



universität  
wien

# DISSERTATION / DOCTORAL THESIS

Titel der Dissertation / Title of the Doctoral Thesis

**„Dissecting the roles of PRC1 and PRC2 in maintenance of ES cell identity“**

verfasst von / submitted by

**Jorge Arturo Zepeda Martínez**

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

**Doctor of Philosophy (PhD)**

Wien, 2018 / Vienna, 2018

Studienkennzahl lt. Studienblatt /

Dissertationsgebiet lt. Studienblatt /

Betreut von / Supervisor:

A 794 685 490

Molekulare Biologie

Dr. Oliver Bell



# Acknowledgements

I would like to thank Oliver Bell for the supervision and support to carry out this project, especially for the stimulating brainstorming sessions and useful advice on communicating science. I also want to thank Carina Pribitzer for her tremendous help with experiments, and the rest of the lab for fruitful discussions and interactions, as well as everyone at the core campus facilities for experimental support and infrastructure.

I am very grateful to all my friends for their unconditional love and support to get through the harsh moments, and celebrate the joyful times. Thanks to Brian for his savvy mentoring on data analysis, Annika for sharing the excitement and ambition of becoming a better version of oneself, Jakob, Madalena, Jasmin, Matthias, and the rest of the Mol Bio clan for making me feel at home in Austria. Thanks to Jill, Tugrul, Hagar, Karin, Robin and the rest of my friends in Vienna for the uncontrollable laughs and words of encouragement.

I would also like to thank Kenan for always bringing out the bright side of things, especially when I could scarcely see the light.

Finalmente, muchas gracias a mi familia, sobretodo a mis padres por el amor y apoyo que siempre me han dado, sin importar las circunstancias o la distancia. Margarita y Rubén, esta tesis va dedicada a ustedes.





# Abstract

Polycomb Repressive Complexes 1 and 2 (PRC1 and PRC2) are transcriptional repressors essential for differentiation of mouse Embryonic Stem (ES) cells. However, whether they play a role in self-renewal is debated. Genetic deletion of PRC1 and PRC2 subunits have minor effects on gene expression and self-renewal of mouse ES cells. In contrast, double deletion of the paralogous PRC1 catalytic subunits *Ring1* and *Rnf2* results in de-repression of Polycomb target genes and loss of self-renewal capacity. Thus, it remains unclear if indeed Polycomb is required for maintenance of ES cell identity, or alternatively loss of self-renewal in *Ring1/Rnf2* dKOs reflects a PRC1-independent role of the catalytic subunits. PRC1 complexes subdivide into canonical PRC1 (cPRC1) that harbor a chromo box protein (*Cbx*), which binds to PRC2-catalyzed H3K27me<sub>3</sub>, and variant PRC1 (vPRC1) complexes that are characterized by the *Rybp* subunit, which is mutually exclusive with *Cbx*. vPRC1 has been shown to bind un-methylated CpG islands independent of PRC2. Based on their independent modes of recruitment and partially overlapping target genes, I hypothesized that PRC2/cPRC1 and vPRC1 cooperate to maintain ES cell identity by robustly repressing key differentiation genes.

Using a chemical genetic protein degradation system, we show that in mouse ES cells *Eed* and *Rybp*, core subunits of PRC2 and vPRC1, are synthetic lethal, analogous to *Ring1* and *Rnf2* double knockout. Conditional degradation of *Rybp* in *Eed* null cells leads to loss of self-renewal and de-repression of shared Polycomb target genes. Moreover, genome-wide analysis demonstrates that parallel impairment of PRC2 and vPRC1 results in failure to deliver PRC1 catalytic core subunit *Rnf2* to chromatin, which is the underlying cause of target gene de-repression. Unexpectedly, genes with the highest Polycomb-dependent chromatin modifications in wild-type cells are most sensitive to loss of PRC1 recruitment and catalytic activity. In contrast, promoters decorated with bivalent nucleosome modifications of H3K27me<sub>3</sub> and H3K4me<sub>3</sub> are not predictive of transcriptional depression, challenging the model of poised chromatin state at these genes. These results support a model whereby PRC2/cPRC1 and vPRC1 synergize to deliver *Ring1/Rnf2* activity to common target genes to ensure robust maintenance of transcriptional repression of developmental genes, thus safeguarding ES cell identity and self-renewal capacity.

## Zusammenfassung

Polycomb Repressive Complex 1 und 2 (PRC1 und PRC2) sind Transkriptionsrepressor-Komplexe, die für die Differenzierung von Maus-embryonalen Stammzellen (ES Zellen) essentiell sind. Ob sie eine Rolle bei deren Selbsterneuerung von ES Zellen spielen ist jedoch unklar. Die genetische Inaktivierung von PRC1- und PRC2-Untereinheiten hat nur geringe Auswirkungen auf die Genexpression und Selbsterneuerung von ES Zellen. Im Gegensatz dazu führt die doppelte Inaktivierung der paralogen PRC1-katalytischen Untereinheiten *Ring1* und *Rnf2* (*Ring1 / Rnf2* dKO) zur Aktivierung von Differenzierungsgenen und zum Verlust der Selbsterneuerungskapazität. Somit bleibt unklar, ob Polycomb tatsächlich für die Aufrechterhaltung der ES-Zellidentität erforderlich ist, oder ob der Verlust der Selbsterneuerung in *Ring1 / Rnf2* dKOs eine PRC1-unabhängige Rolle der katalytischen Untereinheiten widerspiegelt. PRC1-Komplexe unterteilen sich in kanonische PRC1 Komplexe (cPRC1), die ein Chromobox Protein (*Cbx*) aufweisen das an PRC2-katalysierte H3K27me3 bindet, und Varianten von PRC1 (vPRC1), die Rybp-Untereinheiten enthalten, die sich mit Cbx gegenseitig ausschließen. Auf der anderen Seite ist die Binding von vPRC1 an Zielgene unabhängig von PRC2 und H3K27me3. Basierend auf diesen unabhängigen Rekrutierungsmodi und teilweise überlappenden Zielgenen stellte wir die Hypothese auf, dass PRC2 / cPRC1 und vPRC1 zusammenarbeiten, um die ES-Zellidentität aufrechtzuerhalten, indem sie Schlüsselgene der Differenzierung robust unterdrücken.

Mit Hilfe eines chemisch-genetischen Proteinabbau-Systems konnten wir zeigen, dass in den ES-Zellen von Mäusen *Eed* und *Rybp*, Kernuntereinheiten von PRC2 und vPRC1, synthetisch letal sind, vergleichbar mit *Ring1* und *Rnf2* Doppel-Inaktivierung. Induzierter Abbau von *Rybp* in *Eed* Mutantenzellen führte zum Verlust der Selbsterneuerung und Aktivierung von Polycomb-Zielgenen. Darüber hinaus zeigte unsere genomweite Analyse, dass eine parallele Beeinträchtigung von PRC2 und vPRC1 dazu führt, dass die katalytische Kernuntereinheit *Rnf2* von PRC1 nicht an Chromatin bindet und somit zur Aktivierung von Zielgenen führt. Aktivierung findet hauptsächlich an Genen statt, die in Wildtyp-Zellen hohe Mengen an PRC1- und PRC2 gebunden haben. Im Vergleich dazu führt Bivalenz von H3K4me3 und H3K27me3 nicht präferentiell zur transkriptionellen Aktivierung in Abwesenheit von PRC2 und vPRC1. Diese Ergebnisse unterstützen ein

Modell, in dem PRC2 / cPRC1 und vPRC1 synergistisch *Ring1* / *Rnf2*-Aktivität an Zielgenen induzieren, um eine robuste Aufrechterhaltung der Transkriptionsrepression von Entwicklungsgenen sicherzustellen und die Selbsterneuerungskapazität sowie Identität von ES Zellen sichern.

# Table of Contents

|                                                                                                                                        |           |
|----------------------------------------------------------------------------------------------------------------------------------------|-----------|
| <b>ACKNOWLEDGEMENTS.....</b>                                                                                                           | <b>3</b>  |
| <b>ABSTRACT .....</b>                                                                                                                  | <b>5</b>  |
| <b>ZUSAMMENFASSUNG.....</b>                                                                                                            | <b>6</b>  |
| <b>CHAPTER 1 .....</b>                                                                                                                 | <b>10</b> |
| INTRODUCTION .....                                                                                                                     | 10        |
| 1.1 Epigenetics.....                                                                                                                   | 10        |
| 1.2 Polycomb group of proteins (PcG) are epigenetic regulators .....                                                                   | 11        |
| 1.3 PcG proteins assemble into multi-subunit complexes.....                                                                            | 11        |
| 1.4 PcG proteins are transcriptional repressors.....                                                                                   | 12        |
| 1.5 Diversity of mammalian PcG complexes.....                                                                                          | 12        |
| 1.6 Modes of recruitment to chromatin .....                                                                                            | 13        |
| 1.7 Interplay between PRC1 and PRC2 .....                                                                                              | 14        |
| 1.8 PcG and TrxG proteins regulate CpG island promoters in mammalian genomes....                                                       | 15        |
| 1.9 Embryonic stem cells as a model for cellular differentiation .....                                                                 | 16        |
| 1.10 Role of PcG proteins in differentiation and self-renewal .....                                                                    | 17        |
| 1.11 Aim of the thesis .....                                                                                                           | 18        |
| <b>CHAPTER 2 .....</b>                                                                                                                 | <b>20</b> |
| MATERIALS AND METHODS .....                                                                                                            | 20        |
| 2.2 Handling of haploid mESCs .....                                                                                                    | 20        |
| 2.3 Gene trap cell lines .....                                                                                                         | 20        |
| 2.4 Inversion of gene-trap cassettes.....                                                                                              | 21        |
| 2.5 Genotyping of gene trap cassettes .....                                                                                            | 21        |
| 2.6 CRISPR/Cas9-mediated gene knockout in mESCs .....                                                                                  | 21        |
| 2.7 AID-tagging of endogenous genes.....                                                                                               | 22        |
| 2.8 Western blot .....                                                                                                                 | 22        |
| 2.9 Alkaline phosphatase staining .....                                                                                                | 23        |
| 2.10 Growth curves.....                                                                                                                | 23        |
| 2.11 Chromatin Immunoprecipitation (ChIP).....                                                                                         | 24        |
| 2.12 ChIP-seq library preparation .....                                                                                                | 25        |
| 2.13 Transcriptional profiling of 3' end mRNA (QuantSeq) .....                                                                         | 25        |
| 2.14 ChIP-seq data analysis .....                                                                                                      | 26        |
| 2.15 QuantSeq data analysis .....                                                                                                      | 27        |
| 2.16 Table of antibodies used for Western blot.....                                                                                    | 28        |
| 2.17 Table of targeting constructs and sgRNAs.....                                                                                     | 28        |
| 2.18 Table of cell lines generated .....                                                                                               | 29        |
| <b>CHAPTER 3 .....</b>                                                                                                                 | <b>30</b> |
| RESULTS .....                                                                                                                          | 30        |
| 3.1 PcG proteins, Eed and Rybp, are dispensable for silencing of shared Polycomb target genes in mouse embryonic stem (mES) cells..... | 30        |
| 3.2 Double disruption of PRC2/cPRC1 and vPRC1 leads to loss of self-renewal, phenocopying double depletion of Ring1 and Rnf2. ....     | 39        |

|                   |                                                                                                                                            |           |
|-------------------|--------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| 3.3               | <i>PRC2/cPRC1 and vPRC1 function redundantly to maintain repression of shared target genes.</i>                                            | 45        |
| 3.4               | <i>Defects on Rnf2 targeting underlie de-repression of PcG target genes upon double disruption of PRC2 and vPRC1.</i>                      | 50        |
| 3.5               | <i>Preferential loss of repression at Polycomb target genes with highest PRC1 and PRC2 occupancy.</i>                                      | 57        |
| <b>CHAPTER 4</b>  |                                                                                                                                            | <b>65</b> |
| DISCUSSION        |                                                                                                                                            | 65        |
| 3.1               | <i>Eed and Rybp are dispensable for silencing of shared target genes in mouse embryonic stem (mES) cells.</i>                              | 66        |
| 3.2               | <i>Double disruption of PRC2 and variant PRC1 leads to loss of self-renewal, phenocopying double depletion of Ring1 and Rnf2.</i>          | 67        |
| 3.3               | <i>PRC2/cPRC1 and vPRC1 function redundantly to maintain repression of shared target genes.</i>                                            | 68        |
| 3.4               | <i>Defects of Rnf2 targeting underlie de-repression of PcG target genes upon double disruption of PRC2 and vPRC1.</i>                      | 70        |
| 3.5               | <i>In mES cells, PcG proteins prevent transcriptional activation at target genes with the highest buildups of PRC2 and PRC1 complexes.</i> | 71        |
|                   | <i>A model for the role of Polycomb complexes in maintenance of ES cell identity.</i>                                                      | 73        |
| <b>REFERENCES</b> |                                                                                                                                            | <b>75</b> |

# Chapter 1

## Introduction

### 1.1 Epigenetics

In multicellular organisms, cells give rise to tissues and organs, which perform vastly different functions. While sharing identical genetic information, each cell type is defined by a distinct gene expression profile, which has to be faithfully transmitted through cell division to maintain cell identity from one generation to the next. At the same time, gene expression programs need to be reset during development to acquire new cell fates during differentiation. Consequently, animal development requires epigenetic mechanisms that provide memory of gene expression states while ensuring plasticity for developmental transitions.

In 1942, Conrad Waddington coined the concept of “epigenetics” to describe aspects of development that could not be explained by genetic principles [1], [2]. Today, we refer to “epigenetics” as “... a change in the state of expression of a gene that does not involve a mutation, but that is nevertheless inherited in the absence of the signal (or event) that initiated the change” [3]. For example, a *trans* epigenetic mechanism is the signal-dependent induction of a transcriptional activator, which in turn binds to its own promoter to drive expression even in the absence of the initial external stimulus. Less well understood are feedback loops of *cis* epigenetic mechanisms, which involve chemical modifications on chromosomes. Apart from semi-conservative transmission of DNA methylation, it remains to be shown if and how posttranslational chromatin modifications are inherited through cell division [4].

## **1.2 Polycomb group of proteins (PcG) are epigenetic regulators**

Polycomb group (PcG) proteins are a well-conserved class of chromatin-modifying factors essential for proper differentiation and development [5]. PcG genes were originally identified 40 years ago in *Drosophila* for their important roles in maintaining repression of *Hox* genes [6]. Flies harboring PcG mutations can establish appropriate *Hox* gene expression patterns but display severe transformations of body segment identity at subsequent developmental stage. This indicates that PcG activities maintain gene silencing by epigenetic mechanisms that persist in the absence of the initial signal. Since then, studies in *Drosophila* and vertebrate model systems revealed widespread roles for PcG beyond repression of segmentation genes [7].

## **1.3 PcG proteins assemble into multi-subunit complexes**

Historically, biochemical purifications in *Drosophila* were key to the realization that PcG proteins assemble into two large families of multi-subunits complexes; Polycomb repressive complex 1 (PRC1) and 2 (PRC2) [8], which harbor distinct enzymatic and non-enzymatic activities. The composition of PcG complexes is generally simpler in flies than in mammals, which possess more diverse subunit compositions or PRC complexes [9]. Yet, the activities of core PcG complexes are conserved from flies to humans. PRC2 is responsible for methylation of lysine 27 of histone H3 (H3K27me) [10]-[12], whereas PRC1 mediates mono-ubiquitylation of lysine 119 on histone H2A (H2AK119ub1) [13], as well as compaction of polynucleosomes [14].

In addition, the Polycomb repressive deubiquitinase complex (PR-DUB), originally discovered in *Drosophila* consisting of the deubiquitinase Calypso and the additional sex combs *Asx*, bears histone H2A deubiquitinase activity which is paradoxically important for silencing [15]. The mammalian Calypso homologue BAP1, associates with *Asx* homologues *ASXL1* and *ASXL2* constituting the PR-DUB1 and PR-DUB2 complexes, respectively [16].

## 1.4 PcG proteins are transcriptional repressors

PRC1 and PRC2 generally co-localize at target sites in the genome, where their combined activities create Polycomb domains constituted by the PcG proteins themselves, and their associated H3K27me3 and H2AK119Ub1. How Polycomb complexes actually shut down transcription of their target genes remains poorly understood [17]. On the one hand, H2A ubiquitination was initially thought to be involved in transcriptional repression [18]-[20], but recent evidence has demonstrated that H2A ubiquitination is not necessary for silencing in *Drosophila* or mammals [21], [22]. On the other hand, Trimethylation of H3K27 (H3K27me2/3) directly antagonizes activating acetylation of H3K27 (H3K27ac) [23]. Moreover, PRC1 chromobox proteins Pc (*Drosophila*) or CBX (mammals) can inhibit the acetyltransferase activity of CBP [24].

Possibly, the most prominent mechanism of PcG-mediated silencing is chromatin compaction, which hinders SWI/SNF-mediated chromatin remodeling [25]. PRC1 can mediate chromatin compaction via positively charged region of fly Psc or mammalian CBX2 subunits [26], [27], and Ph/PHC can contribute to chromatin compaction thorough oligomerization of its SAM domain [28]. Moreover, PRC2 has also been shown to promote chromatin compaction [29], [30].

## 1.5 Diversity of mammalian PcG complexes

Polycomb complexes are much more diverse in mammals than in *Drosophila*, particularly PRC1 which further subdivides into canonical (cPRC1) and variant (vPRC1). All PRC1 complexes assemble around a conserved protein core consisting of one of the two paralogous RING1 proteins (*RING1* or *RNF2*), which bear the E3 ligase activity specific for H2AK119, and one of six different Polycomb group ring-finger domain proteins (*PCGF1-6*), that dictates the enzymatic activity of the complex [31].



Broadly, PRC1 can be categorized into canonical and variant complexes which are distinguished by the presence or absence of *CBX* subunits which confer the ability to bind PRC2-deposited H3K27me3. Thus, in addition to *Ring1/Rnf2*, canonical PRC1 (cPRC1) are consist of a chromobox protein (*CBX2*, *CBX4*, *CBX6*, *CBX7* and *CBX8*), either *PCGF2* or *PCGF4*, and a *Polyhomeotic (Ph)* homologous protein (*PHC1-PHC3*). On the other hand, variant PRC1 complexes lack H3K27me3 binding capacity. Instead of a *CBX* domain protein, vPRC1 complexes bear the zinc-finger domain and YY1-binding protein (*RYBP*) or its paralog *YAF2*, which stimulate the catalytic activity of the complex [32]. Furthermore, the choice of *PCGF* (*PCGF1*, *PCGF3*, *PCGF5* or *PCGF6*) proteins creates more sub-complexes that assemble with diverse subunits that modulate their regulatory functions and affinities to chromatin and DNA. For instance, vPRC1.1, defined by *PCGF1*, includes the H3K36 histone demethylase *KDM2B*, whereas vPRC1.3/5/6 utilize the WD40 domain proteins *WDR5* and *DCAF7* as central scaffold [16].

The catalytic activity of mammalian PRC2 resides on one of the SET-domain containing histone methyltransferases enhancer of zeste (*EZH1* or *EZH2*), which by themselves possess little enzymatic activity *in vitro*, but when in a complex with suppressor of zeste (*SUZ12*), embryonic ectoderm development (*EED*) and the Retinoblastoma Binding Proteins (*RBBP4* or *RBBP7*), they form an active methyltransferase complex [17]. Furthermore, a plethora of accessory proteins have been described to modulate PRC2 enzymatic activity and chromatin binding specificity. Two alternative sub-complexes of PRC2 have been identified through proteomic mapping in human cells. PRC2.1 is characterized by the presence of mutually exclusive Polycomb-like homologs (PCLs) *PHF1*, *PHF19* or *MTF2*. While PRC2.2 is defined by the interaction of the zinc-finger proteins *JARID2* and *AEBP2* [16].

## 1.6 Modes of recruitment to chromatin

In *Drosophila*, DNA regulatory elements termed Polycomb responsive elements (PREs), serve as recruiting platforms for PRCs, where DNA-binding transcription factors are essential for delivering PcG proteins to target sites [9]. However, in vertebrate genomes

PREs appear to be absent, and instead PRCs rely on locus-specific and more general targeting mechanisms [17].

Unlike in flies, DNA-binding transcription factors have a limited contribution to recruitment of PcG complexes. For instance, the zinc-finger protein *SNAIL*, the *RUNT*-related transcription factor 1 (*RUNX1*) and the REI-silencing transcription factor (*REST*) have been shown to interact with and recruit PRC1 or PRC2, although only in specific circumstances [33]-[36]. Moreover, long non-coding RNAs (*lncRNAs*) have also been implicated in PRC recruitment to specific loci. For example, the X inactive specific transcript (*XIST*) is required for PRC2 targeting to the inactive X chromosome [37]. In addition, *HOTAIR* and *KCNQ1 lncRNAs* participate in targeting of PcG subunits to the *HOXD* locus or a paternally imprinted region of mouse chromosome 1, respectively [38], [39]. However, the contributions of transcription factors and *lncRNAs* seem to be predominantly locus-specific and thus are unlikely to account for the widespread and tissue-specific localization of Polycomb complexes observed *in vivo* [17].

In mammalian genomes, the most prevalent sequence signature that characterizes Polycomb-occupied DNA is the presence of CpG islands (CGIs), which are DNA regions of approximately 1-2 kb enriched in CpG dinucleotides, that are devoid of DNA methylation and bound activators [9]. PRC1.1-associated Lys-specific demethylase 2B (*KDM2B*) recognizes unmethylated CpG dinucleotides through its CxxC domain. *KDM2B* localizes to all unmethylated CGIs in the genome, however significant high level PRC1 enrichment is only found at the most repressed loci [40]-[42], which suggests that additional mechanisms regulate PRC1 recruitment to CGIs. Furthermore, it has been recently reported that PRC2 can also be targeted to CGIs via its accessory subunits the PCL proteins *PHF1* and *MTF2*, which bind unmethylated CpG dinucleotides through their conserved winged-helix motifs [43], [44].

## 1.7 Interplay between PRC1 and PRC2

PRC1 and PRC2 generally co-occupy target sites throughout the genome, and this has been attributed to the so called “hierarchical” recruitment model discovered in *Drosophila*.

This model states that *de novo* recruited PRC2 catalyzes H3K27me<sub>3</sub>, which is recognized by PRC1 via its chromobox subunit (*Pc/CBX*), leading to H2A ubiquitination and Polycomb chromatin domain formation, thus placing PRC1 targeting and function downstream of PRC2 activity [45], [46]. However, in mammals deletion of PRC2 core subunits reduces, but not completely abrogates PRC1 targeting, and bears little effect on H2AK119Ub1 in mouse embryonic stem cells (mESCs), demonstrating PRC2-independent recruitment and activity of PRC1 [47].

More recently, an “alternative” recruitment model has been proposed where mammalian PRC1 can recruit PRC2 to target sites through recognition of H2AK119Ub1 [17]. This alternative model is based on the observations that tethering PRC1 components to a region of the genome otherwise devoid of Polycomb proteins, or pericentric regions, induces deposition of H2AK119Ub1, subsequent PRC2 recruitment and trimethylation of H3K27 [48], [49]. Furthermore, *JARID2* and *AEBP2* containing PRC2.2 complexes have been shown to bind to and stimulate H3K27 methylation on chromatin containing H2AK119Ub1 *in vitro* [50], suggesting that PRC2.2 may be the connection between PRC1 activity and PRC2 recruitment.

## **1.8 PcG and TrxG proteins regulate CpG island promoters in mammalian genomes**

The evolutionary conserved Polycomb (PcG) and Trithorax (TrxG) groups of proteins compose two antagonistic chromatin-based gene regulatory systems that modulate the expression of genes involved in cell differentiation and developmental processes. While in *Drosophila*, Polycomb responsive elements (PREs) and Trithorax responsive elements (TREs) recruit PcG and TrxG complexes to target loci. In mammalian genomes, unmethylated CGIs act as PRE/TRE-like sequences that are able to recruit PcG and TrxG complexes to chromatin [31]. CGIs overlap with approximately 70% of mammalian promoters and other gene regulatory elements including enhancers [51], [52], and most CGI-associated promoter regions are decorated with H3K4me<sub>3</sub> irrespective of their transcriptional state, which is in line with the fact that most mammalian H3K4me<sub>3</sub>

methyltransferase complexes bear CxxC domain subunits that recognize unmethylated CpG DNA [53].

TrxG proteins encompass a great diversity of complexes that participate in transcriptional activation. Nevertheless, two families that have been shown to directly counteract PcG-mediated gene silencing. On one hand, the switch/sucrose non-fermentable (*SWI/SNF*) in *Drosophila*, and its mammalian homologs Polymorphic *BRG/BRM*-associated factor (*BAF*) and Polybromo-associated *BAF* (*PBAF*), which harbor ATP-dependent chromatin remodeling activities that regulate many genes implicated in proliferation, cell cycle, signaling and are often mis-regulated in cancer [54]. On the other hand, the COMPASS family comprises diverse histone-modifying complexes. The SET1/COMPASS complexes are responsible for the majority of H3K4me3 in mammalian genomes [53], [55]. *MLL1/2*-containing COMPASS-like complexes methylate less than 5% of H3K4me3, yet *MLL2* is deposits H3K4me3 at bivalent promoters of mouse embryonic stem cells, whereas *MLL1* is essential for methylation of *HOX* loci [55]-[57]. In addition, *MLL3/4*-containing COMPASS-like complexes include the histone demethylase *KDM6/UTX*, which removes the PRC2-deposited H3K27me3 [58].

## **1.9 Embryonic stem cells as a model for cellular differentiation**

Pluripotency is the capacity of a cell to give rise to all somatic cell types of an organism. *In vivo*, pluripotency arises from a developing totipotent zygote toward a blastocyst, and can be captured indefinitely *in vitro* by derivation of self-renewing embryonic stem cells (ESCs) from the inner cell mass (ICM) of a developing blastocyst, or through somatic cell reprogramming strategies [59]-[61]. Murine ESCs (mESCs) were originally isolated almost 40 years ago, and since then they have been thoroughly studied. Under permissive culture conditions mESCs can divide and self-renew indefinitely without losing cellular identity, and thus constitute an *in vitro* surrogate model for the naive epiblast, furthermore when provided with the proper stimuli, they can differentiate into virtually any cell type [62].

ESC culture media is generally supplemented with the cytokine leukemia inhibitory factor (LIF), which in combination with an extracellular matrix like gelatin and fetal calf serum, can substitute fibroblast feeder cells to support proliferation and self-renewal [63], [64]. A hallmark of ESC identity is the expression of the transcription factors *POU5F1* (*OCT4*), *SOX2* and *NANOG*, which positively regulate their own promoters, cooperatively activate other genes essential for pluripotency, and silence lineage-specific regulators, thus preventing spontaneous differentiation [65]. ESCs cultured in serum/LIF often activate conflicting signaling pathways, and although such conditions support naïve pluripotency within the population, they also promote transcriptional and functional heterogeneity among the cells [66], [67]. Consequently, ESCs in serum/LIF encompass at least two populations that resemble the preimplantation epiblast, (naïve state), or the more developmentally advanced “primed” state, however these two subpopulations can interchange their identity cycling back and forth naïve states [68]. Further knowledge in the molecular circuitry that underlies pluripotency lead to the development of defined specific culture conditions that shield ESCs from differentiation stimuli, namely two small molecule kinase inhibitors; PD0325901 (Mek kinase) and CHIR99021 (GSK3 kinase). This combination of inhibitors referred to as “2i” helps maintain the naïve pluripotent state of ESCs [69].

## **1.10 Role of PcG proteins in differentiation and self-renewal**

In mice, embryos mutant for PRC2 core subunits (*Ezh2*, *Eed* or *Suz12*) show lethality during early postimplantation stages [70]-[72], while knockout of accessory subunits like *Jarid2* or *Mtf2* exhibit less severe phenotypes [73], [74]. On the PRC1 side, mice deficient for the core catalytic subunit *Rnf2*, though not its paralog *Ring1*, die during gastrulation, whereas embryos mutant for cPRC1 components (*Pcgf2*, *Pcgf4*, *Cbx2* and *Phc1*) are viable though present homeotic transformations [75], and mice lacking vPRC1 components like *Rybp* or *Kdm2b* display lethality at postimplantation stage and defects on neural development, respectively [76], [77]. These observations outline the key role of PRC1 and PRC2 in maintenance of a proper differentiation program during embryonic development.

In mESCs, knockout of most PcG subunits does not affect the self-renewal or the expression of pluripotency markers. For instance, while *Cbx* proteins and *Rybp* are required for proper mESC differentiation [47], [78], [79], their depletion does not compromise self-renewal. In contrast, double deletion of the catalytic subunits of PRC1, *Ring1* and *Rnf2*, strongly impairs self-renewal [42], [78]-[81]. Moreover, mESCs mutant for core PRC2 components (*Ezh2*, *Eed*, *Suz12*) are able to self-renew and proliferate normally, but show strongly compromised differentiation capacity [70], [72], [82], [83]. In addition, mESCs deficient for *Eed* or *Rnf2* could still form teratomas, though smaller than wild-type cells, double knockout mESCs of *Eed/Rnf2* failed to do so [84]. Therefore, although PcG proteins positively regulate ESC state, they are also necessary for proper differentiation.

