# The molecular determinants of Polycomb-mediated gene repression

Nadezda A. Fursova




UNIVERSITY COLLEGE, UNIVERSITY OF OXFORD

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

TRINITY TERM, 2019

# ABSTRACT


**The molecular determinants of Polycomb-mediated gene repression**

Nadezda A. Fursova, University College, Trinity Term 2019

*A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy of the University of Oxford*

Precise spatio-temporal regulation of gene expression is required for development of multicellular organisms. Combinatorial activity of DNA-binding transcription factors has been proposed to drive lineage-specific transcriptional programmes. However, in addition to this, the chromatin template, on which transcription occurs, has emerged as another central player in the control of gene expression. The Polycomb repressive system represents one of the most extensively studied paradigms for chromatin-based gene regulation. It is comprised of two multi-subunit complexes, which catalyse post-translational modifications of histone tails - Polycomb repressive complex 1 (PRC1) mono-ubiquitinates histone H2A at lysine 119 (H2AK119ub1) and PRC2 methylates histone H3 at lysine 27 (H3K27me3). Elucidating the molecular mechanisms of Polycomb-mediated transcriptional silencing has proven challenging, as both PRC1 and PRC2 exist as an array of protein assemblies with distinct targeting modules and regulatory activities. Therefore, the primary determinants that underlie target site selection, Polycomb chromatin domain formation and gene repression by the Polycomb system have until now remained elusive.

Here, I have exploited combinatorial genetic perturbations coupled with quantitative genomic approaches to overcome the multiplicity of Polycomb complexes and uncover the central determinants of Polycomb-mediated gene repression in mouse embryonic stem cells. In contrast to the prevailing dogma in the field, I demonstrated that canonical PRC1, which binds Polycomb target sites in a PRC2-dependent manner and mediates higher-order chromatin structures, contributes little to gene repression. Instead, through systematic interrogation of how H2AK119ub1 is placed and kept in check in the genome, I discovered a remarkable degree of synergy between variant PRC1 complexes, which is central to deposition of H2AK119ub1, Polycomb chromatin domain formation and gene repression. Importantly, my observations suggested that the repressive activity of variant PRC1 complexes may rely on their capacity to catalyse H2AK119ub1. Using a conditional genetic strategy to inactivate PRC1, I directly tested this possibility and discovered that PRC1 catalytic activity is the central molecular determinant of Polycomb-mediated gene repression. In conclusion, through systematic dissection of the Polycomb system, I have provided compelling evidence for a variant PRC1-dependent logic for formation of Polycomb chromatin domains and repression of Polycomb target genes. Together, my findings place PRC1 and its catalytic activity at the forefront of Polycomb-mediated transcriptional silencing, paving the way for a mechanistic understanding of Polycomb system function.

# ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor, Prof. Robert Klose, for his constant support, guidance, enthusiasm, and most of all, patience. I would also like to thank all members of the Klose lab, but especially my scientific partner in crime, Dr. Neil Blackledge, with whom I have enjoyed working in close collaboration throughout my DPhil studies. In addition, I would like to extend my special thanks to Dr. Hamish King for teaching me so much about bioinformatics and with whom I had many stimulating discussions. I am also especially thankful to Dr. Emilia Dimitrova for all the experimental advice and Prof. Haruhiko Koseki for providing several key cell lines that were instrumental to this work. Finally, I would like to thank my friends, my partner Alex and my family, who have been forgiving my long absences from home, for their constant love and support.

This work has been financially supported by several organisations to whom I am incredibly grateful, including Department of Biochemistry (University of Oxford), University College Oxford, Clarendon Fund, Medical Sciences Division (University of Oxford), The Vice-Chancellors' Fund (University of Oxford) and European Research Council.

# DECLARATION OF AUTHORSHIP

*I declare that all work presented in this thesis is my own, unless otherwise acknowledged.*

*This thesis has not been submitted, either partially or in full, for another degree, diploma, certificate or other qualification at this University or at any other institution.*

Nadezda A. Fursova

8th October 2019

# ASSOCIATED PUBLICATIONS

Blackledge NP*, **Fursova NA***, Kelley JR, Huseyin MK, Feldmann A, and Klose RJ. (2020). "PRC1 catalytic activity is central to Polycomb system function". ***Molecular Cell*** 77, 857-874.e9.

Rhodes JDP*, Feldmann A*, Hernandez-Rodriguez B*, Diaz N*, Brown JM, **Fursova NA**, Blackledge NP, Prathapan P, Dobrinic P, Huseyin MK, Szczurek A, Kruse K, Nasmyth KA, Buckle VJ, Vaquerizas JM, and Klose RJ. (2020). "Cohesin disrupts Polycomb-dependent chromosome interactions". ***Cell Reports*** 30: pp. 820-835.e10.

**Fursova NA***, Blackledge NP*, Nakayama M, Ito S, Koseki Y, Farcas AM, King HW, Koseki H, and Klose RJ. (2019). "Synergy between variant PRC1 complexes defines Polycomb-mediated gene repression". ***Molecular Cell*** 74: pp. 1020-1036.e8.

King HW, **Fursova NA**, Blackledge NP, and Klose RJ. (2018). "Polycomb repressive complex 1 shapes the nucleosome landscape but not accessibility at target genes". ***Genome Research*** 28: pp. 1494-507.

Rose NR*, King HW*, Blackledge NP*, **Fursova NA**, Ember KJ, Fischer R, Kessler BM, and Klose RJ. (2016). "RYBP stimulates PRC1 to shape chromatin-based communication between Polycomb repressive complexes". ***ELife*** 5: pp. e18591.

*These authors contributed equally.

# CONTENTS

**Abstract** i

**Acknowledgements** ii

**Declaration of authorship** iii

**Associated Publications** iv

**Contents** v

**List of figures** viii

**List of tables** xi

**List of common abbreviations** xii

**1 Introduction** 1
1.1 Regulation of transcription in eukaryotes . . . 1
1.1.1 General principles and key stages of transcription . . . 1
1.1.2 Chromatin-based regulation of transcription . . . 3
1.1.2.1 Chromatin remodelling . . . 7
1.1.2.2 Histone variants . . . 9
1.1.2.3 Histone modifications . . . 10
1.1.2.4 Higher-order chromatin structures . . . 18
1.2 Polycomb group proteins and their role in gene repression . . . 20
1.2.1 Discovery of Polycomb group proteins . . . 20
1.2.2 PRC1 in mammals . . . 22
1.2.2.1 Canonical PRC1 . . . 24
1.2.2.2 Variant PRC1 . . . 28
1.2.3 PRC2 in mammals . . . 33
1.2.3.1 Catalytic core of PRC2 . . . 33
1.2.3.2 PRC2.1 complexes . . . 36
1.2.3.3 PRC2.2 complexes . . . 37
1.2.4 Polycomb repressive deubiquitinase (PR-DUB) complex in mammals . . . 39
1.3 Polycomb-mediated gene repression: proposed models . . . 42
1.3.1 Polycomb recruitment mechanisms to target sites . . . 42
1.3.1.1 Polycomb recruitment by sequence-specific transcription factors 43
1.3.1.2 Polycomb recruitment by long non-coding RNAs . . . 44
1.3.1.3 Polycomb recruitment by generic DNA-binding factors . . . 45
1.3.1.4 Chromatin-based communication between PRC1 and PRC2 . . . 47
1.3.2 Proposed mechanisms of gene repression . . . 49
1.4 Aims of this thesis . . . 53

**2 Materials and Methods** 54
2.1 Cell lines and culture . . . 54
2.1.1 Cell culture conditions . . . 54
2.1.2 Genome engineering by CRISPR/Cas9-mediated Homology-Directed Repair (HDR) . . . 55
2.1.3 Cell line generation . . . 56

2.1.4 Cell growth curve assay . . . 58
2.2 Protein methods . . . 58
2.2.1 Preparation of nuclear extracts . . . 58
2.2.2 Preparation of histone extracts . . . 59
2.2.3 Co-immunoprecipitation (co-IP) . . . 59
2.3 Chromatin immunoprecipitation . . . 60
2.3.1 Calibrated ChIP-sequencing (cChIP-seq) . . . 60
2.3.2 Native cChIP-sequencing . . . 61
2.4 RNA isolation and sequencing . . . 62
2.4.1 Calibrated nuclear RNA-seq (cnRNA-seq) . . . 62
2.4.2 Calibrated total RNA-seq (cRNA-seq) . . . 63
2.5 Massively parallel sequencing, data processing, normalisation and analysis . . . 64
2.5.1 Massively parallel sequencing . . . 64
2.5.2 Read alignment . . . 64
2.5.3 Data normalisation and visualisation . . . 65
2.5.4 Peak calling . . . 66
2.5.5 Read count quantitation and analysis . . . 67
2.5.5.1 Read count quantitation . . . 67
2.5.5.2 Box plot analysis . . . 67
2.5.5.3 Percentile analysis . . . 68
2.5.5.4 Correlation and statistical analysis . . . 68
2.5.5.5 Heatmap and metaplot analysis . . . 68
2.5.6 Differential enrichment and gene expression analysis . . . 69
2.5.7 Functional annotation of genes . . . 69
2.5.7.1 Annotation of Polycomb target genes . . . 69
2.5.7.2 Annotation of differentially upregulated genes in distinct PCGF conditional knockout ES cell lines . . . 70
2.5.7.3 Annotation of differentially upregulated genes in PRC1$^{CPM}$ and PRC1$^{CKO}$ ES cell lines . . . 71
2.5.7.4 Gene ontology (GO) term enrichment analysis . . . 71
2.5.8 Published datasets used in this study . . . 71

**3 Canonical PRC1 complexes are dispensable for Polycomb-mediated gene repression 73**
3.1 Introduction . . . 73
3.2 Results . . . 74
3.2.1 A mouse ES cell line to conditionally remove canonical PRC1 . . . 74
3.2.2 Canonical PRC1 complexes contribute minimally to Polycomb-mediated gene repression . . . 78
3.2.3 Canonical PRC1 shapes high-level enrichment of RING1B but not H2AK119ub1, PRC2 or H3K27me3 at Polycomb chromatin domains . . . 78
3.3 Discussion . . . 82

**4 Distinct roles of variant PRC1 complexes in Polycomb system function 85**
4.1 Introduction . . . 85
4.2 Results . . . 87
4.2.1 H2AK119ub1 is a pervasive histone modification that is enriched at PRC1-bound sites . . . 87
4.2.2 PCGF1-PRC1 plays a major role in deposition of H2AK119ub1 at PRC1-bound sites . . . 90
4.2.3 PCGF1-PRC1 supports Polycomb chromatin domain formation and is essential but not sufficient for Polycomb-mediated gene repression . . . 93

4.3 Discussion . . . . . 95

**5 Dynamic control of genome-wide H2AK119ub1 and its contribution to gene regulation 99**
5.1 Introduction . . . . . 99
5.2 Results . . . . . 101
5.2.1 PCGF3/5-PRC1 play a major role in deposition of the genome-wide pool of H2AK119ub1 . . . . . 101
5.2.2 Pervasive deposition of H2AK119ub1 by PCGF3/5-PRC1 is not required for Polycomb chromatin domain formation and gene repression . . . . . 103
5.2.3 Pervasive PCGF3/5-mediated H2AK119ub1 is linked to chromosome-wide gene silencing during X-inactivation . . . . . 104
5.2.4 BAP1 deubiquitinase complex is the mammalian PR-DUB that keeps in check genome-wide H2AK119ub1 . . . . . 107
5.2.5 Genome-wide accumulation of H2AK119ub1 following removal of BAP1 is associated with widespread transcriptional changes . . . . . 110
5.3 Discussion . . . . . 112

**6 Synergy between variant PRC1 complexes drives Polycomb-mediated gene repression 117**
6.1 Introduction . . . . . 117
6.2 Results . . . . . 118
6.2.1 Removal of PCGF1/3/5-PRC1 does not completely abolish Polycomb chromatin domain formation and gene repression . . . . . 118
6.2.2 Canonical PRC1 does not compensate for compromised activity of variant PRC1 complexes . . . . . 120
6.2.3 Variant PCGF1/3/5/6-PRC1 complexes define Polycomb chromatin domain formation and repression of Polycomb target genes . . . . . 126
6.3 Discussion . . . . . 133

**7 PRC1 catalytic activity is central to Polycomb system function 136**
7.1 Introduction . . . . . 136
7.2 Results . . . . . 138
7.2.1 A conditional catalytic point mutant system to eliminate PRC1 catalytic activity . . . . . 138
7.2.2 PRC1 catalytic activity facilitates PRC2 binding and deposition of H3K27me3 at Polycomb target sites . . . . . 141
7.2.3 Recruitment of canonical PRC1 to Polycomb target sites relies on PRC1 catalytic activity . . . . . 144
7.2.4 Polycomb target site selection relies on DNA-based sampling by variant PRC1 complexes but not PRC1 catalytic activity . . . . . 146
7.2.5 PRC1 catalytic activity is central to Polycomb-mediated gene repression . 150
7.3 Discussion . . . . . 153

**8 Conclusions and future directions 159**

**A Appendices 167**

**Bibliography 170**

# List of Figures


1.1 Chromatin-based regulation of transcription: histone variants and post-translational modifications of histone tails . . . 4
1.2 Dynamic control of deposition and removal of H3K27 methylation and H2AK119 mono-ubiquitination . . . 14
1.3 Diversity of PRC1 complexes in *Drosophila* and mammals . . . 23
1.4 Two distinct classes of mammalian PRC2 complexes . . . 34
1.5 Hierarchical and alternative models for Polycomb chromatin domain formation in mammals . . . 48

3.1 A *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ mouse ES cell line to conditionally remove canonical PRC1 . 75
3.2 Canonical PRC1 complexes are disrupted in the *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ ES cells following OHT treatment . . . 76
3.3 Binding of canonical PRC1 at Polycomb target sites is lost following removal of PCGF2/4 . . . 77
3.4 Canonical PRC1 complexes are not required for Polycomb target gene repression in ES cells . . . 79
3.5 Canonical PRC1 shapes high-level enrichment of RING1B but not H2AK119ub1, PRC2 or H3K27me3 at Polycomb chromatin domains . . . 80
3.6 H2AK119ub1, SUZ12 and H3K27me3 levels are reduced at a small number of large Polycomb chromatin domains following loss of cPRC1 . . . 81

4.1 H2AK119ub1 is enriched at PRC1-bound sites . . . 88
4.2 H2AK119ub1 is a widespread histone modification that blankets the genome . . 89
4.3 A panel of conditional knockout ES cell lines to study the contribution of individual variant PRC1 complexes to deposition of H2AK119ub1 . . . 90
4.4 PCGF1-PRC1 plays a major role in deposition of H2AK119ub1 at PRC1-bound sites . . . 91
4.5 PCGF6-PRC1 is responsible for deposition of H2AK119ub1 at a subset of germline-specific genes . . . 92
4.6 PCGF1-PRC1 supports Polycomb chromatin domain formation at PRC1-bound sites 93
4.7 PCGF1-PRC1 is essential but not sufficient for Polycomb-mediated gene repression 94

5.1 PCGF3/5-PRC1 pervasively deposit H2AK119ub1 throughout the genome . . . 102
5.2 PCGF3/5-dependent H2AK119ub1 is not required for Polycomb chromatin domain formation . . . 103
5.3 Pervasive deposition of H2AK119ub1 by PCGF3/5-PRC1 contributes little to Polycomb-mediated gene repression . . . 104
5.4 An inducible system to study Xist-mediated chromosome-wide gene silencing . . 105
5.5 Broad accumulation of H2AK119ub1 throughout the Xist-expressing chromosome is associated with transcriptional silencing . . . 106
5.6 PCGF3/5-PRC1 are required for chromosome-wide accumulation of H2AK119ub1 during Xist-mediated chromosome silencing . . . 107
5.7 BAP1 is responsible for removal of H2AK119ub1 throughout the genome . . . 108
5.8 H2AK119ub1 and H3K27me3 levels at Polycomb target sites increase following removal of BAP1 . . . 109
5.9 Genome-wide accumulation of H2AK119ub1 following BAP1 removal is associated with widespread transcriptional changes . . . 110

5.10 Redistribution of H2AK119ub1 and H3K27me3 following BAP1 removal compromises Polycomb-mediated gene repression . . . . . 111

6.1 Polycomb chromatin domains are compromised but not completely eroded following loss of PCGF1/3/5-PRC1 . . . . . 119
6.2 Loss of PCGF1/3/5-PRC1 leads to a widespread but not complete derepression of Polycomb target genes . . . . . 120
6.3 Removal of PCGF2 in addition to PCGF1/3/5 does not further compromise Polycomb chromatin domain formation . . . . . 121
6.4 Removal of PCGF1/3/5 or PCGF1/3/5/2 leads to a similar erosion of Polycomb chromatin domains . . . . . 122
6.5 Derepression of Polycomb target genes following loss of PCGF1/3/5 is not further potentiated by removal of PCGF2 . . . . . 122
6.6 Canonical PRC1 does not compensate for compromised Polycomb target gene repression in the absence of PCGF1/3/5-PRC1 . . . . . 123
6.7 PCGF2/4-PRC1 contribute to repression of a subset of Polycomb target genes . . 124
6.8 Variant PCGF1/3/5/6-PRC1 complexes drive Polycomb chromatin domain formation 125
6.9 A subset of Polycomb chromatin domains are retained following removal of PCGF1/3/5/6-PRC1 . . . . . 126
6.10 Variant PCGF1/3/5/6-PRC1 complexes are the central determinants of Polycomb-mediated gene repression . . . . . 128
6.11 Removal of PCGF1/3/5/6 unmasks the contribution of PCGF2/4-PRC1 complexes to Polycomb-mediated gene repression . . . . . 129
6.12 Synergy between variant PCGF1/3/5/6-PRC1 complexes defines Polycomb-mediated gene repression . . . . . 129
6.13 Removal of PCGF proteins in different combinations or RING1A/B results in overlapping transcriptional changes . . . . . 130
6.14 PCGF1/3/5/6-PRC1 complexes co-occupy promoters of variant PRC1 target genes 131

7.1 A conditional catalytic point mutant system to eliminate PRC1 catalytic activity . 139
7.2 H2AK119ub1 is completely lost throughout the genome following OHT treatment in PRC1$^{CPM}$ cells . . . . . 140
7.3 PRC1 catalytic activity supports stable PRC2 binding and deposition of H3K27me3 at Polycomb target sites . . . . . 142
7.4 Loss of PRC1 catalytic activity compromises binding of both classes of PRC2 complexes . . . . . 143
7.5 Canonical PRC1 preferentially occupies PRC1-bound sites which are highly enriched with PRC2 and H3K27me3 . . . . . 145
7.6 Binding of canonical PRC1 at target sites is lost following removal of PRC1 catalytic activity . . . . . 145
7.7 Variant PRC1 complexes broadly occupy Polycomb target sites independently of PRC1 catalysis . . . . . 147
7.8 Removal of PRC1 catalytic activity unmasks DNA-based sampling of Polycomb target sites by variant PRC1 complexes . . . . . 148
7.9 PRC1 catalytic activity drives high-level PRC1 occupancy at lowly transcribed gene promoters . . . . . 149
7.10 PRC1 catalytic activity is required for Polycomb-mediated gene repression . . . 151
7.11 Gene reactivation following OHT treatment is highly similar in PRC1$^{CPM}$ and PRC1$^{CKO}$ cells . . . . . 152
7.12 A small subset of Polycomb target genes are not fully derepressed following removal of PRC1 catalytic activity . . . . . 153

7.13 Genes that are not fully derepressed following removal of PRC1 catalytic activity retain low-level binding of canonical PRC1 . . . . . 154

A.1 Establishing a catalytically inactive PRC1 complex *in vitro* (courtesy of Jessica Kelley) . . . . . 167
A.2 RING1B$^{I53A/D56K}$ forms stable PRC1 complexes (courtesy of Dr. Neil Blackledge) 168
A.3 PRC1 catalytic activity is required to support long-range chromatin interactions between Polycomb target sites (courtesy of Dr. Angelika Feldmann) . . . . . 169

# List of Tables

2.1 Description of the genotype, cell culture conditions, drug treatments and source of origin for mouse ES cell lines used in this study . . . . . 55
2.2 Description of antibodies used in this study for western blot and cChIP-seq analysis 60
2.3 Accession codes for the previously published datasets used in this study . . . . . 72

# LIST OF COMMON ABBREVIATIONS

| | |
|---|---|
| **4SU** | 4-thiouridine |
| **4SU-seq** | isolation of 4SU pulse-labelled RNA followed by massively parallel sequencing |
| **aa** | amino acid |
| **ac** | acetylation |
| **ARID** | AT-rich interaction domain |
| **ATP** | adenosine triphosphate |
| **bp** | base pair |
| **BioCAP-seq** | biotinylated CXXC affinity purification followed by massively parallel sequencing |
| **BSA** | bovine serum albumin |
| **CGI** | CpG island |
| **ChIP** | chromatin immunoprecipitation |
| **ChIP-seq** | ChIP followed by massively parallel sequencing |
| **cChIP-seq** | spike-in calibrated ChIP-seq |
| **chr** | chromosome |
| **CKO** | conditional knockout |
| **CPM** | conditional point mutant |
| **Cre-ERT2** | Cre recombinase fused to the tamoxifen-inducible estrogen receptor |
| **DNase I** | deoxyribonuclease I |
| **DNMT** | DNA methyltransferase |
| **DOX** | doxycycline |
| **DSG** | disuccinimidyl glutarate |
| **dsDNA** | double-stranded DNA |
| **DUB** | deubiquitinase |
| **ESC** | embryonic stem cells |
| **FBS** | fetal bovine serum |
| **FLC** | flowering locus C |
| **gDNA** | genomic DNA |
| **GTF** | general transcription factor |
| **HAT** | histone acetyltransferase |
| **HDAC** | histone deacetylase |
| **HDR** | homology directed repair |
| **HRP** | horse radish peroxidase |
| **IP** | immunoprecipitation |
| **JmjC** | Jumonji C |
| **kb** | kilobase |
| **LIF** | leukaemia inhibitory factor |
| **MBD** | methyl binding domain |
| **MBT** | malignant brain tumour |
| **me1** | mono-methylation |
| **me2** | di-methylation |
| **me3** | tri-methylation |
| **MEF** | mouse embryonic fibroblasts |
| **MNase** | micrococcal nuclease |
| **NDR** | nucleosome-depleted region |
| **NGS** | next generation sequencing |
| **NMI** | non-methylated CpG island |
| **NPC** | neural progenitor cells |

| | |
|---|---|
| **nucRNA** | nuclear RNA |
| **NuRD** | nucleosome remodelling deacetylase |
| **OHT** | 4-hydroxytamoxifen |
| **PBS** | phosphate buffered saline solution |
| **PcG** | Polycomb group |
| **PCGF** | Polycomb group ring finger |
| **PHD** | plant homeodomain |
| **PIC** | pre-initiation complex |
| **PRC** | Polycomb repressive complex |
| **PTM** | post-translational modification |
| **qPCR** | quantitative real-time PCR |
| **RT-qPCR** | reverse transcription followed by qPCR |
| **RNase** | ribonuclease |
| **RING** | really interesting new gene |
| **RNA-seq** | isolation of RNA followed by massively parallel sequencing |
| **cRNA-seq** | spike-in calibrated RNA-seq |
| **cnRNA-seq** | spike-in calibrated nuclear RNA-seq |
| **RPKM** | reads per kilobase per million |
| **SAM** | sterile alpha motif |
| **SCF** | SKP1-CUL1-Fbox |
| **SDS-PAGE** | SDS-polyacrylamide gel electrophoresis |
| **SET** | su(var)3-9-enhancer of zeste-trithorax |
| **sgRNA** | single guide RNA |
| **TAD** | topologically associating domain |
| **TF** | transcription factor |
| **TrxG** | Trithorax group |
| **TSS** | transcription start site |
| **TTS** | transcription termination site |
| **ub1** | mono-ubiquitination |
| **UIM** | ubiquitin interaction motif |
| **WT** | wild type |
| **ZF** | zinc finger |
| **ZF-CXXC** | zinc finger-CXXC |

# 1

# INTRODUCTION

## 1.1 Regulation of transcription in eukaryotes

### 1.1.1 General principles and key stages of transcription

Precise specification and maintenance of unique gene expression programmes is paramount for development of multicellular organisms. Gene expression starts with transcription of a gene's DNA sequence into RNA, and regulation of this process represents the primary mechanism which controls differential gene activity. In eukaryotes, protein-coding genes are transcribed by RNA polymerase II (Roeder and Rutter 1969; Weinmann et al. 1974), in association with a large number of diverse accessory factors, which regulate different steps of the transcription cycle (reviewed in Cramer 2019; Roeder 2019). Transcription typically initiates at specialised DNA sequences, called core promoters (Haberle and Stark 2018), with formation of a pre-initiation complex (PIC) which consists of RNA polymerase II and general transcription factors (GTFs) (Orphanides et al. 1996; Roeder 1996). RNA polymerase II on its own has low affinity for core promoters and it requires association with GTFs to efficiently bind, melt and properly position DNA in the active site for RNA synthesis (Dyke et al. 1988; Buratowski et al. 1989; Kostrewa et al. 2009; Liu et al. 2010; He et al. 2013b; Louder et al. 2016). Despite high conservation of PIC components in eukaryotes (Mühlbacher et al. 2014), PIC composition and functions are not universal across different genes and developmental stages (DAlessio et al. 2009; Goodrich and Tjian 2010; Levine et al. 2014). In line with this, a number of non-canonical GTFs have been discovered as important players in cell type- and gene-specific regulation of transcription (Dikstein et al. 1996; Deato and Tjian 2007; Parry et al. 2010; Guglielmi et al. 2013).

Although core promoters are sufficient to initiate PIC formation, they support only a basal level of gene expression (Kadonaga 2011; Haberle and Stark 2018). RNA polymerase II activity and/or stability at core promoters needs to be further stimulated by transcription factors and co-activator complexes which bind to *cis*-regulatory genomic elements (Banerji et al. 1981; Benoist and Chambon 1981; Moreau et al. 1981; Payvar et al. 1981; Banerji et al. 1983; Dynan and Tjian 1983;

Meisterernst et al. 1991). In addition to promoter-proximal sequences, which predominate in gene regulation in prokaryotes (Losick 1998; Struhl 1999; Ptashne 2005), transcription of eukaryotic genes can also be activated by distal enhancer elements, which can act on promoters located even several hundred kilobases away (Spitz and Furlong 2012; Shlyueva et al. 2014). Importantly, a particular enhancer can activate transcription of only a specific subset of core promoters and this appears to be driven by compatibility between the transcriptional co-activators recruited to the enhancer and the core promoter (Merli et al. 1996; Ohtsuki et al. 1998; Sharpe et al. 1998; Butler and Kadonaga 2001; Zabidi et al. 2014; Haberle et al. 2019). Finally, although core promoters and distal regulatory elements have been originally proposed to play distinct roles in gene regulation (Butler and Kadonaga 2002), this classical view has been recently challenged by the discovery that RNA polymerase II transcription is more pervasive than previously appreciated (Kapranov et al. 2007; Kim et al. 2010). In fact, genome-wide analysis of promoter/enhancer activity of *cis*-regulatory elements revealed that both promoters and enhancers are able to initiate PIC assembly and stimulate transcription of reporter gene promoters (Arnold et al. 2013; Core et al. 2014; Mikhaylichenko et al. 2018; Henriques et al. 2018). This illustrates that in eukaryotes, promoters and enhancers are more functionally similar than has been previously anticipated, yet how shared and distinct functions of these elements in transcriptional regulation are encoded by their DNA sequences still remains poorly understood.

DNA-binding transcription factors, which recognise specific sequence motifs within *cis*-regulatory elements, play fundamental roles in regulating promoter activity (Spitz and Furlong 2012). Multiple mechanisms have been characterised for how transcription factors facilitate transcription. Some of them were shown to directly stabilise binding of GTFs at core promoters, promoting recruitment of RNA polymerase II (Sawadogo 1985; Horikoshi et al. 1988; Roberts et al. 1993; Rochette-Egly et al. 1997; Ptashne and Gann 1997). However, in most cases, transcription factors recruit transcriptional co-activator complexes, which can either directly stimulate RNA polymerase II activity or make promoters more accessible to the transcriptional machinery (Pugh and Tjian 1990; Albright and Tjian 2000; Thomas and Chiang 2006; Fong et al. 2012). One of the co-activators that directly interacts with RNA polymerase II is Mediator, which is a large protein complex comprising 30 different subunits (Flanagan et al. 1991; Thompson et al. 1993; Fondell et al. 1996; Plaschka et al. 2015; Nozawa et al. 2017). Mediator bridges enhancers that are bound by transcriptional activators with the basal transcriptional machinery at the promoter to

facilitate transition of RNA polymerase II into productive elongation (Kornberg 2005; Malik and Roeder 2010; Allen and Taatjes 2015). Although many of the co-activator complexes, including Mediator, are ubiquitously expressed in different tissues, a number of gene- and cell type-specific co-activator assemblies have been identified, providing an additional mechanism for selective gene regulation (Luo et al. 1992; Ge et al. 1998; Lin et al. 2005; Fong et al. 2012). Ultimately, combinatorial and synergistic activities of multiple transcription factors and co-activators define unique gene expression patterns specifying which tissue, which developmental stage and at what level a gene should be expressed.

Integration of activation signals from multiple *cis*-regulatory elements by core promoters enables RNA polymerase II to initiate productive transcription and escape the promoter (Spitz and Furlong 2012; Cramer 2019). Following this, RNA polymerase II typically pauses 20-60 bp downstream of the promoter, before it proceeds to productive elongation (Fraser et al. 1978; Gariglio et al. 1981; Gilmour and Lis 1986; Rougvie and Lis 1988; Rasmussen and Lis 1993; Muse et al. 2007; Zeitlinger et al. 2007). Promoter-proximal pausing represents an additional regulatory checkpoint enabling a rapid gene induction in response to environmental or differentiation cues (Levine 2011; Adelman and Lis 2012). When paused RNA polymerase II is released from the promoter, it forms an elongation complex (Gnatt et al. 2001; Cramer 2001), which processively extends an RNA transcript before reaching a termination signal, after which RNA polymerase II disassociates from RNA and DNA (Jonkers and Lis 2015; Proudfoot 2016; Porrua et al. 2016). Together, initiation, pausing, elongation and termination depict the key steps of eukaryotic transcription cycle, which have to be precisely coordinated to define the appropriate level of gene activity. However, while regulation of elongation and termination additionally fine-tunes the transcriptional output of genes, the decision to schedule a gene for activation is made at the stage of transcription initiation. Therefore, elucidating the DNA sequence-based logic for combinatorial activity of transcription factors and co-activators at promoters and distal regulatory elements is central to deciphering the fundamental principles of gene regulation in eukaryotes.

### 1.1.2 Chromatin-based regulation of transcription

Sequence-specific DNA-binding transcription factors provided us with the initial understanding of gene regulation both in prokaryotes and eukaryotes. However, in eukaryotic cells, transcription does not take place on naked DNA, as their genomic DNA is assembled into nucleosomes which

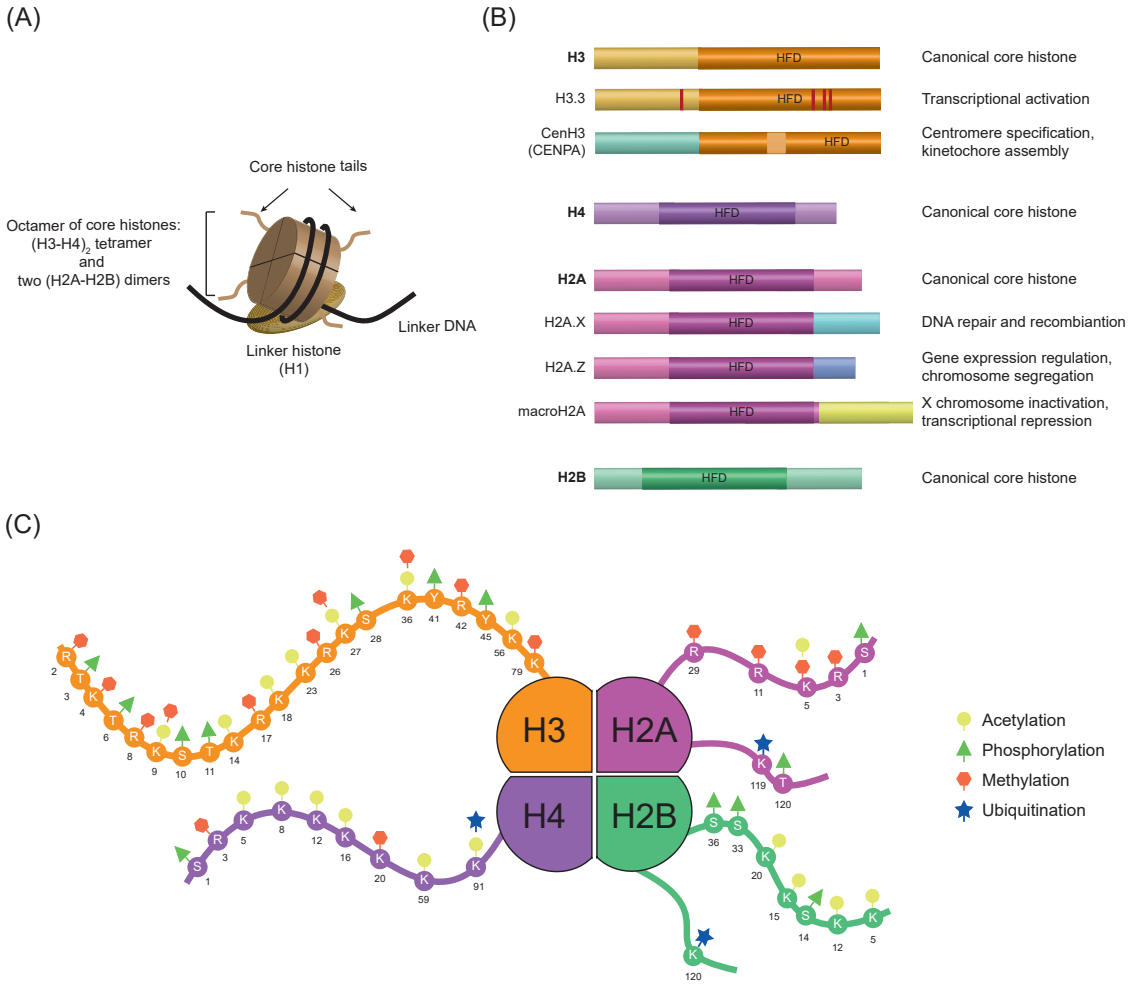



**Figure 1.1: Chromatin-based gene regulation: histone variants and post-translational modifications of histone tails**

(A) Schematic of nucleosome organisation. A typical nucleosome is composed of a histone octamer core, linker DNA and a linker histone H1.

(B) Canonical and variant forms of core histones. Core histones comprise a globular histone-fold domain (HFD) and flexible N- and C-terminal tails. Amino acid residues in the histone H3.3 variant which are different from the canonical histone H3 are highlighted in red. The CENPA variant of histone H3 has a unique N-terminal domain (highlighted in light green), which is different from other core histones. Centromeric localisation of CENPA relies on a small region in its HFD domain which is highlighted in light orange. Variant forms of histone H2A are considerably different from the canonical histone H2A in their C-terminal domains which are highlighted in different colours. MacroH2A has an extended C-terminal domain of an unknown function. The variant forms of histones H2B and H4 are poorly characterised and only the canonical forms are shown. The proposed functions of histone variants are listed on the right. Adapted from Sarma and Reinberg 2005.

(C) Histone tails are subjected to a large variety of post-translational modifications (PTMs). The most common post-translational modifications and their sites in core histones with a proposed role in gene regulation are shown.

are the repeating building units of chromatin (Kornberg and Thomas 1974; Kornberg 1974; Oudet et al. 1975). The nucleosome consists of a 145-147 bp DNA segment wrapped around a histone octamer core, comprising a tetramer of histones $(H3\text{-}H4)_2$ and two dimers of histones H2A-H2B (Figure 1.1A) (Luger et al. 1997; Kornberg and Lorch 1999b; Davey et al. 2002). Neighbouring nucleosomes are separated by linker DNA of variable length (ranging from 20 to 90 bp), depending on species and cell lineage (Szerlong and Hansen 2011). Linker DNA is typically

associated with the linker histone H1 (Noll and Kornberg 1977; Allan et al. 1980; Zhou et al. 1998), which has been proposed to compact nucleosomal arrays into higher-order structures, of which 30 nm chromatin fibre is the most characterised, albeit only *in vitro* (Thoma et al. 1979; Robinson et al. 2006; Robinson and Rhodes 2006). However, the architecture and compaction mechanisms of these secondary chromatin structures, as well as their relevance *in vivo*, have been a matter of active debate in the field for several decades (Widom 1989; Robinson et al. 2006; Razin and Gavrilov 2014). In contrast to the hierarchical model of chromatin folding, recent super-resolution microscopy studies concluded that chromatin exists as highly irregular "bead-on-string" fibres both during interphase and mitosis (Maeshima et al. 2010; Nishino et al. 2012; Maeshima et al. 2014; Ou et al. 2017). Packaging of pliable nucleosomal arrays of variable diameters at different densities was proposed to permit a more dynamic genome organisation, with higher-level compaction achieved via loop extrusion (Fudenberg et al. 2017; Gibcus et al. 2018).

For many years, chromatin has been proposed to play only a structural role in the genome organisation inside the nucleus (Kingston et al. 1996). However, evidence from an increasing number of biochemical and genetic studies led to the realisation that chromatin also plays a fundamental role in gene regulation. Initial support for this idea came from early observations that the degree of chromatin condensation, as evident by the DNA staining intensity, varies across different genomic regions, with some chromosomal segments remaining highly condensed throughout the cell cycle, suggesting that these may represent a "functionally inactive" state of the genome (Heitz 1928; Heitz 1929). Many years later, biochemical studies reported that the nucleosome occupancy at promoters and throughout gene bodies changes upon gene activation independently of replication (Schmid et al. 1992; Adams and Workman 1993; Felsenfeld et al. 1996), with actively transcribed genes displaying a more accessible chromatin configuration as measured by nuclease hypersensitivity (Weintraub and Groudine 1976; Wu et al. 1979; Wu 1980; Levy and Noll 1981; Gross and Garrard 1988). Finally, biochemical isolation of the first chromatin-modifying enzymes (Brownell et al. 1996; Ogryzko et al. 1996; Utley et al. 1998) and chromatin remodelling complexes (Cote et al. 1994; Imbalzano et al. 1994; Kwon et al. 1994) revealed that these activities are often found within classical transcriptional co-activators. Following these initial observations, subsequent advances in chromatin research have made it increasingly clear that chromatin broadly impinges on almost all DNA-templated nuclear processes, including transcription, replication, recombination and DNA repair (reviewed in Groth

et al. 2007; Papamichos-Chronakis and Peterson 2012; House et al. 2014; Hauer and Gasser 2017).

Importantly, direct evidence that nucleosomes act as barriers for transcription came from the *in vitro* transcription assays, which showed that RNA polymerase II is unable to bind core promoters and initiate transcription on chromatinised DNA templates (Knezetic and Luse 1986; Lorch et al. 1987). It was suggested that the nucleosome-imposed transcriptional block may result both from the direct occlusion of regulatory DNA sequences by the nucleosome core and compaction of nucleosomal arrays by unmodified histone tails (Workman and Kingston 1998; Kornberg and Lorch 1999a; An et al. 2002). Therefore, productive initiation of transcription requires nucleosomes at the *cis*-regulatory elements to be evicted or remodelled into a more accessible state and histone tails to be modified to prevent chromatin compaction (Han and Grunstein 1988; Grunstein 1990; Durrin et al. 1991; Grunstein 1997). In line with this idea, addition of transcriptional co-activators with chromatin-modifying and remodelling activities to *in vitro* transcription reactions eliminated nucleosome-mediated inhibition of transcription (Utley et al. 1998; Neely et al. 1999; Kundu et al. 2000; Mizuguchi et al. 2001). Furthermore, histone acetylation was found to be directly linked to increased chromatin accessibility at actively transcribed genomic regions in cells (Sealy and Chalkley 1978; Vidali et al. 1978; Hebbes et al. 1988; Lee et al. 1993; Vettese-Dadey et al. 1996). Importantly, a multitude of post-translational histone modifications associated with gene regulatory elements, and enzymes that place, remove and recognise these modifications have now been characterised (reviewed in Peterson and Laniel 2004; Kouzarides 2007; Ruthenburg et al. 2007; Marmorstein and Trievel 2009). Together, this has suggested that the composition and structure of individual nucleosomes is dynamically regulated and interpreted, and this is fundamental to control of transcription in eukaryotes. Furthermore, higher-order chromatin configuration, including the architecture of nucleosomal arrays and spatial organisation of chromatin within the nucleus, has also been implicated in regulation of eukaryotic transcription (reviewed in Nemeth and Langst 2004; Gilbert and Ramsahoye 2005; Li and Reinberg 2011). In conclusion, it is clear that chromatin imposes significant barriers on all aspects of transcription, which are eliminated through a variety of different mechanisms, including nucleosome remodelling and eviction, histone variant incorporation, histone post-translational modifications and higher-order chromatin structures, which I will briefly discuss in the following sections.

#### 1.1.2.1 Chromatin remodelling

As one of the most stable DNA-protein complexes in the cell (Luger et al. 1997), nucleosomes profoundly limit physical access of the basal transcriptional machinery and associated factors to DNA (Henikoff 2008; Tsompana and Buck 2014). However, this can be overcome by mobilisation of nucleosomes at promoters and distal regulatory elements by chromatin remodelling complexes, which use ATP hydrolysis to disrupt DNA-histone contacts (Clapier and Cairns 2009). Based on their catalytic subunits, chromatin remodellers are generally classified into four evolutionary conserved families - SWI/SNF, ISWI, CHD and INO80, which have versatile functions in regulation of different chromosomal processes (reviewed in Mohrmann and Verrijzer 2005; Corona and Tamkun 2004; Marfella and Imbalzano 2007; Bao and Shen 2007). Importantly, in addition to a conserved helicase-like ATPase domain, chromatin remodellers often incorporate domains which recognise specific histone modifications. Thus, SWI/SNF and RSC complexes contain bromodomains, which bind acetylated histone tails (Hassan et al. 2001; Carey et al. 2006; Awad and Hassan 2008), while CHD complexes contain multiple chromodomains with specificity for H3K4me3 (Pray-Grant et al. 2005; Sims et al. 2005). Furthermore, chromatin remodellers often interact with histone-modifying transcriptional regulators, as exemplified by NuRD complexes, which contain both ATP-dependent chromatin remodelling and histone deacetylase activities (Xue et al. 1998), and functional association between SWI/SNF and SAGA complexes (Sudarsanam 1999; Neely et al. 2002). This suggests that chromatin remodellers and histone-modifying enzymes function in cooperation to regulate transcription in eukaryotes.

Chromatin remodellers can employ different mechanisms to alter chromatin structure, which include changing nucleosome composition (histone dimer eviction or exchange of histone variants), conformation of nucleosomal DNA (nucleosome unwrapping) or nucleosome position (sliding or eviction). A large number of chromatin remodellers have been directly implicated in gene activation (Hirschhorn et al. 1992; Kwon et al. 1994; Armstrong et al. 2002; Simic et al. 2003), as they were shown to promote an accessible chromatin structure at gene regulatory elements, facilitating transcription initiation (Morris et al. 2013; Dieuleveult et al. 2016). One of the first chromatin remodellers shown to promote transcription was the SWI/SNF complex (Happel et al. 1991; Hirschhorn et al. 1992; Muchardt and Yaniv 1993). Interestingly, mammalian SWI/SNF complexes come in a variety of cell type-specific assemblies, which have been proposed to play specialised functions in cellular differentiation during development (Ho and Crabtree

2010). While many chromatin remodellers facilitate transcription, some can also contribute to gene repression, either through assembly of stable nucleosomal arrays at gene promoters, restricting DNA accessibility, or through recruitment of histone modifiers which counteract gene activation. In line with this, chromatin remodelling by CHD3/4, as a part of the NuRD complex, is associated with transcriptional silencing (Denslow and Wade 2007; Ramirez et al. 2012; Basta and Rauchman 2015). Furthermore, members of the ISWI family of chromatin remodellers were shown to prevent antisense transcription at promoters and transcription initiation from cryptic intergenic sites (Whitehouse et al. 2007).

Through their capacity to destabilise nucleosomes, chromatin remodellers were implicated in establishment of nucleosome-depleted regions (NDRs) at promoters and distal regulatory elements, which have been long recognised as a hallmark of active transcription in eukaryotes (Yuan et al. 2005; Mavrich et al. 2008). The absence of nucleosomes has been proposed to contribute to high chromatin accessibility at *cis*-regulatory elements by exposing core promoters and transcription factor binding sites, enabling transcription initiation. Interestingly, accumulating evidence suggests that promoters of highly expressed genes are not completely devoid of nucleosomes. Instead, they are often occupied by highly dynamic nucleosomes which incorporate histone variants destabilising nucleosome structure (Jin et al. 2009; Weiner et al. 2009; Xi et al. 2011; Kubik et al. 2015). This indicates that nucleosomes may adapt a more accessible configuration, which does not interfere with binding of the transcriptional machinery to the promoter DNA, suggesting that nucleosome occupancy and DNA accessibility are not always mutually exclusive (Mieczkowski et al. 2016; Mueller et al. 2017).

Although *in vitro*, transcription elongation is only modestly impeded by nucleosomes (Lorch et al. 1987; Felsenfeld et al. 1996), in cells, it requires additional factors, which facilitate progression of RNA polymerase II through chromatin and re-establish correct nucleosome structure following transcription (Kireeva et al. 2002; Studitsky 2004; Kulaeva et al. 2009). A number of chromatin remodellers and histone chaperones have been implicated in regulation of transcription elongation, based on their co-localisation with RNA polymerase II at sites of active transcription (Simic et al. 2003; Srinivasan 2005; Carey et al. 2006; Murawska et al. 2008; Kuryan et al. 2012). Among them is a histone chaperone FACT (facilitates chromatin transcription), which has been shown to reorganise nucleosomes by directly binding H2A-H2B dimers in an ATP hydrolysis-independent manner (Orphanides et al. 1998; Orphanides et al. 1999; Belotserkovskaya et al. 2003). The

histone-binding activity of FACT has been proposed to contribute to both eviction and reassembly of nucleosomes during transcription (Formosa 2012). In addition, FACT was reported to interact with SWI/SNF and CHD chromatin remodellers, either supporting or counteracting their activities during elongation (Biswas et al. 2006; Takahata et al. 2009). Together, an interplay between nucleosome reorganisation by FACT and remodelling by ATP-dependent enzymes plays a crucial role in promoting transcription elongation, while maintaining ordered chromatin structure.

#### 1.1.2.2 Histone variants

In contrast to the original view that canonical histones represent the universal nucleosome components in eukaryotes (Kornberg and Lorch 1999b), many variant forms of the core histones (apart from histone H4) and the linker histone H1 have been identified (Figure 1.1B) (reviewed in Kamakaka and Biggins 2005; Henikoff and Smith 2015). Some of the histone variants, in particular those of histone H2A, substantially differ from their canonical counterparts due to the extensive alterations in their primary structure or addition of large protein domains. In contrast, H2B and H3 histone variants are typically more similar to canonical histones and have only a few amino acid substitutions (reviewed in Malik and Henikoff 2003). Incorporation of histone variants can confer unique biochemical properties to the nucleosomes. Substitutions in the histone globular domain can impact on intra-nucleosomal histone-histone or histone-DNA interactions, altering nucleosome stability, while differences in the histone tails can change inter-nucleosomal interactions and chromatin compaction, as well as disrupt specific sites for histone post-translational modifications (reviewed in Kamakaka and Biggins 2005). In addition, histone variants can also alter interactions of nucleosomes with histone chaperones and chromatin remodelling complexes, leading to the distinctive timing and/or genomic distribution of their incorporation into chromatin (reviewed in Henikoff and Ahmad 2005). Thus, in striking contrast to canonical histones, incorporation of many histone variants is replication-independent, as they are unable to interact with the canonical histone chaperones (Ahmad and Henikoff 2002; Tagami et al. 2004; Mizuguchi et al. 2004). Furthermore, histone variants were found to mark specific regions of chromosomes, as best exemplified by the CenH3 (CENPA) histone variant, which plays a conserved role in specifying centromere position (Palmer et al. 1991; Malik and Henikoff 2003).

Histone variants H3.3 and H2A.Z, which are highly conserved across all eukaryotes (Malik and Henikoff 2003), have been of a particular interest to the field, because of their poorly understood

but essential role in metazoan development (Faast et al. 2001; Sakai et al. 2009). Both H3.3 and H2A.Z have been previously implicated in gene activation, as they are highly enriched at promoters and distal regulatory elements associated with actively transcribed genes (Ahmad and Henikoff 2002; Chow et al. 2005; Mito et al. 2005; Raisner et al. 2005; Jin et al. 2009; Chen et al. 2013). Incorporation of H3.3 or H2A.Z, individually or in combination, has been reported to decrease the stability of nucleosomes, thereby facilitating productive transcription (Jin and Felsenfeld 2007; Subramanian et al. 2013; Weber et al. 2014). Surprisingly, other studies linked H2A.Z to gene repression both *in vitro* and *in vivo* (Swaminathan et al. 2005; Thakar et al. 2010; Dai et al. 2017). In line with this, in mouse ES cells, H2A.Z is preferentially enriched at the promoters of Polycomb target genes where it contributes to their transcriptional silencing (Creyghton et al. 2008; Surface et al. 2016), potentially via supporting activity and/or recruitment of Polycomb repressive complexes (Hu et al. 2013; Wang et al. 2018c). Together, these observations highlight an essential role of histone variants in regulation of eukaryotic transcription, although the exact mechanisms of how this is achieved require further investigation.

#### 1.1.2.3 Histone modifications

The highly conserved core histone proteins have a similar domain architecture and all share a globular histone-fold domain, around which the DNA is wrapped. Each histone also contains an N-terminal flexible domain, which is highly basic and extends from the surface of the nucleosome. In addition to this, histone H2A has a unique unstructured C-terminal tail (Luger et al. 1997). While the histone-fold domain is essential for the octamer core assembly, the histone tails do not contribute significantly to the structure or stability of nucleosomes, but instead play an essential role in organisation of nucleosomal arrays into higher-order structures. In addition, exposed histone tails are subject to a variety of post-translational modifications (PTMs), including but not limited to, acetylation, methylation, ubiquitination, phosphorylation, sumoylation, ADP-ribosylation, and glycosylation (Figure 1.1C) (Peterson and Laniel 2004; Kouzarides 2007; Zhao and Garcia 2015). Furthermore, adding to the complexity of histone PTMs, recent studies demonstrated that the more structured globular domain of nucleosomes can also be extensively modified (Mersfelder and Parthun 2006; Tropberger and Schneider 2010; Kebede et al. 2014).

Following the initial discovery of histone methylation in calf thymus (Murray 1964), the identity of enzymes that mediate histone modifications remained elusive for more than 30 years. Only in

1996, the first histone acetyltransferase was discovered in *Tetrahymena* as an orthologue of the yeast transcriptional co-activator GCN5 (Brownell et al. 1996). Extensive efforts in the following years have been placed on further characterisation of enzymes that are responsible for other histone modifications, which yielded a detailed account of histone-modifying activities in eukaryotic cells (Sterner and Berger 2000; Zhang and Reinberg 2001; Nowak and Corces 2004; Shilatifard 2006; Nathan et al. 2006; Hassa et al. 2006). Furthermore, most, if not all, histone PTMs were found to be reversible, suggesting that the dynamic nature of histone modifications may underpin high plasticity of transcriptional responses to various internal or external stimuli (Holbert and Marmorstein 2005; Voss and Hager 2013).

Histone PTMs have been proposed to play versatile roles in control of eukaryotic transcription. Histone modifications, which change the residue charge, such as acetylation and phosphorylation, may directly alter the physical properties of individual nucleosomes and nucleosomal arrays (Kingston et al. 1996; Wade et al. 1997; Workman and Kingston 1998). Thus, acetylation of histone tails was shown to disrupt the inter-nucleosome contacts, promoting more open chromatin configuration (Shogren-Knaak et al. 2006). In addition, acetylation of residues in the globular octamer core was found to weaken histone-DNA interactions, reducing the nucleosome stability (Tropberger and Schneider 2010; Cerbo et al. 2014). These observations led to a model, whereby histone acetylation facilitates transcription initiation through increasing chromatin accessibility at *cis*-regulatory elements. However, as histones contain multiple positively charged amino acids, all of which may contribute to intra- and inter- nucleosomal interactions, the extent to which acetylation of a single residue can affect these interactions remains highly debatable. In contrast to their poorly understood structural role, a large body of research has suggested that histone modifications function mainly through recruitment of various effector factors, called readers, including DNA- and histone-modifying enzymes, chromatin remodelling complexes and histone chaperones, which coordinate downstream processes leading to gene activation or silencing (Taverna et al. 2007; Kouzarides 2007). These reader proteins recognise distinct histone modifications through specialised domains, which typically have high affinity for a specific modified histone residue. Histone methylation is recognised by a variety of chromo-like domains, including chromo, tudor and MBD, and structurally unrelated PHD domains, while acetylation is specifically recognised by bromodomains, and phosophorylation is recognised by a family of 14-3-3-domain proteins (reviewed in Bottomley 2004; Cruz et al. 2005; Kouzarides 2007).

Individual histone PTMs or their distinct combinations specify unique chromatin states at different classes of *cis*-regulatory elements and larger genomic domains (The ENCODE Project Consortium 2012). Although the mechanisms behind the locus-specific enrichment of distinct histone PTM combinations are generally poorly understood, this has been proposed to rely on targeting of histone-modifying enzymes to specific genomic sites via DNA-binding transcription factors (Peterson and Laniel 2004; Imhof 2006). In addition, a cross-talk between individual histone modifications may contribute to their coordinated deposition in distinct combinations. In line with this, the activity and/or recruitment of several histone-modifying enzymes was shown to depend on the presence of a second histone modification in the same or adjacent nucleosome (Clements et al. 2003; Nakagawa et al. 2008; Kim et al. 2009b; Kalb et al. 2014). Together, these mechanisms generate a variety of specific patterns of histone modifications in the genome. For example, promoters of transcribed genes display high levels of H3K4me3, while enhancers associated with expressed genes are depleted of H3K4me3 and instead have high levels of H3K4me1 (Barski et al. 2007; Guenther et al. 2007). Histone acetylation is another hallmark of actively transcribed genomic regions, as it is also enriched at promoters and distal regulatory elements of expressed genes (Pokholok et al. 2005; Creyghton et al. 2010; Karmodiya et al. 2012; Taylor et al. 2013). Furthermore, other histone modifications, such as H3K36me3 and H3K79me2/3, show more broad genomic localisation and are placed co-transcriptionally across bodies of highly expressed genes (Krogan et al. 2003; Bannister et al. 2005; Schubeler et al. 2004; Steger et al. 2008).

The multitude of characterised histone PTMs and proteins that specifically recognise them have led to the idea of a "histone code", in which a particular combination of histone PTMs produces a specific functional outcome (Jenuwein and Allis 2001). However, accumulating evidence suggests that different PTMs or their combinations can have a different impact on gene activity, depending on the genomic and cellular context. For instance, H3K9me3 is typically associated with heterochromatin formation and stable gene repression (Nakayama et al. 2001; Schotta et al. 2002; Lehnertz et al. 2003; Martens et al. 2005), but it is also found in the bodies of actively transcribed genes, where it can contribute to their expression (Vakoc et al. 2005). The multifaceted functions of a particular histone modification could rely on the availability of different reader proteins and an interplay with other histone modifications. In conclusion, histone PTMs appear to be differentially interpreted in different cellular contexts, in contrast to the rigid "histone code" model, adding additional plasticity to the chromatin-based regulation of chromosomal processes.

Histone-modifying enzymes have been widely proposed to play fundamental roles in gene regulation, based on severe transcriptional defects which manifest when these proteins are removed (Kouzarides 2007; DesJarlais and Tummino 2016). However, whether histone modifications catalysed by these enzymes directly instruct gene activity still remains a matter of active debate. In fact, histone-modifying enzymes typically reside in large protein complexes which encompass many auxiliary subunits with catalysis-independent activities (DesJarlais and Tummino 2016). Therefore, to directly examine the role of histone modifications in gene regulation, it is necessary to develop approaches for catalytic inactivation of the histone-modifying complexes. However, this will not address a potential contribution of non-histone substrates of these enzymes, which were shown to include transcription factors and co-activators (Sterner and Berger 2000; Ardehali et al. 2017). Resolving this question will require mutating the modification sites in the endogenous histone proteins, which is extremely challenging, if not impossible, in mammals, as they encode a large number of histone genes. Therefore, understanding the role of catalytic activity of histone-modifying complexes in gene regulation still remains a central, yet challenging, question in the field.

Nevertheless, a crucial role of histone-modifying enzymes in control of normal gene expression is indisputable, since loss of their function has been linked to severe developmental disorders and cancer (Chi et al. 2010; Audia and Campbell 2016; Bracken et al. 2019). One of the key players in chromatin-based gene regulation, which is the main focus of this work, is a family of evolutionary conserved Polycomb group (PcG) proteins (Schuettengruber et al. 2017). In mammals, they form two major complexes - Polycomb repressive complex 1 (PRC1) and PRC2, both of which contribute to transcriptional silencing of developmental genes. In the next sections, I will briefly describe histone modifications, which are mediated by the Polycomb complexes.

***H3K27 methylation***

In mammals, addition of mono-, di-, or tri- methylation to lysine 27 of histone H3 is catalysed by the SET domain-containing EZH1/2 (enhancer of zeste homologue 1/2) subunits of PRC2 complexes (Figure 1.2A) (Cao et al. 2002; Kuzmichev et al. 2002; Muller et al. 2002). Importantly, association of EZH1/2 with SUZ12 (suppressor of zeste 12), EED (embryonic ectoderm development) and RbAp46/48 subunits is required for the full PRC2 enzymatic activity (Figure 1.2A) (Kuzmichev et al. 2002; Cao and Zhang 2004; Pasini et al. 2004; Nekrasov et al.

2005), which can be even further potentiated by incorporation of various accessory subunits (Nekrasov et al. 2007; Sarma et al. 2008; Zhang et al. 2011; Choi et al. 2017; Lee et al. 2018; Conway et al. 2018). Interestingly, EED can directly bind the product of PRC2 catalysis, H3K27me3, promoting allosteric stimulation of PRC2 methyltransferase activity (Margueron et al. 2009). Furthermore, the histone modification placed by PRC1, H2AK119ub1, was shown to promote H3K27 tri-methylation by a specific type of PRC2, revealing an intimate cross-talk between the two Polycomb complexes (Kalb et al. 2014). Finally, high density of nucleosomal arrays has also been reported to positively regulate PRC2 catalytic activity (Yuan et al. 2012). In contrast, histone modifications associated with actively transcribed genes, such as H3K4me3 and H3K36me3, were shown to inhibit H3K27 methylation by PRC2 *in vitro* (Schmitges et al. 2011; Yuan et al. 2011).

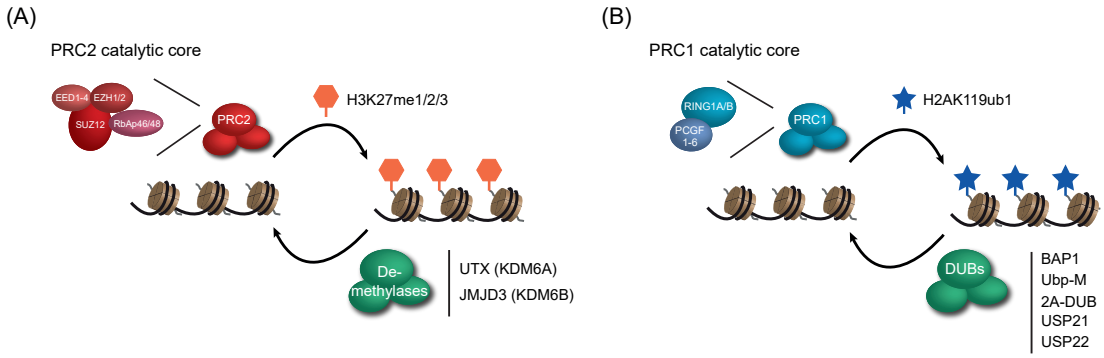



**Figure 1.2: Dynamic control of deposition and removal of H3K27 methylation and H2AK119 mono-ubiquitination**
(A) PRC2 catalyses H3K27 mono-, di- and tri-methylation which are removed by UTX (KDM2A) and JMJD3 (KDM6B) histone demethylases. The composition of PRC2 catalytic core is shown.
(B) PRC1 catalyses H2AK119 mono-ubiquitination which is removed by a number of deubiquitinating enzymes (DUBs), including BAP1, Ubp-M, 2A-DUB, USP21 and USP22. The composition of PRC1 catalytic core is shown.

H3K27me3 was found to be enriched at Polycomb-repressed developmental gene promoters and the inactive X-chromosome in mammals (Plath et al. 2003; Bracken et al. 2006; Boyer et al. 2006; Barski et al. 2007). By analogy with H3K9me3-dependent binding of HP1 leading to heterochromatin formation (Bannister et al. 2001; Danzer and Wallrath 2004), H3K27me3 has been proposed to mediate gene repression through recruitment of CBX-containing canonical PRC1 (Fischle et al. 2003; Wang et al. 2004b), which was shown to compact chromatin and limit activity of chromatin remodellers *in vitro* (Francis et al. 2001; Francis et al. 2004). In agreement with this, a point mutation in histone H3K27 results in a failure to silence classical Polycomb target genes outside of their normal expression domains in *Drosophila* larvae (Pengelly et al. 2013). However, recent studies in mouse ES cells demonstrated that loss of PRC2 causes modest transcriptional

defects, suggesting that Polycomb-mediated gene repression can be also achieved via H3K27me3-independent mechanisms (Riising et al. 2014). Importantly, histone H3K27 methylation is a pervasive histone modification, with approximately 70% of histone H3 being di-methylated, 7% tri-methylated, and 4% mono-methylated in mouse ES cells (Jung et al. 2010; Ferrari et al. 2014). In contrast to the relatively focused enrichment of H3K27me3 at Polycomb-repressed gene promoters, H3K27me2 is placed broadly across transcriptionally silent genomic regions, where it is proposed to protect enhancers from non-specific firing through counteracting H3K27 acetylation (Ferrari et al. 2014). Interestingly, H3K27me1 was shown to accumulate within transcribed gene bodies in a H3K36me3-dependent manner, further facilitating gene expression (Ferrari et al. 2014). These observations indicate that functions of histone H3K27 methylation by PRC2 extend beyond repression of Polycomb target sites, although the mechanisms of how PRC2 pervasive activities contribute to gene regulation throughout the genome require further characterisation in the future.

Methylation of histone H3K27 is a reversible histone modification, with UTX (KDM6A) and JMJD3 (KDM6B) being the main JmjC domain-containing demethylases, which are specific for H3K27me2/me3 (Figure 1.2A) (Agger et al. 2007; Hong et al. 2007). Interestingly, both UTX and JMJD3 were found to associate with MLL H3K4 methyltransferases, potentially as an additional mechanism to counteract PRC2 activity (Cho et al. 2007; Issaeva et al. 2007; Lee et al. 2007; Santa et al. 2007). However, the kinetics of H3K27me3 deposition and removal were shown to be surprisingly slow. Recent studies demonstrated that following DNA replication and concomitant two-fold dilution of parental histones, re-establishment of the parental levels of H3K27me3 on new histones is slow and takes on average more than one cell cycle (Alabert et al. 2015). In addition, a more detailed analysis of the genomic distribution of H3K27me3-modified parental and new histones following S-phase progression revealed that classical highly-enriched Polycomb target sites accumulate H3K27me3 faster than the rest of the genome, yet this still happens slower than for H3K4me3 (Reveron-Gomez et al. 2018). Slow deposition and removal of H3K27me3 has been proposed to filter random fluctuations in activation signals, enabling stable repression of Polycomb target genes, which can be propagated through cell division (Berry et al. 2017).

***H2AK119 mono-ubiquitination***

Ubiquitin is a small 76 amino acid regulatory protein, which can be conjugated to substrates through a three-step process, which includes activation, conjugation and ligation, which are carried

out by E1, E2 and E3 enzymes respectively (Goldstein et al. 1975; Handley et al. 1991; Weake and Workman 2008). First, ubiquitin needs to be activated by an E1 enzyme in an ATP-dependent process, which enables its further conjugation to a cysteine residue in an E2 ubiquitin carrier protein via a reactive thioester bond. The final step of the ubiquitination reaction is catalysed by an E3 ubiquitin ligase, which transfers ubiquitin from the E2 enzyme to a lysine residue in the target protein (Hochstrasser 1996). The vast majority of E3 ubiquitin ligases belong to the RING (really interesting new gene) family (Deshaies and Joazeiro 2009), which is specified by the presence of RING finger protein domain (Freemont et al. 1991) and encoded by more than 600 genes in mammals (Li et al. 2008).

Substrates can be modified with ubiquitin in many ways: addition of one ubiquitin results in mono-ubiquitination, while conjugation of two or more ubiquitin molecules can lead to mono-ubiquitination at multiple sites or poly-ubiquitination, if the ubiquitin itself acts as an acceptor for conjugation (Passmore and Barford 2004). Modification of substrates with poly-ubiquitin chains which are linked via lysine 48 (K48) generally targets proteins for proteasomal degradation. In contrast, mono-ubiquitination or poly-ubiquitination via K63-linked ubiquitin chains has been increasingly recognised as a post-translational modification implicated in regulation of DNA repair, signal transduction, subcellular localization and gene expression. The first protein found to be ubiquitinated in cells was histone H2A, as one of the most abundant proteins in the nucleus (Goldknopf et al. 1975). Histone H2A was shown to be predominantly mono-ubiquitinated at lysine 119 (H2AK119ub1) (Nickel and Davie 1989), and later studies demonstrated that this histone modification is enriched at transcriptionally inactive genomic regions, linking it to gene repression (Wang et al. 2004a; Napoles et al. 2004). In addition to lysine 119, histone H2A can also be mono-ubiquitinated at other sites, including K125/127/129, but these modifications were strongly implicated in control of DNA damage response, rather than gene regulation (Uckelmann and Sixma 2017). Finally, histone H2A variants, H2A.Z and macroH2A.1, were also found to be mono-ubiquitinated at the positions analogous to lysine 119 in histone H2A (Sarcinella et al. 2007; Chu et al. 2006).

In mammals, H2AK119ub1 is catalysed by RING1A (RING1) and RING1B (RNF2) E3 ubiquitin ligases, which are catalytic subunits of PRC1 complexes (Figure 1.2B) (Wang et al. 2004a; Napoles et al. 2004; Endoh et al. 2012). The catalytic core of PRC1 is formed by hetero-dimerisation of RING1A/B with one out of six RING-domain containing PCGF (Polycomb group

ring finger) proteins, which profoundly stimulate inherent RING1A/B E3 ubiquitin ligase activity, despite being incapable of placing H2AK119ub1 on their own (Figure 1.2B) (Wang et al. 2004a; Cao et al. 2005; Buchwald et al. 2006; Taherbhoy et al. 2015). Importantly, the composition of PRC1 catalytic core was shown to define its E3 ubiquitin ligase activity, with PCGF2/4-RING1A/B dimers reported to be less active than PCGF1/3/5/6-RING1A/B dimers in a number of *in vitro* ubiquitination assays, including those that use physiologically relevant nucleosome substrates (Taherbhoy et al. 2015; Rose et al. 2016). Furthermore, PRC1 catalytic activity was found to be further stimulated by incorporation of RYBP/YAF2 auxiliary subunits, which contain a ubiquitin-binding domain (UBD) (Arrigoni et al. 2006; Gao et al. 2012; Rose et al. 2016). In addition to PRC1, 2A-HUB, a component of the co-repressor N-CoR/HDAC1/3 complex, was identified as another mammalian RING-type E3 ubiquitin ligase that can catalyse H2AK119ub1 (Zhou et al. 2008). 2A-HUB was found to mediate H2AK119ub1-dependent repression of a group of chemokine genes in macrophages, potentially through preventing FACT recruitment and blocking transcription elongation. However, in contrast to knockdown of RING1B (Wang et al. 2004a), 2A-HUB depletion did not affect total H2AK119ub1 levels, suggesting that its function is limited to a small subset of target sites.

Histone H2AK119 mono-ubiquitination is an abundant histone modification, which comprises 5-15% of total histone H2A in mammalian cells (Goldknopf et al. 1975; Albright et al. 1979; Matsui et al. 1979). However, very little is known about the distribution of this histone modification outside of Polycomb-repressed developmental genes or the inactive X chromosome, where it was found to be enriched, usually together with H3K27me3 (Napoles et al. 2004; Cao et al. 2005; Boyer et al. 2006; Stock et al. 2007). Interestingly, a recent study in *Drosophila* found that low levels of H2AK118ub1 are deposited outside of classical Polycomb target sites, including active genes and intergenic regions (Lee et al. 2015; Kahn et al. 2016). However, whether widespread H2AK119ub1 contributes to regulation of gene expression or other chromosomal processes remains poorly understood. More importantly, how H2AK119ub1 may contribute to gene repression in general also remains elusive, with some studies arguing that H2AK119ub1 is dispensable for Polycomb-dependent transcriptional silencing (Illingworth et al. 2015; Pengelly et al. 2015). However, recent work from other groups suggests that H2AK119ub1 plays a central role in Polycomb target gene repression (Endoh et al. 2012; Tsuboi et al. 2018). The proposed mechanisms for H2AK119ub1-mediated inhibition of transcription include restraining paused

RNA polymerase II or preventing association of a histone chaperone FACT (Stock et al. 2007; Zhou et al. 2008), as well as recruitment of reader proteins, which are yet to be fully characterised (Richly et al. 2010; Qin et al. 2015b; Zhang et al. 2017). These proposed mechanisms will be discussed in more detail in section 1.3.2.

Finally, mono-ubiquitination of H2AK119 is a highly dynamic histone modification, and its global levels are regulated during the cell cycle, with metaphase chromosomes reported to be completely devoid of H2AK119ub1 (Matsui et al. 1979; Mueller et al. 1985; Joo et al. 2007; Nakagawa et al. 2008). In line with this, in mammals, a number of H2AK119ub1-specific deubiquitinating enzymes (DUBs) have been characterised, including 2A-DUB, USP16 (Ubp-M), USP21, USP22 and BAP1 (Figure 1.2B) (Zhu et al. 2007; Joo et al. 2007; Nakagawa et al. 2008; Yang et al. 2014; Zhao et al. 2008b; Scheuermann et al. 2010). Recent studies linked some of these DUBs to gene activation, with UBP-M shown to promote transcription of *Hox* genes (Joo et al. 2007) and 2A-DUB found to mediate androgen/AR-induced gene activation, in association with p300 histone acetyltransferase (Zhu et al. 2007). Interestingly, in contrast to this, knockout of BAP1 deubiquitinase in *Drosophila* was reported to cause loss of *Hox* gene silencing resulting in Polycomb homeotic transformations (Scheuermann et al. 2010). Nevertheless, it still remains poorly understood how H2AK119ub1 DUBs are recruited to specific genomic loci, as well as why they may be required for gene activation in some cases and repression in others.

#### 1.1.2.4 Higher-order chromatin structures

In addition to the composition and structure of individual nucleosomes, higher-order chromatin organisation can also contribute to regulation of gene expression. Local chromatin compaction, which relies on stable inter-nucleosome interactions, can render gene promoters and distal regulatory elements inaccessible to the basal transcriptional machinery and associated factors. I have already mentioned the role of positively-charged histone tails and their acetylation in modulating the interactions between adjacent nucleosomes (Dorigo et al. 2003; Shogren-Knaak et al. 2006; Kan et al. 2008). In addition to this, association of the linker histone H1 was also shown to promote inter-nucleosome interactions and chromatin condensation, supporting formation of repressive chromatin environments (Thoma et al. 1979; Allan et al. 1981). However, recent studies demonstrated that histone H1 and its numerous isoforms are not distributed uniformly throughout the genome, but instead are enriched at distinct groups of genes, where they can contribute to both

activation and repression of transcription (Shen and Gorovsky 1996; Clausell et al. 2009; Cao et al. 2013; Izzo et al. 2013). Together, these observations suggest that the linker histones may be involved in regulation of local chromatin structure and gene activity at specific genomic loci, rather than global chromatin architecture.

Furthermore, binding of non-histone architectural proteins can also play an important role in regulation of higher-order chromatin structure and gene activity (Li and Reinberg 2011). High mobility group (HMG) proteins were shown to contribute to productive transcription through counteracting chromatin condensation by linker histones (Rochman et al. 2009). Other examples of chromatin architectural factors implicated in gene regulation include HP1 and Polycomb group (PcG) proteins, which bind chromatin at transcriptionally inactive genomic regions through recognition of H3K9me3 or H3K27me3 respectively (Bannister et al. 2001; Lachner et al. 2001; Nakayama et al. 2001; Cao et al. 2002; Min et al. 2003). Binding of these proteins has been shown to compact nucleosomal arrays *in vitro* (Francis et al. 2004; Grau et al. 2011; Azzaz et al. 2014) and support more compact chromatin organisation in cells (Verschure et al. 2005; Boettiger et al. 2016), potentially contributing to gene repression (Eissenberg et al. 1990; Danzer and Wallrath 2004; Lau et al. 2017; Ogiyama et al. 2018). In addition to this, PcG proteins were also shown to mediate long-range chromatin interactions between Polycomb target gene promoters (Schoenfelder et al. 2015; Kundu et al. 2017; Eagen et al. 2017).

Finally, initiation of eukaryotic transcription relies on communication between promoters and enhancers, which are sometimes located hundreds of kilobases away from each other (Spitz and Furlong 2012; Shlyueva et al. 2014). Long-range chromatin interactions between these *cis*-regulatory elements have been proposed to play a central role in the enhancer-dependent gene activation (Marsman and Horsfield 2012; Bickmore 2019). Although the exact mechanisms underlying these long-distance chromatin contacts remain speculative, it was suggested that they are mainly driven by the concerted activity of cohesin and architectural proteins, of which CTCF is the most characterised in mammals (Splinter et al. 2006; Parelho et al. 2008; Kagey et al. 2010; Seitan et al. 2013). However, recent studies found that loss of cohesin or CTCF, while disrupting the higher-order chromatin structures, results in modest transcriptional defects, indicating that the contribution of these processes to gene regulation is perhaps not as extensive as has been previously proposed (Nora et al. 2017; Rao et al. 2017).

## 1.2 Polycomb group proteins and their role in gene repression

### 1.2.1 Discovery of Polycomb group proteins

Metameric segmentation of the anterior-posterior body axis is a hallmark of embryonic development in the bilaterian animals (Tautz 2004; Blair 2008). *Drosophila* embryogenesis has originally served as the major paradigm for studying developmental mechanisms of segmentation, followed by later studies in vertebrates, which uncovered many common features (Kimelman and Martin 2011). The identity of individual segments is determined early in development through the selective expression of homeotic *Hox* genes, which encode homeodomain-containing transcription factors (Garcia-Fernandez 2005; Mallo et al. 2010). In *Drosophila*, *Hox* genes are clustered into two main complexes, antennapedia (ANT-C) and bithorax (BX-C), and the order of genes in these clusters mirrors the order of spatially restricted domains, in which they are expressed along the anterior-posterior axis (Lewis 1978; Kaufman et al. 1980; Ingham 1985b). The boundaries of the *Hox* gene expression domains are established in early embryogenesis by the graded expression of segmentation proteins, including the products of the *bcd*, *gap* and pair-rule genes (Akam 1987; Lehmann and Nüsslein-Volhard 1987; Moazed and OFarrell 1992). Later in development, the patterns of *Hox* gene expression are stably maintained in the absence of the initiating transcription factors, suggesting that additional proteins are responsible for their propagation. Misregulation of *Hox* gene expression leads to homeotic defects, which are characterised by transformation of one body segment into a segment that resembles another segment type. Most striking examples of homeotic transformations include flies with a mutation in the *Antennapedia* gene, which develop legs instead of antennae (Struhl 1981b), or flies with four wings, in which disruption of the *Ultrabithorax* gene causes transformation of halteres into wings (Morata and Kerridge 1982).

Since the homeotic phenotypes are easily recognisable in *Drosophila*, this enabled genetic screens for factors involved in the maintenance of *Hox* gene expression patterns, leading to the discovery of two antagonistic gene families - Polycomb group (PcG) and Trithorax group (TrxG). The first gene of the PcG family, *Polycomb* (*Pc*), was described by Pamela Lewis based on a dominant mutation which produced ectopic sex combs on the legs of adult male flies (Lewis 1947). Thirty years later, Ed Lewis demonstrated that homozygous *Pc* fly mutants die as late embryos and exhibit classical homeotic defects characterised by transformation of anterior embryonic segments into more posterior segments due to the ectopic expression of *Hox* genes (Lewis 1978). In the following

years, genetic analysis identified approximately 40 additional PcG genes, mutations of which either produced similar phenotypes to the loss-of-function mutations in *Pc* or further enhanced the homeotic defects in the *Pc* mutants (Struhl 1981a; Duncan 1982; Ingham 1984; Jürgens 1985; Ayala Alonso et al. 2007).

Subsequent analysis of PcG mutants revealed that, while early embryos have wild-type patterns of *Hox* gene expression, later in development, *Hox* genes become ectopically expressed across all embryonic segments (Struhl and Akam 1985; Simon et al. 1992). These observations led to the idea that PcG proteins function as negative regulators of *Hox* genes by maintaining their silent state in the segments where they were not expressed during early development. In contrast to this, mutations in TrxG genes were found to cause posterior-to-anterior homeotic transformations, due to the reduced expression of *Hox* genes in posterior segments, suggesting that TrxG proteins positively regulate expression of *Hox* genes, antagonising activity of PcG proteins (Ingham 1983; Ingham 1985a; Shearn et al. 1987; Mazo et al. 1990; Tripoulas et al. 1994; Klymenko and Müller 2004). Importantly, many of the identified PcG and TrxG proteins were found to be highly conserved across *Metazoa*, plants and fungi, highlighting their essential role in gene regulation during development of multicellular organisms (Schuettengruber et al. 2017).

Together, PcG and TrxG proteins have been proposed to constitute a cellular memory system, which is responsible for stable inheritance of the established transcriptional states during development (Ringrose and Paro 2004; Steffen and Ringrose 2014). To achieve this, PcG and TrxG proteins bind to the same specialised *cis*-regulatory DNA sequences at the promoters or enhancers of their target genes, called Polycomb response elements (PREs) or sometimes Polycomb/Trithorax response elements (PRE/TREs), due to their dual functionality (Zink and Paro 1989; Chan et al. 1994; Orlando et al. 1998; Tillib et al. 1999). Interestingly, despite high evolutionary conservation of PcG and TrxG proteins and their target genes across *Metazoa*, PREs from different species do not share extensive sequence homology (Bauer et al. 2015). This suggests that the mechanisms behind PcG/TrxG targeting may have undergone more rapid diversification during evolution than the components of these systems.

Extensive biochemical characterisation of PcG and TrxG proteins over the last two decades has revealed that these proteins assemble into large multi-subunit complexes, which modify chromatin to support transcriptional repression or activation of their target genes (reviewed in

Grimaud et al. 2006). Three major classes of PcG protein assemblies, which are conserved from *Drosophila* to mammals, have now been identified - Polycomb repressive complex 1 (PRC1), Polycomb repressive complex 2 (PRC2) and Polycomb repressive deubiquitinase complex (PR-DUB) (reviewed in Chittock et al. 2017). In the next sections, I will describe in more detail the architecture and molecular activities of these Polycomb repressive complexes, with a major focus on the mammalian Polycomb system.

### 1.2.2 PRC1 in mammals

Understanding the mechanisms of Polycomb-mediated gene repression required isolation of functional PcG protein complexes. PRC1 is the first Polycomb complex, which has been biochemically purified from *Drosophila* embryos (Shao et al. 1999; Saurin et al. 2001). The functional core of *Drosophila* PRC1 is composed of four PcG proteins - PC (Polycomb), PH (Polyhomeotic), PSC (Posterior sex combs), and dRING/SCE (sex combs extra) (Figure 1.3, *left*). The reconstituted PRC1 core complex was found to be capable of inhibiting remodelling of nucleosomal arrays and transcription *in vitro* (Shao et al. 1999; Francis et al. 2001; King et al. 2002). Importantly, the overall composition of the PRC1 core complex is conserved throughout evolution, as similar PRC1 assemblies, containing the orthologues of *Drosophila* PC, PH, PSC and dRING subunits, were later purified from mammalian cells (Levine et al. 2002; Vandamme et al. 2011). However, individual PRC1 core subunits have undergone extensive diversification during vertebrate evolution, with mammalian genomes encoding two orthologues of dRING (RING1A/RING1 and RING1B/RNF2), five PC orthologues (CBX2/4/6/7/8), three genes for PH (PHC1-3) and six orthologues of PSC (PCGF1-6) (Schuettengruber et al. 2017).

PRC1 functions as an E3 ubiquitin ligase that catalyses histone H2A mono-ubiquitination (H2AK119ub1) (Wang et al. 2004a; Cao et al. 2005). The catalytic core of PRC1 is composed of a dimer between RING1A/B and one of six PCGF subunits, with RING1A/B being sufficient to support histone H2A mono-ubiquitination and PCGF subunits further stimulating RING1A/B catalytic activity (Wang et al. 2004a; Buchwald et al. 2006; Taherbhoy et al. 2015). Both RING1A/B and PCGF proteins contain an N-terminal RING (really interesting new gene) finger domain, through which they heterodimerise (Satijn and Otte 1999; Buchwald et al. 2006; Li et al. 2006), and a C-terminal RAWUL (RING finger and WD40 ubiquitin-like) domain, through which they interact with accessory subunits (Sanchez-Pulido et al. 2008; Bezsonova et al. 2009).

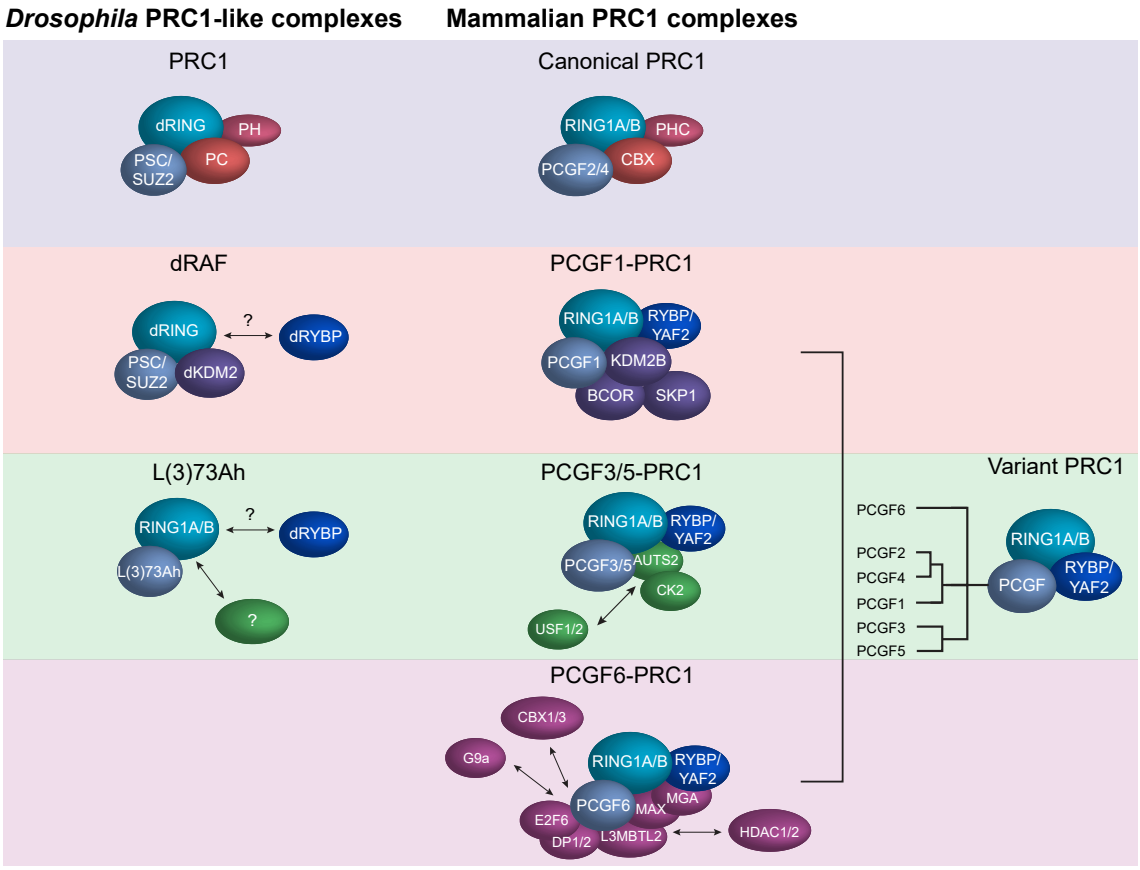



**Figure 1.3: Diversity of PRC1 complexes in *Drosophila* and mammals**
*Left:* A complex comprising four core subunits, PC (Polycomb), PH (Polyhomeotic), PSC (Posterior sex combs) or SUZ2 (Suppressor of zeste 2), and dRING, is traditionally considered to be the only PRC1 complex in *Drosophila*. However, two additional *Drosophila* PRC1-like complexes have been recently identified - dRAF and L(3)73Ah. Both of these complexes lack PC and PH subunits, and may incorporate dRYBP instead, although this interaction has been poorly characterised.
*Right:* In mammals, PRC1 complexes are generally classified into canonical PRC1 and variant PRC1, based on the mutually exclusive incorporation of CBX and RYBP/YAF2 subunits. Canonical PRC1 is the functional orthologue of the *Drosophila* PRC1 complex. Variant PCGF1-PRC1 is closely related to the *Drosophila* dRAF complex, as both of them include KDM2 histone demethylase subunits, and potentially RYBP. Variant PCGF3/5-PRC1 is likely to be homologous to L(3)73Ah-containing PRC1-like complex in *Drosophila*, based on the high sequence similarity between PCGF3/5 and L(3)73Ah proteins. PCGF6-PRC1 is the most divergent variant PRC1 complex which is not closely related to any PRC1-like complex in *Drosophila*. PCGF2 and PCGF4 can also assemble into variant PRC1 complexes, which contain RYBP/YAF2, but surprisingly, no other auxiliary subunits. Arrows indicate non-obligatory auxiliary PRC1 subunits, which were found to associate with the complex more dynamically.

Thus, RAWUL domains of RING1A/B mediate interactions with either CBX (CBX2/4/6/7/8) or RYBP/YAF2 subunits, which are mutually exclusive (Wang et al. 2010). Furthermore, RAWUL domains of PCGF proteins enable selective binding of a variety of auxiliary subunits, which may confer different level of catalytic activity, targeting modalities and regulatory activities to distinct PRC1 protein assemblies (Gao et al. 2012).

Historically, mammalian PRC1 complexes have been classified into two major classes - canonical PRC1 (cPRC1) and non-canonical or variant PRC1 (vPRC1) (Blackledge et al. 2015;

Schuettengruber et al. 2017; Vidal 2019). Canonical PRC1 complexes, which are considered to be the functional orthologues of *Drosophila* PRC1, are defined by the presence of CBX subunits. In contrast, variant PRC1 complexes lack CBX subunits and instead incorporate RYBP or its paralogue YAF2. In the following sections, I will discuss the main differences in the composition, biochemical activities and functions in gene regulation between the two distinct PRC1 classes.

#### 1.2.2.1 Canonical PRC1

Canonical PRC1 is distinguished from variant PRC1 by the incorporation of CBX and PHC subunits. In mammals, PCGF2 (MEL18) and PCGF4 (BMI1) proteins were found to be the only PSC orthologues which can associate with CBX and PHC proteins to form canonical PRC1 complexes (Figure 1.3, *right*) (Gao et al. 2012; Hauri et al. 2016). The unique capacity of PCGF2/4 to bind PHC proteins is mediated by a conserved sequence in their RAWUL domains (Gray et al. 2016), yet how the specificity of PCGF2/4 binding to CBX subunits is achieved still remains poorly understood. Although previous studies have hinted that PCGF2 and PCGF4 may have different affinities for individual paralogues of CBX and PHC subunits (Morey et al. 2012; Morey et al. 2015), PCGF2- and PCGF4-PRC1 complexes appear to have largely identical subunit composition and biochemical properties (Gao et al. 2012; Taherbhoy et al. 2015). Interestingly, PCGF2/4-RING1B heterodimers, which form the catalytic core of canonical PRC1, were shown to have weak E3 ubiquitin ligase activity *in vitro* (Taherbhoy et al. 2015; Rose et al. 2016). Consistent with this, artificial tethering of PCGF4 to an ectopic genomic locus does not result in efficient deposition of H2AK119ub1, suggesting that PCGF2/4-PRC1 complexes may also have limited enzymatic activity *in vivo*. Nevertheless, how much canonical PRC1 complexes contribute to deposition of H2AK119ub1 in cells still remains poorly understood.

In mammals, canonical PRC1 complexes incorporate one of five CBX subunits (CBX2/4/6/7/8), all of which share two conserved domains - an N-terminal chromodomain (chromatin organisation modifier), which can specifically recognise H3K27me3 (Paro and Hogness 1991; Eissenberg 2001; Fischle et al. 2003; Min et al. 2003), and a C-terminal Polycomb repressor (PcR) box domain, which is responsible for binding of CBX proteins to RING1A/B (Schoorlemmer et al. 1997). Recognition of H3K27me3 by CBX subunits has been proposed to drive PRC1 chromatin binding at target sites through a hierarchical recruitment mechanism (Wang et al. 2004b). In agreement with this idea, partial or complete removal of PRC2 results in marked reductions of stable PRC1

occupancy at target sites (Wang et al. 2004b; Morey et al. 2013; Tavares et al. 2012).

Interestingly, despite their high structural similarity, chromodomains of CBX subunits were reported to have different binding preferences for methylated histones, with some of them showing equally high affinity for both H3K27me3 and H3K9me3 *in vitro* (Bernstein et al. 2006b; Kaustov et al. 2010). In addition, chromodomains of some CBX paralogues were also shown to bind RNA and potentially DNA (Bernstein et al. 2006b; Yap et al. 2010; Zhen et al. 2016). In line with their specialised binding activities, individual CBX paralogues were reported to have different dynamics of chromatin binding, occupy unique genomic sites and show distinct patterns of nuclear localisation (Vincenz and Kerppola 2008; Morey et al. 2012; Sowpati et al. 2015; Zhen et al. 2016). Furthermore, single-particle tracking of distinct CBX proteins revealed that while stable chromatin association of CBX7 and CBX8 heavily relies on H3K27me3, the chromatin-bound fractions of CBX2, CBX4 and CBX6 proteins remain largely unchanged in the absence of PRC2 activity (Zhen et al. 2016). Taken together, these observations suggest that multifaceted binding preferences of individual CBX subunits may confer distinct targeting specificity to the individual assemblies of canonical PRC1 complexes.

Further support for the functional specialisation of individual CBX subunits comes from their differential expression patterns during development. In line with this, recent studies found a major switch in the relative abundance of CBX proteins upon differentiation of pluripotent cells (Morey et al. 2012; OLoghlen et al. 2012; Kloet et al. 2016). While CBX7 represents the predominant CBX subunit of canonical PRC1 in mouse ES cells, it becomes substituted by CBX2/4 upon embryonic body differentiation and CBX6 in neural progenitor cells (Morey et al. 2012; Kloet et al. 2016). Interestingly, these observations suggested a high degree of antagonism in expression of individual CBX subunits. In agreement with this, CBX7-containing PRC1 complexes were shown to repress expression of other *Cbx* genes in mouse ES cells (Morey et al. 2012; OLoghlen et al. 2012). Importantly, this was accompanied by the PRC1-dependent repression of the *Cbx7* gene in differentiated cells, highlighting the intricate Polycomb autoregulatory mechanism, which controls the subunit composition of canonical PRC1 complexes during development.

In addition to their role in canonical PRC1 targeting, CBX subunits have been proposed to directly contribute to Polycomb-dependent gene repression by mediating local chromatin compaction. This is based on numerous observations that canonical PRC1 can bind, compact and bridge

chromatin templates *in vitro*, limiting access and activity of SWI/SNF chromatin remodelling complexes (Shao et al. 1999; Francis et al. 2001; Francis et al. 2004; Lavigne et al. 2004; Grau et al. 2011; Beh et al. 2012). Chromatin compaction by canonical PRC1 appears to be independent of PRC1 histone-modifying activity, and instead requires an intrinsically disordered protein domain with high positive charge (Francis et al. 2004; Grau et al. 2011). In *Drosophila*, this domain is accommodated in the PSC protein, which was shown to be sufficient for chromatin compaction *in vitro* (Francis et al. 2004; King et al. 2005; Grau et al. 2011). In mammals, only CBX2 contains an analogous highly basic and disordered protein domain and has the capacity to drive local chromatin compaction (Grau et al. 2011; Lau et al. 2017). Remarkably, the same positively charged domain was shown to drive phase separation of CBX2 proteins *in vitro* and formation of bright CBX2 nuclear foci *in vivo* (Tatavosian et al. 2018; Plys et al. 2019). Importantly, mutations disrupting positively charged residues in the CBX2 compaction domain cause classical PcG anterior-to-posterior transformations of the axial skeleton in mice (Lau et al. 2017). Together, these observations suggest that the CBX2-driven chromatin compaction may play essential roles in Polycomb-mediated gene silencing and body plan specification in mammals. However, given the restricted pattern of CBX2 expression during development and mild homeotic transformations in CBX2-deficient mice, how much chromatin re-organisation by CBX2 contributes to Polycomb target gene repression still remains controversial.

In addition to local chromatin compaction, canonical PRC1 complexes were also implicated in higher-order chromatin folding of individual Polycomb-repressed genomic regions as well as bridging of distant Polycomb target sites (Bickmore et al. 2004; Eskeland et al. 2010b; Schoenfelder et al. 2015; Wani et al. 2016; Boettiger et al. 2016; Kundu et al. 2017; Eagen et al. 2017; Ogiyama et al. 2018). This is based on multiple lines of evidence coming from classical DNA FISH followed by microscopy, super-resolution imaging and chromosome conformation capture (3C) techniques, which demonstrated that Polycomb-occupied genomic loci are organised into compact self-interacting chromatin domains both in *Drosophila* and mammals. These domains often form around clusters of proximal Polycomb target genes, such as *Hox* gene clusters (Bickmore et al. 2004; Eskeland et al. 2010b; Noordermeer et al. 2011), but can also involve long-range interactions between distant Polycomb target sites (Schoenfelder et al. 2015; Bonev et al. 2017). Based on their size, density of interactions and boundary elements, Polycomb-compacted chromatin domains appear to be distinct from extensively characterised topologically associating

domains (TADs), suggesting that they are established via alternative mechanisms (Kundu et al. 2017; Bonev et al. 2017). Indeed, higher-order organisation of Polycomb chromatin domains was found to be driven by PHC subunits of canonical PRC1 complexes (Wani et al. 2016; Boettiger et al. 2016; Kundu et al. 2017). Mechanistically, this is achieved through association of multiple PHC proteins via their conserved C-terminal SAM domains, which were shown to self-polymerise *in vitro* (Kim et al. 2002; Robinson et al. 2012; Isono et al. 2013). Further supporting these observations, PcG proteins were found to form discrete nuclear foci, called Polycomb bodies, which were proposed to reflect high-level occupancy of Polycomb factors at Polycomb-repressed chromatin domains (Saurin et al. 1998; Buchenau et al. 1998; Ren et al. 2008; Cheutin and Cavalli 2012). Importantly, formation of these foci was also shown to rely on PHC subunits (Isono et al. 2013; Wani et al. 2016). Finally, the Polycomb-compacted chromatin domains undergo global decondensation upon gene activation during differentiation, suggesting that the higher-order chromatin folding may be directly linked to transcriptional silencing, yet whether it is the cause or the consequence of gene repression still remains speculative (Noordermeer et al. 2011; Bonev et al. 2017; Kundu et al. 2017).

