min at room temperature (RT). Fixed embryos were blocked at least 4 hours at RT in 0.1% Tween-20 in PBS containing 2%

BSA and 5% normal goat serum, and were then incubated with primary antibodies in blocking solution overnight at 4ºC. Embryos were washed three times for 30 min in 0.1% Tween-20 in PBS containing 2% BSA before application of secondary antibodies for 1 h at RT, followed by three washing steps for 30 min in 0.1% Tween-20 in PBS containing 2% BSA in the dark. Embryos were mounted in Vectashield containing DAPI (Vector Laboratories, H-1200). All primary and secondary antibodies used here are listed in Appendix B: Antibodies.

### 5.4.4 Transcriptional profiling of GV oocytes by RNA-seq

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 vava libraries were sequenced on an Illumina HiSeq instrument. Raw data was processed as described elsewhere (Tippmann et al., 2012). Differetial expression analysis was performed using the raw readcounts and the "edgeR" package in R (see Appendix A: R scripts). Gene ontology enrichment analysis was done with the GENERIC GENE ONTOLOGY (GO) TERM FINDER (Princeton University, USA (Boyle et al., 2004)) and REVIGO (Rudjer Boskovic Institute, Croatia (Supek et al., 2011)).

### 5.4.5 Single-embryo transcriptional profiling of 2-cell embryos

Zygotes from superovulated females, mated with the respective males, were isolated as described above (4.4.1.) and cultured in M16. Embryo collection took place around 48 hours post hCG injection (ca. 36 hpf). First the Zona pelucida was removed with 5 mg/ml Proteinase (Sigma, P5147-1G, in M2 medium). Individual embryos were transferred with a micro-pipette on the inner side of the cap of a 0.5 ml LoBind Eppendorf tube and tube was frozen in liquid nitrogen. Samples were stored at -80°C before downstream processing, following the CEL-Seq protocol (Hashimshony et al., 2012).

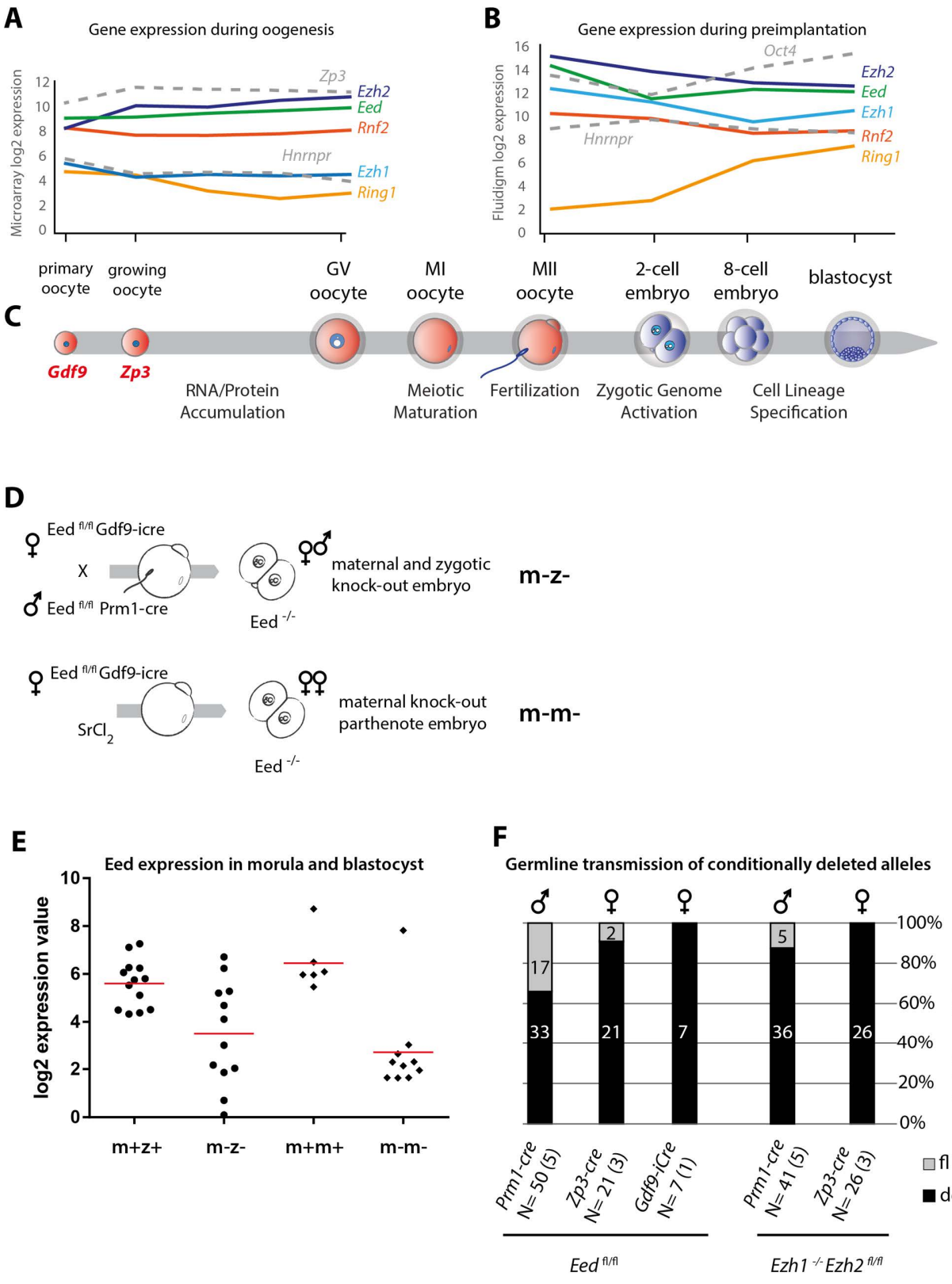

A
Gene expression during oogenesis
Microarray log2 expression
Zp3
Ezh2
Eed
Rnf2
Hnrnpr
Ezh1
Ring1
primary oocyte
growing oocyte
B
Gene expression during preimplantation
Fluidigm log2 expression
Oct4
Ezh2
Eed
Ezh1
Rnf2
Hnrnpr
Ring1
GV oocyte
MI oocyte
MII oocyte
2-cell embryo
8-cell embryo
blastocyst
C
Gdf9
Zp3
RNA/Protein Accumulation
Meiotic Maturation
Fertilization
Zygotic Genome Activation
Cell Lineage Specification
D
Eed fl/fl Gdf9-icre
X
Eed fl/fl Prm1-cre
Eed -/-
maternal and zygotic knock-out embryo
m-z-
Eed fl/fl Gdf9-icre
SrCl2
Eed -/-
maternal knock-out parthenote embryo
m-m-
E
Eed expression in morula and blastocyst
log2 expression value
m+z+
m-z-
m+m+
m-m-
F
Germline transmission of conditionally deleted alleles
Deletion efficiency
100%
80%
60%
40%
20%
0%
17
33
2
21
7
5
36
26
Prm1-cre N= 50 (5)
Zp3-cre N= 21 (3)
Gdf9-iCre N= 7 (1)
Prm1-cre N= 41 (5)
Zp3-cre N= 26 (3)
fl
del
Eed fl/fl
Ezh1 -/- Ezh2 fl/fl

**Figure S 5-2**

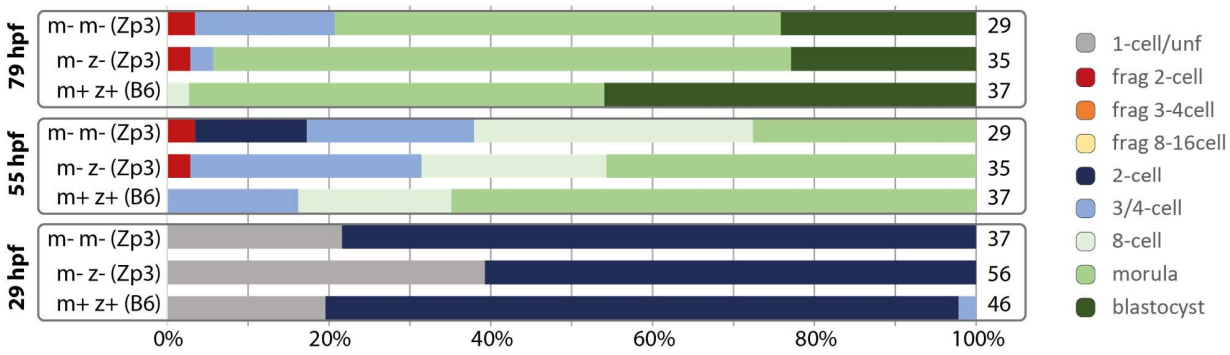



***Figure S 5-1. Developmental progression of Eed Zp3 m-z- and m-m-***

*Comparative time course developmental progression of normally fertilized versus parthenogenic Eed knock-out embryos derived from Eed $^{fl/fl}$ Zp3-Cre females and indicated as m-z- (Zp3) and m-m- (Zp3) respectively. Natural mating in the case of m-z- (Zp3) was done with Eed $^{fl/fl}$ Prm1-cre males. Control m+z+ embryos were isolated from C57BL/6 females mated with C57BL/6 males.*

---

***Figure S 5-2 Conditional knock-out of Eed in the female and male germlines***

***A, B*** *– Expression of core Polycomb genes during oogenesis (A) and preimplantation embryogenesis (B). Data for oogenesis expression is from (Pan et al., 2005), data in (B) is from single-cell qPCR profiling of pre-implantation embryos (Chapter 4).*

***C*** *–Selected stages of oogenesis and preimplantation development, highlighting major developmental events. The promoters of Gdf9 and Zp3 were used as cre-drivers for conditional gene deletion in the oocytes of the primordial follicle at postnatal day 3 or in the growing oocyte prior to the completion of the first meiotic division, respectively.*

***D*** *– Schematic representation of the generation of conditional knock-out embryos. A conditional deletion in the female germline by Zp3-cre or Gdf9-iCre results in knock-out oocytes, which can be fertilized with conditional knock-out sperm (deletion by Prm1-cre) to give homozygous knock-out embryos (m-z-), or alternatively the oocytes can be activated by SrCl2 to give rise to parthenogenic m-m- embryos.*

***E****- Single-embryo RT-qPCR analysis for Eed (normalized to housekeeping genes) in Eed fl/fl Gdf9-iCre m-z- and m-m- embryos, as well as the respective m+z+ and m+m+ controls. Each dot represents a single embryo plotted on a logarithmic normalized expression scale. The red line indicates the mean.*

***F*** *–Breeding experiments testing the conditional cre lines for live birth and deletion of Eed fl/fl or Ezh2 fl/fl alleles. Pups from a cre-positive parent carrying a homozygous conditional allele mated with a wild type partner were genotyped four weeks after birth. The deleted allele was transmitted via the respective germline, depending on the cre-driver, which is indicated above the bars. The number of genotyped pups is given below the bars and the number of parents is given in brackets.*

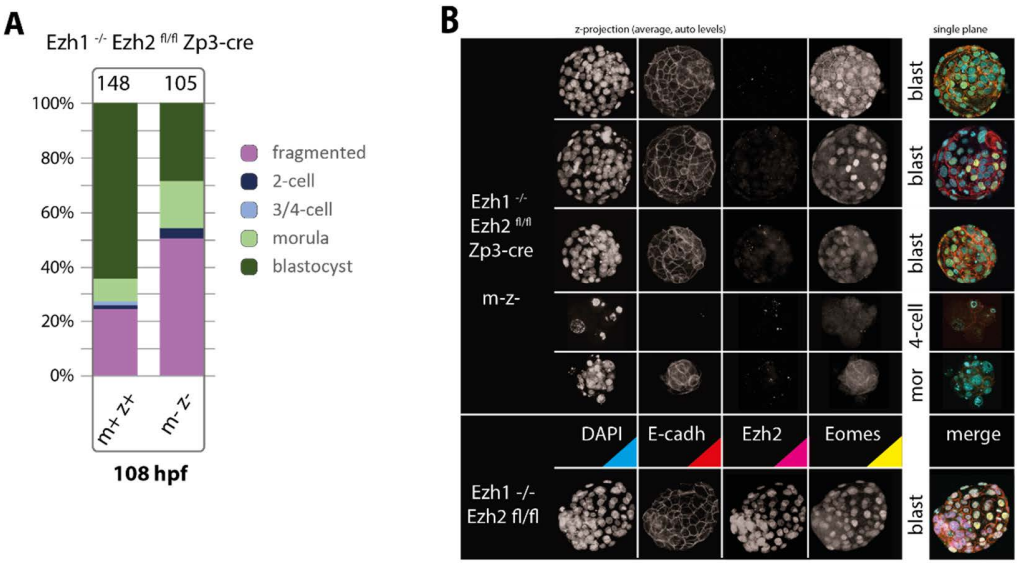



***Figure S 5-3. Characterization of* Ezh1/Ezh2 *double knock-out embryos***

***A** – Developmental progression of Ezh1/Ezh2 double knock-out embryos, compared to Ezh1 single knock-out embryos, scored at 108 hpf. Double knock-out embryos were generated by fertilizing Ezh1 -/- Ezh2 fl/fl Zp3-cre oocytes with Ezh1 -/- Ezh2 fl/fl Prm1-cre sperm. Total number of analysed embryos is given above the graph.*

***B** – Immunofluorescent analysis of Ezh1/Ezh2 double knock-out m-z- embryos, fixed at 100 hpf and stained for DNA, E-cadherin, Ezh2 and Eomes. The m-z- embryos display high levels of heterogeneity in terms of morphology and expression of the trophectoderm marker Eomes. The top three rows represent three different blastocyst embryos, below them is an arrested 4-cell embryo and the fifth embryo is a fragmented morula. As a reference, the lowest panel shows a representative control Ezh1 -/- embryo.*

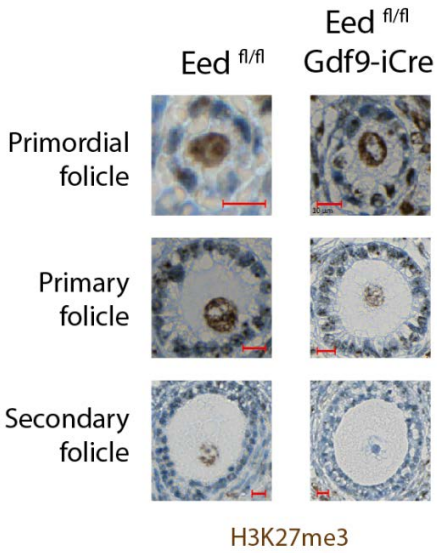



***Figure S 5-4 Histological analysis of ovaries***

*Representative images of primordial, primary and secondary follicles, stained for H3K27me3.*

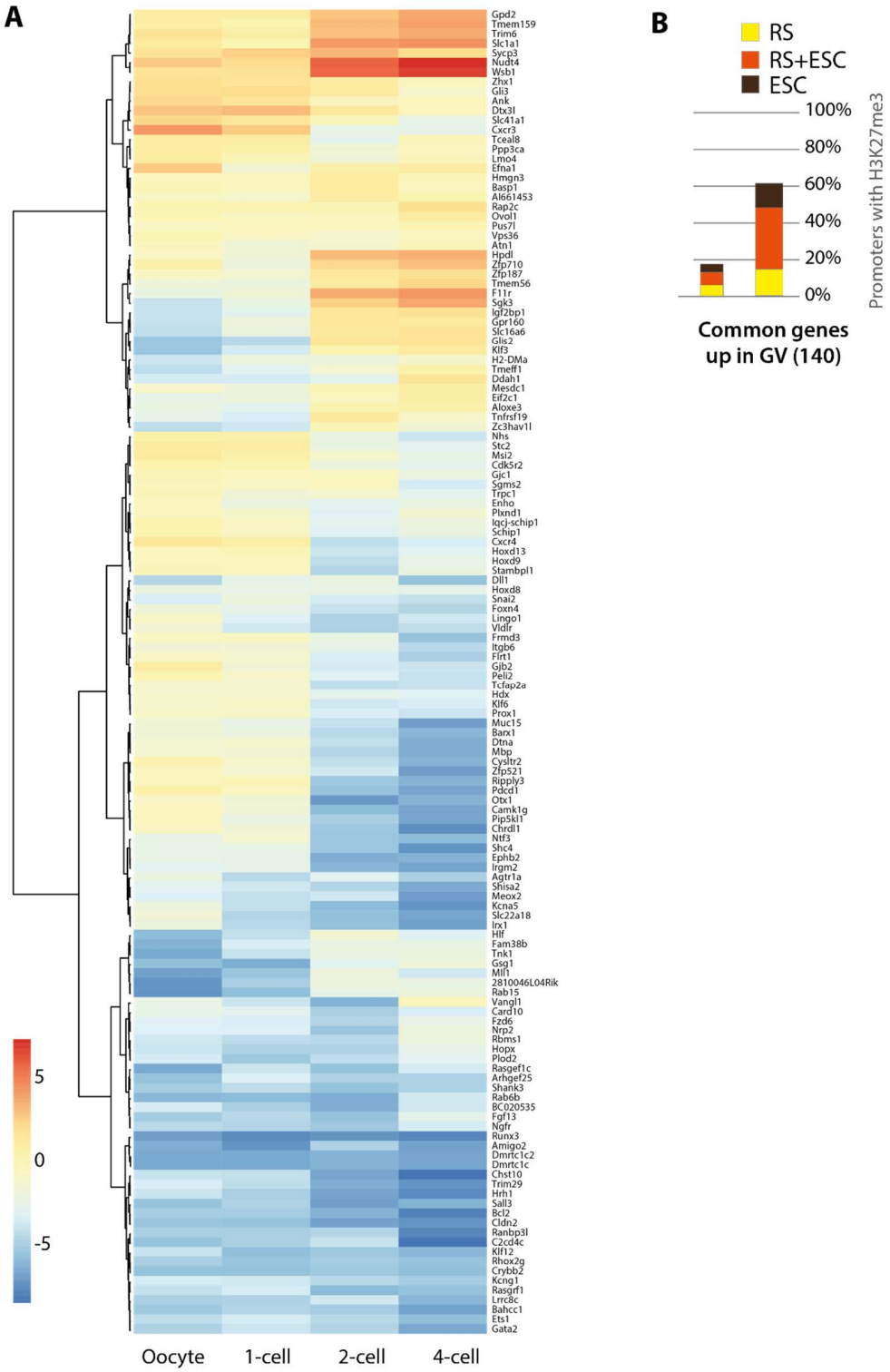



***Figure S 5-5. Analysis of genes commonly upregulated in Polycomb-deficient GV oocytes***

***A*** *– Clustered heatmap showing the expression in wild type oocyte, 1-cell, 2-cell and 4-cell embryos for the 140 genes commonly upregulated in Eed $^{fl/fl}$ Gdf9-iCre, Eed $^{fl/fl}$ Zp3-cre and Ring1 $^{-/-}$ Rnf2 $^{fl/fl}$ Zp3-cre GV oocytes versus the respective control. The expression profiles are from the Park et al. dataset (Park et al., 2013a).*

***B*** *- Proportion of the 140 genes having H3K27me3-positive promoters in RS and ESC.*

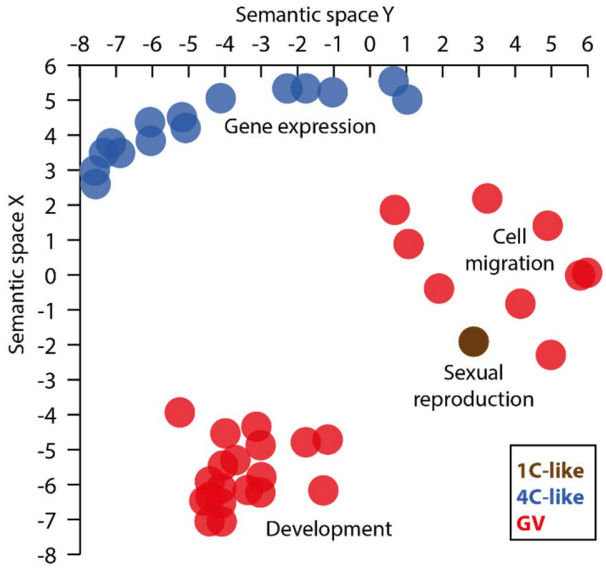



***Figure S 5-6. Gene ontology analysis of genes upregulated in Eed deficient oocytes and embryos***

*Gene ontology (GO) enrichment analysis was performed with the three sets of genes upregulated in Eed $^{fl/fl}$ Gdf9-iCre GV oocytes, 1C-like Eed knock-out 2-cell embryos and 4C-like Eed knock-out embryos versus the respective controls. GO term enrichment analysis was performed relative to the respective set of genes expressed in wild type GV oocytes or 2-cell embryos ((Park et al., 2013a). The list of enriched GO terms was processed with REVIGO in order to identify similar GO terms. The graph shows the non-redundant GO terms enriched in the three sets of upregulated genes plotted such that similar GO terms cluster together. The colours correspond to the set of genes, red is GO terms enriched in the GV set, brown corresponds to the 1C-like set and blue refers to the 4C-like upregulated genes set.*

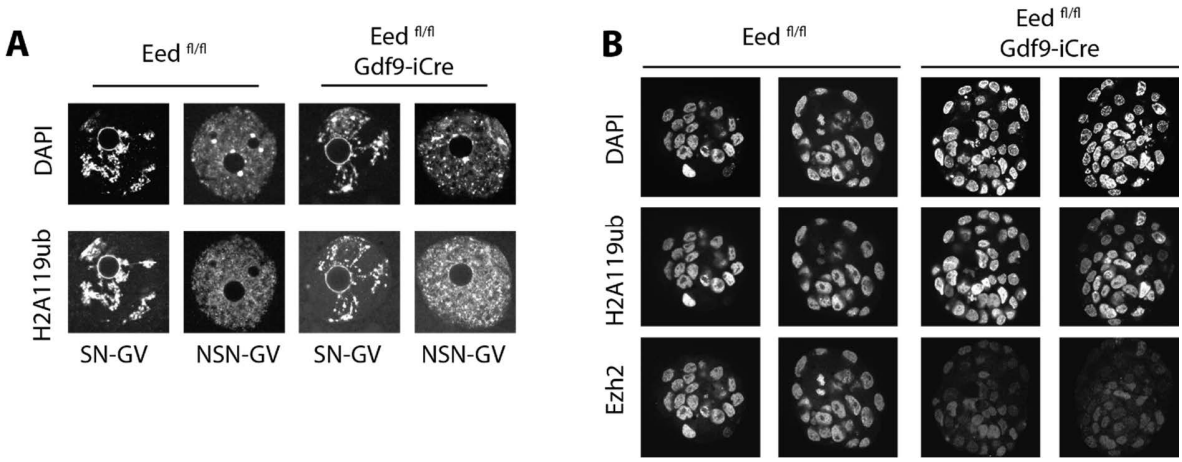



***Figure S 5-7. H2AK119ub in Eed deficient oocytes and embryos***

***A** – Immunofluorescent analysis of GV oocytes stained for H2AK119ub. The figure shows representative images for SN-GV and NSN-GV oocytes from Eed $^{fl/fl}$ and Eed $^{fl/fl}$ Gdf9-iCre females.*

***B** – Immunofluorescent analysis of blastocyst embryos stained for H2AK119ub and Ezh2. Representative images show two Eed $^{fl/fl}$ embryos and two m-z- Eed $^{fl/fl}$ Gdf9-iCre embryos.*

### *Table 5.2. GO term analysis of genes upregulated in Eed-deficient oocytes and 2-cell embryos*

*Gene ontology enrichment analysis of genes upregulated in Eed $^{fl/fl}$ Gdf9-iCre GV oocytes and in Eed $^{fl/fl}$ Gdf9-iCre m-z- 2-cell embryos (1C-like and 4C-like, see text for details). GO term enrichment analysis was performed as described in the Materials and methods section (5.4.5).The table lists the GO term ID and description, the semantic values calculated by REVIGO and used in Figure S 5-6, the enrichment factor (relative to the reference set of all genes expressed under normal conditions in the respective stage), as well as the adjusted P-value for the GO term enrichment. The "DE group" column indicates from the set of differentially expressed genes (GV corresponds to genes upregulated in in Eed $^{fl/fl}$ Gdf9-iCre GV oocytes, 1C-like and 4C-like correspond to genes upregulated in Eed $^{fl/fl}$ Gdf9-iCre m-z- 2-cell embryos).*

| GOterm_ID | GO Term | DE group | SemanticX | SemanticY | Enrichment | AdjPVal |
|---|---|---|---|---|---|---|
| GO:0021783 | preganglionic parasympathetic nervous system development | GV | -4.343 | -3.133 | 15 | 0.00696 |
| GO:0048483 | autonomic nervous system development | GV | -4.873 | -3.016 | 11 | 0.00956 |
| GO:0021545 | cranial nerve development | GV | -4.535 | -3.987 | 11 | 0.00696 |
| GO:0021675 | nerve development | GV | -6.104 | -4.103 | 10 | 0.00377 |
| GO:0060021 | palate development | GV | -4.789 | -1.771 | 6.8 | 0.00536 |
| GO:0001708 | cell fate specification | GV | -4.711 | -1.167 | 11.33333 | 0.00017 |
| GO:0048706 | embryonic skeletal system development | GV | -5.299 | -3.71 | 7.4 | 0.00148 |
| GO:0001763 | morphogenesis of a branching structure | GV | -5.455 | -4.054 | 5.6 | 0.00023 |
| GO:0048839 | inner ear development | GV | -6.347 | -4.401 | 7.444444 | 3.72E-08 |
| GO:0001822 | kidney development | GV | -5.91 | -4.415 | 4.454545 | 0.00941 |
| GO:0045165 | cell fate commitment | GV | -6.166 | -1.288 | 5.6 | 0.00012 |
| GO:0001655 | urogenital system development | GV | -5.788 | -2.996 | 4.785714 | 8.42E-05 |
| GO:0003002 | regionalization | GV | -7.055 | -4.089 | 3.75 | 0.00847 |
| GO:0007389 | pattern specification process | GV | -7.048 | -4.422 | 4.285714 | 3.35E-06 |
| GO:0001501 | skeletal system development | GV | -6.53 | -4.108 | 3.526316 | 0.0077 |
| GO:0048732 | gland development | GV | -6.153 | -3.37 | 3.944444 | 0.00092 |
| GO:0061458 | reproductive system development | GV | -6.227 | -3.024 | 3.761905 | 0.00051 |
| GO:0016477 | cell migration | GV | -0.014 | 5.79 | 2.48 | 0.00272 |
| GO:0072358 | cardiovascular system development | GV | -6.465 | -4.567 | 2.953488 | 3.07E-05 |
| GO:0051674 | localization of cell | GV | 0.06 | 5.983 | 2.351852 | 0.0052 |
| GO:0040011 | locomotion | GV | 1.861 | 0.682 | 2.31746 | 0.00152 |
| GO:0006928 | cellular component movement | GV | 0.886 | 1.062 | 2.109589 | 0.00828 |
| GO:0051239 | regulation of multicellular organismal process | GV | -0.823 | 4.144 | 1.925234 | 0.00208 |
| GO:0032502 | developmental process | GV | -2.288 | 4.983 | 1.671937 | 1.34E-06 |
| GO:0044707 | single-multicellular organism process | GV | -3.935 | -5.243 | 1.718045 | 1.83E-08 |
| GO:0032501 | multicellular organismal process | GV | 1.409 | 4.896 | 1.682482 | 7.42E-08 |
| GO:0019953 | sexual reproduction | 1C-like | -1.91 | 2.857 | 3.192212 | 0.00535 |
| GO:0006414 | translational elongation | 4C-like | 5.019 | 1.029 | 10 | 0.00097 |
| GO:0016072 | rRNA metabolic process | 4C-like | 5.645 | 0.649 | 4.9 | 0.00049 |
| GO:0022613 | ribonucleoprotein complex biogenesis | 4C-like | 2.188 | 3.231 | 4.045455 | 5.01E-07 |
| GO:0034660 | ncRNA metabolic process | 4C-like | 5.227 | -1.032 | 3.708333 | 6.67E-06 |

| GOterm_ID | GO Term | DE group | SemanticX | SemanticY | Enrichment | AdjPVal |
|---|---|---|---|---|---|---|
| GO:0006412 | translation | 4C-like | 5.336 | -2.279 | 4.027778 | 2.42E-12 |
| GO:0006396 | RNA processing | 4C-like | 5.338 | -1.768 | 2.816327 | 1.19E-06 |
| GO:0010467 | gene expression | 4C-like | 5.056 | -4.117 | 1.679856 | 1.25E-09 |
| GO:0009058 | biosynthetic process | 4C-like | 2.6 | -7.563 | 1.492908 | 0.00011 |
| GO:0046483 | heterocycle metabolic process | 4C-like | 4.513 | -5.184 | 1.450311 | 9.57E-05 |
| GO:0006725 | cellular aromatic compound metabolic process | 4C-like | 4.206 | -5.083 | 1.44582 | 9.95E-05 |
| GO:1901360 | organic cyclic compound metabolic process | 4C-like | 3.837 | -6.044 | 1.436364 | 0.00014 |
| GO:0006807 | nitrogen compound metabolic process | 4C-like | 2.984 | -7.995 | 1.373913 | 0.0024 |
| GO:0043170 | macromolecule metabolic process | 4C-like | 4.372 | -6.066 | 1.397229 | 8.96E-07 |
| GO:0044237 | cellular metabolic process | 4C-like | 3.49 | -7.334 | 1.321637 | 3.45E-06 |
| GO:0044238 | primary metabolic process | 4C-like | 3.763 | -7.14 | 1.287671 | 0.00015 |
| GO:0071704 | organic substance metabolic process | 4C-like | 3.491 | -6.89 | 1.271698 | 0.0002 |
| GO:0008152 | metabolic process | 4C-like | -0.39 | 1.901 | 1.248227 | 0.00036 |

# Chapter 6. Manuscript in preparation: Interplay between PRC1 and PRC2

Peter Nestorov*,†, Julia Hacker*,†, Zichuan Liu*, Antoine H.F.M. Peters*,†,2

*Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland †Faculty of Sciences, University of Basel, Basel, Switzerland [1]Equal contribution.

