Mechanisms of Xist-dependent Polycomb recruitment

Mafalda Sofia Fernandes Almeida

Merton College

University of Oxford



A thesis submitted for the degree of

Doctor of Philosophy in Biochemistry

ABSTRACT

The Polycomb repressive complexes PRC1 and PRC2 play a central role in developmental gene regulation, functioning, at least in part, by catalyzing specific histone modifications, H2AK119u1 and H3K27me3, respectively. The hierarchical model of Polycomb recruitment proposes that PRC2 is recruited to target loci by sequence specific factors and/or non-coding RNAs, and that PRC1 recruitment occurs subsequently as a consequence of recognition of H3K27me3 by the PRC1 subunit CBX. However, non-canonical PRC1 complexes, defined by the presence of different PCGF subunits, lack the CBX subunit and are therefore recruited independently of PRC2 activity. Moreover, recent evidence has shown that H2AK119u1 catalyzed by non-canonical PRC1 can directly recruit PRC2, establishing a reverse model of Polycomb recruitment. I analyzed Polycomb recruitment using a classical model system, X chromosome inactivation mediated by Xist RNA. Previous reports suggested that Polycomb recruitment to the inactive X chromosome is initiated by a direct interaction of PRC2 subunits with Xist RNA, and that PRC1 is then recruited via interaction of the CBX subunit with PRC2 mediated H3K27me3. This model has been challenged, since it does not account for all experimental observations. In this thesis, I provide evidence that Polycomb recruitment by Xist RNA is initiated by deposition of H2AK119u1 by non-canonical PRC1 complexes. PRC2 is therefore recruited indirectly, presumably by binding to H2AK119u1. I also show that all non-canonical PRC1 complexes are enriched on the inactive chromosome, albeit with different dynamic behaviors. PCGF3- and PCGF5-PRC1 complexes, which

have a slower dynamics on the inactive chromosome as determined by FRAP, are necessary for the initiation of Polycomb domain formation on the inactive chromosome, presumably through direct recruitment by Xist RNA or through an adaptor protein that mediates this interaction. Subsequent enrichment of other non-canonical PRC1 complexes relies predominantly on H2AK119u1 recognition by RYBP/YAF2, which also contributes to the enrichment of the more dynamic fraction of PCGF3- and PCGF5-PRC1 complexes on the inactive chromosome. Altogether, these observations overturn existing models for Polycomb recruitment by Xist RNA, and thereby, provide new insights into the dynamics and mechanisms for Polycomb recruitment by non-coding RNAs.

ACKNOWLEDGMENTS

I would like to thank Neil for the opportunity to be part of his lab and for giving me such an interesting and challenging project. Thank you for all the supervision and support, scientific discussions and for giving me the freedom to pursue my own ideas. Thank you for the suggestions and comments on my thesis.

A special thank you goes to Tatyana, for all her scientific support, for sharing all her knowledge and expertise with me and for creating such a good lab environment to work in, but most importantly for being there for me since the very first moment when I decided to move to Oxford. Thank you for your friendship, for your patience, for your optimism and for caring so much about my wellbeing.

Thank you to everyone on the third floor of Biochemistry for being so nice and creating such a great atmosphere not only to work in, but also to make friends. A special thanks to past and present members of the Brockdorff lab for all the support throughout the years, especially to those that somehow contributed to this work, particularly Greta, Andrea and Tatyana. Thank you to Dave, for teaching me so many things in the lab, as well as for the fun moments in tissue culture and in the lab. Thanks for your optimism and for encouraging me to do better! Thank you to my office and bench neighbors, Tianyi, Benoit and Vincenzo, for being such great fun to work with and for your friendship. An extra thanks goes to Benoit and Dave for comments on this thesis.

A big thank you to my friends and housemates, Al and Sabrina, for distracting me from science when I needed it and for being so understanding about my crazy scientist schedule. Thanks for taking care of me during these last months and for being such great fun to live with.

This thesis would not be possible without my friends' support and understanding. A big thank you to my friends from Viseu for always being there for me: Rita, Nuno, Rui, Pedro, Sofia, Tiago, Maria, Cappelle, Joana and Filipa. Thank you for still making me feel home there! To the friends I made in Porto for going the extra mile to keep in touch despite the distance: Joel, Joana, Zé, João, Ana, Vera e André. A special thanks to Joana S. and Maria for the long conversations and surprise visits to the UK. Thank you to all the GABBA 15th family, especially my "Cucurus" Diana, Gonçalo, Joana and Arezes. Thank you Arezes for being my support in Oxford this past year, and for making it a lot more fun. I would like to thank all my Portuguese friends in Oxford for being such a fun crew to hang out with. A special thanks to the "Embaixada": Marcelo, Miguel and Inês. You all have a special place in my heart!

Thank you Nick for always being there for me at all levels throughout this journey. Thank you for all comments on my thesis and for your patience to always listen to my ideas and concerns. Thanks for making my life happier! Your unconditional support is invaluable and I can't wait for our future adventures together.

Agradeço também aos meus pais, Américo e Cristina, por acreditarem sempre em mim e nas minhas capacidades, e me fazerem querer ser melhor a cada dia, enquanto pessoa e enquanto cientista. Obrigada por me apoiarem sempre nas minhas decisões, mesmo quando isso implica estar a muitos quilómetros de distância. Muito obrigada por tudo! Nada disto seria possível sem vocês e por isso esta tese é também vossa.

TABLE OF CONTENTS

Abstract ..... **I**

Acknowledgments ..... **III**

Table of Contents ..... **V**

List of Figures ..... **X**

List of Tables ..... **XIV**

List of Abbreviations ..... **XV**

Chapter 1: Introduction **1**

1.1. Epigenetics and chromatin 1

1.1.1. Chromatin structure 2

1.1.1.2. Euchromatin and its epigenetic marks 3

1.1.1.3. Heterochromatin and its epigenetic marks 4

1.1.1.3.1. Constitutive Heterochromatin 5

1.1.1.3.2. Facultative Heterochromatin 6

1.2. Polycomb repressive complexes 8

1.2.1. Discovery of Polycomb Group genes in *Drosophila* 8

1.2.2. Composition and architecture of PcG complexes in *Drosophila* and mammals 10

1.2.2.1. Polycomb Repressive Complex 1: canonical and non-canonical complexes 11

1.2.2.2. Polycomb Repressive Complex 2: core and auxiliary subunits 16

1.2.3. Function of Polycomb complexes 20

1.2.3.1. Gene silencing by PRC1 and PRC2 mediated histone modifications 21
1.2.3.2. A role for Polycomb complexes in chromatin compaction and nuclear organization 23
1.2.3.3. Polycomb function in gene activation 28
1.2.4. Mechanisms of *de novo* recruitment and propagation of Polycomb complexes 29
1.2.4.1. Lessons from *Drosophila* 30
1.2.4.2. The quest for a mammalian PRE 32
1.2.4.3. Recruitment of Polycomb by interaction with RNA 36
1.2.4.4. Reverse model of Polycomb recruitment and propagation of the marks 37
1.3. X chromosome inactivation 42
1.3.1. The process of X chromosome inactivation 42
1.3.2. Xist-mediated silencing and heterochromatinization of the Xi 45
1.3.3. Polycomb contribution to Xi silencing 47
1.3.4. Polycomb recruitment to the Xi 50
1.4. Aims of this study 55
Chapter 2: Materials and Methods 56
2.1. Plasmids and cloning 56
2.2. Mouse embryonic stem cell lines 58
2.2.1. Conditional mESC lines production and derivation 58
2.2.2. Mouse ESC lines and growing conditions 59
2.2.3. Doxycyline and tamoxifen treatments 59
2.2.4. Generation of stable cell lines expressing transgenes 60

2.2.5. Targeting of rtTA-H cassette into Rosa26 locus 61
2.2.6. Knockout of *Rybp*, *Yaf2*, *Pcgf5* and *Pcgf6* using CRISPR/Cas9-mediated mutagenesis 62
2.2.7. Generation of Xist inducible clones 63
2.3. Colony screening by Southern blot 64
2.3.1. Preparation of DNA from 96 well plate for Southern blot 64
2.3.2. Southern blot 65
2.4. Imaging techniques 66
2.4.1. Live cell imaging 66
2.4.2. FRAP and FRAP analysis 66
2.4.3. Immunofluorescence 67
2.4.4. RNA-FISH 69
2.5. General Biochemistry 70
2.5.1. Preparation of whole cell lysates and nuclear extracts 70
2.5.2. Immunoprecipitation 71
2.5.3. Preparation of histone extracts 71
2.5.4. Western blot 72
Chapter 3: Role of PRC1 and H2AK119u1 in PRC2 recruitment by Xist RNA 74
3.1. Introduction 74
3.2. Results 75
3.2.1. Generation of an inducible Xist system in *Ring1A*$^{-/-}$ *Ring1B*$^{fl/fl}$ to assess *de novo* Polycomb recruitment 75
3.2.2. PRC1 activity is required for Xist-mediated recruitment of PRC2... 78

3.3. Discussion 82

Chapter 4: Localization and dynamics of PRC1 complexes on the inactive chromosome domain 86

4.1. Introduction 86

4.2. Results 88

4.2.1. Generation and characterization of stable mESC lines expressing eGFP fused to PRC1 subunits 88

4.2.2. Recruitment of non-canonical PRC1 complexes to the inactive chromosome domain 93

4.2.3. Dynamics of different PRC1 complexes in the nucleus and the inactive chromosome domain 96

4.3. Discussion 104

Chapter 5: The role of RYBP/YAF2 in PRC1 recruitment to the inactive chromosome 109

5.1. Introduction 109

5.2. Results 111

5.2.1. Generation and characterization of *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ mESC lines with an inducible Xist transgene 111

5.2.2. Recruitment of non-canonical PRC1 complexes to the inactive chromosome domain in the absence of RYBP/YAF2 115

5.2.3. Dynamics of RYBP in the nucleus and the inactive chromosome domain in wild-type cells 119

5.2.4. Visualization of non-canonical PRC1 complexes present on the inactive chromosome domain in the absence of RYBP/YAF2 123

5.2.5. Mechanism of RYBP-dependent recruitment of non-canonical PRC1 complexes 124

5.3. Discussion 130

Chapter 6: Dissecting the role of PCGF proteins in the initiation of PRC1 recruitment to the inactive chromosome 135

6.1. Introduction 135

6.2. Results 138

6.2.1. Generation of an inducible Xist system in *Pcgf* conditional knockout mESC to assess *de novo* Polycomb recruitment 138

6.2.2. Generation of *Pcgf5* and *Pcgf6* constitutive knockout mESC to assess *de novo* Xist-mediated PRC1 recruitment 141

6.2.3. Analysis of single *Pcgf* knockout on Xist-mediated Polycomb recruitment 145

6.2.4. Analysis of double Pcgf3/Pcgf5 knockouts effect in de novo Xist-mediated Polycomb recruitment 154

6.2.5. Analysis of other double knockouts in de novo Xist-mediated Polycomb recruitment 159

6.3. Discussion 162

Chapter 7: General Discussion 168

7.1. PCGF3/PCGF5-PRC1 variants and Xist-mediated Polycomb recruitment 169

7.2. Relevance of H2AK119u1 in Polycomb domain formation 174

Chapter 8: Bibliography 177

Chapter 9: Appendix 222

LIST OF FIGURES

Figure 1 – Extra sex combs phenotype in a heterozygous *Pc* mutant in *Drosophila* 10

Figure 2 – Canonical PRC1 composition in mouse and in *Drosophila* 13

Figure 3 – Diversity of non-canonical PRC1 complexes in mouse and in *Drosophila* 16

Figure 4 – Diversity of PRC2 complexes in mouse and in *Drosophila* 20

Figure 5 – Hierarchical model of Polycomb recruitment 32

Figure 6 – Reverse model of Polycomb recruitment 40

Figure 7 – Xist RNA visualization and *Xist* gene locus 44

Figure 8 – Establishment of an inducible Xist system in *Ring1A*$^{-/-}$ *Ring1B*$^{fl/fl}$ mESCs 77

Figure 9 – Global effects of deletion of *Ring1B* in *Ring1A*$^{-/-}$ *Ring1B*$^{fl/fl}$ mESCs… 79

Figure 10 – Immunofluorescence for H3K27me3, EZH2 and JARID2 in *Ring1A*$^{-/-}$ *Ring1B*$^{fl/fl}$ cells in the presence or absence of RING1B following Xist induction for 24 h 80

Figure 11 – RNA-FISH for Xist RNA in *Ring1A*$^{-/-}$ *Ring1B*$^{fl/fl}$ cells and illustrating that deletion of genes encoding RING1A/B has no effect on Xist domain formation 81

Figure 12 – Histogram summarizing the percentage of cells with domains for Xist, H3K27me3, EZH2 and JARID2 in tamoxifen (+OHT) treated and untreated cells after Xist induction for 24 hours (+dox) 82

Figure 13 – Establishment of stable mESC lines expressing eGFP-PCGF proteins 91
Figure 14 – Establishment of stable mESC lines expressing eGFP-RYBP and eGFP-RING1B proteins 92
Figure 15 – Xist dependent localisation of non-canonical PRC1 complexes by live cell imaging..... 94
Figure 16 – Xist dependent localisation of non-canonical PRC1 complexes assayed by immunofluorescence..... 95
Figure 17 – eGFP-PCGF proteins show different dynamics in the nucleoplasm and the inactive chromosome domain 99
Figure 18 – Fluorescence recovery after photobleaching curves for eGFP-PCGF proteins in the nucleoplasm and the inactive chromosome domain .. 100
Figure 19 – Quantitative analysis of FRAP data using a bi-exponential function 103
Figure 20 – Establishment and characterization of *Rybp*$^{-/-}$*Yaf2*$^{-/-}$ mESC line 113
Figure 21 – Immunofluorescence analysis of Polycomb marks on the inactive chromosome domain over a Xist induction timecourse in *Rybp*$^{-/-}$*Yaf2*$^{-/-}$ compared to wild-type mESCs..... 114
Figure 22 – Interaction of eGFP-PCGF proteins with endogenous RING1B in *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ mESCs 115
Figure 23 – Xist dependent localisation of non-canonical PRC1 complexes in *Rybp*$^{-/-}$*Yaf2*$^{-/-}$ mESCs by live cell imaging 117
Figure 24 – Immunofluorescence analysis of *Rybp*$^{-/-}$*Yaf2*$^{-/-}$ mESCs expressing eGFP-PCGF fusion proteins 118

Figure 25 – Xist dependent localisation of eGFP-RING1B in *Rybp*$^{-/-}$*Yaf2*$^{-/-}$ mESCs by live cell imaging 119
Figure 26 – eGFP-RYBP shows very fast turnover in the nucleoplasm and the inactive chromosome domain 120
Figure 27 – Quantitative analysis of FRAP data for eGFP-RYBP using a bi-exponential function 122
Figure 28 – Immunofluorescence analysis of *Rybp*$^{-/-}$*Yaf2*$^{-/-}$ mESCs expressing eGFP-PCGF fusion proteins using perm/fix protocol 124
Figure 29 – Expression of mCherry-RYBP in *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ mESCs rescues enrichment of eGFP-PCGF proteins on the inactive chromosome domain .. 126
Figure 30 – Expression of mCherry-RYBP TF-AA mutant in *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ mESCs does not rescue enrichment of eGFP-PCGF proteins on the inactive chromosome domain 129
Figure 31 – Neighbor Joining phylogenetic tree of mouse PCGF proteins and their *Drosophila* homologs amino acid sequences 136
Figure 32 – Characterization of *Pcgf1*$^{fl/fl}$ and *Pcgf3*$^{fl/fl}$ mESC lines 139
Figure 33 – Engineering of an inducible Xist system encompassing BglSL (Bgl stem loops) for recognition by BglG-mCherry in conditional *Pcgf* knockout mESCs 140
Figure 34 – Schematic of mESC lines generated using CRISPR/Cas9-mediated gene editing 142
Figure 35 – Strategy for CRISPR/Cas9 nickase mediated knockout of *Pcgf5* 143
Figure 36 – Strategy for CRISPR/Cas9 nuclease mediated knockout of *Pcgf6* 145

Figure 37 – Schematic representation of the experiment indicating timepoints for deletion of genes encoding PCGF proteins (+OHT) in conditional knockout mESCs and Xist transgene induction (+dox) 146
Figure 38 – Analysis of *Pcgf1*$^{fl/fl}$ mESCs 148
Figure 39 – Analysis of *Pcgf3*$^{fl/fl}$ mESCs 150
Figure 40 – Analysis of *Pcgf5* single knockout mESCs. 152
Figure 41 – Analysis of *Pcgf6* single knockout mESCs 154
Figure 42 – Analysis of *Pcgf3*/*Pcgf5* double knockout mESCs 155
Figure 43 – Analysis of *Pcgf3*/*Pcgf5* double knockout mESCs with different random Xist integrations 158
Figure 44 – Analysis of other double knockouts generated in *Pcgf1*$^{fl/fl}$ Xist E5 cell line 160
Figure 45 – Analysis of *Pcgf3*/*Pcgf6* double knockout mESCs 162
Figure 46 – Model summarising key steps in Polycomb recruitment by Xist RNA 169
Figure 47 – Fluorescence recovery after photobleaching curves for eGFP-RING1B in the nucleoplasm and the inactive chromosome domain 222
Figure 48 – Effect of *Pcgf3/5* knockout in Polycomb recruitment by Xist RNA *in vivo* 223

LIST OF TABLES

Table 1 – Primers used to amplify inserts for LIC cloning and for mutagenesis. 57

Table 2 – Primers used for genotype of conditional knockout mESCs after tamoxifen treatment 60

Table 3 – CRISPR/Cas9 sgRNAs target sequences 63

Table 4 – Primers used for genotype and sequencing of CRISPR/Cas9 generated mutant mESCs..... 63

Table 5 – Antibodies used in this thesis..... 73

Table 6 – *p*-values resulting from Student t-test for pairwise comparisons of scoring data for H2AK119u1 and H3K27me3 domains overlapping with Xist in all cell lines generated in *Pcgf1*$^{fl/fl}$ background 224

Table 7 – *p*-values resulting from Student t-test for pairwise comparisons of scoring data for H2AK119u1 and H3K27me3 domains overlapping with Xist in all cell lines generated in *Pcgf3*$^{fl/fl}$ background..... 224

LIST OF ABBREVIATIONS

AEBP2	AE Binding Protein 2
ANRIL	Antisense Noncoding RNA in the INK4 locus
ANTC	Antennapedia Complex
ATRX	Alpha Thalassemia/mental Retardation X-linked
AUTS2	Autism Susceptibility Sandidate 2
BCOR	BCL-6 co-repressor
BCORL1	BCL-6 co-repressor-like protein 1
BglSL	Bgl Stem Loop
BMI-1	B Lymphoma Mo-MLV Insertion Region 1 Homolog
bp	base pair
BSA	Bovine Serum Albumin
BXC	Bithorax Complex
C10ORF12	Chromosome 10 Open Reading Frame 12
C17ORF96	Chromosome 17 Open Reading Frame 96
CAF1	Chromatin Assembly Factor 1
CBP	CREB Binding Protein
CBX	Chromobox Protein
cDNA	Complementary DNA
CDYL	Chromodomain Y-like Protein
ChIP	Chromatin Immunoprecipitation
ChIP-seq	Chromatin Immunoprecipitation Sequencing
CK2	Casein Kinase 2
CpG	Cytosine-phosphate-Guanine

CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DAPI	4′,6-diamidino-2-phenylindole
dKDM2B	*Drosophila* Lysine (K)-specific demethylase 2B
dKO	Double knockout
DNA	Deoxyribonucleic acid
DNMT3A/B	DNA Methyltransferase 3A/B
dox	Doxycycline
dRAF	dRing Associated Factors
dRING	*Drosophila* Really Interesting New Gene protein as known as Sce for Sex combs extra
E(Z)	Enhancer of Zeste
ECL	Enhanced chemiluminescence
EED	Embryonic Ectoderm Development
eGFP	Enhanced Green Fluorescent Protein
EPOP	Elongin BC and polycomb repressive complex 2 associated Protein
ESC	Extra Sex Combs
ESCL	Extra Sex Combs Like
EZH1	Enhancer Of Zeste 1
EZH2	Enhancer Of Zeste 2
FRAP	Fluorescence Recovery After Photobleaching
GFP	Green Fluorescent Protein
H2AK118u1	Monoubiquitylation of Lysine 118 of Histone H2A
H2AK119u1	Monoubiquitylation of Lysine 119 of Histone H2A
H3K27	Lysine 27 of Histone H3

H3K27ac	Acetylation of Lysine 27 of Histone H3
H3K27me1	Monomethylation of Lysine 27 of Histone H3
H3K27me2	Dimethylation of Lysine 27 of Histone H3
H3K27me3	Trimethylation of Lysine 27 of Histone H3
H3K36me3	Trimethylation of Lysine 36 of Histone H3
H3K4me3	Trimethylation of Lysine 4 of Histone H3
H3K9	Lysine 9 of Histone H3
H3K9me2	Dimethylation of Lysine 9 of Histone H3
H3K9me3	Trimethylation of Lysine 9 of Histone H3
H4K20me1	Monomethylation of Lysine 20 of Histone H4
HDAC3	Histone Deacetylase 3
HnRNPK	Heterogeneous Nuclear Ribonucleoprotein K
HOTAIR	HOX transcript antisense RNA
Hox	Homeotic
HP1	Heterochromatin Protein 1
IP	Immunoprecipitation
JARID1A	Jumonji/ARID Domain-Containing Protein 1A
JARID2	Jumonji/ARID Domain-Containing Protein 2
KDM2B	Lysine (K)-specific demethylase 2B
KDM6A	Lysine (K)-specific demethylase 6A
KDM6B	Lysine (K)-specific demethylase 6B
LBR	Lamin B Receptor
MEFs	Mouse embryonic fibroblasts
MEL18	Melanoma nuclear protein 18
mESCs	Mouse embryonic stem cells

MLL3	Mixed-lineage Leukemia Protein 3
MPH	Mouse Polyhomeotic
MS	Mass spectrometry
MTF2	Metal response element binding transcription factor 2
ncRNA	Non-coding RNA
NZF-UBD	Zinc-Finger Ubiquitin Binding Domain
OHT	Tamoxifen
PAM	Protospacer adjacent motif
Pc	Polycomb
PcG	Polycomb Group
PCGF	Polycomb group RING finger Protein
PCL	Polycomb Like Protein
PCR	Polymerase Chain Reaction
Ph	Polyhomeotic
PHC	Polyhomeotic-like Protein
PHF1	PHD Finger Protein 1
PHF19	PHD Finger Protein 19
PHO	Pleiohomeotic
PhoRC	Pho-repressive complex
PRC1	Polycomb Repressive Complex 1
PRC2	Polycomb Repressive Complex 2
PRE	Polycomb Response Element
PSC	Posterior Sex Combs
PTM	Post-translational modification
RAWUL	RING finger- and WD40-associated ubiquitin-like

RbAp46/48	Retinoblastoma associated protein 46/48
RBM15	RNA Binding Motif Protein 15
REST	RE1-silencing transcription factor
RING1A	RING finger protein 1
RING1B	RING finger protein 1B
RNA	Ribonucleic acid
RNA-FISH	Fluorescent in situ hybridization targeting ribonucleic acid
rtTA	Reverse Tetracycline-controlled Transactivator
RUNX1	Runt-related transcription factor 1
RYBP	RING1 and YY1-binding protein
SAF-A	Scaffold Attachment Factor A
SAM	Sterile Alpha Motif
SCE	Sex Combs Extra
Scr	Sex combs reduced locus
SFMBT	Scm-Like With Four Mbt Domains
sgRNA	Short Guide RNA
SKP1	S-Phase Kinase-Associated Protein 1
SMCHD1	Structural Maintenance Of Chromosomes Hinge Domain Containing Protein 1
SNAIL1	Snail Family Transcriptional Repressor 1
SPEN	Split ends
SU(Z)2	Suppressor of Zeste 2
SUZ12	Suppressor Of Zeste 12
SWI/SNF	SWItch/Sucrose Non-Fermentable
TAD	Topologically Associating Domain

TBP	TATA-Binding Protein
TF	Transcription Factor
TS	Trophoblast Stem
UIM	Ubiquitin-Interacting Motif
WT	Wild-type
WTAP	Wilms Tumor 1 Associated Protein
XCI	X chromosome inactivation
Xi	Inactive X chromosome
Xist	X-inactive specific transcript
YAF2	YY1 Associated Factor 2
YY1	Yin Yang 1 Transcriptional Repressor
ZF-CXXC	Zinc-finger CXXC Domain

CHAPTER 1: General Introduction

1.1. Epigenetics and chromatin

In multicellular organisms, regulation of transcription programs is crucial to generate a diversity of cell types (200 to 300 in mammals) during development (Grossniklaus and Paro 2014). Transcriptional regulation is achieved through both genetic and epigenetic mechanisms. Epigenetic mechanisms do not rely on changes on the DNA sequence itself, and once triggered by an initial stimulus, can be maintained across subsequent cell divisions to preserve cell type specific transcription patterns. These epigenetic processes, while responsible for cellular memory, are dynamic and reversible upon exposure to new developmental or environmental cues. This is particularly important during development, contributing to the plasticity observed during formation of different tissues (Jaenisch and Bird 2003; Bernstein et al. 2007).

In eukaryotes the DNA molecule is packaged in the cell nucleus in an organized macromolecular structure called chromatin. This nucleocomplex was first identified by Walther Flemming as the stainable material in the nucleus visible under the microscope (reviewed in Trojer and Reinberg 2007). Later, X-ray crystallography studies characterized its basic unit, the nucleosome, which consists of 147bp of DNA wrapped around an octamer of histones H2A, H2B, H3 and H4 (Luger et al. 1997). Arrays of nucleosomes interspersed with linker DNA and histone H1 constitute chromatin, a functional

entity whose dynamic organization reflects the incoming developmental or environmental cue to regulate gene expression without changes in the DNA sequence. The epigenetic mechanisms responsible for shaping chromatin include chemical modifications of DNA, such as cytosine methylation and hydroxymethylation, histone post-translational modifications, expression of non-coding RNAs, replacement of canonical histones by histone variants, and RNA chemical modifications. Epigenetic factors such as readers, writers and erasers of these epigenetic marks also contribute to maintain transcriptional programs in specific cell types (reviewed in Zhang et al. 2015).

1.1.1. Chromatin structure

Chromatin can be broadly subdivided into two main types: active euchromatin and silent heterochromatin (reviewed in Kouzarides 2007). Emil Heitz first coined these terms in 1928, after observing that specific chromosome regions were differentially stained in the nucleus of a cell depending on their condensation state, but that this staining showed a conserved pattern across individual cells of the same type (reviewed in Trojer and Reinberg 2007). The regions that retain high condensation in interphase nuclei and therefore showed strong staining were named heterochromatin, whereas chromosome segments that due to their highly decondensed state did not stain well, were called euchromatin. Heitz also suggested that these different chromosomal regions would be related to gene repression and activity, respectively (reviewed in Passarge 1979). Further genome-wide studies characterizing the distribution of epigenetic marks such as histone post-translational

modifications and DNA methylation helped to define euchromatin and heterochromatin as functional domains related to differential gene transcription.

1.1.1.2. Euchromatin and its epigenetic marks

Active euchromatin is characterized by open, highly decondensed chromatin in interphase cells. Even before the advent of high-throughput sequencing techniques, euchromatin regions were shown to represent only 20% of the genome in interphase calf lymphocytes, even though they accounted for the bulk of messenger RNA synthesized in the cell nucleus (Frenster et al. 1963). The open chromatin conformation characteristic of euchromatin contributes to its transcription since it is highly accessible to transcription factor (TF) binding, RNA polymerase and other transcription machinery.

Chromatin immunoprecipitation (ChIP) techniques coupled with sequencing have been crucial in the characterization of histone post-translational modifications (PTMs) that compose the chromatin signature of euchromatic domains. Actively transcribed genes are generally marked by H3K4me3 and histone H3 and H4 lysine acetylation in the promoter region, which moreover has reduced nucleosome occupancy (Mikkelsen et al. 2007; Lee et al. 2004). DNA methylation is also differentially associated with different chromatin environments. DNA methylation levels at promoters inversely correlate with enrichment of H3K4me3 in mammalian cells, therefore active gene promoters show depletion of DNA methylation (Meissner et al. 2008). Conversely, gene

bodies generally have high DNA methylation levels. There is evidence that this is linked to the presence of H3K36me3 which is enriched over the gene body and is highly correlated with transcription levels (Mikkelsen et al. 2007; Maunakea et al. 2010; Baubec et al. 2015). H3K36me3 was suggested to have a role in preventing pervasive transcription (Carrozza et al. 2005). Other histone marks such as H3K79me3 (Ng et al. 2003) and H2BK120ub1 are also enriched in the bodies of active gene in eukaryotes. Studies of H2B ubiquitylation in yeast suggest that this mark contributes to reassembly of nucleosomes during transcription elongation (Fleming et al. 2008; Batta et al. 2011), as well as to deposition of H3K4me3 in the promoter region (Sun and Allis 2002) and H3K79me3 over the gene body. However, the role of H2B ubiquitylation in the deposition of other histone marks is still a topic of debate (Shahbazian et al. 2005; Nakanishi et al. 2009; Vethantham et al. 2012). Active enhancers, which are marked by H3K27ac and H4K20me1, are also part of euchromatic regions of the genome and are important to regulate gene activity. The presence of specific combinations of histone PTMs correlates strongly with certain genomic features and several studies have been able to use the chromatin signatures to predict functional regions such as promoters or enhancers of active genes (Hon et al. 2009; Heintzman et al. 2007).

1.1.1.3. Heterochromatin and its epigenetic marks

Heterochromatin regions are compact and transcriptionally repressed at interphase. These chromosome regions are generally gene-poor, enriched for H3K9me2/3 or H3K27me3, late replicating and frequently associated with the

nuclear lamina (Peric-Hupkes et al. 2010; Guelen et al. 2008; Ryba et al. 2010). The association with nuclear lamina correlates with gene repression, emphasizing the role of nuclear organization in transcription (Peric-Hupkes et al. 2010; Reddy et al. 2008). Advances in electron microscopy allowed heterochromatin to be subdivided into constitutive or facultative depending on the degree of compaction (reviewed in Trojer and Reinberg 2007) and gene repression (permanent or conditional, respectively) (Beck et al. 2010).

1.1.1.3.1. Constitutive Heterochromatin

Constitutive heterochromatin is associated with repetitive DNA for example at centromeres, pericentromeric regions, transposons and telomeres in mammalian cells (Peters et al. 2001; Mikkelsen et al. 2007). Molecular characteristics of these regions include enrichment of H3K9me3 concomitantly with histone deacetylation and high DNA methylation levels, together establishing a fully and permanently repressive chromatin environment refractory to transcription factor and RNA polymerase binding. The HP1 family of proteins, which are able to recognize H3K9me3 through their chromodomain, are enriched over constitutive heterochromatic regions, and play an important role in maintaining the repressive chromatin environment. HP1 proteins have been suggested to contribute to the perpetuation of the locked chromatin state by recruiting H3K9 histone methyltransferase SUV3-9 (Bannister et al. 2001; Lachner et al. 2001), and *de novo* DNA methyltransferases (DNMT3A/B) that are responsible for CpG DNA methylation (Lehnertz et al. 2003). HP1 proteins are also thought to contribute

to the compaction and high-ordered structural organization of constitutive heterochromatin (Thiru et al. 2004). DNA methylation can be recognized by methyl-CpG binding proteins, for example MeCP2, which contributes to recruitment of SUV3-9 (Fuks et al. 2003). H4K20me3 is also enriched in constitutive heterochromatin domains and was shown to contribute to nucleosome compaction (Lu et al. 2008). Several chromatin writers and readers contribute to feedback mechanisms that guarantee the transcriptional repression of constitutive heterochromatin. Cells lacking H3K9 methyltransferases SUV3-9-H1/H2 have been shown to acquire features of facultative heterochromatin such as H3K27me3, presumably to help maintain the repressed state (Peters et al. 2003).

1.1.1.3.2. Facultative Heterochromatin

Facultative heterochromatin has a more transient nature than constitutive heterochromatin and is more commonly found in higher eukaryotes. This type of chromatin organization is found in genes for which transcriptional status can change upon a specific stimulus or relocalization in the nucleus, for instance, during development (reviewed by Trojer and Reinberg 2007). Examples of facultative heterochromatin in mammals are autosomal imprinted genes, the inactive X chromosome in females, and developmental regulator genes such as the Homeotic (*Hox*) genes, or tissue-specific genes that are expressed in a temporal/spatial specific manner.

Facultative heterochromatin domains are highly dynamic and undergo major changes during development (Mohn et al. 2008). They are mainly characterized by the accumulation of the histone modifications H3K27me3 and H2AK119u1 deposited by the Polycomb repressive complexes PRC2 and PRC1, respectively. H3K27me3 prevents acetylation of the same histone residue, a mark of actively transcribed chromatin, by the histone acetyltransferase p300. Replacement of canonical H2A by the variant macroH2A is a feature of heterochromatin in the inactive X chromosome (Costanzi and Pehrson 1998) as well as at imprinted genes (Choo et al. 2006). MacroH2A increases internucleosome contacts inhibiting TF binding and chromatin remodeling (Doyen et al. 2006). Histone H1 was also suggested to have a role in the local compaction of heterochromatin regions (reviewed by Trojer and Reinberg 2007).

In mouse embryonic stem cells (mESCs), H3K27me3 occupied promoters are often co-occupied by H3K4me3 deposited by Trithorax proteins at what are referred to as bivalent promoters (Azuara et al. 2006; Bernstein et al. 2006a). This co-occupancy is thought to contribute to a poised transcriptional state that allows for activation of developmental genes linked to a change of cell fate (reviewed in Voigt et al. 2013). The mechanisms by which Polycomb proteins are recruited, contribute to silencing and propagate H3K27me3/H2AK119u1 marks through cell division to maintain correct cell identity through development will be discussed in more detail below.

1.2. Polycomb repressive complexes

1.2.1. Discovery of Polycomb Group genes in *Drosophila*

In *Drosophila*, the identity of body segments is determined by homeotic (*Hox*) genes located in the *Hox* clusters Antennapedia (ANTC) and Bithorax (BXC). *Hox* genes encode for transcription factors that are required for specific gene regulation in each segment of the developing fly, therefore, playing a crucial role in the determination of the anterior-posterior body axis of the adult. Genes in the *Hox* clusters are initially activated or repressed through the combined activity of maternally inherited products, as well as transcription factors synthesized in the early embryo (Jackle et al. 1992). A memory mechanism ensures that the expression pattern of *Hox* genes is faithfully maintained throughout development in the absence of the initiating factors (reviewed in Grossniklaus and Paro 2014).

Gain or loss-of-function mutations in *Hox* genes lead to homeotic transformations in which a specific body segment changes identity and acquires structures or functions characteristic of a different segment. Similar homeotic transformations were also observed in *Drosophila* with mutations outside of the *Hox* cluster (reviewed in Kennison 2004). The first of these mutants to be described was the *Polycomb* (*Pc*) gene mutant that led to ectopic expression of genes in the BXC cluster, suggesting that *Pc* would encode for a *trans*-acting repressor of *Hox* genes (Lewis 1978). For instance, *Pc* heterozygous mutant male flies acquire partial ectopic sex combs in their

second and third pair of legs, resembling the sex combs on the first pair of legs in wild-type male flies (Figure 1), due to failure to repress the *Hox* gene *Scr* (*Sex combs reduced*) (reviewed in Kennison 2004). Other genes outside of the *Hox* cluster in which mutations produce similar homeotic transformations were subsequently annotated as Polycomb family genes (Jürgens 1985; Duncan 1982). A study of *Esc* (*Extra sex combs*) mutants showed that this gene encodes for a factor responsible for the maintenance of the repressive state of *Hox* genes rather than the specification of which genes are expressed in each body segment (Struhl and Akam 1985). Analysis of mutations that lead to loss of function of *Hox* genes led also to the identification of the Trithorax group of genes which are Polycomb antagonists, required to maintain gene activation (as reviewed in Kennison 1995; Schuettengruber et al. 2007; Geisler and Paro 2015).

The isolation of Polycomb genes in *Drosophila* led to the identification of orthologues with similar function across different species, including mammals (van Lohuizen et al. 1991; van der Lugt et al. 1994; Coré et al. 1997), the nematode *Caenorhabditis elegans* (Bender et al. 2004; Holdeman et al. 1998) and the plant model organism *Arabidopsis thaliana* (reviewed in Kim and Sung 2014).

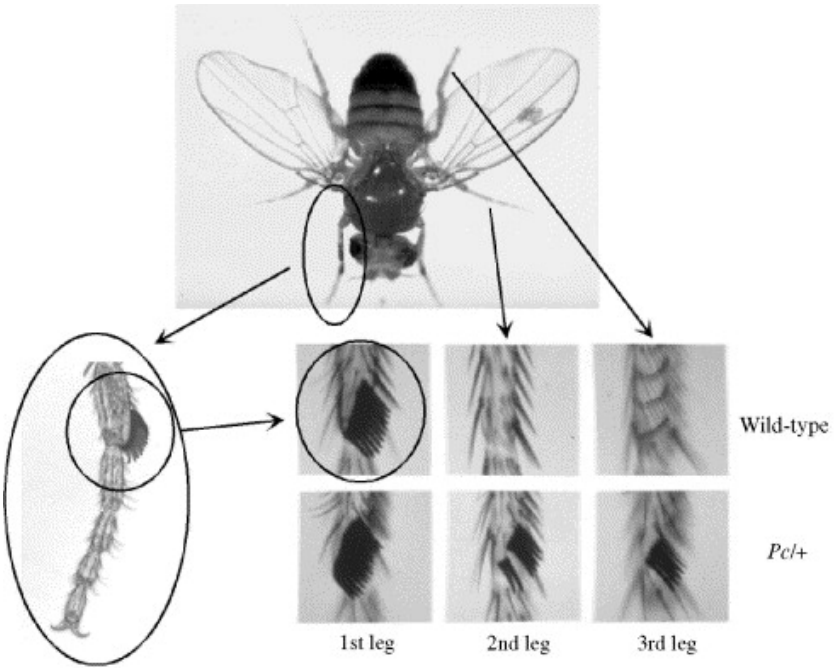


Figure 1: Extra sex combs phenotype in a heterozygous *Pc* mutant in *Drosophila.* In *Drosophila*, sex comb teeth are only found on the first leg of a wild-type individual as shown on the top panel. Mutations in Polycomb group genes, namely *Pc* gene, result in derepression of the *Hox* gene *Scr* and consequent ectopic sex combs on the second and third legs of male individuals. This phenotype has been used to identify new PcG genes in *Drosophila*. (Figure reproduced from Kennison, 2004).

1.2.2. Composition and architecture of PcG complexes in *Drosophila* and mammals

Since the initial description of Polycomb genes in *Drosophila*, extensive biochemical characterization of Polycomb proteins have been carried out across different species. Insights into the composition and structure of different Polycomb complexes contributed to the understanding of how these chromatin-repressors accomplish their functions. Two major multiprotein

complexes, PRC1 (Polycomb Repressive Complex 1) and PRC2 (Polycomb Repressive Complex 2) are the best characterized and the most functionally conserved in mammals and in *Drosophila*. In this section I will review the composition and structure of these mammalian and *Drosophila* Polycomb complexes. PhoRC (Pho Repressive Complex), a two subunit Polycomb complex, was also identified in *Drosophila*. PhoRC function seems not to be paralleled in mammals, despite the presence of some homologous proteins. However, the mechanism of Polycomb recruitment via PhoRC in *Drosophila* provided the first clues for research of recruitment mechanisms in mammals as described below.

1.2.2.1. Polycomb Repressive Complex 1: canonical and non-canonical complexes

Polycomb Repressive Complex 1 (PRC1) is the most heterogeneous of the major Polycomb complexes, notably in mammals in which each Polycomb subunit present in *Drosophila* has up to six different homologs, resulting in a myriad of possible subunit combinations. The only core component present in all different types of PRC1 complexes is the catalytic subunit RING1B or its close homolog RING1A (corresponding to SCE [Sex comb extra], also called dRING, in *Drosophila*). This subunit plays an architectural role, as it constitutes the core to which other proteins bind to assemble into a PRC1 complex. RING1A/B catalyzes deposition of a monoubiquitin moiety on lysine 119 (lysine 118 in flies) of histone H2A (H2AK119u1 in mammals or H2AK118u1 in flies) through interaction with a conjugate E2 ligase (Figures 2

and 3) (Wang et al. 2004b; de Napoles et al. 2004; McGinty et al. 2014). RING1A/B activity requires its heterodimerization with one of six PCGF proteins, PCGF1-6 (homologs of PSC [Posterior sex combs] and SU(Z)2 [Suppressor 2 of zeste] in *Drosophila*) through their highly conserved N-terminal RING domains to form an E3 ligase (Elderkin et al. 2007; Cao et al. 2005). The presence of different PCGF proteins in the complex modulates its activity as observed *in vitro* (Taherbhoy et al. 2015) and *in vivo* (Blackledge et al. 2014). A study published in 1995 noted that the *l(3)73Ah* gene in *Drosophila* has significant homology with *Psc* and *Su(z)2* (Irminger-Finger and Nöthiger 1995), but it was not until recently that L(3)73AH was shown to contribute to H2AK118u1 deposition in *Drosophila* cells. This protein is highly homologous to mammalian PCGF3 (Lee et al. 2015).

The heterodimer between RING1A/B and one of the PCGF proteins can complex into two different types of PRC1 complexes: canonical and non-canonical. Canonical PRC1 complexes are defined by the presence of one of the five CBX proteins present in mammals (CBX2/4/6/7/8) homologs of PC (Polycomb) protein in *Drosophila* (Figure 2). All of these CBX proteins contain a chromodomain that is able to bind methylated histone H3, albeit with different affinities as determined *in vitro* (Bernstein et al. 2006b). The ability of CBX proteins to bind H3K27me3 was suggested to mediate recruitment of canonical PRC1 complexes, implying that PRC2 is recruited to target sites first (Wang et al. 2004b). CBX7 is the most highly expressed homolog in mESCs, repressing expression of other CBX proteins that are only upregulated upon differentiation into specific lineages (Morey et al. 2012).

These findings suggest that different CBX proteins might play different roles, even though they bind to overlapping sites across the genome (Pemberton et al. 2014). Canonical PRC1 complexes also include one of the three polyhomeotic proteins PHC1, PHC2 and PHC3, homologs of Polyhomeotic (PH) in *Drosophila* (Figure 2). Polyhomeotic proteins can oligomerize through their Sterile Alpha Motif (SAM) domain contributing to PRC1 clustering in subnuclear structures called Polycomb bodies (Isono et al. 2013; Wani et al. 2016). Canonical PRC1 complexes include either PCGF2 or PCGF4 (the closest homologs to the *Drosophila* PSC and SU(Z)2) subunits, but not other PCGF proteins (Figure 2) (Gao et al. 2012; Hauri et al. 2016).