Despite the detailed characterisation of their biochemical activities, the exact contribution of individual subunits of canonical PRC1 complexes to Polycomb-mediated gene silencing still remains incompletely understood. While mice deficient for RING1A and RING1B cease development at the two-cell stage (Posfai et al. 2012), mice which lack individual canonical PRC1 subunits can survive until birth (Lugt et al. 1994; Akasaka et al. 1996; Takihara et al. 1997; Core et al. 1997; Forzati et al. 2012; Liu et al. 2013). However, they still exhibit pronounced developmental defects, including retarded growth, homeotic transformations of the axial skeleton and gross malformations of limbs. Interestingly, the type and severity of the developmental abnormalities considerably varies between knockout models for different subunits or even paralogues of the same subunit, suggesting that they may play non-overlapping roles during development. For example, in addition to classical homeotic transformations, CBX2-null mice display a unique sex reversal phenotype, whereby male knockout mice develop female gonads, while female knockout mice do not develop ovaries (Core et al. 1997; Katoh-Fukui et al. 1998). Furthermore, in addition to posterior transformations of the axial skeleton, which are observed both in PCGF2- and PCGF4-deficient mice, loss of PCGF4 also causes neurological defects (Lugt et al. 1994; Akasaka et al. 1996). Strikingly, while single PCGF2 and PCGF4

knockout mice are viable, the combined deletion of PCGF2 and PCGF4 causes embryonic lethality at day 9.5 (Akasaka et al. 2001). This is similar to more severe developmental defects in double PHC1/PHC2 knockout mice, which also die at E9.5, as compared to single knockout of these subunits (Takihara et al. 1997; Isono et al. 2005). Finally, in addition to early embryonic development, canonical PRC1 complexes were also shown to play fundamental roles in maintaining adult stem cell identity (reviewed in Sauvageau and Sauvageau 2010; Luis et al. 2012; Vidal and Starowicz 2017; Carlo et al. 2018).

#### 1.2.2.2 Variant PRC1

In addition to the extensively characterised canonical PRC1, a number of diverse PRC1 complexes, known as non-canonical or variant, have been recently identified in mammals (Figure 1.3, *right*) (reviewed in Turner and Bracken 2013; Bajusz et al. 2018). Variant PRC1 complexes are defined by the presence of the RYBP (or its paralogue YAF2) subunit, which is incorporated into the complex in place of CBX (Wang et al. 2010; Gao et al. 2012; Tavares et al. 2012). In mammals, all six PCGF paralogues can assemble into variant PRC1 complexes (PCGF1/2/3/4/5/6-PRC1), which have unique composition of auxiliary subunits, specified by the identity of the PCGF protein (Gao et al. 2012; Hauri et al. 2016). Furthermore, as already mentioned, the inherent E3 ubiquitin ligase activity of distinct PRC1 assemblies also depends on the PCGF component of the catalytic core. Thus, PCGF-RING1A/B dimers containing PCGF proteins exclusively found in variant PRC1 (PCGF1/3/5/6) were shown to be more active at catalysing H2AK119ub1 *in vitro* than those containing PCGF2/4 (Rose et al. 2016). Importantly, RYBP can further stimulate the catalytic activity of all PCGF-RING1B dimers *in vitro*, and this was shown to contribute to shaping normal H2AK119ub1 levels at Polycomb target sites in mouse ES cells (Gao et al. 2012; Tavares et al. 2012; Rose et al. 2016). Further supporting these observations, artificial tethering of variant PRC1 complexes to an ectopic genomic site was shown to drive efficient accumulation of H2AK119ub1 (Blackledge et al. 2014). However, the exact contribution of individual variant PRC1 complexes to deposition of H2AK119ub1 in the genome still has not been systematically dissected.

Since variant PRC1 complexes lack CBX subunits, their chromatin binding is unlikely to rely on H3K27me3, consistent with largely unperturbed RYBP and H2AK119ub1 levels at Polycomb target sites in PRC2-null mouse ES cells (Tavares et al. 2012). It has been originally speculated that RYBP/YAF2 may function as an alternative targeting module for variant PRC1 complexes.

In agreement with this possibility, RYBP was shown to interact *in vitro* with YY1, which is the mammalian orthologue of *Drosophila* Pho proteins (Brown et al. 1998; Garcia et al. 1999; Brown 2003). Pho is a DNA-binding subunit of the *Drosophila*-specific PcG complex PhoRC (Pho repressive complex), which was shown to contribute to recruitment of PRC1 and PRC2 to PREs (Klymenko et al. 2006; Frey et al. 2016). However, the role of Pho/YY1 in Polycomb-mediated gene repression appears to be not conserved in mammals, as YY1 does not associate with Polycomb repressive complexes (Kloet et al. 2016; Hauri et al. 2016) and does not localise at Polycomb target sites (Mendenhall et al. 2010). Moreover, RYBP itself was found to only modestly contribute to stable chromatin occupancy of PRC1 and Polycomb target gene repression in mouse ES cells (Hisada et al. 2012; Rose et al. 2016), suggesting the existence of alternative RYBP-independent mechanisms. In agreement with this, distinct variant PRC1 complexes incorporate a variety of unique DNA-binding modules and transcriptional regulators, which I will discuss in more detail below.

PCGF1-PRC1 was originally characterised as BCOR complex, comprising a BCOR transcriptional co-repressor, a KDM2B histone demethylase, and a SKP1 E3 ubiquitin ligase, in addition to PGCF1, RING1A/B and RYBP (Gearhart et al. 2006). This complex is closely related to the dRAF non-canonical PRC1 complex in *Drosophila*, which also contains a KDM2 histone demethylase subunit (Lagarou et al. 2008) (Figure 1.3). In mammals, there are two KDM2 proteins, KDM2A and KDM2B, both of which can catalyse removal of H3K36me1/2 through their Jumonji C (JmjC) demethylase domains (Tsukada et al. 2005; He et al. 2008), but only KDM2B was found to interact with PCGF1-PRC1 (Farcas et al. 2012; Wu et al. 2013; He et al. 2013a; Hauri et al. 2016). However, the contribution of the demethylase activity of KDM2 proteins to their function in gene regulation is poorly understood. Some studies proposed that H3K36me1/2 demethylation by KDM2A/B may support a transcriptionally permissive state at their target gene promoters (Blackledge et al. 2010; Blackledge and Klose 2011; Liang et al. 2012), since H3K36me2 was reported to inhibit transcription initiation (Carrozza et al. 2005; Li et al. 2009a). In contrast to this, work from other groups demonstrated that H3K36me2/3 antagonises histone H3K27 tri-methylation by PRC2 both *in vitro* (Yuan et al. 2011; Schmitges et al. 2011) and *in vivo* (Dorighi and Tamkun 2013; Popovic et al. 2014; Lu et al. 2016; Streubel et al. 2018), suggesting that KDM2B may indirectly promote PRC2 catalytic activity at target sites by catalysing removal of H3K36me2. However, recent findings from our lab revealed that

KDM2 proteins constrain transcription of their target genes in mouse ES cells, but this does not rely on their demethylase activity (Turberfield et al. 2019).

In addition, KDM2B contains a ZF-CXXC domain which binds non-methylated CpG-dinucleotides *in vitro* (Voo et al. 2000; Lee et al. 2001), allowing KDM2B to broadly occupy non-methylated CpG islands (NMIs) genome-wide (Farcas et al. 2012). In mammals, the vast majority of Polycomb target sites overlap with NMIs (Ku et al. 2008), suggesting that KDM2B may serve as a direct molecular link between NMIs and Polycomb recruitment. Indeed, recent studies demonstrated that KDM2B is responsible for targeting of PCGF1-PRC1 complexes to NMI-associated developmental gene promoters in mouse ES cells (Farcas et al. 2012; Wu et al. 2013; He et al. 2013a), and this relies on the ZF-CXXC domain of KDM2B (Blackledge et al. 2014). Importantly, deletion of this domain in mouse ES cells leads to reductions in H2AK119ub1 and PRC2 levels at the promoters of Polycomb target genes, as well as their partial reactivation, suggesting that PCGF1-PRC1 recruitment by KDM2B contributes to the formation of repressive Polycomb chromatin domains (Blackledge et al. 2014). In line with this, homozygous deletion of the ZF-CXXC domain of KDM2B in mice leads to homeotic transformations of the axial skeleton and embryonic lethality (Blackledge et al. 2014).

Finally, KDM2B was found to interact with the SKP1 (S-Phase kinase-associated protein 1) protein through its F-box domain (Farcas et al. 2012; Wong et al. 2016). This suggested that KDM2B may assemble into SKP-Cullin-Fbox (SCF) E3 ubiquitin ligase complexes, which promote proteasomal degradation of target proteins (Cardozo and Pagano 2004). However, KDM2B-SKP1 complexes lack Cullin subunits, which normally bridge F-box proteins with E3 ubiquitin ligases (Farcas et al. 2012). Instead, KDM2B itself was reported to possess inherent E3 ubiquitin ligase activity through its potentially misannotated PHD domain (Janzer et al. 2012; Han et al. 2016), although the contribution of this putative catalytic activity of KDM2B towards Polycomb system function remains unknown.

PCGF2 and PCGF4 were also shown to form RYBP-containing variant PRC1 complexes, in addition to CBX-containing canonical PRC1 (Gao et al. 2012; Tavares et al. 2012; Morey et al. 2013; Hauri et al. 2016). Variant PCGF2/4-PRC1 were found to be more active at catalysing H2AK119ub1 *in vitro* compared to their canonical counterparts, but were still less active than variant-specific PCGF1/3/5/6-PRC1 complexes (Gao et al. 2012; Tavares et al. 2012; Rose et al.

2016). The unusual feature of variant PCGF2/4-PRC1 complexes is that they comprise only RING1A/B, PCGF2/4 and RYBP/YAF2 subunits and do not incorporate any other accessory proteins (Hauri et al. 2016). Therefore, it was proposed that these complexes may represent transient protein assemblies that will mature into functional canonical PRC1 complexes. Clearly, future work is required to understand how widely variant PCGF2/4-containing PRC1 complexes may contribute to Polycomb system function in mammals.

PCGF3 and PCGF5 associate with the same auxiliary subunits and assemble into nearly identical PRC1 complexes, which are closely related to the L(3)73Ah non-canonical PRC1 complex in *Drosophila* (Figure 1.3) (Gao et al. 2012; Lee et al. 2015; Hauri et al. 2016). Based on their high similarity, PCGF3/5-PRC1 complexes have been proposed to play partially redundant roles in gene regulation. In line with this, while PCGF3- or PCGF5-deficient mice are viable and have no developmental defects, double knockout of PCGF3/5 results in female-specific embryonic lethality at E9.5–E10.5 (Almeida et al. 2017). Interestingly, this severe phenotype is likely caused by the defects in X chromosome inactivation observed in PCGF3/5-null mice, suggesting an essential role of PCGF3/5-PRC1 in this process. Indeed, recent studies demonstrated that PCGF3/5-PRC1 are recruited by the Xist RNA to the Xist-expressing X chromosome, where they initiate accumulation of Polycomb complexes and their histone modifications and promote chromosome-wide transcriptional silencing (Almeida et al. 2017; Pintacuda et al. 2017; Nesterova et al. 2019). However, beyond X-inactivation, how PCGF3/5-PRC1 are targeted to specific genomic sites on autosomes remains unclear, since they lack any stably associated subunits with DNA- or RNA-binding activities (Gao et al. 2012; Hauri et al. 2016). Interestingly, sequence-specific transcription factors USF1/2 were recently found to dynamically interact with PCGF3/5-PRC1 in mouse ES cells (Scelfo et al. 2019). However, the contribution of these transcription factors to recruitment of PCGF3/5-PRC1 appears to be limited to a small number of target sites, suggesting that they are unlikely to represent the general targeting mechanism for PCGF3/5-PRC1 complexes.

Surprisingly, in contrast to their role in Xist-mediated transcriptional silencing, PCGF3/5-PRC1 have been generally implicated in gene activation, as these complexes preferentially occupy promoters of actively transcribed genes and can stimulate transcription of a luciferase reporter gene when artificially tethered to its promoter (Gao et al. 2014; Zhao et al. 2017a; Yao et al. 2018; Scelfo et al. 2019). Mechanistically, the function of PCGF3/5-PRC1 complexes as transcriptional activators may be mediated via activities of their CK2 (casein kinase 2) and AUTS2

(autism susceptibility candidate 2) accessory subunits. CK2 was shown to inhibit histone H2A mono-ubiquitination by PCGF5-RING1B-AUTS2 complex *in vitro* through phosphorylation of RING1B, while AUTS2 may promote transcription initiation through recruitment of p300/CBP co-activator complex (Gao et al. 2014). However, it remains unclear how widely these activities of PCGF3/5-PRC1 complexes contribute to gene regulation, especially given that association of AUTS2 with these complexes appears to be cell type-specific (Hauri et al. 2016; Zhao et al. 2017a).

PCGF6-PRC1 is the first variant PRC1 complex characterised in mammalian cells, originally known by the name of E2F6.com-1 based on the association with E2F6 transcription factor (Ogawa et al. 2002). In addition to E2F6, PCGF6-PRC1 integrates a variety of sequence-specific DNA-binding subunits, including MAX, MGA and DP1/2 transcription factors, and a histone-binding protein L3MBTL2 (Trojer et al. 2011; Qin et al. 2012; Gao et al. 2012; Hauri et al. 2016; Endoh et al. 2017). A heterodimer between MGA, which contains T-box and bHLH (basic helix-loop-helix) DNA-binding domains, and MAX can recognise E-box sequences, while a heterodimer between E2F6 and DP1/2 transcription factors can recognise the E2F consensus motif (Gaubatz et al. 1998; Hurlin et al. 1999; Jolma et al. 2013). L3MBTL2 contains four MBT domains, through which it can bind mono- and di-methylated tails of histones H3 and H4 *in vitro* (Guo et al. 2009; Yoo et al. 2010; Stielow et al. 2013). All of these targeting modules were shown to extensively co-localise with PCGF6 and RING1B in the genome and cooperatively contribute to stable chromatin binding of PCGF6-PRC1 at target sites (Endoh et al. 2017; Stielow et al. 2018; Scelfo et al. 2019).

In addition to DNA- and histone-binding targeting modules, PCGF6-PRC1 also associates with HDAC1/2 histone deacetylases, G9a/GLP histone methyltransferases and CBX1/3 proteins (also known as HP1$\beta$ and HP1$\gamma$) (Trojer et al. 2011; Gao et al. 2012; Hauri et al. 2016; Endoh et al. 2017), all of which have been previously implicated in transcriptional silencing (Denslow and Wade 2007; Shankar et al. 2013; Canzio et al. 2014). Although these accessory subunits may confer additional H2AK119ub1-independent repressive activities to PCGF6-PRC1 complexes, facilitating target gene repression, their exact contribution to PCGF6-PRC1 function remains poorly understood. In mouse ES cells, PCGF6-PRC1 complexes are recruited to promoters of germline-specific genes, where they elicit transcriptional silencing (Zdzieblo et al. 2014; Endoh et al. 2017; Zhao et al. 2017b). Importantly, repression of germ cell-related genes by PCGF6-PRC1 is essential for self-renewal of ES cells (Endoh et al. 2017; Zhao et al. 2017b) and male germ lineage specification in mice (Pohlers et al. 2005). Moreover, despite being viable, PCGF6-

null mice are not born at the normal Mendelian ratio and display homeotic transformations of the axial skeleton (Endoh et al. 2017). This indicates that in addition to its specialised role in control of the germline-specific gene expression, PCGF6-PRC1 also contributes to transcriptional silencing of classical Polycomb target genes during embryonic development.

Diversification of PcG proteins during vertebrate evolution has led to expansion of the repertoire of mammalian PRC1 complexes. Extensive biochemical characterisation of a multitude of distinct PRC1 protein assemblies has provided important insights into the complexity of gene regulation by the Polycomb system in mammals, and this still represents an area of great interest in the field. Nevertheless, despite a detailed account of their biochemical activities and interaction partners, the exact contribution of individual PRC1 complexes to deposition of H2AK119ub1, Polycomb target site selection and Polycomb-mediated gene repression has remained elusive.

### 1.2.3 PRC2 in mammals

#### 1.2.3.1 Catalytic core of PRC2

The catalytic core of PRC2 is highly conserved across *Metazoa*, as well as plants and some fungi, and in contrast to PRC1, its components underwent little diversification during vertebrate evolution. In mammals, PRC2 catalytic core comprises four core subunits, including EZH1 or EZH2 (enhancer of zeste), SUZ12 (suppressor of zeste 12), EED1-4 (embryonic ectoderm development) and a histone-binding factor RbAp46/48 (also known as RBBP7/4) (Figure 1.4, *top*) (Cao et al. 2002; Czermin et al. 2002; Muller et al. 2002; Kuzmichev et al. 2002). PRC2 is responsible for placing all mono-, di- and tri- methylation of histone H3K27 (H3K27me1/2/3) throughout the genome, which are abundant histone modifications with distinct genomic distributions and functions in gene regulation (discussed in section 1.1.2.3) (Schuettengruber et al. 2017; Laugesen et al. 2019). PRC2 plays an essential role in normal embryonic development, as mice deficient for PRC2 core catalytic subunits fail to complete gastrulation and die at E7.5-E8.5 (Faust et al. 1995; OCarroll et al. 2001; Pasini et al. 2004).

EZH1/2 methyltransferases modify histone H3K27 in a stepwise manner from the mono-methylated to the tri-methylated state through their conserved SET domain (Jenuwein et al. 1998). However, on their own, EZH1/2 adopt a default auto-inhibitory configuration and require the interaction with at least two PRC2 core subunits, EED and SUZ12, for efficient catalysis and

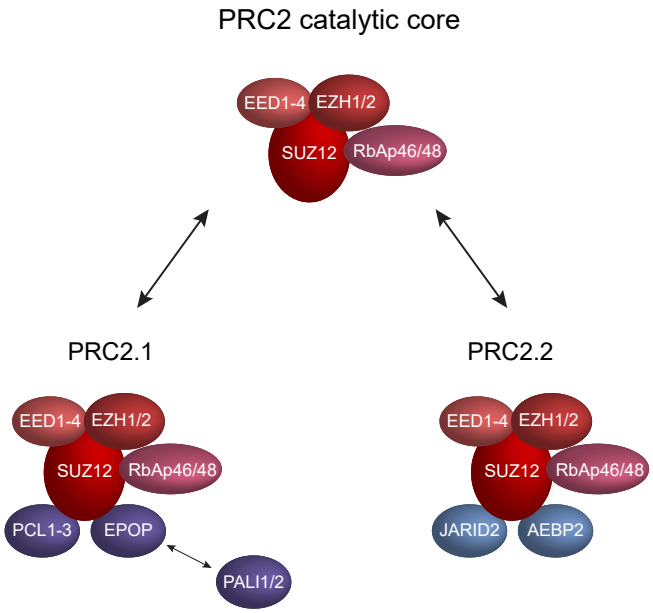



**Figure 1.4: Two distinct classes of mammalian PRC2 complexes**
*Top:* The catalytic core of PRC2 comprises EZH1/2, SUZ12, EED and RbAp46/48 proteins. EZH1/2 are catalytic subunits of PRC2 that catalyse H3K27me1/2/3. SUZ12 bridges all core catalytic subunits but also interacts with accessory proteins. EED binds H3K27me3 and this contributes to allosteric activation of EZH1/2. RbAp46/48 is a histone chaperone which promotes binding of the PRC2 catalytic core to nucleosomes. Association between all of these components is required for the full histone methyltransferase activity of the complex.
*Bottom:* Based on the unique composition of accessory subunits, PRC2 complexes are classified into PRC2.1 and PRC2.2. PRC2.1 contains DNA-binding PCL1-3 subunits and either EPOP or PALI1/2 subunits, which are mutually exclusive. PRC2.2 is specified by the presence of AEBP2 and JARID2 subunits, with JARID2 being able to specifically recognise H2AK119ub1.

*in vivo* stability (Cao and Zhang 2004; Pasini et al. 2004; Jiao and Liu 2015). The two EZH1/2 paralogues are incorporated into PRC2 complexes in a mutually exclusive manner and have both shared and unique functions in gene regulation. EZH1 and EZH2 display distinct cell type-specific expression patterns, with EZH2 being expressed predominantly in actively proliferating cells, and EZH1 found both in dividing and non-dividing cells (Margueron et al. 2008; Stojic et al. 2011). Furthermore, they have distinct substrate specificity and capacity to catalyse different states of histone H3K27 methylation, with the methyltransferase activity of EZH1-PRC2 shown to be less efficient that that of EZH2-PRC2 both *in vitro* and *in vivo* (Margueron et al. 2008; Lee et al. 2018; Lavarone et al. 2019). In line with this, in the absence of EZH2, EZH1-PRC2 is only sufficient to support normal global levels of H3K27me1, but not H3K27me2 and H3K27me3, failing to efficiently counteract non-specific H3K27 acetylation at promoters and distal regulatory elements (Shen et al. 2008; Lavarone et al. 2019). However, EZH1-PRC2 is able to sustain H3K27me3 enrichment at Polycomb target gene promoters in EZH2-null mouse ES cells, consistent with a less severe phenotype of these cells comparing to cells lacking EED (Boyer et al. 2006; Shen et al. 2008; Lavarone et al. 2019). This indicates that EZH1 can at least partially compensate for

loss of EZH2, suggesting that in addition to their specialised roles in regulation of histone H3K27 methylation and gene expression, EZH1 and EZH2 have also retained overlapping functions.

In mammals, all four translation isoforms of EED (EED1-4) contain a characteristic $\beta$-propeller structure formed by seven WD40 domains, through which EED can bind H3K27me3 (Montgomery et al. 2007; Margueron et al. 2009; Xu et al. 2010). Direct recognition of the product of PRC2 catalysis by EED is thought to contribute to allosteric activation of PRC2 histone methyltransferase activity towards unmodified H3K27 residue (Hansen et al. 2008; Margueron et al. 2009; Xu et al. 2010; Jiao and Liu 2015). This positive feedback mechanism has been proposed to underlie stable maintenance of H3K27me3 during cell cycle progression, through parental H3K27me3-modified histones facilitating PRC2 catalysis on newly incorporated histones (Hansen et al. 2008; Margueron et al. 2009). In addition to EED, SUZ12 also stabilises the catalytic core of PRC2 and stimulates the inherently low histone methyltransferase activity of EZH1/2 via its C-terminal VEFS domain (Cao and Zhang 2004; Pasini et al. 2004; Rai et al. 2013; Jiao and Liu 2015). Interestingly, the N-terminal portion of SUZ12 protein, which does not interact with EZH1/2, has been proposed to target PRC2 activity to specific genomic loci, as it can localise at PRC2 target sites independently of other catalytic core components, likely through interactions with accessory factors (Hojfeldt et al. 2018). Finally, incorporation of RbAp46/48 further enhances histone methyltransferase activity of the minimal EZH1/2-SUZ12-EED PRC2 catalytic core through promoting its association with nucleosomes (Pasini et al. 2004; Nekrasov et al. 2005; Nowak et al. 2011). In contrast to the rest of core subunits, which are specific to PRC2, RbAp46/48 was also found to support activity of other histone-modifying and chromatin remodelling complexes, including NURF and NuRD (Zhang et al. 1999; Barak et al. 2003).

In addition to the core catalytic subunits, PRC2 integrates a variety of auxiliary components, which can further modulate PRC2 catalytic activity, chromatin binding and function in gene silencing (Smits et al. 2013; Grijzenhout et al. 2016; Kloet et al. 2016; Hauri et al. 2016). Systematic proteomics analysis of the PRC2 subunit interaction network uncovered two major classes of distinct PRC2 assemblies, which are specified by incorporation of the mutually exclusive accessory subunits (Figure 1.4, *bottom*) (Grijzenhout et al. 2016; Hauri et al. 2016). PRC2.1 complexes are defined by the presence of one of three PCL (Polycomb-like) paralogues (PCL1/2/3 or PHF1/MTF2/PHF19), but also contain recently characterised EPOP and PALI1/2 subunits (Smits et al. 2013; Alekseyenko et al. 2014; Conway et al. 2018), while PRC2.2 specifically

associates with AEBP2 and JARID2 (Peng et al. 2009; Pasini et al. 2010; Grijzenhout et al. 2016). Interestingly, functional specialisation of distinct PRC2 complexes appears to be highly conserved in evolution, as separate PCL- and JARID2-containing PRC2 assemblies have been also identified in *Drosophila* (Nekrasov et al. 2007; Herz et al. 2012). In the following sections, I will discuss the contribution of these accessory subunits to regulation of PRC2 activities.

#### 1.2.3.2 PRC2.1 complexes

In mammals, PRC2.1 complexes integrate one of three PCL subunits, PCL1 (PHF1), PCL2 (MTF2) or PCL3 (PHF19), which contain several conserved domains with characterised DNA- or histone-binding activities. PCL proteins have been previously reported to contribute to chromatin binding of PRC2 complexes at target sites (Casanova et al. 2011; Hunkapiller et al. 2012), although the mechanisms of how this could be achieved remained elusive. Recent genetic and structural studies in mouse ES cells revealed that the winged-helix (WH) domain of PCL1-3 proteins specifically recognises non-methylated CpG-rich DNA sequences, enabling PRC2 recruitment to target sites, many of which overlap with NMIs (Ku et al. 2008; Li et al. 2017; Perino et al. 2018). In line with this, in mouse ES cells lacking the most abundant PCL paralogue, PCL2 (Smits et al. 2013; Kloet et al. 2016), or PCL1-3 together, stable occupancy of PRC2.1 complex at target sites was largely lost, resulting in a major reduction of PRC2 catalytic core association with chromatin at these sites, as well as deposition of H3K27me3 (Li et al. 2017; Perino et al. 2018; Healy et al. 2019; Højfeldt et al. 2019). Interestingly, in addition to its role in PRC2 targeting, DNA-binding activity of PCL1 was also shown to prolong PRC2 residence time on chromatin, promoting catalysis of H3K27me3 *in vitro* (Sarma et al. 2008; Choi et al. 2017).

PCL proteins can also recognise H3K36me3 through their conserved Tudor domains (Ballare et al. 2012; Brien et al. 2012; Cai et al. 2013). H3K36me3-binding activity of PCL1 and PCL3 has been proposed to promote their association with previously actively transcribed genes upon differentiation of mouse ES cells, when these genes need to be transitioned into a Polycomb-repressed state (Ballare et al. 2012; Brien et al. 2012). However, the contribution of H3K36me3-recognition by PCLs to PRC2 recruitment in mouse ES cells remains unclear, in particular, given that the most abundant PCL paralogue in these cells, PCL2, has been predicted to have substantially lower affinity for H3K36me3 than PCL1 and PCL3 (Musselman et al. 2012).

In addition to PCL proteins, PRC2.1 was recently found to associate with EPOP (Elongin BC

and Polycomb repressive complex 2-associated protein) or vertebrate-specific PALI1/2 (PRC2 associated LCOR isoform 1 or 2) accessory subunits, which are incorporated into the complex in a mutually exclusive manner (Smits et al. 2013; Alekseyenko et al. 2014; Hauri et al. 2016; Conway et al. 2018). In mouse ES cells, EPOP is more highly expressed than PALI1/2, but upon differentiation towards neural progenitors, PALI1/2 become more abundant (Kloet et al. 2016). Although both EPOP and PALI1/2 can stimulate PRC2 methyltransferase activity *in vitro* (Zhang et al. 2011; Conway et al. 2018), in mouse ES cells, EPOP appears to be dispensable for normal levels of H3K27me3 at target sites (Beringer et al. 2016), while the contribution of PALI1/2 to H3K27me3 is modest (Conway et al. 2018). Interestingly, EPOP has been implicated in supporting transcriptionally permissive state at PRC2 target gene promoters via its interaction with the B/C subunits of the Elongin complex, which is known to positively regulate transcription elongation (Beringer et al. 2016; Liefke et al. 2016). In contrast to this, another study found that PRC2 methylates Elongin A subunit, and this is required for transcriptional silencing of a subset of Polycomb target genes (Ardehali et al. 2017). Therefore, the relationship between the functions of PRC2 and the Elongin complex in gene regulation remains elusive. Finally, in addition to PRC2, PALI1/2 were also found to interact with other chromatin-based transcriptional regulators, including CTBP1/2 transcriptional co-repressors, G9A histone methyltransferase, SET protein and USP11 and USP22 deubiquitinases (Conway et al. 2018). Therefore, while the PCL accessory subunits appear to play an essential role in targeting of PRC2.1 complexes, the function of PALI1/2 and EPOP components may be to confer specialised gene regulatory activities to PRC2.1 by bridging it with additional transcriptional regulators.

#### 1.2.3.3 PRC2.2 complexes

A class of PRC2.2 complexes is specified by the presence of AEBP2 (Adipocyte enhancer-binding protein 2) and JARID2 (Jumonji- and ARID-domain-containing protein 2) accessory subunits (Kim et al. 2009a; Peng et al. 2009; Shen et al. 2009; Smits et al. 2013; Grijzenhout et al. 2016; Hauri et al. 2016; Conway et al. 2018). Both AEBP2 and JARID2 were shown to stimulate histone methyltransferase activity of the PRC2 catalytic core *in vitro*, although via distinct mechanisms (Cao and Zhang 2004; Li et al. 2010; Son et al. 2013; Lee et al. 2018). Recent studies suggested that JARID2-dependent stimulation of EZH1/2 catalytic activity may be analogous to EED-mediated allosteric activation, and at least in part rely on methylation of

JARID2 at lysine 116 (K116me3), which mimics H3K27me3 (Sanulli et al. 2015; Jiao and Liu 2015; Kasinath et al. 2018). In contrast, AEBP2 is required for enhanced nucleosome binding of the core complex and acts synergistically with JARID2 to stimulate PRC2 catalytic activity (Lee et al. 2018; Kasinath et al. 2018). However, loss of AEBP2 and/or JARID2 in mouse ES cells minimally affects H3K27me3 levels both globally and specifically at Polycomb target sites (Grijzenhout et al. 2016; Conway et al. 2018; Healy et al. 2019; Højfeldt et al. 2019).

Although JARID2 is a founding member of a family of JmjC domain-containing histone demethylases, it lacks the essential residues for cofactor binding and is predicted to be catalytically inactive, suggesting that its function is independent of histone demethylation (Klose et al. 2006). Instead, JARID2 has been proposed to contribute to PRC2 targeting, as disruption of JARID2 in mouse ES cells impairs normal occupancy of PRC2 core catalytic subunits at target sites (Shen et al. 2009; Peng et al. 2009; Landeira et al. 2010; Pasini et al. 2010; Healy et al. 2019; Højfeldt et al. 2019). Initially, it has been proposed that JARID2 guides PRC2 recruitment through DNA-binding activity of its ARID and ZF domains, which were shown to have a weak preference for CG-rich sequences *in vitro* (Kim et al. 2003; Kim et al. 2004; Li et al. 2010). However, the relevance of this binding mode for PRC2 targeting in cells remains unclear. Instead, accumulating evidence suggests that JARID2 may contribute to stable PRC2 chromatin occupancy through direct recognition of H2AK119ub1 via its UIM (ubiquitin interaction motif) domain (Kalb et al. 2014; Cooper et al. 2016). Furthermore, JARID2 binding to H2AK119ub1-modified nucleosomes was shown to stimulate H3K27 tri-methylation by PRC2 *in vitro*, further supporting a central role of H2AK119ub1 in promoting both occupancy and catalytic activity of PRC2 (Kalb et al. 2014).

AEBP2 contains three Gli-Krüppel C2H2-type ZF domains and was reported to bind DNA sequences with a highly degenerate motif *in vitro* (He et al. 1999; Kim et al. 2009a). However, AEBP2 and its potential DNA-binding activity appear to minimally contribute to PRC2 targeting in mouse ES cells, as knockout of AEBP2 did not lead to reductions in occupancy of PRC2 core subunits at Polycomb target sites (Grijzenhout et al. 2016). Interestingly, in contrast to this, H3K27me3 levels at target sites were modestly elevated in the absence of AEBP2 comparing to wild-type cells (Grijzenhout et al. 2016; Conway et al. 2018). This suggests that AEBP2 may constrain deposition of H3K27me3 in mouse ES cells, despite its capacity to stimulate PRC2 catalytic activity *in vitro* (Cao and Zhang 2004; Lee et al. 2018). In agreement with this, AEBP2-deficient mice display posterior-to-anterior transformations of the axial skeleton,

which are characteristic of mutations in TrxG genes (Grijzenhout et al. 2016). The aberrant PRC2 activity in AEBP2-null cells has been proposed to result from reorganisation of PRC2 auxiliary subunits into a more catalytically active hybrid PRC2 complex comprising both JARID2 and PCL2 (Grijzenhout et al. 2016). These observations indicate that AEBP2 may play an important role in specifying the mutually exclusive auxiliary subunit composition of PRC2.1 and PRC2.2.

Importantly, PRC2.1 and PRC2.2 complexes co-occupy largely the same target sites in the genome, suggesting that they may both contribute to stable occupancy of the PRC2 catalytic core and enrichment of H3K27me3 at these genomic loci (Liefke and Shi 2015; Grijzenhout et al. 2016; Perino et al. 2018; Healy et al. 2019; Højfeldt et al. 2019). However, what is the relative contribution of the PRC2.1- and PRC2.2-specific targeting and regulatory activities at Polycomb target sites has been unclear. Recent evidence from combinatorial genetic perturbations of PRC2 accessory subunits in mouse ES cells demonstrated that PCL1-3 and JARID2 represent the major targeting modules of PRC2.1 and PRC2.2 respectively, and their combined activity is required to fully define specific enrichment of PRC2 and H3K27me3 at target sites (Healy et al. 2019; Højfeldt et al. 2019). Interestingly, although in the absence of all accessory subunits, the catalytic core of PRC2 can no longer stably bind and deposit high levels of H3K27me3 at Polycomb target genomic regions, it is still able to place normal levels of H3K27me3 throughout the genome (Højfeldt et al. 2019). This indicates that the main function of PRC2 auxiliary components is not to support integrity and catalytic activity of PRC2 core complexes, but instead to direct stable PRC2 chromatin binding and efficient deposition of H3K27me3 to specific target sites.

### 1.2.4 Polycomb repressive deubiquitinase (PR-DUB) complex in mammals

Polycomb repressive deubiquitinase (PR-DUB) complex was initially characterised in *Drosophila*, as a complex comprising a ubiquitin carboxyl-terminal hydrolase Calypso (dBap1) and a PcG protein ASX (additional sex combs) (Scheuermann et al. 2010). Calipso shows minimal H2AK119ub1-specific deubiquitinase (DUB) activity *in vitro* and requires further activation by the ASX subunit, which promotes Calipso binding to ubiquitin (Scheuermann et al. 2010; Sahtoe et al. 2016). Disruption of PR-DUB subunits leads to a dramatic increase in global H2AK119ub1 levels and somewhat surprisingly derepression of *Hox* genes in *Drosophila* (Scheuermann et al. 2010). In line with this, both *Calipso* and *Asx* were previously identified as PcG genes in genetic screens, as their loss-of-function mutations cause Polycomb homeotic transformations in flies (Jürgens 1985;

Soto et al. 1995; Ayala Alonso et al. 2007). This suggests that deubiquitination of H2AK119ub1 by PR-DUB may be required for Polycomb-mediated gene repression, although the mechanism of how this could be achieved remains unclear. Intriguingly, inactivating mutations in the *Asx* gene were linked to both PcG and TrxG phenotypes in flies, indicating that ASX could play a dual role in activation and silencing of *Hox* genes (Sinclair et al. 1998; Milne et al. 1999).

The catalytic core of PR-DUB complex is highly conserved in mammals, where it comprises an orthologue of Calipso, BAP1 (BRCA1-associated protein 1), and one of three ASX orthologues, ASXL1-3 (Scheuermann et al. 2010). Mammalian BAP1-ASXL complex was also shown to catalyse H2AK119ub1 deubiquitination *in vitro*, with ASXL stimulating the inherently low DUB activity of BAP1 (Scheuermann et al. 2010; Sahtoe et al. 2016). However, in addition to ASXL, BAP1 associates with a large variety of auxiliary factors, including HCF-1 transcriptional regulator, FOXK1/2 and YY1 transcription factors, MBD5/6 methyl-binding domain proteins, KDM1B histone demethylase and OGT O-GlcNAc transferase (Sowa et al. 2009; Dey et al. 2012; Kloet et al. 2016; Hauri et al. 2016). Furthermore, BAP1 was shown to deubiquitinate a large number of protein substrates, in addition to H2AK119ub1, regulating their global protein levels and activity at specific genomic loci (Machida et al. 2009; Misaghi et al. 2009; Dey et al. 2012; Lee et al. 2014; Zarrizi et al. 2014). Through the diverse activities of its interaction partners, BAP1 has been implicated in regulation of many cellular processes, including DNA damage response, DNA replication and cell cycle progression (Carbone et al. 2013). While previous studies have primarily focussed on the role of BAP1 in cell proliferation, BAP1 function in gene regulation, especially in the context of Polycomb-mediated gene repression, has remained poorly characterised. Furthermore, to what extent H2AK119ub1 deubiquitination by BAP1 may contribute to gene regulation also requires a more detailed investigation, in particular, since multiple H2AK119ub1-specific deubiquitinases have been identified in mammalian cells (Joo et al. 2007; Zhu et al. 2007; Nakagawa et al. 2008; Sowa et al. 2009; Yang et al. 2014).

While in *Drosophila*, PR-DUB components were shown to occupy PREs together with PRC1 (Scheuermann et al. 2010), the distribution of mammalian PR-DUB complexes in the genome is poorly characterised. Genomic profiling of chromatin binding of PR-DUB accessory subunits in human cells revealed that they broadly co-localise with PRC1 at a large number of target gene promoters, despite PR-DUB being moderately enriched at actively transcribed genes (Hauri et al. 2016). BAP1 associates with several DNA-binding factors, including FOXK1/2 and YY1, which

were previously shown to direct BAP1 recruitment to a specific subset of gene promoters (Yu et al. 2010; Ji et al. 2014). Nevertheless, the general principles underlying BAP1 target site selection in the genome still remain poorly understood.

Interestingly, mammalian PR-DUB complex incorporates OGT subunit, which catalyses addition of O-linked N-acetylglucosamine (O-GlcNAc) to a large number of nuclear and cytoplasmic proteins, including RNA polymerase II and other transcriptional regulators (Hart et al. 2011; Kloet et al. 2016; Hauri et al. 2016). Although a *Drosophila* orthologue of OGT, Ogt/Sxc, does not form a complex with Calipso and ASX, it has been previously characterised as a PcG protein based on the classical Polycomb phenotype of *ogt/sxc* mutant flies (Ingham 1984). In line with this, Ogt/Sxc and O-GlcNAc-modified proteins were found to be enriched at PREs, while cells lacking Ogt/Sxc failed to silence Polycomb target genes (Gambetta et al. 2009). Furthermore, PH subunit of PRC1 was identified as one of the Ogt/Sxc substrates, with O-GlcNAcylation reported to prevent non-productive aggregation of PH proteins (Gambetta et al. 2009; Gambetta and Müller 2014). Together, these findings support an important role of Ogt/Sxc in Polycomb-mediated gene repression in *Drosophila*. However, although OGT was shown to be essential for early embryonic development and ES cell viability in mammals (Shafi et al. 2000; ODonnell et al. 2004), its contribution to Polycomb system function remains elusive.

Understanding the contribution of mammalian PR-DUB to gene regulation and how it is linked to activities of Polycomb complexes is particularly important in the light of the essential function of BAP1 as a tumour suppressor. BAP1 has received a lot of medical interest in the past, as it was found to be commonly mutated in an increasing number of cancers of diverse origins, including uveal melanoma, pleural mesothelioma, renal cell carcinoma, ovarian and breast cancers (Dey et al. 2012; Carbone et al. 2013; Qin et al. 2015a; He et al. 2019b). Furthermore, inactivating mutations of BAP1 in germline were found to cause a tumour predisposition syndrome (BAP1-TPDS), generally associated with high tumour aggressiveness and poor prognosis (Wiesner et al. 2011; Masoomian et al. 2018). Interestingly, some of the BAP1-deficient cancers were reported to show sensitivity to inhibition of Polycomb activity, suggesting potential antagonism between BAP1 and the Polycomb system (LaFave et al. 2015; Wang et al. 2018a). Therefore, a more detailed understanding of BAP1 function in the context of the Polycomb system might provide an insight into the pathogenesis of malignant transformations caused by BAP1 dysfunction and inform potential therapeutic strategies.

## 1.3 Polycomb-mediated gene repression: proposed models

### 1.3.1 Polycomb recruitment mechanisms to target sites

In mammals, PRC1 and PRC2 typically co-occupy promoters of target genes, many of which have essential functions in the control of normal embryonic development (Bernstein et al. 2006a; Boyer et al. 2006; Bracken et al. 2006). Combined PRC1 and PRC2 activity at their target sites leads to formation of Polycomb chromatin domains, which are highly enriched with Polycomb protein complexes and Polycomb-mediated histone modifications (Mikkelsen et al. 2007). However, how Polycomb target sites are selected and how Polycomb complexes reinforce each other's activity and occupancy to form Polycomb chromatin domains is still a matter of active debate.

Our initial understanding of Polycomb targeting mechanisms came from the discovery of Polycomb response elements (PREs) in *Drosophila* (reviewed in Müller and Kassis 2006; Ringrose and Paro 2007). A small number of PREs were originally identified as specialised *cis*-regulatory genomic elements, which can maintain repression of reporter *Hox* genes outside of their normal expression domains through recruitment of PcG proteins (Simon et al. 1993; Chan et al. 1994; Chiang et al. 1995; Gindhart et al. 1995). Subsequently, genome-wide profiling of chromatin binding of PcG proteins uncovered several hundred additional PREs in the *Drosophila* genome (Nègre et al. 2006; Schwartz et al. 2006; Tolhuis et al. 2006). Surprisingly, computational analysis of these genomic elements revealed that they do not share a conserved sequence consensus, but instead contain multiple recognition motifs for known DNA-binding proteins (Ringrose et al. 2003; Okulski et al. 2011). These often include motifs which are recognised by GAGA factor (GAF) and Pipsqueak (PSQ) proteins (Horard et al. 2000; Mulholland et al. 2003), Pho/Phol subunits of PhoRC complex (discussed in section 1.2.2.2) (Oktaba et al. 2008), HMG-like dorsal switch protein 1 (DSP1) (Deardin et al. 2005), Grainy head (Grh) (Blastyak et al. 2006) and members of the SP1/Krüppel-like factor (KLF) family (Brown et al. 2005). While most of these factors were found to associate with Polycomb protein complexes only at a subset of target sites, binding of Pho appeared to be the universal feature of PREs (Kwong et al. 2008; Schuettengruber et al. 2009). Based on these initial observations, the hierarchical mechanism of Polycomb recruitment was proposed, whereby sequence-specific DNA-binding proteins, often including Pho/Phol, target PRC2 to PREs, where it places H3K27me3, creating a docking site for PRC1 via PC-dependent recognition of H3K27me3 (Wang et al. 2004b) (Figure 1.5A). Interestingly, it

was also shown that Pho can directly bind PRC1 subunits, suggesting that PRC2 may not always be required for PRC1 recruitment (Mohd-Sarip et al. 2002; Kassis and Kennison 2010; Frey et al. 2016).

While the sequence composition of PREs and DNA-binding factors responsible for Polycomb targeting have been characterised reasonably well in *Drosophila*, the analogous elements in mammals have proved to be highly elusive. Although a limited number of mammalian PRE-like elements have been recently identified, how Polycomb complexes are recruited to these sites has remained poorly understood (Sing et al. 2009; Woo et al. 2010). Extrapolating from *Drosophila*, combinatorial recognition of specific DNA motifs by transcription factors has been proposed to underlie Polycomb target site selection in mammals. However, the majority of the *Drosophila* transcription factors involved in recognition of PREs are not conserved in mammals, with the only exception being an orthologue of Pho/Phol, YY1, which does not typically bind Polycomb target sites and thus is unlikely to drive Polycomb recruitment (Mendenhall et al. 2010).

#### 1.3.1.1 Polycomb recruitment by sequence-specific transcription factors

Several sequence-specific DNA-binding proteins have been identified as accessory subunits of variant PRC1 complexes. PCGF6-PRC1 integrates four different transcription factors, which form two separate DNA-binding modules, MAX/MGA and E2F6/DP1/2, which recognise distinct DNA motifs (discussed in section 1.2.2.2) (Gao et al. 2012; Hauri et al. 2016; Endoh et al. 2017). In mouse ES cells, the combined activity of these targeting modules was shown to drive recruitment of PCGF6-PRC1 complexes to promoters of germline-specific genes to elicit gene repression (Endoh et al. 2017; Scelfo et al. 2019). In line with this, depletion of MAX/MGA caused a marked reduction in PCGF6-PRC1 binding at their target sites, which was further decreased by additional depletion of E2F6 (Endoh et al. 2017; Scelfo et al. 2019). Importantly, this effect was specific for germ cell-related genes, as PRC1 occupancy at classical Polycomb target sites was largely unperturbed following loss of MAX/MGA (Endoh et al. 2017). Collectively, these findings indicate that targeting of PCGF6-PRC1 complexes relies on sequence-specific DNA-binding activities of their accessory subunits. Similarly, variant PCGF3/5-PRC1 complexes were recently found to interact with USF1/2 transcription factors (Scelfo et al. 2019). USF1/2 recognise a variant E-box DNA motif (Kiermaier et al. 1999), which was reported to be enriched at a small subset of Polycomb target sites uniquely bound by PCGF3/5-PRC1. Depletion of USF1/2

results in a nearly complete loss of PCGF3 binding at these target sites, indicating that USF1/2 transcription factors may represent the targeting module of PCGF3/5-PRC1 (Scelfo et al. 2019).

Furthermore, candidate-based approaches found several additional DNA-binding factors, which dynamically associate with Polycomb complexes. A transcription factor REST was shown to directly interact with PRC1 subunits, enabling their PRC2-independent recruitment to a subset of Polycomb-regulated neuronal genes in mouse ES cells (Ren and Kerppola 2011; Dietrich et al. 2012). Analogous mechanism has been reported for RUNX1 transcription factor, which directs PRC1 binding to shared target sites in differentiated cells (Yu et al. 2012). Finally, REST and a SNAIL family of transcription factors were also implicated in transient PRC2 recruitment and deposition of H3K27me3 at a specialised group of genes in neuronal progenitors (Arnold et al. 2012). Together, these observations indicate that sequence-specific transcription factors can contribute to Polycomb recruitment, but their contribution is often limited to specialised target genes or specialised developmental contexts.

#### 1.3.1.2 Polycomb recruitment by long non-coding RNAs

In addition to transcription factors, multiple long non-coding RNAs (lncRNAs) have been previously implicated in directing locus-specific binding of Polycomb complexes. Initial insights into the role of lncRNAs in Polycomb recruitment came from studies of X chromosome inactivation in mammals (reviewed in Brockdorff 2017). Initiation of X-inactivation is accompanied by accumulation of Polycomb complexes and their histone modifications on the silenced X chromosome, and this strictly depends on the Xist RNA (Plath et al. 2003; Napoles et al. 2004; Kohlmaier et al. 2004). Early studies proposed that various subunits of PRC2 can directly bind Xist through recognition of a stem-loop structure in the A-repeat element, which is critical for gene silencing (Wutz et al. 2002; Zhao et al. 2008a; Kanhere et al. 2010). However, more recently, systematic deletion analysis of Xist RNA revealed that a different B/C-repeat-containing region of Xist is necessary and sufficient to drive Polycomb recruitment and efficient chromosome-wide silencing through its interaction with PCGF3/5-PRC1 (Almeida et al. 2017; Pintacuda et al. 2017; Colognori et al. 2019; Nesterova et al. 2019). These findings suggested that accumulation of PRC2 complexes at the inactive X chromosome follows recruitment of PCGF3/5-PRC1 and is mediated via recognition of H2AK119ub1 by JARID2 rather than direct binding to Xist RNA. Importantly, binding of PCGF3/5-PRC1 complexes to the minimal Polycomb interaction domain of the Xist

RNA is also not direct and relies on the RNA-binding protein hnRNPK (Pintacuda et al. 2017).

In addition to Xist, several other lncRNAs were implicated in Polycomb recruitment to specific genomic loci in mammals. This includes the human ANRIL lncRNA, which was shown to drive PRC2-dependent silencing of INK4B tumour suppressor gene via interaction with SUZ12 (Kotake et al. 2011), the HOTAIR lncRNA expressed from the human HOXC locus, which was reported to repress the HOXD locus *in trans* by interacting with PRC2 subunits (Rinn et al. 2007), and the Kcnqt1ot1 ncRNA expressed from the partially imprinted locus in mice, which was shown to recruit both PRC2 and G9A for lineage-specific transcriptional silencing of the *Kcnq1* locus (Pandey et al. 2008). However, whether in these cases recruitment of Polycomb complexes relies on direct binding to lncRNAs or whether it is a consequence of lncRNA-induced transcriptional silencing still remains largely unclear. In the light of more recent findings showing that PRC2 pervasively binds nascent RNA without any sequence specificity (Davidovich et al. 2013; Kaneko et al. 2013; Beltran et al. 2016), the Polycomb field has now largely moved away from the idea that individual lncRNAs can direct recruitment of Polycomb complexes to specific genomic regions (reviewed in Brockdorff 2013).

#### 1.3.1.3 Polycomb recruitment by generic DNA-binding factors

While transcription factors and lncRNAs may underpin locus-specific Polycomb targeting mechanisms, they are unable to fully explain the widespread binding of Polycomb complexes in the large mammalian genomes. Genome-wide profiling of Polycomb-occupied target sites in mouse and human ES cells revealed that they are almost exclusively associated with non-methylated CpG islands (NMIs) (Bernstein et al. 2006a; Mikkelsen et al. 2007; Tanay et al. 2007; Ku et al. 2008). Found at approximately 70% of gene promoters in mammals, NMIs are genomic regions, typically 1-2 kb in size, with elevated C+G content and high density of non-methylated CpG-dinucleotides compared to the rest of the genome (Blackledge and Klose 2011; Deaton and Bird 2011). Given the remarkable overlap between these gene regulatory elements and genomic localisation of Polycomb complexes, it has been speculated that NMIs may play a direct role in Polycomb recruitment in mammals. In agreement with this possibility, several studies demonstrated that synthetic CpG-rich DNA sequences or genomic NMIs depleted of motifs recognised by activating transcription factors can drive *de novo* formation of ectopic Polycomb chromatin domains (Mendenhall et al. 2010; Lynch et al. 2012; Wachter et al. 2014; Jermann et al.

2014). Together, these observations have led to the proposal that NMIs may function as the long sought-after mammalian PREs.

A lot of effort in the past has been placed on elucidating the mechanisms which may be responsible for NMI recognition by Polycomb complexes. A molecular link between NMIs and Polycomb recruitment has first emerged from the discovery that KDM2B, an accessory subunit of variant PCGF1-PRC1 complexes, broadly engages with NMIs genome-wide through direct binding to non-methylated CpG-dinucleotides via its ZF-CXXC domain (discussed in section 1.2.2.2) (Farcas et al. 2012; He et al. 2013a; Wu et al. 2013). Importantly, recent studies have provided direct evidence that the inherent DNA-binding activity of KDM2B contributes to recruitment of PCGF1-PRC1 complexes to Polycomb target gene promoters in mouse ES cells (Blackledge et al. 2014). Strikingly, similar DNA-binding activity specific for non-methylated CpG-dinucleotides has been recently characterised for PRC2.1 auxiliary components PCL1-3, which were also shown to shape PRC2.1 occupancy and H3K27me3 at target sites (discussed in section 1.2.3.2) (Li et al. 2017; Perino et al. 2018; Healy et al. 2019; Højfeldt et al. 2019). Together, these observations reveal that both PCGF1-PRC1 and PRC2.1 complexes can independently sample NMIs throughout the genome via direct recognition of non-methylated CpG-rich DNA sequences by their KDM2B and PCL accessory subunits.

However, despite KDM2B and PCL proteins broadly occupying all NMIs throughout the genome, high-level enrichment of PRC1, PRC2 and their associated histone modifications is only achieved at a subset of NMI-associated target gene promoters which have low transcriptional activity. In line with this, our and other labs have previously proposed that transcription may antagonise formation of Polycomb chromatin domains (Klose et al. 2013; Jermann et al. 2014; Riising et al. 2014). In support of this idea, PRC2 has been recently shown to promiscuously bind nascent RNA transcripts *in vitro* and *in vivo* (Davidovich et al. 2013; Kaneko et al. 2013), resulting in inhibition of both PRC2 catalytic activity and stable chromatin binding at target sites (Cifuentes-Rojas et al. 2014; Kaneko et al. 2014; Beltran et al. 2016; Beltran et al. 2019). Furthermore, histone modifications associated with active transcription, including H3K4me3 and H3K36me3, were previously shown to inhibit histone H3K27 methylation by PRC2 (Schmitges et al. 2011; Yuan et al. 2011). Together, these observations support a model, in which generic DNA-binding activities of Polycomb complexes allow them to broadly sample NMIs throughout the genome to assess the transcriptional state of associated genes, while additional mechanisms are required to

promote accumulation of these complexes at sites lacking active transcription.

#### 1.3.1.4 Chromatin-based communication between PRC1 and PRC2

The sequence-specific and generic DNA-binding mechanisms described above underpin how mammalian PRC1 and PRC2 complexes independently select their target sites in the genome. However, PRC1 and PRC2 do not function autonomously and instead typically co-occupy target sites, where they are thought to cooperate to form repressive Polycomb chromatin domains (Mikkelsen et al. 2007). Extensive co-localisation of PRC1 and PRC2 complexes in the genome has been initially attributed to the hierarchical recruitment pathway, which was originally described in *Drosophila* (Wang et al. 2004b). It postulates that *de novo* recruitment of PRC2 initiates formation of Polycomb chromatin domains through deposition of H3K27me3, which is recognised by PC/CBX-containing PRC1 complexes finally placing H2AK119ub1 (Figure 1.5A). The hierarchical recruitment model has been later adopted as a central mechanism for Polycomb chromatin domain formation in mammals, based on the observations that stable PRC1 occupancy at target sites is reduced following loss of PRC2 (Cao et al. 2002; Cao et al. 2005; Boyer et al. 2006; Tavares et al. 2012). Furthermore, in line with this model, H2AK119ub1 and PRC1 levels were reported to closely mirror changes in H3K27me3 caused by depletion or over-expression of H3K27me3-specific histone demethylases (Lee et al. 2007). Finally, additional support for a direct role of H3K27me3 in PRC1 recruitment comes from the surprising finding that a viral SET (vSET) protein can recruit CBX8 subunit of PRC1 to repress transcription of a reporter gene through its histone H3K27 methyltransferase activity (Mujtaba et al. 2008).

However, the universality of the hierarchical model has been challenged by an increasing amount of evidence showing that PRC1 binding at target sites can be independent of PRC2 and H3K27me3 (Schoeftner et al. 2006; Pasini et al. 2007; Leeb et al. 2010). In agreement with this, recent work in mouse ES cells demonstrated that removal of PRC2 does not completely abrogate PRC1 occupancy at target sites, leaving global levels of H2AK119ub1 largely unperturbed (Tavares et al. 2012). Remarkably, these observations suggest that variant PRC1 complexes, which lack CBX subunits recognising H3K27me3, can bind and place H2AK119ub1 at target sites upstream of the PRC2-initiated hierarchical pathway, which has been eliminated in the PRC2-null cells.

Even more strikingly, variant PRC1 complexes have recently been found to drive recruitment of PRC2 and deposition of H3K27me3 through an H2AK119ub1-dependent mechanism (Blackledge

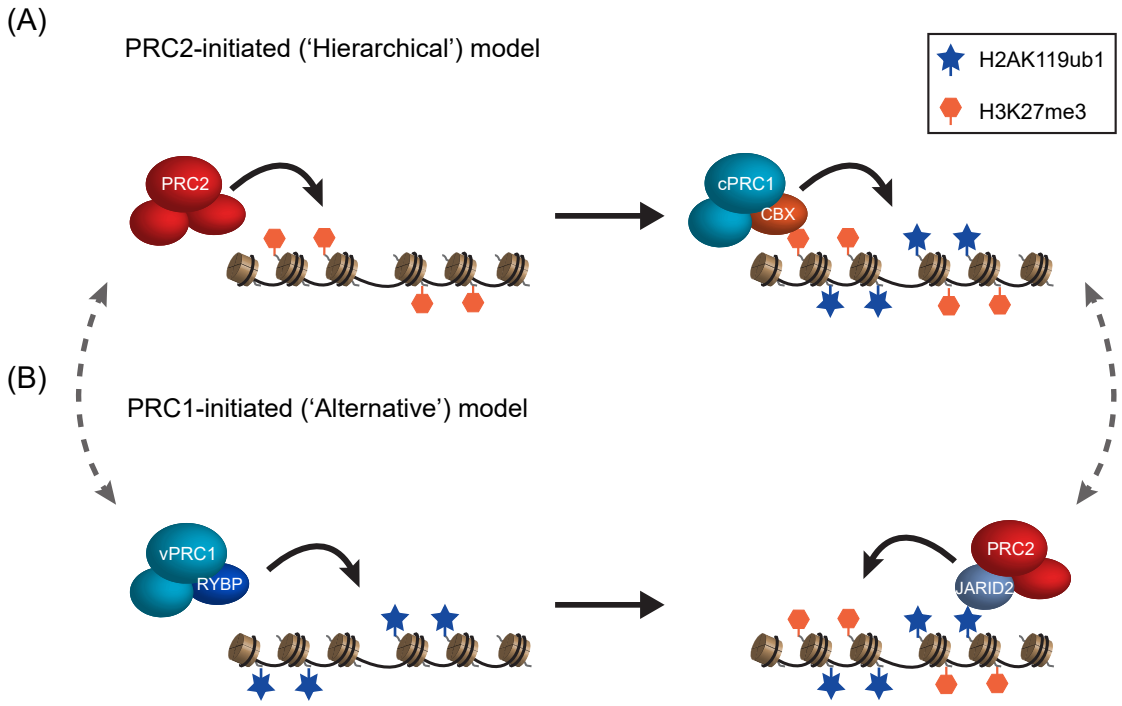



**Figure 1.5: Hierarchical and alternative models for Polycomb chromatin domain formation in mammals**
(A) In the PRC2-initiated hierarchical model, PRC2 is recruited to target sites first, where it places H3K27me3. Canonical PRC1 complexes can recognise H3K27me3 through their CBX subunits, thereby enabling PRC1 binding and deposition of H2AK119ub1 at target sites.
(B) In the PRC1-initiated alternative model, variant PRC1 complexes, which contain RYBP in place of CBX, bind target sites independently of PRC2 and H3K27me3. Variant PRC1 places H2AK119ub1, which can be directly recognised by JARID2 subunit of PRC2.2 complexes. This promotes PRC2 binding and deposition of H3K27me3 at target sites.
The two proposed mechanisms are not mutually exclusive and could reinforce each other to shape Polycomb chromatin domains in different developmental contexts. Deposition of H2AK119ub1 by canonical PRC1 complexes in (A) will promote binding of PRC2 complexes, as H2AK119ub1 can be recognised by JARID2. Deposition of H3K27me3 by PRC2 in (B) will promote binding of canonical PRC1, as H3K27me3 can be recognised by CBX.

et al. 2014), revealing an alternative to the prevailing hierarchical model pathway for Polycomb chromatin domain formation in mammals (Figure 1.5B). This ground-breaking discovery was based on the observation that artificial tethering of variant PRC1-specific subunits or a minimal PRC1 catalytic domain is sufficient to initiate *de novo* formation of a Polycomb chromatin domain at an ectopic genomic site (Blackledge et al. 2014). This was further corroborated by findings in another system, in which PRC1 was recruited to pericentric heterochromatin genomic regions (Cooper et al. 2014). Importantly, this alternative PRC1-initiated recruitment mechanism appears to play an essential role in Polycomb chromatin domain formation at native Polycomb target sites as well, since removal of PRC1 and concomitant loss of H2AK119ub1 in mouse ES cells was shown to cause marked reductions in the levels of PRC2 core subunits and H3K27me3 at target sites (Blackledge et al. 2014).

Importantly, deposition of H2AK119ub1 by variant PRC1 complexes appears to play a key role

in formation of Polycomb chromatin domains. In line with this, artificial tethering of canonical PRC1, which has an inherently weak E3 ubiquitin ligase activity (Taherbhoy et al. 2015; Rose et al. 2016), or an inactivated minimal PRC1 catalytic domain did not lead to accumulation of H2AK119ub1 and failed to initiate recruitment of PRC2 and deposition of H3K27me3 (Blackledge et al. 2014). This suggested that direct recognition of H2AK119ub1 may underpin PRC1-initiated recruitment of PRC2 complexes. In agreement with this possibility, histone H2A mono-ubiquitination was shown to stimulate nucleosome binding and H3K27 methyltransferase activity of JARID2-AEBP2-containing PRC2.2 complexes *in vitro* (Kalb et al. 2014). It was later discovered that JARID2 protein can recognise H2AK119ub1 via its UIM domain, providing a mechanistic link between PRC1-mediated deposition of H2AK119ub1 and recruitment of PRC2.2 complexes (Cooper et al. 2016). Further supporting this idea, ablation of JARID2 in mouse ES cells was shown to compromise stable binding of PRC2 core subunits and deposition of H3K27me3 at Polycomb target sites (Pasini et al. 2010; Healy et al. 2019; Højfeldt et al. 2019). Together, these observations indicate that the functional interplay between PRC1 and PRC2 in mammals is more complex than has been previously anticipated, as it involves multiple chromatin-based feedback mechanisms, cooperation between which is required to shape high-level enrichment of Polycomb complexes and their associated histone modifications at target sites.

Finally, in addition to the chromatin-based communication, PRC1 and PRC2 may directly recruit each other to target sites through physical association of their components. In line with this, EED subunit of PRC2 was proposed to contribute to PRC1 targeting via direct interaction with RING1B and PCGF4 proteins (Cao et al. 2014). Furthermore, as mentioned previously, *Drosophila* PhoRC complex was shown to directly recruit PRC1 to PREs, independently of PRC2 and H3K27me3 (Mohd-Sarip et al. 2002; Kassis and Kennison 2010; Frey et al. 2016). Together, these observations suggest an additional layer of complexity in the functional cross-talk between Polycomb complexes, which does not rely on their histone-modifying activities.

### 1.3.2 Proposed mechanisms of gene repression

Polycomb complexes function together to stably maintain repressed transcriptional states of their target genes throughout development. However, the molecular mechanisms by which this is achieved have remained elusive. Importantly, previous genetic and molecular studies of the Polycomb system have primarily ascribed its repressive activity to PRC1 complexes. In contrast,

whether the contribution of PRC2 to Polycomb system function extends beyond H3K27me3-mediated recruitment of canonical PRC1 is still actively debated and requires further investigation. Admittedly, distinguishing between the PRC1-dependent and PRC2-dependent gene repression mechanisms *in vivo* is challenging, considering that PRC1 and PRC2 activities at target sites are closely intertwined.

Regardless of whether Polycomb chromatin domain formation is initiated by variant PRC1 or PRC2, it culminates in recruitment of canonical PRC1 complexes, which have been proposed to act as central determinants of transcriptional silencing by the Polycomb system (Wang et al. 2004b; Simon and Kingston 2013). Some studies suggested that canonical PRC1 may elicit gene repression by directly blocking activity of transcriptional co-activators, such as histone acetyltransferase CBP (Tie et al. 2016), or through direct interactions with transcriptional repressors, such as REST (Ren and Kerppola 2011). However, the prevailing view in the field is that canonical PRC1 complexes function through isolating Polycomb target genes into distinct higher-order chromatin structures, which are compact and refractory to transcription. In support of this model, canonical PRC1 complexes were shown to drive local chromatin condensation *in vitro* (Francis et al. 2004; Lavigne et al. 2004; Grau et al. 2011) and promote association of Polycomb-occupied chromatin domains across a wide range of genomic distances in cells (described in section 1.2.2.1) (Eskeland et al. 2010b; Eskeland et al. 2010a; Isono et al. 2013; Schoenfelder et al. 2015; Boettiger et al. 2016; Kundu et al. 2017; Bonev et al. 2017; Lau et al. 2017). Importantly, PRC1 E3 ubiquitin ligase activity has been reported to be dispensable for both local and higher-order chromatin compaction by canonical PRC1 (Francis et al. 2004; Eskeland et al. 2010b; Kundu et al. 2017). These observations led to the hypothesis that while H2AK119ub1 may contribute to high-level occupancy of Polycomb complexes at target sites via a PRC1-initiated recruitment pathway, it is not required for Polycomb-mediated gene repression, which is primarily driven by chromatin compaction. In agreement with this possibility, recent studies showed that partial inactivation of PRC1 catalysis in mice results in milder developmental defects and later embryonic lethality than complete loss of PRC1 (Illingworth et al. 2015; Cohen et al. 2018).

However, how chromatin compaction by canonical PRC1 may interfere with transcription remains poorly understood. Initially, condensation of nucleosomal arrays has been proposed to limit binding and/or activity of transcriptional activators at *cis*-regulatory elements. In line with this, canonical PRC1 was shown to inhibit chromatin remodelling by SWI/SNF complexes *in*

*vitro* (Francis et al. 2001; Lavigne et al. 2004; Grau et al. 2011). Furthermore, Polycomb-occupied NMIs were found to have reduced chromatin accessibility, despite NMIs being generally associated with open and highly accessible chromatin conformation (Calabrese et al. 2012; Kelly et al. 2012; Beck et al. 2014; King et al. 2018). However, this gene repression model has been challenged by the observations that canonical PRC1 does not abrogate binding of transcriptional activators, GTFs and RNA polymerase II to gene promoters *in vitro* or *in vivo* (Breiling et al. 2001; King et al. 2002; Stock et al. 2007; Brookes et al. 2012; Lehmann et al. 2012). More importantly, recent work from our lab demonstrated that removal of PRC1 and/or PRC2 does not result in major increases in chromatin accessibility at promoters of Polycomb target genes, indicating that Polycomb complexes do not define low chromatin accessibility at these sites (King et al. 2018). Together, these findings suggest that chromatin compaction by canonical PRC1 complexes does not render Polycomb target gene promoters inaccessible to the transcriptional machinery, and this is unlikely to constitute the primary mechanism of gene repression.

In addition to compaction of Polycomb-occupied chromatin domains, transcriptional silencing by the Polycomb system has been proposed to rely on histone H2A mono-ubiquitination interfering with the activity of the transcriptional machinery. In line with this idea, incorporation of H2AK119ub1 blocks transcription of chromatin templates *in vitro* (Nakagawa et al. 2008; Aihara et al. 2016). Mechanistically, addition of a bulk ubiquitin moiety to histone H2A may have an inhibitory effect on different stages of transcription, either directly or via recruitment of H2AK119ub1 reader proteins with repressive activities. Previous *in vitro* analysis of the transcription efficiency on H2AK119ub1-modified chromatin templates suggested that H2AK119ub1 impedes transcription initiation, but not elongation, through inhibition of histone H3K4 di- and tri-methylation (Nakagawa et al. 2008). However, recent evidence has argued that transcription elongation may also be regulated by histone H2A mono-ubiquitination. In line with this, deposition of H2AK119ub1 was found to prevent recruitment of the transcription elongation factor FACT, blocking release of RNA polymerase II into productive elongation (Zhou et al. 2008). In addition, H2AK119ub1 was also proposed to promote transcriptional pausing, as Polycomb target gene promoters in mouse ES cells were found to be enriched with Ser5P-modified non-productive RNA polymerase II, which was released into elongation following loss of PRC1 (Stock et al. 2007; Brookes et al. 2012). Furthermore, it was shown that catalytic activity of many H3K36-specific methyltransferases, associated with productive elongation, is inhibited on H2AK119ub1-

modified nucleosome substrates (Yuan et al. 2013). Finally, H2AK119ub1 may also support more compact architecture of nucleosomal arrays, as it was shown to promote association of the linker histone H1 with nucleosomes (Jason et al. 2004; Zhu et al. 2007).

Alternatively, PRC1-mediated gene repression may rely on activities of reader proteins, which can specifically recognise H2AK119ub1. Among a few characterised H2AK119ub1 readers is a subunit of the chromatin remodelling complex RSF (remodelling and spacing factor), which was proposed to contribute to Polycomb target gene repression through promoting regular spacing of promoter nucleosomal arrays and binding of the linker histone H1 (Zhang et al. 2017). In addition, another recent study found that the UIM domain of DNMT1 DNA methyltransferase can bind H2AK119ub1 *in vitro*, which led to the speculation that H2AK119ub1 may direct DNA methylation by DNMT1 to PRC1 target sites (Qin et al. 2015b). However, the physiological relevance of this interaction remains unclear, considering that in mouse ES cells, the majority of Polycomb target sites, which are enriched with H2AK119ub1, correspond to non-methylated CpG islands (Ku et al. 2008). Finally, in human cells, ZRF1 (zuotin-related factor 1) was identified as another potential reader of H2AK119ub1, which may contribute to activation of Polycomb-repressed genes during differentiation through displacement of Polycomb complexes (Richly et al. 2010). Together, these observations reveal that although H2AK119ub1-reader proteins may contribute to Polycomb system function in specialised contexts, they are unlikely to underpin general principles of gene regulation by the Polycomb system.

In conclusion, in the last decade, several potential mechanisms for Polycomb-mediated gene repression have emerged from extensive biochemical and genetic characterisation of the Polycomb system. All of these mechanisms can be broadly grouped into two distinct modes of repression, depending on whether they rely on canonical PRC1-mediated chromatin compaction, for which PRC1 catalytic activity is dispensable, or H2AK119ub1-mediated inhibition of productive transcription. The latter may primarily involve variant PRC1 complexes, which appear to be more active at catalysing H2AK119ub1 than canonical PRC1 both *in vitro* and *in vivo* (Blackledge et al. 2014; Taherbhoy et al. 2015; Rose et al. 2016). This multiplicity of putative Polycomb repressive activities, resulting from the immense variety of distinct Polycomb complexes, has hindered our understanding of Polycomb system function in gene regulation. Therefore, uncovering the primary mechanisms underlying Polycomb-mediated transcriptional silencing still remains a central outstanding question in the field.

## 1.4 Aims of this thesis

In eukaryotes, chromatin structure and post-translational histone modifications play fundamental roles in transcriptional control of gene expression. The Polycomb repressive system represents one of the most heavily studied paradigms for chromatin-based gene regulation. In mammals, it encompasses two major complexes - Polycomb repressive complex 1 (PRC1), which mono-ubiquitinates histone H2A at lysine 119 (H2AK119ub1), and PRC2, which methylates histone H3 at lysine 27 (H3K27me1/2/3), with PRC1 considered to play a central role in Polycomb-mediated gene repression. However, the primary mechanisms of how PRC1 elicits transcriptional silencing and whether this relies on PRC1 catalytic activity have remained elusive and still actively debated in the field. In part, this is due to an immense variety of distinct protein assemblies comprising PRC1 in mammals, which integrate unique targeting modules and regulatory activities. Therefore, the central aim of this thesis is to systematically dissect the contribution of individual PRC1 complexes towards Polycomb system function in order to uncover the central molecular determinants of Polycomb-mediated gene repression.

To address this question, I exploit a panel of conditional knockout mouse ES cell lines, in which distinct PRC1 complexes can be removed individually or in combination, which I interrogate using quantitative genomics approaches. I first examine the contribution of individual canonical (Chapter 3) and variant (Chapter 4) PRC1 complexes to deposition of H2AK119ub1, Polycomb chromatin domain formation and Polycomb target gene repression. Following detailed characterisation of the previously unappreciated pervasive distribution of H2AK119ub1 in the genome (Chapter 4), I investigate the mechanisms that control the levels of genome-wide H2AK119ub1 and then determine their contribution to gene regulation (Chapter 5). Having established that loss of individual PRC1 complexes does not fully compromise Polycomb system function (Chapters 3 and 4), I next investigate whether cooperation between distinct PRC1 complexes is required for Polycomb chromatin domain formation and gene repression (Chapter 6). Finally, in Chapter 7, I use a conditional genetic strategy to specifically ablate PRC1 catalytic activity in order to directly examine its role in Polycomb system function. Together, my findings provide compelling evidence for a variant PRC1-dependent logic, which relies on PRC1 catalysis, for Polycomb-mediated gene repression in mammals. Ultimately, this work paves the way for elucidating the molecular mechanisms of how Polycomb complexes counteract transcription.

# 2

# MATERIALS AND METHODS

## 2.1 Cell lines and culture

### 2.1.1 Cell culture conditions

Unless otherwise indicated, male mouse embryonic stem (ES) cells were grown on gelatin-coated plates at 37°C and 5% CO2, in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 15% fetal bovine serum (Labtech), 2 mM L-glutamine (Life Technologies), 1× penicillin-streptomycin solution (Life Technologies), 1× non-essential amino acids (Life Technologies), 0.5 mM beta-mercaptoethanol (Life Technologies), and 10 ng/ml leukemia inhibitory factor. *Pcgf6*$^{fl/fl}$ ES cells and *Mus domesticus* (129S1) × *Mus castaneus* F1 hybrid ES cells were grown on a monolayer of mitomycin-inactivated SNLP feeders (STO mouse fibroblasts expressing Neomycin, Puromycin resistance and LIF genes) in otherwise the same conditions as other mouse ES cells. Cells were fed every 24-48 hrs, and passaged using trypsin-EDTA (0.25%, Gibco). For plating or collecting a certain number of cells, a Countess II cell counter (Life Technologies) with default settings was used to count cells. Prior to harvesting *Pcgf6*$^{fl/fl}$ and *Mus domesticus* (129S1) × *Mus castaneus* F1 hybrid ES cells, the feeders were depleted by pre-plating trypsinised cells for 30 mins at 37°C on plates not coated with gelatin, and discarding attached cells. To induce conditional removal of BAP1, PRC1 complexes or PRC1 catalytic activity, cells were treated with 800 nM 4-hydroxytamoxifen (OHT) for the times indicated in Table 2.1. To induce expression of the full-length Xist transgene which was integrated into the *Mus castaneus* allele of chromosome 3 under a TetO promoter, *Mus domesticus* (129S1) × *Mus castaneus* F1 hybrid ES cells (clone P4D7F4) (Pintacuda et al. 2017) were treated with 1.5 $\mu$g/ml doxycycline for 72 hours. The full list of mouse ES cell lines used in this study together with the growth and drug treatment conditions is provided in Table 2.1.

*Drosophila* S2 (SG4) cells were grown adhesively at 25°C in Schneider's Drosophila Medium (Life Technologies), supplemented with 1× penicillin-streptomycin solution (Life Technologies) and 10% heat-inactivated fetal bovine serum (Labtech). Human HEK293T cells were grown at

| Name | Genotype | Feeders | Drug treatment | Source |
|---|---|---|---|---|
| *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ | *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$; *Rosa26::CreERT2* | w/o feeders | 72 hr OHT | This study |
| *Pcgf1*$^{fl/fl}$ | *Pcgf1*$^{fl/fl}$; *Rosa26::CreERT2* | w/o feeders | 72 hr OHT | (Almeida et al. 2017) |
| *Pcgf3*$^{fl/fl}$ | *Pcgf3*$^{fl/fl}$; *Rosa26::CreERT2* | w/o feeders | 72 hr OHT | (Almeida et al. 2017) |
| *Pcgf5*$^{fl/fl}$ | *Pcgf5*$^{fl/fl}$; *Rosa26::CreERT2* | w/o feeders | 72 hr OHT | Provided by Prof. H. Koseki, RIKEN |
| *Pcgf6*$^{fl/fl}$ | *Pcgf6*$^{fl/fl}$; *Rosa26::CreERT2* | on feeders | 72 hr OHT | (Endoh et al. 2017) |
| *Pcgf3/5*$^{fl/fl}$ | *Pcgf3/5*$^{fl/fl}$; *Rosa26::CreERT2* | w/o feeders | 72 hr OHT | Provided by Prof. Haruhiko Koseki, RIKEN |
| *Pcgf1/3/5*$^{fl/fl}$ | *Pcgf1/3/5*$^{fl/fl}$; *Rosa26::CreERT2* | w/o feeders | 72 hr OHT | Provided by Prof. Haruhiko Koseki, RIKEN |
| *Pcgf1/3/5/2*$^{fl/fl}$ | *Pcgf1/3/5/2*$^{fl/fl}$; *Rosa26::CreERT2* | w/o feeders | 72 hr OHT | This study |
| *Pcgf1/3/5/6*$^{fl/fl}$ | *Pcgf1/3/5/6*$^{fl/fl}$; *Rosa26::CreERT2* | w/o feeders | 72 hr OHT | This study |
| *Pcgf4*$^{-/-}$;*Pcgf1/3/5/2*$^{fl/fl}$ | *Pcgf4-/-;Pcgf1/3/5/2*$^{fl/fl}$; *Rosa26::CreERT2* | w/o feeders | 72 hr OHT | This study |
| *Bap1*$^{fl/fl}$ | *Bap1*$^{fl/fl}$; *Rosa26::CreERT2* | w/o feeders | 96 hr OHT | This study |
| *Mus domesticus* (129S1) × *Mus castaneus* F1 with full-length Xist transgene | C57BL/6JJcl x 129/SvJcl with Xist-WT$^{\text{DOXinducible}}$ (P4D7-F4) | on feeders | 72 hr DOX | (Pintacuda et al. 2017) |
| PRC1$^{\text{CKO}}$ | *Ring1a*$^{-/-}$;*Ring1b*$^{fl/fl}$; *Rosa26::CreERT2* | w/o feeders | 72 hr OHT | Generated by Dr. Neil Blackledge, Oxford |
| PRC1$^{\text{CPM}}$ | *Ring1a*$^{I50A/D53K}$; *Ring1B*$^{(WT->I53A/D56K)fl/fl}$; *Rosa26::CreERT2* | w/o feeders | 72 hr OHT | Generated by Dr. Neil Blackledge, Oxford |

**Table 2.1:** Description of the genotype, cell culture conditions, drug treatments and source of origin for mouse ES cell lines used in this study

37°C and 5% CO2, in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (Labtech), 2 mM L-glutamine (Life Technologies), 1x penicillin-streptomycin solution (Life Technologies), and 0.5 mM beta-mercaptoethanol (Life Technologies). SV40-immortalized mouse embryonic fibroblasts (MEFs) (Endoh et al. 2012; Jullien et al. 2017) were grown in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 7% FBS (Labtech) and 1× penicillin-streptomycin solution (Life Technologies) and maintained in culture for up to 10 passages. All cells used in this study were regularly tested for the presence of mycoplasma.

### 2.1.2 Genome engineering by CRISPR/Cas9-mediated Homology-Directed Repair (HDR)

The pSptCas9(BB)-2A-Puro(PX459)-V2.0 vector was obtained from Addgene (CatNo: 62988) and sgRNAs were designed using the CRISPOR online tool (http://crispor.tefor.net). Targeting constructs with appropriate homology arms were generated by Gibson assembly using the Gibson Assembly Master Mix kit (New England Biolabs), or in the case of single LoxP sites with 150 bp homology arms, purchased from GeneArt Gene Synthesis (Invitrogen). In all instances, targeting

constructs were designed such that Cas9 recognition sites were disrupted by the presence of the LoxP site. One well of a 6-well plate of 60-70% confluent mouse ES cells were transfected with 0.5 $\mu$g of each Cas9 guide, and 2 $\mu$g of a targeting construct (where appropriate) using Lipofectamine 3000 (ThermoFisher) according to manufacturer's guidelines. The following morning, cells were passaged at a range of densities and subjected to selection with 1 $\mu$g/ml puromycin for 48 hrs to eliminate any non-transfected cells. Approximately a week later, individual clones were isolated and expanded in 96-well plates. Positive clones with the desired genomic modification were identified using PCR screens of genomic DNA and validated using Sanger sequencing of PCR products.

### 2.1.3 Cell line generation

Mouse ES cells were targeted to produce conditional alleles for *Pcgf1*, *Pcgf3* (Almeida et al. 2017), *Pcgf5*, and *Pcgf6* (Endoh et al. 2017). For each gene, parallel LoxP sites were inserted to flank a critical portion of the gene, removal of which will lead to a frame-shift in the coding sequence, so a functional protein can no longer be produced. For *Pcgf1* - the LoxP sites flanked exons 2-7, for *Pcgf3* - exon 4 containing ATG start codon was flanked, for *Pcgf5* - exon 2 containing ATG start codon was flanked, and for *Pcgf6* - exons 2-3 were flanked. $Pcgf1^{fl/fl}$, $Pcgf3^{fl/fl}$ and $Pcgf5^{fl/fl}$ mouse lines were established by injecting targeted ES cells into 8-cell stage embryos to create chimeric mouse lines. $Pcgf1^{fl/fl}$, $Pcgf3^{fl/fl}$ and $Pcgf5^{fl/fl}$ mouse lines were crossed with *Rosa26::CreERT2* transgenic mice purchased from Artemis Pharmaceuticals, and further intercrossed to generate $Pcgf3/5^{fl/fl}$, and $Pcgf1/3/5^{fl/fl}$ mice. $Pcgf1^{fl/fl}$, $Pcgf3^{fl/fl}$, $Pcgf5^{fl/fl}$, $Pcgf3/5^{fl/fl}$, and $Pcgf1/3/5^{fl/fl}$ ES cells, all encoding Rosa26::CreERT2, were derived from blastocysts as described in Endoh et al. 2008. $Pcgf6^{fl/fl}$ cells were generated as described previously (Endoh et al. 2017). Unless otherwise specified, mouse ES cells were initially maintained on a monolayer of mitomycin-inactivated fibroblast feeders, before being adapted to grow under feeder-free conditions on plates coated with gelatin. All animal experiments were carried out according to the in-house guidelines for the care and use of laboratory animals at the RIKEN, Yokohama, Japan.

$Pcgf4^{-/-};Pcgf2^{fl/fl}$ ES cells were derived from *Rosa26::CreERT2* ES cells by a two-step process. First, exon 2 of both *Pcgf4* alleles was deleted using two flanking Cas9 guides. A validated $Pcgf4^{-/-}$ ES cell clone was then subjected to a second round of genomic engineering to flank the first coding exon of *Pcgf2* (exon 2) with LoxP sites in a parallel orientation using homology arms

of approximately 1 kb and appropriate Cas9 guides. *Pcgf4*$^{-/-}$*;Pcgf2*$^{fl/fl}$ clones were analysed by both RT-qPCR and western blot to confirm removal of PCGF2 in response to OHT treatment.

Both *Pcgf1/3/5/2*$^{fl/fl}$ and *Pcgf1/3/5/6*$^{fl/fl}$ ES cells were derived from the *Pcgf1/3/5*$^{fl/fl}$ cell line by a two-step process, in which upstream and downstream LoxP sites were sequentially inserted using targeting constructs with 150 bp homology arms and appropriate Cas9 guides. For *Pcgf2*, the LoxP sites flanked the first coding exon (exon 2), while for *Pcgf6* exons 2-3 were flanked. Correctly targeted clones were subjected to OHT treatment and analysed by RT-qPCR and western blot to validate PCGF2 or PCGF6 removal. *Pcgf4*$^{-/-}$*;Pcgf1/3/5/2*$^{fl/fl}$ ES cells were derived from the *Pcgf1/3/5/2*$^{fl/fl}$ cell line by deleting exon 2 of both *Pcgf4* alleles using two Cas9 guides flanking the deletion region (as described above for *Pcgf4*$^{-/-}$*;Pcgf2*$^{fl/fl}$ ES cells).

*Bap1*$^{fl/fl}$ ES cells were derived from an E14 mouse ES cell line by a two-step process. First, parallel loxP sites flanking exon 4 were inserted using a targeting construct with homology arms of approximately 1 kb and appropriate Cas9 guides. Second, *CreERT2* was inserted into the *Rosa26* locus using a Rosa26-specific Cas9 guide. Removal of BAP1 in response to OHT treatment in *Bap1*$^{fl/fl}$ ES cells was confirmed using RT-qPCR and western blot analysis.

PRC1$^{CKO}$ (*Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$) ES cells were derived from E14 mouse ES cells. Exons 1-3 of *Ring1a* were first deleted using two Cas9 guides flanking the deletion region, and in parallel *CreERT2* was inserted into the *Rosa26* locus using a Rosa26-specific Cas9 guide. *Ring1a*$^{-/-}$*;Rosa26::CreERT2* ES cells were then subjected to two rounds of genome editing, in which LoxP sites were inserted in a parallel orientation to flank the first coding exon of *Ring1b*. *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ ES cells were validated by RT-qPCR and western blot to confirm constitutive deletion of RING1A and conditional removal of RING1B in response to OHT treatment.

The isogenic PRC1$^{CPM}$ ES cell line was generated using a targeting construct comprising exon 3 of *Ring1b* in forward orientation (flanked by 100 bp of *Ring1b* intron 2/intron 3) followed by a mutant copy of exon 3 (encoding I53A and D56K mutations) in reverse orientation (flanked by splice donor and acceptor sites from mouse *IgE* gene). Both the wild-type and mutant versions of exon 3 were codon optimized at wobble positions to minimize sequence similarity, preventing hairpin formation and allowing the two to be easily distinguished. The exon 3 wild-type/mutant pair was flanked by doubly inverted LoxP/Lox2272 sites and 1 kb homology arms. The RING1B$^{CPM}$ targeting construct was transfected into E14 ES cells in combination with three different Cas9

guides specific for the *Ring1b* locus. Correctly targeted homozygous clones were identified by PCR screening, followed by RT-PCR and sequencing to check for splicing defects. *CreERT2* was then inserted into the *Rosa26* locus using a Rosa26-specific Cas9 guide, and using a similar approach, the I50A/D53K mutation was constitutively knocked-in into both copies of endogenous *Ring1a*. The final PRC1$^{CPM}$ cell line was validated by PCR, RT-qPCR and western blot to confirm constitutive mutation of *Ring1a* and conditional conversion of *Ring1b* into a mutated form following OHT treatment. PRC1$^{CKO}$ and PRC1$^{CPM}$ were kindly provided by Dr. Neil Blackledge. Sequences of all primers used for cloning of targeting constructs and Cas9 guides, as well as for PCR validation of positive clones were deposited into the Klose lab primer database.

### 2.1.4 Cell growth curve assay

For a growth curve assay, $4 \times 10^5$ *Bap1$^{fl/fl}$* ES cells, untreated (UNT) and following 96 hr tamoxifen treatment (OHT), were plated per one well of a 6-well plate in two technical replicates (day 0). Cells were trypsinised and counted in technical duplicates every 24 hr for five following days for a total of five independent biological replicates. For this assay, only live cells were counted using a haemocytometer, with cell viability assessed by staining with 0.4% Trypan Blue (ThermoScientific) labelling dead cells.

## 2.2 Protein methods

The full list of antibodies used in this study, together with dilutions used for western blot analysis, is provided in Table 2.2.

### 2.2.1 Preparation of nuclear extracts

For nuclear extraction, ES cells were washed with 1× PBS and resuspended in 10 volumes of Buffer A (10 mM Hepes pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT, 0.5 mM PMSF and protease inhibitor cocktail (Roche)). After 10 min incubation on ice, cells were recovered by centrifugation at 1500 g for 5 min and resuspended in 3 volumes of Buffer A supplemented with 0.1% NP-40. The released nuclei were pelleted by centrifugation at 1500 g for 5 min, followed by resuspension in 1 volume of Buffer C (5 mM Hepes (pH 7.9), 26% glycerol, 1.5 mM $MgCl_2$, 0.2 mM EDTA, protease inhibitor cocktail (Roche) and 0.5 mM DTT) supplemented with 400 mM NaCl. The extraction was allowed to proceed on ice for 1 hr with occasional agitation, then the

nuclei were pelleted by centrifugation at 16,000 g for 20 min at 4°C. The supernatant was taken as the nuclear extract. The Bradford protein assay was used to compare protein concentrations across samples. Western blot analysis of nuclear extracts was performed using HRP-conjugated secondary antibodies (1:5000). HRP-activity was detected by incubating with equal volumes of ECL solution I (2.5 mM luminal (Sigma), 0.396 mM p-coumaric acid (Sigma) and 100 mM Tris HCl pH 8.5) and ECL solution II (5.6 mM $H_2O_2$ and 100 mM Tris HCl pH 8.5), and imaged by exposing to X-Ray film (GE Healthcare).

### 2.2.2 Preparation of histone extracts

For histone extraction, ES cells were washed with RSB supplemented with 20 mM NEM, incubated on ice for 10 min in RSB with 0.5% NP-40 and 20 mM NEM, pelleted by centrifugation at 800 g for 5 min and incubated in 2.5 mM $MgCl_2$, 0.4 M HCl and 20 mM NEM on ice for 30 min. After that, cells were pelleted by centrifugation at 16,000 g at 4°C for 20 min, the supernatant recovered and precipitated on ice with 25% TCA for 30 min, followed by centrifugation at 16,000 g for 15 min at 4°C to recover histones. Following two acetone washes, the histones were resuspended in 150 $\mu$l 1xSDS loading buffer and boiled at 95°C for 5 min. Finally, any insoluble precipitate was pelleted by centrifugation at 16,000 g for 15 min at 4°C and the soluble fraction retained as the histone extract. Histone concentrations across samples were compared by Coomassie Blue staining following SDS-PAGE. Semi-quantitative western blot analysis of histone extracts was performed using LI-COR IRDye secondary antibodies (1:15,000) and imaging was done using the LI-COR Odyssey Fc system. To measure the changes in bulk H2AK119ub1 levels, the relative signal of H2AK119ub1 to H3 or H4 histones was quantified.

### 2.2.3 Co-immunoprecipitation (co-IP)

For co-immunoprecipitation, 400 $\mu$g of nuclear extract was added to BC150 buffer (150 mM KCl, 10% glycerol, 50 mM HEPES pH 7.9, 0.5 mM EDTA, 0.5 mM DTT) with 1× protease inhibitor cocktail (Roche) to a total volume of 550 $\mu$l. A 50 $\mu$l Input sample was retained, and 5 $\mu$g of anti-RING1B antibody (Atsuta et al. 2001) (see Table 2.2) was added to the remaining 500 $\mu$l of sample. IP reactions were then incubated overnight at 4°C. Immunoprecipitated material was collected with Protein A agarose beads and washed four times in 1 ml of BC150 buffer. Following the final wash step, beads were directly resuspended in 100 $\mu$l of 1× SDS loading buffer (2% SDS,

0.1 M Tris pH 6.8, 0.1 M DTT, 10% glycerol, 0.1% bromophenol blue) and placed at 95°C for 5 mins. 1× SDS loading buffer was similarly added to Input samples which were also incubated at 95°C for 5 mins, prior to SDS-PAGE and western blot analysis.

| Antibody | Source | Western blot (Dilution) | ChIP ($\mu$l per IP) |
|---|---|---|---|
| Rabbit monoclonal anti-H2AK119ub1 | CST Cat#:8240 | 1:2500 | 5 $\mu$l |
| Rabbit polyclonal anti-H3K27me3 | In house (Rose et al. 2016) | N/A | 5 $\mu$l |
| Rabbit monoclonal anti-H3K27me3 | CST Cat#:9733 | 1:2500 | N/A |
| Rabbit polyclonal anti-H3K4me3 | In house (Farcas et al. 2012) | N/A | 3 $\mu$l |
| Mouse monoclonal anti-H3 | CST Cat#:3638 | 1:5000 | N/A |
| Mouse monoclonal anti-H4 | CST Cat#:2935 | 1:5000 | N/A |
| Rabbit polyclonal anti-H2A | Millipore 07-146 | 1:5000 | N/A |
| Rabbit monoclonal anti-RING1B | CST Cat#:5694 | N/A | 3 $\mu$l |
| Mouse monoclonal anti-RING1B | Atsuta et al. 2001 | 1:1000 | N/A |
| Rabbit polyclonal anti-PCGF6 | LifeSpan BioSciences LS-C482495 | 1:1000 | 3 $\mu$l |
| Rabbit monoclonal anti-PCGF3+PCGF5 | Abcam ab201510 | 1:1000 | N/A |
| Rabbit polyclonal anti-PCGF2 (Mel-18) | Santa Cruz sc-10744 | 1:1000 | 3 $\mu$l |
| Rabbit monoclonal anti-PCGF4 (Bmi1) | CST Cat#:6964 | 1:1000 | N/A |
| Rabbit polyclonal anti-PCGF1 | In house (Blackledge et al. 2014) | 1:1000 | 5 $\mu$l |
| Mouse monoclonal anti-PHC1 | CST Cat#:13768 | 1:1000 | 6 $\mu$l |
| Rabbit polyclonal anti-CBX7 | Millipore 07-981 | 1:1000 | N/A |
| Rabbit polyclonal anti-CBX7 | Abcam ab21873 | N/A | 4 $\mu$l |
| Rabbit polyclonal anti-KDM2B | In house (Farcas et al. 2012) | 1:1000 | N/A |
| Rabbit polyclonal anti-RYBP (DEDAF) | Millipore AB3637 | 1:1000 | N/A |
| Rabbit monoclonal anti-EZH2 | CST Cat#:5246 | 1:2000 | N/A |
| Rabbit monoclonal anti-SUZ12 | CST Cat#:3737 | 1:2000 | 3 $\mu$l |
| Rabbit monoclonal anti-JARID2 | CST Cat#:13594 | 1:1000 | 3 $\mu$l |
| Rabbit polyclonal anti-PCL2 (M96) | GenWay GWB-FA7207 | 1:1000 | 2 $\mu$l |
| Rabbit monoclonal anti-BRG1 | Abcam ab110641 | 1:5000 | N/A |
| Rabbit monoclonal anti-BAP1 | CST #13271 | 1:1000 | N/A |

**Table 2.2: Description of antibodies used in this study for western blot and cChIP-seq analysis**
Dilution of the antibodies for western blot and amount of the antibody used per one IP (in $\mu$l) for ChIP-seq is specified.

## 2.3 Chromatin immunoprecipitation

### 2.3.1 Calibrated ChIP-sequencing (cChIP-seq)

For cChIP-seq analysis of PRC1 and PRC2 subunits, 5 x $10^7$ mouse ES cells (both untreated and following OHT treatment) were mixed with 2 x $10^6$ human HEK293T cells. Cells were resuspended in 10 ml 1× PBS and crosslinked at 25°C with 2 mM DSG (Thermo Scientific) for

45 minutes, and then with 1% formaldehyde (methanol-free, Thermo Scientific) for a further 15 minutes. Reactions were quenched with 125 mM glycine. Crosslinked cells were incubated in lysis buffer (50 mM HEPES pH 7.9, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP40, 0.25% Triton X-100) for 10 minutes at 4°C. Released nuclei were washed (10 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA) for 5 minutes at 4°C. Chromatin was then resuspended in 1 ml of sonication buffer (10 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na deoxycholate, 0.5% N-lauroylsarcosine) and sonicated for 28-30 minutes using a BioRuptor Pico (Diagenode), shearing genomic DNA to an average size of 0.5 kb. Following sonication, Triton X-100 was added to a final concentration of 1%.