[2]Corresponding author: e-mail address: antoine.peters@fmi.ch

## 6.1 Introduction

The canonical Polycomb interaction model suggests a hierarchical link between PRC1 and PRC2, where PRC2 is upstream and its enzymatic activity is required for the proper targeting of PRC1 to the promoters of Polycomb-regulated genes, which ultimately leads to transcriptional silencing. However, recent work by several research groups has added new links and variants to the world of Polycomb (Comet and Helin, 2014; Schwartz and Pirrotta, 2013), most notably a reversed hierarchical model where PRC2 is recruited by PRC1 (Blackledge et al., 2014). The interplay between PcG proteins has mainly been studied biochemically in *in vitro* systems. The relevance of these findings needs yet to be shown *in vivo* and therefore I investigated the connection between PRC1 and PRC2 in mouse preimplantation embryos. I used a conditional *Ezh2/Rnf2* double knock-out (DKO) model corresponding to the *Ezh1/Ezh2* DKO and *Ring1/Rnf2* DKO models that have already been described ((Posfai et al., 2012) and Chapter 5).

## 6.2 Results

### 6.2.1 Phenotypic analysis of Ezh2/Rnf2-deficient preimplantation embryos

In order to study the compound effect of PRC1 and PRC2 in preimplantation development, I used conditional *Ezh2* and *Rnf2* double knock-out mice. The presence of *Ring1*, the homolog of *Rnf2* leads to a rescue of the maternal PRC1 effect (Posfai et al., 2012). The conditional alleles for *Ezh2* and *Rnf2* were deleted early during oogenesis with the *Zp3-cre*, and to create double deficient zygotes, I used males carrying the conditional alleles and expressing *Prm1-cre* late in

spermatogenesis (*Ezh2* $^{fl/fl}$ *Rnf2* $^{fl/fl}$ in combination with *Zp3-cre* and *Prm1-cre*, for the female and male germline deletion respectively). The developmental progression of *Ezh2/Rnf2* m-z- zygotes revealed a severe block of development around the 4-cell stage (Figure 6.1). In particular, around 60% of the m-z-embryos died at the 3-/4-cell stage, another 20% remained arrested in the 2-/4-cell stage and only 20% reached the compacted morula stage. The corresponding no-cre control group showed normal pre-implantation development for more than 50% of the embryos. However, some 40% of *Ezh2/Rnf2* control embryos died before morula compaction at the 8-/16-cell stage. I speculate that the reason might be the heterozygosity of the control group, i.e. the embryos carry a non-deleted maternal and a deleted paternal allele for each *Ezh2* or *Rnf2*. The deleted allele was intentionally transmitted through the male germline, becausePrm1-cre has a reduced efficiency of around 80% (Chapter 5). By using a heterozygous allele constellation in combination with the Prm1-cre, the theoretical efficiency for homozygous m-z- knock-out is 90%. This corresponds to the 10% *Ezh2/Rnf2* m-z- embryos that make it to the morula stage. Finally, an external control group of pure C57/BL6 embryos was included in the experimental setting and developed to blastocyst at a 90% rate, which indicates that the culture conditions and the procedure were optimal.

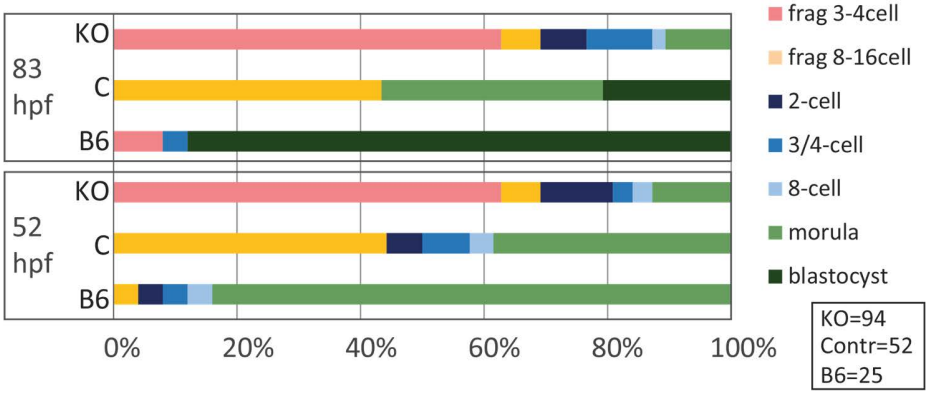



***Figure 6.1 Developmental progression Ezh2/Rnf2 m-z-***

*Developmental progression of Ezh2/Rnf2 double knock-out embryos (m-z-) derived from Ezh2* $^{fl/fl}$ *Rnf2* $^{fl/fl}$ *Zp3-Cre or cre-negative mothers, indicated as* ***KO*** *and* ***C*** *respectively and fertilized in both cases by natural mating with Ezh2* $^{fl/fl}$ *Rnf2* $^{fl/fl}$ *Prm1-Cre males. As an external control, I used wild type C57BL/6 embryos (obtained from C57BL/6 females mated with C57BL/6 males), indicated as* ***B6****.Development was scored at 52 and 83 hours post fertilization (64 and 95 hours post hCG injection). The number of analysed embryos is given in the box on the right.*

## 6.2.2 Characterization of the maternal effect of Ezh2/Rnf2 depletion

### 6.2.2.1 Validation of the model by single-oocyte qPCR

The severe developmental defects observed in preimplantation embryos upon deletion of *Ezh2* and *Rnf2* in the female germline could be due to the maternal function of Polycomb proteins, or in combination with the paternal deletion, could also be due to an embryonic PcG function. In order to describe the maternal effect, I characterized fully grown GV oocytes depleted for *Ezh2* and *Rnf2*. First, I validated the knock-out model by measuring the expression levels of core PcG genes in single GV oocytes (Figure 6.2 and paragraph 6.4.3). I did not detect a signal for *Rnf2* in any of the 27 analysed *Ezh2/Rnf2* DKO oocytes, indicating complete and 100%-efficient deletion for *Rnf2*. For Ezh2, I could detect a very weak signal in six of the 27 DKO oocytes (around 500-fold lower than the wild type levels of *Ezh2* in this stage), while the remaining 21 samples were negative. From the analysed PcG genes, *Eed* and to some extend also *Jarid2* were upregulated in the oocytes that had no *Rnf2* and *Ezh2* expression. The six oocytes with residual *Ezh2* levels had normal levels for *Ezh1, Eed, Suz12, Jarid2* and *Bmi1* and reduced levels for *Ring1a*.

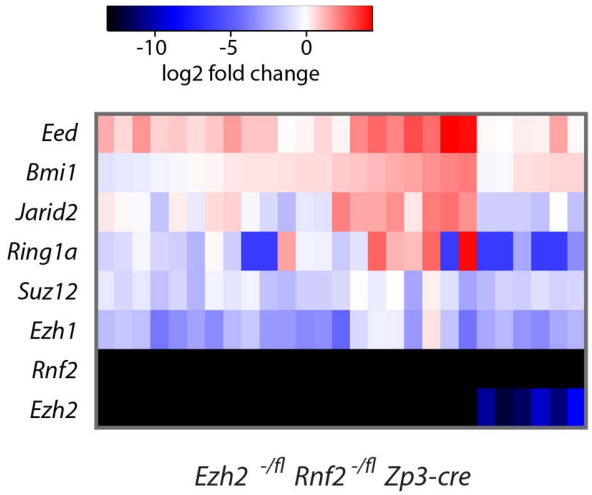



***Figure 6.2 Changes in expression of Polycomb genes in Ezh2/Rnf2 DKO oocytes***

*Differential expression of PcG genes in Ezh2/Rnf2 DKO GV oocytes, relative to C57BL/6 wild type oocytes. Colour range indicates level of misregulation – black: not detected in DKO, blue: strongly downregulated; with: no change; red: strongly upregulated. RT-qPCR data was normalized as described in Materials and methods (6.4.3).*

### 6.2.2.2 Immunofluorescent analysis of H3K27me3 levels and active transcription

I also analysed the transcriptional shutdown and the levels of H3K27me3 by immunofluorescent staining in GV oocytes (Figure 6.3). The global H3K27me3 levels were decreased in the *Ezh2/Rnf2* DKO oocytes compared to control oocytes (Figure 6.3A-D). However, the reduction in global H3K27me3 was weaker than in the *Ezh1/Ezh2* DKO. Furthermore, preliminary results indicate that the PRC1-dependent H2AK119ub chromatin mark is not dramatically changed in the knock-out versus the control condition (experiment performed by Julia Hacker, data not shown). These findings suggest that the PRC1- and PRC2-mediated chromatin modifications are still maintained in the absence of the main enzymatic components Rnf2 and Ezh2 respectively, supposedly by the activity of Ezh1 and Ring1, which is consistent with previous studies (Posfai et al., 2012; Puschendorf et al., 2008).

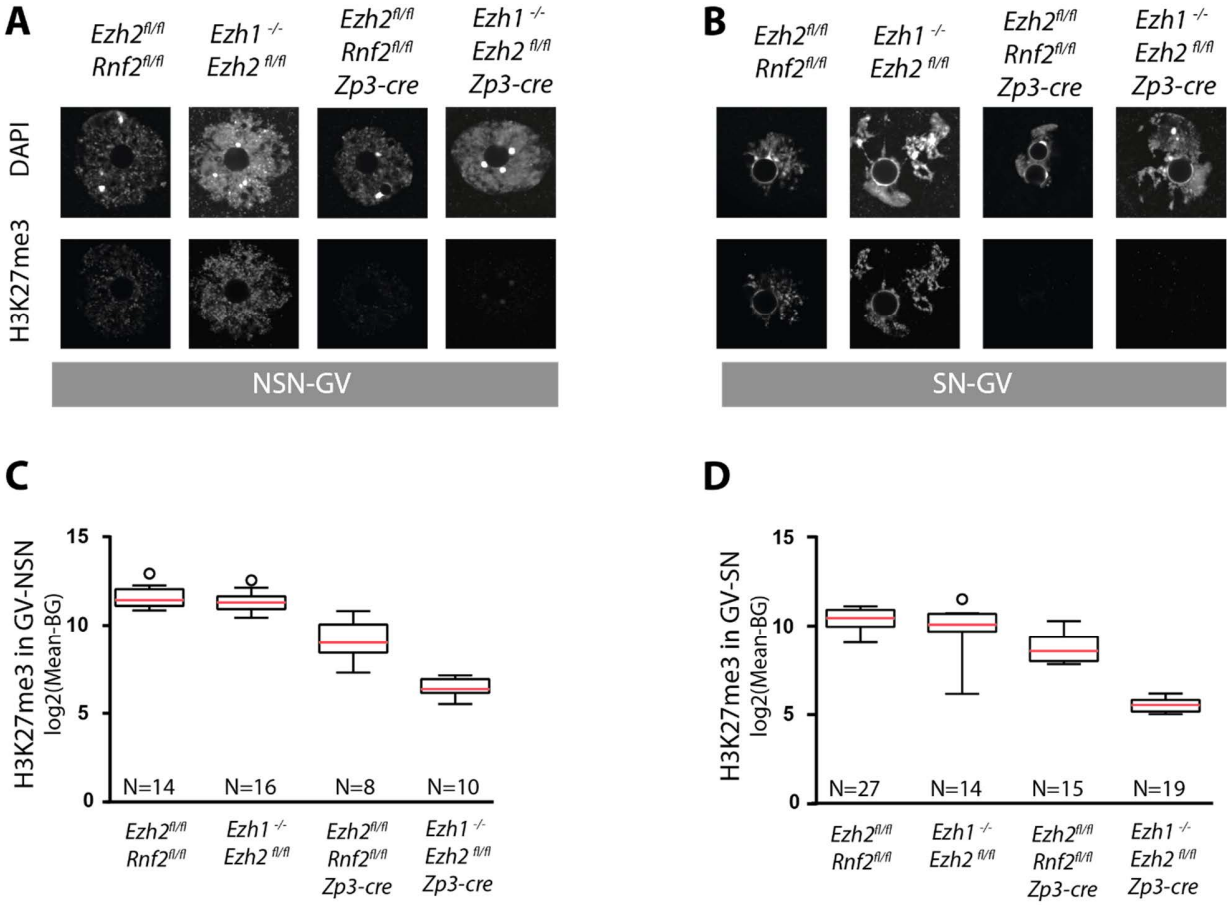



**Figure 6-3.**

***Figure 6.3. Immunofluorescent analysis of H3K27me3 levels in GV oocytes***

*Immunofluorescent staining and quantification of signal intensity for H3K27me3 in NSN-GV oocytes (A, C) and SN-GV oocytes (B, D) for the following genotypes: Ezh2 $^{fl/fl}$ Rnf2 $^{fl/fl}$ (control oocytes), Ezh1 $^{-/-}$ Ezh2 $^{fl/fl}$ (Ezh1 single knock-out), Ezh2 $^{fl/fl}$ Rnf2 $^{fl/fl}$ Zp3-Cre (Ezh2/Rnf2 double knock-out), and Ezh1 $^{-/-}$ Ezh2 $^{fl/fl}$ Zp3-cre (Ezh1/Ezh2 double knock-out). Upper panels in A and B show the DNA stained with DAPI, lower panels show the H3K27me3 signal under constant imaging conditions. Quantification was done by measuring the mean intensity for H3K27me3 in the nucleus (nuclear area determined according to the DAPI staining), normalization was done 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 the NSN-GV (C) and SN-GV (D). The number of analysed oocytes per genotype and stage is given on the graphs.*

### 6.2.2.3 Transcriptional profiling in GV oocytes

Since PcG proteins act mainly as transcriptional regulators, I sought to analyse the effect on the accumulated RNA in GV oocytes deficient for *Ezh2* and *Rnf2*. The RNA-seq experiment revealed very few genes that differed significantly between the control and knock-out condition (*Ezh2 $^{fl/fl}$ Rnf2 $^{fl/fl}$* and *Ezh2 $^{fl/fl}$ Rnf2 $^{fl/fl}$ Zp3-Cre* respectively). There were in total 56 misregulated genes (using the following cut-offs: FDR<0.05, log2FC>1, log2CPM>0), of which 21 were downregulated and 35 upregulated (Figure 6.4 and Table 6.1). Among the top downregulated genes were the conditionally targeted Ezh2 and Rnf2, which still show some expression, however only a truncated transcript could be detected (i.e. there were no reads mapping to the targeted exons, Figure 6.5). The *Zp3-cre* transgene contains part of the *Mt1* gene, which is otherwise not expressed during oogenesis and thus makes *Mt1* the strongest upregulated transcript in *Zp3-cre* expressing oocytes (de Vries et al., 2000). The mild transcriptional changes are visible both on the MA-plot (Figure 6.4A) and in the PCA plot (Figure 6.4B), where the knock-out samples intermingle with the control samples. Figure 6.4C shows the intersection between the upregulated genes in oocytes upon depletion of PRC1 (*Ring1/Rnf2* DKO), PRC2 (*Ezh1/Ezh2* DKO) or *Ezh2* and *Rnf2*. There strongest changes are observed in PRC1 DKO with more than 2000 upregulated transcripts, while PRC2 DKO leads to overexpression of 177 genes. Compared to this numbers, the Polycomb-driven derepression observed in *Ezh2/Rnf2* knock-out oocytes is negligible. The eight common genes significantly upregulated in all Polycomb knock-out oocytes are *Sall3, Tceal8, Bmp7, Tnc, Tnfrsf19, Chrdl1, Nhs* and *Mt1*, suggesting that these genes are core targets of *PRC1* and *PRC2* in mouse oocytes (see also Chapter 5).

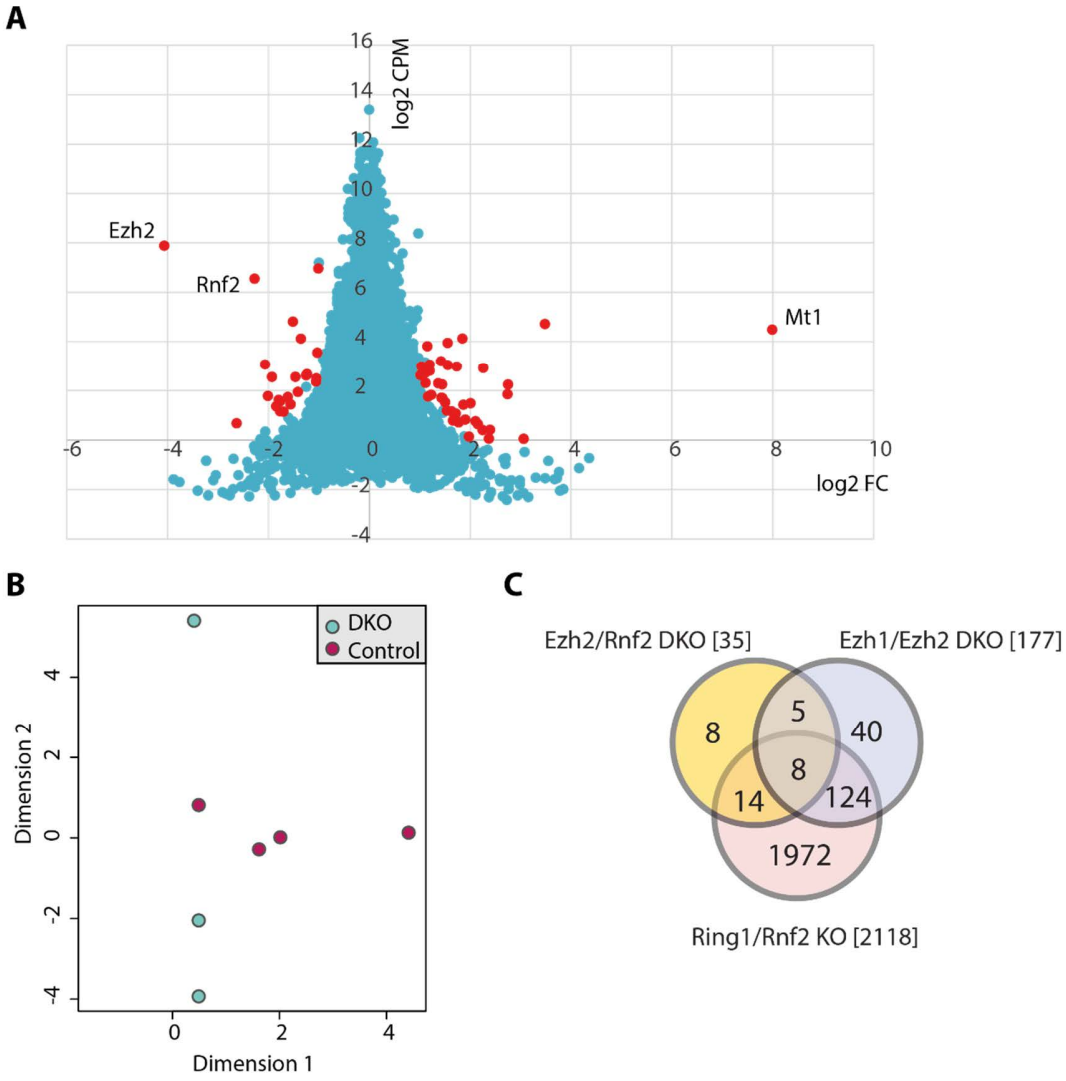



***Figure 6.4 Transcriptional profiling of GV oocytes***

***A*** *– MA plot (logarithmic expression value, log2CPM, plotted against the logarithmic fold-change value, log2FC) displaying the differential expression in Ezh2/Rnf2 DKO oocytes versus control oocytes (Ezh2 $^{fl/fl}$ Rnf2 $^{fl/fl}$ Zp3-Cre and Ezh2 $^{fl/fl}$ Rnf2 $^{fl/fl}$ respectively). Statistically significant differential expression (FDR<0.05, log2FC>1, log2CPM>0) is highlighted in red.*

***B*** *– Principle component analysis (PCA) for the four control samples (Ezh2 $^{fl/fl}$ Rnf2 $^{fl/fl}$) and the three Ezh2/Rnf2 DKO samples (Ezh2 $^{fl/fl}$ Rnf2 $^{fl/fl}$ Zp3-Cre) in the RNA-seq experiment.*

***C*** *– Overlap between the significantly upregulated genes in Ezh2/Rnf2 DKO oocytes, (Ezh2 $^{fl/fl}$ Rnf2 $^{fl/fl}$ Zp3-Cre), Ezh1/Ezh2 DKO (Ezh1 $^{-/-}$ Ezh2 $^{fl/fl}$ Zp3-Cre) and Ring1/Rnf2 DKO (Ring1$^{-/-}$ Rnf2 $^{fl/fl}$ Zp3-Cre). Total number of upregulated genes for each genotype is given in brackets.*

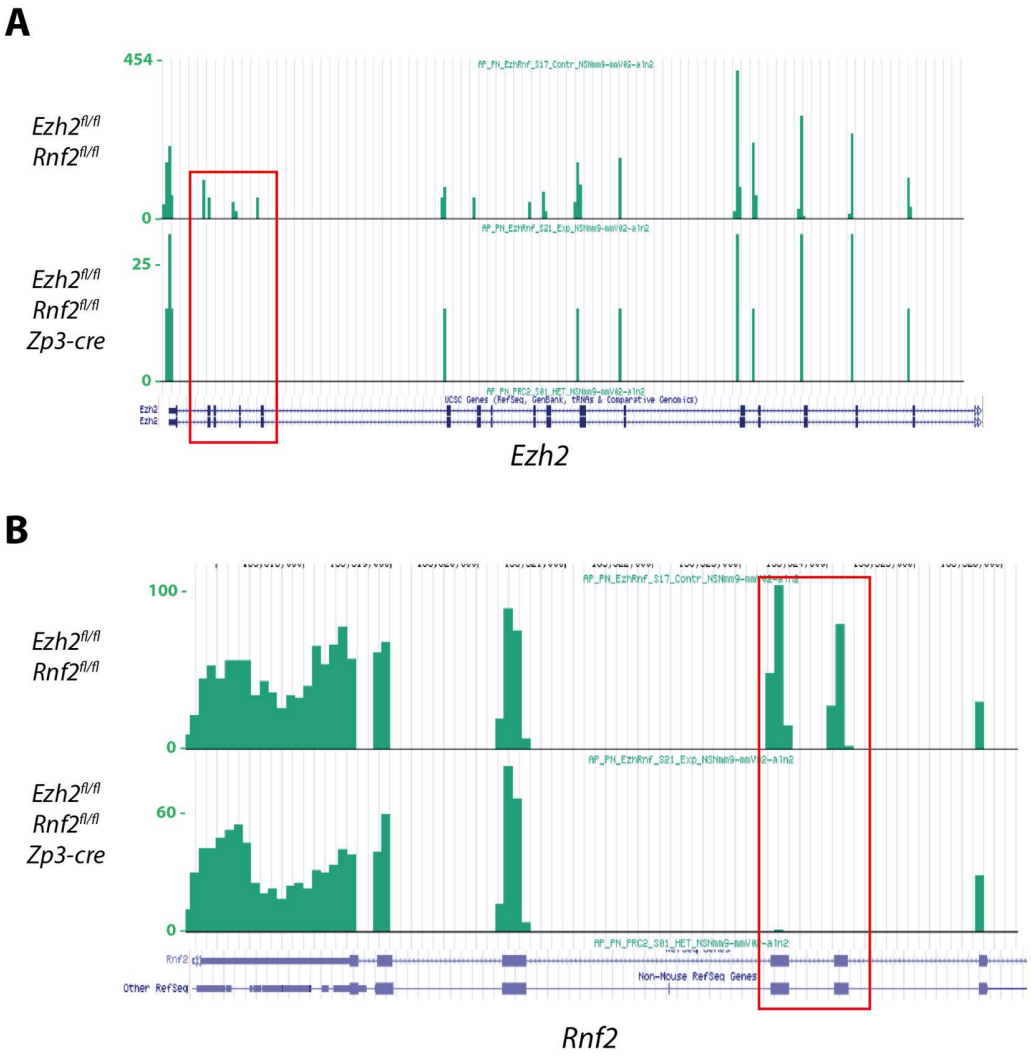



***Figure 6.5. RNA-seq reads mapping to Ezh2 and Rnf2***

*Snapshot from the Genome Browser (UCSC) showing the genomic loci for Ezh2 (**A**) and Rnf2 (**B**) and the reads mapping to the respective region in control GV oocytes (Ezh2 $^{fl/fl}$ Rnf2 $^{fl/fl}$) and in Ezh2/Rnf2 DKO oocytes (Ezh2 $^{fl/fl}$ Rnf2 $^{fl/fl}$ Zp3-Cre). Note that the scales between control and knock-out are different, indicating a global downregulation of Ezh2 and Rnf2 upon deletion in early oocytes. The deleted exons are highlighted with red boxes.*

Next, I performed whole-data enrichment analysis in an attempt to identify functionally related genes that show a common trend of expressional change (Figure 6.6). As a first approach, I used the gene set enrichment analysis (GSEA), which highlighted two specific groups of genes that were underrepresented in the DKO over the control condition. The only two gene sets that were downregulated in the DKO oocytes were the RNA Pol I promoter opening set of histone genes

and the ribosome biosynthesis gene set. In fact, these two sets are functionally related, making the discovery even more significant. An alternative approach looking for the activation of specific pathways (Ingenuity, Qiagen) highlighted the Cell Death pathway as aberrantly activated in *Ezh2/Rnf2* DKO oocytes.