## 1.11 Aim of the thesis

Polycomb group of proteins (PcG) are transcriptional repressors essential for proper embryonic development and cellular differentiation of mouse Embryonic Stem (mESCs) cells. However, whether they play a role in self-renewal is unclear. Genetic deletion of genes encoding PRC1 and PRC2 core components often result in embryonic lethality or developmental defects (see 1.10). However, depletion of most PRC1 and PRC2 subunits have minor effects on self-renewal of mESCs, with the exception of *Ring1/Rnf2* double depletion, which results into massive de-repression of Polycomb target genes and loss of self-renewal [81]. Thus, it remains unclear whether PRC1 and PRC2 are indeed required for maintenance of ES cell identity.

PRC1 and PRC2 co-localized at unmethylated CpG islands (CGIs), and given the recent findings showing that subsets of PRC2 and vPRC1 harbor subunits that specifically recognize unmethylated CpG DNA (see 1.6), and that PRC1 and PRC2 can signal each other for recruitment via its associated histone modifications (see 1.7). I hypothesized that PRC2 and vPRC1 act redundantly at target Polycomb target promoters in order to maintain robust silencing of key differentiation genes.

The aim of this thesis is to dissect the contributions of PcG complexes (vPRC1, cPRC1 and PRC2) to the observed distribution of PcG proteins in mES cells, and to investigate the role of PcG complexes in the regulation of self-renewal and mESC identity.

More specifically, we have utilized a chemical genetic protein degradation system, to conditionally control PRC2 or variant PRC1 function in PcG mutant mESCs, which allowed us to ask the following questions:

- What is the contribution of PRC2 function to recruitment of cPRC1 and vPRC1?
- How does impairment of vPRC1 via Rybp knockout affect targeting of cPRC1 and PRC2?
- How does disruption of PRC2, vPRC2 or both in parallel affect:
  - self-renewal capacity
  - the expression of target genes
  - recruitment of PRC1 catalytic core subunit *Rnf2* to chromatin

# Chapter 2

## Materials and methods

### 2.1 Cell culture conditions of mESCs

All mESCs used in this study were derived from haploid HMSc2 termed AN3-12 [85] and cultivated without feeders in high-glucose-DMEM supplemented with 13.5% fetal bovine serum (Sigma), 10 mM HEPES pH 7.4, 2 mM GlutaMAX (Gibco), 1 mM sodium pyruvate (Sigma), 100 U penicillin/ml (Sigma), 0.1 mg streptomycin/ml (Sigma), 1x non-essential amino acids (Sigma), 50 mM beta-mercaptoethanol (Gibco) and recombinant LIF (37°C, 5% CO<sub>2</sub>).

### 2.2 Handling of haploid mESCs

All cell lines were kept haploid only for the genetic manipulation steps to promote homozygous modifications. For all biochemical and molecular assays diploid cells were used. Haploid cells spontaneously become diploid in culture and can be isolated via Fluorescence-activated cell sorting (FACS). Isolation of haploid mESCs entails staining genomic DNA without killing the cells by incubating with 20 µg/ml Hoechst 33342 (Thermo Scientific Fisher) for 30 min at 37 °C and 5 % CO<sub>2</sub> prior to FACS.

### 2.3 Gene trap cell lines

A haploid gene trap mES cell line derived from AN3-12 cells (as described in [86]) containing a retrovirus-based enhanced gene trap (Retro-EGT) inserted between exons 2 and 3 of the *Eed* gene was used as a parental Eed-KO cell line. *Eed* expression was rescued by inverting the Retro-EGT (see Fig. 3.2.1 B) and Ring1 gene was knocked out using CRISPR-Cas9 technology to obtain Ring1-KO cells (Fig. 3.2.1 A). *Eed/Suz12* double knockout cell lines were generated by knocking *Suz12* out via CRISPR-Cas9 on the Eed-KO parental line (Fig 3.2.4 A).



## 2.4 Inversion of gene-trap cassettes

To invert the gene-trap cassettes, mES cells were transiently transfected with a FlpO recombinase and DsRed expressing plasmid, using Amaxa nucleofector. The cells were FACS-sorted for mCherry+ cells 24–48h after nucleofection, then sorted again after 48–72h for mCherry- cells to avoid stable integration of the plasmid. Then 1,000 cells were seeded on 15-cm tissue culture dishes, 8–10 days after sorting, colonies were picked and subjected to genotyping.

## 2.5 Genotyping of gene trap cassettes

Genomic DNA was isolated by ethanol precipitation of mES cell lysates and subsequently used for PCR genotyping. Genotyping was carried out by PCR amplification using three primers; first fwd: CGACCTCGAGTACCACCACACT, inv fwd: AAACGACGGGATCCGCCATGTCA and common rev: TATCCAGCCCTCACTCCTTCTCT [86]. On unmodified gene trap DNA, first fwd and common rev primers yields a product of 343 bp, while an inverted gene trap produces a 443 bp band with the inv fwd and common rev primers. Further gel electrophoresis analysis of the PCR products determines the orientation of the gene trap of both alleles.

## 2.6 CRISPR/Cas9-mediated gene knockout in mESCs

CRISPR single guide RNAs (gRNAs) were designed using the online tool of the Zhang lab (<http://crispr.mit.edu>) and cloned into a modified lentiCRISPRv2 (one vector system) expression vector expressing hSpCas9, the chimeric gRNA and a fluorescent protein reporter (GFP or BFP). Two gRNAs were generated to excise a predicted functionally relevant region of the gene of interest and each sgRNA was clones in a GFP or BFP lentiCRISPRv2 vector. To obtain Knockout mES cells the two complementary sgRNAs were transiently transfected using Amaxa nucleofector. The cells were FACS-sorted for GFP+ and BFP+ cells 24–48h after nucleofection, then sorted again after 48–72h for GFP- and BFP- cells to avoid stable integrations of the plasmids. Then 1,000 cells were seeded on 15-cm tissue culture dishes, 8–10 days after sorting, colonies were picked and subjected to PCR genotyping.

## 2.7 AID-tagging of endogenous genes

Before introducing the AID tag, mESCs were transduced with a lentiviral vector expressing SFFV driven OsTir1 (F-box receptor Tir1) with a puromycin resistance (SOP). One day after transduction cells were selected with Puromycin ( $1\text{ }\mu\text{g/mL}$ ) for 5 days. Then, the cells were transiently co-transfected with a cutting vector expressing hSpCas9, a sgRNA targeting the gene of interest and a BFP, and a plasmid containing the Auxin inducible degron (AID) tag followed by a 3xFlag tag and a GFP, flanked by two 500 bp homology arms to each side. 24-48h after transfection the cells were FACS-sorted for GFP+ and BFP+ cells, then sorted again after 48-72h for GFP+ and BFP- cells to avoid stable integrations of the cutting plasmid. Then 1,000 cells were seeded on 15-cm tissue culture dishes, 8-10 days after sorting, colonies were picked and subjected to PCR genotyping and further characterization by sanger sequencing and western blot.

## 2.8 Western blot

Nuclear extract from  $10 \times 10^6$  mESCs was obtained by lysis in Buffer A (final: 25 mM Hepes pH 7.6, 5 mM  $\text{MgCl}_2$ , 25 mM KCl, 0.05 mM EDTA, 10 % Glycerol, 1 mM DTT, 1 mM PMSF, 1x Complete Mini protease inhibitor) followed by collection in RIPA buffer (final: 150 mM NaCl, 1% triton, 0.5% sodium deoxy- cholate, 0.1% SDS, 50 mM Tris pH 8.0). Nuclear extracts were homogenized by sonication in a Diagenode Bioruptor and concentration was determined by Bradford assay (Biorad).  $20\text{ }\mu\text{g/lane}$  total protein was run on Novex Life Technology NuPAGE 4-12 % Bis-Tris gels in Invitrogen NuPAGE MES SDS Running Buffer and transferred on a Merck Chemicals Immobilon-P Membrane (PVDF  $45\text{ }\mu\text{m}$ ). The membrane was blocked (5 % non-fat dry milk in 1x PBS, 0.1 % Tween 20) and incubated with primary antibodies. Finally, the membrane was incubated with corresponding secondary HRP coupled antibodies (5 % non-fat dry milk in 1x PBS, 0.1 % Tween 20), developed using Clarity Western ECL Substrate (Biorad) and imaged by a ChemiDoc XRS+ Imaging system (Biorad).

## **2.9 Alkaline phosphatase staining**

1,000 cells were grown on gelatin-coated glass cover slips inside 10cm cell culture dishes to allow sparse colonies to form. The cells were seeded in regular mESC medium (see 2.1) and allowed to adhere to the coverslips, then after 12-16h the medium was exchanged with either regular mESC medium or with medium containing 250  $\mu$ M Auxin (indole-3-acetic acid - IAA). At day 5, coverslips were washed with 100 mM TRIS pH 8 prior to Alkaline phosphatase (AP) activity assay, which was performed using the VECTOR Blue Alkaline Phosphatase (Blue AP) Substrate Kit (Vector Laboratories - VECSK-5300) according to the manufacturer's instructions. After AP staining, colonies were fixed by incubating with Histofix solution (4% formaldehyde in phosphate-buffered saline - PBS) overnight. Next day, coverslips were washed with 1x with PBS and mounted onto glass slides using Fluorescence Mounting Medium (Dako S3023) and left to dry for at least 16 hours. Images from representative colonies were acquired with a bright field microscope (Axioplan2) using a 10x objective. The Panoramic FLASH 250 II (3DHISTECH) wide field microscope was utilized to scan whole slides in order to obtain sufficient images for quantification of staining intensity and morphology parameters. The images obtained from scanning whole slides were processed with a custom workflow created by Thomas Lendl using Definiens Developer image analysis software.

## **2.10 Growth curves**

3300 cells per well were seeded in regular mESC medium (see 2.1) on flat bottom 96 well plates in 4 replicates. The cells were allowed to adhere to the wells for 4-6h, then the medium was exchanged with either regular mESC medium or with medium containing 250  $\mu$ M IAA. The cell count was measured after 24, 48 and 72h by flow cytometry on a LSR Fortessa (BD Biosciences). Cells were prepared for flow cytometry counting by washing 1x with PBS, then incubating with 20uL Trypsin to dissociate the colonies and stop the reaction by adding 100uL of mESC medium. Live cells were measured in a volume of 50uL per well.

## 2.11 Chromatin Immunoprecipitation (ChIP)

For Chromatin Immunoprecipitation, 30-50 x 10<sup>6</sup> mESCs were trypsinized for 6-8 min prior to quenching with FBS containing ES cell medium. 25 x 10<sup>6</sup> mES cells were collected, washed in once in 1x PBS and crosslinked with formaldehyde at a final concentration of 1 % for 7 min. The crosslinking was stopped on ice and with glycine at final 0.125 M concentration. The crosslinked cells were pelleted by centrifugation for 5 min at 1200g at 4 °C. Nuclei were prepared by washes with NP-Rinse buffer 1 (final: 10 mM Tris pH 8.0, 10 mM EDTA pH 8.0, 0.5 mM EGTA, 0.25% Triton X-100) followed by NP-Rinse buffer 2 (final: 10 mM Tris pH 8.0, 1 mM EDTA, 0.5 mM EGTA, 200 mM NaCl). Afterwards the cells were prepared for shearing by sonication by two washes with Covaris shearing buffer (final: 1 mM EDTA pH 8.0, 10 mM Tris-HCl pH 8.0, 0.1% SDS) and resuspension of the nuclei in 0.9 mL Covaris shearing buffer (with 1x protease inhibitors complete mini (Roche)). The nuclei were sonicated for 15 min (Duty factor 5.0; PIP 140.0; Cycles per Burst 200; at 4°C) in 1 ml Covaris glass cap tubes using a Covaris E220 High Performance Focused Ultrasonicator. Input samples were prepared from 25  $\mu$ L sonicated lysate. Therefore, chromatin was RNase A and Proteinase K digested and crosslink reversed overnight at 65 °C. DNA was then precipitated and shearing of DNA was confirmed to be between 500-1000 bp by agarose gel electrophoresis. Crude chromatin lysate was further processed by spinning at 20000 g at 4 °C for 15 min and supernatant used for ChIP. An equivalent of 50  $\mu$ g DNA for PcG proteins and H3K4me3 or 100  $\mu$ g DNA for H3K27me3 and H2AK119ub1, was incubated overnight in 1x IP buffer (final: 50 mM HEPES/KOH pH 7.5, 300 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% DOC, 0.1% SDS), with following antibodies at 4 °C on a rotating wheel: 0.5  $\mu$ L H3K27me3 (Diagenode, C15410195), 3  $\mu$ L Ring1B (Cell Signaling, D22F2), 1.5  $\mu$ L *Suz12* (Cell Signaling, D39F6), 1.5  $\mu$ L H3K4me3 (Millipore, 05-745R), 2  $\mu$ L *Cbx7* (Abcam, ab21873), 1.5  $\mu$ L RYBP (Sigma Aldrich, PRS2227), 1.5  $\mu$ L H2AK119ub (Cell Signaling, D27C4). The overnight IPs were incubated with BSA-preblocked Protein G coupled Dynabeads (Thermo Fisher Scientific) for more than 6h at 4 °C on a rotating wheel. IPs were subsequently washed 5x with 1x IP buffer (final: 50 mM HEPES/KOH pH 7.5, 300 mM NaCl, 1 mM EDTA, 1% Triton-X100, 0.1% DOC, 0.1% SDS), or 1.5x IP buffer for H3K27me3 and H2AK119ub1, 3x with DOC buffer (10 mM Tris pH 8, 0.25 mM LiCl, 1 mM EDTA, 0.5% NP40, 0.5% DOC) and 1x with TE (+50 mM NaCl). The DNA was then eluted 2x with 150  $\mu$ L Elution buffer (final: 1% SDS,

0.1 M NaHCO<sub>3</sub>) for 20 min each at 65 °C. The eluate was treated with RNase A and Proteinase K and crosslink reversed overnight at 65 °C. The IP DNA was PCIA extracted and precipitated and quantified using qPCR on a CFX Connect Real-Time PCR Detection System (Biorad) as a quality control before proceeding to library preparation.

## **2.12 ChIP-seq library preparation**

Libraries were prepared using the NEXTflex ChIP-Seq kit (Bio Scientific) following the “No size-selection cleanup” protocol and doubling the incubation times for all enzymatic steps. Each sample of ChIPed DNA was end-repaired and ligated to unique barcoded adaptors to produce individual libraries. Libraries corresponding to samples to be directly compared to each other (e.g. +IAA vs -IAA) were pooled together and purified using 1 volume of Agencourt AMPure XP (Beckman Coulter). The pooled libraries were eluted with 25 µL of elution buffer (NEXTflex ChIP-Seq kit) and amplified using the KAPA Real-Time Library Amplification Kit (peqlab) following the kit instructions. Finally, the amplified libraries were size-selected to fragments of 200-800 bp by running them on 1.5% agarose gel, and staining with 1x SYBR Gold (Thermo Fischer) to visualize the DNA on a blue light LED screen and cut the appropriate fragments. The size selected libraries were gel purified with the Monarch DNA Gel extraction kit (NEB).

## **2.13 Transcriptional profiling of 3' end mRNA (QuantSeq)**

75,000 cells per well were seeded on 24 well plates in regular mESC medium (see 2.1), in triplicates. The cells were allowed to adhere to the wells for 4-6h, then the medium was exchanged and cells grown for additional 48 hours. After adhered to the wells, the cells were grown in either regular mESC medium or in medium containing 250 µM IAA added for a duration of 6, 24 or 48h before collecting all samples at the same time (reverse time course). Then, the wells were washed 1x with PBS and cells were immediately lysed with buffer RLT (RNeasy kit) and total RNA was isolated using the RNeasy kit (Qiagen) following the on-column DNase digestion protocol. The concentration of the RNA was measured by Nanodrop and the quality of the RNA was assessed by running the RNA samples on a Fragment Analyzer (Advanced Analytical Technologies) using the RNA Analysis Kit (Advanced Analytical Technologies). Once the quality was checked, the RNA

was submitted to the Next Generation Sequencing facility at the VBCF for QuantSeq FWD library preparation and high throughput sequencing.

## 2.14 ChIP-seq data analysis

**Processing of the raw reads:** Standard quality control was carried out with FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>), then the reads were mapped to the mm10 genome using bowtie2 (<http://bowtie-bio.sourceforge.net/bowtie2/index.shtml>) generating BAM files as output, and peaks were called using MACS2 (<https://github.com/taoliu/MACS>).

**GC bias correction:** GC-bias files were calculated using the computeGCBias tool from deepTools (<https://deeptools.readthedocs.io/en/develop/index.html>), called peaks were excluded from the GC bias calculation to avoid taking into account CpG islands (CGIs) enriched by PcG proteins and histone modifications during the ChIP reaction. This way the GC bias calculation reflects the technical biases rather than the ChIP enrichment of CGI DNA.

**GC correction:** correctGCBias (deepTools) was used to generate GC-corrected BAM files from the original BAM files and the GC-bias files. Then biological replicates were merged using SAMtools (<http://samtools.sourceforge.net/>).

**Visualization:** Coverage tracks (bigWig files) were generated with bamCoverage (deepTools) normalizing the signal to Reads Per Kilobase per Million mapped reads (RPKM) and extending the reads to 200 bp. Then heat maps to visualize the enrichments were done with computeMatrix and plotHeatmap both from deepTools. IGV as used to browse the coverage track along the genome and screenshots of coverage tracks were generated using the profile tool from fluff (<http://fluff.readthedocs.io/en/latest/index.html>).

## 2.15 QuantSeq data analysis

**Processing of the raw reads:** Standard quality control was carried out with FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>), then SlamDunk (<http://t-neumann.github.io/slamdunk/docs.html>) was used to map the reads (mm10 genome) and generate count tables. A bed file with custom annotations of the 3' UTR regions of AN3-12 mESCs was kindly provided by Veronika Herzog (Ameres Lab) to improve mapping and counting of the reads.

**Differential Expression Analysis:** Count tables were loaded into R (<https://www.r-project.org/>) and all subsequent analysis were carried out using RStudio (<https://www.rstudio.com/>) within the Tidyverse (<https://www.tidyverse.org/>) data manipulation framework. The DESeq2 (<https://bioconductor.org/packages/release/bioc/html/DESeq2.html>) package was used to calculate pairwise differential expression between different conditions, always using all three replicates per condition.

**Visualization:** All plots to explore and visualize the data were generated by in-house generated scripts using a combination of R packages: ggplot, cowplot, VennDiagram and RColorBrewer, statistical tests were performed and added using ggpubr.

**Gene Ontology Analysis:** Analysis and visualization Gene Ontology (GO) enrichment of up and down-regulated genes was performed using clusterProfiler (<https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html>), in combination with the org.Mm.eg.db genome wide annotation for mouse (<https://bioconductor.org/packages/release/data/annotation/html/org.Mm.eg.db.html>).

## 2.16 Table of antibodies used for Western blot

| Name                                | Cat #   | Vendor          | Host   | Dilution |
|-------------------------------------|---------|-----------------|--------|----------|
| <i>Rnf2</i>                         | 5694S   | Cell Signaling  | rabbit | 1:1,000  |
| <i>Cbx7</i>                         | 07-981  | Merck Millipore | rabbit | 1:1,000  |
| <i>Rybp</i>                         | 41787S  | Cell Signaling  | rabbit | 1:1,000  |
| <i>Eed</i>                          | 61203   | Active Motif    | mouse  | 1:500    |
| <i>LaminB1</i>                      | ab16048 | Abcam           | rabbit | 1:5,000  |
| <i>Ring1</i>                        | 2820S   | Cell Signaling  | rabbit | 1:1,000  |
| Goat anti-Rabbit IgG (H+L), HRP     | 656120  | Invitrogen      | goat   | 1:10,000 |
| Anti-Mouse IgG (H+L), HRP Conjugate | W4021   | Promega         | goat   | 1:2,500  |

## 2.17 Table of targeting constructs and sgRNAs

| Construct                                                    | sgRNA sequence (5' - 3') |
|--------------------------------------------------------------|--------------------------|
| CP045_pBluescript SK-<br>_Ring1B_homology_arms_AID_GFP_Nterm |                          |
| CP046_pBluescript SK-<br>_Eed_homology_arms_AID_GFP_Nterm    |                          |
| CP047_pBluescript SK-<br>_Rybp_homology_arms_AID_GFP_Cterm   |                          |
| CP048_pBluescript SK-<br>_Suz12_homology_arms_AID_GFP_Cterm  |                          |
| CP049_sgRNA_Ring1B_Nterm_BFP_Cas9_CP35                       | GCACAGCCTGAGACATTTCT     |
| CP051_sgRNA_Rybp_Cterm_BFP_Cas9_CP35                         | GAATCTTTCTGAGATTGCACA    |
| CP052_sgRNA_Suz12_Cterm_BFP_Cas9_CP35                        | GCAGTGTCTGTTCAAAACATG    |
| CP053_RRL-SFFV-OsTir1(noXhoI)-3x-Myc-T2A-Puro (SOP)_Zuber    |                          |
| jz031_lentiCRISPRv2-gfp_Suz12-KO-gRNA_1_BsmBI                | CCAATAAGACAAGTCCCTAC     |
| jz030_lentiCRISPRv2-bfp_Suz12-KO-gRNA_2_BsmBI                | CTGTTTAGAGTAACTCGTCC     |
| jz037_lentiCRISPRv2-gfp_Ring1A-KO-gRNA_1_BsmBI               | TCTTTTGTTGTTGCAGGCC      |
| jz036_lentiCRISPRv2-bfp_Ring1A-KO-gRNA_2_BsmBI               | ATCGTCACCGCCCTGCGGAG     |



## 2.18 Table of cell lines generated

| Cell line name | Genotype                                               |
|----------------|--------------------------------------------------------|
| AN3-12         | Wild type (parental)                                   |
| 515D10-E6      | Eed gene trap knockout (Eed-KO)                        |
| 515D10-A12     | Eed-KO; Rybp-AID                                       |
| 515D10-C3      | Eed-KO; Rybp-AID                                       |
| AN3E9-A1       | Rybp CRISPR knockout (Rybp-KO)                         |
| AN3E9-A8       | Rybp-KO; Suz12-AID                                     |
| AN3E9-B11      | Rybp-KO; Suz12-AID                                     |
| D10B3C10-G9    | Eed gene trap rescue; Ring1 CRISPR knockout (Ring1-KO) |
| D10B3C10-B10   | Ring1-KO; AID-Rnf2                                     |
| D10B3C10-C2    | Ring1-KO; AID-Rnf2                                     |
| D10B8-B4       | Eed-KO; Suz12 CRISPR knockout (KO); Rybp-AID           |
| D10B8B4-A4     | Eed-Rescue; Suz12 CRISPR knockout (KO); Rybp-AID       |

# Chapter 3

## Results

### 3.1 PcG proteins, *Eed* and *Rybp*, are dispensable for silencing of shared Polycomb target genes in mouse embryonic stem (mES) cells.

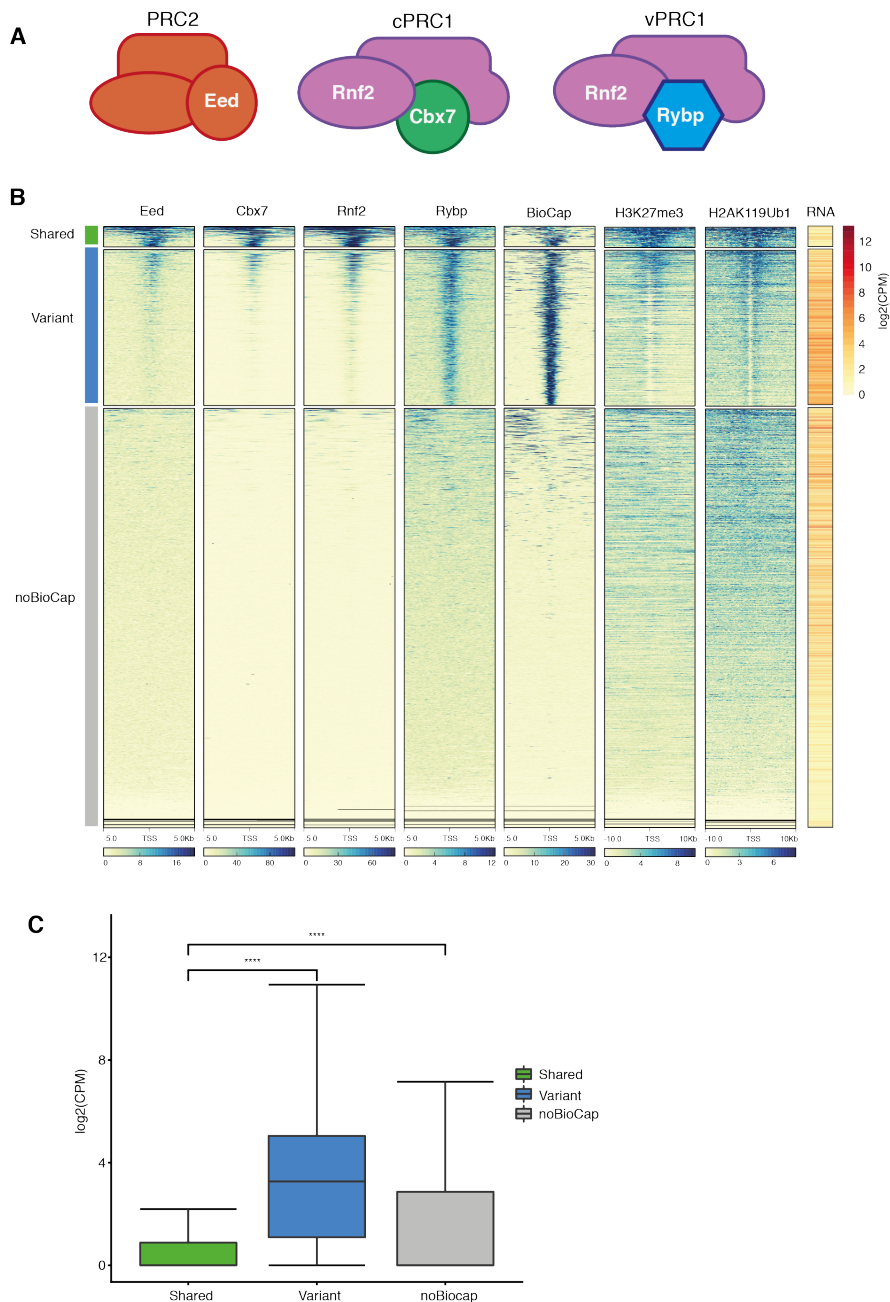
To investigate the relationship of the genomic occupancy and target gene regulation by Polycomb Repressive Complexes we used Chromatin Immunoprecipitation coupled to high throughput sequencing (ChIP-seq) to map the genome-wide distribution of PRC1 and PRC2 core subunits and their respective histone modifications in mouse embryonic stem (mES) cells. In addition, we performed QuantSeq (see Material and Methods) to profile global gene expression based on quantification of 3' end mRNAs (Fig. 3.1.1 B). Since it was previously shown that Polycomb group of proteins (PcG) occupancy negatively correlates with gene expression levels [87], we decided to focus my analyses on promoters adjacent to all Ref-seq (mm10) annotated transcription start sites (TSS). ChIP-seq enrichment of PcG subunits was mapped five kilo-bases (kb) upstream and downstream of the TSS whereas Polycomb-dependent histone modifications, H3K27me3 and H2AK119ub1, were mapped 10kb upstream and downstream of the TSS. It has been previously established that PcG proteins localize to un-methylated CpG islands (CGIs) [88] which overlap with approximately 70% of gene promoters in mammalian genomes [89]. To identify un-methylated CGI DNA in mES cells, we referred to BioCap sequencing data [90] as a blueprint for un-methylated CGIs overlapping with the set of promoters.

To classify TSSs based on the distribution of different Polycomb Complexes, we performed k-means clustering (k=3) of ChIP-seq signal for core subunits of PRC2 (*Eed*), canonical PRC1 (*Cbx7*), variant PRC1 (vPRC1), core PRC1 (*Rnf2*) and un-methylated CGIs (BioCap). This analysis gave rise to three distinct clusters. TSSs in cluster 1 are un-methylated CGIs which show high enrichment for subunits of canonical (cPRC1) and variant PRC1 (vPRC1) as well as PRC2. This cluster will be referred to as “Shared” class of genes/TSSs (965 TSSs). TSSs in cluster 2 overlap with CGIs and are similarly enriched

in vPRC1 subunit *Rybp*. In contrast, this cluster is characterized by only low signals for PRC2 and cPRC1 based on *Eed* and *Cbx7* binding, respectively. Consequently, we will refer to cluster 2 as the variant-specific or “Variant” class of genes/TSSs (7,333 TSSs). Lastly, since the third cluster shows neither enrichment for PcG subunits nor BioCap, we will refer to it as “noBioCap” (19,712 TSSs).

Importantly, the observed differences in promoter occupancy by different PcG complexes argue for distinct targeting mechanisms of cPRC1, PRC2 and vPRC1. Moreover, Variant target genes are significantly more transcriptionally active than Shared genes (Fig. 3.1.1 C) suggesting that PcG complexes also differ in their modes of gene regulation. To elucidate the regulation of cPRC1, PRC2 and vPRC1, all further analyses will be discussed in the framework of “Shared” and “Variant” Polycomb target genes.