For ChIP, sonicated chromatin was diluted 10-fold in ChIP dilution buffer (1% Triton X-100, 1 mM EDTA, 20 mM Tris-HCl pH 8.0, 150 mM NaCl) and pre-cleared for 1 hr using Protein A agarose beads (Repligen) blocked with 1 mg/ml BSA and 1 mg/ml yeast tRNA. For each ChIP reaction, 1 ml of diluted and pre-cleared chromatin was incubated overnight with the appropriate amount of antibody (see Table 2.2 for details). Antibody-bound chromatin was captured using blocked protein A agarose for 1 hr at 4°C and collected by centrifugation. ChIP washes were performed as described previously (Farcas et al. 2012). ChIP DNA was eluted in elution buffer (1% SDS, 0.1 M NaHCO3) and cross-links were reversed overnight at 65°C with 200 mM NaCl and 2 $\mu$l RNase A (Sigma). A matched input sample (corresponding to 10% of original ChIP reaction) was treated identically. The following day, ChIP samples and Inputs were incubated with Proteinase K (Sigma) for 1.5 hrs at 56°C and purified using ChIP DNA Clean and Concentrator Kit (Zymo Research).

cChIP-seq libraries for both ChIP and Input samples were prepared using NEBNext Ultra DNA Library Prep Kit for Illumina following manufacturer's guidelines. Samples were indexed using NEBNext Multiplex Oligos. The average size and concentration of all libraries was analysed using 2100 Bioanalyzer High Sensitivity DNA Kit (Agilent) followed by qPCR using SensiMix SYBR (Bioline, UK) and KAPA Illumina DNA standards (Roche).

### 2.3.2 Native cChIP-sequencing

For native cChIP-seq analysis of histone modifications, 5 x $10^7$ mouse ES cells (both untreated and following OHT/DOX treatment) were mixed with 2 x $10^7$ *Drosophila* SG4 cells in 1× PBS. Mixed cells were pelleted and nuclei were released by resuspending in ice cold lysis buffer (10 mM Tris-

HCl pH 8.0, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% NP40, 5 mM N-ethylmaleimide). Nuclei were then washed, and resuspended in 1 ml of MNase digestion buffer (10 mM Tris-HCl pH 8.0, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% NP40, 0.25 M sucrose, 3 mM CaCl2, 10 mM N-ethylmaleimide, 1× protease inhibitor cocktail (Roche)). Each sample was incubated with 200 units of MNase (Fermentas) at 37°C for 5 min, followed by the addition of 4 mM EDTA to halt MNase digestion. Following centrifugation at 1500 g for 5 min at 4°C, the supernatant (S1) was retained. The remaining pellet was incubated with 300 $\mu$l of nucleosome release buffer (10 mM Tris-HCl pH 7.5, 10 mM NaCl, 0.2 mM EDTA, 1× protease inhibitor cocktail (Roche), 10 mM N-ethylmaleimide) at 4°C for 1 hr, passed five times through a 27G needle using a 1 ml syringe, and spun at 1500 g for 5 min at 4°C. The second supernatant (S2) was collected and combined with corresponding S1 sample from above. A small amount of S1/S2 DNA was purified and visualized on a 1.5% agarose gel to confirm digestion to mostly mono-nucleosomes.

For ChIP, S1/S2 nucleosomes were diluted 10-fold in native ChIP incubation buffer (70 mM NaCl, 10 mM Tris pH 7.5, 2 mM $MgCl_2$, 2 mM EDTA, 0.1% Triton X-100, 1x protease inhibitor cocktail (Roche), 10 mM N-ethylmaleimide (NEM)), and 1 ml aliquots were made. Each ChIP reaction was then incubated overnight at 4°C with the appropriate amount of antibody (see Table 2.2 for details). Antibody-bound nucleosomes were captured using protein A agarose (Repligen) beads, pre-blocked in native ChIP incubation buffer supplemented with 1 mg/ml BSA and 1 mg/ml yeast tRNA for 1 hr at 4°C and collected by centrifugation. Immunoprecipitated material was washed four times with Native ChIP wash buffer (20 mM Tris pH 7.5, 2 mM EDTA, 125 mM NaCl, 0.1% Triton X-100) and once with Tris-EDTA buffer (10 mM Tris pH 8.0, 1 mM EDTA). ChIP DNA was eluted using 100 $\mu$l of elution buffer (1% SDS, 0.1 M NaHCO3), and then purified using ChIP DNA Clean and Concentrator kit (Zymo Research). For each ChIP sample, DNA from a matched Input control (corresponding to 10% of original ChIP reaction) was also purified. Native cChIP-seq library preparation was performed as described above for cChIP-seq.

## 2.4 RNA isolation and sequencing

### 2.4.1 Calibrated nuclear RNA-seq (cnRNA-seq)

For isolation of RNA for cnRNA-seq, $1 \times 10^7$ mouse ES cells (both untreated and following OHT treatment) were mixed with $4 \times 10^6$ *Drosophila* SG4 cells in 1× PBS. Nuclei were isolated in 1

ml HS Lysis buffer (50 mM KCl, 10 mM $MgSO_4.7H_20$, 5 mM HEPES, 0.05% NP40 (IGEPAL CA630)), 1 mM PMSF, 3 mM DTT) for 1 min at room temperature. They were then recovered by centrifugation at 1000 × g for 5 min at 4°C, followed by a total of three washes with ice-cold RSB buffer (10 mM NaCl, 10 mM Tris pH 8.0, 3 mM $MgCl_2$). Nuclei integrity was assessed using 0.4% Trypan Blue staining (ThermoScientific). Next, nuclei were resuspended in 1 ml of TRIzol reagent (ThermoScientific) and RNA was extracted according to the manufacturer's protocol, followed by treatment with the TURBO DNA-free Kit (ThermoScientific). Quality of RNA was assessed using 2100 Bioanalyzer RNA 6000 Pico kit (Agilent). Next, RNA samples were depleted of rRNA using the NEBNext rRNA Depletion kit (NEB). RNA-seq libraries were prepared using the NEBNext Ultra Directional RNA Library Prep kit (NEB). To quantitate the consistency of spike-in cell mixing for each individual sample, a small aliquot of nuclei was used to isolate genomic DNA using phenol-chloroform extraction. This was followed by sonication of DNA for 13-15 min using a BioRuptor Pico sonicator (Diagenode), shearing genomic DNA to an average size of less than 1 kb. Libraries from sonicated genomic DNA were constructed as described above for cChIP-seq.

### 2.4.2 Calibrated total RNA-seq (cRNA-seq)

For cRNA-seq, 1 x $10^7$ mouse ES cells (both untreated and following OHT treatment) were mixed with 4 x $10^6$ *Drosophila* SG4 cells in 600 $\mu$l of 1× PBS. For RNA extraction, 400 $\mu$l of cells was used (corresponding to 6.7 x $10^6$ mouse ES cells), and for DNA extraction the remaining 200 $\mu$l of cells was used (corresponding to 3.3 x $10^6$ mouse ES cells). RNA extraction was performed using RNeasy Mini kit (Qiagen) following manufacturer's guidelines, and 10 $\mu$g was subjected to Turbo DNase (ThermoFisher) treatment to remove any contaminating DNA. Quality of RNA was assessed using 2100 Bioanalyzer RNA 6000 Pico kit (Agilent). Next, RNA samples were depleted of rRNA using the NEBNext rRNA Depletion kit (NEB). RNA-seq libraries were prepared using the NEBNext Ultra II Directional RNA Library Prep kit (NEB). To quantitate the consistency of spike-in cell mixing for each individual sample, a matched sample of cells was used to isolate genomic DNA using Quick-DNA miniprep kit (Zymo Research). Libraries from gDNA were prepared using NEBNext Ultra II FS kit (NEB) following manufacturer's guidelines.

## 2.5 Massively parallel sequencing, data processing, normalisation and analysis

### 2.5.1 Massively parallel sequencing

Libraries for cChIP-seq were sequenced as 40 bp paired-end reads on Illumina NextSeq 500 platform for a minimum of two but typically at least three biological replicates. RNA and DNA libraries for cnRNA-seq and cRNA-seq were sequenced as 80 bp paired-end reads on the Illumina NextSeq 500 platform in biological triplicate. The majority of NGS data included in this thesis was generated on the sequencing platform located at the Department of Zoology, University of Oxford with the help of Amanda Williams.

### 2.5.2 Read alignment

For calibrated ChIP-seq, paired-end reads were aligned to the concatenated mouse and spike-in genome sequences (mm10+dm6 for native cChIP-seq, and mm10+hg19 for cross-linked cChIP-seq) using Bowtie 2 (Langmead and Salzberg 2012) with the "–no-mixed" and "–no-discordant" options. Reads that were mapped more than once were discarded, followed by removal of PCR duplicates with SAMTools (Li et al. 2009b) for native cChIP-seq or Sambamba for cross-linked cChIP-seq (Tarasov et al. 2015).

For cRNA-seq (both total and nuclear), first, paired-end reads were aligned using Bowtie 2 (with "–very-fast", "–no-mixed" and "–no-discordant" options) against the concatenated mm10 and dm6 rDNA genomic sequence (GenBank: BK000964.3 and M21017.1) and reads mapping to rDNA were discarded. All unmapped reads were then aligned against the concatenated mm10 and dm6 genome sequences using STAR (Dobin et al. 2013). Finally, reads that failed to map using STAR were aligned against the mm10+dm6 concatenated genome using Bowtie 2 (with "–sensitive-local", "–no-mixed" and "–no-discordant" options) to improve mapping of introns. Uniquely aligned reads from the last two steps were combined for further analysis. PCR duplicates were removed using Sambamba. For the corresponding gDNA-seq experiments, paired-end read alignment and processing was carried out as described above for cChIP-seq.

Allele-specific analysis of H2AK119ub1 cChIP-seq in the *Mus domesticus* (129S1) × *Mus castaneus* F1 hybrid ES cell line with inducible full-length Xist transgene was performed as

described previously with minor alterations (Pintacuda et al. 2017). Briefly, paired-end reads were aligned using STAR (with "–outFilterMultimapNmax 1", "–outFilterMismatchNmax 2", "–alignEndsType EndToEnd" parameters) against the concatenated mm10+dm6 genome, in which polymorphic SNPs for 129S1 and CAST mouse strains were N-masked. Reads that mapped more than once were discarded and PCR duplicates were removed with Sambamba.

### 2.5.3 Data normalisation and visualisation

For visualisation of cChIP-seq and cRNA-seq data and annotation of genomic regions with read counts, uniquely aligned mm10 reads were randomly subsampled using factors that reflected the total number of dm6 (or hg19) reads in each sample. To account for any minor variation in spike-in cell mixing in different biological replicates, the downsampling factors were additionally corrected using the ratio of dm6 (or hg19)/mm10 total read counts in corresponding Input samples. For cnRNA-seq, corresponding genomic DNA-seq samples were used to calculate the ratio of dm6/mm10 total read counts. The following formula was used for calculation of downsampling factors:

$$Downsampling\ factor = \alpha \times \frac{1}{N(spikeIn\ in\ ChIP/RNA)} \times \frac{N(spikeIn\ in\ Input)}{N(mouse\ in\ Input)}$$

where

*N (spikeIn in ChIP/RNA)* - total number of reads aligning to the spike-in genome for each ChIP-seq/RNA-seq sample

*N (spikeIn in Input)* - total number of reads aligning to the spike-in genome in the corresponding Input sample

*N (mouse in Input)* - total number of reads aligning to the mouse genome in the corresponding Input sample

$\alpha$ - coefficient applied for all files normalised together so the value of the largest downsampling factor equals 1

For allele-specific analysis of H2AK119ub1 cChIP-seq, reads uniquely mapped to the mm10 genome were randomly subsampled using downsampling factors calculated based on the spike-in calibration as described above. SNPsplit (Krueger and Andrews 2016) was used with the paired-end mode to separate reads specifically mapping to the 129S1 and CAST alleles.

After normalisation, read coverages for individual biological replicates were compared across regions of interest (RING1B peaks for cChIP-seq or gene bodies for cRNA-seq) using multiBamSummary and plotCorrelation functions from deepTools suite (Ramirez et al. 2014). For each condition, biological replicates correlated well with each other (Pearson correlation coefficient>0.9) and were merged for downstream applications. Genome coverage tracks were generated using the pileup function from MACS2 (Zhang et al. 2008) for cChIP-seq and genomeCoverageBed from BEDTools (Quinlan 2014) for cRNA-seq and visualised using the UCSC genome browser (Kent et al. 2002). BigwigCompare from deeptools (v3.1.1) (Ramirez et al. 2014) was used to make differential genome coverage tracks.

### 2.5.4 Peak calling

For Chapters 3-6 (with the exception of section 5.2.4), to identify PRC1- or PRC2-bound regions, I used cChIP-seq data for RING1B and SUZ12 respectively, which were obtained from *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ and *Pcgf1/3/5*$^{fl/fl}$ ES cells. RING1B and SUZ12 peak sets were generated using MACS2 ("BAMPE" and "–broad" options specified), with corresponding Input samples used for background normalisation. For each factor, a set of peaks identified in all biological replicates and in both cell lines was used for further analysis. In total, 8833 RING1B peaks and 5965 SUZ12 peaks were identified. Classical Polycomb chromatin domains were defined by k-means clustering of ± 10 kb windows around the centre of RING1B peaks, which was based on RING1B and SUZ12 cChIP-seq signal in untreated *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ ES cells using deeptools. This generated three clusters, two of which displayed high enrichment of both RING1B and SUZ12 and were combined into one set of genomic intervals referred to as classical Polycomb chromatin domains (n = 2096).

For Chapter 5 (section 5.2.4) and Chapter 7, genomic regions bound by RING1B, SUZ12 and PCGF2 were also identified by peak calling using MACS2 ("BAMPE" and "–broad" options specified), with corresponding Input samples used as a control. For RING1B and SUZ12, peaks were called using merged biological replicates from untreated PRC1$^{CPM}$ and PRC1$^{CKO}$ cells for RING1B and SUZ12 cChIP-seq respectively, and only peaks identified in both cell lines were selected. For PCGF2, individual replicates from untreated PRC1$^{CPM}$ ES cells were used, and only peaks identified in all biological replicates were taken forward. In total, I identified 18,643 RING1B peaks, 7438 SUZ12 peaks and 3680 PCGF2 peaks. Based on the overlap between these

peak sets, Polycomb (PcG)-occupied regions were defined as RING1B peaks that overlap SUZ12 peaks (n = 7074), and PCGF2 target sites as RING1B peaks overlapping PCGF2 peaks (n = 3568).

For all peak sets, peaks overlapping with a custom-build set of blacklisted genomic regions were discarded to remove sequencing artefacts. For RING1B-bound regions, RING1B cChIP-seq in PRC1$^{CKO}$ ES cells was used to filter out peaks which showed no significant loss of RING1B signal following OHT treatment (p-adj$<$0.05 and FC$>$2-fold). To characterise the genome-wide blanket of H2AK119ub1, a set of 100 kb windows spanning the genome (n = 27,348) was generated using makewindows function from BEDtools. NMIs (n = 27,047) were identified using MACS2 peak calling from BioCAP-seq data (Long et al. 2013) as regions which are enriched with non-methylated CpG-rich DNA.

### 2.5.5 Read count quantitation and analysis

#### 2.5.5.1 Read count quantitation

For cChIP-seq, genomic regions of interest were annotated with read counts using a custom-made Perl script for Chapters 3-6 or multiBamSummary from deeptools ( "–outRawCounts") for Chapter 7. For comparative box plot or correlation analysis, read counts from merged spike-in normalised replicates were used, while for differential enrichment analysis, read counts from individual biological replicates prior to spike-in normalisation were obtained. For differential gene expression analysis in Chapters 3-6 and Chapter 7, the same respective read count quantitation approaches were applied for a custom-built, non-redundant mm10 gene set. Briefly, mm10 refGene genes were filtered to remove very short genes with poor sequence mappability and highly similar transcripts, which resulted in the final non-redundant set of 20,633 genes.

#### 2.5.5.2 Box plot analysis

Normalised read counts and log2 fold changes for different genomic intervals were visualised using custom R scripts and ggplot2. For box plot analysis of cChIP-seq signal for different factors before and after treatment, read counts were normalised to the genomic region size (in kb) and median value of cChIP-seq signal in untreated cells, log2 of which was set to 1, unless otherwise specified. For box plots comparing H2AK119ub1 enrichment at RING1B-bound sites and 100 kb genomic windows, read density is shown relative to the median signal at RING1B-bound sites in untreated cells. Size of the boxes for box plots indicates interquartile range (IQR) and whiskers

extend by 1.5×IQR.

#### 2.5.5.3 Percentile analysis

To study the relationship between the levels of different factors/histone modifications and gene expression at RING1B-bound genomic regions, the RING1B-bound sites that overlapped with gene promoters were divided into percentiles based on the expression level of the associated gene in untreated PRC1$^{CPM}$ cells. For each percentile, mean read density normalised to the genomic region size (in kb) was plotted relative to the mean read density in the fiftieth percentile, together with loess regression trend lines.

#### 2.5.5.4 Correlation and statistical analysis

*Ggcor* function from the *GGally* (v1.4.0) R package was used to generate a correlation matrix for the association of RING1B occupancy before and after OHT treatment in PRC1$^{CPM}$ cells with a set of chromatin features and occupancy of other Polycomb factors. All correlation analyses in this study used Pearson correlation coefficient to measure the strength of the association between the variables and were visualised using scatterplots coloured by density with stat_density2d. Linear regression lines were plotted using stat_poly_eq function from the *ggpmisc* (v0.3.1) R package, together with the model's $R^2$ coefficient of determination. Student's t-test and Wilcoxon rank sum statistical tests were also performed in R with samples considered to be independent and two-sided alternative hypothesis, unless otherwise specified.

#### 2.5.5.5 Heatmap and metaplot analysis

For cChIP-seq, computeMatrix and plotProfile/plotHeatmap from deeptools were used to perform metaplot and heatmap analysis of read density at regions of interest. For each cChIP-seq dataset and each cell line, the maximum read density around the peak centre in untreated cells was set to 1. For chromosome-wide density plots, read coverage in 250 kb bins was calculated using a custom R script utilising *GenomicRanges*, *GenomicAlignments* and *Rsamtools* Bioconductor packages (Huber et al. 2015) and visualised using *ggplot2*.

### 2.5.6 Differential enrichment and gene expression analysis

To identify significant changes in cChIP-seq and cRNA-seq, a custom R script incorporating spike-in calibration into DESeq2 (Love et al. 2014) analysis was used. To achieve this, read counts from the spike-in genome at a control set of intervals were supplied to calculate DESeq2 size factors for normalisation of raw mm10 read counts, similarly to the approach described in (Taruttis et al. 2017). For spike-in normalisation of cRNA-seq, unique dm6 genes from refGene were used to calculate the size factors. For differential enrichment analysis of cChIP-seq in Chapters 3-6, unique dm6 refGene genes were also used for native cChIP-seq, while a set of hg19 CpG island regions obtained from UCSC Table Browser was used for RING1B and SUZ12 cChIP-seq. In Chapter 7, spike-in normalisation size factors were calculated using 10 kb windows (for cChIP-seq of PRC1/PRC2 factors) or 1 kb windows (for cChIP-seq of histone modifications) spanning the spike-in genome. Prior to quantitation, spike-in reads were pre-normalised using the spike-in/mouse read ratio derived from the corresponding Input (for cChIP-seq) or genomic DNA-seq (for cRNA-seq) sample in order to account for minor variations in mixing of mouse and spike-in cells. For a change to be called significant, a threshold of p-adj $<0.05$ and fold change $>1.5$ was applied unless otherwise specified. Log2-fold change values were visualised using R and *ggplot2* with MA plots and violin plots. For MA plots, density of the data points across y-axis is shown to reflect the general direction of gene expression changes. UpSet plots were generated using the UpSetR package available on CRAN (Lex et al. 2014).

### 2.5.7 Functional annotation of genes

#### 2.5.7.1 Annotation of Polycomb target genes

Mouse genes from a custom-build non-redundant set (n = 20,633) were classified into three categories based on the overlap of their promoters with non-methylated CpG islands (NMI), as well as RING1B- and SUZ12- bound sites. For Chapters 3-6 (with the exception of section 5.2.5), PcG-occupied genes were defined as genes, for which corresponding TSS regions (± 2500 bp) overlapped with stringent RING1B and SUZ12 peak sets identified in *Pcgf4*$^{-/-}$*;Pcgf2*$^{fl/fl}$ and *Pcgf1/3/5*$^{fl/fl}$ ES cells. These genes (n = 4868) represent *bona fide* Polycomb target genes. Furthermore, genes that did not overlap with genomic regions bound by both PRC1 and PRC2 but contain an NMI were classified as non-PcG-occupied genes (n = 9898). Finally, genes that did not

overlap with an NMI were referred to as non-NMI genes (n = 5867). To compare gene expression changes across distinct PCGF conditional knockout cell lines, I isolated PRC1 target genes (n = 2071), which were defined as a subset of PcG-occupied genes showing a statistically significant increase in gene expression following removal of PRC1 in $PRC1^{CKO}$ cells.

For Chapter 5 (section 5.2.5) and Chapter 7, a similar approach for gene classification was used with minor alterations. First, all genes with promoters (TSS ± 2500 bp) not overlapping with NMIs were referred to as non-NMI genes (n = 5326). Genes that contained NMIs at their promoters were further sub-divided into PcG-occupied genes (n = 6283), if their promoters also overlapped with both RING1B and SUZ12-bound sites defined in $PRC1^{CKO}$ and $PRC1^{CPM}$ cells, and non-PcG-occupied genes (n = 9024), if they did not. For genes with several described transcripts and promoters, a complete mm10 refGene gene set was used for classification, with the overlap of at least one promoter with the feature of interest being required to assign a gene into PcG/non-PcG occupied categories. Using this new classification, PRC1-repressed genes (n = 2482) were defined as a subset of PcG-occupied genes which showed a statistically significant increase in gene expression following removal of PRC1 in $PRC1^{CKO}$ cells.

#### 2.5.7.2 Annotation of differentially upregulated genes in distinct PCGF conditional knockout ES cell lines

To compare occupancy of individual PCGF proteins at promoters of variant PRC1-regulated genes, PCGF1-sensitive genes were defined as genes significantly upregulated following OHT treatment in $Pcgf1^{fl/fl}$ cells. PCGF1/3/5-sensitive genes were defined as the set of genes significantly upregulated following OHT treatment in $Pcgf1/3/5^{fl/fl}$ cells, excluding the PCGF1-sensitive set of genes. Similarly, PCGF1/3/5/6-sensitive genes were defined as the set of genes significantly upregulated following OHT treatment in $Pcgf1/3/5/6^{fl/fl}$ cells, excluding the PCGF1- and PCGF1/3/5-sensitive sets of genes.

To identify genes that were further reactivated following removal of PCGF2 in combination with PCGF1/3/5 in $Pcgf1/3/5/2^{fl/fl}$ cells, I performed differential analysis comparing gene expression in the OHT-treated $Pcgf1/3/5^{fl/fl}$ and $Pcgf1/3/5/2^{fl/fl}$ cells. This was done using a traditional DESeq2 approach without spike-in normalisation. PRC1-bound sites that retained RING1B following removal of PCGF1/3/5/6 were identified as sites with no significant decrease in RING1B signal by cChIP-seq (sites with p-adj$<$0.05 and fold change$<$0 were excluded), which overlapped with

promoters (TSS ± 2500 bp) of PRC1 target genes characterised earlier. To identify genes that depend on PCGF6 for normal level of H2AK119ub1 at their promoters, PRC1-bound sites that displayed a significant decrease (p-adj<0.05 and fold change<0) in H2AK119ub1 following OHT treatment in *Pcgf6*$^{fl/fl}$ cells were overlapped with promoters (TSS ± 2500 bp) of genes from a custom-build non-redundant mm10 set.

#### 2.5.7.3 Annotation of differentially upregulated genes in PRC1$^{CPM}$ and PRC1$^{CKO}$ ES cell lines

In order to identify genes that were differentially affected by removal of PRC1 or specifically PRC1 catalytic activity, two complementary approaches were used. First, I isolated genes that were expressed at significantly lower levels (p-adj<0.05 and fold change>1.5) in the OHT-treated PRC1$^{CPM}$ as compared to the OHT-treated PRC1$^{CKO}$ cells. In addition, I also identified genes, for which the magnitude of expression changes following OHT treatment was significantly smaller in PRC1$^{CPM}$ cells than in PRC1$^{CKO}$ cells, using DESeq2 design that included the interaction term between the cell line and treatment factors. In both cases, a custom R script incorporating spike-in calibration into DESeq2 analysis was used. Combination of these two approaches has yielded a total of 241 PRC1-repressed genes that were derepressed to a smaller extent in PRC1$^{CPM}$ cells as compared to PRC1$^{CKO}$ cells following OHT treatment.

#### 2.5.7.4 Gene ontology (GO) term enrichment analysis

HOMER (Heinz et al. 2010) was used to perform gene ontology (GO) term enrichment analysis for differentially regulated genes or genes associated with differentially enriched genomic regions, with the complete non-redundant mm10 gene set used as a background.

### 2.5.8 Published datasets used in this study

Description of the previously published datasets which were used for analysis in this study is provided in Table 2.3.

| Dataset | Accession code | Reference |
|---|---|---|
| BioCAP-seq | GSE43512 | Long et al. 2013 |
| RNA-seq in *Pcgf6*$^{fl/fl}$ ES cells (untreated and OHT-treated) | GSE84480 | Endoh et al. 2017 |
| 4sU RNA-seq in *Mus domesticus* (129S1) x *Mus castaneus* F1 ESCs with full-length Xist transgene (untreated and DOX-treated) | GSE103370 | Pintacuda et al. 2017 |
| 4sU RNA-seq in ES cells following RA-induced differentiation | GSE98756 | Dimitrova et al. 2018 |
| KDM2B ChIP-seq | GSE55698 | Blackledge et al. 2014 |
| PCGF5 ChIP-seq | GSE107377 | Yao et al. 2018 |
| PCGF6 and L3MBTL2 ChIP-seq for Chapter 6 | ArrayExpress E-MTAB-6007 | Stielow et al. 2018 |
| MAX ChIP-seq | GSE48175 | Krepelova et al. 2014 |
| H2AK119ub1 ChIP-seq in *iXist-Chr3-TIR-Pcgf3/5DKO* ES cells (untreated and DOX-treated) | GSE119607 | Nesterova et al. 2019 |

**Table 2.3:** Accession codes for the previously published datasets used in this study

# 3

# CANONICAL PRC1 COMPLEXES ARE DISPENSABLE FOR POLYCOMB-MEDIATED GENE REPRESSION

## 3.1 Introduction

A family of evolutionary conserved Polycomb group (PcG) proteins are chromatin-based transcriptional repressors which are essential for normal regulation of gene expression during development. The first PcG gene, *Polycomb* (*Pc*), was discovered more than 70 years ago in *Drosophila* as a repressor of homeotic genes, which control normal body plan specification (Lewis 1947; Lewis 1978). Genetic and biochemical dissection of the PcG protein family later revealed that PC interacts with other PcG proteins, PH, PSC and dRING, to form what is now known as canonical PRC1 complex (cPRC1) (Jürgens 1985; Gunster et al. 1997; Shao et al. 1999; Saurin et al. 2001; Francis et al. 2001; Levine et al. 2002). Subsequently, it was demonstrated that the PC protein contains a chromodomain which can recognise H3K27me3 (Paro and Hogness 1991; Eissenberg 2001; Fischle et al. 2003; Min et al. 2003), and PH contains a SAM domain which can self-polymerase to form higher-order chromatin structures (Kim et al. 2002; Robinson et al. 2012; Isono et al. 2013). Based on these early observations, the Polycomb system has been proposed to function through a PRC2-initiated hierarchical pathway, in which initial recruitment of PRC2 to target genes leads to deposition of H3K27me3 and binding of canonical PRC1 (Cao et al. 2002; Wang et al. 2004b). In the context of this model, canonical PRC1 complexes elicit gene repression through formation of compact higher-order chromatin structures at Polycomb target gene promoters, rendering them inaccessible to the transcriptional machinery (Francis et al. 2001; Francis et al. 2004; Grau et al. 2011; Eskeland et al. 2010a; Isono et al. 2013; Schoenfelder et al. 2015; Boettiger et al. 2016; Kundu et al. 2017; Lau et al. 2017). However, the extent to which this cPRC1-centred model could explain gene repression by the Polycomb system, especially in mammalian cells, still remains to be defined.

Previous attempts to determine the role of canonical PRC1 complexes in Polycomb system

function in mammals revealed that a relatively small number of Polycomb target genes were derepressed in cPRC1-deficient cells (Morey et al. 2012; Morey et al. 2015; Kundu et al. 2017). Although these observations did not support the central role of canonical PRC1 in Polycomb-mediated gene repression, the classical cPRC1-centred model is still viewed as one of the key mechanisms of gene silencing by the Polycomb system (Aranda et al. 2015; Talbert et al. 2019). However, previous studies were inherently limited as they relied on constitutive knockdown or knockout systems to remove individual cPRC1 subunits. In addition to being insufficient to completely ablate canonical PRC1, these approaches also involved prolonged culture of cPRC1-deficient cells, and this may have masked the full effect of cPRC1 perturbations due to the compensatory selection. Here, I set out to overcome these limitations in order to define the contribution of canonical PRC1 to Polycomb-mediated gene repression. To achieve this, I establish a new mouse ES cell line, in which removal of all canonical PRC1 complexes can be induced. I then exploit this conditional knockout system, in combination with quantitative genomic approaches, to examine the primary transcriptional changes following loss of canonical PRC1. Finally, I compare the effect of canonical PRC1 removal on gene expression to the transcriptional defects that arise when all PRC1 complexes are lost. This approach will allow me to test the cPRC1-centric model of gene repression by the Polycomb system.

## 3.2 Results

### 3.2.1 A mouse ES cell line to conditionally remove canonical PRC1

To define the role of canonical PRC1 in Polycomb-mediated gene repression, I set out to generate a new mouse ES cell line in which removal of canonical PRC1 complexes can be induced. I chose mouse ES cells as a model system for my studies, because they are highly amenable to genetic manipulations and PcG proteins have been extensively characterised in these cells. In mammals, canonical PRC1 can form around either PCGF2 or PCGF4 subunits, with PCGF2 accounting for the majority of canonical PRC1 complexes in mouse ES cells (Figure 3.1A) (Kloet et al. 2016; Hauri et al. 2016). Although expression of PCGF4 in this cell type is normally low, it increases when PCGF2 is depleted (Morey et al. 2015). Therefore, in order to mitigate against any possible compensation from PCGF4 and achieve complete removal of canonical PRC1, I generated a cell line in which the *Pcgf4* gene is constitutively deleted and knockout of the *Pcgf2*

gene is conditionally induced via addition of tamoxifen (OHT) (*Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$) (Figure 3.1B).

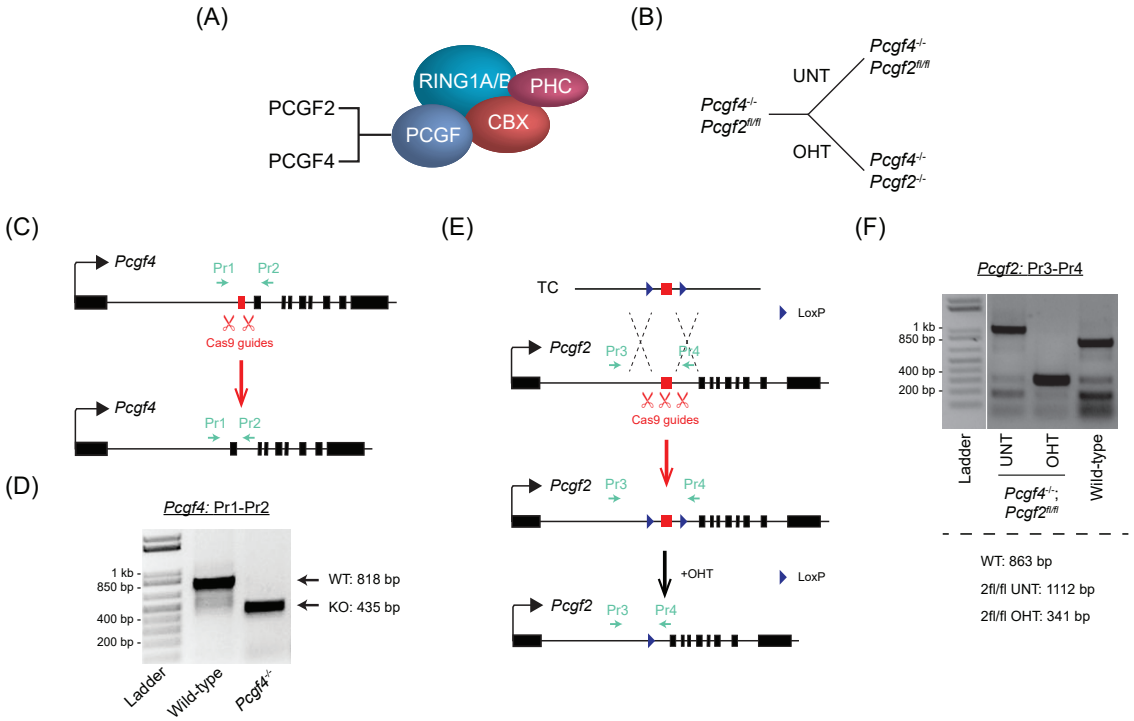



**Figure 3.1: A *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ mouse ES cell line to conditionally remove canonical PRC1**
(A) A schematic showing the major components of canonical PRC1 complexes.
(B) A schematic of the *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ conditional knockout ES cell line in which addition of tamoxifen (OHT) leads to removal of PCGF2 and loss of all canonical PRC1 complexes.
(C) A schematic of the CRISPR/Cas9-based strategy used to constitutively knockout both alleles of the endogenous *Pcgf4* gene. Positions of two sgRNA target sites (Cas9 guides) flanking the first coding exon of the *Pcgf4* gene (exon 2, *in red*) are indicated. Also shown are positions of screening primers (Pr1 and Pr2) used to identify clones in which the desired deletion was introduced in both alleles.
(D) Agarose gel showing the results of analytical PCR of genomic DNA using primers described in (C) for wild-type cells and the selected *Pcgf4*$^{-/-}$ clone. Expected sizes of the PCR products for the wild-type (WT) and *Pcgf4* knockout (KO) cells are shown on the right. DNA ladder is shown on the left.
(E) A schematic of the CRISPR/Cas9-based strategy used to generate homozygous conditional alleles of the endogenous *Pcgf2* gene. Cas9 guides were designed against target sites flanking the first coding exon of the *Pcgf2* gene (exon 2, *in red*) to introduce DNA double-strand breaks. An HDR template harbouring two parallel loxP sites flanking the target exon was supplied (TC - targeting construct). Also shown are positions of screening primers (Pr3 and Pr4) used to identify clones in which successful knock-in of loxP sites into both *Pcgf2* alleles allowed for inducible deletion of the target exon upon addition of OHT.
(F) Agarose gel showing the results of analytical PCR of genomic DNA using primers described in (E) for the wild-type cells and the selected *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ clone before and after addition of tamoxifen (OHT). Expected sizes of the PCR products for the wild-type (WT) and *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ cells (untreated and OHT-treated, 2fl/fl UNT and 2fl/fl OHT respectively) are shown at the bottom. DNA ladder is shown on the left.

To achieve this, cells which express a tamoxifen-inducible ERT2-Cre recombinase from the *Rosa26* locus were subjected to two rounds of CRISPR/Cas9-assisted genome editing. Initially, the first coding exon of both *Pcgf4* alleles was constitutively deleted (Figure 3.1C). Next, the first coding exon of both *Pcgf2* alleles was flanked by parallel loxP sites (Figure 3.1E). In this system, addition of tamoxifen allows ERT2-Cre to translocate into the nucleus where it mediates recombination between the loxP sites, resulting in deletion of the *Pcgf2* exon. Both in the case of

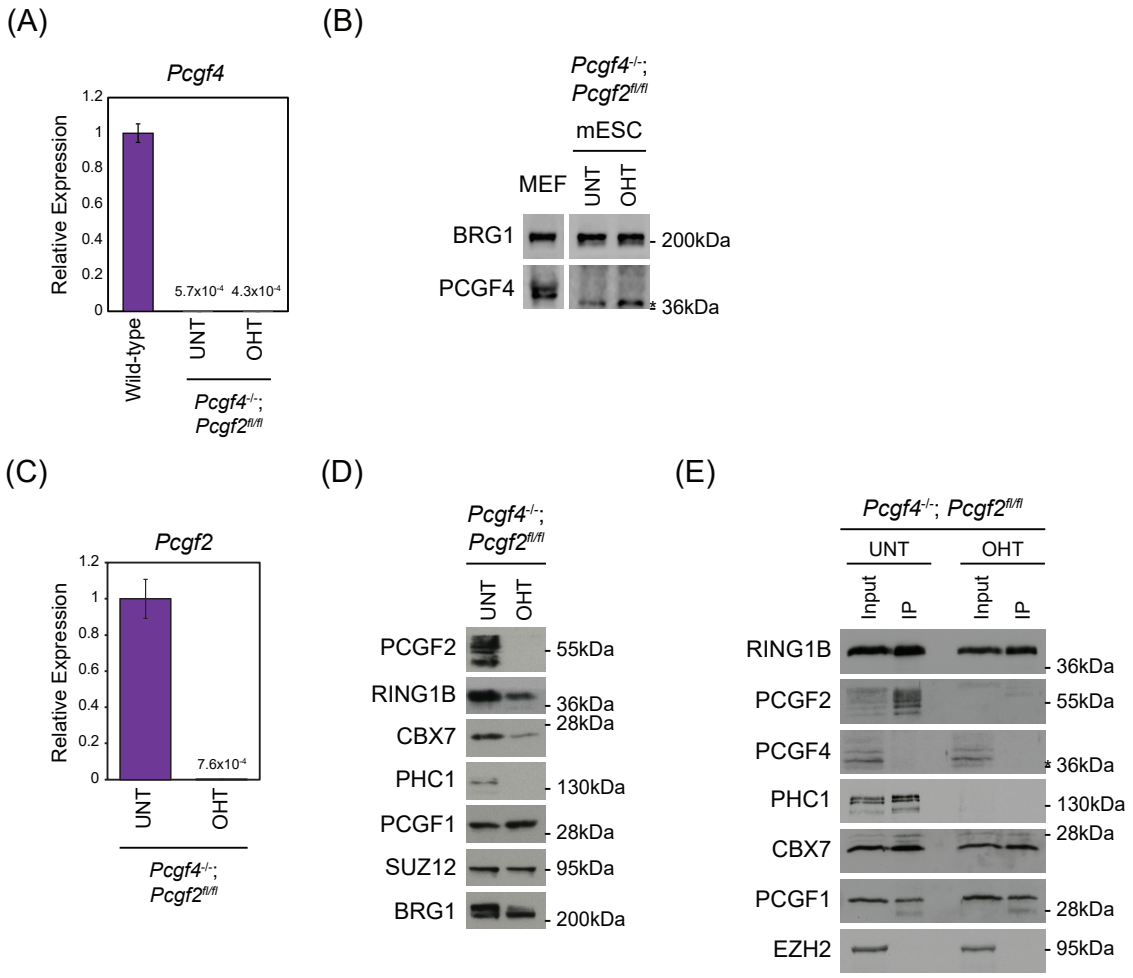



**Figure 3.2: Canonical PRC1 complexes are disrupted in the *Pcgf4$^{-/-}$;Pcgf2$^{fl/fl}$* ES cells following OHT treatment**
(A) RT-qPCR analysis comparing expression levels of the *Pcgf4* first coding exon in the wild-type parental cells and *Pcgf4$^{-/-}$;Pcgf2$^{fl/fl}$* cells (untreated and OHT-treated). Gene expression is normalised to *b-actin* and is shown relative to the average expression in wild-type cells. Error bars represent SEM (n=2).
(B) Western blot analysis of PCGF4 in MEF cells and *Pcgf4$^{-/-}$;Pcgf2$^{fl/fl}$* ES cells before and after addition of OHT. BRG1 is shown as a loading control. Asterisk indicates a non-specific band detected by the anti-PCGF4 antibody.
(C) RT-qPCR analysis comparing expression levels of the *Pcgf2* first coding exon in untreated and OHT-treated *Pcgf4$^{-/-}$;Pcgf2$^{fl/fl}$* cells. Gene expression is normalised to *b-actin* and is shown relative to the average expression in untreated cells. Error bars represent SEM (n=2).
(D) Western blot analysis of PRC1 subunits (cPRC1-specific - PCGF2, CBX7, PHC1, shared - RING1B, vPRC1-specific - PCGF1) and PRC2 subunits (SUZ12) in untreated and OHT-treated *Pcgf4$^{-/-}$;Pcgf2$^{fl/fl}$* cells. BRG1 is shown as a loading control.
(E) Input and immunoprecipitation of RING1B from the *Pcgf4$^{-/-}$;Pcgf2$^{fl/fl}$* cells (UNT and OHT-treated) followed by western blot analysis with antibodies against PCGF2, PCGF4, PHC1 and CBX7 (canonical PRC1-specific), PCGF1 (variant PRC1-specific) and EZH2 (PRC2 component that does not interact with RING1B).

*Pcgf2* and *Pcgf4*, deletion of the first coding exon leads to a frame-shift in the coding sequence, so a functional protein can no longer be produced. Successful deletion (for *Pcgf4*) or targeting (for *Pcgf2*) of both alleles was confirmed by analytical PCR of genomic DNA and sequencing of the PCR products (Figure 3.1D and F).

Having validated the genotype of the *Pcgf4$^{-/-}$;Pcgf2$^{fl/fl}$* cell line, I wanted to confirm that tamoxifen treatment results in a complete loss of canonical PRC1 complexes in these cells. Initially, I verified

that the first coding exon of *Pcgf4* was not expressed in this cell line by RT-qPCR (Figure 3.2A). PCGF4 protein expression in mouse ES cells is so low that I was unable to detect it by western blot in the wild-type or *Pcgf4* knockout cells (Figure 3.2B). This was not simply due to the poor performance of antibodies in western blot, as I could detect a robust signal for PCGF4 in mouse embryonic fibroblasts (MEFs), which express higher levels of PCGF4 than ES cells (Figure 3.2B).

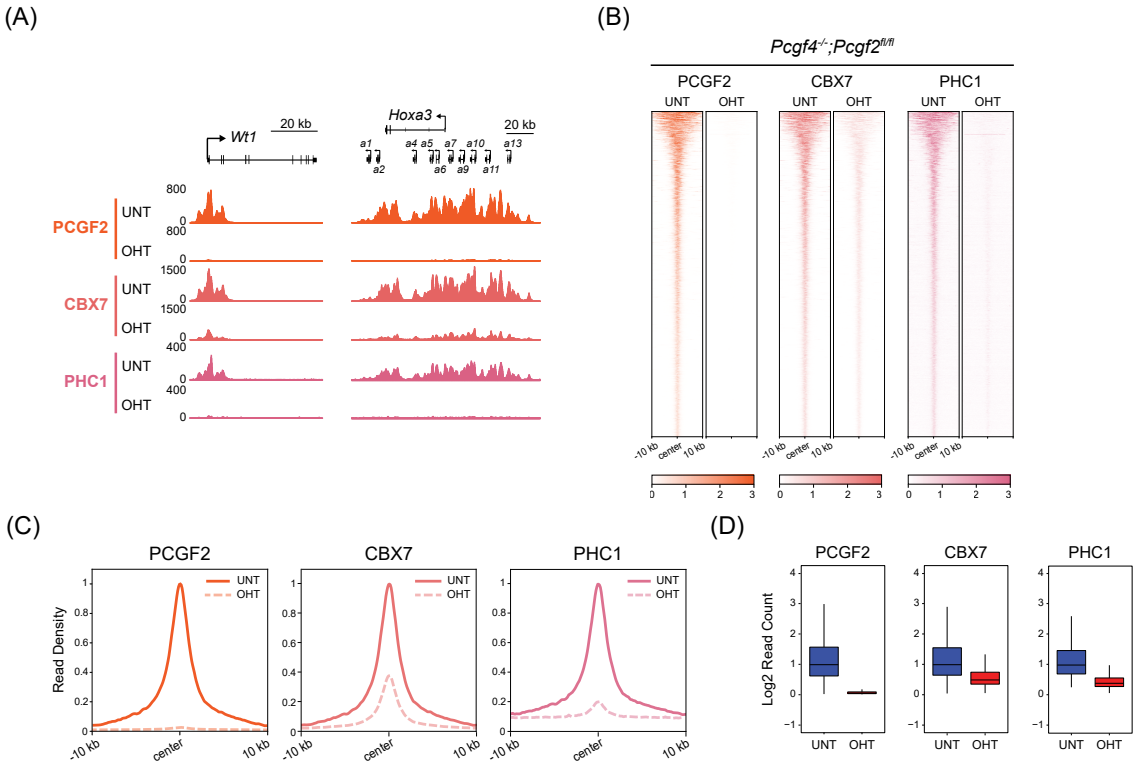



**Figure 3.3: Binding of canonical PRC1 at Polycomb target sites is lost following removal of PCGF2/4**
(A) Genomic snapshots of typical Polycomb target genes showing cChIP-seq for cPRC1-specific subunits, PCGF2, CBX7 and PHC1, in *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ cells (untreated and OHT-treated).
(B) Heatmap analysis of PCGF2, CBX7 and PHC1 cChIP-seq in *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ cells (untreated and OHT-treated) at classical Polycomb chromatin domains (n = 2096) sorted based on RING1B occupancy in untreated cells.
(C) Metaplot analysis of cChIP-seq for cPRC1-specific subunits (PCGF2, CBX7 and PHC1) at classical Polycomb chromatin domains (n = 2096) in *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ cells (untreated and OHT-treated).
(D) Box plots of the normalised cChIP-seq signal for PCGF2, CBX7 and PHC1 at classical Polycomb chromatin domains in *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ cells (untreated and OHT-treated).

Next, I analysed expression of PCGF2 in the *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ cells following 72 hours of OHT treatment. This confirmed that the first coding exon of *Pcgf2* was no longer expressed as measured by RT-qPCR and PCGF2 protein could not be detected by western blot (Figure 3.2C and D). In addition, the protein levels of other subunits of canonical PRC1, including RING1B, were also considerably reduced (Figure 3.2D). Furthermore, immunoprecipitation of RING1B from the OHT-treated cells revealed that, although RING1B could still interact with CBX7, it no longer associated with PCGF2 or PHC1 (Figure 3.2E), indicating that canonical PRC1 complexes have been disrupted. Finally, to confirm that removal of PCGF2/4 completely abolished binding

of canonical PRC1 at Polycomb target sites, I profiled genomic occupancy of cPRC1-specific subunits using calibrated chromatin immunoprecipitation sequencing (cChIP-seq) (Bonhoure et al. 2014; Hu et al. 2015; Orlando et al. 2014). This revealed that in the OHT-treated cells, PCGF2 and PHC1 no longer occupied Polycomb target sites and occupancy of CBX7 was considerably reduced (Figure 3.3A-D). Together, these results demonstrate that removal of PCGF2/4 subunits disrupts canonical PRC1 complexes and abolishes their binding at target genes, indicating that PCGF2/4 are central to the formation and stability of canonical PRC1 in mouse ES cells.

### 3.2.2 Canonical PRC1 complexes contribute minimally to Polycomb-mediated gene repression

Having developed a system in which canonical PRC1 complexes can be conditionally removed, I set out to test whether these complexes are essential for Polycomb-mediated gene repression in mouse ES cells. To this end, I carried out calibrated nuclear RNA-seq (cnRNA-seq) in the cPRC1 conditional knockout cell line before and after OHT treatment. To compare the primary transcriptional changes following loss of canonical PRC1 to the effect of removing PRC1 completely, I also performed cnRNA-seq in the previously characterised *Ring1a*$^{-/-}$;*Ring1b*$^{fl/fl}$ ES cell line (Endoh et al. 2008; Blackledge et al. 2014). In this system, conditional knockout of PRC1 resulted in a dramatic derepression of several thousand genes (Figure 3.4A and D), with Polycomb target genes being preferentially affected (Figure 3.4B and C). In striking contrast, disruption of canonical PRC1 led to minimal gene expression changes (Figure 3.4E and F), with only a small number of Polycomb target genes showing a significant increase in expression (Figure 3.4G and H). Together, these observations demonstrate that although PRC1 is essential for repression of Polycomb target genes in mouse ES cells, canonical PRC1 complexes contribute minimally to this process.

### 3.2.3 Canonical PRC1 shapes high-level enrichment of RING1B but not H2AK119ub1, PRC2 or H3K27me3 at Polycomb chromatin domains

The unexpected finding that removal of canonical PRC1 had a minimal effect on repression of Polycomb target genes (Figure 3.4) suggested that some of the remaining PRC1 and/or PRC2 activities must be sufficient for transcriptional silencing. Therefore, I used cChIP-seq to identify features of Polycomb chromatin domains which were retained in the absence of PCGF2/4.

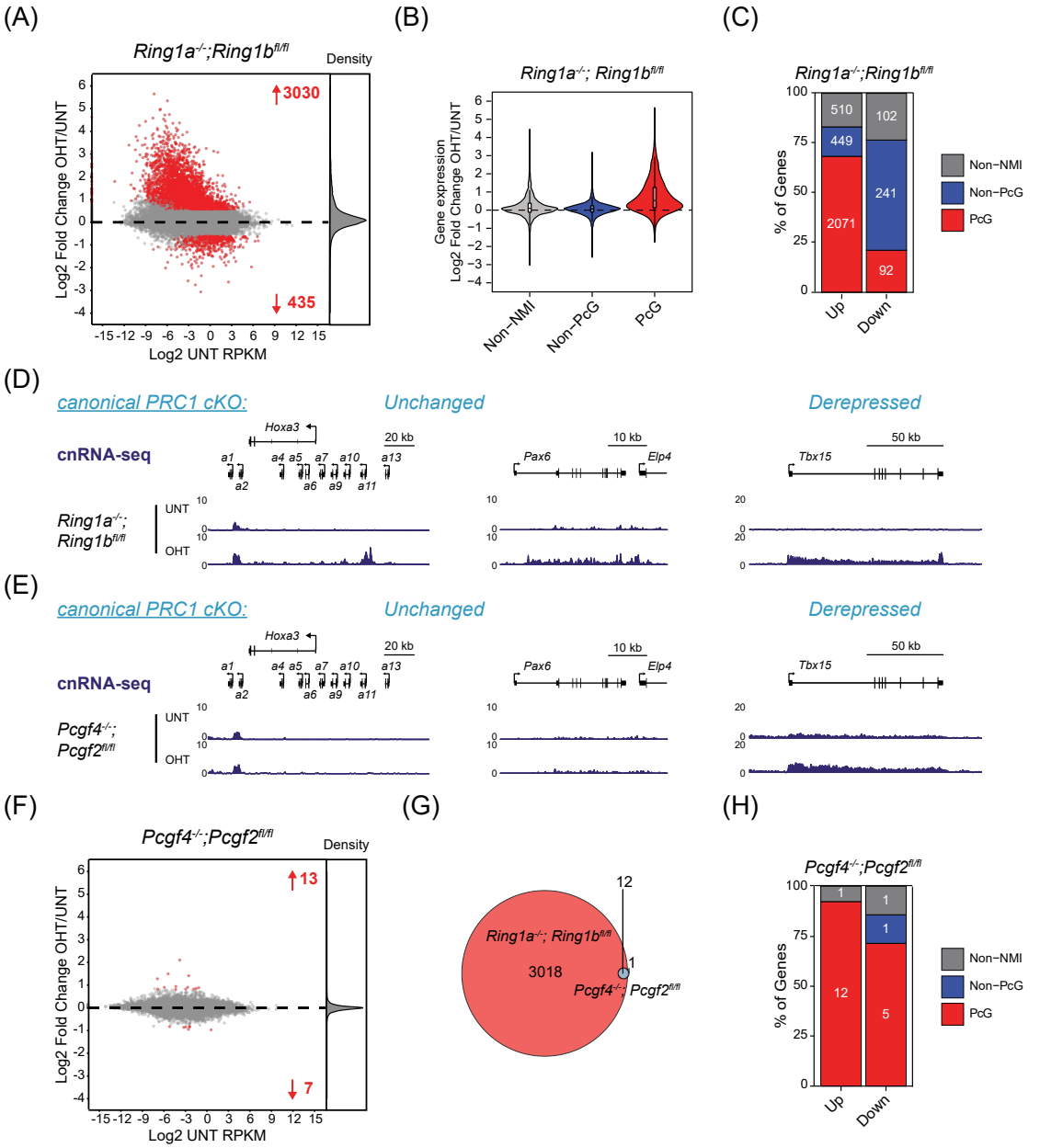



**Figure 3.4: Canonical PRC1 complexes are not required for Polycomb target gene repression in ES cells**

(A) An MA-plot of log2 fold changes in gene expression (cnRNA-seq) in *Ring1a*$^{-/-}$;*Ring1b*$^{fl/fl}$ cells following OHT treatment. Significant gene expression changes (p-adj<0.05 and FC>1.5-fold) are shown in red. Density of gene expression changes is shown on the right.

(B) A violin plot comparing log2 fold changes in expression in *Ring1a*$^{-/-}$;*Ring1b*$^{fl/fl}$ cells following OHT treatment for three classes of genes - Non-NMI, Non-PcG- and PcG- occupied. Non-NMI genes lack a non-methylated CpG island (NMI) at their promoter. PcG- and Non-PcG- occupied genes have an NMI promoter that is either bound or not bound by both PRC1 and PRC2.

(C) A bar plot illustrating the distribution of gene expression changes (p-adj<0.05 and FC>1.5-fold) between the gene classes described in (B) (Non-NMI, Non-PcG, PcG) in *Ring1a*$^{-/-}$;*Ring1b*$^{fl/fl}$ cells following OHT treatment.

(D) A genomic snapshot of gene expression (cnRNA-seq) in *Ring1a*$^{-/-}$;*Ring1b*$^{fl/fl}$ cells (untreated and OHT-treated) for typical Polycomb target genes which showed no change (Unchanged) or a significant increase in expression (Derepressed) following removal of canonical PRC1 complexes.

(E) As in (D) for *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ cells.

(F) As in (A) for *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ cells.

(G) A Venn diagram showing an overlap of significantly derepressed genes (p-adj<0.05 and FC>1.5-fold) in *Ring1a*$^{-/-}$;*Ring1b*$^{fl/fl}$ and *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ cells following OHT treatment.

(H) As in (C) for *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ cells.

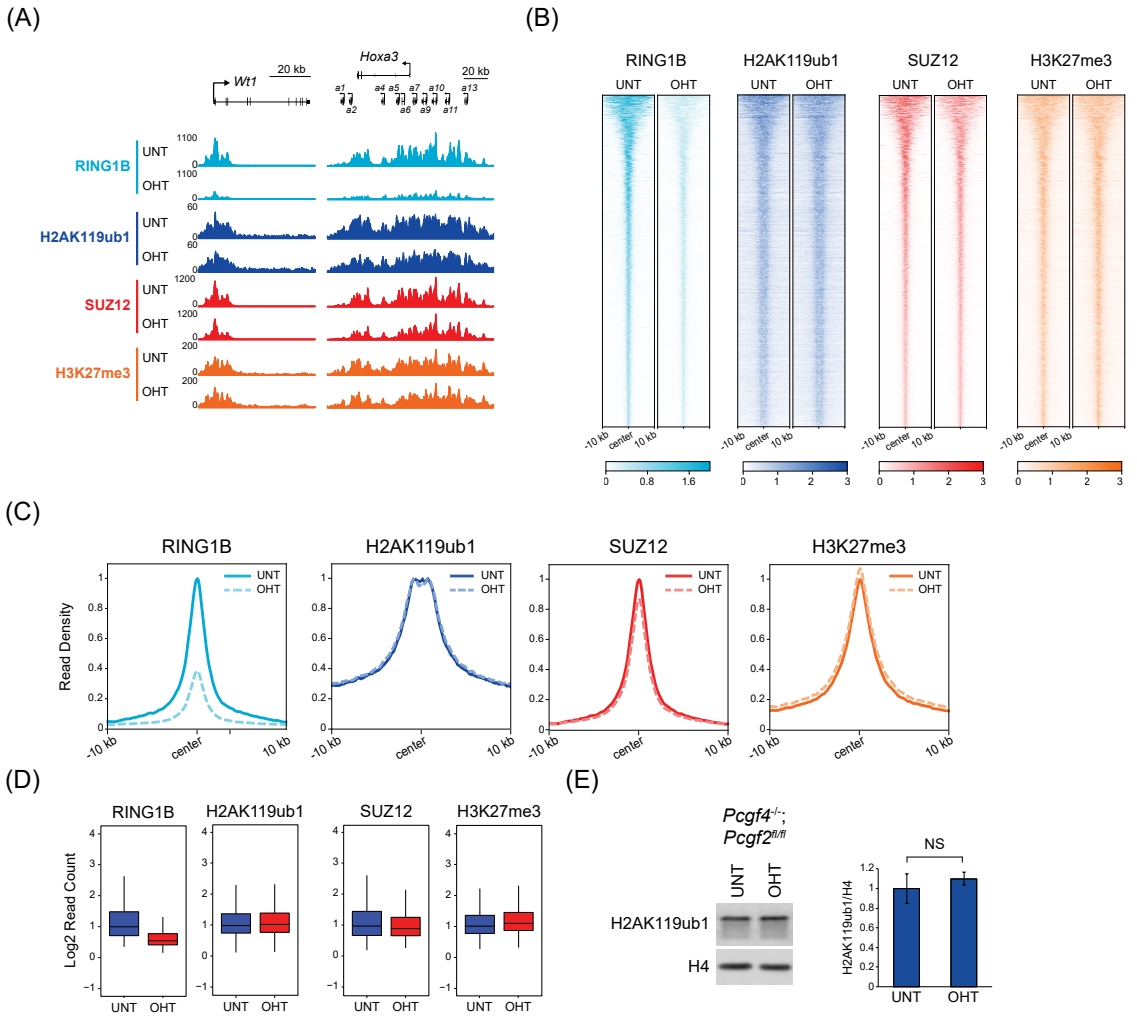



**Figure 3.5: Canonical PRC1 shapes high-level enrichment of RING1B but not H2AK119ub1, PRC2 or H3K27me3 at Polycomb chromatin domains**

(A) Genomic snapshots of typical Polycomb target genes showing cChIP-seq for RING1B, H2AK119ub1, SUZ12 and H3K27me3 in *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ cells (untreated and OHT-treated).

(B) Heatmap analysis of cChIP-seq for RING1B, H2AK119ub1, SUZ12 and H3K27me3 in *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ cells (untreated and OHT-treated) at classical Polycomb chromatin domains (n = 2096) sorted based on RING1B occupancy in untreated cells.

(C) Metaplot analysis of cChIP-seq for RING1B, H2AK119ub1, SUZ12 and H3K27me3 at classical Polycomb chromatin domains (n = 2096) in *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ cells (untreated and OHT-treated).

(D) Box plots of the normalised cChIP-seq signal for RING1B, H2AK119ub1, SUZ12 and H3K27me3 at classical Polycomb chromatin domains in *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ cells (untreated and OHT-treated).

(E) Western blot analysis of H2AK119ub1 (with histone H4 as a loading control) in *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ cells (untreated and OHT-treated). Also shown is quantification of H2AK119ub1 levels relative to histone H4. Error bars represent SEM (n=3). NS indicates not statistically significant changes (paired Student's t-test p-value>0.05).

First, I examined the effect that loss of canonical PRC1 had on the overall PRC1 occupancy at classical Polycomb chromatin domains, using cChIP-seq for RING1B as a proxy. This revealed a major reduction in RING1B binding at these sites following OHT treatment in *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ cells (Figure 3.5A-D), consistent with the previous PCGF2 knockdown experiments (Morey et al. 2015). This indicates that canonical PRC1 complexes are responsible for shaping high-level RING1B occupancy at classical Polycomb chromatin domains, yet this is not required for

repression of Polycomb target genes.

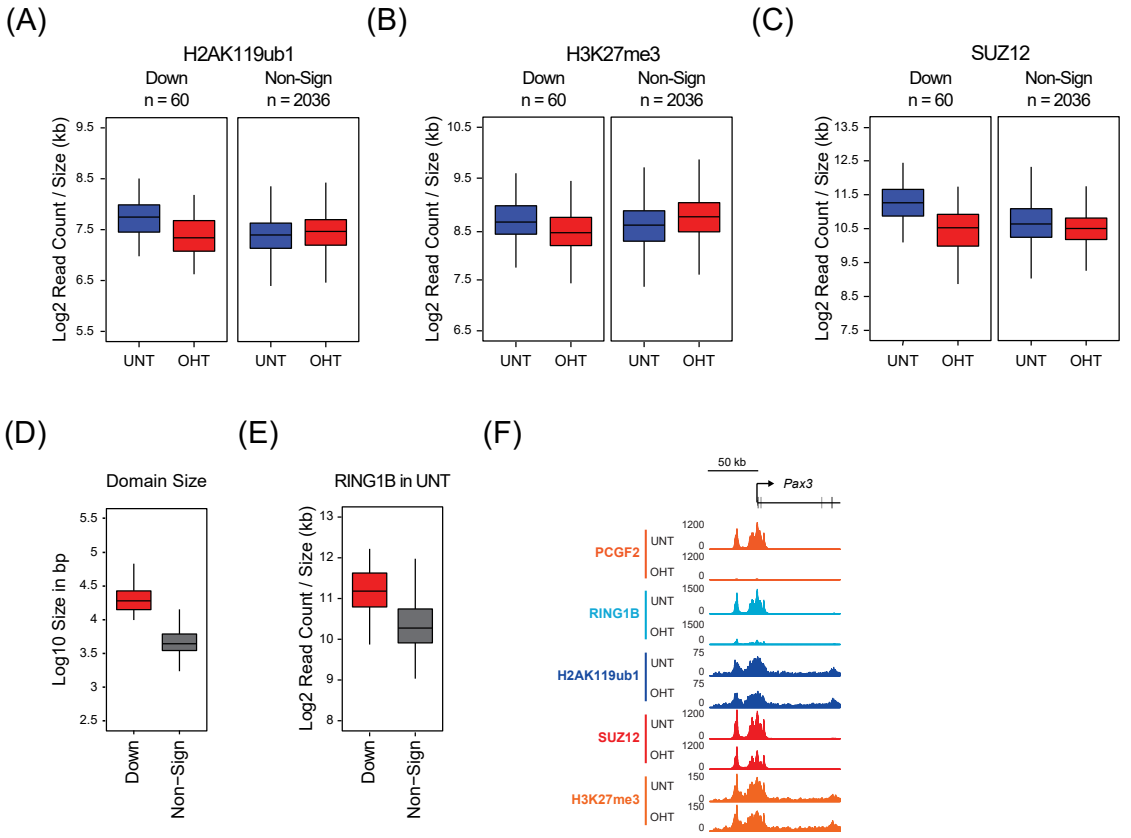



**Figure 3.6: H2AK119ub1, SUZ12 and H3K27me3 levels are reduced at a small number of large Polycomb chromatin domains following loss of cPRC1**
(A) Box plots of the cChIP-seq signal for H2AK119ub1 in untreated and OHT-treated *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ cells at classical Polycomb chromatin domains divided into sites that show a significant decrease (p-adj<0.05 and FC<0) or no change in H2AK119ub1 levels following loss of cPRC1.
(B) As in (A) for H3K27me3 cChIP-seq.
(C) As in (A) for SUZ12 cChIP-seq.
(D) Box plots of the sizes of classical Polycomb chromatin domains divided into sites that show a significant decrease (p-adj<0.05 and FC<0) or no change in H2AK119ub1 levels following loss of cPRC1.
(E) Box plots of the cChIP-seq signal for RING1B in untreated *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ cells at classical Polycomb chromatin domains divided into sites that show a significant decrease (p-adj<0.05 and FC<0) or no change in H2AK119ub1 levels following loss of cPRC1.
(F) Genomic snapshot showing cChIP-seq for PCGF2, RING1B, H2AK119ub1, SUZ12 and H3K27me3 in *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ cells (untreated and OHT-treated) for a classical Polycomb chromatin domain at the promoter of *Pax3* gene which shows a significant reduction in H2AK119ub1 and PRC2 occupancy/activity following loss of cPRC1.

Having found that PCGF2/4 removal caused a major reduction in RING1B binding at Polycomb chromatin domains, I asked whether this had any effect on deposition of H2AK119ub1. Interestingly, global H2AK119ub1 levels as quantified by bulk histone western blot remained unchanged following OHT treatment in *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$ cells (Figure 3.5E). Furthermore, cChIP-seq for H2AK119ub1 showed that the levels of this histone modification at Polycomb chromatin domains were largely unaffected following loss of PCGF2/4 (Figure 3.5A-D), indicating that canonical PRC1 contributes little to deposition of H2AK119ub1 in mouse ES cells. Finally,

recognition of H2AK119ub1 by PRC2 has been proposed to shape PRC2 occupancy and H3K27me3 at target sites (Blackledge et al. 2014; Cooper et al. 2014; Kalb et al. 2014; Cooper et al. 2016). In line with this, removal of canonical PRC1 also had a minimal effect on the levels of H3K27me3 and the PRC2 core subunit, SUZ12, at Polycomb chromatin domains (Figure 3.5A-D).

Interestingly, at a small number of sites, removal of canonical PRC1 did lead to a minor reduction in H2AK119ub1 levels, which was accompanied by reductions in SUZ12 binding and H3K27me3 (Figure 3.6A-C and F). The affected sites tended to be large Polycomb chromatin domains highly enriched with Polycomb protein complexes (Figure 3.6D-F). It is therefore possible that high-level occupancy of canonical PRC1 can contribute to deposition of H2AK119ub1 and formation of Polycomb chromatin domains, despite cPRC1 inherently weak catalytic activity (Gao et al. 2012; Taherbhoy et al. 2015; Rose et al. 2016). Nevertheless, this effect was limited to a small number of sites (n = 60), with the vast majority of classical Polycomb chromatin domains being completely unaffected following loss of canonical PRC1 (Figure 3.6A-C). Together, these findings demonstrate that, although canonical PRC1 complexes play a central role in defining RING1B occupancy at Polycomb chromatin domains, they contribute little to deposition of H2AK119ub1 and PRC2 recruitment/activity, as well as repression of Polycomb target genes.

## 3.3 Discussion

The Polycomb repressive system represents a central paradigm for chromatin-based gene regulation. Nevertheless, the fundamental mechanisms that govern Polycomb-mediated gene silencing remain poorly understood. A widely adopted model in the field postulates that the Polycomb system functions through a PRC2-initiated pathway which culminates in recruitment of canonical PRC1 complexes to target genes, where they elicit transcriptional silencing. In Chapter 3, I have tested this model by probing the role of canonical PRC1 in Polycomb-mediated gene repression in mouse ES cells. I discovered that conditional ablation of canonical PRC1 has little effect on Polycomb target gene repression, in striking contrast to the widespread transcriptional defects that manifest when all PRC1 is removed. In addition to this, I also found that, while canonical PRC1 complexes define stable RING1B binding at classical Polycomb target sites, they contribute little to deposition of H2AK119ub1 and Polycomb chromatin domain formation. Together, these findings strongly indicate that the prevailing dogma in the field that places canonical PRC1 at the centre of Polycomb-mediated gene repression must be revisited.

In line with PRC1 being essential for self-renewal of mouse ES cells (Endoh et al. 2008), I showed that conditional removal of the PRC1 core catalytic subunits, RING1A/B, results in severe transcriptional defects, with thousands of Polycomb target genes becoming derepressed. In contrast to this, I found that loss of canonical PRC1 complexes leads to only minimal gene expression changes. This contrasts previous studies which reported that constitutive RNAi knockdown or knockout of individual canonical PRC1 subunits leads to derepression of several hundred genes (Morey et al. 2012; Morey et al. 2015; Kundu et al. 2017). However, less than a half of the transcriptional changes observed in these experiments corresponded to genes directly bound by Polycomb complexes. Unexpectedly high proportion of non-Polycomb target genes being misregulated likely represents secondary effects of the prolonged culture of cPRC1-deficient cells or off-target effects of RNAi. Importantly, although I found only a small number of genes which were derepressed following conditional removal of canonical PRC1, the majority of these were Polycomb target genes. This highlights the importance of using conditional knockout strategies to ensure that primary transcriptional defects caused by removal of a protein of interest are examined.

Further support for the idea that canonical PRC1 does not play a central role in Polycomb-mediated gene repression, at least at early stages in development, comes from the phenotypes of knockout mouse models. RING1B knockout mice fail to complete gastrulation (Voncken et al. 2003) and mice deficient for RING1A and RING1B cease development at the two-cell stage (Posfai et al. 2012). In contrast, mice which are deficient for individual subunits of canonical PRC1 complexes are viable, but exhibit homeotic transformations of different severity (Lugt et al. 1994; Akasaka et al. 1996; Takihara et al. 1997; Core et al. 1997; Forzati et al. 2012; Liu et al. 2013). Strikingly, double knockout of PCGF2/4 or PHC1/2 in mice does lead to embryonic lethality, albeit at a later developmental stage (E9.5) than combined deletion of RING1A/B (Akasaka et al. 2001; Isono et al. 2005). Together, these observations suggest that canonical PRC1 is not essential for the establishment of gene expression patterns early in embryogenesis, but may contribute to their maintenance later in development. In agreement with this possibility, the composition of PRC1 complexes changes during development. Canonical PRC1-specific subunits become more abundant upon differentiation of ES cells, potentially reflecting a shift towards cPRC1-dependent mechanisms of gene repression in more committed cell types (Morey et al. 2015; Kloet et al. 2016). Furthermore, canonical PRC1 plays essential roles in regulation of proliferation and differentiation of adult stem cells and more committed cells (reviewed in Sauvageau and Sauvageau 2010;

Luis et al. 2012). In contrast to pluripotent ES cells, the genomic organisation of these more slowly dividing cells has been shown to be less dynamic (Delgado-Olguin and Recillas-Targa 2011; Murtha et al. 2015). Therefore, in these cell types, the compact higher-order chromatin structures mediated by canonical PRC1 may be more stable and thus may further reinforce the fidelity of Polycomb-mediated gene repression. Moving forward, it will be important to utilise the conditional knockout approaches described in Chapter 3 to examine the contribution of canonical PRC1 complexes to gene regulation in other developmental and cellular contexts.

Polycomb-mediated gene repression has been previously associated with formation of Polycomb chromatin domains, which accumulate high levels of PRC1, PRC2 and their respective histone modifications (Mikkelsen et al. 2007). In Chapter 3, I demonstrated that PCGF2/4 shape high-level RING1B occupancy at these classical Polycomb target sites, consistent with the previous studies in cPRC1-deficient cells which examined a small selection of target gene promoters by ChIP-qPCR (Morey et al. 2015; Kundu et al. 2017). However, I found that these major reductions in RING1B binding at Polycomb target gene promoters were not associated with gene derepression, indicating that cPRC1-dependent high-level RING1B occupancy is downstream of gene repression.

Surprisingly, despite the major reductions in RING1B binding, removal of PCGF2/4 had little effect on deposition of H2AK119ub1 at classical Polycomb chromatin domains. This is in line with previous work in the PRC2-null mouse ES cells which demonstrated that, although RING1B occupancy at target sites was dramatically reduced, presumably due to the loss of canonical PRC1 binding, this had little effect on the levels of H2AK119ub1 (Tavares et al. 2012). Furthermore, I found that H3K27me3 and stable PRC2 binding at Polycomb chromatin domains also remained largely unchanged following removal of PCGF2/4. Together with the minimal transcriptional defects in PCGF2/4-deficient cells, these observations suggested that the remaining activities associated with Polycomb chromatin domains must be responsible for Polycomb-mediated transcriptional silencing. Importantly, constitutive deletion of PRC2 and a concomitant loss of H3K27me3 did not result in widespread reactivation of Polycomb target genes, indicating that neither PRC2 nor H3K27me3 can be the driving force behind Polycomb-mediated gene repression (Riising et al. 2014). Therefore, some activity within the remaining PRC1 complexes must be sufficient to drive repression of Polycomb target genes. Since H2AK119ub1 was unaffected following removal of canonical PRC1, in the next chapters, I will investigate whether H2AK119ub1 may be directly linked to gene repression.

# 4

# DISTINCT ROLES OF VARIANT PRC1 COMPLEXES IN POLYCOMB SYSTEM FUNCTION

## 4.1 Introduction

In mammals, further diversification of the PcG protein family has led to expansion of the repertoire of Polycomb repressive complexes. In addition to canonical PRC1, a multitude of biochemically distinct PRC1 complexes, known as non-canonical or variant PRC1 (vPRC1), have been described (Gao et al. 2012; Hauri et al. 2016). Out of six mammalian PCGF proteins, PCGF1/3/5/6 dimerise with RING1A/B to exclusively form variant PRC1 complexes which have unique auxiliary subunits. In contrast to canonical PRC1, variant PRC1 complexes lack CBX subunits and are recruited to target gene promoters independently of PRC2 and H3K27me3 (Gao et al. 2012; Tavares et al. 2012; Morey et al. 2013; Hauri et al. 2016). Targeting of variant PRC1 has been proposed to rely on recognition of specific DNA sequences, as several variant PRC1 subunits contain DNA-binding domains (Farcas et al. 2012; He et al. 2013a; Wu et al. 2013; Endoh et al. 2017; Stielow et al. 2018; Scelfo et al. 2019). Furthermore, in place of CBX subunits, variant PRC1 complexes incorporate RYBP (or its paralogue YAF2) (Gao et al. 2012; Morey et al. 2013; Hauri et al. 2016), which stimulates catalytic activity of these complexes both *in vitro* and *in vivo* (Gao et al. 2012; Taherbhoy et al. 2015; Rose et al. 2016). The unique targeting and regulatory modules have been proposed to confer non-redundant functions to individual variant PRC1 protein assemblies (Sowpati et al. 2015). Thus, it has been suggested that distinct variant PRC1 complexes are recruited to non-overlapping sets of target genes, at which they contribute to different transcriptional outcomes, including, in some cases, gene activation (Gao et al. 2012; Gao et al. 2014; Zhao et al. 2017a; Yao et al. 2018). However, despite a detailed account of their biochemical properties and interaction partners, the exact contribution of individual variant PRC1 complexes to deposition of H2AK119ub1 and repression of Polycomb target genes remains poorly understood.

Histone H2A mono-ubiquitination is an abundant histone modification that decorates 5-10% of

histone H2A in mammalian cells (Goldknopf et al. 1975; Albright et al. 1979; Matsui et al. 1979). However, very little is known about the distribution of this histone modification in the genome outside of Polycomb target gene promoters, which in total account for less than 1% of nucleosomes in the genome. For many years, profiling the genomic distribution of H2AK119ub1 has remained one of the greatest technical challenges in the field. This has limited genome-scale analyses of the Polycomb system, with PRC2 and H3K27me3 genomic localisation typically used to identify Polycomb target sites. Furthermore, an inability to study how the product of PRC1 catalysis is distributed in the genome has hindered our understanding of how PRC1 catalytic activity is regulated *in vivo* and how it contributes to silencing of Polycomb target genes.

In Chapter 3, I established that canonical PRC1 is not required for Polycomb-mediated gene repression in mouse ES cells, indicating that activities of the remaining Polycomb complexes must be sufficient to inhibit transcription. Previously, it has been shown that ablation of PRC2 in mouse ES cells has little impact on expression of Polycomb target genes (Riising et al. 2014), suggesting that the remaining variant PRC1 complexes must be responsible for gene repression. Since H2AK119ub1 at Polycomb target sites was unaffected following removal of canonical PRC1, I reasoned that this could be related to transcriptional silencing. Therefore, pinpointing which variant PRC1 complexes deposit H2AK119ub1 in the genome could provide clues as to which PRC1 components are required for Polycomb-mediated gene repression. To achieve this, here I use a quantitative genomic assay to generate high-quality genome-wide profiles of H2AK119ub1, which I then extensively validate. Next, I go on to systematically dissect the contribution of individual variant PRC1 complexes to deposition of H2AK119ub1 throughout the genome, using a panel of conditional knockout mouse ES cell lines, in which each variant PRC1 complex can be removed on its own. After I identify which variant PRC1 complexes place the majority of H2AK119ub1 at Polycomb target gene promoters, I then examine whether these complexes are also required for Polycomb-mediated gene repression. This approach enables me to overcome the complexity of the mammalian PRC1 system as it provides a simple way to screen for potential repressive activity of individual variant PRC1 complexes.

## 4.2 Results

### 4.2.1 H2AK119ub1 is a pervasive histone modification that is enriched at PRC1-bound sites

In recent years, it has become clear that calibration of ChIP-seq with an exogenous spike-in genome is required to accurately characterise the genomic distribution of histone modifications, particularly when these modifications are abundant and placed broadly throughout the genome (Orlando et al. 2014; Bonhoure et al. 2014; Hu et al. 2015). Given that H2AK119ub1 is abundant and has been challenging to profile using conventional ChIP-seq in the past, I utilised the calibrated ChIP-seq (cChIP-seq) approach to provide a detailed description of H2AK119ub1 localisation in the genome. In parallel, I used the same cChIP-seq strategy for RING1B as a proxy for total PRC1 occupancy to compare the distribution of H2AK119ub1 to PRC1 binding in the genome. To validate the obtained genomic profiles, I performed cChIP-seq for H2AK119ub1 and RING1B in the *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ mouse ES cells, in which addition of tamoxifen (OHT) leads to a complete loss of all PRC1 complexes and H2AK119ub1 (Figure 4.1A and B) (Endoh et al. 2008; Schoenfelder et al. 2015). Following OHT treatment in these cells, RING1B and H2AK119ub1 cChIP-seq signal should be lost and any remaining signal will reflect the background level of non-specific binding of antibodies.

Using this quantitative cChIP-seq approach, I found that RING1B binding in the genome is more widespread than has been previously appreciated and is not limited to classical Polycomb chromatin domains (Figure 4.1D-E). In total, I have identified 8833 genomic sites which were significantly enriched with RING1B, and this signal was completely lost when RING1B was removed following OHT treatment in *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ cells (Figure 4.1C-D). The majority of RING1B-bound sites (herein also referred to as PRC1-bound sites) overlapped with non-methylated CpG islands (NMIs) that were associated with gene promoters (6798 out of 8833 sites, 77%), consistent with the proposed role of NMIs as Polycomb response elements (PREs) in mammals (Blackledge et al. 2015; Bauer et al. 2015). Importantly, H2AK119ub1 was also enriched at PRC1-bound sites and scaled well with RING1B occupancy (Figure 4.1C-E). A proportion of PRC1 target sites which showed the highest enrichment of RING1B and H2AK119ub1 represent classical Polycomb chromatin domains as they were also enriched with PRC2 and H3K27me3 (Figure 4.1E, Q4 group). These sites were associated with promoters

of lowly transcribed genes which became derepressed following PRC1 removal (Figure 4.1F). In contrast, the majority of PRC1-bound sites were characterised by low to moderate levels of RING1B and H2AK119ub1 and corresponded to promoters of more actively transcribed genes that were less susceptible to reactivation following removal of PRC1 (Figure 4.1D-F).

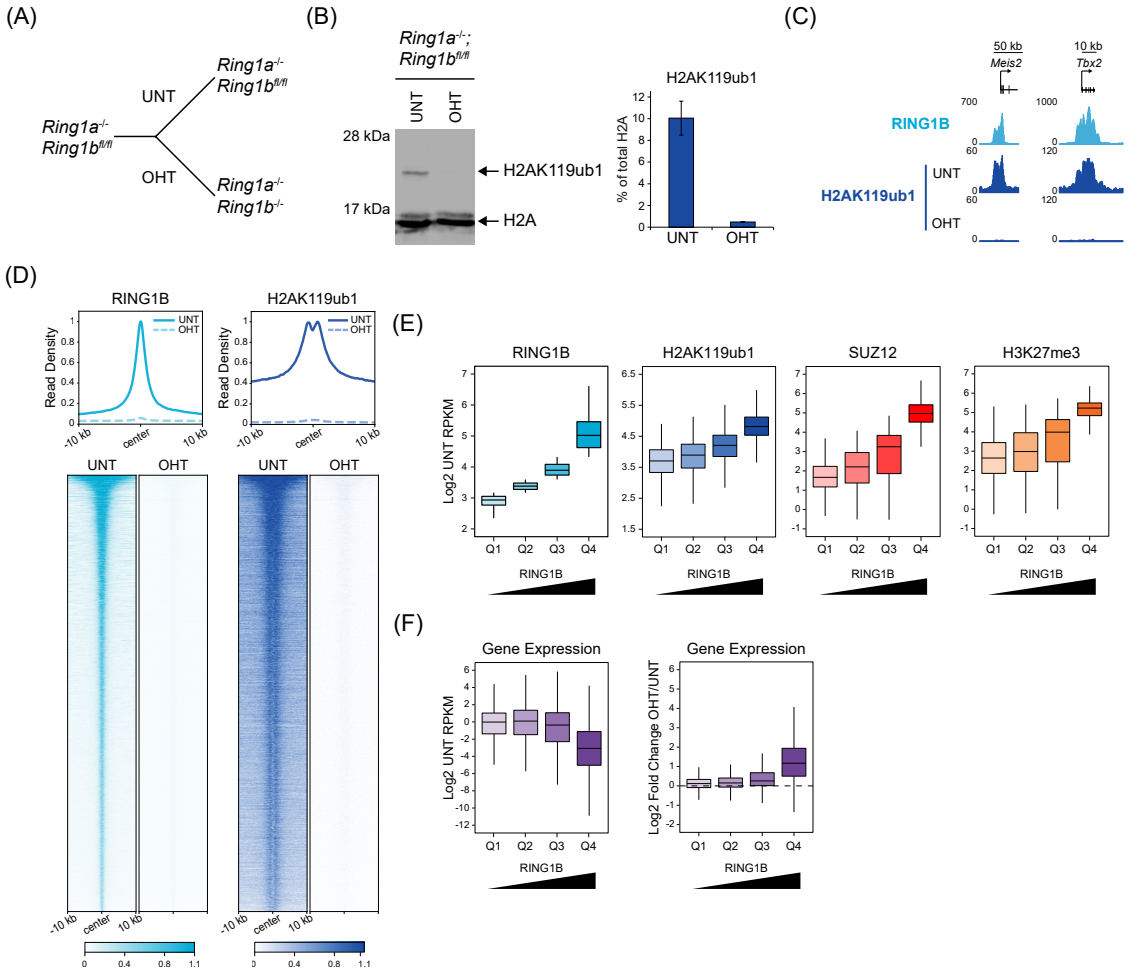



**Figure 4.1: H2AK119ub1 is enriched at PRC1-bound sites**

(A) A schematic of the *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ ES cell line in which addition of tamoxifen (OHT) leads to removal of RING1B and loss of all PRC1.

(B) *Left panel:* Western blot of H2AK119ub1 in *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ cells (untreated and OHT-treated), using an antibody against total histone H2A. *Right panel:* Quantification of H2AK119ub1 levels relative to total histone H2A from the western blot analysis on the left. Error bars represent SEM (n = 3).

(C) Genomic snapshots of typical PRC1-bound sites illustrating RING1B and H2AK119ub1 cChIP-seq in *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ cells (untreated and OHT-treated).

(D) Metaplot and heatmap analysis of RING1B and H2AK119ub1 cChIP-seq at PRC1-bound sites (n = 8833) in *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ cells (untreated and OHT-treated). Genomic regions were sorted based on RING1B occupancy in untreated cells.

(E) Box plots of RING1B, H2AK119ub1, SUZ12 and H3K27me3 cChIP-seq signal in untreated *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ cells at PRC1-bound gene promoters divided into quartiles based on RING1B occupancy (from lowest Q1 to highest Q4).

(F) Box plots of gene expression in untreated cells and log2 fold gene expression changes following OHT treatment in *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ cells for the quartiles of PRC1-bound gene promoters as in (E).

In addition to the punctate enrichment of H2AK119ub1 at PRC1-bound sites, I found that H2AK119ub1 was also placed more globally throughout the genome, albeit at a much lower level

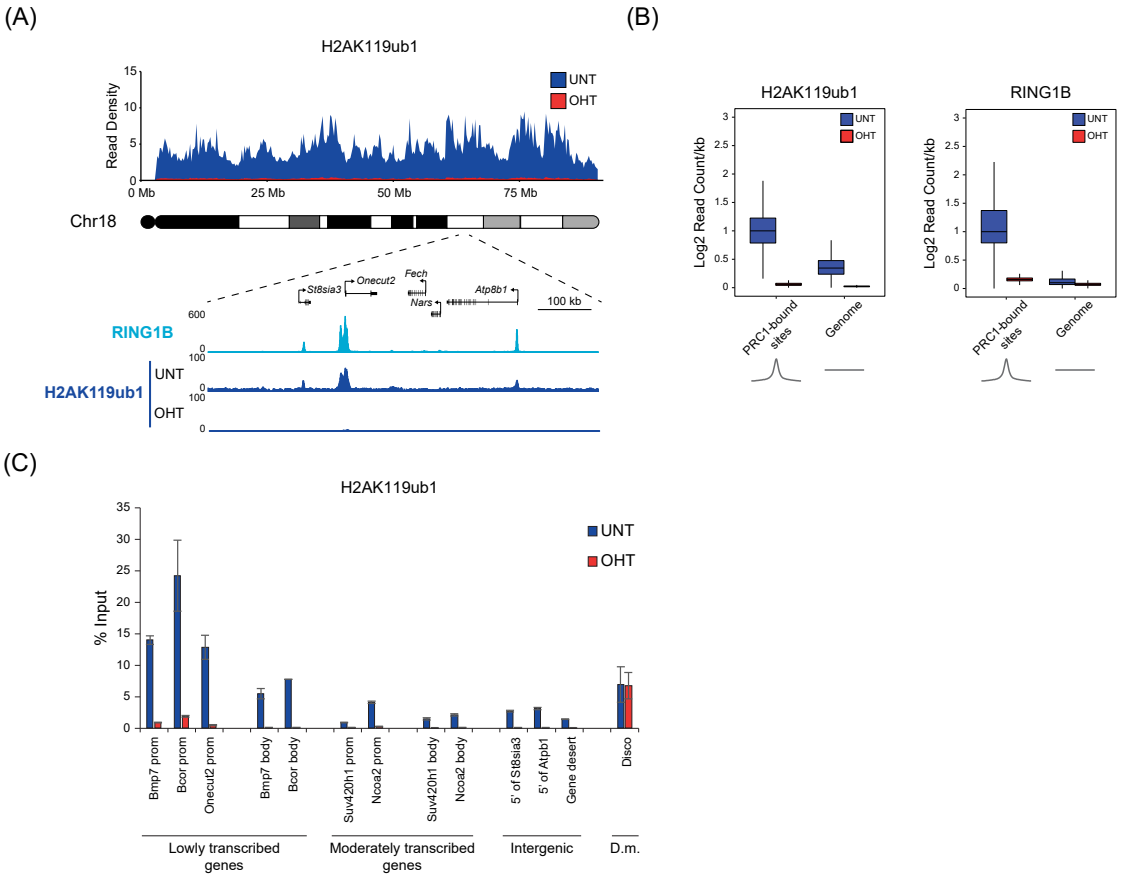



**Figure 4.2: H2AK119ub1 is a widespread histone modification that blankets the genome**
(A) A chromosome density plot showing H2AK119ub1 cChIP-seq across chromosome 18 in *Ring1a*$^{-/-}$;*Ring1b*$^{fl/fl}$ cells (untreated and OHT-treated), together with an expanded genomic snapshot of the chromosome 18 region which also shows RING1B cChIP-seq.
(B) Box plots of the normalised H2AK119ub1 and RING1B cChIP-seq signal at PRC1-bound sites and 100 kb genomic windows in *Ring1a*$^{-/-}$;*Ring1b*$^{fl/fl}$ cells (untreated and OHT-treated). This shows that there is an appreciable amount of H2AK119ub1 throughout the genome, outside of PRC1-bound sites, which is lost after removal of RING1A/B.
(C) ChIP-qPCR analysis of H2AK119ub1 enrichment in untreated and OHT-treated *Ring1a*$^{-/-}$;*Ring1b*$^{fl/fl}$ cells at genes with different expression levels and intergenic regions. Spike-in with *Drosophila* SG4 cells was included and enrichment of H2AK118ub1 at a Polycomb target gene in *Drosophila* (D.m.) was analysed as a control for ChIP efficiency. Error bars represent SEM for at least two biological replicates.

(Figure 4.2A and B). The widespread distribution of H2AK119ub1 in the genome was evident when I visualised the entire chromosome or focused on genomic regions between sites highly enriched with RING1B and H2AK119ub1 (Figure 4.2A). Importantly, this low-level ubiquitous H2AK119ub1 was completely lost following OHT treatment, confirming that this signal depends on PRC1 (Figure 4.2A and B). The observed pervasive distribution of H2AK119ub1 in the genome is in line with previous biochemical analysis which indicated that approximately 10% of histone H2A is mono-ubiquitinated (Goldknopf et al. 1975; Albright et al. 1979; Matsui et al. 1979) and is consistent with my quantifications based on western blot analysis of bulk histones (Figure 4.1B). Interestingly, no cChIP-seq signal above background was detected for RING1B outside of the highly enriched sites, suggesting that although RING1B must broadly engage with the

genome to place H2AK119ub1, this genome-wide binding appears to be transient (Figure 4.2B). Finally, H2AK119ub1 cChIP-seq was further validated at several genomic loci using the ChIP-qPCR approach which did not rely on spike-in calibration (Figure 4.2C). This confirmed that H2AK119ub1 was highly enriched at the promoters of lowly transcribed genes, but was also evident at the actively transcribed genes and intergenic genomic regions, which were not bound by RING1B (Figure 4.2C). In conclusion, this rigorous analysis of H2AK119ub1 distribution in the genome revealed two distinct pools of H2AK119ub1: a punctate enrichment at sites stably bound by RING1B and a low-level global H2AK119ub1 blanket that covers the genome.

### 4.2.2 PCGF1-PRC1 plays a major role in deposition of H2AK119ub1 at PRC1-bound sites

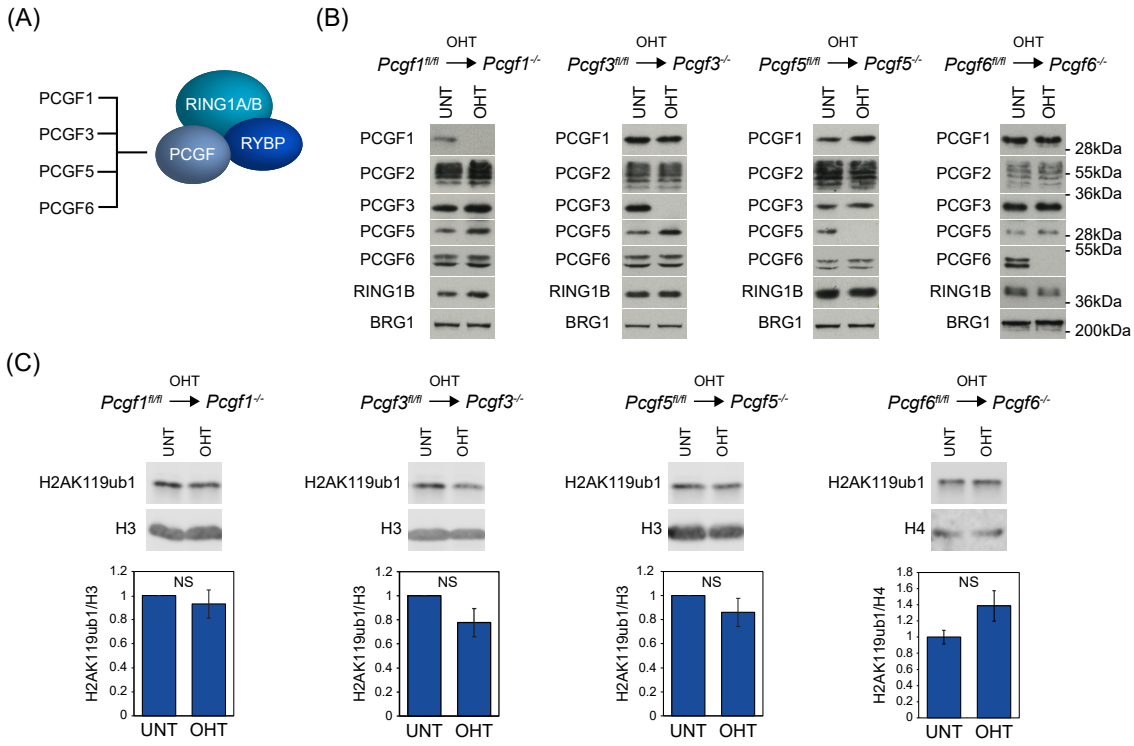



**Figure 4.3: A panel of conditional knockout ES cell lines to study the contribution of individual variant PRC1 complexes to deposition of H2AK119ub1**
(A) A schematic of variant PRC1 complexes that form around PCGF1, PCGF3, PCGF5 and PCGF6.
(B) Western blot analysis of PCGF proteins and RING1B in *Pcgf1*$^{fl/fl}$, *Pcgf3*$^{fl/fl}$, *Pcgf5*$^{fl/fl}$, and *Pcgf6*$^{fl/fl}$ cells before (UNT) and after tamoxifen treatment (OHT). BRG1 is shown as a loading control.
(C) Western blot analysis of H2AK119ub1 (with histone H3 or H4 as a loading control) in *Pcgf1*$^{fl/fl}$, *Pcgf3*$^{fl/fl}$, *Pcgf5*$^{fl/fl}$, and *Pcgf6*$^{fl/fl}$ cells (untreated and OHT-treated). Quantification of H2AK119ub1 levels relative to histone H3 (or H4) is shown below. Error bars represent SEM (n=2). NS indicates not statistically significant changes (paired Student's t-test p-value>0.05).