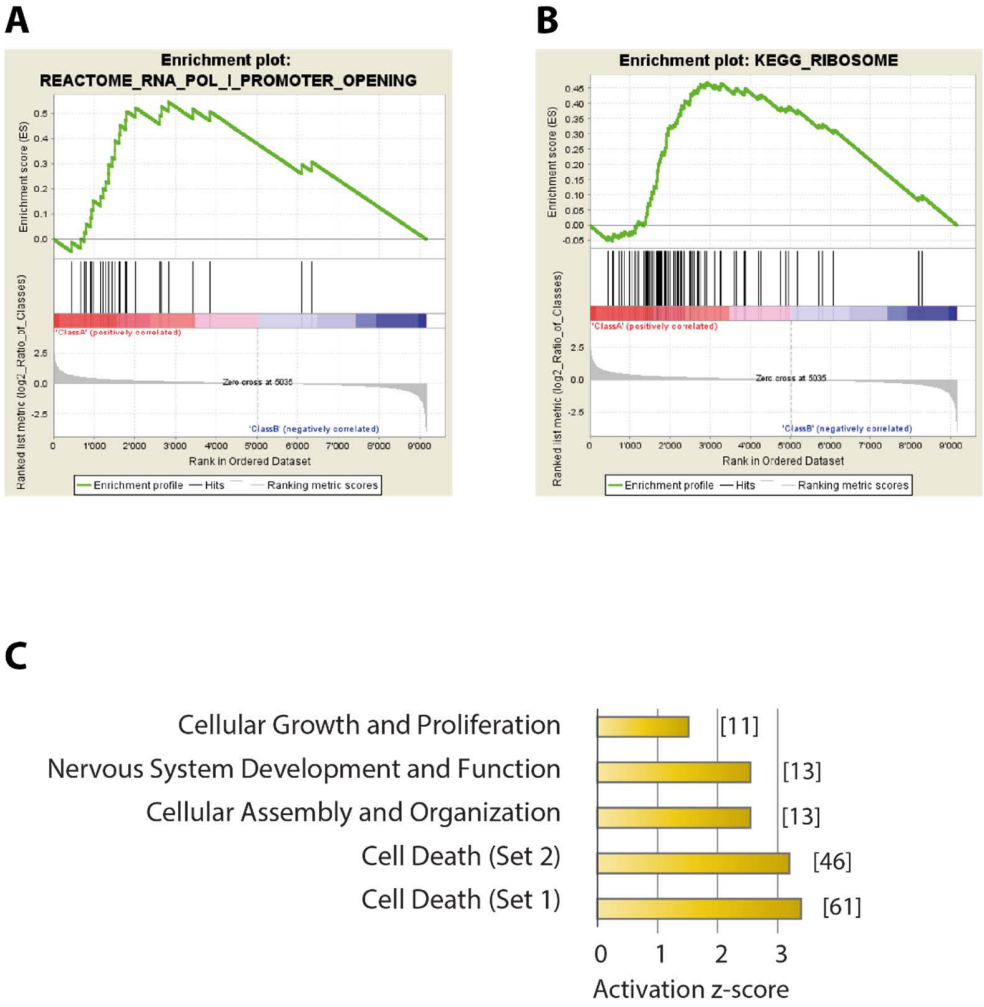



***Figure 6.6 Gene expression enrichment analyses***

***A, B*** *–Gene set enrichment analysis (GSEA) identified two sets of genes that were depleted in Ezh2/Rnf2 DKO (Ezh2 $^{fl/fl}$ Rnf2 $^{fl/f}$ Zp3-cre)oocytes compared to control oocytes (Ezh2 $^{fl/fl}$ Rnf2 $^{fl/fl}$).The plot represents the enrichment in a ranked order from downregulated in DKO (red) to upregulated in DKO (blue). The genes comprising a given gene set are marked as lines and an enrichment profile is displayed as a green line. The gene set in (**A**) comprises mostly the histone genes, while the gene set in (**B**) features multiple genes involved in ribosome biosynthesis.*

***C*** *– Pathway analysis based on the differential expression between Ezh2/Rnf2 DKO and control oocytes. The bar chart shows the pathway that are activated in DKO versus control oocytes. The number of genes in each pathway is given in brackets. The z-score indicates the significance of the finding (higher z-score corresponds to higher significance).*

## 6.3 Discussion

The current work highlights the importance of the interplay between PRC1 and PRC2. Conditional deletion of either Ezh2 or Rnf2 using the maternal Zp3-cre and the paternal Prm1-cre does not disturb preimplantation development (Puschendorf et al., 2008). However, conditional deletion of both genes together causes major defects in DKO embryos around the 8-cell stage. This happens in the presence of the homologous members of PRC1 and PRC2, Ring1 and Ezh1 respectively, which have been shown to back up the function of the main players Rnf2 and Ezh2 in their absence ((Posfai et al., 2012) and Chapter 5). Therefore, the Polycomb activity has not been completely lost in the Ezh2/Rnf2 DKO embryos, which is evident also from the persisting global H3K27me3 and H2AK119ub modifications in mutant oocytes and embryos. Yet, Ezh2 and Rnf2 play a more important role in embryogenesis and development than their counterparts Ezh1 and Ring1, which is evident from the lethality around gastrulation of Rnf2 and Ezh2 mutant embryos (O'Carroll et al., 2001; Voncken et al., 2003) on one hand, and the viability and fertility of Ezh1 and Ring1 single mutant mice on the other (consistently observed in different projects in the group of Antoine Peters, including my own work, and also reported by (Ezhkova et al., 2011) and (Lorente et al., 2000)). Given these observations, it is very likely that the Ezh2/Rnf2 deletion causes reduced Polycomb activity and that Ezh2 and Rnf2 are important for the interaction between the two complexes in early embryos, in particular around the 4-/8-cell stage when the phenotype becomes visible. Interestingly, a similar observation regarding the compound effect of a PRC1 and PRC2 deletion has been made in embryonic stem cells, where single Eed KO or Rnf2 KO do not disturb stemness and differentiation, but a double Eed/Rnf2 knock-out ESC line fails differentiation and has reduced self-renewal potential (Leeb et al., 2010).

Preimplantation development is a complex process that depends initially on the maternally provided factors and later on the coordination and activation of the two parental genomes. Any phenotype observed between the 2-cell and 8-cell stage could be due to the maternal or zygotic effect of a given gene of interest. I did not observe any significant developmental problems in fully grown oocytes depleted for Ezh2 and Rnf2. Furthermore, the transcriptional

profiling identified only a few misregulated genes, which is a strong indication that the observed embryonic phenotype is of zygotic origin. However, previous work on the maternal effect of PRC1 has shown that massive changes of transcription do not affect oogenesis and become apparent only when the protein is produced upon fertilization (Posfai et al., 2012). Therefore, it cannot be completely excluded that some of the few aberrant transcripts in the mutant oocytes cause a developmental defect in embryos. Particularly because, gene enrichment analysis of the RNA-seq data highlighted a few pathways related with cell death, which may be triggered once the message is translated into protein. This is something that needs to be explored further.

## 6.4 Materials and methods

### 6.4.1 Mice

The *Ezh2* and *Rnf2* conditional allele mice were generated as described previously (Puschendorf et al., 2008).

### 6.4.2 Embryo culture, immunofluorescent analysis and RNA-seq

The experimental procedures were described above as follows: embryo culture (4.4.1), immunofluorescent analysis (5.4.3) and transcriptional profiling of GV oocytes (5.4.4). In addition, RNA-seq data was analysed for enrichment of gene sets using the online platform for gene set enrichment analysis (GSEA) by the Broad Institute (Subramanian et al., 2005). Pathway enrichment analysis was performed using the Ingenuity IPA software (Qiagen).

### 6.4.3 Single oocyte RT-qPCR

Single oocytes were processed for RNA amplification and cDNA production as described above (4.4.2). Pre-amplified cDNA samples were assayed by regular qPCR using SYBR® Select Master

Mix (Life technologies, 4472908). The qPCR primers for *Ezh2* were selected from the targeted exons, such that any detected signal would correspond to wild type transcript and not truncated RNA (Ezh2_FP_ex17_2088: GGATGAAGCAGACAGAAGAGGA; Ezh2_RP_ex19_2279: GCAAAGATGCCTATCCTGTGG). Primer sequences for all other genes are given in Appendix C: Fluidigm qPCR primers. The Ct values were converted into expression values by subtracting each Ct value from the assumed detection threshold of 30 (30-Ct), followed by normalization to the reference gene *Hnrnpr* for each sample. Differential fold-change was calculated relative to C57BL/6 control oocytes.

***Table 6.1. Differentially expressed genes in Ezh2/Rnf2 DKO oocytes***

*List of the 56 genes up- or downregulated in Ezh2/Rnf2 DKO versus control oocytes (Ezh2 $^{fl/fl}$ Rnf2 $^{fl/fl}$ Zp3 and Ezh2 $^{fl/fl}$ Rnf2 $^{fl/fl}$ respectively). The following cut-off was used for defining differential expression: FDR<0.05, log2FC>1, log2CPM>0. The list is ordered by ascending logarithmic fold-change values.*

| Symbol | Gene ID | log2FC | log2CPM | FDR |
|---|---|---|---|---|
| *Ezh2* | 14056 | -4.06302 | 7.889919 | 6.23E-87 |
| *Tie1* | 21846 | -2.6288 | 0.673027 | 0.000282 |
| *Rnf2* | 19821 | -2.27452 | 6.554492 | 5.65E-76 |
| *Fbxo16* | 50759 | -2.06543 | 3.066027 | 1.55E-11 |
| *Rbm34* | 52202 | -2.00799 | 1.783047 | 0.000455 |
| *Slc7a8* | 50934 | -1.92721 | 2.550389 | 4.91E-06 |
| *Adam7* | 11500 | -1.84361 | 1.363816 | 0.002259 |
| *Frmd6* | 319710 | -1.79381 | 1.617685 | 0.008761 |
| *Pak3* | 18481 | -1.76403 | 1.566357 | 0.000482 |
| *Acsm4* | 233801 | -1.69885 | 1.143488 | 0.021106 |
| *Rnase6* | 78416 | -1.61275 | 1.737764 | 0.001217 |
| *Pdlim3* | 53318 | -1.55824 | 1.440116 | 0.010434 |
| *Cttnbp2* | 30785 | -1.51265 | 4.815512 | 5.06E-16 |
| *Hist1h1e* | 50709 | -1.4643 | 2.550863 | 0.001059 |
| *Coro2b* | 235431 | -1.35247 | 4.122052 | 0.000394 |
| *Naip6* | 17952 | -1.24894 | 2.614856 | 0.007631 |
| *Pnp2* | 667034 | -1.23459 | 2.671722 | 0.000556 |
| *Ccdc122* | 108811 | -1.05185 | 2.364682 | 0.021106 |
| *G2e3* | 217558 | -1.04834 | 2.494402 | 0.035561 |
| *Rad51l1* | 19363 | -1.02872 | 3.550745 | 0.00251 |
| *Abcc4* | 239273 | -1.00834 | 6.965405 | 1.72E-15 |
| *Bhlhb9* | 70237 | 3.062378 | 0.049295 | 0.003739 |

| Symbol | Gene ID | log2FC | log2CPM | FDR |
|---|---|---|---|---|
| *Sall3* | 20689 | 2.372389 | 0.056167 | 0.02134 |
| *Ldb2* | 16826 | 1.977623 | 0.137605 | 0.02134 |
| *Gbp8* | 76074 | 2.248963 | 0.397478 | 0.013962 |
| *Cox4i2* | 84682 | 2.395294 | 0.408704 | 0.003739 |
| *Tceal8* | 66684 | 2.155843 | 0.641063 | 0.009803 |
| *Sult1c1* | 20888 | 1.775331 | 0.713868 | 0.009076 |
| *Nfatc1* | 18018 | 1.771954 | 0.726954 | 0.010329 |
| *Fshr* | 14309 | 1.661471 | 0.783061 | 0.027123 |
| *Gm8817* | 667794 | 1.905555 | 0.824881 | 0.005427 |
| *Nsg2* | 18197 | 1.716079 | 1.069922 | 0.002507 |
| *Gng2* | 14702 | 1.629346 | 1.157664 | 0.027123 |
| *Ptprq* | 237523 | 1.539944 | 1.189914 | 0.026433 |
| *Hormad1* | 67981 | 1.86657 | 1.422351 | 0.024198 |
| *Epcam* | 17075 | 2.008286 | 1.486029 | 0.000582 |
| *Pdha2* | 18598 | 1.506706 | 1.524251 | 0.011602 |
| *Sox13* | 20668 | 1.452029 | 1.693641 | 0.012027 |
| *Cspg5* | 29873 | 1.431397 | 1.70709 | 0.004437 |
| *39510* | 381270 | 1.167121 | 1.759521 | 0.014837 |
| *Bmx* | 12169 | 1.228316 | 1.82666 | 0.013658 |
| *Trim6* | 94088 | 2.741698 | 1.855415 | 4.72E-11 |
| *Klf12* | 16597 | 1.444951 | 2.250616 | 7.73E-05 |
| *Traf6* | 22034 | 1.370477 | 2.295981 | 0.0164 |
| *Bmp7* | 12162 | 1.116095 | 2.309193 | 0.026611 |
| *Ccnb1ip1* | 239083 | 1.016393 | 2.625511 | 0.010433 |
| *Ptgs2* | 19225 | 1.199269 | 2.811571 | 0.031862 |
| *Tnc* | 21923 | 2.259941 | 2.922354 | 6.59E-13 |
| *Frk* | 14302 | 1.557352 | 3.036619 | 0.000111 |
| *Chst9* | 71367 | 1.198982 | 3.059496 | 0.002144 |
| *Tnfrsf19* | 29820 | 1.425187 | 3.20139 | 1.04E-05 |
| *Chrdl1* | 83453 | 1.153566 | 3.814625 | 0.000895 |
| *Kcnk18* | 332396 | 1.55611 | 3.950006 | 0.02029 |
| *Nhs* | 195727 | 1.848877 | 4.130598 | 5.06E-16 |
| *Mt1* | 17748 | 7.990697 | 4.493159 | 3.2E-175 |
| *Uggt2* | 66435 | 3.484661 | 4.72283 | 7.99E-84 |

# Chapter 7. General discussion and outlook

## 7.1 Main findings

I have studied the role and dynamics of Polycomb repressive mechanism during preimplantation and germ cell development. Most notably, I discovered a dosage-dependent requirement for maternally inherited H3K27me3 and a role for PRC2 during ZGA, as well as in morula and blastocyst embryos (Chapter 5). Next, I found evidence for the interplay between PRC1 and PRC2 based on a requirement for simultaneous activity of *Ezh2* and *Rnf2* in 2-cell to 8-cell embryos (Chapter 6). I performed a quantitative, single-cell expression profiling describing the preimplantation transcriptional profiles of genes coding for chromatin modifiers, which highlighted a number of maternal and zygotic factors, as well as a few lineage-specific genes (Chapter 4). Finally, I contributed to a project that discovered a dosage-dependent role for PRC1 in the maintenance and sex-specific development of PGCs (Chapter 3).

## 7.2 Polycomb function in the pluripotency life cycle

The idea that preimplantation and germ cell development can be considered as part of one process, which is characterized by the retention of pluripotency was proposed by Harry Leitch and Austin Smith in 2012 (Leitch and Smith, 2013). The epiblast lineage of the E4.5 blastocyst is considered as the ground state of pluripotency, i.e. “naïve pluripotency”. Later, in early post-implantation embryos prior to gastrulation, the epiblast lineage still possesses a reduced pluripotency potential, termed as “primed pluripotency”, referring to the cell fate priming of the late epiblast cells. Around this time in development, between E6.5 and E7.5, the induction of the primordial germ cells takes place. The PGCs are induced by BMP signalling from the extra-embryonic ectoderm that recruits epiblast cells to commit to the germ line. The cell of origin of the PGCs represents a potential break in the cycle of pluripotency, as some authors propose that PGCs are recruited from cells committed to the mesoderm lineage and thus the induction

of PGCs from these cells would be a form of cell fate reprograming (Hayashi and Surani, 2009). However, the differentiated state of precursor PGCs is still under debate (Leitch and Smith, 2013). Nevertheless, it is clear that upon induction, PGCs rapidly and effectively repress somatic gene expression programs and transcriptionally resemble late epiblast cells (Kurimoto et al., 2008; Magnúsdóttir et al., 2013). After induction, PGCs undergo a specification process and migrate to the embryonic gonads, where they remain arrested for the rest of prenatal development. The pluripotency cycle is closed by the maturation of the gametes to oocytes and spermatozoa, followed by fertilization and the formation of the totipotent zygote. The process spans across two generations and involves two major rounds of epigenetic reprograming, one in preimplantation embryos and one in the developing PGCs.

Besides resetting and *de novo* establishment of DNA methylation, there are also dynamic changes in histone methylation, as described above (Chapter 1). Interestingly, two essential factors involved in the induction of PGCs are histone methyltransferases *Prdm1* (*Blimp1*) and *Prdm14*. In respect to Polycomb repression, it seems to be required in multiple instances during the pluripotency cycle - for the maintenance of early PGC fate (Chapter 3), for antagonizing external cues in the female PGCs (Chapter 3), for regulating the maternal RNA pool in oocytes and proper zygotic genome activation (Chapter 5) (Posfai et al., 2012), as well as during the zygotic control of preimplantation development (Chapter 5, Chapter 6). These findings, underscore the importance of chromatin-based gene regulation and highlight the role of Polycomb repression in safeguarding the pluripotent state. It has to be noted that this is in contrast to the observations from *in vitro* studies showing that ESC can maintain the pluripotent state in the absence of PRC2 (Leeb et al., 2010; Pasini et al., 2007; Shen et al., 2008). However, ESC are cultured in a medium containing factors that promote the pluripotent state, while PGCs, oocytes and epiblast cells are exposed to diverse signalling cues, not necessarily compatible with a naïve pluripotent state. Thus PRC2 needs to be in place and counteract differentiation signals, which would activate *Hox* genes and lineage-specific factors (Chapter 5). Furthermore, it became evident that PRC1 and PRC2 serve a parallel and partially redundant function in preimplantation embryos (Chapter 6), which is in line with the severe phenotype observed in *Eed/Rnf2* DKO ESC compared to the respective single knock-out cell lines (Leeb et al., 2010).

Based on the current state of knowledge, I can conclude that Polycomb-mediated repression plays an essential role in establishing and maintaining pluripotency *in vivo*.

## 7.3 Polycomb repression in the context of the GRN

Each cell is characterized by its transcriptional state, that is the state of the genome-encoded GRN (Huang, 2012a). Germ cells and preimplantation embryos are no exception to this principle and in light of the above mentioned notion of the pluripotency cycle (Leitch and Smith, 2013), it can be suggested that germ cells, gametes and the totipotent/pluripotent embryonic cells share similarities and key features of their respective GRN states. In other words, these different cell types may reside in a set of similar, "quasi-pluripotent" GRN states that can give rise to all other states upon differentiation. The GRN state is a function of the gene expression levels of all genes in a given cell, thus each factor that can influence transcription has a potential impact on the GRN state. Polycomb proteins mediate transcriptional repression and thus are an integral part of the regulatory factors defining a GRN state. Moreover, it has been shown that Polycomb proteins primarily repress lineage-specific transcription factors ("developmental genes"), which are not compatible with the current GRN state of a cell (Bernstein et al., 2006a; Mohn et al., 2008). In ESC, there is a special class of genes that are marked by both H3K27me3 and H3K4me3 and are kept in a repressed state. These genes have been termed "bivalent" and comprise lineage-specific transcription factors that are silenced in pluripotent ESC, but would be required upon differentiation into a given lineage (Bernstein et al., 2006a). The suggested model posits that bivalent genes are kept in a poised expression state, which would allow rapid activation upon induction of differentiation. In accordance with the suggestion that PGCs and ESCs occupy similar GRN states, which are reflected also by the histone modification profiles, a recent genome-wide study shows a strong overlap between bivalent genes in PGCs and ESC (Sachs et al., 2013). Thus a picture emerges, where Polycomb complexes serve a stabilizing role by increasing the required energy for a shift in the GRN state. As illustrated in Figure 1.1, a given GRN state can be described as a basin of attraction – once has reached the basin, it would require significant energy to leave it and move to another basin, which makes the GRN state stable.

And this is where I see the role of Polycomb repressive mechanisms – by silencing key developmental genes that would be required for a GRN shift, they increase the energy required to leave a GRN state and thus make the latter more stable. If Polycomb function is abolished, this would lead to destabilization of the GRN stage. This could eventually cause a transcriptional shift towards another GRN state, or a misregulation incompatible with the cellular program. However, a Polycomb-deficient cell could also maintain its GRN if the environment and the current transcriptional program are stable and reinforce the given GRN state, which is the case in PRC2-deficient ESC cultured under conditions that stimulate the pluripotent state. On the other hand, Polycomb repression represents a barrier, but not an irreversible block and thus can be overcome. An example for such a situation is the activation of the Polycomb-repressed master regulator MyoD in human fibroblasts upon induction with exogenous MyoD (Taberlay et al., 2011). Here, I report a range of phenotypes and transcriptional outcomes that depend on the dosage of Polycomb activity. This dosage-dependency could be explained with the stabilization of the GRN state by Polycomb – by lowering the Polycomb-mediated repression, the transcriptional state gets destabilized, however only below a given threshold and in the presence of extrinsic or intrinsic cues, this leads to developmental arrest or apoptosis. Furthermore, maternal deficiency for PRC2 leads to derepression of several hundred genes in oocytes, but this does not prevent preimplantation development in the presence of zygotic PRC2. Therefore the destabilized GRN state in the oocyte could be brought under control before the misregulation has reached a state incompatible with development. When both maternal and zygotic PRC2 are absent, the aberrant changes accumulate to a level that prevents normal development. Strongly reduced levels of PRC2/H3K27me3 in the oocyte and total absence of enzymatic activity over several cell divisions in embryos are still compatible with development. I interpret this as a safety mechanism during the pluripotency cycle, in which the Polycomb barrier is set much higher than actually required in order to stabilize and protect the GRN state. This could be also true for other GRN states and especially could also occur in terminally differentiated cells, where a given set of genes has been silenced over multiple generations and GRN transitions.

## 7.4 Single-cell heterogeneity

An important characteristic of cells is the non-genetic heterogeneity, which is defined as the phenotypic variability within a population of genetically identical cells (Huang, 2009). In the context of GRNs, this means that cells in the same GRN state could still differ from each other by their gene expression. This is possible because GRN states are not a single point in the multidimensional space, but rather occupy a certain volume that comprises a virtually infinite number of "sub-states" (some of which are more stable than others). Thus a seemingly homogenous cell population might be actually oscillating between sub-states at the single-cell level over time. Indeed, this phenomenon has been described in ESC (Chambers et al., 2007; Hayashi et al., 2008) and has even been suggested as a mechanism for priming lineage choices in multipotent cells (Chang et al., 2008). In line with this notion, a recent study reveals stochastic gene expression variations at the single-cell level in the ICM of the mouse blastocyst (Ohnishi et al., 2014).

I have observed phenotypic and transcriptional heterogeneity at the single-cell/single-embryo level in multiple instances in the current work. Furthermore, the variability observed in Polycomb-deficient embryos was consistently higher than the fluctuations in the respective control groups. Thus, it is logical to conclude that removing a major stabilizer of the pluripotency GRN state would lead to increased transcriptional noise and ultimately to a variety of phenotypic outcomes. In other words, the effect of abolishing Polycomb in one embryo might not be the same as the effect observed in another embryo, simply due to the stochastic nature of the GRN fluctuations (where all possible states within a given range of the quasi-potential can be occupied with equal probability).

## 7.5 Final conclusion and outlook

My work underscores the importance of Polycomb repressive mechanisms, and PRC2 in particular, as key regulators of gene expression in development. Reduced or abolished Polycomb function unleashes a variety of cellular responses depending on timing and dosage

of the repressive activity. This is presumably caused by destabilization of the GRN state and amplification of transcriptional noise. The idea that Polycomb is primarily acting as a repressor of transcriptional noise, came to me through my discussions with Mark Ptashne (Memorial Sloan–Kettering Cancer Center, New York, USA), who is currently investigating the mechanism of eukaryotic transcriptional regulation from a nucleosome perspective (Berrozpe et al., 2013).

What happens to the GRN state of a cell upon division? The daughter cell would inherit most of the key GRN regulators that were present in the mother cell, including transcription factors, signalling molecules, coding and non-coding RNA. Furthermore, an integral part of the GRN state is the chromatin conformation, which is mainly characterized by the chromatin organization and histone modifications. There are two main possibilities how the daughter cell could obtain the chromatin state of the mother cell – it could be actively established on new nucleosomes during replication, based solely on the presence of a histone mark (self-perpetuation), or it could be erased and re-established by *de novo* targeting of the histone modifiers to chromatin (through DNA-binding proteins, non-coding RNAs and other GRN factors) (Moazed, 2011). These two possibilities represent the difference between an inherited chromatin state and a chromatin state that is governed by the core GRN factors. My results suggest that a PRC2-mediated chromatin feature (H3K27me3) may be transmitted and required for the maintenance of a stable GRN states between two generations, i.e. from the gametes to the zygote. Although such a chromatin feature may be inherited over several cell divisions in the absence of the enzymatic complex (Gaydos et al., 2014; Puschendorf et al., 2008), there is still little evidence for an independent role of histone modifications in transgenerational epigenetic inheritance, as defined above (Chapter 1, i.e. stable transmission of a trait over multiple generations, this process). By definition, transgenerational epigenetic inheritance should persist in the absence of the initial signal or an active recruiter (such as Oct4 that acts as a recruiter of PRC2 in ESCs (Lee et al., 2006)). However, there seems to be a set of Polycomb target genes that bear H3K27me3 in all cell types of the pluripotency cycle (germ cells, gametes and pluripotent ESCs). Furthermore, some of these genes display specific retention of H3K27me3 and remain repressed upon reduction (but not full ablation) of PRC2 activity (Shen et al., 2008). These considerations and findings open the possibility for this set of genes to be

part of a transgenerational epigenetic inheritance mechanism that ensures the integrity of the pluripotent state vie PRC2-mediated repression.

An important aspect of Polycomb regulation that potentially has an impact on cancer mechanisms, is the impact of Polycomb repression on non-genetic heterogeneity of the system. Reduced or abolished Polycomb activity would allow the cells of a given population to occupy many more sub-states within a stable GRN state, which may facilitate the transition of some cells to a different stable GRN. This could be illustrated by the popcorn example – if one prepares popcorn in a deep pot, few if any of the popcorn will pop out of the pot. In contrast, if popcorn is prepared in a pan, most of it will end up outside the pan. Thus, lowering the constraints to shift from one GRN state to another may derail normal the developmental program and lead to abnormal cell fates that potentially give rise to cancer (Huang, 2012b). Indeed, chromatin modifying complexes, and Polycomb group proteins in particular, have been often associated with cancer (Albert and Helin, 2010).

There are several open questions that would be of particular interest in order to better understand the mechanism of Polycomb regulation in the pluripotency cycle. First, as far as technology allows, single-cell approaches need to be applied to assess, quantify and model the heterogeneity arising in Polycomb-deficient embryos. Next, the molecular link and hierarchy between PRC1 and PRC2 is still not entirely clear and is potentially changing in the different stages of the pluripotency cycle, thus it has to be addressed in a systematic way by altering the ratio between PRC1 and PRC2 in the various stages of development. It became evident that Polycomb mechanisms are dosage-dependent and therefore, it would be of interest to study the dosage effects by altering both PRC2 and H3K27-specific demethylases in various stages of development. Finally, we are still lacking a method to identify the Polycomb target genes in oocytes and early embryos, so developing single-cell or low-input ChIP would bring direct evidence for the molecular role of PRC2 in early embryos and also allow the comparison to PGCs and ESCs.

# V. References


Abe, K., Naruse, C., Kato, T., Nishiuchi, T., Saitou, M., and Asano, M. (2011). Loss of heterochromatin protein 1 gamma reduces the number of primordial germ cells via impaired cell cycle progression in mice. Biol. Reprod. *85*, 1013–1024.

Abel, J., Eskeland, R., Raffa, G.D., Kremmer, E., and Imhof, A. (2009). Drosophila HP1c is regulated by an auto-regulatory feedback loop through its binding partner Woc. PloS One *4*, e5089.

Agger, K., Christensen, J., Cloos, P.A.C., and Helin, K. (2008). The emerging functions of histone demethylases. Curr. Opin. Genet. Dev. *18*, 159–168.

Ahmed, K., Dehghani, H., Rugg-Gunn, P., Fussner, E., Rossant, J., and Bazett-Jones, D.P. (2010). Global chromatin architecture reflects pluripotency and lineage commitment in the early mouse embryo. PloS One *5*, e10531.

Ajduk, A., Biswas Shivhare, S., and Zernicka-Goetz, M. (2014). The basal position of nuclei is one pre-requisite for asymmetric cell divisions in the early mouse embryo. Dev. Biol. *392*, 133–140.

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

Akasaka, T., Lohuizen, M. van, Lugt, N. van der, Mizutani-Koseki, Y., Kanno, M., Taniguchi, M., Vidal, M., Alkema, M., Berns, A., and Koseki, H. (2001). 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 *128*, 1587–1597.

Albert, M., and Helin, K. (2010). Histone methyltransferases in cancer. Semin. Cell Dev. Biol. *21*, 209–220.

Albert, M., and Peters, A.H.F.M. (2009). Genetic and epigenetic control of early mouse development. Curr. Opin. Genet. Dev. *19*, 113–121.

Alder, O., Lavial, F., Helness, A., Brookes, E., Pinho, S., Chandrashekran, A., Arnaud, P., Pombo, A., O'Neill, L., and Azuara, V. (2010). Ring1B and Suv39h1 delineate distinct chromatin states at bivalent genes during early mouse lineage commitment. Dev. Camb. Engl. *137*, 2483–2492.

Aldiri, I., and Vetter, M.L. (2012). PRC2 during vertebrate organogenesis: A complex in transition. Dev. Biol. *367*, 91–99.

Alizadeh, Z., Kageyama, S.-I., and Aoki, F. (2005). Degradation of maternal mRNA in mouse embryos: Selective degradation of specific mRNAs after fertilization. Mol. Reprod. Dev. *72*, 281–290.

Allshire, R.C., Nimmo, E.R., Ekwall, K., Javerzat, J.P., and Cranston, G. (1995). Mutations derepressing silent centromeric domains in fission yeast disrupt chromosome segregation. Genes Dev. *9*, 218–233.

Anderson, A.E., Karandikar, U.C., Pepple, K.L., Chen, Z., Bergmann, A., and Mardon, G. (2011). The enhancer of trithorax and polycomb gene Caf1/p55 is essential for cell survival and patterning in Drosophila development. Dev. Camb. Engl. *138*, 1957–1966.

Aoki, F., Worrad, D.M., and Schultz, R.M. (1997). Regulation of transcriptional activity during the first and second cell cycles in the preimplantation mouse embryo. Dev. Biol. *181*, 296–307.

Arney, K.L., Bao, S., Bannister, A.J., Kouzarides, T., and Surani, M.A. (2002). Histone methylation defines epigenetic asymmetry in the mouse zygote. Int. J. Dev. Biol. *46*, 317–320.

Arnold, P., Schöler, A., Pachkov, M., Balwierz, P., Jørgensen, H., Stadler, M.B., van Nimwegen, E., and Schübeler, D. (2012). Modeling of epigenome dynamics identifies transcription factors that mediate Polycomb targeting. Genome Res.

Artus, J., Piliszek, A., and Hadjantonakis, A.-K. (2011). The primitive endoderm lineage of the mouse blastocyst: sequential transcription factor activation and regulation of differentiation by Sox17. Dev. Biol. *350*, 393–404.