Two distinct models for the mode of PcG complex recruitment have been proposed. The “Hierarchical model” poses that PRC1 is recruited to chromatin through interaction with PRC2-dependent H3K27me3. This interaction involves recognition of the histone mark by chromodomain-containing subunit of the *Cbx* family, of which *Cbx7* is the most highly expressed in mouse embryonic stem cells (mESCs) [78]. The second model, also referred to as “Alternative” model states that affinity of accessory subunits of vPRC1 and PRC2 to unmethylated CGIs, i.e. *Kdm2b* (vPRC1) and *Mtf2/Phf1* (PRC2), promote recruitment to target gene promoters (see 1.6). In light of these two models, we decided to investigate the individual contributions of PRC2 and vPRC1 to mediating the targeting of PcG complexes and their histone modifications in mESCs. To that end, we mapped core subunits of PRC2, cPRC1 and vPRC1, as well as H3K27me3 and H2AK119ub1 in wild type, *Eed* knockout and *Rybp* knockout cells. *Eed* is necessary for assembly and catalytic function of the PRC2 complex [70], [83], whereas *Rybp* is exclusive to all vPRC1 complexes, apart from *Ring1/Rnf2*, and it has been shown to be mutually exclusive with the *Cbx* subunits as it binds the same surface on the *Ring1/Rnf2* protein [32], [47], [91].

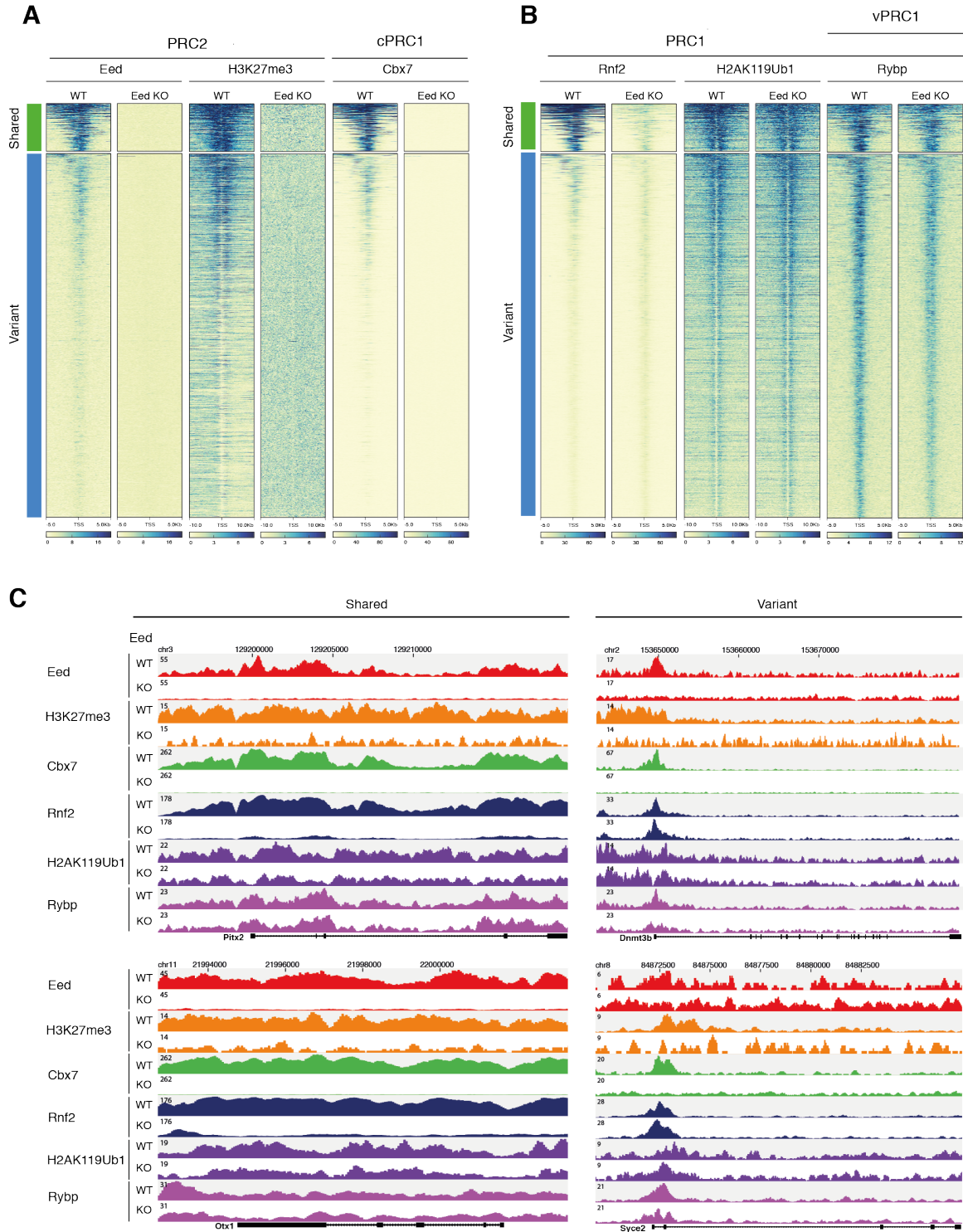


**Figure 3.1.1: Canonical and variant PRC1 partially overlap with PRC2 at silent CGI promoters in mouse ES cells. (A)** Schematic representation of core subunits of PRC2, canonical and variant PRC1 complexes. ChIP sequencing analyses of Eed reflects PRC2 distribution. Genomic *Rnf2* enrichment reflects all PRC1 binding sites while *Cbx7* and *Rybp* respectively indicate the distribution of canonical and variant complexes. **(B)** Heat maps display ChIP-seq signal (FPKM) of core PcG proteins, their associated histone modifications, and gene expression (QuantSeq). TSSs were classified (K-means clustering) into three categories: Shared (965 TSSs), Variant (7,333 TSSs) or noBioCap (19,712 TSSs). Within each category, TSSs were ranked according to the enrichment of canonical PRC1 (*Rnf2* and *Cbx7*), variant PRC1 (*Rybp*), PRC2 (*Eed*) subunits and un-methylated CGIs (BioCap). **(C)** Boxplot quantification of gene expression (QuantSeq) at the three categories from B. Significance (p-value) calculated using Wilcoxon test (\*\*\*\*:  $p < 2.2 \times 10^{-16}$ ).

## **PRC2 function is critical for cPRC1 but not vPRC1 targeting**

To study the role of cPRC1 and vPRC1 in silencing of shared target genes, we sought to analyze changes in PcG complex occupancy and activity in *Eed* and Rybp KO cells. To characterize *Eed* KO mESCs, we obtained a gene trap mES cell line that harbors a gene trap cassette between exons 2 and 3 of the *Eed* gene. The gene trap contains a splice acceptor and a poly-A signal such that it leads to early termination of the transcript, and therefore *Eed* loss of function. Furthermore, the gene trap cassette is flanked by F3/FRT and incompatible loxP/lox5171 recombination sites, such that the cassette can be inverted to rescue *Eed* transcription (Fig 3.2.1 B).

Several important aspects need to be considered in order to ensure unbiased, quantitative comparison of ChIP-seq enrichments between wild-type and mutant mESCs. Bias towards GC-rich sequences is a known problem for high throughput sequencing methodologies including ChIP-seq [92]. Moreover, we have observed that libraries with vastly different DNA concentrations, as it happens when the epitope is decreased globally, tend to amplify to different degrees, even when amplified in parallel for the same number of cycles (data not shown). This can make side-by-side comparisons challenging and potentially create artifacts. To circumvent this problem, we ligated barcoded adaptors to the ChIP-enriched genomic DNA fragments, then pooled untreated with IAA treated samples for each genotype, and proceeded to library amplification (see Methods). The intrinsic bias of short read sequencing towards high GC containing fragments is specially concerning for Polycomb protein and histone modification ChIPs because PcG proteins localize to CGIs. Thus, to make sure that the ChIP-seq signal reflects real enrichment of ChIPed DNA rather than a sequencing bias, we calculated the GC bias per replicate, excluding the peaks (PcG enriched sites), and used it to generate GC bias corrected files (see Methods).



**Figure 3.1.2: PRC2 function is required for recruitment of canonical PRC1, but not variant PRC1. (A) and (B)** Heat maps display enrichments of core subunits of PRC2, canonical PRC1, and variant PRC1 and their respective histone modifications in wild-type and *Eed* KO mESCs. **(C)** Screen shots of ChIP-seq tracks for PRC2/cPRC1 and vPRC1 subunits and respective histone modifications at selected genes representative of the Shared or Variant categories in wild-type and *Eed* KO mESCs.

ChIP-seq analysis revealed that genetic disruption of *Eed* (KO) leads to global loss of H3K27me3, completely abolishing *Cbx7* binding, and drastically reducing *Rnf2* occupancy (Fig. 3.1.1 A and C). These results are in line with the “hierarchical model” of H3K27me3-dependent recruitment of PRC1 and consistent with previous reports [46], [93]. Interestingly, despite the dramatic loss of cPRC1 at Shared target gene, deposition of H2AK119ub1 remained largely unaffected (Fig. 3.1.1 B and C). Unlike *Cbx7*, occupancy of vPRC1-specific core subunit *Rybp* is only mildly affected at shared and variant target gene promoters in *Eed* KO mESCs. This is consistent with persistence of H2AK119ub1, *Rybp* has been proposed to enhance catalytic Rnf2 activity within vPRC1 [32]. Hence, in agreement with previous studies, we conclude that vPRC1 targeting is independent of PRC2 and cPRC1 [47], [49].

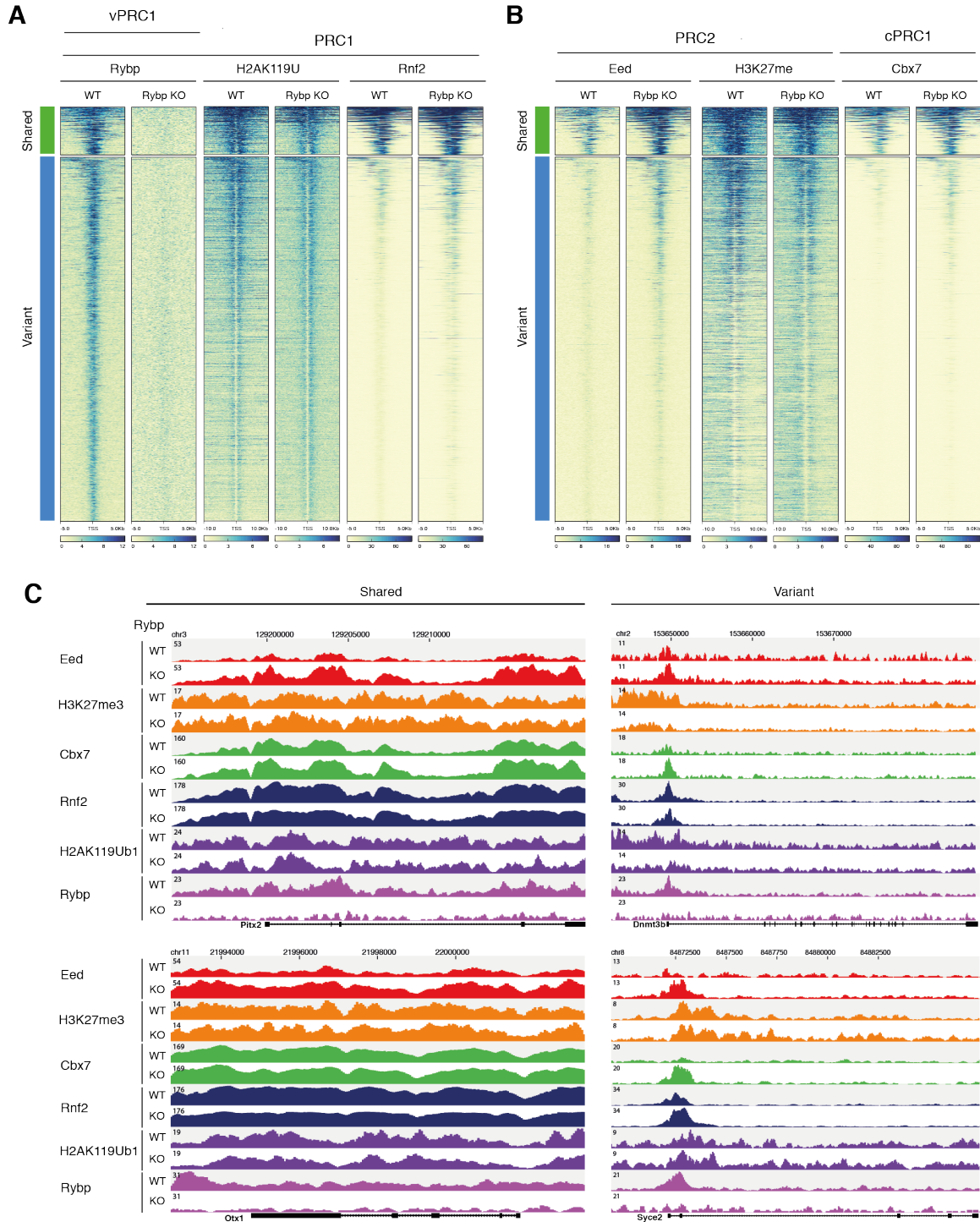
### ***Rybp* loss impairs vPRC1-dependent H2AK119ub1 catalysis but is dispensable for PRC2 and cPRC1 targeting**

We employed CRISPR-Cas9 gene editing targeting the *Rybp* gene and successfully isolated mESCs with loss-of-function (LOF) mutations. *Rybp* KO cells show reduced levels of H2AK119ub1, in line with previous observations that vPRC1 complexes are more catalytically active than cPRC1 [32]. Targeting of PRC2 and cPRC1 is not disrupted by loss of *Rybp*. If anything, we observed that *Cbx7*, *Eed* and *Rnf2* show a slight increase in occupancy. Furthermore, deposition of H3K27me3 remained largely unchanged supporting the notion that PRC2/cPRC1 and vPRC1 rely on distinct targeting mechanism. Thus, we conclude that loss of *Rybp* disrupts assembly of vPRC1 and impairs its catalytic activity at shared and variant target genes.

### **Disruption of cPRC1 or vPRC1 function does not affect repression of shared target genes**

To investigate the consequences of individual loss-of-function of *Eed* or *Rybp* (sKO) on shared and variant-specific target gene regulation, we calculated differential expression of sKO vs wild type (WT) mES cells (Fig. 3.1.4). *Rybp* KO cells have more differentially expressed genes than *Eed* KO (Fig 3.1.4 A and B), which is not surprising given that *Rybp* occupies more TSSs than *Eed* (Fig. 3.1.1 B). Only around 10% (181) of differentially

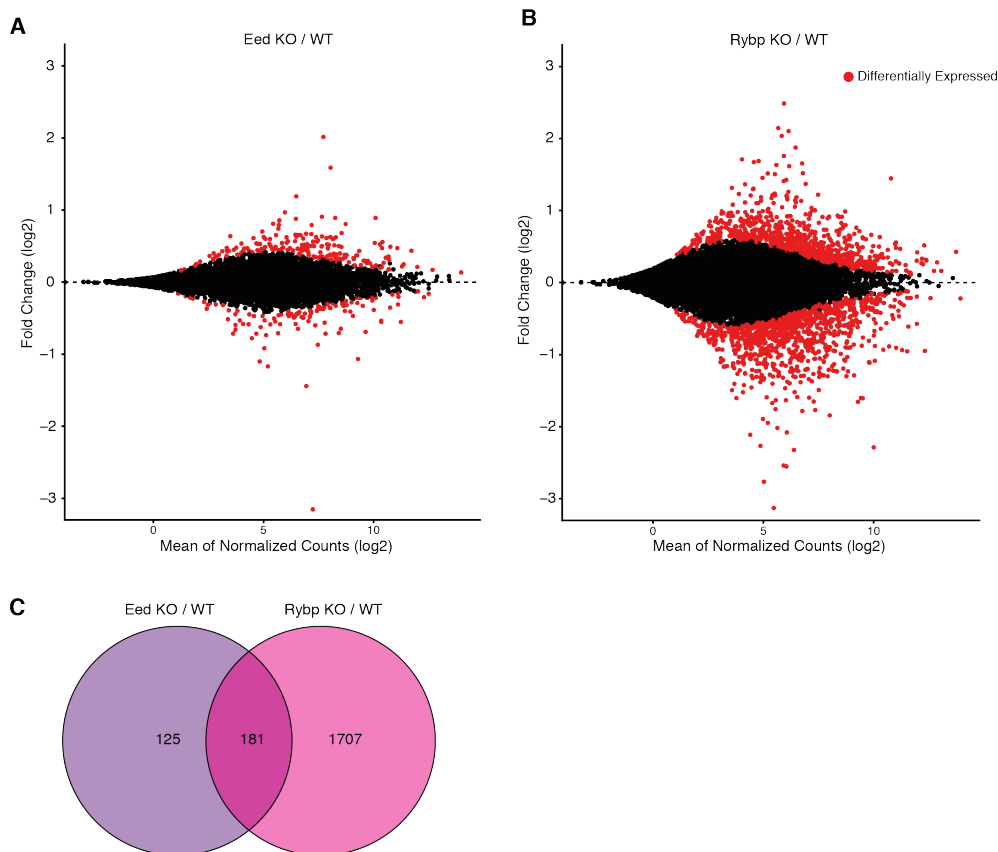




**Figure 3.1.3: Binding of PRC2 and cPRC1 is increased at shared target genes in response to deletion of vPRC1 core subunit Rybp.** Heat maps displaying enrichments of variant PRC1 (**A**), PRC2 and canonical PRC1 (**B**) subunits and histone modifications in wild type vs *Rybp* KO cells. (**C**) Screen shots of ChIP-seq tracks of PRC2/cPRC1 and vPRC1 proteins and histone modifications at selected genes representative of the Shared or Variant categories in wild type or *Rybp* KO mESCs.

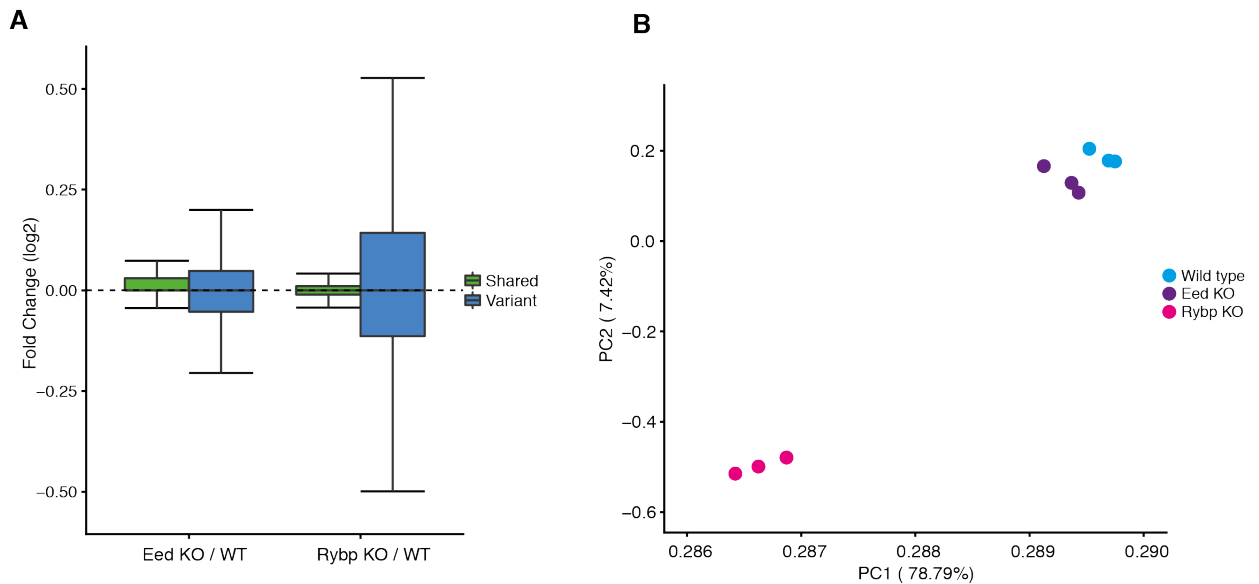


expressed genes in *Rybp* KO cells are shared with *Eed* KO. In contrast, in *Eed* KO more than half of all differentially expressed genes can be assigned to shared target genes (Fig. 3.1.4 C). To discriminate the impact of *Eed* and *Rybp* KO, we plotted the expression changes of sKO over WT cells at Shared and Variant Polycomb target genes. Disruption of cPRC1 or vPRC1 had very limited effect on the repression of Shared genes. By comparison, Variant genes show up and down-regulation, particularly in case of *Rybp* KO mES cells. Finally, we performed principal component analysis (PCA) on whole expression profiles comparing wild-type and sKO cells to assess the global differences in gene expression among the three genotypes. All three experimental replicates of wild-type and *Eed* KO cells cluster closely together indicating high similarity of global gene expression. As expected based, experimental replicates of *Rybp* KO mES cells cluster away from the other two genotypes reflecting the global differences in gene expression at Variant target genes (Fig 3.1.5.B).



**Figure 3.1.4. Single knockouts of *Eed* or *Rybp* lead transcriptional changes of different magnitude, but partially overlapping:** MA plots displaying differential expression (log2 fold change) between *Eed* KO and wild-type mESCs (A) or *Rybp* KO and wild-type mESCs (B). Differentially expressed genes are colored in red (padj < 0.1). (C) Overlap of differentially expressed genes between *Eed* KO / WT and *Rybp* KO / WT.

Taken together these results show that variant and canonical PRC1 complexes co-occupy a subset of transcriptionally silent CGI promoters. However, cPRC1 and vPRC1 are recruited to their targets by different mechanisms. Recruitment of cPRC1 largely depends on PRC2 function, whereas vPRC1 relies on different mechanisms, presumably the recognition of un-methylated CGI's by *Kdm2b* or other components of the vPRC1 complexes. At the transcriptional level, neither inactivation of PRC2 by knockout of *Eed*, nor disruption of vPRC1 by loss of *Rybp* result into de-repression of Shared target genes. These results raise the possibility that silencing of shared target genes in mES cells relies on independent repressive activities of PRC2/cPRC1 and vPRC1. The limited impact of individual cPRC1 or vPRC1 disruption on H2AK119ub1 at Shared target genes suggests that interference with canonical PRC1 is functionally compensated by variant PRC1 and vice versa. In extension, we hypothesize that simultaneous disruption of *Eed* and *Rybp* would abrogate Polycomb repression and result in de-repression of Shared targets akin to double deletion of *Ring1* and *Rnf2* [81].

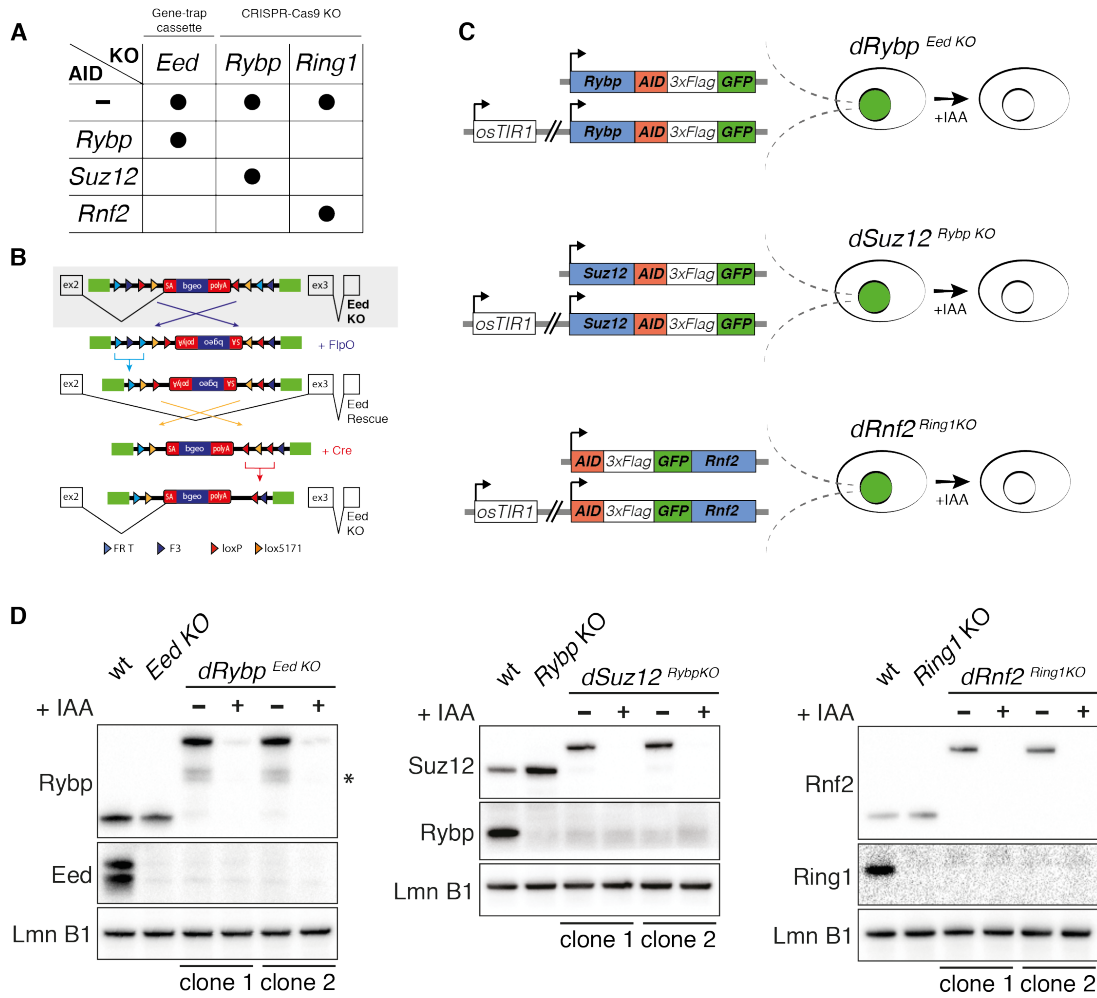


**Figure 3.1.5. Impairment of PRC2 or vPRC1, by knockout of *Eed* or *Rybp*, affects Variant, but not Shared Polycomb target genes: (A)** Boxplots indicating the gene expression changes (log2 fold change) between wild-type and *Eed* KO mESCs or wild-type and *Rybp* KO mESCs at Shared and Variant target genes. **(B)** Principal component analysis (PCA) of variance-stabilized transformed QuantSeq read counts for wild-type, *Eed* KO mESCs and *Rybp* KO mESCs.

### 3.2 Double disruption of PRC2/cPRC1 and vPRC1 leads to loss of self-renewal, phenocopying double depletion of *Ring1* and *Rnf2*.

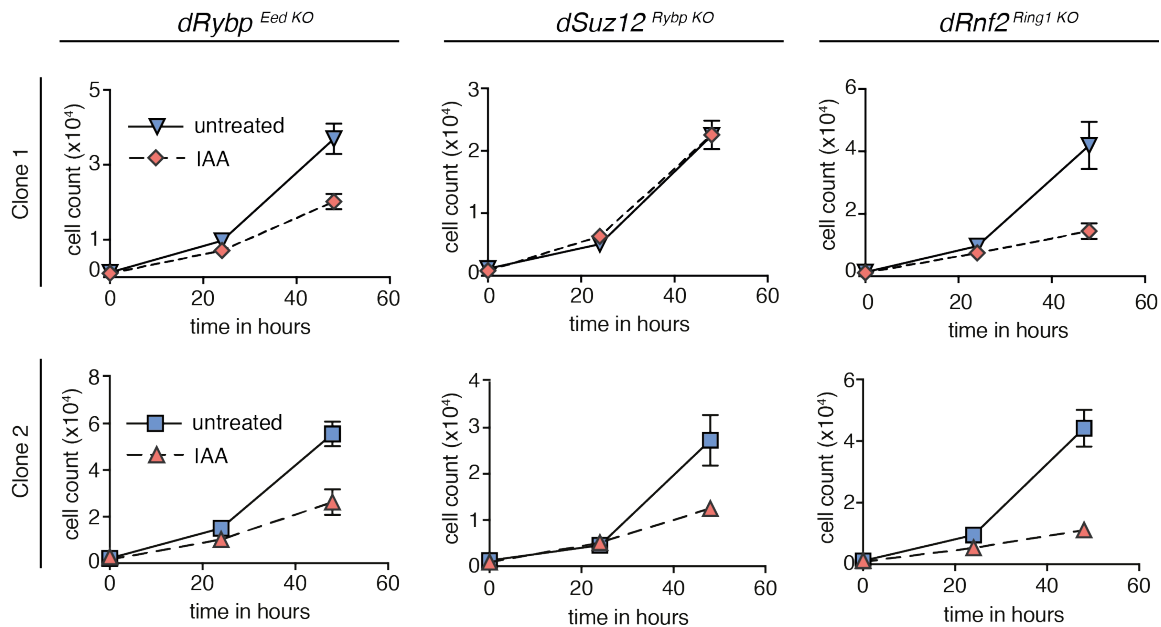
To test the hypothesis of compensation between PRC2/cPRC1 and vPRC1 to silence Shared target genes, we set out to simultaneously disrupt PRC2 and vPRC1 complexes in mES cells. We employed CRISPR-Cas9 gene editing to generate *Rybp* LOF mutations in *Eed* KO mES cells. Despite successful gene editing as determined by PCR genotyping, we failed to recover homozygous *Eed* and *Rybp* double KO mES cells (data not shown) indicating a synthetic genetic interaction between these two Polycomb genes. To interrogate the potential synthetic lethality between PRC2/cPRC1 and vPRC1 at the level of gene expression and chromatin changes, we decided to take a conditional approach involving small molecule-induced, synchronous target protein degradation (Fig. 3.2.1 A). The Auxin-inducible degron (AID) system has been shown to direct target protein degradation via the proteasome within less than one hour [94], [95].