Having uncovered two distinct pools of H2AK119ub1 in the genome, I set out to determine which PRC1 complexes deposit these and whether this contributes to gene repression. Given

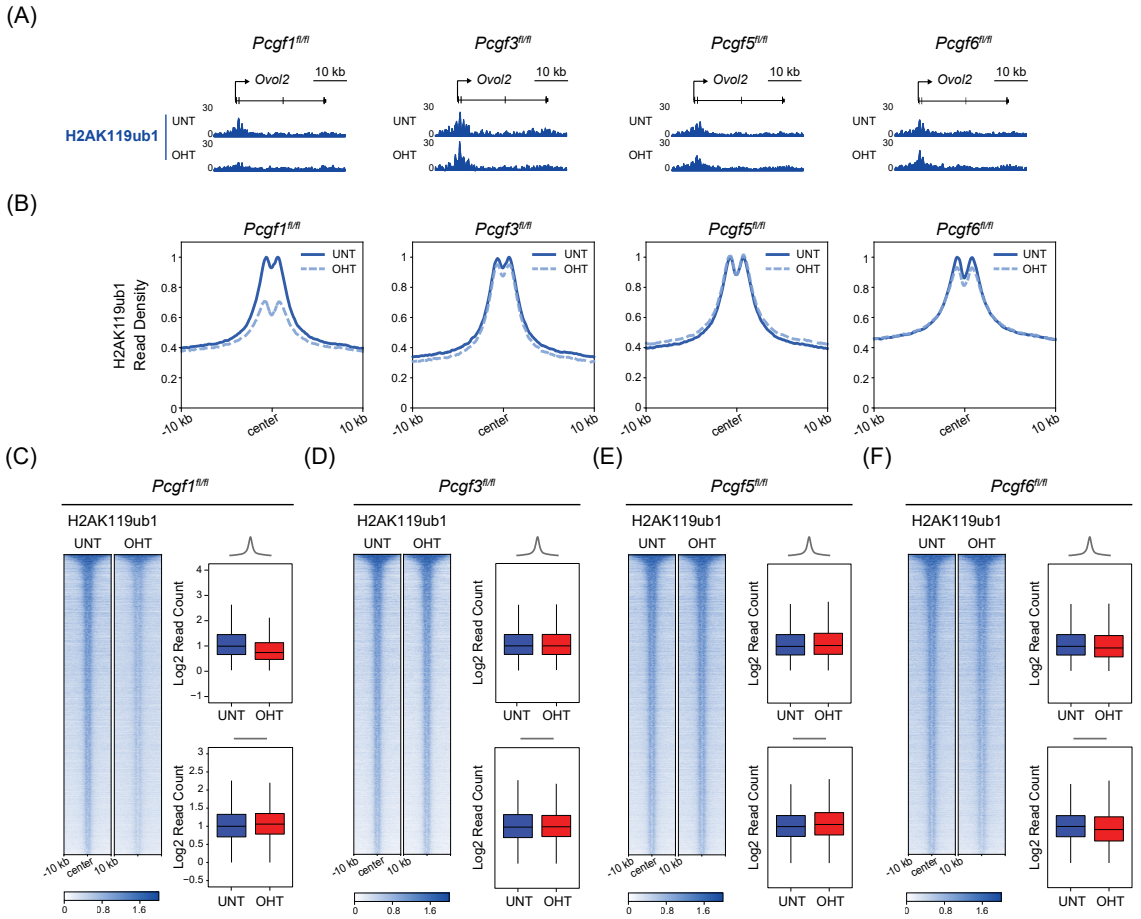



**Figure 4.4: PCGF1-PRC1 plays a major role in deposition of H2AK119ub1 at PRC1-bound sites**
(A) Genomic snapshots of a typical Polycomb target gene showing H2AK119ub1 cChIP-seq in *Pcgf1*$^{fl/fl}$, *Pcgf3*$^{fl/fl}$, *Pcgf5*$^{fl/fl}$, and *Pcgf6*$^{fl/fl}$ cells (untreated and OHT-treated).
(B) Metaplot analysis of H2AK119ub1 cChIP-seq at PRC1-bound sites (n = 8833) in *Pcgf1*$^{fl/fl}$, *Pcgf3*$^{fl/fl}$, *Pcgf5*$^{fl/fl}$, and *Pcgf6*$^{fl/fl}$ cells (untreated and OHT-treated).
(C) *Left panel:* Heatmap analysis of H2AK119ub1 cChIP-seq at PRC1-bound sites (n = 8833) in *Pcgf1*$^{fl/fl}$ cells (untreated and OHT-treated). Genomic intervals are sorted as in Figure 4.1. *Right panel:* Box plots of the normalised H2AK119ub1 cChIP-seq signal at PRC1-bound sites (⋀) and 100 kb genomic windows (—) in *Pcgf1*$^{fl/fl}$ cells (untreated and OHT-treated).
(D) As in (C) for *Pcgf3*$^{fl/fl}$ cells.
(E) As in (C) for *Pcgf5*$^{fl/fl}$ cells.
(F) As in (C) for *Pcgf6*$^{fl/fl}$ cells.

that removal of PCGF2/4-containing PRC1 complexes had a minimal effect on H2AK119ub1 and gene repression (Chapter 3), I focused my attention on variant PRC1 complexes which form around PCGF1, PCGF3, PCGF5, and PCGF6 (Figure 4.3A) (Gao et al. 2012; Hauri et al. 2016). To systematically dissect the role of distinct PRC1 complexes in shaping H2AK119ub1 throughout the genome, I characterised a panel of conditional knockout mouse ES cell lines in which addition of tamoxifen leads to deletion of individual PCGF proteins (Figure 4.3B). Using this series of cell lines, I first examined whether removal of individual variant PRC1 complexes had any global effect on H2AK119ub1 levels by western blot analysis of bulk histones. However, I found that

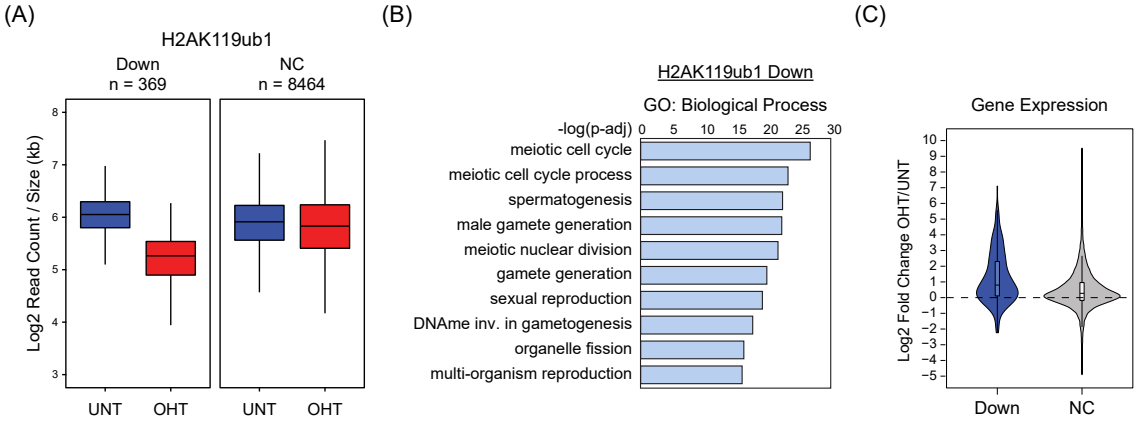



**Figure 4.5: PCGF6-PRC1 is responsible for deposition of H2AK119ub1 at a subset of germline-specific genes**
(A) Box plots of the H2AK119ub1 cChIP-seq signal at PRC1-bound sites which showed a significant reduction (n = 369, p-adj<0.05, Down) or no change (NC) in H2AK119ub1 levels following OHT treatment in the untreated and OHT-treated *Pcgf6*$^{fl/fl}$ cells.
(B) A gene ontology analysis of Biological Process term enrichment for genes showing a significant reduction (p-adj<0.05) in H2AK119ub1 at the promoter in *Pcgf6*$^{fl/fl}$ cells following OHT treatment.
(C) A violin plot comparing log2 fold expression changes for genes showing a significant reduction (p-adj<0.05) (Down) or no change (NC) in H2AK119ub1 at the promoter in *Pcgf6*$^{fl/fl}$ cells following OHT treatment. RNA-seq data is from Endoh et al. 2017.

total levels of H2AK119ub1 remained largely unchanged following OHT treatment in the PCGF conditional knockout cell lines (Figure 4.3C), suggesting that the pervasive genome-wide pool, which accounts for the majority of H2AK119ub1, was not majorly affected.

Given that individual variant PRC1 complexes contributed minimally to deposition of bulk H2AK119ub1, I was interested to explore whether instead these complexes were responsible for placing H2AK119ub1 at specific genomic loci. Therefore, I performed cChIP-seq for H2AK119ub1 before and after removal of either PCGF1, PCGF3, PCGF5 or PCGF6. This revealed that ablation of PCGF1 on its own led to a major but punctate reduction in H2AK119ub1 specifically at PRC1-bound sites (5284 sites (60%) showed a significant reduction) (Figure 4.4A-C). This is in agreement with previous work which showed that PCGF1-PRC1 is targeted to Polycomb-occupied NMIs via a ZF-CXXC domain of its KDM2B subunit (Farcas et al. 2012; He et al. 2013a; Wu et al. 2013). In contrast to PCGF1, removal of either PCGF3 or PCGF5 had a minor non-significant effect on H2AK119ub1 both at PRC1-bound sites and throughout the genome (Figure 4.4A-B and D-E). Finally, following PCGF6 removal, levels of H2AK119ub1 at PRC1-bound sites and throughout the genome were largely unaffected (Figure 4.4A-B and F). Interestingly, I did find a small number of sites (n = 369) at which H2AK119ub1 was significantly reduced in the absence of PCGF6 (Figure 4.5A). These PCGF6-dependent sites were preferentially associated with promoters of germ cell-specific genes (Figure 4.5B),

which became derepressed following PCGF6 removal (Figure 4.5C). These observations link PCGF6-mediated transcriptional silencing and deposition of H2AK119ub1, further supporting the previously proposed role of PCGF6 in repressing genes that are specific to the germ cell lineage (Endoh et al. 2017; Zhao et al. 2017b). In conclusion, through systematic removal of individual PCGF proteins, I found that PCGF1-containing PRC1 complexes play a major role in placing H2AK119ub1 at PRC1-bound sites.

### 4.2.3 PCGF1-PRC1 supports Polycomb chromatin domain formation and is essential but not sufficient for Polycomb-mediated gene repression

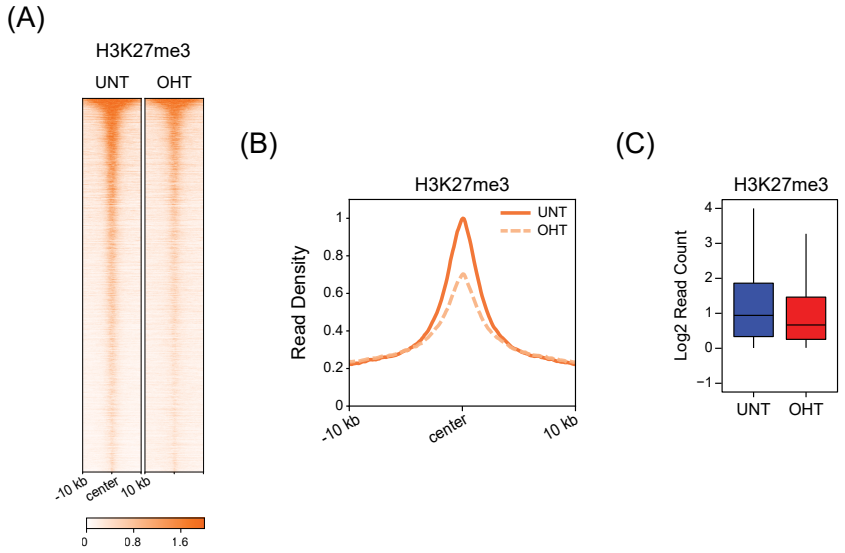



**Figure 4.6: PCGF1-PRC1 supports Polycomb chromatin domain formation at PRC1-bound sites**
(A) Heatmap analysis of H3K27me3 cChIP-seq at PRC1-bound sites (n = 8833) in *Pcgf1*$^{fl/fl}$ cells (untreated and OHT-treated). Genomic intervals are sorted as in Figure 4.1.
(B) Metaplot analysis of H3K27me3 cChIP-seq at PRC1-bound sites (n = 8833) in *Pcgf1*$^{fl/fl}$ cells (untreated and OHT-treated).
(C) Box plots of the normalised H3K27me3 cChIP-seq signal at PRC1-bound sites in *Pcgf1*$^{fl/fl}$ cells (untreated and OHT-treated).

Having established that PCGF1-PRC1 makes a major contribution to deposition of H2AK119ub1 at PRC1-bound sites, I asked whether this complex is also required to support PRC2 binding/activity and formation of Polycomb chromatin domains at these genomic regions. First of all, I confirmed that H3K27me3 was enriched across all PRC1-bound sites in untreated ES cells (Figures 4.6A-B), indicating that PRC1 and PRC2 function at the same target sites in the genome (herein referred to as Polycomb target sites), consistent with previous observations (Bernstein et al. 2006a; Boyer et al. 2006; Bracken et al. 2006; Mikkelsen et al. 2007). Remarkably, in agreement with the proposed role for H2AK119ub1 in shaping PRC2 occupancy (Blackledge et al. 2014; Cooper et al. 2014), removal of PCGF1 resulted in a considerable reduction in H3K27me3 at these

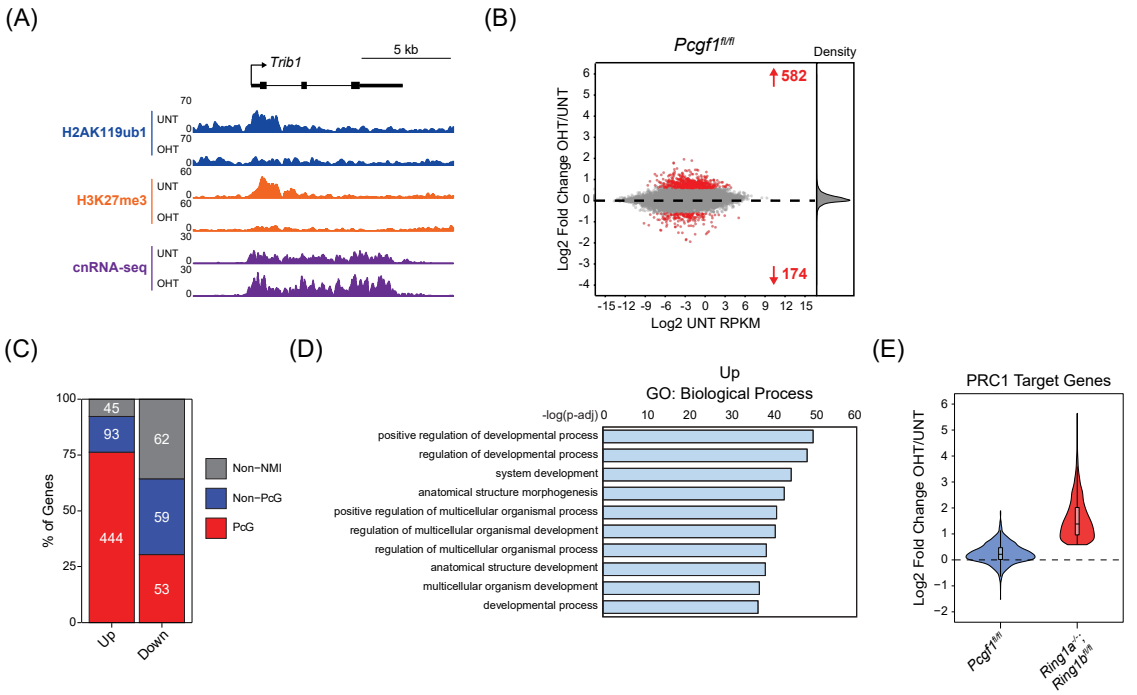



**Figure 4.7: PCGF1-PRC1 is essential but not sufficient for Polycomb-mediated gene repression**
(A) Genomic snapshot of a typical Polycomb target gene showing H2AK119ub1 and H3K27me3 cChIP-seq, as well as gene expression (cnRNA-seq) in *Pcgf1*$^{fl/fl}$ cells (untreated and OHT-treated).
(B) An MA-plot of log2 fold changes in gene expression (cnRNA-seq) in *Pcgf1*$^{fl/fl}$ cells following OHT treatment. Significant gene expression changes (p-adj<0.05 and FC>1.5-fold) are shown in red. Density of gene expression changes is shown on the right.
(C) A bar plot illustrating the distribution of gene expression changes (p-adj<0.05 and FC>1.5-fold) in cnRNA-seq between the gene classes (Non-NMI, Non-PcG, PcG) described in Figure 3.4B for *Pcgf1*$^{fl/fl}$ cells following OHT treatment.
(D) A gene ontology analysis of Biological Process term enrichment for genes showing a significant increase in expression (p-adj<0.05 and FC>1.5-fold) in *Pcgf1*$^{fl/fl}$ cells following OHT treatment (n = 582).
(E) A violin plot comparing log2 fold expression changes for PRC1 target genes in *Pcgf1fl/fl* and *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ cells following OHT treatment.

sites (4095 sites (46%) showed a significant decrease), which was comparable to the observed reduction in H2AK119ub1 (Figures 4.6A-C). Furthermore, expression of several hundred genes was significantly increased when PCGF1 was removed (Figures 4.7A-B). Importantly, the majority of these transcriptional changes corresponded to Polycomb target genes involved in control of embryonic development (Figures 4.7C-D), indicating that PCGF1-PRC1 plays an essential role in Polycomb-mediated gene repression in mouse ES cells. However, both the number of derepressed genes and the magnitude of gene reactivation following PCGF1 removal were considerably smaller than following removal of all PRC1 complexes in *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ cells (Figure 4.7E). Together, these observations demonstrate that although PCGF1-PRC1 is essential to place normal H2AK119ub1 levels and reinforce Polycomb chromatin domain formation at Polycomb target sites, its loss does not recapitulate the full extent of Polycomb target gene derepression following simultaneous removal of all PRC1 complexes.

## 4.3 Discussion

In vertebrates, PRC1 and PRC2 each comprise an array of biochemically distinct protein assemblies with unique targeting modules and regulatory activities, the exact contribution of which to Polycomb-mediated gene repression is poorly understood (reviewed in Scelfo et al. 2014; Schuettengruber et al. 2017; Bajusz et al. 2018; Vidal 2019). For years, the multiplicity of the Polycomb system has hampered our understanding of the primary mechanisms that underlie transcriptional silencing. In Chapter 4, I sought to overcome this complexity by developing a simple screen to systematically dissect the contribution of distinct PRC1 complexes to gene repression. Previously, it has been shown that PRC2 is largely dispensable for Polycomb-mediated gene repression in mouse ES cells (Riising et al. 2014). In light of these observations, my findings in Chapter 3 suggested that H2AK119ub1 and variant PRC1 complexes may be responsible for gene repression. Therefore, here I utilised a quantitative cChIP-seq approach to first characterise the genomic distribution of H2AK119ub1 and then identify which PRC1 complexes make the largest contribution to shaping H2AK119ub1 in the genome. Unexpectedly, this revealed that H2AK119ub1 is a pervasive histone modification that is enriched at PRC1-bound sites but also blankets the genome at a low level. Next, through systematic removal of individual variant PRC1 complexes, I uncovered that PCGF1-PRC1 plays a major role in placing H2AK119ub1 and stimulating H3K27me3 at Polycomb target sites. However, although PCGF1-containing PRC1 complexes were essential for proper silencing of a subset of Polycomb target genes, they were not sufficient to fully explain Polycomb-mediated gene repression in mouse ES cells.

Polycomb repressive complexes are typically considered to function at their target gene promoters, where they deposit high levels of H2AK119ub1 and H3K27me3. However, Polycomb-associated chromatin modifications have been shown to be extremely abundant in cells, and this cannot be solely explained by the activity of PRC1 and PRC2 at gene promoters, as these regions constitute less than 1% of the genome (Goldknopf et al. 1975; Albright et al. 1979; Matsui et al. 1979; Peters et al. 2003; Voigt et al. 2012; Ferrari et al. 2014; Lee et al. 2015). In mouse ES cells, PRC2 pervasively modifies chromatin throughout the genome by catalysing different H3K27 methylation states, which in total account for more than 80% of histone H3 (Ferrari et al. 2014). These pervasive chromatin-modifying activities of PRC2 have been proposed to counteract inappropriate activation of enhancers and transcription of intergenic regions (Ferrari et al. 2014; Lee et al. 2015).

Similarly, according to early biochemical quantifications in mammalian cells, around 10% of histone H2A is mono-ubiquitinated at lysine 119 (Goldknopf et al. 1975; Albright et al. 1979; Matsui et al. 1979). However, the genomic distribution of this histone modification outside of the classical Polycomb chromatin domains has remained poorly characterised. In *Drosophila*, moderate enrichment of histone H2A mono-ubiquitination was observed at a small number of intergenic regions and active gene promoters (Lee et al. 2015; Kahn et al. 2016). In line with this, when I characterised the genome-wide profile of H2AK119ub1 in mouse ES cells, I found that in addition to being enriched at PRC1-bound sites, especially those associated with Polycomb-repressed gene promoters, H2AK119ub1 is also placed more globally throughout the genome, albeit at lower levels. Interestingly, I did not detect a corresponding genome-wide binding pattern for RING1B, indicating that although PRC1 must interact with chromatin throughout the genome to place H2AK119ub1, this binding is likely transient and thereby not effectively captured by ChIP-seq. Moving forward, an important challenge remains to uncover the mechanisms by which these distinct pools of H2AK119ub1 are deposited and what role they play in regulation of gene expression and other chromatin-based processes.

In addition to canonical histone H2A, PRC1 also mono-ubiquitinates histone H2A variants (Ogawa et al. 2005; Sarcinella et al. 2007; Ismail et al. 2013). Interestingly, one of these histone variants, H2A.Z, is highly enriched at promoters of Polycomb target genes in mouse ES cells, where it contributes to gene silencing, and this is crucial for proper ES cell differentiation (Creyghton et al. 2008; Wang et al. 2018c). Although H2A.Z mono-ubiquitination is also enriched at Polycomb target sites in ES cells (Surface et al. 2016), the exact contribution of this chromatin modification to gene repression by the Polycomb system remains poorly understood. In the future, it would be interesting to tackle this question by extending the quantitative cChIP-seq analysis that I have validated here for H2AK119ub1 to characterise the genomic distribution of the ubiquitinated forms of H2A.Z and other histone H2A variants using specific antibodies against these modifications.

Although *in vitro*, variant PRC1-specific PCGF-RING1B dimers have similar E3 ubiquitin ligase activity, *in vivo* they form complexes with different auxiliary subunits, which may additionally modulate their catalytic activity and specify targeting to distinct genomic regions (Gao et al. 2012; Taherbhoy et al. 2015; Hauri et al. 2016). Artificial tethering of PCGF1/3/5/6 but not PCGF2/4 to an ectopic genomic site led to accumulation of H2AK119ub1, indicating that variant PRC1 complexes can place H2AK119ub1 *in vivo* (Blackledge et al. 2014; Endoh et al. 2017).

However, where and how individual PRC1 complexes normally contribute to deposition of H2AK119ub1 and gene repression in the genome has until now remained poorly understood. Here, I utilised H2AK119ub1 cChIP-seq as a simple screen to identify genomic regions where individual variant PRC1 complexes function. I found that PCGF1-PRC1, which contains a ZF-CXXC domain protein KDM2B, plays a major role in placing H2AK119ub1 at Polycomb target sites, the majority of which overlap with non-methylated CpG islands (Ku et al. 2008). For PCGF3/5-PRC1, which do not stably incorporate any subunits with characterised DNA-binding activities (Hauri et al. 2016), I did not find specific genomic sites at which H2AK119ub1 was significantly affected following removal of these complexes. In line with the minimal effect on H2AK119ub1 that has been previously reported for constitutive single deletions of *Pcgf3* and *Pcgf5* genes (Zhao et al. 2017a), this suggests that PCGF3- and PCGF5-PRC1 complexes may be partially redundant. Finally, I found that PCGF6-PRC1 is required for deposition of H2AK119ub1 at promoters of germline-specific genes, which were shown to be enriched with DNA-binding motifs of MAX/MGA transcription factors (Endoh et al. 2017; Scelfo et al. 2019). Together, these observations suggest that variant PRC1 complexes rely on recognition of DNA sequence motifs or features to select target sites where they deposit high levels of H2AK119ub1. However, how much different DNA-based mechanisms contribute to shaping stable occupancy of variant PRC1 complexes at different genomic regions and whether there are additional sequence-independent mechanisms that can also contribute to this process needs to be further explored.

Based on my findings in Chapter 3, I hypothesised that H2AK119ub1 may be related to repression of Polycomb target genes. Here, I demonstrated that PCGF1- and PCGF6-PRC1 complexes both contribute to deposition of H2AK119ub1 at Polycomb target sites, albeit to a different extent. However, do these complexes also contribute to transcriptional silencing at their target genomic regions? Earlier studies reported that PCGF1 and PCGF6 act as activators of transcription in mouse ES cells as expression of a large number of genes was significantly reduced in PCGF1 and PCGF6 knockout cells (Yan et al. 2017; Yang et al. 2016). However, the interpretation of these experiments is complicated by the fact that constitutive deletion was used to perturb PCGF1/PCGF6-PRC1 complexes and transcriptional changes were analysed using traditional genomic approaches which are inherently non-quantitative. In contrast, more recent studies relying on conditional knockout approaches showed that PCGF6-PRC1 occupies promoters of germ cell-specific genes, where it is required for gene repression, in line with my hypothesis

(Endoh et al. 2017). In the case of PCGF1-PRC1, I found that in addition to its essential role in placing H2AK119ub1, PCGF1-PRC1 also supports formation of Polycomb chromatin domains and is required for repression of a subset of Polycomb target genes. This is in agreement with earlier reports of PCGF1 repressive activity in luciferase reporter assays and previous work from our lab showing that artificial tethering of PCGF1 was sufficient to initiate recruitment of PRC2 and deposition of H3K27me3 at an ectopic genomic region (Gong et al. 2004; Blackledge et al. 2014). Therefore, taken together, my observations provide additional evidence that links PRC1-mediated gene repression to deposition of H2AK119ub1.

Systematic removal of individual PCGF proteins did not recapitulate the effects of removing RING1A/B, suggesting that cooperation between multiple variant PRC1 complexes may be required to shape H2AK119ub1 in the genome and repress transcription. Indeed, although it has been proposed that distinct PRC1 complexes have specialised and non-overlapping roles in gene regulation (Gao et al. 2012; Sowpati et al. 2015), several studies have reported that some PRC1 complexes can cooperate to silence shared target genes (Bel et al. 1998; Zhao et al. 2017a; Almeida et al. 2017). In line with this idea, some PCGF proteins have highly similar primary structures and form nearly identical PRC1 complexes (Sowpati et al. 2015; Hauri et al. 2016). Therefore, to uncover any potential cooperation between individual PRC1 complexes, in the next chapters, I will investigate how simultaneous removal of PCGF proteins in different combinations affects deposition of H2AK119ub1, formation of Polycomb chromatin domains and repression of Polycomb target genes. Identifying which PRC1 complexes function together to define Polycomb-mediated gene repression and which of their specific shared activities are required to achieve this will help to shed the light on the molecular determinants of transcriptional silencing by the Polycomb system.

# 5

# DYNAMIC CONTROL OF GENOME-WIDE H2AK119UB1 AND ITS CONTRIBUTION TO GENE REGULATION

## 5.1 Introduction

In recent years, an increasing number of studies have suggested that the Polycomb system functions more broadly throughout the genome than has been previously anticipated (Ferrari et al. 2014; Lee et al. 2015; Kahn et al. 2016). In addition to repressed Polycomb target genes, PRC1 has been shown to bind promoters of actively transcribed genes, although the functional consequences of this binding remain poorly understood (Schwartz et al. 2010; Schaaf et al. 2013; Pherson et al. 2017; Cohen et al. 2018). In Chapter 4, I demonstrated that the way PRC1 engages with the genome is even more widespread than this, as H2AK119ub1 is deposited broadly throughout the genome, including gene bodies and intergenic regions, albeit at a lower level than at classical Polycomb target sites. However, whether this genome-wide pool of H2AK119ub1 contributes to regulation of gene expression or other chromosomal processes, or is simply a by-product of non-specific PRC1 activity remains largely unknown. Therefore, in order to address this important question, the mechanisms that control the levels of pervasive H2AK119ub1 throughout the genome first need to be defined.

Systematic removal of individual PCGF proteins had no major effect on the genome-wide pool of H2AK119ub1 (Chapter 4), suggesting that PRC1 complexes functionally cooperate to place this histone modification throughout the genome. I reasoned that the most closely related PCGF proteins are more likely to form biochemically similar PRC1 complexes and underpin any potential cooperation. Apart from PCGF2 and PCGF4, which form PRC1 complexes that contribute little to deposition of H2AK119ub1, PCGF3 and PCGF5 proteins share the highest sequence similarity (Figure 5.1A) (Sowpati et al. 2015) and assemble into nearly identical protein complexes (Gao et al. 2012; Hauri et al. 2016). Functional redundancy between PCGF3- and PCGF5-PRC1 has been previously reported in the context of X chromosome inactivation, as both

of them can initiate Xist-mediated recruitment of Polycomb complexes and chromosome-wide transcriptional silencing (Almeida et al. 2017; Pintacuda et al. 2017). During X-inactivation, PCGF3/5-PRC1 support accumulation of H2AK119ub1 on the silenced X chromosome as evident by immunofluorescence microscopy (Almeida et al. 2017). However, where exactly H2AK119ub1 is placed on the X chromosome in the course of the chromosome-wide gene repression and how these two processes are related remains poorly understood. Furthermore, since none of the proteins that stably associate with PCGF3/5-PRC1 contain known DNA-binding domains (Gao et al. 2012; Hauri et al. 2016), the mechanisms through which PCGF3/5-PRC1 complexes recognise their target sites in the genome are largely unknown. Finally, there is currently no consensus on the role of PCGF3/5-PRC1 in gene regulation, as some studies report that they act as transcriptional activators (Gao et al. 2014; Zhao et al. 2017a; Yao et al. 2018), while others show that these complexes are required for gene silencing (Almeida et al. 2017; Pintacuda et al. 2017).

In addition to PRC1-mediated deposition, the levels of H2AK119ub1 throughout the genome are controlled by deubiquitinating enzymes (DUBs) that can specifically remove this histone modification (reviewed in Vissers et al. 2008; Belle and Nijnik 2014). In *Drosophila*, a complex between two PcG proteins, Calipso and ASX, known as Polycomb repressive deubiquitinase (PR-DUB), is responsible for deubiquitination of H2AK118ub1 (Scheuermann et al. 2010). Surprisingly, mutations in PR-DUB subunits lead to loss of Polycomb target gene silencing and homeotic transformations in flies (Scheuermann et al. 2010). The requirement for PR-DUB activity, which is supposed to counteract PRC1, in Polycomb-mediated gene repression is somewhat counterintuitive, and a mechanistic explanation for how this activity could contribute to transcriptional silencing by the Polycomb system is still missing. Mammalian orthologue of PR-DUB complex forms around BAP1 and is also capable of catalysing H2AK119ub1 deubiquitination *in vitro* (Jensen et al. 1998; Scheuermann et al. 2010; Hauri et al. 2016; Kloet et al. 2016). However, in mammals, it remains unclear whether BAP1 is also the major H2AK119ub1 DUB *in vivo*, since multiple DUB complexes with specificity for H2AK119ub1 have been characterised (Nijman et al. 2005; Joo et al. 2007; Zhu et al. 2007; Nakagawa et al. 2008; Sowa et al. 2009; Yang et al. 2014). BAP1 plays an important role in transcriptional control of cell proliferation and is often mutated in cancer (reviewed in Carbone et al. 2013). A large repertoire of chromatin binders and modifiers, as well as transcription factors and cell cycle regulators associate with BAP1 in mammals (Machida et al. 2009; Misaghi et al. 2009; Yu et al. 2010; Ji et al. 2014;

Balasubramani et al. 2015; Hauri et al. 2016; Kloet et al. 2016). This has made it difficult to discern to what extent BAP1 function in gene regulation relies on its H2AK119ub1 DUB activity. Finally, although BAP1 has been reported to antagonise the Polycomb system in the context of cancer cells (LaFave et al. 2015; Wang et al. 2018a), the role of BAP1 in counteracting Polycomb activity during normal embryonic development remains poorly understood.

Despite the growing evidence that PRC1 may function more broadly throughout the genome than previously appreciated, whether this widespread PRC1 activity contributes to regulation of gene expression or other chromosomal processes remains poorly understood. In Chapter 5, I aim to uncover the mechanisms that define deposition and removal of the genome-wide pool of H2AK119ub1 and then determine if they broadly contribute to transcriptional regulation throughout the genome. First, I ask whether cooperation between specific PRC1 complexes is responsible for deposition of low-level H2AK119ub1 across the genome. To address this question, I establish a mouse ES cell line in which highly similar PCGF3/5-PRC1 complexes can be removed together and examine the effect of their combined loss on genome-wide H2AK119ub1, Polycomb chromatin domain formation and gene repression. Furthermore, since PCGF3/5-PRC1 have been previously implicated in X-inactivation, I investigate whether H2AK119ub1 accumulation during Xist-mediated chromosome silencing is punctate or pervasive and examine how this is linked to transcriptional repression. Finally, I develop a mouse ES cell line in which BAP1 deubiquitinase complex can be conditionally removed and investigate whether its loss leads to an increase in H2AK119ub1 levels in the genome. I then examine the functional consequences of BAP1 removal on gene regulation, particularly focusing on the antagonism between BAP1 and the Polycomb system.

## 5.2 Results

### 5.2.1 PCGF3/5-PRC1 play a major role in deposition of the genome-wide pool of H2AK119ub1

Through systematic knockout analysis of individual PCGF proteins, I found that the genome-wide pool of H2AK119ub1 must be placed by the combined activity of several PRC1 complexes (Chapter 4). Since PCGF3-PRC1 and PCGF5-PRC1 associate with the same auxiliary subunits and biochemically are almost identical (Gao et al. 2012; Hauri et al. 2016), I reasoned that

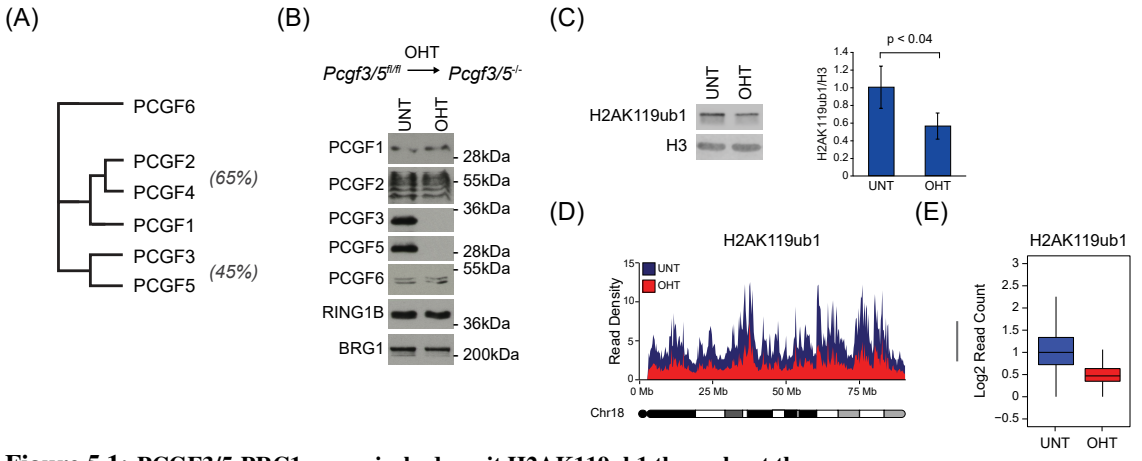



**Figure 5.1: PCGF3/5-PRC1 pervasively deposit H2AK119ub1 throughout the genome**
(A) Neighbour-joining phylogenetic tree of amino acid sequences for mouse PCGF proteins. Percentage of identity for PCGF2/4 and PCGF3/5 proteins is shown in brackets on the right.
(B) Western blot analysis of PCGF proteins and RING1B in *Pcgf3/5*$^{fl/fl}$ cells before (UNT) and after tamoxifen treatment (OHT). BRG1 is shown as a loading control.
(C) Western blot analysis of H2AK119ub1 in *Pcgf3/5*$^{fl/fl}$ cells (untreated and OHT-treated). Histone H3 is shown as a loading control. Also shown is quantification of H2AK119ub1 levels relative to histone H3. Error bars represent SEM (n=3). P-value denotes statistical significance calculated by a paired Student's t-test.
(D) A chromosome density plot showing H2AK119ub1 cChIP-seq across chromosome 18 in *Pcgf3/5*$^{fl/fl}$ cells (untreated and OHT-treated).
(E) Box plot analysis of the normalised H2AK119ub1 cChIP-seq signal at 100 kb genomic windows (—) in *Pcgf3/5*$^{fl/fl}$ cells (untreated and OHT-treated).

they were likely to be at least partially functionally redundant and likely to cooperate to deposit pervasive H2AK119ub1. Therefore, to mitigate against possible compensation between these complexes, I established a *Pcgf3/5*$^{fl/fl}$ mouse ES cell line in which simultaneous removal of PCGF3 and PCGF5 can be conditionally induced (Figure 5.1B). Remarkably, in contrast to knockout of PCGF3 or PCGF5 individually, which had no effect on H2AK119ub1 (Chapter 4), loss of PCGF3/5 together resulted in an almost 50% reduction in bulk H2AK119ub1 levels as shown by western blot analysis of histone extracts (Figure 5.1C). Furthermore, a major and uniform depletion of H2AK119ub1 throughout the genome was also evident by cChIP-seq when H2AK119ub1 distribution was visualised at the chromosome level (Figure 5.1D) or quantified in large windows across the genome (Figure 5.1E). Together, these analyses revealed that PCGF3/5-PRC1 deposit the majority of low-level H2AK119ub1 genome-wide. Importantly, this also demonstrated that the double PCGF3/5 conditional knockout system can be used to specifically manipulate the levels of H2AK119ub1 blanket, enabling me to address its role in Polycomb system function and gene regulation more generally.

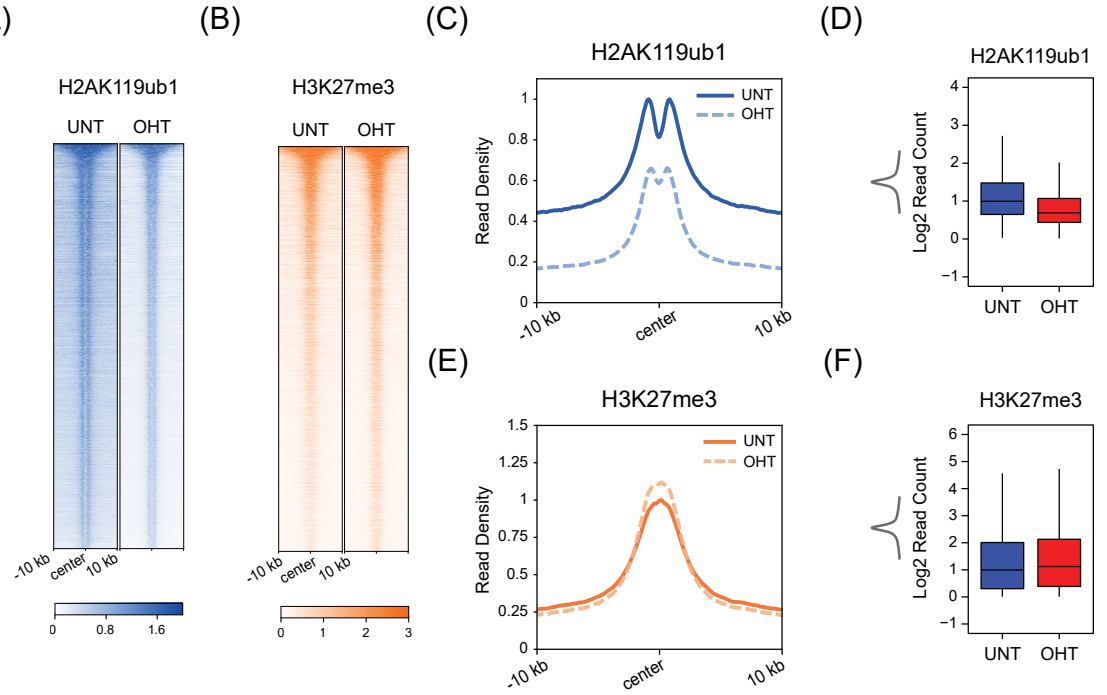



**Figure 5.2: PCGF3/5-dependent H2AK119ub1 is not required for Polycomb chromatin domain formation**
(A) Heatmap analysis of H2AK119ub1 cChIP-seq at PRC1-bound sites (n = 8833) in *Pcgf3/5*$^{fl/fl}$ cells (untreated and OHT-treated). Genomic intervals are sorted as in Figure 4.1.
(B) Heatmap analysis of H3K27me3 cChIP-seq at PRC1-bound sites (n = 8833) in *Pcgf3/5*$^{fl/fl}$ cells (untreated and OHT-treated). Genomic intervals are sorted as in Figure 4.1.
(C) Metaplot analysis of H2AK119ub1 cChIP-seq at PRC1-bound sites (n = 8833) in *Pcgf3/5*$^{fl/fl}$ cells (untreated and OHT-treated).
(D) Box plot analysis of the normalised H2AK119ub1 cChIP-seq signal at PRC1-bound sites (⋀) in *Pcgf3/5*$^{fl/fl}$ cells (untreated and OHT-treated).
(E) Metaplot analysis of H3K27me3 cChIP-seq at PRC1-bound sites (n = 8833) in *Pcgf3/5*$^{fl/fl}$ cells (untreated and OHT-treated).
(F) Box plot analysis of the normalised H3K27me3 cChIP-seq signal at PRC1-bound sites (⋀) in *Pcgf3/5*$^{fl/fl}$ cells (untreated and OHT-treated).

### 5.2.2 Pervasive deposition of H2AK119ub1 by PCGF3/5-PRC1 is not required for Polycomb chromatin domain formation and gene repression

Having established that removal of PCGF3/5 leads to a global reduction in H2AK119ub1 throughout the genome, I next examined whether H2AK119ub1 levels were also reduced at highly enriched Polycomb target sites. Following OHT treatment in *Pcgf3/5*$^{fl/fl}$ cells, I observed a largely uniform reduction in H2AK119ub1 across PRC1-bound sites and their flanking regions (Figures 5.2A and C). This indicates that PCGF3/5-PRC1 complexes deposit H2AK119ub1 broadly throughout the genome, without specifically discriminating between Polycomb target sites and the rest of the genome. Although the absolute levels of H2AK119ub1 decreased at PRC1-bound sites following PCGF3/5 removal, the punctate enrichment of this histone modification was retained (Figures 5.2A, C and D, and Figure 5.3A). Strikingly, enrichment of H3K27me3

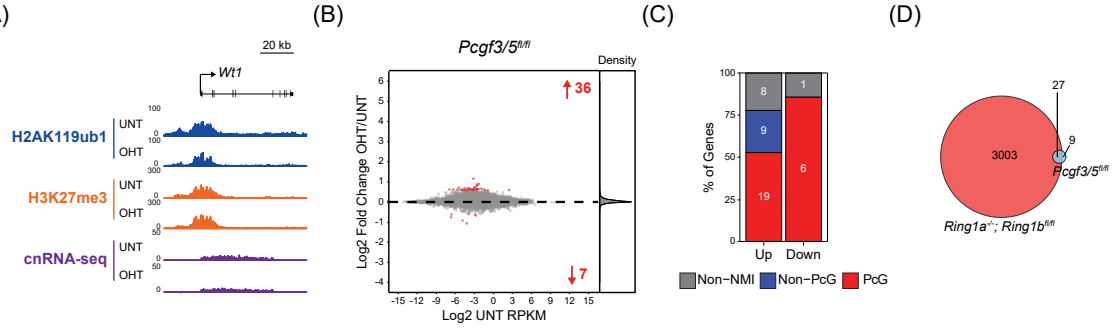



**Figure 5.3: Pervasive deposition of H2AK119ub1 by PCGF3/5-PRC1 contributes little to Polycomb-mediated gene repression**
(A) Genomic snapshot of a typical Polycomb target gene showing H2AK119ub1 and H3K27me3 cChIP-seq, as well as gene expression (cnRNA-seq) in *Pcgf3/5*$^{fl/fl}$ cells (untreated and OHT-treated).
(B) An MA-plot of log2 fold changes in gene expression (cnRNA-seq) in *Pcgf3/5*$^{fl/fl}$ cells following OHT treatment. Significant gene expression changes (p-adj<0.05 and FC>1.5-fold) are shown in red. Density of gene expression changes is shown on the right.
(C) A bar plot illustrating the distribution of gene expression changes (p-adj<0.05 and FC>1.5-fold) in cnRNA-seq between the gene classes (Non-NMI, Non-PcG, PcG) described in Figure 3.4B for *Pcgf3/5*$^{fl/fl}$ cells following OHT treatment.
(D) A Venn diagram showing an overlap of significantly derepressed genes (p-adj<0.05 and FC>1.5-fold) in *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ and *Pcgf3/5*$^{fl/fl}$ cells following OHT treatment.

at these sites was also largely unaffected and even slightly increased (Figures 5.2B, E and F, and Figure 5.3A), suggesting that the relative enrichment of H2AK119ub1 rather than its absolute level facilitates PRC2 binding/activity at Polycomb target sites. Importantly, only a small number of Polycomb target genes were derepressed following removal of PCGF3/5 and the resulting genome-wide depletion of H2AK119ub1 (Figures 5.3A-D). Finally, I found no evidence to indicate that PCGF3/5-PRC1 function as transcriptional activators, which has been previously reported using luciferase reporter assays and constitutive knockout systems (Gao et al. 2014; Zhao et al. 2017a; Yao et al. 2018; Scelfo et al. 2019) (Figure 5.3B). Together, these observations demonstrate that pervasive deposition of H2AK119ub1 by PCGF3/5-PRC1 is not required to support formation of Polycomb chromatin domains or repression of Polycomb target genes.

### 5.2.3 Pervasive PCGF3/5-mediated H2AK119ub1 is linked to chromosome-wide gene silencing during X-inactivation

Although PCGF3/5-PRC1 were not required for classical Polycomb target gene repression, they have been shown to support accumulation of H2AK119ub1 and recruitment of Polycomb complexes at the silenced X chromosome (Almeida et al. 2017; Pintacuda et al. 2017). Therefore, I wondered whether pervasive deposition of H2AK119ub1 by PCGF3/5-PRC1 may be utilised for gene repression in specialised contexts, such as chromosome-wide transcriptional silencing during X-inactivation. To test this possibility, I examined where H2AK119ub1 is accumulated during

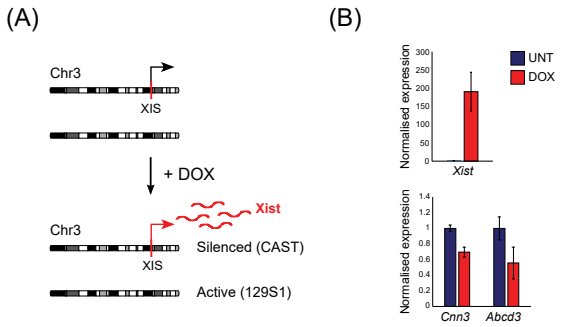



**Figure 5.4: An inducible system to study Xist-mediated chromosome-wide gene silencing**
(A) A schematic of the *Mus domesticus* (129S1) × *Mus castaneus* F1 hybrid ES cells, in which a doxycycline (DOX)-inducible Xist transgene is integrated into one allele of chromosome 3. DOX treatment leads to Xist-mediated silencing of the allele harbouring the Xist integration site (XIS) (silenced allele, from *Mus castaneus*, CAST), but not the other allele (active, from *Mus domesticus*, 129S1).
(B) RT-qPCR gene expression analysis for Xist transgene and two neighbouring genes before (UNT) and after DOX treatment. Expression is normalised to *b-actin* and shown relative to the average expression in UNT cells. Error bars show SEM (n = 3).

Xist-mediated chromosome inactivation and how this is related to gene repression. To this end, I performed allele-specific cChIP-seq analysis for H2AK119ub1 in an interspecies hybrid mouse ES cell line, in which a DOX-inducible Xist transgene has been integrated into a single autosomal allele (Figures 5.4A-B) (Pintacuda et al. 2017). I found that induction of Xist expression resulted in broad accumulation of H2AK119ub1 throughout the Xist-expressing chromosome (Figures 5.5A-D), with no preference for gene promoters or sites already occupied by PRC1 (Figures 5.5E-F).

The widespread distribution of H2AK119ub1 on the silenced chromosome, although of much higher level, closely resembled genome-wide deposition of H2AK119ub1 by PCGF3/5-PRC1 (Figure 5.1D), in line with the previously proposed role of these complexes in H2AK119ub1 accumulation on the inactive X chromosome (Almeida et al. 2017). The requirement for PCGF3/5-PRC1 in broad acquisition of H2AK119ub1 during X-inactivation is further supported by recent studies showing that, in cells lacking PCGF3/5, Xist induction fails to elevate H2AK119ub1 levels both at pre-existing PRC1-bound sites (Figure 5.6B) and more broadly throughout the Xist-expressing chromosome (Figure 5.6A-B) (Nesterova et al. 2019). Most importantly, spreading of H2AK119ub1 across the silenced chromosome scaled with the distance from the Xist integration site (XIS) (Figure 5.5A) and showed a good correlation with gene repression (Figure 5.5G), which also decreased further away from the XIS (Figure 5.5A). Together, these findings indicate that, in contrast to classical Polycomb-mediated gene repression, which is associated with punctate H2AK119ub1 at target gene promoters, large-scale transcriptional silencing during Xist-mediated chromosome inactivation is linked to pervasive deposition of H2AK119ub1 by PCGF3/5-PRC1.

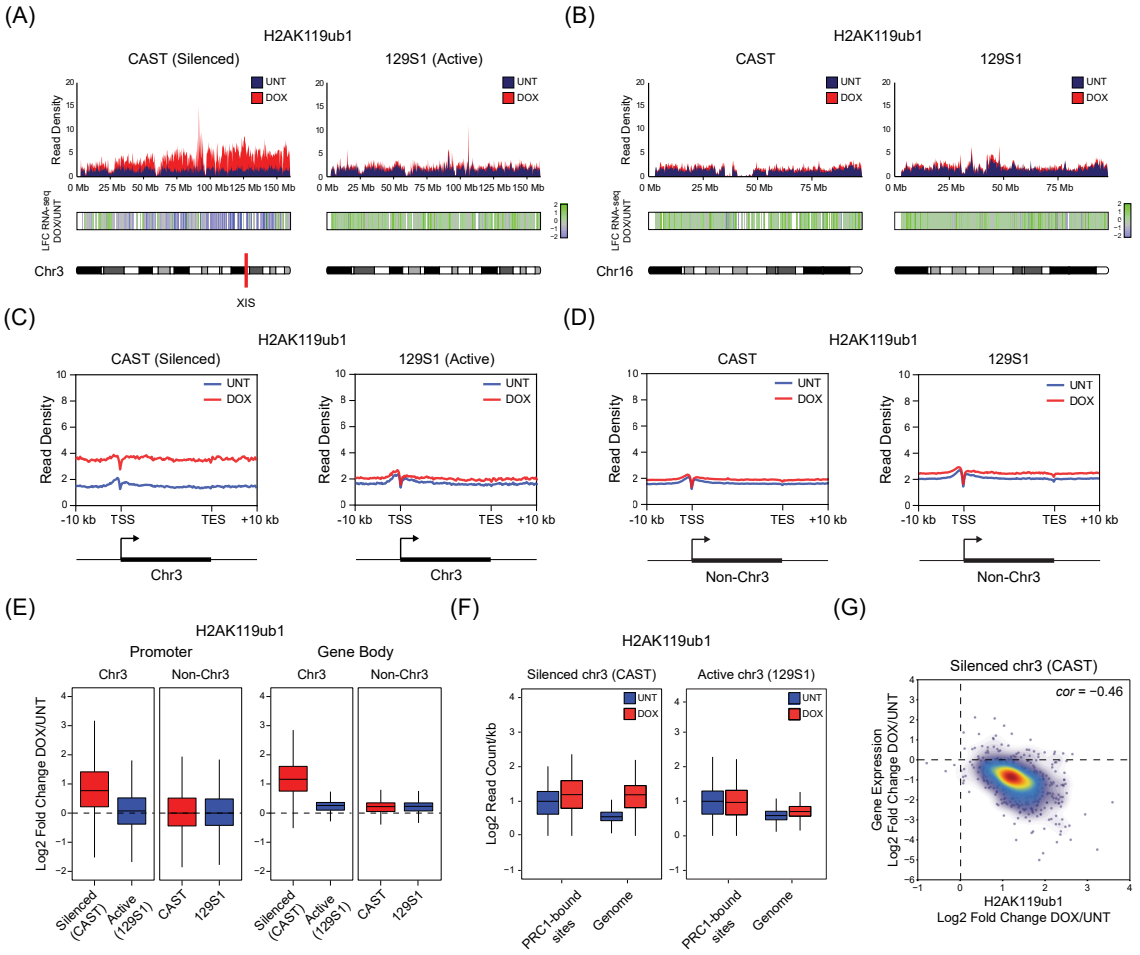



**Figure 5.5: Broad accumulation of H2AK119ub1 throughout the Xist-expressing chromosome is associated with transcriptional silencing**

(A) Chromosome density plots of H2AK119ub1 cChIP-seq across two alleles of chromosome 3 in the system described in Figure 5.4A before (UNT) and after DOX treatment. Below is a heatmap of corresponding log2 fold changes in gene expression upon addition of DOX. Bins containing no expressed genes are shown in white. XIS indicates position of the Xist integration site. The gene expression data from (Pintacuda et al. 2017) was used.

(B) As in (A) for two alleles of chromosome 16, which do not carry the *Xist* transgene.

(C) Allele-specific metaplot analysis of H2AK119ub1 cChIP-seq for genes located on chromosome 3 in the system described in Figure 5.4A before (UNT) and after DOX treatment. Genes were scaled to the same length and aligned at their TSS and TES.

(D) As in (C) for genes located on chromosomes other than chromosome 3.

(E) Allele-specific box plot analysis of log2 fold changes in H2AK119ub1 cChIP-seq at promoters (TSS ± 500 bp) or bodies of genes located on chromosome 3 (Chr3) or any other chromosome (Non-Chr3) following DOX treatment in the system described in Figure 5.4A.

(F) Allele-specific box plot analysis of normalised H2AK119ub1 cChIP-seq signal at PRC1-bound sites and 100 kb genomic windows for the silenced (*Mus castaneus*) or active (*Mus domesticus*) allele of chromosome 3 in the system described in Figure 5.4A before (UNT) and after DOX treatment.

(G) Scatter plot analysis comparing log2 fold changes in H2AK119ub1 and gene expression for all expressed 250 kb genomic bins across the silenced (*Mus castaneus*) allele of chromosome 3 following DOX treatment in the system described in Figure 5.4A. *Cor* denotes a Pearson correlation coefficient.

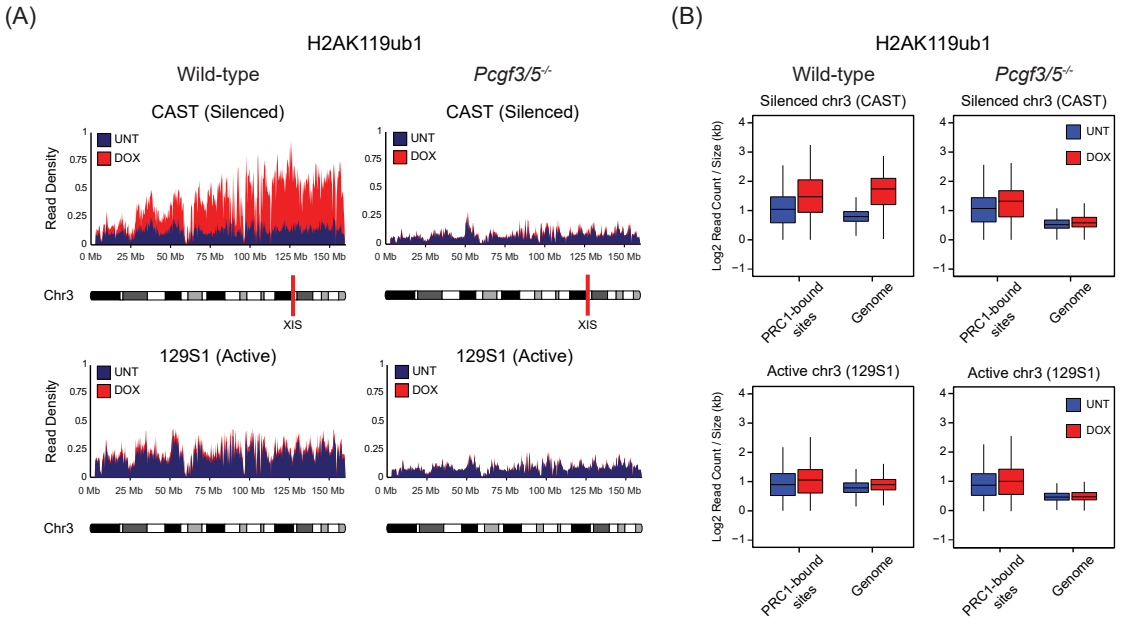



**Figure 5.6: PCGF3/5-PRC1 are required for chromosome-wide accumulation of H2AK119ub1 during Xist-mediated chromosome silencing**

(A) Chromosome density plots of H2AK119ub1 cChIP-seq across two alleles of chromosome 3 in wild-type (left) and *Pcgf3/5*$^{-/-}$ (right) *Mus domesticus* (129S1) × *Mus castaneus* F1 hybrid ES cells described in Figure 5.4A before (UNT) and after DOX treatment. XIS indicates position of the Xist integration site. The data from Nesterova et al. 2019 was used for this analysis.

(B) Allele-specific box plot analysis of normalised H2AK119ub1 cChIP-seq signal at PRC1-bound sites and 100 kb genomic windows for the silenced (*Mus castaneus*) or active (*Mus domesticus*) allele of chromosome 3 in wild-type (left) and *Pcgf3/5*$^{-/-}$ (right) *Mus domesticus* (129S1) × *Mus castaneus* F1 hybrid ES cells described in Figure 5.4A before (UNT) and after DOX treatment. The data from Nesterova et al. 2019 was used for this analysis.

### 5.2.4 BAP1 deubiquitinase complex is the mammalian PR-DUB that keeps in check genome-wide H2AK119ub1

Although H2AK119ub1 is placed by PCGF3/5-PRC1 pervasively throughout the genome, its levels outside of Polycomb target sites are low and do not appear to impinge on transcription of genes. Keeping the genome-wide H2AK119ub1 pool at this low level may play an important role in global regulation of gene expression by limiting non-target repressive activities of the Polycomb system. Consistent with this idea, multiple H2AK119ub1-specific DUBs have been characterised in mammals, including BAP1 deubiquitinase complex, which is an orthologue of the evolutionary conserved PR-DUB complex (Joo et al. 2007; Zhu et al. 2007; Nakagawa et al. 2008; Scheuermann et al. 2010; Hauri et al. 2016). Previous studies have reported that exogenous overexpression of BAP1 leads to a global depletion of H2AK119ub1 as measured by western blot analysis, suggesting that BAP1 may be responsible for keeping in check the genome-wide pool of H2AK119ub1 (Balasubramani et al. 2015). Therefore, I decided to test whether removal of BAP1 would cause a widespread increase in H2AK119ub1 levels throughout the genome and unmask

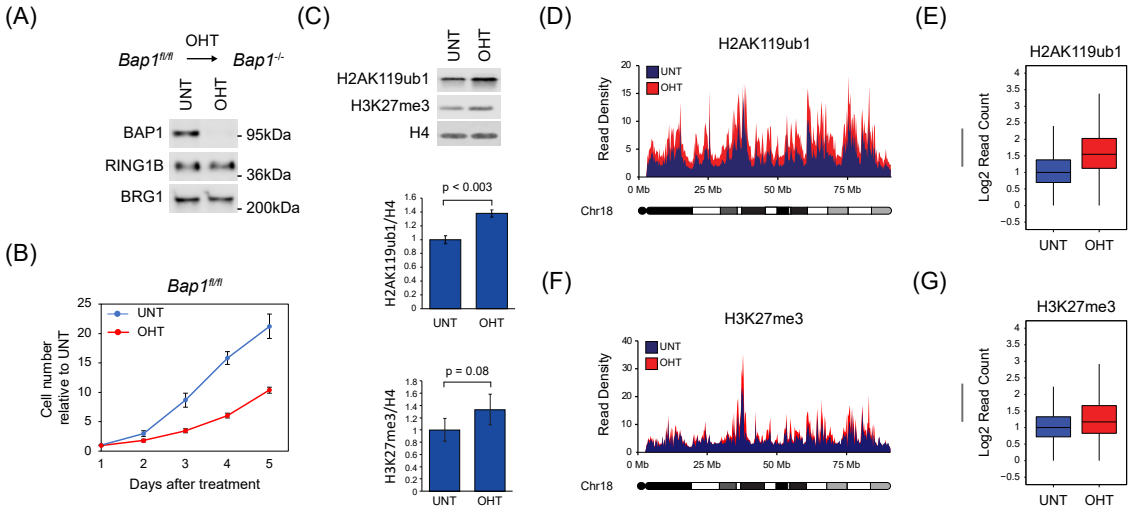



**Figure 5.7: BAP1 is responsible for removal of H2AK119ub1 throughout the genome**
(A) Western blot analysis of BAP1 and RING1B in *Bap1*$^{fl/fl}$ cells before (UNT) and after tamoxifen treatment (OHT). BRG1 is shown as a loading control.
(B) Growth curve for untreated (UNT) and OHT-treated *Bap1*$^{fl/fl}$ cells cultured for up to 5 days following initial OHT treatment for 96 hrs. Cell number is shown relative to the number of untreated cells on day 1 (day 0 corresponds to 96 hr after beginning OHT treatment). Error bars represent SEM (n=5).
(C) Western blot analysis of H2AK119ub1 and H3K27me3 in *Bap1*$^{fl/fl}$ cells (untreated and OHT-treated). Histone H4 is shown as a loading control. Quantification of H2AK119ub1 and H3K27me3 levels relative to histone H4 is shown below. Error bars represent SEM (n=3). P-value denotes statistical significance calculated by a paired Student's t-test.
(D) A chromosome density plot showing H2AK119ub1 cChIP-seq across chromosome 18 in *Bap1*$^{fl/fl}$ cells (untreated and OHT-treated).
(E) Box plot analysis of the normalised H2AK119ub1 cChIP-seq signal at 100 kb genomic windows (—) in *Bap1*$^{fl/fl}$ cells (untreated and OHT-treated).
(F) A chromosome density plot showing H3K27me3 cChIP-seq across chromosome 18 in *Bap1*$^{fl/fl}$ cells (untreated and OHT-treated).
(G) Box plot analysis of the normalised H3K27me3 cChIP-seq signal at 100 kb genomic windows (—) in *Bap1*$^{fl/fl}$ cells (untreated and OHT-treated).

a potential contribution of this histone modification to gene regulation. To this end, I generated a conditional knockout *Bap1*$^{fl/fl}$ mouse ES cell line (Figure 5.7A) using the same targeting and screening strategy as has been described previously (Chapter 3). In agreement with observations in other cell types (Machida et al. 2009; Pan et al. 2015), proliferation of BAP1-deficient ES cells declined drastically following prolonged culture (Figure 5.7B). Importantly, to minimise any secondary effects, in my experiments, I chose an earlier OHT treatment time point, at which BAP1 protein was no longer detected by western blot (Figure 5.7A), but proliferation of OHT-treated cells was not yet grossly perturbed (Figure 5.7B, 96 hr OHT treatment, day 0).

First, I asked whether BAP1 was required to control the genome-wide blanket of H2AK119ub1. Following removal of BAP1, I observed a pronounced increase in the global levels of H2AK119ub1 by western blot (Figure 5.7C), which was also evident and appeared to be largely uniform throughout the genome by cChIP-seq (Figures 5.7D-E). I then examined how Polycomb chromatin domains were affected following this global increase in H2AK119ub1. Similarly to

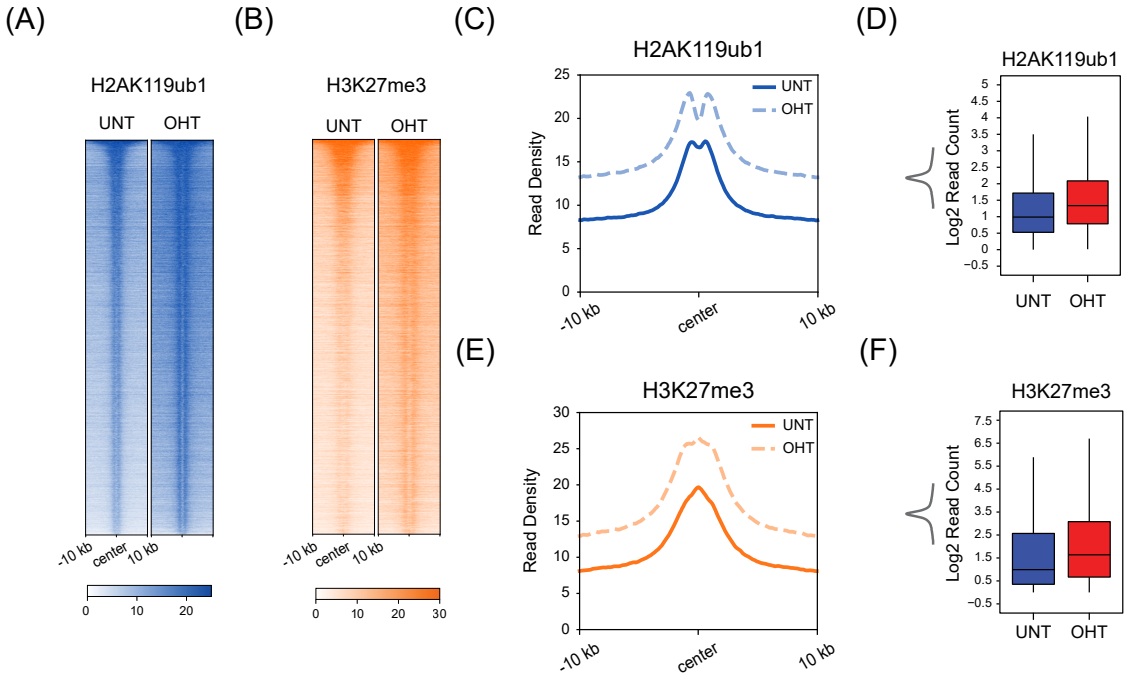



**Figure 5.8: H2AK119ub1 and H3K27me3 levels at Polycomb target sites increase following removal of BAP1**
(A) Heatmap analysis of H2AK119ub1 cChIP-seq at PRC1-bound sites (n = 18,643) in *Bap1*$^{fl/fl}$ cells (untreated and OHT-treated). Genomic intervals are sorted based on RING1B levels in wild-type ES cells.
(B) Heatmap analysis of H3K27me3 cChIP-seq at PRC1-bound sites (n = 18,643) in *Bap1*$^{fl/fl}$ cells (untreated and OHT-treated). Genomic intervals are sorted based on RING1B levels in wild-type ES cells.
(C) Metaplot analysis of H2AK119ub1 cChIP-seq at PRC1-bound sites (n = 18,643) in *Bap1*$^{fl/fl}$ cells (untreated and OHT-treated).
(D) Box plots of the normalised H2AK119ub1 cChIP-seq signal at PRC1-bound sites (⋀) in *Bap1*$^{fl/fl}$ cells (untreated and OHT-treated).
(E) Metaplot analysis of H3K27me3 cChIP-seq at PRC1-bound sites (n = 18,643) in *Bap1*$^{fl/fl}$ cells (untreated and OHT-treated).
(F) Box plots of the normalised H3K27me3 cChIP-seq signal at PRC1-bound sites (⋀) in *Bap1*$^{fl/fl}$ cells (untreated and OHT-treated).

the rest of the genome, H2AK119ub1 was also elevated around PRC1-bound sites (Figures 5.8A, C and D), albeit on average to a smaller extent than genome-wide (compare Figure 5.7E and Figure 5.8D). Remarkably, this was accompanied by a concomitant increase in H3K27me3 around these regions (Figures 5.8B, E and F), in agreement with the previous findings from our lab that H2AK119ub1 facilitates PRC2 binding/activity at Polycomb target sites (Blackledge et al. 2014; Rose et al. 2016). Interestingly, H3K27me3 was also increased broadly throughout the genome, although this effect was more modest than for H2AK119ub1 (Figures 5.7F-G). Together, these observations demonstrate a conserved role for BAP1 in regulation of H2AK119ub1 levels genome-wide and provide additional support for H2AK119ub1-dependent communication between PRC1 and PRC2.

### 5.2.5 Genome-wide accumulation of H2AK119ub1 following removal of BAP1 is associated with widespread transcriptional changes

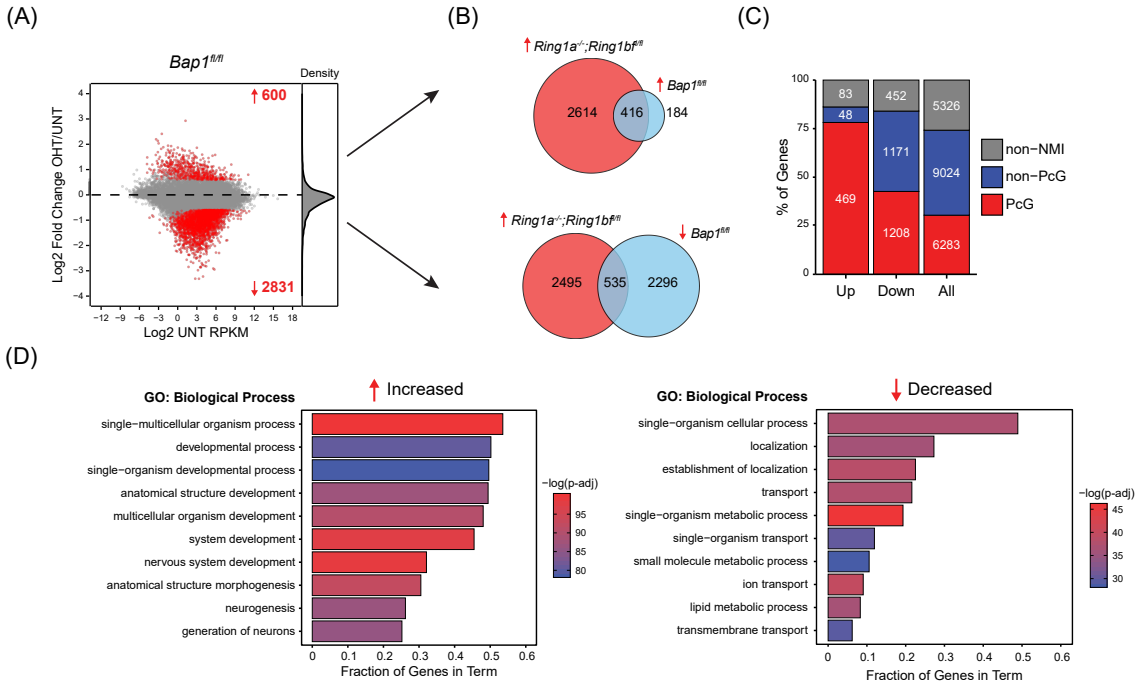



**Figure 5.9: Genome-wide accumulation of H2AK119ub1 following BAP1 removal is associated with widespread transcriptional changes**
(A) An MA-plot of log2 fold changes in gene expression (cnRNA-seq) in *Bap1*$^{fl/fl}$ cells following OHT treatment. Significant gene expression changes (p-adj<0.05 and FC>1.5-fold) are shown in red. Density of gene expression changes is shown on the right.
(B) Venn diagrams showing an overlap of significantly derepressed genes (p-adj<0.05 and FC>1.5-fold) in *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ cells and genes with significantly increased (*top*) or decreased (*bottom*) expression (p-adj<0.05 and FC>1.5-fold) in *Bap1*$^{fl/fl}$ cells following OHT treatment.
(C) A bar plot illustrating the distribution of gene expression changes (p-adj<0.05 and FC>1.5-fold) in cnRNA-seq between the gene classes (Non-NMI, Non-PcG, PcG) described in Figure 3.4B for *Bap1*$^{fl/fl}$ cells following OHT treatment.
(D) A gene ontology analysis of Biological Process term enrichment for genes showing a significant increase (*left*) or decrease (*right*) in expression (p-adj<0.05 and FC>1.5-fold) in *Bap1*$^{fl/fl}$ cells following OHT treatment.

Having established that loss of BAP1 leads to accumulation of H2AK119ub1 and H3K27me3 at Polycomb target sites and more broadly throughout the genome, I asked what effect this had on gene expression. To address this question, I performed cnRNA-seq in *Bap1*$^{fl/fl}$ cells before and after OHT treatment. This revealed that removal of BAP1 results in widespread transcriptional silencing, with approximately 6500 genes showing at least 20% reduction in expression, out of which expression of 2831 genes was reduced significantly (Figure 5.9A). Nevertheless, despite the overall negative direction of transcriptional changes following BAP1 loss, there were also 600 genes which displayed a significant increase in expression (Figure 5.9A). Unexpectedly, when I looked at which type of genes were affected by removal of BAP1, I found that the upregulated genes were enriched with Polycomb target genes (Figure 5.9C), suggesting that

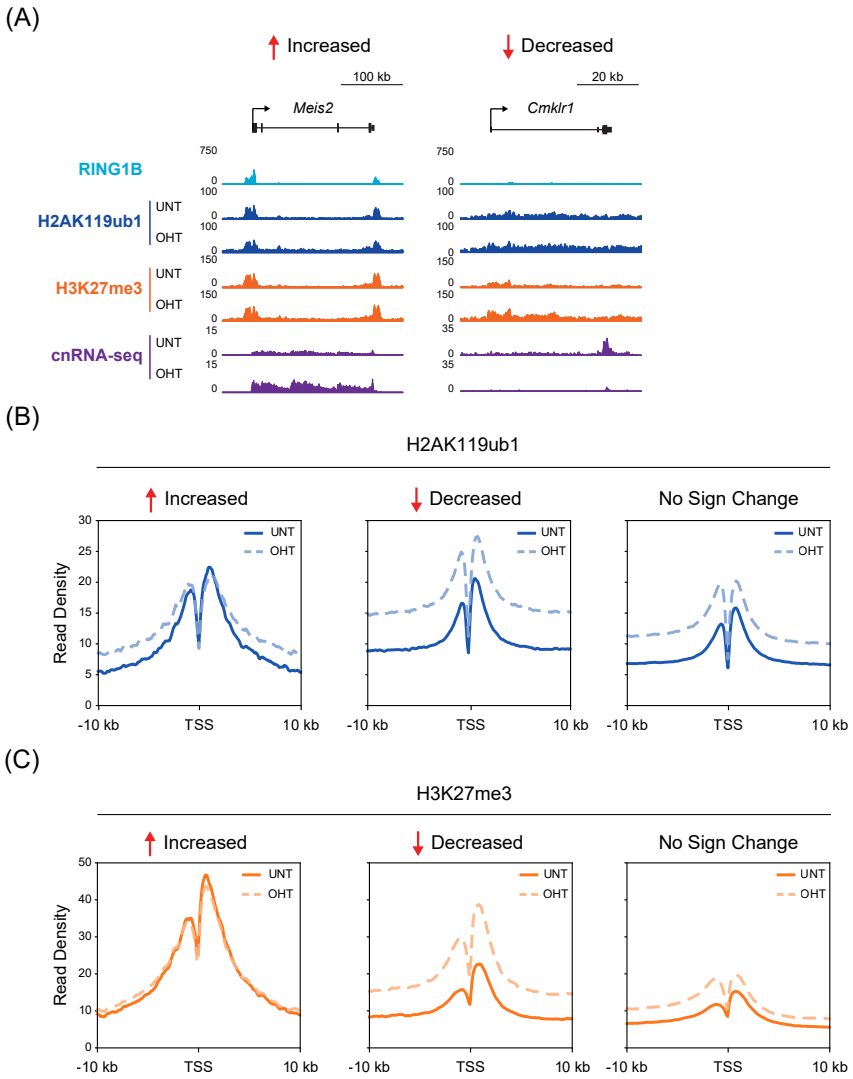



**Figure 5.10: Redistribution of H2AK119ub1 and H3K27me3 following BAP1 removal compromises Polycomb-mediated gene repression**

(A) Genomic snapshots showing H2AK119ub1 and H3K27me3 cChIP-seq, as well as cnRNA-seq in *Bap1*$^{fl/fl}$ cells (untreated and OHT-treated) for genes which display a significant increase (*left*) or decrease (*right*) in expression following OHT treatment. Also shown is RING1B cChIP-seq in wild-type ES cells.

(B) Metaplot analysis of H2AK119ub1 cChIP-seq in *Bap1*$^{fl/fl}$ cells (untreated and OHT-treated) at the promoters of genes which display a significant increase, decrease or no change in expression (p-adj<0.05 and FC>1.5-fold) following OHT treatment.

(C) Metaplot analysis of H3K27me3 cChIP-seq in *Bap1*$^{fl/fl}$ cells (untreated and OHT-treated) at the promoters of genes which display a significant increase, decrease or no change in expression (p-adj<0.05 and FC>1.5-fold) following OHT treatment.

BAP1 contributes to Polycomb-mediated gene repression. This surprising observation was further supported by the fact that many of the genes upregulated in cells lacking BAP1 were implicated in regulation of developmental processes (Figure 5.9D) and almost 70% of these genes were also derepressed following removal of PRC1 (Figure 5.9B). In contrast to this, Polycomb target genes were not over-represented among the genes which showed reduced expression in BAP1-deficient cells, in line with the majority of these genes not being occupied (Figure 5.9C) or repressed

by Polycomb complexes in ES cells (Figure 5.9B and D). Although somewhat counterintuitive, these observations are consistent with complex phenotypes of mutations in *Drosophila* PR-DUB subunits, which cause both posterior-to-anterior (characteristic of mutations in TrxG genes) and anterior-to-posterior (characteristic of mutations in PcG genes) homeotic transformations in flies (Jürgens 1985; Sinclair et al. 1998; Milne et al. 1999; Ayala Alonso et al. 2007).

Intrigued by the fact that removal of BAP1 had an opposite effect on expression of Polycomb target and non-Polycomb target genes, I decided to examine more closely changes in Polycomb-associated histone modifications at the promoters of genes which were significantly misregulated in BAP1-deficient cells. Consistent with a general increase in H2AK119ub1 and H3K27me3 throughout the genome, the levels of these histone modifications were also elevated at the promoters of downregulated genes in cells lacking BAP1 (Figure 5.10A-C). Surprisingly, in contrast to this, H2AK119ub1 and H3K27me3 levels were largely unaffected at the promoters of upregulated genes following BAP1 removal (Figure 5.10A-C). Furthermore, comparing the levels of H2AK119ub1 and H3K27me3 across different groups of genes revealed that in untreated cells their levels were generally higher at the promoters of upregulated genes, in line with these genes being enriched with Polycomb targets (Figure 5.10B-C). However, H2AK119ub1 and H3K27me3 levels became comparable between up- and down-regulated gene promoters following removal of BAP1, revealing redistribution of Polycomb complexes in the genome (Figure 5.10B-C). Based on these observations, I propose that BAP1 is required to counteract pervasive activities of Polycomb repressive complexes throughout the genome to focus them at Polycomb target sites. In the absence of BAP1, PRC1 and PRC2 are titrated away from their target gene promoters, which leads to a partial loss of Polycomb-mediated gene repression. Finally, increased genome-wide deposition of H2AK119ub1 and H3K27me3 outside of Polycomb target sites is associated with global transcriptional silencing in BAP1-deficient cells, suggesting that pervasive H2AK119ub1 needs to be limited to levels at which it does not interfere with gene expression.

## 5.3 Discussion

It is now becoming widely recognised that activities of Polycomb repressive complexes extend beyond their well-known function in silencing of developmental genes. In line with this idea, in Chapter 4, I demonstrated that PRC1 places H2AK119ub1 pervasively throughout the genome. However, which PRC1 complexes contribute to the genome-wide deposition of H2AK119ub1

and whether this plays any role in gene regulation by the Polycomb system has been unclear. In Chapter 5, I discovered that PCGF3/5-PRC1 complexes are mainly responsible for uniform deposition of low-level H2AK119ub1 throughout the genome, which is, however, not required for Polycomb chromatin domain formation and gene repression. I then investigated whether increasing global H2AK119ub1 levels could lead to widespread transcriptional silencing. I first addressed this question in the context of Xist-mediated chromosome inactivation, in which PCGF3/5-PRC1 have been previously implicated (Almeida et al. 2017; Pintacuda et al. 2017). I found that H2AK119ub1 is not placed specifically at punctate Polycomb chromatin domains associated with gene promoters but instead accumulates uniformly throughout the Xist-expressing chromosome, and this is closely linked to gene repression. Finally, I developed a system to conditionally remove BAP1 deubiquitinase complex, which is the mammalian orthologue of PR-DUB, and demonstrated that function of this complex in controlling H2AK119ub1 genomic levels is conserved in mammals. Using this system, I found that an increase in genome-wide H2AK119ub1 following BAP1 removal was associated with both widespread silencing of non-Polycomb target genes and reactivation of some Polycomb target genes, likely due to redistribution of Polycomb repressive activities in the genome. In conclusion, these observations revealed that pervasive H2AK119ub1 has to be tightly controlled, as an inappropriate increase in its levels leads to widespread transcriptional defects. However, in specialised contexts, broad deposition of H2AK119ub1 by PCGF3/5-PRC1 can be utilised as an efficient way to achieve large-scale transcriptional silencing.

The role of mammalian PCGF3/5-PRC1 in deposition of genome-wide H2AK119ub1 is conserved in evolution, as the *Drosophila* orthologue of these complexes, L(3)73Ah, has been reported to place the blanket of H2AK118ub1 outside of classical PREs (Lee et al. 2015; Kahn et al. 2016). In agreement with my observations, L(3)73Ah-dependent pervasive H2AK118ub1 was not required to support PRC2 activity or Polycomb-mediated transcriptional silencing in *Drosophila* (Lee et al. 2015). Since genome-wide activities of PRC2 have been proposed to counteract inappropriate transcription from intergenic regions (Ferrari et al. 2014; Lee et al. 2015), I was intrigued by the possibility that pervasive H2AK119ub1 could play a similar global role in gene silencing. Indeed, in *Drosophila*, PRC1 modestly contributes to repression of intergenic transcription, yet this does not rely on L(3)73Ah (Enderle et al. 2010; Lee et al. 2015). However, I found no signs of global gene activation or inappropriate expression of intergenic regions in cnRNA-seq data from

PCGF3/5- or RING1A/B-deficient cells, in line with no evidence in the literature to indicate that PRC1 mediates genome-wide transcriptional silencing in mammals. Interestingly, PRC1 has been previously proposed to repress several classes of repeat elements during mammalian development (Leeb et al. 2010; Hisada et al. 2012; He et al. 2019a). Investigating which PRC1 complexes may contribute to transcriptional control of these genomic elements and whether this relies on pervasive H2AK119ub1 is an exciting question to be explored in the future.

In contrast to other PRC1 complexes, PCGF3/5-PRC1 do not stably incorporate factors which can recognise specific DNA or chromatin features (Hauri et al. 2016; Kloet et al. 2016). Therefore, it is still unclear how PCGF3/5-PRC1 complexes engage with the genome and whether there are mechanisms for recruitment of these complexes to specific genomic loci. Interestingly, a recent study found that the USF1/2 transcription factors dynamically associate with PCGF3/5-PRC1 and recruit these complexes to a small number of target sites (Scelfo et al. 2019). In addition to transcription factors, it has been proposed that lncRNAs could also contribute to targeting of PCGF3/5-PRC1 complexes to the silenced X chromosome during X-inactivation (Almeida et al. 2017). However, recent work showed that PCGF3/5-PRC1 do not directly interact with the Xist RNA and instead require an additional RNA-binding protein hnRNPK for recruitment (Pintacuda et al. 2017). Although these findings provide some insight into the mechanisms by which PCGF3/5-PRC1 can be targeted to specific genomic sites, they do not explain how these complexes interact with the genome more broadly to place H2AK119ub1. One possibility is that there is no specialised genome sampling mechanism, and instead the affinity of PCGF3/5-PRC1 for nucleosome substrates is sufficient for their broad but transient engagement with the genome. Alternatively, the widespread activity of PCGF3/5-PRC1 complexes in the genome may be facilitated by their RYBP subunit, which has been reported to have affinity for H2AK119ub1 *in vitro* (Arrigoni et al. 2006; Fereres et al. 2014).

Pervasive PCGF3/5-dependent H2AK119ub1 contributes little to gene regulation when kept at low levels. However, its broad distribution suggests that genome-wide H2AK119ub1 may still contribute to regulation of large-scale chromosomal processes that are distinct from gene regulation. In fact, H2AK119ub1 has been previously implicated in control of DNA replication, DNA repair and cell cycle progression (Ismail et al. 2010; Lee et al. 2014; Bravo et al. 2015; Uckelmann and Sixma 2017). Here, I have established two genetic systems to manipulate genomic levels of H2AK119ub1, which will enable me to further probe the role of H2AK119ub1 blanket

in regulation of different aspects of the genome function. The *Pcgf3/5*$^{fl/fl}$ system is particularly well suited for addressing these questions as conditional knockout of PCGF3/5 leads to only minor transcriptional changes. This would make it possible to disentangle the direct effects of H2AK119ub1 depletion on various chromosomal processes from the secondary effects of gene reactivation in other PRC1 perturbation systems. Intriguingly, in BAP1-deficient cells, I observed a severe reduction in cell proliferation, suggesting that a failure to efficiently remove H2AK119ub1 may interfere with cell cycle progression. However, there are many potential explanations for why cell growth might be affected and all of them should be carefully considered. In fact, a recent study has suggested that the most likely cause of the cell growth defects in BAP1-deficient cells is increased apoptosis rather than problems with cell cycle (He et al. 2019b).

Although controlled accumulation of H2AK119ub1 can be utilised for gene silencing in specialised contexts, such as X-inactivation, a global increase in H2AK119ub1 throughout the genome could have deleterious effects on transcription of many genes. In line with this idea, elevated levels of H2AK119ub1 following removal of BAP1 were associated with widespread gene repression. Similar large-scale transcriptional inhibition has been recently reported in BAP1-null human HAP1 cells (Campagne et al. 2019), supporting my conclusion that BAP1 plays an essential role in promoting gene expression throughout the genome. However, it still remains unclear whether BAP1 functions through counteracting the Polycomb repressive system and whether this relies directly on its H2AK119ub1 DUB activity. This is further complicated by the fact that BAP1 assembles into a large multi-subunit complex which incorporates several factors with known functions in transcriptional regulation (Kloet et al. 2016; Hauri et al. 2016). Therefore, future work should focus on investigating whether transcriptional changes following removal of BAP1 are directly caused by the changes in Polycomb-associated histone modifications. If gene silencing in BAP1-deficient cells requires activities of Polycomb repressive complexes, then removing PRC1/PRC2, in addition to BAP1, should result in an epistatic phenotype that rescues BAP1-dependent gene expression changes. Remarkably, using this approach, two recent studies have revealed that PRC1 and specifically PRC1 catalytic activity was responsible for transcriptional defects in BAP1-deficient cells, while the contribution of PRC2 to gene silencing was minimal (Campagne et al. 2019; He et al. 2019b). This strongly suggests that BAP1 counteracts PRC1, presumably through H2AK119ub1 deubiquitination, to focus Polycomb repressive activities at target sites and protect the rest of the genome from pervasive silencing by the Polycomb system.

Finally, additional support for the antagonism between BAP1 and the Polycomb system comes from the fact that some of the BAP1-deficient cancers are sensitive to inhibition of Polycomb complexes (LaFave et al. 2015; Wang et al. 2018a). Therefore, understanding the function of BAP1 in the context of the Polycomb system and whether loss of BAP1 DUB activity contributes to malignant transformations might provide an insight into the cause of BAP1-associated cancers and inform potential therapeutic strategies.

In Chapters 4 and 5, I discovered two distinct modes of H2AK119ub1 deposition in the genome, which are controlled independently by distinct PRC1 complexes. PCGF1-PRC1 is responsible for shaping punctate enrichment of H2AK119ub1 at Polycomb target sites and PCGF3/5-PRC1 place low-level H2AK119ub1 uniformly throughout the genome. While activity of PCGF3/5-PRC1 accounts for the bulk of genomic H2AK119ub1, it is largely dispensable for Polycomb-mediated gene repression. In contrast, PCGF1-PRC1 places only a small fraction of total H2AK119ub1, but plays an important role in gene repression by the Polycomb system. However, removal of either PCGF1 or PCGF3/5 did not fully compromise Polycomb target gene silencing, indicating that PRC1 complexes must cooperate to support gene repression. Despite incorporating different modules with unique biochemical activities, both PCGF1-PRC1 and PCGF3/5-PRC1 independently contribute to shaping H2AK119ub1 at Polycomb target sites, suggesting that the shared catalytic activity of these complexes may underpin their potential cooperation which is required to fully repress Polycomb target genes. To test this hypothesis, in the next chapters I will determine which PRC1 complexes have to be removed together to phenocopy a complete loss of PRC1 (Chapter 6) and then directly test the role of PRC1 catalytic activity in Polycomb-mediated gene repression (Chapter 7).

# 6

# SYNERGY BETWEEN VARIANT PRC1 COMPLEXES DRIVES POLYCOMB-MEDIATED GENE REPRESSION

## 6.1 Introduction

Functional interrogation of PRC1 system in Chapters 3-5 revealed that removal of individual PCGF proteins did not recapitulate the transcriptional defects caused by a complete loss of PRC1, and neither did combined removal of highly similar PRC1 complexes. This suggested that formation of Polycomb chromatin domains and Polycomb target gene repression rely on cooperation between distinct PRC1 complexes. However, given that PCGF proteins contain unique evolutionary conserved domains and occupy mutually exclusive regions in the genome, individual PRC1 complexes have been proposed to have non-overlapping functions in regulation of Polycomb target genes (Sowpati et al. 2015; Gao et al. 2012). In line with this idea, PCGF6-PRC1 has been shown to play a unique role in preventing premature expression of germ cell-related genes (Endoh et al. 2017; Zhao et al. 2017b), while PCGF5 has been implicated in activation of neuronal genes (Gao et al. 2014; Yao et al. 2018; Wang et al. 2018b). Although these observations do not completely rule out the possibility of functional cooperation between PRC1 complexes, the current consensus in the field is that individual PRC1 complexes regulate genes independently via unique mechanisms.

In Chapter 3, I discovered that canonical PRC1 contributes little to deposition of H2AK119ub1 and repression of Polycomb target genes in mouse ES cells, suggesting that variant PRC1 complexes must be sufficient to support PRC1-mediated transcriptional silencing. Although the genomic distribution of variant PRC1 subunits has been poorly characterised in these cells, I have demonstrated that PCGF1- and PCGF3/5-PRC1 both contribute to shaping H2AK119ub1 at Polycomb target sites (Chapters 4 and 5), indicating that these complexes function at the same genomic loci. Furthermore, despite their highly diverse subunit composition, all variant PRC1 complexes appear to be more active at catalysing H2AK119ub1 than canonical PRC1,

and can be further stimulated by variant PRC1-specific RYBP/YAF2 subunits (Gao et al. 2012; Taherbhoy et al. 2015; Rose et al. 2016). Together, these observations support the possibility that variant PRC1 complexes may cooperate to deposit appropriate H2AK119ub1 levels at their target gene promoters and drive gene repression. Furthermore, very little is known about functional interactions between variant and canonical PRC1, and whether these complexes can compensate for each other. Although both PRC1 classes co-occupy many shared genomic loci (Morey et al. 2013; Stielow et al. 2018), they have been proposed to function independently and contribute to distinct modes of transcriptional silencing by the Polycomb system (Morey et al. 2013; Aranda et al. 2015). However, whether these modes of repression are functionally linked and whether they can reinforce each other has remained poorly understood.

In Chapter 6, I utilise a series of combinatorial genetic perturbations to determine whether Polycomb-mediated gene repression relies on cooperation between distinct PRC1 complexes. First, I investigate whether variant PRC1 complexes function synergistically to deposit H2AK119ub1 and silence Polycomb target genes. Next, I ask whether canonical PRC1 complexes can compensate for compromised activity of variant PRC1 complexes in gene regulation. Ultimately, through identifying the PCGF proteins which are sufficient to account for PRC1 repressive activity in mouse ES cells, I aim to uncover the central determinants of Polycomb-mediated gene repression.

## 6.2 Results

### 6.2.1 Removal of PCGF1/3/5-PRC1 does not completely abolish Polycomb chromatin domain formation and gene repression

PCGF1-PRC1 and PCGF3/5-PRC1 independently contribute to deposition of H2AK119ub1 at Polycomb target sites, but removing these complexes individually did not explain the full extent of Polycomb-mediated gene repression (Chapters 4 and 5). Therefore, I reasoned that PCGF1/3/5-PRC1 complexes may function synergistically to support gene repression. To test this possibility, I characterised a triple conditional knockout *Pcgf1/3/5*$^{fl/fl}$ ES cell line, in which simultaneous removal of PCGF1/3/5 can be induced (Figure 6.1A). Strikingly, loss of PCGF1/3/5-PRC1 resulted in a dramatic reduction in total H2AK119ub1 levels as evident by western blot (Figure 6.1B). In line with this, cChIP-seq analysis revealed that H2AK119ub1 was majorly depleted throughout

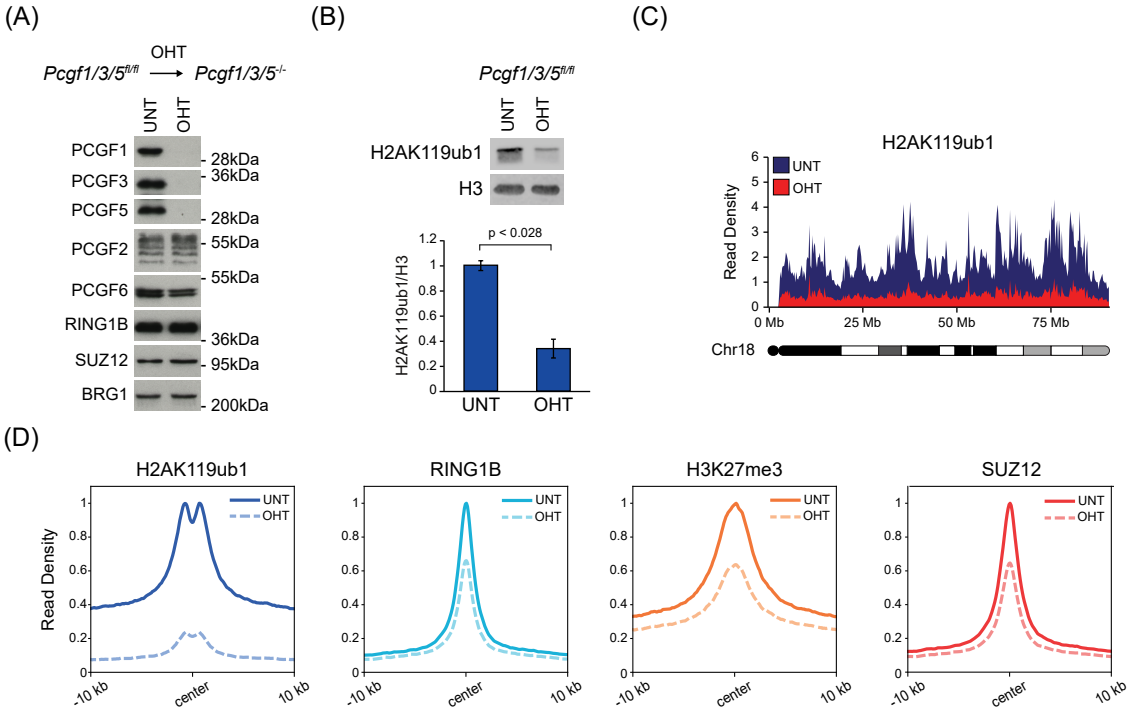



**Figure 6.1: Polycomb chromatin domains are compromised but not completely eroded following loss of PCGF1/3/5-PRC1**

(A) Western blot analysis of PCGF proteins, RING1B and SUZ12 in *Pcgf1/3/5*$^{fl/fl}$ cells before (UNT) and after tamoxifen treatment (OHT). BRG1 is shown as a loading control.

(B) Western blot analysis of H2AK119ub1 in *Pcgf1/3/5*$^{fl/fl}$ cells (untreated and OHT-treated). Histone H3 is shown as a loading control. Also shown is quantification of H2AK119ub1 levels relative to histone H3. Error bars represent SEM (n=3). P-value denotes statistical significance calculated by a paired Student's t-test.

(C) A chromosome density plot showing H2AK119ub1 cChIP-seq across chromosome 18 in *Pcgf1/3/5*$^{fl/fl}$ cells (untreated and OHT-treated).