Aubert, D., Chen, L., Moon, Y.H., Martin, D., Castle, L.A., Yang, C.H., and Sung, Z.R. (2001). EMF1, a novel protein involved in the control of shoot architecture and flowering in Arabidopsis. Plant Cell *13*, 1865–1875.

Aucott, R., Bullwinkel, J., Yu, Y., Shi, W., Billur, M., Brown, J.P., Menzel, U., Kioussis, D., Wang, G., Reisert, I., et al. (2008). HP1-beta is required for development of the cerebral neocortex and neuromuscular junctions. J. Cell Biol. *183*, 597–606.

Ayoub, N., Jeyasekharan, A.D., Bernal, J.A., and Venkitaraman, A.R. (2008). HP1-beta mobilization promotes chromatin changes that initiate the DNA damage response. Nature *453*, 682–686.

Ayyanathan, K., Lechner, M.S., Bell, P., Maul, G.G., Schultz, D.C., Yamada, Y., Tanaka, K., Torigoe, K., and Rauscher, F.J., 3rd (2003). Regulated recruitment of HP1 to a euchromatic gene induces mitotically heritable, epigenetic gene silencing: a mammalian cell culture model of gene variegation. Genes Dev. *17*, 1855–1869.

Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jørgensen, H.F., John, R.M., Gouti, M., Casanova, M., Warnes, G., Merkenschlager, M., et al. (2006). Chromatin signatures of pluripotent cell lines. Nat. Cell Biol. *8*, 532–538.

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

Baumann, C., and De La Fuente, R. (2011). Role of Polycomb Group Protein Cbx2/M33 in Meiosis Onset and Maintenance of Chromosome Stability in the Mammalian Germline. Genes *2*, 59–80.

Baumbusch, L.O., Thorstensen, T., Krauss, V., Fischer, A., Naumann, K., Assalkhou, R., Schulz, I., Reuter, G., and Aalen, R.B. (2001). The Arabidopsis thaliana genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes. Nucleic Acids Res. *29*, 4319–4333.

Bergsmedh, A., Donohoe, M.E., Hughes, R.-A., and Hadjantonakis, A.-K. (2011). Understanding the Molecular Circuitry of Cell Lineage Specification in the Early Mouse Embryo. Genes *2*, 420–448.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006a). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315–326.

Bernstein, E., Duncan, E.M., Masui, O., Gil, J., Heard, E., and Allis, C.D. (2006b). Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. Mol. Cell. Biol. *26*, 2560–2569.

Berrozpe, G., Bryant, G.O., Warpinski, K., and Ptashne, M. (2013). Regulation of a Mammalian Gene Bearing a CpG Island Promoter and a Distal Enhancer. Cell Rep. *4*, 445–453.

Bilodeau, S., Kagey, M.H., Frampton, G.M., Rahl, P.B., and Young, R.A. (2009). SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. Genes Dev. *23*, 2484–2489.

Birve, A., Sengupta, A.K., Beuchle, D., Larsson, J., Kennison, J.A., Rasmuson-Lestander A, and Müller, J. (2001). Su(z)12, a novel Drosophila Polycomb group gene that is conserved in vertebrates and plants. Dev. Camb. Engl. *128*, 3371–3379.

Black, J.C., Allen, A., Van Rechem, C., Forbes, E., Longworth, M., Tschöp, K., Rinehart, C., Quiton, J., Walsh, R., Smallwood, A., et al. (2010). Conserved antagonism between JMJD2A/KDM4A and HP1γ during cell cycle progression. Mol. Cell *40*, 736–748.

Blackledge, N.P., Farcas, A.M., Kondo, T., King, H.W., McGouran, J.F., Hanssen, L.L.P., Ito, S., Cooper, S., Kondo, K., Koseki, Y., et al. (2014). Variant PRC1 Complex-Dependent H2A Ubiquitylation Drives PRC2 Recruitment and Polycomb Domain Formation. Cell *157*, 1445–1459.

Bouyer, D., Roudier, F., Heese, M., Andersen, E.D., Gey, D., Nowack, M.K., Goodrich, J., Renou, J.-P., Grini, P.E., Colot, V., et al. (2011). Polycomb Repressive Complex 2 Controls the Embryo-to-Seedling Phase Transition. PLoS Genet *7*, e1002014.

Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K., et al. (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature *441*, 349–353.

Boyle, E.I., Weng, S., Gollub, J., Jin, H., Botstein, D., Cherry, J.M., and Sherlock, G. (2004). GO::TermFinder--open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. Bioinforma. Oxf. Engl. *20*, 3710–3715.

Bracken, A.P., Dietrich, N., Pasini, D., Hansen, K.H., and Helin, K. (2006). Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. Genes Dev. *20*, 1123–1136.

Braig, M., Lee, S., Loddenkemper, C., Rudolph, C., Peters, A.H.F.M., Schlegelberger, B., Stein, H., Dörken, B., Jenuwein, T., and Schmitt, C.A. (2005). Oncogene-induced senescence as an initial barrier in lymphoma development. Nature *436*, 660–665.

Bratzel, F., López-Torrejón, G., Koch, M., Del Pozo, J.C., and Calonje, M. (2010). Keeping Cell Identity in Arabidopsis Requires PRC1 RING-Finger Homologs that Catalyze H2A Monoubiquitination. Curr. Biol. *20*, 1853–1859.

Breen, T.R., and Duncan, I.M. (1986). Maternal expression of genes that regulate the bithorax complex of Drosophila melanogaster. Dev. Biol. *118*, 442–456.

Brower-Toland, B., Riddle, N.C., Jiang, H., Huisinga, K.L., and Elgin, S.C.R. (2009). Multiple SET methyltransferases are required to maintain normal heterochromatin domains in the genome of Drosophila melanogaster. Genetics *181*, 1303–1319.

Brown, J.P., Bullwinkel, J., Baron-Lühr, B., Billur, M., Schneider, P., Winking, H., and Singh, P.B. (2010). HP1gamma function is required for male germ cell survival and spermatogenesis. Epigenetics Chromatin *3*, 9.

Bruce, A.W., and Zernicka-Goetz, M. (2010). Developmental control of the early mammalian embryo: competition among heterogeneous cells that biases cell fate. Curr. Opin. Genet. Dev. *20*, 485–491.

Brykczynska, U., Hisano, M., Erkek, S., Ramos, L., Oakeley, E.J., Roloff, T.C., Beisel, C., Schübeler, D., Stadler, M.B., and Peters, A.H.F.M. (2010). Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. Nat. Struct. Mol. Biol. *17*, 679–687.

Buchwald, G., van der Stoop, P., Weichenrieder, O., Perrakis, A., van Lohuizen, M., and Sixma, T.K. (2006). Structure and E3-ligase activity of the Ring-Ring complex of polycomb proteins Bmi1 and Ring1b. EMBO J. *25*, 2465–2474.

Bultman, S.J., Gebuhr, T.C., Pan, H., Svoboda, P., Schultz, R.M., and Magnuson, T. (2006). Maternal BRG1 regulates zygotic genome activation in the mouse. Genes Dev. *20*, 1744–1754.

Bulut-Karslioglu, A., Perrera, V., Scaranaro, M., de la Rosa-Velazquez, I.A., van de Nobelen, S., Shukeir, N., Popow, J., Gerle, B., Opravil, S., Pagani, M., et al. (2012). A transcription factor-based mechanism for mouse heterochromatin formation. Nat. Struct. Mol. Biol.

Burton, A., and Torres-Padilla, M.-E. (2010). Epigenetic reprogramming and development: a unique heterochromatin organization in the preimplantation mouse embryo. Brief. Funct. Genomics *9*, 444–454.

Burton, A., Muller, J., Tu, S., Padilla-Longoria, P., Guccione, E., and Torres-Padilla, M.-E. (2013). Single-Cell Profiling of Epigenetic Modifiers Identifies PRDM14 as an Inducer of Cell Fate in the Mammalian Embryo. Cell Rep. *5*, 687–701.

Cammas, F., Janoshazi, A., Lerouge, T., and Losson, R. (2007). Dynamic and selective interactions of the transcriptional corepressor TIF1 beta with the heterochromatin protein HP1 isotypes during cell differentiation. Differ. Res. Biol. Divers. *75*, 627–637.

Campos, E.I., and Reinberg, D. (2009). Histones: annotating chromatin. Annu. Rev. Genet. *43*, 559–599.

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science *298*, 1039–1043.

Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., and Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. Nature *450*, 1230–1234.

Chang, H.H., Hemberg, M., Barahona, M., Ingber, D.E., and Huang, S. (2008). Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. Nature *453*, 544–547.

Chazaud, C., Yamanaka, Y., Pawson, T., and Rossant, J. (2006). Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. Dev. Cell *10*, 615–624.

Checchi, P.M., and Engebrecht, J. (2011). Caenorhabditis elegans Histone Methyltransferase MET-2 Shields the Male X Chromosome from Checkpoint Machinery and Mediates Meiotic Sex Chromosome Inactivation. PLoS Genet *7*, e1002267.

Cheutin, T., and Cavalli, G. (2014). Polycomb silencing: from linear chromatin domains to 3D chromosome folding. Curr. Opin. Genet. Dev. *25*, 30–37.

Cheutin, T., McNairn, A.J., Jenuwein, T., Gilbert, D.M., Singh, P.B., and Misteli, T. (2003). Maintenance of stable heterochromatin domains by dynamic HP1 binding. Science *299*, 721–725.

Chin, H.G., Estève, P.-O., Pradhan, M., Benner, J., Patnaik, D., Carey, M.F., and Pradhan, S. (2007). Automethylation of G9a and its implication in wider substrate specificity and HP1 binding. Nucleic Acids Res. *35*, 7313–7323.

Cho, C., Willis, W.D., Goulding, E.H., Jung-Ha, H., Choi, Y.C., Hecht, N.B., and Eddy, E.M. (2001). Haploinsufficiency of protamine-1 or -2 causes infertility in mice. Nat. Genet. *28*, 82–86.

Cho, S., Park, J.S., and Kang, Y.-K. (2011). Dual functions of histone-lysine N-methyltransferase Setdb1 protein at promyelocytic leukemia-nuclear body (PML-NB): maintaining PML-NB structure and regulating the expression of its associated genes. J. Biol. Chem. *286*, 41115–41124.

Cho, S., Park, J.S., Kwon, S., and Kang, Y.-K. (2012). Dynamics of Setdb1 expression in early mouse development. Gene Expr. Patterns GEP *12*, 213–218.

Chuva de Sousa Lopes, S.M., Hayashi, K., Shovlin, T.C., Mifsud, W., Surani, M.A., and McLaren, A. (2008). X Chromosome Activity in Mouse XX Primordial Germ Cells. PLoS Genet *4*, e30.

Clough, E., Moon, W., Wang, S., Smith, K., and Hazelrigg, T. (2007). Histone methylation is required for oogenesis in Drosophila. Dev. Camb. Engl. *134*, 157–165.

Comet, I., and Helin, K. (2014). Revolution in the Polycomb hierarchy. Nat. Struct. Mol. Biol. *21*, 573–575.

Consortium, T.E.P. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

Coré, N., Bel, S., Gaunt, S.J., Aurrand-Lions, M., Pearce, J., Fisher, A., and Djabali, M. (1997). Altered cellular proliferation and mesoderm patterning in Polycomb-M33-deficient mice. Dev. Camb. Engl. *124*, 721–729.

Coustham, V., Li, P., Strange, A., Lister, C., Song, J., and Dean, C. (2012). Quantitative modulation of polycomb silencing underlies natural variation in vernalization. Science *337*, 584–587.

Couteau, F., Guerry, F., Muller, F., and Palladino, F. (2002). A heterochromatin protein 1 homologue in Caenorhabditis elegans acts in germline and vulval development. EMBO Rep. *3*, 235–241.

Crevillén, P., Yang, H., Cui, X., Greeff, C., Trick, M., Qiu, Q., Cao, X., and Dean, C. (2014). Epigenetic reprogramming that prevents transgenerational inheritance of the vernalized state. Nature *advance online publication*.

Crick, F. (1970). Central Dogma of Molecular Biology. Nature *227*, 561–563.

Dambacher, S., Hahn, M., and Schotta, G. (2010). Epigenetic regulation of development by histone lysine methylation. Heredity *105*, 24–37.

Daujat, S., Weiss, T., Mohn, F., Lange, U.C., Ziegler-Birling, C., Zeissler, U., Lappe, M., Schübeler, D., Torres-Padilla, M.-E., and Schneider, R. (2009). H3K64 trimethylation marks heterochromatin

and is dynamically remodeled during developmental reprogramming. Nat. Struct. Mol. Biol. *16*, 777–781.

Davidson, E.H., and Erwin, D.H. (2006). Gene Regulatory Networks and the Evolution of Animal Body Plans. Science *311*, 796–800.

Daxinger, L., and Whitelaw, E. (2012). Understanding transgenerational epigenetic inheritance via the gametes in mammals. Nat. Rev. Genet. *13*, 153–162.

Denholtz, M., Bonora, G., Chronis, C., Splinter, E., de Laat, W., Ernst, J., Pellegrini, M., and Plath, K. (2013). Long-Range Chromatin Contacts in Embryonic Stem Cells Reveal a Role for Pluripotency Factors and Polycomb Proteins in Genome Organization. Cell Stem Cell *13*, 602–616.

Dietrich, J.-E., and Hiiragi, T. (2007). Stochastic patterning in the mouse pre-implantation embryo. Development *134*, 4219–4231.

Dietrich, N., Lerdrup, M., Landt, E., Agrawal-Singh, S., Bak, M., Tommerup, N., Rappsilber, J., Södersten, E., and Hansen, K. (2012). REST-mediated recruitment of polycomb repressor complexes in mammalian cells. PLoS Genet. *8*, e1002494.

Dinant, C., and Luijsterburg, M.S. (2009). The emerging role of HP1 in the DNA damage response. Mol. Cell. Biol. *29*, 6335–6340.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380.

Dodge, J.E., Kang, Y.-K., Beppu, H., Lei, H., and Li, E. (2004). Histone H3-K9 methyltransferase ESET is essential for early development. Mol. Cell. Biol. *24*, 2478–2486.

Dong, K.B., Maksakova, I.A., Mohn, F., Leung, D., Appanah, R., Lee, S., Yang, H.W., Lam, L.L., Mager, D.L., Schübeler, D., et al. (2008). DNA methylation in ES cells requires the lysine methyltransferase G9a but not its catalytic activity. EMBO J. *27*, 2691–2701.

Duncan, I.M. (1982). Polycomblike: A Gene That Appears to Be Required for the Normal Expression of the Bithorax and Antennapedia Gene Complexes of DROSOPHILA MELANOGASTER. Genetics *102*, 49–70.

Dura, J.M., Brock, H.W., and Santamaria, P. (1985). Polyhomeotic: a gene of Drosophila melanogaster required for correct expression of segmental identity. Mol. Gen. Genet. MGG *198*, 213–220.

Dura, J.-M., Randsholt, N.B., Deatrick, J., Erk, I., Santamaria, P., Freeman, J.D., Freeman, S.J., Weddell, D., and Brock, H.W. (1987). A complex genetic locus, polyhomeotic, is required for segmental specification and epidermal development in D. melanogaster. Cell *51*, 829–839.

Ebbs, M.L., and Bender, J. (2006). Locus-specific control of DNA methylation by the Arabidopsis SUVH5 histone methyltransferase. Plant Cell *18*, 1166–1176.

Eissenberg, J.C., Morris, G.D., Reuter, G., and Hartnett, T. (1992). The heterochromatin-associated protein HP-1 is an essential protein in Drosophila with dosage-dependent effects on position-effect variegation. Genetics *131*, 345–352.

Endoh, M., Endo, T.A., Endoh, T., Fujimura, Y., Ohara, O., Toyoda, T., Otte, A.P., Okano, M., Brockdorff, N., Vidal, M., et al. (2008). Polycomb group proteins Ring1A/B are functionally linked to the core transcriptional regulatory circuitry to maintain ES cell identity. Dev. Camb. Engl. *135*, 1513–1524.

Endoh, M., Endo, T.A., Endoh, T., Isono, K.-I., Sharif, J., Ohara, O., Toyoda, T., Ito, T., Eskeland, R., Bickmore, W.A., et al. (2012). Histone H2A Mono-Ubiquitination Is a Crucial Step to Mediate PRC1-Dependent Repression of Developmental Genes to Maintain ES Cell Identity. PLoS Genet. *8*, e1002774.

Epsztejn-Litman, S., Feldman, N., Abu-Remaileh, M., Shufaro, Y., Gerson, A., Ueda, J., Deplus, R., Fuks, F., Shinkai, Y., Cedar, H., et al. (2008). De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. Nat. Struct. Mol. Biol. *15*, 1176–1183.

Erhardt, S., Su, I. -hsi., Schneider, R., Barton, S., Bannister, A.J., Perez-Burgos, L., Jenuwein, T., Kouzarides, T., Tarakhovsky, A., and Surani, M.A. (2003). Consequences of the depletion of zygotic and embryonic enhancer of zeste 2 during preimplantation mouse. Development *130*, 4235–4248.

Erkek, S., Hisano, M., Liang, C.-Y., Gill, M., Murr, R., Dieker, J., Schübeler, D., Vlag, J. van der, Stadler, M.B., and Peters, A.H.F.M. (2013). Molecular determinants of nucleosome retention at CpG-rich sequences in mouse spermatozoa. Nat. Struct. Mol. Biol. *20*, 868–875.

Estève, P.-O., Chin, H.G., Smallwood, A., Feehery, G.R., Gangisetty, O., Karpf, A.R., Carey, M.F., and Pradhan, S. (2006). Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication. Genes Dev. *20*, 3089–3103.

Eymery, A., Callanan, M., and Vourc'h, C. (2009). The secret message of heterochromatin: new insights into the mechanisms and function of centromeric and pericentric repeat sequence transcription. Int. J. Dev. Biol. *53*, 259–268.

Ezhkova, E., Lien, W.-H., Stokes, N., Pasolli, H.A., Silva, J.M., and Fuchs, E. (2011). EZH1 and EZH2 cogovern histone H3K27 trimethylation and are essential for hair follicle homeostasis and wound repair. Genes Dev. *25*, 485–498.

Farcas, A.M., Blackledge, N.P., Sudbery, I., Long, H.K., McGouran, J.F., Rose, N.R., Lee, S., Sims, D., Cerase, A., Sheahan, T.W., et al. (2012). KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands. eLife *1*.

Faust, C., Lawson, K.A., Schork, N.J., Thiel, B., and Magnuson, T. (1998). The Polycomb-group gene eed is required for normal morphogenetic movements during gastrulation in the mouse embryo. Dev. Camb. Engl. *125*, 4495–4506.

Feldman, N., Gerson, A., Fang, J., Li, E., Zhang, Y., Shinkai, Y., Cedar, H., and Bergman, Y. (2006). G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. Nat. Cell Biol. *8*, 188–194.

Filion, G.J., van Bemmel, J.G., Braunschweig, U., Talhout, W., Kind, J., Ward, L.D., Brugman, W., de Castro, I.J., Kerkhoven, R.M., Bussemaker, H.J., et al. (2010). Systematic Protein Location Mapping Reveals Five Principal Chromatin Types in Drosophila Cells. Cell *143*, 212–224.

Fischle, W., Wang, Y., Jacobs, S.A., Kim, Y., Allis, C.D., and Khorasanizadeh, S. (2003). Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. Genes Dev. *17*, 1870–1881.

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

Fisher, C.L., Lee, I., Bloyer, S., Bozza, S., Chevalier, J., Dahl, A., Bodner, C., Helgason, C.D., Hess, J.L., Humphries, R.K., et al. (2010). Additional sex combs-like 1 belongs to the enhancer of trithorax and polycomb group and genetically interacts with Cbx2 in mice. Dev. Biol. *337*, 9–15.

Fodor, B.D., Kubicek, S., Yonezawa, M., O'Sullivan, R.J., Sengupta, R., Perez-Burgos, L., Opravil, S., Mechtler, K., Schotta, G., and Jenuwein, T. (2006). Jmjd2b antagonizes H3K9 trimethylation at pericentric heterochromatin in mammalian cells. Genes Dev. *20*, 1557–1562.

Font-Burgada, J., Rossell, D., Auer, H., and Azorín, F. (2008). Drosophila HP1c isoform interacts with the zinc-finger proteins WOC and Relative-of-WOC to regulate gene expression. Genes Dev. *22*, 3007–3023.

Freitag, M., Hickey, P.C., Khlafallah, T.K., Read, N.D., and Selker, E.U. (2004). HP1 is essential for DNA methylation in neurospora. Mol. Cell *13*, 427–434.

Frietze, S., O'Geen, H., Blahnik, K.R., Jin, V.X., and Farnham, P.J. (2010). ZNF274 recruits the histone methyltransferase SETDB1 to the 3' ends of ZNF genes. PloS One *5*, e15082.

Fritsch, L., Robin, P., Mathieu, J.R.R., Souidi, M., Hinaux, H., Rougeulle, C., Harel-Bellan, A., Ameyar-Zazoua, M., and Ait-Si-Ali, S. (2010). A subset of the histone H3 lysine 9 methyltransferases Suv39h1, G9a, GLP, and SETDB1 participate in a multimeric complex. Mol. Cell *37*, 46–56.

Frohnhöfer, H.G., and Nüsslein-Volhard, C. (1986). Organization of anterior pattern in the Drosophila embryo by the maternal gene bicoid. Nature *324*, 120–125.

Gao, H., Yu, Z., Bi, D., Jiang, L., Cui, Y., Sun, J., and Ma, R. (2007). Akt/PKB interacts with the histone H3 methyltransferase SETDB1 and coordinates to silence gene expression. Mol. Cell. Biochem. *305*, 35–44.

Gao, Z., Zhang, J., Bonasio, R., Strino, F., Sawai, A., Parisi, F., Kluger, Y., and Reinberg, D. (2012). PCGF Homologs, CBX Proteins, and RYBP Define Functionally Distinct PRC1 Family Complexes. Mol. Cell *45*, 344–356.

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

Gaudin, V., Libault, M., Pouteau, S., Juul, T., Zhao, G., Lefebvre, D., and Grandjean, O. (2001). Mutations in LIKE HETEROCHROMATIN PROTEIN 1 affect flowering time and plant architecture in Arabidopsis. Dev. Camb. Engl. *128*, 4847–4858.

Gaydos, L.J., Wang, W., and Strome, S. (2014). H3K27me and PRC2 transmit a memory of repression across generations and during development. Science *345*, 1515–1518.

Gearhart, M.D., Corcoran, C.M., Wamstad, J.A., and Bardwell, V.J. (2006). Polycomb group and SCF ubiquitin ligases are found in a novel BCOR complex that is recruited to BCL6 targets. Mol. Cell. Biol. *26*, 6880–6889.

Gelbart, W., and Emmert, D. (2011). Calculation of RPKM to generate quantitative expression data: read-length values for modENCODE developmental timecourse RNA-Seq data.

Gerberding, M., Browne, W.E., and Patel, N.H. (2002). Cell lineage analysis of the amphipod crustacean Parhyale hawaiensis reveals an early restriction of cell fates. Dev. Camb. Engl. *129*, 5789–5801.

Gerstein, M.B., Lu, Z.J., Nostrand, E.L.V., Cheng, C., Arshinoff, B.I., Liu, T., Yip, K.Y., Robilotto, R., Rechtsteiner, A., Ikegami, K., et al. (2010). Integrative Analysis of the Caenorhabditis elegans Genome by the modENCODE Project. Science *330*, 1775–1787.

Gill, M.E., Erkek, S., and Peters, A.H. (2012). Parental epigenetic control of embryogenesis: a balance between inheritance and reprogramming? Curr. Opin. Cell Biol. *24*, 387–396.

Gkountela, S., and Clark, A.T. (2014). A big surprise in the little zygote: the curious business of losing methylated cytosines. Cell Stem Cell *15*, 393–394.

Goodrich, J., Puangsomlee, P., Martin, M., Long, D., Meyerowitz, E.M., and Coupland, G. (1997). A Polycomb-group gene regulates homeotic gene expression in Arabidopsis. Publ. Online 06 March 1997 Doi101038386044a0 *386*, 44–51.

Graveley, B.R., Brooks, A.N., Carlson, J.W., Duff, M.O., Landolin, J.M., Yang, L., Artieri, C.G., van Baren, M.J., Boley, N., Booth, B.W., et al. (2011). The developmental transcriptome of Drosophila melanogaster. Nature *471*, 473–479.

Grossniklaus, U., Vielle-Calzada, J.-P., Hoeppner, M.A., and Gagliano, W.B. (1998). Maternal Control of Embryogenesis by MEDEA, a Polycomb Group Gene in Arabidopsis. Science *280*, 446–450.

Guil, S., Soler, M., Portela, A., Carrère, J., Fonalleras, E., Gómez, A., Villanueva, A., and Esteller, M. (2012). Intronic RNAs mediate EZH2 regulation of epigenetic targets. Nat. Struct. Mol. Biol. *19*, 664–670.

Guo, G., Huss, M., Tong, G.Q., Wang, C., Li Sun, L., Clarke, N.D., and Robson, P. (2010). Resolution of Cell Fate Decisions Revealed by Single-Cell Gene Expression Analysis from Zygote to Blastocyst. Dev. Cell *18*, 675–685.

Gutiérrez, L., Oktaba, K., Scheuermann, J.C., Gambetta, M.C., Ly-Hartig, N., and Müller, J. (2012). The role of the histone H2A ubiquitinase Sce in Polycomb repression. Development *139*, 117–127.

Guttman, M., Donaghey, J., Carey, B.W., Garber, M., Grenier, J.K., Munson, G., Young, G., Lucas, A.B., Ach, R., Bruhn, L., et al. (2011). lincRNAs act in the circuitry controlling pluripotency and differentiation. Nature *477*, 295–300.

Gyory, I., Wu, J., Fejér, G., Seto, E., and Wright, K.L. (2004). PRDI-BF1 recruits the histone H3 methyltransferase G9a in transcriptional silencing. Nat. Immunol. *5*, 299–308.

Hajkova, P., Ancelin, K., Waldmann, T., Lacoste, N., Lange, U.C., Cesari, F., Lee, C., Almouzni, G., Schneider, R., and Surani, M.A. (2008). Chromatin dynamics during epigenetic reprogramming in the mouse germ line. Nature *452*, 877–881.

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

Han, Z., Xing, X., Hu, M., Zhang, Y., Liu, P., and Chai, J. (2007). Structural Basis of EZH2 Recognition by EED. Structure *15*, 1306–1315.

Hanna, J.H., Saha, K., and Jaenisch, R. (2010). Pluripotency and Cellular Reprogramming: Facts, Hypotheses, Unresolved Issues. Cell *143*, 508–525.

Hannah-Alava, A. (1958). Developmental Genetics of the Posterior Legs in Drosophila Melanogaster. Genetics *43*, 878–905.

Hansen, K.H., Bracken, A.P., Pasini, D., Dietrich, N., Gehani, S.S., Monrad, A., Rappsilber, J., Lerdrup, M., and Helin, K. (2008). A model for transmission of the H3K27me3 epigenetic mark. Nat. Cell Biol. *10*, 1291–1300.

Hasegawa, K., Zhang, P., Wei, Z., Pomeroy, J.E., Lu, W., and Pera, M.F. (2010). Comparison of Reprogramming Efficiency Between Transduction of Reprogramming Factors, Cell–Cell Fusion, and Cytoplast Fusion. STEM CELLS *28*, 1338–1348.

Hashimshony, T., Wagner, F., Sher, N., and Yanai, I. (2012). CEL-Seq: Single-Cell RNA-Seq by Multiplexed Linear Amplification. Cell Rep. *2*, 666–673.

Hathaway, N.A., Bell, O., Hodges, C., Miller, E.L., Neel, D.S., and Crabtree, G.R. (2012). Dynamics and memory of heterochromatin in living cells. Cell *149*, 1447–1460.

Hayashi, K., and Surani, M.A. (2009). Resetting the Epigenome beyond Pluripotency in the Germline. Cell Stem Cell *4*, 493–498.

Hayashi, K., de Sousa Lopes, S.M.C., Tang, F., Lao, K., and Surani, M.A. (2008). Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. Cell Stem Cell *3*.

Hayashi, M.T., Takahashi, T.S., Nakagawa, T., Nakayama, J., and Masukata, H. (2009). The heterochromatin protein Swi6/HP1 activates replication origins at the pericentromeric region and silent mating-type locus. Nat. Cell Biol. *11*, 357–362.

Heard, E., and Martienssen, R.A. (2014). Transgenerational Epigenetic Inheritance: Myths and Mechanisms. Cell *157*, 95–109.

Van der Heijden, G.W., Dieker, J.W., Derijck, A.A.H.A., Muller, S., Berden, J.H.M., Braat, D.D.M., van der Vlag, J., and de Boer, P. (2005). Asymmetry in histone H3 variants and lysine methylation between paternal and maternal chromatin of the early mouse zygote. Mech. Dev. *122*, 1008–1022.

Van der Heijden, G.W., Derijck, A.A.H.A., Pósfai, E., Giele, M., Pelczar, P., Ramos, L., Wansink, D.G., van der Vlag, J., Peters, A.H.F.M., and de Boer, P. (2007). Chromosome-wide nucleosome replacement and H3.3 incorporation during mammalian meiotic sex chromosome inactivation. Nat. Genet. *39*, 251–258.