*Eed* KO mES cells were transduced with F-box protein *osTir1* (*TIR1*) and endogenous *Rybp* was C-terminally tagged with the Auxin-inducible degron (AID) (*dRybp*<sup>*Eed* KO</sup>) (Fig. 3.2.1 A, B and C). Conversely, in *Rybp* KO mES cells, we transduced *TIR1* and modified endogenous *Suz12* with a N-terminal AID tag (*dSuz12*<sup>*Rybp* KO</sup>) (Fig. 3.2.1 A, C). Genetic targeting of the *Suz12* gene in *Rybp* KO mES cells was selected for convenience since *Eed* protein contains four alternative translation start sites. Finally, we generated *Ring1* KO mES cells and *Ring1* mES cells with the AID inserted in the *Rnf2* gene (*dRnf2*<sup>*Ring1* KO</sup>) (see Methods) (Fig. 3.2.1 A, C). Two independent clones were isolated per genotype (see Methods), and target protein degradation at 250  $\mu$ M IAA was confirmed by Western blot. Fig. 3.2.1 D shows the IAA mediated degradation of AID-tagged proteins in each of the two clones selected upon 24 hours of IAA treatment, as well as the absence of *Eed*, *Rybp* and *Ring1* at the protein level in *dRybp*<sup>*Eed* KO</sup>, *dSuz12*<sup>*Rybp* KO</sup> and *dRnf2*<sup>*Ring1* KO</sup> clones respectively, along wild-type control cells.



**Figure 3.2.1: Experimental strategy to test functional redundancy between cPRC1 and vPRC1 in mESCs.** (A) Design table of the constructed single and double mutant mESCs. Endogenous genes (*Rybp*, *Suz12* and *Rnf2*) were genetically modified in haploid, single KO mESCs by inserting an in-frame Auxin-inducible degron (AID). Single loss-of-function KO mutations were generated either by gene-trap cassette insertion (*Eed* KO) or CRISPR-Cas9 gene editing (*Rybp* KO and *Ring1* KO). Single KO cell lines serve as controls. (B) Schematic representation of the gene-trap *Eed* KO cell line. The gene-trap cassette, which contains a splice acceptor site (SA) and a poly-A termination signal, was inserted in haploid mouse ES cells between exons 2 and 3 of the *Eed* gene. The splice acceptor site can be reversed using non-compatible FRT/F3 sites (triangles) to restore *Eed* expression, and inverted once more using loxP/lox5171 sites to disrupt *Eed* expression (scheme adapted from [86]). (C) Schematic representation of AID tagging in single KO mESCs expressing the F-box receptor *TIR1* transgene. Endogenous *Rybp* and *Suz12* were genetically modified by introducing a DNA sequence encoding the AID degron, a 3xFLAG epitope and GFP (*dRybp*<sup>*Eed* KO</sup> and *dSuz12*<sup>*Rybp* KO</sup>). In addition, we generated *Ring1* KO mESCs harboring an AID-3xFLAG-GFP in front of endogenous *Rnf2* (*dRnf2*<sup>*Ring1* KO</sup>). +IAA indicates addition of Auxin and results in proteasome-mediated degradation of the GFP fusion protein. (D) Western blot analysis of wild-type, single KO mESCs and two independent mESC clones of *dRybp*<sup>*Eed* KO</sup>, *dSuz12*<sup>*Rybp* KO</sup> and *dRnf2*<sup>*Ring1* KO</sup>. IAA treatment (250  $\mu$ M) for 24 hours results in loss of AID-tagged protein. \* indicates unspecific band in *dRybp*<sup>*Eed* KO</sup>. Lamin B1 (Lmn B1) serves as loading control.

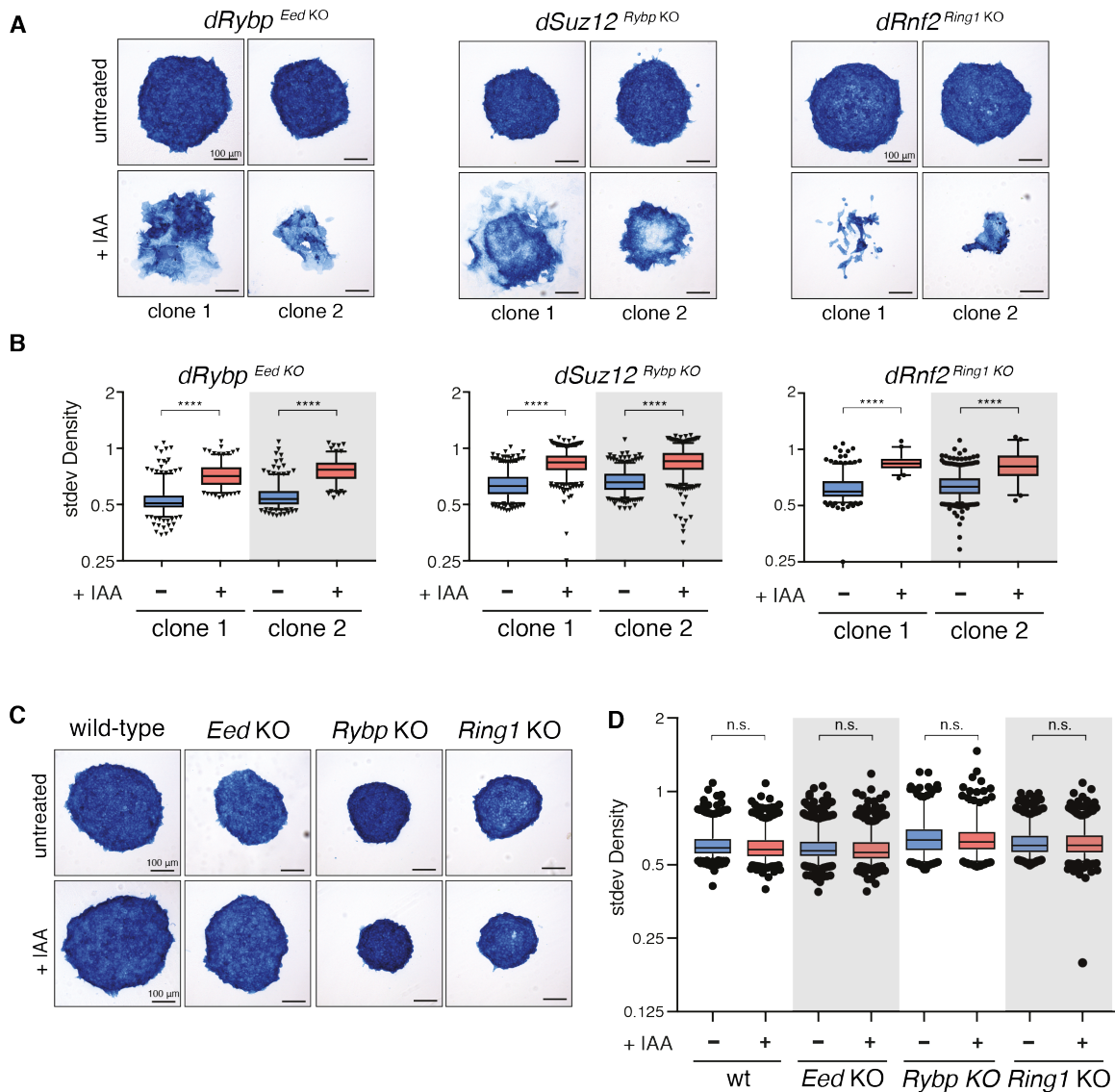
Next, we performed cell growth assays in combination with Alkaline Phosphatase (AP) staining to investigate whether double disruption of *Eed* and *Rybp* would impair self-renewal capacity, as suggested by our failure to obtain double KO mES cells for *Eed* and *Rybp*. Fig. 3.2.2 depicts growth curves measuring the total cell count every 24 hours, as measured by flow cytometry (see Methods). IAA treatment impairs proliferation of the *dRybp*<sup>*Eed* KO</sup> and *dRnf2*<sup>*Ring1* KO</sup> clones as early as 24 hours, whereas one clone of *dSuz12*<sup>*Rybp* KO</sup> displayed reduced cell count until 48 hours, while the other clone did not show any measurable difference even after 48 hours (Fig. 3.2.2).



**Figure 3.2.2: Disruption of *Rybp* in *Eed* KO mESCs or *Suz12* in *Rybp* KO mESCs resembles loss of self-renewal upon depletion of *Rnf2* in *Ring1* KO mESCs.** (A) Growth curves showing two independent serum-cultivated mESC clones of each AID-tagged genotype, either grown in absence (untreated) or presence of 250 $\mu$ M IAA. Live cells were quantified by flow cytometry every 24 hours. Displayed is the median and standard deviation of four replicates.

Alkaline phosphatase (AP) is expressed in most cell types, nevertheless pluripotent stem cells are characterized for high expression [96]. Thus, AP staining is a widely used method to phenotypically distinguish pluripotent from differentiated cells. Fig. 3.2.3 shows AP staining images, obtained by light microscopy, of AID-tagged colonies (A) and wild-type or single KO controls (C) grown in the presence or absence of IAA for five days. IAA treatment disrupts the round dome-shaped morphology, typical of mES cells, in all AID-

tagged clones, but not in the wild-type or sKO controls. Furthermore, quantitative analysis of the standard deviation of the AP staining optical density (stdev Density), which measures the degree of heterogeneity of the AP staining within each colony, shows that IAA treatment increases the apparent differentiation of the cells within the colonies for all AID-tagged clones, but not in the wild-type or sKO controls (Fig. 3.2.3 B and D).



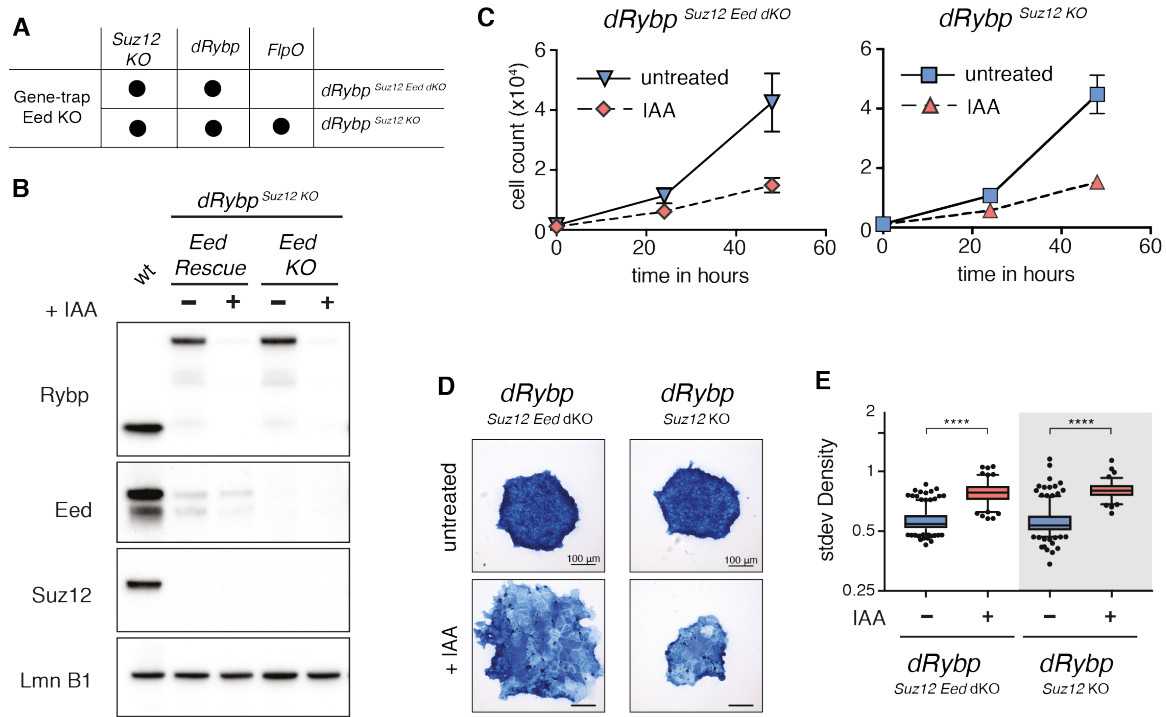
**Figure 3.2.3: Disruption of *Rybp* in *Eed KO* mESCs or *Suz12* in *Rybp KO* mESCs resembles changes in ESC morphology upon depletion of *Rnf2* in *Ring1 KO* mESCs. (A) and (C) Alkaline phosphatase staining of mESC colonies cultured in absence (untreated) or presence or 250 $\mu$ M IAA for 5 days. (B) and (D) stdev Density quantifies degree of mESC colony heterogeneity in response to five days of IAA treatment. Statistical significance calculated using unpaired t-test.**

Conditional genetic deletion of *Rnf2* in *Ring1* KO mES cells has been shown to result in loss of self-renewal and upregulation of differentiation marker genes with 48 hours [81]. Based on similar phenotypic changes in mES cell proliferation and morphology between IAA-treated *dRybp*<sup>*Eed* KO</sup> and *dRnf2*<sup>*Ring1* KO</sup>, we conclude that simultaneous loss of *Eed* and *Rybp* lead to loss of self-renewal resembling previous reports of *Ring1* and *Rnf2* genetic interaction [81]. However, the fact that the phenotype in *dSuz12*<sup>*Rybp* KO</sup> is not as strong as in *dRybp*<sup>*Eed* KO</sup> suggests three possible explanations: 1) incomplete degradation of *Suz12*-AID, 2) increased tolerance to PRC2 disruption of *Rybp* KO cells, due to transcriptional differences to wild-type and *Eed* KO cells (Fig 3.1.5. B), or 3) functional differences between *Eed* and *Suz12* as proposed by a study that shows *Eed* being part of a subset of PRC1 complexes [97].

In order to test if the different phenotype is linked to incomplete *Suz12* degradation, we sought to genetically disrupt *Suz12* in *Rybp* KO mESCs using the CRISPR-Cas9 system. However, similar to CRISPR targeting of *Rybp* in *Eed* KO mESCs, we failed to obtain homozygous *Rybp/Suz12* double knockout cells (data not shown). Next, we tested whether genetic LOF mutation of *Eed* or *Suz12* leads to similar phenotypes in response to *Rybp* degradation. Taking advantage of the conditional gene trap in *Eed* KO mES cells, we sought to restore *Eed* expression by inverting the gene trap cassette to a non-disruptive position. In *Eed* KO mES cells, we successfully generated *Suz12* LOF mutations using CRISPR-Cas9 gene editing. Phenotypically, *Eed/Suz12* dKO mES cells do not differ from *Eed* sKO mES cells. Next, we targeted endogenous *Rybp* introducing the AID tag to obtain conditional *Rybp*-AID mES cells in the background of *Eed/Suz12* double KO (*dRybp*<sup>*Suz12/Eed* dKO</sup>). Finally, we inverted the gene trap cassette to restore *Eed* expression to generate conditional *Rybp*-AID cells in the background of *Suz12* single KO cells (*dRybp*<sup>*Suz12* KO</sup>) as depicted on Fig. 3.2.4 A. Western blot analysis shows the IAA-induced degradation of *Rybp* and the protein levels of *Eed* and *Suz12*. *Eed* levels are significantly lower in the *Eed* rescue cell line (Fig 3.2.4 B), which likely reflects *Eed* protein instability not assembled in a PRC2 complex, as we have previously observed. Alternatively, it is also possible that the gene trap could impair *Eed* transcription to some extent.



Having these two cell lines, both deficient for PRC2, but lacking only *Suz12* or *Eed* and *Suz12* (Fig. 3.2.4 B), we characterized the consequences of IAA-mediated *Rybp* degradation on proliferation and self-renewal. Figure Fig 3.2.4 C-E shows our results from cell growth assays and AP staining in both cell lines revealing similar phenotypes, comparable to the phenotype in the *dRybp*<sup>*Eed* KO</sup> cells. Taken together, we conclude that double disruption of PRC2 and vPRC1 leads to loss of self-renewal, phenocopying the loss of *Ring1* and *Rnf2*, and that the milder phenotype observed in *dSuz12*<sup>*Rybp* KO</sup> cells may be due to inherent transcriptional differences of *Rybp* KO cells to wild-type and *Eed* KO cells.

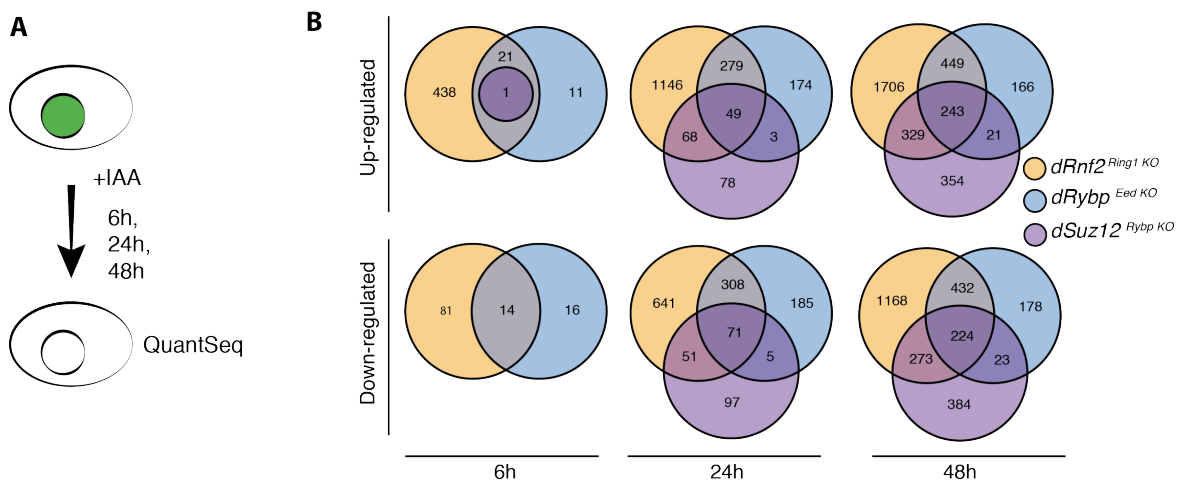


**Figure 3.2.4: *Eed* and *Suz12* knockouts do not display additive effects upon Auxin-induced *Rybp* degradation.** (A) Design table of the mutant mESC lines to compare the impact of *Rybp* depletion in the background of loss-of-function mutations in *Suz12* alone or *Eed* and *Suz12*. *Suz12* was genetically inactivated using the CRISPR-Cas9 gene editing in *Eed* gene-trap KO mESCs. Subsequently, endogenous *Rybp* was modified by introducing the AID degron as in Fig 3.2.1-C (*dRybp*<sup>*Suz12 Eed dKO*</sup>). Finally, the *Eed* gene-trap cassette was inverted using the FlpO recombinase to restore *Eed* expression (*dRybp*<sup>*Eed* KO</sup>). (B) Western blots showing protein expression of *Eed* and *Suz12* in both mutant mESC lines, as well as *Rybp* degradation after 24 hours of IAA treatment. (C) Growth curves displaying the proliferation phenotypes in both cell lines. (D) Alkaline phosphatase staining of representative mESC colonies cultured in absence (untreated) or presence or 250 $\mu$ M IAA for 5 days. (E) stdev Density quantifies degree of mESC colony heterogeneity in response to five days of IAA treatment. Statistical significance calculated using unpaired t-test.



### 3.3 PRC2/cPRC1 and vPRC1 function redundantly to maintain repression of shared target genes.

Next, we aimed to gain insight to the consequences of disrupting PRC2 and vPRC1 on gene expression. Double depletion of *Ring1* and *Rnf2* causes de-repression of PcG target genes and down-regulation of pluripotent stem cell markers [81], therefore we hypothesized that double depletion of *Eed* and *Rybp* or *Rybp* and *Suz12* would yield similar gene expression changes. Taking advantage of the synchronicity and high temporal resolution of Auxin-induced degradation, we collected cells to analyze gene expression changes at 6h, 24h and 48h of IAA treatment (Fig. 3.3.1 A). We utilized 3' mRNA quantification using QuantSeq to profile gene expression. Quantseq requires 10-times less reads as regular mRNA-seq to accurately quantify polyadenylated RNAs [98]. This approach facilitated parallel analysis of wild-type and untagged single KO controls, as well as, two clones for each AID-tagged genotype, all in parallel and in triplicates. The AID-tagged cell lines underwent a number of genetic manipulation steps that may alter the expression of some genes in a clone specific manner. Therefore, to minimize these effects, we calculated differential gene expression between untreated and IAA treated samples always using the two independent clones per genotype, this way the adjusted p-values reflect the differences between treatments, rather than clonal differences.



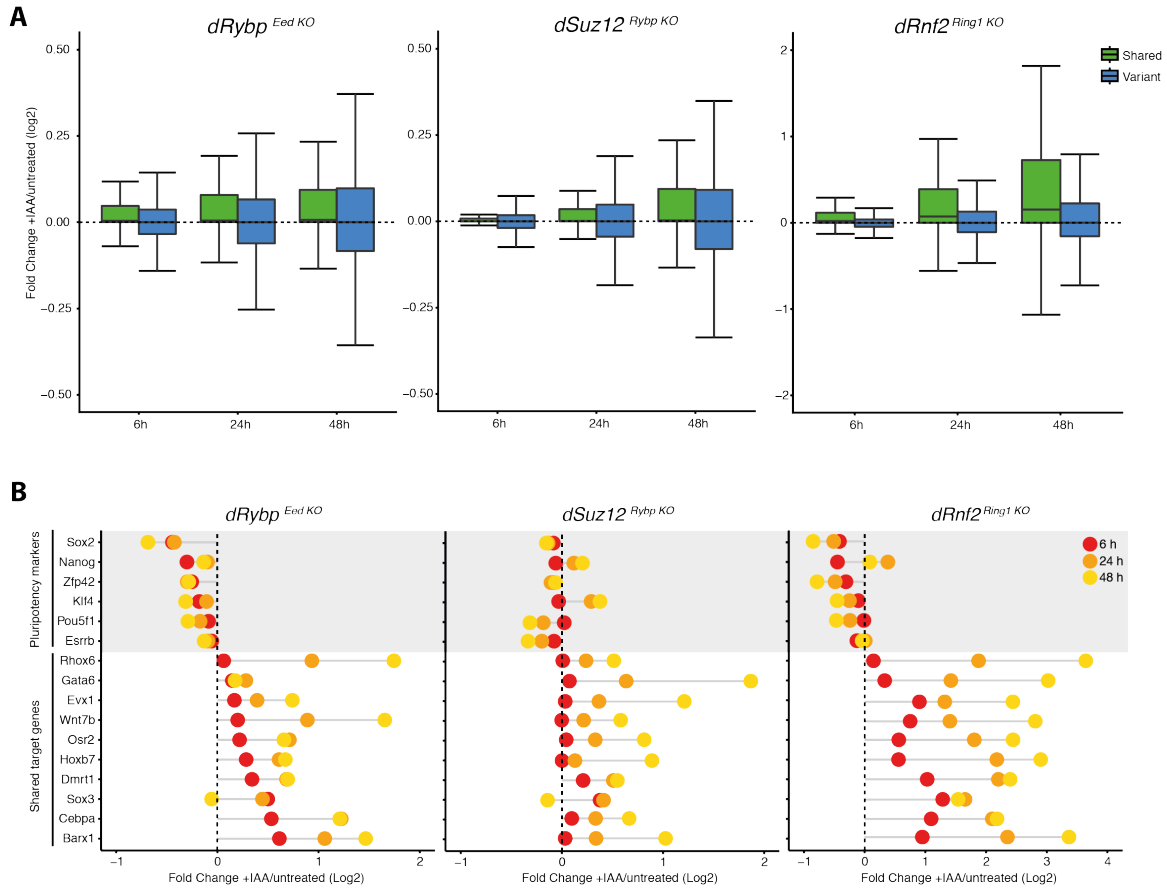
**Figure 3.3.1: Time course analysis of differential expression in the double depleted mESC lines. (A)** Scheme of the experimental design of gene expression analysis. AID-tagged mESC lines were treated with IAA for 6, 24 or 48 hours along with the untreated controls in a reverse time course, then collected at the same time for RNA extraction and QuantSeq library preparation. **(B)**

Venn diagrams representing the overlap of differentially expressed genes among the three genotypes compared untreated vs IAA-treated mESC lines at each time point. Differential gene expression ( $p_{adj} < 0.1$ ) of each genotype was calculated taking into account two clones per genotype, each sequenced in triplicates.

Figure 3.3.1. B shows the overlap of differentially expressed genes among the three genotypes. IAA-treated *dRybp*<sup>Eed KO</sup> and *dSuz12*<sup>Rybp KO</sup> exhibit substantial overlap with IAA-treated *dRnf2*<sup>Ring1 KO</sup>, particularly after the 24-hour time point. In addition, *dRybp*<sup>Eed KO</sup> and *dSuz12*<sup>Rybp KO</sup> also share a significant overlap in differentially expressed genes.

Our hypothesis of compensation between PRC2/cPRC1 and vPRC1 predicts that *Eed/Rybp* or *Rybp/Suz12* double depletions should cause de-repression of Shared target genes. To investigate this possibility, we plotted the fold-change in gene expression of all Shared or Variant target genes at each time point, and for all three genotypes. As displayed in Fig. 3.3.2 A, the gene expression changes in response to IAA treatment in *dRybp*<sup>Eed KO</sup> and particularly *dSuz12*<sup>Rybp KO</sup>, are less pronounced and appear to be delayed compared to *dRnf2*<sup>Ring1 KO</sup>. However, all three genotypes show up-regulation of Shared target genes, whereas Variant target genes are de-regulated in both directions, consistent with the compensation hypothesis.

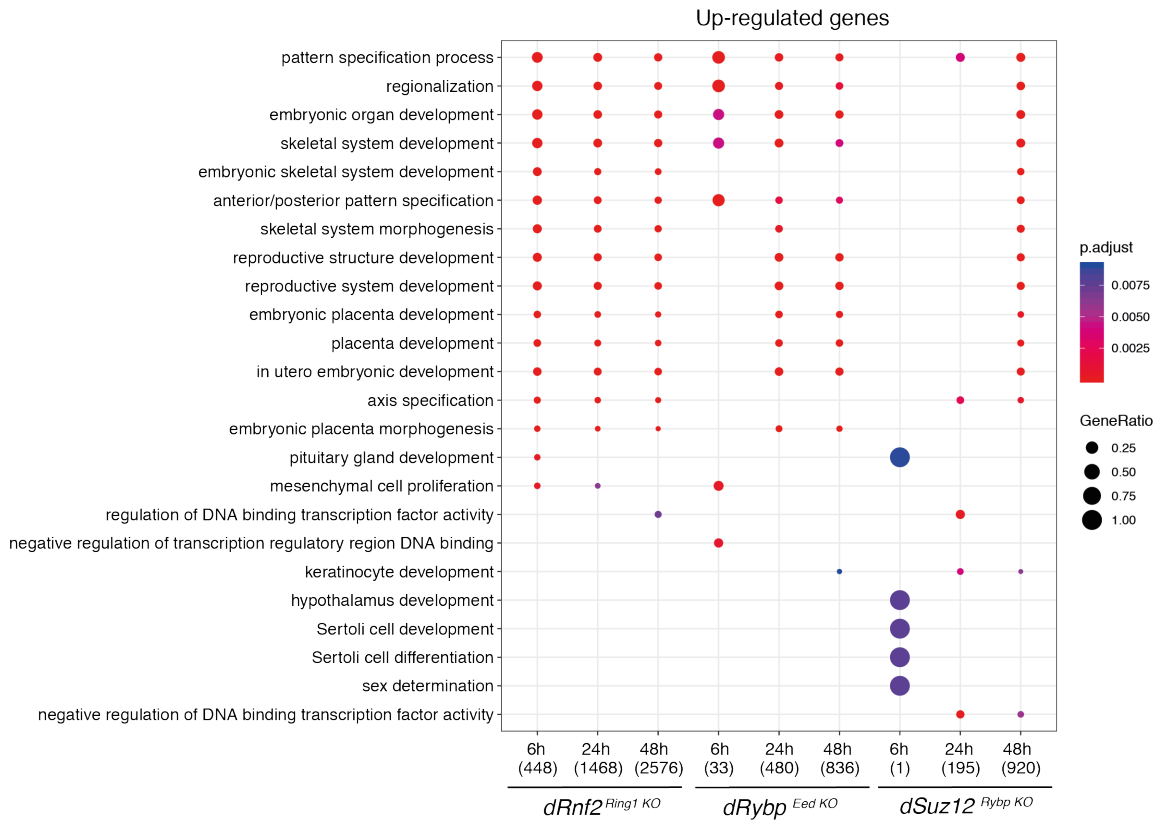
Next, we examined the expression changes of key pluripotency marker genes at each time point (Fig. 3.3.2 B gray background). Both *dRybp*<sup>Eed KO</sup> and *dRnf2*<sup>Ring1 KO</sup> display down-regulation of *Sox2*, *Zfp42* (*Rex1*), *Klf4* and *Pou5f1* (*Oct4*), whereas *dSuz12*<sup>Rybp KO</sup> only shows significant down-regulation on *Pou5f1* and *Esrbb*, which goes in line with the milder loss of self-renewal phenotype observed for *dSuz12*<sup>Rybp KO</sup>. We also looked at a few examples of Shared PcG targets that are differentiation markers, they show gradual up-regulation, again with *dRnf2*<sup>Ring1 KO</sup> showing the strongest de-repression, followed by *dRybp*<sup>Eed KO</sup>, and *dSuz12*<sup>Rybp KO</sup> showing delayed de-repression (Fig 3.3.2. B, white background).