(D) Metaplot analysis of cChIP-seq for H2AK119ub1, RING1B, H3K27me3 and SUZ12 at PRC1-bound sites (n = 8833) in *Pcgf1/3/5*$^{fl/fl}$ cells (untreated and OHT-treated).

the genome (Figure 6.1C), and levels of this histone modification at Polycomb target sites were also dramatically decreased following OHT treatment (Figures 6.1D and 6.2A). In addition to a massive reduction in H2AK119ub1, binding of RING1B and SUZ12, as well as H3K27me3, was also compromised at Polycomb target sites, indicating a major but incomplete erosion of Polycomb chromatin domains following removal of PCGF1/3/5 (Figures 6.1D and 6.2A). Importantly, in contrast to the modest transcriptional changes in PCGF1- and especially PCGF3/5-deficient cells, removal of PCGF1/3/5 in combination resulted in a dramatic derepression of a large number of Polycomb target genes (Figures 6.2A-C). However, the widespread transcriptional reactivation caused by removal of PCGF1/3/5-PRC1 still did not completely recapitulate the extent of Polycomb target gene derepression observed in RING1A/B-deficient cells (Figure 6.2D). Nevertheless, combined loss of PCGF1/3/5-PRC1 complexes had a greater-than-additive effect on deposition of H2AK119ub1, Polycomb chromatin domain formation and gene repression compared to removal of PCGF1-PRC1 and PCGF3/5-PRC1 individually. This indicates that

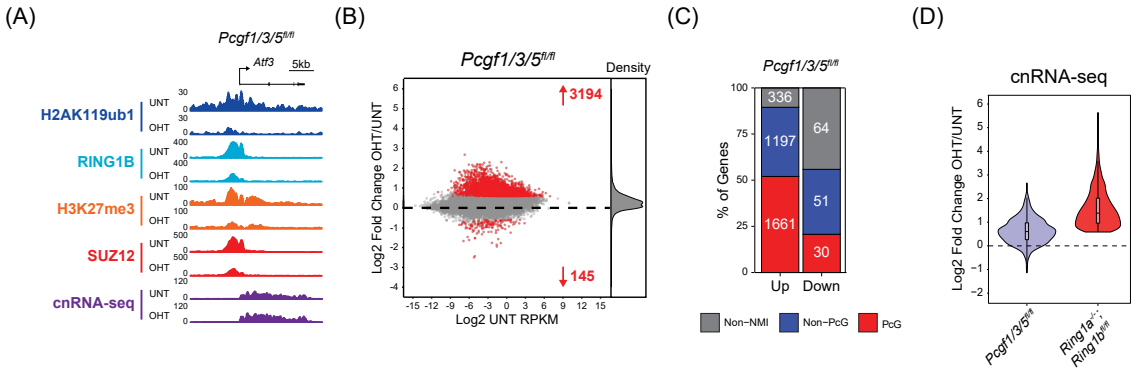



**Figure 6.2: Loss of PCGF1/3/5-PRC1 leads to a widespread but not complete derepression of Polycomb target genes**
(A) Genomic snapshot of a typical Polycomb target gene, showing cnRNA-seq and cChIP-seq for PRC1 (H2AK119ub1 and RING1B) and PRC2 (H3K27me3 and SUZ12) in *Pcgf1/3/5*$^{fl/fl}$ cells (untreated and OHT-treated).
(B) An MA-plot of log2 fold changes in gene expression (cnRNA-seq) in *Pcgf1/3/5*$^{fl/fl}$ cells following OHT treatment. Significant gene expression changes (p-adj<0.05 and FC>1.5-fold) are shown in red. Density of gene expression changes is shown on the right.
(C) A bar plot illustrating the distribution of gene expression changes (p-adj<0.05 and FC>1.5-fold) in cnRNA-seq between the gene classes (Non-NMI, Non-PcG, PcG) described in Figure 3.4B for *Pcgf1/3/5*$^{fl/fl}$ cells following OHT treatment.
(D) A violin plot comparing log2 fold changes in PRC1 target gene expression (cnRNA-seq) in *Pcgf1/3/5*$^{fl/fl}$ and *Ring1a*$^{-/-}$;*Ring1b*$^{fl/fl}$ cells following OHT treatment.

cooperation between these variant PRC1 complexes plays an important role in Polycomb system function.

### 6.2.2 Canonical PRC1 does not compensate for compromised activity of variant PRC1 complexes

In the absence of PCGF1/3/5, the remaining PRC1-dependent gene repression could be achieved via the activity of PCGF2/4- and/or PCGF6-containing PRC1 complexes. Since H3K27me3 was partially retained at Polycomb target sites following PCGF1/3/5 removal, I reasoned that canonical PCGF2/4-PRC1 could still bind at these genomic regions and compensate for compromised activity of variant PRC1 complexes. To examine the contribution of PCGF2-PRC1 to silencing of Polycomb target genes in the absence of PCGF1/3/5-PRC1 complexes, I generated a quadruple conditional knockout *Pcgf1/3/5/2*$^{fl/fl}$ ES cell line (Figure 6.3A). However, removal of PCGF2 in combination with PCGF1/3/5 caused no further reduction in H2AK119ub1 at the majority of Polycomb target sites (Figures 6.3D, 6.4A-B and 6.5A) or more broadly throughout the genome (Figures 6.3B-C), despite RING1B binding being even more dramatically reduced (Figures 6.3D, 6.4A-B and 6.5A). Similarly to H2AK119ub1, reductions in SUZ12 binding and H3K27me3 were highly comparable between PCGF1/3/5- and PCGF1/3/5/2-deficient cells (Figures 6.3D, 6.4A-B, and 6.5A). In agreement with my earlier observations (Chapter 3), this demonstrates that although

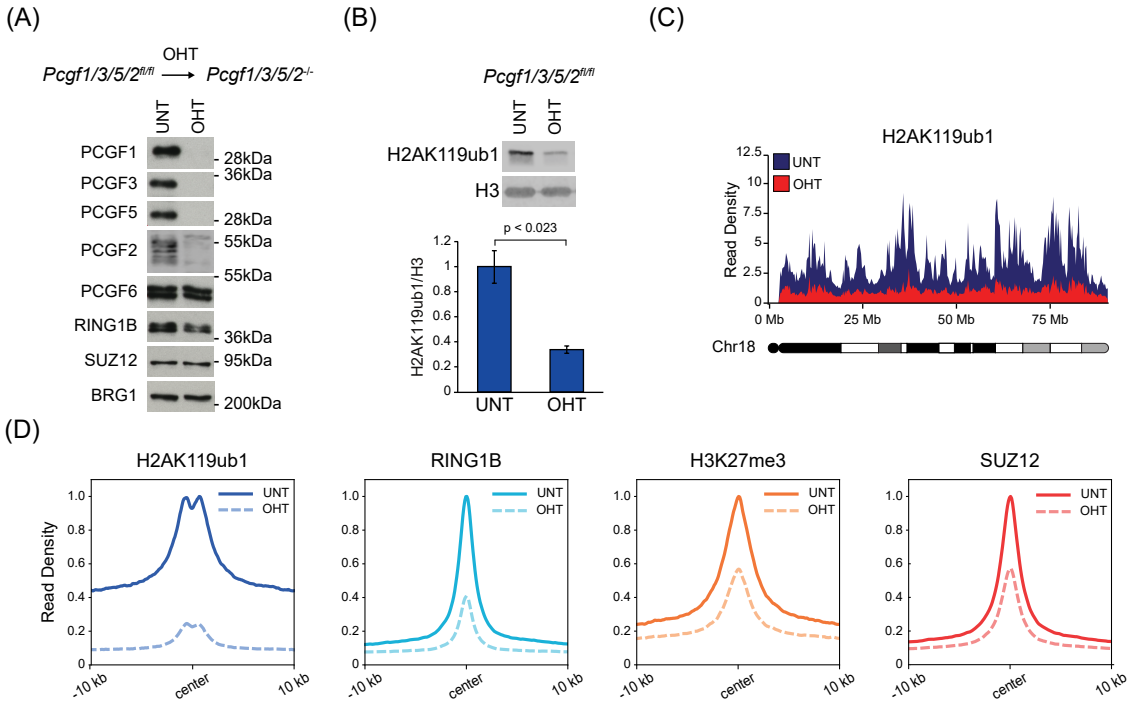



**Figure 6.3: Removal of PCGF2 in addition to PCGF1/3/5 does not further compromise Polycomb chromatin domain formation**

(A) Western blot analysis of PCGF proteins, RING1B and SUZ12 in *Pcgf1/3/5/2*$^{fl/fl}$ cells before (UNT) and after tamoxifen treatment (OHT). BRG1 is shown as a loading control.

(B) Western blot analysis of H2AK119ub1 in *Pcgf1/3/5/2*$^{fl/fl}$ cells (untreated and OHT-treated). Histone H3 is shown as a loading control. Also shown is quantification of H2AK119ub1 levels relative to histone H3. Error bars represent SEM (n=3). P-value denotes statistical significance calculated by a paired Student's t-test.

(C) A chromosome density plot showing H2AK119ub1 cChIP-seq across chromosome 18 in *Pcgf1/3/5/2*$^{fl/fl}$ cells (untreated and OHT-treated).

(D) Metaplot analysis of cChIP-seq for H2AK119ub1, RING1B, H3K27me3 and SUZ12 at PRC1-bound sites (n = 8833) in *Pcgf1/3/5/2*$^{fl/fl}$ cells (untreated and OHT-treated).

canonical PRC1 complexes play a central role in shaping RING1B occupancy, they contribute little to deposition of H2AK119ub1 and PRC1-dependent stimulation of PRC2 binding/activity at Polycomb target sites.

Strikingly, combined removal of PCGF2 and PCGF1/3/5 led to only a slight increase in the extent of Polycomb target gene derepression as compared to PCGF1/3/5-deficient cells (Figures 6.2A, 6.5A-C and 6.6A), which was highly similar between the two cell lines (Figure 6.6B) and still failed to phenocopy a complete loss of PRC1 (Figure 6.6A). This indicated that even when PRC1-mediated gene repression was severely compromised following removal of variant PCGF1/3/5-PRC1 complexes, canonical PRC1 still contributed very little to silencing of the vast majority of Polycomb target genes. Finally, to mitigate against any potential compensation from PCGF4 in these experiments, I constitutively deleted the *Pcgf4* gene in *Pcgf1/3/5/2*$^{fl/fl}$ cells. Importantly, when I examined expression of several Polycomb target genes which were not fully derepressed

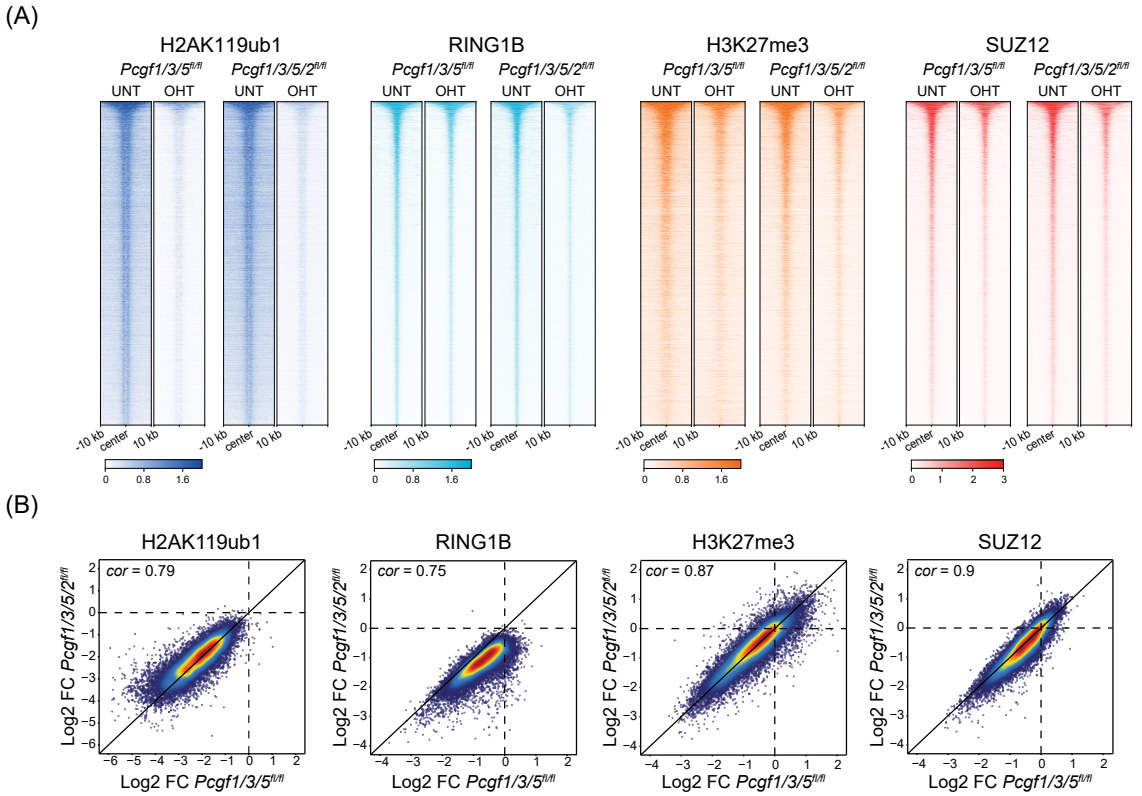



**Figure 6.4: Removal of PCGF1/3/5 or PCGF1/3/5/2 leads to a similar erosion of Polycomb chromatin domains**
(A) Heatmap analysis of H2AK119ub1, RING1B, H3K27me3 and SUZ12 cChIP-seq at PRC1-bound sites (n = 8833) in *Pcgf1/3/5*$^{fl/fl}$ (*left*) and *Pcgf1/3/5/2*$^{fl/fl}$ (*right*) cells (untreated and OHT-treated). Genomic intervals are sorted as in Figure 4.1.
(B) Scatter plot analysis comparing log2 fold changes in H2AK119ub1, RING1B, H3K27me3 and SUZ12 cChIP-seq at PRC1-bound sites between *Pcgf1/3/5*$^{fl/fl}$ and *Pcgf1/3/5/2*$^{fl/fl}$ cells following OHT treatment. *Cor* denotes a Pearson correlation coefficient.

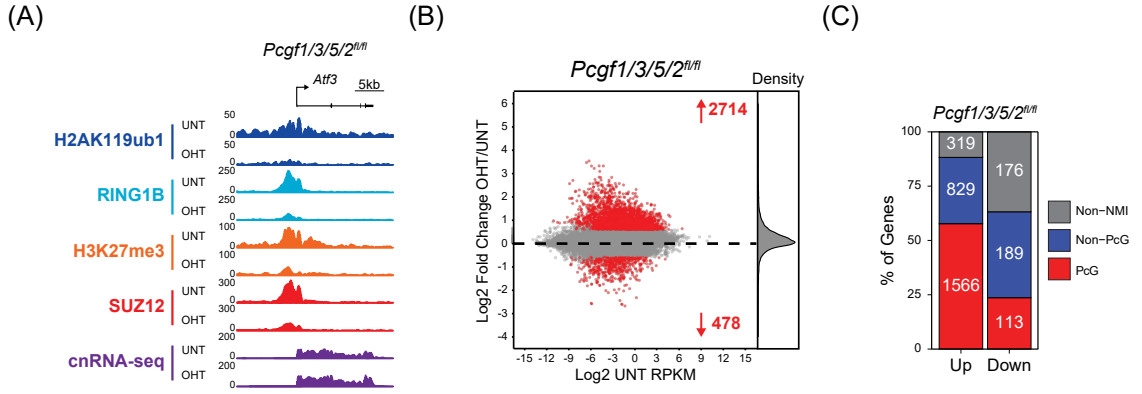



**Figure 6.5: Derepression of Polycomb target genes following loss of PCGF1/3/5 is not further potentiated by removal of PCGF2**
(A) Genomic snapshot of a typical Polycomb target gene, showing cnRNA-seq and cChIP-seq for PRC1 (H2AK119ub1 and RING1B) and PRC2 (H3K27me3 and SUZ12) in *Pcgf1/3/5/2*$^{fl/fl}$ cells (untreated and OHT-treated).
(B) An MA-plot of log2 fold changes in gene expression (cnRNA-seq) in *Pcgf1/3/5/2*$^{fl/fl}$ cells following OHT treatment. Significant gene expression changes (p-adj$<$0.05 and FC$>$1.5-fold) are shown in red. Density of gene expression changes is shown on the right.
(C) A bar plot illustrating the distribution of gene expression changes (p-adj$<$0.05 and FC$>$1.5-fold) in cnRNA-seq between the gene classes (Non-NMI, Non-PcG, PcG) described in Figure 3.4B for *Pcgf1/3/5/2*$^{fl/fl}$ cells following OHT treatment.

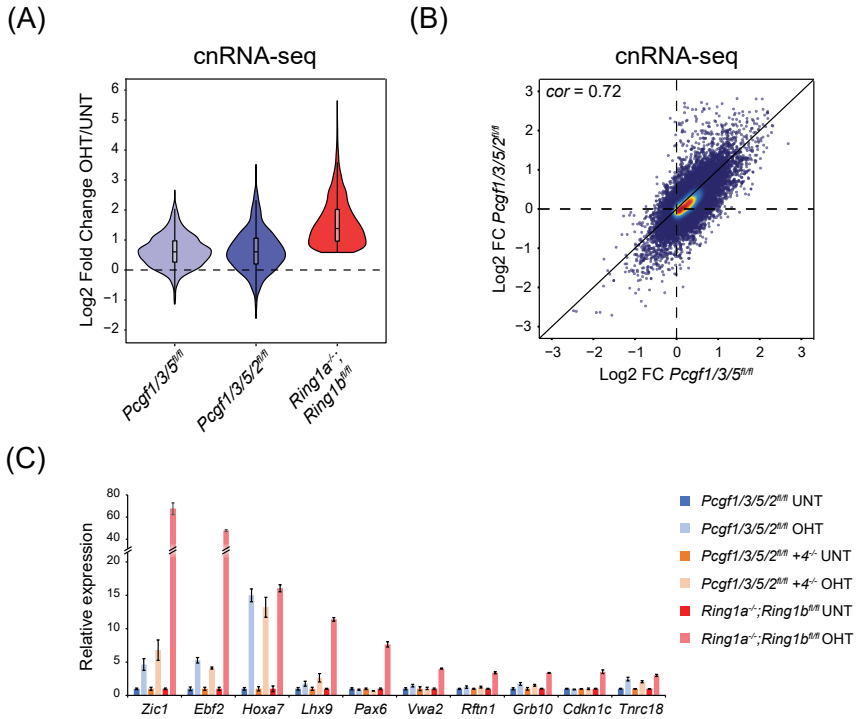



**Figure 6.6: Canonical PRC1 does not compensate for compromised Polycomb target gene repression in the absence of PCGF1/3/5-PRC1**
(A) A violin plot comparing log2 fold changes in PRC1 target gene expression (cnRNA-seq) in *Pcgf1/3/5*$^{fl/fl}$, *Pcgf1/3/5/2*$^{fl/fl}$ and *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ cells following OHT treatment.
(B) A scatter plot analysis comparing log2 fold changes in gene expression (cnRNA-seq) between *Pcgf1/3/5*$^{fl/fl}$ and *Pcgf1/3/5/2*$^{fl/fl}$ cells following OHT treatment. *Cor* denotes a Pearson correlation coefficient.
(C) RT-qPCR expression analysis for a panel of Polycomb target genes in *Pcgf1/3/5/2*$^{fl/fl}$, *Pcgf4*$^{-/-}$*;Pcgf1/3/5/2*$^{fl/fl}$ and *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ cells (untreated and OHT-treated). Gene expression is normalised to *b-actin* and is shown relative to the average expression in untreated cells for each individual cell line. Error bars show SEM (n=3).

following PCGF1/3/5/2 removal by RT-qPCR, I observed no additional reactivation of these genes following OHT treatment in *Pcgf4*$^{-/-}$*;Pcgf1/3/5/2*$^{fl/fl}$ cells (Figure 6.6C). Together, these findings further demonstrate that PCGF2/4-PRC1 complexes play a minor role in Polycomb-mediated gene repression, even when the function of variant PRC1 is substantially perturbed.

Interestingly, I identified a small group of 137 PRC1-repressed genes, enriched with developmental regulators, which were further reactivated when PCGF2 was removed in combination with PCGF1/3/5 compared to PCGF1/3/5-deficient cells (Figures 6.7A-C). Polycomb chromatin domains, which were associated with promoters of these genes, were larger and occupied by higher levels of RING1B than domains at typical PRC1-repressed gene promoters (Figure 6.7D). Strikingly, H2AK119ub1, SUZ12 and H3K27me3 were further reduced at the promoters of these genes in PCGF1/3/5/2-deficient cells as compared to PCGF1/3/5-deficient cells (Figure 6.7E). This indicates that at this specialised group of Polycomb target genes, PCGF2/4-PRC1 complexes can contribute to deposition of H2AK119ub1, PRC2 recruitment/activity and

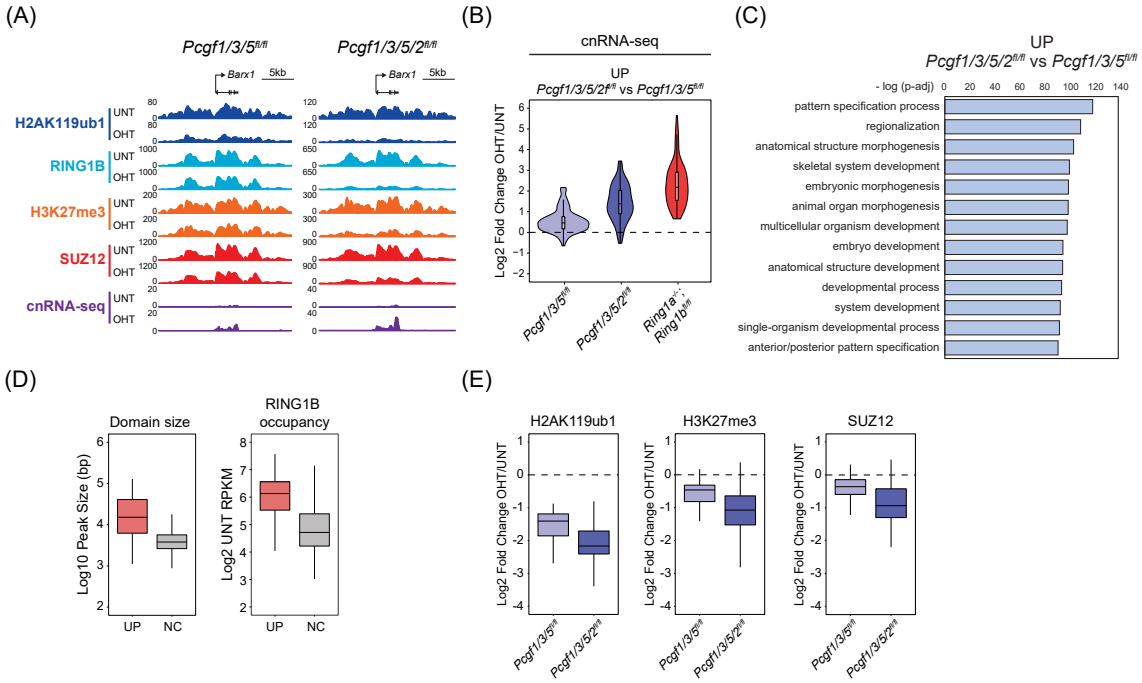



**Figure 6.7: PCGF2/4-PRC1 contribute to repression of a subset of Polycomb target genes**
(A) Genomic snapshot of a PRC1 target gene that is further derepressed when PCGF2 is removed in addition to PCGF1/3/5, showing cChIP-seq for PRC1 (H2AK119ub1 and RING1B) and PRC2 (H3K27me3 and SUZ12), as well as cnRNA-seq in *Pcgf1/3/5*$^{fl/fl}$ (*left*) and *Pcgf1/3/5/2*$^{fl/fl}$ (*right*) cells (untreated and OHT-treated).
(B) A violin plot comparing log2 fold changes in expression of PRC1 target genes that are further derepressed when PCGF2 is removed in addition to PCGF1/3/5 (n = 137) in *Pcgf1/3/5*$^{fl/fl}$, *Pcgf1/3/5/2*$^{fl/fl}$ and *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ cells following OHT treatment.
(C) A gene ontology analysis of Biological Process term enrichment for genes showing a further significant increase in expression (p-adj<0.05 and FC>1.5-fold) in *Pcgf1/3/5/2*$^{fl/fl}$ comparing to *Pcgf1/3/5*$^{fl/fl}$ cells following OHT treatment.
(D) A box plot analysis of the size (*left*) and RING1B occupancy in untreated cells (*right*) for Polycomb chromatin domains associated with promoters of PRC1 target genes showing a further increase (UP, n = 137) or no change (NC, n = 1934) in expression when PCGF2 is removed in addition to PCGF1/3/5.
(E) A box plot analysis comparing log2 fold changes in H2AK119ub1, H3K27me3 and SUZ12 cChIP-seq at Polycomb chromatin domains associated with promoters of PRC1 target genes that are further derepressed when PCGF2 is removed in addition to PCGF1/3/5 (n = 137) in *Pcgf1/3/5*$^{fl/fl}$ and *Pcgf1/3/5/2*$^{fl/fl}$ cells following OHT treatment.

transcriptional silencing, although this contribution is modest and becomes unmasked only when the repressive activity of variant PRC1 complexes is severely compromised. It is possible that high-level occupancy of canonical PCGF2/4-PRC1 at the promoters of these genes allows these complexes to partially support formation of Polycomb chromatin domains and Polycomb-mediated gene repression, despite their inherently weak E3 ubiquitin ligase activity (Gao et al. 2012; Taherbhoy et al. 2015; Rose et al. 2016). Alternatively, more catalytically active RYBP-containing PCGF2/4-PRC1 complexes (Gao et al. 2012; Rose et al. 2016) may cooperate with other variant PRC1 complexes to silence this subset of Polycomb target genes.

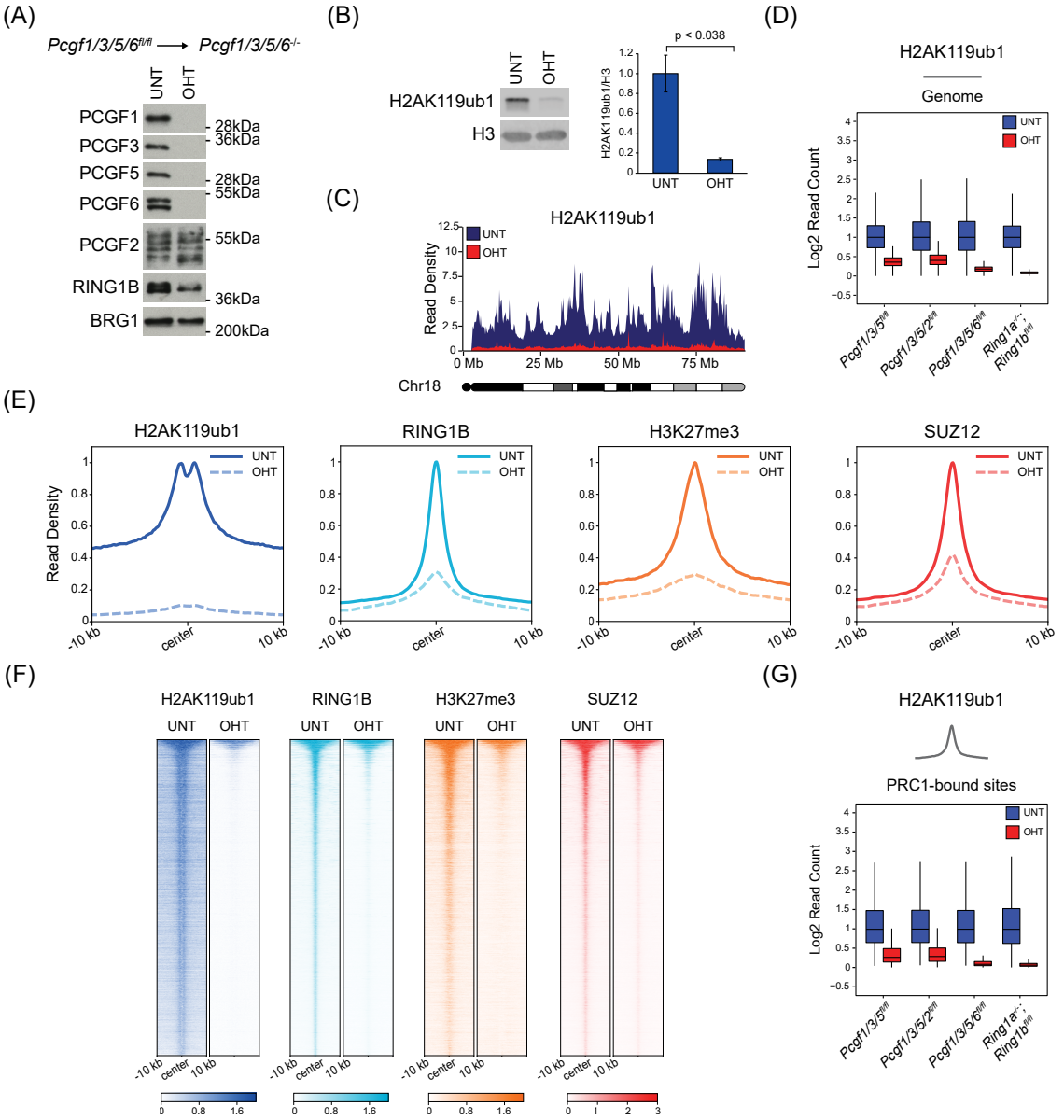



**Figure 6.8: Variant PCGF1/3/5/6-PRC1 complexes drive Polycomb chromatin domain formation**

(A) Western blot analysis of PCGF proteins and RING1B in *Pcgf1/3/5/6*$^{fl/fl}$ cells before (UNT) and after tamoxifen treatment (OHT). BRG1 is shown as a loading control.

(B) Western blot analysis of H2AK119ub1 in *Pcgf1/3/5/6*$^{fl/fl}$ cells (untreated and OHT-treated). Histone H3 is shown as a loading control. Also shown is quantification of H2AK119ub1 levels relative to histone H3. Error bars represent SEM (n=3). P-value denotes statistical significance calculated by a paired Student's t-test.

(C) A chromosome density plot showing H2AK119ub1 cChIP-seq across chromosome 18 in *Pcgf1/3/5/6*$^{fl/fl}$ cells (untreated and OHT-treated).

(D) A box plot analysis of the normalised H2AK119ub1 cChIP-seq signal at 100 kb genomic windows (—) in *Pcgf1/3/5*$^{fl/fl}$, *Pcgf1/3/5/2*$^{fl/fl}$, *Pcgf1/3/5/6*$^{fl/fl}$ and *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ cells (untreated and OHT-treated).

(E) Metaplot analysis of cChIP-seq for H2AK119ub1, RING1B, H3K27me3 and SUZ12 at PRC1-bound sites (n = 8833) in *Pcgf1/3/5/6*$^{fl/fl}$ cells (untreated and OHT-treated).

(F) Heatmap analysis of H2AK119ub1, RING1B, H3K27me3 and SUZ12 cChIP-seq at PRC1-bound sites (n = 8833) in *Pcgf1/3/5/6*$^{fl/fl}$ cells (untreated and OHT-treated). Genomic intervals are sorted as in Figure 4.1.

(G) A box plot analysis of the normalised H2AK119ub1 cChIP-seq signal at PRC1-bound sites (⋀) in *Pcgf1/3/5*$^{fl/fl}$, *Pcgf1/3/5/2*$^{fl/fl}$, *Pcgf1/3/5/6*$^{fl/fl}$ and *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ cells (untreated and OHT-treated).

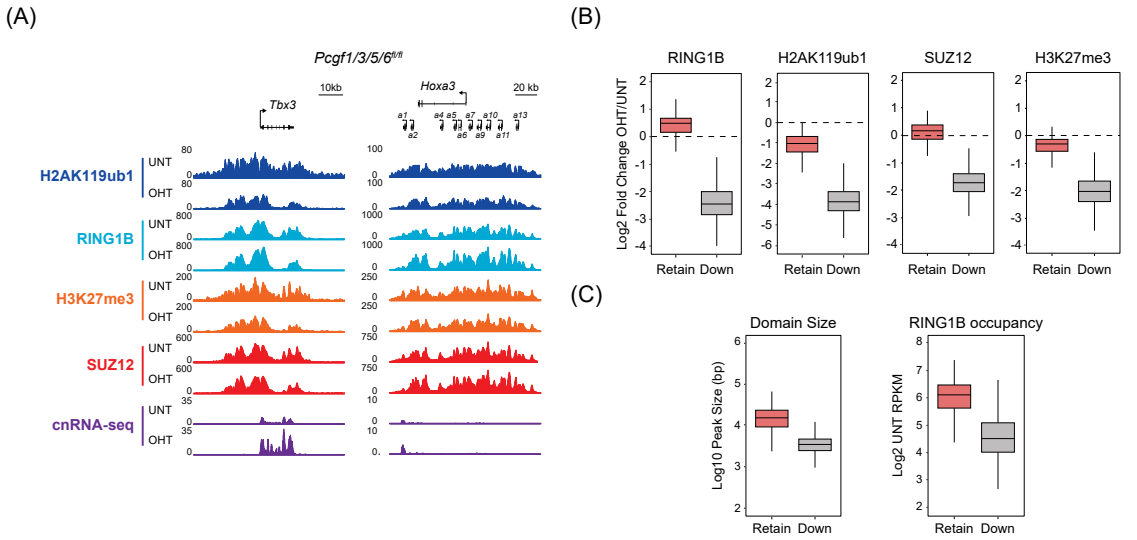



**Figure 6.9: A subset of Polycomb chromatin domains are retained following removal of PCGF1/3/5/6-PRC1**
(A) Genomic snapshot of Polycomb target sites which retain RING1B binding following removal of PCGF1/3/5/6, showing PRC1 (H2AK119ub1 and RING1B) and PRC2 (H3K27me3 and SUZ12) cChIP-seq, as well as cnRNA-seq in *Pcgf1/3/5/6*$^{fl/fl}$ cells (untreated and OHT-treated).
(B) A box plot analysis of log2 fold changes in RING1B, H2AK119ub1, SUZ12 and H3K27me3 cChIP-seq at Polycomb chromatin domains associated with PRC1 target gene promoters which show no reduction (Retain, n=354) or a significant decrease (Down, n=1717) in RING1B occupancy in *Pcgf1/3/5/6*$^{fl/fl}$ cells following OHT treatment.
(C) A box plot analysis of the size (*left*) and RING1B occupancy in untreated cells (*right*) for Polycomb chromatin domains associated with PRC1 target gene promoters which show no reduction (Retain, n=354) or a significant decrease (Down, n=1717) in RING1B occupancy in *Pcgf1/3/5/6*$^{fl/fl}$ cells following OHT treatment.

### 6.2.3 Variant PCGF1/3/5/6-PRC1 complexes define Polycomb chromatin domain formation and repression of Polycomb target genes

In cells lacking PCGF1/3/5, PCGF2/4-PRC1 did not account for the remaining Polycomb-mediated gene repression, suggesting that PCGF6, the only PCGF protein expressed in ES cells which I have not removed, must be responsible for the residual repressive activity of PRC1. This was somewhat surprising as PCGF6-PRC1 has been previously reported to play a specialised role in silencing of germline-specific genes (Suzuki et al. 2016; Endoh et al. 2017; Zhao et al. 2017b). Nevertheless, to examine the possibility that PCGF6 may contribute to Polycomb-mediated gene repression more broadly, I developed a quadruple conditional knockout *Pcgf1/3/5/6*$^{fl/fl}$ ES cell line (Figure 6.8A). Remarkably, when PCGF6 was removed in combination with PCGF1/3/5, the residual H2AK119ub1 throughout the genome (Figures 6.8B-D) and punctate H2AK119ub1 at Polycomb target sites (Figures 6.8E-G and Figure 6.10A) were almost completely lost. Importantly, the effect of PCGF1/3/5/6 removal on H2AK119ub1 exceeded the combined contributions of individual PRC1 complexes (Figure 4.4), even when cooperation between PCGF1/3/5-PRC1 complexes (Figure 6.1) was taken into account, indicating that PCGF1/3/5/6-

PRC1 complexes function synergistically to shape H2AK119ub1 in the genome. Finally, in parallel to the dramatic loss of H2AK119ub1, levels of RING1B, SUZ12 and H3K27me3 at Polycomb target sites were more profoundly reduced following PCGF1/3/5/6 removal, indicating a further erosion of Polycomb chromatin domains (Figures 6.8E-F and Figure 6.10A).

Interestingly, in PCGF1/3/5/6-deficient cells, remaining PRC1 and PRC2, as well as their histone modifications, were not distributed uniformly across Polycomb target sites, but were enriched at large Polycomb chromatin domains bound by high levels of Polycomb factors in untreated cells (Figure 6.8F and Figure 6.9C). In total, I identified 354 PRC1-repressed genes which retained largely unchanged levels of RING1B, SUZ12 and H3K27me3 at their promoters following removal of PCGF1/3/5/6, while H2AK119ub1 was reduced at these sites, but to a smaller extent than at typical PRC1-repressed gene promoters (Figures 6.9A and B). This small group of genes was enriched with genes implicated in the control of developmental processes, including such classical developmental regulators as *Hox* genes (Figures 6.9A and 6.11B). Furthermore, they were highly reminiscent of the genes at which PCGF2/4-PRC1 complexes contributed to transcriptional silencing in the absence of PCGF1/3/5 (Figure 6.7). In agreement with my observations in PCGF1/3/5/2-deficient cells, this further suggested that high-level occupancy of PCGF2/4-PRC1 complexes may be sufficient to support deposition of H2AK119ub1 and Polycomb chromatin domain formation at a small number of Polycomb target sites.

Having uncovered a widespread and major erosion of Polycomb chromatin domains in PCGF1/3/5/6-deficient cells, I next asked what effect removal of PCGF1/3/5/6 had on gene expression. Surprisingly, I found that loss of PCGF1/3/5/6-PRC1 resulted in a global but modest increase in expression of many non-Polycomb target genes (Figures 6.10B, C and E), which was not evident in RING1A/B-deficient cells (Figure 6.10E). One explanation for this widespread transcriptional activation could be that PCGF1/3/5/6 proteins have repressive activities which are independent of RING1A/B. For example, PCGF6 interacts with HP1, and PCGF1 interacts with KDM2B (Farcas et al. 2012; Endoh et al. 2017), both of which have been previously implicated in PRC1-independent modes of gene repression (Jones et al. 2000; Schultz et al. 2002; Turberfield et al. 2019). Alternatively, it is possible that variant PCGF1/3/5/6-PRC1 complexes may counteract a PCGF2/4-dependent activity that stimulates gene expression. Indeed, several recent studies have proposed that PRC1 may facilitate as well as inhibit transcription (Frangini et al. 2013; Creppe et al. 2014; Gao et al. 2014; Morey et al. 2015; Cohen et al. 2018).

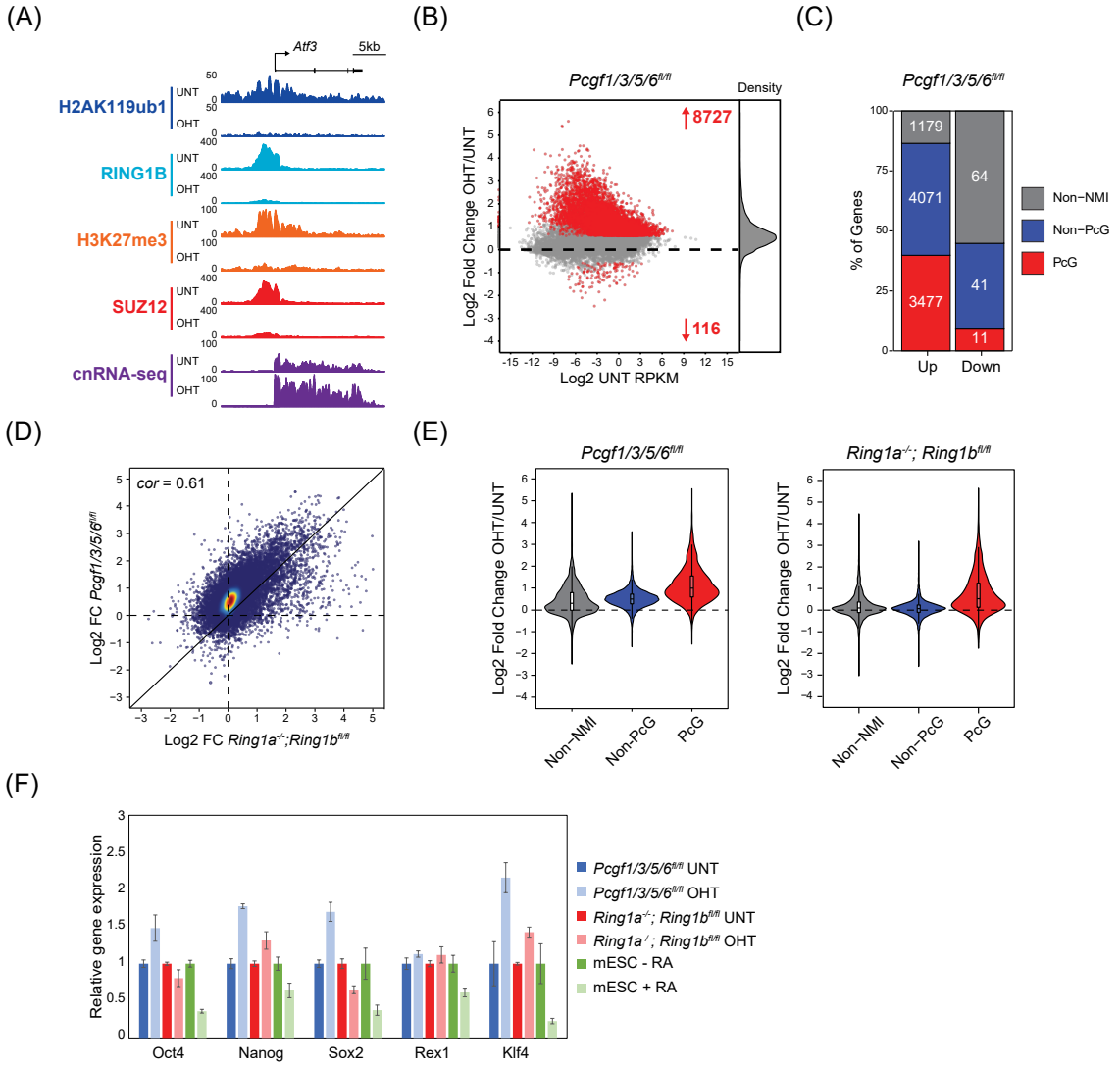



**Figure 6.10: Variant PCGF1/3/5/6-PRC1 complexes are the central determinants of Polycomb-mediated gene repression**

(A) Genomic snapshot of a typical Polycomb target gene, showing cnRNA-seq and cChIP-seq for PRC1 (H2AK119ub1 and RING1B) and PRC2 (H3K27me3 and SUZ12) in *Pcgf1/3/5/6*$^{fl/fl}$ cells (untreated and OHT-treated).
(B) An MA-plot of log2 fold changes in gene expression (cnRNA-seq) in *Pcgf1/3/5/6*$^{fl/fl}$ cells following OHT treatment. Significant gene expression changes (p-adj<0.05 and FC>1.5-fold) are shown in red. Density of gene expression changes is shown on the right.
(C) A bar plot illustrating the distribution of gene expression changes (p-adj<0.05 and FC>1.5-fold) in cnRNA-seq between the gene classes (Non-NMI, Non-PcG, PcG) described in Figure 3.4B for *Pcgf1/3/5/6*$^{fl/fl}$ cells following OHT treatment.
(D) A scatter plot analysis comparing log2 fold changes in gene expression in *Pcgf1/3/5/6*$^{fl/fl}$ and *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ cells following OHT treatment. *Cor* denotes a Pearson correlation coefficient.
(E) A violin plot analysis of log2 fold gene expression changes in *Pcgf1/3/5/6*$^{fl/fl}$ (*left*) and *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ (*right*) cells following OHT treatment for different classes of genes (Non-NMI, Non-PcG and PcG) described in Figure 3.4B.
(F) Expression of key pluripotency factors in *Pcgf1/3/5/6*$^{fl/fl}$ and *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ cells (untreated and OHT-treated), as well as mouse ES cells before (-RA) and after 72 hr retinoic acid treatment (+RA) to induce differentiation. Read counts from cnRNA-seq (*Pcgf1/3/5/6*$^{fl/fl}$ and *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ cells) or 4sU RNA-seq (RA-treated mouse ES cells, from Dimitrova et al. 2018) are shown relative to the corresponding untreated cells. Error bars show SEM (n=3).

Nevertheless, when I focused on Polycomb target genes, I found that they were dramatically derepressed following simultaneous removal of PCGF1/3/5/6-PRC1 complexes (Figures 6.10A, C and E). Importantly, gene reactivation in PCGF1/3/5/6-deficient cells largely phenocopied

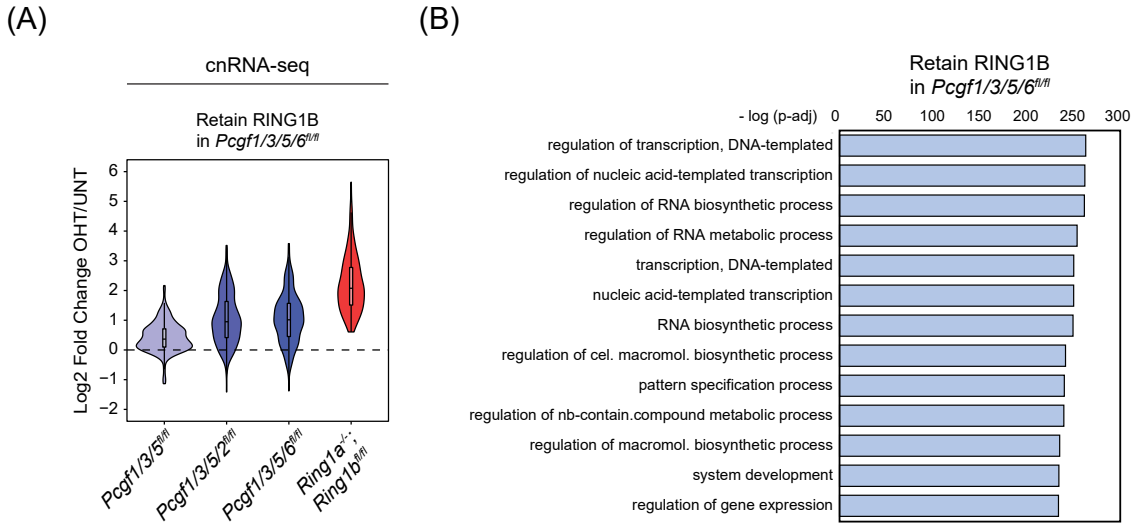



**Figure 6.11: Removal of PCGF1/3/5/6 unmasks the contribution of PCGF2/4-PRC1 complexes to Polycomb-mediated gene repression**
(A) A violin plot analysis of log2 fold expression changes for PRC1 target genes that retain RING1B binding at their promoter upon PCGF1/3/5/6 removal (n = 354) in *Pcgf1/3/5*$^{fl/fl}$, *Pcgf1/3/5/2*$^{fl/fl}$, *Pcgf1/3/5/6*$^{fl/fl}$ and *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ cells following OHT treatment.
(B) A gene ontology analysis of Biological Process term enrichment for PRC1 target genes that retain RING1B binding at their promoters following PCGF1/3/5/6 removal (n = 354).

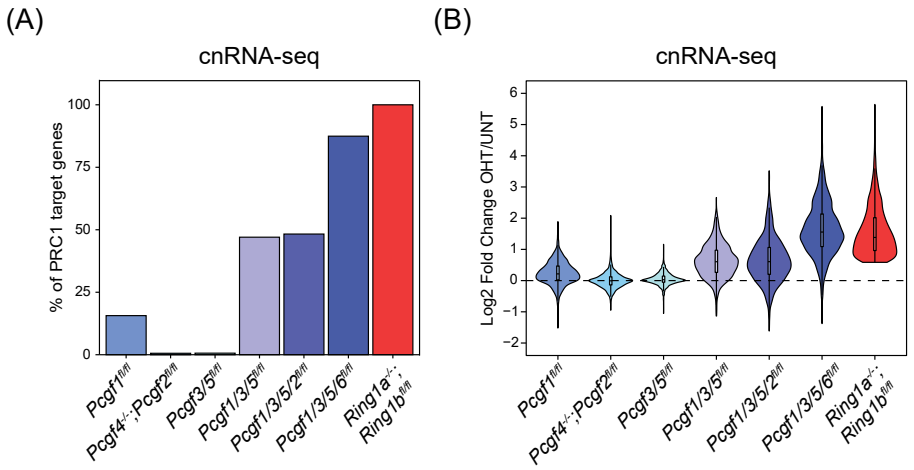



**Figure 6.12: Synergy between variant PCGF1/3/5/6-PRC1 complexes defines Polycomb-mediated gene repression**
(A) A bar plot comparing the proportion of PRC1 target genes that become derepressed (p-adj<0.05 and FC>1.5-fold) across the panel of PRC1 perturbation ES cell lines.
(B) A violin plot analysis comparing log2 fold changes in expression of PRC1 target genes across the panel of PRC1 perturbation ES cell lines.

a complete loss of PRC1 (Figure 6.10D), with approximately 87% of Polycomb target genes which were derepressed following RING1A/B removal being also derepressed to a similar extent following removal of PCGF1/3/5/6 (Figures 6.12A-B). Importantly, these gene expression changes were caused by loss of PRC1 activity and not ES cell differentiation, as both PCGF1/3/5/6- and RING1A/B-deficient cells still expressed key pluripotency factors (Figure 6.10F). Together, these observations provide compelling new evidence for a central role of variant PRC1 complexes in

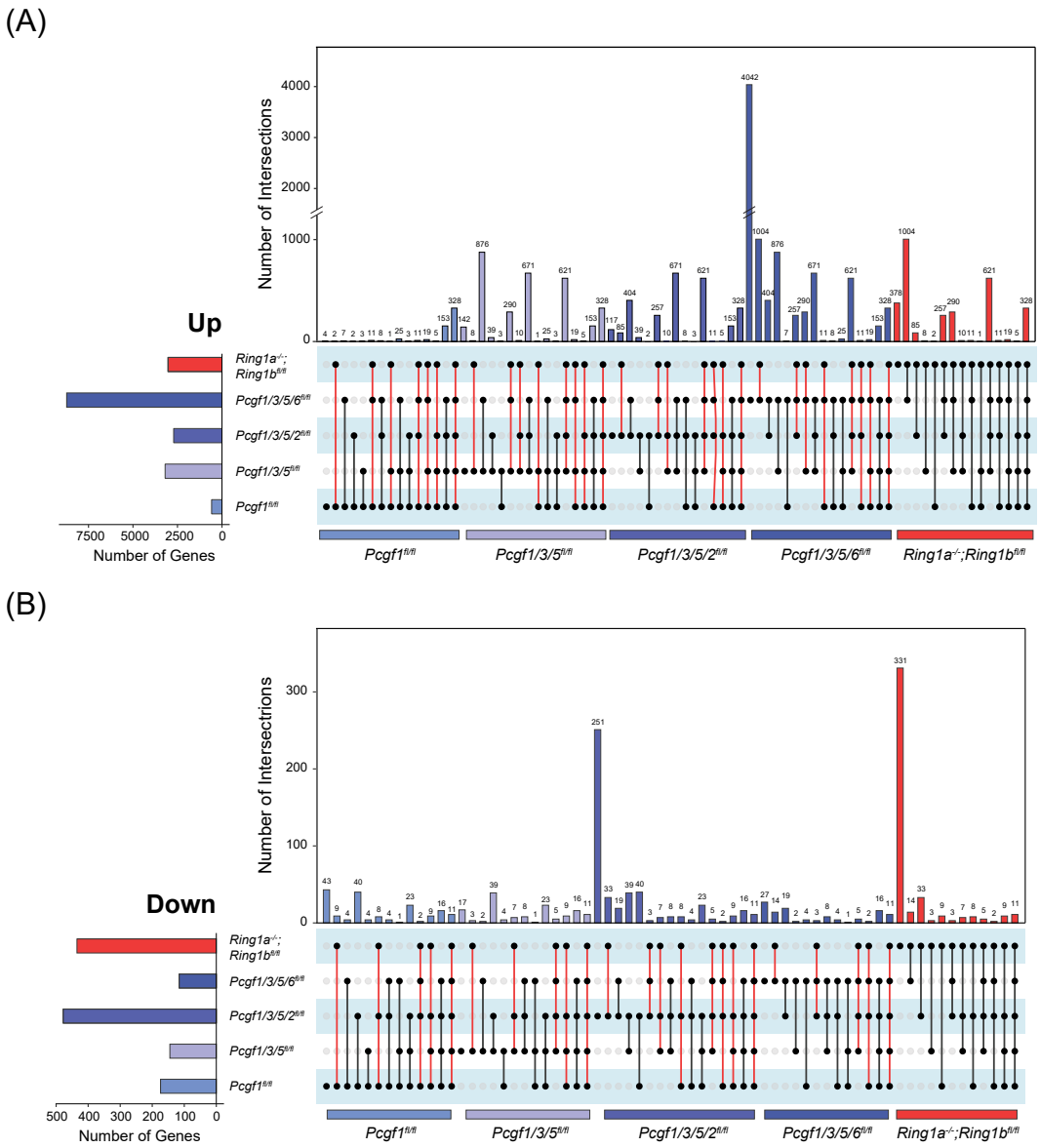



**Figure 6.13: Removal of PCGF proteins in different combinations or RING1A/B results in overlapping transcriptional changes**
(A) An UpSet plot for significantly upregulated genes in *Pcgf1*$^{fl/fl}$, *Pcgf1/3/5*$^{fl/fl}$, *Pcgf1/3/5/2*$^{fl/fl}$, *Pcgf1/3/5/6*$^{fl/fl}$ and *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ cells following OHT treatment. The bar chart on the left shows the number of genes in each set. All non-empty intersections for each set were grouped together and sorted by the number of the overlapped sets. Intersections of each gene set with the set of genes which are upregulated in *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ cells following OHT treatment are highlighted in red.
(B) As in (A) for significantly downregulated genes.

Polycomb-mediated gene repression. Interestingly, a small number of Polycomb target genes, which retained Polycomb chromatin domains following PCGF1/3/5/6 removal, were not fully derepressed in PCGF1/3/5/6-deficient cells as compared to cells completely lacking PRC1 (Figure 6.11A). This indicates that PCGF2/4-PRC1 can contribute to repression of this small subset of Polycomb target genes in addition to PCGF1/3/5/6-PRC1, making these genes more robust to transcriptional perturbations. However, this contrasts the vast majority of Polycomb target genes,

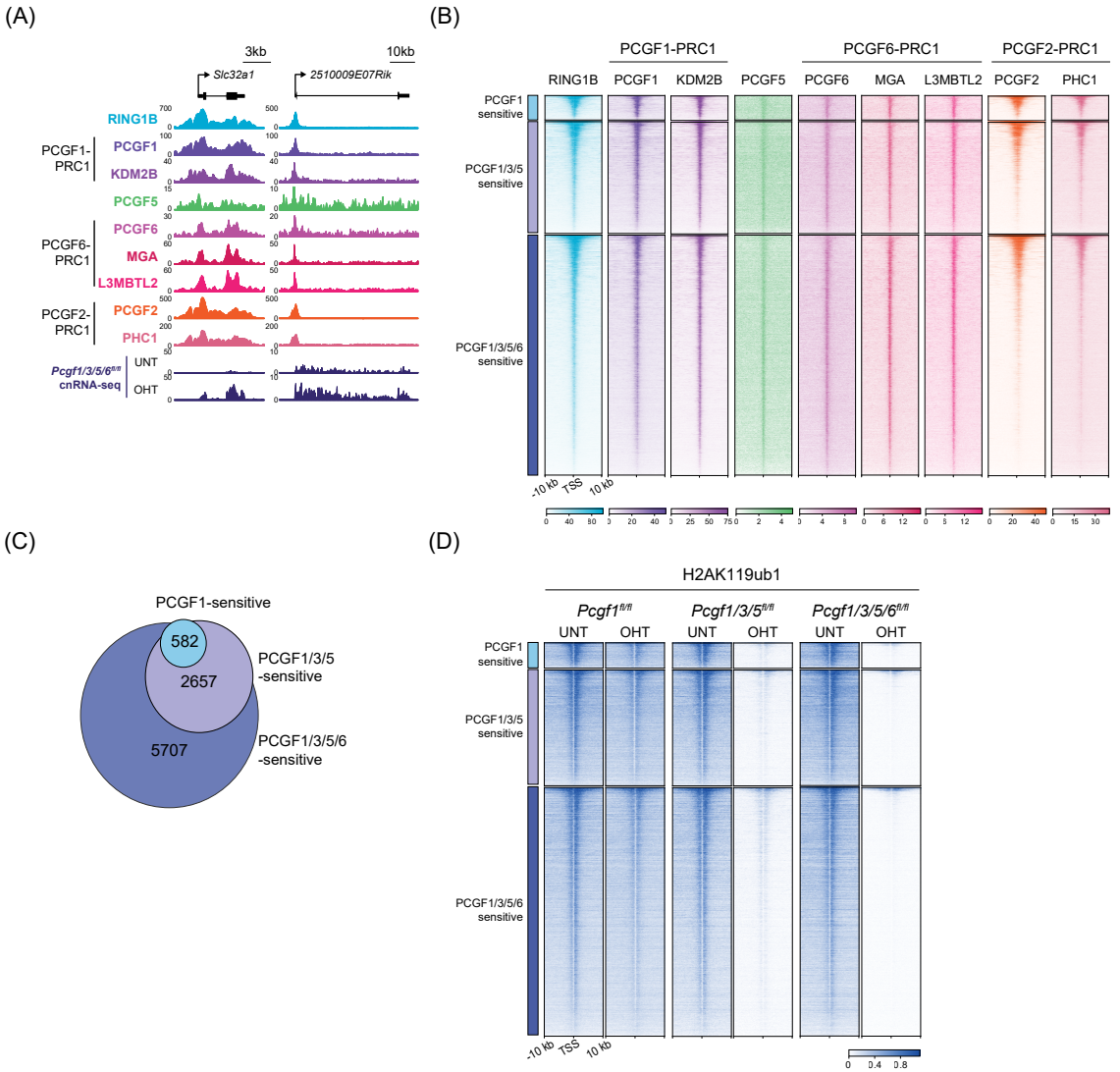



**Figure 6.14: PCGF1/3/5/6-PRC1 complexes co-occupy promoters of variant PRC1 target genes**
(A) Genomic snapshots of two Polycomb target genes, showing ChIP-seq for RING1B and specific subunits of variant and canonical PRC1 complexes in wild-type ES cells, as well as cnRNA-seq in *Pcgf1/3/5/6*$^{fl/fl}$ (untreated and OHT-treated).
(B) Heatmaps of ChIP-seq data for PRC1 factors shown in (A) at the promoters of variant PRC1-regulated genes (PCGF1-sensitive genes, n = 582; PCGF1/3/5-sensitive genes, n = 2657; PCGF1/3/5/6-sensitive genes, n = 5707) sorted by RING1B occupancy. Classification of genes into PCGF1-, PCGF1/3/5- and PCGF1/3/5/6- sensitive genes is described in (C).
(C) A Venn diagram illustrating the classification of variant PRC1-regulated genes used to plot the heatmaps in (B) and (D). PCGF1-sensitive genes are the genes that were significantly upregulated following OHT treatment in *Pcgf1*$^{fl/fl}$ cells. PCGF1/3/5-sensitive genes are the genes that were significantly upregulated following OHT treatment in *Pcgf1/3/5*$^{fl/fl}$ cells, excluding the PCGF1-sensitive genes. Similarly, PCGF1/3/5/6-sensitive genes are the genes that were significantly upregulated following OHT treatment in *Pcgf1/3/5/6*$^{fl/fl}$ cells, excluding the PCGF1- and PCGF1/3/5-sensitive genes.
(D) Heatmaps of H2AK119ub1 cChIP-seq in *Pcgf1*$^{fl/fl}$, *Pcgf1/3/5*$^{fl/fl}$ and *Pcgf1/3/5/6*$^{fl/fl}$ cells (untreated and OHT-treated) at the promoters of variant PRC1-regulated genes. Classification of genes into PCGF1-, PCGF1/3/5- and PCGF1/3/5/6-sensitive genes is described in (C).

which rely almost entirely on repressive activities of variant PCGF1/3/5/6-PRC1 complexes.

Comparing Polycomb target gene derepression across the panel of PRC1 perturbation cell lines highlighted a remarkably high degree of synergy between variant PRC1 complexes, which

underpins Polycomb-mediated transcriptional silencing (Figures 6.12A-B). Removal of PCGF1 led to a moderate reactivation of Polycomb target genes, whereas removal of PCGF2 or PCGF3/5 had a minimal effect on their expression. Simultaneous knockout of PCGF1/3/5 resulted in a synergistic increase in Polycomb target gene derepression when compared to removal of PCGF1 and PCGF3/5 individually. However, removal of PCGF2 in addition to PCGF1/3/5 failed to further potentiate gene derepression. Instead, combined removal of PCGF6 and PCGF1/3/5 was required to achieve largely the same level of gene reactivation as observed in RING1A/B-deficient cells, indicating that PCGF1/3/5/6-PRC1 synergise to drive repression of Polycomb target genes. In addition to the widespread gene reactivation, expression of some genes was reduced following removal of different combinations of PCGF proteins or RING1A/B. However, in contrast to the extensively overlapping sets of upregulated genes, these downregulated genes were less numerous and showed very little overlap between different cell lines (Figures 6.13A-B), suggesting that they most likely represent secondary effects of different PRC1 perturbations.

Finally, to complement the systematic gene expression analysis across PCGF conditional knockout cell lines, I examined the distribution of specific PRC1 subunits at the promoters of variant PRC1 target genes. This revealed that PCGF1/3/5/6-PRC1 complexes extensively co-localise at the promoters of the genes that they regulate (Figures 6.14A-C), in addition to broadly contributing to deposition of H2AK119ub1 at these sites (Figure 6.14D). This was particularly clear for PCGF1-PRC1 and PCGF6-PRC1, which displayed sharp promoter-specific enrichment at the majority of variant PRC1-regulated genes (Figures 6.14A-B). However, enrichment of PCGF5-PRC1 at these regions was considerably more modest (Figure 6.14B), suggesting that these complexes engage with variant PRC1 target gene promoters more transiently, in line with their role in pervasive deposition of H2AK119ub1 throughout the genome. Supporting these observations, occupancy of RING1B, a core component of all PRC1 complexes, was also evident throughout the complement of the derepressed gene promoters (Figure 6.14B). However, in contrast to variant PRC1, distribution of canonical PRC1 was restricted to only a subset of gene promoters, which were bound by high levels of RING1B (Figure 6.14B). Therefore, despite the inherent limitations of comparing ChIP-seq signals derived from different antibodies, these genomic profiles of PCGF proteins are consistent with the central role of cooperation between variant PRC1 complexes in direct regulation of their target genes.

## 6.3 Discussion

Through combinatorial genetic perturbations of PCGF proteins, in Chapter 6, I discovered the central determinants of Polycomb-mediated gene repression in mouse ES cells. I demonstrated that, while cooperation between PCGF1/3/5-PRC1 complexes shapes the majority of H2AK119ub1 at Polycomb target sites and genome-wide, their removal is not sufficient to completely abolish Polycomb chromatin domain formation and repression of Polycomb target genes. I then generated PCGF1/3/5/2- and PCGF1/3/5/6-quadruple conditional knockout cell lines in order to establish which of the remaining PRC1 complexes were responsible for the residual PRC1 repressive activity in PCGF1/3/5-deficient cells. I found that, although PCGF2/4-containing PRC1 complexes can contribute to formation of Polycomb chromatin domains and transcriptional silencing at a small subset of Polycomb target genes, they are unable to fully compensate for the compromised PCGF1/3/5-PRC1 activity. Instead, I discovered that synergy between PCGF1/3/5/6-containing variant PRC1 complexes drives deposition of H2AK119ub1, Polycomb chromatin domain formation and Polycomb-mediated gene repression.

Canonical PRC1 complexes, which bind chromatin through recognition of H3K27me3, have been previously proposed to mediate repression of Polycomb target genes (reviewed in Croce and Helin 2013; Aranda et al. 2015). However, I have demonstrated that conditional removal of PCGF2/4-PRC1 either on their own or in combination with variant PCGF1/3/5-PRC1 complexes leads to minor transcriptional changes, in line with previous work showing that PRC2 is largely dispensable for Polycomb-mediated gene repression in ES cells (Riising et al. 2014). This raises an interesting question: what could be the role of canonical PRC1 at Polycomb target genes, if variant PRC1 complexes are largely sufficient for silencing of these genes? One possibility is that canonical PRC1 complexes may reinforce the fidelity of variant PRC1-dependent gene repression, potentially by mediating higher-order chromatin interactions between Polycomb target gene promoters (Isono et al. 2013; Vieux-Rochas et al. 2015; Schoenfelder et al. 2015; Boettiger et al. 2016; Bonev et al. 2017; Eagen et al. 2017; Kundu et al. 2017; Ogiyama et al. 2018). In line with this possibility, artificial tethering of high levels of the canonical PRC1-specific CBX7 subunit to a reporter gene has been shown to support more stable transcriptional silencing than tethering of a variant PRC1-specific RYBP subunit (Moussa et al. 2019). Efficient maintenance of repression would be particularly advantageous during lineage specification in development,

when diverse transcriptional activation signals have to be constrained. In agreement with these ideas, canonical PRC1 has been previously shown to play an important role in regulation of gene expression programmes in more committed cell types at later developmental stages (Luis et al. 2012; Morey et al. 2015; Kloet et al. 2016; Cohen et al. 2018).

The extensive synergy between variant PRC1 complexes in Polycomb-mediated gene repression was striking, particularly given that mutually exclusive genomic binding of exogenously expressed PCGF proteins has been previously used as evidence for their distinct roles in gene regulation (Gao et al. 2012). In addition, functional specialisation of individual PRC1 complexes has been recently reported by another study, which characterised genomic localisation of endogenous PCGF proteins and concluded that they display high target specificity and virtually no cooperation (Scelfo et al. 2019). However, the specificity was based on modest enrichment of individual PCGF proteins at distinct genomic loci, while their binding appeared to be widespread across all Polycomb target sites, in agreement with my observations. Nevertheless, a combination of this systematic genomic profiling and my functional interrogation of PCGF proteins has provided a comprehensive account of distinct PRC1 complexes and their activities in the genome, with variant PRC1 complexes showing an extensive overlap. Together, these observations indicate that variant PRC1 complexes have retained a shared capacity to identify classical Polycomb target sites, where they cooperate to deposit H2AK119ub1, in addition to evolving distinct targeting activities that support their unique functions at specialised groups of genes, such as germline-related genes in the case of PCGF6-PRC1 (Endoh et al. 2017). Preserving the primal PRC1 function would reinforce the robustness of transcriptional silencing by the Polycomb system, while providing it with the versatility to accommodate new functions during evolutionary diversification. This functional redundancy between variant PRC1 complexes may be particularly advantageous during early embryonic development, when the precise control of gene expression is paramount.

Understanding what drives synergy between variant PRC1 complexes is required to elucidate the molecular mechanisms by which PRC1 counteracts transcription. The only activity shared between otherwise biochemically distinct variant PRC1 complexes is their E3 ubiquitin ligase activity, raising an intriguing possibility that synergy between variant PRC1 complexes may rely on their capacity to catalyse H2AK119ub1. In line with this idea, my findings provided a strong link between the punctate pool of highly enriched H2AK119ub1, which is placed by PCGF1/3/5/6-PRC1, and Polycomb-mediated gene repression. While partial loss of H2AK119ub1 at Polycomb

target gene promoters was associated with a partial derepression of these genes, complete reactivation of Polycomb target genes was only achieved when all punctate H2AK119ub1 was erased following simultaneous removal of variant PRC1 complexes. Interestingly, genes which escaped a complete reactivation in PCGF1/3/5/6-deficient cells retained enrichment of H2AK119ub1 at their promoters, which was sufficient to support Polycomb chromatin domain formation. Together, these observations suggest that H2AK119ub1 may play a central role in Polycomb-mediated gene repression. However, there still remains an active debate in the field whether catalysis by PRC1 is required to silence Polycomb target genes. While some studies have suggested only a modest contribution (Eskeland et al. 2010b; Pengelly et al. 2015; Illingworth et al. 2015; Kundu et al. 2017; Cohen et al. 2018), other studies have provided evidence that H2AK119ub1 is absolutely essential for gene repression (Endoh et al. 2012). Moving forward, an important challenge remains to directly test the role of PRC1 catalytic activity and H2AK119ub1 in key aspects of Polycomb system function. To this end, in Chapter 7, I will systematically dissect the requirement for PRC1 catalytic activity in Polycomb target site selection, formation of Polycomb chromatin domains and gene repression.

# 7

# PRC1 CATALYTIC ACTIVITY IS CENTRAL TO POLYCOMB SYSTEM FUNCTION

## 7.1 Introduction

Based on early studies in *Drosophila*, it has been widely postulated that the Polycomb system functions through a PRC2-initiated hierarchical pathway, in which initial recruitment of PRC2 via DNA-binding factors (Herranz et al. 2008; Dietrich et al. 2012; Li et al. 2017; Perino et al. 2018) or lncRNAs (Brockdorff 2013) leads to deposition of H3K27me3 at Polycomb target sites. This histone modification is recognised by CBX subunits of canonical PRC1 complexes (Cao et al. 2002; Min et al. 2003; Fischle et al. 2003; Wang et al. 2004b), which can support local chromatin compaction and long-range interactions between Polycomb target gene promoters (Francis et al. 2004; Grau et al. 2011; Boettiger et al. 2016; Kundu et al. 2017). Together, these canonical PRC1 activities have been proposed to drive Polycomb-mediated gene repression (Simon and Kingston 2009; Croce and Helin 2013). However, my findings in Chapters 3-6 have challenged this classical PRC2-initiated model, as they showed that canonical PRC1 was largely dispensable for silencing of Polycomb target genes in mouse ES cells. Instead, synergy between variant PRC1 complexes appeared to be paramount, arguing in favour of the PRC1-initiated mechanism for Polycomb chromatin domain formation, which has been recently proposed by our and other labs (Blackledge et al. 2014; Cooper et al. 2014; Kalb et al. 2014). In the context of this alternative model, variant PRC1 complexes rely on DNA-binding activities rather than recognition of H3K27me3 (Farcas et al. 2012; He et al. 2013a; Wu et al. 2013; Blackledge et al. 2014; Endoh et al. 2017; Stielow et al. 2018; Scelfo et al. 2019) to sample putative Polycomb target sites, where in the absence of active transcription, they place high levels of H2AK119ub1, which drives accumulation of PRC2 and H3K27me3, and nucleation of canonical PRC1 binding.

Despite extensive efforts to understand the molecular principles of gene regulation by the Polycomb system, there still remains an active debate in the field as to whether PRC1- or PRC2-initiated pathways underpin Polycomb chromatin domain formation and gene repression

(Simon and Kingston 2013; Steffen and Ringrose 2014; Blackledge et al. 2015; Schuettengruber et al. 2017; Talbert et al. 2019; Loubiere et al. 2019). A fundamental distinction between these alternative models lies in the requirement for PRC1 catalytic activity and H2AK119ub1 in Polycomb system function. In the classical PRC2-initiated model, catalysis by PRC1 has been proposed to be dispensable for PRC2 recruitment, long-range chromatin interactions and repression of Polycomb target genes (Eskeland et al. 2010b; Illingworth et al. 2015; Kundu et al. 2017). In contrast, in the PRC1-initiated model, H2AK119ub1 is the central determinant that drives Polycomb chromatin domain formation and gene repression (Blackledge et al. 2014; Cooper et al. 2014). Although several studies have previously attempted to test the requirement for PRC1 catalytic activity in Polycomb-mediated transcriptional silencing, they have come to contradicting conclusions, with some suggesting that H2AK119ub1 is not required for gene repression (Eskeland et al. 2010b; Illingworth et al. 2015; Pengelly et al. 2015; Kundu et al. 2017) and others providing evidence for its essential role (Endoh et al. 2012). Therefore, the exact contribution of PRC1 catalytic activity and H2AK119ub1 to Polycomb system function still remains elusive and, as a consequence, so do the molecular mechanisms that underpin how the Polycomb system counteracts transcription.

In Chapter 7, I aim to systematically dissect the role of PRC1 catalytic activity in Polycomb system function. To achieve this, I characterise a conditional catalytic point mutant system in which PRC1 can be converted into a catalytically inactive form, while the complex itself remains intact. Using this system in combination with a PRC1 conditional knockout cell line, I directly compare the effect of PRC1 catalytic inactivation to the effect of PRC1 removal on different aspects of Polycomb system function. First, I investigate whether PRC1 catalytic activity underpins Polycomb target site selection or formation of Polycomb chromatin domains. To this end, I examine whether PRC2 binding and H3K27me3, as well as canonical PRC1 occupancy at Polycomb target sites are affected following removal of PRC1 catalytic activity. Next, I investigate how loss of PRC1 catalysis affects overall PRC1 occupancy and distribution of individual variant PRC1 complexes across Polycomb target sites. Finally, I define the contribution of PRC1 catalytic activity to Polycomb-mediated transcriptional silencing. This systematic interrogation of the requirement for PRC1 catalysis in Polycomb system function enables me to determine whether Polycomb chromatin domain formation and gene repression primarily rely on the PRC1- or PRC2-initiated mechanism, helping to finally resolve the ongoing debate in the field.

## 7.2 Results

### 7.2.1 A conditional catalytic point mutant system to eliminate PRC1 catalytic activity

Dissecting the contribution of PRC1 catalytic activity to Polycomb system function required mutations that can inactivate PRC1 catalysis without affecting the stability of PRC1 complexes. Previous work studying the role of PRC1 catalytic activity in mouse ES cells relied on a RING1B$^{I53A}$ mutation, which has been predicted to disrupt the interaction between RING1B and the E2 ubiquitin-conjugating enzyme which is necessary for ubiquitin transfer (Buchwald et al. 2006; Elderkin et al. 2007; Eskeland et al. 2010b; Illingworth et al. 2015; Kundu et al. 2017). However, this mutation is hypomorphic as RING1B$^{I53A}$-PRC1 still retains residual E3 ubiquitin ligase activity both *in vitro* (Figure A.1, Buchwald et al. 2006) and *in vivo* (Tsuboi et al. 2018). Based on the structure of the minimal catalytic domain of PRC1 with E2, several additional mutations have been proposed to disrupt the RING1B-E2 interaction interface (Buchwald et al. 2006; Bentley et al. 2011; McGinty et al. 2014). Among them, was RING1B$^{D56K}$, which has been shown to render PRC1 catalytically inactive both *in vitro* (Figure A.1, Bentley et al. 2011; Taherbhoy et al. 2015) and *in vivo* (Tsuboi et al. 2018), particularly in combination with a RING1B$^{I53A}$ mutation. Importantly, incorporation of RING1B$^{I53A/D56K}$ into PRC1 complexes had no effect on their stability (Figure A.2, courtesy of Neil Blackledge). Therefore, based on the previous biochemical and structural characterisation of PRC1 (Buchwald et al. 2006; Bentley et al. 2011; McGinty et al. 2014; Taherbhoy et al. 2015) and recent unpublished data from our lab (Figures A.1 and A.2), I identified a highly validated mutant form of RING1B, which completely lacks catalytic activity but forms stable PRC1 complexes.

To study transcriptional defects that manifest in the absence of PRC1, in Chapter 3, I have characterised a *Ring1a*$^{-/-}$*;Ring1b*$^{fl/fl}$ conditional knockout ES cell line (PRC1$^{CKO}$). It enables OHT-inducible removal of PRC1 and H2AK119ub1 (Figures 7.1C-E), which is important, since cells constitutively lacking RING1A/B can not be maintained in culture (Endoh et al. 2008). Therefore, to directly compare the effect of PRC1 removal to the effect of PRC1 catalytic inactivation, I required an isogenic conditional PRC1 catalytic point mutant (PRC1$^{CPM}$) ES cell line, in which conversion of PRC1 complexes into a catalytically inactive form can be induced (Figure 7.1A). To achieve this, a cassette, in which a wild-type exon encoding the E2 interaction domain in the sense

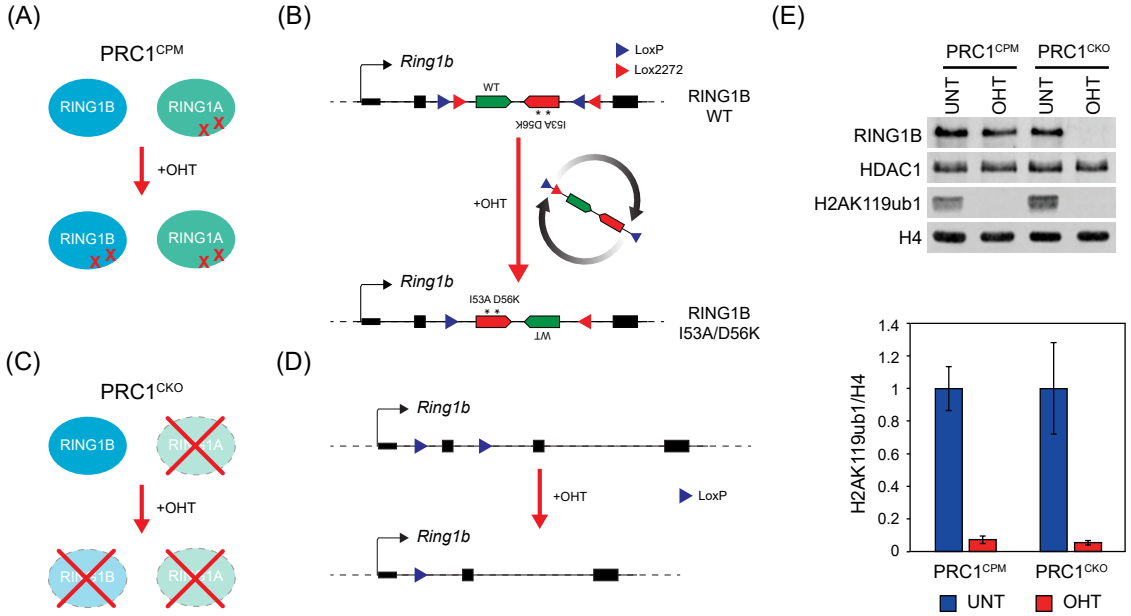



**Figure 7.1: A conditional catalytic point mutant system to eliminate PRC1 catalytic activity**
(A) A schematic of the PRC1$^{CPM}$ cell line, in which RING1A is constitutively inactivated (RING1A$^{I50A/D53K}$) and RING1B can be converted into a catalytically inactive form (RING1B$^{I53A/D56K}$) by addition of OHT.
(B) A schematic of the engineered endogenous *Ring1b* locus in PRC1$^{CPM}$ cells before (UNT) and after OHT treatment, showing inverted double LoxP/Lox2272 sites flanking a cassette encoding a wild-type and mutant version of the RING1B-E2 interaction domain.
(C) A schematic of the PRC1$^{CKO}$ cell line, in which RING1A is constitutively deleted and removal of RING1B can be induced by addition of OHT.
(D) A schematic of the engineered endogenous *Ring1b* locus in PRC1$^{CKO}$ cells before (UNT) and after OHT treatment, showing parallel LoxP sites flanking the first coding exon (exon 2).
(E) Western blot analysis of RING1B (HDAC1 is shown as a loading control) and H2AK119ub1 (H4 is shown as a loading control) in PRC1$^{CPM}$ and PRC1$^{CKO}$ cells before (UNT) and after OHT treatment. Also shown is quantification of H2AK119ub1 levels relative to histone H4. Error bars represent SEM (n=3).

orientation is followed by an I53A/D56K mutant version of this exon in the antisense orientation, has been inserted into both endogenous alleles of the *Ring1b* gene (Figure 7.1B). This wild-type/mutant exon pair was flanked by inverted double LoxP/Lox2272 sites (Figure 7.1B), and the cells were engineered to express ERT2-Cre recombinase from the *Rosa 26* locus. In this PRC1$^{CPM}$ system, addition of OHT leads to an inversion of the wild-type/mutant exon cassette, which results in the mutant exon being expressed and a catalytically inactive RING1B$^{I53A/D56K}$ protein being produced (Figures 7.1A, B and E). Furthermore, RING1A-containing PRC1 complexes can also contribute to deposition of H2AK119ub1 in mouse ES cells, despite RING1A being expressed at much lower levels than RING1B (Endoh et al. 2008; Leeb et al. 2010; Kloet et al. 2016). Therefore, to eliminate RING1A contribution, both endogenous alleles of the *Ring1a* gene were constitutively inactivated by introducing RING1A$^{I50A/D53K}$ mutations, which correspond to the inactivating catalytic mutations in RING1B (Figure 7.1A).

To validate the PRC1$^{CPM}$ system, I first confirmed that RING1B and H2AK119ub1 levels were

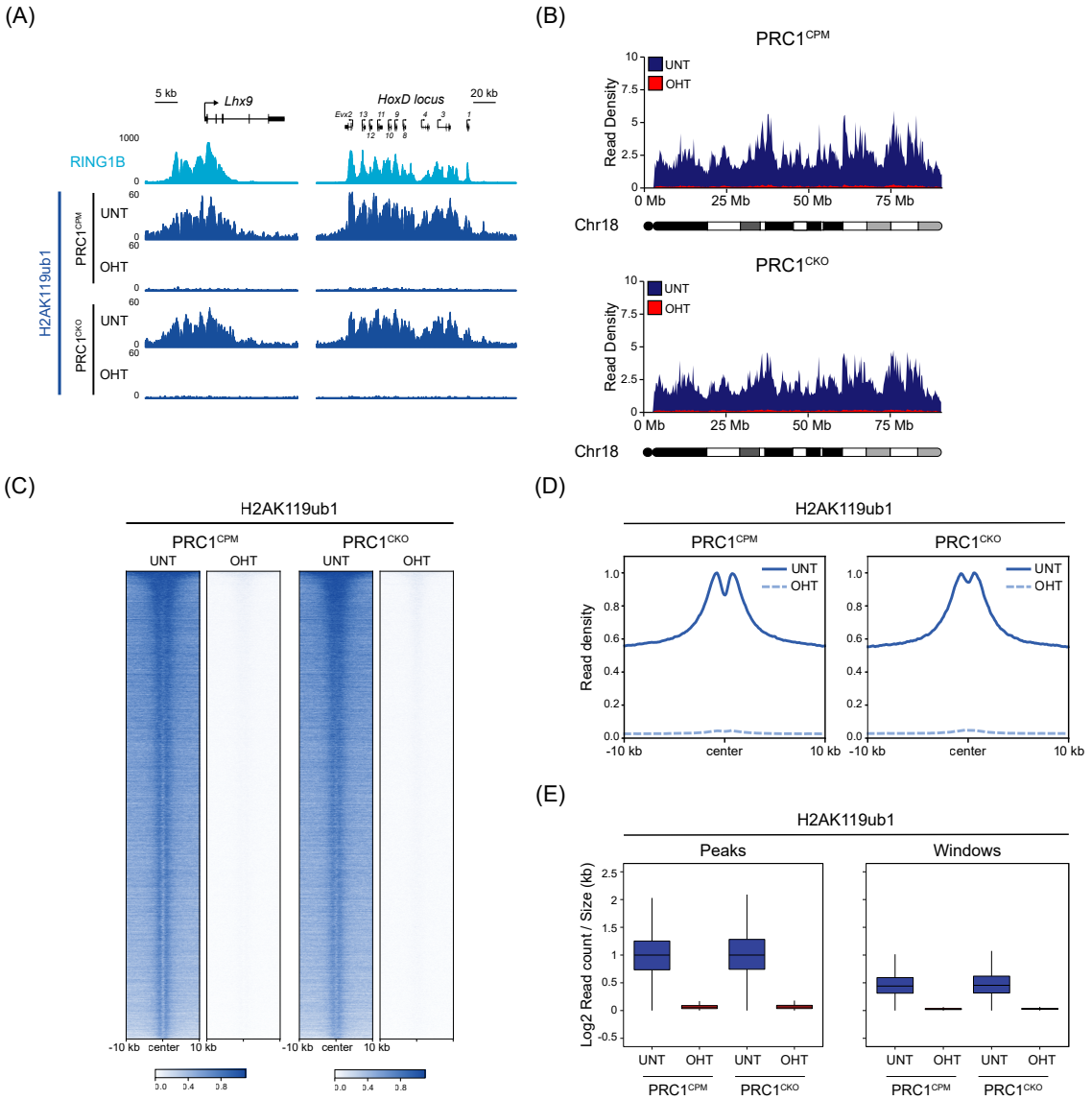



**Figure 7.2: H2AK119ub1 is completely lost throughout the genome following OHT treatment in PRC1$^{CPM}$ cells**
(A) Genomic snapshots of typical Polycomb target genes showing RING1B cChIP-seq in untreated PRC1$^{CKO}$ cells and H2AK119ub1 cChIP-seq in PRC1$^{CPM}$ and PRC1$^{CKO}$ cells (untreated and OHT-treated).
(B) Chromosome density plots showing H2AK119ub1 cChIP-seq across chromosome 18 in untreated and OHT-treated PRC1$^{CPM}$ (*top*) and PRC1$^{CKO}$ (*bottom*) cells.
(C) Heatmap analysis of H2AK119ub1 cChIP-seq at PRC1-bound sites (n = 18,643) in PRC1$^{CPM}$ and PRC1$^{CKO}$ cells (untreated and OHT-treated). Genomic intervals are sorted based on RING1B occupancy in untreated PRC1$^{CKO}$ cells.
(D) Metaplot analysis of H2AK119ub1 cChIP-seq at PRC1-bound sites (n = 18,643) in PRC1$^{CPM}$ and PRC1$^{CKO}$ cells (untreated and OHT-treated).
(E) Box plot analysis of the normalised H2AK119ub1 cChIP-seq signal at PRC1-bound sites and 100 kb genomic windows in PRC1$^{CPM}$ and PRC1$^{CKO}$ cells (untreated and OHT-treated). Read counts/kb were normalised to the median signal at PRC1-bound sites in untreated cells for each cell line individually.

similar in untreated PRC1$^{CPM}$ and PRC1$^{CKO}$ cells (Figure 7.1E), indicating that manipulations with the *Ring1b* locus in PRC1$^{CPM}$ cells did not interfere with normal processing of the *Ring1b* transcript and production of a catalytically competent RING1B protein. Importantly, following OHT treatment in PRC1$^{CPM}$ cells, RING1B protein levels remained largely unchanged, whereas H2AK119ub1 was completely lost as quantified by western blot analysis of histone extracts

(Figure 7.1E). Genome-wide erasure of H2AK119ub1 was further corroborated by cChIP-seq analysis, which demonstrated that H2AK119ub1 was no longer found at highly enriched Polycomb target sites (Figures 7.2A, C and D) or more broadly throughout the genome (Figures 7.2B and E). Importantly, loss of H2AK119ub1 following OHT treatment in PRC1$^{CPM}$ cells was identical to that observed in PRC1$^{CKO}$ cells, which was evident both by western blot and cChIP-seq analysis (Figures 7.1E and 7.2A-E), while PRC1$^{CKO}$ cells also lost PRC1 complexes (Figure 7.1E). Together, a combination of the isogenic conditional catalytic point mutant (PRC1$^{CPM}$) and conditional knockout (PRC1$^{CKO}$) ES cell lines (Figures 7.1A and C) now enables me to test the PRC1-initiated model for Polycomb chromatin domain formation and finally resolve whether PRC1 catalytic activity is required for Polycomb system function.

### 7.2.2 PRC1 catalytic activity facilitates PRC2 binding and deposition of H3K27me3 at Polycomb target sites

Our lab has recently discovered that removal of PRC1 results in a major reduction in PRC2 and H3K27me3 levels at Polycomb target gene promoters (Blackledge et al. 2014; Rose et al. 2016). Based on these observations and work from other labs, it has been proposed that Polycomb chromatin domains are formed through a PRC1-initiated mechanism which relies on recognition of H2AK119ub1 by PRC2 (Cooper et al. 2014; Kalb et al. 2014; Cooper et al. 2016). My findings in Chapter 6 further support this idea, as I uncovered a strong correlation between reductions in H2AK119ub1 levels and erosion of Polycomb chromatin domains following removal of variant PRC1 complexes. Nevertheless, it still remains unclear whether PRC1 catalytic activity directly contributes to PRC2 binding and H3K27me3 deposition at Polycomb target sites. To address this question, I performed cChIP-seq analysis for SUZ12 (a core PRC2 subunit) and H3K27me3 in untreated and OHT-treated PRC1$^{CPM}$ and PRC1$^{CKO}$ cells. I found that in untreated cells, SUZ12 binding mostly occurred at sites which were also occupied by RING1B (96% of SUZ12-bound sites overlapped with RING1B-bound sites), while RING1B binding appeared to be more broad (Figure 7.3B). Strikingly, following loss of PRC1 catalytic activity in the OHT-treated PRC1$^{CPM}$ cells, both SUZ12 and H3K27me3 levels were dramatically reduced at Polycomb-occupied sites (Figures 7.3A, C-E and G-H). Importantly, these changes closely recapitulated the extent of Polycomb chromatin domain erosion following complete removal of PRC1 in the OHT-treated PRC1$^{CKO}$ cells (Figures 7.3A and C-I), indicating that PRC1 catalytic activity is required for

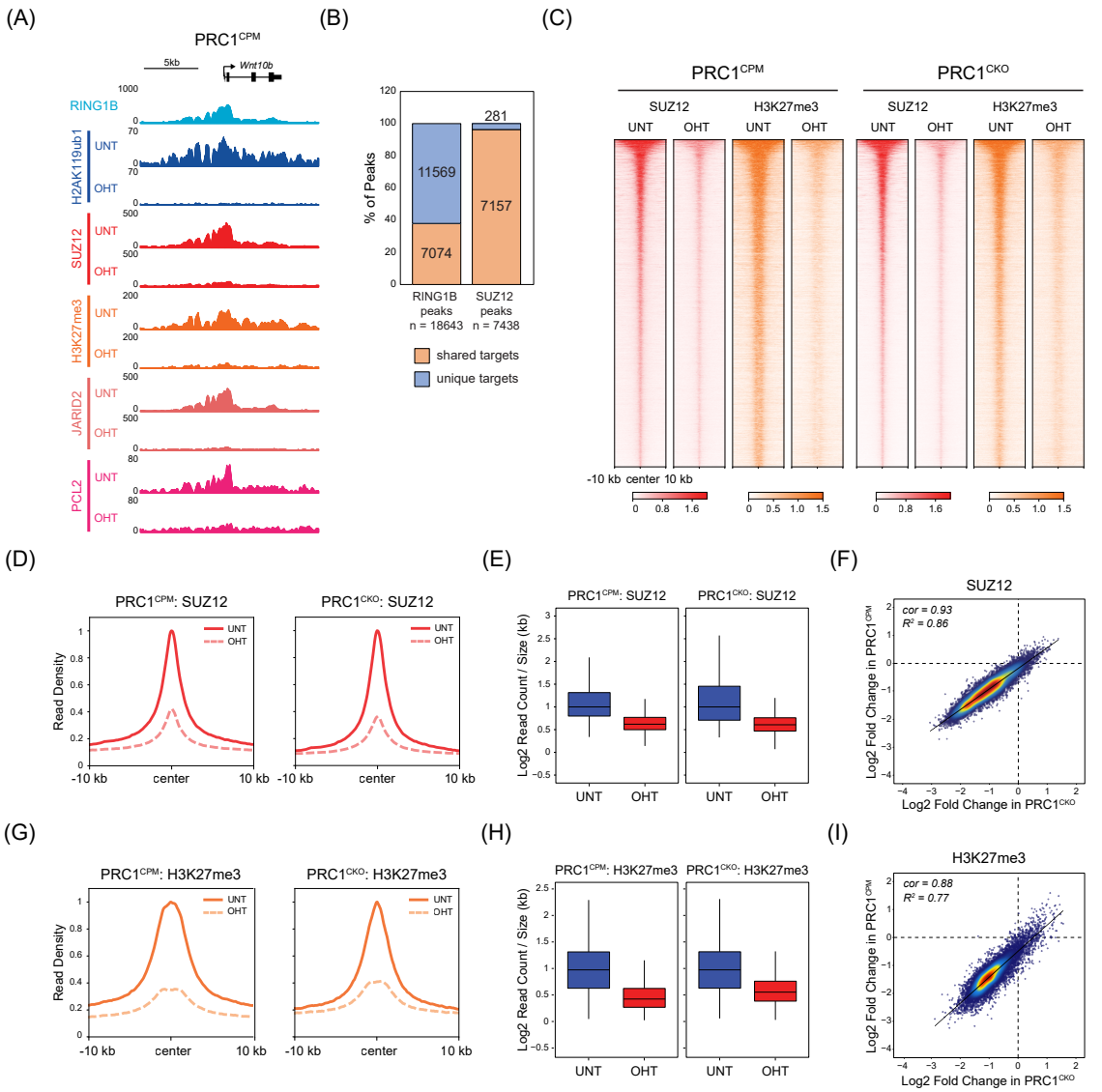



**Figure 7.3: PRC1 catalytic activity supports stable PRC2 binding and deposition of H3K27me3 at Polycomb target sites**

(A) Genomic snapshot of a typical Polycomb target gene showing cChIP-seq for H2AK119ub1, SUZ12, H3K27me3, JARID2 and PCL2 in untreated and OHT-treated PRC1$^{CPM}$ cells. Also shown is RING1B cChIP-seq in untreated cells.
(B) Bar plot showing the overlap between RING1B- and SUZ12-bound genomic regions. RING1B-bound sites that are also occupied by SUZ12 are herein referred to as Polycomb-occupied sites (n = 7074).
(C) Heatmap analysis of cChIP-seq for SUZ12 and H3K27me3 at Polycomb-occupied sites (n = 7074) in PRC1$^{CPM}$ and PRC1$^{CKO}$ cells (untreated and OHT-treated). Genomic intervals are sorted based on RING1B occupancy in untreated PRC1$^{CKO}$ cells.
(D) Metaplot analysis of SUZ12 cChIP-seq at Polycomb-occupied sites (n = 7074) in PRC1$^{CPM}$ and PRC1$^{CKO}$ cells (untreated and OHT-treated).
(E) Box plot analysis of the normalised SUZ12 cChIP-seq signal at Polycomb-occupied sites in untreated and OHT-treated PRC1$^{CPM}$ and PRC1$^{CKO}$ cells.
(F) Scatter plot analysis comparing log2 fold changes in SUZ12 cChIP-seq signal at Polycomb-occupied sites in PRC1$^{CPM}$ versus PRC1$^{CKO}$ cells following OHT treatment. $R^2$ represents a coefficient of determination for linear regression, and *cor* denotes a Pearson correlation coefficient.
(G) Metaplot analysis of H3K27me3 cChIP-seq at Polycomb-occupied sites (n = 7074) in PRC1$^{CPM}$ and PRC1$^{CKO}$ cells (untreated and OHT-treated).
(H) Box plot analysis of the normalised H3K27me3 cChIP-seq signal at Polycomb-occupied sites for untreated and OHT-treated PRC1$^{CPM}$ and PRC1$^{CKO}$ cells.
(I) Scatter plot analysis comparing log2 fold changes in H3K27me3 cChIP-seq signal at Polycomb-occupied sites in PRC1$^{CPM}$ versus PRC1$^{CKO}$ cells following OHT treatment. $R^2$ represents a coefficient of determination for linear regression, and *cor* denotes a Pearson correlation coefficient.