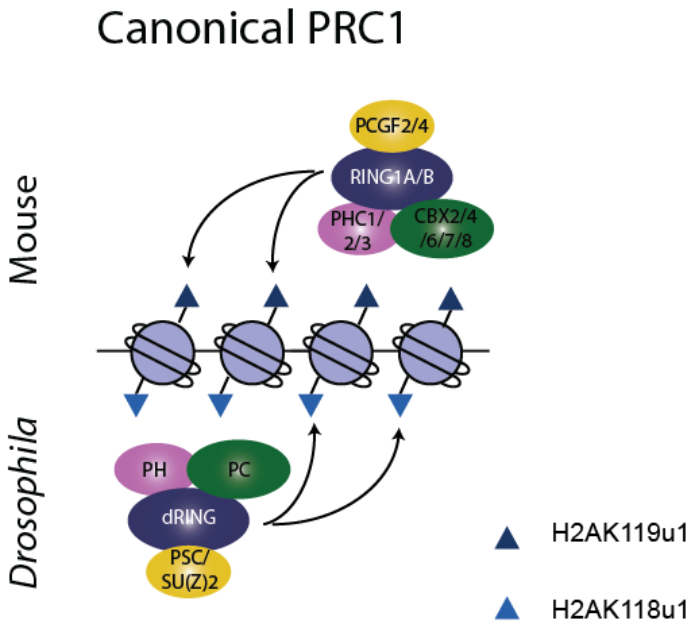


Figure 2: Canonical PRC1 composition in mouse and in *Drosophila*. Subunits of canonical PRC1 complex in mouse are shown on top. This complex catalyzes H2AK119u1. Homolog proteins in *Drosophila* are shown in the same colors at the bottom and catalyze the correspondent PTM, H2AK118u1. Names of alternative mutually exclusive subunits are separated by slashes. (Figure adapted from Bauer et al., 2016)

Non-canonical PRC1 complexes consist of the catalytic dimer RING1A/B and one of the six PCGF proteins (PCGF1-6) in complex with RYBP or YAF2 (the two homologs of dRYBP in *Drosophila*) (Figure 3) that compete with CBX proteins for the C-terminal binding surface in RING1B (Czypionka et al. 2007; Wang et al. 2010; Gao et al. 2012; Tavares et al. 2012). RYBP localizes to Polycomb targets in mammalian cells (Tavares et al. 2012; Morey et al. 2013), and contributes to the activity of non-canonical PRC1 complexes (Rose et al. 2016). Contrary to what was observed in *Drosophila*, depletion of RYBP is embryonically lethal in mouse (Pirity et al. 2005). The function of dRYBP is not fully established, but the observation that *dRYBP* mutants do not present classical homeotic transformations, and show no upregulation of genes in the *Hox* cluster, suggested an alternative role outside of the Polycomb complex. In fact, mutation of *dRYBP* concomitantly with mutations in Polycomb or Trithorax components enhances both phenotypes, suggesting a role in the maintenance of both active and inactive transcriptional states in *Drosophila* (Gonzalez et al. 2008). These observations suggest that RYBP role might not be fully conserved across species.

The PCGF subunit is important in defining different identities of PRC1 complexes (Figure 3) (Gao et al. 2012). This is particularly relevant since different PCGF proteins have been shown to interact specifically with other proteins that contribute to PRC1 recruitment (Farcas et al. 2012; He et al. 2013; Trojer et al. 2011) or function (Gao et al. 2014) (as further discussed in sections 1.2.3. and 1.2.4.). The most widely described example in mammals, and that is also represented in *Drosophila*, is the non-canonical PCGF1-PRC1

complex. In *Drosophila*, this complex (called dRAF for dRING associated factors) includes the core proteins PSC and dRING as well as the H3K36me2 demethylase dKDM2B (Figure 3) (Lagarou et al. 2008). This demethylase was shown to contribute towards H3K36me2 demethylation, as well as to enhance core PRC1 catalytic activity (Lagarou et al. 2008). Recent studies in mammals identified KDM2B as the direct homolog for dKDM2B, and characterized this as a component of PCGF1-PRC1. KDM2B is required for PCGF1-PRC1 recruitment to unmethylated CpG islands (Farcas et al. 2012; He et al. 2013; Wu et al. 2013), GC-rich regions with high proportion of CpG dinucleotides, found around the majority of gene promoters (reviewed in Deaton and Bird 2011). Importantly, since *Drosophila* lacks DNA methylation, which counteracts KDM2B mediated recruitment of PCGF1-PRC1 to CpG islands in mammals, the role of dRAF is not clear. Moreover, not all unmethylated CpG islands that are occupied by KDM2B in mammals show concomitant Polycomb occupancy, suggesting that other chromatin features might be necessary for KDM2B-mediated recruitment of PRC1 (Farcas et al. 2012; He et al. 2013).

The mammalian PCGF1-PRC1 complex also includes the proteins BCOR/BCORL1 and SKP1. The interaction of BCOR/BCORL1 with the PCGF1 RAWUL domain was shown to be crucial to generate a wide interface for KDM2B interaction (Wong et al. 2016). This example suggests that the RAWUL domain of different PCGF proteins probably constitutes an interface for specific interactions that might be relevant in conferring specific recruitment and/or activity to PRC1 subcomplexes (Junco et al. 2013).

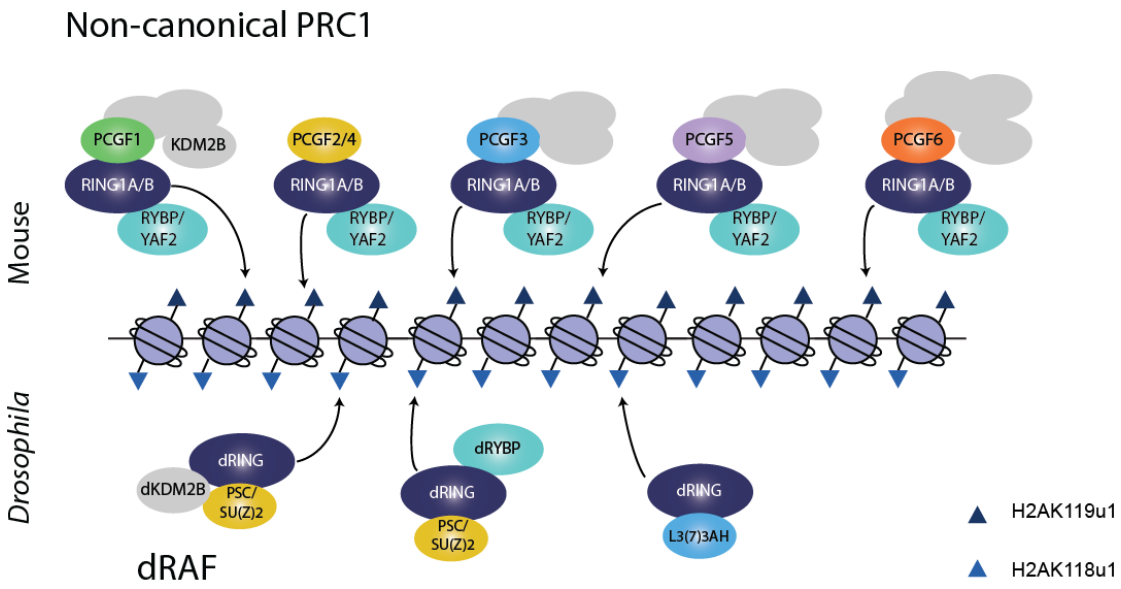


Figure 3: Diversity of non-canonical PRC1 complexes in mouse and in *Drosophila*. Subunits of non-canonical PRC1 variants in mouse are shown on top. Each variant is defined by the presence of a specific PCGF protein and its interactors (depicted in grey). PRC1 complex catalyzes H2AK119u1 in mouse. Homolog proteins in *Drosophila* are shown in the same colors at the bottom and catalyze the correspondent PTM, H2AK118u1. Names of alternative components are separated by slashes. (Figure adapted from Bauer et al., 2016)

1.2.2.2. Polycomb Repressive Complex 2: core and auxiliary subunits

Polycomb Repressive Complex 2 (PRC2) subunits are more conserved than PRC1 both in sequence and in diversity between flies and mammals (reviewed in Margueron and Reinberg 2011). The core subunits of the PRC2 multiprotein complex are the catalytic subunit EZH1 or EZH2 (in *Drosophila* represented only by Enhancer of zeste [E(Z)]), four alternative isoforms of EED (Extra sex combs [ESC] or Extra sex combs like [ESCL] in *Drosophila*),

SUZ12 (Suppressor 12 of Zeste [SU(Z)12] in *Drosophila*) and RbAp46/48 (called Chromatin assembly factor 1 [CAF1] or p55 subunit in *Drosophila*) (Figure 4). Recent crystal structures of the PRC2 complex highlighted the relevance of each core subunit (Jiao and Liu 2015; Justin et al. 2016).

EZH2, the main catalytic component has methyltransferase activity mediated by its SET domain, placing mono-, di- and trimethylation on lysine 27 of histone H3 (H3K27me1/2/3) (Ferrari et al. 2014). This protein can be replaced in the PRC2 complex by EZH1, which is expressed later during development and is also catalytically active (Ezhkova et al. 2011; Shen et al. 2008). However, several studies reported different functions for PRC2 complexes that include either EZH1 or EZH2, suggesting that alternative subunit composition might impact function, as seen for PRC1 complexes (reviewed in Aranda et al. 2015). The four different isoforms of EED in mammals contribute to the diversity of mammalian PRC2. However, here it is not clear if different isoforms define unique functions (Kuzmichev et al. 2004; Martin et al. 2006). EED has been shown to play a central structural role in the context of PRC2 since it interacts with EZH2 and RbAp46/48. Further evidence supports its crucial role, as both EED and SUZ12 are indispensable for EZH2 catalytic activity through maintenance of the positioning and structure of the SET activation loop of EZH2 (Jiao and Liu 2015). EED was also suggested to play a crucial role in propagating H3K27me3 through chromatin domains by recruiting the PRC2 complex to sites with previously placed H3K27me3 (Margueron et al. 2009). This is mediated by the recognition of H3K27me3 through the EED C-terminal portion. Consistent with the important role of

these subunits in the PRC2 complex, *Eed* (Schumacher et al. 1996), *Ezh2* (O'Carroll et al. 2001) and *Suz12* (Pasini et al. 2004) knockouts in mouse are embryonically lethal between days E7.5 and E8.5.

RbAp46/48 were initially biochemically identified as part of the PRC2 complex. These proteins are also present in other chromatin-related complexes with distinct functions, for example, the nucleosome remodeling and histone deacetylase NuRD complex (reviewed in Hennig et al. 2005; Alqarni et al. 2014). RbAp46/48 were shown to be part of a chromatin assembly complex and can bind H3-H4 tetramers to promote *de novo* nucleosome assembly on DNA (Verreault et al. 1996; Murzina et al. 2008; Zhang et al. 2013). RbAp46/48 were suggested to have a generic role in chromatin modifying complexes, likely mediating H3/H4 binding (Nekrasov et al. 2005). A recent genetic analysis of the *Drosophila* homolog *Caf1* shows that mutant flies have a phenotype resembling that resulting from *Hox* gene misregulation. Moreover, *Caf1* mutation enhances the previously described *Pc* phenotype in flies carrying both mutations (Anderson et al. 2011).

Besides the core PRC2 subunits, several accessory proteins have been identified both in *Drosophila* and in mammals (shown in green in Figure 4). These substoichiometric components contribute to the diversity of PRC2 complexes with roles in recruitment of the complex to target sites (Casanova et al. 2011; Pasini et al. 2010a; Peng et al. 2009; Kim et al. 2009; Kalb et al. 2014; Beringer et al. 2016; Cooper et al. 2016), but also in modulation of its histone methyltransferase activity (Hunkapiller et al. 2012; Sarma et al. 2008;

Li et al. 2011; Peng et al. 2009; Kalb et al. 2014; Grijzenhout et al. 2016). Recent studies have converged in defining two main variants of PRC2 complexes in mammals, here referred to as PRC2.1 and PRC2.2 (Figure 4) (Alekseyenko et al. 2014; Grijzenhout et al. 2016; Hauri et al. 2016). PRC2.1 includes one of three PCL proteins found in mammals (PHF1, MTF2 and PHF19 also known as PCL1, 2 and 3, respectively) and C10ORF12 or C17ORF96 (also known as EPOP). PRC2.2 includes the accessory subunits JARID2 and AEBP2. Homologs of some of these substoichiometric components have also been identified in *Drosophila* (bottom part of Figure 4). Previous studies assessed the role of each of the substoichiometric subunits individually. In light of the new classification of PRC2 subcomplexes it will be relevant to understand if auxiliary subunits act synergistically to confer specific functions to PRC2.

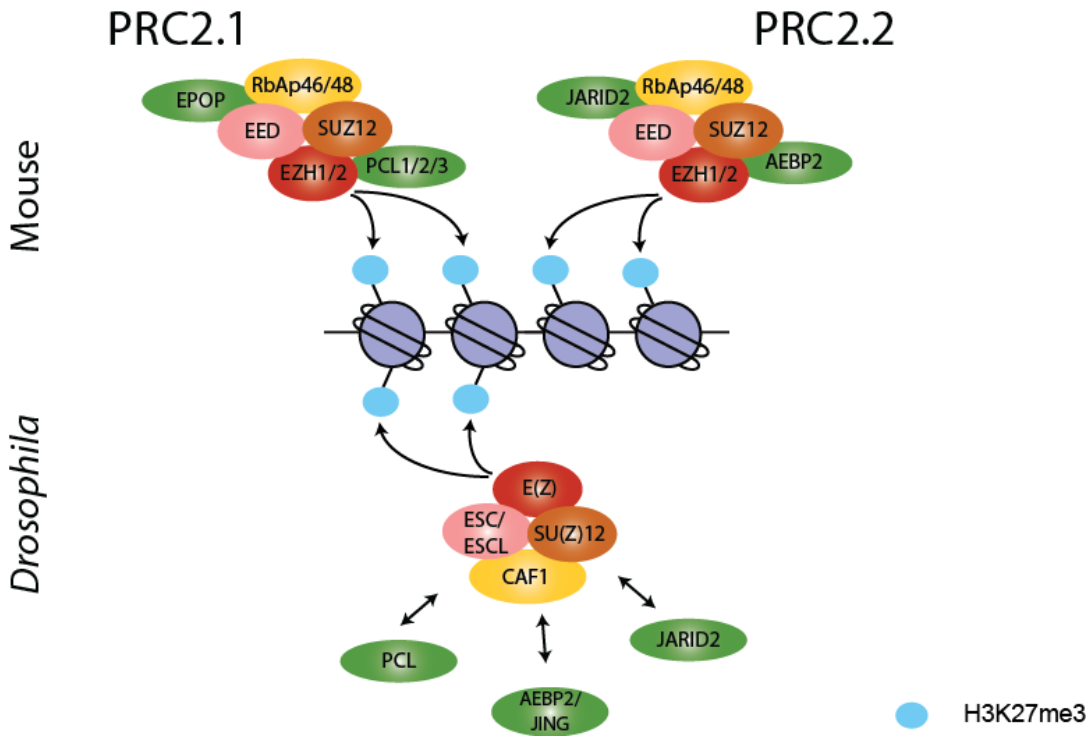


Figure 4: Diversity of PRC2 complexes in mouse and in *Drosophila*. Substoichiometric components of PRC2 (shown in green) define two alternative subcomplexes, PRC2.1 and PRC2.2 in mouse, as shown at the top. Some of these components are also present in *Drosophila* and interact with the four conserved core components EED, SUZ12, EZH1/2 and RbAp46/48 (homolog core components are shown in the same color for mouse and *Drosophila*). PRC2 complexes catalyze H3K27me3 in both species. (Figure adapted from Bauer et al., 2016 and Grijzenhout et al., 2016).

1.2.3. Function of Polycomb complexes

As previously mentioned, Polycomb genes were identified through their role in maintaining the silent chromatin state of the *Hox* gene cluster in *Drosophila*. Since then, many paralog genes were identified in vertebrates and the function of Polycomb proteins in repressing the *Hox* cluster and many other

developmental regulator genes was shown to be crucial for mammalian development. The mechanisms through which PRC1 and PRC2 contribute to gene activity have been extensively investigated, but are not yet fully understood and might be dependent on the cell type, differentiation state and/or organism.

1.2.3.1. Gene silencing by PRC1 and PRC2 mediated histone modifications

PRC1 and PRC2 catalyze the post-translational histone modifications H2AK119u1 and H3K27me3, respectively. Moreover, PRC2 can also mono- and dimethylate H3K27 in regions of the genome to which PRC2 components are not stably bound (Ferrari et al. 2014). Enrichment of both H2AK119u1 and H3K27me3 occurs mostly over silent gene promoters (Endoh et al. 2012; Tavares et al. 2012; Sarma et al. 2008; Kloet et al. 2016), while H3K27me2 is distributed in large chromatin domains including intergenic and intragenic regions and representing 70% of the total H3 (Ferrari et al. 2014). H3K27me1 is associated with gene activation and it will be discussed in section 1.2.3.3.

H2AK119u1 and H3K27me3 enrichment have been suggested to interfere with RNA polymerase II function. Stock et al. (2007) showed that H2AK119u1 present at the promoter of bivalent genes is required to prevent transcriptional elongation by RNA polymerase II. Loss of RING1B in mESCs leads to transcriptional activation of bivalent genes due to release of RNA Polymerase II (Stock et al. 2007). A similar observation was reported in *Drosophila* (Wang

et al. 2004a). Importantly, an independent study on the role of H2AK119u1 in macrophages provided more direct evidence that RNA polymerase II stalling at promoters is related to the histone mark and not to PRC1 complex occupancy (Zhou et al. 2008). H3K27me3 enrichment has also been correlated with the presence of poised RNA polymerase II at gene promoters, as two demethylases UTX and JMJD3 (also known as KDM6A and KDM6B) seem to contribute to transcriptional elongation partially through the removal of methylation (Seenundun et al. 2010; Chen et al. 2012). A study in *Drosophila* using histone H3 mutants that mimic the unmethylated H3K27 form of H3 showed that methylation is required for silencing of Polycomb targets (Pengelly et al. 2013), confirming previous observations made in *E(z)* mutant flies (Cao et al. 2002), but emphasizing the role of H3K27me3 over PRC2 occupancy. Similar studies would be important in mammals to better understand the role of H3K27me3 in gene silencing, as Stock et al. (2007) reported that this mark is retained at reactivated bivalent gene promoters in the absence of H2AK119u1. Generating a mammalian H3K27 mutant to address the role of H3K27me3 would be challenging due to the multiple H3 genes occurring in clusters on different chromosomes (as discussed in Moazed 2013). However, a mutation identified in pediatric glioblastomas that results in a mutant histone with a methionine replacing lysine 27 on histone H3 (H3K27M) was shown to lead to an overall reduction of H3K27me3 (Lewis et al. 2013). Further investigation in this context might provide valuable information on the role of H3K27me3 for gene silencing.

Besides direct interference with RNA Polymerase II function, both PRC1 and PRC2 were proposed to regulate transcription by preventing the deposition of histone modifications linked to gene activation. Studies in *Drosophila* and in mammals point to the inhibition of binding of the histone acetyltrasferase CBP to Polycomb targets, preventing deposition of the activation linked modification H3K27ac (Pasini et al. 2010b; Tie et al. 2016). Both complexes were also suggested to block H3K4me3 deposition, either by recruiting the demethylase JARID1A to Polycomb targets in the case of PRC2 (Pasini et al. 2008), or by direct inhibition of the histone methyltransferase MLL3 by H2AK119u1 (Nakagawa et al. 2008).

Even though H3K27me2 constitutes most of the total H3K27 methylation in mESCs, its role has been largely overlooked. Ferrari et al. (2012) showed that this mark is crucial to prevent aberrant accumulation of H3K27ac over enhancer elements preventing their activation and hence upregulation of their gene targets. This finding provided a new mechanism for how PRC2 impacts on gene transcription, i.e. through H3K27me2 regulation of *cis* regulatory elements.

1.2.3.2. A role for Polycomb complexes in chromatin compaction and nuclear organization

While the role of H3K27me3 and PRC2 in maintaining transcriptional repression seems to depend mostly on its catalytic activity, there is evidence that PRC1 has functions independent of catalyzing H2AK119u1 (or

H2AK118u1 in *Drosophila*). While *Drosophila* mutants expressing H3R27 that cannot acquire methylation showed classical Polycomb homeotic transformations (Pengelly et al. 2013), in the corresponding mutant for the H2A residue where PRC1 deposits ubiquitin (H2AK118R mutant) there are no major changes in the expression of *Hox* genes (Pengelly et al. 2015). This suggests that H2AK118u1 is dispensable for the repression of *Hox* genes. However, H2A mutant flies, as well as *Sce* mutants that are unable to deposit ubiquitin due to impaired catalytic activity of PRC1, do not develop properly, suggesting that H2AK118u1 might be required to regulate expression of other developmental-related genes. Importantly, complete deletion of *Sce* leads to aberrant expression of *Hox* genes, pointing towards an alternative PRC1-dependent mechanism responsible for gene repression (Pengelly et al. 2015). Similar observations were made in mice carrying a mutation that renders RING1B catalytically inactive (Illingworth et al. 2015). However, a major caveat of this latter study is that RING1A, the close homolog of RING1B, is still expressed, potentially contributing to low levels of H2AK119u1 that impact in gene expression, at least of some genomic targets. Even though RING1B catalytic mutant embryos develop further than *Ring1B* knockout (Voncken et al. 2003), their development is arrested around E16.5 suggesting that H2AK119u1 does play a role in developmental gene regulation (Illingworth et al. 2015).

Both of the aforementioned studies point to a non-catalytic repressive function for PRC1. In fact, previous experiments *in vitro* (Francis et al. 2004) and *in vivo* (Eskeland et al. 2010) suggested that PRC1 is able to mediate chromatin

compaction at gene targets independently of its catalytic activity and of H3K27me3. This is possibly linked to the interference of PRC1 with chromatin remodelers of the SWI/SNF family demonstrated *in vitro* (Shao et al. 1999). The identity of the PRC1 subunit responsible for chromatin compaction is still under debate, with some *in vitro* experiments suggesting that it might be different in mammals and *Drosophila* (Grau et al. 2011; Lo and Francis 2010). A study using mammalian cells proposed that PHC proteins, namely PHC1 and PHC2, contribute to chromatin condensation and gene silencing through their ability to promote clustering of PRC1 complexes, emphasizing the role of canonical PRC1 complexes for silencing independently of H2AK119u1 deposition (Isono et al. 2013). Chromatin compaction by PRC1 is also likely to contribute to gene silencing through activation of PRC2, since this complex has been shown to be more active on compacted chromatin (Yuan et al. 2012). EZH1, a component of PRC2 was suggested to contribute to chromatin compaction, but its role during development is still not clear as this protein is mostly expressed in non-proliferative adult tissues (Margueron et al. 2008).

Recent interest in understanding the relevance of nuclear architecture and its impact on gene transcription showed that metazoan genomes are organized in Topologically Associated Domains (TADs). TADs correspond to genomic regions in the nucleus within which sequences interact with high frequency. TADs are defined by boundaries that prevent interactions between genomic regions in different TADs. TADs are characterized by the presence of specific histone marks and usually encompass genes in the same transcriptional state (Dixon et al. 2012; Dowen et al. 2014). A specific group of TADs that include

genes repressed during development are marked by Polycomb occupancy and its histone marks. Therefore, several studies have attempted to understand the role of Polycomb, namely PRC2, in organizing the genome in TADs. However, analysis of TADs in mESCs as well as on the inactive X chromosome showed that the genome does not undergo major conformational changes in the absence of PRC2 and H3K27me3, not even on the inactive X chromosome where PRC2 is highly enriched (Denholtz et al. 2013; Nora et al. 2012). On the other hand, studies in *Drosophila* showed that Polycomb TADs have specific features that are lost upon loss of function of the *Ph* gene, whose product is known to contribute to gene repression through PRC1 clustering (Boettiger et al. 2016; Wani et al. 2016). The role of PRC1 in the formation of mammalian TADs is still to be explored.

Due to the ability of Polycomb proteins to compact chromatin, it is not surprising that these complexes might contribute to chromatin looping inside each TAD. Studies of the *Hox* cluster genes in *Drosophila* were the first to characterize physical long distance interactions between PRE (Polycomb Response Elements) sequences as well as between PREs and promoters, establishing a high-order nuclear structure related to Polycomb occupancy (Lanzuolo et al. 2007; Comet et al. 2011). This chromatin looping results in the formation of Polycomb bodies, visible by microscopy in *Drosophila* (Gonzalez et al. 2014), and positioning of Polycomb targets in these regions stabilizes the maintenance of gene silencing (Lanzuolo et al. 2007; Bantignies et al. 2011). Polycomb proteins were proposed to contribute to the three-dimensional chromatin organization of its targets, which fits with the

observation that non-Polycomb target silent genes form different networks of co-associations suggesting that genome organization is not a direct result of transcriptional output (Bantignies et al. 2011; Gonzalez et al. 2014).

Polycomb proteins have also been shown to play a role in establishing chromatin interactions in mammals. Loss of the PRC2 subunit EED in mESCs was shown to disrupt both intrachromosomal and interchromosomal interactions (Denholtz et al. 2013). A recent publication showed that depletion of PRC1 subunits RING1A/B contributes to loss of promoter-promoter interactions between Polycomb target genes resulting in their upregulation. The same authors also showed that enhancer-promoter interactions are not disrupted in the absence of RING1B, even though PRC1 plays a role in keeping these enhancers in a silent state (Schoenfelder et al. 2015). Other studies suggested roles for loop formation mediated by PRC1 and PRC2 at specific gene targets (Kondo et al. 2014; Terranova et al. 2008; Tiwari et al. 2008).

In sum, there is evidence for PRC1 and PRC2 mediated gene repression occurring independently of the catalytic activity of the complexes, even though the mechanisms still need to be further characterized. Given the myriad of Polycomb complexes present in mammalian cells, it will also be important to disentangle the role of each component/subcomplex in these alternative silencing mechanisms.

1.2.3.3. Polycomb function in gene activation

Polycomb complexes were first identified for their role in maintaining transcriptional repression. However, a few examples in the literature suggest that particular subcomplexes might act as transcriptional activators in specific cellular contexts, by mechanisms that are, in most of the cases, not yet fully understood.

As previously mentioned, PRC2 catalyzes all three H3K27 methylation states (me1/2/3), which are concentrated over distinct regions of the genome. While H3K27me2/3 correlate with gene silencing, H3K27me1 is deposited over the gene bodies of highly expressed genes co-occupied by H3K36me3. Depletion of H3K27me1 over these genes leads to a decrease in gene expression, suggesting that PRC2 can have a role in active gene transcription potentially through increasing nucleosome mobility during elongation (Ferrari et al. 2014).

EZH1 and MEL18 (PCGF2) were proposed to contribute to gene activation during cardiac differentiation (Mousavi et al. 2012; Morey et al. 2015), even though the mechanisms by which this occurs are unclear (as discussed in Riising and Helin 2012). The replacement of CBX7-PRC1 by CBX8-PRC1 during differentiation was also linked to gene activation (Creppe et al. 2014). Recently, two publications showed that EPOP, a substoichiometric component of PRC2 is required to fine-tune transcription at bivalent promoters (Beringer et al. 2016; Liefke et al. 2016). EPOP binds directly to ElonginBC

and prevents excessive PRC2 binding and complete transcriptional repression. Instead, genes occupied by EPOP-PRC2 at their promoters are expressed at low levels, which was proposed to be important to maintain transcriptional plasticity in ES cells (Beringer et al. 2016).

The most well described mechanism for Polycomb-mediated gene activation was recently published and concerns a non-canonical PRC1 subcomplex including PCGF5, AUTS2 and CK2. AUTS2 and CK2 are crucial for the transcriptional activation function of this PRC1 complex. First, CK2 inhibits the catalytic activity of PRC1 by phosphorylating its catalytic subunit RING1B preventing H2AK119u1 deposition. Additionally, AUTS2 recruits the histone acetyltransferase P300, which leads to robust gene activation. Depletion of AUTS2 in the brain, where it is usually highly expressed during development, leads to neurodevelopmental defects in mouse (Gao et al. 2014).

1.2.4. Mechanisms of *de novo* recruitment and propagation of Polycomb complexes

Understanding how Polycomb complexes are recruited to their targets has been a topic of extensive investigation. However, answering this question has been challenging, particularly in mammalian systems where there seem to be numerous mechanisms involved, some of which are specific to certain genomic loci and others with a more broad function. Importantly, these studies also helped to understand how Polycomb complexes maintain a silent

transcriptional state from one cell generation to the next, through propagation of their marks.

1.2.4.1. Lessons from *Drosophila*

Polycomb Response Elements (PREs) were identified as DNA sequences required for Polycomb-mediated repression of *Hox* genes in *Drosophila* (Simon et al. 1993; Busturia et al. 1997). More recently, the use of chromatin immunoprecipitation revealed that the Polycomb complexes PRC1 and PRC2 associate with these elements to mediate repression, and that PRE-mediated Polycomb recruitment occurs genome-wide (Papp and Müller 2006; Oktaba et al. 2008; Schuettengruber et al. 2009). PREs are DNA elements that encompass binding motifs for several DNA binding factors. However, none of the PRC1 or PRC2 components in *Drosophila* has a known DNA-binding activity, suggesting that PRC1 and PRC2 recruitment to PREs is mediated by a DNA-binding protein that interacts with core Polycomb complexes. Many DNA-binding factors were identified in PREs and their role in recruitment of Polycomb complexes as well as gene silencing has been widely studied. However, mutations for many of these DNA-binding factors do not produce a classic Polycomb phenotype, raising some questions about their function as Polycomb recruiters (as reviewed in Kassis and Brown 2013; Müller and Kassis 2006). The DNA-binding protein PHO is the exception as *Pho* mutant flies resemble *Polycomb* mutants, and the PHO protein has DNA-binding activity and is enriched over PREs. This protein was shown to complex with SFMBT to form the Polycomb complex PhoRC in *Drosophila*. Due to its

interaction with PRC2, PHO was proposed to be the main Polycomb recruiter in *Drosophila* (Wang et al. 2004b) even though other factors might modulate binding and activity in each PRE. PHO was also proposed to directly interact with PRC1 (Mohd-Sarip et al. 2002) but this idea was subsequently disregarded because Wang et al. (2004b) observed that PRC1 enrichment at PRE was lost upon depletion of PRC2. These findings led to the acceptance of a hierarchical model of Polycomb recruitment. According to this model PRC2 is recruited to PREs by direct interaction with PHO and catalyzes H3K27me3 deposition. The methylation mark is then recognized by the chromodomain of the Polycomb subunit of PRC1 (Figure 5).

This model was recently reappraised (Kahn et al. 2016) taking into account lessons from Polycomb recruitment in mammals (as referred to in section 1.2.4.4.) as well as the evidence that PRC1 subunits do, in fact, directly interact with PhoRC (Frey et al. 2016). Concepts learned from the models initially proposed for Polycomb recruitment in *Drosophila* guided research conducted in mammals for many years, as will be discussed in the next section.

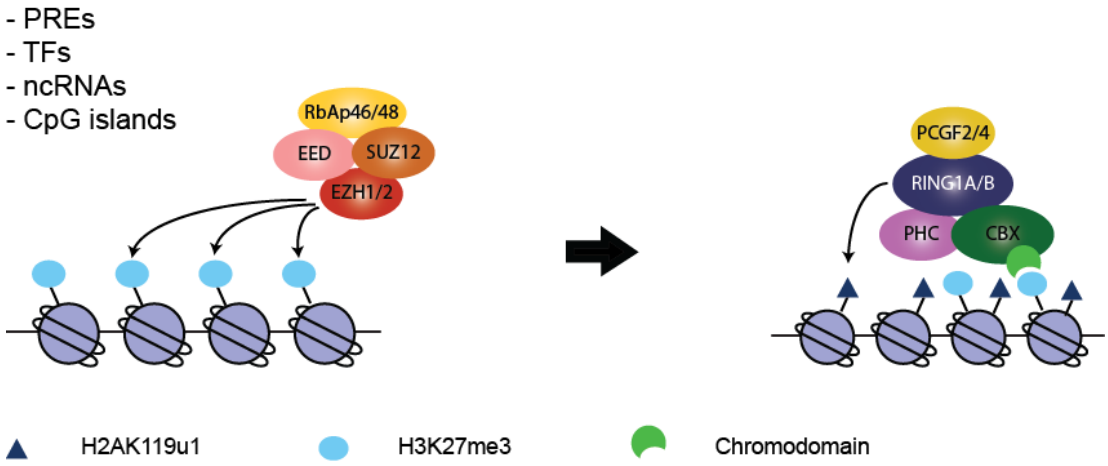


Figure 5: Hierarchical model of Polycomb recruitment. According to this model, PRC2 is the first complex to be recruited to chromatin by the action of PREs (in *Drosophila*), transcription factors, non-coding RNAs or binding of substoichiometric subunits to CpG islands, and deposits H3K27me3. This histone PTM is then recognized by CBX proteins chromodomain bringing PRC1 to the same target regions to place H2AK119u1.

1.2.4.2. The quest for a mammalian PRE

With the discovery of PHO-mediated DNA binding recruiting Polycomb complexes to PREs in *Drosophila*, many laboratories attempted to identify the corresponding mammalian PRE as well as the Polycomb subunits that potentially mediate DNA binding in mammals (as reviewed in Bauer et al. 2015). Given the proposed hierarchical model of recruitment, attention was initially focused on understanding how PRC2 is recruited to chromatin as this was expected to determine PRC1 binding through recognition of H3K27me3. This model was recently revised (see section 1.2.4.4.) although information

obtained in the search for a mammalian PRE is still relevant to fully understand Polycomb recruitment, as summarized in this section.

The direct homolog of PHO in mammals is the transcription factor YY1. YY1 is able to rescue the *Drosophila Pho* mutant phenotype and Polycomb-mediated repression (Atchison et al. 2003; Srinivasan et al. 2005). However, genome-wide analysis in mESCs indicated that there is little overlap between YY1 and Polycomb complexes suggesting that this mechanism of targeting is not widely conserved (Mendenhall et al. 2010). It cannot be ruled out that YY1 facilitates the recruitment of Polycomb to a few targets in mammalian cells, for example those that have been proposed to have a PRE (Sing et al. 2009; Woo et al. 2010; Mendenhall et al. 2010).

Transcription factors other than YY1 have been implicated in Polycomb targeting in mammals. Examples include REST and RUNX1, which have been shown to interact with Polycomb subunits and to be necessary for PRC1 binding at a small subset of genomic loci (Dietrich et al. 2012; Yu et al. 2012). SNAIL1 was proposed to have a similar function in PRC2 recruitment (Herranz et al. 2008). The transcription factors E2F6, MAX and MGA were found in a subtype of non-canonical PRC1 containing PCGF6 and they might play a role in its recruitment, but this has not been fully investigated (Gao et al. 2012; Ogawa et al. 2002). E2F6 was shown to have a role in recruitment of PCGF6-PRC1 though its interaction with L3MBTL2 (Trojer et al. 2011). Even though all these TFs might play a role in Polycomb recruitment to specific loci limited to certain stages of development or to specific cell types, their

expression and distribution is insufficient to explain how Polycomb complexes are specifically recruited to all of their target genes in all cell types and developmental stages.

Efforts to characterize potential mammalian PRE sequences led to the observation that Polycomb targets in mESCs correlate with CpG islands, characterized by the presence of GC-rich elements without any conserved transcription factor binding-motifs (Mendenhall et al. 2010). Other GC-rich regions outside CpG islands often accumulate DNA methylation, which is refractory to Polycomb binding (Reddington et al. 2013).

Binding of Polycomb to CpG islands may reflect binding preferences for CpG islands specific chromatin features. Notably, several chromatin features correlate with or are required to promote PRC2 recruitment: gene silencing (Riising et al. 2014), deacetylation of H3K27 (Morey et al. 2008; Reynolds et al. 2012; Schmitges et al. 2011), histone variant H3.3 incorporation at bivalent promoters (Banaszynski et al. 2013) and binding and activity of the H3K9 histone methyltransferases G9a/GLP (Mozzetta et al. 2014). This suggests that Polycomb recruitment in mammals is more complex than was suggested from studies in *Drosophila*, and that many factors might be involved.

In mammals, two Polycomb subunits have been suggested to directly bind DNA and to contribute to Polycomb recruitment: JARID2, a substoichiometric PRC2 subunit (Kim et al. 2003) and KDM2B, a histone demethylase that

interacts with a subtype of non-canonical PRC1 containing PCGF1 (Farcas et al. 2012). Both were proposed to have affinity for CpG islands.

JARID2 was suggested to preferentially bind GC-rich DNA based on *in vitro* experiments (Li et al. 2010). Genome-wide ChIP-seq analysis demonstrated that JARID2 is present at a proportion of PRC2 sites, and that it is required for PRC2 recruitment to these sites (Pasini et al. 2010a; Landeira et al. 2010). However, it is not clear if this recruitment is mediated by a direct interaction between JARID2 and DNA *in vivo*. Indeed, a recent publication identified a ubiquitin-interacting motif (UIM) on the N-terminus of JARID2 and that is required to mediate Polycomb recruitment through recognition of H2AK119u1 (Cooper et al. 2016). This argues against the DNA binding ability of JARID2 being important for PRC2 recruitment.

KDM2B has a ZF-CxxC domain, which binds to unmethylated CpG dinucleotides present in CpG islands (Farcas et al. 2012). In a subset of CpG islands KDM2B overlaps with and is necessary for Polycomb occupancy, although not all KDM2B-bound regions are enriched for Polycomb proteins (Farcas et al. 2012; He et al. 2013; Wu et al. 2013; Blackledge et al. 2014). These observations added support to the idea that CpG islands constitute mammalian PREs, even though it is likely that there is more than one mechanism of Polycomb recruitment to these targets (reviewed in Bauer et al. 2015).

1.2.4.3. Recruitment of Polycomb by interaction with RNA

The inactive X chromosome in cells of female mammals is a classic example of facultative heterochromatin characterized by the enrichment of Polycomb complexes, PRC1 and PRC2, and their histone marks, H2AK119u1 and H3K27me3 (Plath et al. 2003; Silva et al. 2003; de Napoles et al. 2004). The process of X chromosome inactivation is triggered by expression of a long non-coding RNA, Xist (X inactive specific transcript), which was shown to directly mediate PRC2 recruitment (Zhao et al. 2008) through interaction with the substoichiometric component JARID2 (da Rocha et al. 2014) (further discussed in section 1.3.4.). Even though this observation has been challenged over the years (reviewed in Brockdorff 2013), it gave rise to the idea that Polycomb could be recruited through a direct interaction with RNA. Since then, several cases have been reported in which different RNAs contribute to recruitment of PRC1 and PRC2 to specific gene targets in *cis* (as on the inactive X chromosome) or in *trans*. A systematic screen identified a few hundred non-coding RNAs (ncRNA) transcribed from the human *Hox* cluster, from which HOTAIR (*Hox* transcript antisense RNA) was shown to bind PRC2. This ncRNA acts in *trans*, so even though it is transcribed from the *HoxC* locus, it is necessary to recruit PRC2 and establish silencing of the *HoxD* locus located on a different chromosome (Rinn et al. 2007; Tsai et al. 2010). Another example is the RNA ANRIL expressed from the *INK4b/ARF/INK4a* locus. ANRIL RNA binds to CBX7, a component of canonical PRC1 (Yap et al. 2010) previously shown to bind RNA *in vitro* (Bernstein et al. 2006b). CBX7 is also able to bind H3K27me3 and in this

example, its ability to bind to both the histone mark and ANRIL is crucial to regulate transcription of the *INK4b/ARF/INK4a* locus in *cis* (Yap et al. 2010). A more generic RNA mediated targeting mechanism was proposed based on evidence that PRC2 binds to short RNA transcripts from the 5' end of repressed Polycomb target genes (Kanhere et al. 2010). However, PRC2 seems to bind RNA in a promiscuous manner, as it was also reported to bind transcripts from active genes. In this case, the interaction with nascent transcripts was suggested to inhibit PRC2 deposition of H3K27me3, providing a mechanism that counteracts Polycomb deposition of the histone mark instead of Polycomb recruitment (Davidovich et al. 2013; Kaneko et al. 2013; Kaneko et al. 2014). Further investigation will be necessary to unveil the importance of RNAs in Polycomb *de novo* targeting.

1.2.4.4. Reverse model of Polycomb recruitment and propagation of the marks

The hierarchical model of recruitment (Figure 5) initially proposed in *Drosophila* (Wang et al. 2004b) and then supported by several studies in mammals (Boyer et al. 2006; Cao et al. 2005) led the direction of research into mechanisms of Polycomb recruitment for almost a decade. However, over the years several observations challenged the idea that PRC2 is always the first complex to be recruited, and that PRC1 recruitment is exclusively dependent on H3K27me3.

Studies of Polycomb mark deposition on the inactive X chromosome showed that in *Eed*$^{-/-}$ mESCs in which H3K27me3 is fully depleted, RING1B and H2AK119u1 are still enriched over the inactive X chromosome, even though other components of canonical PRC1 complexes are not enriched (Schoeftner et al. 2006). Also, global depletion of H3K27me3 in this cell line did not result in global changes of H2AK119u1 levels (Schoeftner et al. 2006; Tavares et al. 2012). Similar observations for RING1B localization to gene targets in the absence of PRC2 were reported by Leeb et al. (2010).

Importantly, the identification and biochemical characterization of non-canonical PRC1 complexes containing RYBP instead of CBX proteins was crucial to understand how RING1B could be enriched on the inactive X chromosome in the absence of H3K27me3 (Tavares et al. 2012). In addition, the authors showed that PRC2 independent recruitment is not restricted to the inactive X chromosome, as H2AK119u1 is still present at target gene promoters in *Eed*$^{-/-}$ mESCs (Tavares et al. 2012).

These studies provided the first evidence that PRC1 could be recruited to chromatin independently of PRC2 and H3K27me3. Involvement of non-canonical PRC1 complexes in this pathway is consistent with the fact that they do not include the H3K27me3 reader protein CBX (Gao et al. 2012; Tavares et al. 2012). The recent report of KDM2B as an interactor of PCGF1-PRC1 and its role in the recruitment of PRC1 to unmethylated CpG islands (as described in section 1.2.4.2.) (Farcas et al. 2012; He et al. 2013; Wu et al. 2013; Blackledge et al. 2014) provided the first mechanistic insight for PRC2-

independent recruitment of PRC1. As previously mentioned, the fact that the interaction of PRC1 with KDM2B is mediated by PCGF1 also raised the possibility that different PRC1 variants could be recruited to chromatin by diverse mechanisms conferred by the specific PCGF protein present in the complex (Junco et al. 2013). This idea would fit with genome-wide ChIP-seq data that suggests that different PRC1 variants including different PCGF proteins have a distinct distribution pattern across the genome (Gao et al. 2012). However, the mechanisms that determine the recruitment and differential distribution of these complexes have not been extensively investigated.