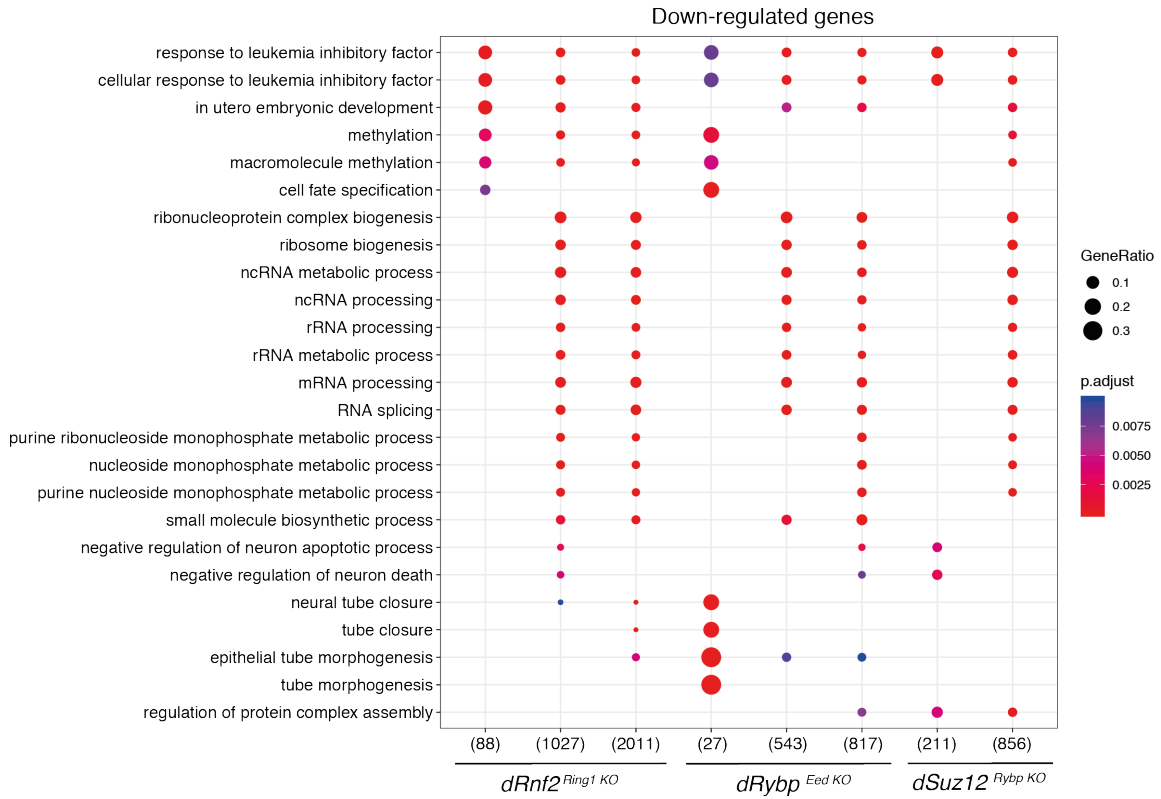
**Figure 3.3.2: Simultaneous perturbation of cPRC1 and vPRC1 targeting results in de-repression of Shared target genes. (A)** Box plots of gene expression changes (log2 fold change) of *dRybp<sup>Eed KO</sup>*, *dSuz12<sup>Rybp KO</sup>* and *dRnf2<sup>Ring1 KO</sup>* at Shared and Variant genes after IAA treatment for 6h, 24h and 48h. **(B)** Dot charts display dynamic expression changes after 6, 24 and 48 hours of IAA treatment at selected mESC pluripotency markers (gray shading), or Shared Polycomb target genes.

Having established that double depletion of *Ring1/Rnf2*, *Eed/Rybp* and *Rybp/Suz12* exhibit similar patterns of gene expression changes, we performed Gene Ontology (GO) enrichment analysis of up and down-regulated genes to determine the biological function of the set of differentially expressed genes. Figure 3.3.3 displays a comparison among the enriched functional GO categories (biological process) of up-regulated genes at each time point, for the three genotypes. While the enriched categories remain constant at all time points in IAA-treated *dRnf2<sup>Ring1 KO</sup>*, they vary for the other two genotypes, until they converge at 24 hours in *dRybp<sup>Eed KO</sup>* and 48 hours in *dRybp<sup>Suz12 KO</sup>*. The most significantly enriched GO categories are related to patterning, embryonic and extra embryonic development, in agreement with Polycomb repression of master developmental regulator

genes [5]. Similarly, enriched functional GO categories of down-regulated genes for the three genotypes converge at 24 hours for *dRnf2*<sup>Ring1 KO</sup> and *dRybp*<sup>Eed KO</sup>, while *dSuz12*<sup>Rybp KO</sup> resembles the other two until 48 hours (Fig. 3.3.4). Common enriched GO categories at 48 hours are mostly related to RNA processing and metabolism, hinting towards down-regulation of housekeeping programs that may explain the observed proliferation defects.



**Figure 3.3.3: GO enrichment analysis of up-regulated genes upon simultaneous perturbation cPRC1 and vPRC1 targeting. (A)** Comparison of Gene Ontology (GO) categories in response to IAA treatment of *dRybp*<sup>Eed KO</sup>, *dSuz12*<sup>Rybp KO</sup> and *dRnf2*<sup>Ring1 KO</sup> at 6h, 24h and 48h. The x-axis shows number of up-regulated genes (padj < 0.1) at each time point used to calculate the GO profiles. The color of the dots indicates the significance of the GO term enrichment, and the size represents the ratio of genes belonging to each category. **(B,C,D)** Cnet plots showing the association of genes involved in the top 5 most significantly enriched GO categories per genotype at 48 hours.



**Figure 3.3.4: GO enrichment analysis of down-regulated genes upon simultaneous perturbation cPRC1 and vPRC1 targeting.** (A) Comparison of Gene Ontology (GO) categories in response to IAA treatment of *dRybp*<sup>Eed KO</sup>, *dSuz12*<sup>Rybp KO</sup> and *dRnf2*<sup>Ring1 KO</sup> at 6h, 24h and 48h. The x-axis shows number of down-regulated genes (padj < 0.1) used to calculate the GO profiles. The color of the dots indicates the significance of the GO term enrichment, and the size represents the ratio of genes belonging to each category. (B,C,D) Cnet plots showing the association of genes involved in the top 5 most significantly enriched GO categories per genotype at 48 hours.

In conclusion, this analysis demonstrates PRC2/cPRC1 and vPRC1 act redundantly to maintain silencing of shared target genes which control cell-type specification and drive loss-of-self renewal when derepressed in mES cells.

### 3.4 Defects on *Rnf2* targeting underlie de-repression of PcG target genes upon double disruption of PRC2 and vPRC1.

Disruption of PRC2 activity impairs cPRC1 targeting (Fig. 3.1.2), and *Rybp* depletion tampers with vPRC1 complex stability [32], and possibly also with its recruitment. Thus, the defects in Polycomb-mediated silencing caused by disruption of PRC2 and vPRC1 result from failure to recruit the catalytic PRC1 activity (*Ring1/Rnf2-Pcgf*) to Shared target gene promoters. To investigate the impact of dual disruption of *Eed* and *Rybp* on PRC1 targeting and catalysis, we used ChIP-seq to map the occupancy of *Rnf2*, and H2AK119ub1 upon IAA treatment.

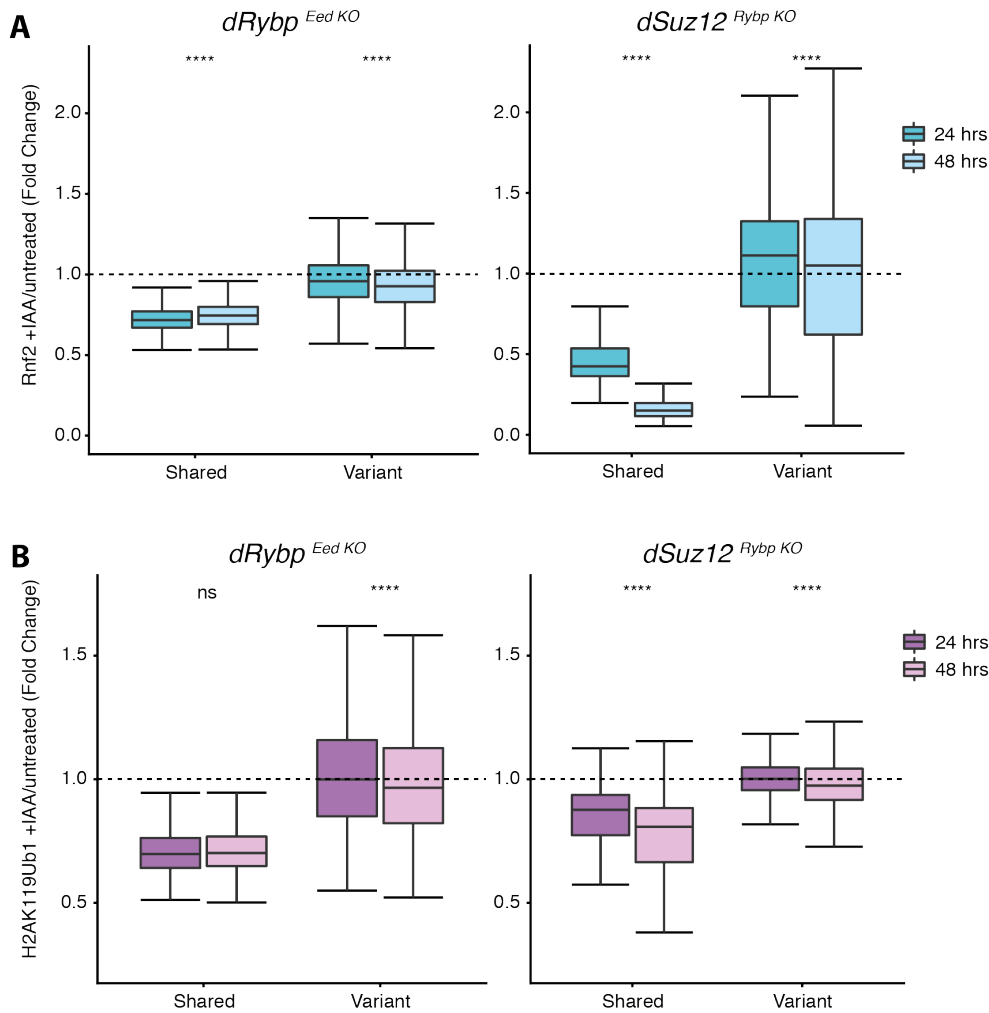
Fig. 3.4.1 A shows heat maps comparing the occupancy of *Rnf2* and H2AK119ub1 at Shared and Variant TSSs in cells untreated or treated with IAA for 24 hours. All three genotypes show substantial decrease in ChIP enrichment for *Rnf2* and H2AK119ub1. As expected, Western blot analysis in *dRnf2*<sup>*Ring1 KO*</sup> confirms efficient Auxin-induced *Rnf2* depletion within 24 hours of IAA treatment and reveals the strongest reduction in H2AK119ub1 (Fig. 3.2.1 D).

To quantify the differences between untreated and IAA treated samples, we calculated the normalized ChIP-seq signal (RPKM) within defined windows centered at the TSS, +/- 5kb for *Rnf2* and +/-10kb for H2AK119ub1 (same regions as displayed on the heat maps). Fig. 3.4.1 depicts box plot quantifications comparing untreated to IAA treated samples for *Rnf2* (B) and H2AK119ub1 (C), showing that the signal reduction is significantly greater at Shared compared to Variant targets in all genotypes.

Because *dSuz12*<sup>*Rybp KO*</sup> cells exhibit a delayed transcriptional response, compared to *dRybp*<sup>*Eed KO*</sup> cells, we suspected that at 48 hours of IAA treatment the enrichment levels of *Rnf2* and H2AK119ub1 would drop even further. To test this idea, we performed ChIP-seq for *Rnf2* and H2AK119ub1 in *dRybp*<sup>*Eed KO*</sup> and *dSuz12*<sup>*Rybp KO*</sup> cells treated with IAA for 48 hours, along with untreated controls.



In order to compare ChIP-seq signal between the two time points, we calculated the fold-changes between untreated and IAA treated samples at each TSS-centered window. Fig. 3.4.2 A shows that in *dRybp<sup>Eed KO</sup>* changes in *Rnf2* occupancy are similar between 24 (blue) and 48 hours (light blue), whereas in *dSuz12<sup>Rybp KO</sup>* *Rnf2* binding is further decreased as compared to the 24-hour time point. Furthermore, while in *dRybp<sup>Eed KO</sup>* both Shared and Variant TSSs exhibit *Rnf2* decrease, in *dSuz12<sup>Rybp KO</sup>* only Shared TSSs show signal reduction, and the Variant TSSs display a slight increase in *Rnf2* signal. As expected, similar changes are observed with H2AK119ub1 (Fig 3.4.2.B). Hence, similar to changes in gene expression, we observe a delay in *Rnf2* targeting defects which is likely to account for the phenotypic differences in mESC proliferation in *dSuz12<sup>Rybp KO</sup>* in response to IAA treatment.

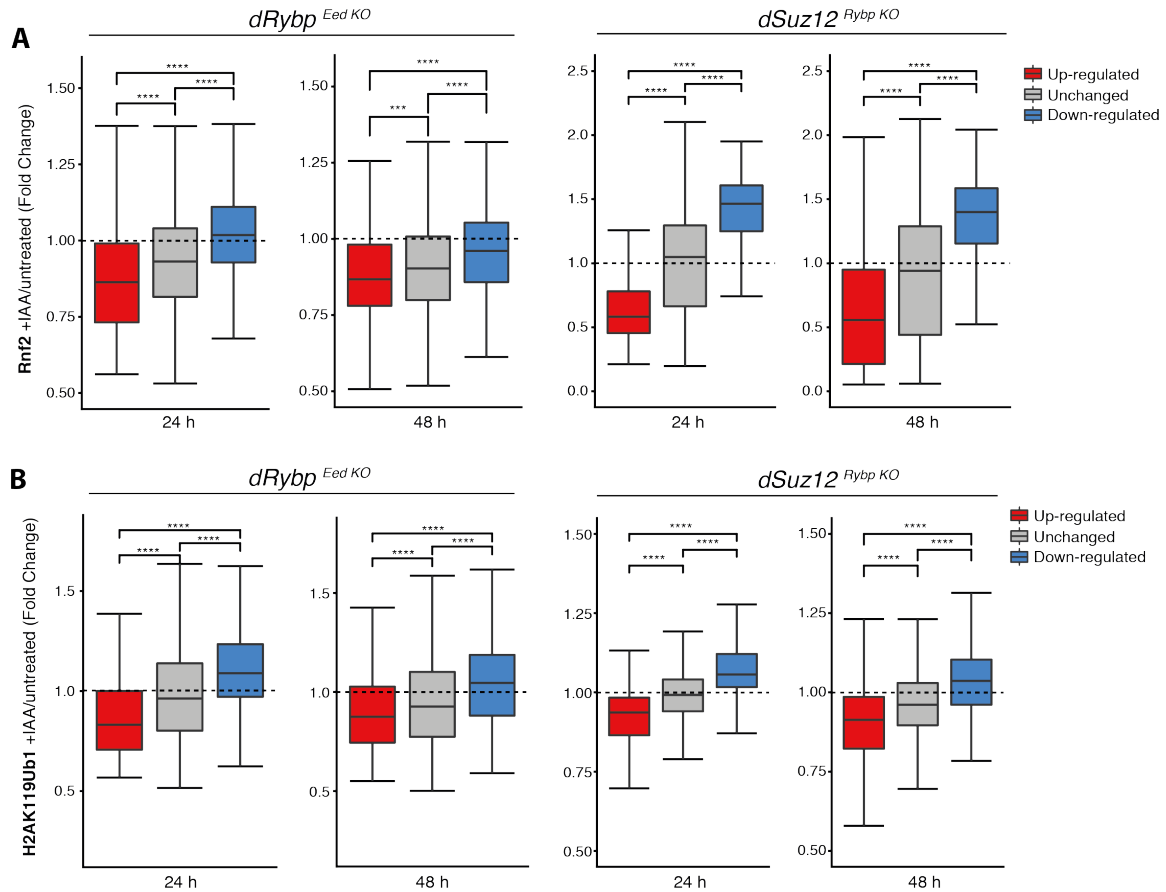


**Figure 3.4.2: *dRybp<sup>Eed KO</sup>* and *dSuz12<sup>Rybp KO</sup>* show differential kinetics in deregulation of *Rnf2* targeting and H2AK119ub1 deposition. (A)** Boxplots show the ratio of *Rnf2* ChIP-seq signal at Shared and Variant TSSs after 24h and 48h IAA treatment in *dRybp<sup>Eed KO</sup>* and *dSuz12<sup>Rybp KO</sup>*



mESCs. **(B)** Boxplots show the ratio of H2AK119ub1 ChIP-seq signal at Shared and Variant TSSs before and after 24h and 48h IAA treatment in *dRybp<sup>Eed KO</sup>* and *dSuz12<sup>Rybp KO</sup>* mESCs.

Next, we sought to address whether increased gene expression generally predicts greater loss of *Rnf2* and H2AK119ub1 signal in response to IAA treatment. Based on differential expression analysis at 24 and 48 hours, Shared and Variant genes were classified into up-regulated, unchanged and down-regulated categories. Fig. 3.4.3 shows that in *dRybp<sup>Eed KO</sup>* and *dSuz12<sup>Rybp KO</sup>*, genes upregulated in response to 24h and 48h IAA treatment exhibit significantly lower levels of *Rnf2* (A) and H2AK119ub1 (B) compared to unchanged and down-regulated genes. Unexpectedly, in *dSuz12<sup>Rybp KO</sup>* the signal for *Rnf2* and H2AK119ub1 increases at down-regulated genes, whereas in *dRybp<sup>Eed KO</sup>* *Rnf2* only slightly increases at 24h, but not at 48 hours and H2AK119ub1 shows a small increase at both time points, suggesting a redistribution of *Rnf2* specific to *dSuz12<sup>Rybp KO</sup>*.



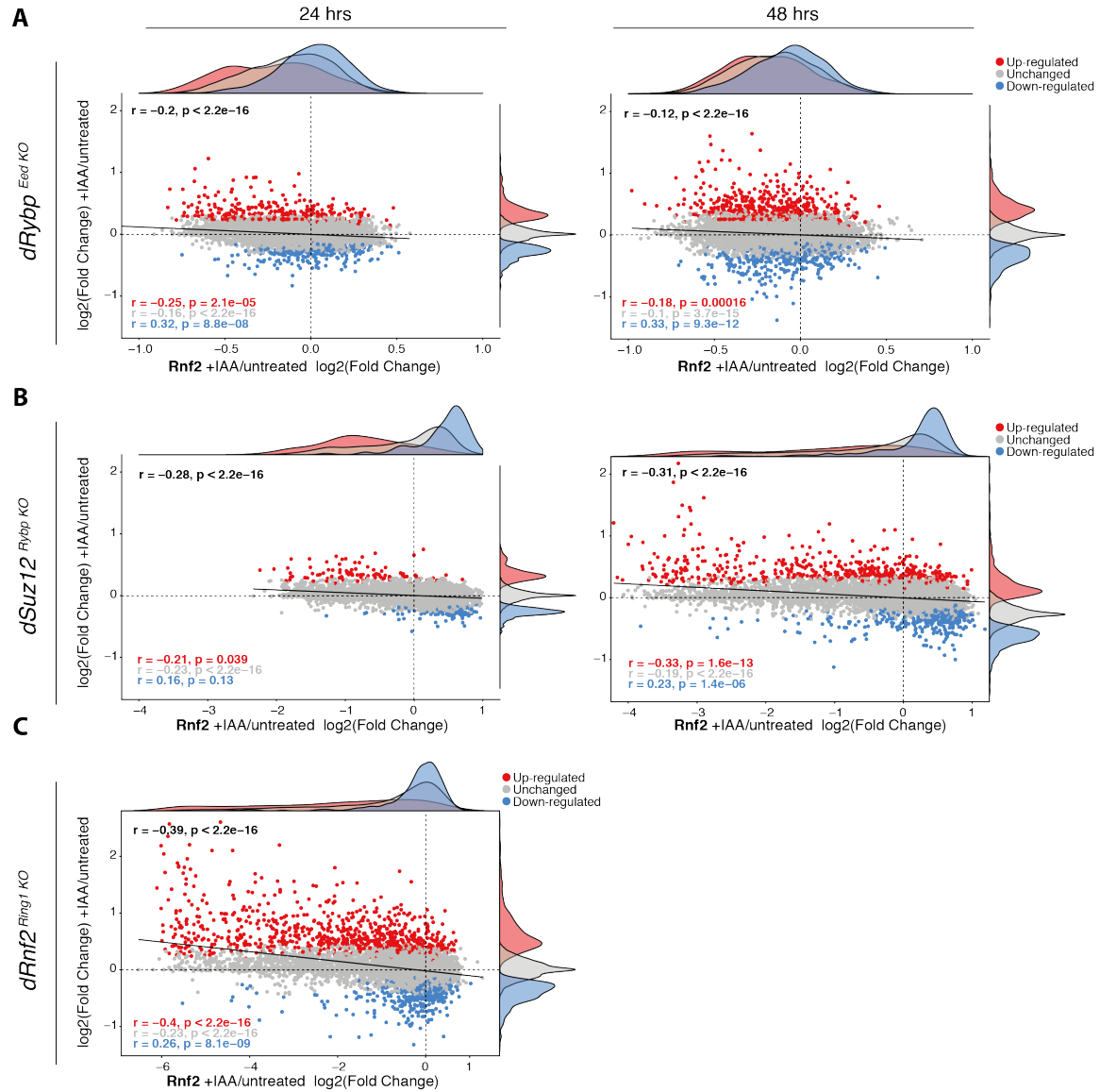
**Figure 3.4.3: Loss of transcriptional silencing double depletion of *Eed* and *Rybp* or *Rybp* and *Suz12* is linked to reduced targeting of *Rnf2* and H2AK119ub1 deposition. (A)** Boxplots show the ratio of *Rnf2* ChIP-seq signal at up-regulated, unchanged and down-regulated genes after 24h and 48h IAA treatment in *dRybp<sup>Eed KO</sup>* and *dSuz12<sup>Rybp KO</sup>* mESCs. Classification of genes according to the differential gene expression data at each time point (padj < 0.1). **(B)** Boxplots show

the ratio of H2AK119ub1 ChIP-seq signal at up-regulated, unchanged and down-regulated genes after 24h and 48h IAA treatment in *dRybp<sup>Eed KO</sup>* and *dSuz12<sup>Rybp KO</sup>* mESCs. Classification of genes according to the differential gene expression data at each time point (padj <0.1). Significance (p-value) calculated using Wilcoxon test (\*\*\*\*:  $p < 2.2 \times 10^{-16}$ ).

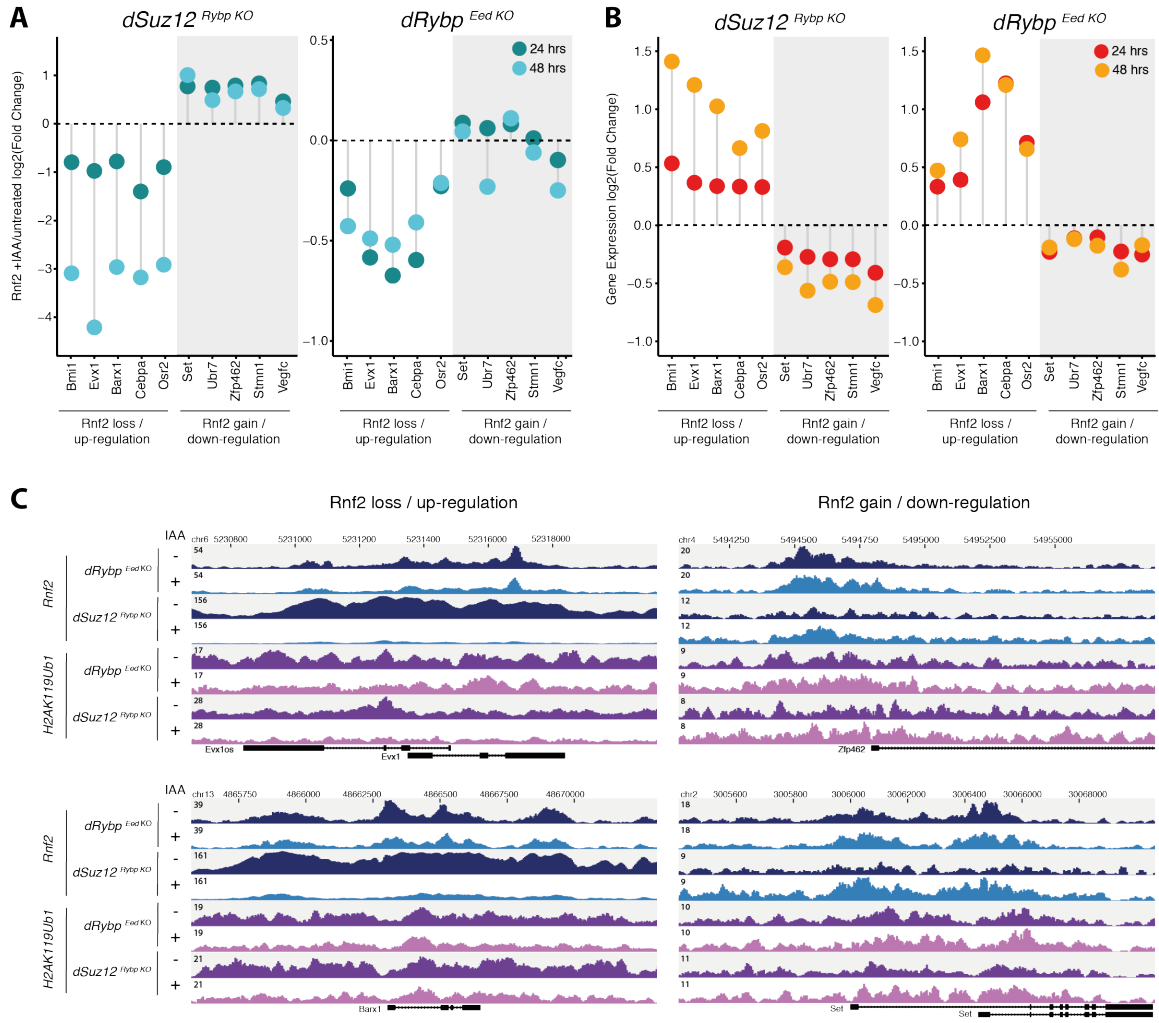
Box plots are a good way of representing the shape of the distribution, the median, and its variability. However, outliers can skew the overall distribution, particularly with few data points, potentially leading to erroneous conclusions. Therefore, in order to investigate whether there is a significant correlation between the change in *Rnf2* occupancy and gene expression at PcG target genes, we decided to directly compare fold change values of *Rnf2* enrichment at the TSS, versus fold change gene expression values of the corresponding gene. Variables were plotted in log2 scale so that positive values indicate gain of signal, and negative ones correspond to loss of signal. The scatter plots in Fig. 3.4.4 show that for all three genotypes there is a small but significant negative correlation (black lines and R values) between gene expression changes and *Rnf2* enrichment changes at all genes. This trend is more pronounced, if we split the correlations into up-regulated (red), unchanged (gray) and down-regulated (blue) sets of genes. Up-regulated genes show the highest positive correlations, followed by the unchanged genes, and down-regulated genes display positive correlations.

Similar to Fig. 3.4.3 A, where in *dSuz12<sup>Rybp KO</sup>* down-regulated genes show an increase in *Rnf2* signal upon IAA treatment, Fig. 3.4.4 B shows that in *dSuz12<sup>Rybp KO</sup>* the majority of down-regulated genes gain *Rnf2*. This trend is not observed at down-regulated genes in IAA-treated *dRybp<sup>Eed KO</sup>* mES cells. This inverse correlation is apparent at five selected genes that exhibit loss of *Rnf2* and transcriptional up-regulation, and five genes which show gain of *Rnf2* and gene down-regulation. Fig. 3.4.5 follows the changes in *Rnf2* occupancy (A) and gene expression (B), of individual genes belonging to each category, at 24 and 48 hours of IAA treatment. Similar to the scatterplots from Fig. 3.4.4, only *dSuz12<sup>Rybp KO</sup>* shows a clear gain of *Rnf2* which is accompanied with transcriptional down-regulation, whereas both *dRybp<sup>Eed KO</sup>* and *dSuz12<sup>Rybp KO</sup>* genotypes exhibit loss of *Rnf2* coupled to transcriptional up-regulation upon IAA treatment. Moreover, the *Rnf2* occupancy changes are also mirrored by changes in H2AK119Ub1 deposition, reflecting not only altered targeting, but also catalytic activity (Fig. 3.4.5 C). Taken together, this data

demonstrates that double disruption of PRC2 and vPRC1 causes de-repression of Shared target genes, and mis-regulation of Variant target genes due to defects on *Rnf2* targeting.



**Figure 3.4.4: Changes in PRC1 occupancy inversely correlate with expression changes.** Scatterplots display pairwise correlation comparing changes in *Rnf2* occupancy with gene expression changes after 24h and 48h of IAA treatment in *dRybp<sup>Eed KO</sup>*, *dSuz12<sup>Rybp KO</sup>* and *dRnf2<sup>Ring1 KO</sup>* mESCs. Solid black lines represent linear regressions based on all data points. Colored legends correspond to Pearson correlation coefficients, and p-values from all (black), up-regulated (red), unchanged (gray) or down-regulated (blue) genes. **(A)** Pairwise correlation of changes in *dRybp<sup>Eed KO</sup>* at 24h and 48h. **(B)** Pairwise correlation of changes in *dSuz12<sup>Rybp KO</sup>* at 24h and 48h. **(C)** Pairwise correlation of changes in *dRnf2<sup>Ring1 KO</sup>* at 24h.