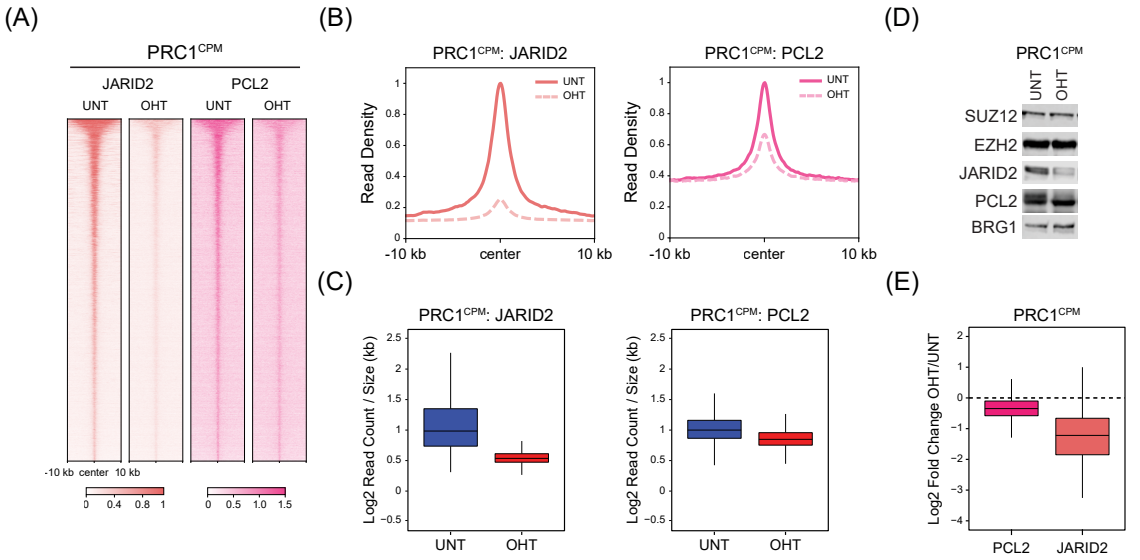



**Figure 7.4: Loss of PRC1 catalytic activity compromises binding of both classes of PRC2 complexes**
(A) Heatmap analysis of cChIP-seq for JARID2 and PCL2 at Polycomb-occupied sites (n = 7074) in PRC1$^{CPM}$ cells (untreated and OHT-treated). Genomic intervals are sorted as in Figure 7.3C.
(B) Metaplot analysis of cChIP-seq for JARID2 and PCL2 at Polycomb-occupied sites (n = 7074) in PRC1$^{CPM}$ cells (untreated and OHT-treated).
(C) Box plot analysis of the normalised PCL2 and JARID2 cChIP-seq signal at Polycomb-occupied sites in untreated and OHT-treated PRC1$^{CPM}$ cells.
(D) Western blot analysis of PRC2 subunits (SUZ12, EZH2, JARID2 and PCL2) in untreated and OHT-treated PRC1$^{CPM}$ cells. BRG1 is shown as a loading control.
(E) Box plot analysis of log2 fold changes in PCL2 and JARID2 cChIP-seq signal at Polycomb-occupied sites in PRC1$^{CPM}$ cells following OHT treatment.

PRC2 binding and deposition of H3K27me3 at Polycomb target sites.

Interestingly, low levels of PRC2 and H3K27me3 were still evident across Polycomb target sites following OHT treatment both in PRC1$^{CPM}$ and PRC1$^{CKO}$ cells (Figures 7.3C, D and G), suggesting that in addition to PRC1-dependent binding, alternative mechanisms may be responsible for PRC2 recruitment to these sites. Indeed, two distinct classes of PRC2 complexes with unique targeting modules have been recently characterised (reviewed in Laugesen et al. 2019; Yu et al. 2019; Mierlo et al. 2019). PRC2.1 incorporates PCL subunits, which have been shown to bind non-methylated CpG-rich DNA and methylated histone tails (Ballare et al. 2012; Brien et al. 2012; Musselman et al. 2012; Cai et al. 2013; Li et al. 2017; Perino et al. 2018), while PRC2.2 contains JARID2 subunit, which has been implicated in direct recognition of H2AK119ub1 (Kalb et al. 2014; Cooper et al. 2016). Therefore, I wondered whether PRC1 catalytic activity was required for recruitment of PRC2.2 but not PRC2.1 complexes to Polycomb target sites. To test this hypothesis, I performed cChIP-seq analysis for specific subunits of the two PRC2 classes (PCL2 for PRC2.1 and JARID2 for PRC2.2) in untreated and OHT-treated PRC1$^{CPM}$ cells.

In line with the proposed role of JARID2 as a reader of H2AK119ub1 (Kalb et al. 2014; Cooper

et al. 2016), JARID2 occupancy at Polycomb-occupied sites was almost completely lost following OHT treatment in PRC1$^{CPM}$ cells (Figures 7.4A-C). Interestingly, this was accompanied by a marked reduction in JARID2 protein levels (Figure 7.4D), suggesting that chromatin binding may be required to stabilise JARID2 protein. Surprisingly, occupancy of PCL2 at Polycomb-occupied sites was also reduced in the OHT-treated PRC1$^{CPM}$ cells (Figures 7.4A-C), albeit to a lesser extent than for JARID2 (Figure 7.4E), while PCL2 protein levels were largely unaffected (Figure 7.4D). Although PCL subunits have been proposed to drive PRC2.1 occupancy through specific binding to NMIs (Li et al. 2017; Perino et al. 2018), recognition of H3K27me3 by the PRC2 core subunit, EED, stimulates PRC2 catalytic activity, and this has been also proposed to directly or indirectly promote recruitment of PRC2 to target sites (Hansen et al. 2008; Margueron et al. 2009; He et al. 2017; Oksuz et al. 2018; Youmans et al. 2018). Therefore, it is possible that following removal of PRC1 catalytic activity, a dramatic decrease in H3K27me3 levels at Polycomb-occupied sites (Figures 7.3A, C, G and H) may have caused reductions in PCL2 occupancy (Figure 7.4A-C). Furthermore, the remaining PRC2.1 binding and H3K27me3 may also explain the residual JARID2 occupancy at Polycomb-occupied sites in the OHT-treated PRC1$^{CPM}$ cells (Figure 7.4A-C). Together, these observations demonstrate that, although low levels of PRC2.1 and PRC2.2 can be recruited to Polycomb target sites independently of PRC1 catalysis, high-level occupancy of both PRC2 classes and H3K27me3 deposition at these sites requires PRC1 catalytic activity.

### 7.2.3 Recruitment of canonical PRC1 to Polycomb target sites relies on PRC1 catalytic activity

Having established that removal of PRC1 catalytic activity leads to a dramatic reduction in H3K27me3 levels at Polycomb target sites (Figure 7.3), I reasoned that this may compromise recruitment of canonical PRC1 complexes which recognise H3K27me3 (Cao et al. 2002; Min et al. 2003; Wang et al. 2004b). To test whether this was the case, I carried out cChIP-seq analysis for PCGF2, a core subunit of canonical PRC1, before and after OHT treatment in PRC1$^{CPM}$ cells. In untreated cells, PCGF2 occupancy was enriched at PRC1-bound sites with high levels of PRC2 and H3K27me3 (Figures 7.5A-B), consistent with the PRC2-dependent mechanism for canonical PRC1 recruitment (Wang et al. 2004b). Importantly, following loss of PRC1 catalytic activity in the OHT-treated PRC1$^{CPM}$ cells, PCGF2 binding at target sites was almost completely lost (Figures 7.6A-D). Furthermore, this was associated with marked reductions in protein levels of

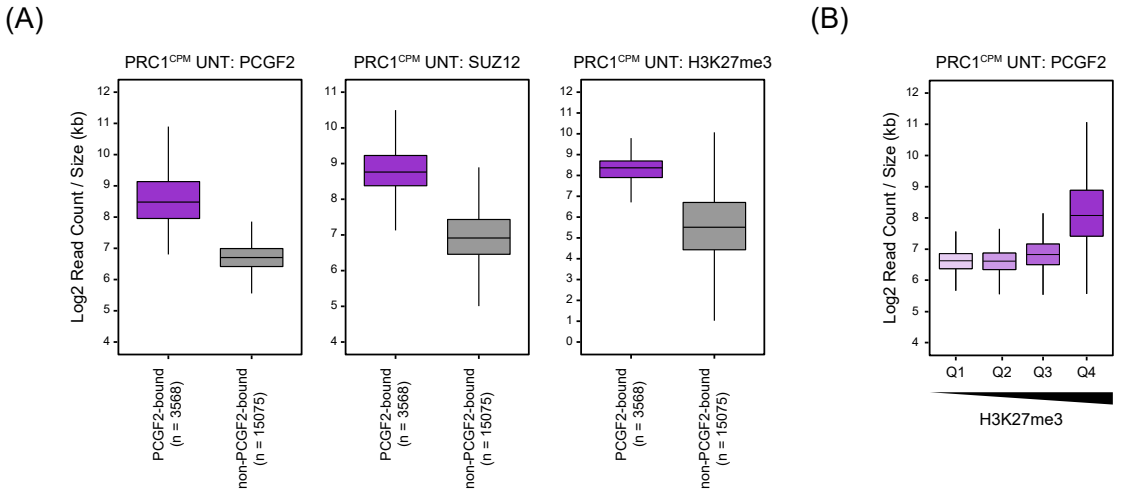



**Figure 7.5: Canonical PRC1 preferentially occupies PRC1-bound sites which are highly enriched with PRC2 and H3K27me3**

(A) Box plot analysis of cChIP-seq for PCGF2, SUZ12 and H3K27me3 in untreated PRC1$^{CPM}$ cells at PRC1-bound sites (n = 18,643) divided into sites that are occupied (herein referred to as PCGF2 target sites) or not occupied by PCGF2.

(B) Box plot analysis of cChIP-seq for PCGF2 in untreated PRC1$^{CPM}$ cells at PRC1-bound sites divided into quartiles (from lowest Q1 to highest Q4) based on H3K27me3 levels in untreated cells.

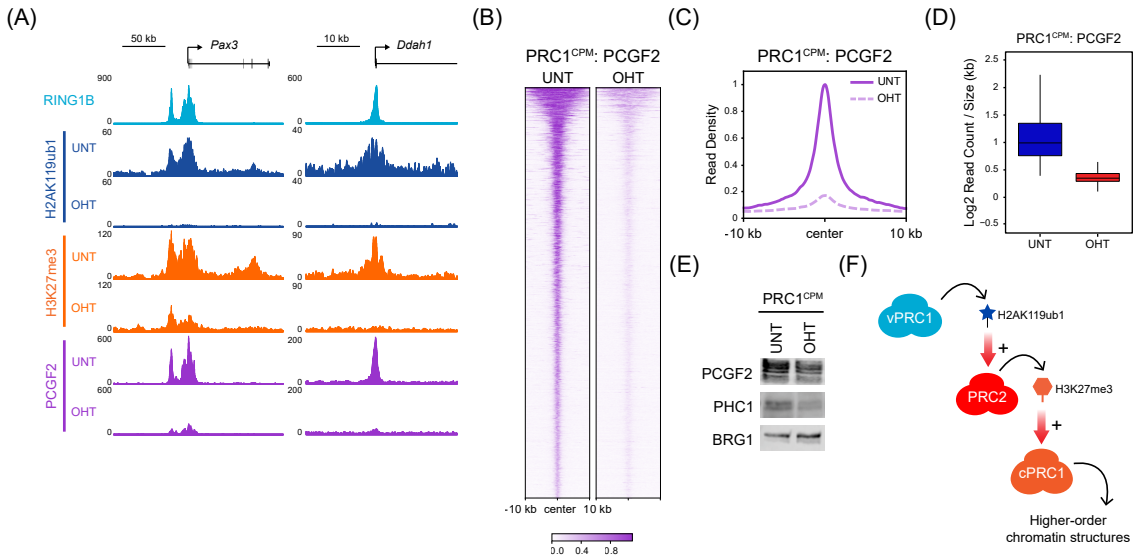



**Figure 7.6: Binding of canonical PRC1 at target sites is lost following removal of PRC1 catalytic activity**

(A) Genomic snapshots of typical Polycomb target genes showing cChIP-seq for H2AK119ub1, H3K27me3 and PCGF2 in PRC1$^{CPM}$ cells (untreated and OHT-treated). Also shown is RING1B cChIP-seq in untreated cells.

(B) Heatmap analysis of PCGF2 cChIP-seq at PCGF2 target sites (n = 3568, as in Figure 7.5A) in PRC1$^{CPM}$ cells (untreated and OHT-treated). Genomic intervals are sorted based on RING1B occupancy in untreated PRC1$^{CKO}$ cells.

(C) Metaplot analysis of PCGF2 cChIP-seq at PCGF2 target sites (n = 3568) in PRC1$^{CPM}$ cells (untreated and OHT-treated).

(D) Box plot analysis of the normalised cChIP-seq signal for PCGF2 at PCGF2 target sites in untreated and OHT-treated PRC1$^{CPM}$ cells.

(E) Western blot analysis of canonical PRC1 components (PCGF2 and PHC1) in PRC1$^{CPM}$ cells (untreated and OHT-treated). BRG1 is shown as a loading control.

(F) A schematic summarising the PRC1-initiated model for Polycomb chromatin domain formation: deposition of H2AK119ub1 by variant PRC1 drives recruitment of PRC2 which places H3K27me3 at Polycomb target sites; Recognition of H3K27me3 facilitates binding of canonical PRC1 which mediates formation of higher-order chromatin interactions between these genomic regions.

canonical PRC1-specific subunits (Figure 7.6E), suggesting that canonical PRC1 complexes are destabilised when not bound to chromatin. Finally, unpublished work from our lab revealed that long-range chromatin interactions between Polycomb target sites, which have been proposed to rely on canonical PRC1 (Kundu et al. 2017), were largely abolished following removal of PRC1 catalytic activity (Figure A.3, courtesy of Dr. Angelika Feldmann). Together, these observations further support a PRC1-initiated recruitment model, in which histone H2A mono-ubiquitination by PRC1 drives a sequence of downstream events, including PRC2 binding, deposition of H3K27me3 and recruitment of canonical PRC1, which culminate in the formation of Polycomb chromatin domains that can engage in long-range genomic interactions (Figure 7.6F).

### 7.2.4 Polycomb target site selection relies on DNA-based sampling by variant PRC1 complexes but not PRC1 catalytic activity

Having established that PRC1 and its catalytic activity drive Polycomb chromatin domain formation, I reasoned that elucidating the mechanisms that underpin PRC1 binding in the genome will help me to understand how target sites which form Polycomb chromatin domains are selected. To this end, I analysed genome-wide distribution of RING1B using cChIP-seq in untreated and OHT-treated PRC1$^{CPM}$ cells and compared it to the binding of other Polycomb factors and chromatin features. In untreated cells, RING1B occupancy correlated strongly with the levels of PCGF2 and PRC2 (SUZ12 and H3K27me3) (Figures 7.7A-B and 7.8A-B), consistent with a dominant role of canonical PRC1 in defining high-level RING1B binding in ES cells (Chapters 3 and 6). However, following removal of PRC1 catalytic activity and the resulting loss of PRC2 and canonical PRC1 binding (Figures 7.3 and 7.6), RING1B occupancy at sites with the highest enrichment of PRC1 and PRC2 was dramatically reduced (Figures 7.7A-D). Strikingly, at a large number of sites that were bound by low or moderate levels of PRC1 and PRC2 in untreated cells, RING1B occupancy remained unchanged or even increased in OHT-treated PRC1$^{CPM}$ cells (Figures 7.7A-D) and now correlated only modestly with canonical PRC1 binding or H3K27me3 (Figures 7.7A-B and 7.8A-B). In agreement with my previous findings, this demonstrates that canonical PRC1 complexes, which are recruited to chromatin in a PRC2-dependent manner, account for the majority of RING1B occupancy at Polycomb target sites in ES cells (Chapters 3 and 6), but this is a downstream consequence of PRC1 catalysis and not the primary mechanism for Polycomb target site selection.

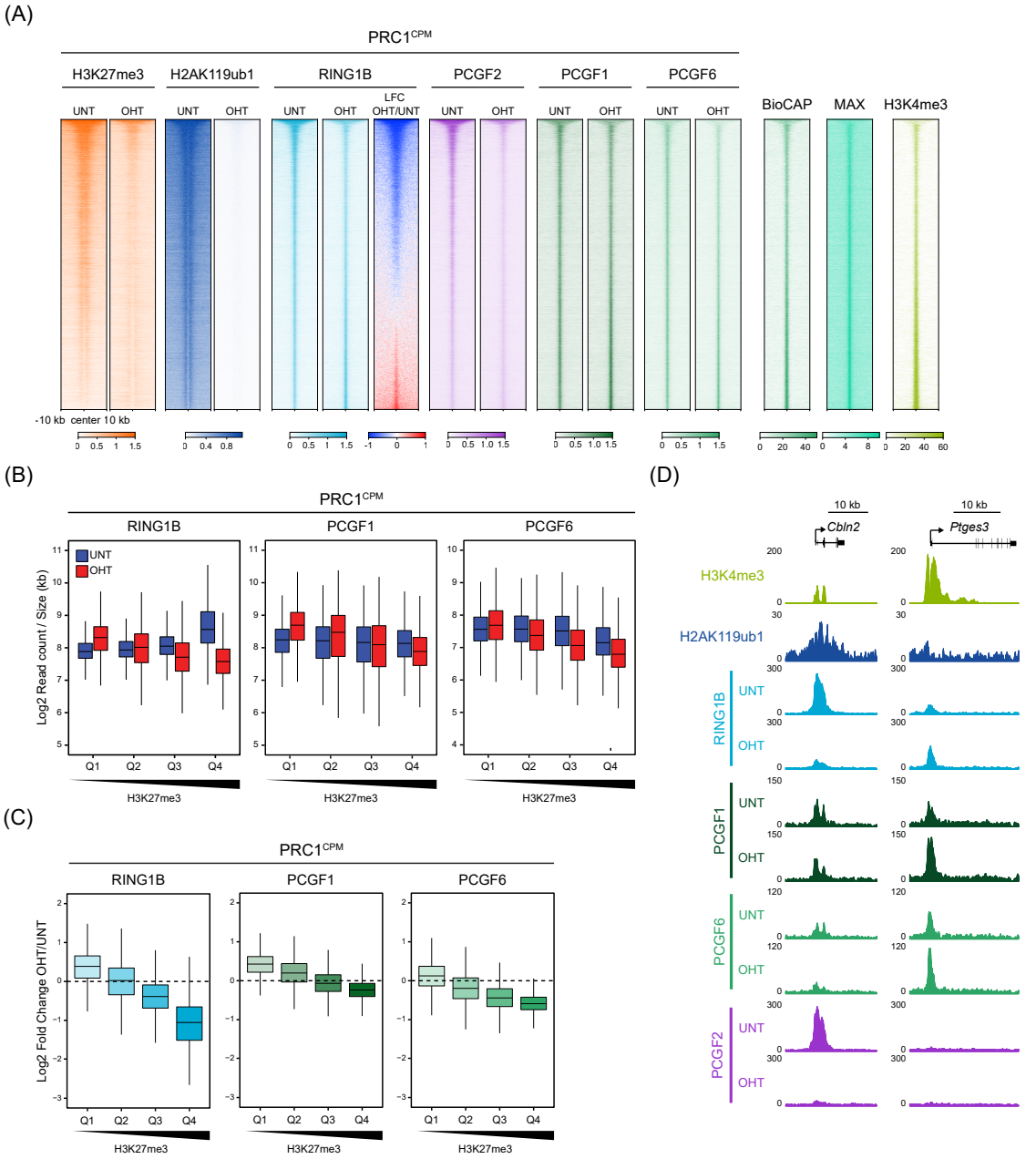



**Figure 7.7: Variant PRC1 complexes broadly occupy Polycomb target sites independently of PRC1 catalysis**
(A) Heatmap analysis of cChIP-seq for H3K27me3, H2AK119ub1, RING1B, PCGF2, PCGF1 and PCGF6 at PRC1-bound sites (n = 18,643) in PRC1$^{CPM}$ cells (untreated and OHT-treated). Also shown is BioCAP-seq (as a measure of non-methylated CpG-rich DNA) and ChIP-seq for MAX and H3K4me3 in wild-type ES cells. Genomic intervals are sorted based on log2 fold changes in RING1B binding following OHT treatment in PRC1$^{CPM}$ cells.
(B) Box plot analysis of cChIP-seq signal for RING1B, PCGF1 and PCGF6 in untreated and OHT-treated PRC1$^{CPM}$ cells at PRC1-bound sites divided into quartiles based on H3K27me3 levels in untreated cells (from lowest Q1 to highest Q4).
(C) Box plot analysis comparing log2 fold changes in cChIP-seq signal for RING1B, PCGF1 and PCGF6 following OHT treatment in PRC1$^{CPM}$ cells at PRC1-bound sites divided into quartiles based on H3K27me3 levels in untreated cells (from lowest Q1 to highest Q4).
(D) Genomic snapshots of two genes with different expression levels showing cChIP-seq for RING1B, PCGF1, PCGF6 and PCGF2 in PRC1$^{CPM}$ cells (untreated and OHT-treated). Also shown is H2AK119ub1 cChIP-seq in untreated cells and H3K4me3 ChIP-seq in wild-type ES cells.

To understand what is responsible for PRC1 recruitment to target sites in the absence of PRC1 catalytic activity, I examined genomic distribution of variant PCGF1- and PCGF6-

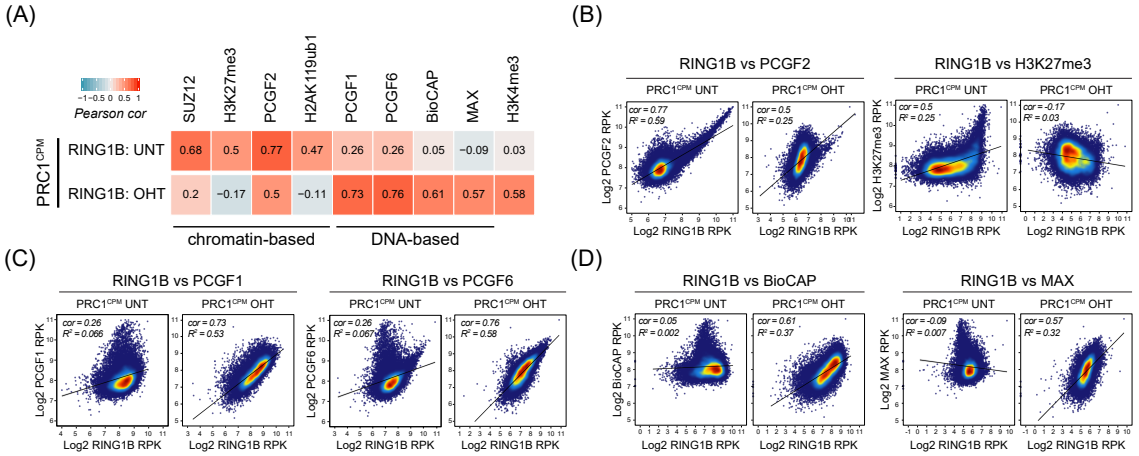



**Figure 7.8: Removal of PRC1 catalytic activity unmasks DNA-based sampling of Polycomb target sites by variant PRC1 complexes**
(A) Correlation of RING1B cChIP-seq signal with cChIP-seq signal for PRC2 (SUZ12 and H3K27me3) and PRC1 (PCGF2, PCGF1, PCGF6 and H2AK119ub1) in untreated and OHT-treated PRC1$^{CPM}$ cells. Also shown is correlation with BioCAP-seq, as well as MAX and H3K4me3 ChIP-seq, in wild-type ES cells.
(B) Scatter plot analysis showing the relationship between RING1B cChIP-seq signal and PCGF2 or H3K27me3 cChIP-seq signal at PRC1-bound sites in PRC1$^{CPM}$ cells (untreated and OHT-treated). $R^2$ represents a coefficient of determination for linear regression and *cor* denotes a Pearson correlation coefficient.
(C) As in (B) for RING1B and PCGF1 cChIP-seq (*left*) or RING1B and PCGF6 cChIP-seq (*right*).
(D) Scatter plot analysis showing the relationship between RING1B cChIP-seq signal in PRC1$^{CPM}$ cells (untreated and OHT-treated) with BioCAP-seq (*left*) or MAX ChIP-seq signal (*right*) in wild-type ES cells at PRC1-bound sites. $R^2$ represents a coefficient of determination for linear regression and *cor* denotes a Pearson correlation coefficient.

PRC1 complexes, both of which contain DNA-binding modules. The PCGF1-PRC1 complex incorporates KDM2B, a ZF-CXXC domain protein that binds non-methylated CpG-rich DNA (Farcas et al. 2012; He et al. 2013a; Wu et al. 2013), while PCGF6-PRC1 incorporates MAX/MGA and E2F6/DP1 DNA-binding transcription factors (Hauri et al. 2016; Endoh et al. 2017). In untreated PRC1$^{CPM}$ cells, cChIP-seq for PCGF1 and PCGF6 revealed a broad and generally uniform binding of both factors across PRC1-occupied sites (Figures 7.7A-B), which correlated poorly with RING1B levels (Figures 7.8A and C). Furthermore, occupancy of PCGF1 and PCGF6 was only modestly affected following loss of PRC1 catalytic activity (Figures 7.7A-D), indicating that binding of variant PCGF1- and PCGF6-PRC1 complexes is largely independent and upstream of PRC1 catalysis and Polycomb chromatin domain formation. Importantly, in the OHT-treated PRC1$^{CPM}$ cells, RING1B occupancy now showed a strong correlation with PCGF1 and PCGF6 binding, as well as with non-methylated CpG-rich DNA and MAX levels (Figures 7.8A, C and D). Together, these observations strongly suggest that DNA-binding activities inherent to variant PRC1 complexes constitute the primary targeting mechanisms that underpin target site selection by PRC1. Interestingly, a small reduction in PCGF1 and PCGF6 occupancy was evident following removal of PRC1 catalytic activity, particularly in the case of PCGF6 (Figures 7.7A-C). This

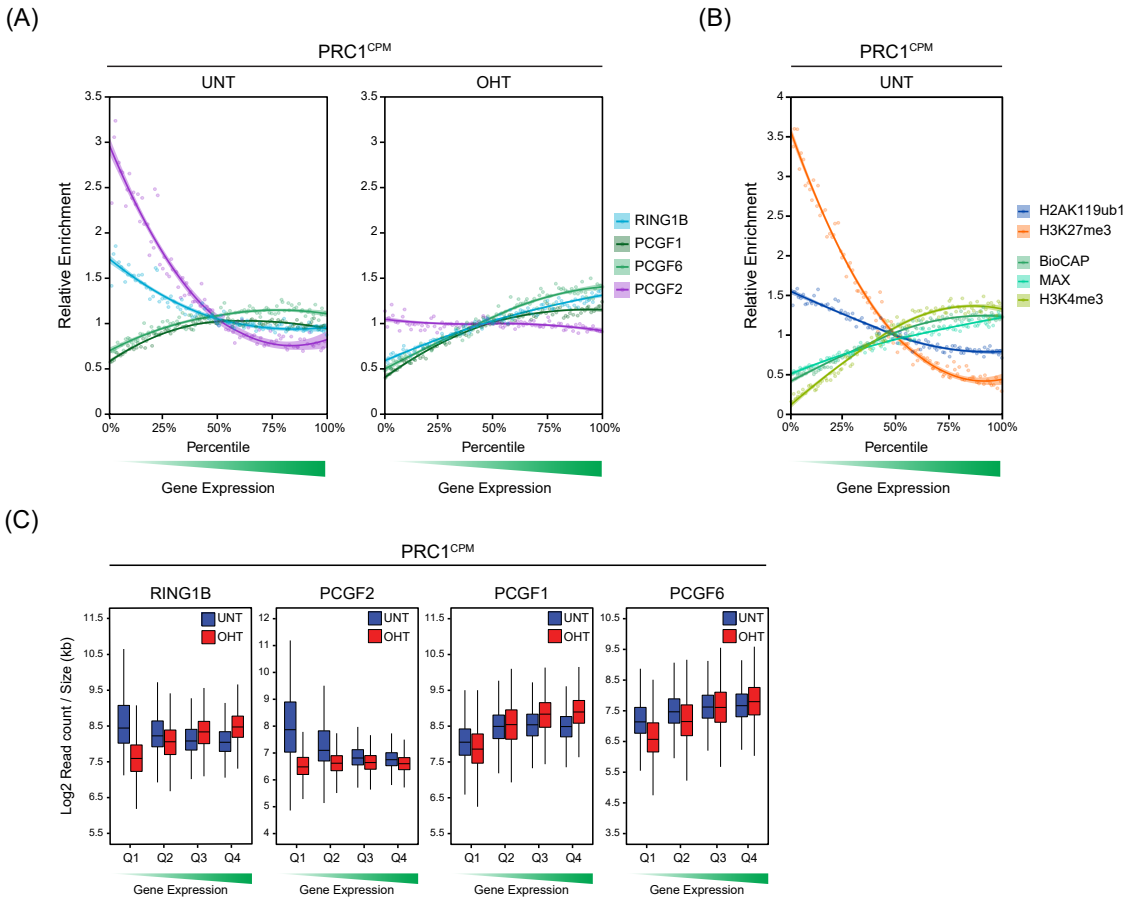



**Figure 7.9: PRC1 catalytic activity drives high-level PRC1 occupancy at lowly transcribed gene promoters**
(A) Relative enrichment of RING1B, PCGF1, PCGF6 and PCGF2 cChIP-seq in PRC1$^{CPM}$ cells (untreated and OHT-treated) at PRC1-bound sites divided into percentiles based on the expression level of the associated gene in untreated PRC1$^{CPM}$ cells. For each factor, the enrichment is shown relative to the fiftieth percentile. Lines represent smoothed conditional means based on loess local regression fitting.
(B) Relative enrichment of H2AK119ub1 and H3K27me3 cChIP-seq in untreated PRC1$^{CPM}$ cells and H3K4me3 and MAX ChIP-seq, as well as BioCAP-seq signal, in wild-type ES cells at PRC1-bound sites divided into percentiles based on the expression level of the associated gene in untreated PRC1$^{CPM}$ cells. For each factor, the enrichment is shown relative to the fiftieth percentile. Lines represent smoothed conditional means based on loess local regression fitting.
(C) Box plot analysis of RING1B, PCGF2, PCGF1 and PCGF6 cChIP-seq signal in untreated and OHT-treated PRC1$^{CPM}$ cells at PRC1-bound sites divided into quartiles based on the expression level of the associated gene in untreated PRC1$^{CPM}$ cells (from lowest Q1 to highest Q4).

suggests that in addition to the primary DNA-based mode of recruitment, occupancy of variant PRC1 complexes at target sites may be further stabilised by recognition of Polycomb-associated histone modifications, although the contribution of this chromatin-based mode of binding is modest. In line with this idea, variant PRC1-specific RYBP subunit contains a ubiquitin-binding domain, which has been proposed to recognise H2AK119ub1 (Arrigoni et al. 2006) and contribute to PRC1 recruitment in specialised contexts, such as X-inactivation (Almeida et al. 2017).

While variant PRC1 complexes occupy many sites in the genome, Polycomb chromatin domains form only at a small subset of PRC1-bound genomic loci (Figure 7.7A), suggesting that something

may be counteracting their formation at the remaining sites. Active transcription has been previously proposed to interfere with binding and/or activity of Polycomb repressive complexes (Schmitges et al. 2011; Klose et al. 2013; Jermann et al. 2014; Riising et al. 2014; Beltran et al. 2016). In support of this idea, in untreated PRC1$^{CPM}$ cells, the highest levels of RING1B and PCGF2 were associated with promoters of lowly transcribed genes, which were enriched with H3K27me3 and H2AK119ub1, but not H3K4me3, a hallmark of expressed genes (Guenther et al. 2007; Barski et al. 2007; Benayoun et al. 2014) (Figures 7.7A and D, and 7.9A-C). In contrast, RING1B target sites which were associated with more actively transcribed genes showed low levels of H2AK119ub1, PRC2, H3K27me3 and canonical PRC1, despite being occupied by similar levels of PCGF1- and PCGF6-PRC1 complexes (Figures 7.7A and D, and 7.9A-C). Importantly, high-level enrichment of RING1B at lowly transcribed gene promoters was lost following OHT treatment in PRC1$^{CPM}$ cells, unmasking an underlying uniform distribution of RING1B across target sites, which now mirrored occupancy of PCGF1 and PCGF6 (Figures 7.7A and 7.9A-C). Together these observations are consistent with a model in which variant PCGF1- and PCGF6-PRC1 complexes broadly engage with, or sample, putative Polycomb target sites via their DNA-binding subunits, which, in the absence of counteracting features associated with active transcription, can lead to high-level H2AK119ub1 deposition and subsequent Polycomb chromatin domain formation.

### 7.2.5 PRC1 catalytic activity is central to Polycomb-mediated gene repression

Having demonstrated that Polycomb chromatin domain formation relies on catalysis by PRC1, I set out to address the ultimate outstanding question of whether PRC1 catalytic activity is also required for Polycomb-mediated gene repression. To this end, I carried out calibrated RNA-seq (cRNA-seq) analysis to directly compare gene expression changes following OHT treatment in the isogenic PRC1$^{CPM}$ and PRC1$^{CKO}$ cells. In agreement with my calibrated nuclear RNA-seq analysis of nascent transcription (Chapter 3), I found that removal of PRC1 resulted in dramatic derepression of Polycomb target genes in the OHT-treated PRC1$^{CKO}$ cells (Figures 7.10A-B), supporting a central role for PRC1 in Polycomb-mediated gene repression in ES cells. Remarkably, following inactivation of PRC1 catalysis in the OHT-treated PRC1$^{CPM}$ cells, a similar number of Polycomb target genes were reactivated (Figures 7.10A-B), indicating that PRC1 catalytic activity is essential for silencing of these genes. Importantly, the magnitude

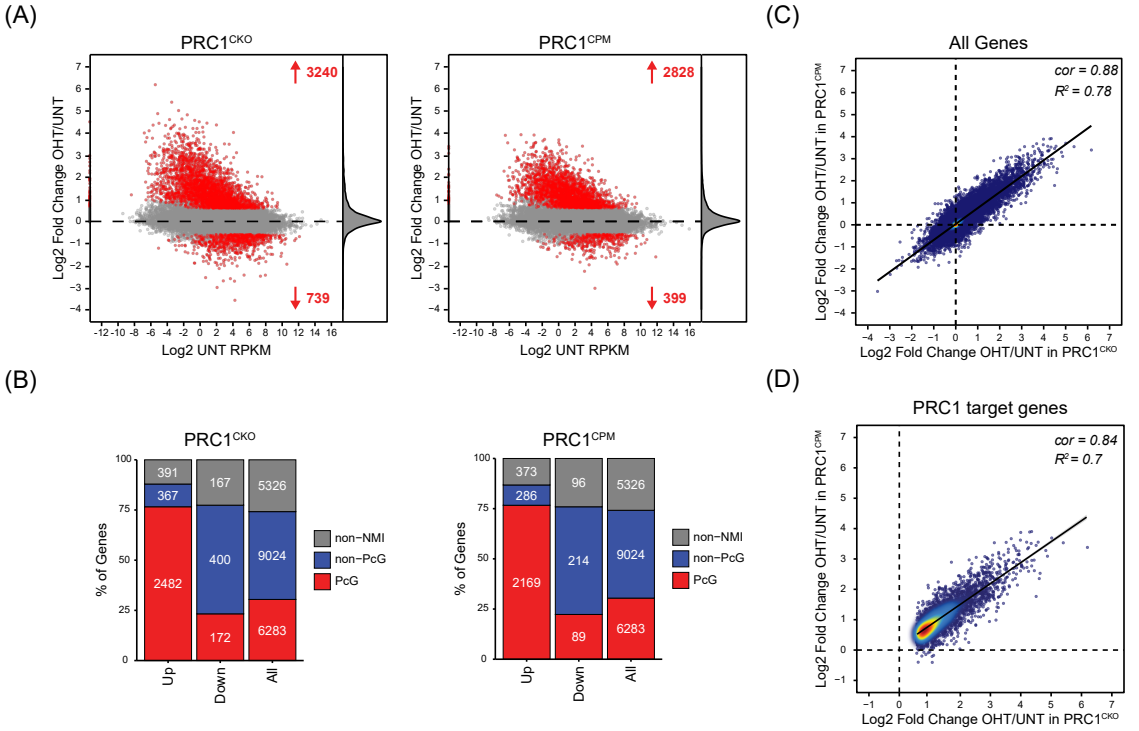



**Figure 7.10: PRC1 catalytic activity is required for Polycomb-mediated gene repression**
(A) MA-plots of log2 fold changes in gene expression (cRNA-seq) in PRC1$^{CKO}$ and PRC1$^{CPM}$ cells following OHT treatment. Significant gene expression changes (p-adj<0.05 and FC>1.5-fold) are shown in red. Density of gene expression changes is shown on the right.
(B) Bar plots illustrating the distribution of gene expression changes (p-adj<0.05 and FC>1.5-fold) in cRNA-seq between the gene classes (Non-NMI, Non-PcG, PcG) described in Figure 3.4B for PRC1$^{CKO}$ and PRC1$^{CPM}$ cells following OHT treatment.
(C) Scatter plot analysis comparing log2 fold expression changes (cRNA-seq) for all genes in PRC1$^{CKO}$ and PRC1$^{CPM}$ cells following OHT treatment. $R^2$ represents a coefficient of determination for linear regression and *cor* denotes a Pearson correlation coefficient.
(D) Scatter plot analysis comparing log2 fold expression changes (cRNA-seq) for PRC1 target genes in PRC1$^{CKO}$ and PRC1$^{CPM}$ cells following OHT treatment. $R^2$ represents a coefficient of determination for linear regression and *cor* denotes a Pearson correlation coefficient.

of transcriptional changes following OHT treatment, both for all genes and specifically for PRC1 target genes, strongly correlated between the PRC1$^{CPM}$ and PRC1$^{CKO}$ cell lines (Figures 7.10C-D). Together, these observations reveal that removal of PRC1 catalytic activity closely recapitulates the gene expression defects that manifest when PRC1 is completely lost.

Although the majority of derepressed genes following OHT treatment were shared between PRC1$^{CPM}$ and PRC1$^{CKO}$ cells, some genes were classified as derepressed only in one cell line (Figure 7.11A, *top panel*). Importantly, the cell line-specific misregulated genes were reactivated considerably less compared to the shared misregulated genes but to the same level following OHT treatment in both cell lines (Figure 7.11A, *bottom panel*). This suggests that most of the differences between PRC1$^{CPM}$ and PRC1$^{CKO}$ cell lines are due to the higher sensitivity of small gene expression changes to the choice of the significance cut-off rather than different

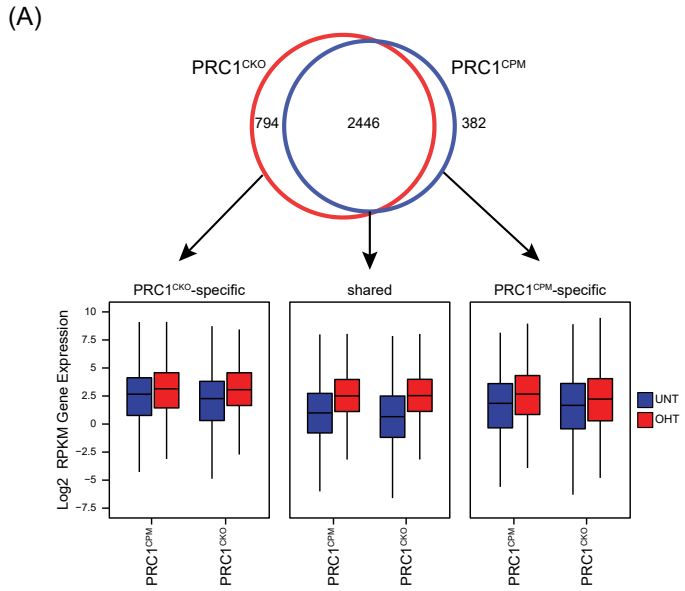



**Figure 7.11: Gene reactivation following OHT treatment is highly similar in PRC1$^{CPM}$ and PRC1$^{CKO}$ cells**
(A) *Top panel:* Venn diagram showing the overlap between genes which were derepressed (p-adj<0.05 and FC>1.5-fold) in PRC1$^{CKO}$ (*red*) and PRC1$^{CPM}$ (*blue*) cells following OHT treatment. *Bottom panel:* Box plot analysis comparing gene expression (cRNA-seq) in PRC1$^{CPM}$ and PRC1$^{CKO}$ cells (untreated and OHT-treated) for the three groups of derepressed genes defined in the Venn diagram above: PRC1$^{CKO}$-specific, shared between PRC1$^{CKO}$ and PRC1$^{CPM}$ cells, and PRC1$^{CPM}$-specific.

effects of PRC1 catalytic inactivation or removal on transcriptional silencing. However, I did find 241 PRC1 target genes which were significantly less reactivated following OHT treatment in PRC1$^{CPM}$ comparing to PRC1$^{CKO}$ cells (Figures 7.12B-E). Promoters of these genes were associated with large Polycomb chromatin domains which were highly enriched with canonical PRC1 and H3K27me3 in untreated cells (Figures 7.12C and 7.13A-C), reminiscent of the sites at which canonical PRC1 could partially compensate for the loss of variant PRC1 complexes and contribute to gene silencing (Chapter 6). In line with this, low levels of canonical PRC1 binding and H3K27me3 were retained at the promoters of these genes in the OHT-treated PRC1$^{CPM}$ cells (Figures 7.12C and 7.13A-B). Nevertheless, repression of this small subset of genes still heavily relied on PRC1 catalysis, as removal of canonical PRC1 alone caused only a modest increase in their expression (Figure 7.12D). Importantly, the rest of the PRC1 target genes were reactivated to the same level in the OHT-treated PRC1$^{CPM}$ and PRC1$^{CKO}$ cells (Figure 7.12A and E). Therefore, through systematic analysis of the requirement for PRC1 catalysis in Polycomb system function, I discovered that PRC1 catalytic activity is the central determinant of Polycomb-mediated gene repression in mouse ES cells.

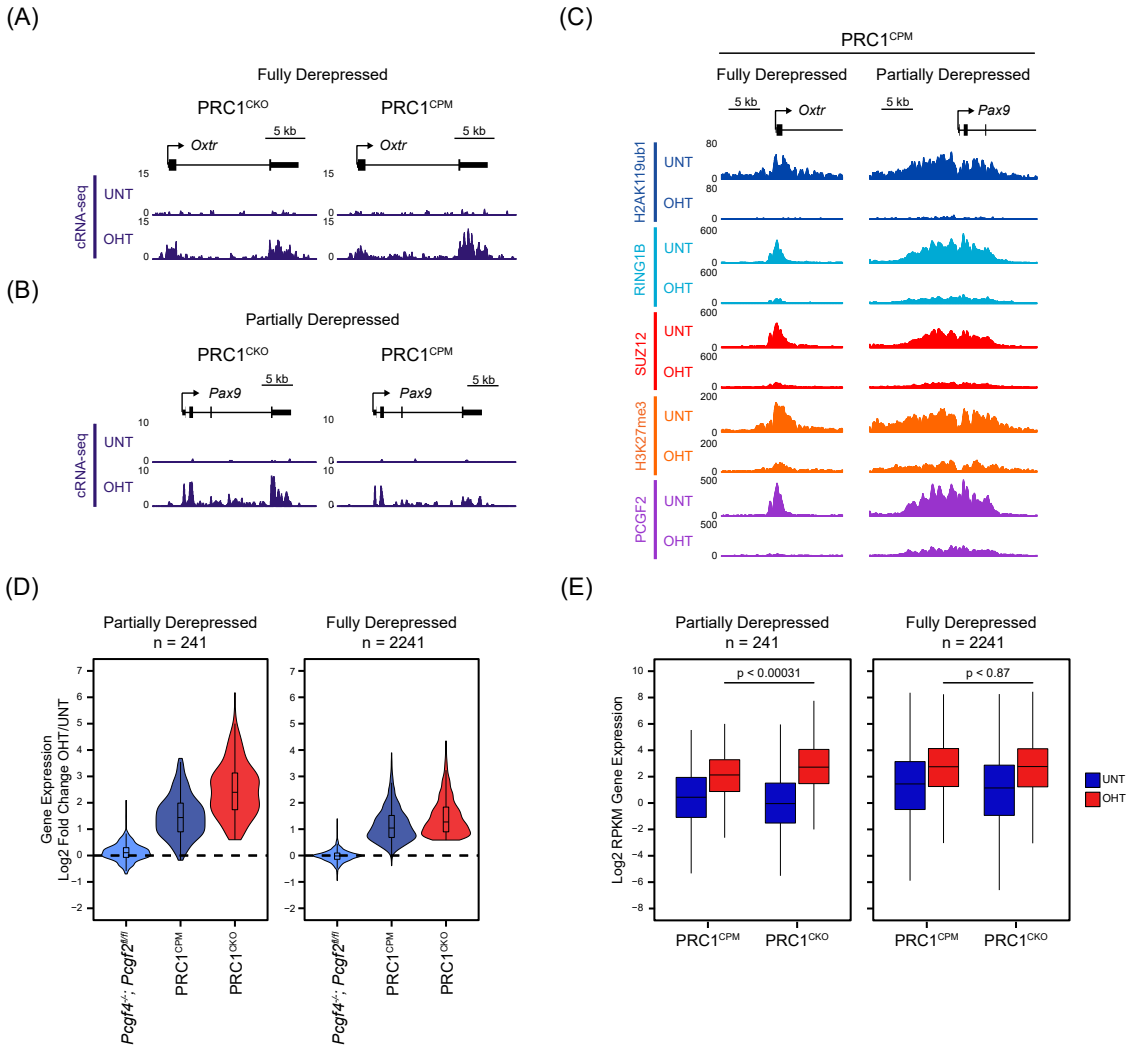



**Figure 7.12: A small subset of Polycomb target genes are not fully derepressed following removal of PRC1 catalytic activity**

(A) Genomic snapshot of a PRC1 target gene that is fully derepressed following loss of PRC1 catalysis as compared to a complete loss of PRC1, showing gene expression (cRNA-seq) in PRC1$^{CKO}$ and PRC1$^{CPM}$ cells (untreated and OHT-treated).

(B) Genomic snapshot of a PRC1 target gene that is partially derepressed following loss of PRC1 catalysis as compared to a complete loss of PRC1, showing gene expression (cRNA-seq) in PRC1$^{CKO}$ and PRC1$^{CPM}$ cells (untreated and OHT-treated).

(C) Genomic snapshots of the genes described in (A) and (B), showing cChIP-seq for PRC1 (RING1B, H2AK119ub1 and PCGF2) and PRC2 (SUZ12 and H3K27me3) in PRC1$^{CPM}$ cells (untreated and OHT-treated).

(D) Violin plot analysis comparing log2 fold changes in expression of PRC1 target genes which are partially (n = 241) or fully derepressed (n = 2241) following loss of PRC1 catalysis as compared to a complete loss of PRC1 in *Pcgf4*$^{-/-}$;*Pcgf2*$^{fl/fl}$, PRC1$^{CPM}$ and PRC1$^{CKO}$ cell lines.

(E) Box plot analysis of expression levels (cRNA-seq) for PRC1 target genes that are partially (n = 241) or fully derepressed (n = 2241) following loss of PRC1 catalysis as compared to a complete loss of PRC1 in untreated and OHT-treated PRC1$^{CPM}$ and PRC1$^{CKO}$ cells. P-values denote statistical significance calculated by a Wilcoxon rank sum test.

## 7.3 Discussion

Polycomb-mediated transcriptional silencing represents a central paradigm for chromatin-based gene regulation, yet whether histone-modifying activities of Polycomb complexes are required for

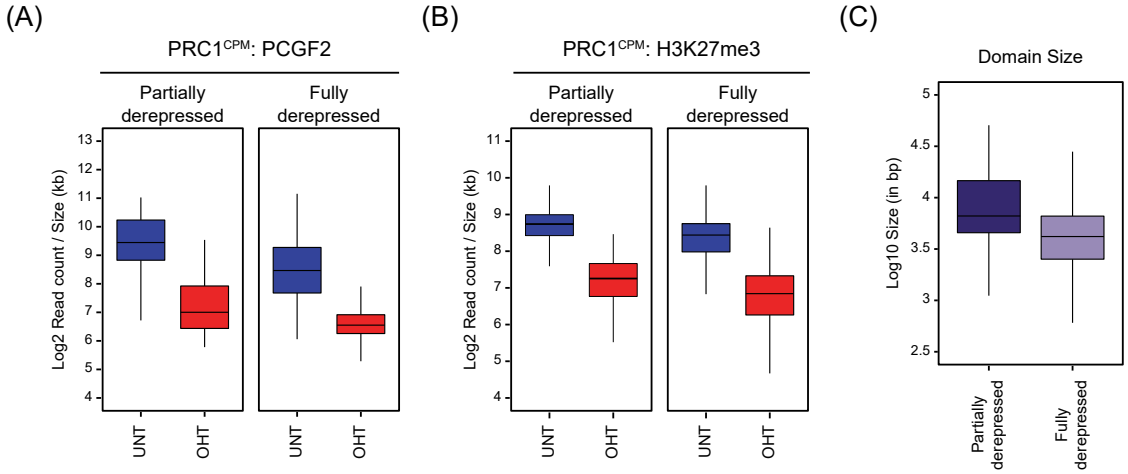



**Figure 7.13: Genes that are not fully derepressed following removal of PRC1 catalytic activity retain low-level binding of canonical PRC1**
(A) Box plot analysis of PCGF2 cChIP-seq signal in PRC1$^{CPM}$ cells (untreated and OHT-treated) at PRC1-bound sites associated with the promoters of PRC1 target genes which are either partially (n = 241) or fully derepressed (n = 2241) following loss of PRC1 catalysis as compared to a complete loss of PRC1.
(B) As in (A) for H3K27me3 cChIP-seq.
(C) Box plot analysis comparing sizes of PRC1-bound sites overlapping with the promoters of PRC1 target genes which are either partially (n = 241) or fully derepressed (n = 2241) following loss of PRC1 catalysis as compared to a complete loss of PRC1.

gene repression remains controversial. To address this conceptual gap in our understanding of Polycomb system function, in Chapter 7, I exploited a conditional genetic strategy to completely ablate PRC1 catalytic activity, while leaving PRC1 complexes intact. Using this approach, I demonstrated that catalysis by PRC1 facilitates PRC2 binding and deposition of H3K27me3 at Polycomb target sites, which leads to recruitment of canonical PRC1 complexes and formation of Polycomb chromatin domains. In addition, I showed that variant PRC1 complexes broadly occupy Polycomb target sites independently of PRC1 catalysis, indicating that their inherent DNA-binding activities may represent the primary mechanism for Polycomb target site selection. Finally, I established that PRC1 catalytic activity is absolutely essential for Polycomb-mediated gene repression. Together, these observations reveal a central requirement for PRC1 catalysis in Polycomb system function and provide new compelling evidence which supports a PRC1-initiated model for Polycomb chromatin domain formation and gene repression.

Previous attempts to evaluate the contribution of PRC1 E3 ligase activity to Polycomb system function have relied on a RING1B$^{I53A}$ mutation to inactivate PRC1 catalysis (Eskeland et al. 2010b; Endoh et al. 2012; Illingworth et al. 2015; Kundu et al. 2017; Cohen et al. 2018). When this mutation was introduced into mouse ES cells, only a small number of Polycomb target genes became derepressed (Illingworth et al. 2015) and higher-order chromatin interactions between

classical Polycomb target sites were unaffected (Eskeland et al. 2010b; Kundu et al. 2017). In addition, RING1B$^{I53A}$ mice exhibited a less severe embryonic phenotype than *Ring1b*$^{-/-}$ animals (Illingworth et al. 2015), suggesting that PRC1 catalytic activity is largely dispensable for both Polycomb-mediated gene repression and early embryonic development. However, RING1B$^{I53A}$ is a hypomorphic mutation, as it does not render PRC1 completely catalytically inactive *in vitro* (Buchwald et al. 2006, and unpublished observations from our lab) and a substantial amount of H2AK119ub1 is retained in the cells which carry this mutation (Illingworth et al. 2015; Kundu et al. 2017), even in the absence of RING1A (Tsuboi et al. 2018). This indicates that, in the previous experiments, PRC1 catalytic activity has not been completely ablated, strongly suggesting that the conclusions which have been made using the RING1B$^{I53A}$ mutation have to be revisited. Importantly, recent work from other groups, which utilised different RING1B mutations in the RING1A-null background to achieve a complete loss of PRC1 catalysis, concluded that PRC1 catalytic activity is essential for Polycomb-mediated gene repression, in agreement with my findings (Tsuboi et al. 2018; Tamburri et al. 2019).

While I have discovered that catalysis by PRC1 is central to Polycomb system function, my analysis has been limited to mouse ES cells, which represent a pluripotent stage in early embryonic development (Nichols and Smith 2010; Niwa 2010). Interestingly, a recent study has used a neuronal cell fate restriction model to examine the requirement for PRC1 catalytic activity in gene repression (Tsuboi et al. 2018). Based on the expression analysis of a small selection of neurogenic genes, this study concluded that PRC1 catalytic activity is required for gene repression during early stages of neuronal lineage differentiation, while in late-stage neural progenitor cells (NPCs), it is dispensable for PRC1-mediated transcriptional silencing (Tsuboi et al. 2018). In agreement with another study, which examined the role of PRC1 catalytic activity in skin development and adult stem cell specification (Cohen et al. 2018), this suggests that catalysis-independent mechanisms, which may rely on formation of higher-order chromatin structures by canonical PRC1, contribute to Polycomb target gene repression in more committed cell types. Moving forward, it will be important to use the point mutations which completely inactivate PRC1 catalysis in combination with quantitative genomic approaches to further explore how widely PRC1 catalytic activity contributes to Polycomb system function at different stages of mammalian development. Furthermore, these approaches should be also applied in other model organisms, such as *Drosophila*, where the contribution of PRC1 catalysis to gene regulation and development

has been proposed to be more limited (Pengelly et al. 2015).

Uncovering the mechanisms of Polycomb target site selection and Polycomb chromatin domain formation is crucial to understand the logic for how the Polycomb system functions. In mammals, non-methylated CpG islands (NMIs) have been proposed to define genomic occupancy of Polycomb repressive complexes, as binding of PRC1 and PRC2 is enriched at these elements (Mikkelsen et al. 2007; Ku et al. 2008) and ectopic integration of CG-rich DNA into the genome is sufficient to drive *de novo* formation of Polycomb chromatin domains (Mendenhall et al. 2010; Lynch et al. 2012; Jermann et al. 2014). Targeting of the Polycomb system to NMIs has been proposed to rely on direct recognition of non-methylated CpG-rich DNA by a ZF-CXXC domain-containing KDM2B subunit of PCGF1-PRC1 (Farcas et al. 2012; He et al. 2013a; Wu et al. 2013). Moreover, PCGF6-PRC1 can also contribute to shaping PRC1 occupancy at Polycomb target sites as it incorporates several DNA-binding modules with specificity for DNA motifs which are enriched at NMIs (Rozenberg et al. 2008; Jaeger et al. 2010; Deaton and Bird 2011; Bell and Vertino 2017; Endoh et al. 2017; Stielow et al. 2018; Scelfo et al. 2019). In Chapter 7, I found compelling evidence suggesting that the inherent DNA-binding activities of variant PCGF1- and PCGF6-PRC1 complexes enable them to broadly sample putative Polycomb target sites, and this represents the primary DNA-based mechanism for Polycomb target site selection in mammals.

In contrast to PCGF1 and PCGF6, RING1B occupancy in mouse ES cells is highly enriched at the promoters of lowly transcribed genes, and so are PRC2 and Polycomb-associated histone modifications. This suggests that additional mechanisms, which do not rely solely on recognition of DNA sequences, specify where Polycomb chromatin domains are formed. My findings in Chapter 7 support a PRC1-initiated model for Polycomb chromatin domain formation, whereby in the absence of active transcription, high-level deposition of H2AK119ub1 drives PRC2 binding and activity, as well as high-level enrichment of PRC1, which relies on H3K27me3-dependent binding of canonical PRC1 complexes. This is in line with previous studies showing that the Polycomb system responds to, as opposed to instructs, the transcriptional state of a gene (Hosogane et al. 2013; Riising et al. 2014; Berrozpe et al. 2017). Together, this suggests that Polycomb repressive complexes are not recruited to actively transcribed genes to turn off their transcription but instead accumulate at genes which are no longer expressed, possibly to protect them against stochastic gene activation events, which can compromise maintenance of cell type-specific transcriptional programmes. Importantly, in addition to the PRC1-initiated pathway for

Polycomb chromatin domain formation, alternative mechanisms may be responsible for targeting of Polycomb repressive complexes to NMIs. Indeed, PCL subunits of PRC2.1 complexes have been recently shown to have affinity for non-methylated CpG-rich DNA (Li et al. 2017; Perino et al. 2018), suggesting that PRC2 complexes may also broadly sample putative Polycomb target sites, independently of PRC1 and its catalytic activity.

Using the PRC1$^{CPM}$ system, I found that in the absence of chromatin-based stimulation, binding of catalytically inactive PRC1 complexes was uniform across PRC1 target gene promoters. Strikingly, similarly widespread distribution of RING1B across gene regulatory elements has been previously observed in several cancer cell lines and adult stem cells (Loubiere et al. 2016; Chan et al. 2018; Cohen et al. 2018). In these systems, RING1B occupancy did not correlate with H3K27me3 or gene expression and was enriched at promoters of actively transcribed genes (Loubiere et al. 2016; Chan et al. 2018; Cohen et al. 2018). Interestingly, the most abundant type of PRC1 in these cells are canonical PRC1 complexes, which have limited E3 ubiquitin ligase activity (Gao et al. 2012; Taherbhoy et al. 2015; Rose et al. 2016). This suggests that in specialised contexts during normal embryonic development or malignant transformations, the composition of PRC1 complexes and thus their catalytic activity may be regulated to support different modes of PRC1 targeting, which may underpin different modes of gene regulation by the Polycomb system.

Although here I have established that PRC1 catalytic activity is essential for Polycomb-mediated gene repression, the exact mechanism of how it contributes to transcriptional silencing remains largely unknown. Building upon the traditional hierarchical model, one may argue that the main role of PRC1-mediated H2AK119ub1 is to facilitate PRC2 binding, deposition of H3K27me3 and recruitment of canonical PRC1 complexes, which are required to compact chromatin and elicit gene repression. However, my findings in Chapter 3, together with earlier work from other labs (Riising et al. 2014), suggest that this is unlikely to be the case, since removal of canonical PRC1 or PRC2 has little effect on Polycomb-mediated gene repression in mouse ES cells. Another possibility is that H2AK119ub1 itself may directly or indirectly counteract the activity of transcriptional machinery (Stock et al. 2007; Zhou et al. 2008). In line with this idea, incorporation of H2AK119ub1 blocks transcription of chromatin templates *in vitro* (Nakagawa et al. 2008; Aihara et al. 2016), potentially by directly interfering with deposition of histone modifications associated with high level of transcriptional activity (Nakagawa et al. 2008; Yuan et al. 2013; Aihara et al. 2016) or by promoting binding of the linker histone H1, which can

inhibit transcription (Jason et al. 2004; Zhu et al. 2007; Fyodorov et al. 2017). Alternatively, H2AK119ub1 may indirectly mediate transcriptional silencing by recruiting H2AK119ub1-reader proteins that elicit gene repression (Richly et al. 2010; Kalb et al. 2014; Qin et al. 2015b; Cooper et al. 2016; Zhang et al. 2017; Ali et al. 2018). Finally, it cannot be ruled out that instead of H2AK119ub1, ubiquitination of histone H2A variants (Surface et al. 2016) or other non-histone substrates (Ben-Saadon et al. 2006) may be the driving force behind gene repression. Clearly, establishing the mechanisms by which PRC1 catalytic activity impinges on transcription should be a major focus of future studies. Nevertheless, my findings place PRC1 and its catalytic activity at the forefront of Polycomb-mediated gene repression, paving the way for a mechanistic understanding of Polycomb system function.

# 8

# CONCLUSIONS AND FUTURE DIRECTIONS

Polycomb repressive complex 1 (PRC1) and PRC2 are evolutionary conserved transcriptional repressors, which modify chromatin and control normal gene expression during development of multicellular organisms (reviewed in Schuettengruber et al. 2017; Vidal 2019). While a significant amount of effort has been placed in the past on understanding how Polycomb complexes are recruited to target sites, the molecular mechanisms of Polycomb-mediated transcriptional silencing have not been extensively investigated and thereby still remained poorly understood (reviewed in Blackledge et al. 2015; Laugesen et al. 2019; Mierlo et al. 2019; Vidal 2019). One of the major hurdles in uncovering central determinants of gene repression by the Polycomb system has been an immense variety of distinct Polycomb complexes with unique targeting modules and auxiliary subunits, some of which have activities that are independent of catalysis (reviewed in Bajusz et al. 2018; Mierlo et al. 2019). Moreover, accumulating evidence suggests that histone-modifying enzymes do no not always rely on their catalytic activity to regulate gene expression (Illingworth et al. 2015; Dorighi et al. 2017; Bao and Jia 2019; Turberfield et al. 2019). Therefore, whether histone modifications which are mediated by the Polycomb system are central to its function in gene silencing has also remained a matter of active debate in the field.

Here, I have examined the fundamental principles of Polycomb system function using mouse ES cells as a model system, since components of Polycomb complexes have been extensively characterised in these cells. To this end, I focused my analysis on PRC1, which has been shown to play a central role in Polycomb-mediated gene repression in mouse ES cells (Endoh et al. 2012; Riising et al. 2014), and developed a panel of conditional knockout cell lines, in which distinct PRC1 complexes can be removed individually or in combination. Together with quantitative genomic approaches, this enabled me to systematically dissect the role of individual PRC1 complexes in Polycomb system function. I discovered that, in contrast to prevailing views, canonical PRC1 is dispensable for silencing of Polycomb target genes (Chapter 3), while a remarkable degree of synergy between variant PRC1 complexes drives deposition of H2AK119ub1, formation of Polycomb chromatin domains and gene repression (Chapter 6). In

addition, I uncovered distinct contributions of individual variant PRC1 complexes to deposition of H2AK119ub1 and transcriptional silencing (Chapters 4 and 5). I found that PCGF1- and PCGF6-PRC1 play a major role in shaping punctate enrichment of H2AK119ub1 at Polycomb target sites and repress developmental genes driving specification of distinct cell lineages (Chapter 4). In contrast, PCGF3/5-PRC1 place H2AK119ub1 broadly throughout the genome and are employed in chromosome-wide gene silencing during X-inactivation (Chapter 5). Interestingly, I also demonstrated that the genome-wide blanket of H2AK119ub1 is kept in check by BAP1 deubiquitinase, which protects non-target genes from Polycomb-mediated repression (Chapter 5). Most importantly, all of my observations have pointed towards a central role of H2AK119ub1 in Polycomb system function. Using conditional genetic inactivation of PRC1 catalysis, I directly tested this hypothesis and discovered that PRC1 catalytic activity is absolutely essential for Polycomb chromatin domain formation and repression of Polycomb target genes (Chapter 7).

My findings revealed that catalytic activity of variant PRC1 complexes drives both Polycomb chromatin domain formation and gene repression (Chapters 6 and 7), yet how these two processes are related has remained unclear. Formation of Polycomb chromatin domains culminates in recruitment of canonical PRC1, which has been proposed to mediate compact higher-order chromatin structures at promoters of Polycomb target genes, rendering them inaccessible to the transcriptional machinery (Simon and Kingston 2013; Yu et al. 2019; Illingworth 2019). This reflects a classical view on Polycomb system function, according to which Polycomb chromatin domain formation represents the central mechanism behind gene repression. However, our lab has previously demonstrated that at least in mouse ES cells, Polycomb repressive complexes do not limit chromatin accessibility at target gene promoters (King et al. 2018). In addition to this, I have found that canonical PRC1 complexes contribute little to gene repression (Chapter 3). Although together these observations clearly indicate that silencing of Polycomb target genes is not driven by canonical PRC1, it still could rely on other factors which are associated with Polycomb chromatin domains. One possibility is that PRC2 and/or H3K27me3 may directly contribute to Polycomb-mediated gene repression. However, previous studies reported that constitutive deletion of PRC2 has no effect on self-renewal of mouse ES cells and causes minimal transcriptional defects (Pasini et al. 2007; Chamberlain et al. 2008; Shen et al. 2008; Riising et al. 2014). In agreement with later embryonic lethality of PRC2-null mice compared to PRC1-null mice (Faust et al. 1995; OCarroll et al. 2001; Voncken et al. 2003; Pasini et al. 2004; Posfai et al. 2012), this argues that

PRC2 and, as a consequence, hierarchical formation of Polycomb chromatin domains in general, is unlikely to drive Polycomb target gene repression at early developmental stages. Instead, Polycomb chromatin domains and their higher-order chromatin organisation may reinforce the fidelity of gene repression by the Polycomb system later in development, in agreement with an essential role of PRC2 and canonical PRC1 in lineage specification in mammals (reviewed in Sauvageau and Sauvageau 2010; Luis et al. 2012; Vidal and Starowicz 2017; Carlo et al. 2018).

Nevertheless, there still remain many open questions about the exact contribution of PRC2 to Polycomb system function, as earlier studies have suggested that PRC2 plays a major role in gene regulation and functions at least in part redundantly with PRC1 to support Polycomb-mediated gene repression (Boyer et al. 2006; Leeb et al. 2010). In line with this idea, in mouse ES cells, ablation of PRC2 in addition to PRC1 has been reported to further potentiate reactivation of Polycomb target genes caused by loss of PRC1 alone (Leeb et al. 2010; King et al. 2018). Similar conclusions have been made by genetic studies in *Drosophila*, which showed that while PRC1 plays a dominant role in control of early embryonic development with little compensation from PRC2 (Soto et al. 1995), functions of PRC1 and PRC2 appear to be largely redundant at later developmental stages (Beuchle et al. 2001). Furthermore, I demonstrated that a small amount of PRC2 remains bound to target sites in the absence of PRC1 and/or its catalytic activity (Chapter 7), suggesting that PRC2 can be recruited to Polycomb target genes and thus contribute to their silencing independently of PRC1. Therefore, it is possible that at target sites, which are shared between PRC1 and PRC2, PRC2 may at least partially compensate for PRC1 loss, while at unique sites, which are occupied by PRC2 alone, it may also have specialised functions in gene regulation (Boyer et al. 2006; Kahn et al. 2016). In conclusion, future work is clearly required to define the exact contribution of PRC2 complexes to Polycomb-mediated gene repression in ES cells and more committed cell types. To achieve this, conditional genetic perturbations combined with quantitative genomic approaches described here should be used to measure the primary transcriptional defects which arise following PRC2 removal.

Although functional consequences of Polycomb chromatin domain formation remain elusive, it is clear that recruitment of high levels of PRC1 and PRC2 does not play an instructive role in gene regulation (Chapter 3 and Riising et al. 2014). This challenges a model which has proposed that Polycomb complexes are recruited to target sites by sequence-specific DNA-binding factors to actively drive transcriptional silencing (Beisel and Paro 2011; Simon and Kingston 2013; Schwartz

and Pirrotta 2013). This model has been based on early genetic studies in *Drosophila* (Chan et al. 1994; Poux et al. 2001b; Mohd-Sarip et al. 2002; Sengupta et al. 2004) and more recent work in mammals showing that artificial tethering of various PcG proteins to reporter gene promoters leads to their repression (Hansen et al. 2008; Margueron et al. 2008; Gao et al. 2014; Bintu et al. 2016; Scelfo et al. 2019). In contrast, my observations provided compelling evidence for the responsive model of Polycomb system function, whereby high-level recruitment of Polycomb factors does not drive but follows transcriptional repression, which most likely results from loss of activation signals (Klose et al. 2013). In agreement with this idea, I demonstrated that binding of variant PRC1 complexes is not limited to the promoters of genes that they repress but instead is evident across all Polycomb target sites and does not always lead to transcriptional silencing (Chapters 6 and 7). In line with recent findings from our and other groups, my work suggests that variant PRC1 complexes dynamically sample putative Polycomb target sites through their inherent DNA-binding activities (Farcas et al. 2012; Wu et al. 2013; He et al. 2013a; Endoh et al. 2017; Stielow et al. 2018), enabling the Polycomb system to identify lowly transcribed or silenced genes at which to form Polycomb chromatin domains via the chromatin-based feedback mechanisms.

While here I have examined the principles of Polycomb recruitment and how it is linked to gene repression in mouse ES cells, it remains unclear whether my findings represent the universal mechanism for Polycomb targeting across different model systems. In *Drosophila*, early genetic studies have reported that artificial tethering of PcG proteins upstream of a tissue-specific enhancer does not overrule the enhancer-specific pattern of the transgene expression, despite uniform binding of PcG proteins to the tethering site in all tissues (Poux et al. 2001a). This suggests that, in *Drosophila*, recruitment of Polycomb complexes does not dictate the transcriptional output of target genes, in agreement with the responsive model. Furthermore, in *Arabidopsis*, Polycomb complexes have been shown to bind the FLC locus, which controls flowering, only following an initial wave of cold-induced transcriptional silencing, further supporting the idea that the Polycomb system facilitates but does not initiate gene repression (Swiezewski et al. 2009; Costa and Dean 2019). Nevertheless, considering that the responsive and instructive models are not mutually exclusive, it is possible that both of them may be utilised to drive Polycomb recruitment in different organisms and perhaps different developmental contexts. Elucidating the relative contribution of these proposed mechanisms and uncovering new determinants of Polycomb targeting in normal and disease states is an exciting avenue for future research.

If Polycomb complexes and their associated chromatin modifications accumulate at genes which are already minimally transcribed, then what exactly is the contribution of the Polycomb system to gene repression? One possibility is that Polycomb complexes contribute to stable maintenance of the silenced transcriptional states over long time-scales of mammalian development, during which cells are subjected to differentiation cues which vary in nature and magnitude. This requires faithfully propagating the repressed gene states through cell divisions and increasing the threshold for transcriptional activation to protect against stochastic transcription factor binding events. In agreement with these ideas, I demonstrated that, although a large number of Polycomb target genes are derepressed following removal of PRC1 in mouse ES cells, the magnitude of gene derepression is generally quite modest (Chapter 7). This indicates that at the majority of target genes, Polycomb counteracts only a basal level of transcriptional activation, which is insufficient to drive high-level cell type-specific gene expression. Moving forward, a direct comparison of the expression levels for lineage markers in PRC1-deficient ES cells and the relevant differentiated cell types is required to place the functional consequences of PRC1 ablation into the appropriate physiological context. Finally, in this work, I have utilised a quantitative RNA-seq approach to measure transcriptional defects following a series of PRC1 perturbations. However, an important limitation of this method is that it does not quantify the absolute levels of individual transcripts but instead estimates fold-changes in gene expression which are extremely sensitive to the initial gene abundance. Therefore, in future studies, it will be crucial to employ orthogonal approaches, such as single-molecule RNA FISH and single-particle RNA tracking, to measure the actual number of transcripts which are produced for Polycomb target genes in individual cells before and after removal of various components of the Polycomb system.

In mammals, a large repertoire of Polycomb repressive complexes gives rise to an extremely complex gene regulatory system with intricate feedback mechanisms and potential redundancy between individual components (reviewed in Chittock et al. 2017; Bajusz et al. 2018; Vidal 2019). However, previous studies primarily focused on examining the role of distinct Polycomb complexes in isolation and perhaps this is why they have failed to uncover the central determinants of Polycomb-mediated gene repression. In my work, I was able to overcome the Polycomb system complexity through combinatorial genetic perturbations of PCGF proteins, which define biochemically distinct PRC1 assemblies. Through this systematic analysis, I found that, while loss of PRC1 complexes individually could not fully compromise Polycomb-mediated gene

repression (Chapter 4), simultaneous removal of all variant PRC1 complexes was sufficient to largely phenocopy the Polycomb target gene reactivation following a complete loss of PRC1 (Chapter 6). These observations revealed two fundamentally important features of the mammalian Polycomb system. First, since distinct variant PRC1 complexes can compensate for each other and independently support some level of Polycomb target gene repression, this indicates that they perform at least partially overlapping functions, in contrast to what has been previously proposed (Gao et al. 2012; Sowpati et al. 2015). Therefore, despite diversification of the Polycomb system during vertebrate evolution, variant PRC1 complexes have maintained a certain degree of functional redundancy, which likely facilitates the robustness of lineage-specific transcriptional programmes, while still allowing the flexibility for the individual complexes to acquire specialised functions in gene regulation.

Second and most strikingly, I discovered that simultaneous removal of all variant PRC1 complexes had a greater-than-additive effect on expression of Polycomb target genes comparing to loss of the individual complexes, revealing a remarkable degree of synergy in the Polycomb system (Chapter 6). Combined ablation of variant PRC1 complexes also had a synergistic effect on H2AK119ub1 at Polycomb target gene promoters, which was strongly associated with silencing of these genes. This suggests that synergy between variant PRC1 complexes, which defines Polycomb-mediated gene repression, is likely driven by their shared capacity to catalyse H2AK119ub1. Therefore, it is possible that cooperation between variant PRC1 complexes enables them to place H2AK119ub1 at Polycomb target sites at sufficiently high levels, which above a certain threshold can support formation of Polycomb chromatin domains and interfere with transcriptional activation. In agreement with this idea, I demonstrated that PRC1 catalytic activity is central to gene repression by the Polycomb system (Chapter 7). However, the mechanistic understanding of how PRC1 catalysis counteracts the activity of RNA polymerase II transcriptional machinery is still missing. Moreover, accumulating evidence suggests that modifications of non-histone substrates, such as transcription factors and co-activator complexes, can play important roles in regulation of gene expression (Sterner and Berger 2000; Barrett and Wood 2008; Pradhan et al. 2009; Ardehali et al. 2017). Therefore, future work should establish whether H2AK119ub1 directly contributes to Polycomb-mediated gene silencing, and if it does, focus on identifying the mechanisms through which H2AK119ub1 may impinge on different stages of transcription.

Synergy between variant PRC1 complexes is an integral part of the Polycomb system function,

but what are the advantages of this built-in cooperativity for gene regulation? As my findings suggest, in synergistic repressive systems, a concerted action of multiple complexes is required to elicit a full degree of transcriptional silencing, and removal of a single component can lead to a disproportionately large reduction in the system repressive capacity. This implies that when a system which integrates synergy-based feedback is subjected to a robust activation signal, it would not respond gradually to the strength of the signal but instead undergo a switch-like transition from a silent to active transcriptional state, once a certain activation threshold is achieved. Remarkably, accumulating evidence suggests that many developmental genes are regulated through similar bistable switch mechanisms, which likely result from the opposing activities of Polycomb and Trithorax regulatory systems at these sites (Klose et al. 2013; Hoffmann et al. 2017; Sneppen and Ringrose 2019). The switch-like behaviour of developmental genes in response to differentiation cues is crucial for lineage specification, as it enables sharp transitions between cell type-specific transcriptional programmes, which ultimately orchestrate formation of highly-specialised anatomical structures. Furthermore, since the bistable systems can tolerate a certain level of stochastic perturbations before transitioning to another state, this provides the basis for the digital chromatin-based transcriptional memory, which supports robust maintenance of established cellular identities. In contrast to this, analogue regulatory systems, in which the transcriptional response is proportional to the activation signal, are sensitive to random fluctuations in the levels of transcriptional activators and therefore cannot be relied upon in the cell fate decision-making process.

Cooperativity between multiple DNA-binding transcription factors or transcriptional co-activators has been long recognised as a key determinant driving a robust on-off switch of gene activation (Carey 1998; Wang et al. 1999; Courey 2001; Struhl 2001; Lebrecht et al. 2005; Val et al. 2008; Heinz et al. 2010). Mechanistic understanding of this phenomenon comes primarily from the mathematical modelling of cooperative DNA binding of multiple transcription factors or a single factor to multiple binding sites (Veitia 2003; Šrobár 2008; Segal et al. 2008; Gutierrez et al. 2009; He et al. 2010; Gutierrez et al. 2012). Based on this theoretical analysis, synergistic binding of transcription factors has been predicted to promote switch-like transcriptional response. This has been later validated in a number of experimental studies which examined gene regulatory properties of the synthetic *cis*-regulatory DNA elements with a different number and/or position of transcription factor binding sites (Burz et al. 1998; Klinkenberg et al. 2006; Jolma et al. 2015). In

line with this, the same transcription factor can control both switch-like and graded transcriptional behaviour at different genomic loci, depending on whether the distribution of DNA-binding sites supports cooperative or independent binding (Klinkenberg et al. 2005; Giorgetti et al. 2010).

However, much less is known about synergistic interactions between chromatin-modifying transcriptional regulators and their contribution to control of transcriptional bistability. Previous theoretical studies have suggested that cooperativity between histone modifications or complexes that control their levels is required for bistable switch-like behaviour of gene regulatory systems (Kelemen et al. 2010; He et al. 2010; Binder et al. 2011), yet experimental validation of these predictions has been missing. Recent work has demonstrated that at least some chromatin-modifiers, including Polycomb complexes, define a switch-like transcriptional response of synthetic reporter genes (Bintu et al. 2016), but to what extent this relies on the chromatin-mediated synergy in these regulatory systems has remained unclear. Complementing these findings, my work here has uncovered extensive cooperativity within the Polycomb system, which is driven by PRC1 catalytic activity and possibly H2AK119ub1 and is central to Polycomb-mediated gene repression. Together, these observations provide compelling evidence that the chromatin-based transcriptional synergy may be an integral component of bistable gene regulatory systems which underpin normal embryonic development.

# A

# APPENDICES

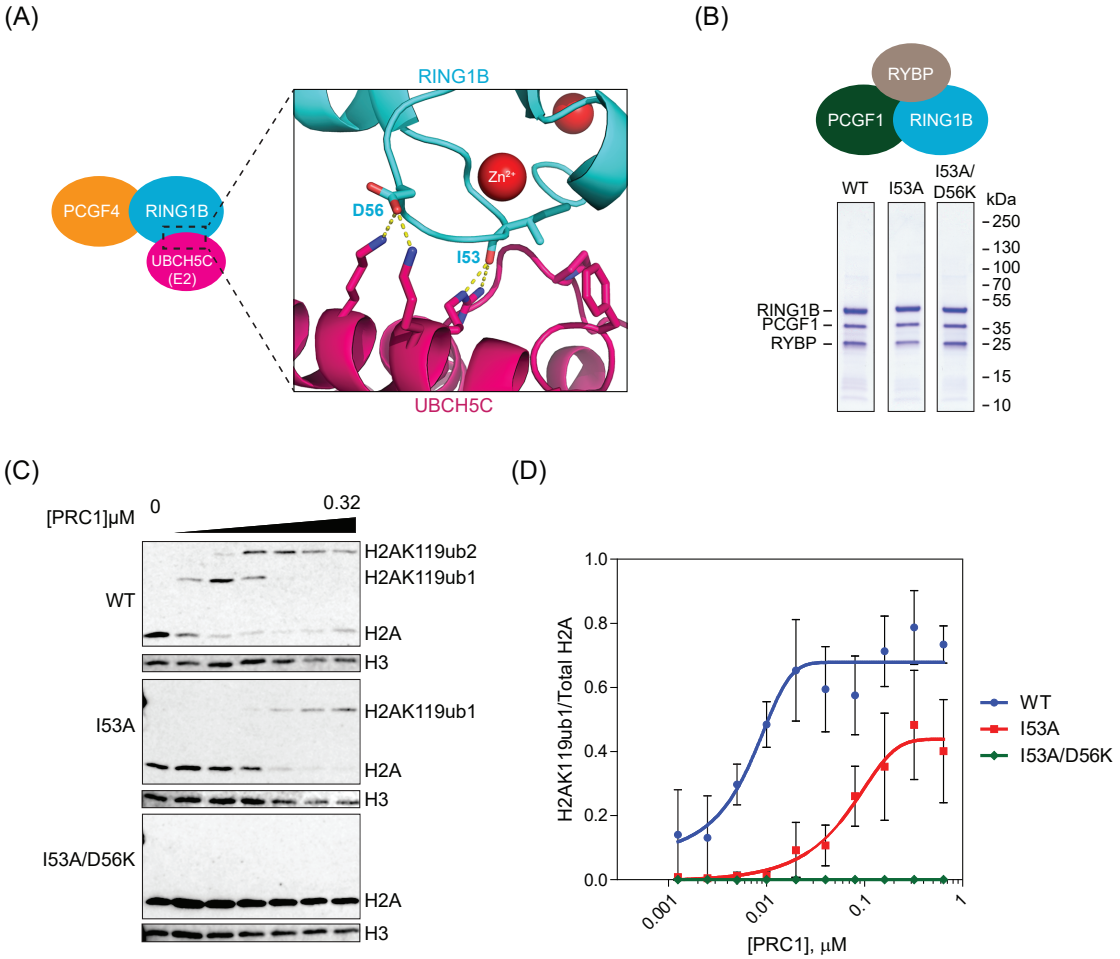



**Figure A.1: Establishing a catalytically inactive PRC1 complex *in vitro* (courtesy of Jessica Kelley)**
(A) A schematic of the crystal structure illustrating the RING1B-PCGF4 dimer bound to the E2 ubiquitin-conjugating enzyme (UBCH5C) (PDB ID: 3RPG) (Bentley et al. 2011). The key amino acids that mediate RING1B-E2 interaction are indicated.
(B) Coomassie-stained gels of affinity-purified RING1B-PCGF1-RYBP complexes incorporating wild-type or mutated forms of RING1B.
(C) *In vitro* E3 ubiquitin ligase assays carried out using reconstituted nucleosomes. Conversion of unmodified histone H2A to more slowly migrating ubiquitinated forms was measured by western blot with an H2A-specific antibody and an H3-specific antibody was used as a control.
(D) Quantification of the mean fraction of histone H2A converted to H2AK119ub1 across a wide range of PRC1 concentrations from *in vitro* E3 ubiquitin ligase assays in (C). Error bars show SEM (n≥2).

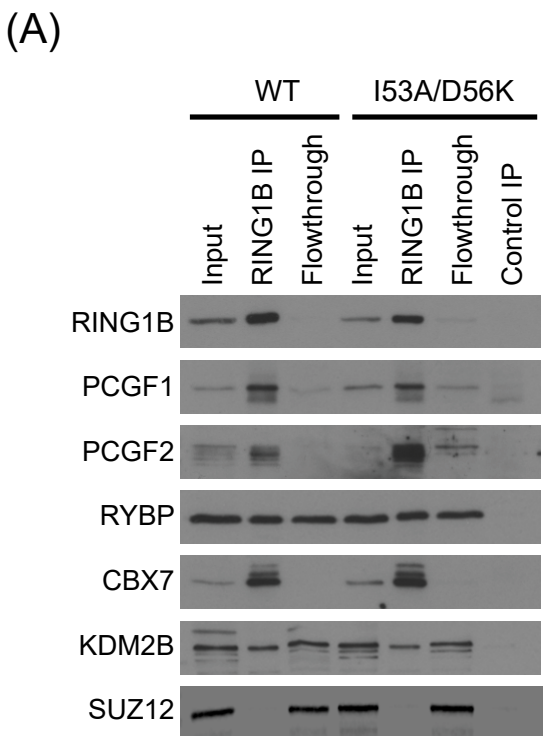



**Figure A.2: RING1B$^{I53A/D56K}$ forms stable PRC1 complexes (courtesy of Dr. Neil Blackledge)**
(A) Immunoprecipitation of RING1B from RING1B$^{WT}$ or RING1B$^{I53A/D56K}$ cells followed by western blot analysis for canonical PRC1 and variant PRC1 subunits as indicated. Western blot for SUZ12 (a core PRC2 subunit) was used as a negative control. For RING1B$^{I53A/D56K}$ cells, a control IP was performed with an isotype control antibody.

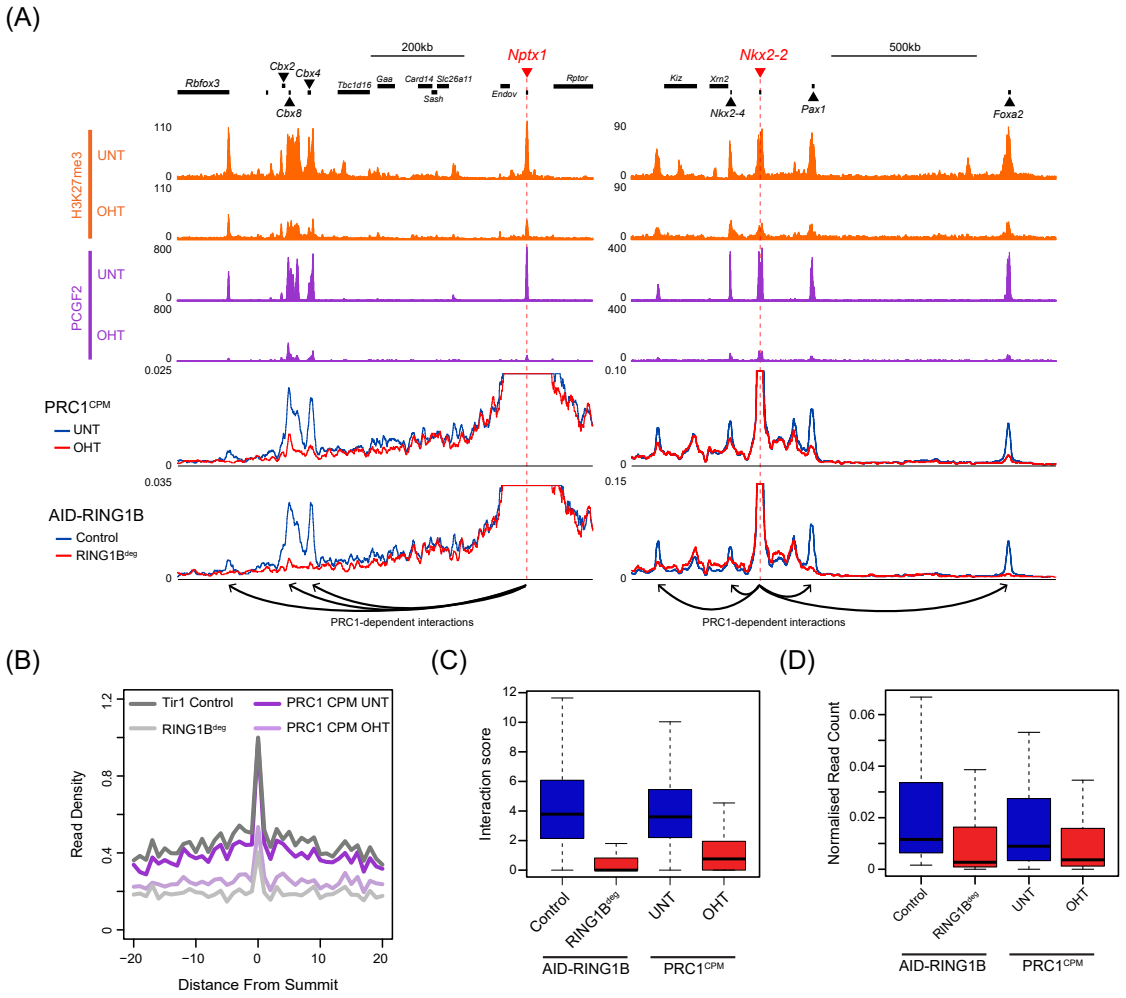



**Figure A.3: PRC1 catalytic activity is required to support long-range chromatin interactions between Polycomb target sites (courtesy of Dr. Angelika Feldmann)**

(A) Genomic snapshots of two regions around classical Polycomb target sites used as baits in CaptureC analysis (bait capture probe positions are marked by dashed red lines). H3K27me3 and PCGF2 cChIP-seq data in PRC1$^{CPM}$ cells (untreated and OHT-treated) are shown at the top. Below is the bait interaction landscape measured by CaptureC in PRC1$^{CPM}$ cells (untreated and OHT-treated) and auxin-treated *Ring1a*$^{-/-}$*;AID-RING1B* (RING1B$^{deg}$) cells (relative to the Control cell line with intact PRC1). Long-range PRC1-dependent interactions are illustrated by arrows.

(B) Metaplot analysis of the normalised mean read count from CaptureC for the interactions between the bait Polycomb target regions and other PCGF2 target sites (n = 130) in auxin-treated *Ring1a*$^{-/-}$*;AID-RING1B* (RING1B$^{deg}$) cells (relative to the Control cell line with intact PRC1) and PRC1$^{CPM}$ cells (untreated and OHT-treated).

(C) Box plot analysis of CHiCAGO scores for the interactions between the bait Polycomb target regions and other PCGF2 target sites (n = 130) in auxin-treated *Ring1a*$^{-/-}$*;AID-RING1B* (RING1B$^{deg}$) cells (relative to the Control cell line with intact PRC1) and PRC1$^{CPM}$ cells (untreated and OHT-treated).

(D) Box plot analysis of the normalised mean read count from CaptureC for the interactions between the bait Polycomb target regions and other PCGF2 target sites (n = 130) in auxin-treated *Ring1a*$^{-/-}$*;AID-RING1B* (RING1B$^{deg}$) cells (relative to the Control cell line with intact PRC1) and PRC1$^{CPM}$ cells (untreated and OHT-treated).

# BIBLIOGRAPHY


Adams, C. C. and Workman, J. L. (1993). "Nucleosome displacement in transcription". *Cell* 72: pp. 305–308.

Adelman, K. and Lis, J. T. (2012). "Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans". *Nature Reviews Genetics* 13: pp. 720–731.

Agger, K., Cloos, P. A. C., Christensen, J., Pasini, D., Rose, S., Rappsilber, J., Issaeva, I., Canaani, E., Salcini, A. E., and Helin, K. (2007). "UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development". *Nature* 449: pp. 731–734.

Ahmad, K. and Henikoff, S. (2002). "The Histone Variant H3.3 Marks Active Chromatin by Replication-Independent Nucleosome Assembly". *Molecular Cell* 9: pp. 1191–1200.

Aihara, H., Nakagawa, T., Mizusaki, H., Yoneda, M., Kato, M., Doiguchi, M., Imamura, Y., Higashi, M., Ikura, T., Hayashi, T., Kodama, Y., Oki, M., Nakayama, T., Cheung, E., Aburatani, H., Takayama, K. I., Koseki, H., Inoue, S., Takeshima, Y., and Ito, T. (2016). "Histone H2A T120 Phosphorylation Promotes Oncogenic Transformation via Upregulation of Cyclin D1". *Mol Cell* 64: pp. 176–188.

Akam, M. (1987). "The molecular basis for metameric pattern in the Drosophila embryo". *Development* 101: p. 1.

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

Akasaka, T., Lohuizen, M. van, Lugt, N. van der, Mizutani-Koseki, Y., Kanno, M., Taniguchi, M., Vidal, M., Alkema, M., Berns, A., and Koseki, H. (2001). "Mice doubly deficient for the Polycomb Group genes Mel18 and Bmi1 reveal synergy and requirement for maintenance but not initiation of Hox gene expression". *Development* 128: pp. 1587–97.

Alabert, C., Barth, T. K., Reverón-Gómez, N., Sidoli, S., Schmidt, A., Jensen, O. N., Imhof, A., and Groth, A. (2015). "Two distinct modes for propagation of histone PTMs across the cell cycle". *Genes & Development* 29: pp. 585–590.

Albright, S. C., Nelson, P. P., and Garrard, W. T. (1979). "Histone molar ratios among different electrophoretic forms of mono- and dinucleosomes". *J Biol Chem* 254: pp. 1065–73.

Albright, S. R. and Tjian, R. (2000). "TAFs revisited: more data reveal new twists and confirm old ideas". *Gene* 242: pp. 1–13.

Alekseyenko, A. A., Gorchakov, A. A., Kharchenko, P. V., and Kuroda, M. I. (2014). "Reciprocal interactions of human C10orf12 and C17orf96 with PRC2 revealed by BioTAP-XL cross-linking and affinity purification". *Proceedings of the National Academy of Sciences* 111: pp. 2488–2493.

Ali, M. A., Strickfaden, H., Lee, B. L., Spyracopoulos, L., and Hendzel, M. J. (2018). "RYBP Is a K63-Ubiquitin-Chain-Binding Protein that Inhibits Homologous Recombination Repair". *Cell Reports* 22: pp. 383–395.

Allan, J., Cowling, G. J., Harborne, N., Cattini, P., Craigie, R., and Gould, H. (1981). "Regulation of the higher-order structure of chromatin by histones H1 and H5." *J Cell Biol* 90: p. 279.

Allan, J., Hartman, P. G., Crane-Robinson, C., and Aviles, F. X. (1980). "The structure of histone H1 and its location in chromatin". *Nature* 288: pp. 675–679.

Allen, B. L. and Taatjes, D. J. (2015). "The Mediator complex: a central integrator of transcription". *Nature Reviews Molecular Cell Biology* 16: pp. 155–166.

Almeida, M., Pintacuda, G., Masui, O., Koseki, Y., Gdula, M., Cerase, A., Brown, D., Mould, A., Innocent, C., Nakayama, M., Schermelleh, L., Nesterova, T. B., Koseki, H., and Brockdorff, N. (2017). "PCGF3/5-PRC1 initiates Polycomb recruitment in X chromosome inactivation". *Science* 356: pp. 1081–1084.

An, W., Palhan, V. B., Karymov, M. A., Leuba, S. H., and Roeder, R. G. (2002). "Selective Requirements for Histone H3 and H4 N Termini in p300-Dependent Transcriptional Activation from Chromatin". *Molecular Cell* 9: pp. 811–821.

Aranda, S., Mas, G., and Croce, L. D. (2015). "Regulation of gene transcription by Polycomb proteins". *Science Advances* 1: e1500737.

Ardehali, M. B., Anselmo, A., Cochrane, J. C., Kundu, S., Sadreyev, R. I., and Kingston, R. E. (2017). "Polycomb Repressive Complex 2 Methylates Elongin A to Regulate Transcription". *Molecular Cell* 68: 872–884.e6.

Armstrong, J. A., Papoulas, O., Daubresse, G., Sperling, A. S., Lis, J. T., Scott, M. P., and Tamkun, J. W. (2002). "The Drosophila BRM complex facilitates global transcription by RNA polymerase II". *The EMBO Journal* 21: pp. 5245–5254.

Arnold, C. D., Gerlach, D., Stelzer, C., Boryn, L. M., Rath, M., and Stark, A. (2013). "Genome-Wide Quantitative Enhancer Activity Maps Identified by STARR-seq". *Science* 339: pp. 1074–1077.

Arnold, P., Scholer, A., Pachkov, M., Balwierz, P. J., Jorgensen, H., Stadler, M. B., Nimwegen, E. van, and Schubeler, D. (2012). "Modeling of epigenome dynamics identifies transcription factors that mediate Polycomb targeting". *Genome Research* 23: pp. 60–73.

Arrigoni, R., Alam, S. L., Wamstad, J. A., Bardwell, V. J., Sundquist, W. I., and Schreiber-Agus, N. (2006). "The Polycomb-associated protein Rybp is a ubiquitin binding protein". *FEBS Letters* 580: pp. 6233–6241.

Atsuta, T., Fujimura, S., Moriya, H., Vidal, M., Akasaka, T., and Koseki, H. (2001). "Production of monoclonal antibodies against mammalian Ring1B proteins". *Hybridoma* 20: pp. 43–6.

Audia, J. E. and Campbell, R. M. (2016). "Histone Modifications and Cancer". *Cold Spring Harbor Perspectives in Biology* 8: a019521.

Awad, S. and Hassan, A. H. (2008). "The Swi2/Snf2 Bromodomain Is Important for the Full Binding and Remodeling Activity of the SWI/SNF Complex on H3- and H4-acetylated Nucleosomes". *Annals of the New York Academy of Sciences* 1138: pp. 366–375.