Hekimoglu, B., and Ringrose, L. (2009). Non-coding RNAs in polycomb/trithorax regulation. RNA Biol. *6*, 129–137.

Hirasawa, R., Chiba, H., Kaneda, M., Tajima, S., Li, E., Jaenisch, R., and Sasaki, H. (2008). Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. Genes Dev. *22*, 1607–1616.

Hisada, K., Sánchez, C., Endo, T.A., Endoh, M., Román-Trufero, M., Sharif, J., Koseki, H., and Vidal, M. (2012). RYBP represses endogenous retroviruses and preimplantation- and germ line-specific genes in mouse embryonic stem cells. Mol. Cell. Biol. *32*, 1139–1149.

Hogarth, C.A., Mitchell, D., Evanoff, R., Small, C., and Griswold, M. (2011). Identification and expression of potential regulators of the mammalian mitotic-to-meiotic transition. Biol. Reprod. *84*, 34–42.

Holec, S., and Berger, F. (2012). Polycomb group complexes mediate developmental transitions in plants. Plant Physiol. *158*, 35–43.

Honda, S., and Selker, E.U. (2008). Direct interaction between DNA methyltransferase DIM-2 and HP1 is required for DNA methylation in Neurospora crassa. Mol. Cell. Biol. *28*, 6044–6055.

Honda, S., Lewis, Z.A., Huarte, M., Cho, L.Y., David, L.L., Shi, Y., and Selker, E.U. (2010). The DMM complex prevents spreading of DNA methylation from transposons to nearby genes in Neurospora crassa. Genes Dev. *24*, 443–454.

Horáková, A.H., Bártová, E., Galiová, G., Uhlírová, R., Matula, P., and Kozubek, S. (2010). SUV39h-independent association of HP1 beta with fibrillarin-positive nucleolar regions. Chromosoma *119*, 227–241.

Huang, S. (2009). Non-genetic heterogeneity of cells in development: more than just noise. Development *136*, 3853–3862.

Huang, S. (2012a). The molecular and mathematical basis of Waddington's epigenetic landscape: A framework for post-Darwinian biology? BioEssays *34*, 149–157.

Huang, S. (2012b). Tumor progression: Chance and necessity in Darwinian and Lamarckian somatic (mutationless) evolution. Prog. Biophys. Mol. Biol. *110*, 69–86.

Huang, J., and Berger, S.L. (2008). The emerging field of dynamic lysine methylation of non-histone proteins. Curr. Opin. Genet. Dev. *18*, 152–158.

Huntley, S., Baggott, D.M., Hamilton, A.T., Tran-Gyamfi, M., Yang, S., Kim, J., Gordon, L., Branscomb, E., and Stubbs, L. (2006). A comprehensive catalog of human KRAB-associated zinc finger genes: insights into the evolutionary history of a large family of transcriptional repressors. Genome Res. *16*, 669–677.

Iqbal, K., Jin, S.-G., Pfeifer, G.P., and Szabó, P.E. (2011). Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. Proc. Natl. Acad. Sci. *108*, 3642–3647.

Isono, K., Endo, T.A., Ku, M., Yamada, D., Suzuki, R., Sharif, J., Ishikura, T., Toyoda, T., Bernstein, B.E., and Koseki, H. (2013). SAM Domain Polymerization Links Subnuclear Clustering of PRC1 to Gene Silencing. Dev. Cell *26*, 565–577.

Isono, K.-I., Fujimura, Y.-I., Shinga, J., Yamaki, M., O-Wang, J., Takihara, Y., Murahashi, Y., Takada, Y., Mizutani-Koseki, Y., and Koseki, H. (2005). Mammalian polyhomeotic homologues Phc2 and Phc1 act in synergy to mediate polycomb repression of Hox genes. Mol. Cell. Biol. *25*, 6694–6706.

Ito, S., D'Alessio, A.C., Taranova, O.V., Hong, K., Sowers, L.C., and Zhang, Y. (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature *466*, 1129–1133.

Iyengar, S., and Farnham, P.J. (2011). KAP1 protein: an enigmatic master regulator of the genome. J. Biol. Chem. *286*, 26267–26276.

Jackson, J.P., Lindroth, A.M., Cao, X., and Jacobsen, S.E. (2002). Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature *416*, 556–560.

Jackson, J.P., Johnson, L., Jasencakova, Z., Zhang, X., PerezBurgos, L., Singh, P.B., Cheng, X., Schubert, I., Jenuwein, T., and Jacobsen, S.E. (2004). Dimethylation of histone H3 lysine 9 is a critical mark for DNA methylation and gene silencing in Arabidopsis thaliana. Chromosoma *112*, 308–315.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science *293*, 1074–1080.

Jeon, Y., Sarma, K., and Lee, J.T. (2012). New and Xisting regulatory mechanisms of X chromosome inactivation. Curr. Opin. Genet. Dev. *22*, 62–71.

Jermann, P., Hoerner, L., Burger, L., and Schübeler, D. (2014). Short sequences can efficiently recruit histone H3 lysine 27 trimethylation in the absence of enhancer activity and DNA methylation. Proc. Natl. Acad. Sci. U. S. A. *111*, E3415–E3421.

Johnson, L.M., Bostick, M., Zhang, X., Kraft, E., Henderson, I., Callis, J., and Jacobsen, S.E. (2007). The SRA methyl-cytosine-binding domain links DNA and histone methylation. Curr. Biol. CB *17*, 379–384.

Jones, P.A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat. Rev. Genet. *13*, 484–492.

Jürgens, G. (1985). A group of genes controlling the spatial expression of the bithorax complex in Drosophila. Publ. Online 11 July 1985 Doi101038316153a0 *316*, 153–155.

Kanhere, A., Viiri, K., Araújo, C.C., Rasaiyaah, J., Bouwman, R.D., Whyte, W.A., Pereira, C.F., Brookes, E., Walker, K., Bell, G.W., et al. (2010). Short RNAs are transcribed from repressed Polycomb target genes and interact with Polycomb Repressive Complex-2. Mol. Cell *38*, 675–688.

Karimi, M.M., Goyal, P., Maksakova, I.A., Bilenky, M., Leung, D., Tang, J.X., Shinkai, Y., Mager, D.L., Jones, S., Hirst, M., et al. (2011). DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. Cell Stem Cell *8*, 676–687.

Kato, Y., Kato, M., Tachibana, M., Shinkai, Y., and Yamaguchi, M. (2008). Characterization of Drosophila G9a in vivo and identification of genetic interactants. Genes Cells Devoted Mol. Cell. Mech. *13*, 703–722.

Katoh-Fukui, Y., Tsuchiya, R., Shiroishi, T., Nakahara, Y., Hashimoto, N., Noguchi, K., and Higashinakagawa, T. (1998). Male-to-female sex reversal in M33 mutant mice. Nature *393*, 688–692.

Katz, A., Oliva, M., Mosquna, A., Hakim, O., and Ohad, N. (2004). FIE and CURLY LEAF polycomb proteins interact in the regulation of homeobox gene expression during sporophyte development. Plant J. *37*, 707–719.

Katz, D.J., Edwards, T.M., Reinke, V., and Kelly, W.G. (2009). A C. elegans LSD1 demethylase contributes to germline immortality by reprogramming epigenetic memory. Cell *137*, 308–320.

Kaustov, L., Ouyang, H., Amaya, M., Lemak, A., Nady, N., Duan, S., Wasney, G.A., Li, Z., Vedadi, M., Schapira, M., et al. (2011). Recognition and specificity determinants of the human cbx chromodomains. J. Biol. Chem. *286*, 521–529.

Kellum, R., and Alberts, B.M. (1995). Heterochromatin protein 1 is required for correct chromosome segregation in Drosophila embryos. J. Cell Sci. *108 ( Pt 4)*, 1419–1431.

Kellum, R., Raff, J.W., and Alberts, B.M. (1995). Heterochromatin protein 1 distribution during development and during the cell cycle in Drosophila embryos. J. Cell Sci. *108 ( Pt 4)*, 1407–1418.

Ketel, C.S., Andersen, E.F., Vargas, M.L., Suh, J., Strome, S., and Simon, J.A. (2005). Subunit Contributions to Histone Methyltransferase Activities of Fly and Worm Polycomb Group Complexes. Mol. Cell. Biol. *25*, 6857–6868.

Khalil, A.M., Boyar, F.Z., and Driscoll, D.J. (2004). Dynamic histone modifications mark sex chromosome inactivation and reactivation during mammalian spermatogenesis. Proc. Natl. Acad. Sci. U. S. A. *101*, 16583–16587.

Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., Kim, J., Aryee, M.J., Ji, H., Ehrlich, L.I.R., et al. (2010). Epigenetic memory in induced pluripotent stem cells. Nature *467*, 285–290.

Kim, S.Y., Lee, J., Eshed-Williams, L., Zilberman, D., and Sung, Z.R. (2012). EMF1 and PRC2 Cooperate to Repress Key Regulators of Arabidopsis Development. PLoS Genet *8*, e1002512.

Klattenhoff, C., Xi, H., Li, C., Lee, S., Xu, J., Khurana, J.S., Zhang, F., Schultz, N., Koppetsch, B.S., Nowosielska, A., et al. (2009). The Drosophila HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. Cell *138*, 1137–1149.

Klauke, K., Radulović, V., Broekhuis, M., Weersing, E., Zwart, E., Olthof, S., Ritsema, M., Bruggeman, S., Wu, X., Helin, K., et al. (2013). Polycomb Cbx family members mediate the balance between haematopoietic stem cell self-renewal and differentiation. Nat. Cell Biol. *15*, 353–362.

Klose, R.J., Kallin, E.M., and Zhang, Y. (2006). JmjC-domain-containing proteins and histone demethylation. Nat. Rev. Genet. *7*, 715–727.

Koch, C.M., Honemann-Capito, M., Egger-Adam, D., and Wodarz, A. (2009). Windei, the Drosophila Homolog of mAM/MCAF1, Is an Essential Cofactor of the H3K9 Methyl Transferase dSETDB1/Eggless in Germ Line Development. PLoS Genet *5*, e1000644.

Köhler, C., Hennig, L., Bouveret, R., Gheyselinck, J., Grossniklaus, U., and Gruissem, W. (2003). Arabidopsis MSI1 is a component of the MEA/FIE Polycomb group complex and required for seed development. EMBO J. *22*, 4804–4814.

Kotake, T., Takada, S., Nakahigashi, K., Ohto, M., and Goto, K. (2003). Arabidopsis TERMINAL FLOWER 2 gene encodes a heterochromatin protein 1 homolog and represses both FLOWERING LOCUS T to regulate flowering time and several floral homeotic genes. Plant Cell Physiol. *44*, 555–564.

Kouzarides, T. (2007a). Chromatin modifications and their function. Cell *128*, 693–705.

Kouzarides, T. (2007b). Chromatin modifications and their function. Cell *128*, 693–705.

Kouzminova, E., and Selker, E.U. (2001). dim-2 encodes a DNA methyltransferase responsible for all known cytosine methylation in Neurospora. EMBO J. *20*, 4309–4323.

Krauss, V. (2008). Glimpses of evolution: heterochromatic histone H3K9 methyltransferases left its marks behind. Genetica *133*, 93–106.

Ku, M., Koche, R.P., Rheinbay, E., Mendenhall, E.M., Endoh, M., Mikkelsen, T.S., Presser, A., Nusbaum, C., Xie, X., Chi, A.S., et al. (2008). Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. PLoS Genet. *4*, e1000242.

Kuhlmann, M., and Mette, M.F. (2012). Developmentally non-redundant SET domain proteins SUVH2 and SUVH9 are required for transcriptional gene silencing in Arabidopsis thaliana. Plant Mol. Biol. *79*, 623–633.

Kurimoto, K., Yabuta, Y., Ohinata, Y., Shigeta, M., Yamanaka, K., and Saitou, M. (2008). Complex genome-wide transcription dynamics orchestrated by Blimp1 for the specification of the germ cell lineage in mice. Genes Dev. *22*, 1617–1635.

Kurzhals, R.L., Tie, F., Stratton, C.A., and Harte, P.J. (2008). Drosophila ESC-Like can substitute for ESC and becomes required for Polycomb silencing if ESC is absent. Dev. Biol. *313*, 293–306.

Kuzmichev, A., Jenuwein, T., Tempst, P., and Reinberg, D. (2004). Different Ezh2-Containing Complexes Target Methylation of Histone H1 or Nucleosomal Histone H3. Mol. Cell *14*, 183–193.

Kuzmichev, A., Margueron, R., Vaquero, A., Preissner, T.S., Scher, M., Kirmizis, A., Ouyang, X., Brockdorff, N., Abate-Shen, C., Farnham, P., et al. (2005). Composition and histone substrates of polycomb repressive group complexes change during cellular differentiation. Proc. Natl. Acad. Sci. U. S. A. *102*, 1859–1864.

Kwon, S.H., and Workman, J.L. (2011). The changing faces of HP1: From heterochromatin formation and gene silencing to euchromatic gene expression. BioEssays *33*, 280–289.

Kwon, S.H., Florens, L., Swanson, S.K., Washburn, M.P., Abmayr, S.M., and Workman, J.L. (2010). Heterochromatin protein 1 (HP1) connects the FACT histone chaperone complex to the phosphorylated CTD of RNA polymerase II. Genes Dev. *24*, 2133–2145.

Kwong, C., Adryan, B., Bell, I., Meadows, L., Russell, S., Manak, J.R., and White, R. (2008). Stability and Dynamics of Polycomb Target Sites in Drosophila Development. PLoS Genet *4*, e1000178.

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature *410*, 116–120.

Ladewig, J., Koch, P., and Brüstle, O. (2013). Leveling Waddington: the emergence of direct programming and the loss of cell fate hierarchies. Nat. Rev. Mol. Cell Biol. *14*, 225–236.

Lagarou, A., Mohd-Sarip, A., Moshkin, Y.M., Chalkley, G.E., Bezstarosti, K., Demmers, J.A.A., and Verrijzer, C.P. (2008). dKDM2 couples histone H2A ubiquitylation to histone H3 demethylation during Polycomb group silencing. Genes Dev. *22*, 2799–2810.

Lan, Z.-J., Xu, X., and Cooney, A.J. (2004). Differential oocyte-specific expression of Cre recombinase activity in GDF-9-iCre, Zp3cre, and Msx2Cre transgenic mice. Biol. Reprod. *71*, 1469–1474.

Lapthanasupkul, P., Feng, J., Mantesso, A., Takada-Horisawa, Y., Vidal, M., Koseki, H., Wang, L., An, Z., Miletich, I., and Sharpe, P.T. (2012). Ring1a/b polycomb proteins regulate the mesenchymal stem cell niche in continuously growing incisors. Dev. Biol. *367*, 140–153.

Larsson, A.S., Landberg, K., and Meeks-Wagner, D.R. (1998). The TERMINAL FLOWER2 (TFL2) gene controls the reproductive transition and meristem identity in Arabidopsis thaliana. Genetics *149*, 597–605.

Laurent, L., Wong, E., Li, G., Huynh, T., Tsirigos, A., Ong, C.T., Low, H.M., Kin Sung, K.W., Rigoutsos, I., Loring, J., et al. (2010). Dynamic changes in the human methylome during differentiation. Genome Res. *20*, 320–331.

Law, J.A., and Jacobsen, S.E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat. Rev. Genet. *11*, 204–220.

Lawrence, P.A., and Levine, M. (2006). Mosaic and regulative development: two faces of one coin. Curr. Biol. *16*, R236–R239.

Lee, T.I., Jenner, R.G., Boyer, L.A., Guenther, M.G., Levine, S.S., Kumar, R.M., Chevalier, B., Johnstone, S.E., Cole, M.F., Isono, K., et al. (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. Cell *125*, 301–313.

Leeb, M., Pasini, D., Novatchkova, M., Jaritz, M., Helin, K., and Wutz, A. (2010). Polycomb complexes act redundantly to repress genomic repeats and genes. Genes Dev. *24*, 265–276.

Lehnertz, B., Ueda, Y., Derijck, A.A.H.A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T., and Peters, A.H.F.M. (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr. Biol. CB *13*, 1192–1200.

Leinhart, K., and Brown, M. (2011). SET/MYND Lysine Methyltransferases Regulate Gene Transcription and Protein Activity. Genes *2*, 210–218.

Leitch, H.G., and Smith, A. (2013). The mammalian germline as a pluripotency cycle. Development *140*, 2495–2501.

Leung, D.C., Dong, K.B., Maksakova, I.A., Goyal, P., Appanah, R., Lee, S., Tachibana, M., Shinkai, Y., Lehnertz, B., Mager, D.L., et al. (2011). Lysine methyltransferase G9a is required for de novo DNA

methylation and the establishment, but not the maintenance, of proviral silencing. Proc. Natl. Acad. Sci. U. S. A. *108*, 5718–5723.

Lewis, E.B. (1978). A gene complex controlling segmentation in Drosophila. Publ. Online 07 Dec. 1978 Doi101038276565a0 *276*, 565–570.

Lewis, P.H. (1947). Polycomb:PC. Drosoph. Inf. Serv. *21*, 69.

Lewis, Z.A., Honda, S., Khlafallah, T.K., Jeffress, J.K., Freitag, M., Mohn, F., Schübeler, D., and Selker, E.U. (2009). Relics of repeat-induced point mutation direct heterochromatin formation in Neurospora crassa. Genome Res. *19*, 427–437.

Lewis, Z.A., Adhvaryu, K.K., Honda, S., Shiver, A.L., and Selker, E.U. (2010). Identification of DIM-7, a protein required to target the DIM-5 H3 methyltransferase to chromatin. Proc. Natl. Acad. Sci. U. S. A. *107*, 8310–8315.

Li, L., Zheng, P., and Dean, J. (2010a). Maternal control of early mouse development. Development *137*, 859–870.

Li, Q., Wang, X., Lu, Z., Zhang, B., Guan, Z., Liu, Z., Zhong, Q., Gu, L., Zhou, J., Zhu, B., et al. (2010b). Polycomb CBX7 directly controls trimethylation of histone H3 at lysine 9 at the p16 locus. PloS One *5*, e13732.

Li, X., Isono, K.-I., Yamada, D., Endo, T.A., Endoh, M., Shinga, J., Mizutani-Koseki, Y., Otte, A.P., Casanova, M., Kitamura, H., et al. (2011). Mammalian polycomb-like Pcl2/Mtf2 is a novel regulatory component of PRC2 that can differentially modulate polycomb activity both at the Hox gene cluster and at Cdkn2a genes. Mol. Cell. Biol. *31*, 351–364.

Lin, C.-H., Li, B., Swanson, S., Zhang, Y., Florens, L., Washburn, M.P., Abmayr, S.M., and Workman, J.L. (2008). Heterochromatin protein 1a stimulates histone H3 lysine 36 demethylation by the Drosophila KDM4A demethylase. Mol. Cell *32*, 696–706.

Lin, C.-H., Paulson, A., Abmayr, S.M., and Workman, J.L. (2012). HP1a Targets the Drosophila KDM4A Demethylase to a Subset of Heterochromatic Genes to Regulate H3K36me3 Levels. PloS One *7*, e39758.

Liu, H., Kim, J.-M., and Aoki, F. (2004). Regulation of histone H3 lysine 9 methylation in oocytes and early pre-implantation embryos. Development *131*, 2269–2280.

Liu, Y., Taverna, S.D., Muratore, T.L., Shabanowitz, J., Hunt, D.F., and Allis, C.D. (2007). RNAi-dependent H3K27 methylation is required for heterochromatin formation and DNA elimination in Tetrahymena. Genes Dev. *21*, 1530–1545.

Lo, S.M., Follmer, N.E., Lengsfeld, B.M., Madamba, E.V., Seong, S., Grau, D.J., and Francis, N.J. (2012). A Bridging Model for Persistence of a Polycomb Group Protein Complex through DNA Replication In Vitro. Mol. Cell *46*, 784–796.

Loh, Y.-H., Zhang, W., Chen, X., George, J., and Ng, H.-H. (2007). Jmjd1a and Jmjd2c histone H3 Lys 9 demethylases regulate self-renewal in embryonic stem cells. Genes Dev. *21*, 2545–2557.

Lohmann, F., Loureiro, J., Su, H., Fang, Q., Lei, H., Lewis, T., Yang, Y., Labow, M., Li, E., Chen, T., et al. (2010). KMT1E mediated H3K9 methylation is required for the maintenance of embryonic stem cells by repressing trophectoderm differentiation. Stem Cells Dayt. Ohio *28*, 201–212.

Lorente, M. del M., Marcos-Gutierrez, C., Perez, C., Schoorlemmer, J., Ramirez, A., Magin, T., and Vidal, M. (2000). Loss- and gain-of-function mutations show a polycomb group function for Ring1A in mice. Development *127*, 5093–5100.

Lucia, F.D., Crevillen, P., Jones, A.M.E., Greb, T., and Dean, C. (2008). A PHD-Polycomb Repressive Complex 2 triggers the epigenetic silencing of FLC during vernalization. Proc. Natl. Acad. Sci. *105*, 16831–16836.

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

Van der Lugt, N.M., Domen, J., Linders, K., van Roon, M., Robanus-Maandag, E., te Riele, H., van der Valk, M., Deschamps, J., Sofroniew, M., and van Lohuizen, M. (1994). Posterior transformation, neurological abnormalities, and severe hematopoietic defects in mice with a targeted deletion of the bmi-1 proto-oncogene. Genes Dev. *8*, 757–769.

Luo, M., Bilodeau, P., Koltunow, A., Dennis, E.S., Peacock, W.J., and Chaudhury, A.M. (1999). Genes controlling fertilization-independent seed development in Arabidopsis thaliana. Proc. Natl. Acad. Sci. U. S. A. *96*, 296–301.

Magnani, L., Eeckhoute, J., and Lupien, M. (2011). Pioneer factors: directing transcriptional regulators within the chromatin environment. Trends Genet. *27*, 465–474.

Magnúsdóttir, E., Dietmann, S., Murakami, K., Günesdogan, U., Tang, F., Bao, S., Diamanti, E., Lao, K., Gottgens, B., and Azim Surani, M. (2013). A tripartite transcription factor network regulates primordial germ cell specification in mice. Nat. Cell Biol. *15*, 905–915.

Maison, C., Bailly, D., Roche, D., Montes de Oca, R., Probst, A.V., Vassias, I., Dingli, F., Lombard, B., Loew, D., Quivy, J.-P., et al. (2011). SUMOylation promotes de novo targeting of HP1α to pericentric heterochromatin. Nat. Genet. *43*, 220–227.

Mak, W., Nesterova, T.B., Napoles, M. de, Appanah, R., Yamanaka, S., Otte, A.P., and Brockdorff, N. (2004). Reactivation of the Paternal X Chromosome in Early Mouse Embryos. Science *303*, 666–669.

Maksakova, I.A., Goyal, P., Bullwinkel, J., Brown, J.P., Bilenky, M., Mager, D.L., Singh, P.B., and Lorincz, M.C. (2011). H3K9me3-binding proteins are dispensable for SETDB1/H3K9me3-dependent retroviral silencing. Epigenetics Chromatin *4*, 12.

Mansour, A.A., Gafni, O., Weinberger, L., Zviran, A., Ayyash, M., Rais, Y., Krupalnik, V., Zerbib, M., Amann-Zalcenstein, D., Maza, I., et al. (2012). The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming. Nature *488*, 409–413.

Margueron, R., and Reinberg, D. (2011). The Polycomb complex PRC2 and its mark in life. Nature *469*, 343–349.

Margueron, R., Li, G., Sarma, K., Blais, A., Zavadil, J., Woodcock, C.L., Dynlacht, B.D., and Reinberg, D. (2008). Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. Mol. Cell *32*, 503–518.

Margueron, R., Justin, N., Ohno, K., Sharpe, M.L., Son, J., Drury, W.J., 3rd, Voigt, P., Martin, S.R., Taylor, W.R., De Marco, V., et al. (2009). Role of the polycomb protein EED in the propagation of repressive histone marks. Nature *461*, 762–767.

Matsui, T., Leung, D., Miyashita, H., Maksakova, I.A., Miyachi, H., Kimura, H., Tachibana, M., Lorincz, M.C., and Shinkai, Y. (2010). Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. Nature *464*, 927–931.

Mayer, W., Smith, A., Fundele, R., and Haaf, T. (2000). Spatial separation of parental genomes in preimplantation mouse embryos. J. Cell Biol. *148*, 629–634.

McQuilton, P., St Pierre, S.E., and Thurmond, J. (2012). FlyBase 101--the basics of navigating FlyBase. Nucleic Acids Res. *40*, D706–D714.

Meglicki, M., Teperek-Tkacz, M., and Borsuk, E. (2012). Appearance and heterochromatin localization of HP1α in early mouse embryos depends on cytoplasmic clock and H3S10 phosphorylation. Cell Cycle Georget. Tex *11*, 2189–2205.

Meglio, T.D., Kratochwil, C.F., Vilain, N., Loche, A., Vitobello, A., Yonehara, K., Hrycaj, S.M., Roska, B., Peters, A.H.F.M., Eichmann, A., et al. (2013). Ezh2 Orchestrates Topographic Migration and Connectivity of Mouse Precerebellar Neurons. Science *339*, 204–207.

Meilhac, S.M., Adams, R.J., Morris, S.A., Danckaert, A., Le Garrec, J.-F., and Zernicka-Goetz, M. (2009). Active cell movements coupled to positional induction are involved in lineage segregation in the mouse blastocyst. Dev. Biol. *331*, 210–221.

Meister, P., Schott, S., Bedet, C., Xiao, Y., Rohner, S., Bodennec, S., Hudry, B., Molin, L., Solari, F., Gasser, S.M., et al. (2011). Caenorhabditis elegans Heterochromatin protein 1 (HPL-2) links developmental plasticity, longevity and lipid metabolism. Genome Biol. *12*, R123.

Mendenhall, E.M., Koche, R.P., Truong, T., Zhou, V.W., Issac, B., Chi, A.S., Ku, M., and Bernstein, B.E. (2010). GC-rich sequence elements recruit PRC2 in mammalian ES cells. PLoS Genet. *6*, e1001244.

Metzler-Guillemain, C., Luciani, J., Depetris, D., Guichaoua, M.R., and Mattei, M.G. (2003). HP1beta and HP1gamma, but not HP1alpha, decorate the entire XY body during human male meiosis. Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol. *11*, 73–81.

Minc, E., Allory, Y., Worman, H.J., Courvalin, J.C., and Buendia, B. (1999). Localization and phosphorylation of HP1 proteins during the cell cycle in mammalian cells. Chromosoma *108*, 220–234.

Mis, J., Ner, S.S., and Grigliatti, T.A. (2006). Identification of three histone methyltransferases in Drosophila: dG9a is a suppressor of PEV and is required for gene silencing. Mol. Genet. Genomics MGG *275*, 513–526.

Miura, A., Nakamura, M., Inagaki, S., Kobayashi, A., Saze, H., and Kakutani, T. (2009). An Arabidopsis jmjC domain protein protects transcribed genes from DNA methylation at CHG sites. EMBO J. *28*, 1078–1086.

Moazed, D. (2011). Mechanisms for the Inheritance of Chromatin States. Cell *146*, 510–518.

Mohn, F., Weber, M., Rebhan, M., Roloff, T.C., Richter, J., Stadler, M.B., Bibel, M., and Schübeler, D. (2008). Lineage-Specific Polycomb Targets and De Novo DNA Methylation Define Restriction and Potential of Neuronal Progenitors. Mol. Cell *30*, 755–766.

Müller, J., Hart, C.M., Francis, N.J., Vargas, M.L., Sengupta, A., Wild, B., Miller, E.L., O'Connor, M.B., Kingston, R.E., and Simon, J.A. (2002). Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. Cell *111*, 197–208.

Nakahigashi, K., Jasencakova, Z., Schubert, I., and Goto, K. (2005). The Arabidopsis heterochromatin protein1 homolog (TERMINAL FLOWER2) silences genes within the euchromatic region but not genes positioned in heterochromatin. Plant Cell Physiol. *46*, 1747–1756.

Nakamura, T., Liu, Y.-J., Nakashima, H., Umehara, H., Inoue, K., Matoba, S., Tachibana, M., Ogura, A., Shinkai, Y., and Nakano, T. (2012). PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. Nature *486*, 415–419.

Naruse, C., Fukusumi, Y., Kakiuchi, D., and Asano, M. (2007). A novel gene trapping for identifying genes expressed under the control of specific transcription factors. Biochem. Biophys. Res. Commun. *361*, 109–115.

Naumann, K., Fischer, A., Hofmann, I., Krauss, V., Phalke, S., Irmler, K., Hause, G., Aurich, A.-C., Dorn, R., Jenuwein, T., et al. (2005). Pivotal role of AtSUVH2 in heterochromatic histone methylation and gene silencing in Arabidopsis. EMBO J. *24*, 1418–1429.