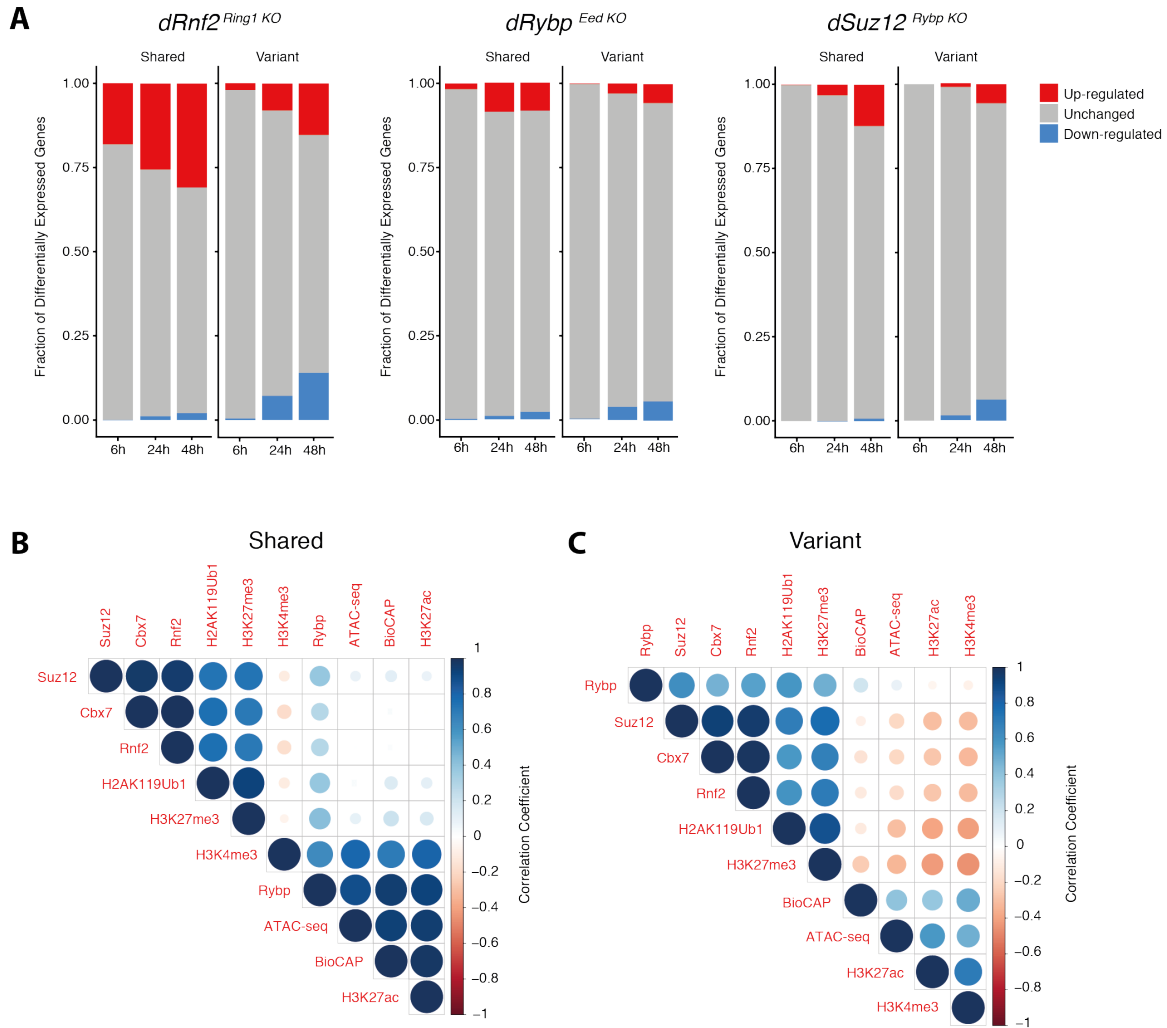


**Figure 3.4.5: Selected genes illustrating the inverse relation between *Rnf2* occupancy and gene expression. (A) Dot charts display changes in *Rnf2* occupancy after 24h and 48h of IAA treatment. The x-axis shows changes in *Rnf2* occupancy at selected genes which are either up-regulated (left) or down-regulated (right). (B) Dot charts display RNA expression changes (Quant-seq) after 24h and 48h of IAA treatment at the same genes as in (A). (C) Screen shots display ChIP-seq enrichment of *Rnf2* and H2AK119ub1 at two upregulated and two down-regulated genes.**

### 3.5 Preferential loss of repression at Polycomb target genes with highest PRC1 and PRC2 occupancy.

Our results show that simultaneous disruption of PRC2 and vPRC1 results in de-repression of only a fraction of shared target genes. Similarly, double *Rnf2* depletion in *Ring1* KO mESCs does not impact silencing all Shared target genes after 48 hours of IAA treatment. Fig. 3.5.1 A displays the fraction of differentially expressed Shared and Variant genes at each time point, and for all three genotypes. In all cases the fraction of Shared genes that respond to IAA treatment is higher than for the Variant ones. Shared genes are, for the most part, up-regulated, while Variant genes show equivalent fractions of genes becoming up and down-regulated. These observations suggest that Shared and Variant genes may be regulated differently, possibly reflecting the different occupancy levels for different PcG subunits (Fig. 3.1.1 B).

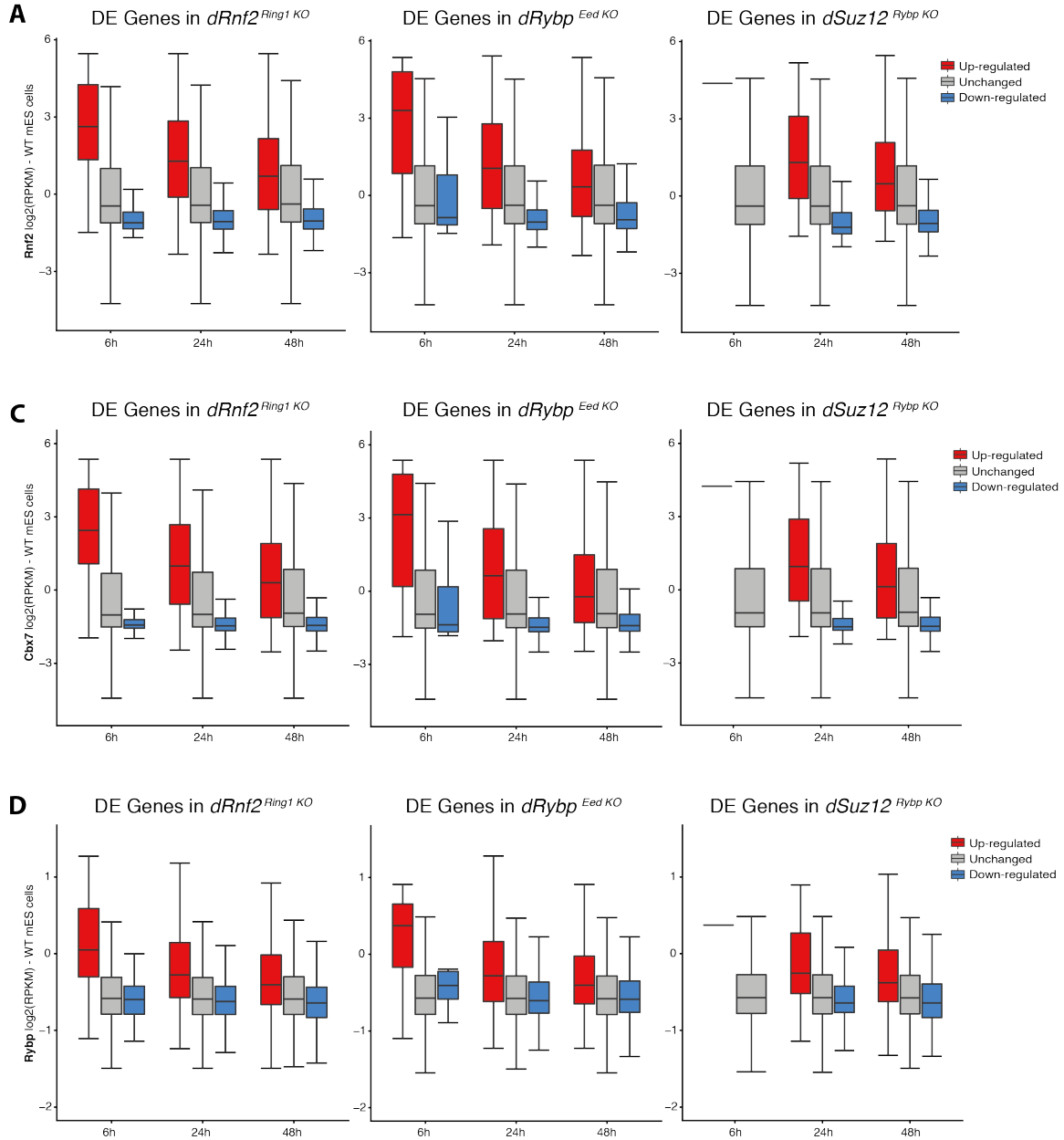
To identify chromatin features predictive of transcriptional responses to PRC2 and vPRC1 double depletion, we calculated Pearson correlations among different subunits and histone modifications in wild-type mESCs representing PRC2 (*Suz12* and H3K27me3), all PRC1 (*Rnf2* and H2AK119Ub1), cPRC1 (*Cbx7*), vPRC1 (*Rybp*), un-methylated CGI DNA (BioCap), Trithorax (TrxG) activity (H3K4me3 and H3K27ac), and chromatin accessibility (ATAC-seq). Fig. 3.1.1 shows correlation matrixes for such chromatin features at Shared (B) or Variant (C) TSSs. In both TSS subsets PcG proteins and histone modifications form one defined cluster, while TrxG histone modifications cluster together with BioCap and ATAC-seq. At Shared TSSs PcG and TrxG features are highly correlated within each cluster, but show barely any significant correlation to one another, while at Variant TSSs PcG and TrxG histone modifications show significant negative correlations. This suggests that at Shared TSSs, gene silencing may be controlled mainly by PcG proteins, whereas at Variant promoters there may be a more complex interplay between PcG and TrxG activities that regulates their gene expression.



**Figure 3.5.1: Acute disruption of cPRC1 and vPRC1 reveals a critical role for Polycomb in silencing a fraction of shared target genes in mouse ES cells. (A)** Bar plots show the fraction of differentially expressed genes at Variant and Shared genes after 6h, 24h and 48h of IAA treatment. **(B)** and **(C)** Pearson correlation matrixes of different chromatin features associated with TSSs of Shared and Variant genes.

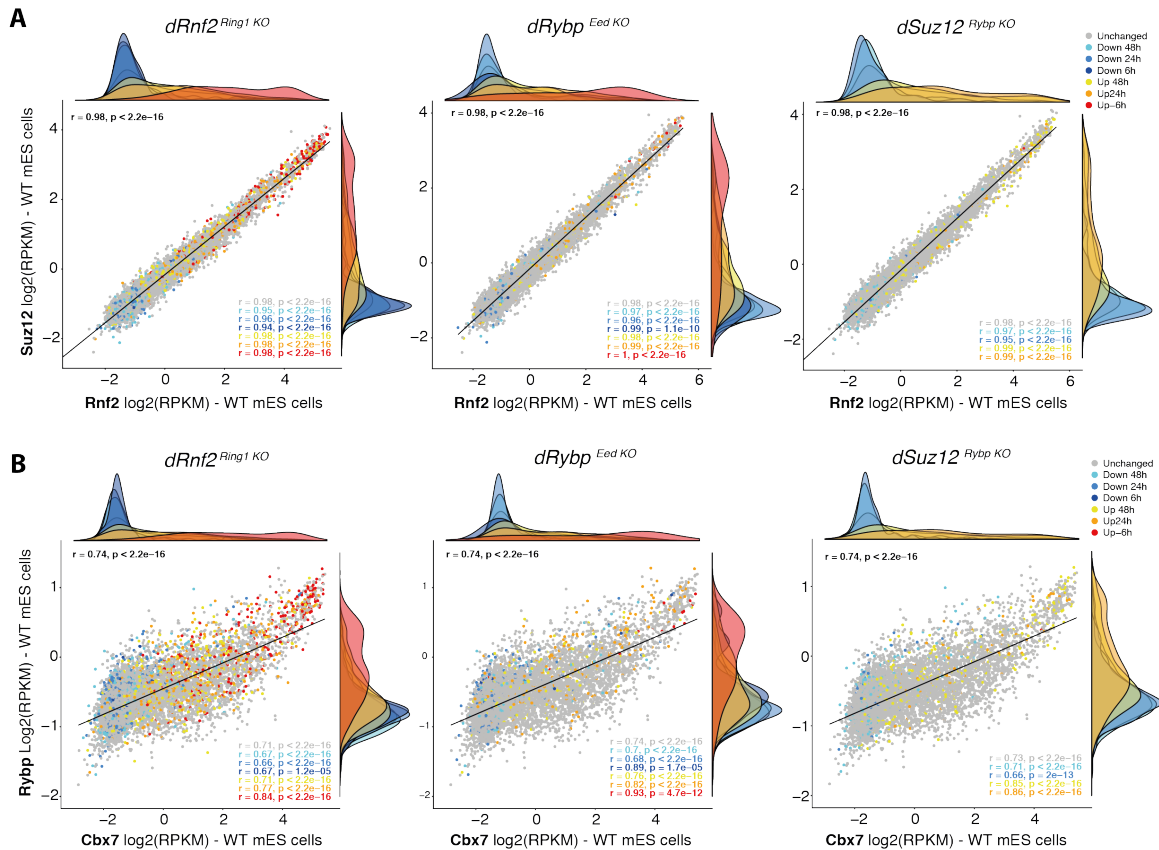
Given the negative correlation between *Rnf2* occupancy and gene expression upon disruption of PRC2 and vPRC1, we tested if PcG protein occupancy predicts target genes that are preferentially upregulation upon IAA treatment. First, we quantified the occupancy of *Rnf2* at the TSSs of Shared and Variant genes in wild-type mES cells. The box plots in Fig. 3.5.2 A show the *Rnf2* enrichment levels grouped by the differential gene expression profiles at 6, 24 and 48 hours in the three genotypes. Interestingly, for all genotypes, and across all time points, up-regulated genes display the highest levels of *Rnf2*, followed by the unchanged genes, and the down-regulated genes having the least. This is agreement with the observations that up-regulated genes show the largest decrease in *Rnf2* upon

IAA treatment in all AID-tagged genotypes (Fig. 3.4.3 A). *Cbx7* occupancy shows a very similar pattern to *Rnf2* (Fig. 3.5.2 B), while *Rybp* occupancy at up-regulated genes also displays the highest enrichment, but the distinction between unchanged and down-regulated is less pronounced (Fig. 3.5.2 C).



**Figure 3.5.2: mESC genes dependent on Polycomb silencing have high occupancy of canonical and variant PRC1 core subunits.** Boxplots compare ChIP-seq enrichment in wild-type mESCs of *Rnf2* (A), *Cbx7* (B) or *Rybp* (C) at upregulated (red), unchanged (gray) and down-regulated (blue) genes after 6h, 24h and 48h of IAA treatment in *dRybp*<sup>Eed KO</sup>, *dSuz12*<sup>Rybp KO</sup> and *dRnf2*<sup>Ring1 KO</sup> mESCs.

To investigate the correspondence between PRC2 and PRC1 occupancy at TSSs of genes that respond to disruption of PRC2/vPRC1 or *Ring1/Rnf2* at different time points, we classified the PcG target genes corresponding to their differential expression status (up-regulated, unchanged or down-regulated) at the earliest time point at which they are called differentially expressed. Fig. 3.5.3 A shows a high, almost perfect, correspondence between *Rnf2* (PRC1) and *Suz12* (PRC2) occupancy at PcG targets in wild type cells. The earliest up-regulated genes show the highest occupancy for *Rnf2* and *Suz12*, in all three genotypes, and the Pearson correlation between *Rnf2* and *Suz12* are highest for earlier up-regulated genes. Fig. 3.5.3 B shows a similar analysis applied to *Cbx7* (cPRC1) and *Rybp* (vPRC1), where the overall correspondence is lower for all genes compared to *Rnf2* and *Suz12*, but it is significantly higher in up-regulated genes compared to unchanged or down-regulated, and within up-regulated genes the correlation is higher for the earlier time points, in all three genotypes. This analysis indicates that in wild type mES cells PRC2, cPRC1 and vPRC1 cooperatively maintain silencing of target genes which they co-occupy at high levels.

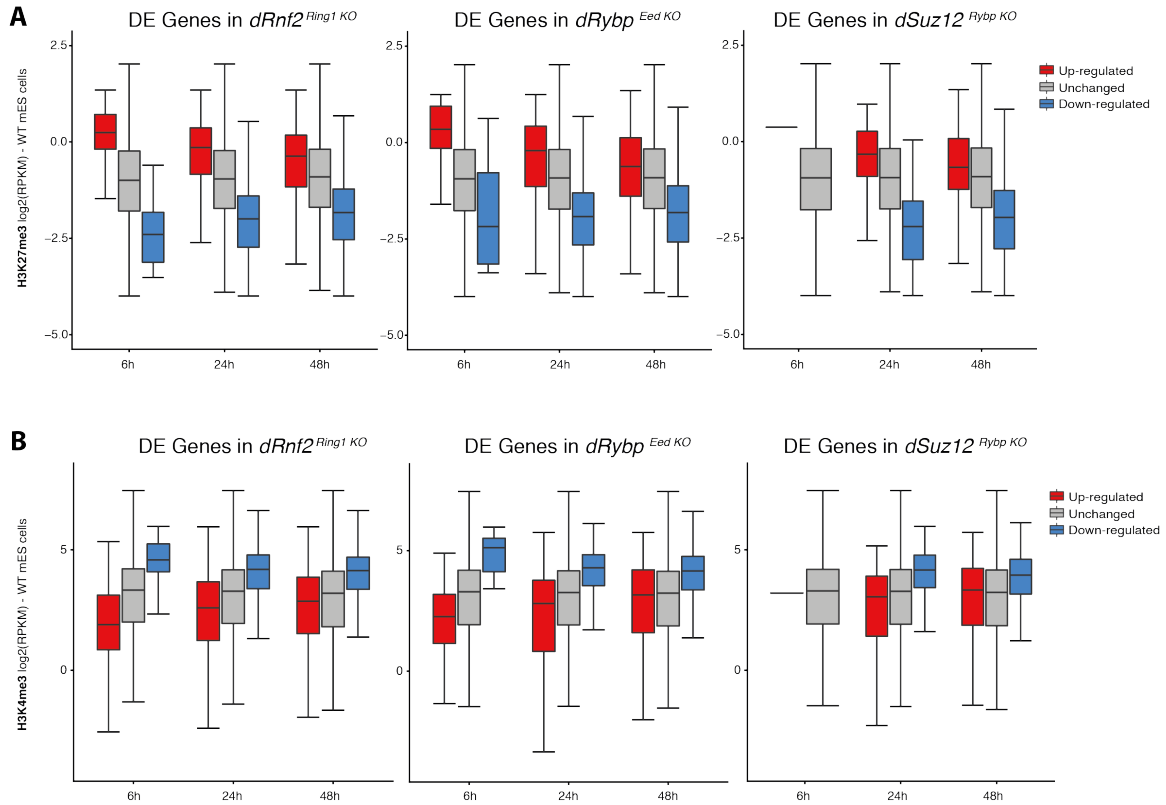


**Figure 3.5.3: mESC genes dependent on Polycomb silencing have highest occupancy of PRC1 and PRC2.** Scatter plots display pairwise comparisons of ChIP-seq signals in wild type



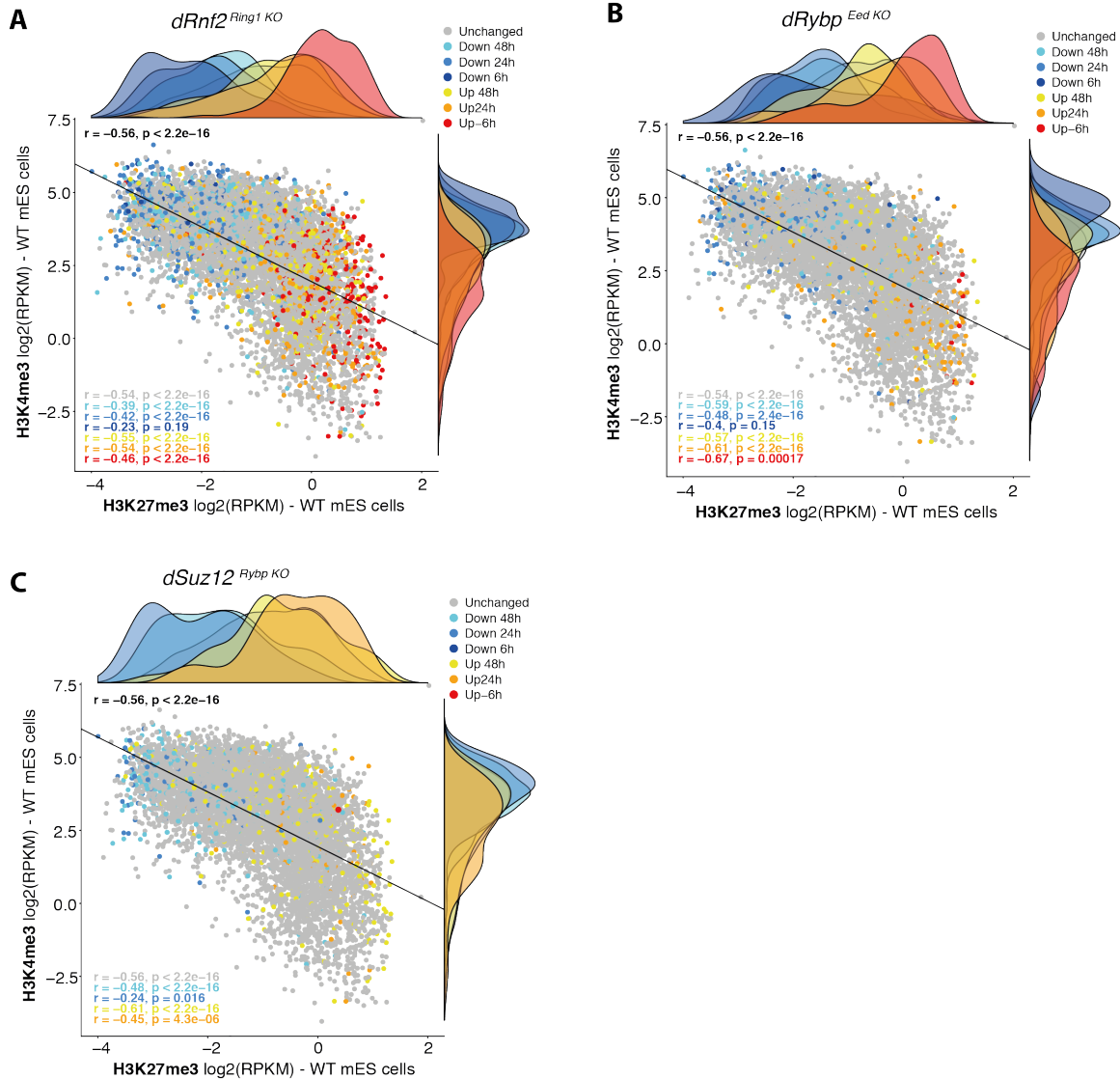
mESCs between *Rnf2* and *Suz12* **(A)** and *Cbx7* with *Rybp* **(B)**. Coloring scheme corresponds to genes up-regulated, down-regulated or unchanged after 6h, 24h or 48h of IAA treatment. Solid black lines represent linear regressions based on all data points. Pearson correlation coefficients and p-values calculated for all data points (black), or for each category of genes (in colors). Density plots above and on the right visualize the distribution of different gene categories.

In mES cells, several promoters show a bivalent chromatin signature being decorated with repressive H3K27me3 and activating H3K4me3 [88], [99]. These “bivalent” promoters are thought to be in a poised state, and thus can be rapidly activated or repressed upon developmental stimuli [100], [101]. We speculated that bivalent genes could be more prone for de-repression upon disruption of *Ring1/Rnf2* or PRC2/vPRC1, as TrxG complexes would face no counteraction by PcG activity. To examine this possibility, we quantified the ChIP-seq signal for H3K27me3 and H3K4me3 at the TSSs of PcG target genes in wild type mESCs. Fig. 3.5.4 A shows that H3K27me3 enrichment is the highest at genes that become up-regulated upon disruption of PRC2 and vPRC1, unchanged genes display medium levels of H3K27me3, and up-regulated genes present the lowest levels. In contrast, H3K4me3 shows the opposite pattern (Fig. 3.5.4 B) where up-regulated genes have the lowest enrichment, followed by unchanged genes and downregulated genes display the highest levels.



**Figure 3.5.4: mESC promoters dependent on Polycomb silencing have high levels of H3K27me3 and low levels H3K4me3.** Boxplots compare ChIP-seq enrichment in wild-type mESCs of H3K27me3 (A) and H3K4me3 (B) at upregulated (red), unchanged (grey) and down-regulated (blue) genes after 6h, 24h and 48h of IAA treatment in *dRybp*<sup>Eed KO</sup>, *dSuz12*<sup>Rybp KO</sup> and *dRnf2*<sup>Ring1 KO</sup> mESCs.

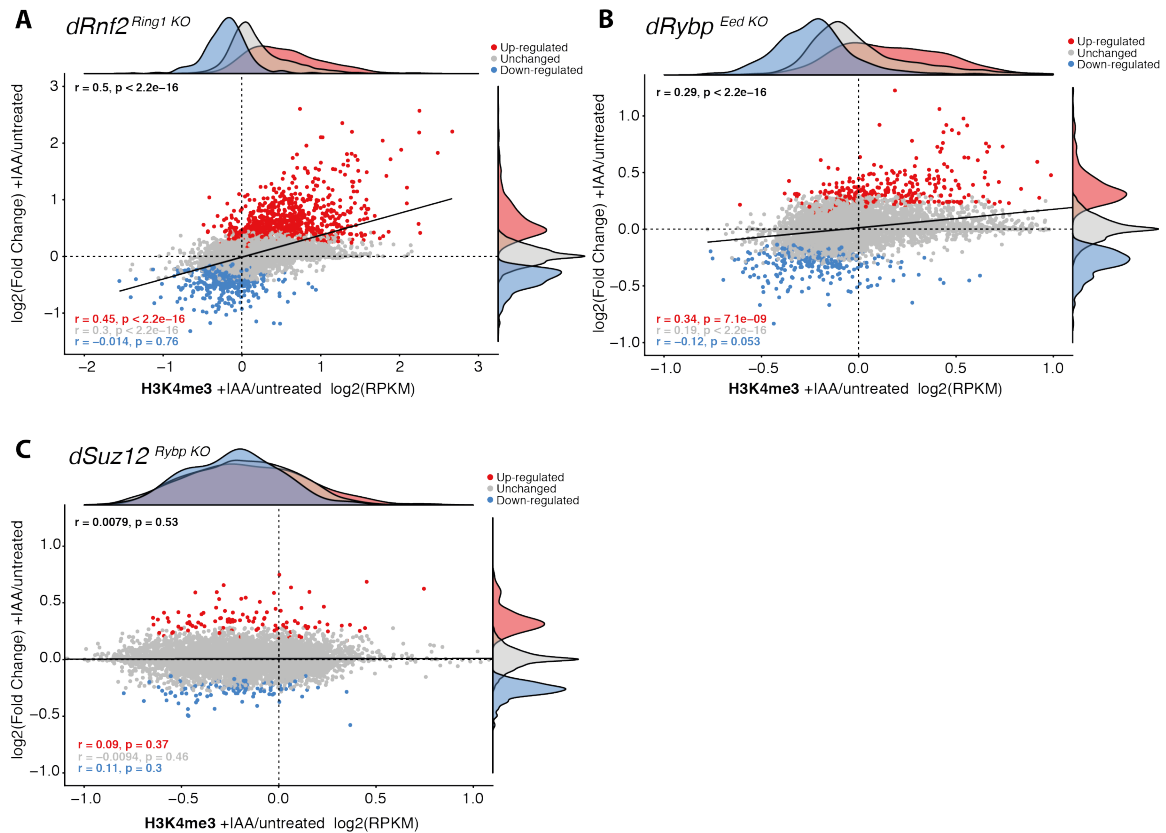
Next, we looked at the correspondence between H3K27me3 and H3K4me3 in wild-type mESCs, to investigate whether H3K4me3 enrichment would be predictive of transcriptional up-regulation at genes decorated with H3K27me3. The scatterplots in Fig. 3.5.5 A-C show that even though there are genes with high H3K27me3 and H4K4me3 (bivalent), they are not preferably up-regulated at early time points. Rather, up-regulated genes tend to have high H3K27me3 and low H3K4me3, whereas down-regulated genes are characterized for low H3K27me3 and high H3K4me3. Furthermore, the Pearson correlation between H3K27me3 and H3K4me4 is negative (-0.56) for all PcG target TSSs, yet it becomes even more negative for up-regulated genes, compared to down-regulated ones, indicating that H3K4me3 does not predispose PcG target genes to become depressed upon disruption of *Ring1/Rnf2* or PRC2/vPRC1.