Following up on these observations, two recent publications proposed a complete reversal of the paradigm for Polycomb recruitment. In one study it was found that tethering PRC1 components to a gene desert leads to accumulation of PRC2 and H3K27me3, and moreover PRC2 is also able to recognize and bind to H2AK119u1 deposited by non-canonical PRC1 (Blackledge et al. 2014). This data was supported by an independent study using a similar tethering system directing H2AK119u1 to pericentric heterochromatin in mESCs (Cooper et al. 2014). Supporting the idea that PRC2 recruitment by PRC1 is mediated through H2AK119u1, Kalb et al. (2014) showed that PRC2 complexes including AEBP2 and JARID2 have a high affinity for nucleosomes containing the ubiquitin mark. The authors also showed that the presence of H2AK119u1 stimulates PRC2 catalytic activity. More recently, JARID2 was identified as the component directly responsible for H2AK119u1 recognition (Cooper et al. 2016).

Together these observations led to the proposal of a reversed hierarchical model for Polycomb recruitment (reviewed by Comet and Helin 2014; Blackledge et al. 2015) suggesting that non-canonical PRC1 would be the first complexes to localize to target genes (Figure 6). PRC1 mediated H2AK119u1 can then be recognized by JARID2, leading to the recruitment of PRC2 (Figure 6) (Cooper et al. 2016). PRC2 deposition of H3K27me3 (Kalb et al. 2014) will stimulate the recruitment of canonical CBX-containing PRC1 complexes, whose function needs to be reevaluated in more detail as these complexes seem to have low H2AK119 E3 ligase activity (Blackledge et al. 2014; Taherbhoy et al. 2015; Rose et al. 2016).

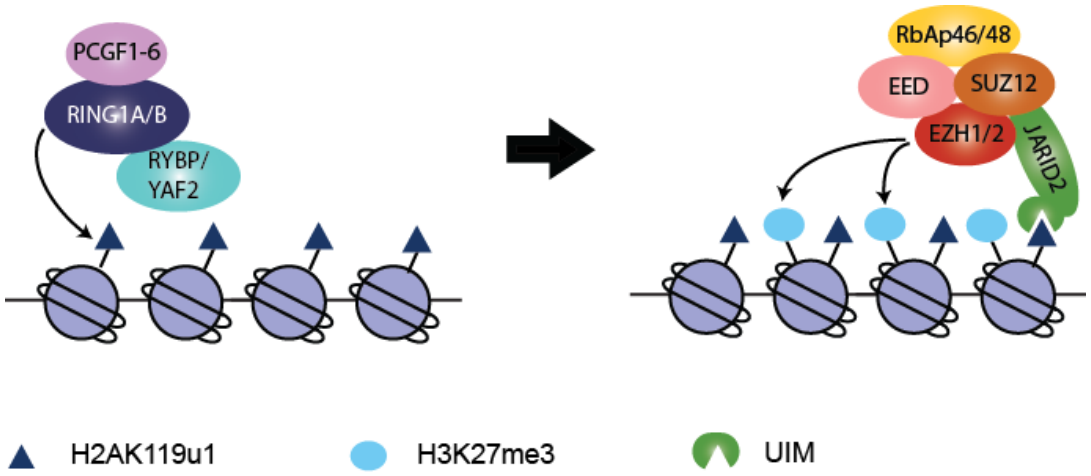


Figure 6: Reverse model of Polycomb recruitment. According to this model, non-canonical PRC1 variants that include the RYBP/YAF2 subunit are the first Polycomb proteins to be recruited to chromatin. This complex catalyzes H2AK119u1, which can be recognized by JARID2 ubiquitin-interacting motif (UIM) to recruit PRC2 complex to target genes, where it places H3K27me3.

Whilst the reverse model implies that PRC1 could be targeted first in some circumstances, there is as yet no clear evidence that this is the case in a physiological context *in vivo*. Indeed, it is plausible that reciprocal recognition of H2AK119u1 and H3K27me3 by PRC2 and PRC1, respectively, is instead relevant for the propagation of Polycomb domains and maintenance of transcriptional silencing through cell division.

A similar idea was put forward previously based on initial evidence that PRC2 recognizes its own mark H3K27me3 (Hansen et al. 2008). Recognition of H3K27me3 mediated by the core subunit EED is required for allosteric activation of PRC2 and deposition of H3K27me3 in newly synthesized histones. The deposition of this histone mark, even though not necessary to trigger silencing, is required for its maintenance (Hansen et al. 2008; Margueron et al. 2009). The non-canonical PRC1 subunit RYBP was shown to be able to bind H2AK119u1 (Arrigoni et al. 2006). Similarly to PRC2, the recognition of H2AK119u1 might contribute to PRC1 propagation *in vivo*, although this idea has never been fully tested.

The insights gained from studying mammalian systems may have relevance in *Drosophila* with some of the proposed mechanisms applying to both systems, even if mediated by different factors (Kahn et al. 2016). New studies in *Drosophila* will be required to fully understand which component mediates PRC2 recruitment by H2AK118u1 as well as to better understand the role of the newly identified Polycomb component L(3)73AH and its catalytic activity outside of PREs.

1.3. X chromosome inactivation

1.3.1. The process of X chromosome inactivation

Female mammals have two X chromosomes, in contrast to males that have an X and a Y chromosome. Accordingly, a mechanism for dosage compensation evolved to silence X-linked genes on one of the two chromosomes in mammalian females. This process is known as X chromosome inactivation (XCI). XCI was first proposed by Mary Lyon who, based on observations in mice deduced that one of the X chromosomes, randomly selected in each cell early during mouse embryogenesis, would become inactive (Lyon 1961). She further proposed that the inactive X (Xi) corresponds to a more densely compacted chromosome visible in the nucleus of female cells, previously reported by others (Barr and Bertram 1949).

In mice, there are two waves of X chromosome inactivation: imprinted and random. Imprinted X chromosome inactivation is established exclusively on the paternal X chromosome, in all cells of the zygote starting at the 4-cell stage (Patrat et al. 2009). Imprinted X inactivation persists throughout development only in extraembryonic tissues, while random X chromosome inactivation occurs in epiblast cells of the embryo after reactivation of the imprinted inactive X (Mak et al. 2004; Okamoto et al. 2004). Random X chromosome inactivation occurs shortly after implantation (E5.5 to E6.5), and once established, is stable and inherited throughout subsequent cell divisions (Mak et al. 2004; Okamoto et al. 2004).

Classical genetic studies determined that a single *cis*-acting X inactivation center controls the XCI process (reviewed in Rastan and Brown 2009). The long non-coding RNA Xist (X-inactive specific transcript) was then identified as the master regulator of this process in humans and in mice (Brown et al. 1991; Brown et al. 1992; Brockdorff et al. 1991; Brockdorff et al. 1992). In mice, Xist is a spliced, capped, polyadenylated, 17kb long non-coding RNA that is expressed from the X chromosome and coats it in *cis*, bypassing nuclear export pathways, to promote silencing (Brown et al. 1992; Clemson et al. 1996). Xist forms a cloud over the inactive X chromosome that can be visualized by RNA-FISH (Brown et al. 1992) or by live cell imaging (Ng et al. 2011) (Figure 7A). Recent work aimed at understanding how Xist RNA spreads over the chromosome by mapping its localization at high-resolution has suggested that Xist spreads through interactions with DNA regions that are in close proximity in the three-dimensional space with the *Xist* locus (Engreitz et al. 2016) binding preferentially to active, gene-rich regions first (Simon et al. 2013).

Xist expression has been shown to be necessary for the establishment of X chromosome inactivation (Penny et al. 1996; Marahrens et al. 1997) but not its maintenance (Csankovszki et al. 1999). Its expression is developmentally regulated by many factors in *cis* and in *trans* that ensure the inactivation of only one X chromosome in a tightly regulated manner during development (as reviewed in Dupont and Gribnau 2013; Galupa and Heard 2015). Xist expression and coating triggers a cascade of events that contribute to the heterochromatinization of the inactive X chromosome (as reviewed in

Chaligné and Heard 2014; Pinter 2016) (further described in section 1.3.2.). Understanding how these signals are coordinated and integrated during development to establish a transcriptionally inactive heterochromatic X chromosome in females is important not only in the context of XCI, but also because XCI is a classical model for understanding epigenetic regulation of the genome during development.

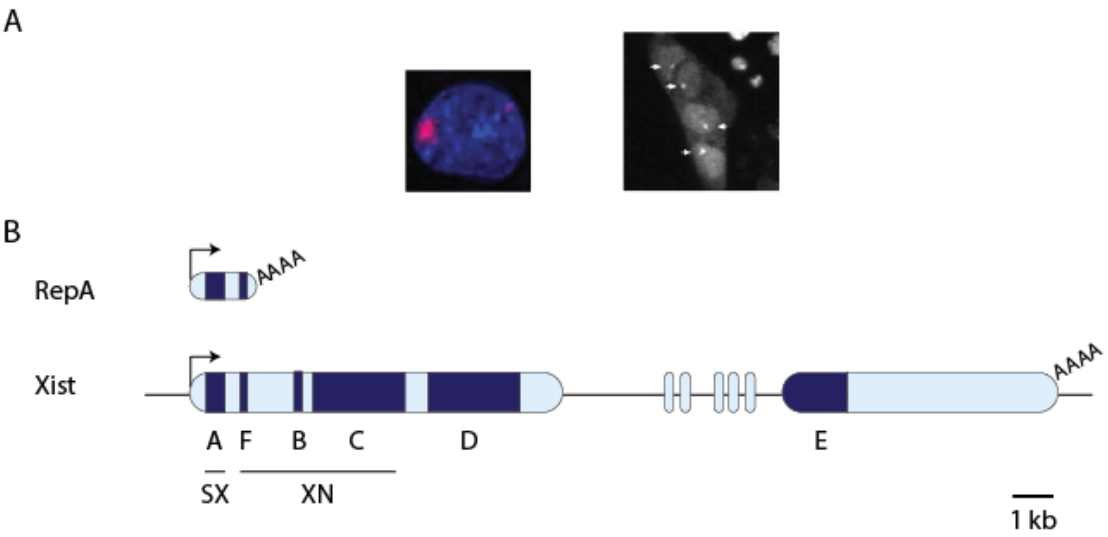


Figure 7: Xist RNA visualization and *Xist* gene locus. (A) Xist RNA, the master regulator of the XCI process, is expressed, coats the X chromosome in *cis* and can be visualized by RNA-FISH (on the left) and by live cell imaging (on the right). (Figure reproduced from Ng et al., 2011 under license CC BY-NC-SA 3.0). (B) Schematic representing the *Xist* gene locus and the repetitive regions of Xist RNA. On top, the ncRNA RepA that is transcribed from the same locus and has been suggested to be required for Polycomb recruitment to the inactive X chromosome. Tandem repetitive functional regions of Xist are labeled A to F and shown in dark blue. The fragment marked SX that includes the A-repeat region is required for chromosome silencing and the XN region was proposed to be necessary for Polycomb recruitment. (Figure adapted from Brockdorff, 2013)

1.3.2. Xist-mediated silencing and heterochromatinization of the Xi

The mechanisms by which Xist achieves X chromosome silencing, either directly or through recruitment of chromatin or RNA binding proteins, have been extensively studied. Since Xist was identified in mouse and in humans it was clear that part of its sequence was constituted by several conserved tandem repeat regions (Brown et al. 1992; Brockdorff et al. 1991). This conservation led to the idea that these regions could have some functional relevance in the process of XCI. Indeed, repeat A, located at the 5' end of Xist, was shown to be required for silencing, even though its deletion had no impact on Xist localization (Wutz et al. 2002) or recruitment of Polycomb proteins to the inactive X chromosome (Kohlmaier et al. 2004; da Rocha et al. 2014). Although the Wutz et al. (2002) study was pivotal in highlighting the potential role of different regions of Xist RNA to silence the X chromosome, the authors could not draw clear conclusions on the mechanisms involved in Xist-mediated silencing as their assay to evaluate silencing relied on assessing cell death upon silencing of the single X chromosome in male mESCs. Further studies were then essential in determining some of the mechanisms by which different Xist regions contribute to silencing for example by promoting Xist localization or the recruitment of important silencing factors (Sarma et al. 2010; Senner et al. 2011; Minks et al. 2013; da Rocha et al. 2014; McHugh et al. 2015).

To better understand the mechanisms involved in XCI, several studies characterized the process from the initiation of Xist expression and

heterochromatinization of the X chromosome to the maintenance phase in which chromosome silencing becomes independent of continued Xist expression. Monoallelic expression of Xist RNA occurs concomitantly with loss of euchromatic marks H3K4me3, H3K36me3 and acetylation of histone H3 and H4 as well as exclusion of RNA polymerase II and general transcription factors from the X chromosome that will become inactive (Heard et al. 2001; Chaumeil et al. 2002; Okamoto et al. 2004; Mak et al. 2004). Three-dimensional super-resolution microscopy analysis of the inactive chromosome reveals a collapse of the active nuclear compartment upon Xist expression (Smeets et al. 2014). The exclusion of RNA polymerase II and deacetylation of histones is now attributed, at least in part, to a complex containing SPEN (McHugh et al. 2015), a newly identified direct interactor of the Xist A-repeat region that is required for XCI (Chu et al. 2015; McHugh et al. 2015; Minajigi et al. 2015; Moindrot et al. 2015; Monfort et al. 2015). SPEN mediates recruitment and activation of the histone deacetylase HDAC3 to the Xist territory promoting the initiation of XCI (McHugh et al. 2015). Other factors identified as important for X chromosome silencing include RBM15 and WTAP (Moindrot et al. 2015) that constitute part of the RNA methylation machinery that contribute to Xist methylation (Patil et al. 2016). LBR was also shown to be an important player in the process of XCI, contributing to silencing by tethering the X chromosome to the nuclear lamina and enabling Xist spreading to active gene regions of the X chromosome (Chen et al. 2016). The timing by which these events happen during the initiation of XCI is still not fully resolved. The initiation phase of XCI is also accompanied by recruitment of Polycomb complexes, PRC1 and PRC2, and deposition of the

inactive marks H2AK119u1 and H3K27me3 on the inactive X chromosome (Plath et al. 2004; de Napoles et al. 2004; Plath et al. 2003; Silva et al. 2003). The mechanisms of recruitment and function of these complexes in the context of XCI will be discussed in the next section. H4K20me1 and H3K9me2 also accumulate on the Xi at this stage (Kohlmaier et al. 2004; Heard et al. 2001; Rougeulle et al. 2004).

As the XCI process proceeds, the Xi undergoes full compaction and becomes a distinct cytological structure in the female nucleus, the Barr body (Eils et al. 1996; Rego et al. 2008). Incorporation of the histone variant macroH2A (Costanzi and Pehrson 1998; Pasque et al. 2011), and methylation of CpG islands lock the silent status of the X chromosome in what is referred to as the maintenance phase. The protein SMCHD1 is required for DNA methylation and maintenance of the inactive status. However, the mechanism underpinning this process is unknown (Blewitt et al. 2008; Gendrel et al. 2012; Brideau et al. 2015). CDYL, a reader of H3K27me3 and H3K9me2 histone modifications, is also enriched on the inactive chromosome in the maintenance phase of XCI and was proposed to contribute to the propagation of H3K9me2 mark (Escamilla-Del-Arenal et al. 2013).

1.3.3. Polycomb contribution to Xi silencing

Several studies were conducted to address the role of each Polycomb complex in X chromosome silencing using PRC1 and PRC2 knockouts. Historically, PRC2 was the most widely studied Polycomb complex in the

context of XCI. Interest in PRC2 arose from the observation that *Eed* knockout mice are unable to effectively silence the X chromosome in the extra embryonic trophoblast giant cell lineage (Wang et al. 2001). Subsequently, Silva et al. (2003) reported that this silencing defect is not restricted to the extra embryonic lineage, as *Eed* knockout embryos, which do not accumulate H3K27me3 on the inactive X chromosome in the cells of the embryo proper, showed defective H3 deacetylation and consequent sporadic reactivation of X-linked genes. Although both studies pointed towards a role for PRC2 in Xist-mediated gene silencing, it remained unclear if its function was equally relevant in different cell lineages. Later work proposed that PRC2 is important to keep the silent X chromosome state in trophoblast stem (TS) cells only upon differentiation and attempted to reconcile the previous observations by suggesting a new role for PRC2 in blocking transcriptional reactivation signals during development rather than in establishing the transcriptionally silent state (Kalantry et al. 2006). The authors suggested that PRC2 might also have a role in the maintenance of random X chromosome inactivation upon differentiation of the cells from the embryonic lineage, although experiments to test this were not performed. In fact, other studies in mESCs using an inducible Xist transgene system in *Eed* knockout cells showed that PRC2 is not necessary to establish or maintain an inactive chromosome state triggered by Xist expression (Schoeftner et al. 2006). Consistent with this finding, expression of silencing deficient Xist harboring an A-repeat deletion (SX mutant) (Figure 7B) does not abolish PRC2 recruitment. Moreover, deletion of the downstream XN region (encompassing F-, B- and C-repeats) (Figure 7B), later shown to be required for PRC2 recruitment, had no effect on X

chromosome silencing as determined by analysis of cell death (Wutz et al. 2002; Kohlmaier et al. 2004; da Rocha et al. 2014). Due to the non-quantitative nature of this assay, the precise contribution of Polycomb to Xist-mediated silencing remains elusive.

PRC1 was shown to accumulate on the Xi and deposit H2AK119u1 through the redundant activity of RING1A and RING1B (de Napoles et al. 2004; Fang et al. 2004). Schoeftner et al. (2006) observed that H2AK119u1 could accumulate upon Xist expression even in *Eed* knockout cells, suggesting that PRC1 could play an important role in establishing and maintaining silencing, compensating for the absence of PRC2. To address this, Leeb and Wutz (2007) analyzed RING1B depleted mESCs encompassing the same inducible Xist transgene system as *Eed* knockout cells. The authors observed that in the absence of RING1B in mESCs, H2AK119u1 does not accumulate upon Xist expression (Leeb and Wutz 2007) contrary to what was observed in mouse fibroblasts in which depletion of both RING1A and RING1B is required to deplete the histone mark deposition (de Napoles et al. 2004). Analysis of gene silencing in this cell line led Leeb and Wutz to suggest that similarly to PRC2, PRC1 is also not necessary to initiate or maintain X chromosome inactivation (Leeb and Wutz 2007). However, this conclusion needs to be cautiously interpreted as RING1A, the close homolog of RING1B is present in these cells and might contribute to Xist-mediated silencing. To date, the effect of complete loss of PRC1 on XCI has not been fully investigated.

Additional studies will be necessary to fully understand how Polycomb complexes accomplish their function on the Xi and how they coordinate with other mechanisms of Xist-mediated silencing.

1.3.4. Polycomb recruitment to the Xi

Polycomb recruitment depends on Xist expression as Xist knockout in differentiated cells (Zhang et al. 2007) or switch off of an inducible Xist transgene in mESCs (Kohlmaier et al. 2004) result in loss of Polycomb accumulation on the inactive chromosome. *A priori* there are three possible mechanisms by which Xist could recruit Polycomb complexes: 1) Polycomb proteins bind directly to Xist RNA, or 2) one or several Xist direct binder proteins serve as adaptors for Polycomb, mediating its recruitment, or 3) Xist mediated silencing generates a permissive chromatin environment for Polycomb recruitment.

The first proposed mechanism received wide support in the X inactivation field. Early studies showing that EED localization to Giemsa-light bands on the inactive X chromosome at metaphase resembles Xist distribution (Duthie 1999; Mak et al. 2002) and that PRC2 recruitment is totally dependent on Xist expression (Kohlmaier et al. 2004) provided the first clues to suggest that PRC2 may accumulate on the Xi through a direct interaction with Xist RNA. Zhao et al. (2008) provided biochemical evidence to support a direct interaction. Specifically, they reported an interaction between EZH2, the catalytic subunit of PRC2, and a 1.6kb ncRNA RepA transcribed from the 5'

end of the *Xist* locus (Figure 7B). They also found that RepA is necessary to establish H3K27me3 on the Xi. In subsequent studies a different PRC2 subunit SUZ12 was also suggested to contribute to PRC2 binding to Xist RNA (Maenner et al. 2010; Kanhere et al. 2010). This proposal was further reinforced by studies that showed that H3K27me3 deposition on the Xi closely correlates with sites where Xist localizes (Engreitz et al. 2013; Simon et al. 2013; Pinter et al. 2012). More recently the chromatin remodeling protein ATRX was proposed to facilitate loading of PRC2 onto RepA RNA (Sarma et al. 2014). Finally, the PRC2 accessory protein JARID2 has been implicated in mediating the interaction of PRC2 with Xist RNA (da Rocha et al. 2014).

The aforementioned findings, together with the accepted hierarchical model for PRC1 recruitment to chromatin mediated by H3K27me3 recognition by the chromodomain in CBX proteins (Cao et al. 2002; Fischle et al. 2003), prevailed as an explanation for Polycomb domain formation on the Xi. However, this model has been challenged by several more recent observations. First, the fact that expression of Xist lacking the A-repeat region, which encompasses the RepA sequence, leads to H3K27me3 deposition (Kohlmaier et al. 2004; da Rocha et al. 2014), suggests this Xist region is non-essential for Polycomb domain formation. Moreover, deletion of B- and F- repeats on Xist RNA impairs PRC2 recruitment, again challenging the proposed role of the A-repeat (da Rocha et al. 2014). Additionally, Xist can be detected in mouse female pre implantation embryos from the 4-cell stage, but PRC2 proteins and H3K27me3 on the Xi could not be detected until the 16-cell stage arguing against a direct interaction between PRC2 and Xist

RNA (Okamoto et al. 2004). Finally, recent studies aiming to systematically characterize Xist RNA interactors using two independent quantitative MS approaches (Chu et al. 2015; McHugh et al. 2015) failed to identify any PRC2 components as direct partners of Xist.

In an effort to directly test the proposed interaction of PRC2 with Xist RNA, Cerase et al. (2014) analyzed the spatial relationship of PRC2 proteins with Xist RNA using 3D-SIM super resolution microscopy. The authors concluded that the separation between Xist and PRC2 is inconsistent with a direct interaction, although an independent study using STORM microscopy came to the opposite conclusion (Sunwoo et al. 2015).

It should be noted that the prevailing model also fails to explain how PRC1 and H2AK119u1 are enriched upon Xist expression in the absence of PRC2 (Schoeftner et al. 2006; Tavares et al. 2012). To date, the possibility of PRC1 directly interacting with Xist was never fully addressed even though some PRC1 components were identified in screens for Xist interactors (Chu et al. 2015; Minajigi et al. 2015).

The aforementioned findings that challenge the idea that PRC2 interacts directly with Xist RNA do not completely rule out that a Xist RNA binding protein acts as an adaptor to recruit Polycomb. In favor of the adaptor protein model is the fact that Polycomb domain formation only occurs if Xist is expressed in a specific time window during differentiation of mESCs, suggesting that this timeframe might coincide with the expression of a specific

protein required for Polycomb recruitment (Kohlmaier et al. 2004; Schoeftner et al. 2006). Specific candidates for this function include the Xist RNA binding proteins SPEN, RBM15 and HnRNPK, which when depleted result in reduced levels of Polycomb marks on the Xi (McHugh et al. 2015; Monfort et al. 2015; Moindrot et al. 2015; Chu et al. 2015). These observations were based on knockdown experiments, and it is therefore unclear if total absence of these proteins would result in a complete loss of Polycomb on the Xi.

The third putative mechanism for Xist-mediated Polycomb recruitment is based on the idea that Polycomb is able to read the silent chromatin environment, in this case established by Xist RNA. Transcriptional silencing has been previously shown to precede PRC2 deposition at gene promoters (Riising et al. 2014). However, deletion of the A-repeat region of Xist that impairs X chromosome silencing does not cause loss of PRC2 from the inactive domain (Kohlmaier et al. 2004; Pullirsch et al. 2010). Also, displacement of Xist in mouse embryonic fibroblasts (MEFs) leads to loss of PRC2 without immediate reactivation of the X chromosome, suggesting that the silent state, even when fully established in differentiated cells, is not enough for Polycomb deposition (Sarma et al. 2010). PRC1 accumulation on the Xi was also shown to be independent of its silencing (Schoeftner et al. 2006). These observations argue against the role of a responsive model for Polycomb recruitment to the Xi. However, further experiments will be necessary to entirely rule out this model as differential accumulation of euchromatin marks coincide with differential Polycomb recruitment in different models. Deletion of the A-repeat region of Xist induces deacetylation of H4

and loss of H3K4me2/3 despite its inability to silence (Kohlmaier et al. 2004; Pullirsch et al. 2010). On the other hand, Xist deletion in fibroblasts, in which the silencing state is irreversible and independent of Xist expression (Wutz and Jaenisch 2000), causes gain of H4 acetylation and H3K4me3 on the X chromosome (Diaz-Perez et al. 2006). These discrepancies in the presence of euchromatic marks and silencing state between different models might account for differential Polycomb recruitment.

Altogether the arguments presented here are not sufficient to support a unifying model for Xist-mediated Polycomb recruitment and further studies are necessary to resolve this question.

1.4. Aims of this study

Understanding *de novo* Polycomb recruitment during development is an important step forward in unveiling Polycomb function. Taking advantage of established strategies to mimic the initiation of X chromosome inactivation in mESCs in culture, I set out to better understand the mechanisms involved in Polycomb recruitment.

The specific aims of this study are:

- To investigate the role of PRC1 in recruiting PRC2 to the inactive chromosome
- To clarify which non-canonical PRC1 complexes are enriched on the inactive chromosome and characterize their dynamics in this context compared to the nucleoplasm
- To understand the role of RYBP/YAF2 and the ubiquitin-binding domain in the recruitment of non-canonical PRC1 complexes
- To determine the contribution of each non-canonical PRC1 complex in the initiation of *de novo* Polycomb recruitment to the Xist domain

CHAPTER 2: Materials and Methods

2.1. Plasmids and cloning

Coding sequences for PCGF1,2,3,5,6, RYBP and RING1B proteins were amplified from cDNA prepared from 129/1 wild-type cell line using primers listed in Table 1. Primers contained an overhang to allow for digestion and cloning. These sequences were inserted by LIC cloning (Aslanidis and de Jong 1990) into a modified pCAG-IRES-Puro mammalian expression plasmid between the coding sequence for a N-terminal enhanced GFP (eGFP) and a 3' IRES sequence that precedes puromycin resistant gene. This vector was kindly provided by Dr. David Brown from Prof. Rob Klose's lab. Mutagenesis of RYBP sequence into RYBP TF-AA was performed using QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent Genomics) according to manufacturer instructions and primers listed below. RYBP and RYBP TF-AA sequences were inserted by LIC cloning into a modified pCAG-IRES-Blast mammalian expression plasmid between the coding sequence for a N-terminal mCherry and a 3' IRES sequence that precedes the blasticidin resistance gene. pCAG-mCherry-IRES-Blast was modified from a pCAG-mCherry-IRES-Neo vector provided from Dr. David Brown, using a traditional PCR and enzyme digestion cloning strategy. Plasmids expressing FLAG tagged RING1B WT and RING1B I53A were obtained from the Koseki lab (Endoh et al. 2012). Puromycin resistance cassette was replaced by Blasticidin resistance cassette using a traditional PCR and enzyme digestion cloning strategy in the RING1B plasmids.

The plasmid used to target rtTA-H (Reverse Tetracycline-Controlled Transactivator - Hygromycin) into the Rosa26 locus, as well as the plasmid including Tet-ON inducible full-length Xist transgene (pTREXist) for random integration were available in the lab and have been previously used to generate other cell lines in the lab (Tang et al. 2010). The plasmid including a Tet-ON inducible full-length Xist transgene with a tandem array of 18 repeats of Bgl stem loop motif after exon 7 (pTREXist BglSLEx7) was also available in the lab (Moindrot et al. 2015). A slightly modified version of the plasmid containing the BglG-mCherry sequence (Moindrot et al. 2015) was generated by PCR and traditional restriction enzyme cloning of the BglG-mCherry sequence into a plasmid without loxP sites, for insertion of this plasmid into conditional knockout cell lines. After each cloning step, each plasmid sequence was confirmed by Sanger sequencing.

Table 1: Primers used to amplify inserts for LIC cloning and for mutagenesis.

LIC Cloning primer	Sequence (5'-3')
PCGF1_LIC_F	TACTTCCAATCCatgaggcttcggaaccagc
PCGF1_LIC_R	TATCCACCTTTACTGctacctcctcttctctttcac
PCGF2_LIC_F	TACTTCCAATCCatgcatcggaccacacgga
PCGF2_LIC_R	TATCCACCTTTACTGtcaagggggcaaggagc
PCGF3_LIC_F	TACTTCCAATCCatgttgaccaggaagattaaac
PCGF3_LIC_R	TATCCACCTTTACTGtcaagggggcaaggagc
PCGF5_LIC_F	TACTTCCAATCCatggctacccaaaggaaacac
PCGF5_LIC_R	TATCCACCTTTACTGctacccaaaatcaattctcggt
PCGF6_LIC_F	TACTTCCAATCCatggacgaggctgagacgg
PCGF6_LIC_R	TATCCACCTTTACTGtcaagttattttcaaaggagaaac
RYBP_LIC_F	TACTTCCAATCCaccatgaccatgggcgac
RYBP_LIC_R	TATCCACCTTTACTtcagaaagattcatcattc
Mutagenesis primer	**Sequence (5'-3'; mutations in capital and bold)**
RYBP_TF-AA_F	gattgtagcgtctgc**G**cc**GC**taggaacagcgccgaa
RYBP_TF-AA_R	ttcggcgctgttccta**GC**gg**C**gcagacgctacaatc

2.2. Mouse embryonic stem cell lines

2.2.1. Conditional mESC lines production and derivation

M1 ES cells were derived from the cross between C57BL/6JJcl and 129^{+Ter}/SvJcl and used to produce a conditional knockout allele for *Pcgf1* and *Pcgf3* in Haruhiko Koseki's lab. The arms of homology in targeting vectors span 2936 bp for the 5' arm and 3260 bp for the 3' arm of *Pcgf1*, 3051 bp for the 5' arm and 3608 bp for the 3' arm of *Pcgf3*, and 3018 bp for the 5' arm and 3506 bp for the 3' arm of *Pcgf5*, respectively. For *Pcgf1*, a Neomycin resistant gene flanked with FRT and loxP sequences was inserted at intron 7 and a loxP sequence at intron 1. For *Pcgf3*, a Neomycin resistance cassette flanked with FRT and loxP sequences was inserted at intron 4 and a loxP sequence at intron 3. To establish $Pcfg1^{fl/fl}$ and $Pcgf3^{fl/fl}$ mouse lines, targeted ES cells were injected into mouse 8-cell stage embryos to create chimeric mouse lines, then those chimeric mouse lines were mated with B6;SJL-Tg(ACTFLPe)9205Dym/J mice (The Jackson Laboratory) to remove the Neomycin resistance cassette. ES cell lines harbouring the *Rosa26:CreERT2* allele were derived from matings between $Pcgf1^{fl/fl}$ and $Pcgf3^{fl/fl}$ lines *and Rosa26:CreERT2* line that was purchased from Artemis Pharmaceuticals (Seibler et al. 2003).

2.2.2. Mouse ESC lines and growing conditions

Cell lines used in this study were 3E-H (derivative of 3E (Tang et al. 2010) with puromycin resistance cassette exchanged for a hygromycin resistance cassette) (Moindrot et al. 2015), *Pcgf1*$^{fl/fl}$, *Pcgf3*$^{fl/fl}$ (derived for this study) and *Ring1A*$^{-/-}$ *Ring1B*$^{fl/fl}$ (Endoh et al. 2008) mESCs.

All ES cells were grown and passaged in feeder-dependent conditions on gelatinized plates with ESC medium at 37°C in a 5% CO_2 incubator. Mitomycin-inactivated SNLP (STO mouse fibroblasts expressing neomycin, puromycin resistance and Lif genes) were used as feeders. ESC medium consists of D-MEM without sodium pyruvate (Life Technologies, Invitrogen) supplemented with 10% Fetal Calf Serum (Seralab), 1x non-essential amino acids (Life Technologies, Invitrogen), 2mM L-glutamine (Life Technologies, Invitrogen), 50µM β-mercaptoethanol (Life Technologies, Invitrogen), 1x Penicillin-Streptomycin (Life Technologies, Invitrogen) and 1mL of Leukemia Inhibitory Factor-conditioned medium made in-house.

2.2.3. Doxycyline and tamoxifen treatments

Xist expression driven by the Tet-ON promoter was induced by adding doxycycline (1 to 1.5µg/mL) to the culture medium for 6 to 24 hours depending on the experiment.

Conditional deletion of *Ring1B*, *Pcgf1* and *Pcgf3* was carried out by adding 800nM 4-hydroxytamoxifen to the culture medium for 48 hours. To generate *Pcgf3*$^{-/-}$ *Pcgf5*$^{-/-}$ mESCs (constitutive knockout), cells were treated with 800nM 4-hydroxytamoxifen for 96 hours. Efficient deletion of the gene was confirmed by PCR (using primers in Table 2) unless a good antibody was available to confirm protein depletion by Western blot.

Table 2: Primers used for genotype of conditional knockout mESCs after tamoxifen treatment.

Genotype primer	Sequence (5'-3')
PCGF1_F1	GCTGGACTTGGACGCTTGCGC
PCGF1_F2	AGTAAGCCCATCATGTTACGG
PCGF1_R1	GTACCTAAGAGGGAGATGGGC
PCGF3_F1	ATAAGATGAGATGGGATGGGC
PCGF3_F2	GTTGGTCTGTCAGTCAAGCCG
PCGF3_R1	ACGCCTCCAGGTGATCCATAC

2.2.4. Generation of stable cell lines expressing transgenes

All ES cells expressing transgenes in this thesis are stable expressing cell lines. To generate stable ES cell lines, cells were plated in 6-well plates on feeders at a density of 1-2x10^6 cells/well, one day before transfection. Medium was replaced with antibiotics-free medium before transfection. Cells were transfected using Lipofectamine 2000 (Life Technologies, Invitrogen), according to the manufacturer's instructions. After 24 hours, transfected cells were passaged to 90mm gelatinized Petri dishes with feeders. Puromycin (1.5 to 2.5µg/mL) or blasticidin (10µg/mL) selection was applied 48 hours after lipofection. Cells were grown for 10-12 days under selection with medium

being changed every day. Individual ES colonies were picked and expanded for screening.

2.2.5. Targeting of rtTA-H cassette into *Rosa26* locus

A modified version of pR26/P'nlsrtTA construct where the puromycin resistance cassette was replaced by a hygromycin resistance cassette (Moindrot et al. 2015) was electroporated into *Pcgf1*$^{fl/fl}$, *Pcgf3*$^{fl/fl}$ and *Ring1A*$^{-/-}$ *Ring1B*$^{fl/fl}$ to target rtTA-H into the constitutively active *Rosa26* locus as previously described (Tang et al. 2010). Briefly, a $2x10^7$ cells pellet was washed twice with PBS and then resuspended in 0.8mL of PBS. Ten to 15 µg of the linearized targeting plasmid were added to the cell suspension and mixed by pipetting. The suspension was transferred to an electroporation cuvette placed on ice and incubated for 5 min. Cells were electroporated using a Gene Pulser Xcell machine (BioRad) in a 0.4cm electrode gap cuvette using 200V/960µF. After electroporation cells were kept on ice for 5 min to recover and then transferred to a Petri dish with feeders. Cells were fed the next day and left to grow another 24 hours before adding hygromycin to the medium. Cells were kept under selection for 10 to 12 days, being fed everyday. Individual colonies were picked to a 96 well plate and expanded for freezing and screening. Efficient targeting was verified by Southern blot analysis using two different unique sequence probes.

2.2.6. Knockout of *Rybp*, *Yaf2*, *Pcgf5* and *Pcgf6* using CRISPR/Cas9-mediated mutagenesis

sgRNAs targeting *Rybp* and *Yaf2* were cloned into plasmid pX459 (Ran et al. 2013b). Target sequences are shown in Table 3. 3E-H mouse ES cells were plated in feeder-dependent conditions on a 6-well plate one day before transfection. On the day of transfection, medium was replaced with antibiotic-free medium. Cells in each well were transfected with 2µg of each plasmid using Lipofectamine 2000 (Life Technologies, Invitrogen), according to manufacturer's instructions. Cells were plated at different densities in three Petri dishes with feeders, 24 hours after transfection. The next day, medium was replaced by ES medium with puromycin (2µg/mL) for 48 hours, after which cells were grown in ES medium until ES colonies were ready to be picked. Screening of knockout clones was achieved by Western blot or PCR from genomic DNA when the first was not possible. The clone selected for experiments was characterized using PCR followed by Sanger sequencing of the PCR product (using primers in Table 4).

sgRNAs targeting *Pcgf5* cloned into plasmid pX462 (Ran et al. 2013a) were provided by Dr. Andrew Bassett (Genome Engineering Oxford). sgRNA targeting *Pcgf6* was cloned into pX459 (Ran et al. 2013b). Target sequences are shown in Table 3. *Pcgf5* and *Pcgf6* were mutagenized in *Pcgf1*$^{fl/fl}$ rtTA-H and *Pcgf3*$^{fl/fl}$ rtTA-H using a similar strategy as described for *Rybp*/*Yaf2*. Colonies were screened by PCR and the sequence of relevant clones was characterized by Sanger sequencing (using primers in Table 4).

Table 3: CRISPR/Cas9 sgRNAs target sequences.

sgRNA target sequences	Sequence (5'-3'; PAM sequence in bold)
RYBP	CGCCCGGCAGGTACCTGGTC**GGG**
YAF2	GCCGAAGCCCGCCTCGGATG**AGG**
PCGF5_1	ATAGCCTTTACAGATGTAGC**AGG**
PCGF5_2	GATCAAGCCCACGACAGTGA**CGG**
PCGF6	AGAGGCGCGCTTGTTTTCCG**TGG**

Table 4: Primers used for genotype and sequencing of CRISPR/Cas9 generated mutant mESCs.

Genotype primer	Sequence (5'-3')
RYBP_F	CCGGTCGTCGTCCTCTTATT
RYBP_R	CCCCGGGTCTCCATCTTAG
YAF2_F	CGTGTCCGGTTGTGTCCT
YAF2_R	ACATCGCACATCATGCACTT
PCGF5_F1	TGTTTACAGAGAGGAAGCGCC
PCGF5_R1	TGGCCTTGGTACACATATAGC
PCGF5_gDNA_F1	GAGCAGCTTTGCATCTCCTT
PCGF5_gDNA_R1	AGAAAGCGGAGGACACAGAA
PCGF5_gDNA_F2	GCTCCTCTGCTTCAACTGCT
PCGF5_gDNA_R2	CCAGACCAAAACAGAGCACA
PCGF6_gDNA_F1	CTTGCCTCGCTCTCAACTCG
PCGF6_gDNA_R1	CTCCTCGTCCTCGGAGTCTG
PCGF6_gDNA_F2	TCTCCTGAACAGCACACTGC
PCGF6_gDNA_R2	GCAGGGTCGTCGATACCTC

2.2.7. Generation of *Xist* inducible clones

Pcgf1$^{fl/fl}$, *Pcgf3*$^{fl/fl}$ and *Pcgf3*$^{fl/fl}$ *Pcgf5*$^{-/-}$ ES cells with rtTA-H were co-transfected with pTREXist BglSLEx7 and the plasmid containing the fusion between the N-terminal RNA recognition motif of BglG and mCherry protein together with a neomycin resistance cassette as described (Moindrot et al. 2015). Stable cell lines were established using selection with 400µg/mL G418

and screened by RNA-FISH for the presence of a single Xist cloud after doxycycline induction. Immunofluorescence against mCherry was used to identify clones with BglG-mCherry-Xist clouds after 24 hours of doxycycline treatment.

To establish an inducible Xist transgenic cell line, *Ring1A*$^{-/-}$ *Ring1B*$^{fl/fl}$ carrying a rtTA-H targeted into *Rosa26* locus was cotransfected with pTREXist vector (Tang et al. 2010) and a PGKpuro cassette-containing plasmid. Cells were selected using 1.5µg/mL puromycin until individual ES colonies were picked. Colonies were screened by RNA-FISH for the presence of a single Xist cloud after doxycycline induction for 24 hours.

2.3. Colony screening by Southern blot

2.3.1. Preparation of DNA from 96 well plate for Southern blot

DNA was prepared from a 96 well plate were colonies were grown without feeders. Cells were washed with PBS and then lysed using Lysis Solution (10mM NaCl, 10mM Tris pH 7.5, 10mM EDTA pH8.0, 0.5% Sodium lauroyl sarcosinate) including freshly added Proteinase K (final concentration 1mg/mL). Cells were incubated in this solution overnight in a rotating water bath at 55º C. The next day, 2 volumes of cold 100% ethanol were added to precipitated DNA. For this purpose the lysate was left at -20º C for at least 30 min. The supernatant was carefully removed and DNA was carefully washed three times with cold 70% ethanol solution. DNA was air-dried for 10 min and

an enzymatic digestion mixture without enzyme was added to dissolve DNA. The plate was placed on a rotating platform at 37º C for 2 hours. After 2 hours the same volume of the enzymatic digestion mixture was added including the appropriate enzyme (BamHI or ApaI).

2.3.2. Southern blot

The enzymatically digested DNA samples were run overnight on a low percentage agarose gel. The samples were then prepared for Southern blot by transferring the gel into a depurination solution (0.125M HCl) for 15 min and subsequently to a denaturation solution (0.5M NaOH, 1.5M NaCl) for 45 min and a neutralization buffer (0.5M Tris, 1.5M NaCl at a final pH 8) for 30 min. DNA was blotted overnight by capillarity into a Immobilon NY+ membrane (Millipore). The next day the membrane was briefly washed with 2x SSC, air-dried and DNA was crosslinked to the membrane using a SpectroLinker XL-1500 UV Crosslinker (Spectronics Corporation) optimal crosslinking function. The membrane was incubated with pre-warmed 65º C Dextran hybridization buffer. Probe was radioactively labeled with ATP-alpha-32P using Prime-It II Random Primer Labeling Kit (Stratagene) following the manufacturer's instructions. The labeled probe was purified using Illustra MicroSpin G50 Columns (GE Healthcare) and added to the hybridization mixture in which the membrane was placed. The hybridization was performed overnight at 65º C. After hybridization the membrane was washed twice in a low stringency washing solution (2xSSC, 1% SDS) before exposure to a

phosphorimaging screen. The screen was scanned after overnight exposure using a Fujifilm FLA-7000 image reader.