Nekrasov, M., Klymenko, T., Fraterman, S., Papp, B., Oktaba, K., Köcher, T., Cohen, A., Stunnenberg, H.G., Wilm, M., and Müller, J. (2007). Pcl-PRC2 is needed to generate high levels of H3-K27 trimethylation at Polycomb target genes. EMBO J. *26*, 4078–4088.

Nestorov, P., Tardat, M., and Peters, A.H.F.M. (2013a). H3K9/HP1 and Polycomb: Two Key Epigenetic Silencing Pathways for Gene Regulation and Embryo Development. Curr. Top. Dev. Biol. *104*, 243–291.

Nestorov, P., Battke, F., Levesque, M.P., and Gerberding, M. (2013b). The maternal transcriptome of the crustacean Parhyale hawaiensis is inherited asymmetrically to invariant cell lineages of the ectoderm and mesoderm. PloS One *8*, e56049.

Ng, R.K., Dean, W., Dawson, C., Lucifero, D., Madeja, Z., Reik, W., and Hemberger, M. (2008). Epigenetic restriction of embryonic cell lineage fate by methylation of Elf5. Nat. Cell Biol. *10*, 1280–1290.

Ng, S.-Y., Johnson, R., and Stanton, L.W. (2012). Human long non-coding RNAs promote pluripotency and neuronal differentiation by association with chromatin modifiers and transcription factors. EMBO J. *31*, 522–533.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Schöler, H., and Smith, A. (1998). Formation of Pluripotent Stem Cells in the Mammalian Embryo Depends on the POU Transcription Factor Oct4. Cell *95*, 379–391.

Nishio, H., and Walsh, M.J. (2004). CCAAT displacement protein/cut homolog recruits G9a histone lysine methyltransferase to repress transcription. Proc. Natl. Acad. Sci. U. S. A. *101*, 11257–11262.

Nishioka, N., Inoue, K., Adachi, K., Kiyonari, H., Ota, M., Ralston, A., Yabuta, N., Hirahara, S., Stephenson, R.O., Ogonuki, N., et al. (2009). The Hippo signaling pathway components Lats and Yap pattern Tead4 activity to distinguish mouse trophectoderm from inner cell mass. Dev. Cell *16*, 398–410.

Nowak, A.J., Alfieri, C., Stirnimann, C.U., Rybin, V., Baudin, F., Ly-Hartig, N., Lindner, D., and Müller, C.W. (2011). Chromatin-modifying complex component Nurf55/p55 associates with histones H3 and H4 and polycomb repressive complex 2 subunit Su(z)12 through partially overlapping binding sites. J. Biol. Chem. *286*, 23388–23396.

O'Carroll, D., Scherthan, H., Peters, A.H., Opravil, S., Haynes, A.R., Laible, G., Rea, S., Schmid, M., Lebersorger, A., Jerratsch, M., et al. (2000). Isolation and characterization of Suv39h2, a second histone H3 methyltransferase gene that displays testis-specific expression. Mol. Cell. Biol. *20*, 9423–9433.

O'Carroll, D., Erhardt, S., Pagani, M., Barton, S.C., Surani, M.A., and Jenuwein, T. (2001). The polycomb-group gene Ezh2 is required for early mouse development. Mol. Cell. Biol. *21*, 4330–4336.

Ohad, N., Yadegari, R., Margossian, L., Hannon, M., Michaeli, D., Harada, J.J., Goldberg, R.B., and Fischer, R.L. (1999). Mutations in FIE, a WD polycomb group gene, allow endosperm development without fertilization. Plant Cell *11*, 407–416.

Ohnishi, Y., Huber, W., Tsumura, A., Kang, M., Xenopoulos, P., Kurimoto, K., Oleś, A.K., Araúzo-Bravo, M.J., Saitou, M., Hadjantonakis, A.-K., et al. (2014). Cell-to-cell expression variability followed by signal reinforcement progressively segregates early mouse lineages. Nat. Cell Biol. *16*, 27–37.

Ohno, K., McCabe, D., Czermin, B., Imhof, A., and Pirrotta, V. (2008). ESC, ESCL and their roles in Polycomb Group mechanisms. Mech. Dev. *125*, 527–541.

Okada, Y., Scott, G., Ray, M.K., Mishina, Y., and Zhang, Y. (2007). Histone demethylase JHDM2A is critical for Tnp1 and Prm1 transcription and spermatogenesis. Nature *450*, 119–123.

Okada, Y., Tateishi, K., and Zhang, Y. (2010). Histone demethylase JHDM2A is involved in male infertility and obesity. J. Androl. *31*, 75–78.

Okamoto, I., Otte, A.P., Allis, C.D., Reinberg, D., and Heard, E. (2004). Epigenetic Dynamics of Imprinted X Inactivation During Early Mouse Development. Science *303*, 644–649.

Onder, T.T., Kara, N., Cherry, A., Sinha, A.U., Zhu, N., Bernt, K.M., Cahan, P., Marcarci, B.O., Unternaehrer, J., Gupta, P.B., et al. (2012). Chromatin-modifying enzymes as modulators of reprogramming. Nature *483*, 598–602.

Ooga, M., Inoue, A., Kageyama, S., Akiyama, T., Nagata, M., and Aoki, F. (2008). Changes in H3K79 methylation during preimplantation development in mice. Biol. Reprod. *78*, 413–424.

Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W., and Walter, J. (2000). Active demethylation of the paternal genome in the mouse zygote. Curr. Biol. CB *10*, 475–478.

Paik, W.K., Paik, D.C., and Kim, S. (2007). Historical review: the field of protein methylation. Trends Biochem. Sci. *32*, 146–152.

Palacios, I.M., and Johnston, D.S. (2001). GETTING THE MESSAGE ACROSS: The Intracellular Localization of mRNAs in Higher Eukaryotes. Annu. Rev. Cell Dev. Biol. *17*, 569–614.

Pal-Bhadra, M., Leibovitch, B.A., Gandhi, S.G., Rao, M., Bhadra, U., Birchler, J.A., and Elgin, S.C.R. (2004). Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery. Science *303*, 669–672.

Pan, H., O'brien, M.J., Wigglesworth, K., Eppig, J.J., and Schultz, R.M. (2005). Transcript profiling during mouse oocyte development and the effect of gonadotropin priming and development in vitro. Dev. Biol. *286*, 493–506.

Park, S.-J., Komata, M., Inoue, F., Yamada, K., Nakai, K., Ohsugi, M., and Shirahige, K. (2013a). Inferring the choreography of parental genomes during fertilization from ultralarge-scale whole-transcriptome analysis. Genes Dev. *27*, 2736–2748.

Park, S.-J., Komata, M., Inoue, F., Yamada, K., Nakai, K., Ohsugi, M., and Shirahige, K. (2013b). Inferring the choreography of parental genomes during fertilization from ultralarge-scale whole-transcriptome analysis. Genes Dev. *27*, 2736–2748.

Pasini, D., Bracken, A.P., Jensen, M.R., Lazzerini Denchi, E., and Helin, K. (2004). Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. EMBO J. *23*, 4061–4071.

Pasini, D., Bracken, A.P., Hansen, J.B., Capillo, M., and Helin, K. (2007). The polycomb group protein Suz12 is required for embryonic stem cell differentiation. Mol. Cell. Biol. *27*, 3769–3779.

Patrat, C., Okamoto, I., Diabangouaya, P., Vialon, V., Baccon, P.L., Chow, J., and Heard, E. (2009). Dynamic changes in paternal X-chromosome activity during imprinted X-chromosome inactivation in mice. Proc. Natl. Acad. Sci. *106*, 5198–5203.

Pengelly, A.R., Copur, Ö., Jäckle, H., Herzig, A., and Müller, J. (2013). A histone mutant reproduces the phenotype caused by loss of histone-modifying factor Polycomb. Science *339*, 698–699.

Peters, A.H., O'Carroll, D., Scherthan, H., Mechtler, K., Sauer, S., Schöfer, C., Weipoltshammer, K., Pagani, M., Lachner, M., Kohlmaier, A., et al. (2001). Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. Cell *107*, 323–337.

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

Phillips, M.D., and Shearn, A. (1990). Mutations in polycombeotic, a Drosophila polycomb-group gene, cause a wide range of maternal and zygotic phenotypes. Genetics *125*, 91–101.

Piacentini, L., Fanti, L., Negri, R., Del Vescovo, V., Fatica, A., Altieri, F., and Pimpinelli, S. (2009). Heterochromatin protein 1 (HP1a) positively regulates euchromatic gene expression through RNA transcript association and interaction with hnRNPs in Drosophila. PLoS Genet. *5*, e1000670.

Pikó, L., and Clegg, K.B. (1982). Quantitative changes in total RNA, total poly(A), and ribosomes in early mouse embryos. Dev. Biol. *89*, 362–378.

Pinheiro, I., Margueron, R., Shukeir, N., Eisold, M., Fritzsch, C., Richter, F.M., Mittler, G., Genoud, C., Goyama, S., Kurokawa, M., et al. (2012). Prdm3 and Prdm16 are H3K9me1 Methyltransferases Required for Mammalian Heterochromatin Integrity. Cell *150*, 948–960.

Piotrowska-Nitsche, K., and Zernicka-Goetz, M. (2005). Spatial arrangement of individual 4-cell stage blastomeres and the order in which they are generated correlate with blastocyst pattern in the mouse embryo. Mech. Dev. *122*, 487–500.

Pirity, M.K., Locker, J., and Schreiber-Agus, N. (2005). Rybp/DEDAF is required for early postimplantation and for central nervous system development. Mol. Cell. Biol. *25*, 7193–7202.

Pirrotta, V., and Li, H.-B. (2012). A view of nuclear Polycomb bodies. Curr. Opin. Genet. Dev. *22*, 101–109.

Plath, K., Fang, J., Mlynarczyk-Evans, S.K., Cao, R., Worringer, K.A., Wang, H., de la Cruz, C.C., Otte, A.P., Panning, B., and Zhang, Y. (2003). Role of histone H3 lysine 27 methylation in X inactivation. Science *300*, 131–135.

Plusa, B., Piliszek, A., Frankenberg, S., Artus, J., and Hadjantonakis, A.-K. (2008). Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. Dev. Camb. Engl. *135*, 3081–3091.

Posfai, E., Kunzmann, R., Brochard, V., Salvaing, J., Cabuy, E., Roloff, T.C., Liu, Z., Tardat, M., Lohuizen, M. van, Vidal, M., et al. (2012). Polycomb function during oogenesis is required for mouse embryonic development. Genes Dev. *26*, 920–932.

Probst, A.V., Santos, F., Reik, W., Almouzni, G., and Dean, W. (2007). Structural differences in centromeric heterochromatin are spatially reconciled on fertilisation in the mouse zygote. Chromosoma *116*, 403–415.

Probst, A.V., Dunleavy, E., and Almouzni, G. (2009). Epigenetic inheritance during the cell cycle. Nat. Rev. Mol. Cell Biol. *10*, 192–206.

Ptashne, M., and Gann, A. (2002). Genes & Signals (Cold Spring Harbor Laboratory Press).

Puschendorf, M., Terranova, R., Boutsma, E., Mao, X., Isono, K., Brykczynska, U., Kolb, C., Otte, A.P., Koseki, H., Orkin, S.H., et al. (2008). PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos. Nat. Genet. *40*, 411–420.

Qin, J., Whyte, W.A., Anderssen, E., Apostolou, E., Chen, H.-H., Akbarian, S., Bronson, R.T., Hochedlinger, K., Ramaswamy, S., Young, R.A., et al. (2012). The Polycomb Group Protein L3mbtl2 Assembles an Atypical PRC1-Family Complex that Is Essential in Pluripotent Stem Cells and Early Development. Cell Stem Cell.

Quenneville, S., Verde, G., Corsinotti, A., Kapopoulou, A., Jakobsson, J., Offner, S., Baglivo, I., Pedone, P.V., Grimaldi, G., Riccio, A., et al. (2011). In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. Mol. Cell *44*, 361–372.

Quenneville, S., Turelli, P., Bojkowska, K., Raclot, C., Offner, S., Kapopoulou, A., and Trono, D. (2012). The KRAB-ZFP/KAP1 System Contributes to the Early Embryonic Establishment of Site-Specific DNA Methylation Patterns Maintained during Development. Cell Rep. *2*, 766–773.

Rando, O.J. (2012). Combinatorial complexity in chromatin structure and function: revisiting the histone code. Curr. Opin. Genet. Dev. *22*, 148–155.

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., et al. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature *406*, 593–599.

Ren, X., and Kerppola, T.K. (2011). REST interacts with Cbx proteins and regulates polycomb repressive complex 1 occupancy at RE1 elements. Mol. Cell. Biol. *31*, 2100–2110.

Richly, H., Aloia, L., and Di Croce, L. (2011). Roles of the Polycomb group proteins in stem cells and cancer. Cell Death Dis. *2*, e204.

Ringrose, L., and Paro, R. (2007). Polycomb/Trithorax response elements and epigenetic memory of cell identity. Dev. Camb. Engl. *134*, 223–232.

Rinn, J.L., Kertesz, M., Wang, J.K., Squazzo, S.L., Xu, X., Brugmann, S.A., Goodnough, L.H., Helms, J.A., Farnham, P.J., Segal, E., et al. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell *129*, 1311–1323.

Rosnoblet, C., Vandamme, J., Völkel, P., and Angrand, P.-O. (2011). Analysis of the human HP1 interactome reveals novel binding partners. Biochem. Biophys. Res. Commun. *413*, 206–211.

Rossant, J., and Tam, P.P.L. (2009). Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse. Development *136*, 701–713.

Rothbart, S.B., Krajewski, K., Nady, N., Tempel, W., Xue, S., Badeaux, A.I., Barsyte-Lovejoy, D., Martinez, J.Y., Bedford, M.T., Fuchs, S.M., et al. (2012). Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation. Nat. Struct. Mol. Biol.

Rowe, H.M., Jakobsson, J., Mesnard, D., Rougemont, J., Reynard, S., Aktas, T., Maillard, P.V., Layard-Liesching, H., Verp, S., Marquis, J., et al. (2010). KAP1 controls endogenous retroviruses in embryonic stem cells. Nature *463*, 237–240.

Roy, S., Ernst, J., Kharchenko, P.V., Kheradpour, P., Negre, N., Eaton, M.L., Landolin, J.M., Bristow, C.A., Ma, L., Lin, M.F., et al. (2010). Identification of Functional Elements and Regulatory Circuits by Drosophila modENCODE. Science *330*, 1787–1797.

Sachs, M., Onodera, C., Blaschke, K., Ebata, K.T., Song, J.S., and Ramalho-Santos, M. (2013). Bivalent Chromatin Marks Developmental Regulatory Genes in the Mouse Embryonic Germline In Vivo. Cell Rep.

Saha, B., Home, P., Ray, S., Larson, M., Paul, A., Rajendran, G., Behr, B., and Paul, S. (2013). EED and KDM6B Coordinate First Mammalian Cell Lineage Commitment to Ensure Embryo Implantation. Mol. Cell. Biol.

Saitou, M., and Yamaji, M. (2012). Primordial Germ Cells in Mice. Cold Spring Harb. Perspect. Biol. *4*, a008375.

Saitou, M., Kagiwada, S., and Kurimoto, K. (2012). Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells. Dev. Camb. Engl. *139*, 15–31.

Sampath, S.C., Marazzi, I., Yap, K.L., Sampath, S.C., Krutchinsky, A.N., Mecklenbräuker, I., Viale, A., Rudensky, E., Zhou, M.-M., Chait, B.T., et al. (2007). Methylation of a histone mimic within the histone methyltransferase G9a regulates protein complex assembly. Mol. Cell *27*, 596–608.

Sanchez-Pulido, L., Devos, D., Sung, Z.R., and Calonje, M. (2008). RAWUL: A new ubiquitin-like domain in PRC1 Ring finger proteins that unveils putative plant and worm PRC1 orthologs. BMC Genomics *9*, 308.

Santos, F., Hendrich, B., Reik, W., and Dean, W. (2002). Dynamic reprogramming of DNA methylation in the early mouse embryo. Dev. Biol. *241*, 172–182.

Santos, F., Peters, A.H., Otte, A.P., Reik, W., and Dean, W. (2005). Dynamic chromatin modifications characterise the first cell cycle in mouse embryos. Dev. Biol. *280*, 225–236.

Sarraf, S.A., and Stancheva, I. (2004). Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. Mol. Cell *15*, 595–605.

Sasaki, T., Kobayashi, A., Saze, H., and Kakutani, T. (2012). RNAi-independent de novo DNA methylation revealed in Arabidopsis mutants of chromatin remodeling gene DDM1. Plant J. Cell Mol. Biol. *70*, 750–758.

Savla, U., Benes, J., Zhang, J., and Jones, R.S. (2008). Recruitment of Drosophila Polycomb-group proteins by Polycomblike, a component of a novel protein complex in larvae. Dev. Camb. Engl. *135*, 813–817.

Saze, H., Shiraishi, A., Miura, A., and Kakutani, T. (2008). Control of genic DNA methylation by a jmjC domain-containing protein in Arabidopsis thaliana. Science *319*, 462–465.

Scheuermann, J.C., de Ayala Alonso, A.G., Oktaba, K., Ly-Hartig, N., McGinty, R.K., Fraterman, S., Wilm, M., Muir, T.W., and Müller, J. (2010). Histone H2A deubiquitinase activity of the Polycomb repressive complex PR-DUB. Nature *465*, 243–247.

Schmitges, F.W., Prusty, A.B., Faty, M., Stützer, A., Lingaraju, G.M., Aiwazian, J., Sack, R., Hess, D., Li, L., Zhou, S., et al. (2011). Histone methylation by PRC2 is inhibited by active chromatin marks. Mol. Cell *42*, 330–341.

Schotta, G., Ebert, A., Krauss, V., Fischer, A., Hoffmann, J., Rea, S., Jenuwein, T., Dorn, R., and Reuter, G. (2002). Central role of Drosophila SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. EMBO J. *21*, 1121–1131.

Schotta, G., Ebert, A., and Reuter, G. (2003). SU(VAR)3-9 is a conserved key function in heterochromatic gene silencing. Genetica *117*, 149–158.

Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B., and Cavalli, G. (2007). Genome regulation by polycomb and trithorax proteins. Cell *128*, 735–745.

Schuettengruber, B., Ganapathi, M., Leblanc, B., Portoso, M., Jaschek, R., Tolhuis, B., van Lohuizen, M., Tanay, A., and Cavalli, G. (2009). Functional Anatomy of Polycomb and Trithorax Chromatin Landscapes in Drosophila Embryos. PLoS Biol *7*, e1000013.

Schultz, R.M. (2002). The molecular foundations of the maternal to zygotic transition in the preimplantation embryo. Hum. Reprod. Update *8*, 323–331.

Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G., and Rauscher, F.J., 3rd (2002). SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. Genes Dev. *16*, 919–932.

Schwaiger, M., Kohler, H., Oakeley, E.J., Stadler, M.B., and Schübeler, D. (2010). Heterochromatin protein 1 (HP1) modulates replication timing of the Drosophila genome. Genome Res. *20*, 771–780.

Schwartz, Y.B., and Pirrotta, V. (2013). A new world of Polycombs: unexpected partnerships and emerging functions. Nat. Rev. Genet. *14*, 853–864.

Schwartz, Y.B., Kahn, T.G., Nix, D.A., Li, X.-Y., Bourgon, R., Biggin, M., and Pirrotta, V. (2006). Genome-wide analysis of Polycomb targets in Drosophila melanogaster. Nat. Genet. *38*, 700–705.

Seki, Y., Yamaji, M., Yabuta, Y., Sano, M., Shigeta, M., Matsui, Y., Saga, Y., Tachibana, M., Shinkai, Y., and Saitou, M. (2007). Cellular dynamics associated with the genome-wide epigenetic reprogramming in migrating primordial germ cells in mice. Development *134*, 2627–2638.

Seum, C., Bontron, S., Reo, E., Delattre, M., and Spierer, P. (2007a). Drosophila G9a is a nonessential gene. Genetics *177*, 1955–1957.

Seum, C., Reo, E., Peng, H., Rauscher, F.J., 3rd, Spierer, P., and Bontron, S. (2007b). Drosophila SETDB1 is required for chromosome 4 silencing. PLoS Genet. *3*, e76.

Sewalt, R.G.A.B., Lachner, M., Vargas, M., Hamer, K.M., den Blaauwen, J.L., Hendrix, T., Melcher, M., Schweizer, D., Jenuwein, T., and Otte, A.P. (2002). Selective interactions between vertebrate polycomb homologs and the SUV39H1 histone lysine methyltransferase suggest that histone H3-K9 methylation contributes to chromosomal targeting of Polycomb group proteins. Mol. Cell. Biol. *22*, 5539–5553.

Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T.A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K., et al. (2007). The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature *450*, 908–912.

Shaver, S., Casas-Mollano, J.A., Cerny, R.L., and Cerutti, H. (2010). Origin of the polycomb repressive complex 2 and gene silencing by an E(z) homolog in the unicellular alga Chlamydomonas. Epigenetics Off. J. DNA Methylation Soc. *5*, 301–312.

Shen, X., Liu, Y., Hsu, Y.-J., Fujiwara, Y., Kim, J., Mao, X., Yuan, G.-C., and Orkin, S.H. (2008). EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. Mol. Cell *32*, 491–502.

Simon, J.A., and Kingston, R.E. (2009). Mechanisms of Polycomb gene silencing: knowns and unknowns. Nat. Rev. Mol. Cell Biol. *10*, 697–708.

Sing, A., Pannell, D., Karaiskakis, A., Sturgeon, K., Djabali, M., Ellis, J., Lipshitz, H.D., and Cordes, S.P. (2009). A vertebrate Polycomb response element governs segmentation of the posterior hindbrain. Cell *138*, 885–897.

Slifer, E.H. (1942). A mutant stock of Drosophila with extra sex-combs. J. Exp. Zool. *90*, 31–40.

Smallwood, A., Estève, P.-O., Pradhan, S., and Carey, M. (2007). Functional cooperation between HP1 and DNMT1 mediates gene silencing. Genes Dev. *21*, 1169–1178.

Smallwood, A., Hon, G.C., Jin, F., Henry, R.E., Espinosa, J.M., and Ren, B. (2012). CBX3 regulates efficient RNA processing genome-wide. Genome Res. *22*, 1426–1436.

Smith, Z.D., Chan, M.M., Mikkelsen, T.S., Gu, H., Gnirke, A., Regev, A., and Meissner, A. (2012). A unique regulatory phase of DNA methylation in the early mammalian embryo. Nature *484*, 339–344.

Song, J., Angel, A., Howard, M., and Dean, C. (2012). Vernalization – a cold-induced epigenetic switch. J. Cell Sci.

Song, J.-J., Garlick, J.D., and Kingston, R.E. (2008). Structural basis of histone H4 recognition by p55. Genes Dev. *22*, 1313–1318.

Sonne-Hansen, K., Nielsen, M., and Byskov, A.G. (2003). Oocyte number in newborn mice after prenatal octylphenol exposure. Reprod. Toxicol. *17*, 59–66.

Sripathy, S.P., Stevens, J., and Schultz, D.C. (2006). The KAP1 corepressor functions to coordinate the assembly of de novo HP1-demarcated microenvironments of heterochromatin required for KRAB zinc finger protein-mediated transcriptional repression. Mol. Cell. Biol. *26*, 8623–8638.

Stabell, M., Eskeland, R., Bjørkmo, M., Larsson, J., Aalen, R.B., Imhof, A., and Lambertsson, A. (2006). The Drosophila G9a gene encodes a multi-catalytic histone methyltransferase required for normal development. Nucleic Acids Res. *34*, 4609–4621.

Van Steensel, B. (2011). Chromatin: constructing the big picture. EMBO J. *30*, 1885–1895.

Struhl, G., and Brower, D. (1982). Early role of the esc+ gene product in the determination of segments in Drosophila. Cell *31*, 285–292.

Strumpf, D., Mao, C.-A., Yamanaka, Y., Ralston, A., Chawengsaksophak, K., Beck, F., and Rossant, J. (2005). Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. Dev. Camb. Engl. *132*, 2093–2102.

Studencka, M., Konzer, A., Moneron, G., Wenzel, D., Opitz, L., Salinas-Riester, G., Bedet, C., Krüger, M., Hell, S.W., Wisniewski, J.R., et al. (2012). Novel roles of Caenorhabditis elegans heterochromatin protein HP1 and linker histone in the regulation of innate immune gene expression. Mol. Cell. Biol. *32*, 251–265.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. *102*, 15545–15550.

Sugimoto, M., and Abe, K. (2007). X chromosome reactivation initiates in nascent primordial germ cells in mice. PLoS Genet. *3*, e116.

Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011). REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. PLoS ONE *6*, e21800.

Tabansky, I., Lenarcic, A., Draft, R.W., Loulier, K., Keskin, D.B., Rosains, J., Rivera-Feliciano, J., Lichtman, J.W., Livet, J., Stern, J.N.H., et al. (2013). Developmental bias in cleavage-stage mouse blastomeres. Curr. Biol. CB *23*, 21–31.

Taberlay, P.C., Kelly, T.K., Liu, C.-C., You, J.S., De Carvalho, D.D., Miranda, T.B., Zhou, X.J., Liang, G., and Jones, P.A. (2011). Polycomb-repressed genes have permissive enhancers that initiate reprogramming. Cell *147*, 1283–1294.

Tachibana, M., Sugimoto, K., Nozaki, M., Ueda, J., Ohta, T., Ohki, M., Fukuda, M., Takeda, N., Niida, H., Kato, H., et al. (2002). G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. Genes Dev. *16*, 1779–1791.

Tachibana, M., Nozaki, M., Takeda, N., and Shinkai, Y. (2007). Functional dynamics of H3K9 methylation during meiotic prophase progression. EMBO J. *26*, 3346–3359.

Tachibana, M., Matsumura, Y., Fukuda, M., Kimura, H., and Shinkai, Y. (2008). G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. EMBO J. *27*, 2681–2690.

Tadros, W., and Lipshitz, H.D. (2009). The maternal-to-zygotic transition: a play in two acts. Development *136*, 3033–3042.

Takada, Y., Naruse, C., Costa, Y., Shirakawa, T., Tachibana, M., Sharif, J., Kezuka-Shiotani, F., Kakiuchi, D., Masumoto, H., Shinkai, Y., et al. (2011). HP1γ links histone methylation marks to meiotic synapsis in mice. Dev. Camb. Engl. *138*, 4207–4217.

Takagi, N., and Sasaki, M. (1975). Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. Nature *256*, 640–642.

Takahashi, K., and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. Cell *126*, 663–676.

Takihara, Y., Tomotsune, D., Shirai, M., Katoh-Fukui, Y., Nishii, K., Motaleb, M.A., Nomura, M., Tsuchiya, R., Fujita, Y., Shibata, Y., et al. (1997). Targeted disruption of the mouse homologue of the Drosophila polyhomeotic gene leads to altered anteroposterior patterning and neural crest defects. Dev. Camb. Engl. *124*, 3673–3682.

Tamaru, H., and Selker, E.U. (2001). A histone H3 methyltransferase controls DNA methylation in Neurospora crassa. Nature *414*, 277–283.

Tan, M.H., Au, K.F., Leong, D.E., Foygel, K., Wong, W.H., and Yao, M.W. (2013). An Oct4-Sall4-Nanog network controls developmental progression in the pre-implantation mouse embryo. Mol. Syst. Biol. *9*.

Tavares, L., Dimitrova, E., Oxley, D., Webster, J., Poot, R., Demmers, J., Bezstarosti, K., Taylor, S., Ura, H., Koide, H., et al. (2012). RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3. Cell *148*, 664–678.

Terranova, R., Yokobayashi, S., Stadler, M.B., Otte, A.P., van Lohuizen, M., Orkin, S.H., and Peters, A.H.F.M. (2008). Polycomb Group Proteins Ezh2 and Rnf2 Direct Genomic Contraction and Imprinted Repression in Early Mouse Embryos. Dev. Cell *15*, 668–679.

Thorstensen, T., Fischer, A., Sandvik, S.V., Johnsen, S.S., Grini, P.E., Reuter, G., and Aalen, R.B. (2006). The Arabidopsis SUVR4 protein is a nucleolar histone methyltransferase with preference for monomethylated H3K9. Nucleic Acids Res. *34*, 5461–5470.

Thorstensen, T., Grini, P.E., and Aalen, R.B. (2011). SET domain proteins in plant development. Biochim. Biophys. Acta *1809*, 407–420.

Tie, F., Furuyama, T., Prasad-Sinha, J., Jane, E., and Harte, P.J. (2001). The Drosophila Polycomb Group proteins ESC and E(Z) are present in a complex containing the histone-binding protein p55 and the histone deacetylase RPD3. Dev. Camb. Engl. *128*, 275–286.

Tie, F., Prasad-Sinha, J., Birve, A., Rasmuson-Lestander, Å., and Harte, P.J. (2003). A 1-Megadalton ESC/E(Z) Complex from Drosophila That Contains Polycomblike and RPD3. Mol. Cell. Biol. *23*, 3352–3362.