Ayala Alonso, A. G. de, Gutierrez, L., Fritsch, C., Papp, B., Beuchle, D., and Müller, J. (2007). "A Genetic Screen Identifies Novel Polycomb Group Genes in Drosophila". *Genetics* 176: pp. 2099–2108.

Azzaz, A. M., Vitalini, M. W., Thomas, A. S., Price, J. P., Blacketer, M. J., Cryderman, D. E., Zirbel, L. N., Woodcock, C. L., Elcock, A. H., Wallrath, L. L., and Shogren-Knaak, M. A. (2014). "Human Heterochromatin Protein 1 alpha Promotes Nucleosome Associations That Drive Chromatin Condensation". *Journal of Biological Chemistry* 289: pp. 6850–6861.

Bajusz, I., Kovács, G., and Pirity, M. (2018). "From Flies to Mice: The Emerging Role of Non-Canonical PRC1 Members in Mammalian Development". *Epigenomes* 2: p. 4.

Balasubramani, A., Larjo, A., Bassein, J. A., Chang, X., Hastie, R. B., Togher, S. M., Lähdesmäki, H., and Rao, A. (2015). "Cancer-associated ASXL1 mutations may act as gain-of-function mutations of the ASXL1-BAP1 complex". *Nature Communications* 6.

Ballare, C., Lange, M., Lapinaite, A., Martin, G. M., Morey, L., Pascual, G., Liefke, R., Simon, B., Shi, Y., Gozani, O., Carlomagno, T., Benitah, S. A., and Croce, L. D. (2012). "Phf19 links methylated Lys36 of histone H3 to regulation of Polycomb activity". *Nature Structural & Molecular Biology* 19: pp. 1257–1265.

Banerji, J., Olson, L., and Schaffner, W. (1983). "A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes". *Cell* 33: pp. 729–740.

Banerji, J., Rusconi, S., and Schaffner, W. (1981). "Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences". *Cell* 27: pp. 299–308.

Bannister, A. J., Schneider, R., Myers, F. A., Thorne, A. W., Crane-Robinson, C., and Kouzarides, T. (2005). "Spatial Distribution of Di- and Tri-methyl Lysine 36 of Histone H3 at Active Genes". *Journal of Biological Chemistry* 280: pp. 17732–17736.

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

Bao, K. and Jia, S. (2019). "Noncatalytic Function of a JmjC Domain Protein Disrupts Heterochromatin". *Epigenetics Insights* 12: p. 251686571986224.

Bao, Y. and Shen, X. (2007). "INO80 subfamily of chromatin remodeling complexes". *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 618: pp. 18–29.

Barak, O., Lazzaro, M. A., Lane, W. S., Speicher, D. W., Picketts, D. J., and Shiekhattar, R. (2003). "Isolation of human NURF: a regulator of Engrailed gene expression". *The EMBO Journal* 22: pp. 6089–6100.

Barrett, R. M. and Wood, M. A. (2008). "Beyond transcription factors: The role of chromatin modifying enzymes in regulating transcription required for memory". *Learning & Memory* 15: pp. 460–467.

Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). "High-Resolution Profiling of Histone Methylations in the Human Genome". *Cell* 129: pp. 823–837.

Basta, J. and Rauchman, M. (2015). "The nucleosome remodeling and deacetylase complex in development and disease". *Translational Research* 165: pp. 36–47.

Bauer, M., Trupke, J., and Ringrose, L. (2015). "The quest for mammalian Polycomb response elements: are we there yet?" *Chromosoma* 125: pp. 471–496.

Beck, S., Lee, B.-K., Rhee, C., Song, J., Woo, A. J., and Kim, J. (2014). "CpG island-mediated global gene regulatory modes in mouse embryonic stem cells". *Nature Communications* 5.

Beh, L. Y., Colwell, L. J., and Francis, N. J. (2012). "A core subunit of Polycomb repressive complex 1 is broadly conserved in function but not primary sequence". *Proceedings of the National Academy of Sciences* 109: E1063–E1071.

Beisel, C. and Paro, R. (2011). "Silencing chromatin: comparing modes and mechanisms". *Nature Reviews Genetics* 12: pp. 123–135.

Bel, S., Core, N., Djabali, M., Kieboom, K., Lugt, N. Van der, Alkema, M., and Van Lohuizen, M. (1998). "Genetic interactions and dosage effects of Polycomb group genes in mice". *Development* 125: pp. 3543–3551.

Bell, J. S. K. and Vertino, P. M. (2017). "Orphan CpG islands define a novel class of highly active enhancers". *Epigenetics* 12: pp. 449–464.

Belle, J. I. and Nijnik, A. (2014). "H2A-DUBbing the mammalian epigenome: Expanding frontiers for histone H2A deubiquitinating enzymes in cell biology and physiology". *The International Journal of Biochemistry & Cell Biology* 50: pp. 161–174.

Belotserkovskaya, R., Oh, S., Bondarenko, V. A., Orphanides, G., Studitsky, V. M., and Reinberg, D. (2003). "FACT Facilitates Transcription-Dependent Nucleosome Alteration". *Science* 301: p. 1090.

Beltran, M., Tavares, M., Justin, N., Khandelwal, G., Ambrose, J., Foster, B. M., Worlock, K. B., Tvardovskiy, A., Kunzelmann, S., Herrero, J., Bartke, T., Gamblin, S. J., Wilson, J. R., and Jenner, R. G. (2019). "G-tract RNA removes Polycomb repressive complex 2 from genes". *Nature Structural & Molecular Biology*.

Beltran, M., Yates, C. M., Skalska, L., Dawson, M., Reis, F. P., Viiri, K., Fisher, C. L., Sibley, C. R., Foster, B. M., Bartke, T., Ule, J., and Jenner, R. G. (2016). "The interaction of PRC2 with RNA or chromatin is mutually antagonistic". *Genome Research* 26: pp. 896–907.

Ben-Saadon, R., Zaaroor, D., Ziv, T., and Ciechanover, A. (2006). "The Polycomb Protein Ring1B Generates Self Atypical Mixed Ubiquitin Chains Required for Its In Vitro Histone H2A Ligase Activity". *Molecular Cell* 24: pp. 701–711.

Benayoun, B. A., Pollina, E. A., Ucar, D., Mahmoudi, S., Karra, K., Wong, E. D., Devarajan, K., Daugherty, A. C., Kundaje, A. B., Mancini, E., Hitz, B. C., Gupta, R., Rando, T. A., Baker, J. C., Snyder, M. P., Cherry, J. M., and Brunet, A. (2014). "H3K4me3 Breadth Is Linked to Cell Identity and Transcriptional Consistency". *Cell* 158: pp. 673–688.

Benoist, C. and Chambon, P. (1981). "In vivo sequence requirements of the SV40 early promoter region". *Nature* 290: pp. 304–310.

Bentley, M. L., Corn, J. E., Dong, K. C., Phung, Q., Cheung, T. K., and Cochran, A. G. (2011). "Recognition of UbcH5c and the nucleosome by the Bmi1/Ring1b ubiquitin ligase complex". *The EMBO Journal* 30: pp. 3285–3297.

Beringer, M., Pisano, P., Carlo, V. D., Blanco, E., Chammas, P., Vizán, P., Gutiérrez, A., Aranda, S., Payer, B., Wierer, M., and Croce, L. D. (2016). "EPOP Functionally Links Elongin and Polycomb in Pluripotent Stem Cells". *Molecular Cell* 64: pp. 645–658.

Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., Jaenisch, R., Wagschal, A., Feil, R., Schreiber, S. L., and Lander, E. S. (2006a). "A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells". *Cell* 125: pp. 315–326.

Bernstein, E., Duncan, E. M., Masui, O., Gil, J., Heard, E., and Allis, C. D. (2006b). "Mouse Polycomb Proteins Bind Differentially to Methylated Histone H3 and RNA and Are Enriched in Facultative Heterochromatin". *Mol. Cell. Biol.* 26: p. 2560.

Berrozpe, G., Bryant, G. O., Warpinski, K., Spagna, D., Narayan, S., Shah, S., and Ptashne, M. (2017). "Polycomb Responds to Low Levels of Transcription". *Cell Reports* 20: pp. 785–793.

Berry, S., Dean, C., and Howard, M. (2017). "Slow Chromatin Dynamics Allow Polycomb Target Genes to Filter Fluctuations in Transcription Factor Activity". *Cell Systems* 4: 445–457.e8.

Beuchle, D., Struhl, G., and Muller, J. (2001). "Polycomb group proteins and heritable silencing of Drosophila Hox genes". *Development* 128: p. 993.

Bezsonova, I., Walker, J. R., Bacik, J. P., Duan, S., Dhe-Paganon, S., and Arrowsmith, C. H. (2009). "Ring1B Contains a Ubiquitin-Like Docking Module for Interaction with Cbx Proteins," *Biochemistry* 48: pp. 10542–10548.

Bickmore, W. A., M. A. H. Y., N. L., and Chambeyron, S. (2004). "Do Higher-Order Chromatin Structure and Nuclear Reorganization Play a Role in Regulating Hox Gene Expression during Development?" *Cold Spring Harbor Symposia on Quantitative Biology* 69: pp. 251–258.

Bickmore, W. (2019). "Higher-order chromatin folding and gene regulation". *The FASEB Journal* 27: pp. 456.1–456.1.

Binder, H., Steiner, L., Rohlf, T., Prohaska, S., and Galle, J. (2011). "Transcriptional memory emerges from cooperative histone modifications". *Nature Precedings*.

Bintu, L., Yong, J., Antebi, Y. E., McCue, K., Kazuki, Y., Uno, N., Oshimura, M., and Elowitz, M. B. (2016). "Dynamics of epigenetic regulation at the single-cell level". *Science* 351: pp. 720–724.

Biswas, D., Dutta-Biswas, R., Mitra, D., Shibata, Y., Strahl, B. D., Formosa, T., and Stillman, D. J. (2006). "Opposing roles for Set2 and yFACT in regulating TBP binding at promoters". *The EMBO Journal* 25: pp. 4479–4489.

Blackledge, N. P., Farcas, A. M., Kondo, T., King, H. W., McGouran, J. F., Hanssen, L. L., Ito, S., Cooper, S., Kondo, K., Koseki, Y., Ishikura, T., Long, H. K., Sheahan, T. W., Brockdorff, N., Kessler, B. M., Koseki, H., and Klose, R. J. (2014). "Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation". *Cell* 157: pp. 1445–59.

Blackledge, N. P., Rose, N. R., and Klose, R. J. (2015). "Targeting Polycomb systems to regulate gene expression: modifications to a complex story". *Nat Rev Mol Cell Biol* 16: pp. 643–649.

Blackledge, N. P. and Klose, R. (2011). "CpG island chromatin". *Epigenetics* 6: pp. 147–152.

Blackledge, N. P., Zhou, J. C., Tolstorukov, M. Y., Farcas, A. M., Park, P. J., and Klose, R. J. (2010). "CpG Islands Recruit a Histone H3 Lysine 36 Demethylase". *Molecular Cell* 38: pp. 179–190.

Blair, S. S. (2008). "Segmentation in animals". *Current Biology* 18: R991–R995.

Blastyak, A., Mishra, R. K., Karch, F., and Gyurkovics, H. (2006). "Efficient and Specific Targeting of Polycomb Group Proteins Requires Cooperative Interaction between Grainyhead and Pleiohomeotic". *Molecular and Cellular Biology* 26: pp. 1434–1444.

Boettiger, A. N., Bintu, B., Moffitt, J. R., Wang, S., Beliveau, B. J., Fudenberg, G., Imakaev, M., Mirny, L. A., Wu, C. T., and Zhuang, X. (2016). "Super-resolution imaging reveals distinct chromatin folding for different epigenetic states". *Nature* 529: pp. 418–22.

Bonev, B., Mendelson Cohen, N., Szabo, Q., Fritsch, L., Papadopoulos, G. L., Lubling, Y., Xu, X., Lv, X., Hugnot, J. P., Tanay, A., and Cavalli, G. (2017). "Multiscale 3D Genome Rewiring during Mouse Neural Development". *Cell* 171: 557–572 e24.

Bonhoure, N., Bounova, G., Bernasconi, D., Praz, V., Lammers, F., Canella, D., Willis, I. M., Herr, W., Hernandez, N., Delorenzi, M., and Cycli, X. C. (2014). "Quantifying ChIP-seq data: a spiking method providing an internal reference for sample-to-sample normalization". *Genome Res* 24: pp. 1157–68.

Bottomley, M. J. (2004). "Structures of protein domains that create or recognize histone modifications". *EMBO reports* 5: pp. 464–469.

Boyer, L. A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L. A., Lee, T. I., Levine, S. S., Wernig, M., Tajonar, A., Ray, M. K., Bell, G. W., Otte, A. P., Vidal, M., Gifford, D. K., Young, R. A., and Jaenisch, R. (2006). "Polycomb complexes repress developmental regulators in murine embryonic stem cells". *Nature* 441: pp. 349–53.

Bracken, A. P., Brien, G. L., and Verrijzer, C. P. (2019). "Dangerous liaisons: interplay between SWI/SNF, NuRD, and Polycomb in chromatin regulation and cancer". *Genes & Development*.

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

Bravo, M., Nicolini, F., Starowicz, K., Barroso, S., Cales, C., Aguilera, A., and Vidal, M. (2015). "Polycomb RING1A- and RING1B-dependent histone H2A monoubiquitylation at pericentromeric regions promotes S-phase progression". *J Cell Sci* 128: pp. 3660–71.

Breiling, A., Turner, B. M., Bianchi, M. E., and Orlando, V. (2001). "General transcription factors bind promoters repressed by Polycomb group proteins". *Nature* 412: pp. 651–655.

Brien, G. L., Gambero, G., OConnell, D. J., Jerman, E., Turner, S. A., Egan, C. M., Dunne, E. J., Jurgens, M. C., Wynne, K., Piao, L., Lohan, A. J., Ferguson, N., Shi, X., Sinha, K. M., Loftus, B. J., Cagney, G., and Bracken, A. P. (2012). "Polycomb PHF19 binds H3K36me3 and recruits PRC2 and demethylase NO66 to embryonic stem cell genes during differentiation". *Nature Structural & Molecular Biology* 19: pp. 1273–1281.

Brockdorff, N. (2013). "Noncoding RNA and Polycomb recruitment". *RNA* 19: pp. 429–442.

Brockdorff, N. (2017). "Polycomb complexes in X chromosome inactivation". *Philosophical Transactions of the Royal Society B: Biological Sciences* 372: p. 20170021.

Brookes, E., Santiago, I. de, Hebenstreit, D., Morris, K. J., Carroll, T., Xie, S. Q., Stock, J. K., Heidemann, M., Eick, D., Nozaki, N., Kimura, H., Ragoussis, J., Teichmann, S. A., and Pombo, A. (2012). "Polycomb Associates Genome-wide with a Specific RNA Polymerase II Variant, and Regulates Metabolic Genes in ESCs". *Cell Stem Cell* 10: pp. 157–170.

Brown, J. L. (2003). "The Drosophila pho-like gene encodes a YY1-related DNA binding protein that is redundant with pleiohomeotic in homeotic gene silencing". *Development* 130: pp. 285–294.

Brown, J. L., Grau, D. J., DeVido, S. K., and Kassis, J. A. (2005). "An Sp1/KLF binding site is important for the activity of a Polycomb group response element from the Drosophila engrailed gene". *Nucleic Acids Research* 33: pp. 5181–5189.

Brown, J., Mucci, D., Whiteley, M., Dirksen, M.-L., and Kassis, J. A. (1998). "The Drosophila Polycomb Group Gene pleiohomeotic Encodes a DNA Binding Protein with Homology to the Transcription Factor YY1". *Molecular Cell* 1: pp. 1057–1064.

Brownell, J. E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D. G., Roth, S. Y., and Allis, C. D. (1996). "Tetrahymena Histone Acetyltransferase A: A Homolog to Yeast Gcn5 Linking Histone Acetylation to Gene Activation". *Cell* 84: pp. 843–851.

Buchenau, P., Hodgson, J., Strutt, H., and Arndt-Jovin, D. J. (1998). "The Distribution of Polycomb-Group Proteins During Cell Division and Development in DrosophilaEmbryos: Impact on Models for Silencing". *The Journal of Cell Biology* 141: pp. 469–481.

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

Buratowski, S., Hahn, S., Guarente, L., and Sharp, P. A. (1989). "Five intermediate complexes in transcription initiation by RNA polymerase II". *Cell* 56: pp. 549–561.

Burz, D. S., Rivera-Pomar, R., Jackle, H., and Hanes, S. D. (1998). "Cooperative DNA-binding by Bicoid provides a mechanism for threshold-dependent gene activation in the Drosophila embryo". *The EMBO Journal* 17: pp. 5998–6009.

Butler, J. E. F. and Kadonaga, J. T. (Oct. 2001). "Enhancer-promoter specificity mediated by DPE or TATA core promoter motifs". *Genes & Development* 15: pp. 2515–2519.

Butler, J. E. F. and Kadonaga, J. T. (Oct. 2002). "The RNA polymerase II core promoter: a key component in the regulation of gene expression". *Genes & Development* 16: pp. 2583–2592.

Cai, L., Rothbart, S. B., Lu, R., Xu, B., Chen, W.-Y., Tripathy, A., Rockowitz, S., Zheng, D., Patel, D. J., Allis, C. D., Strahl, B. D., Song, J., and Wang, G. G. (2013). "An H3K36 Methylation-Engaging Tudor Motif of Polycomb-like Proteins Mediates PRC2 Complex Targeting". *Molecular Cell* 49: pp. 571–582.

Calabrese, J. M., Sun, W., Song, L., Mugford, J. W., Williams, L., Yee, D., Starmer, J., Mieczkowski, P., Crawford, G. E., and Magnuson, T. (2012). "Site-Specific Silencing of Regulatory Elements as a Mechanism of X Inactivation". *Cell* 151: pp. 951–963.

Campagne, A., Lee, M.-K., Zielinski, D., Michaud, A., Corre, S. L., Dingli, F., Chen, H., Shahidian, L. Z., Vassilev, I., Servant, N., Loew, D., Pasmant, E., Postel-Vinay, S., Wassef, M., and Margueron, R. (2019). "BAP1 complex promotes transcription by opposing PRC1-mediated H2A ubiquitylation". *Nature Communications* 10.

Canzio, D., Larson, A., and Narlikar, G. J. (2014). "Mechanisms of functional promiscuity by HP1 proteins". *Trends in Cell Biology* 24: pp. 377–386.

Cao, K., Lailler, N., Zhang, Y., Kumar, A., Uppal, K., Liu, Z., Lee, E. K., Wu, H., Medrzycki, M., Pan, C., Ho, P.-Y., Cooper, G. P., Dong, X., Bock, C., Bouhassira, E. E., and Fan, Y. (2013). "High-Resolution Mapping of H1 Linker Histone Variants in Embryonic Stem Cells". *PLoS Genetics* 9. Ed. by W. A. Bickmore: e1003417.

Cao, Q., Wang, X., Zhao, M., Yang, R., Malik, R., Qiao, Y., Poliakov, A., Yocum, A. K., Li, Y., Chen, W., Cao, X., Jiang, X., Dahiya, A., Harris, C., Feng, F. Y., Kalantry, S., Qin, Z. S., Dhanasekaran, S. M., and Chinnaiyan, A. M. (2014). "The central role of EED in the orchestration of polycomb group complexes". *Nature Communications* 5.

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R. S., and Zhang, Y. (2002). "Role of histone H3 lysine 27 methylation in Polycomb-group silencing". *Science* 298: pp. 1039–43.

Cao, R., Tsukada, Y. ichi, and Zhang, Y. (2005). "Role of Bmi-1 and Ring1A in H2A Ubiquitylation and Hox Gene Silencing". *Molecular Cell* 20: pp. 845–854.

Cao, R. and Zhang, Y. (2004). "SUZ12 Is Required for Both the Histone Methyltransferase Activity and the Silencing Function of the EED-EZH2 Complex". *Molecular Cell* 15: pp. 57–67.

Carbone, M., Yang, H., Pass, H. I., Krausz, T., Testa, J. R., and Gaudino, G. (2013). "BAP1 and cancer". *Nature Reviews Cancer* 13: pp. 153–159.

Cardozo, T. and Pagano, M. (2004). "The SCF ubiquitin ligase: insights into a molecular machine". *Nature Reviews Molecular Cell Biology* 5: pp. 739–751.

Carey, M. (1998). "The Enhanceosome and Transcriptional Synergy". *Cell* 92: pp. 5–8.

Carey, M., Li, B., and Workman, J. L. (2006). "RSC Exploits Histone Acetylation to Abrogate the Nucleosomal Block to RNA Polymerase II Elongation". *Molecular Cell* 24: pp. 481–487.

Carlo, V. D., Mocavini, I., and Croce, L. D. (2018). "Polycomb complexes in normal and malignant hematopoiesis". *The Journal of Cell Biology* 218: pp. 55–69.

Carrozza, M. J., Li, B., Florens, L., Suganuma, T., Swanson, S. K., Lee, K. K., Shia, W.-J., Anderson, S., Yates, J., Washburn, M. P., and Workman, J. L. (2005). "Histone H3 Methylation by Set2 Directs Deacetylation of Coding Regions by Rpd3S to Suppress Spurious Intragenic Transcription". *Cell* 123: pp. 581–592.

Casanova, M., Preissner, T., Cerase, A., Poot, R., Yamada, D., Li, X., Appanah, R., Bezstarosti, K., Demmers, J., Koseki, H., and Brockdorff, N. (2011). "Polycomblike 2 facilitates the recruitment of PRC2 Polycomb group complexes to the inactive X chromosome and to target loci in embryonic stem cells". *Development* 138: pp. 1471–1482.

Cerbo, V. D., Mohn, F., Ryan, D. P., Montellier, E., Kacem, S., Tropberger, P., Kallis, E., Holzner, M., Hoerner, L., Feldmann, A., Richter, F. M., Bannister, A. J., Mittler, G., Michaelis, J., Khochbin, S., Feil, R., Schuebeler, D., Owen-Hughes, T., Daujat, S., and Schneider, R. (2014). "Acetylation of histone H3 at lysine 64 regulates nucleosome dynamics and facilitates transcription". *eLife* 3.

Chamberlain, S. J., Yee, D., and Magnuson, T. (2008). "Polycomb Repressive Complex 2 Is Dispensable for Maintenance of Embryonic Stem Cell Pluripotency". *Stem Cells* 26: pp. 1496–1505.

Chan, C., Rastelli, L., and Pirrotta, V. (1994). "A Polycomb response element in the Ubx gene that determines an epigenetically inherited state of repression." *The EMBO Journal* 13: pp. 2553–2564.

Chan, H. L., Beckedorff, F., Zhang, Y., Garcia-Huidobro, J., Jiang, H., Colaprico, A., Bilbao, D., Figueroa, M. E., LaCava, J., Shiekhattar, R., and Morey, L. (2018). "Polycomb complexes associate with enhancers and promote oncogenic transcriptional programs in cancer through multiple mechanisms". *Nature Communications* 9.

Chen, P., Zhao, J., Wang, Y., Wang, M., Long, H., Liang, D., Huang, L., Wen, Z., Li, W., Li, X., Feng, H., Zhao, H., Zhu, P., Li, M., Wang, Q. f., and Li, G. (2013). "H3.3 actively marks enhancers and primes gene transcription via opening higher-ordered chromatin". *Genes & Development* 27: pp. 2109–2124.

Cheutin, T. and Cavalli, G. (2012). "Progressive Polycomb Assembly on H3K27me3 Compartments Generates Polycomb Bodies with Developmentally Regulated Motion". *PLoS Genetics* 8. Ed. by N. Perrimon: e1002465.

Chi, P., Allis, C. D., and Wang, G. G. (2010). "Covalent histone modifications - miswritten, misinterpreted and mis-erased in human cancers". *Nature Reviews Cancer* 10: pp. 457–469.

Chiang, A., OConnor, M. B., Paro, R., Simon, J., and Bender, W. (1995). "Discrete Polycomb-binding sites in each parasegmental domain of the bithorax complex". *Development* 121: p. 1681.

Chittock, E. C., Latwiel, S., Miller, T. C., and Müller, C. W. (2017). "Molecular architecture of polycomb repressive complexes". *Biochemical Society Transactions* 45: pp. 193–205.

Cho, Y.-W., Hong, T., Hong, S., Guo, H., Yu, H., Kim, D., Guszczynski, T., Dressler, G. R., Copeland, T. D., Kalkum, M., and Ge, K. (2007). "PTIP Associates with MLL3- and MLL4-containing Histone H3 Lysine 4 Methyltransferase Complex". *Journal of Biological Chemistry* 282: pp. 20395–20406.

Choi, J., Bachmann, A. L., Tauscher, K., Benda, C., Fierz, B., and Müller, J. (2017). "DNA binding by PHF1 prolongs PRC2 residence time on chromatin and thereby promotes H3K27 methylation". *Nature Structural & Molecular Biology* 24: pp. 1039–1047.

Chow, C.-M., Georgiou, A., Szutorisz, H., Silva, A. M. e, Pombo, A., Barahona, I., Dargelos, E., Canzonetta, C., and Dillon, N. (2005). "Variant histone H3.3 marks promoters of transcriptionally active genes during mammalian cell division". *EMBO reports* 6: pp. 354–360.

Chu, F., Nusinow, D. A., Chalkley, R. J., Plath, K., Panning, B., and Burlingame, A. L. (2006). "Mapping Post-translational Modifications of the Histone Variant MacroH2A1 Using Tandem Mass Spectrometry". *Mol Cell Proteomics* 5: p. 194.

Cifuentes-Rojas, C., Hernandez, A. J., Sarma, K., and Lee, J. T. (2014). "Regulatory Interactions between RNA and Polycomb Repressive Complex 2". *Molecular Cell* 55: pp. 171–185.

Clapier, C. R. and Cairns, B. R. (2009). "The Biology of Chromatin Remodeling Complexes". *Annual Review of Biochemistry* 78: pp. 273–304.

Clausell, J., Happel, N., Hale, T. K., Doenecke, D., and Beato, M. (2009). "Histone H1 Subtypes Differentially Modulate Chromatin Condensation without Preventing ATP-Dependent Remodeling by SWI/SNF or NURF". *PLOS ONE* 4: e0007243.

Clements, A., Poux, A. N., Lo, W.-S., Pillus, L., Berger, S. L., and Marmorstein, R. (2003). "Structural Basis for Histone and Phosphohistone Binding by the GCN5 Histone Acetyltransferase". *Molecular Cell* 12: pp. 461–473.

Cohen, I., Zhao, D., Bar, C., Valdes, V. J., Dauber-Decker, K. L., Nguyen, M. B., Nakayama, M., Rendl, M., Bickmore, W. A., Koseki, H., Zheng, D., and Ezhkova, E. (2018). "PRC1 Fine-tunes Gene Repression and Activation to Safeguard Skin Development and Stem Cell Specification". *Cell Stem Cell* 22: 726–739 e7.

Colognori, D., Sunwoo, H., Kriz, A. J., Wang, C.-Y., and Lee, J. T. (2019). "Xist Deletional Analysis Reveals an Interdependency between Xist RNA and Polycomb Complexes for Spreading along the Inactive X". *Molecular Cell* 74: 101–117.e10.

Consortium, E. N. C. O. D. E. (2012). "An integrated encyclopedia of DNA elements in the human genome". *Nature* 489: p. 57.

Conway, E., Jerman, E., Healy, E., et al. (2018). "A Family of Vertebrate-Specific Polycombs Encoded by the LCOR/LCORL Genes Balance PRC2 Subtype Activities". *Mol Cell* 70: 408–421 e8.

Cooper, S., Dienstbier, M., Hassan, R., Schermelleh, L., Sharif, J., Blackledge, N. P., De Marco, V., Elderkin, S., Koseki, H., Klose, R., Heger, A., and Brockdorff, N. (2014). "Targeting polycomb to pericentric heterochromatin in embryonic stem cells reveals a role for H2AK119u1 in PRC2 recruitment". *Cell Rep* 7: pp. 1456–1470.

Cooper, S., Grijzenhout, A., Underwood, E., Ancelin, K., Zhang, T., Nesterova, T. B., Anil-Kirmizitas, B., Bassett, A., Kooistra, S. M., Agger, K., Helin, K., Heard, E., and Brockdorff, N. (2016). "Jarid2 binds mono-ubiquitylated H2A lysine 119 to mediate crosstalk between Polycomb complexes PRC1 and PRC2". *Nat Commun* 7: p. 13661.

Core, L. J., Martins, A. L., Danko, C. G., Waters, C. T., Siepel, A., and Lis, J. T. (2014). "Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers". *Nature Genetics* 46: pp. 1311–1320.

Core, N., Bel, S., Gaunt, S. J., Aurrand-Lions, M., Pearce, J., Fisher, A., and Djabali, M. (1997). "Altered cellular proliferation and mesoderm patterning in Polycomb-M33-deficient mice". *Development* 124: pp. 721–9.

Corona, D. F. and Tamkun, J. W. (2004). "Multiple roles for ISWI in transcription, chromosome organization and DNA replication". *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* 1677: pp. 113–119.

Costa, S. and Dean, C. (2019). "Storing memories: the distinct phases of Polycomb-mediated silencing of Arabidopsis FLC". *Biochemical Society Transactions* 47: pp. 1187–1196.

Cote, J, Quinn, J, Workman, J., and Peterson, C. (1994). "Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex". *Science* 265: pp. 53–60.

Courey, A. J. (2001). "Cooperativity in transcriptional control". *Current Biology* 11: R250–R252.

Cramer, P. (2001). "Structural Basis of Transcription: RNA Polymerase II at 2.8 Angstrom Resolution". *Science* 292: pp. 1863–1876.

Cramer, P. (2019). "Organization and regulation of gene transcription". *Nature* 573: pp. 45–54.

Creppe, C., Palau, A., Malinverni, R., Valero, V., and Buschbeck, M. (2014). "A Cbx8-containing polycomb complex facilitates the transition to gene activation during ES cell differentiation". *PLoS Genet* 10: e1004851.

Creyghton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., Hanna, J., Lodato, M. A., Frampton, G. M., Sharp, P. A., Boyer, L. A., Young, R. A., and Jaenisch, R. (2010). "Histone H3K27ac separates active from poised enhancers and predicts developmental state". *Proceedings of the National Academy of Sciences* 107: pp. 21931–21936.

Creyghton, M. P., Markoulaki, S., Levine, S. S., Hanna, J., Lodato, M. A., Sha, K., Young, R. A., Jaenisch, R., and Boyer, L. A. (2008). "H2AZ Is Enriched at Polycomb Complex Target Genes in ES Cells and Is Necessary for Lineage Commitment". *Cell* 135: pp. 649–661.

Croce, L. D. and Helin, K. (2013). "Transcriptional regulation by Polycomb group proteins". *Nature Structural & Molecular Biology* 20: pp. 1147–1155.

Cruz, X. de la, Lois, S., Sanchez-Molina, S., and Martinez-Balbas, M. A. (2005). "Do protein motifs read the histone code?" *BioEssays* 27: pp. 164–175.

Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A., and Pirrotta, V. (2002). "Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites". *Cell* 111: pp. 185–96.

Dai, X., Bai, Y., Zhao, L., Dou, X., Liu, Y., Wang, L., Li, Y., Li, W., Hui, Y., Huang, X., Wang, Z., and Qin, Y. (2017). "H2A.Z Represses Gene Expression by Modulating Promoter Nucleosome Structure and Enhancer Histone Modifications in Arabidopsis". *Molecular Plant* 10: pp. 1274–1292.

DAlessio, J. A., Wright, K. J., and Tjian, R. (2009). "Shifting Players and Paradigms in Cell-Specific Transcription". *Molecular Cell* 36: pp. 924–931.

Danzer, J. R. and Wallrath, L. L. (2004). "Mechanisms of HP1-mediated gene silencing in Drosophila". *Development* 131: p. 3571.

Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W., and Richmond, T. J. (2002). "Solvent Mediated Interactions in the Structure of the Nucleosome Core Particle at 1.9A Resolution". 319: pp. 1097–1113.

Davidovich, C., Zheng, L., Goodrich, K. J., and Cech, T. R. (2013). "Promiscuous RNA binding by Polycomb repressive complex 2". *Nature Structural & Molecular Biology* 20: pp. 1250–1257.

Deardin, J., Rappailles, A., Cuvier, O., Grimaud, C., Decoville, M., Locker, D., and Cavalli, G. (2005). "Recruitment of Drosophila Polycomb group proteins to chromatin by DSP1". *Nature* 434: pp. 533–538.

Deato, M. D. E. and Tjian, R. (2007). "Switching of the core transcription machinery during myogenesis". *Genes & Development* 21: pp. 2137–2149.

Deaton, A. M. and Bird, A. (2011). "CpG islands and the regulation of transcription". *Genes & Development* 25: pp. 1010–1022.

Delgado-Olguin, P. and Recillas-Targa, F. (2011). "Chromatin structure of pluripotent stem cells and induced pluripotent stem cells". *Briefings in Functional Genomics* 10: pp. 37–49.

Denslow, S. A. and Wade, P. A. (2007). "The human Mi-2/NuRD complex and gene regulation". *Oncogene* 26: p. 5433.

Deshaies, R. J. and Joazeiro, C. A. P. (2009). "RING Domain E3 Ubiquitin Ligases". *Annu. Rev. Biochem.* 78: pp. 399–434.

DesJarlais, R. and Tummino, P. J. (2016). "Role of Histone-Modifying Enzymes and Their Complexes in Regulation of Chromatin Biology". *Biochemistry* 55: pp. 1584–1599.

Dey, A., Seshasayee, D., Noubade, R., et al. (2012). "Loss of the Tumor Suppressor BAP1 Causes Myeloid Transformation". *Science* 337: pp. 1541–1546.

Dietrich, N., Lerdrup, M., Landt, E., Agrawal-Singh, S., Bak, M., Tommerup, N., Rappsilber, J., Södersten, E., and Hansen, K. (2012). "REST-Mediated Recruitment of Polycomb Repressor Complexes in Mammalian Cells". *PLoS Genetics* 8. Ed. by H. D. Madhani: e1002494.

Dieuleveult, M. de, Yen, K., Hmitou, I., et al. (2016). "Genome-wide nucleosome specificity and function of chromatin remodellers in ES cells". *Nature* 530: pp. 113–116.

Dikstein, R., Zhou, S., and Tjian, R. (1996). "Human TAFII105 Is a Cell Type–Specific TFIID Subunit Related to hTAFII130". *Cell* 87: pp. 137–146.

Dimitrova, E., Kondo, T., Feldmann, A., Nakayama, M., Koseki, Y., Konietzny, R., Kessler, B. M., Koseki, H., and Klose, R. J. (2018). "FBXL19 recruits CDK-Mediator to CpG islands of developmental genes priming them for activation during lineage commitment". *Elife* 7.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T. R. (2013). "STAR: ultrafast universal RNA-seq aligner". *Bioinformatics* 29: pp. 15–21.

Dorighi, K. M., Swigut, T., Henriques, T., Bhanu, N. V., Scruggs, B. S., Nady, N., Still, C. D., Garcia, B. A., Adelman, K., and Wysocka, J. (2017). "Mll3 and Mll4 Facilitate Enhancer RNA Synthesis and Transcription from Promoters Independently of H3K4 Monomethylation". *Molecular Cell* 66: 568–576.e4.

Dorighi, K. M. and Tamkun, J. W. (2013). "The trithorax group proteins Kismet and ASH1 promote H3K36 dimethylation to counteract Polycomb group repression in Drosophila". *Development* 140: p. 4182.

Dorigo, B., Schalch, T., Bystricky, K., and Richmond, T. J. (2003). "Chromatin Fiber Folding: Requirement for the Histone H4 N-terminal Tail". *Journal of Molecular Biology* 327: pp. 85–96.

Duncan, I. M. (1982). "Polycomblike: a gene that appears to be required for the normal expression of the bithorax and antennapedia gene complexes of Drosophila melanogaster". eng. *Genetics* 102: pp. 49–70.

Durrin, L. K., Mann, R. K., Kayne, P. S., and Grunstein, M. (1991). "Yeast histone H4 N-terminal sequence is required for promoter activation in vivo". *Cell* 65: pp. 1023–1031.

Dyke, M. V., Roeder, R., and Sawadogo, M (1988). "Physical analysis of transcription preinitiation complex assembly on a class II gene promoter". *Science* 241: pp. 1335–1338.

Dynan, W. S. and Tjian, R. (1983). "The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter". *Cell* 35: pp. 79–87.

Eagen, K. P., Aiden, E. L., and Kornberg, R. D. (2017). "Polycomb-mediated chromatin loops revealed by a subkilobase-resolution chromatin interaction map". *Proc Natl Acad Sci U S A* 114: pp. 8764–8769.

Eissenberg, J. C., James, T. C., Foster-Hartnett, D. M., Hartnett, T., Ngan, V., and Elgin, S. C. (1990). "Mutation in a heterochromatin-specific chromosomal protein is associated with

suppression of position-effect variegation in Drosophila melanogaster." *Proceedings of the National Academy of Sciences* 87: pp. 9923–9927.

Eissenberg, J. C. (2001). "Molecular biology of the chromo domain: an ancient chromatin module comes of age". *Gene* 275: pp. 19–29.

Elderkin, S., Maertens, G. N., Endoh, M., Mallery, D. L., Morrice, N., Koseki, H., Peters, G., Brockdorff, N., and Hiom, K. (2007). "A Phosphorylated Form of Mel-18 Targets the Ring1B Histone H2A Ubiquitin Ligase to Chromatin". *Molecular Cell* 28: pp. 107–120.

Enderle, D., Beisel, C., Stadler, M. B., Gerstung, M., Athri, P., and Paro, R. (2010). "Polycomb preferentially targets stalled promoters of coding and noncoding transcripts". *Genome Research* 21: pp. 216–226.

Endoh, M., Endo, T. A., Endoh, T., Fujimura, Y., Ohara, O., Toyoda, T., Otte, A. P., Okano, M., Brockdorff, N., Vidal, M., and Koseki, H. (2008). "Polycomb group proteins Ring1A/B are functionally linked to the core transcriptional regulatory circuitry to maintain ES cell identity". *Development* 135: pp. 1513–24.

Endoh, M., Endo, T. A., Endoh, T., Isono, K., Sharif, J., Ohara, O., Toyoda, T., Ito, T., Eskeland, R., Bickmore, W. A., Vidal, M., Bernstein, B. E., and Koseki, H. (2012). "Histone H2A mono-ubiquitination is a crucial step to mediate PRC1-dependent repression of developmental genes to maintain ES cell identity". *PLoS Genet* 8: e1002774.

Endoh, M., Endo, T. A., Shinga, J., Hayashi, K., Farcas, A., Ma, K. W., Ito, S., Sharif, J., Endoh, T., Onaga, N., Nakayama, M., Ishikura, T., Masui, O., Kessler, B. M., Suda, T., Ohara, O., Okuda, A., Klose, R., and Koseki, H. (2017). "PCGF6-PRC1 suppresses premature differentiation of mouse embryonic stem cells by regulating germ cell-related genes". *Elife* 6.

Eskeland, R., Freyer, E., Leeb, M., Wutz, A., and Bickmore, W. A. (2010a). "Histone Acetylation and the Maintenance of Chromatin Compaction by Polycomb Repressive Complexes". *Cold Spring Harbor Symposia on Quantitative Biology* 75: pp. 71–78.

Eskeland, R., Leeb, M., Grimes, G. R., Kress, C., Boyle, S., Sproul, D., Gilbert, N., Fan, Y., Skoultchi, A. I., Wutz, A., and Bickmore, W. A. (2010b). "Ring1B compacts chromatin structure and represses gene expression independent of histone ubiquitination". *Mol Cell* 38: pp. 452–64.

Faast, R., Thonglairoam, V., Schulz, T. C., Beall, J., Wells, J. R., Taylor, H., Matthaei, K., Rathjen, P. D., Tremethick, D. J., and Lyons, I. (2001). "Histone variant H2A.Z is required for early mammalian development". *Current Biology* 11: pp. 1183–1187.

Farcas, A. M., Blackledge, N. P., Sudbery, I., Long, H. K., McGouran, J. F., Rose, N. R., Lee, S., Sims, D., Cerase, A., Sheahan, T. W., Koseki, H., Brockdorff, N., Ponting, C. P., Kessler, B. M., and Klose, R. J. (2012). "KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands". *Elife* 1: e00205.

Faust, C., Schumacher, A., Holdener, B., and Magnuson, T. (1995). "The eed mutation disrupts anterior mesoderm production in mice". *Development* 121: p. 273.

Felsenfeld, G., Boyes, J., Chung, J., Clark, D., and Studitsky, V. (1996). "Chromatin structure and gene expression." *Proceedings of the National Academy of Sciences* 93: pp. 9384–9388.

Fereres, S., Simón, R., Mohd-Sarip, A., Verrijzer, C. P., and Busturia, A. (2014). "dRYBP Counteracts Chromatin-Dependent Activation and Repression of Transcription". *PLoS ONE* 9. Ed. by R. Mantovani: e113255.

Ferrari, K. J., Scelfo, A., Jammula, S., Cuomo, A., Barozzi, I., Stützer, A., Fischle, W., Bonaldi, T., and Pasini, D. (2014). "Polycomb-Dependent H3K27me1 and H3K27me2 Regulate Active Transcription and Enhancer Fidelity". *Molecular Cell* 53: pp. 49–62.

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

Flanagan, P. M., Kelleher, R. J., Sayre, M. H., Tschochner, H., and Kornberg, R. D. (1991). "A mediator required for activation of RNA polymerase II transcription in vitro". *Nature* 350: pp. 436–438.

Fondell, J. D., Ge, H., and Roeder, R. G. (1996). "Ligand induction of a transcriptionally active thyroid hormone receptor coactivator complex." *Proceedings of the National Academy of Sciences* 93: pp. 8329–8333.

Fong, Y. W., Cattoglio, C., Yamaguchi, T., and Tjian, R. (2012). "Transcriptional regulation by coactivators in embryonic stem cells". *Trends in Cell Biology* 22: pp. 292–298.

Formosa, T. (2012). "The role of FACT in making and breaking nucleosomes". *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1819: pp. 247–255.

Forzati, F., Federico, A., Pallante, P., Abbate, A., Esposito, F., Malapelle, U., Sepe, R., Palma, G., Troncone, G., Scarfò, M., Arra, C., Fedele, M., and Fusco, A. (2012). "CBX7 is a tumor suppressor in mice and humans". *Journal of Clinical Investigation* 122: pp. 612–623.

Francis, N. J., Kingston, R. E., and Woodcock, C. L. (2004). "Chromatin Compaction by a Polycomb Group Protein Complex". *Science* 306: p. 1574.

Francis, N. J., Saurin, A. J., Shao, Z., and Kingston, R. E. (2001). "Reconstitution of a Functional Core Polycomb Repressive Complex". *Molecular Cell* 8: pp. 545–556.

Frangini, A., Sjoberg, M., Roman-Trufero, M., Dharmalingam, G., Haberle, V., Bartke, T., Lenhard, B., Malumbres, M., Vidal, M., and Dillon, N. (2013). "The aurora B kinase and the polycomb protein ring1B combine to regulate active promoters in quiescent lymphocytes". *Mol Cell* 51: pp. 647–61.

Fraser, N. W., Sehgal, P. B., and Darnell, J. E. (1978). "DRB-induced premature termination of late adenovirus transcription". *Nature* 272: pp. 590–593.

Freemont, P. S., Hanson, I. M., and Trowsdale, J. (1991). "A novel gysteine-rich sequence motif". *Cell* 64: pp. 483–484.

Frey, F., Sheahan, T., Finkl, K., Stoehr, G., Mann, M., Benda, C., and Müller, J. (2016). "Molecular basis of PRC1 targeting to Polycomb response elements by PhoRC". eng. *Genes & development* 30: pp. 1116–1127.

Fudenberg, G., Abdennur, N., Imakaev, M., Goloborodko, A., and Mirny, L. A. (2017). "Emerging Evidence of Chromosome Folding by Loop Extrusion". *Cold Spring Harbor Symposia on Quantitative Biology* 82: pp. 45–55.

Fyodorov, D. V., Zhou, B.-R., Skoultchi, A. I., and Bai, Y. (2017). "Emerging roles of linker histones in regulating chromatin structure and function". *Nature Reviews Molecular Cell Biology* 19: pp. 192–206.

Gambetta, M. C. and Müller, J. (2014). "O-GlcNAcylation Prevents Aggregation of the Polycomb Group Repressor Polyhomeotic". *Developmental Cell* 31: pp. 629–639.

Gambetta, M. C., Oktaba, K., and Müller, J. (2009). "Essential Role of the Glycosyltransferase Sxc/Ogt in Polycomb Repression". *Science* 325: pp. 93–96.

Gao, Z., Lee, P., Stafford, J. M., Schimmelmann, M. von, Schaefer, A., and Reinberg, D. (2014). "An AUTS2-Polycomb complex activates gene expression in the CNS". *Nature* 516: pp. 349–54.

Gao, Z., Zhang, J., Bonasio, R., Strino, F., Sawai, A., Parisi, F., Kluger, Y., and Reinberg, D. (2012). "PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes". *Mol Cell* 45: pp. 344–56.

Garcia, E., Marcos-Gutiérrez, C., Mar Lorente, M. del, Moreno, J. C., and Vidal, M. (1999). "RYBP, a new repressor protein that interacts with components of the mammalian Polycomb complex, and with the transcription factor YY1". *The EMBO Journal* 18: pp. 3404–3418.

Garcia-Fernandez, J. (2005). "The genesis and evolution of homeobox gene clusters". *Nature Reviews Genetics* 6: pp. 881–892.

Gariglio, P., Bellard, M., and Chambon, P. (1981). "Clustering of RNA polymerase B molecules in the 5-prime moiety of the adult beta-globin gene of hen erythrocytes". *Nucleic Acids Research* 9: pp. 2589–2598.

Gaubatz, S., Wood, J. G., and Livingston, D. M. (1998). "Unusual proliferation arrest and transcriptional control properties of a newly discovered E2F family member, E2F-6". *Proceedings of the National Academy of Sciences* 95: pp. 9190–9195.

Ge, H., Si, Y., and Roeder, R. G. (1998). "Isolation of cDNAs encoding novel transcription coactivators p52 and p75 reveals an alternate regulatory mechanism of transcriptional activation". *The EMBO Journal* 17: pp. 6723–6729.

Gearhart, M. D., Corcoran, C. M., Wamstad, J. A., and Bardwell, V. J. (2006). "Polycomb Group and SCF Ubiquitin Ligases Are Found in a Novel BCOR Complex That Is Recruited to BCL6 Targets". *Molecular and Cellular Biology* 26: pp. 6880–6889.

Gibcus, J. H., Samejima, K., Goloborodko, A., Samejima, I., Naumova, N., Nuebler, J., Kanemaki, M. T., Xie, L., Paulson, J. R., Earnshaw, W. C., Mirny, L. A., and Dekker, J. (2018). "A pathway for mitotic chromosome formation". *Science* 359: eaao6135.

Gilbert, N. and Ramsahoye, B. (2005). "The relationship between chromatin structure and transcriptional activity in mammalian genomes". *Briefings in Functional Genomics* 4: pp. 129–142.

Gilmour, D. S. and Lis, J. T. (1986). "RNA polymerase II interacts with the promoter region of the noninduced hsp70 gene in Drosophila melanogaster cells." *Molecular and Cellular Biology* 6: pp. 3984–3989.

Gindhart, J. G., Jr, and Kaufman, T. C. (1995). "Identification of Polycomb and trithorax group responsive elements in the regulatory region of the Drosophila homeotic gene Sex combs reduced". eng. *Genetics* 139: pp. 797–814.

Giorgetti, L., Siggers, T., Tiana, G., Caprara, G., Notarbartolo, S., Corona, T., Pasparakis, M., Milani, P., Bulyk, M. L., and Natoli, G. (2010). "Noncooperative Interactions between Transcription Factors and Clustered DNA Binding Sites Enable Graded Transcriptional Responses to Environmental Inputs". *Molecular Cell* 37: pp. 418–428.

Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A., and Kornberg, R. D. (June 2001). "Structural Basis of Transcription: An RNA Polymerase II Elongation Complex at 3.3 A Resolution". *Science* 292: p. 1876.

Goldknopf, I. L., Taylor, C. W., Baum, R. M., Yeoman, L. C., Olson, M. O., Prestayko, A. W., and Busch, H (1975). "Isolation and characterization of protein A24, a "histone-like" non-histone chromosomal protein." *Journal of Biological Chemistry* 250: pp. 7182–7187.

Goldstein, G., Scheid, M., Hammerling, U., Schlesinger, D. H., Niall, H. D., and Boyse, E. A. (1975). "Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells". *Proc Natl Acad Sci USA* 72: p. 11.

Gong, Y., Wang, X., Liu, J., Shi, L., Yin, B., Peng, X., Qiang, B., and Yuan, J. (2004). "NSPc1, a mainly nuclear localized protein of novel PcG family members, has a transcription repression activity related to its PKC phosphorylation site at S183". *FEBS Letters* 579: pp. 115–121.

Goodrich, J. A. and Tjian, R. (2010). "Unexpected roles for core promoter recognition factors in cell-type-specific transcription and gene regulation". *Nature Reviews Genetics* 11: pp. 549–558.

Grau, D. J., Chapman, B. A., Garlick, J. D., Borowsky, M., Francis, N. J., and Kingston, R. E. (2011). "Compaction of chromatin by diverse Polycomb group proteins requires localized regions of high charge". *Genes Dev* 25: pp. 2210–21.

Gray, F., Cho, H. J., Shukla, S., He, S., Harris, A., Boytsov, B., Jaremko, u., Jaremko, M., Demeler, B., Lawlor, E. R., Grembecka, J., and Cierpicki, T. (2016). "BMI1 regulates PRC1 architecture and activity through homo- and hetero-oligomerization". *Nature Communications* 7: p. 13343.

Grijzenhout, A., Godwin, J., Koseki, H., Gdula, M. R., Szumska, D., McGouran, J. F., Bhattacharya, S., Kessler, B. M., Brockdorff, N., and Cooper, S. (2016). "Functional analysis of

AEBP2, a PRC2 Polycomb protein, reveals a Trithorax phenotype in embryonic development and in ESCs". *Development* 143: pp. 2716–2723.

Grimaud, C., Nègre, N., and Cavalli, G. (2006). "From genetics to epigenetics: the tale of Polycomb group and trithorax group genes". *Chromosome Research* 14: pp. 363–375.

Gross, D. S. and Garrard, W. T. (1988). "Nuclease Hypersensitive Sites in Chromatin". *Annual Review of Biochemistry* 57: pp. 159–197.

Groth, A., Rocha, W., Verreault, A., and Almouzni, G. (2007). "Chromatin Challenges during DNA Replication and Repair". *Cell* 128: pp. 721–733.

Grunstein, M. (1990). "Histone Function in Transcription". *Annual Review of Cell Biology* 6: pp. 643–676.

Grunstein, M. (1997). "Histone acetylation in chromatin structure and transcription". *Nature* 389: pp. 349–352.

Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R., and Young, R. A. (2007). "A Chromatin Landmark and Transcription Initiation at Most Promoters in Human Cells". *Cell* 130: pp. 77–88.

Guglielmi, B., Rochelle, N. L., and Tjian, R. (2013). "Gene-Specific Transcriptional Mechanisms at the Histone Gene Cluster Revealed by Single-Cell Imaging". *Molecular Cell* 51: pp. 480–492.

Gunster, M. J., Satijn, D. P., Hamer, K. M., Blaauwen, J. L. den, Bruijn, D de, Alkema, M. J., Lohuizen, M van, Driel, R van, and Otte, A. P. (1997). "Identification and characterization of interactions between the vertebrate polycomb-group protein BMI1 and human homologs of polyhomeotic." *Molecular and Cellular Biology* 17: pp. 2326–2335.

Guo, Y., Nady, N., Qi, C., Allali-Hassani, A., Zhu, H., Pan, P., Adams-Cioaba, M. A., Amaya, M. F., Dong, A., Vedadi, M., Schapira, M., Read, R. J., Arrowsmith, C. H., and Min, J. (2009). "Methylation-state-specific recognition of histones by the MBT repeat protein L3MBTL2". *Nucleic Acids Research* 37: pp. 2204–2210.

Gutierrez, P. S., Monteoliva, D., and Diambra, L. (2009). "Role of cooperative binding on noise expression". *Physical Review E* 80.

Gutierrez, P. S., Monteoliva, D., and Diambra, L. (2012). "Cooperative Binding of Transcription Factors Promotes Bimodal Gene Expression Response". *PLoS ONE* 7. Ed. by F. Emmert-Streib: e44812.

Haberle, V., Arnold, C. D., Pagani, M., Rath, M., Schernhuber, K., and Stark, A. (2019). "Transcriptional cofactors display specificity for distinct types of core promoters". *Nature* 570: pp. 122–126.

Haberle, V. and Stark, A. (2018). "Eukaryotic core promoters and the functional basis of transcription initiation". *Nature Reviews Molecular Cell Biology* 19: pp. 621–637.

Han, M. and Grunstein, M. (1988). "Nucleosome loss activates yeast downstream promoters in vivo". *Cell* 55: pp. 1137–1145.

Han, X.-R., Zha, Z, Yuan, H.-X., Feng, X, Xia, Y.-K., Lei, Q.-Y., Guan, K.-L., and Xiong, Y (2016). "KDM2B/FBXL10 targets c-Fos for ubiquitylation and degradation in response to mitogenic stimulation". *Oncogene* 35: pp. 4179–4190.

Handley, P. M., Mueckler, M., Siegel, N. R., Ciechanover, A., and Schwartz, A. L. (1991). "Molecular cloning, sequence, and tissue distribution of the human ubiquitin-activating enzyme E1." *Proceedings of the National Academy of Sciences* 88: pp. 258–262.

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

Happel, A. M., Swanson, M. S., and Winston, F. (1991). "The SNF2, SNF5 and SNF6 genes are required for Ty transcription in Saccharomyces cerevisiae". eng. *Genetics* 128: pp. 69–77.

Hart, G. W., Slawson, C., Ramirez-Correa, G., and Lagerlof, O. (2011). "Cross Talk Between O-GlcNAcylation and Phosphorylation: Roles in Signaling, Transcription, and Chronic Disease". *Annual Review of Biochemistry* 80: pp. 825–858.

Hassa, P. O., Haenni, S. S., Elser, M., and Hottiger, M. O. (2006). "Nuclear ADP-Ribosylation Reactions in Mammalian Cells: Where Are We Today and Where Are We Going?" *Microbiology and Molecular Biology Reviews* 70: pp. 789–829.

Hassan, A. H., Neely, K. E., and Workman, J. L. (2001). "Histone Acetyltransferase Complexes Stabilize SWI/SNF Binding to Promoter Nucleosomes". *Cell* 104: pp. 817–827.

Hauer, M. H. and Gasser, S. M. (2017). "Chromatin and nucleosome dynamics in DNA damage and repair". *Genes & Development* 31: pp. 2204–2221.

Hauri, S., Comoglio, F., Seimiya, M., Gerstung, M., Glatter, T., Hansen, K., Aebersold, R., Paro, R., Gstaiger, M., and Beisel, C. (2016). "A High-Density Map for Navigating the Human Polycomb Complexome". *Cell Rep* 17: pp. 583–595.

He, G.-P., Kim, S., and Ro, H.-S. (1999). "Cloning and Characterization of a Novel Zinc Finger Transcriptional Repressor". *Journal of Biological Chemistry* 274: pp. 14678–14684.

He, J., Shen, L., Wan, M., Taranova, O., Wu, H., and Zhang, Y. (2013a). "Kdm2b maintains murine embryonic stem cell status by recruiting PRC1 complex to CpG islands of developmental genes". *Nat Cell Biol* 15: pp. 373–84.

He, J., Fu, X., Zhang, M., He, F., Li, W., Abdul, M. M., Zhou, J., Sun, L., Chang, C., Li, Y., Liu, H., Wu, K., Babarinde, I. A., Zhuang, Q., Loh, Y.-H., Chen, J., Esteban, M. A., and Hutchins, A. P. (2019a). "Transposable elements are regulated by context-specific patterns of chromatin marks in mouse embryonic stem cells". *Nature Communications* 10.

He, J., Kallin, E. M., Tsukada, Y. ichi, and Zhang, Y. (2008). "The H3K36 demethylase Jhdm1b/Kdm2b regulates cell proliferation and senescence through p15-Ink4b". *Nature Structural & Molecular Biology* 15: pp. 1169–1175.

He, M., Chaurushiya, M. S., Webster, J. D., Kummerfeld, S., Reja, R., Chaudhuri, S., Chen, Y.-J., Modrusan, Z., Haley, B., Dugger, D. L., Eastham-Anderson, J., Lau, S., Dey, A., Caothien, R., Roose-Girma, M., Newton, K., and Dixit, V. M. (2019b). "Intrinsic apoptosis shapes the tumor spectrum linked to inactivation of the deubiquitinase BAP1". *Science* 364: pp. 283–285.

He, X., Samee, M. A. H., Blatti, C., and Sinha, S. (2010). "Thermodynamics-Based Models of Transcriptional Regulation by Enhancers: The Roles of Synergistic Activation, Cooperative Binding and Short-Range Repression". *PLoS Computational Biology* 6. Ed. by U. Ohler: e1000935.

He, Y., Fang, J., Taatjes, D. J., and Nogales, E. (2013b). "Structural visualization of key steps in human transcription initiation". *Nature* 495: pp. 481–486.

He, Y., Selvaraju, S., Curtin, M. L., et al. (2017). "The EED protein–protein interaction inhibitor A-395 inactivates the PRC2 complex". *Nature Chemical Biology* 13: pp. 389–395.

Healy, E., Mucha, M., Glancy, E., Fitzpatrick, D. J., Conway, E., Neikes, H. K., Monger, C., Mierlo, G. V., Baltissen, M. P., Koseki, Y., Vermeulen, M., Koseki, H., and Bracken, A. P. (2019). "PRC2.1 and PRC2.2 Synergize to Coordinate H3K27 Trimethylation". *Molecular Cell*.

Hebbes, T. R., Thorne, A. W., and Crane-Robinson, C. (1988). "A direct link between core histone acetylation and transcriptionally active chromatin". eng. *The EMBO journal* 7: pp. 1395–1402.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., Cheng, J. X., Murre, C., Singh, H., and Glass, C. K. (2010). "Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities". *Mol Cell* 38: pp. 576–89.

Heitz, E. (1928). "Das Heterochromatin der Moose". *Jhrb. Wiss. Botanik* 69: pp. 762–818.

Heitz, E. (1929). "Heterochromatin, Chromocentren, Chromomeren (Vorlaufige Mitteilung)". *Ber. Dtsch. Bot. Ges.* 47: pp. 274–284.

Henikoff, S. (2008). "Nucleosome destabilization in the epigenetic regulation of gene expression". *Nature Reviews Genetics* 9: pp. 15–26.

Henikoff, S. and Ahmad, K. (2005). "Assembly of variant histones into chromatin". *Annual Review of Cell and Developmental Biology* 21: pp. 133–153.

Henikoff, S. and Smith, M. M. (2015). "Histone Variants and Epigenetics". *Cold Spring Harbor Perspectives in Biology* 7: a019364.

Henriques, T., Scruggs, B. S., Inouye, M. O., Muse, G. W., Williams, L. H., Burkholder, A. B., Lavender, C. A., Fargo, D. C., and Adelman, K. (2018). "Widespread transcriptional pausing and elongation control at enhancers". *Genes & Development* 32: pp. 26–41.

Herranz, N., Pasini, D., Diaz, V. M., Franci, C., Gutierrez, A., Dave, N., Escriva, M., Hernandez-Munoz, I., Croce, L. D., Helin, K., Herreros, A. G. de, and Peiro, S. (2008). "Polycomb Complex 2 Is Required for E-cadherin Repression by the Snail1 Transcription Factor". *Molecular and Cellular Biology* 28: pp. 4772–4781.

Herz, H.-M., Mohan, M., Garrett, A. S., Miller, C., Casto, D., Zhang, Y., Seidel, C., Haug, J. S., Florens, L., Washburn, M. P., Yamaguchi, M., Shiekhattar, R., and Shilatifard, A. (2012). "Polycomb Repressive Complex 2-Dependent and -Independent Functions of Jarid2 in Transcriptional Regulation in Drosophila". *Molecular and Cellular Biology* 32: pp. 1683–1693.

Hirschhorn, J. N., Brown, S. A., Clark, C. D., and Winston, F. (1992). "Evidence that SNF2/SWI2 and SNF5 activate transcription in yeast by altering chromatin structure." *Genes & Development* 6: pp. 2288–2298.

Hisada, K., Sanchez, C., Endo, T. A., Endoh, M., Roman-Trufero, M., Sharif, J., Koseki, H., and Vidal, M. (2012). "RYBP Represses Endogenous Retroviruses and Preimplantation- and Germ Line-Specific Genes in Mouse Embryonic Stem Cells". *Molecular and Cellular Biology* 32: pp. 1139–1149.

Ho, L. and Crabtree, G. R. (2010). "Chromatin remodelling during development". *Nature* 463: pp. 474–484.

Hochstrasser, M. (1996). "Ubiquitin-dependent protein degradation". *Annual Review of Genetics* 30: pp. 405–439.

Hoffmann, A., Sportelli, V., Ziller, M., and Spengler, D. (2017). "Switch-Like Roles for Polycomb Proteins from Neurodevelopment to Neurodegeneration". *Epigenomes* 1: p. 21.

Hojfeldt, J. W., Laugesen, A., Willumsen, B. M., Damhofer, H., Hedehus, L., Tvardovskiy, A., Mohammad, F., Jensen, O. N., and Helin, K. (2018). "Accurate H3K27 methylation can be established de novo by SUZ12-directed PRC2". *Nature Structural & Molecular Biology* 25: pp. 225–232.

Højfeldt, J. W., Hedehus, L., Laugesen, A., Tatar, T., Wiehle, L., and Helin, K. (2019). "Non-core Subunits of the PRC2 Complex Are Collectively Required for Its Target-Site Specificity". *Molecular Cell*.

Holbert, M. A. and Marmorstein, R. (2005). "Structure and activity of enzymes that remove histone modifications". *Current Opinion in Structural Biology* 15: pp. 673–680.

Hong, S., Cho, Y.-W., Yu, L.-R., Yu, H., Veenstra, T. D., and Ge, K. (2007). "Identification of JmjC domain-containing UTX and JMJD3 as histone H3 lysine 27 demethylases". *Proceedings of the National Academy of Sciences* 104: pp. 18439–18444.

Horard, B., Tatout, C., Poux, S., and Pirrotta, V. (2000). "Structure of a Polycomb Response Element and In Vitro Binding of Polycomb Group Complexes Containing GAGA Factor". *Molecular and Cellular Biology* 20: pp. 3187–3197.

Horikoshi, M., Hai, T., Lin, Y.-S., Green, M. R., and Roeder, R. G. (1988). "Transcription factor ATF interacts with the TATA factor to facilitate establishment of a preinitiation complex". *Cell* 54: pp. 1033–1042.

Hosogane, M., Funayama, R., Nishida, Y., Nagashima, T., and Nakayama, K. (2013). "Ras-Induced Changes in H3K27me3 Occur after Those in Transcriptional Activity". *PLoS Genetics* 9. Ed. by G. S. Barsh: e1003698.

House, N. C. M., Koch, M. R., and Freudenreich, C. H. (2014). "Chromatin modifications and DNA repair: beyond double-strand breaks". *Frontiers in Genetics* 5.

Hu, B., Petela, N., Kurze, A., Chan, K. L., Chapard, C., and Nasmyth, K. (2015). "Biological chromodynamics: a general method for measuring protein occupancy across the genome by calibrating ChIP-seq". *Nucleic Acids Res* 43: e132.

Hu, G., Cui, K., Northrup, D., Liu, C., Wang, C., Tang, Q., Ge, K., Levens, D., Crane-Robinson, C., and Zhao, K. (2013). "H2A.Z Facilitates Access of Active and Repressive Complexes to Chromatin in Embryonic Stem Cell Self-Renewal and Differentiation". *Cell Stem Cell* 12: pp. 180–192.

Huber, W., Carey, V. J., Gentleman, R., et al. (2015). "Orchestrating high-throughput genomic analysis with Bioconductor". *Nat Methods* 12: pp. 115–21.

Hunkapiller, J., Shen, Y., Diaz, A., Cagney, G., McCleary, D., Ramalho-Santos, M., Krogan, N., Ren, B., Song, J. S., and Reiter, J. F. (2012). "Polycomb-Like 3 Promotes Polycomb Repressive Complex 2 Binding to CpG Islands and Embryonic Stem Cell Self-Renewal". *PLoS Genetics* 8. Ed. by G. Cavalli: e1002576.

Hurlin, P. J., Steingrìmsson, E., Copeland, N. G., Jenkins, N. A., and Eisenman, R. N. (1999). "Mga, a dual-specificity transcription factor that interacts with Max and contains a T-domain DNA-binding motif". *The EMBO Journal* 18: pp. 7019–7028.

Illingworth, R. S., Moffat, M., Mann, A. R., Read, D., Hunter, C. J., Pradeepa, M. M., Adams, I. R., and Bickmore, W. A. (2015). "The E3 ubiquitin ligase activity of RING1B is not essential for early mouse development". *Genes Dev* 29: pp. 1897–902.

Illingworth, R. S. (2019). "Chromatin folding and nuclear architecture: PRC1 function in 3D". *Current Opinion in Genetics & Development* 55: pp. 82–90.

Imbalzano, A. N., Kwon, H., Green, M. R., and Kingston, R. E. (1994). "Facilitated binding of TATA-binding protein to nucleosomal DNA". *Nature* 370: pp. 481–485.

Imhof, A. (2006). "Epigenetic regulators and histone modification". *Briefings in Functional Genomics* 5: pp. 222–227.

Ingham, P. W. (1983). "Differential expression of bithorax complex genes in the absence of the extra sex combs and trithorax genes". *Nature* 306: pp. 591–593.

Ingham, P. W. (1984). "A gene that regulates the bithorax complex differentially in larval and adult cells of Drosophila". *Cell* 37: pp. 815–823.

Ingham, P. W. (1985a). "A clonal analysis of the requirement for the trithorax gene in the diversification of segments in Drosophila". *Journal of Embryology and Experimental Morphology* VOL. 89: pp. 349–365.

Ingham, P. W. (1985b). "Genetic Control of the Spatial Pattern of Selector Gene Expression in Drosophila". *Cold Spring Harbor Symposia on Quantitative Biology* 50: pp. 201–208.

Ismail, I. H., Andrin, C., McDonald, D., and Hendzel, M. J. (2010). "BMI1-mediated histone ubiquitylation promotes DNA double-strand break repair". *J Cell Biol* 191: pp. 45–60.

Ismail, I. H., McDonald, D., Strickfaden, H., Xu, Z., and Hendzel, M. J. (2013). "A Small Molecule Inhibitor of Polycomb Repressive Complex 1 Inhibits Ubiquitin Signaling at DNA Double-strand Breaks". *Journal of Biological Chemistry* 288: pp. 26944–26954.

Isono, K., Endo, T. A., Ku, M., Yamada, D., Suzuki, R., Sharif, J., Ishikura, T., Toyoda, T., Bernstein, B. E., and Koseki, H. (2013). "SAM domain polymerization links subnuclear clustering of PRC1 to gene silencing". *Dev Cell* 26: pp. 565–77.

Isono, K., Fujimura, Y., Shinga, J., Yamaki, M., J, O. W., Takihara, Y., Murahashi, Y., Takada, Y., Mizutani-Koseki, Y., and Koseki, H. (2005). "Mammalian polyhomeotic homologues Phc2 and Phc1 act in synergy to mediate polycomb repression of Hox genes". *Mol Cell Biol* 25: pp. 6694–706.

Issaeva, I., Zonis, Y., Rozovskaia, T., Orlovsky, K., Croce, C. M., Nakamura, T., Mazo, A., Eisenbach, L., and Canaani, E. (2007). "Knockdown of ALR (MLL2) Reveals ALR Target Genes and Leads to Alterations in Cell Adhesion and Growth". *Mol. Cell. Biol.* 27: p. 1889.

Izzo, A., Kamieniarz-Gdula, K., Ramirez, F., Noureen, N., Kind, J., Manke, T., Steensel, B. van, and Schneider, R. (2013). "The Genomic Landscape of the Somatic Linker Histone Subtypes H1.1 to H1.5 in Human Cells". *Cell Reports* 3: pp. 2142–2154.

Jaeger, S. A., Chan, E. T., Berger, M. F., Stottmann, R., Hughes, T. R., and Bulyk, M. L. (2010). "Conservation and regulatory associations of a wide affinity range of mouse transcription factor binding sites". *Genomics* 95: pp. 185–195.

Janzer, A., Stamm, K., Becker, A., Zimmer, A., Buettner, R., and Kirfel, J. (2012). "The H3K4me3 Histone Demethylase Fbxl10 Is a Regulator of Chemokine Expression, Cellular Morphology, and the Metabolome of Fibroblasts". *Journal of Biological Chemistry* 287: pp. 30984–30992.

Jason, L. J. M., Finn, R. M., Lindsey, G., and Ausió, J. (2004). "Histone H2A Ubiquitination Does Not Preclude Histone H1 Binding, but It Facilitates Its Association with the Nucleosome". *Journal of Biological Chemistry* 280: pp. 4975–4982.

Jensen, D. E., Proctor, M., Marquis, S. T., Gardner, H. P., Ha, S. I., Chodosh, L. A., Ishov, A. M., Tommerup, N., Vissing, H., Sekido, Y., Minna, J., Borodovsky, A., Schultz, D. C., Wilkinson, K. D., Maul, G. G., Barlev, N., Berger, S. L., Prendergast, G. C., and Rauscher, F. J. (1998). "BAP1: a novel ubiquitin hydrolase which binds to the BRCA1 RING finger and enhances BRCA1-mediated cell growth suppression". *Oncogene* 16: pp. 1097–1112.

Jenuwein, T., Laible, G., Dorn, R., and Reuter, G. (1998). "SET domain proteins modulate chromatin domains in eu- and heterochromatin". *Cellular and Molecular Life Sciences CMLS* 54: pp. 80–93.

Jenuwein, T. and Allis, C. D. (2001). "Translating the Histone Code". *Science* 293: p. 1074.

Jermann, P., Hoerner, L., Burger, L., and Schubeler, D. (2014). "Short sequences can efficiently recruit histone H3 lysine 27 trimethylation in the absence of enhancer activity and DNA methylation". *Proceedings of the National Academy of Sciences* 111: E3415–E3421.

Ji, Z., Mohammed, H., Webber, A., Ridsdale, J., Han, N., Carroll, J. S., and Sharrocks, A. D. (2014). "The forkhead transcription factor FOXK2 acts as a chromatin targeting factor for the BAP1-containing histone deubiquitinase complex". *Nucleic Acids Research* 42: pp. 6232–6242.

Jiao, L. and Liu, X. (2015). "Structural basis of histone H3K27 trimethylation by an active polycomb repressive complex 2". eng. *Science (New York, N.Y.)* 350: aac4383–aac4383.

Jin, C. and Felsenfeld, G. (2007). "Nucleosome stability mediated by histone variants H3.3 and H2A.Z". *Genes & Development* 21: pp. 1519–1529.

Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K., and Felsenfeld, G. (2009). "H3.3/H2A.Z double variant-containing nucleosomes mark nucleosome-free regions of active promoters and other regulatory regions". *Nature Genetics* 41: pp. 941–945.

Jolma, A., Yan, J., Whitington, T., Toivonen, J., Nitta, K. R., Rastas, P., Morgunova, E., Enge, M., Taipale, M., Wei, G., Palin, K., Vaquerizas, J. M., Vincentelli, R., Luscombe, N. M., Hughes, T. R., Lemaire, P., Ukkonen, E., Kivioja, T., and Taipale, J. (2013). "DNA-Binding Specificities of Human Transcription Factors". *Cell* 152: pp. 327–339.

Jolma, A., Yin, Y., Nitta, K. R., Dave, K., Popov, A., Taipale, M., Enge, M., Kivioja, T., Morgunova, E., and Taipale, J. (2015). "DNA-dependent formation of transcription factor pairs alters their binding specificity". *Nature* 527: pp. 384–388.

Jones, D. O., Cowell, I. G., and Singh, P. B. (2000). "Mammalian chromodomain proteins: their role in genome organisation and expression". *Bioessays* 22: pp. 124–137.

Jonkers, I. and Lis, J. T. (2015). "Getting up to speed with transcription elongation by RNA polymerase II". *Nature Reviews Molecular Cell Biology* 16: pp. 167–177.

Joo, H.-Y., Zhai, L., Yang, C., Nie, S., Erdjument-Bromage, H., Tempst, P., Chang, C., and Wang, H. (2007). "Regulation of cell cycle progression and gene expression by H2A deubiquitination". *Nature* 449: pp. 1068–1072.

Jürgens, G. (1985). "A group of genes controlling the spatial expression of the bithorax complex in Drosophila". *Nature* 316: pp. 153–155.

Jullien, J., Vodnala, M., Pasque, V., Oikawa, M., Miyamoto, K., Allen, G., David, S. A., Brochard, V., Wang, S., Bradshaw, C., Koseki, H., Sartorelli, V., Beaujean, N., and Gurdon, J. (2017). "Gene Resistance to Transcriptional Reprogramming following Nuclear Transfer Is Directly Mediated by Multiple Chromatin-Repressive Pathways". *Molecular Cell* 65: 873–884.e8.

Jung, H. R., Pasini, D., Helin, K., and Jensen, O. N. (2010). "Quantitative Mass Spectrometry of Histones H3.2 and H3.3 inSuz12-deficient Mouse Embryonic Stem Cells Reveals Distinct, Dynamic Post-translational Modifications at Lys-27 and Lys-36". *Molecular & Cellular Proteomics* 9: pp. 838–850.

Kadonaga, J. T. (2011). "Perspectives on the RNA polymerase II core promoter". *Wiley Interdisciplinary Reviews: Developmental Biology* 1: pp. 40–51.

Kagey, M. H., Newman, J. J., Bilodeau, S., Zhan, Y., Orlando, D. A., Berkum, N. L. van, Ebmeier, C. C., Goossens, J., Rahl, P. B., Levine, S. S., Taatjes, D. J., Dekker, J., and Young, R. A. (2010). "Mediator and cohesin connect gene expression and chromatin architecture". *Nature* 467: pp. 430–435.

Kahn, T. G., Dorafshan, E., Schultheis, D., Zare, A., Stenberg, P., Reim, I., Pirrotta, V., and Schwartz, Y. B. (2016). "Interdependence of PRC1 and PRC2 for recruitment to Polycomb Response Elements". *Nucleic Acids Res* 44: pp. 10132–10149.

Kalb, R., Latwiel, S., Baymaz, H. I., Jansen, P. W., Muller, C. W., Vermeulen, M., and Muller, J. (2014). "Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression". *Nat Struct Mol Biol* 21: pp. 569–71.

Kamakaka, R. T. and Biggins, S. (2005). "Histone variants: deviants?" *Genes & Development* 19: pp. 295–316.

Kan, P.-Y., Caterino, T. L., and Hayes, J. J. (2008). "The H4 Tail Domain Participates in Intra- and Internucleosome Interactions with Protein and DNA during Folding and Oligomerization of Nucleosome Arrays". *Molecular and Cellular Biology* 29: pp. 538–546.

Kaneko, S., Son, J., Bonasio, R., Shen, S. S., and Reinberg, D. (2014). "Nascent RNA interaction keeps PRC2 activity poised and in check". *Genes & Development* 28: pp. 1983–1988.

Kaneko, S., Son, J., Shen, S. S., Reinberg, D., and Bonasio, R. (2013). "PRC2 binds active promoters and contacts nascent RNAs in embryonic stem cells". *Nature Structural & Molecular Biology* 20: pp. 1258–1264.

Kanhere, A., Viiri, K., Araujo, C. C., Rasaiyaah, J., Bouwman, R. D., Whyte, W. A., Pereira, C. F., Brookes, E., Walker, K., Bell, G. W., Pombo, A., Fisher, A. G., Young, R. A., and Jenner, R. G. (2010). "Short RNAs Are Transcribed from Repressed Polycomb Target Genes and Interact with Polycomb Repressive Complex-2". *Molecular Cell* 38: pp. 675–688.

Kapranov, P., Cheng, J., Dike, S., et al. (2007). "RNA Maps Reveal New RNA Classes and a Possible Function for Pervasive Transcription". *Science* 316: pp. 1484–1488.

Karmodiya, K., Krebs, A. R., Oulad-Abdelghani, M., Kimura, H., and Tora, L. (2012). "H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells". *BMC Genomics* 13: p. 424.

Kasinath, V., Faini, M., Poepsel, S., Reif, D., Feng, X. A., Stjepanovic, G., Aebersold, R., and Nogales, E. (2018). "Structures of human PRC2 with its cofactors AEBP2 and JARID2". *Science* 359: pp. 940–944.

Kassis, J. A. and Kennison, J. A. (2010). "Recruitment of Polycomb Complexes: a Role for SCM". *Molecular and Cellular Biology* 30: pp. 2581–2583.

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

Kaufman, T. C., Lewis, R., and Wakimoto, B. (1980). "Cytogenetic Analysis of Chromosome 3 in DROSOPHILA MELANOGASTER: The Homoeotic Gene Complex in Polytene Chromosome Interval 84a-B". eng. *Genetics* 94: pp. 115–133.

Kaustov, L., Ouyang, H., Amaya, M., Lemak, A., Nady, N., Duan, S., Wasney, G. A., Li, Z., Vedadi, M., Schapira, M., Min, J., and Arrowsmith, C. H. (2010). "Recognition and Specificity Determinants of the Human Cbx Chromodomains". *Journal of Biological Chemistry* 286: pp. 521–529.

Kebede, A. F., Schneider, R., and Daujat, S. (2014). "Novel types and sites of histone modifications emerge as players in the transcriptional regulation contest". *FEBS Journal* 282: pp. 1658–1674.

Kelemen, J. Z., Ratna, P., Scherrer, S., and Becskei, A. (2010). "Spatial Epigenetic Control of Mono- and Bistable Gene Expression". *PLoS Biology* 8. Ed. by A. Levchenko: e1000332.

Kelly, T. K., Liu, Y., Lay, F. D., Liang, G., Berman, B. P., and Jones, P. A. (2012). "Genome-wide mapping of nucleosome positioning and DNA methylation within individual DNA molecules". *Genome Research* 22: pp. 2497–2506.

Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., and Haussler, D. (2002). "The human genome browser at UCSC". *Genome Res* 12: pp. 996–1006.

Kiermaier, A., Gawn, J. M., Desbarats, L., Saffrich, R., Ansorge, W., Farrell, P. J., Eilers, M., and Packham, G. (1999). "DNA binding of USF is required for specific E-box dependent gene activation in vivo". *Oncogene* 18: pp. 7200–7211.

Kim, C. A., Gingery, M., Pilpa, R. M., and Bowie, J. U. (2002). "The SAM domain of polyhomeotic forms a helical polymer". *Nat Struct Biol* 9: pp. 453–7.

Kim, H., Kang, K., and Kim, J. (2009a). "AEBP2 as a potential targeting protein for Polycomb Repression Complex PRC2". *Nucleic Acids Research* 37: pp. 2940–2950.

Kim, J., Guermah, M., McGinty, R. K., Lee, J.-S., Tang, Z., Milne, T. A., Shilatifard, A., Muir, T. W., and Roeder, R. G. (2009b). "RAD6-Mediated Transcription-Coupled H2B Ubiquitylation Directly Stimulates H3K4 Methylation in Human Cells". *Cell* 137: pp. 459–471.

Kim, T. g., Chen, J., Sadoshima, J., and Lee, Y. (2004). "Jumonji Represses Atrial Natriuretic Factor Gene Expression by Inhibiting Transcriptional Activities of Cardiac Transcription Factors". *Molecular and Cellular Biology* 24: pp. 10151–10160.

Kim, T.-G., Kraus, J. C., Chen, J., and Lee, Y. (2003). "JUMONJI, a Critical Factor for Cardiac Development, Functions as a Transcriptional Repressor". *Journal of Biological Chemistry* 278: pp. 42247–42255.

Kim, T.-K., Hemberg, M., Gray, J. M., Costa, A. M., Bear, D. M., Wu, J., Harmin, D. A., Laptewicz, M., Barbara-Haley, K., Kuersten, S., Markenscoff-Papadimitriou, E., Kuhl, D., Bito, H., Worley, P. F., Kreiman, G., and Greenberg, M. E. (2010). "Widespread transcription at neuronal activity-regulated enhancers". *Nature* 465: pp. 182–187.

Kimelman, D. and Martin, B. L. (2011). "Anterior-posterior patterning in early development: three strategies". *Wiley Interdisciplinary Reviews: Developmental Biology* 1: pp. 253–266.

King, H. W., Fursova, N. A., Blackledge, N. P., and Klose, R. J. (2018). "Polycomb repressive complex 1 shapes the nucleosome landscape but not accessibility at target genes". *Genome Research* 28: pp. 1494–1507.

King, I. F. G., Emmons, R. B., Francis, N. J., Wild, B., Muller, J., Kingston, R. E., and Wu, C. (2005). "Analysis of a Polycomb Group Protein Defines Regions That Link Repressive Activity on Nucleosomal Templates to In Vivo Function". *Molecular and Cellular Biology* 25: pp. 6578–6591.

King, I. F. G., Francis, N. J., and Kingston, R. E. (2002). "Native and Recombinant Polycomb Group Complexes Establish a Selective Block to Template Accessibility To Repress Transcription In Vitro". *Molecular and Cellular Biology* 22: pp. 7919–7928.

Kingston, R. E., Bunker, C. A., and Imbalzano, A. N. (1996). "Repression and activation by multiprotein complexes that alter chromatin structure." *Genes & Development* 10: pp. 905–920.

Kireeva, M. L., Walter, W., Tchernajenko, V., Bondarenko, V., Kashlev, M., and Studitsky, V. M. (2002). "Nucleosome Remodeling Induced by RNA Polymerase II". *Molecular Cell* 9: pp. 541–552.

Klinkenberg, L. G., Mennella, T. A., Luetkenhaus, K., and Zitomer, R. S. (2005). "Combinatorial Repression of the Hypoxic Genes ofSaccharomyces cerevisiaeby DNA Binding Proteins Rox1 and Mot3". *Eukaryotic Cell* 4: pp. 649–660.

Klinkenberg, L. G., Webb, T., and Zitomer, R. S. (2006). "Synergy among Differentially Regulated Repressors of the Ribonucleotide Diphosphate Reductase Genes ofSaccharomyces cerevisiae". *Eukaryotic Cell* 5: pp. 1007–1017.

Kloet, S. L., Makowski, M. M., Baymaz, H. I., Voorthuijsen, L. van, Karemaker, I. D., Santanach, A., Jansen, P., Di Croce, L., and Vermeulen, M. (2016). "The dynamic interactome and genomic targets of Polycomb complexes during stem-cell differentiation". *Nat Struct Mol Biol* 23: pp. 682–690.

Klose, R. J., Cooper, S., Farcas, A. M., Blackledge, N. P., and Brockdorff, N. (2013). "Chromatin sampling–an emerging perspective on targeting polycomb repressor proteins". *PLoS Genet* 9: e1003717.

Klose, R. J., Kallin, E. M., and Zhang, Y. (2006). "JmjC-domain-containing proteins and histone demethylation". *Nature Reviews Genetics* 7: pp. 715–727.

Klymenko, T. and Müller, J. (2004). "The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins". *EMBO reports* 5: pp. 373–377.

Klymenko, T., Papp, B., Fischle, W., Köcher, T., Schelder, M., Fritsch, C., Wild, B., Wilm, M., and Müller, J. (2006). "A Polycomb group protein complex with sequence-specific DNA-binding and selective methyl-lysine-binding activities". *Genes & Development* 20: pp. 1110–1122.

Knezetic, J. A. and Luse, D. S. (1986). "The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II in vitro". *Cell* 45: pp. 95–104.

Kohlmaier, A., Savarese, F., Lachner, M., Martens, J., Jenuwein, T., and Wutz, A. (2004). "A Chromosomal Memory Triggered by Xist Regulates Histone Methylation in X Inactivation". *PLoS Biology* 2. Ed. by P. Becker: e171.

Kornberg, R. D. (1974). "Chromatin Structure: A Repeating Unit of Histones and DNA". *Science* 184: pp. 868–871.

Kornberg, R. D. (2005). "Mediator and the mechanism of transcriptional activation". *Trends in Biochemical Sciences* 30: pp. 235–239.

Kornberg, R. D. and Lorch, Y. (1999a). "Chromatin-modifying and -remodeling complexes". *Current Opinion in Genetics & Development* 9: pp. 148–151.

Kornberg, R. D. and Lorch, Y. (1999b). "Twenty-Five Years of the Nucleosome, Fundamental Particle of the Eukaryote Chromosome". *Cell* 98: pp. 285–294.

Kornberg, R. D. and Thomas, J. O. (1974). "Chromatin Structure: Oligomers of the Histones". *Science* 184: pp. 865–868.

Kostrewa, D., Zeller, M. E., Armache, K.-J., Seizl, M., Leike, K., Thomm, M., and Cramer, P. (2009). "RNA polymerase II–TFIIB structure and mechanism of transcription initiation". *Nature* 462: pp. 323–330.

Kotake, Y, Nakagawa, T, Kitagawa, K, Suzuki, S, Liu, N, Kitagawa, M, and Xiong, Y (2011). "Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15INK4B tumor suppressor gene". *Oncogene* 30: pp. 1956–1962.

Kouzarides, T. (2007). "Chromatin Modifications and Their Function". *Cell* 128: pp. 693–705.

Krepelova, A., Neri, F., Maldotti, M., Rapelli, S., and Oliviero, S. (2014). "Myc and Max Genome-Wide Binding Sites Analysis Links the Myc Regulatory Network with the Polycomb and the Core Pluripotency Networks in Mouse Embryonic Stem Cells". *PLoS ONE* 9. Ed. by S. Cotterill: e88933.

Krogan, N. J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D. P., Beattie, B. K., Emili, A., Boone, C., Shilatifard, A., Buratowski, S., and Greenblatt, J. (2003). "Methylation of Histone H3 by Set2 in Saccharomyces cerevisiae Is Linked to Transcriptional Elongation by RNA Polymerase II". *Molecular and Cellular Biology* 23: pp. 4207–4218.

Krueger, F. and Andrews, S. R. (2016). "SNPsplit: Allele-specific splitting of alignments between genomes with known SNP genotypes". *F1000Res* 5: p. 1479.

Ku, M., Koche, R. P., Rheinbay, E., Mendenhall, E. M., Endoh, M., Mikkelsen, T. S., Presser, A., Nusbaum, C., Xie, X., Chi, A. S., Adli, M., Kasif, S., Ptaszek, L. M., Cowan, C. A., Lander, E. S., Koseki, H., and Bernstein, B. E. (2008). "Genomewide Analysis of PRC1 and PRC2 Occupancy Identifies Two Classes of Bivalent Domains". *PLoS Genetics* 4. Ed. by B. van Steensel: e1000242.

Kubik, S., Bruzzone, M. J., Jacquet, P., Falcone, J.-L., Rougemont, J., and Shore, D. (2015). "Nucleosome Stability Distinguishes Two Different Promoter Types at All Protein-Coding Genes in Yeast". *Molecular Cell* 60: pp. 422–434.

Kulaeva, O. I., Gaykalova, D. A., Pestov, N. A., Golovastov, V. V., Vassylyev, D. G., Artsimovitch, I., and Studitsky, V. M. (2009). "Mechanism of chromatin remodeling and recovery during passage of RNA polymerase II". *Nature Structural & Molecular Biology* 16: pp. 1272–1278.

Kundu, S., Ji, F., Sunwoo, H., Jain, G., Lee, J. T., Sadreyev, R. I., Dekker, J., and Kingston, R. E. (2017). "Polycomb Repressive Complex 1 Generates Discrete Compacted Domains that Change during Differentiation". *Mol Cell* 65: 432–446 e5.

Kundu, T. K., Palhan, V. B., Wang, Z., An, W., Cole, P. A., and Roeder, R. G. (2000). "Activator-Dependent Transcription from Chromatin In Vitro Involving Targeted Histone Acetylation by p300". *Molecular Cell* 6: pp. 551–561.

Kuryan, B. G., Kim, J., Tran, N. N. H., Lombardo, S. R., Venkatesh, S., Workman, J. L., and Carey, M. (2012). "Histone density is maintained during transcription mediated by the chromatin remodeler RSC and histone chaperone NAP1 in vitro". *Proceedings of the National Academy of Sciences* 109: pp. 1931–1936.

Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2002). "Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein". *Genes Dev* 16: pp. 2893–905.

Kwon, H., Imbalzano, A. N., Khavari, P. A., Kingston, R. E., and Green, M. R. (1994). "Nucleosome disruption and enhancement of activator binding by a human SWI/SNF complex". *Nature* 370: pp. 477–481.

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

Lachner, M., OCarroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). "Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins". *Nature* 410: pp. 116–120.

LaFave, L. M., Béguelin, W., Koche, R., et al. (2015). "Loss of BAP1 function leads to EZH2-dependent transformation". *Nature Medicine* 21: pp. 1344–1349.

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

Landeira, D., Sauer, S., Poot, R., et al. (2010). "Jarid2 is a PRC2 component in embryonic stem cells required for multi-lineage differentiation and recruitment of PRC1 and RNA Polymerase II to developmental regulators". *Nature Cell Biology* 12: pp. 618–624.

Langmead, B. and Salzberg, S. L. (2012). "Fast gapped-read alignment with Bowtie 2". *Nat Methods* 9: pp. 357–9.

Lau, M. S., Schwartz, M. G., Kundu, S., Savol, A. J., Wang, P. I., Marr, S. K., Grau, D. J., Schorderet, P., Sadreyev, R. I., Tabin, C. J., and Kingston, R. E. (2017). "Mutation of a nucleosome compaction region disrupts Polycomb-mediated axial patterning". *Science* 355: pp. 1081–1084.

Laugesen, A., Højfeldt, J. W., and Helin, K. (2019). "Molecular Mechanisms Directing PRC2 Recruitment and H3K27 Methylation". *Molecular Cell* 74: pp. 8–18.

Lavarone, E., Barbieri, C. M., and Pasini, D. (2019). "Dissecting the role of H3K27 acetylation and methylation in PRC2 mediated control of cellular identity". *Nature Communications* 10.

Lavigne, M., Francis, N. J., King, I. F., and Kingston, R. E. (2004). "Propagation of silencing; recruitment and repression of naive chromatin in trans by polycomb repressed chromatin". *Mol Cell* 13: pp. 415–25.

Lebrecht, D., Foehr, M., Smith, E., Lopes, F. J. P., Vanario-Alonso, C. E., Reinitz, J., Burz, D. S., and Hanes, S. D. (2005). "Bicoid cooperative DNA binding is critical for embryonic patterning in Drosophila". *Proceedings of the National Academy of Sciences* 102: pp. 13176–13181.

Lee, C.-H., Holder, M., Grau, D., Saldana-Meyer, R., Yu, J.-R., Ganai, R. A., Zhang, J., Wang, M., LeRoy, G., Dobenecker, M.-W., Reinberg, D., and Armache, K.-J. (2018). "Distinct Stimulatory Mechanisms Regulate the Catalytic Activity of Polycomb Repressive Complex 2". *Molecular Cell* 70: 435–448.e5.

Lee, D. Y., Hayes, J. J., Pruss, D., and Wolffe, A. P. (1993). "A positive role for histone acetylation in transcription factor access to nucleosomal DNA". *Cell* 72: pp. 73–84.

Lee, H. G., Kahn, T. G., Simcox, A., Schwartz, Y. B., and Pirrotta, V. (2015). "Genome-wide activities of Polycomb complexes control pervasive transcription". *Genome Res* 25: pp. 1170–81.

Lee, H. S., Lee, S. A., Hur, S. K., Seo, J. W., and Kwon, J. (2014). "Stabilization and targeting of INO80 to replication forks by BAP1 during normal DNA synthesis". *Nat Commun* 5: p. 5128.

Lee, J.-H., Voo, K. S., and Skalnik, D. G. (2001). "Identification and Characterization of the DNA Binding Domain of CpG-binding Protein". *Journal of Biological Chemistry* 276: pp. 44669–44676.

Lee, M. G., Villa, R., Trojer, P., Norman, J., Yan, K.-P., Reinberg, D., Croce, L. D., and Shiekhattar, R. (2007). "Demethylation of H3K27 Regulates Polycomb Recruitment and H2A Ubiquitination". *Science* 318: pp. 447–450.

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

Lehmann, L., Ferrari, R., Vashisht, A. A., Wohlschlegel, J. A., Kurdistani, S. K., and Carey, M. (2012). "Polycomb Repressive Complex 1 (PRC1) Disassembles RNA Polymerase II Preinitiation Complexes". *Journal of Biological Chemistry* 287: pp. 35784–35794.

Lehmann, R. and Nüsslein-Volhard, C. (1987). "hunchback, a gene required for segmentation of an anterior and posterior region of the Drosophila embryo". *Developmental Biology* 119: pp. 402–417.

Lehnertz, B., Ueda, Y., Derijck, A. A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T., and Peters, A. H. (2003). "Suv39h-Mediated Histone H3 Lysine 9 Methylation Directs DNA Methylation to Major Satellite Repeats at Pericentric Heterochromatin". *Current Biology* 13: pp. 1192–1200.

Levine, M. (2011). "Paused RNA Polymerase II as a Developmental Checkpoint". *Cell* 145: pp. 502–511.

Levine, M., Cattoglio, C., and Tjian, R. (2014). "Looping Back to Leap Forward: Transcription Enters a New Era". *Cell* 157: pp. 13–25.

Levine, S. S., Weiss, A., Erdjument-Bromage, H., Shao, Z., Tempst, P., and Kingston, R. E. (2002). "The core of the polycomb repressive complex is compositionally and functionally conserved in flies and humans". *Mol Cell Biol* 22: pp. 6070–8.

Levy, A. and Noll, M. (1981). "Chromatin fine structure of active and repressed genes". *Nature* 289: pp. 198–203.

Lewis, E. B. (1978). "A gene complex controlling segmentation in Drosophila". *Nature* 276: pp. 565–70.

Lewis, P. (1947). "New mutants report". *Dros. Inf. Serv.* 21.

Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R., and Pfister, H. (2014). "UpSet: Visualization of Intersecting Sets". *IEEE Trans Vis Comput Graph* 20: pp. 1983–92.

Li, B., Jackson, J., Simon, M. D., Fleharty, B., Gogol, M., Seidel, C., Workman, J. L., and Shilatifard, A. (2009a). "Histone H3 Lysine 36 Dimethylation (H3K36me2) Is Sufficient to Recruit the Rpd3s Histone Deacetylase Complex and to Repress Spurious Transcription". *Journal of Biological Chemistry* 284: pp. 7970–7976.

Li, G., Margueron, R., Ku, M., Chambon, P., Bernstein, B. E., and Reinberg, D. (2010). "Jarid2 and PRC2, partners in regulating gene expression". *Genes & Development* 24: pp. 368–380.

Li, G. and Reinberg, D. (2011). "Chromatin higher-order structures and gene regulation". *Current Opinion in Genetics & Development* 21: pp. 175–186.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009b). "The Sequence Alignment/Map format and SAMtools". *Bioinformatics* 25: pp. 2078–9.

Li, H., Liefke, R., Jiang, J., Kurland, J. V., Tian, W., Deng, P., Zhang, W., He, Q., Patel, D. J., Bulyk, M. L., Shi, Y., and Wang, Z. (2017). "Polycomb-like proteins link the PRC2 complex to CpG islands". *Nature* 549: pp. 287–291.