2.4. Imaging techniques

2.4.1. Live cell imaging

Cells were plated in gelatinized µ-Slide 8 well slides or 35mm µ-Dishes (Ibidi), one day before Xist induction by doxycycline. Before imaging, cells were carefully washed with PBS and phenol red-free medium (Lonza) containing a final concentration of 2mM sodium pyruvate (Life Technologies, Invitrogen) and 12.5mM HEPES (Life Technologies, Invitrogen) was added. Live cell imaging was performed on an IX81 Olympus microscope connected to a Spinning Disk Confocal system (UltraView VoX PerkinElmer) equipped with an EMCCD camera (ImagEM, Hamamatsu Photonics) in a 25^{o}C heated chamber. All the pictures and z-stack movies were captured using a PlanApo 60x/ 1.4 N.A. oil-immersion objective heated to 37^{o}C, using Volocity software. Images were captured with a 100ms exposure at 10% laser power and a detector sensitivity of 201. Z-stacks were captured every 0.8µm. All captured images were analyzed and processed using Fiji software.

2.4.2. FRAP and FRAP analysis

Fluorescence Recovery After Photobleaching (FRAP) was performed in the same setup described for live cell imaging. Photobleaching of a 1.25µm radius

spot was performed by one bleach pulse of 50ms using 488nm laser at 100% intensity. Bleaching occurred after 100 pre-bleach images. After bleaching, 500 images were captured at a maximum speed rate (approximately 8fps). Each image was captured with a 100ms exposure using 15% laser power and a detector sensitivity of 201. Experiment was repeated on three different days for the same stable cell line.

Fluorescence recovery was calculated normalizing for fluorescence before bleaching, background intensity and fluorescence loss due to photobleaching (Mueller et al. 2012; Sprague et al. 2004). To determine the half recovery time $t_{1/2}$ and the percentage of immobile fraction (1-*v*) under the described conditions, FRAP data for each cell was fit to the following biexponential equation:

$$FRAP(t) = v - Ae^{-t/c} - Be^{-t/d}$$

where *t* is time, *v* is the plateau of the FRAP curve and *A*,*B*,*c* and *d* are fitted parameters (Mueller et al. 2012).

2.4.3. Immunofluorescence

Cells were plated on slides at least one day before the experiment. On the day of the experiment, cells on slides were washed with PBS and then fixed with 2% formaldehyde for 15 min followed by 5 min of permeabilization using 0.4% Triton X-100. Cells were briefly washed with PBS before blocking with a 0.2% PBS-based solution of fish gelatin for three periods of 5 min. Primary antibody dilutions were prepared in fish gelatin solution with 5% normal goat

or normal donkey serum depending on the secondary antibodies used. Primary antibody dilutions are listed below. Cells on slides were incubated with primary antibody (summarized in Table 5) for 1-2 hours in a humid chamber at room temperature. Slides were washed three times in fish gelatin solution. Secondary antibodies were diluted 1:400 in fish gelatin solution and incubated with cells on slides for 1h30min in a humid chamber at room temperature. After incubation, slides were washed twice with fish gelatin and one time with PBS before mounting using Vectashield mounting medium with 4′,6-diamidino-2-phenylindole (DAPI). Excess of mounting medium was removed and the coverslips were sealed to slides using nail varnish.

For the protocol using permeabilization before fixation, cells were plated in 6-well plates. On the day of the experiment, cells were washed with PBS, trypsinized and counted. Tubes containing $5x10^5$ cells were prepared and these cells were permeabilized by resuspension in 1mL of 0.4% Triton X-100 for 5 min. After 5 min 1mL of 4% formaldehyde solution was added to fix cells (in a final solution of 2% formaldehyde) for 15 min. Cells were then cytospun onto slides. Slides were washed with PBS. Blocking, staining and mounting was performed as described above.

Cells were analyzed on an inverted fluorescence Axio Observer Z.1 microscope using a PlanApo 63x/1.4 N.A. oil-immersion objective. Images were acquired using AxioVision software. Optimal exposure time for each field and channel was automatically determined by the software. A minimum of 100

cells was scored per slide and each experiment was scored in a minimum of 3 replicates.

2.4.4. RNA-FISH

For the experiment in *Ring1A*$^{-/-}$ *Ring1B*$^{fl/fl}$ ES cells, cells were plated on an extra slide in parallel with cells plated for immunofluorescence. After washing twice with PBS, cells were fixed for 10 min with 2.6% formaldehyde followed by permeabilization with 0.42% Triton X-100 for 5 min at 4°C. After fixation, cells on slides were washed twice with 70% ethanol and dehydrated using a series of concentrated ethanol solutions (80%, 95% and 100%). Slides were briefly air dried before incubation with the probe. RNA probe was generated from an 18 kb fragment spanning the whole Xist transcript using a nick translation kit (Abbott Molecular) as previously described (Moindrot et al. 2015). Labelled probe (1.5 µl) was coprecipitated with 10 µg salmon sperm DNA, 1/10 volume 3 M sodium acetate (pH 5.2) and 3 volumes ethanol. After washing in 75% ethanol, the pellet was dried, resuspended in 6 µl formamide and denatured at 75°C for 7 min before flash cooling on ice. Probe was diluted in 6 µl 2x hybridisation buffer (5x SSC, 12.5% dextran sulfate, 2.5 mg/ml BSA (NEB)), added to the slide and incubated overnight at 37°C in a humid chamber. After incubation, slides were washed three times with a solution of 2xSSC/50%formamide followed by three washes with 2xSSC in a water bath at 42°C. Slides were mounted and sealed as for immunofluorescence.

2.5. General Biochemistry

2.5.1. Preparation of whole cell lysates and nuclear extracts

Total cell lysates were prepared from ES clones growing without feeders. Cells were washed with PBS and then lysed using Laemmli Buffer (0.2M Tris pH 6.8, 2% SDS, 20% glycerol, 10% β-mercaptoethanol, 0.2% Bromophenol Blue). Cell lysates were snap frozen and kept at -20^{o}C until further use.

Nuclear extracts for immunoprecipitation and Western blot analyses were prepared according to a modified version of the method described previously (Dignam et al. 1983). Briefly, cells were lysed in buffer A (10mM HEPES pH 7.9, 1.5mM MgCl2, 10mM KCl), incubated on ice for 10 min and recovered by centrifugation. Cells were then resuspended in 3 volumes of buffer A and nuclei were released using a Dounce homogenizer. Recovered nuclei were resuspended in buffer C (5mM HEPES pH 7.9, 26% glycerol, 1.5mM MgCl2, 0.2mM EDTA, 250mM NaCl and complete protease inhibitors (Roche)). Salt concentration was increased to 400mM NaCl and extraction was performed on ice for 1 hour. Nuclei were pelleted at 13,000rpm for 20 min at 4^{o}C and the supernatant was collected as nuclear extract. Samples concentration was measured using Bradford assay (Bio-Rad) according to manufacturers' instructions.

2.5.2. Immunoprecipitation

One milligram of nuclear extract (400mM NaCl) was diluted into buffer C without NaCl to adjust to a final concentration of 150mM NaCl in 1100 µl. All samples were treated with benzonase for 30 min rotating at 4 ^{o}C. After centrifugation to remove any precipitated protein, 100µl of the supernatant were collected as input. The remaining supernatant was transferred to a new protein lobind tube and incubated with 1µl of anti-GFP antibody (ab290, Abcam) rotating for 3 hours at 4 ^{o}C. Protein A magnetic dynabeads (Life Technologies) were pre-incubated with 1.5 mg/ml BSA for 2 to 3 hours. 40µl of beads were added to each IP and incubated rotating for an extra hour at 4^{o}C. IP samples were washed five times in Wash Buffer (300mM NaCl, 20mM HEPES pH 7.9, 0.1%Tween-20, 5% glycerol) and then resuspended in 2x Laemmli Buffer and boiled for 10 min. Samples were stored at -20 ^{o}C until further analysis by Western blot.

2.5.3. Preparation of histone extracts

Histone extracts were prepared from ES cells grown without feeders. The cells were spun down and washed once with PBS. The cell pellet was then resuspended in cold PBS containing complete protease inhibitors (Roche) and incubated on ice for 10 min before a 4 min centrifugation at 235g at 4ºC. The pellet was then resuspended in 1mL of cold 0.2M H_2SO_4, incubated for 30 min and spun at 20,200g for 2 min at 4ºC. The supernatant was transferred to a new 1.5mL tube and trichloroacetic acid (TCA) was added to

a final concentration of 25% to precipitate the histones during a 30 min incubation on ice. Histone extracts were pelleted at 20,200g for 10 min at 4ºC. The pellet was washed twice with cold acetone for 10 min on ice. Histones were pelleted at 20,200g for 10 min at 4ºC, briefly air-dried and dissolved in 100mM Tris pH7.6. Samples were kept at -20ºC until further use. Before western blot analysis, samples were prepared by adding 2x Laemmli Buffer and boiled for 10 min before loading.

2.5.4. Western blot

Total cell lysate, nuclear extract proteins, or histone extracts were resolved on an SDS-PAGE separating gel with a stacking gel. After gel separation, proteins were transferred to an Amersham Hybond-P PVDF or nitrocellulose membrane using a semi-dry system (12V for 45 minutes or 17V for 1hour in the case of histones). Membranes were washed with PBS and then blocked for at least an hour in Blocking Buffer (100mM Tris pH7.5, 0.9% NaCl, 0.1% Tween, 5% non-fat milk) on a shaker at room temperature. Primary antibodies (summarized in Table 5) were prepared at the appropriate dilution in Blocking Buffer and membranes were incubated with antibody solutions overnight shaking at 4^{o}C. Membranes were washed 3 times 10 min each in Blocking Buffer and then incubated with secondary antibodies for an hour at room temperature. Secondary antibodies (either fluorescent-dye -1:15000- or horseradish peroxidase-conjugated antibodies -1:2000) were diluted in Blocking Buffer. After incubation, membranes were washed with Blocking Buffer twice, TBST (100mM Tris pH7.5, 0.9% NaCl, 0.1% Tween) and PBS,

15 min each. The result was visualized either directly on ODYSSEY Fc Dual Mode Imaging System (LI-COR Biosciences Ltd, UK) for fluorescent-dye conjugated secondary antibodies or using enhanced chemiluminescence (ECL) solution for horseradish-peroxidase secondary antibodies.

Table 5: Antibodies used in this thesis.

Antibody	Raised in	Source	Reference/ Cat. No.	Dilution for IF	Dilution for WB
H2AK119u1	Rabbit	Cell Signaling	mAb #8240	1:500-1:1000	1:1000
H3K27me3	Mouse	Active Motif	61017	1:500-1:1000	1:1000
GFP	Chicken	Abcam	ab13970	1:500	-
GFP	Rabbit	Abcam	Ab290	-	1:1000
mCherry	Rabbit	Source BioScience	ABE3523	1:500	-
mCherry	Sheep	Gift from F. Barr	-	1:500	-
RING1B	Mouse	Gift from H. Koseki	-	-	1:3000
RYBP	Rabbit	Millipore	AB3637	-	1:500
PCGF1	Rabbit	Gift from R. Klose	-	-	1:500
CBX7	Rabbit	Abcam	ab21873	-	1:1000
EZH2	Rabbit	Cell Signaling	mAb #5246	1:500	1:1000
SUZ12	Rabbit	Cell Signaling	mAb #3737	-	1:1000
JARID2	Rabbit	Novus Biologicals	NB100-2214	1:500	-
Histone H3 (loading control)	Rabbit	Abcam	ab1791	-	1:1000
Histone H4 (loading control)	Rabbit	Abcam	ab7311	-	1:1000
Lamin B (loading control)	Goat	Santa Cruz	sc6216 and sc6217	-	1:1000
TBP (loading control)	Mouse	Abcam	ab51841	-	1:1000
Tubulin (loading control)	Rabbit	Cell Signaling	#2144	-	1:1000

CHAPTER 3: Role of PRC1 and H2AK119u1 in PRC2 recruitment by Xist RNA

3.1. Introduction

Xist expression and spreading over the X chromosome initiates transcriptional inactivation in female mammalian cells, leading to early loss of histone marks associated with euchromatin and accumulation of inactive marks such as H3K27me3 and H2AK119u1, deposited by Polycomb proteins. As previously discussed, there are several mechanisms by which Polycomb proteins can be recruited to the Xi. The hierarchical model of recruitment, according to which PRC2 is directly recruited by direct interaction with Xist RNA preceding PRC1 recruitment, was widely supported. However, this hierarchical model of recruitment has been challenged on several levels, including the finding that in PRC2 depleted cells (*Eed*$^{-/-}$ mESCs) RING1B, the catalytic subunit of PRC1 (Schoeftner et al. 2006), as well as RYBP, a recently identified exclusive component of non-canonical PRC1 complexes were seen enriched on the inactive chromosome (Tavares et al. 2012). These data contribute towards the idea that PRC1 can be recruited independently of PRC2 and H3K27me3 recognition.

Recent studies on the mechanism of Polycomb domain formation came to the unexpected conclusion that these non-canonical PRC1 complexes, besides being recruited independently of PRC2/H3K27me3, are able to initiate Polycomb domain formation by promoting H2AK119u1-dependent recruitment

of PRC2 (Blackledge et al. 2014; Cooper et al. 2014). The finding that JARID2 has a UIM mediating binding to H2AK119u1 leads to the idea that recognition of this histone modification might underlie JARID2 mediated recruitment of PRC2 to the Xi (Cooper et al. 2016), contrary to the direct interaction of this substoichiometric PRC2 component with Xist as recently suggested (da Rocha et al. 2014). This evidence points to the possibility that PRC1 functions as the initiator of Polycomb recruitment. In this chapter, I set up a system to address the relevance of PRC1 and H2AK119u1 in *de novo* Polycomb domain formation in a physiological model, Xist-mediated chromosome inactivation in mESCs.

3.2. Results

3.2.1. Generation of an inducible Xist system in *Ring1A*$^{-/-}$ *Ring1B*$^{fl/fl}$ to assess *de novo* Polycomb recruitment

Considering the recent evidence that PRC2 can be recruited to chromatin by recognition of PRC1 mediated H2AK119u1, I set out to investigate if this mechanism is relevant for PRC2 recruitment in the context of Xist RNA expression. To address this question I used a mouse embryonic stem cell line (*Ring1A*$^{-/-}$ *Ring1B*$^{fl/fl}$) in which *Ring1A*, one of the catalytic subunits of PRC1 complexes, is constitutively deleted, while the functional homolog *Ring1B* can be knocked-out through the use of a tamoxifen inducible Cre recombinase system (Endoh et al. 2008). Addition of tamoxifen to the cell culture medium for 48h-72h results in a total depletion of RING1B at the protein level, and

extensive depletion of global H2AK119u1 levels as assessed by Western blot (Endoh et al. 2008).

I set out to check how PRC2 recruitment by Xist RNA is affected in the absence of PRC1 and H2AK119u1, by engineering a doxycycline inducible Xist transgene in this cell line. The doxycycline Tet-on system has two components: a transactivator rtTA (reverse tetracycline-controlled transactivator) and a full-length Xist cDNA preceded by a tetracycline-dependent promoter. This inducible system has been used previously to mimic the initiation of X chromosome inactivation in mESCs (Lee and Jaenisch 1997; Wutz and Jaenisch 2000). The transactivator rtTA was targeted to the constitutively expressed *Rosa26* locus in *Ring1A*$^{-/-}$ *Ring1B*$^{fl/fl}$ mESCs, as confirmed by Southern blot (Figure 8A). Full-length Xist cDNA under the tetracycline-dependent promoter was randomly integrated in the genome. Addition of doxycycline (+dox) to the culture medium leads to the activation of Xist expression under the tetracycline-dependent promoter. A clone presenting single Xist clouds in the nucleus of the majority of cells as visualized by RNA-FISH (Figure 8B) was selected for further analysis.

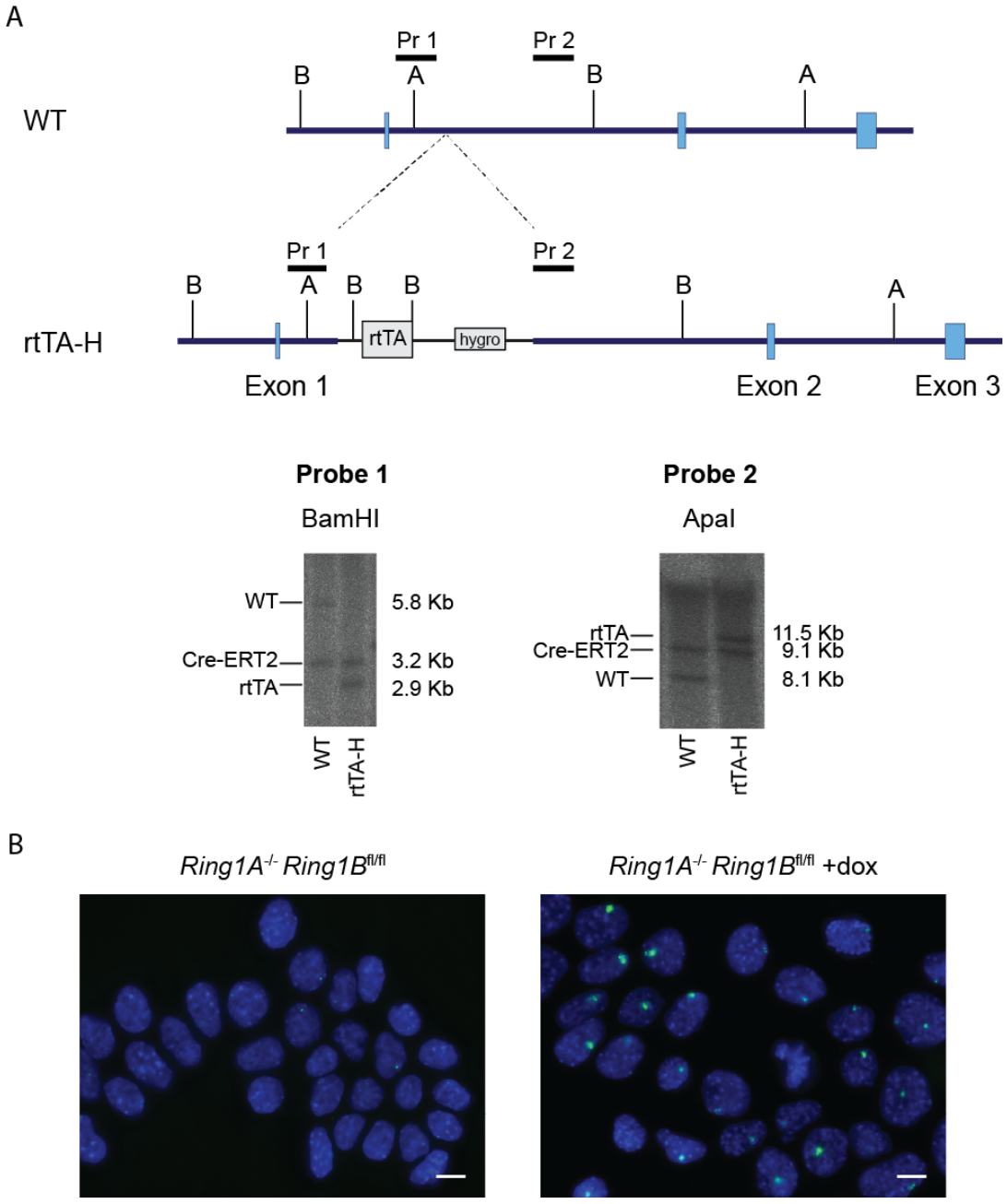


Figure 8: Establishment of an inducible Xist system in *Ring1A*$^{-/-}$ *Ring1B*$^{fl/fl}$ mESCs. (A) Schematic representation of the strategy used to target rtTA-H into the *Rosa26* locus and confirmation by Southern blot. Pr1 and Pr2 represent two different probes used for Southern blot using different restriction enzyme digestions. B and A represent BamHI and ApaI digestion sites, respectively. Bottom panels show Southern blots using Pr1 and Pr2 to confirm targeting of rtTA-H. (B) RNA-FISH for Xist RNA showing formation of a single cloud in the nucleus of most cells after 24 hours of doxycycline treatment. Scale bar is 10μm.

3.2.2. PRC1 activity is required for Xist-mediated recruitment of PRC2

Deletion of *Ring1B* was initiated 24 hours before Xist induction as shown in Figure 9A. This timing was important as it allows assessment of the hierarchy of events in *de novo* recruitment mediated by Xist RNA while excluding the potential interference of positive feedback loops involved in the recruitment of Polycomb proteins (described in section 1.2.4.4.). Efficient depletion of RING1B and H2AK119u1 reduction were confirmed by Western blot (Figure 9B). Even though no global changes of H3K27me3 levels or EZH2 expression were detected upon *Ring1B* deletion (Figure 9B), accumulation of this mark and recruitment of PRC2 to the inactive chromosome domain as evaluated by immunofluorescence of EZH2 (catalytic subunit of PRC2) and JARID2, were largely abolished. Immuno-staining in these cells is more similar to controls in which Xist RNA expression was not induced (*Ring1A*$^{-/-}$ *Ring1B*$^{fl/fl}$ and *Ring1A*$^{-/-}$ *Ring1B*$^{fl/fl}$ +OHT in Figure 10A). In the approximately 10% of cells in which H3K27me3 domains were still detectable (Figure 12), the size of the domain is visibly smaller and the intensity of the domain is more similar to the overall nuclear staining than in control cells that express Xist in the presence of RING1B (*Ring1A*$^{-/-}$ *Ring1B*$^{fl/fl}$ +dox in Figure 10B).

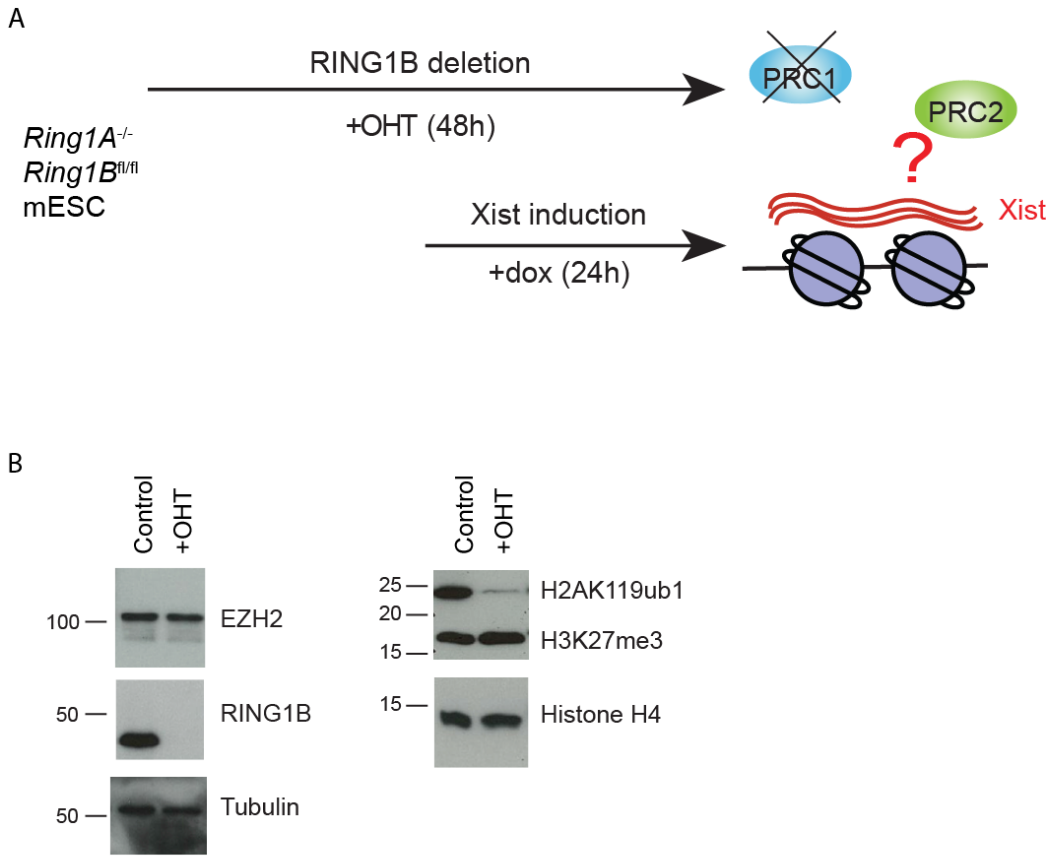


Figure 9: Global effects of deletion of *Ring1B* in *Ring1A*$^{-/-}$ *Ring1B*$^{fl/fl}$ mESCs. (A) Schematic representation of the experimental setup. Tamoxifen (+OHT) addition to the cell medium promotes deletion of Ring1B. Doxycycline (+dox) induces expression of Xist transgene. (B) Western blots of Polycomb proteins (EZH2 and RING1B), and Polycomb mediated histone modifications (H2AK119u1 and H3K27me3), before and 48h after tamoxifen (+OHT) addition to induce Cre-ERT2-mediated deletion of Ring1B. Loading controls are Tubulin and histone H4.

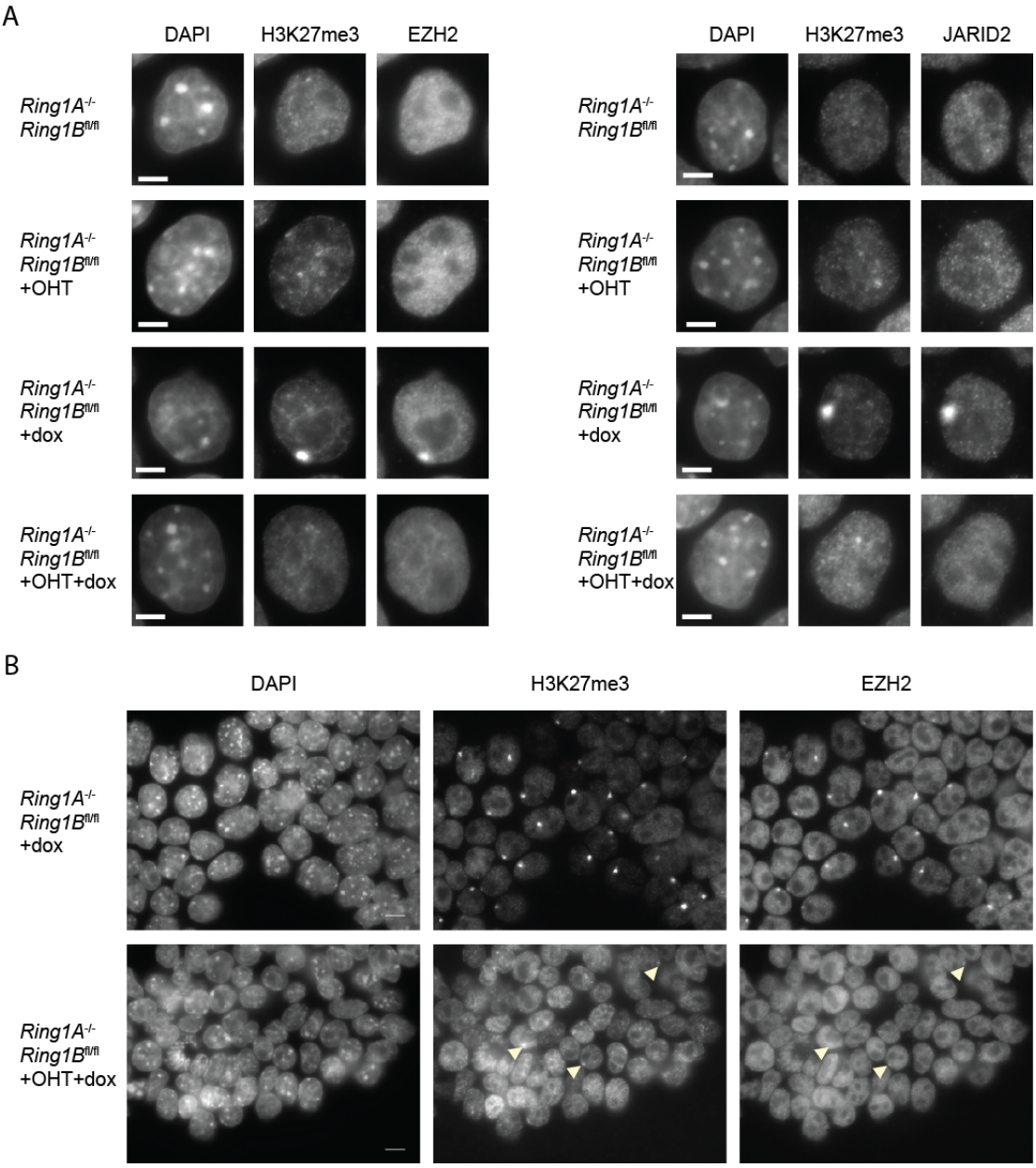


Figure 10: Immunofluorescence for H3K27me3, EZH2 and JARID2 in *Ring1A*$^{-/-}$ *Ring1B*$^{fl/fl}$ cells in the presence or absence of RING1B following Xist induction for 24 h. (A) Immuno-staining for H3K27me3 and EZH2 (left panel) and H3K27me3 and JARID2 (right panel) in *Ring1A*$^{-/-}$ *Ring1B*$^{fl/fl}$ mESCs. Tamoxifen (+OHT) induces depletion of RING1B and doxycycline (+dox) is used to induce Xist RNA expression. Scale bar is 5µm. (B) Single plane snapshot representing H3K27me3 and EZH2 domains (yellow arrows) still visible in a small proportion of cells in *Ring1A*$^{-/-}$ *Ring1B*$^{fl/fl}$ +OHT+dox comparing to control *Ring1A*$^{-/-}$ *Ring1B*$^{fl/fl}$ +dox. Scale bar is 10µm.

To confirm that the lack of PRC2 and H3K27me3 accumulation in characteristic inactive chromosome foci is not a result of Xist cloud formation defects, I performed RNA-FISH in parallel to each immunofluorescence assay. In the majority of cells Xist RNA still forms clearly detectable clouds similar to the control (Figure 11), as quantified in Figure 12. The probe used for this experiment is also able to detect Tsix RNA expressed from the endogenous X chromosome, visible as a punctate signal and shown in the RNA-FISH picture for the control cell lines in which Xist was not induced using doxycycline (*Ring1A*$^{-/-}$ *Ring1B*$^{fl/fl}$ and *Ring1A*$^{-/-}$ *Ring1B*$^{fl/fl}$ +OHT) (Figure 11). This punctate signal is also present in cells in which Xist was induced.

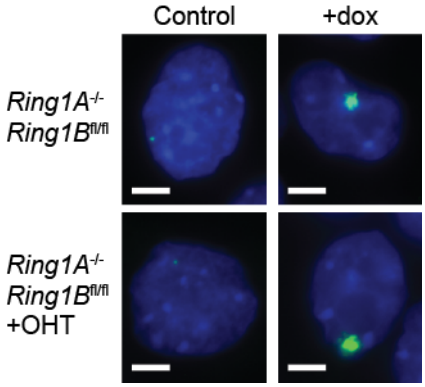


Figure 11: RNA-FISH for Xist RNA in *Ring1A*$^{-/-}$ *Ring1B*$^{fl/fl}$ cells and illustrating that deletion of genes encoding RING1A/B has no effect on Xist domain formation.

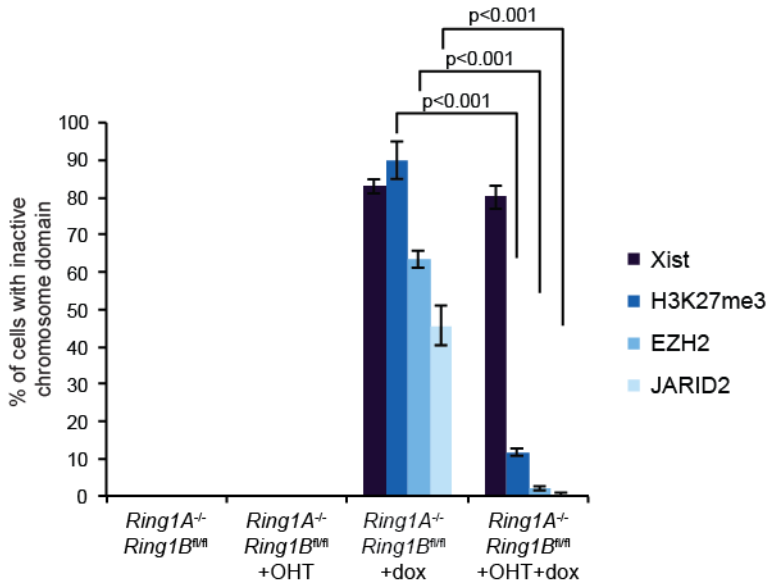


Figure 12: Bar chart summarizing the percentage of cells with domains for Xist, H3K27me3, EZH2 and JARID2 in tamoxifen (+OHT) treated and untreated cells after Xist induction for 24 hours (+dox). A minimum of 800 cells was counted in 4 biological repeats. Error bars indicate standard deviation.

3.3. Discussion

Analysis of PRC2 recruitment by Xist RNA in the context of RING1A/B depletion provides the first evidence for PRC1-dependent PRC2 recruitment in a physiological context. In summary, I was able to show that Xist RNA expression is not sufficient to recruit PRC2 in the absence of PRC1/H2AK119u1, suggesting that recognition of this complex, or the mark that it places on chromatin, is crucial for H3K27me3 deposition.

An independent experiment performed by Dr. Andrea Cerase in the Brockdorff lab, supports that Xist-mediated PRC2 recruitment is driven by recognition of H2AK119u1 and not the PRC1 complex itself. Briefly, expression of Xist RNA in cells treated with MG132, a proteasome inhibitor, which depletes global H2AK119u1 from chromatin without affecting PRC1 composition (Tavares et al. 2012), showed that PRC2 recruitment and H3K27me3 deposition are impaired (unpublished data). This fits well with prior data demonstrating that tethering of PRC1 complexes that are ineffective in placing H2AK119u1, does not lead to PRC2 recruitment (Blackledge et al. 2014).

The ability to induce PRC1 depletion before Xist induction allowed assessment of the efficiency of *de novo* PRC2 recruitment in a setting in which PRC1/H2AK119u1 are substantially depleted, thereby discounting positive feedback mechanisms. However, it is important to mention that even in this assay, a proportion of cells (around 10%) showed H3K27me3 accumulation in what seems to be a small inactive domain. I interpret this as a result of the fact that RING1B depletion does not occur synchronously and to the same extent in all cells, hence being present in enough levels in these few cells to allow H2AK119u1 Xist-mediated deposition and some PRC2 accumulation (mostly undetectable) leading to H3K27me3 domain formation.

Importantly, this experiment overturns the currently accepted model for Polycomb recruitment mediated by Xist RNA. Expression of full length Xist RNA was not enough to recruit PRC2, contradicting results from Zhao et al. (2008). It is important to emphasize that the different nature of the

experiments performed (*in vitro* vs. *in vivo*) might account for the differences in the result. I cannot rule out, based on my results, that PRC2 could potentially interact with Xist RNA, but if that is the case, this interaction is not sufficient to trigger Polycomb accumulation over the inactive chromosome in the absence of PRC1/H2AK119u1.

Da Rocha et al. (2014) reported that JARID2 is required to recruit PRC2 to the Xi and that this is mediated by the XN region of Xist RNA, encompassing B- and F-repeats. Unpublished findings from the Brockdorff lab have further shown that deleting the XN region of Xist RNA leads to loss of H2AK119u1, in addition to H3K27me3 (G. Pintacuda, personal communication). Additionally, a recent publication reported an UIM in JARID2 that was shown to contribute to its recruitment to H2AK119u1 sites (Cooper et al. 2016). These recent observations resolve the apparent discrepancy in the interpretation of the XN deletion and the findings reported in this chapter, thus confirming that loss of PRC2 recruitment by Xist RNA lacking the XN region was a consequence of failure to recruit PRC1. This also pinpoints the XN region of Xist RNA as being required for PRC1 recruitment.

In the context of Xist-dependent Polycomb domains, JARID2 recruitment to H2AK119u1 does not fully explain PRC2 accumulation as some H3K27me3 deposition is retained following JARID2 knockdown (da Rocha et al. 2014) or knockout (Cooper et al. 2016). Further studies will be important to define the alternative mechanisms that might contribute to PRC2 recruitment through recognition of H2AK119u1 on the inactive chromosome.

My results define PRC1 complex(es) as being required to initiate Polycomb domain formation upon Xist expression. Which one of the multiple PRC1 complexes are recruited to the inactive chromosome domain and how are they recruited are the focus of the next chapters in this thesis.

CHAPTER 4: Localization and dynamics of PRC1 complexes on the inactive chromosome domain

4.1. Introduction

Evidence that PRC2 recruitment to chromatin in response to Xist RNA expression (as shown in Chapter 3) requires prior PRC1 mediated H2AK119u1 deposition led me to investigate PRC1 complexes in the context of X inactivation.

It has been previously shown that RING1B, the catalytic subunit of PRC1, localizes to the Xi early after initiation of random X inactivation in postimplantation mouse embryos, as well as to the imprinted inactive X in trophoectoderm and inner cell mass cells of blastocyst stage embryos (de Napoles et al. 2004). Consistent with RING1B activity in the context of PRC1 complexes, H2AK119u1 also accumulates on the X chromosome both in differentiating female stem cells and in mouse fibroblasts (de Napoles et al. 2004). RING1A and RING1B were shown to have overlapping function in depositing mono-ubiquitin on H2AK119 (de Napoles et al. 2004). Other PRC1 components such as MEL-18 (known as PCGF2) and BMI-1 (PCGF4), MPH1 and MPH2, CBX2 and CBX4, and the alternative catalytic subunit RING1A were reported to be enriched on the Xi in differentiating female mouse ES cells, mouse ES cells expressing a Xist transgene, TS cells, and terminally differentiated mouse fibroblasts (de Napoles et al. 2004; Plath et al. 2004; Schoeftner et al. 2006). Interestingly, temporal accumulation of these proteins

differs over the course of differentiation, potentially due to expression changes or different dynamics in differentiating cells (de Napoles et al. 2004; Plath et al. 2004; Schoeftner et al. 2006).

While together these published data unequivocally demonstrate that PRC1 is recruited to the inactive X chromosome, catalyzing H2AK119u1 on underlying chromatin, recently it has been shown that a multiplicity of PRC1 complexes are present in mammalian cells. Besides the canonical PRC1 complexes that are recruited to chromatin through docking of CBX proteins chromodomain to H3K27me3 (Cao et al. 2002; Fischle et al. 2003), a myriad of non-canonical complexes that do not include any CBX proteins was recently described (Tavares et al. 2012; Gao et al. 2012). Instead of CBX, each of these non-canonical PRC1 complexes includes the protein RYBP (or its homolog YAF2), which competes with CBX proteins to bind the C-terminal region of RING1B (Wang et al. 2010). The fact that these non-canonical complexes could be recruited to chromatin independently of H3K27me3 provides a potential explanation for previous observations that PRC1 could be recruited and function in PRC2 depleted cells (*Eed*$^{-/-}$), both at selected gene targets (Leeb et al. 2010) as well as on the inactive X chromosome (Schoeftner et al. 2006). These results have been confirmed genome-wide by ChIP-seq analysis of RING1B as well as on the inactive chromosome by staining of RYBP protein in *Eed*$^{-/-}$ mESCs (Tavares et al. 2012). Together these observations provided the first evidence that, in addition to canonical PRC1 complexes containing PCGF2/4, MPH and CBX proteins, non-canonical PRC1 complexes are also enriched on Xist-decorated chromatin.

Further characterization of non-canonical PRC1 complexes showed that there are at least six variants defined by the presence of one of six different PCGF proteins (PCGF1-6), which bind to RING1B N-terminus in a mutually exclusive manner (Gao et al. 2012). Importantly, both PCGF2 and PCGF4 were shown to be able to assemble into CBX-containing canonical PRC1 complexes as well as RYBP-containing non-canonical PRC1 (Gao et al. 2012). On the other hand, PCGF1, 3, 5 and 6 were exclusively found in the context of non-canonical PRC1 (Gao et al. 2012), and the presence of these complexes on the Xi has never been addressed. Of note, a recent study provided evidence that non-canonical PRC1 complexes represent the majority of PRC1 complexes in mESCs (Kloet et al. 2016). In this chapter I describe experiments to investigate if different PCGF proteins are enriched over Xist domains in live cells (Ng et al. 2011), using a cell line containing an inducible Xist transgene stably integrated on chromosome 17 (Moindrot et al. 2015).

4.2. Results

4.2.1. Generation and characterization of stable mESC lines expressing eGFP fused to PRC1 subunits

To elucidate which non-canonical PRC1 variants are enriched on the inactive chromosome upon Xist RNA expression, I made use of a mESC line, 3E-H, in which an inducible Xist transgene is expressed from chromosome 17 after doxycycline induction. The advantages of this inducible Xist system were introduced in chapter 3. Classical studies showing that specific factors (for

example, PRC2 components, SAF-A and others) and histone modifications (H3K27me3, H2AK119u1) are enriched on the Xi often make use of specific antibodies that allow visualization of proteins/ modifications in fixed cells. However, sub-nuclear resolution detection of many PRC1 components, the PCGF proteins in particular, is not currently possible using available antibodies. Thus, I decided to take an alternative approach, expressing PCGF proteins fused with enhanced green fluorescent protein (eGFP), that would both allow me to visualize PRC1 components in live cells as well as perform immunofluorescence assays using the same antibody for all of the proteins of interest, discounting the variability of detection of these proteins with different antibodies using differently optimized protocols. Accordingly I set out to express PCGF1, PCGF2, PCGF3, PCGF5 and PCGF6, as well as RING1B and RYBP as controls, fused with eGFP on the N-terminus.

The expression constructs were assembled using cDNA amplified from a 129/1 mESC line using specific primers for each transcript of interest (see Chapter 2), which was inserted by LIC cloning into the pCAG-eGFP-IRES-Puro mammalian expression vector depicted in Figure 13A. I then transfected the plasmids into 3E-H mESCs to establish stable cell lines expressing each protein of interest (eGFP-PCGF1, eGFP-PCGF2, eGFP-PCGF3, eGFP-PCGF5 and eGFP-PCGF6). RYBP and RING1B, which are known to be enriched on the inactive X (de Napoles et al. 2004; Schoeftner et al. 2006; Tavares et al. 2012), were tagged with eGFP in the same way (Figure 14).