**Figure 3.5.5: mESC promoters dependent on Polycomb silencing have high levels of H3K27me3 and low levels H3K4me3.** Scatter plots display pairwise comparisons of ChIP-seq signals in wild type mESCs between H3K27me3 and H3K4me3 (**A,B,C**), and between H3K27me3 and H3K27ac (**D**). Coloring scheme corresponds to genes up-regulated, down-regulated or unchanged at 6h, 24h or 48h of IAA treatment. Solid black lines represent linear regressions from all data points. Pearson correlation coefficients and p-values calculated for all data points (black), or for each category of genes (in colors). Density plots above and on the right visualize the distribution of different gene categories.

TrxG activity, as measured by H3K4me3, does not poise Polycomb target genes for transcriptional activation upon removal of PcG activity, but it may increase as a consequence of de-repression caused by disruption of PcG activity. To investigate this possibility, we performed ChIP-seq for H3K4me3 on the AID-tagged cell lines untreated or treated with IAA for 24 hours. Fig. 3.5.6 A-C shows pairwise comparisons between gene

expression changes and H3K4me3 occupancy changes at PcG target genes, in each of the AID-tagged cell lines. Gene expression and H3K4me3 display positive correlations in *dRnf2*<sup>Ring1 KO</sup> and *dRybp*<sup>Eed KO</sup> (black lines and r values), where up-regulated genes gain H3K4me3 and down-regulated genes lose it. Furthermore, up-regulated genes display the highest positive correlations, whereas down-regulated genes show slight negative correlations. However, in *dSuz12*<sup>Rybp KO</sup> there is no significant correlation between gene expression and H3K4me3 changes, possibly because the transcriptional response is delayed compared to the other two genotypes. Collectively, this data demonstrates that bivalency does not facilitate gene activation upon PcG failure, and that H3K4me3 deposition increase as a consequence of transcriptional up-regulation that results from disruption of PcG activity.



**Figure 3.5.6: H3K4me3 is increased at up-regulated genes following disruption of cPRC1 and vPRC1.** Scatter plots display pairwise comparisons of changes in H3K4me3 (A,B,C) or H2AK119ub1 (D) relative to changes in gene expression after 24h of IAA treatment in *dRybp*<sup>Eed KO</sup>, *dSuz12*<sup>Rybp KO</sup> and *dRnf2*<sup>Ring1 KO</sup> mESCs. Solid black lines represent linear regressions based on all data points, colored legends correspond to Pearson correlation coefficients and p-values based on all (black), up-regulated (red), unchanged (gray) or down-regulated (blue) genes after 24h of IAA treatment. Density plots above and on the right visualize the distribution of different gene categories.

# Chapter 4

## Discussion

Loss-of-function (LOF) mutations in genes encoding PRC1 and PRC2 core components often result in embryonic lethality or in developmental defects [5]. In insects and vertebrates, repressive Polycomb-dependent chromatin modifications are thought to stabilize cell fate decisions by preventing inappropriate expression of cell-type specific master regulatory genes. Consistent with its essential functions in differentiation, mouse embryonic stem cells (mESCs) with LOF mutations in PRC2 core subunits (*Eed* and *Suz12*) fail to differentiate *in vitro* as expression of pluripotency genes is not efficiently silenced [82], [102]. In contrast, it is debated whether Polycomb regulation plays a role in ES cell self-renewal. Genetic deletion of most PRC1 and PRC2 subunits yields only minor changes in mES cell gene expression [103]. However, loss of both paralogous PRC1 catalytic subunits *Ring1* and *Rnf2* results in de-repression of Polycomb target genes and failure to maintain mESC self-renewal [81]. Thus, it is unclear if PRC1 and PRC2 repressive functions are restricted to cell differentiation or if they are also required for target gene repression to ensure maintenance of ESC identity.

In order to address the role of Polycomb in maintenance of self-renewal of mESCs, we have utilized a chemical genetic protein degradation system, to conditionally control PRC2 or variant PRC1 function in PcG mutant mESCs. Genome-wide analysis of PcG protein occupancy and target gene expression revealed that PRC2 is essential for canonical PRC1 recruitment to target TSSs. In the absence of canonical PRC1, variant PRC1 persists mediating repressive activity at the Shared set of target genes, independently of PRC2. In contrast, simultaneous disruption of PRC2 and vPRC1 function disrupts redundant targeting of PRC1 catalytic core subunit *Rnf2* to chromatin and triggers loss of silencing and mES cell self-renewal. Our results suggest that in mESCs, PRC2/cPRC1 and vPRC1 cooperate to target *Ring1/Rnf2* activity to Shared target genes and thus ensure robust maintenance of transcriptional repression of master developmental regulators, thus safeguarding ES cell identity.

### 3.1 *Eed* and *Rybp* are dispensable for silencing of shared target genes in mouse embryonic stem (mES) cells.

In agreement with previous reports that recruitment of *Cbx7* depends on H3K27me3 [78] and that *Rnf2* binding is reduced in PRC2-deficient mES cells [47], inactivation of PRC2 complex by *Eed* KO lead to complete dislodgement of *Cbx7* and dramatic decrease in *Rnf2* binding at all PcG target TSSs (Fig. 3.1.2 A). Yet, at the transcriptional level loss of *Eed* induces only small changes to vPRC1-specific genes (Variant), while Shared target genes co-occupied by PRC2, cPRC1 and vPRC1 remain virtually unaffected (Fig. 3.1.5 A). Based on the observation that *Rybp* recruitment and deposition of H2AK119ub are not affected by *Eed* KO (Fig. 3.1.2 B), we hypothesized that vPRC1 complexes may compensate for the loss of canonical PRC1 to maintain silencing of Shared target genes.

Variant PRC1 constitute a diverse family of complexes with different accessory subunits. Thus, genetic analysis of variant function poses a significant challenge. In addition to the catalytic core composed of *Ring1/Rnf2* and one of four variant-specific *Pcgf* paralogs, *Rybp* is present in all vPRC1 complexes [16], [47], [80]. It has been shown to stabilize assembly of the complex and stimulate its ubiquitylation activity *in vitro* [32]. Thus, *Rybp* KO creates an opportunity to study compromised vPRC1 complexes.

We find that *Rybp* KO leads to reduction of H2AK119ub1 at Shared and Variant target TSSs (Fig. 3.1.3). Interesting, our experiments reveal an increase in *Rnf2*, *Eed* and *Cbx7* binding in response to *Rybp* LOF. While mostly Variant targets are de-regulated, *Rybp* KO does not lead to de-repression of Shared target genes, similar to *Eed* KO and consistent with potential functional compensation between canonical and variant PRC1 (Fig. 3.1.5 A). To explore this functional redundancy, we sought to generate conditional double mutant mES cells to ablate both cPRC1 and vPRC1 in parallel.

### 3.2 Double disruption of PRC2 and variant PRC1 leads to loss of self-renewal, phenocopying double depletion of *Ring1* and *Rnf2*.

Genetic deletion of most PcG genes does not impair self-renewal or the expression of pluripotency genes in mES cells (see 1.10). In fact, only double deletion of *Ring1* and *Rnf2* (PRC1 catalytic core) leads to strong de-repression of PcG target genes and loss of self-renewal [81]. Thus, having conditional *Rnf2* mES cell lines in the background of *Ring1* KO, provides a paradigm to study complete inactivation of all PRC1 complexes.

The AID (Auxin inducible degron) system (Fig. 3.2.1 A and C) offered a powerful approach to study the consequences of synchronous and acute PcG protein degradation on the regulation of chromatin modifications and gene expression in mES cells. Unlike single knockout mES cells, induced double depletion of *Eed* and *Rybp* results into loss of self-renewal, similar to double depletion of *Ring1* and *Rnf2* (Fig. 3.2.3). This result supports the idea that PRC2/cPRC1 and vPRC1 play partially overlapping roles to silence master regulators of differentiation, and thus maintain pluripotency state. Furthermore, it excludes that the loss of self-renewal in *Ring1/Rnf2* dKOs reflects a PRC1-independent role of the catalytic subunits.

Since degradation of *Rybp* on *Eed* KO cells impairs self-renewal, it would be expected to observe a similar phenotype upon degradation of *Suz12* on *Rybp* KO cells. However, it is not entirely the case as only one clone of the two *dSuz12<sup>Rybp KO</sup>* cells displayed proliferation defects (Fig. 3.2.2), and both clones showed less severe changes in mES colony morphology and Alkaline Phosphatase staining compared to *dRybp<sup>Eed KO</sup>* cells (Fig. 3.2.3 A and B). Nevertheless, *Rybp* and *Suz12* stable double knockout mES cell lines could not be derived, suggesting that *Rybp/Suz12* double depletion does impair self-renewal. Given that *Rybp* KO mES cells differ from wild-type and *Eed* KO mES cells displaying upregulation of germ-cell specific marker genes (Fig. 3.1.5 C), we speculate that *Rybp* KO mES cells trigger additional mechanisms to suppress activation of shared differentiation genes and that the order of PRC2 and vPRC1 disruption is relevant for loss of ES cell identity.

Alternatively, the discrepancy between *Eed/Rybp* and *Rybp/Suz12* double depletion phenotypes could be explained by *Eed* having additional functions outside of PRC2, as *Eed* has been reported to co-purify with PRC1 complexes [97]. mES cell lines with conditional *Rybp*-AID in the background of *Suz12* single KO or *Eed/Suz12* double KO showed very similar loss of self-renewal phenotypes, which are comparable to *Eed/Rybp* double depletion. This suggests that PRC2 inactivation by knockout of *Eed*, *Suz12* or both are equivalent in terms of being synthetic lethal with *Rybp*, and supports the idea that the ameliorated loss of self-renewal observed in *Rybp/Suz12* double depletion may be explained by the different transcriptional states of *Rybp* KO mES cell at the time of *Suz12* degradation. Alternatively, *Rybp* KO cells may be able to adapt to *Suz12* degradation more effectively than *Eed/Rybp* depleted cells, as *dSuz12<sup>Rybp KO</sup>* cells exhibit delayed defects in *Rnf2* targeting (Fig. 3.4.2 A) and de-repression of Shared PcG targets (Fig. 3.3.2 A) compared to *dRybp<sup>Eed KO</sup>*.

### 3.3 PRC2/cPRC1 and vPRC1 function redundantly to maintain repression of shared target genes.

Parallel impairment of PRC2 and vPRC1 causes a gradual mis-regulation of several genes, reaching almost 2000 genes after 48 hours of IAA treatment, which is more or less the time that it takes unperturbed cells to undergo three cell divisions. Despite the phenotypic differences observed between *dRybp<sup>Eed KO</sup>* and *dSuz12<sup>Rybp KO</sup>*, they share around 30% of up and down-regulated genes, in addition ~80% of *dRybp<sup>Eed KO</sup>* and ~60% of *dSuz12<sup>Rybp KO</sup>* mis-regulated genes overlap with those in *dRnf2<sup>Ring1 KO</sup>* after 48 hours of IAA treatment (Fig. 3.3.1 B). Thus, we conclude that double disruption of PRC2 and vPRC1 induces similar gene expression changes to *Ring1* and *Rnf2* double depletion, likely reflecting ablation of PRC1 activity.

In *dRybp<sup>Eed KO</sup>* and *dSuz12<sup>Rybp KO</sup>* Shared target genes become de-repressed upon IAA treatment (Fig. 3.3.2 A), in contrast to *Eed* and *Rybp* single KOs that display minor changes at Shared genes, which suggests that PcG complexes act redundantly to silence Shared genes. On the other hand, either single KOs or *dRybp<sup>Eed KO</sup>* and *dSuz12<sup>Rybp KO</sup>* treated with IAA display up and down-regulation at Variant genes (Fig 3.1.5 A and 3.3.2



A). This suggests that Variant genes are more susceptible to impairment of Polycomb complexes, possibly because they are bound primarily by only vPRC1 subunits, and that PcG depletion may allow recruitment of counteracting activating chromatin regulators such as MLL2-containing or other Trithorax (TrxG) complexes to these loci.

Key pluripotency marker genes *Sox2*, *Zfp42* (*Rex1*), *Klf4* and *Pou5f1* (*Oct4*) showed gradual down-regulation in *dRybp*<sup>Eed KO</sup> and *dRnf2*<sup>Ring1 KO</sup> upon IAA treatment, matching the loss of self-renewal observed in Fig. 3.2.2 and Fig. 3.2.3 A, while *dSuz12*<sup>Rybp KO</sup> only showed slight down-regulation of *Pou5f1* and *Esrrb* consistent with its less severe self-renewal phenotype (Fig. 3.2.2 and Fig. 3.2.3 A). In contrast, Shared PcG targets are progressively up-regulated in all three genotypes, although with slower kinetics in *dSuz12*<sup>Rybp KO</sup> (Fig 3.3.2).

Gene Ontology (GO) enrichment analysis revealed that mis-regulated genes in all three genotypes share enriched biological processes at 48 hours of IAA treatment, which is the time point at which there are more differentially expressed genes. Up-regulated genes at 48 hours display a good overlap of enriched GO categories related to embryonic and extra embryonic development, patterning and regionalization, consistent with the known role of PcG proteins to control master developmental regulators [75]. On the other hand, down-regulated genes, show more diverse enrichment of GO categories, however at 48 hours all three genotypes display an overlap of housekeeping categories related to RNA metabolism, ribosome and ribonucleoprotein biogenesis. This analysis reveals a common transcriptional signature that comes as a consequence of double disruption of PRC2 and vPRC1, which is analogous to inactivation of all PRC1 activity.

From this data, we cannot pinpoint the exact cause of self-renewal impairment that results from disrupting PcG activity in mESCs. However, the *dRnf2*<sup>Ring1 KO</sup> *dRybp*<sup>Eed KO</sup> and *dSuz12*<sup>Rybp KO</sup> cell lines, developed in this study, constitute great tools for disentangling the molecular events that connect up-regulation of developmental regulators, downregulation of pluripotency genes and down-regulation of housekeeping genes, that ultimately lead to loss of self-renewal in mESCs. For instance, nascent RNA profiling at earlier time points should help discriminate primary from secondary effects with high

temporal resolution, and thus draw causality links between the earliest waves of mis-regulated genes to later gene expression states. Furthermore, these tools make it possible to define not only the requirements of PcG proteins to ES cell identity, but also the mechanisms by which they participate in maintaining cellular homeostasis at the ES cell state, and direct precise transitioning during developmental progression.

### **3.4. Defects of *Rnf2* targeting underlie de-repression of PcG target genes upon double disruption of PRC2 and vPRC1.**

Since double depletion of *Eed/Rybp* and *Rybp/Suz12* result in de-repression of PcG target genes, similar to *Ring1/Rnf2* double depletion, we speculated that the underlying cause would be failure to deliver the catalytic core of PRC1 (*Ring1/Rnf2*) to target TSSs. As expected, analysis of ChIP-seq data revealed that *Rnf2* occupancy is significantly decreased at PcG targets upon parallel disruption of PRC2 and vPRC1 (Fig. 3.4.1 A and B).

Low levels of *Rnf2* in *Eed* KO cells are sufficient to mediate deposition of H2AK119Ub1 to levels comparable with wild-type cells (Fig. 3.1.2 B). Yet in *dRybp<sup>Eed KO</sup>*, degradation of *Rybp* induces further reduction of *Rnf2* and H2AK119ub1, concomitant with the observed up-regulation of target genes. *dRnf2<sup>Ring1 KO</sup>* shows complete loss of *Rnf2* signal in response to IAA treatment, as Western blot analysis shows full degradation of AID-*Rnf2* protein after 24 hours of IAA treatment (Fig. 3.2.1 D). The decrease in H2AK119ub1 mirrors that of *Rnf2* in all three cell lines, indicating that not only targeting is affected, but also catalytic activity.

*Rnf2* and H2AK119ub1 occupancy is reduced gradually in *dSuz12<sup>Rybp KO</sup>*, as the 48-hour time point displays further reduction compared to 24 hours. This is concordant with PcG target de-repression, which is stronger at 48 than at 24 hours. Moreover, the reduction of *Rnf2* and H2AK119ub1 signal in *dRybp<sup>Eed KO</sup>* at 24 hours of IAA treatment is very comparable to the 48-hour time point, nevertheless PcG targets show further de-repression at 48 than at 24 hours. This implies that defects in *Rnf2* targeting precede gene de-repression.

*Rnf2* occupancy at target TSSs and gene expression are inversely correlated. As shown in Fig. 3.4.3 the bulk of up-regulated genes, at either 24 or 48 hours of IAA treatment, show the strongest reduction in *Rnf2* and H2AK119ub1, followed by unchanged, and down-regulated genes. Moreover, fold changes in gene expression negatively correlate with fold changes of *Rnf2* occupancy. This is true for all three AID tagged cell lines at 24 and 48 hours of IAA treatment (Fig 3.4.4).

Unexpectedly, down-regulated genes in *dSuz12<sup>Rybp KO</sup>* displays gain of *Rnf2* and H2AK119ub1 upon IAA treatment (Fig. 3.4.3). This observation is not an artifact of outliers, as Fig. 3.4.4 B shows that most down-regulated genes shift towards gain of *Rnf2* signal. Closer examination of a few candidate genes confirms the gain of *Rnf2* and transcriptional down-regulation (Fig. 3.4.5 A) at 24 and 48 hours of IAA treatment (Fig. 3.4.5 B). Furthermore, increased levels of H2AK119ub1 mirror the gain of *Rnf2* around the TSS of candidate genes in *dSuz12<sup>Rybp KO</sup>* but not in *dRybp<sup>Eed KO</sup>* (Fig. 3.4.5 B). This increase in *Rnf2* occupancy at down-regulated genes in *dSuz12<sup>Rybp KO</sup>*, although small, supports the notion that redistribution of *Rnf2* can change the transcriptional status of target genes. In light of these results, we conclude that defects on *Rnf2* targeting are causal for the de-repression of PcG target genes observed upon double disruption of PRC2 and vPRC1.

### **3.5 In mES cells, PcG proteins prevent transcriptional activation at target genes with the highest buildups of PRC2 and PRC1 complexes.**

Given that *Rnf2* dislodgment from TSSs leads to de-repression of target genes, it would be reasonable to expect the majority of target genes become up-regulated upon disruption of PRC2/vPRC1 or *Ring1/Rnf2* depletion. However, only a fraction of Shared target genes are up-regulated after 48 hours of *Ring1/Rnf2* co-depletion (~30%) or PRC2/vPRC1 disruption (~10%). In contrast, Variant targets show up and down-regulation in similar numbers *Ring1/Rnf2* (~15%) and PRC2/vPRC1 (~5%) (Fig. 3.5.1 A).

ChIP-seq analysis in wild-type mES cells revealed that the subset of de-repressed target genes are characterized by higher levels Polycomb proteins and repressive histone modifications (Fig. 3.5.2). Consistent with this trend, our time-course analysis of gene

expression changes shows that genes de-repressed within six hours have higher occupancy of PcG proteins compared to 24 hours and 48 hours. Furthermore, pairwise comparisons between *Rnf2* and *Suz12*, as well as *Cbx7* and *Rybp*, show the strongest positive correlation at up-regulated genes compared to unchanged and down-regulated (Fig. 3.5.3).

Taking these results into account, we conclude that targeting of PRC1 catalytic core (*Ring1/Rnf2*) depends on the combined activities of PRC2/cPRC1 and vPRC1. Given that vPRC1 interacts directly with chromatin, it may initiate recruitment and transcriptional repression at target genes, where PRC2-dependent H3K27me3 can signal for cPRC1 recruitment to create a self-reinforcing system that is robust to environmental and genetic perturbations.

In mES cells, several CGI TSSs harbor PcG-dependent H3K27me3 and TrxG-dependent H3K4me3 marks (bivalent domains), and it is thought that bivalency allows for a rapid transition from silencing to activation of target genes upon developmental stimuli [31]. We wondered whether the presence of activating H3K4me4 at PcG target genes would prime those bivalent genes for activation upon depletion of *Ring1/Rnf2* or disruption of PRC2/vPRC1.

Wild-type occupancy of H3K27me3 is the highest at TSSs of genes that become de-repressed upon disruption of PRC2/vPRC1 or *Ring1/Rnf2* depletion, whereas H3K4me3 is depleted at those TSSs, compared to unchanged and down-regulated genes (Fig. 3.5.4). Pairwise comparisons of H3K27me3 and H3K4me3 show overall negative correlations at PcG targets, which are stronger for up-regulated compared to down-regulated genes (Fig. 3.5.5). In summary, up-regulated genes are characterized by high H3K27me3 and low H3K4me3, while down-regulated genes harbor medium-to-high H3K4me3 and low H3K27me3 (Fig 3.5.5).

Thus, bivalency does not facilitate gene activation upon ablation of PcG complexes. Nevertheless, H3K4me3 deposition increases at genes that become up-regulated, and decreases at down-regulated genes, after 24 hours of *Ring1/Rnf2* or *Eed/Rybp* double

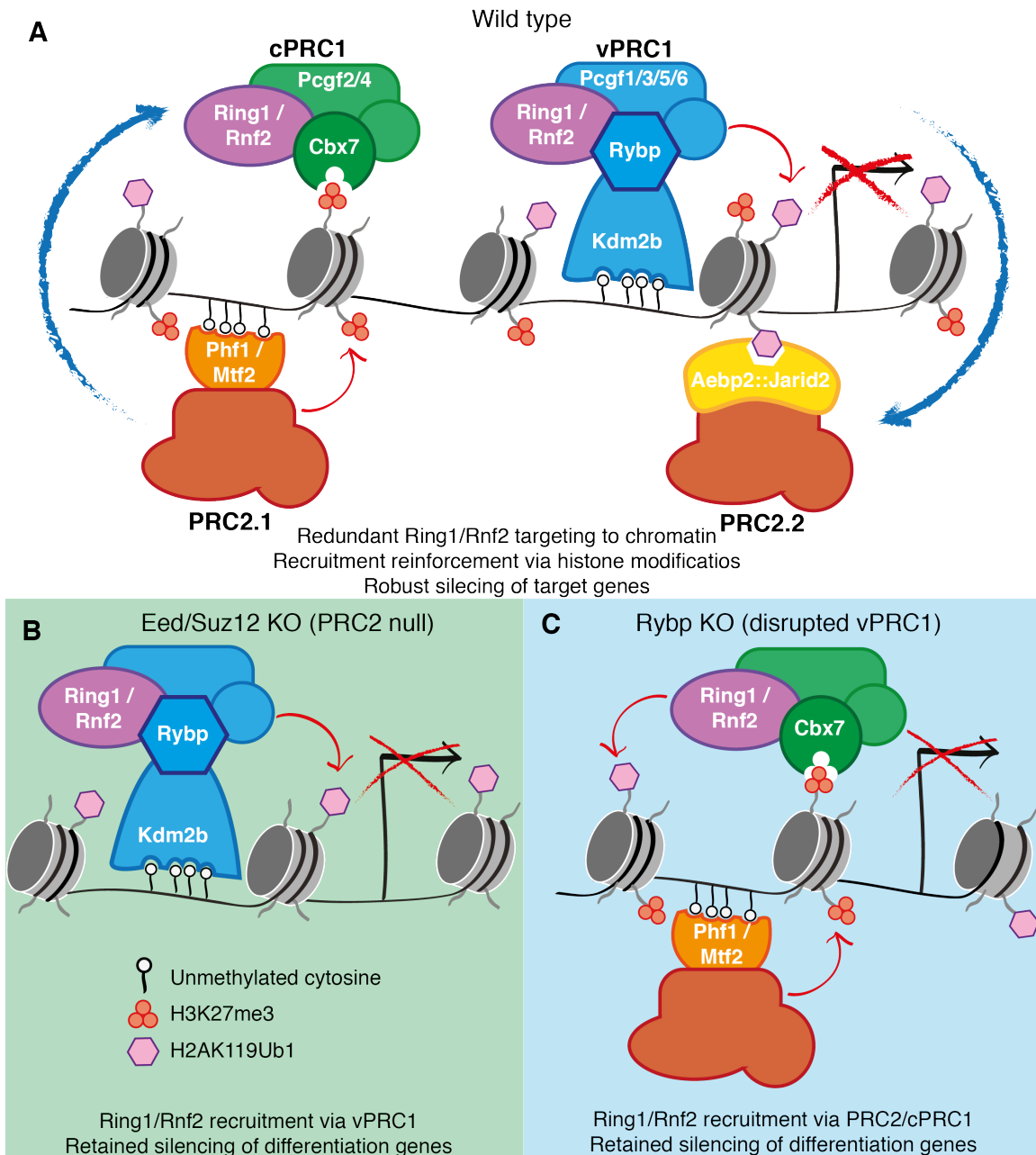
depletions, (Fig. 3.5.6 A and B). Unexpectedly, *Rybp/Suz12* double depletion leads to a general reduction of H3K4me3 irrespective of transcriptional changes (Fig. 3.5.6 C), which likely reflects the fact that after 24 hours *Rnf2* targeting defects and gene expression changes are less dramatic than after 48 hours of *Rybp/Suz12* double depletion.

In summary, these results demonstrate that bivalency does not facilitate gene activation in the absence of Polycomb-mediated silencing. Instead, Trithorax activity, as measured by H3K4me3, increases at de-repressed genes, which is not completely surprising as disrupting Polycomb function is quite distinct from differentiation stimuli. Thus, We speculate that it is not sufficient to inactivate the repressive Polycomb machinery, but it may be necessary to express/recruit specific transcription factors to activate the unresponsive PcG targets.

### **A model for the role of Polycomb complexes in maintenance of ES cell identity**

The above results reveal that in mES cells, PRC2/cPRC1 and vPRC1 play partially redundant roles in delivering the PRC1 catalytic core (*Ring1/Rnf2*) to exert its repressive capacity on target genes. Therefore, Polycomb targeting depends on the combined activities of canonical and variant PRC1 as well as PRC2 to maintain repression of master regulators of development (Fig. 4.1).

Surprisingly, target genes most sensitive to Polycomb loss bear the highest buildups of PcG proteins. While it remains unknown how this highly occupied Polycomb domains are formed, we speculate that vPRC1 and PRC2 are initially nucleated at un-methylated CGIs, given that vPRC1 recognizes un-methylated CGIs via its Kdm2b subunit, and Polycomb-like (PCL) proteins MTF2 and PFH1 have been recently reported to recruit PRC2 to un-methylated CGIs. Then PRC2-dependent H3K27me3 can signal for cPRC1 recruitment and thus create a self-reinforcing system that is robust to environmental and genetic perturbations, in this way ensuring maintenance of transcriptional repression of master regulators of developmental, this ensures the safeguarding ES cell identity and self-renewal capacity.



**Figure 4.1: Model of PcG complexes synergistic recruitment in mESCs.** (A) Schematic depiction of PcG complexes recruitment to, and reinforcing interactions on chromatin. PRC2.1 can recognize unmethylated CGIs via the PCL subunits Phf1 or Mtf2, where it deposits H3K27me3. On the other hand, vPRC1.1 can also bind to unmethylated CGIs through the Kdm2b subunit and catalyze H2AK119Ub1. Moreover, cPRC1 is recruited to chromatin via recognition of H3K27me3 by Cbx7, while PRC1-deposited H2AK119Ub1 is recognized the Aebp2 and Jarid2 containing PRC2.2. This way, large Polycomb chromatin domains are formed over target sites to ensure robust silencing of differentiation driver genes. (B) In PRC2 null mESCs, recruitment to CGIs is maintained by vPRC1.1 which delivers the PRC1 catalytic core (Ring1/Rnf2) ensuring silencing of most target genes. (C) Targeting of Ring1/Rnf2 to CGIs is retained in Rybp KO (disrupted vPRC1) mESCs by PRC2.1 and cPRC1, maintaining repression of Shared (PRC2/cPRC1 and vPRC1) target genes. Therefore, PRC2 null or vPRC1 impaired mESCs can retain ES cell identity and self-renewal capacity.