Tippmann, S.C., Ivanek, R., Gaidatzis, D., Schöler, A., Hoerner, L., van Nimwegen, E., Stadler, P.F., Stadler, M.B., and Schübeler, D. (2012). Chromatin measurements reveal contributions of synthesis and decay to steady-state mRNA levels. Mol. Syst. Biol. *8*, 593.

Torok, D., Ching, R.W., and Bazett-Jones, D.P. (2009). PML nuclear bodies as sites of epigenetic regulation. Front. Biosci. J. Virtual Libr. *14*, 1325–1336.

Torres-Padilla, M.E., and Zernicka-Goetz, M. (2006). Role of TIF1alpha as a modulator of embryonic transcription in the mouse zygote. J. Cell Biol. *174*, 329–338.

Torres-Padilla, M.-E., Parfitt, D.-E., Kouzarides, T., and Zernicka-Goetz, M. (2007). Histone arginine methylation regulates pluripotency in the early mouse embryo. Nature *445*, 214–218.

Towbin, B.D., González-Aguilera, C., Sack, R., Gaidatzis, D., Kalck, V., Meister, P., Askjaer, P., and Gasser, S.M. (2012). Step-Wise Methylation of Histone H3K9 Positions Heterochromatin at the Nuclear Periphery. Cell *150*, 934–947.

Trojer, P., Cao, A.R., Gao, Z., Li, Y., Zhang, J., Xu, X., Li, G., Losson, R., Erdjument-Bromage, H., Tempst, P., et al. (2011). L3MBTL2 protein acts in concert with PcG protein-mediated monoubiquitination of H2A to establish a repressive chromatin structure. Mol. Cell *42*, 438–450.

Tsai, M.-C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J.K., Lan, F., Shi, Y., Segal, E., and Chang, H.Y. (2010). Long Noncoding RNA as Modular Scaffold of Histone Modification Complexes. Science *329*, 689–693.

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

Turck, F., Roudier, F., Farrona, S., Martin-Magniette, M.-L., Guillaume, E., Buisine, N., Gagnot, S., Martienssen, R.A., Coupland, G., and Colot, V. (2007). Arabidopsis TFL2/LHP1 Specifically Associates with Genes Marked by Trimethylation of Histone H3 Lysine 27. PLoS Genet *3*, e86.

Turner, B.M. (2002). Cellular Memory and the Histone Code. Cell *111*, 285–291.

Turner, J.M.A. (2007). Meiotic sex chromosome inactivation. Dev. Camb. Engl. *134*, 1823–1831.

Vakoc, C.R., Mandat, S.A., Olenchock, B.A., and Blobel, G.A. (2005). Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. Mol. Cell *19*, 381–391.

Vassilev, A., Kaneko, K.J., Shu, H., Zhao, Y., and DePamphilis, M.L. (2001). TEAD/TEF transcription factors utilize the activation domain of YAP65, a Src/Yes-associated protein localized in the cytoplasm. Genes Dev. *15*, 1229–1241.

Veiseth, S.V., Rahman, M.A., Yap, K.L., Fischer, A., Egge-Jacobsen, W., Reuter, G., Zhou, M.-M., Aalen, R.B., and Thorstensen, T. (2011). The SUVR4 Histone Lysine Methyltransferase Binds Ubiquitin and Converts H3K9me1 to H3K9me3 on Transposon Chromatin in Arabidopsis. PLoS Genet *7*, e1001325.

Vermaak, D., and Malik, H.S. (2009). Multiple roles for heterochromatin protein 1 genes in Drosophila. Annu. Rev. Genet. *43*, 467–492.

Vermaak, D., Henikoff, S., and Malik, H.S. (2005). Positive selection drives the evolution of rhino, a member of the heterochromatin protein 1 family in Drosophila. PLoS Genet. *1*, 96–108.

Verschure, P.J., van der Kraan, I., de Leeuw, W., van der Vlag, J., Carpenter, A.E., Belmont, A.S., and van Driel, R. (2005). In vivo HP1 targeting causes large-scale chromatin condensation and enhanced histone lysine methylation. Mol. Cell. Biol. *25*, 4552–4564.

Voigt, P., LeRoy, G., Drury III, W.J., Zee, B.M., Son, J., Beck, D.B., Young, N.L., Garcia, B.A., and Reinberg, D. (2012). Asymmetrically Modified Nucleosomes. Cell *151*, 181–193.

Volpe, A.M., Horowitz, H., Grafer, C.M., Jackson, S.M., and Berg, C.A. (2001). Drosophila rhino encodes a female-specific chromo-domain protein that affects chromosome structure and egg polarity. Genetics *159*, 1117–1134.

Voncken, J.W., Roelen, B.A.J., Roefs, M., de Vries, S., Verhoeven, E., Marino, S., Deschamps, J., and van Lohuizen, M. (2003). Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. Proc. Natl. Acad. Sci. U. S. A. *100*, 2468–2473.

De Vries, W.N., Binns, L.T., Fancher, K.S., Dean, J., Moore, R., Kemler, R., and Knowles, B.B. (2000). Expression of Cre recombinase in mouse oocytes: A means to study maternal effect genes. Genesis *26*, 110–112.

Waddington, C.H. (1942). CANALIZATION OF DEVELOPMENT AND THE INHERITANCE OF ACQUIRED CHARACTERS : Abstract : Nature. Nature *150*.

Waddington, C.H. (1957). Waddington's book The Strategy of the Genes (Allen & Unwin).

Wagschal, A., Sutherland, H.G., Woodfine, K., Henckel, A., Chebli, K., Schulz, R., Oakey, R.J., Bickmore, W.A., and Feil, R. (2008). G9a histone methyltransferase contributes to imprinting in the mouse placenta. Mol. Cell. Biol. *28*, 1104–1113.

Wang, H., and Dey, S.K. (2006). Roadmap to embryo implantation: clues from mouse models. Nat. Rev. Genet. *7*, 185–199.

Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R.S., and Zhang, Y. (2004). Role of histone H2A ubiquitination in Polycomb silencing. Nature *431*, 873–878.

Wang, J., Mager, J., Chen, Y., Schneider, E., Cross, J.C., Nagy, A., and Magnuson, T. (2001). Imprinted X inactivation maintained by a mouse Polycomb group gene. Nat. Genet. *28*, 371–375.

Wang, J., Zhang, M., Zhang, Y., Kou, Z., Han, Z., Chen, D.-Y., Sun, Q.-Y., and Gao, S. (2010). The histone demethylase JMJD2C is stage-specifically expressed in preimplantation mouse embryos and is required for embryonic development. Biol. Reprod. *82*, 105–111.

Wang, X., Pan, L., Wang, S., Zhou, J., McDowell, W., Park, J., Haug, J., Staehling, K., Tang, H., and Xie, T. (2011). Histone H3K9 trimethylase Eggless controls germline stem cell maintenance and differentiation. PLoS Genet. *7*, e1002426.

Wennekamp, S., Mesecke, S., Nédélec, F., and Hiiragi, T. (2013). A self-organization framework for symmetry breaking in the mammalian embryo. Nat. Rev. Mol. Cell Biol. *14*, 452–459.

Whetstine, J.R., Nottke, A., Lan, F., Huarte, M., Smolikov, S., Chen, Z., Spooner, E., Li, E., Zhang, G., Colaiacovo, M., et al. (2006). Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. Cell *125*, 467–481.

Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. Cell *153*, 307–319.

De Wit, E., Greil, F., and van Steensel, B. (2007). High-resolution mapping reveals links of HP1 with active and inactive chromatin components. PLoS Genet. *3*, e38.

Woo, C.J., Kharchenko, P.V., Daheron, L., Park, P.J., and Kingston, R.E. (2010). A region of the human HOXD cluster that confers polycomb-group responsiveness. Cell *140*, 99–110.

Wossidlo, M., Nakamura, T., Lepikhov, K., Marques, C.J., Zakhartchenko, V., Boiani, M., Arand, J., Nakano, T., Reik, W., and Walter, J. (2011). 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. Nat. Commun. *2*, 241.

Wu, G., and Schöler, H.R. (2014). Role of Oct4 in the early embryo development. Cell Regen. *3*, 7.

Xie, H., Xu, J., Hsu, J.H., Nguyen, M., Fujiwara, Y., Peng, C., and Orkin, S.H. (2014). Polycomb Repressive Complex 2 Regulates Normal Hematopoietic Stem Cell Function in a Developmental-Stage-Specific Manner. Cell Stem Cell.

Xu, L., and Shen, W.-H. (2008). Polycomb silencing of KNOX genes confines shoot stem cell niches in Arabidopsis. Curr. Biol. CB *18*, 1966–1971.

Xu, C., Bian, C., Yang, W., Galka, M., Ouyang, H., Chen, C., Qiu, W., Liu, H., Jones, A.E., MacKenzie, F., et al. (2010). Binding of different histone marks differentially regulates the activity and specificity of polycomb repressive complex 2 (PRC2). Proc. Natl. Acad. Sci. *107*, 19266–19271.

Xu, M., Wang, W., Chen, S., and Zhu, B. (2012). A model for mitotic inheritance of histone lysine methylation. EMBO Rep. *13*, 60–67.

Yagi, R., Kohn, M.J., Karavanova, I., Kaneko, K.J., Vullhorst, D., DePamphilis, M.L., and Buonanno, A. (2007). Transcription factor TEAD4 specifies the trophectoderm lineage at the beginning of mammalian development. Dev. Camb. Engl. *134*, 3827–3836.

Yamamizu, K., Fujihara, M., Tachibana, M., Katayama, S., Takahashi, A., Hara, E., Imai, H., Shinkai, Y., and Yamashita, J.K. (2012). Protein Kinase A Determines Timing of Early Differentiation through Epigenetic Regulation with G9a. Cell Stem Cell *10*, 759–770.

Yamanaka, S., and Blau, H.M. (2010). Nuclear reprogramming to a pluripotent state by three approaches. Nature *465*, 704–712.

Yang, L., Lin, C., Liu, W., Zhang, J., Ohgi, K.A., Grinstein, J.D., Dorrestein, P.C., and Rosenfeld, M.G. (2011). ncRNA- and Pc2 Methylation-Dependent Gene Relocation between Nuclear Structures Mediates Gene Activation Programs. Cell *147*, 773–788.

Yap, K.L., Li, S., Muñoz-Cabello, A.M., Raguz, S., Zeng, L., Mujtaba, S., Gil, J., Walsh, M.J., and Zhou, M.-M. (2010). Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. Mol. Cell *38*, 662–674.

Yeap, L.-S., Hayashi, K., and Surani, M.A. (2009). ERG-associated protein with SET domain (ESET)-Oct4 interaction regulates pluripotency and represses the trophectoderm lineage. Epigenetics Chromatin *2*, 12.

Yokobayashi, S., Liang, C.-Y., Kohler, H., Nestorov, P., Liu, Z., Vidal, M., van Lohuizen, M., Roloff, T.C., and Peters, A.H.F.M. (2013). PRC1 coordinates timing of sexual differentiation of female primordial germ cells. Nature *495*, 236–240.

Yoon, J., Lee, K.-S., Park, J.S., Yu, K., Paik, S.-G., and Kang, Y.-K. (2008). dSETDB1 and SU(VAR)3-9 sequentially function during germline-stem cell differentiation in Drosophila melanogaster. PloS One *3*, e2234.

Yuan, P., Han, J., Guo, G., Orlov, Y.L., Huss, M., Loh, Y.-H., Yaw, L.-P., Robson, P., Lim, B., and Ng, H.-H. (2009). Eset partners with Oct4 to restrict extraembryonic trophoblast lineage potential in embryonic stem cells. Genes Dev. *23*, 2507–2520.

Yuan, W., Xu, M., Huang, C., Liu, N., Chen, S., and Zhu, B. (2011). H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. J. Biol. Chem. *286*, 7983–7989.

Yun, J.-Y., Weigel, D., and Lee, I. (2002). Ectopic Expression of SUPERMAN Suppresses Development of Petals and Stamens. Plant Cell Physiol. *43*, 52–57.

Zamudio, N.M., Chong, S., and O'Bryan, M.K. (2008). Epigenetic regulation in male germ cells. Reprod. Camb. Engl. *136*, 131–146.

Zaret, K.S., Watts, J., Xu, J., Wandzioch, E., Smale, S.T., and Sekiya, T. (2008). Pioneer factors, genetic competence, and inductive signaling: programming liver and pancreas progenitors from the endoderm. Cold Spring Harb. Symp. Quant. Biol. *73*, 119–126.

Zeng, F., Baldwin, D.A., and Schultz, R.M. (2004). Transcript profiling during preimplantation mouse development. Dev. Biol. *272*, 483–496.

Zeng, W., Ball, A.R., Jr, and Yokomori, K. (2010). HP1: heterochromatin binding proteins working the genome. Epigenetics Off. J. DNA Methylation Soc. *5*, 287–292.

Zhang, K., Mosch, K., Fischle, W., and Grewal, S.I.S. (2008). Roles of the Clr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. Nat. Struct. Mol. Biol. *15*, 381–388.

Zhao, B., Ye, X., Yu, J., Li, L., Li, W., Li, S., Yu, J., Lin, J.D., Wang, C.-Y., Chinnaiyan, A.M., et al. (2008). TEAD mediates YAP-dependent gene induction and growth control. Genes Dev. *22*, 1962–1971.

Zhao, J., Ohsumi, T.K., Kung, J.T., Ogawa, Y., Grau, D.J., Sarma, K., Song, J.J., Kingston, R.E., Borowsky, M., and Lee, J.T. (2010). Genome-wide identification of Polycomb-associated RNAs by RIP-seq. Mol. Cell *40*, 939–953.

Zhao, M., Shirley, C.R., Hayashi, S., Marcon, L., Mohapatra, B., Suganuma, R., Behringer, R.R., Boissonneault, G., Yanagimachi, R., and Meistrich, M.L. (2004). Transition nuclear proteins are required for normal chromatin condensation and functional sperm development. Genes. N. Y. N 2000 *38*, 200–213.

Zheng, D., Zhao, K., and Mehler, M.F. (2009). Profiling RE1/REST-mediated histone modifications in the human genome. Genome Biol. *10*, R9.

Zofall, M., and Grewal, S.I.S. (2006). Swi6/HP1 recruits a JmjC domain protein to facilitate transcription of heterochromatic repeats. Mol. Cell *22*, 681–692.

Zuo, X., Sheng, J., Lau, H.-T., McDonald, C.M., Andrade, M., Cullen, D.E., Bell, F.T., Iacovino, M., Kyba, M., Xu, G., et al. (2012). Zinc finger protein ZFP57 requires its co-factor to recruit DNA methyltransferases and maintains DNA methylation imprint in embryonic stem cells via its transcriptional repression domain. J. Biol. Chem. *287*, 2107–2118.

# VI. Acknowledgments

I would like to thank Antoine Peters for giving me the opportunity to work on an exciting research project in his laboratory and for giving me the freedom required to explore new scientific grounds, as well as for his supervision and guidance.

I would like to thank the members of my PhD committee, Rolf Zeller and Filippo Rijli. I would like also to thank Frederick Meins who agreed to chair my PhD defense.

I would like to express my thanks to all current and past members of the Peters lab for creating a great, stimulating and joyful working atmosphere. Especially, I would like to thank my office mates Rico Kunzmann, Mathieu Tardat, Susanne Kaspar, Vijay Tiwari, Rabih Murr, Philip Jermann, Florian Lienert and more recently Julia Hacker, Philipp Bammer, Darko Barisic and Sylvia Domcke. We had many fun moments and also some inspiring scientific discussions together. I would also like to thank Shihori Yokobayashi, Nathalie Veron and Remi Terranova for their help and advice during my PhD.

I would like to thank Dirk Schübeler and his lab members for many discussions and helpful suggestions.

I acknowledge all the excellent facilities at the FMI. I thank all the past and current members of the Functional Genoimics facility and especially Tim Roloff, the Bioinformatics group, the FAIM facility and the Histology facility.

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

I thank my wife Anna and my parents for for giving me constant motivation and support during my studies. I am also very happy to have become a father in the last year of my PhD – I thank Anna and Sophia for all the lovely moments that we had and will have together.

# VII. Appendix A: R scripts

## Annotated PCA

```
Tab <-  as.matrix(read.csv('Dataset.csv', header=T, row.names=1))
pca <- prcomp(Tab)
pdf("20121105.pdf", height=6, width=5, pointsize=12)
pchs=c(8,19)
cols=c("red","green", "blue","black")
plot(pca$x, pch=pchs[as.factor(unlist(strsplit(rownames(Tab), c("[rt]\\."),perl=TRUE))[c(T,F)])],
col=cols[as.factor(unlist(strsplit(rownames(Tab), c("_"),perl = TRUE))[c(F,T,F)])], ylim=c(-20,30))
legend(x="topright", pch=pchs, col=c("black"), legend=levels(as.factor(unlist(strsplit(rownames(Tab),
c("[rt]\\."),perl = TRUE))[c(T,F)])))
legend(x="topleft", pch=22, col=cols, legend=levels(as.factor(unlist(strsplit(rownames(Tab), c("_"),perl =
TRUE))[c(F,T,F)])))
dev.off()
# plot with rainbow colors:
cols=c(rainbow(6))
plot(TabL, pch=1, col=cols[as.factor(unlist(strsplit(rownames(TabL), c("_"),perl = TRUE))[c(F,T)])])
legend(x="topleft", pch=1, col=cols, legend=levels(as.factor(unlist(strsplit(rownames(TabL), c("_"),perl =
TRUE))[c(F,T)])))
# label with sample/gene names:
library(calibrate)
textxy(TabL[,1], TabL[,2], rownames(TabL))
# export loadings:
write.table(pca$rotation, "test.txt")
```

## RNA-seq analysis with edgeR

```
library(edgeR)
x <- read.csv("raw_Ezh2Rnf2NSN_int.csv",row.names="Refseq")
group <- factor(c(1,1,2,2,2))
y <- DGEList(counts=x,group=group)
y <- estimateCommonDisp(y)
y <- estimateTagwiseDisp(y)
et <- exactTest(y)
topTags(et)
summary(de <- decideTestsDGE(et, p=0.05, adjust="BH"))
plotMDS(y)
plotBCV(y)
detags <- rownames(y)[as.logical(de)]
plotSmear(et, smooth.scatter=TRUE, de.tags=detags)
resultsTbl <- topTags( et , n = nrow( et$table ) )$table
colnames( resultsTbl ) <- c( "logConc" , "logFC" , "pVal.Cmn" , "adj.pVal.Cmn" )
write.table( resultsTbl , file = "resultsTbl" , sep = "," , row.names = TRUE )
```

## Violin plot

Package “vioplots”

## Annotated heatmap

```
> k27 <- as.matrix(read.csv("20140808_H3K27_ordered_hclust.csv",row.names=1))
> k27[k27 > 5] = 5
> k27[k27 < -10] =-10
> aheatmap(k27, color = "-RdYlBu2:100", border_color = NA, scale = "none", Rowv = NA, Colv =
TRUE, annCol = c(stage[,1]), distfun = "euclidean", hclustfun = "ward")
```

# VIII. Appendix B: Antibodies

| Antibody | Host organism | Dilution | Supplier, Product number |
|---|---|---|---|
| Anti-CDX2-88 | Mouse | 100 | BioGenex, CDX2-88, MU392A-UC |
| Anti-E-cadherin | Rat | 500 | Life technologies (Invitrogen) 13-1900 |
| Anti-Eomes | Rabbit | 200 | Abcam 23345 |
| Anti-Ezh2 | Mouse | 200 | Leica (Novocastra) NCL-L-EZH2 |
| Anti-gamma H2AX | Rabbit | 500 | Millipore 05-636 |
| Anti-H2AK119Ub | Rabbit | 500 | Cell signaling 8240 |
| Anti-H3K27me3 | Rabbit | 500 | Millipore (Upstate) 07-449 |
| Anti-H3K27me3 | Rabbit | 500 | Cell signaling 9733 |
| Anti-H3K4me3 | Rabbit | 500 | Millipore (Upstate) 07-473 |
| Anti-H3K9me3 | Rabbit | 500 | Millipore 07-442 |
| Anti-RNA Pol II | Mouse | 200 | Covance MMS 126R |
| Alexa Fluor® 488 Goat Anti-Mouse IgG (H+L) | Goat | 500 | Life technologies (Molecular Probes) A-11001 |
| Alexa Fluor® 568 Goat Anti-Rabbit IgG (H+L) | Goat | 500 | Life technologies (Molecular Probes) A-11011 |
| Alexa Fluor® 647 Goat Anti-Rat IgG (H+L) | Goat | 500 | Life technologies (Molecular Probes) A-21247 |

# IX. Appendix C: Fluidigm qPCR primers

*The table below lists the 192 genes assayed in the Fluidigm-based single-cell profiling of mouse preimplantation embryos (Chapter 4). The table includes the primer sequences, as well as an extended gene name, giving information on the molecular function of each gene. The extended gene name first shows the official gene name, followed by an alternative gene name (where applicable), the main molecular activity (CR – chromatin remodelling; DNA – DNA binding; DNAme – DNA methylation; HDM – histone demethylation; HR – histone binding; HMT – histone methylation; HK – housekeeping; SIG – signalling), important functional domains (SET – SET histone methyltransferase domain; ZnF – Zn-finger domain; ATPase; Helicase; Kinase; CD – chromodomain; JmjC/JmjD - Jumonji C/D domains) and finally the associated histone modification. Genes with "#N/A" given instead of an Extended Gene Name were omitted from the analysis due to poor performance of the primers (this primers may not work in certain applications).*

| Gene Name | Extended Gene Name | qPCR Forward Primer | qPCR Reverse Primer |
|---|---|---|---|
| *2410016O06Rik* | 2410016O06Rik_NO66_HDM_JmjC_H3K4_H3K36 | ACCCAAGTCCACATGCTTCA | TCCACCGTGTGATACAGGAAC |
| *Asf1a* | ASF1A_CR | CCCGTTCCAGTTCGAGATCA | ACTTTCTGCAGAGCCCACATA |
| *Asf1b* | ASF1B_CR | GACCCAGAGCTTCGAGAGAA | GGTTAGAGGCCAAGATGTTCC |
| *Ash1l* | Ash1l_HMT_SET_ZnF_H3K4_H3K36 | TGAGGCCAGATTCATCAACCA | CAGTCCAATACGGTACACTCCA |
| *Ash2l* | Ash2l_HMT_H3K4_H3K36 | GGACAAAGCAGAGAAGAGCCTA | AAGCCACACCCTGATTGACA |
| *Atrx* | Atrx_CR_ATPase_Helicase_ZnF | ACACCCTTCACTCAAAGTCC | TCCATCTGAGTCACGGCTA |
| *Bmi1* | Bmi1_HR_ZnF_PRC1_H3K27_H2AK119ub | CCTGTGTGGAGGGTACTTCA | TGCTGGTCTCCAAGTAACGTA |
| *Brdt* | Brdt_CR | CGCCCTTAAACATGGCACAA | CGCCTCTTCACACCCTTTGTA |
| *Carm1* | Carm1_HMT_H3R17 | GAGAGCTACCTCCATGCCAAA | GGTGCGAGGTGGACATCA |
| *Cbx1* | Cbx1_HR_CD_HP1b_H3K9 | GGAGAGGAAAGCAAACCAAAGAA | TAATCCGCTCTGGCTCCAAA |
| *Cbx2* | #N/A | ACCCGAGGCTGCTCCTA | GGGTCTCTTGCCTCTCTTCC |
| *Cbx3* | Cbx3_HR_CD_HP1g_H3K9 | GTTGTACATGTCCCCTTTGGAA | AGCTCTATCTGCCTCTAGCC |
| *Cbx4* | #N/A | GCAGAGTGGAGTATCTGGTGAA | AAGATGTTCTCCTCTGGTTCCC |
| *Cbx5* | Cbx5_HR_CD_HP1a_H3K9 | CAACAGATTCCTGCGGTGAC | CAAGAACCAGGTCAGCTTCA |
| *Cbx6* | Cbx6_HR_CD_PRC1_H3K27_H2AK119ub | CGGGTCATTGGGAAGAGCAA | AGAGCGCAAATGTGCCAAAC |
| *Cbx7* | Cbx7_HR_CD_PRC1_H3K27_H2AK119ub | GTGGAGAGCATCCGGAAGAA | CCATCCTTTCCACTTCACCAGATA |
| *Cbx8* | #N/A | GGACGCATGGAATATCTCGTGAA | AGGAGGCGAGCATCCAGAATA |
| *Cdk1* | Cdk1_SIG_Kinase | AAGTACCTGGACTCCATCCC | TCCCTGGAGGATTTGGTGTAA |
| *Cdx2* | Cdx2_DNA | CGATACATCACCATCAGGAGGAA | TGGCTCTGCGGTTCTGAAA |
| *Clock* | Clock_SIG | ACGTTCACTCAGGACAGACA | ACAAGCTACAGGAGCAGTCA |
| *Cox7a2* | #N/A | CCTCCTCTACAGAGCCACAA | GGGAAATGCAGCCATAGCTAA |
| *Creb3l2* | Creb3l2_DNA | GAGACGAGCTTGGAGAAGTCA | TGAGTGTTTCGTTCCCTTCCA |
| *Ctcf* | Ctcf_DNA_ZnF | AAGCCTCCGAAGCCAACAAA | ACTGCAAAGCTCACACTGGAA |
| *Cxxc1* | Cxxc1_DNA_ZnF | GTGAAGGTGAAGCACGTGAA | TCTGTCGATGCCGTTTGTAC |
| *Daxx* | Daxx_CR | TCAGTCTCTCCGGCAGTGTA | GTCTGAAAGCACGATGATCTCC |
| *Dazl* | Dazl | GCTGGAGAGCAGAGGAGTTA | CAGCTCCTGGATCAACTTCAC |
| *Dek* | Dek_CR | ACCAGATGAACTTAGAAATCTACACA | CTGACCCACGTTCTTCTTCA |
| *Dicer1* | Dicer1 | TGCTCGAGATGGAACCAGAAA | ACTTCCACGGTGACTCTGAC |
| *Dlx1* | Dlx1_DNA | TCCTTGGGACTCACACAGAC | CCCCGCCTTGCTTCATCA |
| *Dmap1* | Dmap1_DNAme | CCAAAAGAAGGAGGCTGAGAA | GTGACACCTGCCGACTTAAA |
| *Dnmt1* | #N/A | AGCCATTGGCCTGGAGATTA | GCAGCCTCCTCTTTTGCTTTA |
| *Dnmt3a* | Dnmt3a_DNAme_ZnF | CGCCAGAAGTGCAGAAACA | AATGAAGAGTGGGTGCTCCA |
| *Dnmt3b* | Dnmt3b_DNAme_ZnF | GACGTCCGGAAAATCACCAA | GATCATTGCATGGGCTTCCA |
| *Dot1l* | #N/A | GCGGCTGTGTGACAAATACA | TCAGCTTCATGGGCTGTGTA |
| *Dppa1* | Dppa1_DNA | GTTTGGAGCACCACTGTGAC | GGAGGGTTGTGTTGGGATCA |