Expression levels of all eGFP tagged proteins were assessed by Western blot using anti-GFP antibody as shown in Figure 13C and Figure 14A. eGFP-PCGF3 was consistently the lowest expressed fusion protein in all analyzed clones. eGFP-PCGF2 and eGFP-PCGF6 showed a slower migrating band on the Western blot besides their expected size bands (Figure 13C). This observation is consistent with the fact that both PCGF2 and PCGF6 are phosphorylated in mouse and human cells (Elderkin et al. 2007; Akasaka et al. 2002) and suggests that this post-translational modification may be present in the eGFP-tagged version of these proteins. Immunostaining using anti-GFP antibody showed that, as expected, all the eGFP fusion proteins localize predominantly to the cell nucleus (Figure 13B).

A co-immunoprecipitation assay was used to confirm that eGFP-fusion proteins form PRC1 complexes in the cells, showing specific interaction with RING1B but not with EZH2, a component of the PRC2 complex. A stable cell line expressing eGFP only was used as a negative control in the immunoprecipitation experiment. Immunoprecipitation of eGFP on its own does not pulldown RING1B or EZH2, showing that the interaction between eGFP-PCGFs and RING1B is specific to the PCGF proteins (Figure 13D).

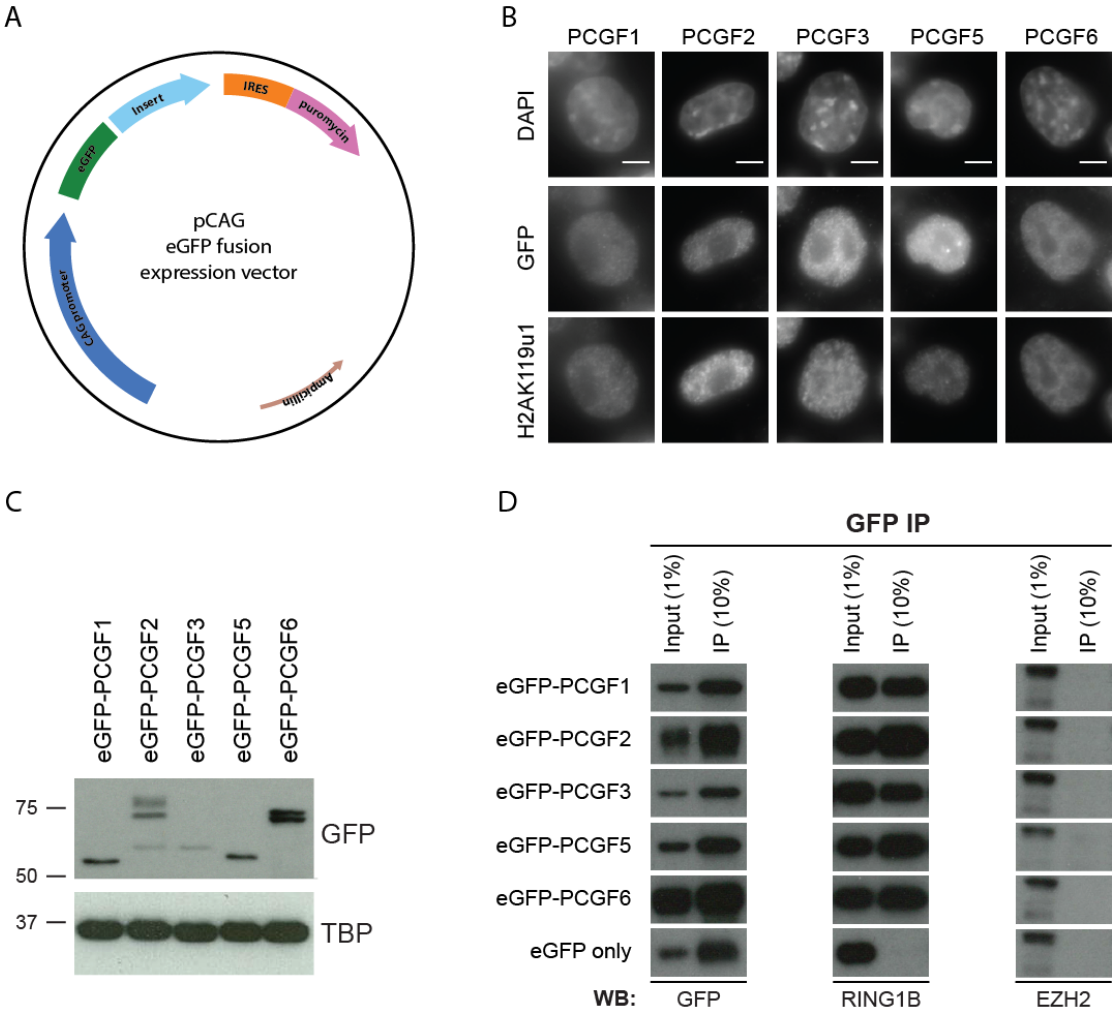


Figure 13: Establishment of stable mESC lines expressing eGFP-PCGF proteins. (A) Schematic representation of pCAG-eGFP-IRES-puromycin plasmid used for expression of eGFP-PCGF proteins in mouse cells. (B) Immunofluorescence using anti-GFP antibody showing nuclear staining for all five eGFP-PCGF proteins. H2AK119u1 staining is used as a nuclear marker. Scale bar is 5µm. (C) Western blot illustrating levels of eGFP-PCGF fusion proteins in stable cell lines. TATA-binding protein (TBP) is used as loading control. (D) Western blot illustrating co-immunoprecitation of endogenous RING1B with eGFP-PCGF fusion proteins using anti-GFP antibody and nuclear extracts from mESC stable lines. A stable cell line expressing eGFP only and EZH2 Western blot were used as negative controls for the co-immunoprecipitation experiment.

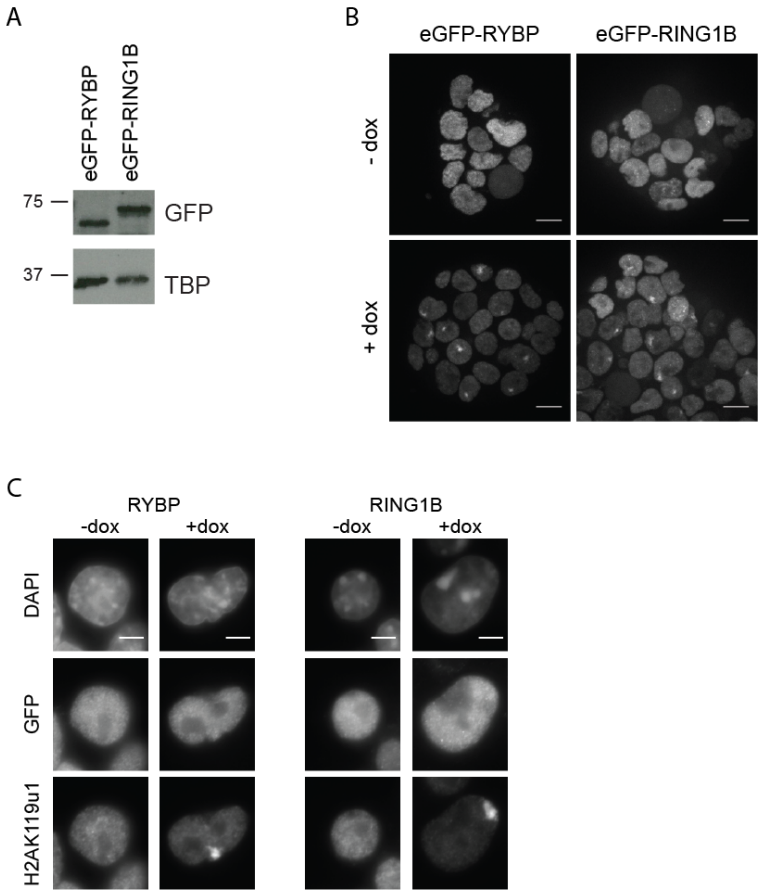


Figure 14: Establishment of stable mESC lines expressing eGFP-RYBP and eGFP-RING1B proteins. (A) Western blot illustrating levels of eGFP-RYBP and eGFP-RING1B fusion proteins in stable cell lines. TATA-binding protein (TBP) is used as loading control. (B) Live-cell imaging illustrating nuclear distribution of eGFP-RYBP and eGFP-RING1B proteins in in 3E-H mESC line. Xist-resembling domains are visible both for eGFP-RYBP and eGFP-RING1B after Xist induction (+dox) for 24 hours. Images are maximum intensity projections of six consecutive z-stacks. Scale bar is 10μm. (C) Immunofluorescence using anti-GFP antibody showing nuclear staining for eGFP-RYBP and eGFP-RING1B proteins. Neither protein shows enrichment over the inactive chromosome domain by immunofluorescence (+dox). H2AK119u1 staining is used as a nuclear marker and a surrogate marker for the inactive chromosome. Scale bar is 5μm.

4.2.2. Recruitment of non-canonical PRC1 complexes to the inactive chromosome domain

Using the aforementioned cell lines I went on to test which non-canonical PRC1 complexes are recruited to the inactive chromosome domain using live cell imaging. Spinning disk confocal microscopy was applied as it allows fast and sensitive imaging of dynamic processes inside living cells. As expected, eGFP-RYBP and eGFP-RING1B both showed enrichment in one discrete focus per nucleus upon doxycycline induction of Xist RNA expression, validating this technique for detection of proteins enriched over the inactive chromosome domain (Figure 14B). Surprisingly, using immunofluorescence (using anti-GFP antibody) neither of these proteins could be detected overlapping with H2AK119u1 foci (here used as surrogate marker for the inactive chromosome) (Figure 14C). Possible reasons for this are discussed below.

To clarify which of the non-canonical PRC1 complexes are enriched on the inactive chromosome domain I performed the same assays in the stable clones expressing eGFP-PCGF fusion proteins. Live imaging of these cell lines showed that all of the non-canonical PRC1 complexes (represented by eGFP-PCGF1/3/5/6) and in addition eGFP-PCGF2, which can be present in both canonical and non-canonical complexes, are recruited to the inactive chromosome domain after induction of Xist RNA expression (+dox). Images, represented in Figure 15, were obtained 24 hours after Xist RNA induction, but accumulation of all eGFP-PCGF proteins on the inactive chromosome

region is visible as early as 6 hours after Xist induction (data not shown). The characteristic inactive chromosome domain foci are not visible in control cells (-dox) (Figure 15, top row). eGFP-PCGF2 and eGFP-RING1B show a very distinctive distribution in the nucleus accumulating in small bright speckles, both in control and in Xist expressing cells (Figure 14 and Figure 15). Similar speckles were previously observed in MEFs using live cell imaging of RING1B and are thought to be the result of polymerization of PHC, an exclusive component of canonical PRC1 complexes (Isono et al. 2013).

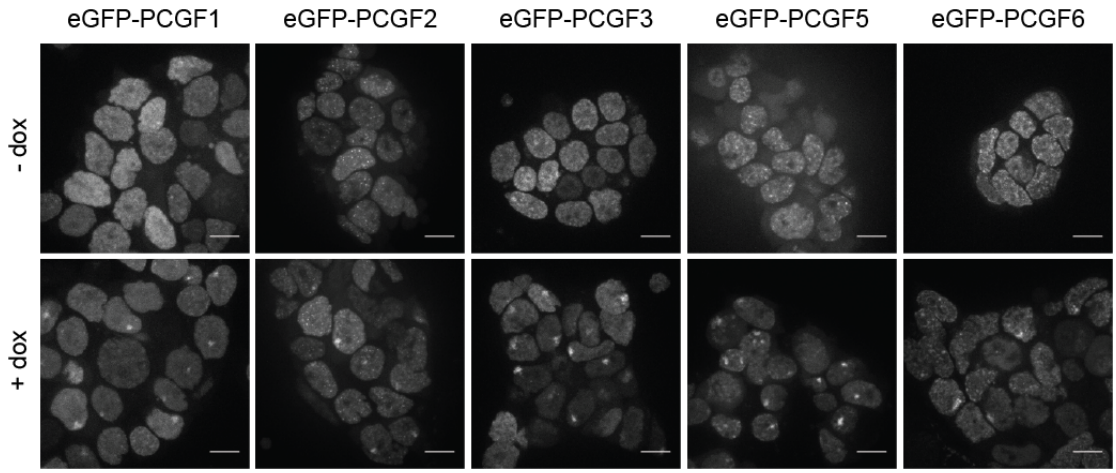


Figure 15: Xist dependent localisation of non-canonical PRC1 complexes by live cell imaging. Live imaging of mESCs illustrating focal enrichment of eGFP-PCGF fusion proteins following induction of Xist RNA for 24h in 3E-H mESC line (+dox). Focal enrichment is not seen in uninduced cells (-dox). Images are maximum intensity projections of six consecutive z-stacks. Scale bar is 10µm.

Enrichment of eGFP-PCGF proteins over the inactive chromosome domain was also seen by immunofluorescence, but to a different extent depending on the protein. eGFP-PCGF3 and eGFP-PCGF5 were detected overlapping with

H2AK119u1 domains in a large proportion of cells, while eGFP-PCGF2 and eGFP-PCGF6 could only be seen in a low percentage of cells. eGFP-PCGF1 enrichment could not be detected using immunofluorescence (Figure 16).

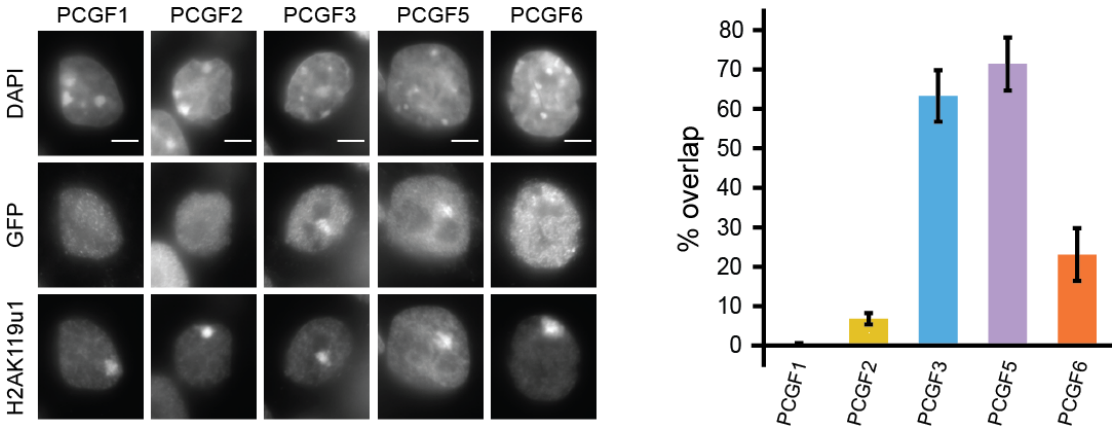


Figure 16: Xist dependent localisation of non-canonical PRC1 complexes assayed by immunofluorescence. Images show representative single cells stained for GFP and H2AK119u1 after 24h of Xist induction. DNA was counterstained with DAPI. Scale bar is 5µm. The bar chart on the right represents percentage of cells in which focal localisation of H2AK119u1 and eGFP-PCGF fusion proteins coincides (% overlap). Data was obtained from scoring a minimum of 400 cells in at least four replicates. Error bars indicate standard deviation between replicates.

A prior study demonstrated that proteins that interact with chromatin with fast dynamics could be difficult to detect using formaldehyde fixation compared with live cell imaging (Schmiedeberg et al. 2009). With this in mind I hypothesized that the discrepancy between the results in live cell imaging and immunofluorescence may indicate that the different eGFP fusion proteins have different chromatin-binding dynamics. This can be due to intrinsic differences between variant PRC1 complexes or potentially due to alternative binding modes to the inactive chromosome, allowing for different degrees of

fixation when using formaldehyde in the immunofluorescence protocol. These differences will be further explored in the next sections.

4.2.3. Dynamics of different PRC1 complexes in the nucleus and the inactive chromosome domain

Differences in protein dynamics in the nucleus might depend on their abundance or interactions within the nucleus, both with other proteins as well as with chromatin or RNA (reviewed in Meshorer and Misteli 2006; Phair et al. 2004; Daigle and Ellenberg 2007). Fluorescence Recovery After Photobleaching (FRAP) experiments have proven useful to measure protein turnover in the nucleus and to characterize their dynamics in different chromatin compartments (euchromatin vs. heterochromatin) or in different chromatin states (pluripotent vs. differentiated cells) (Cheutin et al. 2003; Meshorer et al. 2006; Ren et al. 2008). By bleaching the GFP signal over a small region of the nucleus using a short pulse of a high intensity laser and measuring its recovery in the same region over time, one can calculate how fast fluorescent molecules of the GFP tagged protein move into the bleached area. Measuring the percentage of the initial pre-bleach fluorescence that is recovered over a certain period of time also provides information about which proportion of the protein is stably bound to the region of interest in the timeframe of the experiment (bleached molecules that remain bound to the region of interest thus not being replaced by fluorescent ones).

To investigate if there is an intrinsic difference in dynamics between different PRC1 complexes, I set out to conduct FRAP measurements in the nucleoplasm of the stable cell lines expressing eGFP-PCGF proteins. The same experiment was also performed over the inactive chromosome domain to clarify if potential differences in dynamics are acquired over the inactive chromosome compartment, potentially due to acquisition of a new specific binding mode to chromatin when compared with elsewhere in the genome.

Conditions for the FRAP experiment were optimized during a trial experiment with a smaller number of cells (around 5 per cell line), varying the duration of the high intensity laser pulse used for bleaching (25ms, 50ms and 100ms), and in some cases testing two different clones for the same cell line (data not shown). The size of the bleaching spot (radius=1.25μm) and the laser intensity (100%) were kept unchanged across the trial experiment. Clones that allowed imaging with a better signal-to-noise ratio were preferred as they allow for less noise in the fluorescence recovery curves, without affecting the overall recovery. Bleach depth (percentage of fluorescence that is bleached in the region of interest) was compared between the three different bleaching pulse durations but no significant differences were observed. A bleach depth of >85% over the inactive domain and >65% on the nucleoplasm were achieved in this trial. The fact that I was unable to reach a higher bleach depth on the nucleoplasm, even using a longer bleach duration suggests that this has to do with a general high mobility, or at least with the presence of a fast recovery fraction for each protein. A lag between bleaching and capturing the first frame after bleaching (0.005 sec) allows for some fluorescence recovery

that cannot be captured by the microscope camera, leading to an underestimation of the bleach depth. I considered that these conditions were satisfactory to continue with the experiment, using 50ms bleach duration. This trial suggested that eGFP-PCGF turnover is very fast so all frames were acquired at the maximum speed rate (approximately 8 fps in the current setup).

Analysis of average fluorescence recovery curves for all eGFP-PCGF proteins measured in random points in the nucleoplasm (outside of the nucleolus), show that their turnover occurs on a timescale of seconds (Figures 17 and 18). In the first 5 seconds after photobleaching, all fusion proteins already recovered to >80% of initial intensity (Figures 17 and 18). The stably bound fraction represents only a very small proportion of these proteins in the nucleoplasm, as in the timescale of the experiment all eGFP fused proteins recovered to 95% or more of the initial fluorescence (Figure 18). Both of these observations reflect the expected hyperdynamic behavior of chromatin binders previously reported in pluripotent ES cells (Meshorer et al. 2006).

The same analysis was performed for eGFP-PCGF1, eGFP-PCGF2, eGFP-PCGF3, eGFP-PCGF5 and eGFP-PCGF6, enriched over the inactive chromosome region, approximately 22-26h after expression of Xist RNA in mESCs (Figures 17 and 18). Interestingly, the different eGFP-PCGF proteins showed distinct dynamic behavior over the inactive chromosome domain when compared to the nucleoplasm. Analysis of average fluorescence recovery curves of the domain compared to the nucleoplasm showed that all the proteins present an initial slower recovery rate in the inactive domain

region (Figure 18). The stably bound fraction within the domain is greater than in the nucleoplasm, especially for eGFP-PCGF3 and eGFP-PCGF5 (Figure 18).

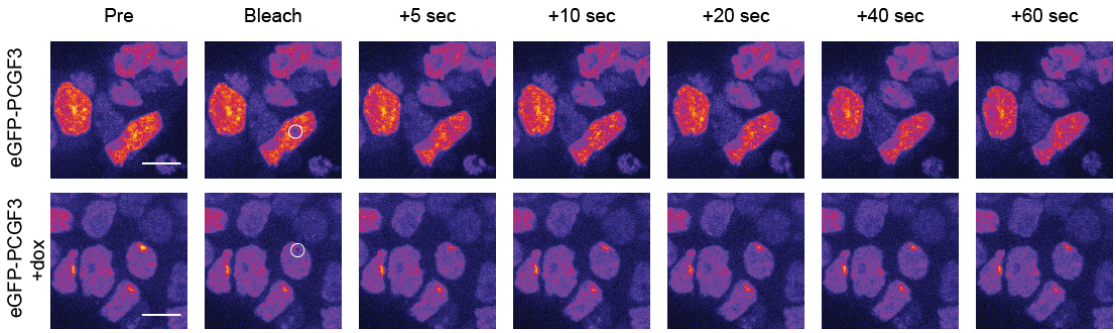


Figure 17: eGFP-PCGF proteins show different dynamics in the nucleoplasm and the inactive chromosome domain. Images are representative of a FRAP experiment for 3E-H cell line expressing eGFP-PCGF3. Fluorescence on the bleached region (indicated by a circle on bleach snapshot) recovers to pre-bleached levels on the nucleoplasm (top row) but not on the inactive chromosome domain (+dox, bottom row). Scale bar is 10µm.

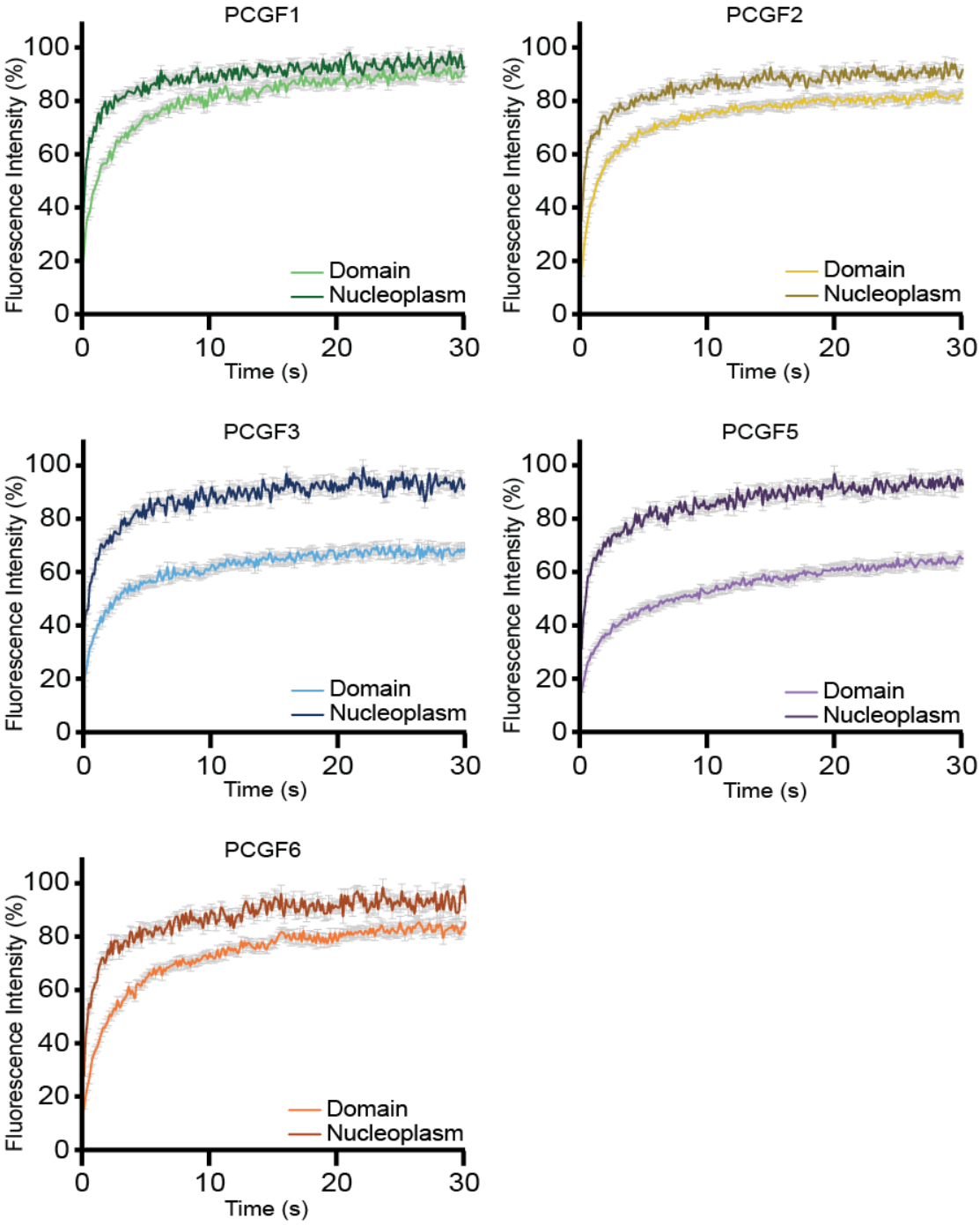


Figure 18: Fluorescence recovery after photobleaching curves for eGFP-PCGF proteins in the nucleoplasm and the inactive chromosome domain. Average fluorescence recovery curves for all eGFP-PCGF proteins over the first 30 seconds after photobleaching show an initial slower recovery over the inactive chromosome domain when compared to the nucleoplasm. A minimum of 24 cells was analyzed for each region of interest. Grey bars represent standard error.

For quantitative comparison of protein mobility, individual FRAP curves were fitted using a biexponential function (Figures 19A and 19B). This allowed me to calculate the half time of recovery ($t_{1/2}$) for each protein in each context (Figure 19C), as well as the immobile fraction over the timescale of the experiment (Figure 19D). All the quantitative FRAP analysis was done in collaboration with Dr. David Brown, a postdoc from Professor Rob Klose's lab.

Pairwise comparison of $t_{1/2}$ values for all eGFP-PCGF proteins in the inactive domain with the nucleoplasm using Mann-Whitney U test showed that their half time of recovery is significantly slower in the inactive chromosome domain compared to the nucleoplasm ($p<0.01$ for eGFP-PCGF2; $p<0.001$ for eGFP-PCGF1, eGFP-PCGF3, eGFP-PCGF5 and eGFP-PCGF6) (Figure 19C). This suggests that all these proteins have specific interaction modes within the inactive chromosome domain (either with other proteins or with Xist RNA directly), or that there is an enrichment of chromatin binding sites in the inactive chromosome when compared to the rest of the nucleus. Further comparisons between different eGFP-PCGF within the domain suggest that these specific interactions have a stronger effect on the mobility of eGFP-PCGF5 and eGFP-PCGF6 ($p<0.001$ in all pairwise comparisons with other eGFP-PCGF proteins using Mann-Whitney U test). eGFP-PCGF3 turnover is also affected compared to eGFP-PCGF1/2 ($p<0.05$), but not to the same extent as eGFP-PCGF5/6 (Figures 19A and 19C). Even though some of these proteins show different intrinsic turnover rates on the nucleoplasm ($t_{1/2}$ for PCGF1 is faster than PCGF2/3/5; PCGF6 is intermediate between these two groups) (Figures 19B and 19C), these differences are much more significant

when comparing PCGF proteins within the domain ($t_{1/2}$ for PCGF5 is significantly slower than all other proteins; PCGF6 and PCGF3 are marginally different from each other, but both of them are significantly slower than PCGF1 and PCGF2 for which turnover rate is indistinguishable in the domain) (Figures 19A and 19C).

Comparison of the immobile fraction for each protein shows that a significant proportion of eGFP-PCGF2/3/5 is more stably bound to the inactive chromosome than to elsewhere in the nucleus ($p<0.001$ in pairwise comparisons between inactive domain and nucleoplasm) (Figure 19D). Within the inactive chromosome domain, eGFP-PCGF3 and eGFP-PCGF5 are the proteins that appear to be most stably bound to chromatin ($p<0.001$ for eGFP-PCGF3 vs. eGFP-PCGF1/2/6 and eGFP-PCGF5 vs. eGFP-PCGF1/6; $p<0.05$ eGFP-PCGF5 vs. eGFP-PCGF2) (Figures 19A and 19D). In the nucleoplasm, eGFP-PCGF5 and eGFP-PCGF1 represent the most different proteins in terms of immobile fraction ($p<0.01$), but their immobile fractions are not significantly different from other eGFP-PCGF proteins (Figures 19B and 19D). This establishes a range of immobile fraction for all eGFP-PCGF in the nucleoplasm that varies from nearly 0 to 9% percent (median).

These results show that eGFP-PCGF5 is more stably bound to the inactive domain (median immobile fraction around 27%) and that its mobile fraction turns over more slowly in this chromatin compartment than in the rest of the nucleus. eGFP-PCGF6 also exhibits a slower turnover rate in the inactive chromosome domain, but does not form a stably bound fraction, contrary to

eGFP-PCGF3 that, despite having a faster turnover than eGFP-PCGF6, presents an immobile fraction corresponding to around 28% (median) of the total protein in the bleached area.

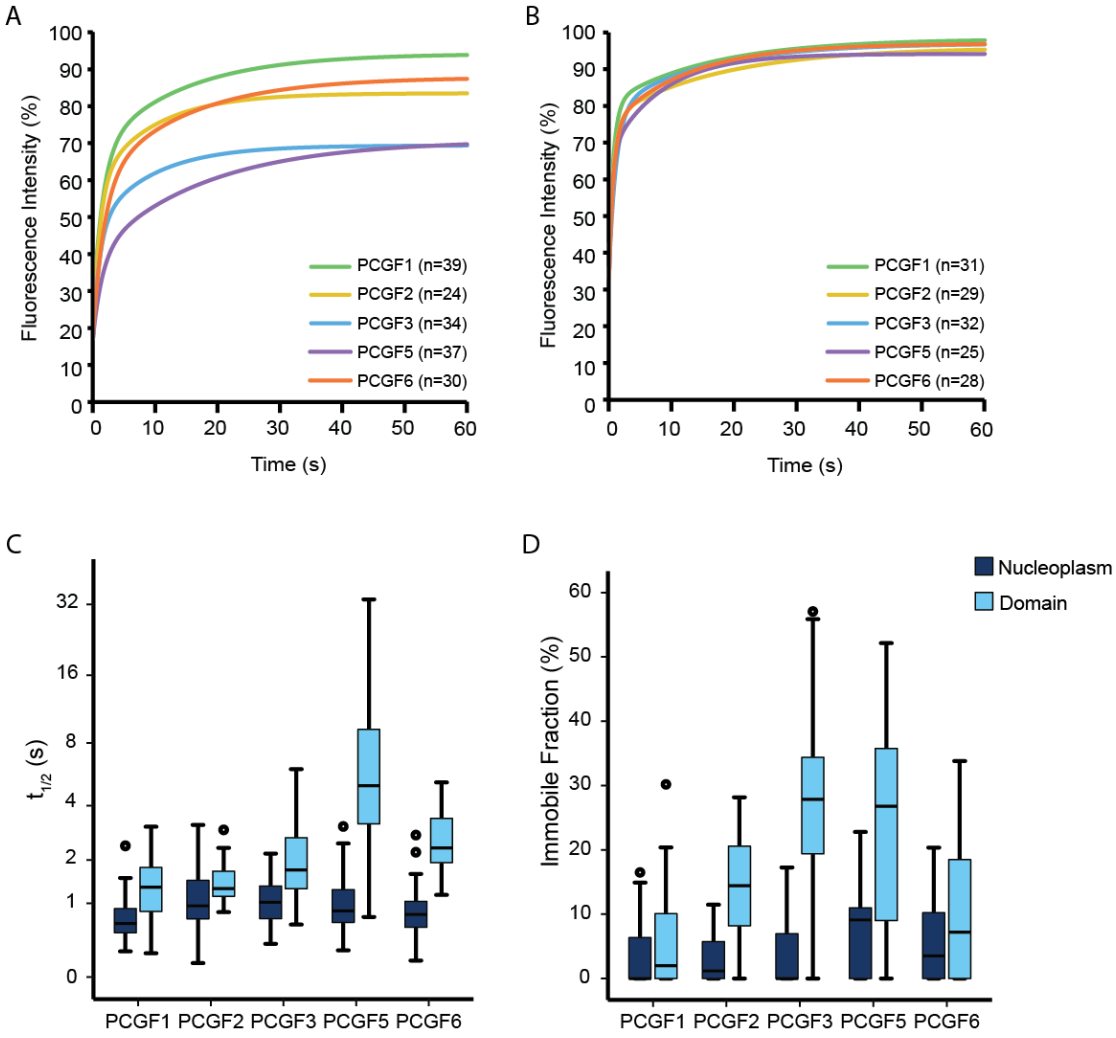


Figure 19: Quantitative analysis of FRAP data using a bi-exponential function. Fluorescence recovery data for each eGFP-PCGF protein in single cells was fitted to a bi-exponential model for quantitative analysis. Average fitted curves for each protein (A) on the inactive chromosome domain and (B) on the nucleoplasm are shown. Number of cells analyzed is showed in brackets (n). (C) Distribution of half-time ($t_{1/2}$) recovery for each protein in each region of interest is represented in a boxplot. $t_{1/2}$ is shown in seconds and represented on a $\log_2$ scale. (D) Distribution of percentage of eGFP-PCGF fusion proteins in the immobile fraction 60 seconds after

photobleaching calculated for the nucleoplasm and the inactive chromosome domain.

4.3. Discussion

In summary, I have shown that all of the non-canonical PRC1 complexes are enriched on the inactive chromosome after Xist expression is induced. Analysis of FRAP curves for each eGFP-PCGF protein suggests that all non-canonical PRC1 complexes have a significantly slower turnover on the inactive chromosome domain when compared to their dynamics on the nucleoplasm. Interestingly, for two out of the five PCGF proteins, namely PCGF3 and PCGF5, approximately 30% of the protein observed at the inactive domain remains stably bound to this compartment over 60 sec after fluorescence bleaching. I also observed a strong correlation between protein dynamics as determined by FRAP and the ability to detect these proteins by immunofluorescence on the inactive chromosome domain.

Live cell imaging of mESCs expressing eGFP-PCGF proteins showed, as expected, that all of them mainly localize to the nucleus, being excluded from the nucleolar compartment. Importantly, these tagged proteins were shown to retain specific interactions with endogenous RING1B, indicating that eGFP-PCGF fusions provide a valuable tool for studying PRC1 complexes inside the cells. Also, clustering of eGFP-RING1B and eGFP-PCGF2 in speckles resembling Polycomb bodies in the nucleus, as previously reported by others (Isono et al. 2013), reinforce the idea that these fusion proteins are

assembled and distributed similarly to their endogenous counterparts. Furthermore, use of the eGFP tag allows a systematic and unbiased comparison between proteins using the same antibody.

Expression of Xist RNA triggered enrichment of all five eGFP-PCGF proteins over the inactive chromosome domain. This provided evidence that all variants of non-canonical PRC1 complexes are recruited to Xist-coated chromatin. However, it is important to emphasize that the use of live cell imaging was crucial to make this observation, since analysis of the same cells using immunofluorescence only, would lead to different conclusions. By immunofluorescence, eGFP-PCGF1 is not detected on the inactive domain, and eGFP-PCGF2 and eGFP-PCGF6 are only seen in a small proportion of cells. Although I cannot rule out that optimization of the protocol could lead to different results, it seems likely that the differences between live cell and fixed cell analysis is related to chromatin-binding dynamics affecting the ability to formaldehyde fix a specific protein over the Xist territory. It is important to emphasize that the same antibody was used to detect all five eGFP-PCGF proteins using the same protocol, and that detection of eGFP-PCGF3 and eGFP-PCGF5 on the inactive domain are unlikely to result from protein overexpression since these two fusions, especially eGFP-PCGF3 was the lowest expressed of the five transgenes.

GFP fusion proteins were previously used to address dynamics of canonical PRC1 components (CBX proteins and BMI-1) in the nucleoplasm or over Polycomb bodies in mammalian cells (Ren et al. 2008; Hernandez-Munoz et

al. 2005), as well as for factors enriched on the Xi (including CBX7, a canonical PRC1 component) (Fackelmayer 2004; Escamilla-Del-Arenal et al. 2013). All eGFP-PCGF proteins analyzed in this study show a very fast turnover (median $t_{1/2}$ below 2 seconds) and small immobile fractions (median below 10%) in the nucleoplasm. This fits well with FRAP measurements obtained for other chromatin binders, including CBX proteins, that show a similar hyperdynamic behavior in ES cells (Meshorer et al. 2006; Phair and Misteli 2000; Ren et al. 2008). It has been proposed that transient binding to chromatin is an important feature that contributes to maintain the plasticity in undifferentiated pluripotent cells (Meshorer et al. 2006).

Importantly, previous analysis by FRAP of a protein that accumulates on the Xi, CDYL, provided an important clue about how its mode of binding to the inactive chromosome, with recognition of H3K27me3 and H3K9me3 affects interaction dynamics when compared to the rest of the nucleus (Escamilla-Del-Arenal et al. 2013). The change of CDYL dynamics on the Xi is similar to what was observed in this study for eGFP-PCGF3 and eGFP-PCGF5, which present an immobile fraction, i.e. proportion of protein that remains bound to the bleached area over the course of the FRAP experiment (60 sec in this case), of around 30% on the inactive chromosome domain compared to only 10% or less in the nucleoplasm. Also, it is important to note that all eGFP-PCGF proteins analyzed presented a slower turnover over the inactive domain when compared to the nucleoplasm, especially eGFP-PCGF5 and eGFP-PCGF6. It is then likely that a combination of a slower turnover rate with a larger immobile fraction of a protein over the Xist territory could

contribute to its ability to be fixed and visualized using immunofluorescence. Indeed, as mentioned above, a previous study of the DNA binding protein MeCP2, in its wild-type and mutant forms, suggested that a protein requires a minimum residency time to be effectively formaldehyde cross-linked to DNA (Schmiedeberg et al. 2009). This provides a potential explanation for the detection by immunofluorescence of eGFP-PCGF3 and 5, and also for eGFP-PCGF6 to a certain extent. The detection of eGFP-CBX proteins on the Xi using the same methodology (Bernstein et al. 2006b) suggests that canonical PRC1 could have slower turnover than some non-canonical PRC1 which is in agreement with FRAP data for CBX7 (estimated $t_{1/2}$ on the Xi = 4.47 ± 0.96 sec) (Escamilla-Del-Arenal et al. 2013).

Examination of FRAP recovery curves suggests that all eGFP-PCGF proteins have a fast and a slow recovering fraction, notably in the context of Xist-coated chromatin. The fast fraction is responsible for the initial recovery of fluorescence and occurs within the first seconds after bleaching, while the slow fraction is responsible for the later fluorescence recovery and reflects a less mobile, more stably chromatin-bound fraction of the protein. This has been described as a common feature of chromatin binders (Phair et al. 2004) as well as for BMI-1 (PCGF4) at Polycomb bodies (Hernandez-Munoz et al. 2005). However, the mechanistic biological details contributing to each one of these behaviors are not clear and could be a result of different interactors/ recruiters for each fraction of the protein or availability of binding sites/ stability of the interaction with chromatin. Findings in the following chapters

shed some light on the basis of these behaviors of eGFP-PCGF fusion proteins.

CHAPTER 5: The role of RYBP/YAF2 in PRC1 recruitment to the inactive chromosome

5.1. Introduction

In the previous chapter, I showed that all non-canonical PRC1 complexes localize to Xist-coated chromatin albeit with different dynamics. In addition to the catalytic core subunits RING1A or RING1B, all of these non-canonical PRC1 complexes include a subunit called RYBP (or its homolog YAF2) (Gao et al. 2012). The presence of RYBP/YAF2 distinguishes non-canonical PRC1 from canonical PRC1 complexes. RYBP/YAF2 bind to the same region of RING1B as CBX proteins (Wang et al. 2010; García 1999), and thereby defines a alternative type of PRC1 complex that is not recruited to chromatin through recognition of H3K27me3 (Tavares et al. 2012; Hisada et al. 2012). Accordingly, depletion of PRC2 (*Eed*$^{-/-}$ mESCs) or canonical PRC1 (through knockdown of CBX7) has little or no effect on RYBP occupancy at gene targets (Tavares et al. 2012; Morey et al. 2013; Hisada et al. 2012).

The mechanisms by which RYBP-containing complexes are recruited and function on chromatin has been a point of debate. *In vitro* studies of H2A ubiquitylation activity of PRC1, comparing RYBP or CBX-containing complexes, gave conflicting results, with one study suggesting that both variants are equally active (Tavares et al. 2012) while another study reports an increase in activity of complexes that include RYBP compared to those with the CBX subunit (Gao et al. 2012). More recent studies suggest that

activity of PRC1 *in vitro* is also dependent on the other subunits in the complex, namely different PCGF proteins (Rose et al. 2016; Taherbhoy et al. 2015), but that in all cases RYBP/YAF2 stimulates activity (Rose et al. 2016). Attempts to address this further led Rose *et al.* (2016) to suggest that *in vivo* as well as *in vitro*, RYBP is important for activity but not recruitment of PRC1 complexes. This was because knockout of *Rybp*/*Yaf2* led to decreases in H2AK119u1 but had no effect on RING1B occupancy (Rose et al. 2016). However, these results are conflicting with ChIP data from Morey et al. (2013) that shows that RYBP-target sites show lower levels of H2AK119u1 than CBX-PRC1 targets. These authors also report a reduction of H2AK119u1 both in CBX7 as well as in RYBP knockdown cells.

It is also important to consider that a role for RYBP in PRC1 recruitment has been suggested (Arrigoni et al. 2006). RYBP has a highly conserved Zinc-Finger Ubiquitin Binding Domain (NZF-UBD) on its N-terminus. This fragment was shown to recognize ubiquitin and promote RYBP binding to H2AK119u1, although *in vitro* the binding was fairly weak (479 ± 4 μM) (Arrigoni et al. 2006).

I set out to address the relevance of RYBP/YAF2 in the context of Xist-mediated polycomb recruitment, focusing on the contribution of these proteins towards PRC1 complex stability, dynamics, and overall recruitment to the inactive chromosome.

5.2. Results

5.2.1. Generation and characterization of *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ mESC lines with an inducible Xist transgene

To explore the role of RYBP and its homologue YAF2 in Xist-mediated PRC1 recruitment, I generated a double knockout mESC line. A single annotated coding transcript is indicated for both *Rybp* and *Yaf2* genes in the mouse genome. Thus, I decided to design one guide RNA (gRNA) for CRISPR/Cas9 nuclease mediated cleavage to induce deletions in close proximity to the translation start site of each gene (Figure 20A). I generated these deletions in the parental cell line 3E-H (used in chapter 4) that includes an inducible Xist transgene stably integrated on chromosome 17. After screening several clones by PCR around the expected cleavage site (red arrow in Figure 20A) followed by Sanger sequencing of the PCR products, I selected a double homozygous mutant clone (Figure 20A). For *Rybp*, deletion of 33bp results in a frameshift mutation that alters the protein sequence after amino acid 5. Knockout of *Rybp* was confirmed to disrupt protein expression by Western blot (Figure 20B). For *Yaf2*, a 13bp deletion (Figure 20A) results in a frameshift mutation that changes the protein sequence after amino acid 15. This knockout could not be confirmed on the protein level, as antibodies available for mouse YAF2 protein were not obtainable. The selected mESC clone grew somewhat slower than the parental cell line, but nevertheless retained apparent mESC morphology.