Li, W., Bengtson, M. H., Ulbrich, A., Matsuda, A., Reddy, V. A., Orth, A., Chanda, S. K., Batalov, S., and Joazeiro, C. A. P. (2008). "Genome-Wide and Functional Annotation of Human E3 Ubiquitin Ligases Identifies MULAN, a Mitochondrial E3 that Regulates the Organelle's Dynamics and Signaling". *PLOS ONE* 3: e1487.

Li, Z., Cao, R., Wang, M., Myers, M. P., Zhang, Y., and Xu, R.-M. (2006). "Structure of a Bmi-1-Ring1B Polycomb Group Ubiquitin Ligase Complex". *Journal of Biological Chemistry* 281: pp. 20643–20649.

Liang, G., He, J., and Zhang, Y. (2012). "Kdm2b promotes induced pluripotent stem cell generation by facilitating gene activation early in reprogramming". *Nature Cell Biology* 14: pp. 457–466.

Liefke, R., Karwacki-Neisius, V., and Shi, Y. (2016). "EPOP Interacts with Elongin BC and USP7 to Modulate the Chromatin Landscape". *Molecular Cell* 64: pp. 659–672.

Liefke, R. and Shi, Y. (2015). "The PRC2-associated factor C17orf96 is a novel CpG island regulator in mouse ES cells". *Cell Discovery* 1.

Lin, J., Handschin, C., and Spiegelman, B. M. (2005). "Metabolic control through the PGC-1 family of transcription coactivators". *Cell Metabolism* 1: pp. 361–370.

Liu, B., Liu, Y.-F., Du, Y.-R., Mardaryev, A. N., Yang, W., Chen, H., Xu, Z.-M., Xu, C.-Q., Zhang, X.-R., Botchkarev, V. A., Zhang, Y., and Xu, G.-L. (2013). "Cbx4 regulates the proliferation of thymic epithelial cells and thymus function". *Development* 140: pp. 780–788.

Liu, X., Bushnell, D. A., Wang, D., Calero, G., and Kornberg, R. D. (2010). "Structure of an RNA Polymerase II-TFIIB Complex and the Transcription Initiation Mechanism". *Science* 327: pp. 206–209.

Long, H. K., Sims, D., Heger, A., Blackledge, N. P., Kutter, C., Wright, M. L., Grutzner, F., Odom, D. T., Patient, R., Ponting, C. P., and Klose, R. J. (2013). "Epigenetic conservation at gene regulatory elements revealed by non-methylated DNA profiling in seven vertebrates". *Elife* 2: e00348.

Lorch, Y., LaPointe, J. W., and Kornberg, R. D. (1987). "Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones". *Cell* 49: pp. 203–210.

Losick, R. (1998). "Summary: Three Decades after Sigma". *Cold Spring Harbor Symposia on Quantitative Biology* 63: pp. 653–666.

Loubiere, V., Delest, A., Thomas, A., Bonev, B., Schuettengruber, B., Sati, S., Martinez, A.-M., and Cavalli, G. (2016). "Coordinate redeployment of PRC1 proteins suppresses tumor formation during Drosophila development". *Nature Genetics* 48: pp. 1436–1442.

Loubiere, V., Martinez, A.-M., and Cavalli, G. (2019). "Cell Fate and Developmental Regulation Dynamics by Polycomb Proteins and 3D Genome Architecture". *BioEssays* 41: p. 1800222.

Louder, R. K., He, Y., López-Blanco, J. R., Fang, J., Chacón, P., and Nogales, E. (2016). "Structure of promoter-bound TFIID and model of human pre-initiation complex assembly". *Nature* 531: pp. 604–609.

Love, M. I., Huber, W., and Anders, S. (2014). "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2". *Genome Biol* 15: p. 550.

Lu, C., Jain, S. U., Hoelper, D., et al. (2016). "Histone H3K36 mutations promote sarcomagenesis through altered histone methylation landscape". *Science* 352: pp. 844–849.

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

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

Luis, N. M., Morey, L., Croce, L. D., and Benitah, S. A. (2012). "Polycomb in Stem Cells: PRC1 Branches Out". *Cell Stem Cell* 11: pp. 16–21.

Luo, Y., Fujii, H., Gerster, T., and Roeder, R. G. (1992). "A novel B cell-derived coactivator potentiates the activation of immunoglobulin promoters by octamer-binding transcription factors". *Cell* 71: pp. 231–241.

Lynch, M. D., Smith, A. J. H., Gobbi, M. D., Flenley, M., Hughes, J. R., Vernimmen, D., Ayyub, H., Sharpe, J. A., Sloane-Stanley, J. A., Sutherland, L., Meek, S., Burdon, T., Gibbons, R. J., Garrick, D., and Higgs, D. R. (2012). "An interspecies analysis reveals a key role for unmethylated CpG dinucleotides in vertebrate Polycomb complex recruitment". *The EMBO Journal* 31: pp. 317–329.

Machida, Y. J., Machida, Y., Vashisht, A. A., Wohlschlegel, J. A., and Dutta, A. (2009). "The Deubiquitinating Enzyme BAP1 Regulates Cell Growth via Interaction with HCF-1". *Journal of Biological Chemistry* 284: pp. 34179–34188.

Maeshima, K., Hihara, S., and Eltsov, M. (2010). "Chromatin structure: does the 30-nm fibre exist in vivo?" *Current Opinion in Cell Biology* 22: pp. 291–297.

Maeshima, K., Imai, R., Tamura, S., and Nozaki, T. (2014). "Chromatin as dynamic 10-nm fibers". *Chromosoma* 123: pp. 225–237.

Malik, H. S. and Henikoff, S. (2003). "Phylogenomics of the nucleosome". *Nature Structural & Molecular Biology* 10: pp. 882–891.

Malik, S. and Roeder, R. G. (2010). "The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation". *Nature Reviews Genetics* 11: pp. 761–772.

Mallo, M., Wellik, D. M., and Deschamps, J. (2010). "Hox genes and regional patterning of the vertebrate body plan". *Developmental Biology* 344: pp. 7–15.

Marfella, C. G. and Imbalzano, A. N. (2007). "The Chd family of chromatin remodelers". *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 618: pp. 30–40.

Margueron, R., Justin, N., Ohno, K., Sharpe, M. L., Son, J., Drury W. J., r., Voigt, P., Martin, S. R., Taylor, W. R., De Marco, V., Pirrotta, V., Reinberg, D., and Gamblin, S. J. (2009). "Role of the polycomb protein EED in the propagation of repressive histone marks". *Nature* 461: pp. 762–7.

Margueron, R., Li, G., Sarma, K., Blais, A., Zavadil, J., Woodcock, C. L., Dynlacht, B. D., and Reinberg, D. (2008). "Ezh1 and Ezh2 Maintain Repressive Chromatin through Different Mechanisms". *Molecular Cell* 32: pp. 503–518.

Marmorstein, R. and Trievel, R. C. (2009). "Histone modifying enzymes: Structures, mechanisms, and specificities". *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1789: pp. 58–68.

Marsman, J. and Horsfield, J. A. (2012). "Long distance relationships: Enhancer-promoter communication and dynamic gene transcription". *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1819: pp. 1217–1227.

Martens, J. H. A., OSullivan, R. J., Braunschweig, U., Opravil, S., Radolf, M., Steinlein, P., and Jenuwein, T. (2005). "The profile of repeat-associated histone lysine methylation states in the mouse epigenome". *The EMBO Journal* 24: pp. 800–812.

Masoomian, B., Shields, J. A., and Shields, C. L. (2018). "Overview of BAP1 cancer predisposition syndrome and the relationship to uveal melanoma". *Journal of Current Ophthalmology* 30: pp. 102–109.

Matsui, S. I., Seon, B. K., and Sandberg, A. A. (1979). "Disappearance of a structural chromatin protein A24 in mitosis: implications for molecular basis of chromatin condensation". *Proc Natl Acad Sci U S A* 76: pp. 6386–90.

Mavrich, T. N., Jiang, C., Ioshikhes, I. P., Li, X., Venters, B. J., Zanton, S. J., Tomsho, L. P., Qi, J., Glaser, R. L., Schuster, S. C., Gilmour, D. S., Albert, I., and Pugh, B. F. (2008). "Nucleosome organization in the Drosophila genome". *Nature* 453: pp. 358–362.

Mazo, A. M., Huang, D. H., Mozer, B. A., and Dawid, I. B. (1990). "The trithorax gene, a trans-acting regulator of the bithorax complex in Drosophila, encodes a protein with zinc-binding domains." *Proceedings of the National Academy of Sciences* 87: pp. 2112–2116.

McGinty, R. K., Henrici, R. C., and Tan, S. (2014). "Crystal structure of the PRC1 ubiquitylation module bound to the nucleosome". *Nature* 514: pp. 591–596.

Meisterernst, M., Roy, A. L., Lieu, H. M., and Roeder, R. G. (1991). "Activation of class II gene transcription by regulatory factors is potentiated by a novel activity". *Cell* 66: pp. 981–993.

Mendenhall, E. M., Koche, R. P., Truong, T., Zhou, V. W., Issac, B., Chi, A. S., Ku, M., and Bernstein, B. E. (2010). "GC-Rich Sequence Elements Recruit PRC2 in Mammalian ES Cells". *PLoS Genetics* 6. Ed. by H. D. Madhani: e1001244.

Merli, C, Bergstrom, D. E., Cygan, J. A., and Blackman, R. K. (1996). "Promoter specificity mediates the independent regulation of neighboring genes." *Genes & Development* 10: pp. 1260–1270.

Mersfelder, E. L. and Parthun, M. R. (2006). "The tale beyond the tail: histone core domain modifications and the regulation of chromatin structure". *Nucleic Acids Research* 34: pp. 2653–2662.

Mühlbacher, W., Sainsbury, S., Hemann, M., Hantsche, M., Neyer, S., Herzog, F., and Cramer, P. (2014). "Conserved architecture of the core RNA polymerase II initiation complex". *Nature Communications* 5.

Mieczkowski, J., Cook, A., Bowman, S. K., Mueller, B., Alver, B. H., Kundu, S., Deaton, A. M., Urban, J. A., Larschan, E., Park, P. J., Kingston, R. E., and Tolstorukov, M. Y. (2016). "MNase titration reveals differences between nucleosome occupancy and chromatin accessibility". *Nature Communications* 7.

Mierlo, G. van, Veenstra, G. J. C., Vermeulen, M., and Marks, H. (2019). "The Complexity of PRC2 Subcomplexes". *Trends in Cell Biology*.

Mikhaylichenko, O., Bondarenko, V., Harnett, D., Schor, I. E., Males, M., Viales, R. R., and Furlong, E. E. (2018). "The degree of enhancer or promoter activity is reflected by the levels and directionality of eRNA transcription". *Genes & Development* 32: pp. 42–57.

Mikkelsen, T. S., Ku, M., Jaffe, D. B., et al. (2007). "Genome-wide maps of chromatin state in pluripotent and lineage-committed cells". *Nature* 448: pp. 553–560.

Milne, T. A., Sinclair, D. A. R., and Brock, H. W. (1999). "The Additional sex combs gene of Drosophila is required for activation and repression of homeotic loci, and interacts specifically with Polycomb and super sex combs". *Molecular and General Genetics MGG* 261: pp. 753–761.

Min, J., Zhang, Y., and Xu, R.-M. (2003). "Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27". *Genes & Development* 17: pp. 1823–1828.

Misaghi, S., Ottosen, S., Izrael-Tomasevic, A., Arnott, D., Lamkanfi, M., Lee, J., Liu, J., O'Rourke, K., Dixit, V. M., and Wilson, A. C. (2009). "Association of C-Terminal Ubiquitin Hydrolase BRCA1-Associated Protein 1 with Cell Cycle Regulator Host Cell Factor 1". *Molecular and Cellular Biology* 29: pp. 2181–2192.

Mito, Y., Henikoff, J. G., and Henikoff, S. (2005). "Genome-scale profiling of histone H3.3 replacement patterns". *Nature Genetics* 37: pp. 1090–1097.

Mizuguchi, G., Shen, X., Landry, J., Wu, W.-H., Sen, S., and Wu, C. (2004). "ATP-Driven Exchange of Histone H2AZ Variant Catalyzed by SWR1 Chromatin Remodeling Complex". *Science* 303: p. 343.

Mizuguchi, G., Vassilev, A., Tsukiyama, T., Nakatani, Y., and Wu, C. (2001). "ATP-dependent Nucleosome Remodeling and Histone Hyperacetylation Synergistically Facilitate Transcription of Chromatin". *Journal of Biological Chemistry* 276: pp. 14773–14783.

Müller, J. and Kassis, J. A. (2006). "Polycomb response elements and targeting of Polycomb group proteins in Drosophila". *Current Opinion in Genetics & Development* 16: pp. 476–484.

Moazed, D. and OFarrell, P. H. (1992). "Maintenance of the engrailed expression pattern by Polycomb group genes in Drosophila". *Development* 116: p. 805.

Mohd-Sarip, A., Venturini, F., Chalkley, G. E., and Verrijzer, C. P. (2002). "Pleiohomeotic Can Link Polycomb to DNA and Mediate Transcriptional Repression". *Molecular and Cellular Biology* 22: pp. 7473–7483.

Mohrmann, L. and Verrijzer, C. P. (2005). "Composition and functional specificity of SWI2/SNF2 class chromatin remodeling complexes". *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* 1681: pp. 59–73.

Montgomery, N. D., Yee, D., Montgomery, S. A., and Magnuson, T. (2007). "Molecular and Functional Mapping of EED Motifs Required for PRC2-Dependent Histone Methylation". *Journal of Molecular Biology* 374: pp. 1145–1157.

Morata, G. and Kerridge, S. (1982). "The role of position in determining homoeotic gene function in Drosophila". *Nature* 300: pp. 191–192.

Moreau, P., Hen, R., Wasylyk, B., Everett, R., Gaub, M., and Chambon, P. (1981). "The SV40 72 base repair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants". *Nucleic Acids Research* 9: pp. 6047–6068.

Morey, L., Aloia, L., Cozzuto, L., Benitah, S. A., and Di Croce, L. (2013). "RYBP and Cbx7 define specific biological functions of polycomb complexes in mouse embryonic stem cells". *Cell Rep* 3: pp. 60–9.

Morey, L., Santanach, A., Blanco, E., Aloia, L., Nora, E. P., Bruneau, B. G., and Di Croce, L. (2015). "Polycomb Regulates Mesoderm Cell Fate-Specification in Embryonic Stem Cells through Activation and Repression Mechanisms". *Cell Stem Cell* 17: pp. 300–15.

Morey, L., Pascual, G., Cozzuto, L., Roma, G., Wutz, A., Benitah, S. A., and Croce, L. D. (2012). "Nonoverlapping Functions of the Polycomb Group Cbx Family of Proteins in Embryonic Stem Cells". *Cell Stem Cell* 10: pp. 47–62.

Morris, S. A., Baek, S., Sung, M.-H., John, S., Wiench, M., Johnson, T. A., Schiltz, R. L., and Hager, G. L. (2013). "Overlapping chromatin-remodeling systems collaborate genome wide at dynamic chromatin transitions". *Nature Structural & Molecular Biology* 21: pp. 73–81.

Moussa, H. F., Bsteh, D., Yelagandula, R., Pribitzer, C., Stecher, K., Bartalska, K., Michetti, L., Wang, J., Zepeda-Martinez, J. A., Elling, U., Stuckey, J. I., James, L. I., Frye, S. V., and Bell, O. (2019). "Canonical PRC1 controls sequence-independent propagation of Polycomb-mediated gene silencing". *Nature Communications* 10.

Muchardt, C. and Yaniv, M. (1993). "A human homologue of Saccharomyces cerevisiae SNF2/SWI2 and Drosophila brm genes potentiates transcriptional activation by the glucocorticoid receptor". eng. *The EMBO journal* 12: pp. 4279–4290.

Mueller, B., Mieczkowski, J., Kundu, S., Wang, P., Sadreyev, R., Tolstorukov, M. Y., and Kingston, R. E. (2017). "Widespread changes in nucleosome accessibility without changes in nucleosome occupancy during a rapid transcriptional induction". *Genes & Development* 31: pp. 451–462.

Mueller, R. D., Yasuda, H., Hatch, C. L., Bonner, W. M., and Bradbury, E. M. (1985). "Identification of ubiquitinated histones 2A and 2B in Physarum polycephalum. Disappearance of these proteins at metaphase and reappearance at anaphase." *Journal of Biological Chemistry* 260: pp. 5147–5153.

Mujtaba, S., Manzur, K. L., Gurnon, J. R., Kang, M., Etten, J. L. V., and Zhou, M.-M. (2008). "Epigenetic transcriptional repression of cellular genes by a viral SET protein". *Nature Cell Biology* 10: pp. 1114–1122.

Mulholland, N. M., King, I. F. G., and Kingston, R. E. (2003). "Regulation of Polycomb group complexes by the sequence-specific DNA binding proteins Zeste and GAGA". *Genes & Development* 17: pp. 2741–2746.

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

Murawska, M., Kunert, N., Vugt, J. van, Langst, G., Kremmer, E., Logie, C., and Brehm, A. (2008). "dCHD3, a Novel ATP-Dependent Chromatin Remodeler Associated with Sites of Active Transcription". *Molecular and Cellular Biology* 28: pp. 2745–2757.

Murray, K. (1964). "The Occurrence of e-N-Methyl Lysine in Histones". *Biochemistry* 3: pp. 10–15.

Murtha, M., Strino, F., Tokcaer-Keskin, Z., Bayin, N. S., Shalabi, D., Xi, X., Kluger, Y., and Dailey, L. (2015). "Comparative FAIRE-seq Analysis Reveals Distinguishing Features of the Chromatin Structure of Ground State- and Primed-Pluripotent Cells". *STEM CELLS* 33: pp. 378–391.

Muse, G. W., Gilchrist, D. A., Nechaev, S., Shah, R., Parker, J. S., Grissom, S. F., Zeitlinger, J., and Adelman, K. (2007). "RNA polymerase is poised for activation across the genome". *Nature Genetics* 39: pp. 1507–1511.

Musselman, C. A., Avvakumov, N., Watanabe, R., Abraham, C. G., Lalonde, M.-E., Hong, Z., Allen, C., Roy, S., Nuñez, J. K., Nickoloff, J., Kulesza, C. A., Yasui, A., Côté, J., and Kutateladze, T. G. (2012). "Molecular basis for H3K36me3 recognition by the Tudor domain of PHF1". *Nature Structural & Molecular Biology* 19: pp. 1266–1272.

Nakagawa, T., Kajitani, T., Togo, S., Masuko, N., Ohdan, H., Hishikawa, Y., Koji, T., Matsuyama, T., Ikura, T., Muramatsu, M., and Ito, T. (2008). "Deubiquitylation of histone H2A activates transcriptional initiation via trans-histone cross-talk with H3K4 di- and trimethylation". *Genes Dev* 22: pp. 37–49.

Nakayama, J.-i., Rice, J. C., Strahl, B. D., Allis, C. D., and Grewal, S. I. S. (2001). "Role of Histone H3 Lysine 9 Methylation in Epigenetic Control of Heterochromatin Assembly". *Science* 292: p. 110.

Napoles, M. de, Mermoud, J. E., Wakao, R., Tang, Y. A., Endoh, M., Appanah, R., Nesterova, T. B., Silva, J., Otte, A. P., Vidal, M., Koseki, H., and Brockdorff, N. (2004). "Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation". *Dev Cell* 7: pp. 663–76.

Nathan, D., Ingvarsdottir, K., Sterner, D. E., Bylebyl, G. R., Dokmanovic, M., Dorsey, J. A., Whelan, K. A., Krsmanovic, M., Lane, W. S., Meluh, P. B., Johnson, E. S., and Berger, S. L. (2006). "Histone sumoylation is a negative regulator in Saccharomyces cerevisiae and shows dynamic interplay with positive-acting histone modifications". eng. *Genes & development* 20: pp. 966–976.

Neely, K. E., Hassan, A. H., Brown, C. E., Howe, L., and Workman, J. L. (2002). "Transcription Activator Interactions with Multiple SWI/SNF Subunits". *Molecular and Cellular Biology* 22: pp. 1615–1625.

Neely, K. E., Hassan, A. H., Wallberg, A. E., Steger, D. J., Cairns, B. R., Wright, A. P., and Workman, J. L. (1999). "Activation Domain–Mediated Targeting of the SWI/SNF Complex to Promoters Stimulates Transcription from Nucleosome Arrays". *Molecular Cell* 4: pp. 649–655.

Nègre, N., Hennetin, J., Sun, L. V., Lavrov, S., Bellis, M., White, K. P., and Cavalli, G. (2006). "Chromosomal Distribution of PcG Proteins during Drosophila Development". *PLoS Biology* 4. Ed. by O. J. Rando: e170.

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

Nekrasov, M., Wild, B., and Müller, J. (2005). "Nucleosome binding and histone methyltransferase activity of Drosophila PRC2". *EMBO reports* 6: pp. 348–353.

Nemeth, A. and Langst, G. (2004). "Chromatin higher order structure: Opening up chromatin for transcription". *Briefings in Functional Genomics* 2: pp. 334–343.

Nesterova, T. B., Wei, G., Coker, H., Pintacuda, G., Bowness, J. S., Zhang, T., Almeida, M., Bloechl, B., Moindrot, B., Carter, E. J., Rodrigo, I. A., Pan, Q., Bi, Y., Song, C.-X., and Brockdorff, N. (2019). "Systematic allelic analysis defines the interplay of key pathways in X chromosome inactivation". *Nature Communications* 10.

Nichols, J. and Smith, A. (2010). "The origin and identity of embryonic stem cells". *Development* 138: pp. 3–8.

Nickel, B. E. and Davie, J. R. (1989). "Structure of polyubiquitinated histone H2A". *Biochemistry* 28: pp. 964–968.

Nijman, S. M., Luna-Vargas, M. P., Velds, A., Brummelkamp, T. R., Dirac, A. M., Sixma, T. K., and Bernards, R. (2005). "A Genomic and Functional Inventory of Deubiquitinating Enzymes". *Cell* 123: pp. 773–786.

Nishino, Y., Eltsov, M., Joti, Y., Ito, K., Takata, H., Takahashi, Y., Hihara, S., Frangakis, A. S., Imamoto, N., Ishikawa, T., and Maeshima, K. (2012). "Human mitotic chromosomes consist predominantly of irregularly folded nucleosome fibres without a 30-nm chromatin structure". *The EMBO Journal* 31: pp. 1644–1653.

Niwa, H. (2010). "Mouse ES cell culture system as a model of development". *Development, Growth & Differentiation* 52: pp. 275–283.

Noll, M. and Kornberg, R. D. (1977). "Action of micrococcal nuclease on chromatin and the location of histone H1". *Journal of Molecular Biology* 109: pp. 393–404.

Noordermeer, D., Leleu, M., Splinter, E., Rougemont, J., De Laat, W., and Duboule, D. (2011). "The Dynamic Architecture of Hox Gene Clusters". *Science* 334: p. 222.

Nora, E. P., Goloborodko, A., Valton, A.-L., Gibcus, J. H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L. A., and Bruneau, B. G. (2017). "Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization". *Cell* 169: 930–944.e22.

Nowak, A. J., Alfieri, C., Stirnimann, C. U., Rybin, V., Baudin, F., Ly-Hartig, N., Lindner, D., and Müller, C. W. (2011). "Chromatin-modifying Complex Component Nurf55/p55 Associates with Histones H3 and H4 and Polycomb Repressive Complex 2 Subunit Su(z)12 through Partially Overlapping Binding Sites". *Journal of Biological Chemistry* 286: pp. 23388–23396.

Nowak, S. J. and Corces, V. G. (2004). "Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation". *Trends in Genetics* 20: pp. 214–220.

Nozawa, K., Schneider, T. R., and Cramer, P. (2017). "Core Mediator structure at 3.4  extends model of transcription initiation complex". *Nature* 545: p. 248.

OCarroll, D., Erhardt, S., Pagani, M., Barton, S. C., Surani, M. A., and Jenuwein, T. (2001). "The Polycomb-Group Gene Ezh2 Is Required for Early Mouse Development". *Molecular and Cellular Biology* 21: pp. 4330–4336.

ODonnell, N., Zachara, N. E., Hart, G. W., and Marth, J. D. (2004). "Ogt-Dependent X-Chromosome-Linked Protein Glycosylation Is a Requisite Modification in Somatic Cell Function and Embryo Viability". *Molecular and Cellular Biology* 24: pp. 1680–1690.

Ogawa, H., Ishiguro, K.-i., Gaubatz, S., Livingston, D. M., and Nakatani, Y. (2002). "A Complex with Chromatin Modifiers That Occupies E2F- and Myc-Responsive Genes in G0 Cells". *Science* 296: p. 1132.

Ogawa, Y., Ono, T., Wakata, Y., Okawa, K., Tagami, H., and Shibahara, K i. (2005). "Histone variant mmacroH2A1.2 is mono-ubiquitinated at its histone domain". *Biochemical and Biophysical Research Communications* 336: pp. 204–209.

Ogiyama, Y., Schuettengruber, B., Papadopoulos, G. L., Chang, J. M., and Cavalli, G. (2018). "Polycomb-Dependent Chromatin Looping Contributes to Gene Silencing during Drosophila Development". *Mol Cell* 71: 73–88 e5.

Ogryzko, V. V., Schiltz, R., Russanova, V., Howard, B. H., and Nakatani, Y. (1996). "The Transcriptional Coactivators p300 and CBP Are Histone Acetyltransferases". *Cell* 87: pp. 953–959.

Ohtsuki, S., Levine, M., and Cai, H. N. (1998). "Different core promoters possess distinct regulatory activities in the Drosophila embryo". *Genes & Development* 12: pp. 547–556.

Oksuz, O., Narendra, V., Lee, C.-H., Descostes, N., LeRoy, G., Raviram, R., Blumenberg, L., Karch, K., Rocha, P. P., Garcia, B. A., Skok, J. A., and Reinberg, D. (2018). "Capturing the Onset of PRC2-Mediated Repressive Domain Formation". *Molecular Cell* 70: 1149–1162.e5.

Oktaba, K., Gutiérrez, L., Gagneur, J., Girardot, C., Sengupta, A. K., Furlong, E. E., and Müller, J. (2008). "Dynamic Regulation by Polycomb Group Protein Complexes Controls Pattern Formation and the Cell Cycle in Drosophila". *Developmental Cell* 15: pp. 877–889.

Okulski, H., Druck, B., Bhalerao, S., and Ringrose, L. (2011). "Quantitative analysis of polycomb response elements (PREs) at identical genomic locations distinguishes contributions of PRE sequence and genomic environment". *Epigenetics & Chromatin* 4: p. 4.

OLoghlen, A., Munoz-Cabello, A., Gaspar-Maia, A., et al. (2012). "MicroRNA Regulation of Cbx7 Mediates a Switch of Polycomb Orthologs during ESC Differentiation". *Cell Stem Cell* 10: pp. 33–46.

Orlando, D. A., Chen, M. W., Brown, V. E., Solanki, S., Choi, Y. J., Olson, E. R., Fritz, C. C., Bradner, J. E., and Guenther, M. G. (2014). "Quantitative ChIP-Seq normalization reveals global modulation of the epigenome". *Cell Rep* 9: pp. 1163–70.

Orlando, V., Jane, E. P., Chinwalla, V., Harte, P. J., and Paro, R. (1998). "Binding of trithorax and Polycomb proteins to the bithorax complex: dynamic changes during early Drosophila embryogenesis". eng. *The EMBO journal* 17: pp. 5141–5150.

Orphanides, G, Lagrange, T, and Reinberg, D (1996). "The general transcription factors of RNA polymerase II." *Genes & Development* 10: pp. 2657–2683.

Orphanides, G., LeRoy, G., Chang, C.-H., Luse, D. S., and Reinberg, D. (1998). "FACT, a Factor that Facilitates Transcript Elongation through Nucleosomes". *Cell* 92: pp. 105–116.

Orphanides, G., Wu, W.-H., Lane, W. S., Hampsey, M., and Reinberg, D. (1999). "The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins". *Nature* 400: pp. 284–288.

Ou, H. D., Phan, S., Deerinck, T. J., Thor, A., Ellisman, M. H., and O'Shea, C. C. (2017). "ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells". *Science* 357: eaag0025.

Oudet, P., Gross-Bellard, M., and Chambon, P. (1975). "Electron microscopic and biochemical evidence that chromatin structure is a repeating unit". *Cell* 4: pp. 281–300.

Palmer, D. K., ODay, K., Trong, H. L., Charbonneau, H., and Margolis, R. L. (1991). "Purification of the centromere-specific protein CENP-A and demonstration that it is a distinctive histone." *Proceedings of the National Academy of Sciences* 88: pp. 3734–3738.

Pan, H., Jia, R., Zhang, L., Xu, S., Wu, Q., Song, X., Zhang, H., Ge, S., Xu, X. L., and Fan, X. (2015). "BAP1 regulates cell cycle progression through E2F1 target genes and mediates transcriptional silencing via H2A monoubiquitination in uveal melanoma cells". *The International Journal of Biochemistry & Cell Biology* 60: pp. 176–184.

Pandey, R. R., Mondal, T., Mohammad, F., Enroth, S., Redrup, L., Komorowski, J., Nagano, T., Mancini-DiNardo, D., and Kanduri, C. (2008). "Kcnq1ot1 Antisense Noncoding RNA Mediates Lineage-Specific Transcriptional Silencing through Chromatin-Level Regulation". *Molecular Cell* 32: pp. 232–246.

Papamichos-Chronakis, M. and Peterson, C. L. (2012). "Chromatin and the genome integrity network". *Nature Reviews Genetics* 14: pp. 62–75.

Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H. C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T., Cobb, B. S., Yokomori, K., Dillon, N., Aragon, L., Fisher, A. G., and Merkenschlager, M. (2008). "Cohesins Functionally Associate with CTCF on Mammalian Chromosome Arms". *Cell* 132: pp. 422–433.

Paro, R. and Hogness, D. S. (1991). "The Polycomb protein shares a homologous domain with a heterochromatin-associated protein of Drosophila." *Proceedings of the National Academy of Sciences* 88: pp. 263–267.

Parry, T. J., Theisen, J. W. M., Hsu, J.-Y., Wang, Y.-L., Corcoran, D. L., Eustice, M., Ohler, U., and Kadonaga, J. T. (2010). "The TCT motif, a key component of an RNA polymerase II transcription system for the translational machinery". *Genes & Development* 24: pp. 2013–2018.

Pasini, D., Bracken, A. P., Hansen, J. B., Capillo, M., and Helin, K. (2007). "The Polycomb Group Protein Suz12 Is Required for Embryonic Stem Cell Differentiation". *Molecular and Cellular Biology* 27: pp. 3769–3779.

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

Pasini, D., Cloos, P. A. C., Walfridsson, J., Olsson, L., Bukowski, J.-P., Johansen, J. V., Bak, M., Tommerup, N., Rappsilber, J., and Helin, K. (2010). "JARID2 regulates binding of the Polycomb repressive complex 2 to target genes in ES cells". *Nature* 464: pp. 306–310.

Passmore, L. A. and Barford, D. (2004). "Getting into position: the catalytic mechanisms of protein ubiquitylation". *Biochemical Journal* 379: pp. 513–525.

Payvar, F., Wrange, O., Carlstedt-Duke, J., Okret, S., Gustafsson, J. A., and Yamamoto, K. R. (1981). "Purified glucocorticoid receptors bind selectively in vitro to a cloned DNA fragment whose transcription is regulated by glucocorticoids in vivo." *Proceedings of the National Academy of Sciences* 78: pp. 6628–6632.

Peng, J. C., Valouev, A., Swigut, T., Zhang, J., Zhao, Y., Sidow, A., and Wysocka, J. (2009). "Jarid2/Jumonji Coordinates Control of PRC2 Enzymatic Activity and Target Gene Occupancy in Pluripotent Cells". *Cell* 139: pp. 1290–1302.

Pengelly, A. R., Copur, O., Jackle, H., Herzig, A., and Muller, J. (2013). "A Histone Mutant Reproduces the Phenotype Caused by Loss of Histone-Modifying Factor Polycomb". *Science* 339: pp. 698–699.

Pengelly, A. R., Kalb, R., Finkl, K., and Muller, J. (2015). "Transcriptional repression by PRC1 in the absence of H2A monoubiquitylation". *Genes Dev* 29: pp. 1487–92.

Perino, M., Mierlo, G. van, Karemaker, I. D., Genesen, S. van, Vermeulen, M., Marks, H., Heeringen, S. J. van, and Veenstra, G. J. C. (2018). "MTF2 recruits Polycomb Repressive Complex 2 by helical-shape-selective DNA binding". *Nat Genet* 50: pp. 1002–1010.

Peters, A. H., Kubicek, S., Mechtler, K., O'Sullivan, R. J., Derijck, A. A., Perez-Burgos, L., Kohlmaier, A., Opravil, S., Tachibana, M., Shinkai, Y., Martens, J. H., and Jenuwein, T. (2003). "Partitioning and Plasticity of Repressive Histone Methylation States in Mammalian Chromatin". *Molecular Cell* 12: pp. 1577–1589.

Peterson, C. L. and Laniel, M.-A. (2004). "Histones and histone modifications". *Current Biology* 14: R546–R551.

Pherson, M., Misulovin, Z., Gause, M., Mihindukulasuriya, K., Swain, A., and Dorsett, D. (2017). "Polycomb repressive complex 1 modifies transcription of active genes". *Science Advances* 3: e1700944.

Pintacuda, G., Wei, G., Roustan, C., Kirmizitas, B. A., Solcan, N., Cerase, A., Castello, A., Mohammed, S., Moindrot, B., Nesterova, T. B., and Brockdorff, N. (2017). "hnRNPK Recruits PCGF3/5-PRC1 to the Xist RNA B-Repeat to Establish Polycomb-Mediated Chromosomal Silencing". *Mol Cell* 68: 955–969 e10.

Plaschka, C., Larivière, L., Wenzeck, L., Seizl, M., Hemann, M., Tegunov, D., Petrotchenko, E. V., Borchers, C. H., Baumeister, W., Herzog, F., Villa, E., and Cramer, P. (2015). "Architecture of the RNA polymerase II–Mediator core initiation complex". *Nature* 518: pp. 376–380.

Plath, K., Fang, J., Mlynarczyk-Evans, S. K., Cao, R., Worringer, K. A., Wang, H., Cruz, C. C. de la, Otte, A. P., Panning, B., and Zhang, Y. (2003). "Role of Histone H3 Lysine 27 Methylation in X Inactivation". *Science* 300: p. 131.

Plys, A. J., Davis, C. P., Kim, J., Rizki, G., Keenen, M. M., Marr, S. K., and Kingston, R. E. (2019). "Phase separation of Polycomb-repressive complex 1 is governed by a charged disordered region of CBX2". *Genes & Development*.

Pohlers, M., Truss, M., Frede, U., Scholz, A., Strehle, M., Kuban, R.-J., Hoffmann, B., Morkel, M., Birchmeier, C., and Hagemeier, C. (2005). "A Role for E2F6 in the Restriction of Male-Germ-Cell-Specific Gene Expression". *Current Biology* 15: pp. 1051–1057.

Pokholok, D. K., Harbison, C. T., Levine, S., Cole, M., Hannett, N. M., Lee, T. I., Bell, G. W., Walker, K., Rolfe, P. A., Herbolsheimer, E., Zeitlinger, J., Lewitter, F., Gifford, D. K., and Young, R. A. (2005). "Genome-wide Map of Nucleosome Acetylation and Methylation in Yeast". *Cell* 122: pp. 517–527.

Popovic, R., Martinez-Garcia, E., Giannopoulou, E. G., Zhang, Q., Zhang, Q., Ezponda, T., Shah, M. Y., Zheng, Y., Will, C. M., Small, E. C., Hua, Y., Bulic, M., Jiang, Y., Carrara, M., Calogero, R. A., Kath, W. L., Kelleher, N. L., Wang, J.-P., Elemento, O., and Licht, J. D. (2014). "Histone Methyltransferase MMSET/NSD2 Alters EZH2 Binding and Reprograms the Myeloma Epigenome through Global and Focal Changes in H3K36 and H3K27 Methylation". *PLoS Genetics* 10. Ed. by T. Milne: e1004566.

Porrua, O., Boudvillain, M., and Libri, D. (2016). "Transcription Termination: Variations on Common Themes". *Trends in Genetics* 32: pp. 508–522.

Posfai, E., Kunzmann, R., Brochard, V., Salvaing, J., Cabuy, E., Roloff, T. C., Liu, Z., Tardat, M., Lohuizen, M. van, Vidal, M., Beaujean, N., and Peters, A. H. F. M. (2012). "Polycomb function during oogenesis is required for mouse embryonic development". *Genes & Development* 26: pp. 920–932.

Poux, S., McCabe, D., and Pirrotta, V. (2001a). "Recruitment of components of Polycomb Group chromatin complexes in Drosophila". *Development* 128: p. 75.

Poux, S., Melfi, R., and Pirrotta, V. (2001b). "Establishment of Polycomb silencing requires a transient interaction between PC and ESC". *Genes Dev* 15: pp. 2509–14.

Pradhan, S., Chin, H. G., Esteve, P.-O., and Jacobsen, S. E. (2009). "SET7/9 mediated methylation of non-histone proteins in mammalian cells". *Epigenetics* 4: pp. 383–387.

Pray-Grant, M. G., Daniel, J. A., Schieltz, D., Yates, J. R., and Grant, P. A. (2005). "Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation". *Nature* 433: pp. 434–438.

Proudfoot, N. J. (2016). "Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut". *Science* 352: aad9926–aad9926.

Ptashne, M. (2005). "Regulation of transcription: from lambda to eukaryotes". *Trends in Biochemical Sciences* 30: pp. 275–279.

Ptashne, M. and Gann, A. (1997). "Transcriptional activation by recruitment". *Nature* 386: pp. 569–577.

Pugh, B. and Tjian, R. (1990). "Mechanism of transcriptional activation by Sp1: Evidence for coactivators". *Cell* 61: pp. 1187–1197.

Qin, J., Whyte, W. A., Anderssen, E., Apostolou, E., Chen, H.-H., Akbarian, S., Bronson, R. T., Hochedlinger, K., Ramaswamy, S., Young, R. A., and Hock, H. (2012). "The Polycomb Group Protein L3mbtl2 Assembles an Atypical PRC1-Family Complex that Is Essential in Pluripotent Stem Cells and Early Development". *Cell Stem Cell* 11: pp. 319–332.

Qin, J., Zhou, Z., Chen, W., Wang, C., Zhang, H., Ge, G., Shao, M., You, D., Fan, Z., Xia, H., Liu, R., and Chen, C. (2015a). "BAP1 promotes breast cancer cell proliferation and metastasis by deubiquitinating KLF5". *Nature Communications* 6.

Qin, W., Wolf, P., Liu, N., Link, S., Smets, M., Mastra, F. L., Forné, I., Pichler, G., Hörl, D., Fellinger, K., Spada, F., Bonapace, I. M., Imhof, A., Harz, H., and Leonhardt, H. (2015b). "DNA methylation requires a DNMT1 ubiquitin interacting motif (UIM) and histone ubiquitination". *Cell Research* 25: pp. 911–929.

Quinlan, A. R. (2014). "BEDTools: The Swiss-Army Tool for Genome Feature Analysis". *Curr Protoc Bioinformatics* 47: pp. 11 12 1–34.

Rai, A. N., Vargas, M. L., Wang, L., Andersen, E. F., Miller, E. L., and Simon, J. A. (2013). "Elements of the Polycomb Repressor SU(Z)12 Needed for Histone H3-K27 Methylation, the Interface with E(Z), and In Vivo Function". *Molecular and Cellular Biology* 33: pp. 4844–4856.

Raisner, R. M., Hartley, P. D., Meneghini, M. D., Bao, M. Z., Liu, C. L., Schreiber, S. L., Rando, O. J., and Madhani, H. D. (2005). "Histone Variant H2A.Z Marks the 5′ Ends of Both Active and Inactive Genes in Euchromatin". *Cell* 123: pp. 233–248.

Ramirez, F., Dundar, F., Diehl, S., Gruning, B. A., and Manke, T. (2014). "deepTools: a flexible platform for exploring deep-sequencing data". *Nucleic Acids Res* 42: W187–91.

Ramirez, J., Dege, C., Kutateladze, T. G., and Hagman, J. (2012). "MBD2 and Multiple Domains of CHD4 Are Required for Transcriptional Repression by Mi-2/NuRD Complexes". *Molecular and Cellular Biology* 32: pp. 5078–5088.

Rao, S. S., Huang, S.-C., Hilaire, B. G. S., et al. (2017). "Cohesin Loss Eliminates All Loop Domains". *Cell* 171: 305–320.e24.

Rasmussen, E. B. and Lis, J. T. (1993). "In vivo transcriptional pausing and cap formation on three Drosophila heat shock genes." *Proceedings of the National Academy of Sciences* 90: pp. 7923–7927.

Razin, S. V. and Gavrilov, A. A. (2014). "Chromatin without the 30-nm fiber". *Epigenetics* 9: pp. 653–657.

Ren, X. and Kerppola, T. K. (2011). "REST Interacts with Cbx Proteins and Regulates Polycomb Repressive Complex 1 Occupancy at RE1 Elements". *Molecular and Cellular Biology* 31: pp. 2100–2110.

Ren, X., Vincenz, C., and Kerppola, T. K. (2008). "Changes in the Distributions and Dynamics of Polycomb Repressive Complexes during Embryonic Stem Cell Differentiation". *Molecular and Cellular Biology* 28: pp. 2884–2895.

Reveron-Gomez, N., Gonzaez-Aguilera, C., Stewart-Morgan, K. R., Petryk, N., Flury, V., Graziano, S., Johansen, J. V., Jakobsen, J. S., Alabert, C., and Groth, A. (2018). "Accurate Recycling of Parental Histones Reproduces the Histone Modification Landscape during DNA Replication". *Molecular Cell* 72: 239–249.e5.

Richly, H., Rocha-Viegas, L., Ribeiro, J. D., Demajo, S., Gundem, G., Lopez-Bigas, N., Nakagawa, T., Rospert, S., Ito, T., and Di Croce, L. (2010). "Transcriptional activation of polycomb-repressed genes by ZRF1". *Nature* 468: pp. 1124–8.

Riising, E. M., Comet, I., Leblanc, B., Wu, X., Johansen, J. V., and Helin, K. (2014). "Gene silencing triggers polycomb repressive complex 2 recruitment to CpG islands genome wide". *Mol Cell* 55: pp. 347–60.

Ringrose, L. and Paro, R. (2007). "Polycomb/Trithorax response elements and epigenetic memory of cell identity". *Development* 134: pp. 223–232.

Ringrose, L. and Paro, R. (2004). "Epigenetic Regulation of Cellular Memory by the Polycomb and Trithorax Group Proteins". *Annual Review of Genetics* 38. PMID: 15568982: pp. 413–443.

Ringrose, L., Rehmsmeier, M., Dura, J.-M., and Paro, R. (2003). "Genome-Wide Prediction of Polycomb/Trithorax Response Elements in Drosophila melanogaster". *Developmental Cell* 5: pp. 759–771.

Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., Brugmann, S. A., Goodnough, L. H., Helms, J. A., Farnham, P. J., Segal, E., and Chang, H. Y. (2007). "Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs". *Cell* 129: pp. 1311–1323.

Roberts, S. G. E., Ha, I., Maldonado, E., Reinberg, D., and Green, M. R. (1993). "Interaction between an acidic activator and transcription factor TFIIB is required for transcriptional activation". *Nature* 363: pp. 741–744.

Robinson, A. K., Leal, B. Z., Chadwell, L. V., Wang, R., Ilangovan, U., Kaur, Y., Junco, S. E., Schirf, V., Osmulski, P. A., Gaczynska, M., Hinck, A. P., Demeler, B., McEwen, D. G., and Kim, C. A. (2012). "The Growth-Suppressive Function of the Polycomb Group Protein Polyhomeotic Is Mediated by Polymerization of Its Sterile Alpha Motif (SAM) Domain". *Journal of Biological Chemistry* 287: pp. 8702–8713.

Robinson, P. J. J., Fairall, L., Huynh, V. A. T., and Rhodes, D. (2006). "EM measurements define the dimensions of the "30-nm" chromatin fiber: Evidence for a compact, interdigitated structure". *Proceedings of the National Academy of Sciences* 103: pp. 6506–6511.

Robinson, P. J. and Rhodes, D. (2006). "Structure of the '30nm' chromatin fibre: A key role for the linker histone". *Current Opinion in Structural Biology* 16: pp. 336–343.

Rochette-Egly, C., Adam, S., Rossignol, M., Egly, J.-M., and Chambon, P. (1997). "Stimulation of RAR alpha activation function AF-1 through binding to the general transcription factor TFIIH and phosphorylation by CDK7". *Cell* 90: pp. 97–107.

Rochman, M., Postnikov, Y., Correll, S., Malicet, C., Wincovitch, S., Karpova, T. S., McNally, J. G., Wu, X., Bubunenko, N. A., Grigoryev, S., and Bustin, M. (2009). "The Interaction of NSBP1/HMGN5 with Nucleosomes in Euchromatin Counteracts Linker Histone-Mediated Chromatin Compaction and Modulates Transcription". *Molecular Cell* 35: pp. 642–656.

Roeder, R. G. (1996). "Nuclear RNA polymerases: Role of general initiation factors and cofactors in eukaryotic transcription". *Methods in Enzymology*. Elsevier: pp. 165–171.

Roeder, R. G. (2019). "50 years of eukaryotic transcription: an expanding universe of factors and mechanisms". *Nature Structural & Molecular Biology* 26: pp. 783–791.

Roeder, R. G. and Rutter, W. J. (1969). "Multiple Forms of DNA-dependent RNA Polymerase in Eukaryotic Organisms". *Nature* 224: pp. 234–237.

Rose, N. R., King, H. W., Blackledge, N. P., Fursova, N. A., Ember, K. J., Fischer, R., Kessler, B. M., and Klose, R. J. (2016). "RYBP stimulates PRC1 to shape chromatin-based communication between Polycomb repressive complexes". *Elife* 5.

Rougvie, A. E. and Lis, J. T. (1988). "The RNA polymerase II molecule at the 5-prime end of the uninduced hsp70 gene of D. melanogaster is transcriptionally engaged". *Cell* 54: pp. 795–804.

Rozenberg, J. M., Shlyakhtenko, A., Glass, K., Rishi, V., Myakishev, M. V., FitzGerald, P. C., and Vinson, C. (2008). "All and only CpG containing sequences are enriched in promoters abundantly bound by RNA polymerase II in multiple tissues". *BMC Genomics* 9: p. 67.

Ruthenburg, A. J., Li, H., Patel, D. J., and Allis, C. D. (2007). "Multivalent engagement of chromatin modifications by linked binding modules". *Nature Reviews Molecular Cell Biology* 8: pp. 983–994.

Sahtoe, D. D., Dijk, W. J. van, Ekkebus, R., Ovaa, H., and Sixma, T. K. (2016). "BAP1/ASXL1 recruitment and activation for H2A deubiquitination". *Nature Communications* 7.

Sakai, A., Schwartz, B. E., Goldstein, S., and Ahmad, K. (2009). "Transcriptional and Developmental Functions of the H3.3 Histone Variant in Drosophila". *Current Biology* 19: pp. 1816–1820.

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

Santa, F. D., Totaro, M. G., Prosperini, E., Notarbartolo, S., Testa, G., and Natoli, G. (2007). "The Histone H3 Lysine-27 Demethylase Jmjd3 Links Inflammation to Inhibition of Polycomb-Mediated Gene Silencing". *Cell* 130: pp. 1083–1094.

Sanulli, S., Justin, N., Teissandier, A., Ancelin, K., Portoso, M., Caron, M., Michaud, A., Lombard, B., Rocha, S. T. da, Offer, J., Loew, D., Servant, N., Wassef, M., Burlina, F., Gamblin, S. J., Heard, E., and Margueron, R. (2015). "Jarid2 Methylation via the PRC2 Complex Regulates H3K27me3 Deposition during Cell Differentiation". *Molecular Cell* 57: pp. 769–783.

Sarcinella, E., Zuzarte, P. C., Lau, P. N. I., Draker, R., and Cheung, P. (2007). "Monoubiquitylation of H2A.Z Distinguishes Its Association with Euchromatin or Facultative Heterochromatin". *Molecular and Cellular Biology* 27: pp. 6457–6468.

Sarma, K., Margueron, R., Ivanov, A., Pirrotta, V., and Reinberg, D. (2008). "Ezh2 Requires PHF1 To Efficiently Catalyze H3 Lysine 27 Trimethylation In Vivo". *Molecular and Cellular Biology* 28: pp. 2718–2731.

Sarma, K. and Reinberg, D. (Feb. 2005). "Histone variants meet their match". *Nature Reviews Molecular Cell Biology* 6: pp. 139–149.

Satijn, D. P. and Otte, A. P. (1999). "RING1 interacts with multiple Polycomb-group proteins and displays tumorigenic activity". eng. *Molecular and cellular biology* 19: pp. 57–68.

Saurin, A. J., Shiels, C., Williamson, J., Satijn, D. P., Otte, A. P., Sheer, D., and Freemont, P. S. (1998). "The human polycomb group complex associates with pericentromeric heterochromatin to form a novel nuclear domain". *J Cell Biol* 142: pp. 887–98.

Saurin, A. J., Shao, Z., Erdjument-Bromage, H., Tempst, P., and Kingston, R. E. (2001). "A Drosophila Polycomb group complex includes Zeste and dTAFII proteins". *Nature* 412: pp. 655–660.

Sauvageau, M. and Sauvageau, G. (2010). "Polycomb Group Proteins: Multi-Faceted Regulators of Somatic Stem Cells and Cancer". *Cell Stem Cell* 7: pp. 299–313.

Sawadogo, M (1985). "Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region". *Cell* 43: pp. 165–175.

Scelfo, A., Fernández-Pérez, D., Tamburri, S., Zanotti, M., Lavarone, E., Soldi, M., Bonaldi, T., Ferrari, K. J., and Pasini, D. (2019). "Functional Landscape of PCGF Proteins Reveals Both RING1A/B-Dependent-and RING1A/B-Independent-Specific Activities". *Molecular Cell* 74: 1037–1052.e7.

Scelfo, A., Piunti, A., and Pasini, D. (2014). "The controversial role of the Polycomb group proteins in transcription and cancer: how much do we not understand Polycomb proteins?" *FEBS Journal* 282: pp. 1703–1722.

Schaaf, C. A., Misulovin, Z., Gause, M., Koenig, A., Gohara, D. W., Watson, A., and Dorsett, D. (2013). "Cohesin and Polycomb Proteins Functionally Interact to Control Transcription at Silenced and Active Genes". *PLoS Genetics* 9. Ed. by G. Cavalli: e1003560.

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

Schmid, A., Fascher, K.-D., and Hörz, W. (1992). "Nucleosome disruption at the yeast PHO5 promoter upon PHO5 induction occurs in the absence of DNA replication". *Cell* 71: pp. 853–864.

Schmitges, F. W., Prusty, A. B., Faty, M., et al. (2011). "Histone Methylation by PRC2 Is Inhibited by Active Chromatin Marks". *Molecular Cell* 42: pp. 330–341.

Schoeftner, S., Sengupta, A. K., Kubicek, S., Mechtler, K., Spahn, L., Koseki, H., Jenuwein, T., and Wutz, A. (2006). "Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing". *The EMBO Journal* 25: pp. 3110–3122.

Schoenfelder, S., Sugar, R., Dimond, A., et al. (2015). "Polycomb repressive complex PRC1 spatially constrains the mouse embryonic stem cell genome". *Nat Genet* 47: pp. 1179–1186.

Schoorlemmer, J., Marcos-Gutiérrez, C., Were, F., Martínez, R., García, E., Satijn, D. P., Otte, A. P., and Vidal, M. (1997). "Ring1A is a transcriptional repressor that interacts with the Polycomb-M33 protein and is expressed at rhombomere boundaries in the mouse hindbrain". *The EMBO Journal* 16: pp. 5930–5942.

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

Schubeler, D., MacAlpine, D. M., Scalzo, D., Wirbelauer, C., Kooperberg, C., Leeuwen, F. van, Gottschling, D. E., ONeill, L. P., Turner, B. M., Delrow, J., Bell, S. P., and Groudine, M. (2004). "The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote". eng. *Genes & development* 18: pp. 1263–1271.

Schuettengruber, B., Bourbon, H.-M., Croce, L. D., and Cavalli, G. (2017). "Genome Regulation by Polycomb and Trithorax: 70 Years and Counting". *Cell* 171: pp. 34–57.

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

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

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

Schwartz, Y. B., Kahn, T. G., Stenberg, P., Ohno, K., Bourgon, R., and Pirrotta, V. (2010). "Alternative Epigenetic Chromatin States of Polycomb Target Genes". *PLoS Genetics* 6. Ed. by A. Akhtar: e1000805.

Schwartz, Y. B. and Pirrotta, V. (2013). "A new world of Polycombs: unexpected partnerships and emerging functions". *Nature Reviews Genetics* 14: pp. 853–864.

Sealy, L. and Chalkley, R. (1978). "DNA associated with hyperacetylated histone is preferentially digested by DNase I". *Nucleic Acids Research* 5: pp. 1863–1876.

Segal, E., Raveh-Sadka, T., Schroeder, M., Unnerstall, U., and Gaul, U. (2008). "Predicting expression patterns from regulatory sequence in Drosophila segmentation". *Nature* 451: pp. 535–540.

Seitan, V. C., Faure, A. J., Zhan, Y., McCord, R. P., Lajoie, B. R., Ing-Simmons, E., Lenhard, B., Giorgetti, L., Heard, E., Fisher, A. G., Flicek, P., Dekker, J., and Merkenschlager, M. (2013). "Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments". eng. *Genome research* 23: pp. 2066–2077.

Sengupta, A. K., Kuhrs, A., and Müller, J. (May 2004). "General transcriptional silencing by a Polycomb response element in Drosophila". *Development* 131: p. 1959.

Shafi, R., Iyer, S. P. N., Ellies, L. G., O'Donnell, N., Marek, K. W., Chui, D., Hart, G. W., and Marth, J. D. (2000). "The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny". *Proceedings of the National Academy of Sciences* 97: pp. 5735–5739.

Shankar, S. R., Bahirvani, A. G., Rao, V. K., Bharathy, N., Ow, J. R., and Taneja, R. (2013). "G9a, a multipotent regulator of gene expression". *Epigenetics* 8: pp. 16–22.

Shao, Z., Raible, F., Mollaaghababa, R., Guyon, J. R., Wu, C. T., Bender, W., and Kingston, R. E. (1999). "Stabilization of chromatin structure by PRC1, a Polycomb complex". *Cell* 98: pp. 37–46.

Sharpe, J., Nonchev, S., Gould, A., Whiting, J., and Krumlauf, R. (1998). "Selectivity, sharing and competitive interactions in the regulation of Hoxb genes". *The EMBO Journal* 17: pp. 1788–1798.

Shearn, A., Hersperger, E., and Hersperger, G. (1987). "Genetic studies of mutations at two loci of Drosophila melanogaster which cause a wide variety of homeotic transformations". *Roux's Archives of Developmental Biology* 196: pp. 231–242.

Shen, X., Kim, W., Fujiwara, Y., Simon, M. D., Liu, Y., Mysliwiec, M. R., Yuan, G.-C., Lee, Y., and Orkin, S. H. (2009). "Jumonji Modulates Polycomb Activity and Self-Renewal versus Differentiation of Stem Cells". *Cell* 139: pp. 1303–1314.

Shen, X., Liu, Y., Hsu, Y.-J., Fujiwara, Y., Kim, J., Mao, X., Yuan, G.-C., and Orkin, S. H. (2008). "EZH1 Mediates Methylation on Histone H3 Lysine 27 and Complements EZH2 in Maintaining Stem Cell Identity and Executing Pluripotency". *Molecular Cell* 32: pp. 491–502.

Shen, X. and Gorovsky, M. A. (1996). "Linker Histone H1 Regulates Specific Gene Expression but Not Global Transcription In Vivo". *Cell* 86: pp. 475–483.

Shilatifard, A. (2006). "Chromatin Modifications by Methylation and Ubiquitination: Implications in the Regulation of Gene Expression". *Annual Review of Biochemistry* 75: pp. 243–269.

Shlyueva, D., Stampfel, G., and Stark, A. (2014). "Transcriptional enhancers: from properties to genome-wide predictions". *Nature Reviews Genetics* 15: pp. 272–286.

Shogren-Knaak, M., Ishii, H., Sun, J.-M., Pazin, M. J., Davie, J. R., and Peterson, C. L. (2006). "Histone H4-K16 Acetylation Controls Chromatin Structure and Protein Interactions". *Science* 311: p. 844.

Simic, R., Lindstrom, D. L., Tran, H. G., Roinick, K. L., Costa, P. J., Johnson, A. D., Hartzog, G. A., and Arndt, K. M. (2003). "Chromatin remodeling protein Chd1 interacts with transcription elongation factors and localizes to transcribed genes". eng. *The EMBO journal* 22: pp. 1846–1856.

Simon, J., Chiang, A., and Bender, W. (1992). "Ten different Polycomb group genes are required for spatial control of the abdA and AbdB homeotic products". *Development* 114: p. 493.

Simon, J. A. and Kingston, R. E. (2013). "Occupying chromatin: Polycomb mechanisms for getting to genomic targets, stopping transcriptional traffic, and staying put". *Mol Cell* 49: pp. 808–24.

Simon, J., Chiang, A., Bender, W., Shimell, M. J., and O'Connor, M. (1993). "Elements of the Drosophila Bithorax Complex That Mediate Repression by Polycomb Group Products". *Developmental Biology* 158: pp. 131–144.

Simon, J. A. and Kingston, R. E. (2009). "Mechanisms of Polycomb gene silencing: knowns and unknowns". *Nature Reviews Molecular Cell Biology* 10: pp. 697–708.

Sims, R. J., Chen, C.-F., Santos-Rosa, H., Kouzarides, T., Patel, S. S., and Reinberg, D. (2005). "Human but Not Yeast CHD1 Binds Directly and Selectively to Histone H3 Methylated at Lysine 4 via Its Tandem Chromodomains". *Journal of Biological Chemistry* 280: pp. 41789–41792.

Sinclair, D. A., Milne, T. A., Hodgson, J. W., Shellard, J., Salinas, C. A., Kyba, M., Randazzo, F., and Brock, H. W. (1998). "The Additional sex combs gene of Drosophila encodes a chromatin protein that binds to shared and unique Polycomb group sites on polytene chromosomes". *Development* 125: p. 1207.

Sing, A., Pannell, D., Karaiskakis, A., Sturgeon, K., Djabali, M., Ellis, J., Lipshitz, H. D., and Cordes, S. P. (2009). "A Vertebrate Polycomb Response Element Governs Segmentation of the Posterior Hindbrain". *Cell* 138: pp. 885–897.

Smits, A. H., Jansen, P. W. T. C., Poser, I., Hyman, A. A., and Vermeulen, M. (2013). "Stoichiometry of chromatin-associated protein complexes revealed by label-free quantitative mass spectrometry-based proteomics". *Nucleic Acids Research* 41: e28–e28.

Sneppen, K. and Ringrose, L. (2019). "Theoretical analysis of Polycomb-Trithorax systems predicts that poised chromatin is bistable and not bivalent". *Nature Communications* 10.

Son, J., Shen, S. S., Margueron, R., and Reinberg, D. (2013). "Nucleosome-binding activities within JARID2 and EZH1 regulate the function of PRC2 on chromatin". *Genes & Development* 27: pp. 2663–2677.

Soto, M. C., Chou, T. B., and Bender, W. (1995). "Comparison of germline mosaics of genes in the Polycomb group of Drosophila melanogaster." *Genetics* 140: p. 231.

Sowa, M. E., Bennett, E. J., Gygi, S. P., and Harper, J. W. (2009). "Defining the Human Deubiquitinating Enzyme Interaction Landscape". *Cell* 138: pp. 389–403.

Sowpati, D. T., Ramamoorthy, S., and Mishra, R. K. (2015). "Expansion of the polycomb system and evolution of complexity". *Mech Dev* 138 Pt 2: pp. 97–112.

Spitz, F. and Furlong, E. E. M. (2012). "Transcription factors: from enhancer binding to developmental control". *Nature Reviews Genetics* 13: pp. 613–626.

Splinter, E., Heath, H., Kooren, J., Palstra, R.-J., Klous, P., Grosveld, F., Galjart, N., and Laat, W. de (2006). "CTCF mediates long-range chromatin looping and local histone modification in the b-globin locus". *Genes & Development* 20: pp. 2349–2354.

Srinivasan, S. (2005). "The Drosophila trithorax group protein Kismet facilitates an early step in transcriptional elongation by RNA Polymerase II". *Development* 132: pp. 1623–1635.

Šrobár, F. (2008). "Synergy of feedback mechanisms in gene regulation systems with promoter and repressor transcription factors". *Open Physics* 6.

Steffen, P. A. and Ringrose, L. (2014). "What are memories made of? How Polycomb and Trithorax proteins mediate epigenetic memory". *Nature Reviews Molecular Cell Biology* 15: pp. 340–356.

Steger, D. J., Lefterova, M. I., Ying, L., Stonestrom, A. J., Schupp, M., Zhuo, D., Vakoc, A. L., Kim, J.-E., Chen, J., Lazar, M. A., Blobel, G. A., and Vakoc, C. R. (2008). "DOT1L/KMT4 recruitment and H3K79 methylation are ubiquitously coupled with gene transcription in mammalian cells". eng. *Molecular and cellular biology* 28: pp. 2825–2839.

Sterner, D. E. and Berger, S. L. (2000). "Acetylation of Histones and Transcription-Related Factors". *Microbiology and Molecular Biology Reviews* 64: pp. 435–459.

Stielow, B., Finkernagel, F., Stiewe, T., Nist, A., and Suske, G. (2018). "MGA, L3MBTL2 and E2F6 determine genomic binding of the non-canonical Polycomb repressive complex PRC1.6". *PLoS Genet* 14: e1007193.

Stielow, C., Stielow, B., Finkernagel, F., Scharfe, M., Jarek, M., and Suske, G. (2013). "SUMOylation of the polycomb group protein L3MBTL2 facilitates repression of its target genes". *Nucleic Acids Research* 42: pp. 3044–3058.

Stock, J. K., Giadrossi, S., Casanova, M., Brookes, E., Vidal, M., Koseki, H., Brockdorff, N., Fisher, A. G., and Pombo, A. (2007). "Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells". *Nature Cell Biology* 9: pp. 1428–1435.

Stojic, L., Jasencakova, Z., Prezioso, C., Stützer, A., Bodega, B., Pasini, D., Klingberg, R., Mozzetta, C., Margueron, R., Puri, P. L., Schwarzer, D., Helin, K., Fischle, W., and Orlando, V. (2011). "Chromatin regulated interchange between polycomb repressive complex 2 (PRC2)-Ezh2 and PRC2-Ezh1 complexes controls myogenin activation in skeletal muscle cells". *Epigenetics & Chromatin* 4: p. 16.

Streubel, G., Watson, A., Jammula, S. G., Scelfo, A., Fitzpatrick, D. J., Oliviero, G., McCole, R., Conway, E., Glancy, E., Negri, G. L., Dillon, E., Wynne, K., Pasini, D., Krogan, N. J., Bracken, A. P., and Cagney, G. (2018). "The H3K36me2 Methyltransferase Nsd1 Demarcates PRC2-Mediated H3K27me2 and H3K27me3 Domains in Embryonic Stem Cells". *Molecular Cell* 70: 371–379.e5.

Struhl, G. and Akam, M. (1985). "Altered distributions of Ultrabithorax transcripts in extra sex combs mutant embryos of Drosophila". eng. *The EMBO journal* 4: pp. 3259–3264.

Struhl, G. (1981a). "A gene product required for correct initiation of segmental determination in Drosophila". *Nature* 293: pp. 36–41.

Struhl, G. (1981b). "A homoeotic mutation transforming leg to antenna in Drosophila". *Nature* 292: pp. 635–638.

Struhl, K. (2001). "GENE REGULATION: A Paradigm for Precision". *Science* 293: pp. 1054–1055.

Struhl, K. (1999). "Fundamentally Different Logic of Gene Regulation in Eukaryotes and Prokaryotes". *Cell* 98: pp. 1–4.

Studitsky, V (2004). "Chromatin remodeling by RNA polymerases". *Trends in Biochemical Sciences* 29: pp. 127–135.

Subramanian, V., Mazumder, A., Surface, L. E., Butty, V. L., Fields, P. A., Alwan, A., Torrey, L., Thai, K. K., Levine, S. S., Bathe, M., and Boyer, L. A. (2013). "H2A.Z Acidic Patch Couples Chromatin Dynamics to Regulation of Gene Expression Programs during ESC Differentiation". *PLoS Genetics* 9. Ed. by H. D. Madhani: e1003725.

Sudarsanam, P. (1999). "The nucleosome remodeling complex, Snf/Swi, is required for the maintenance of transcription invivo and is partially redundant with the histone acetyltransferase, Gcn5". *The EMBO Journal* 18: pp. 3101–3106.

Surface, L. E., Fields, P. A., Subramanian, V., Behmer, R., Udeshi, N., Peach, S. E., Carr, S. A., Jaffe, J. D., and Boyer, L. A. (2016). "H2A.Z.1 Monoubiquitylation Antagonizes BRD2 to Maintain Poised Chromatin in ESCs". *Cell Reports* 14: pp. 1142–1155.

Suzuki, A., Hirasaki, M., Hishida, T., Wu, J., Okamura, D., Ueda, A., Nishimoto, M., Nakachi, Y., Mizuno, Y., Okazaki, Y., Matsui, Y., Izpisua Belmonte, J. C., and Okuda, A. (2016). "Loss of MAX results in meiotic entry in mouse embryonic and germline stem cells". *Nat Commun* 7: p. 11056.

Swaminathan, J., Baxter, E. M., and Corces, V. G. (2005). "The role of histone H2Av variant replacement and histone H4 acetylation in the establishment of Drosophila heterochromatin". *Genes & Development* 19: pp. 65–76.

Swiezewski, S., Liu, F., Magusin, A., and Dean, C. (2009). "Cold-induced silencing by long antisense transcripts of an Arabidopsis Polycomb target". *Nature* 462: pp. 799–802.

Szerlong, H. J. and Hansen, J. C. (2011). "Nucleosome distribution and linker DNA: connecting nuclear function to dynamic chromatin structure." *Biochemistry and Cell Biology* 89: pp. 24–34.

Tagami, H., Ray-Gallet, D., Almouzni, G., and Nakatani, Y. (2004). "Histone H3.1 and H3.3 Complexes Mediate Nucleosome Assembly Pathways Dependent or Independent of DNA Synthesis". *Cell* 116: pp. 51–61.

Taherbhoy, A. M., Huang, O. W., and Cochran, A. G. (2015). "BMI1-RING1B is an autoinhibited RING E3 ubiquitin ligase". *Nat Commun* 6: p. 7621.

Takahata, S., Yu, Y., and Stillman, D. J. (2009). "FACT and Asf1 Regulate Nucleosome Dynamics and Coactivator Binding at the HO Promoter". *Molecular Cell* 34: pp. 405–415.

Takihara, Y., Tomotsune, D., Shirai, M., Katoh-Fukui, Y., Nishii, K., Motaleb, M., Nomura, M., Tsuchiya, R., Fujita, Y., Shibata, Y., Higashinakagawa, T., and Shimada, K. (1997). "Targeted disruption of the mouse homologue of the Drosophila polyhomeotic gene leads to altered anteroposterior patterning and neural crest defects". *Development* 124: pp. 3673–3682.

Talbert, P. B., Meers, M. P., and Henikoff, S. (2019). "Old cogs, new tricks: the evolution of gene expression in a chromatin context". *Nature Reviews Genetics* 20: pp. 283–297.

Tamburri, S., Lavarone, E., Fernandez-Perez, D., Zanotti, M., Manganaro, D., Conway, E., and Pasini, D. (2019). "Histone H2AK119 Mono-Ubiquitination is Essential for Polycomb-Mediated Transcriptional Repression".

Tanay, A., ODonnell, A. H., Damelin, M., and Bestor, T. H. (2007). "Hyperconserved CpG domains underlie Polycomb-binding sites". *Proceedings of the National Academy of Sciences* 104: pp. 5521–5526.

Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J., and Prins, P. (2015). "Sambamba: fast processing of NGS alignment formats". *Bioinformatics* 31: pp. 2032–4.

Taruttis, F., Feist, M., Schwarzfischer, P., Gronwald, W., Kube, D., Spang, R., and Engelmann, J. C. (2017). "External calibration with Drosophila whole-cell spike-ins delivers absolute mRNA fold changes from human RNA-Seq and qPCR data". *Biotechniques* 62: pp. 53–61.

Tatavosian, R., Kent, S., Brown, K., Yao, T., Duc, H. N., Huynh, T. N., Zhen, C. Y., Ma, B., Wang, H., and Ren, X. (2018). "Nuclear condensates of the Polycomb protein chromobox 2 (CBX2) assemble through phase separation". *Journal of Biological Chemistry* 294: pp. 1451–1463.

Tautz, D. (2004). "Segmentation". *Developmental Cell* 7: pp. 301–312.

Tavares, L., Dimitrova, E., Oxley, D., Webster, J., Poot, R., Demmers, J., Bezstarosti, K., Taylor, S., Ura, H., Koide, H., Wutz, A., Vidal, M., Elderkin, S., and Brockdorff, N. (2012). "RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3". *Cell* 148: pp. 664–78.

Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D., and Patel, D. J. (2007). "How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers". *Nature Structural & Molecular Biology* 14: pp. 1025–1040.

Taylor, G. C. A., Eskeland, R., Hekimoglu-Balkan, B., Pradeepa, M. M., and Bickmore, W. A. (2013). "H4K16 acetylation marks active genes and enhancers of embryonic stem cells, but does not alter chromatin compaction". *Genome Research* 23: pp. 2053–2065.

Thakar, A., Gupta, P., McAllister, W. T., and Zlatanova, J. (2010). "Histone Variant H2A.Z Inhibits Transcription in Reconstituted Nucleosomes". *Biochemistry* 49: pp. 4018–4026.

Thoma, F., Koller, T., and Klug, A. (1979). "Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin." *J Cell Biol* 83: p. 403.

Thomas, M. C. and Chiang, C.-M. (2006). "The General Transcription Machinery and General Cofactors". *Critical Reviews in Biochemistry and Molecular Biology* 41: pp. 105–178.

Thompson, C. M., Koleske, A. J., Chao, D. M., and Young, R. A. (1993). "A multisubunit complex associated with the RNA polymerase II CTD and TATA-binding protein in yeast". *Cell* 73: pp. 1361–1375.

Tie, F., Banerjee, R., Fu, C., Stratton, C. A., Fang, M., and Harte, P. J. (2016). "Polycomb inhibits histone acetylation by CBP by binding directly to its catalytic domain". *Proceedings of the National Academy of Sciences* 113: E744–E753.

Tillib, S., Petruk, S., Sedkov, Y., Kuzin, A., Fujioka, M., Goto, T., and Mazo, A. (1999). "Trithorax- and Polycomb-Group Response Elements within an UltrabithoraxTranscription Maintenance

Unit Consist of Closely Situated but Separable Sequences". *Molecular and Cellular Biology* 19: pp. 5189–5202.

Tolhuis, B., Muijrers, I., Wit, E. de, Teunissen, H., Talhout, W., Steensel, B. van, and Lohuizen, M. van (2006). "Genome-wide profiling of PRC1 and PRC2 Polycomb chromatin binding in Drosophila melanogaster". *Nature Genetics* 38: pp. 694–699.

Tripoulas, N. A., Hersperger, E., La Jeunesse, D., and Shearn, A. (1994). "Molecular genetic analysis of the Drosophila melanogaster gene absent, small or homeotic discs1 (ash1)". eng. *Genetics* 137: pp. 1027–1038.

Trojer, P., Cao, A. R., Gao, Z., Li, Y., Zhang, J., Xu, X., Li, G., Losson, R., Erdjument-Bromage, H., Tempst, P., Farnham, P. J., and Reinberg, D. (2011). "L3MBTL2 Protein Acts in Concert with PcG Protein-Mediated Monoubiquitination of H2A to Establish a Repressive Chromatin Structure". *Molecular Cell* 42: pp. 438–450.

Tropberger, P. and Schneider, R. (2010). "Going global: Novel histone modifications in the globular domain of H3". *Epigenetics* 5: pp. 112–117.

Tsompana, M. and Buck, M. J. (2014). "Chromatin accessibility: a window into the genome". *Epigenetics & Chromatin* 7.

Tsuboi, M., Kishi, Y., Yokozeki, W., Koseki, H., Hirabayashi, Y., and Gotoh, Y. (2018). "Ubiquitination-Independent Repression of PRC1 Targets during Neuronal Fate Restriction in the Developing Mouse Neocortex". *Developmental Cell* 47: 758–772.e5.

Tsukada, Y. ichi, Fang, J., Erdjument-Bromage, H., Warren, M. E., Borchers, C. H., Tempst, P., and Zhang, Y. (2005). "Histone demethylation by a family of JmjC domain-containing proteins". *Nature* 439: pp. 811–816.

Turberfield, A. H., Kondo, T., Nakayama, M., Koseki, Y., King, H. W., Koseki, H., and Klose, R. J. (2019). "KDM2 proteins constrain transcription from CpG island gene promoters independently of their histone demethylase activity". *Nucleic Acids Research*.

Turner, S. A. and Bracken, A. P. (2013). "A "Complex" Issue: Deciphering the Role of Variant PRC1 in ESCs". *Cell Stem Cell* 12: pp. 145–146.

Uckelmann, M. and Sixma, T. K. (2017). "Histone ubiquitination in the DNA damage response". *DNA Repair (Amst)* 56: pp. 92–101.

Utley, R. T., Ikeda, K., Grant, P. A., Cote, J., Steger, D. J., Eberharter, A., John, S., and Workman, J. L. (1998). "Transcriptional activators direct histone acetyltransferase complexes to nucleosomes". *Nature* 394: pp. 498–502.

Vakoc, C. R., Mandat, S. A., Olenchock, B. A., and Blobel, G. A. (2005). "Histone H3 Lysine 9 Methylation and HP1 gamma Are Associated with Transcription Elongation through Mammalian Chromatin". *Molecular Cell* 19: pp. 381–391.

Val, S. D., Chi, N. C., Meadows, S. M., Minovitsky, S., Anderson, J. P., Harris, I. S., Ehlers, M. L., Agarwal, P., Visel, A., Xu, S.-M., Pennacchio, L. A., Dubchak, I., Krieg, P. A., Stainier, D. Y., and Black, B. L. (2008). "Combinatorial Regulation of Endothelial Gene Expression by Ets and Forkhead Transcription Factors". *Cell* 135: pp. 1053–1064.

Vandamme, J., Völkel, P., Rosnoblet, C., Faou, P. L., and Angrand, P.-O. (2011). "Interaction Proteomics Analysis of Polycomb Proteins Defines Distinct PRC1 Complexes in Mammalian Cells". *Molecular & Cellular Proteomics* 10: p. M110.002642.

Veitia, R. A. (2003). "A sigmoidal transcriptional response: cooperativity, synergy and dosage effects". *Biological Reviews of the Cambridge Philosophical Society* 78: pp. 149–170.

Verschure, P. J., Kraan, I. van der, Leeuw, W. de, Vlag, J. van der, Carpenter, A. E., Belmont, A. S., and Driel, R. van (2005). "In Vivo HP1 Targeting Causes Large-Scale Chromatin Condensation and Enhanced Histone Lysine Methylation". *Molecular and Cellular Biology* 25: pp. 4552–4564.

Vettese-Dadey, M., Grant, P. A., Hebbes, T. R., Crane-Robinson, C., Allis, C. D., and Workman, J. L. (1996). "Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro." *The EMBO Journal* 15: pp. 2508–2518.

Vidal, M. (2019). "Polycomb Assemblies Multitask to Regulate Transcription". *Epigenomes* 3: p. 12.

Vidal, M. and Starowicz, K. (2017). "Polycomb complexes PRC1 and their function in hematopoiesis". *Experimental Hematology* 48: pp. 12–31.

Vidali, G., Boffa, L. C., Bradbury, E. M., and Allfrey, V. G. (1978). "Butyrate suppression of histone deacetylation leads to accumulation of multiacetylated forms of histones H3 and H4 and increased DNase I sensitivity of the associated DNA sequences." *Proceedings of the National Academy of Sciences* 75: pp. 2239–2243.

Vieux-Rochas, M., Fabre, P. J., Leleu, M., Duboule, D., and Noordermeer, D. (2015). "Clustering of mammalian Hox genes with other H3K27me3 targets within an active nuclear domain". *Proc Natl Acad Sci U S A* 112: pp. 4672–7.

Vincenz, C. and Kerppola, T. K. (2008). "Different polycomb group CBX family proteins associate with distinct regions of chromatin using nonhomologous protein sequences". *Proceedings of the National Academy of Sciences* 105: pp. 16572–16577.

Vissers, J. H., Nicassio, F., Lohuizen, M. van, Fiore, P. D., and Citterio, E. (2008). "The many faces of ubiquitinated histone H2A: insights from the DUBs". *Cell Division* 3: p. 8.

Voigt, P., LeRoy, G., Drury, W. J., Zee, B. M., Son, J., Beck, D. B., Young, N. L., Garcia, B. A., and Reinberg, D. (2012). "Asymmetrically Modified Nucleosomes". *Cell* 151: pp. 181–193.

Voncken, J. W., Roelen, B. A. J., Roefs, M., Vries, S. de, Verhoeven, E., Marino, S., Deschamps, J., and Lohuizen, M. van (2003). "Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition". *Proceedings of the National Academy of Sciences* 100: pp. 2468–2473.

Voo, K. S., Carlone, D. L., Jacobsen, B. M., Flodin, A., and Skalnik, D. G. (2000). "Cloning of a Mammalian Transcriptional Activator That Binds Unmethylated CpG Motifs and Shares a CXXC Domain with DNA Methyltransferase, Human Trithorax, and Methyl-CpG Binding Domain Protein 1". *Molecular and Cellular Biology* 20: pp. 2108–2121.

Voss, T. C. and Hager, G. L. (2013). "Dynamic regulation of transcriptional states by chromatin and transcription factors". *Nature Reviews Genetics* 15: pp. 69–81.

Wachter, E., Quante, T., Merusi, C., Arczewska, A., Stewart, F., Webb, S., and Bird, A. (2014). "Synthetic CpG islands reveal DNA sequence determinants of chromatin structure". *eLife* 3.

Wade, P. A., Pruss, D., and Wolffe, A. P. (1997). "Histone acetylation: chromatin in action". *Trends in Biochemical Sciences* 22: pp. 128–132.

Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R. S., and Zhang, Y. (2004a). "Role of histone H2A ubiquitination in Polycomb silencing". *Nature* 431: pp. 873–8.

Wang, J., Ellwood, K., Lehman, A., Carey, M. F., and She, Z.-S. (1999). "A mathematical model for synergistic eukaryotic gene activation". *Journal of Molecular Biology* 286: pp. 315–325.

Wang, L., Brown, J. L., Cao, R., Zhang, Y., Kassis, J. A., and Jones, R. S. (2004b). "Hierarchical recruitment of polycomb group silencing complexes". *Mol Cell* 14: pp. 637–46.

Wang, L., Zhao, Z., Ozark, P. A., Fantini, D., Marshall, S. A., Rendleman, E. J., Cozzolino, K. A., Louis, N., He, X., Morgan, M. A., Takahashi, Y. hei, Collings, C. K., Smith, E. R., Ntziachristos, P., Savas, J. N., Zou, L., Hashizume, R., Meeks, J. J., and Shilatifard, A. (2018a). "Resetting the epigenetic balance of Polycomb and COMPASS function at enhancers for cancer therapy". *Nature Medicine* 24: pp. 758–769.

Wang, Q., Geng, Z., Gong, Y., Warren, K., Zheng, H., Imamura, Y., and Gao, Z. (2018b). "WDR68 is essential for the transcriptional activation of the PRC1-AUTS2 complex and neuronal differentiation of mouse embryonic stem cells". *Stem Cell Research* 33: pp. 206–214.

Wang, R., Taylor, A. B., Leal, B. Z., Chadwell, L. V., Ilangovan, U., Robinson, A. K., Schirf, V., Hart, P. J., Lafer, E. M., Demeler, B., Hinck, A. P., McEwen, D. G., and Kim, C. A.

(2010). "Polycomb Group Targeting through Different Binding Partners of RING1B C-Terminal Domain". *Structure* 18: pp. 966–975.

Wang, Y., Long, H., Yu, J., Dong, L., Wassef, M., Zhuo, B., Li, X., Zhao, J., Wang, M., Liu, C., Wen, Z., Chang, L., Chen, P., Wang, Q. fei, Xu, X., Margueron, R., and Li, G. (2018c). "Histone variants H2A.Z and H3.3 coordinately regulate PRC2-dependent H3K27me3 deposition and gene expression regulation in mES cells". *BMC Biology* 16.

Wani, A. H., Boettiger, A. N., Schorderet, P., Ergun, A., Munger, C., Sadreyev, R. I., Zhuang, X., Kingston, R. E., and Francis, N. J. (2016). "Chromatin topology is coupled to Polycomb group protein subnuclear organization". *Nat Commun* 7: p. 10291.

Weake, V. M. and Workman, J. L. (2008). "Histone Ubiquitination: Triggering Gene Activity". *Molecular Cell* 29: pp. 653–663.

Weber, C. M., Ramachandran, S., and Henikoff, S. (2014). "Nucleosomes Are Context-Specific, H2A.Z-Modulated Barriers to RNA Polymerase". *Molecular Cell* 53: pp. 819–830.

Weiner, A., Hughes, A., Yassour, M., Rando, O. J., and Friedman, N. (2009). "High-resolution nucleosome mapping reveals transcription-dependent promoter packaging". *Genome Research* 20: pp. 90–100.

Weinmann, R., Raskas, H. J., and Roeder, R. G. (1974). "Role of DNA-Dependent RNA Polymerases II and III in Transcription of the Adenovirus Genome Late in Productive Infection". *Proceedings of the National Academy of Sciences* 71: pp. 3426–3430.

Weintraub, H and Groudine, M (1976). "Chromosomal subunits in active genes have an altered conformation". *Science* 193: pp. 848–856.

Whitehouse, I., Rando, O. J., Delrow, J., and Tsukiyama, T. (2007). "Chromatin remodelling at promoters suppresses antisense transcription". *Nature* 450: pp. 1031–1035.

Widom, J (1989). "Toward a Unified Model of Chromatin Folding". *Annual Review of Biophysics and Biophysical Chemistry* 18: pp. 365–395.

Wiesner, T., Obenauf, A. C., Murali, R., et al. (2011). "Germline mutations in BAP1 predispose to melanocytic tumors". *Nature Genetics* 43: pp. 1018–1021.

Wong, S. J., Gearhart, M. D., Taylor, A. B., Nanyes, D. R., Ha, D. J., Robinson, A. K., Artigas, J. A., Lee, O. J., Demeler, B., Hart, P. J., Bardwell, V. J., and Kim, C. A. (2016). "KDM2B Recruitment of the Polycomb Group Complex, PRC1.1, Requires Cooperation between PCGF1 and BCORL1". *Structure* 24: pp. 1795–1801.

Woo, C. J., Kharchenko, P. V., Daheron, L., Park, P. J., and Kingston, R. E. (2010). "A Region of the Human HOXD Cluster that Confers Polycomb-Group Responsiveness". *Cell* 140: pp. 99–110.

Workman, J. L. and Kingston, R. E. (1998). "Alteration of nucleosome structure as a mechanism of transcriptional regulation". *Annual Review of Biochemistry* 67: pp. 545–579.

Wu, C. (1980). "The 5′ ends of Drosophila heat shock genes in chromatin are hypersensitive to DNase I". *Nature* 286: pp. 854–860.

Wu, C., Wong, Y.-C., and Elgin, S. C. (1979). "The chromatin structure of specific genes: II. Disruption of chromatin structure during gene activity". *Cell* 16: pp. 807–814.

Wu, X., Johansen, J. V., and Helin, K. (2013). "Fbxl10/Kdm2b recruits polycomb repressive complex 1 to CpG islands and regulates H2A ubiquitylation". *Mol Cell* 49: pp. 1134–46.

Wutz, A., Rasmussen, T. P., and Jaenisch, R. (2002). "Chromosomal silencing and localization are mediated by different domains of Xist RNA". *Nature Genetics* 30: pp. 167–174.

Xi, Y., Yao, J., Chen, R., Li, W., and He, X. (2011). "Nucleosome fragility reveals novel functional states of chromatin and poises genes for activation". *Genome Research* 21: pp. 718–724.

Xu, C., Bian, C., Yang, W., Galka, M., Ouyang, H., Chen, C., Qiu, W., Liu, H., Jones, A. E., MacKenzie, F., Pan, P., Li, S. S.-C., Wang, H., and Min, J. (2010). "Binding of different histone marks differentially regulates the activity and specificity of polycomb repressive complex 2 (PRC2)". *Proceedings of the National Academy of Sciences* 107: pp. 19266–19271.

Xue, Y., Wong, J., Moreno, G., Young, M. K., Côté, J., and Wang, W. (1998). "NURD, a Novel Complex with Both ATP-Dependent Chromatin-Remodeling and Histone Deacetylase Activities". *Molecular Cell* 2: pp. 851–861.

Yan, Y., Zhao, W., Huang, Y., Tong, H., Xia, Y., Jiang, Q., and Qin, J. (2017). "Loss of Polycomb Group Protein Pcgf1 Severely Compromises Proper Differentiation of Embryonic Stem Cells". *Scientific Reports* 7.

Yang, C.-S., Chang, K.-Y., Dang, J., and Rana, T. M. (2016). "Polycomb Group Protein Pcgf6 Acts as a Master Regulator to Maintain Embryonic Stem Cell Identity". *Scientific Reports* 6.

Yang, W., Lee, Y.-H., Jones, A. E., Woolnough, J. L., Zhou, D., Dai, Q., Wu, Q., Giles, K. E., Townes, T. M., and Wang, H. (2014). "The histone H2A deubiquitinase Usp16 regulates embryonic stem cell gene expression and lineage commitment". *Nature Communications* 5.

Yao, M., Zhou, X., Zhou, J., Gong, S., Hu, G., Li, J., Huang, K., Lai, P., Shi, G., Hutchins, A. P., Sun, H., Wang, H., and Yao, H. (2018). "PCGF5 is required for neural differentiation of embryonic stem cells". *Nat Commun* 9: p. 1463.

Yap, K. L., Li, S., Muñoz-Cabello, A. M., Raguz, S., Zeng, L., Mujtaba, S., Gil, J., Walsh, M. J., and Zhou, M.-M. (2010). "Molecular Interplay of the Noncoding RNA ANRIL and Methylated Histone H3 Lysine 27 by Polycomb CBX7 in Transcriptional Silencing of INK4a". *Molecular Cell* 38: pp. 662–674.

Yoo, J.-Y., Choi, K.-C., Kang, H., Kim, Y. J., Lee, J., Jun, W. J., Kim, M.-J., Lee, Y.-H., Lee, O.-H., and Yoon, H.-G. (2010). "Histone deacetylase 3 is selectively involved in L3MBTL2-mediated transcriptional repression". *FEBS Letters* 584: pp. 2225–2230.

Youmans, D. T., Schmidt, J. C., and Cech, T. R. (2018). "Live-cell imaging reveals the dynamics of PRC2 and recruitment to chromatin by SUZ12-associated subunits". *Genes & Development* 32: pp. 794–805.

Yu, H., Mashtalir, N., Daou, S., Hammond-Martel, I., Ross, J., Sui, G., Hart, G. W., Rauscher, F. J., Drobetsky, E., Milot, E., Shi, Y., and Affar, E. B. (2010). "The Ubiquitin Carboxyl Hydrolase BAP1 Forms a Ternary Complex with YY1 and HCF-1 and Is a Critical Regulator of Gene Expression". *Molecular and Cellular Biology* 30: pp. 5071–5085.

Yu, J.-R., Lee, C.-H., Oksuz, O., Stafford, J. M., and Reinberg, D. (2019). "PRC2 is high maintenance". *Genes & Development*.

Yu, M., Mazor, T., Huang, H., Huang, H.-T., Kathrein, K. L., Woo, A. J., Chouinard, C. R., Labadorf, A., Akie, T. E., Moran, T. B., Xie, H., Zacharek, S., Taniuchi, I., Roeder, R. G., Kim, C. F., Zon, L. I., Fraenkel, E., and Cantor, A. B. (2012). "Direct Recruitment of Polycomb Repressive Complex 1 to Chromatin by Core Binding Transcription Factors". *Molecular Cell* 45: pp. 330–343.

Yuan, G., Ma, B., Yuan, W., Zhang, Z., Chen, P., Ding, X., Feng, L., Shen, X., Chen, S., Li, G., and Zhu, B. (2013). "Histone H2A Ubiquitination Inhibits the Enzymatic Activity of H3 Lysine 36 Methyltransferases". *Journal of Biological Chemistry* 288: pp. 30832–30842.

Yuan, G.-C., Liu, Y.-J., Dion, M. F., Slack, M. D., Wu, L. F., Altschuler, S. J., and Rando, O. J. (2005). "Genome-Scale Identification of Nucleosome Positions in S. cerevisiae". *Science* 309: p. 626.

Yuan, W., Wu, T., Fu, H., Dai, C., Wu, H., Liu, N., Li, X., Xu, M., Zhang, Z., Niu, T., Han, Z., Chai, J., Zhou, X. J., Gao, S., and Zhu, B. (2012). "Dense Chromatin Activates Polycomb Repressive Complex 2 to Regulate H3 Lysine 27 Methylation". *Science* 337: pp. 971–975.

Yuan, W., Xu, M., Huang, C., Liu, N., Chen, S., and Zhu, B. (2011). "H3K36 Methylation Antagonizes PRC2-mediated H3K27 Methylation". *Journal of Biological Chemistry* 286: pp. 7983–7989.

Zabidi, M. A., Arnold, C. D., Schernhuber, K., Pagani, M., Rath, M., Frank, O., and Stark, A. (2014). "Enhancer–core-promoter specificity separates developmental and housekeeping gene regulation". *Nature* 518: pp. 556–559.

Zarrizi, R., Menard, J. A., Belting, M., and Massoumi, R. (2014). "Deubiquitination of $\gamma$-Tubulin by BAP1 Prevents Chromosome Instability in Breast Cancer Cells". *Cancer Research* 74: pp. 6499–6508.

Zdzieblo, D., Li, X., Lin, Q., Zenke, M., Illich, D. J., Becker, M., and Muller, A. M. (2014). "Pcgf6, a polycomb group protein, regulates mesodermal lineage differentiation in murine ESCs and functions in iPS reprogramming". *Stem Cells* 32: pp. 3112–25.

Zeitlinger, J., Stark, A., Kellis, M., Hong, J.-W., Nechaev, S., Adelman, K., Levine, M., and Young, R. A. (2007). "RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo". *Nature Genetics* 39: pp. 1512–1516.

Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nusbaum, C., Myers, R. M., Brown, M., Li, W., and Liu, X. S. (2008). "Model-based analysis of ChIP-Seq (MACS)". *Genome Biol* 9: R137.

Zhang, Y., Ng, H.-H., Erdjument-Bromage, H., Tempst, P., Bird, A., and Reinberg, D. (1999). "Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation". *Genes & Development* 13: pp. 1924–1935.

Zhang, Y. and Reinberg, D. (2001). "Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails". *Genes & Development* 15: pp. 2343–2360.

Zhang, Z., Jones, A. E., Wu, W., et al. (2017). "Role of remodeling and spacing factor 1 in histone H2A ubiquitination-mediated gene silencing". *Proc Natl Acad Sci U S A* 114: E7949–E7958.

Zhang, Z., Jones, A., Sun, C.-W., Li, C., Chang, C.-W., Joo, H.-Y., Dai, Q., Mysliwiec, M. R., Wu, L.-C., Guo, Y., Yang, W., Liu, K., Pawlik, K. M., Erdjument-Bromage, H., Tempst, P., Lee, Y., Min, J., Townes, T. M., and Wang, H. (2011). "PRC2 Complexes with JARID2, MTF2, and esPRC2p48 in ES Cells to Modulate ES Cell Pluripotency and Somatic Cell Reprograming". *STEM CELLS* 29: pp. 229–240.

Zhao, J., Sun, B. K., Erwin, J. A., Song, J.-J., and Lee, J. T. (2008a). "Polycomb Proteins Targeted by a Short Repeat RNA to the Mouse X Chromosome". *Science* 322: pp. 750–756.

Zhao, W., Huang, Y., Zhang, J., Liu, M., Ji, H., Wang, C., Cao, N., Li, C., Xia, Y., Jiang, Q., and Qin, J. (2017a). "Polycomb group RING finger proteins 3/5 activate transcription via an interaction with the pluripotency factor Tex10 in embryonic stem cells". *J Biol Chem* 292: pp. 21527–21537.

Zhao, W., Tong, H., Huang, Y., Yan, Y., Teng, H., Xia, Y., Jiang, Q., and Qin, J. (2017b). "Essential Role for Polycomb Group Protein Pcgf6 in Embryonic Stem Cell Maintenance and a Noncanonical Polycomb Repressive Complex 1 (PRC1) Integrity". *Journal of Biological Chemistry* 292: pp. 2773–2784.

Zhao, Y. and Garcia, B. A. (2015). "Comprehensive Catalog of Currently Documented Histone Modifications". *Cold Spring Harbor Perspectives in Biology* 7: a025064.

Zhao, Y., Lang, G., Ito, S., Bonnet, J., Metzger, E., Sawatsubashi, S., Suzuki, E., Guezennec, X. L., Stunnenberg, H. G., Krasnov, A., Georgieva, S. G., Schüle, R., Takeyama, K.-I., Kato, S., Tora, L., and Devys, D. (2008b). "A TFTC/STAGA Module Mediates Histone H2A and H2B Deubiquitination, Coactivates Nuclear Receptors, and Counteracts Heterochromatin Silencing". *Molecular Cell* 29: pp. 92–101.

Zhen, C. Y., Tatavosian, R., Huynh, T. N., Duc, H. N., Das, R., Kokotovic, M., Grimm, J. B., Lavis, L. D., Lee, J., Mejia, F. J., Li, Y., Yao, T., and Ren, X. (2016). "Live-cell single-molecule tracking reveals co-recognition of H3K27me3 and DNA targets polycomb Cbx7-PRC1 to chromatin". *eLife* 5.

Zhou, W., Zhu, P., Wang, J., Pascual, G., Ohgi, K. A., Lozach, J., Glass, C. K., and Rosenfeld, M. G. (2008). "Histone H2A Monoubiquitination Represses Transcription by Inhibiting RNA Polymerase II Transcriptional Elongation". *Molecular Cell* 29: pp. 69–80.

Zhou, Y.-B., Gerchman, S. E., Ramakrishnan, V., Travers, A., and Muyldermans, S. (1998). "Position and orientation of the globular domain of linker histone H5 on the nucleosome". *Nature* 395: pp. 402–405.

Zhu, P., Zhou, W., Wang, J., Puc, J., Ohgi, K. A., Erdjument-Bromage, H., Tempst, P., Glass, C. K., and Rosenfeld, M. G. (2007). "A Histone H2A Deubiquitinase Complex Coordinating Histone Acetylation and H1 Dissociation in Transcriptional Regulation". *Molecular Cell* 27: pp. 609–621.

Zink, B. and Paro, R. (1989). "In vivo binding pattern of a trans-regulator of homoeotic genes in Drosophila melanogaster". *Nature* 337: pp. 468–471.