| Gene Name | Extended Gene Name | qPCR Forward Primer | qPCR Reverse Primer |
|---|---|---|---|
| *Dppa3* | Dppa3_DNA | TGAAGAGGACGCTTTGGATGA | CCGGGGTTTAGGGTTAGCTT |
| *Dppa4* | #N/A | AAGGCTAAAGCAACGGGGAA | GTCTCAGTGTCTTGTGGTTTTCC |
| *Eed* | Eed_HR_WD_PRC2_H3K27 | TGGAAGGGCACAGAGATGAA | AGAGTGATCCATACCACAGGAC |
| *Ehmt1* | #N/A | AGGAAACCTTGGAAAGTGCTCTA | TACAGCTGCTTTGGGTGGAA |
| *Ehmt2* | Ehmt2_G9a_HMT_SET_H3K9_H3K27 | CCCGGAAAACCATGTCCAAA | GCGGAAATGCTGGACTTCA |
| *Eif1a* | Eif1a_DNA | CAGGCCAGAACCGAAGTACTA | AGAAGATCCACAGGCAGCAA |
| *Elf5* | Elf5_DNA | GCATCAAGAGTCAAGACTGTCAC | ACAGCAGCAAGTCTCTGACA |
| *Eomes* | #N/A | GCGGCAAAGCGGACAATAAC | ATCCAGTGGGAGCCAGTGTTA |
| *Esrrb* | Esrrb_SIG | GGCGTTCTTCAAGAGAACCA | CTCCGTTTGGTGATCTCACA |
| *Ezh1* | Ezh1_HMT_SET_PRC2_H3K27 | TCTGGGGACAAAGACATGCA | GCTCATCTGTTGGCAGCTTTA |
| *Ezh2* | Ezh2_HMT_SET_PRC2_H3K27 | TGATGGAAAAGTGCATGGTGAC | GACCAAGAGCATTTACCAACTCC |
| *Fgf4* | Fgf4_SIG | TGGTGTGCACGCAGACAC | GCCACTCCGAAGATGCTCAC |
| *Fgfr2* | Fgfr2_SIG | TCAAGTGGATGGCTCCTGAA | CACATTAACACCCCGAAGGAC |
| *Fgfr3* | Fgfr3_SIG | AGGATTTAGACCGCATCCTCAC | CCTGGCGAGTACTGCTCAAA |
| *Foxa1* | #N/A | TGAGAGCAACGACTGGAACA | CCGGAGTTCATGTTGCTGAC |
| *Gapdh* | #N/A | AGACGGCCGCATCTTCTT | TTCACACCGACCTTCACCAT |
| *Gata2* | #N/A | CACCCCTAAGCAGAGAAGCAA | TGTGGCACCACAGTTGACA |
| *Gata3* | Gata3_DNA | CCTACCGGGTTCGGATGTAA | CCGCAGTTCACACACTCC |
| *Gata4* | #N/A | GTAATGCCTGCGGCCTCTA | TGGTTTGAATCCCCTCCTTCC |
| *Gata6* | Gata6_DNA | CCCCTCATCAAGCCACAGAA | AGGTAGTGGTTGTGGTGTGAC |
| *H3f3a* | H3f3a_CR | CGGCGTGTGTAGGGGAA | CGAAGGCTGCGAACACAA |
| *H3f3b* | H3f3b_CR | CGGGGTGAAGAAGCCTCA | GGTAACGACGGATCTCTCTCA |
| *Hand1* | Hand1_DNA | CGCGAAAGCAAGCGGAAAA | CGGTGCGCCCTTTAATCC |
| *Hat1* | Hat1_H4K5ac_H4K12ac | GGCAATACAGGCACAGCAA | AGTCACCAGCAGTCGAAGAA |
| *Hdac1* | Hdac1 | TGACATCGTCCTGGCCATC | GCCATCGCCATGGTGAATATC |
| *Hdac4* | Hdac4 | CTGCAGCAGCTCAAGAACAA | ACTCCTGCAGCTTCATCTTCA |
| *Hells* | Hells_Helicase | GAACCGCACAATTGCAAACA | CGTCGTTTTGGTCTTCCAGTA |
| *Hira* | Hira_CR_WD | TGAAGCTCTTGAAGCCAACC | TTGGTCCCATCAGGGTGAATA |
| *Hnrnpr* | Hnrnpr_HK | GAAAAGAAGCTGCGCAGGAA | ATGCACACTCCCAGGTGTTTA |
| *Hoxd10* | #N/A | CCGAAGTGCAGGAGAAGGAA | GCGTTTGGTGCTTGGTGTAA |
| *Hspa8* | #N/A | GCAGCTGGGCCTACACA | GTAGGTGGTGCCGAGATCAA |
| *Id2* | Id2_DNA | GAACACGGACATCAGCATCC | AGCCACAGAGTACTTTGCTATCA |
| *Jarid2* | Jarid2_HMT_JmjC_JmjN_PRC2_H3K27 | AGAAAAGGCAGCATGGGGAA | CCTTAGCCTGGGCATTACCAA |
| *Jhdm1d* | Jhdm1d_HDM_JmjC_ZnF_H3K27me2 _H3K9me2_H4K20me1 | TGCATCAAACCCACCACCTA | TGTTTGGCTGTTGCCATTCC |
| *Jmjd1c* | Jmjd1c_HDM_JmjC_H3K9 | TGGCTTGCTTGGACTGTTAC | CTGGCCTTTCACACACTTCA |
| *Jmjd4* | Jmjd4_HDM_JmjC | TGAGGTCGAAGAATGGAAGGAC | GATCCCTGTGCAGGACTTCA |
| *Jmjd5* | Jmjd5_HDM_JmjC_H3K36 | AGACTGGTCCCAGACTCTCA | AGCAAGGTACCCGACATCC |
| *Jmjd6* | Jmjd6_HDM_JmjC_H3R2_H4R3 | TGTGTGGCACAAGACGGTAA | CTCAGGGTGCTCCTGTTTCAA |
| *Jmjd7* | #N/A | GCTTCTGTCTGACATTGAGTCC | ATCACCCAGCCAGAAGTTCA |
| *Jmjd8* | Jmjd8_HDM_JmjC | TGGGAATTGCAGCAGCAGTA | CAAGATGACGGGCTTGAGGAA |
| *Kdm1a* | Kdm1a_Lsd1_HDM_H3K4_H3K9 | ACGTTTGAAGCCACTCTCCA | AACTGTGAACTCGGTGGACAA |
| *Kdm1b* | Kdm1b_Lsd2_HDM_H3K4 | ACTTACCGCTGTGGCATGAA | CAGATCCTCGGGGAAAGAACAA |
| *Kdm2a* | Kdm2a_Jhdm1a_HDM_JmjC_ZnF_H3K36 | CAGGAAGACAGCTCGGACAA | GCCGTAAGACTTTGGCTTGTAC |
| *Kdm2b* | Kdm2b_Jhdm1b_HDM_JmjC_ZnF_H3K4_H3K36 | GGAGTGTCCGAAGTGTAACCA | GGCAGGTTGGAGGCATACTTA |
| *Kdm3a* | Kdm3a_ Jmjd1a_HDM_JmjC_H3K9 | ATTCGAGCTGTTTCCCACAC | TTTCTCCAAGACTCCCCATCA |
| *Kdm3b* | Kdm3b_Jmjd1b_HDM_H3K9 | CCATGACCCCAGCAACAAAA | TGCACCCCTGAAACTAGCA |
| *Kdm4a* | Kdm4a_Jmjd2a_HDM_JmjC_JmjN_ZnF_H3K9_H3K36 | GTCCAGTGGATGTGAGCAAA | TCCGTTTTCTCCGCTTCTTAC |
| *Kdm4b* | Kdm4b_Jmjd2b_HDM_JmjC_JmjN_ZnF_H3K9 | GACACTCATCTCGCCCATCA | ATCATGAACTCCCCAGCTTCC |
| *Kdm4c* | Kdm4c_Jmjd2c_HDM_JmjC_JmjN_ZnF_H3K9_H3K36 | TGTCCCAGAACGGACACAAA | CCGGTGTCTACAGAAGATGCA |
| *Kdm4d* | #N/A | AGCCTAGGGACTAGAGTTGTCA | GACTGGGGTGCAGCAGAA |
| *Kdm5a* | Kdm5a_Rbp2_HDM_JmjC_JmjN_H3K4 | CTGTGTGTGCAGCCCAAAA | TCATCACAGCCACCATCACA |
| *Kdm5b* | Kdm5b_Jarid1b_HDM_JmjC_JmjN_ZnF_H3K4 | AGCAAGCTGACCGAAGTTCA | TTGCATCTCGTTTCCCTCGAA |
| *Kdm5c* | Kdm5c_Jarid1c_HDM_JmjC_JmjN_ZnF_H3K4 | ATGAGGCTTCGAAGGAACCA | TCTTCATCTCCTCGGGAACAC |
| *Kdm5d* | Kdm5d_Jarid1d_HDM_JmjC_JmjN_ZnF_H3K4 | GAAGCAGAGGCTTGTATTTCCC | GGGTGATTTGCGGTGTTTGTA |

| Gene Name | Extended Gene Name | qPCR Forward Primer | qPCR Reverse Primer |
|---|---|---|---|
| *Kdm6a* | Kdm6a_Utx_HDM_JmjC_H3K27 | TTCGGGTTCGTGAGGTTTCA | AGAGATTCGTAGCAGCGAACA |
| *Kdm6b* | Kdm6b_Jmjd3_HDM_JmjC_H3K27 | GAACCCCATCACCGTCATCA | GTCTTGGTGGAGAAAAGGCCTA |
| *Klf2* | Klf2_DNA_ZnF | CTAAAGGCGCATCTGCGTAC | TTCCAGCCGCATCCTTCC |
| *Klf4* | Klf4_DNA_ZnF | CAGGCTGTGGCAAAACCTATAC | CGTCCCAGTCACAGTGGTAA |
| *Krt19* | Krt19_SIG | GCCACCATTGACAACTCCAA | CAAGGCGTGTTCTGTCTCAA |
| *Krt8* | Krt8_SIG | ACCGACGAGATCAACTTCCTC | ACCACAGACGTGTCTGAGATC |
| *L3mbtl3* | L3mbtl2_DNA_PRC1_H3K27_H2AK119ub | ACCCGCTGTTCATCAGAGTA | CAACTGCTCCAGCCATCAAA |
| *Lats2* | Lats2_SIG_Kinase | ACCCCGAAGTTTGGACCTTA | GCCTGACTCGTTGGCAAAA |
| *Lmnb1* | #N/A | ACATGGAGATCAGCGCCTAC | GAAGGGCTTGGAGAGAGCTTTA |
| *Lrpap1* | Lrpap1 | GGACCCCAGGCTGGAAAA | ACAGCTTGTCCAGCTCTTCA |
| *Mapk1* | Mapk1_SIG_Kinase | CGTTGGTACAGAGCTCCAGAA | TGCAGCCCACAGACCAAATA |
| *Mapk3* | Mapk3_SIG_Kinase | TGGGCCAAGCTCTTTCCTAA | TGCGCTTGTTTGGGTTGAA |
| *Mecom* | Mecom_Prdm3_DNA | AGATCCATGGCAACCAGGAC | TGACAGCATGTGCTTCTCCA |
| *Mll1* | Mll1_HMT_SET_H3K4 | AACAGACTGACCAGCCCAAA | TTTAATCCGGGGTCCTCGAAC |
| *Mll2* | Mll2_HMT_SET_ZnF_H3K4 | TCACCCGTACTGTGTCAACA | CACACACGATACACTCCACAC |
| *Mll3* | Mll3_HMT_SET_H3K4 | ATGGTCCCAGATCGGGTCTTA | GCAGCAACAACGCTGCTAA |
| *Mll5* | Mll5_HMT_SET_ZnF_H3K4 | CCCCAGAATCCACCACAAAA | ACCCACTCTTCCTAGCTTCA |
| *Mtf2* | Mtf2_Pcl2_DNA_PRC2_ZnF_H3K27 | AGCTTGTGTCCAATGCCTTCA | CCAGAACTGCAGACAGAGCAA |
| *Nanog* | Nanog_DNA | TCTGGGAACGCCTCATCAA | GAGGCAGGTCTTCAGAGGAA |
| *Nf2* | Nf2_SIG | GCCTGGCTCAAAATGGACAA | TTGGCCAGGAAGTGAAAGGTA |
| *Nsd1* | Nsd1_HMT_H3K36_H4K20 | GGGGATGTAAGCAGCAAGGATA | TCTTGCTGCAGCTTCCTGAA |
| *Pax3* | #N/A | TTACCAGCCCACGTCTATTCC | GTGTACAGTGCTCGGAGGAA |
| *Pax6* | Pax6_DNA | TATCCCGGGACTTCAGTACCA | TGATGGAGTTGGTGTTCTCTCC |
| *Pax7* | #N/A | TGGCGAGAAGAAAGCCAAAC | CACATCTGAGCCCTCATCCA |
| *Pcgf1* | Pcgf1_DNA_PRC1_ZnF_H3k27_H2AK119ub | TCAGTCCCGAGGCTTAGACA | TGGGCTTTAGAGTGGTCAAAAC |
| *Pdgfra* | Pdgfra_SIG_Kinase | CAAAGGGAGGACGTTCAAGAC | TGCGTCCATCTCCAGATTCA |
| *Pecam1* | Pecam1_SIG | GCACAGTGATGCTGAACAAC | GTCACCTTGGGCTTGGATAC |
| *Phc1* | #N/A | GGTCCCTGCTGCTTTCTACA | GACTCAGCTTTGCGTTTCACA |
| *Phc2* | Phc2_DNA_PRC1_ZnF_H3k27_H2AK119ub | GCCTACAAGTTCAAGCGTTCC | CGTTTGGTGCATCCCACATTA |
| *Phc3* | #N/A | CGGTCCAAACGATTCTGTACTA | GGCTTACGATTCCAACGACTA |
| *Phf19* | Phf19_Pcl3_DNA_PRC2_ZnF_H3K27 | TCCATGAGGCTTGCACACA | ACACACGGAGCAGAAGAACA |
| *Phf2* | Phf2_HDM_JmjC_ZnF_H3K9 | ACATTTCCTCCACAGCCTCA | TCAGGCTGCCAAGTTTTAACC |
| *Phf8* | Phf8_HDM_JmjC_ZnF_H3K9_H3K27_H4K20 | GCATTCGGAGGGAACTTCTTAC | TGTGCTCAGCCTCTTCTCAA |
| *Pou5f1* | Pou5f1_Oct4_DNA | TCCCTACAGCAGATCACTCAC | CGCCGGTTACAGAACCATAC |
| *Prdm1* | #N/A | CGTTCGGTCAGCTCTCCAA | TGCAGGTCTGGCACTTGAAA |
| *Prdm10* | Prdm10_HMT | CTTCTGGGCAAGTGAAAGCA | GAGGCTTGCTTTTCCTCCAA |
| *Prdm11* | #N/A | AACTGAAGGGAAAGCGTGAC | TACTCTTGGCAGGACTCACA |
| *Prdm12* | #N/A | GATAGGCACGAGCATCTTCTACA | CCAGGAAGGTGTTGTGGGAA |
| *Prdm13* | #N/A | ATCCCTACAACTGCGACTCC | CGAAAGGTCCTCCAGCAGTA |
| *Prdm14* | Prdm14_HMT_SET | TGAGAGTCCACTCTGGAGACA | AGTATGCTGGAGGCAGTGAA |
| *Prdm15* | Prdm15_HMT_SET | AGCGCCACAAACTCATCCA | AGCATGTCCTTCCTGGCAAA |
| *Prdm16* | #N/A | TCCGAAACTTCATCGCCAAC | CTGTCCAGGTCTTGGATCTCA |
| *Prdm2* | Prdm2_HMT_SET_ZnF | TACCATCTGCGGTCGACAA | CACCAACAAATGGCCCAAAC |
| *Prdm4* | Prdm4_HMT_SET_ZnF | ACCTACGGACACACCTCAAAA | GAGCCTTCTGCGTGAAAGAC |
| *Prdm5* | #N/A | CTGAAGCGTCACATGATTACCC | TGATCGAGCCTCTTGAAGGAC |
| *Prdm6* | Prdm6_HMT_SET_ZnF | AAGGCACTGTGGAGAACAGAA | TGCCCCTGGGGATATCTATACA |
| *Prdm8* | #N/A | AGTCCACGGACAAGAGAACA | TGAGACCTTCTGAGGAACCA |
| *Rbp2* | Rbp2_CR_PRC2_H3K27 | GGAACCTGGGAAATGGAGAGTAA | CTTGCGGGTGGCAAAATCAA |
| *Ring1* | Ring1_CR_PRC1_ZnF_H3K27_H2AK119ub | ATCCCCTGCTTGTGGAGAAA | GATGATCCACTGTGGCATTCC |
| *Rnf2* | Rnf2_CR_PRC1_ZnF_H3K27_H2AK119ub | CAGGCCCCATCCAACTCTTA | CAACAGTGGCATTGCCTGAA |
| *Runx1* | #N/A | AGAACCAGGTAGCGAGATTCA | ACGGTGATGGTCAGAGTGAA |
| *Satb1* | Satb1_DNA | TAAAACACTCGGGCCATCTCA | TGTTCCACCACGCAGAAAAC |
| *Satb2* | Satb2_DNA | CCAGGAGTTTGGGAGATGGTA | TGAAAGGTTCTCTCGCTCCA |
| *Scmh1* | Scmh1_CR_PRC1_ZnF_H3K27_H2AK119ub | GTGGGCTGGTGTTCTTTGAC | AAGGGGACGGATTCTTTGGAA |

| Gene Name | Extended Gene Name | qPCR Forward Primer | qPCR Reverse Primer |
|---|---|---|---|
| *Scml2* | Scml2_CR_PRC1_ZnF_H3K27_H2AK119ub | GAAACCAGACCTGGGAAACA | TCCAAAGGTTGGGTGAATCC |
| *Setd1a* | Setd1a_HMT_SET_H3K4 | CTTTTCCCTCCCGGTTCCTA | AGCCTCGCAAAAGTTACTTCC |
| *Setd1b* | Setd1b_HMT_SET_H3K4 | CGACCTGCTCAAGTTCAACC | CAGTCGTGAATGTGGCTCTTAC |
| *Setd2* | Setd2_HMT_SET_H3K36 | TCCAGTGAGCTGGCAAAGAA | AAGGATTCAGGCACTGGACAA |
| *Setd3* | Setd3_HMT_SET_H3K36 | ACAGGCTCTACGCCATGAA | TCAGTGGAATGCAACGCAAA |
| *Setd4* | Setd4_HMT_SET | AGGAAGAGGGCTGATGAGCAA | TGTCCGTGGTGAGCAGACA |
| *Setd5* | Setd5_HMT_SET | TCAGGATGGCTTCCTTCTCAAC | GCTTCCGCCGATGAAGTCTA |
| *Setd6* | Setd6_HMT_SET | CTGACTGAAGAGGAGCTAGCC | CCATCCAGCCCGCTCTTTATA |
| *Setd7* | Setd7_HMT_SET_H3K4 | GAAAAGAATGGGCGTGGGAA | AGGGCATCGTCCACGTAATA |
| *Setd8* | Setd7_HMT_SET_H4K20 | AATGCCCGGGGAAACCATTA | CAGCGCTTCGTATAACGTTCC |
| *Setdb1* | Setdb1_Eset_HMT_SET_H3K9 | AGCCTCTGCAATGGAGAAGAA | AACAGGACCTTTGCTCCTCA |
| *Setdb2* | Setdb2_HMT_SET_H3K9 | CCCATTGGTGGCCTTCTTCA | GGCCCCAGCTTCATAACCATAA |
| *Sfmbt1* | Sfmbt1_HR_PRC_PhoRC | GTGCTCCGTTGGCAAGAATA | TGGGAAGGGTAAGGAGCAAA |
| *Sfmbt2* | #N/A | TGCCACACGTGTCCTTCA | TTCTTGTTGGCCACCTCCAA |
| *Sirt1* | Sirt1_CR_PRC2_H3K27_H3K9 | CTGAAAGTGAGACCAGTAGCA | GATGAGGCAAAGGTTCCCTA |
| *Smyd1* | Smyd1_HMT_SET_ZnF | ATCTTTGCGGAGAGGGCTTA | AGCTTCTCCTGCCTCTTGAA |
| *Smyd2* | Smyd2_HMT_SET_ZnF | CAGATACGCACGCAATGTCA | CAGATCTCCAGCAGCTCACTA |
| *Smyd3* | Smyd3_HMT_SET_ZnF | CCATCAAGGCATGTTTCCTCA | GCCGTGGGTCACTTTCATAA |
| *Smyd4* | Smyd4_HMT_SET_ZnF | AGCTTCTCCGAACCGGTAAA | TCTGCTGACAAGAAGCTCTCA |
| *Smyd5* | Smyd5_HMT_SET_ZnF | CACCTTCATTGACCAGCTGTAC | CGAAGAGGCCAGATCCTTCA |
| *Sohlh1* | #N/A | GCAAGCCAGACTCCGGTATA | TGTATCCAGCATCCCAAAGCA |
| *Sox17* | Sox17_DNA | CAGAACCCAGATCTGCACAAC | GCTTCTCTGCCAAGGTCAAC |
| *Sox2* | Sox2_DNA | TGAAGGAGCACCCGGATTATA | CGGGAAGCGTGTACTTATCC |
| *Sox7* | #N/A | ACCCGGACCTGCACAAC | CGCTCTGCCTCATCCACATA |
| *Ssu72* | Ssu72_HK | GGTGTGCTCGAGTAACCAGAA | CAAAGGAGCGGACACTGAAAC |
| *Suv39h1* | Suv39h1_HMT_SET_H3K9 | ACGCTGGAAAAGATCCGAAA | CTCTGCCTCCTCTGAGGTAA |
| *Suv39h2* | Suv39h2_HMT_SET_H3K9 | AGGTGGAGTACTTGTGTGAC | AGAATCTGGCCATCCTTTCC |
| *Suv420h1* | Suv420h1_HMT_SET_H4K20 | GGGCAAGGACACCCTGAA | TGCCCTTCAAATCCCGAGTTA |
| *Suv420h2* | Suv420h2_HMT_SET_H4K20 | CAGGAGGCTGCTCTCAAGAC | CAGAATGGTGAAGCCACTCTCA |
| *Suz12* | Suz12_HMT_PRC2_ZnF_H3K27 | CCACAGCAGGTTCATCTTCAA | TTCCTGCATAGGAGCCATCA |
| *T* | #N/A | ATGCTGCCTGTGAGTCATAAC | GGTACCATTGCTCACAGACC |
| *Tada2a* | Tada2a_DNA | TGCAGAATGGGACTTAAGAGACA | ATCCACCACAGCCATCTTCA |
| *Tada3* | Tada3_DNA | CCTTTGGAGCCCTGACACA | CCCCTGACTCTTTCCCAGAC |
| *Tbx3* | Tbx3_DNA | AGTTTCACAAGCGGGGTACA | CAGTCCAGAGCACCTCACTTTA |
| *Tcl1* | #N/A | CGAAGCTGCGACTCCATGTA | AGTTCAAGCAACATGTCCTCCA |
| *Tead4* | Tead4_DNA | GGCAAGCAAGTGGTGGAGAA | CCGGTGGATGCGGTACAAATA |
| *Tet1* | Tet1_DNAme | AGATGGCTCCAGTTGCTTATCA | ACGCCCCTCTTCATTTCCAA |
| *Tet3* | Tet3_DNAme | ACGCCAGAGAAGATCAAGCA | GACAATCCACCCTTCAGAGACA |
| *Trim28* | Trim28_Tif1b_DNA | CGGGTGAAATACACCAAGGAC | CAGTAGACTGTTCGCTCTCCA |
| *Tspan8* | Tspan8_SIG | GCTGTGGAGCTGTGAAAGAA | CTGCCACTTGCAGAATCAGAA |
| *Ube2e1* | Ube2e1_HK | CCAGCCCTAACCATCTCGAAA | AGTGGCAATACTTCCCACCAA |
| *Uty* | #N/A | TCAGAAGCAAGTGCAGACAC | GGCTTGCAAGGCATCCATA |
| *Whsc1* | Whsc1_Nsd2_HMT_SET_ZnF_H3K4_H3K36_H4K20 | GCCAAGAAAGTGCCAAGCA | GGCTCTCAATCTCCCCGAAATA |
| *Whsc1l1* | Whsc1l1_Nsd3_HMT_SET_ZnF_H3K4_H3K36_H3K27 | CCAGCTTGAGGTCCATTCCA | CCTCTCTGGCTGGTTGCTAAA |
| *Xist* | Xist | TGCTCCTCCGTTACATCAGAC | TTCTTGAGGCAGGAGCACAA |
| *Yap1* | Yap1_SIG | CTGCCCGACTCCTTCTTCAA | CCGCAGTACCTGCATCAGTA |
| *Yy1* | Yy1_DNA_PRC_ZnF_PhoRC | CAAGAACAATAGCTTGCCCTCA | GGTGTGCAGATGCTTTCTCA |
| *Zfp42* | Zfp42_DNA_ZnF | TGGGACACGTGGCAAAAGAA | CAGCACAGTGAGGCGATCC |
| *Zp3* | Zp3_SIG | TACCCATTGAGTGCCGATACC | ACAGTGGCTCTGAAGGGAAC |

# X. Appendix D: Curriculum vitae


**Peter NESTOROV**
*Dipl.-Biochem.*

Nationality: Bulgarian
Date of birth: 18.10.1983
Marital status: Married

Address: Sonnhaldenstrasse 65
4600 Olten, Switzerland
Phone: +41 78 6637358
E-mail: peter.nestorov@yahoo.com


## ▪ Research Experience

▪ 04/2009 - present

**PhD Student and Boehringer Ingelheim Fonds Fellow,**
Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland

- *Scope of research*: studying the role of Polycomb repressive complexes in mouse oogenesis and pre-implantation embryogenesis with a focus on maternal-to-embryonic transition, cell fate specification and dynamics of chromatin states.
- *Major achievements:*
  - Identified a role for PRC2 as regulator of zygotic genome activation. Performed extensive phenotypic and gene expression analysis of mouse oocytes and pre-implantation embryos from conditional knock-out mouse models.
  - Performed comprehensive single-cell analysis of gene expression patterns during pre-implantation development for 156 genes coding for chromatin modifying proteins.

▪ 03/2005 – 03/2009

**Student Research Assistant and Diploma Student,**
Max Planck Institute for Developmental Biology, Tübingen, Germany

- *Scope of research:* studying early embryonic development in the invertebrate *Parhyale hawaiensis* (*Crustacea, Arthropoda*), focusing on lineage specification and segmentation.
- *Major achievements*
  - Established a microarray-based method for the study of invertebrate embryogenesis.
  - Cloned and characterized Wingless family genes in *Parhyale hawaiensis.*

## ▪ Selected Honours and Awards

- 2012 Roche Continents Arts&Science Program.
- 2010-2012 Boehringer Ingelheim Fonds PhD Fellowship.
- 2011 FEBS Young Scientists Forum Fellowship.
- 2009 Best Diploma Thesis Award, German Society for Biochemistry and Molecular Biology (GBM).
- 2004 DAAD Scholarship for a language summer course at the University of Hannover, Germany.

## ▪ Education

- 04/2009 – present: **PhD in Genetics,** University of Basel, Switzerland. *(expected graduation: Autumn 2014)*
  Thesis advisor: Prof. Dr. Antoine HFM Peters
- 10/2004 – 12/2008: **Diploma in Biochemistry,** Eberhard Karls University of Tübingen, Germany.
  *Grade 1.5 (1=highest, 4=lowest positive)*
- 09/2002 – 09/2004: **Studies in Molecular Biology,** Sofia University St. Kliment Ohridski, Bulgaria.
  *Successfully finished 4 semesters (interim grades acknowledged by the University of Tübingen).*
- 09/1998 – 06/2002: **Diploma for completed secondary education** (Matura), Major subject: Chemistry,
  Akad. L. Chakalov National School for Natural Sciences and Mathematics, Sofia, Bulgaria.
  *Grade 5.47 (6=highest, 3=lowest positive)*

## ▪ Skills and Hobbies

- Languages
  - **Bulgarian** (Native), **English** (Fully Proficient), **German** (Fully Proficient), **Russian** (Basic).
- Lab skills
  - **Gene expression**: RT-qPCR (Fluidigm, ABI), RNA-seq (Illumina), microarrays.
  - **Imaging**: brightfield and fluorescent microscopy (Zeiss LSM700; Zeiss AxioScan.Z1).
  - ***In vivo***: setting up and maintaining conditional knock-out mouse lines; certified to work with vertebrate lab animals; working with crustaceans (aquaculture).
  - **Molecular biology**: cloning, protein analysis (SDS-PAGE, Western blot, chromatography), PCR, DNA/RNA analysis (purification and concentration, Bioanalyzer, electrophoresis)
  - **Enzyme kinetics**: kinase assays, analysing enzyme activity/inhibition data.
- IT skills
  - Text editing, presentations, e-mail, organization, reference management:
    **MS Office** (Expert user); **Zotero** (Expert user); **EndNote** (Expert user).
  - Statistical analysis, large data set analysis and data visualization:
    **R** – clustering, PCA, matrix manipulation; scatter plots, heatmaps, Venn diagrams;
    **GraphPad Prism** descriptive statistics, statistical tests, box plots, histograms.
  - Graphics editing, image analysis and processing:
    **Adobe Photoshop and Illustrator** (Advanced user); **ImageJ** (Basic user).
- Hobbies

**Cooking:** creating meals inspired by the Mediterranean cuisine;
**Outdoor activities:** staying close to nature;
**Traveling:** exploring gastronomy, architecture and art.

## ▪ Organizational experience

- 02/2011 – present **PhD Student Representative**, Friedrich Miescher Institute, Basel, Switzerland.
  - Organizing international scientific meetings and career seminars.
- 03/2003 – 09/2004 **President and Board Member**, AEGEE (European Students Association), Sofia, Bulgaria.
  - Organized and managed international student projects.
  - Main co-ordinator of an EU-funded project (YOUTH programme).

## ▪ List of publications


Nestorov P, Tardat M and Peters AHFM. (2013) H3K9/HP1 and Polycomb: two key epigenetic silencing pathways for gene regulation and embryo development. ***Current Topics in Developmental Biology****; 104:243-91.*

Yokobayashi S, Liang CY, Kohler H, Nestorov P, Liu Z, Vidal M, van Lohuizen M, Roloff TC and Peters AHFM. (2013) PRC1 coordinates timing of sexual differentiation of female primordial germ cells. ***Nature****; 495(7440):236-40.*

Nestorov P, Battke F, Levesque MP, Gerberding M. (2013) The maternal transcriptome of the crustacean *Parhyale hawaiensis* is inherited asymmetrically to invariant cell lineages of the ectoderm and mesoderm. ***PLoS One****; 8(2):e56049.*

Valtcheva N, Nestorov P, Beck A, Russwurm M, Hillenbrand M, Weinmeister P and Feil R. (2009) The commonly used cGMP-dependent protein kinase type I (cGKI) inhibitor Rp-8-Br-PET-cGMPS can activate cGKI in vitro and in intact cells. ***Journal of Biological Chemistry****; 284(1):556-62.*