Rybp/*Yaf2* double knockout cells (*Rybp*$^{-/-}$*Yaf2*$^{-/-}$) showed no significant changes in the global levels of H2AK119u1 or H3K27me3, nor of PRC1 (RING1B, CBX7) or PRC2 (EZH2, SUZ12) subunits as evaluated by western blot. This suggests that the stability of other Polycomb subunits does not depend on the presence of RYBP/YAF2 (Figure 20B).

To address if absence of RYBP/YAF2 has any impact on Xist-dependent Polycomb recruitment, I performed immunofluorescence for H2AK119u1 and H3K27me3 at different timepoints after Xist induction by doxycycline (6h, 12h, 24h) in wild-type and *Rybp*$^{-/-}$*Yaf2*$^{-/-}$ cells (Figure 21). Scoring of the percentage of cells showing accumulation of these marks at different timepoints showed a delay in the accumulation of both H2AK119u1 and more prominently H3K27me3 over the inactive chromosome domain in double knockout when compared to wild-type mESCs (Figure 21). Visual inspection of the domains in *Rybp*$^{-/-}$*Yaf2*$^{-/-}$ also suggest that these are in general smaller/less bright when compared to nucleoplasm levels than in wild-type cells (data not shown). Both observations indicate that even though RYBP/YAF2 is not necessary to initiate Xist-mediated Polycomb recruitment, they may play a role in either stabilizing the PRC1 complex or contributing to its enrichment on the inactive chromosome domain or its activity. These options will be further explored in the following sections.

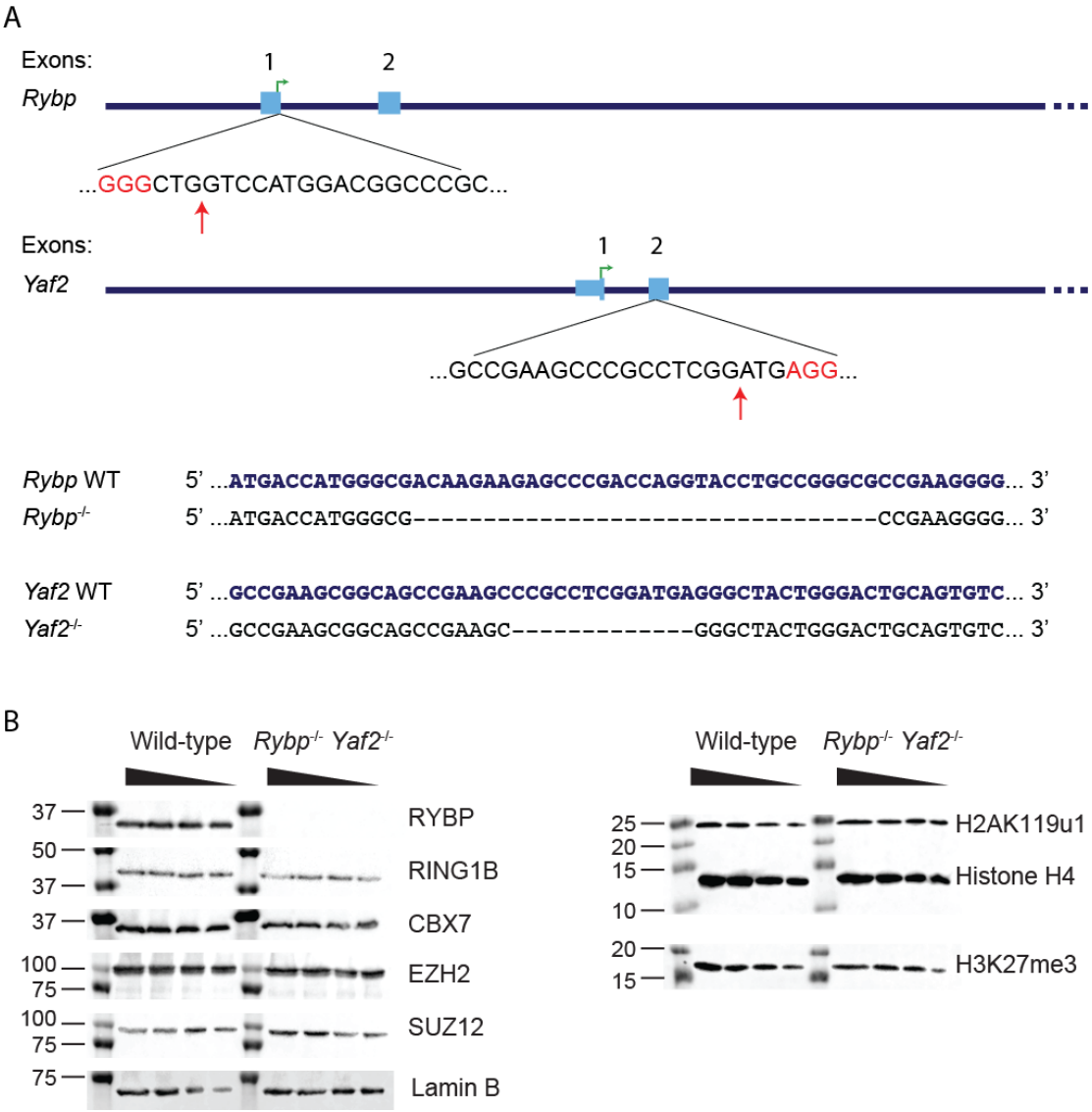


Figure 20: Establishment and characterization of *Rybp*$^{-/-}$*Yaf2*$^{-/-}$ mESC line. (A) Strategy for CRISPR-Cas9 mediated mutagenesis of *Rybp* and *Yaf2* genes. The position of guide RNAs relative to *Rybp* and *Yaf2* is shown, with PAM sequences (red font) and predicted cleavage sites (red arrows) indicated. An alignment of wild-type sequence with frameshift deletions of the clone used in this study is shown below. (B) Western blot analysis of the PRC1 subunits RYBP, RING1B, and CBX7, the PRC2 subunits EZH2 and SUZ12, and the histone modifications H2AK119u1 and H3K27me3 in *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ compared to wild-type mESCs. LAMIN B and histone H4 western blots are included as loading controls. Volumes loaded were 14µl, 10µl, 8µl and 6µl (whole cell extract) or 10µL, 7µL, 5µL and 2µL (histone extract). Molecular weight markers are shown on the left.

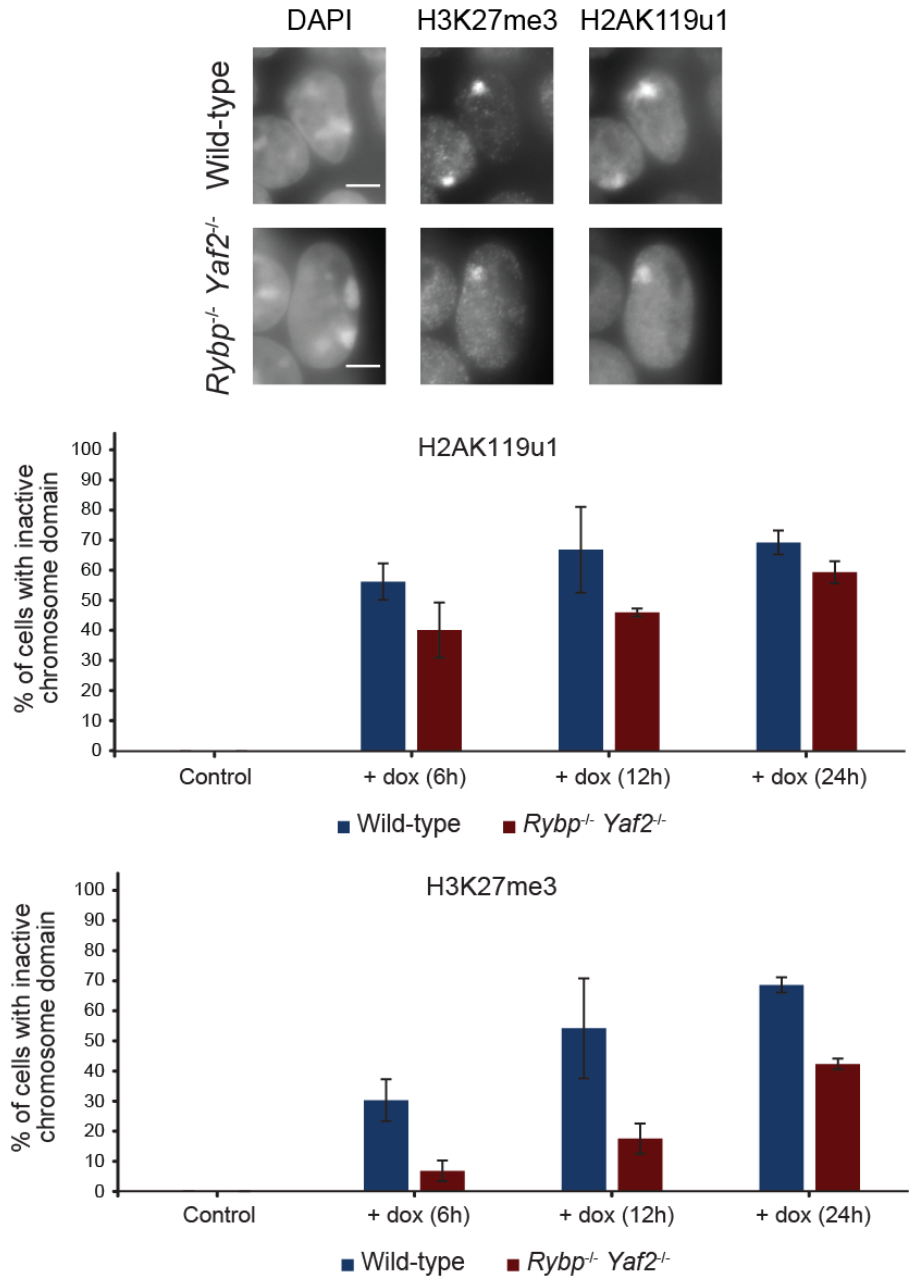


Figure 21: Immunofluorescence analysis of Polycomb marks on the inactive chromosome domain over a Xist induction timecourse in *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ compared to wild-type mESCs. Immunostaining for H2AK119u1 and H3K27me3 marks deposited by Polycomb complexes after Xist induction by doxycycline (+dox) for 6, 12 and 24 hours. Pictures represent cells after 24h of Xist induction. Scale bar is 5µm. Bar chart represents percentage of cells showing H2AK119u1 (top) or H3K27me3 (bottom) domains for each timepoint after Xist induction (+dox) in *Rybp*$^{-/-}$*Yaf2*$^{-/-}$ mESC line compared to wild-type. Scoring data was obtained from a minimum of 300 cells per timepoint in three replicates. Error bars show standard deviation.

5.2.2. Recruitment of non-canonical PRC1 complexes to the inactive chromosome domain in the absence of RYBP/YAF2

The observed delay in the accumulation of H2AK119u1 after Xist expression suggests that RYBP/YAF2 might have a function in this context. RYBP/YAF2 might stabilize the interaction between other non-canonical PRC1 components or contribute to recruitment of PRC1 complexes. To test these hypotheses, I decided to express eGFP-PCGF fusion proteins described in chapter 4 in the *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ mESC line.

A co-immunoprecipitation assay was used to assess if the eGFP-PCGF fusions retained their interaction with RING1B in the absence of RYBP/YAF2. All eGFP-PCGF were shown to interact with RING1B, creating a minimal catalytic PRC1 complex without RYBP/YAF2 (Figure 22).

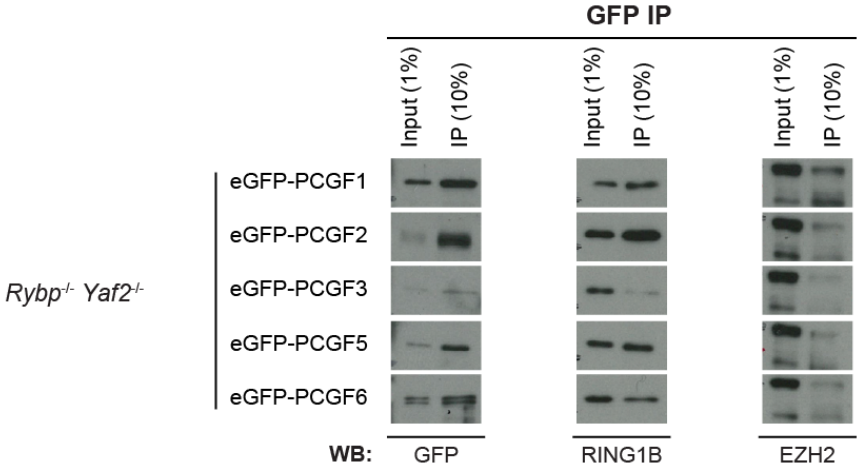


Figure 22: Interaction of eGFP-PCGF proteins with endogenous RING1B in *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ mESCs. Western blot illustrating co-immunoprecitation assay using anti-GFP antibody and nuclear extracts from stable cell lines.

To assess if loss of RYBP/YAF2 has an impact on the recruitment of any of the non-canonical PRC1 complexes, I went on to do live cell imaging in these cell lines. Analysis of the cells before Xist induction show that all eGFP-PCGF proteins localize to the nucleus, excluding the nucleolar region (Figure 23) similar to what was observed in wild-type cells (Figure 15). Nuclear speckles characteristic of eGFP-PCGF2 accumulation in Polycomb bodies are visible in *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ cells (Figure 23). However, none of the eGFP-PCGF proteins were seen enriched in a discrete focus after 24h of doxycycline-mediated Xist induction by live cell imaging (Figure 23). Immunofluorescence analysis using a GFP antibody, similar to what was performed in wild-type cells (Figure 16), also showed no enrichment of the eGFP-PCGF proteins within the H2AK119u1 enriched domain (Figure 24). Both of these observations were unexpected as H2AK119u1 deposition over the inactive chromosome domain is still observed in a large proportion of cells after 24h of Xist induction (Figure 21), demonstrating that PRC1 is still recruited to Xist-coated chromatin.

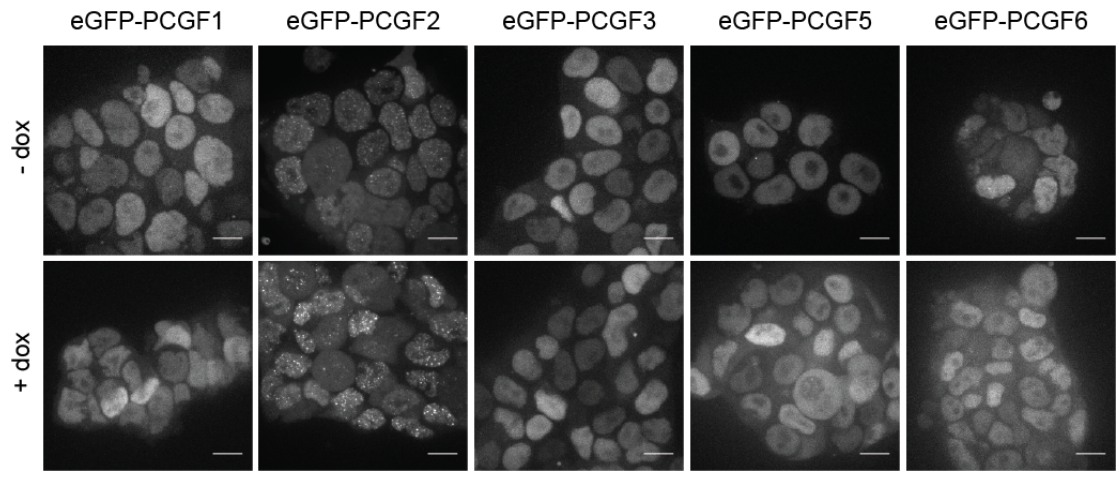


Figure 23: Xist dependent localisation of non-canonical PRC1 complexes in *Rybp*$^{-/-}$*Yaf2*$^{-/-}$ mESCs by live cell imaging. Live imaging of mESCs illustrating absence of focal enrichment of eGFP-PCGF fusion proteins following induction of Xist RNA for 24h in *Rybp*$^{-/-}$*Yaf2*$^{-/-}$ mESC line (+dox). Images are maximum intensity projections of six consecutive z-stacks. Scale bar is 10µm.

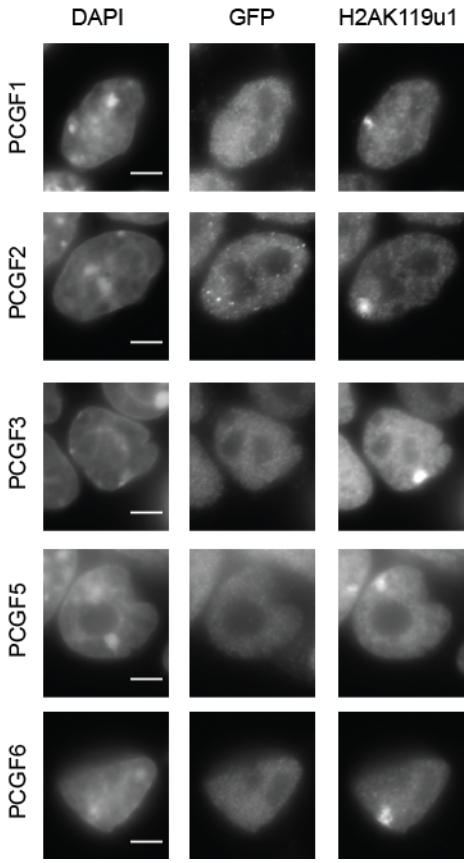


Figure 24: Immunofluorescence analysis of *Rybp*$^{-/-}$*Yaf2*$^{-/-}$ mESCs expressing eGFP-PCGF fusion proteins. Images show representative single cells stained for GFP and H2AK119u1 after 24h of Xist induction. Scale bar is 5µm.

Since RING1B is the catalytic subunit responsible for H2AK119u1 deposition, I decided to test if this protein could be seen enriched over the inactive chromosome region in *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ mESCs. I therefore expressed eGFP-RING1B in *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ cells and performed live cell imaging 24h after Xist induction. No discrete focus resembling the inactive chromosome domain was visible in these cells, although similar to eGFP-PCGF2, eGFP-RING1B still accumulated in small nuclear speckles (Figure 25) as in wild-type cells (Figure 14B). In summary, these results show that H2AK119u1 deposition occurs in *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ mESCs, therefore demonstrating that RYBP and YAF2 are not

necessary to initiate Xist-mediated Polycomb recruitment. However, in the absence of RYBP and YAF2, the ability to detect the other subunits of the complex on the inactive chromosome domain is lost, revealing the importance of RYBP/YAF2 in PRC1 enrichment.

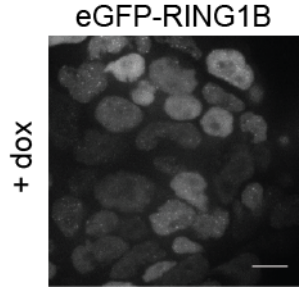


Figure 25: Xist dependent localisation of eGFP-RING1B in *Rybp*$^{-/-}$*Yaf2*$^{-/-}$ mESCs by live cell imaging. Live imaging of mESCs illustrating no focal enrichment of eGFP-RING1B fusion protein following induction of Xist RNA for 24h in *Rybp*$^{-/-}$*Yaf2*$^{-/-}$ mESC line (+dox). Image is a maximum intensity projection of six consecutive z-stacks. Scale bar is 10μm.

5.2.3. Dynamics of RYBP in the nucleus and the inactive chromosome domain in wild-type cells

To understand how RYBP contributes to the enrichment of non-canonical PRC1 complexes, I decided to characterize the dynamics of this protein by taking FRAP measurements in the wild-type cell line generated as a control in Chapter 4 (3E-H mESC lines expressing eGFP-RYBP).

Analysis of average fluorescence recovery curves showed that eGFP-RYBP has a very fast turnover rate both in the nucleoplasm and over the inactive chromosome domain (Figures 26A and 26B), albeit slightly slower in the latter

(Figure 26B). In both cases, there is a very small eGFP-RYBP immobile fraction on the timescale of this experiment, corresponding to less than 3% (median) in the inactive chromosome domain, and to a negligible proportion in the nucleoplasm.

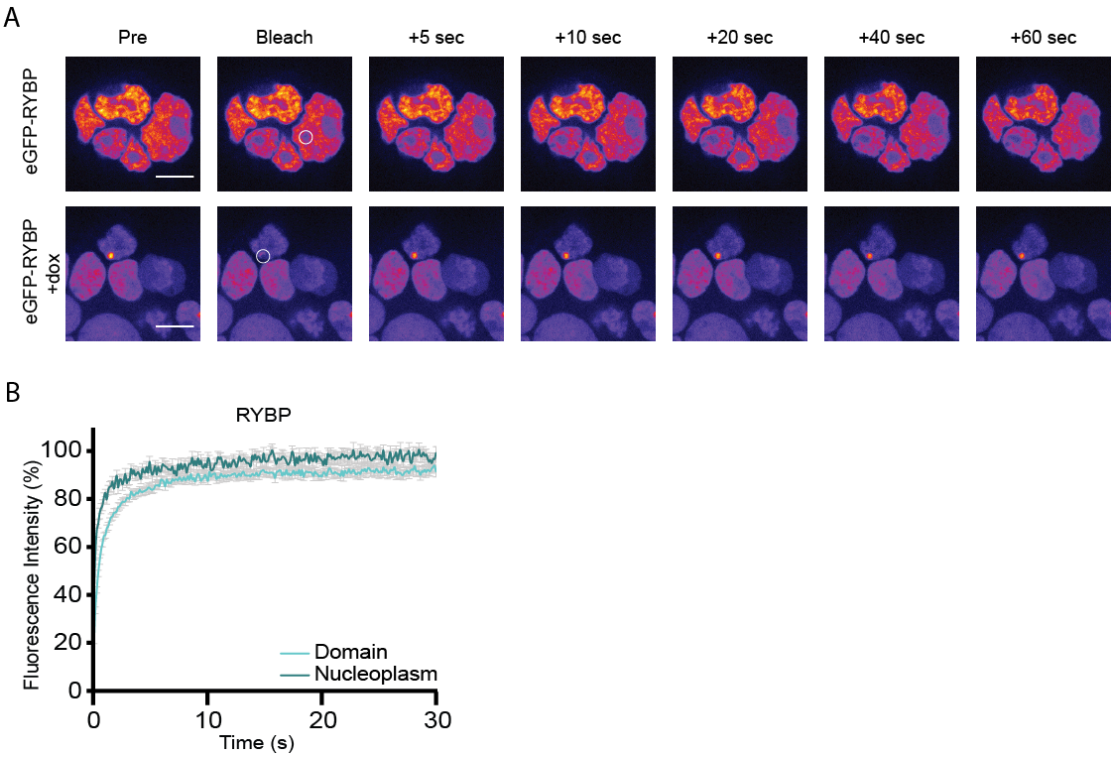


Figure 26: eGFP-RYBP shows very fast turnover in the nucleoplasm and the inactive chromosome domain. (A) Images are representative of a FRAP experiment for 3E-H cell line expressing eGFP-RYBP. Fluorescence on the bleached region (indicated by a circle on bleach snapshot) recovers completely in the nucleoplasm (top row) and almost completely over the inactive chromosome domain (+dox, bottom row). Scale bar is 10µm. (B) Average fluorescence recovery curves for eGFP-RYBP protein over the first 30 seconds after photobleaching show a very similar recovery over the inactive chromosome domain and in the nucleoplasm. A minimum of 31 cells was analyzed for each region of interest. Grey bars represent standard error.

Quantitative analysis of individual FRAP curves fitted to a bi-exponential model show that eGFP-RYBP has very distinct dynamics when compared to the eGFP-PCGF proteins, both in the nucleoplasm as well as in the inactive chromosome domain (Figures 27A and 27B). In the nucleoplasm, eGFP-RYBP has virtually no immobile fraction in the timescale of the experiment as full turnover of the protein occurs within the first 15 seconds after photobleaching (Figures 27B and 27C). Although eGFP-RYBP has a slightly slower turnover over the inactive chromosome domain, this behavior is still very distinct from the eGFP-PCGF proteins ($p<0.001$ for all pairwise comparisons between eGFP-RYBP and all eGFP-PCGF proteins $t_{1/2}$ both in the inactive domain as well as on the nucleoplasm) (Figures 27A and 27D).

This result, together with the fact that RYBP is required for enrichment of all non-canonical PRC1 complexes on the inactive chromosome domain, leads me to hypothesize that RYBP mainly contributes to the fast recovery fraction of each non-canonical PRC1 complex in a wild-type background. This might explain why in *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ mESCs none of the eGFP-PCGF proteins are visible over the inactive chromosome domain, since as shown in the previous chapter the bulk of each protein in the domain is represented by its mobile fraction.

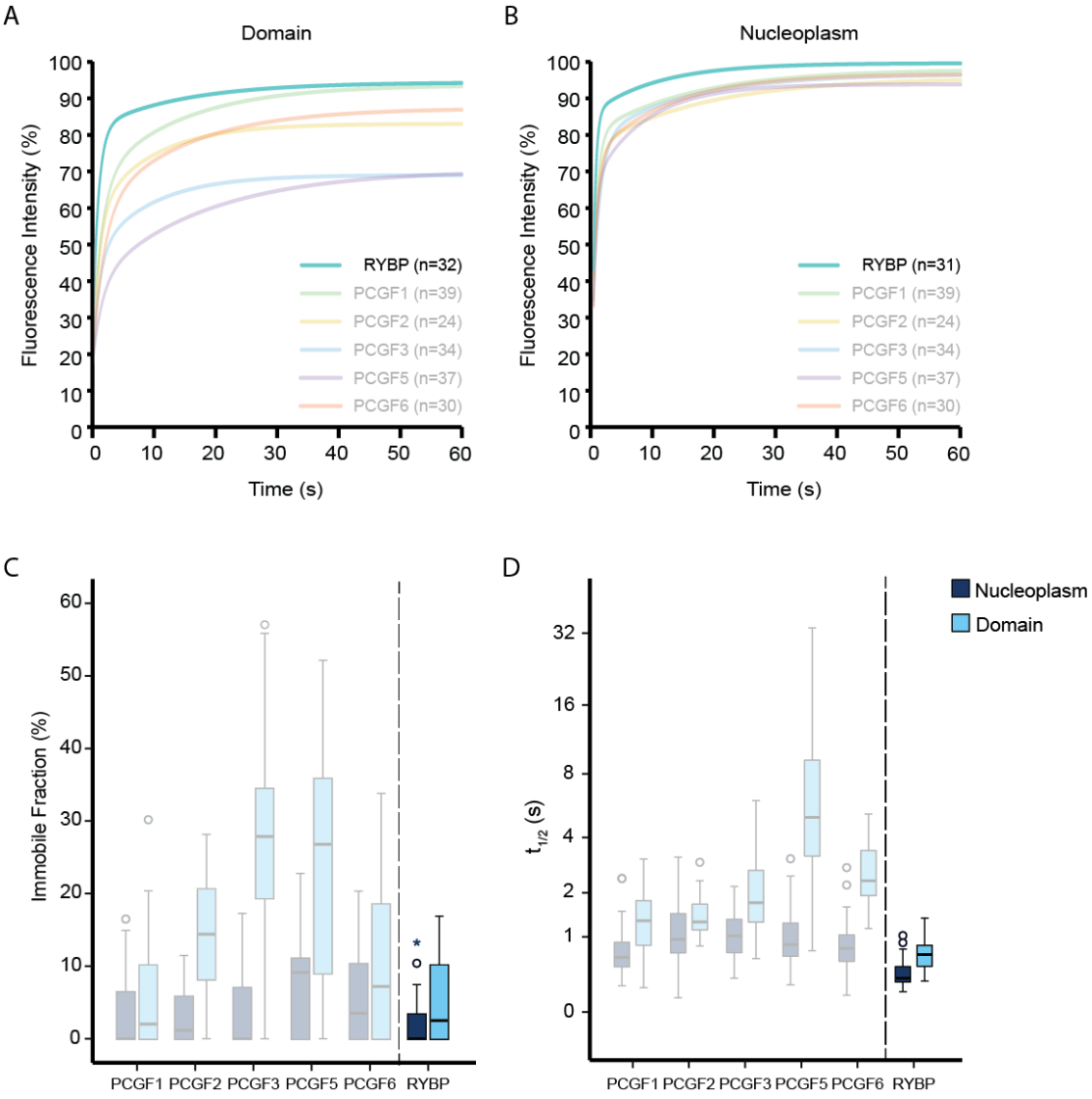


Figure 27: Quantitative analysis of FRAP data for eGFP-RYBP using a bi-exponential function. Fluorescence recovery data for eGFP-RYBP in single cells was fitted to a bi-exponential model for quantitative analysis. Average fitted curves for eGFP-RYBP (A) on the inactive chromosome domain and (B) on the nucleoplasm are shown. eGFP-PCGF data is shown in faded colors for comparison. Number of cells analyzed is showed in brackets (n). (C) Distribution of the proportion of eGFP-RYBP in the immobile fraction 60 seconds after photobleaching calculated for nucleoplasm and Xist RNA domains. Boxplot shows data for eGFP-PCGF proteins (faded colors) for comparison. (D) Distribution of half-time ($t_{1/2}$) recovery for eGFP-RYBP in each region of interest is represented in a boxplot for comparison with eGFP-PCGF proteins (faded colors). $t_{1/2}$ is shown in seconds and represented on a $\log_2$ scale.

5.2.4. Visualization of non-canonical PRC1 complexes present on the inactive chromosome domain in the absence of RYBP/YAF2

Given that eGFP-PCGF fusion proteins enrichment on inactive chromosome domains is undetectable in *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ mESCs, but deposition of H2AK119u1 is nevertheless evident, I spectulate that an increased pool of non-chromatin bound PRC1 complexes in the nucleoplasm might mask inactive chromosome localization. To test if I could capture any of the eGFP-PCGF proteins overlapping with H2AK119u1 I used a protocol to deplete the non-chromatin bound proteins from the nucleus. Specifically I performed immunofluorescence on cells which were permeabilized prior to fixation. Using this approach, H2AK119u1 domains characteristic of Xist foci were still visible and eGFP-PCGF proteins were detected in the nucleus in all cases (Figure 28). Importantly, in a proportion of cells, eGFP-PCGF3 and eGFP-PCGF5 were seen overlapping with H2AK119u1 on the inactive chromosome (Figure 28). Replicates of this experiment gave variable results, probably due to differences in permeabilization and fixation between experiments performed on different days, and consequently average scoring data is not shown. Accordingly, although only eGFP-PCGF3 and eGFP-PCGF5 were seen accumulating on the inactive chromosome domain in *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ cells, I cannot rule out the possibility that PCGF1, PCGF2 and PCGF6 might accumulate to some degree, as immunofluorescence was quite inefficient at capturing their enrichment even in a wild-type cell line (Figure 16).

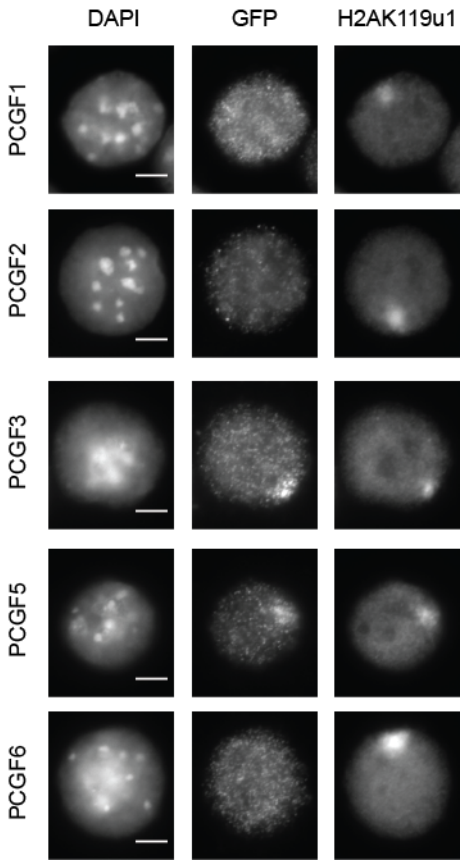


Figure 28: Immunofluorescence analysis of *Rybp*$^{-/-}$*Yaf2*$^{-/-}$ mESCs expressing eGFP-PCGF fusion proteins using perm/fix protocol. Images show representative single cells stained for GFP and H2AK119u1 after 24h of Xist induction using a protocol in which permeabilization precedes fixation to remove the non-chromatin bound fraction of protein from the nucleus. Scale bar is 5µm.

5.2.5. Mechanism of RYBP-dependent recruitment of non-canonical PRC1 complexes

In previous sections I showed that even though RYBP/YAF2 are not necessary for initiation of Xist-mediated Polycomb recruitment, these proteins contribute significantly to PRC1 complex enrichment over the inactive chromosome domain. Therefore, rescuing the expression of RYBP in the

Rybp$^{-/-}$ *Yaf2*$^{-/-}$ cells should revert the phenotype to a wild-type scenario in which eGFP-PCGF proteins accumulate in a discrete focus in the nucleus upon Xist expression. To test this, I generated stable cell lines expressing RYBP tagged with mCherry fluorescent protein on its N-terminus (mCherry-RYBP) in the *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ mESCs expressing different eGFP-PCGF proteins.

Co-immunoprecipitation analysis of eGFP-PCGF interaction with RING1B showed that this interaction remains intact upon expression of mCherry-RYBP in *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ mESCs (Figure 29A).

Live cell imaging of these clones 24 hours after induction of Xist expression showed that mCherry-RYBP is enriched over the inactive chromosome domain (Figure 29B, bottom row), similar to eGFP-RYBP expressed in wild-type cells (Figure 14B). Importantly, expression of mCherry-RYBP in *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ cells leads to enrichment of all five eGFP-PCGF proteins over a discrete focus in the nucleus (Figure 29B, top row), resembling Xist-induced accumulation of these fusion proteins seen in wild-type cells (Figure 15). This result confirms that RYBP contributes to the accumulation of a large fraction of each eGFP-PCGF protein over the inactive chromosome, hence allowing their visualization by live cell imaging.

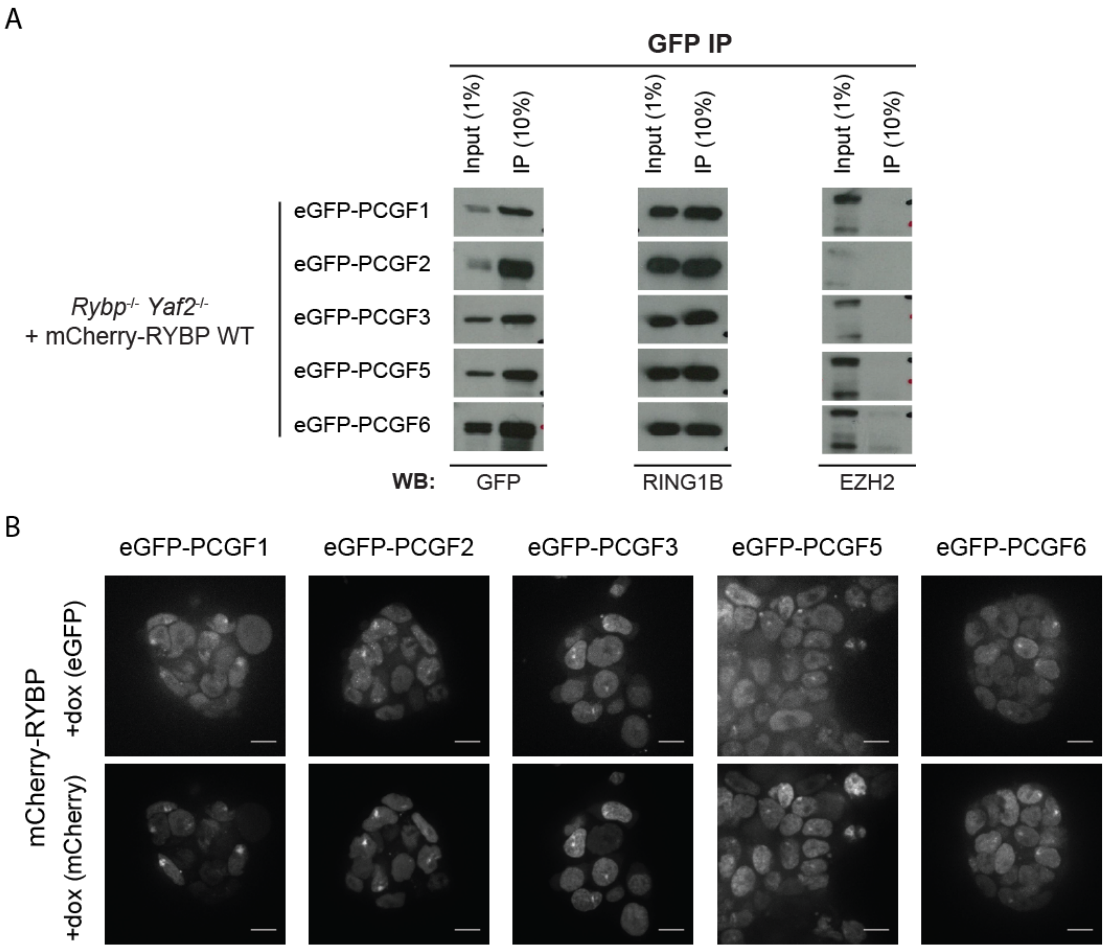


Figure 29: Expression of mCherry-RYBP in *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ mESCs rescues enrichment of eGFP-PCGF proteins on the inactive chromosome domain. (A) Western blot illustrating co-immunoprecitation of endogenous RING1B with eGFP-PCGF fusion proteins using anti-GFP antibody and nuclear extracts from stable *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ mESC lines expressing mCherry-RYBP. (B) Live imaging of mESCs illustrating focal enrichment of mCherry-RYBP and eGFP-PCGF fusion proteins following induction of Xist RNA for 24h in *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ mESC line (+dox). Images are maximum intensity projections of six consecutive z-stacks. Scale bar is 10µm.

Protein alignments between the homologs RYBP/YAF2, mutually exclusive components of the non-canonical PRC1 complexes (García et al. 1999; Wang et al. 2010), showed that the conservation between the two proteins is only

54% in mouse (Figure 30A). The two proteins show a highly conserved region around amino acid positions 161 to 173 of RYBP, which has been shown to be necessary to mediate its interaction with the PRC1 catalytic subunits RING1A and RING1B (García et al. 1999). Another region in the N-terminal portion of RYBP spanning from amino acid 19 to 50 (green box in Figure 30A) was previously shown to be highly conserved across species and to encode for a zinc finger domain with ubiquitin binding activity (NZF-UBD) (Arrigoni et al. 2006). This domain in RYBP was shown to be able to recognize H2AK119u1, and this recognition could be disrupted by a double alanine substitution (TF-AA mutant; TF residues indicated by red arrows in Figure 30A) in the NZF-UBD (Arrigoni et al. 2006). This region is also highly conserved in YAF2, suggesting that both proteins should have the ability to recognize ubiquitin.

To test if H2AK119u1 recognition by RYBP is the mechanism that promotes its enrichment on the inactive chromosome and consequent accumulation of eGFP-PCGF proteins, I stably expressed an N-terminus mCherry-tagged RYBP TF-AA mutant version (mCherry-RYBP TF-AA) in *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ mESCs previously expressing different eGFP-PCGF proteins.

Co-immunoprecipitation analysis shows that expression of mCherry-RYBP TF-AA in *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ mESCs does not interfere with eGFP-PCGF protein interaction with RING1B (Figure 30B).

Analysis of stable clones expressing mCherry-RYBP TF-AA mutant showed that the mutant version of the protein is unable to accumulate on the inactive chromosome to the extent that allows visualization in live cells (Figure 30C, bottom row). Consequently, enrichment of the eGFP-PCGF proteins over the inactive chromosome was not rescued in *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ mESCs expressing mCherry-RYBP TF-AA mutant (Figure 30C, top row). This result suggests that recognition of previously deposited H2AK119u1 by RYBP is required for recruitment of a large fraction of the PRC1 that accumulates on the inactive chromosome.