## References

- [1] C. H. Waddington, "The epigenotype," *Endeavour*, no. 1, pp. 18–20, 1942.
- [2] C. H. Waddington, "CANALIZATION OF DEVELOPMENT AND THE INHERITANCE OF ACQUIRED CHARACTERS," *Nature*, vol. 150, no. 3811, pp. 563–565, Nov. 1942.
- [3] M. Ptashne, "On the use of the word 'epigenetic'.," *Curr. Biol.*, vol. 17, no. 7, pp. R233–6, Apr. 2007.
- [4] R. Bonasio, S. Tu, and D. Reinberg, "Molecular Signals of Epigenetic States," *Science*, vol. 330, no. 6004, pp. 612–616, Oct. 2010.
- [5] L. Aloia, B. Di Stefano, and L. Di Croce, "Polycomb complexes in stem cells and embryonic development.," *Development*, vol. 140, no. 12, pp. 2525–2534, Jun. 2013.
- [6] E. B. Lewis, "A gene complex controlling segmentation in *Drosophila*.," *Nature*, vol. 276, no. 5688, pp. 565–570, Dec. 1978.
- [7] B. Schuettengruber and G. Cavalli, "Recruitment of polycomb group complexes and their role in the dynamic regulation of cell fate choice.," *Development*, vol. 136, no. 21, pp. 3531–3542, Nov. 2009.
- [8] R. Margueron and D. Reinberg, "The Polycomb complex PRC2 and its mark in life.," *Nature*, vol. 469, no. 7330, pp. 343–349, Jan. 2011.
- [9] J. A. Simon and R. E. Kingston, "Occupying chromatin: Polycomb mechanisms for getting to genomic targets, stopping transcriptional traffic, and staying put.," *Mol Cell*, vol. 49, no. 5, pp. 808–824, Mar. 2013.
- [10] R. Cao, L. Wang, H. Wang, L. Xia, H. Erdjument-Bromage, P. Tempst, R. S. Jones, and Y. Zhang, "Role of histone H3 lysine 27 methylation in Polycomb-group silencing.," *Science*, vol. 298, no. 5595, pp. 1039–1043, Nov. 2002.
- [11] J. Müller, C. M. Hart, N. J. Francis, M. L. Vargas, A. Sengupta, B. Wild, E. L. Miller, M. B. O'Connor, R. E. Kingston, and J. A. Simon, "Histone methyltransferase activity of a *Drosophila* Polycomb group repressor complex.," *Cell*, vol. 111, no. 2, pp. 197–208, Oct. 2002.
- [12] B. Czermin, R. Melfi, D. McCabe, V. Seitz, A. Imhof, and V. Pirrotta, "Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites.," *Cell*, vol. 111, no. 2, pp. 185–196, Oct. 2002.
- [13] R. Cao, Y.-I. Tsukada, and Y. Zhang, "Role of Bmi-1 and Ring1A in H2A

- ubiquitylation and Hox gene silencing.,” *Mol Cell*, vol. 20, no. 6, pp. 845–854, Dec. 2005.
- [14] N. J. Francis, R. E. Kingston, and C. L. Woodcock, “Chromatin compaction by a polycomb group protein complex.,” *Science*, vol. 306, no. 5701, pp. 1574–1577, Nov. 2004.
- [15] J. C. Scheuermann, A. G. de Ayala Alonso, K. Oktaba, N. Ly-Hartig, R. K. McGinty, S. Fraterman, M. Wilm, T. W. Muir, and J. Müller, “Histone H2A deubiquitinase activity of the Polycomb repressive complex PR-DUB.,” *Nature*, vol. 465, no. 7295, pp. 243–247, May 2010.
- [16] S. Hauri, F. Comoglio, M. Seimiya, M. Gerstung, T. Glatter, K. Hansen, R. Aebersold, R. Paro, M. Gstaiger, and C. Beisel, “A High-Density Map for Navigating the Human Polycomb Complexome.,” *Cell Rep*, vol. 17, no. 2, pp. 583–595, Oct. 2016.
- [17] N. P. Blackledge, N. R. Rose, and R. J. Klose, “Targeting Polycomb systems to regulate gene expression: modifications to a complex story.,” *Nat Rev Mol Cell Biol*, vol. 16, no. 11, pp. 643–649, Nov. 2015.
- [18] J. K. Stock, S. Giadrossi, M. Casanova, E. Brookes, M. Vidal, H. Koseki, N. Brockdorff, A. G. Fisher, and A. Pombo, “Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells.,” *Nat Cell Biol*, vol. 9, no. 12, pp. 1428–1435, Dec. 2007.
- [19] T. Nakagawa, T. Kajitani, S. Togo, N. Masuko, H. Ohdan, Y. Hishikawa, T. Koji, T. Matsuyama, T. Ikura, M. Muramatsu, and T. Ito, “Deubiquitylation of histone H2A activates transcriptional initiation via trans-histone cross-talk with H3K4 di- and trimethylation.,” *Genes Dev*, vol. 22, no. 1, pp. 37–49, Jan. 2008.
- [20] W. Zhou, P. Zhu, J. Wang, G. Pascual, K. A. Ohgi, J. Lozach, C. K. Glass, and M. G. Rosenfeld, “Histone H2A monoubiquitination represses transcription by inhibiting RNA polymerase II transcriptional elongation.,” *Molecular Cell*, vol. 29, no. 1, pp. 69–80, Jan. 2008.
- [21] R. S. Illingworth, M. Moffat, A. R. Mann, D. Read, C. J. Hunter, M. M. Pradeepa, I. R. Adams, and W. A. Bickmore, “The E3 ubiquitin ligase activity of RING1B is not essential for early mouse development.,” *Genes Dev*, vol. 29, no. 18, pp. 1897–1902, Sep. 2015.
- [22] A. R. Pengelly, R. Kalb, K. Finkl, and J. Müller, “Transcriptional repression by PRC1 in the absence of H2A monoubiquitylation.,” *Genes Dev*, vol. 29, no. 14, pp. 1487–1492, Jul. 2015.
- [23] F. Tie, R. Banerjee, C. A. Stratton, J. Prasad-Sinha, V. Stepanik, A. Zlobin, M. O. Diaz, P. C. Scacheri, and P. J. Harte, “CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing.,” *Development*, vol. 136, no. 18, pp. 3131–3141, Sep.



- 2009.
- [24] F. Tie, R. Banerjee, C. Fu, C. A. Stratton, M. Fang, and P. J. Harte, "Polycomb inhibits histone acetylation by CBP by binding directly to its catalytic domain.," *Proc Natl Acad Sci USA*, vol. 113, no. 6, pp. E744–53, Feb. 2016.
  - [25] Z. Shao, F. Raible, R. Mollaaghababa, J. R. Guyon, C. T. Wu, W. Bender, and R. E. Kingston, "Stabilization of chromatin structure by PRC1, a Polycomb complex.," *Cell*, vol. 98, no. 1, pp. 37–46, Jul. 1999.
  - [26] D. J. Grau, B. A. Chapman, J. D. Garlick, M. Borowsky, N. J. Francis, and R. E. Kingston, "Compaction of chromatin by diverse Polycomb group proteins requires localized regions of high charge.," *Genes Dev*, vol. 25, no. 20, pp. 2210–2221, Oct. 2011.
  - [27] M. S. Lau, M. G. Schwartz, S. Kundu, A. J. Savol, P. I. Wang, S. K. Marr, D. J. Grau, P. Schorderet, R. I. Sadreyev, C. J. Tabin, and R. E. Kingston, "Mutation of a nucleosome compaction region disrupts Polycomb-mediated axial patterning.," *Science*, vol. 355, no. 6329, pp. 1081–1084, Mar. 2017.
  - [28] A. H. Wani, A. N. Boettiger, P. Schorderet, A. Ergun, C. Munger, R. I. Sadreyev, X. Zhuang, R. E. Kingston, and N. J. Francis, "Chromatin topology is coupled to Polycomb group protein subnuclear organization.," *Nat Commun*, vol. 7, p. 10291, Jan. 2016.
  - [29] R. Margueron, G. Li, K. Sarma, A. Blais, J. Zavadil, C. L. Woodcock, B. D. Dynlacht, and D. Reinberg, "Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms.," *Mol Cell*, vol. 32, no. 4, pp. 503–518, Nov. 2008.
  - [30] R. Terranova, S. Yokobayashi, M. B. Stadler, A. P. Otte, M. van Lohuizen, S. H. Orkin, and A. H. F. M. Peters, "Polycomb group proteins Ezh2 and Rnf2 direct genomic contraction and imprinted repression in early mouse embryos.," *Dev Cell*, vol. 15, no. 5, pp. 668–679, Nov. 2008.
  - [31] B. Schuettengruber, H.-M. Bourbon, L. Di Croce, and G. Cavalli, "Genome Regulation by Polycomb and Trithorax: 70 Years and Counting.," *Cell*, vol. 171, no. 1, pp. 34–57, Sep. 2017.
  - [32] N. R. Rose, H. W. King, N. P. Blackledge, N. A. Fursova, K. J. Ember, R. Fischer, B. M. Kessler, and R. J. Klose, "RYBP stimulates PRC1 to shape chromatin-based communication between Polycomb repressive complexes.," *Elife*, vol. 5, p. 524, Oct. 2016.
  - [33] N. Herranz, D. Pasini, V. M. Dıaz, C. Francı, A. Gutierrez, N. Dave, M. Escriva, I. Hernandez-Munoz, L. Di Croce, K. Helin, A. Garcıa de Herreros, and S. Peiro, "Polycomb complex 2 is required for E-cadherin

- repression by the Snail1 transcription factor.,” *Mol Cell Biol*, vol. 28, no. 15, pp. 4772–4781, Aug. 2008.
- [34] N. Dietrich, M. Lerdrup, E. Landt, S. Agrawal-Singh, M. Bak, N. Tommerup, J. Rappsilber, E. Södersten, and K. Hansen, “REST-mediated recruitment of polycomb repressor complexes in mammalian cells.,” *PLoS Genet*, vol. 8, no. 3, p. e1002494, Mar. 2012.
- [35] X. Ren and T. K. Kerppola, “REST interacts with Cbx proteins and regulates polycomb repressive complex 1 occupancy at RE1 elements.,” *Mol Cell Biol*, vol. 31, no. 10, pp. 2100–2110, May 2011.
- [36] M. Yu, T. Mazor, H. Huang, H.-T. Huang, K. L. Kathrein, A. J. Woo, C. R. Chouinard, A. Labadorf, T. E. Akie, T. B. Moran, H. Xie, S. Zacharek, I. Taniuchi, R. G. Roeder, C. F. Kim, L. I. Zon, E. Fraenkel, and A. B. Cantor, “Direct recruitment of polycomb repressive complex 1 to chromatin by core binding transcription factors.,” *Mol Cell*, vol. 45, no. 3, pp. 330–343, Feb. 2012.
- [37] K. Plath, J. Fang, S. K. Mlynarczyk-Evans, R. Cao, K. A. Worringer, H. Wang, C. C. de la Cruz, A. P. Otte, B. Panning, and Y. Zhang, “Role of histone H3 lysine 27 methylation in X inactivation.,” *Science*, vol. 300, no. 5616, pp. 131–135, Apr. 2003.
- [38] C. A. McHugh, C.-K. Chen, A. Chow, C. F. Surka, C. Tran, P. McDonel, A. Pandya-Jones, M. Blanco, C. Burghard, A. Moradian, M. J. Sweredoski, A. A. Shishkin, J. Su, E. S. Lander, S. Hess, K. Plath, and M. Guttman, “The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3.,” *Nature*, vol. 521, no. 7551, pp. 232–236, May 2015.
- [39] J. L. Rinn, M. Kertesz, J. K. Wang, S. L. Squazzo, X. Xu, S. A. Brugmann, L. H. Goodnough, J. A. Helms, P. J. Farnham, E. Segal, and H. Y. Chang, “Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs.,” *Cell*, vol. 129, no. 7, pp. 1311–1323, Jun. 2007.
- [40] A. M. Farcas, N. P. Blackledge, I. Sudbery, H. K. Long, J. F. McGouran, N. R. Rose, S. Lee, D. Sims, A. Cerase, T. W. Sheahan, H. Koseki, N. Brockdorff, C. P. Ponting, B. M. Kessler, and R. J. Klose, “KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands.,” *Elife*, vol. 1, p. e00205, Dec. 2012.
- [41] J. He, L. Shen, M. Wan, O. Taranova, H. Wu, and Y. Zhang, “Kdm2b maintains murine embryonic stem cell status by recruiting PRC1 complex to CpG islands of developmental genes.,” *Nat Cell Biol*, Mar. 2013.
- [42] X. Wu, J. V. Johansen, and K. Helin, “Fbxl10/Kdm2b recruits polycomb repressive complex 1 to CpG islands and regulates H2A ubiquitylation.,”

- Mol Cell*, vol. 49, no. 6, pp. 1134–1146, Mar. 2013.
- [43] H. Li, R. Liefke, J. Jiang, J. V. Kurland, W. Tian, P. Deng, W. Zhang, Q. He, D. J. Patel, M. L. Bulyk, Y. Shi, and Z. Wang, “Polycomb-like proteins link the PRC2 complex to CpG islands,” *Nature*, vol. 549, no. 7671, pp. 287–291, Sep. 2017.
- [44] M. Perino, G. van Mierlo, I. D. Karemaker, S. van Genesen, M. Vermeulen, H. Marks, S. J. van Heeringen, and G. J. C. Veenstra, “MTF2 recruits Polycomb Repressive Complex 2 by helical-shape-selective DNA binding,” *Nat Genet*, vol. 50, no. 7, pp. 1002–1010, May 2018.
- [45] J. Min, Y. Zhang, and R.-M. Xu, “Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27,” *Genes Dev*, vol. 17, no. 15, pp. 1823–1828, Aug. 2003.
- [46] L. Wang, J. L. Brown, R. Cao, Y. Zhang, J. A. Kassiss, and R. S. Jones, “Hierarchical recruitment of polycomb group silencing complexes,” *Molecular Cell*, vol. 14, no. 5, pp. 637–646, Jun. 2004.
- [47] L. Tavares, E. Dimitrova, D. Oxley, J. Webster, R. Poot, J. Demmers, K. Bezstarosti, S. Taylor, H. Ura, H. Koide, A. Wutz, M. Vidal, S. Elderkin, and N. Brockdorff, “RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3,” *Cell*, vol. 148, no. 4, pp. 664–678, Feb. 2012.
- [48] S. Cooper, M. Dienstbier, R. Hassan, L. Schermelleh, J. Sharif, N. P. Blackledge, V. De Marco, S. Elderkin, H. Koseki, R. Klose, A. Heger, and N. Brockdorff, “Targeting polycomb to pericentric heterochromatin in embryonic stem cells reveals a role for H2AK119u1 in PRC2 recruitment,” *Cell Rep*, vol. 7, no. 5, pp. 1456–1470, Jun. 2014.
- [49] N. P. Blackledge, A. M. Farcas, T. Kondo, H. W. King, J. F. McGouran, L. L. P. Hanssen, S. Ito, S. Cooper, K. Kondo, Y. Koseki, T. Ishikura, H. K. Long, T. W. Sheahan, N. Brockdorff, B. M. Kessler, H. Koseki, and R. J. Klose, “Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation,” *Cell*, vol. 157, no. 6, pp. 1445–1459, Jun. 2014.
- [50] R. Kalb, S. Latwiel, H. I. Baymaz, P. W. T. C. Jansen, C. W. Müller, M. Vermeulen, and J. Müller, “Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression,” *Nat Struct Mol Biol*, vol. 21, no. 6, pp. 569–571, Jun. 2014.
- [51] R. Illingworth, A. Kerr, D. Desousa, H. J. P. Ellis, J. Stalker, D. Jackson, C. Clee, R. Plumb, J. Rogers, S. Humphray, T. Cox, C. Langford, and A. Bird, “A novel CpG island set identifies tissue-specific methylation at developmental gene loci,” *PLoS Biol*, vol. 6, no. 1, p. e22, Jan. 2008.

- [52] R. S. Illingworth, U. Gruenewald-Schneider, S. Webb, A. R. W. Kerr, K. D. James, D. J. Turner, C. Smith, D. J. Harrison, R. Andrews, and A. P. Bird, "Orphan CpG islands identify numerous conserved promoters in the mammalian genome.," *PLoS Genet*, vol. 6, no. 9, p. e1001134, Sep. 2010.
- [53] A. Shilatifard, "The COMPA family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis.," *Annu Rev Biochem*, vol. 81, pp. 65–95, 2012.
- [54] C. Hodges, J. G. Kirkland, and G. R. Crabtree, "The Many Roles of BAF (mSWI/SNF) and PBAF Complexes in Cancer.," *Cold Spring Harb Perspect Med*, vol. 6, no. 8, p. a026930, Aug. 2016.
- [55] M. Wu, P. F. Wang, J. S. Lee, S. Martin-Brown, L. Florens, M. Washburn, and A. Shilatifard, "Molecular regulation of H3K4 trimethylation by Wdr82, a component of human Set1/COMPAA.," *Mol Cell Biol*, vol. 28, no. 24, pp. 7337–7344, Dec. 2008.
- [56] P. Wang, C. Lin, E. R. Smith, H. Guo, B. W. Sanderson, M. Wu, M. Gogol, T. Alexander, C. Seidel, L. M. Wiedemann, K. Ge, R. Krumlauf, and A. Shilatifard, "Global analysis of H3K4 methylation defines MLL family member targets and points to a role for MLL1-mediated H3K4 methylation in the regulation of transcriptional initiation by RNA polymerase II.," *Mol Cell Biol*, vol. 29, no. 22, pp. 6074–6085, Nov. 2009.
- [57] D. Hu, A. S. Garruss, X. Gao, M. A. Morgan, M. Cook, E. R. Smith, and A. Shilatifard, "The Mll2 branch of the COMPA family regulates bivalent promoters in mouse embryonic stem cells.," *Nat Struct Mol Biol*, vol. 20, no. 9, pp. 1093–1097, Sep. 2013.
- [58] A. Piunti and A. Shilatifard, "Epigenetic balance of gene expression by Polycomb and COMPA families.," *Science*, vol. 352, no. 6290, p. aad9780, Jun. 2016.
- [59] M. J. Evans and M. H. Kaufman, "Establishment in culture of pluripotential cells from mouse embryos.," *Nature*, vol. 292, no. 5819, pp. 154–156, Jul. 1981.
- [60] G. R. Martin, "Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells.," *Proc Natl Acad Sci USA*, vol. 78, no. 12, pp. 7634–7638, Dec. 1981.
- [61] K. Takahashi and S. Yamanaka, "Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors.," *Cell*, vol. 126, no. 4, pp. 663–676, Aug. 2006.
- [62] J. A. Hackett and M. A. Surani, "Regulatory principles of pluripotency:

- from the ground state up.,” *Cell Stem Cell*, vol. 15, no. 4, pp. 416–430, Oct. 2014.
- [63] A. G. Smith, J. K. Heath, D. D. Donaldson, G. G. Wong, J. Moreau, M. Stahl, and D. Rogers, “Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides.,” *Nature*, vol. 336, no. 6200, pp. 688–690, Dec. 1988.
- [64] R. L. Williams, D. J. Hilton, S. Pease, T. A. Willson, C. L. Stewart, D. P. Gearing, E. F. Wagner, D. Metcalf, N. A. Nicola, and N. M. Gough, “Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells.,” *Nature*, vol. 336, no. 6200, pp. 684–687, Dec. 1988.
- [65] L. Morey, A. Santanach, and L. Di Croce, “Pluripotency and epigenetic factors in mouse embryonic stem cell fate regulation.,” *Mol Cell Biol*, vol. 35, no. 16, pp. 2716–2728, Aug. 2015.
- [66] I. Chambers, J. Silva, D. Colby, J. Nichols, B. Nijmeijer, M. Robertson, J. Vrana, K. Jones, L. Grotewold, and A. Smith, “Nanog safeguards pluripotency and mediates germline development.,” *Nature*, vol. 450, no. 7173, pp. 1230–1234, Dec. 2007.
- [67] Y. Toyooka, D. Shimosato, K. Murakami, K. Takahashi, and H. Niwa, “Identification and characterization of subpopulations in undifferentiated ES cell culture.,” *Development*, vol. 135, no. 5, pp. 909–918, Mar. 2008.
- [68] E. Abranches, E. Bekman, and D. Henrique, “Generation and characterization of a novel mouse embryonic stem cell line with a dynamic reporter of Nanog expression.,” *PLoS ONE*, vol. 8, no. 3, p. e59928, Mar. 2013.
- [69] Q. L. Ying, J. Wray, J. Nichols, L. Battle-Morera, B. Doble, J. Woodgett, P. Cohen, and A. Smith, “The ground state of embryonic stem cell self-renewal.,” *Nature*, vol. 453, no. 7194, pp. 519–523, May 2008.
- [70] C. Faust, K. A. Lawson, N. J. Schork, B. Thiel, and T. Magnuson, “The Polycomb-group gene *ee* is required for normal morphogenetic movements during gastrulation in the mouse embryo.,” *Development*, vol. 125, no. 22, pp. 4495–4506, Nov. 1998.
- [71] D. Pasini, A. P. Bracken, M. R. Jensen, E. Lazzerini Denchi, and K. Helin, “Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity.,” *EMBO J*, vol. 23, no. 20, pp. 4061–4071, Oct. 2004.
- [72] D. O’Carroll, S. Erhardt, M. Pagani, S. C. Barton, M. A. Surani, and T. Jenuwein, “The polycomb-group gene *Ezh2* is required for early mouse development.,” *Mol Cell Biol*, vol. 21, no. 13, pp. 4330–4336, Jul. 2001.

- [73] T. Takeuchi, Y. Yamazaki, Y. Katoh-Fukui, R. Tsuchiya, S. Kondo, J. Motoyama, and T. Higashinakagawa, "Gene trap capture of a novel mouse gene, jumonji, required for neural tube formation.," *Genes Dev*, vol. 9, no. 10, pp. 1211–1222, May 1995.
- [74] S. Wang, F. He, W. Xiong, S. Gu, H. Liu, T. Zhang, X. Yu, and Y. Chen, "Polycomblike-2-deficient mice exhibit normal left-right asymmetry.," *Dev Dyn*, vol. 236, no. 3, pp. 853–861, Mar. 2007.
- [75] L. Morey and K. Helin, "Polycomb group protein-mediated repression of transcription.," *Trends Biochem Sci*, vol. 35, no. 6, pp. 323–332, Jun. 2010.
- [76] M. K. Pirity, J. Locker, and N. Schreiber-Agus, "Rybp/DEDAF is required for early postimplantation and for central nervous system development.," *Mol Cell Biol*, vol. 25, no. 16, pp. 7193–7202, Aug. 2005.
- [77] T. Fukuda, A. Tokunaga, R. Sakamoto, and N. Yoshida, "Fbxl10/Kdm2b deficiency accelerates neural progenitor cell death and leads to exencephaly.," *Mol Cell Neurosci*, vol. 46, no. 3, pp. 614–624, Mar. 2011.
- [78] L. Morey, G. Pascual, L. Cozzuto, G. Roma, A. Wutz, S. A. Benitah, and L. Di Croce, "Nonoverlapping functions of the Polycomb group Cbx family of proteins in embryonic stem cells.," *Cell Stem Cell*, vol. 10, no. 1, pp. 47–62, Jan. 2012.
- [79] K. Hisada, C. Sánchez, T. A. Endo, M. Endoh, M. Román-Trufero, J. Sharif, H. Koseki, and M. Vidal, "RYBP represses endogenous retroviruses and preimplantation- and germ line-specific genes in mouse embryonic stem cells.," *Mol Cell Biol*, vol. 32, no. 6, pp. 1139–1149, Mar. 2012.
- [80] Z. Gao, J. Zhang, R. Bonasio, F. Strino, A. Sawai, F. Parisi, Y. Kluger, and D. Reinberg, "PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes.," *Mol Cell*, vol. 45, no. 3, pp. 344–356, Feb. 2012.
- [81] M. Endoh, T. A. Endo, T. Endoh, Y.-I. Fujimura, O. Ohara, T. Toyoda, A. P. Otte, M. Okano, N. Brockdorff, M. Vidal, and H. Koseki, "Polycomb group proteins Ring1A/B are functionally linked to the core transcriptional regulatory circuitry to maintain ES cell identity.," *Development*, vol. 135, no. 8, pp. 1513–1524, Apr. 2008.
- [82] D. Pasini, A. P. Bracken, J. B. Hansen, M. Capillo, and K. Helin, "The polycomb group protein Suz12 is required for embryonic stem cell differentiation.," *Mol Cell Biol*, vol. 27, no. 10, pp. 3769–3779, May 2007.
- [83] J. Wang, J. Mager, E. Schnedier, and T. Magnuson, "The mouse PcG

- gene eed is required for Hox gene repression and extraembryonic development,” *Mammalian Genome*, vol. 13, no. 9, pp. 493–503, Sep. 2002.
- [84] M. Leeb, D. Pasini, M. Novatchkova, M. Jaritz, K. Helin, and A. Wutz, “Polycomb complexes act redundantly to repress genomic repeats and genes,” *Genes & Development*, vol. 24, no. 3, pp. 265–276, Feb. 2010.
- [85] U. Elling, J. Taubenschmid, G. Wirnsberger, R. O'Malley, S.-P. Demers, Q. Vanhaelen, A. I. Shukalyuk, G. Schmauss, D. Schramek, F. Schnuetgen, H. von Melchner, J. R. Ecker, W. L. Stanford, J. Zuber, A. Stark, and J. M. Penninger, “Forward and reverse genetics through derivation of haploid mouse embryonic stem cells,” *Cell Stem Cell*, vol. 9, no. 6, pp. 563–574, Dec. 2011.
- [86] U. Elling, R. A. Wimmer, A. Leibbrandt, T. Burkard, G. Michlits, A. Leopoldi, T. Micheler, D. Abdeen, S. Zhuk, I. M. Aspalter, C. Handl, J. Liebergesell, M. Hubmann, A.-M. Husa, M. Kinzer, N. Schuller, E. Wetzel, N. van de Loo, J. A. Z. Martinez, D. Estoppey, R. Riedl, F. Yang, B. Fu, T. Dechat, Z. Ivics, C. A. Agu, O. Bell, D. Blaas, H. Gerhardt, D. Hoepfner, A. Stark, and J. M. Penninger, “A reversible haploid mouse embryonic stem cell biobank resource for functional genomics,” *Nature*, vol. 14, no. 7674, p. 357, Sep. 2017.
- [87] A. M. Deaton and A. Bird, “CpG islands and the regulation of transcription,” *Genes & Development*, vol. 25, no. 10, pp. 1010–1022, May 2011.
- [88] M. Ku, R. P. Koche, E. Rheinbay, E. M. Mendenhall, M. Endoh, T. S. Mikkelsen, A. Presser, C. Nusbaum, X. Xie, A. S. Chi, M. Adli, S. Kasif, L. M. Ptaszek, C. A. Cowan, E. S. Lander, H. Koseki, and B. E. Bernstein, “Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains,” *PLoS Genet*, vol. 4, no. 10, p. e1000242, Oct. 2008.
- [89] N. R. Rose and R. J. Klose, “Understanding the relationship between DNA methylation and histone lysine methylation,” *Biochim. Biophys. Acta*, vol. 1839, no. 12, pp. 1362–1372, Dec. 2014.
- [90] N. P. Blackledge, H. K. Long, J. C. Zhou, S. Kriaucionis, R. Patient, and R. J. Klose, “Bio-CAP: a versatile and highly sensitive technique to purify and characterise regions of non-methylated DNA,” *Nucleic Acids Res.*, vol. 40, no. 4, pp. e32–e32, Feb. 2012.
- [91] R. Wang, A. B. Taylor, B. Z. Leal, L. V. Chadwell, U. Ilangoan, A. K. Robinson, V. Schirf, P. J. Hart, E. M. Lafer, B. Demeler, A. P. Hinck, D. G. McEwen, and C. A. Kim, “Polycomb group targeting through different binding partners of RING1B C-terminal domain,” *Structure*, vol. 18,

- no. 8, pp. 966–975, Aug. 2010.
- [92] Y. Benjamini and T. P. Speed, “Summarizing and correcting the GC content bias in high-throughput sequencing.,” *Nucleic Acids Res.*, vol. 40, no. 10, pp. e72–e72, May 2012.
  - [93] L. A. Boyer, K. Plath, J. Zeitlinger, T. Brambrink, L. A. Medeiros, T. I. Lee, S. S. Levine, M. Wernig, A. Tajonar, M. K. Ray, G. W. Bell, A. P. Otte, M. Vidal, D. K. Gifford, R. A. Young, and R. Jaenisch, “Polycomb complexes repress developmental regulators in murine embryonic stem cells.,” *Nature*, vol. 441, no. 7091, pp. 349–353, May 2006.
  - [94] K. Nishimura, T. Fukagawa, H. Takisawa, T. Kakimoto, and M. Kanemaki, “An auxin-based degron system for the rapid depletion of proteins in nonplant cells.,” *Nature Methods*, vol. 6, no. 12, pp. 917–922, Dec. 2009.
  - [95] M. Morawska and H. D. Ulrich, “An expanded tool kit for the auxin-inducible degron system in budding yeast.,” *Yeast*, vol. 30, no. 9, pp. 341–351, Sep. 2013.
  - [96] K. Štefková, J. Procházková, and J. Pacherník, “Alkaline Phosphatase in Stem Cells,” *Stem Cells International*, vol. 2015, no. 10, pp. 1–11, 2015.
  - [97] Q. Cao, X. Wang, M. Zhao, R. Yang, R. Malik, Y. Qiao, A. Poliakov, A. K. Yocum, Y. Li, W. Chen, X. Cao, X. Jiang, A. Dahiya, C. Harris, F. Y. Feng, S. Kalantry, Z. S. Qin, S. M. Dhanasekaran, and A. M. Chinnaiyan, “The central role of EED in the orchestration of polycomb group complexes.,” *Nat Commun*, vol. 5, p. 3127, 2014.
  - [98] P. Moll, M. Ante, A. Seitz, and T. Reda, “QuantSeq 3’ mRNA sequencing for RNA quantification,” *Nature Methods*, vol. 11, no. 12, pp. i–iii, Dec. 2014.
  - [99] P. A. Steffen and L. Ringrose, “What are memories made of? How Polycomb and Trithorax proteins mediate epigenetic memory.,” *Nat Rev Mol Cell Biol*, vol. 15, no. 5, pp. 340–356, May 2014.
  - [100] R. J. Klose, S. Cooper, A. M. Farcas, N. P. Blackledge, and N. Brockdorff, “Chromatin sampling--an emerging perspective on targeting polycomb repressor proteins.,” *PLoS Genet*, vol. 9, no. 8, p. e1003717, Aug. 2013.
  - [101] P. Voigt, W.-W. Tee, and D. Reinberg, “A double take on bivalent promoters.,” *Genes Dev*, vol. 27, no. 12, pp. 1318–1338, Jun. 2013.
  - [102] S. J. Chamberlain, D. Yee, and T. Magnuson, “Polycomb repressive complex 2 is dispensable for maintenance of embryonic stem cell pluripotency.,” *Stem Cells*, vol. 26, no. 6, pp. 1496–1505, Jun. 2008.
  - [103] E. M. Riising, I. Comet, B. Leblanc, X. Wu, J. V. Johansen, and K. Helin, “Gene silencing triggers polycomb repressive complex 2 recruitment to



CpG islands genome wide.,” *Mol Cell*, vol. 55, no. 3, pp. 347–360, Aug. 2014.