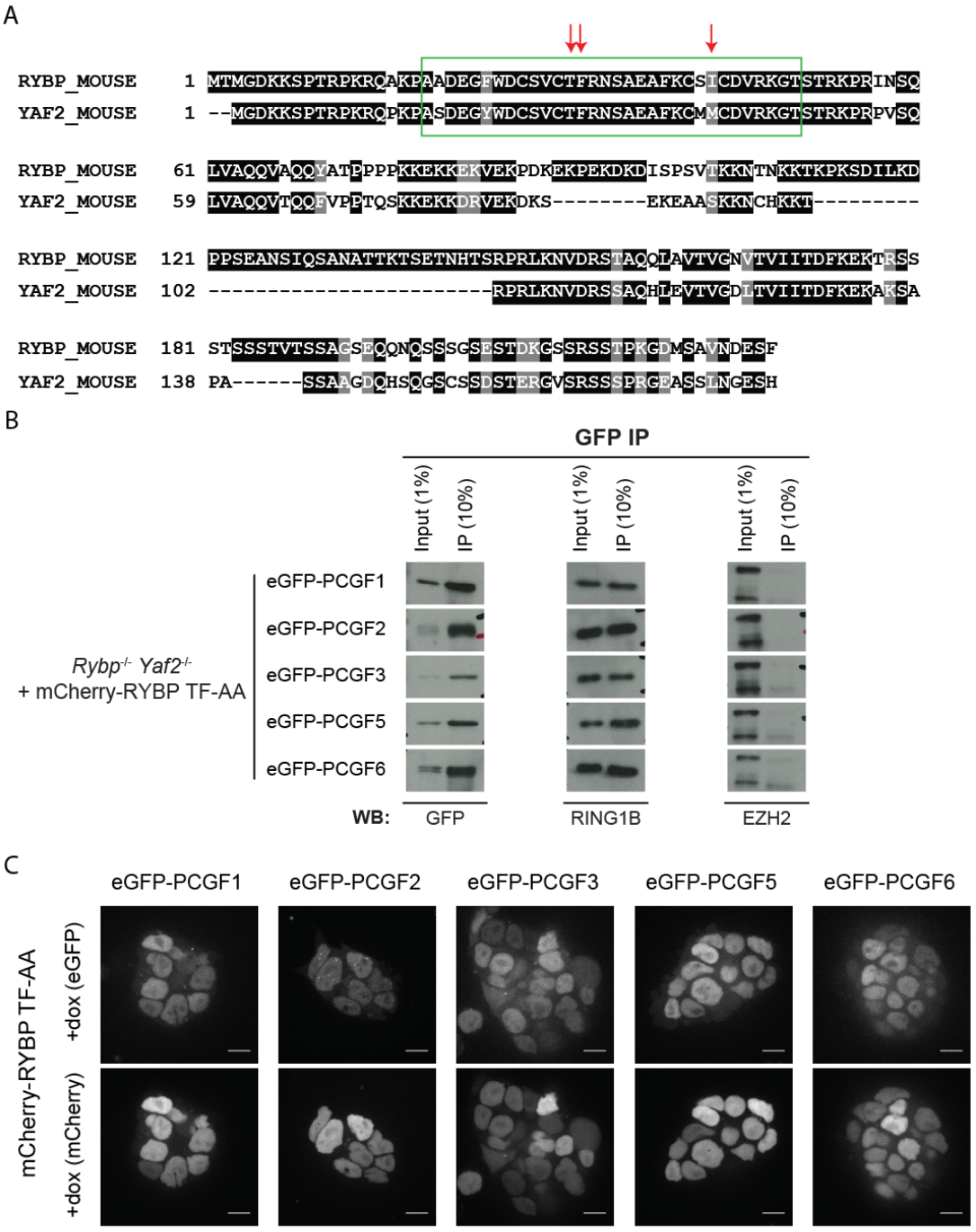


Figure 30: Expression of mCherry-RYBP TF-AA mutant in *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ mESCs does not rescue enrichment of eGFP-PCGF proteins on the inactive chromosome domain. (A) Protein alignments for mouse RYBP and YAF2 proteins showing conservation of a NZF-UBD (green box). Red arrows indicate the critical residues necessary for ubiquitin binding. (B) Western blot illustrating co-immunoprecitation of endogenous RING1B with eGFP-PCGF fusion proteins using anti-GFP antibody and nuclear extracts from stable *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ cell lines expressing mCherry-RYBP TF-AA. (B) Live imaging of

mESCs illustrating no focal enrichment of mCherry-RYBP TF-AA mutant and eGFP-PCGF fusion proteins following induction of Xist RNA for 24h in *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ mESC line (+dox). Images are maximum intensity projections of six consecutive z-stacks. Scale bar is 10µm.

5.3. Discussion

In this chapter, I generated *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ mESCs to investigate the potential role of these proteins on non-canonical PRC1 recruitment mediated by Xist RNA. Importantly, depletion of RYBP and YAF2 did not cause any morphological changes in this mESC line. This is in agreement with what was previously reported for *Rybp*$^{-/-}$ mESCs (Hisada et al. 2012). Analysis of global levels of Polycomb proteins and the repressive histone marks H3K27me3 and H2AK119u1 by western blot showed that no dramatic changes occur in *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ when compared to wild-type mESCs. This is in agreement with recent observations made by Rose et al. (2016) in an independently generated double knockout mESC line (Rose et al. 2016).

Absence of RYBP/YAF2 causes a delay in deposition of Polycomb associated marks H2AK119u1 and H3K27me3 on the inactive chromosome domain, although deposition of these marks is still detectable in a large proportion of cells. The enrichment of Polycomb-mediated histone modifications on the inactive chromosome domain is paradoxical given that enrichment of PRC1 subunits (here represented by eGFP-PCGF proteins and eGFP-RING1B) could not be seen, both in live cells and in immunostained fixed cells.

Importantly, eGFP-PCGF3 and eGFP-PCGF5 could be detected using a modified immunofluorescence protocol. The detection of PCGF3 and PCGF5 enrichment suggests that these two non-canonical complexes are responsible for the visible H2AK119u1 domains in *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ mESCs. However, the role of other non-canonical complexes cannot be ruled out as PCGF1-, PCGF2- and PCGF6-containing complexes are not readily detected in fixed cells, even in a wild-type background (as discussed in Chapter 4).

A question that remains unresolved is why inactive chromosome enrichment of PRC1 subunits is not visible using live cell imaging. This could be attributable to several reasons. In the case of eGFP-PCGF1, eGFP-PCGF2 and eGFP-PCGF6, which have relatively fast dynamics, particularly on the inactive chromosome domain (as shown in Chapter 4), recognition of preexisting H2AK119u1 by the UBD of RYBP/YAF2 may be the solely mechanism for recruitment of these non-canonical PRC1 complexes to the inactive chromosome. The weak binding of RYBP to ubiquitin observed *in vitro* (Arrigoni et al. 2006) fits well with the transient binding of this protein to chromatin observed by FRAP analysis in mESCs as well as with the fast dynamics of the aforementioned non-canonical complexes, which recruitment to the inactive chromosome might be entirely RYBP-dependent. A similar argument can be applied to account for loss of the highly dynamic fraction fraction of eGFP-PCGF3 and eGFP-PCGF5 as they form complexes that also include the RYBP/YAF2 subunit. However, eGFP-PCGF3 and eGFP-PCGF5 also show an immobile fraction that is the largest among eGFP-PCGF proteins representing around 30% of the total protein on the inactive

chromosome domain (as shown in Chapter 4). This fraction is apparently retained on the inactive chromosome domain in the absence of RYBP/YAF2, evidenced by immunofluorescence using permeabilization before fixation, and it is therefore difficult to explain why it cannot be observed in live cell imaging experiments. A possible reason is that depletion of RYBP/YAF2 and consequently absence of this mode of PRC1 recruitment to chromatin, results in a high proportion of PRC1 complexes in the nucleus that are no longer bound to chromatin and this might mask the eGFP-PCGF proteins retained on the inactive chromosome domain, rendering these indistinguishable from the nucleoplasm in live cell imaging.

Related to the previous argument, eGFP-RING1B was also not seen enriched on the inactive chromosome domain in *Rybp*$^{-/-}$ *Yaf2*$^{-/-}$ mESCs, even though H2AK119u1 clearly accumulates there. In coherence with the previous hypothesis that RYBP-dependent recruitment accounts for the fast recovery fraction of each eGFP-PCGF protein on the inactive domain, eGFP-RING1B would only be present on Xist-coated chromatin in complex with the slow recovery fractions of each eGFP-PCGF protein that potentially have a RYBP-independent recruiting mechanism. In that case, there are two hypothetical scenarios to consider: (1) A recent publication showed that PCGF1 and PCGF6 together, account for approximately 80% of all non-canonical PRC1 in mESC (Kloet et al. 2016). Assuming that all PCGF containing complexes have an alternative RYBP-independent mechanism to be recruited to Xist-coated chromatin (this will be addressed in Chapter 6), PCGF1-PRC1 and PCGF6-PRC1 would account for the majority of non-canonical PRC1 on the

inactive domain. However, if we consider that most of these proteins on the inactive domain is represented by their fast mobile fraction, hence RYBP-dependent, the immobile fraction of these proteins on the domain is too small to contribute to detectable eGFP-RING1B enrichment over nucleoplasm levels. (2) On the other hand, PCGF3 and PCGF5-PRC1 complexes contribute to around only 15% of all non-canonical PRC1 complexes containing RING1B in mESC (Kloet et al. 2016). If only up to 30% of these PCGF3/5-PRC1 complexes (corresponding to their immobile fraction on the inactive chromosome) would be recruited independently of RYBP, that means that only a very small proportion of RING1B present in the cell would accumulate on the inactive chromosome, hence possibly rendering it indistinguishable from nucleoplasm levels. The idea that PCGF slow recovery fractions contributes only to a very small proportion of RING1B enriched on the inactive chromosome fits well with FRAP data for eGFP-RING1B that shows that this protein has a fast turnover and a small immobile fraction on the domain in wild-type cells (see Figure 47).

My observations emphasize the importance of protein dynamics in interpreting data from live cell imaging experiments and notably in methods that use formaldehyde cross-linking. The same argument also applies to the interpretation of ChIP-seq datasets. A recent study characterized RING1B and H2AK119u1 occupancy by ChIP-seq in a *Rybp/Yaf2* double knockout mESC line (Rose et al. 2016). The authors saw that depletion of these two proteins leads to reduction of H2AK119u1 at a large proportion of PRC1 target sites even though the majority of them do not lose RING1B binding

(Rose et al. 2016). This is in agreement with ChIP data in RYBP knockdown mESC from Morey *et al.* (2013). In light of the fast dynamics of RYBP on chromatin and the fact that, similarly to what happens on the inactive chromosome domain, RYBP might contribute to the recruitment of a large proportion of non-canonical PRC1 complexes genome-wide, it is reasonable to speculate that most of the RING1B occupancy that is detected in ChIP experiments is contributed by canonical PRC1 complexes, usually recruited to H3K27me3 sites. Hence, deletion of *Rybp/Yaf2* would still impact on H2AK119u1, as observed, but not on RING1B occupancy at the majority of sites.

ChIP data from studies in *Eed*$^{-/-}$ and in CBX7 knockdown mESCs (Tavares et al. 2012; Morey et al. 2013) support the idea that most of RING1B sites visualized by ChIP are contributed by canonical PRC1, as in these cell lines there is a major loss of RING1B occupancy even though H2AK119u1 levels remain unchanged at many target sites. H2AK119u1 losses in *Rybp*/*Yaf2* double knockout cells would be a result of loss of non-canonical PRC1 occupancy as this type of complex has been shown to be more catalytically active both *in vivo* and *in vitro* (Blackledge et al. 2014; Rose et al. 2016; Taherbhoy et al. 2015).

CHAPTER 6: Dissecting the role of PCGF proteins in the initiation of PRC1 recruitment to the inactive chromosome

6.1. Introduction

Previous work by Tavares *et al.* (2012) showed that non-canonical PRC1 complexes are recruited to target sites, including Xi, independently of H3K27me3 deposition. Moreover, experiments described in Chapter 3 place the focus on non-canonical PRC1 complexes as the initiators of Polycomb domain formation on the inactive chromosome. I showed that all of the non-canonical PRC1 variants are enriched over the Xist territory in mouse ES cells, and that this enrichment is dependent on RYBP/YAF2. However, Polycomb dependent deposition of H2AK119u1 does occur on the inactive chromosome domain in the absence of RYBP/YAF2, implying that other mechanisms initiate Xist-mediated recruitment. Experiments in this chapter aim to determine the role of the different non-canonical PRC1 complexes, defined by the presence of PCGF1, PCGF3, PCGF5 and PCGF6 (PCGF proteins exclusively found in non-canonical PRC1 complexes), in the initiation of Polycomb domain formation.

PCGF proteins in mammals are homologs of PSC, a core subunit of PRC1 complexes in *Drosophila*. Two other homologs were also identified in *Drosophila*, SU(Z)2 and L(3)73AH. The function of these two homologs has not been fully addressed. Interestingly, protein sequence alignments show that L(3)73AH is the most similar to mammalian PCGF proteins, with higher

homology with PCGF3 (Figure 31). Moreover, construction of a phylogenetic tree based on protein sequences show a cluster between mammalian PCGF3 and PCGF5 (Figure 31), which were shown to have a similar unique dynamic behavior on the inactive chromosome domain as analyzed by FRAP on Chapter 4. The previously described close homologs PCGF2 and PCGF4 also cluster (Figure 31), as expected.

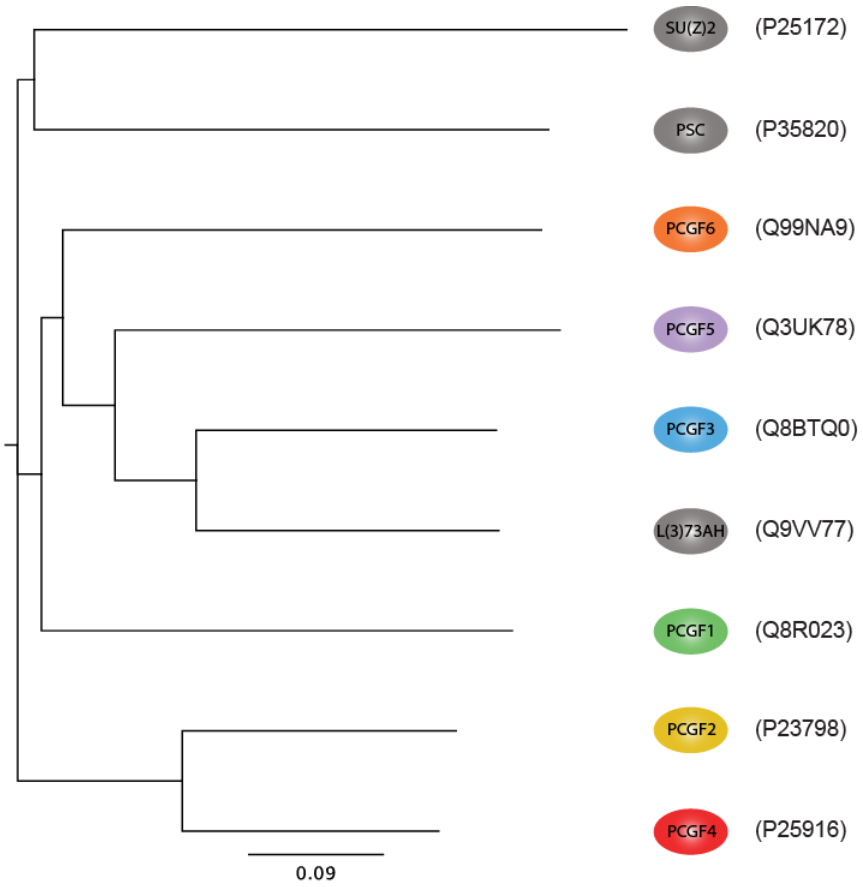


Figure 31: Neighbor Joining phylogenetic tree of mouse PCGF proteins and their *Drosophila* homologs amino acid sequences. Scale represents amino acid substitutions per site. UNIPROT ID of each protein is shown in brackets. *Drosophila* proteins are represented in grey.

Despite the sequence divergence and probable acquisition of specific functions, PCGF1-6 share two very conserved domains: the RING domain on their N-terminus and the RAWUL domain on their C-terminus (Sanchez-Pulido et al. 2008). The RING domain mediates interaction with the catalytic subunits of PRC1, RING1A and RING1B (Buchwald et al. 2006; Li et al. 2006; McGinty et al. 2014), while the RAWUL domain of different PCGF proteins has been shown to mediate interactions with other proteins in a specific manner (Alkema et al. 1997; Junco et al. 2013; Wong et al. 2016). The specificity of interactions with each PCGF protein results in the assembly of biochemically and most likely functionally distinct complexes (Gao et al. 2012). As an example, contrary to the repressive role described for Polycomb proteins, PCGF5 was recently shown to interact with AUTS2 in a transcriptional activator complex (Gao et al. 2014), providing evidence of how specific PCGF interactions can change the function of the complexes. These interactions might also play a role in recruitment to specific targets as has been shown for PCGF1 and PCGF6. PCGF1 assembles in a specific PRC1 complex that includes the proteins BCOR (a direct interactor of the RAWUL domain of PCGF1) and KDM2B. KDM2B has been shown to drive recruitment of this PRC1 variant to unmethylated CpG islands mediated by its zinc-finger CxxC-domain (Farcas et al. 2012; Wu et al. 2013; He et al. 2013). PCGF6 has been shown to interact with transcriptional repressor E2F6 through binding to L3MBTL2, depletion of which leads to H2AK119u1 reduction and increased expression of gene targets (Trojer et al. 2011). Together these findings provide evidence that PCGF proteins are important to define each non-canonical PRC1 complex in terms of function as well as recruitment.

6.2. Results

6.2.1. Generation of an inducible Xist system in *Pcgf* conditional knockout mESC to assess *de novo* Polycomb recruitment

To identify the PRC1 complex(es) that are required to initiate Xist-mediated Polycomb recruitment, I took advantage of two mESC lines with conditional null alleles for *Pcgf1* and *Pcgf3* (kind gift of Dr. H. Koseki, RIKEN, Yokohama, Japan). These cell lines are engineered with a tamoxifen inducible Cre recombinase that drives deletion of exons 2-7 of *Pcgf1* in *Pcgf1*$^{fl/fl}$ mESC (Figure 32A) and deletion of exon 4 of *Pcgf3* in *Pcgf3*$^{fl/fl}$ mESC (Figure 32B). The efficiency of deletion can be confirmed by PCR using the primers marked on the schematics in Figure 32 and as shown in the examples on the right. Both cell lines are male ES cells.

To check if *Pcgf1* or *Pcgf3* deletion would affect Xist dependent Polycomb recruitment, I engineered a doxycycline inducible transgene system similar to that described in Chapter 3. Targeting of rtTA-H to the *Rosa26* locus was performed and screening was achieved using the same strategy, here depicted in Figure 33A. Southern blots show the selected targeted clones (Figure 33A). To make the analysis easier, and to circumvent the need to use RNA-FISH for Xist detection in each cell line, I used a Xist transgene in which Bgl stem loop motifs (BglSL) were inserted after exon 7 of Xist cDNA (Figure 33B). Expressing an RNA recognition motif encoded by the BglG protein fused to mCherry (Declerck et al. 2002; Chen et al. 2009) in the same cells

allows for visualization of Xist foci using an anti-mCherry antibody in the immunofluorescence protocol (Figure 33C) (Moindrot et al. 2015).

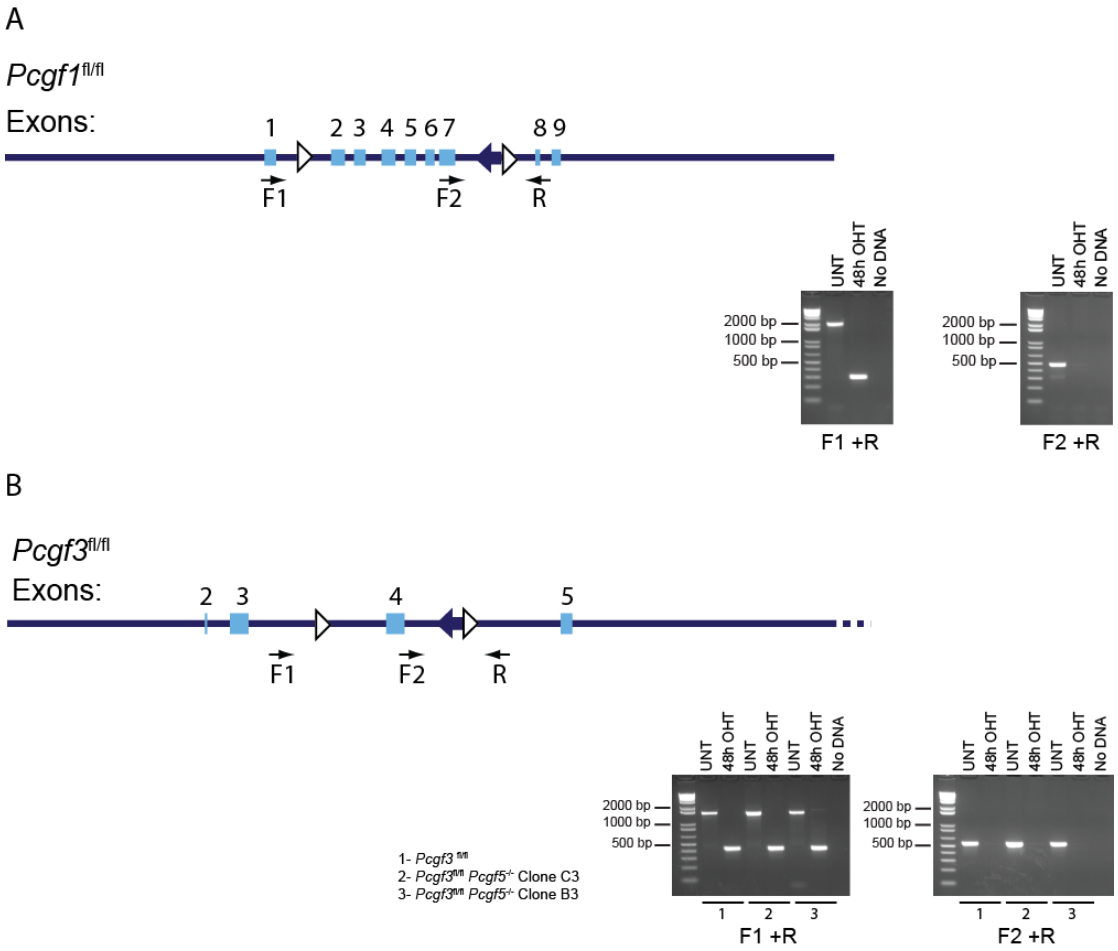


Figure 32: Characterization of *Pcgf1*$^{fl/fl}$ and *Pcgf3*$^{fl/fl}$ mESC lines. (A) Strategy for conditional null alleles of *Pcgf1* (*Pcgf1*$^{fl/fl}$). PCR analysis illustrates deletion of *Pcgf1* in *Pcgf1*$^{fl/fl}$ mouse ES cells. (B) Strategy for conditional null alleles of *Pcgf3* (*Pcgf3*$^{fl/fl}$). PCR analysis illustrates deletion of *Pcgf3* in *Pcgf3*$^{fl/fl}$ mESCs and also in two independent clones, C3 and B3 in which *Pcgf5* was first deleted by CRISPR/Cas9 mediated knockout. Schematics in (A) and (B) show intron/exon structure of the regions of interest with loxP (open triangle) and frt (filled triangles) site specific recombination recognition motifs indicated. PCR primers used to discriminate floxed and deleted alleles are indicated (F1, F2 and R).

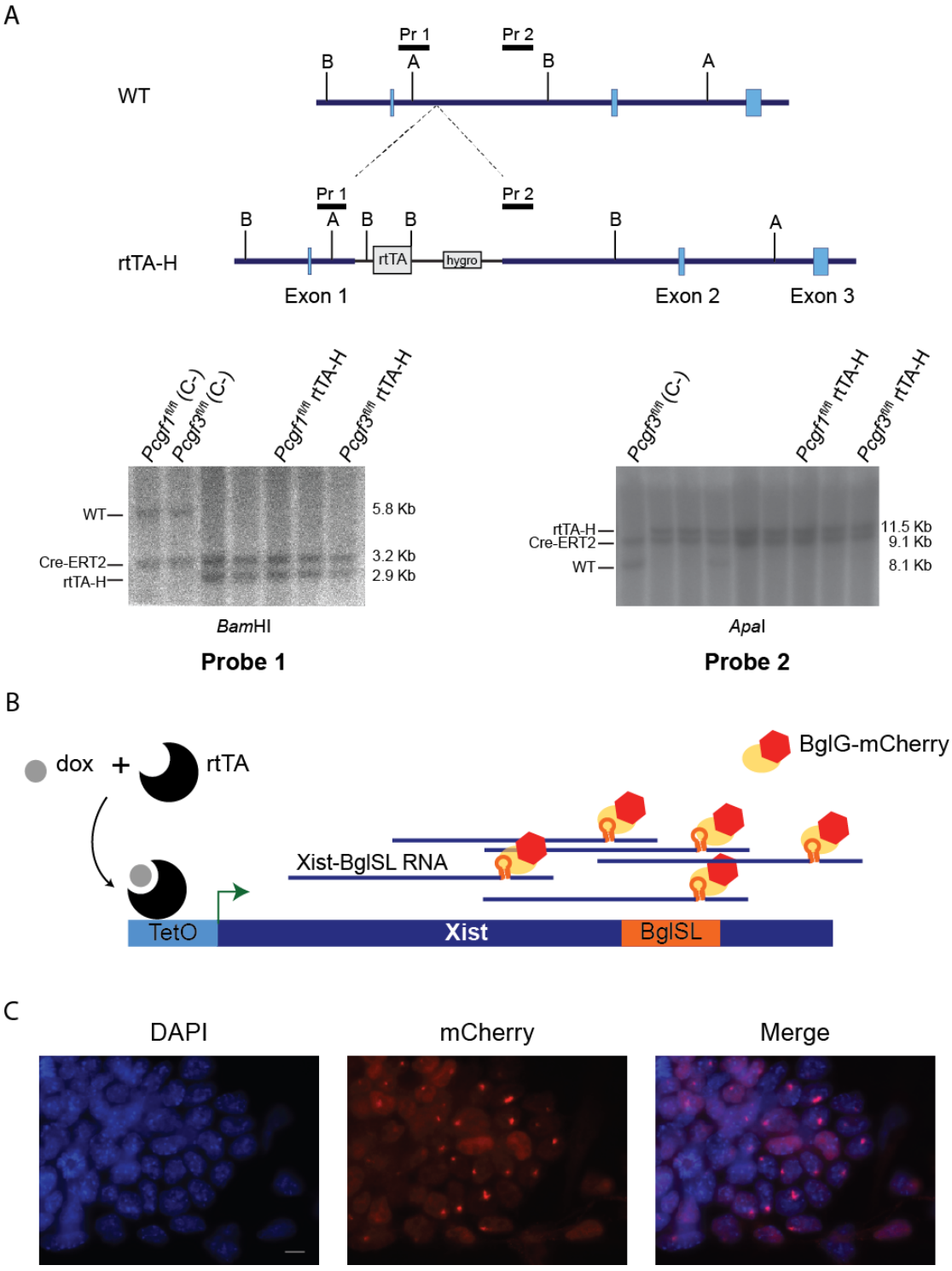


Figure 33: Engineering of an inducible Xist system encompassing BglSL (Bgl stem loops) for recognition by BglG-mCherry in conditional *Pcgf* knockout mESCs. (A) Schematic representation of the strategy used to target rtTA-H into the *Rosa26* locus and confirmation by Southern blot. Pr1 and Pr2 represent two different probes used for Southern blot using different

enzymatic digestions. B and A represent BamHI and ApaI digestion sites, respectively. Bottom panels show Southern blots using Pr1 and Pr2 to confirm targeting of rtTA-H. (B) Schematic representation of the BglG-mCherry system to recognize the inducible Xist RNA containing BglSL. Doxycycline (+dox) is added to the cells to activate TetO promoter by binding to its transactivator (rtTA). Xist RNA containing Bgl stem loops (BglSL) is expressed. These loops are recognized by BglG protein expressed with a mCherry tag allowing for (C) Xist visualization in an immunofluorescence assay. A single Xist cloud is visible in each nucleus. Scale bar is 10μm.

6.2.2. Generation of *Pcgf5* and *Pcgf6* constitutive knockout mESC to assess *de novo* Xist-mediated PRC1 recruitment

To explore the role of complexes including PCGF5 and PCGF6, I used CRISPR/Cas9-mediated mutagenesis to generate knockouts for these proteins in the *Pcgf1* or *Pcgf3* conditional knockout mESCs. All the cell lines generated are depicted in Figure 34. Establishment of single knockouts for *Pcgf5* or *Pcgf6* in the conditional *Pcgf1* or *Pcgf3* knockout mESC allowed me to assess the impact of depleting PCGF5 or PCGF6 on the establishment of H2AK119u1 domains on the inactive chromosome. Moreover, these cell lines can be used to clarify if PCGF proteins have an overlapping function in initiating Xist-mediated Polycomb recruitment by analysis of double knockouts. Importantly, several cell lines were generated for the depletion of both PCGF3 and PCGF5 as these two proteins are close homologs that have similar dynamics over the Xist domain as measured by FRAP (see Chapter 4) and can be seen localizing to the inactive chromosome domain in the absence of RYBP/YAF2 (see Chapter 5).

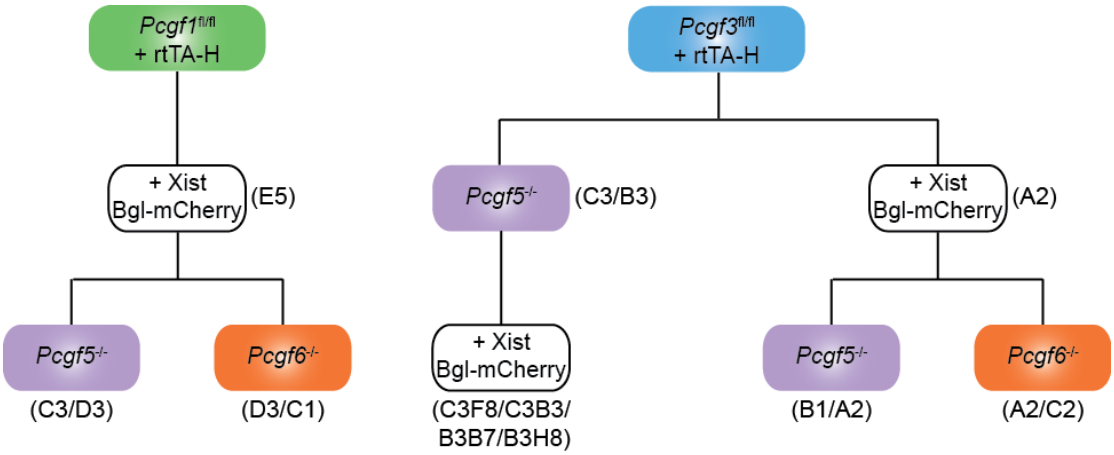


Figure 34: Schematic of mESC lines generated using CRISPR/Cas9-mediated gene editing. *Pcgf1*$^{fl/fl}$ rtTA-H and *Pcgf3*$^{fl/fl}$ rtTA-H characterized in Figures 32 and 33A were used as parental cell lines. Clones selected for each cell line are shown in brackets.

To induce *Pcgf5* mutations, I used a pair of guide RNAs to mediate cleavage of DNA by Cas9 nickase. Expected cleavage sites are indicated by red arrows in Figure 35A. These gRNAs were designed and cloned by Dr. Andrew Bassett at the Genome Engineering Oxford (GEO) facility. Mouse ES cell colonies were screened by PCR followed by Sanger sequencing of relevant PCR products. For each parental cell line, two clones were selected for further analysis (Figure 35B). To confirm that all *Pcgf5* mutants are homozygous, additional PCR characterization using a primer overlapping with the deleted region was performed in clones with deletions larger than 20bp (namely those in the *Pcgf1*$^{fl/fl}$ cell line with Xist transgene). This PCR yielded no product for any of the clones. Cas9-nickase mediated cleavage resulted in DNA insertions and deletions causing several different mutations. All mutations are

characterized in Figure 35B. All *Pcgf5* mutant clones show no apparent defects in cell growth or ES cell morphology.

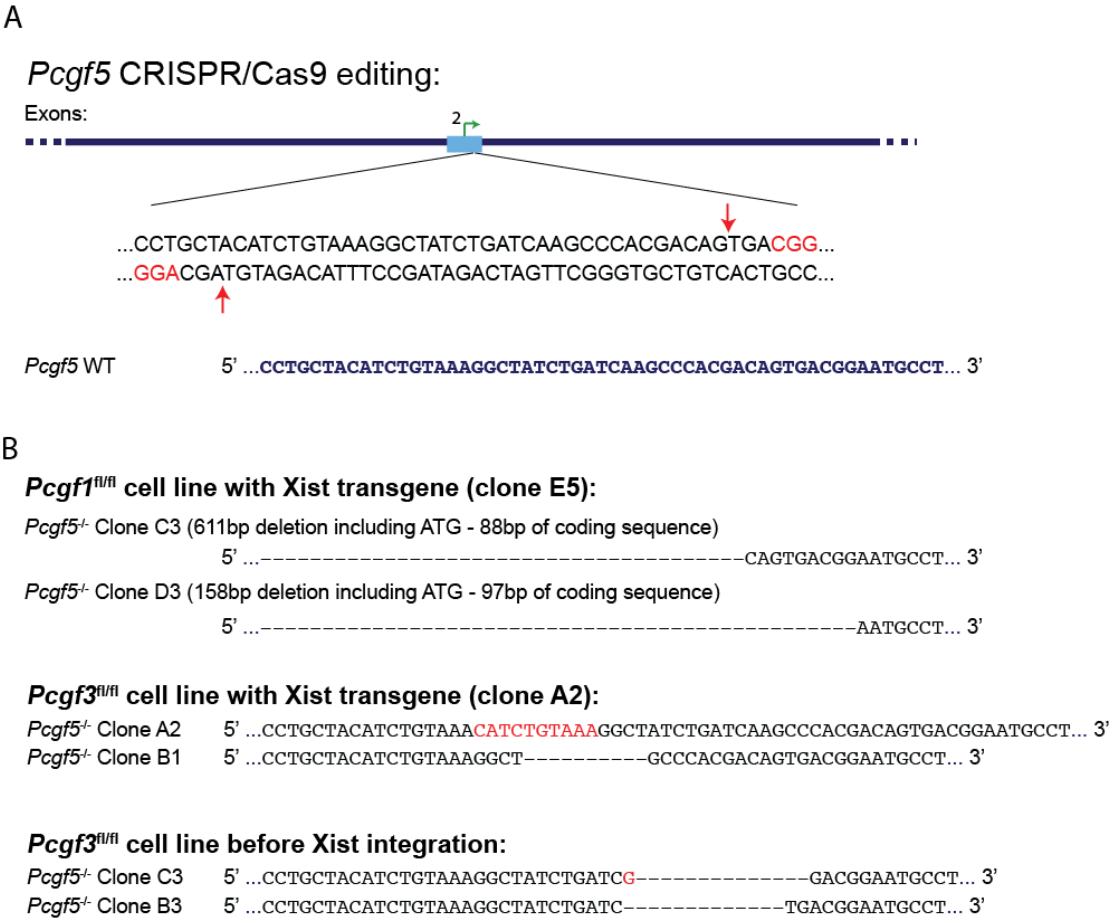


Figure 35: Strategy for CRISPR/Cas9 nickase mediated knockout of *Pcgf5*. (A) Location of guide RNAs (PAM sequences in red) and predicted cleavage sites (red arrows) are shown. Sequence of the wild-type *Pcgf5* around cleavage sites is depicted. (B) Two independent mutant mESC lines in *Pcgf1*$^{fl/fl}$ Xist E5 (C3/D3), *Pcgf3*$^{fl/fl}$ Xist A2 (A2/B1) and *Pcgf3*$^{fl/fl}$ before Xist integration (C3/B3) were selected for further analysis. Mutations were characterized by Sanger sequencing of PCR products and are shown here.

Pcgf6 mutants were generated by Cas9 nuclease mediated cleavage using a single guide RNA. A red arrow in Figure 36A indicates the expected cleavage site. ES colonies were screened by PCR around the expected cleavage site followed by Sanger sequencing of PCR products. Two mutant clones generated in each parental cell line were selected for further analysis (Figure 36B). *Pcgf6* cleavage by Cas9 nuclease generated frameshift deletions in all the selected clones as shown in Figure 36B. Three out of four had homozygous deletions. To confirm that this is the case, clones with deletions larger than 20bp were further analyzed by PCR using one primer inside of the deleted region. This was particularly relevant for the characterization of *Pcgf6*$^{-/-}$ clone C1, which despite appearing to be a homozygous mutant in my first PCR screening, was shown to be heterozygous after further investigation (Figure 36B). PCR using DNA from the other three clones yielded no product. Thorough characterization of Cas9-mediated mutations is particularly important in the case of *Pcgf5* and *Pcgf6* as no good antibodies are available to confirm depletion on the protein level in mouse. *Pcgf6* mutant clones show no apparent defects in cell growth or ES cell morphology.

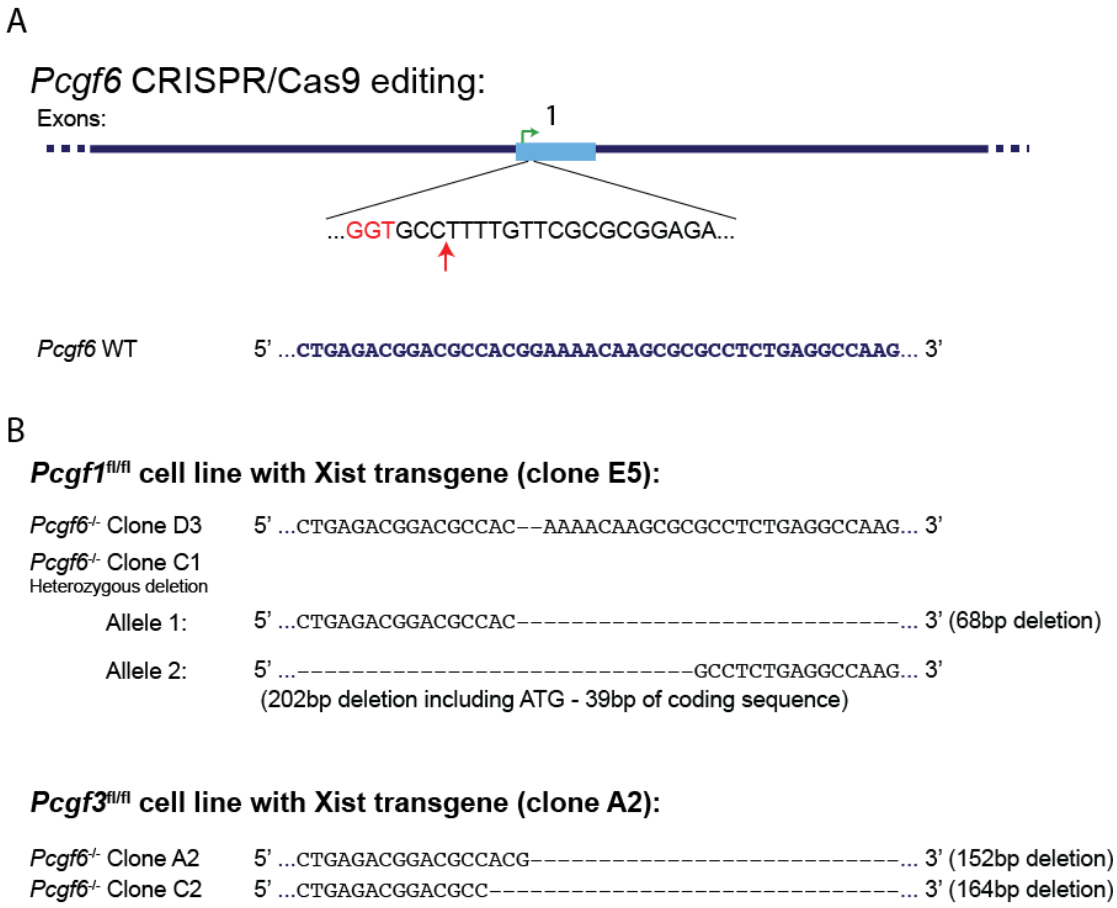


Figure 36: Strategy for CRISPR/Cas9 nuclease mediated knockout of *Pcgf6*. (A) Location of guide RNA (PAM sequences in red) and predicted cleavage sites (red arrows) are shown. Sequence of the wild-type *Pcgf6* around cleavage sites is depicted. (B) Two independent mutant mESC lines in *Pcgf1*$^{fl/fl}$ Xist E5 (D3/C1) and *Pcgf3*$^{fl/fl}$ Xist A2 (A2/C2) were selected for further analysis. Mutations were characterized by Sanger sequencing of PCR products and are shown here.

6.2.3. Analysis of single *Pcgf* knockout on Xist-mediated Polycomb recruitment

To evaluate the importance of PCGF1 and PCGF3 in the initiation of Polycomb recruitment mediated by Xist RNA, I used the experimental setup

shown in Figure 37. Adding tamoxifen (OHT) to the culture medium promotes Cre recombinase mediated deletion of the gene of interest. The deletion was initiated 24 hours before Xist RNA induction to allow for depletion of the protein of interest (PCGF1 or PCGF3) before Xist mediated Polycomb recruitment is promoted. Xist expression was then induced for 24 hours, still in the presence of tamoxifen, after which an immunofluorescence assay was performed. As the cell lines generated to assess the importance of PCGF5 and PCGF6 are constitutive knockouts, no tamoxifen treatment was required. Therefore, these cell lines were only exposed to doxycycline for 24 hours to induce Xist expression before the immunofluorescence assay. Staining of either H2AK119u1 or H3K27me3 overlapping with mCherry was quantified to measure the impact of losing each PCGF protein (PCGF1, PCGF3, PCGF5, PCGF6) on the recruitment of PRC1 and PRC2. For CRISPR/Cas9 generated knockouts (PCGF5 and PCGF6), an average of at least three replicates in two clones originating from the same parental cell line was compared to the results obtained in the parental cell line (*Pcgf1*$^{fl/fl}$ Xist clone E5 and *Pcgf3*$^{fl/fl}$ Xist clone A2) after Xist RNA expression.

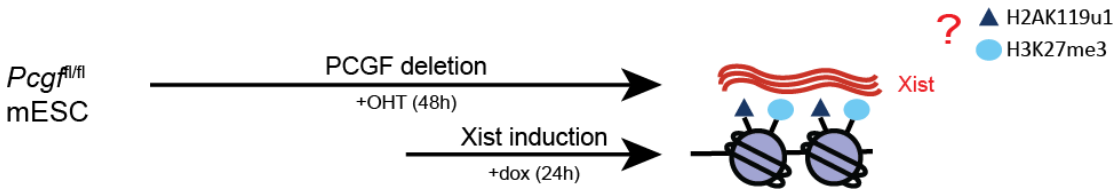


Figure 37: Schematic representation of the experiment indicating timepoints for deletion of genes encoding PCGF proteins (+OHT) in conditional knockout mESCs and Xist transgene induction (+dox). Immunofluorescence for H2AK119u1 and H3K27me3 co-stained with Xist-BglG-mCherry was used as a readout.

Deletion of *Pcgf1* seems to have no significant impact on deposition of H2AK119u1 or H3K27me3 over the Xist domains as illustrated in Figure 38A. Quantification of the number of cells that show H2AK119u1 and H3K27me3 overlapping with Xist-BglG-mCherry is unchanged when compared to control cells in which only Xist RNA was induced (+OHT+dox vs. +dox) (Figure 38B). I carried out further characterization of this cell line by Western blot. PCGF1 depletion was confirmed on the protein level. However, *Pcgf1* deletion (48h OHT) didn't cause any major changes in protein expression of other PRC1 (RYBP, RING1B, CBX7) or PRC2 (EZH2) subunits (Figure 38C, left panel). H2AK119u1 and H3K27me3 global levels were also assessed and remain apparently unchanged in PCGF1 depleted cells (48h OHT) (Figure 38C, right panel).

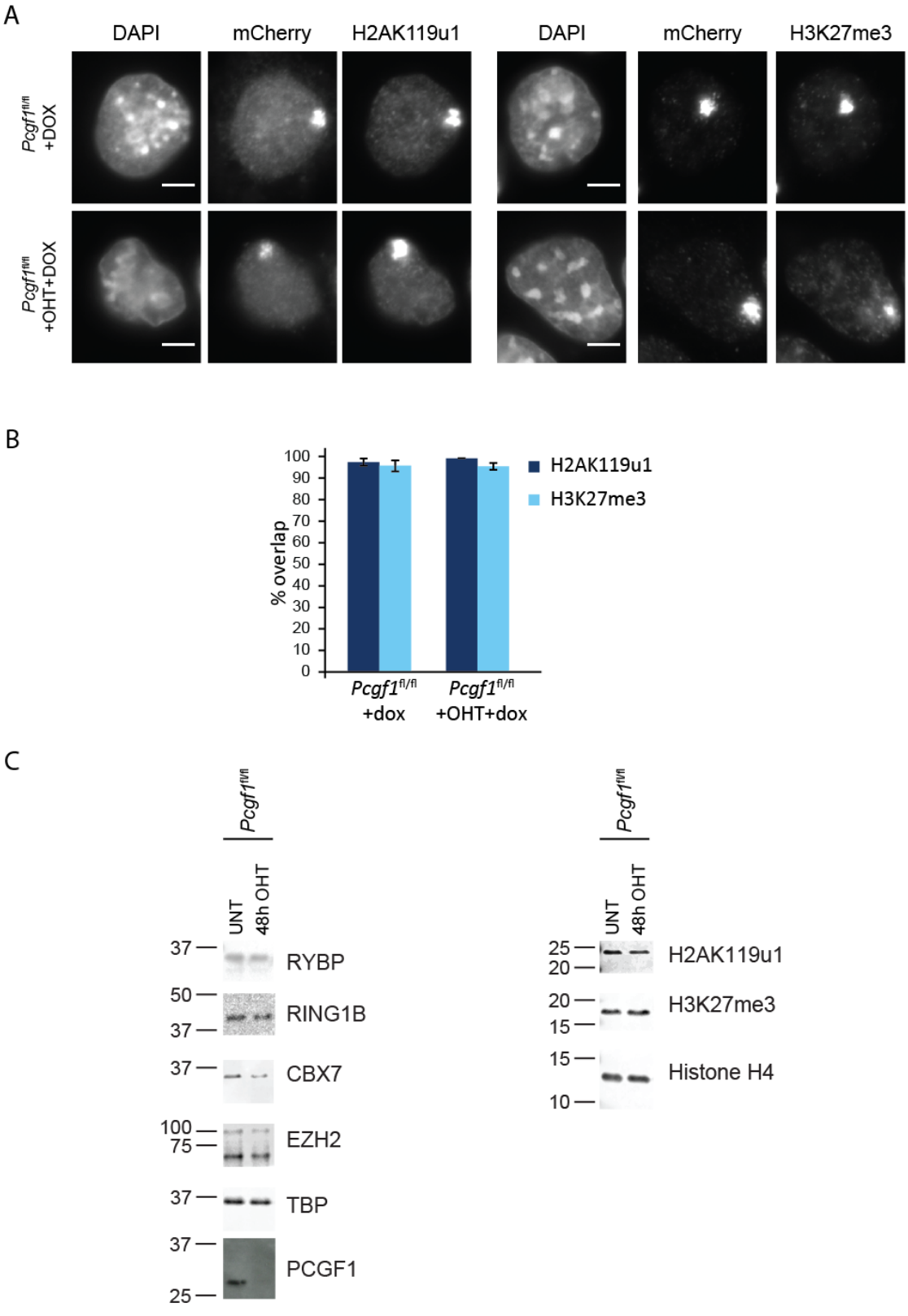
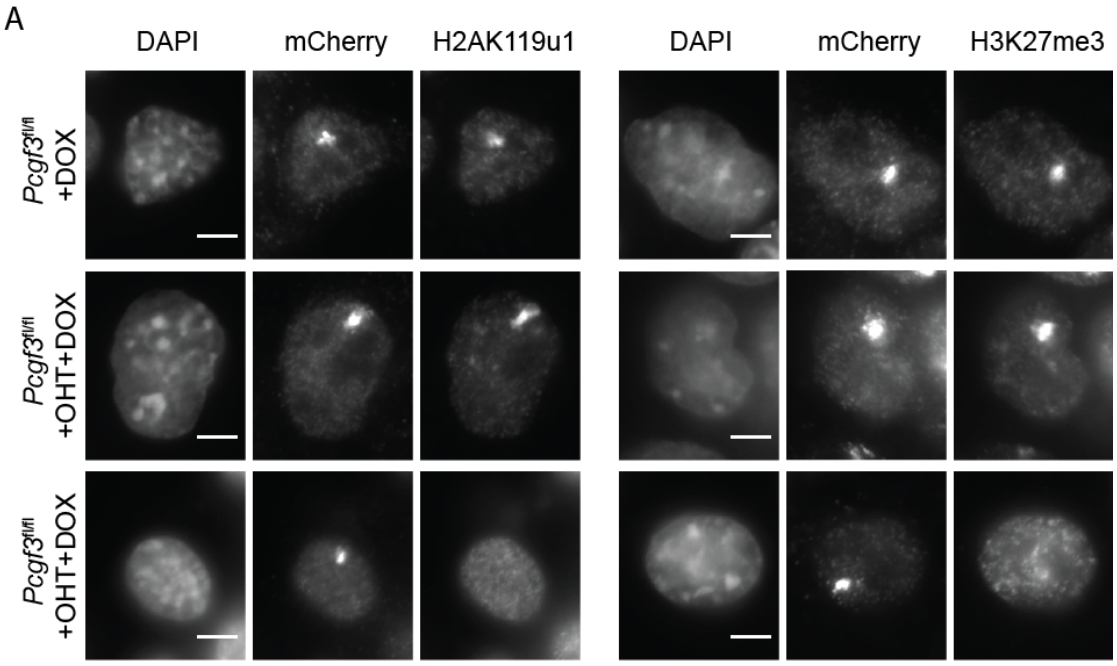
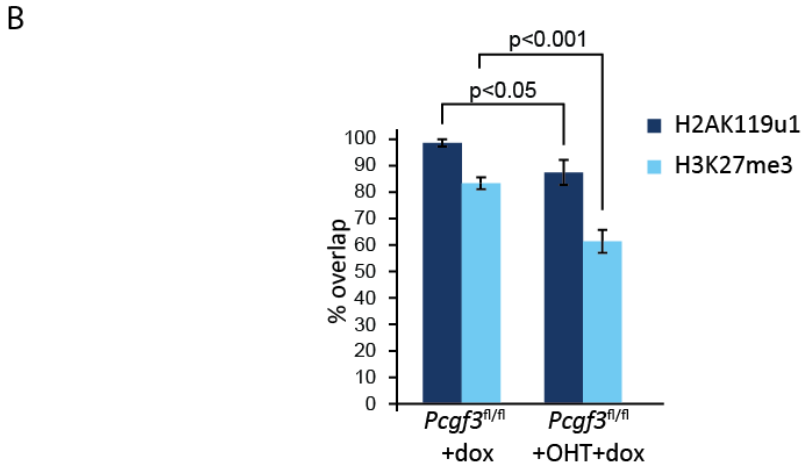


Figure 38: Analysis of *Pcgf1*$^{fl/fl}$ mESCs. (A) Immunofluorescence analysis of Xist-BglG-mCherry (mCherry) and H2AK119u1/H3K27me3 in the presence and absence of PCGF1. Scale bar is 5µm. (B) Bar chart represents

percentage of mCherry domains with staining for the histone modifications. Scoring data was obtained from a minimum of 300 cells in three replicates. Error bars show standard deviation. (C) Western blot analysis of PRC1 (RYBP, RING1B, CBX7) and PRC2 (EZH2) subunits, and H3K27me3 and H2AK119u1 histone modifications, before (UNT) and after (48h OHT) deletion of *Pcgf1*. PCGF1 deletion was confirmed by Western blot. Loading controls are TBP and histone H4. Molecular weight markers are shown on the left.

A similar analysis was conducted in *Pcgf3*$^{fl/fl}$ mESC line. In this experiment, deletion of *Pcgf3* led to a significant decrease in the number of cells showing H2AK119u1 domains overlapping with Xist-BglG-mCherry domains, and an even more accentuated reduction of H3K27me3 domains (Figure 39B). Cells representing control cell line (+dox) and examples of both the presence and absence of H2AK119u1/H3K27me3, upon *Pcgf3* deletion before Xist expression (+OHT+dox), are shown in Figure 39A. Further analysis of this cell line by Western blot shows that the reduction of H2AK119u1/H3K27me3 on the inactive domain is not due to changes in protein levels of PRC1 (RYBP, RING1B and CBX7) or PRC2 (EZH2) components (Figure 39C, left panel). Global levels of H2AK119u1 and H3K27me3 are also unchanged (Figure 39C, right panel), suggesting that PCGF3 might play a specific role in the initiation of Xist-mediated Polycomb recruitment.


A
DAPI
mCherry
H2AK119u1
DAPI
mCherry
H3K27me3
Pcgf3fl/fl +DOX
Pcgf3fl/fl +OHT+DOX
Pcgf3fl/fl +OHT+DOX


B
p<0.001
p<0.05
H2AK119u1
H3K27me3
% overlap
100
90
80
70
60
50
40
30
20
10
0
Pcgf3fl/fl +dox
Pcgf3fl/fl +OHT+dox

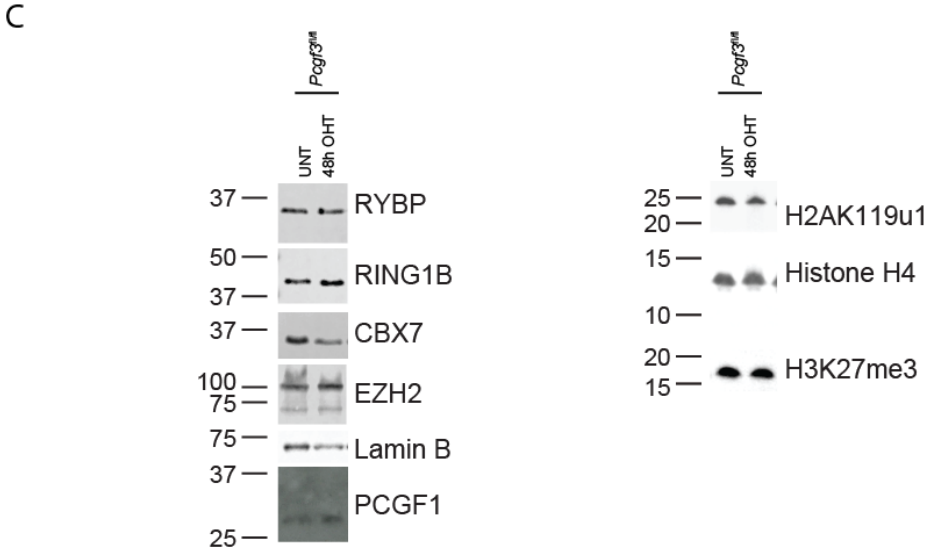

C
Pcgf3fl/fl
UNT
48h OHT
37
RYBP
50
RING1B
37
37
CBX7
100
75
EZH2
75
Lamin B
37
PCGF1
25
Pcgf3fl/fl
UNT
48h OHT
25
20
H2AK119u1
15
Histone H4
10
20
H3K27me3
15

Figure 39: Analysis of *Pcgf3*$^{fl/fl}$ mESCs. (A) Immunofluorescence analysis of Xist-BglG-mCherry (mCherry) and H2AK119u1/H3K27me3 in the presence and absence of PCGF3. Scale bar is 5µm. (B) Bar chart represents percentage of mCherry domains with staining for the histone modifications. Scoring data was obtained from a minimum of 300 cells in three replicates. Error bars show standard deviation. (C) Western blot analysis of PRC1 (RYBP, RING1B, CBX7, PCGF1) and PRC2 (EZH2) subunits, and H3K27me3 and H2AK119u1 histone modifications, before (UNT) and after (48h OHT) deletion of *Pcgf1*. Loading controls are LAMIN B and histone H4. Molecular weight markers are shown on the left.

Analysis of *Pcgf5* knockout in *Pcgf1*$^{fl/fl}$ Xist clone E5 background shows that depletion of PCGF5 has no effect on the number of cells showing H2AK119u1 domains (Figure 40B). However, there is a significant decrease of around 25% in the percentage of cells showing H3K27me3 accumulation over the Xist-Bgl-mCherry territory (Figure 40B). This result is consistent between the two clones tested (clone C3 and D3 characterized in Figure 35B). The same analysis was carried out in *Pcgf5* knockout in *Pcgf3*$^{fl/fl}$ Xist clone A2 mESC. In the two clones analyzed (clone A2 and B1 characterized in Figure 35B) there was no change in the number of cells showing H2AK119u1 or H3K27me3 accumulation over the Xist-Bgl-mCherry domain (Figure 40B). I conclude that PCGF5 is not required for initiation of Polycomb recruitment to the Xist territory, even though it might have an impact on H3K27me3 accumulation. The latter observation needs to be cautiously interpreted since there is a discrepancy between cell lines.

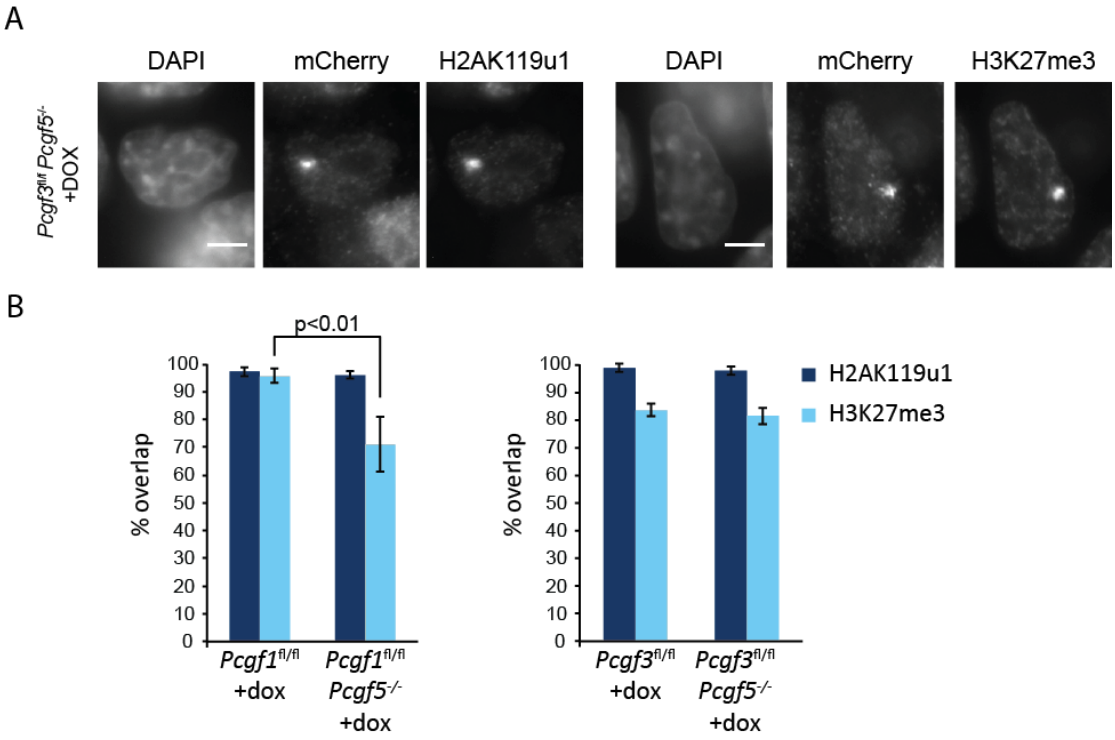


Figure 40: Analysis of *Pcgf5* single knockout mESCs. (A) Immunofluorescence analysis of Xist-BglG-mCherry (mCherry) and H2AK119u1/H3K27me3 in the absence of PCGF5. Scale bar is 5µm. (B) Bar charts represent percentage of mCherry domains with staining for the histone modifications in *Pcgf5*$^{-/-}$ in *Pcgf1*$^{fl/fl}$ Xist E5 and *Pcgf3*$^{fl/fl}$ Xist A2 cells lines. Scoring data is the average calculated from a minimum of 300 cells in three replicates for two different knockout clones. *Pcgf1*$^{fl/fl}$ +dox and *Pcgf3*$^{fl/fl}$ +dox are shown for comparison. Error bars show standard deviation.

Analysis of *Pcgf6* knockout in *Pcgf1*$^{fl/fl}$ Xist clone E5 mESCs indicated that depletion of PCGF6 might cause a small decrease both in the number of cells with H2AK119u1 as well as H3K27me3 domains (Figure 41). However, the difference in the number of cells with accumulation of the marks over the Xist-BglG-mCherry territory is not significantly different between *Pcgf1*$^{fl/fl}$ Xist clone E5 and *Pcgf6* knockout cells after 24 hours of Xist induction (Figure 41B).

Consistently, analysis of *Pcgf6* knockout in *Pcgf3*$^{fl/fl}$ Xist clone A2 mESCs shows no significant changes in the percentage of cells with H2AK119u1 domains (Figure 41B). The percentage of cells with H3K27me3 over the Xist-BglG-mCherry territory seems to increase in the *Pcgf6*$^{-/-}$ knockout clones. Even though this difference is statistically significant, it is likely to be related to variability in H3K27me3 staining performed at different times rather than a real biological effect since *Pcgf3*$^{fl/fl}$ Xist clone A2 (Figure 41B), used as a control (+dox) shows a lower proportion of cells with H3K27me3 domains than expected in a wild-type background. Also this difference is not observed in *Pcgf6*$^{-/-}$ generated in *Pcgf1*$^{fl/fl}$ Xist E5 clone (Figure 41B). Nonetheless, the current observations led me to conclude that PCGF6 is not required for initiation of Polycomb recruitment to the Xist territory and it has no impact on H2AK119u1/H3K27me3 accumulation.

Altogether these experiments showed that PCGF3, and to a lesser extent PCGF5, are the only two PCGF proteins that impact on the accumulation of Polycomb deposited marks on the inactive chromosome domain. However, none of the *Pcgf* deletions resulted in a complete loss of PRC1 or PRC2 from the inactive chromosome, therefore suggesting that some of these proteins might be recruited through alternative paths or that some of them might have an overlapping function and be able to compensate for each other.

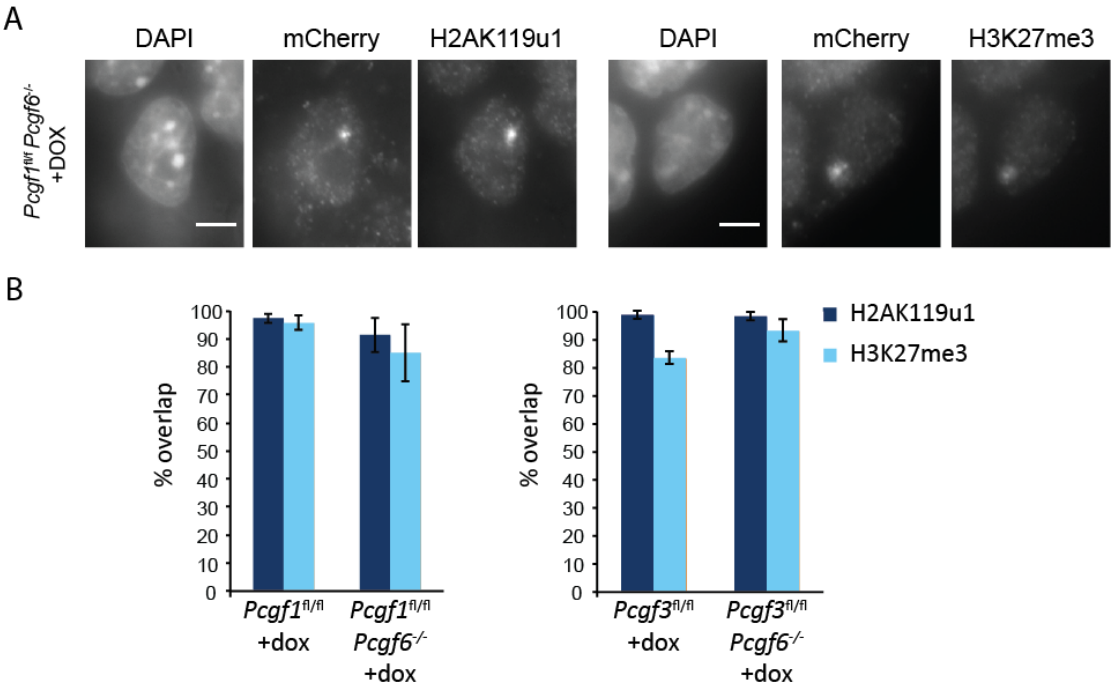


Figure 41: Analysis of *Pcgf6* single knockout mESCs. (A) Immunofluorescence analysis of Xist-BglG-mCherry (mCherry) and H2AK119u1/H3K27me3 in the absence of PCGF6. Scale bar is 5µm. (B) Bar charts represent percentage of mCherry domains with staining for the histone modifications in *Pcgf6*$^{-/-}$ in *Pcgf1*$^{fl/fl}$ Xist E5 and *Pcgf3*$^{fl/fl}$ Xist A2 cells lines. Scoring data is the average calculated from a minimum of 300 cells in three replicates for two different knockout clones. *Pcgf1*$^{fl/fl}$ +dox and *Pcgf3*$^{fl/fl}$ +dox are shown for comparison. Error bars show standard deviation.

6.2.4. Analysis of double *Pcgf3/Pcgf5* knockouts effect in *de novo* Xist-mediated Polycomb recruitment

To test if PCGF3 and PCGF5, which are close homologs, function redundantly in initiating recruitment of Polycomb to the inactive chromosome I used mESCs in which I generated mutations in *Pcgf5* gene in *Pcgf3*$^{fl/fl}$ Xist A2

(characterized in Figure 35B). Clones A2 and B1 previously used to address the effect of *Pcgf5* deletion were tamoxifen treated to delete *Pcgf3* before Xist RNA induction, according to the experimental setup shown in Figure 37. Analysis of the number of cells showing H2AK119u1 and H3K27me3 accumulation overlapping with the Xist-BglG-mCherry cloud showed that depletion of PCGF3 and PCGF5 in combination leads to a dramatic reduction of the number of cells with Polycomb domains (Figure 42). This result suggests that PCGF3 and PCGF5 function redundantly to initiate Polycomb recruitment to the inactive chromosome domain.

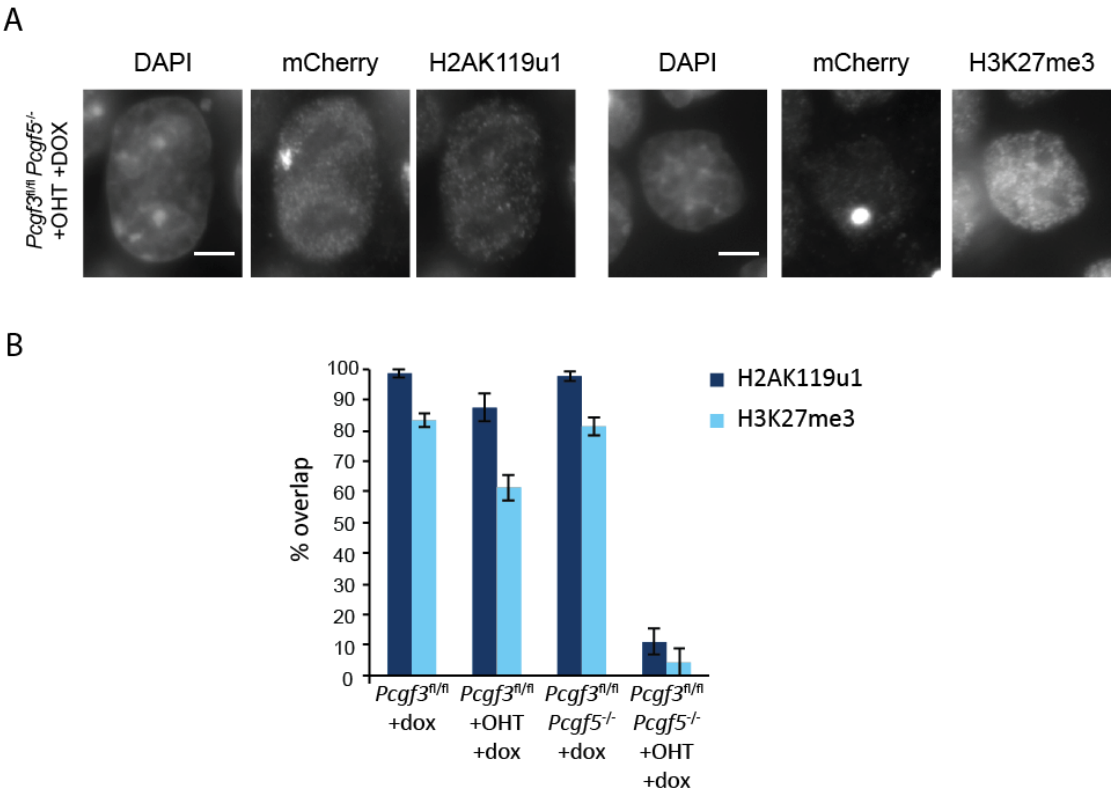


Figure 42: Analysis of *Pcgf3*/*Pcgf5* double knockout mESCs. (A) Immunofluorescence analysis of Xist-BglG-mCherry (mCherry) and H2AK119u1/H3K27me3 in cells lacking PCGF3 and PCGF5. Scale bar is 5µm. (B) Bar chart represents percentage of mCherry domains with staining for the histone modifications in *Pcgf5*$^{-/-}$ *Pcgf3*$^{fl/fl}$ Xist A2 cells line after 48 hours of tamoxifen treatment. Scoring data is the average calculated from a

minimum of 300 cells in three replicates for two different knockout clones. *Pcgf3*$^{fl/fl}$ +dox and *Pcgf3/Pcgf5* single knockouts data is shown for comparison. Error bars show standard deviation. Complete statistical analysis by Student t-test comparing all different groups can be found in Table 7.

To confirm this result, I generated an alternative cell line in which *Pcgf5* knockout was created in *Pcgf3*$^{fl/fl}$ mESCs before the Xist-BglG-mCherry system was integrated, as shown in the schematic in Figure 34. I selected clones C3 and B3 homozygous mutants (Figure 35B) in which I randomly integrate Xist transgenes. Two integrants (Clones C3B3/C3F8 and B3B7/B3H8) were analyzed for each *Pcgf5* deletion (C3 and B3). Using the same experimental setup (Figure 37) and analysis of H2AK119u1 and H3K27me3 domains (Figures 43A and 43B), I confirmed the results previously obtained in *Pcgf5* knockouts generated in *Pcgf3*$^{fl/fl}$ Xist A2 cell line (Figure 42). Further characterization of cells lacking PCGF3 and PCGF5 by Western blot showed that deletion of these genes does not result in any major changes in the expression of other PRC1 (RYBP, RING1B, CBX7) and PRC2 (EZH2) subunits (Figure 43C, left panel). Likewise H2AK119u1 and H3K27me3 global levels seem largely unchanged (Figure 43C, right panel).

A small proportion of cells show accumulation of the inactive chromatin marks H2AK119u1 and H3K27me3 over the inactive chromosome marked by Xist-BglG-mCherry after Cre-recombinase deletion of *Pcgf3* for 48 hours in *Pcgf5* knockout cells (Figure 43B). This could be due to incomplete Cre-mediated deletion of *Pcgf3* (even though this was not detected by PCR – Figure 32B) and/or the stability of previously translated protein in the nucleus. With this in

mind I went on to derive stable long-term *Pcgf3/Pcgf5* knockout mESCs lines in order to induce Xist expression at a later timepoint, when the proteins will be entirely depleted.

Constitutive null mESCs (*Pcgf3*$^{-/-}$ *Pcgf5*$^{-/-}$) were generated by treating 2 clones (C3F8 and B3B7) with tamoxifen for 96 hours. Between 48h and 72h of tamoxifen treatment, these cell lines underwent significant cell death. Presumably this occurs as a response to the simultaneous loss of PCGF3 and PCGF5, as the same is not seen in *Pcgf1*$^{fl/fl}$ mESCs. A similar behavior was observed in *Pcgf3*$^{fl/fl}$ when tamoxifen treated for longer than 48h (data not shown) suggesting that depletion of PCGF3 in particular has an impact on cell viability and a potential role outside of X inactivation. Despite the cell death, I was able to expand *Pcgf3*$^{-/-}$ *Pcgf5*$^{-/-}$ mESCs that survived the treatment. These cells retain apparently normal ES cell morphology and growth characteristics. PCR analysis using primers in Figure 32B confirmed that these surviving cells had deleted *Pcgf3*.

Analysis of constitutive *Pcgf3*$^{-/-}$ *Pcgf5*$^{-/-}$ after 24h of Xist induction by doxycycline showed that while Xist-BglG-mCherry domain formation is normal, there is no accumulation of H2AK119u1 or H3K27me3 in any cell of any of the two clones (Figure 43B). This result confirms that PCGF3 and PCGF5 have overlapping functions to initiate Polycomb recruitment by Xist RNA.

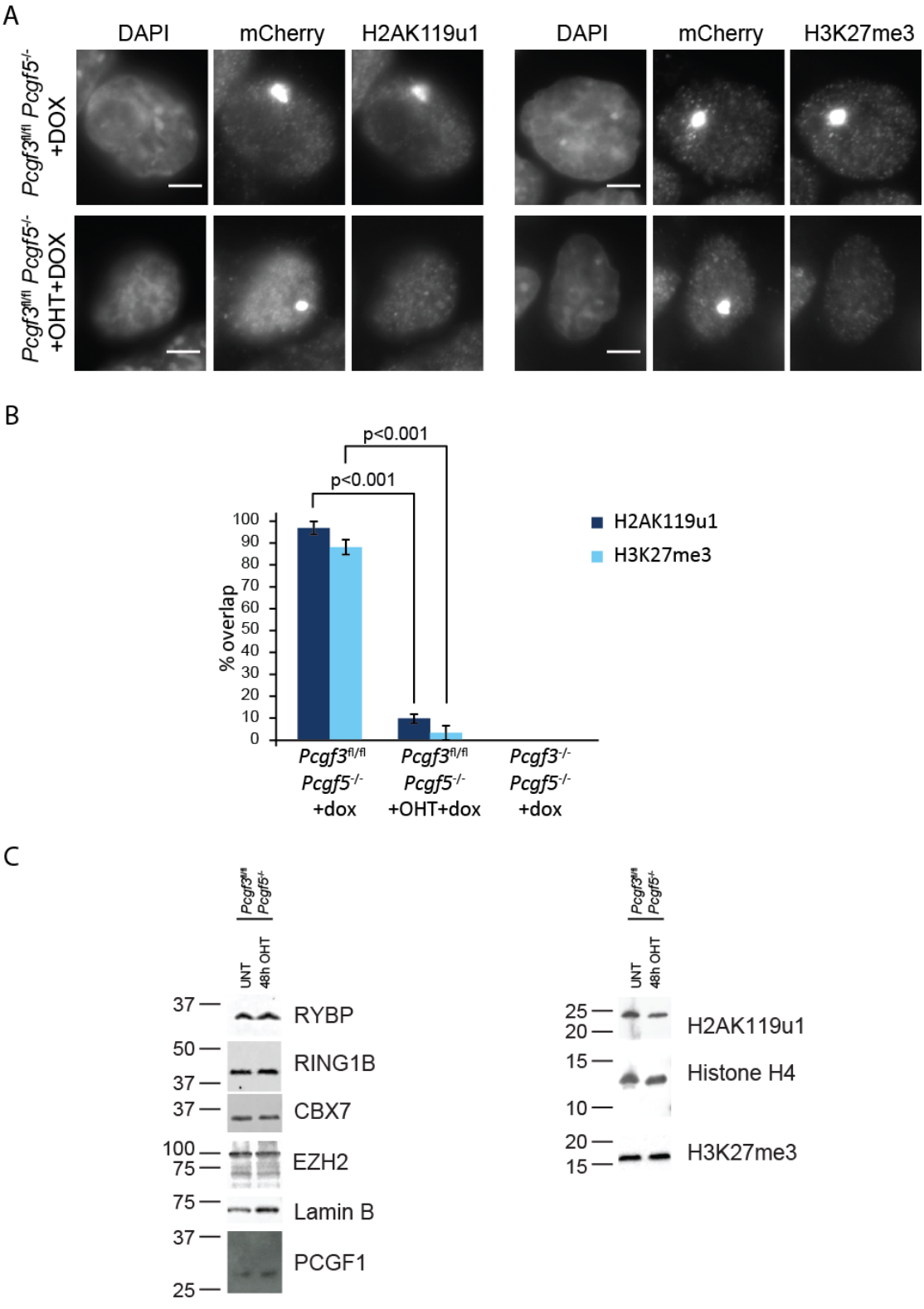


Figure 43: Analysis of *Pcgf3*/*Pcgf5* double knockout mESCs with different random Xist integrations. (A) Immunofluorescence analysis of Xist-BglG-mCherry (mCherry) and H2AK119u1/H3K27me3 in cells lacking PCGF3 and PCGF5. Scale bar is 5µm. (B) Bar chart represents percentage of

mCherry domains with staining for the histone modifications in *Pcgf5*$^{-/-}$ *Pcgf3*$^{fl/fl}$ cells line after 48 hours of tamoxifen treatment and in constitutive *Pcgf3*$^{-/-}$ *Pcgf5*$^{-/-}$ mESC. Scoring data is the average calculated from a minimum of 300 cells in three replicates for two different knockout clones. Error bars show standard deviation. (C) Western blot analysis of PRC1 (RYBP, RING1B, CBX7, PCGF1) and PRC2 (EZH2) subunits, and H3K27me3 and H2AK119u1 histone modifications, before (UNT) and after (48h OHT) deletion of *Pcgf3* in *Pcgf5*$^{-/-}$ mESCs. Loading controls are LAMIN B and histone H4.

6.2.5. Analysis of other double knockouts in *de novo* Xist-mediated Polycomb recruitment

Pcgf5 and *Pcgf6* knockout in *Pcgf1*$^{fl/fl}$ Xist clone E6 mESCs and *Pcgf6* knockout in *Pcgf3*$^{fl/fl}$ Xist clone A2 mESCs were tested to confirm that recruitment of Polycomb to the Xist domain is specifically mediated by PCGF3 and PCGF5.

Depletion of PCGF1 in the absence of PCGF5 or PCGF6 resulted in no significant changes in the percentage of cells with H2AK119u1 accumulation over the Xist-Bgl-mCherry domain (Figure 44). Interestingly, both these scenarios result in a statistically significant decrease in the proportion of cells with H3K27me3 domains compared to wild-type control. In *Pcgf1*$^{-/-}$ *Pcgf5*$^{-/-}$ there is a visible contribution from the absence of PCGF5 to this decrease in H3K27me3 that is then accentuated by PCGF1 depletion (Figure 44). In *Pcgf1*$^{-/-}$ *Pcgf6*$^{-/-}$ mESCs, it seems that the H3K27me3 decrease results from the cumulative effect of depleting PCGF1 and PCGF6, since lack of both of

these proteins individually does not have a significant impact in the proportion of cells with H3K27me3 domains on the inactive chromosome (Figure 44). The decrease in H3K27me3 domains in *Pcgf1*$^{-/-}$ *Pcgf6*$^{-/-}$ is only modest when compared to the decrease observed in *Pcgf1*$^{-/-}$ *Pcgf5*$^{-/-}$.

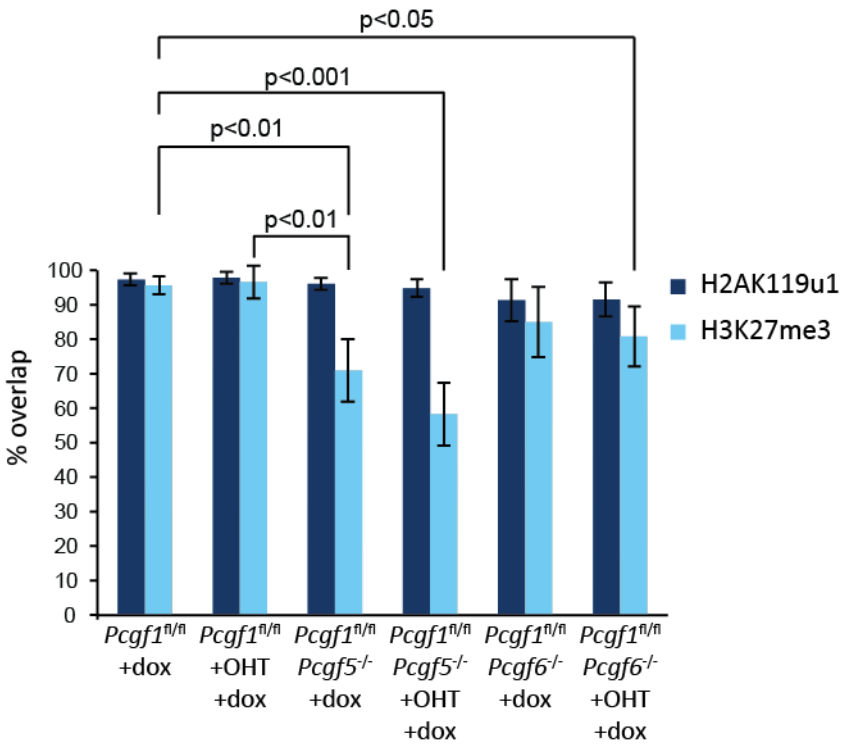


Figure 44: Analysis of other double knockouts generated in *Pcgf1*$^{fl/fl}$ Xist E5 cell line. Bar chart summarizes percentage of cells with H2AK119u1/H3K27me3 over the Xist-BglG-mCherry domain in *Pcgf1*/*Pcgf5* and *Pcgf1*/*Pcgf6* double knockout mESCs. Single knockouts data is shown for comparison. Scoring data is the average calculated from a minimum of 300 cells in three replicates for two different knockout clones. Error bars show standard deviation. Complete statistical analysis by Student t-test comparing all different groups can be found in Table 6.

Analysis of *Pcgf3*$^{-/-}$ *Pcgf6*$^{-/-}$ mESCs showed that deletion of these two genes has no impact on H2AK119u1 deposition on the inactive chromosome domain

when compared to *Pcgf3*$^{fl/fl}$ Xist A2 control. This might seem surprising taking into account that deletion of *Pcgf3* (Figure 45) leads to a reduction in H2AK119u1 deposition. However, the observed decrease was small and on the limit of statistical significance, so immunostaining variability might account for the fact that there is no significant effect in *Pcgf3*$^{-/-}$ *Pcgf6*$^{-/-}$ cells. Analysis of H3K27me3 shows that there is a decrease in the proportion of cells showing accumulation of this mark when compared to *Pcgf3*$^{fl/fl}$ Xist A2 control (Figure 45). In light of the data from single knockouts it is reasonable to assume that this reduction is mostly attributable to loss of PCGF3 (*Pcgf3* single knockout data, Figure 45).

Altogether these results suggest that depletion of any two PCGF proteins (*Pcgf1*$^{-/-}$ *Pcgf5*$^{-/-}$, *Pcgf1*$^{-/-}$ *Pcgf6*$^{-/-}$ and *Pcgf3*$^{-/-}$ *Pcgf6*$^{-/-}$) have an impact on H3K27me3 deposition over the Xist-BglG-mCherry domain. However, in all the double knockout cells lines except *Pcgf3*$^{-/-}$ *Pcgf5*$^{-/-}$, the proportion of cells with deposition of H2AK119u1 over the Xist-BglG-mCherry domain is on the whole unaffected.

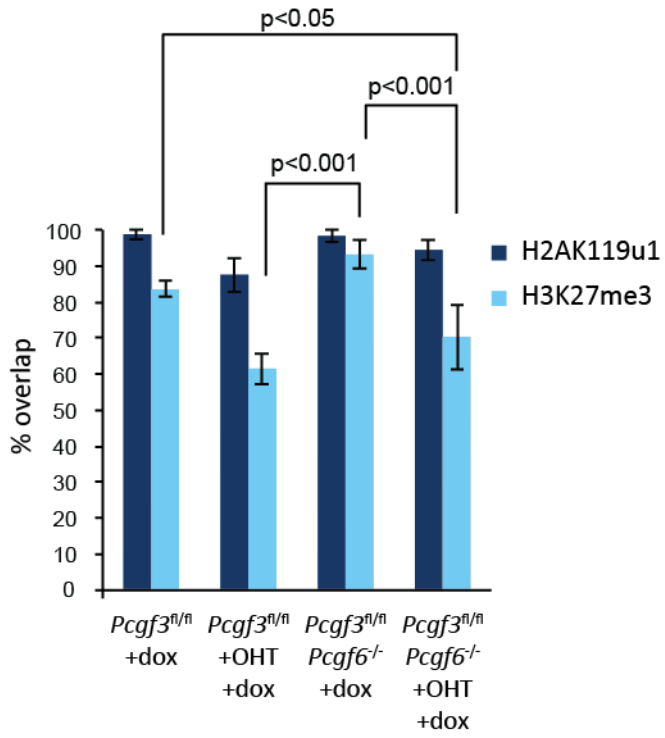


Figure 45: Analysis of *Pcgf3*/*Pcgf6* double knockout mESCs. Bar chart summarizes percentage of cells with H2AK119u1/H3K27me3 over the Xist-BglG-mCherry domain in *Pcgf3*/*Pcgf6* double knockout mESCs. Error bars show standard deviation. Complete statistical analysis by Student t-test comparing all different groups can be found in Table 7.

6.3. Discussion

In this chapter, I investigated the role of PCGF proteins in the initiation of Polycomb domain formation upon Xist expression in mESCs. Importantly, single knockout of PCGF proteins highlighted that depletion of mainly PCGF3, but also PCGF5, has an impact on Polycomb deposition on the inactive chromosome territory, contrary to what happens in cells depleted of PCGF1 or PCGF6. Interestingly, depletion of PCGF3, or PCGF3 and PCGF5 in the

same cells, led to a cell viability crisis suggesting that these proteins or at least PCGF3 has an important role in mESCs. Interestingly, no major changes in PRC1 or PRC2 subunit levels or global H2AK119u1/H3K27me3 levels were observed by Western blot, so further investigation is needed to understand what causes loss of cell viability upon deletion of *Pcgf3* both in wild-type as well as in *Pcgf5* knockout cells. This will be important to unveil specific functions for non-canonical PRC1 complexes containing PCGF3 or PCGF5. Interestingly, PCGF3 is the closest homologue to L(3)73AH in *Drosophila*, a poorly characterized PSC homolog that has been shown to be responsible for around 70% of total H2AK118u1 (Lee et al. 2015) and for which depletion is lethal in flies (Belote et al. 1990).

Even though depletion of PCGF3 leads to a decrease in the number of cells showing H2AK119u1 and concomitantly H3K27me3 over the Xist domain, this loss only occurs in a small proportion of cells suggesting that a compensatory mechanism for Polycomb domain formation is in place in most of the cells analyzed. Importantly, the fact that H3K27me3 is lost in approximately the same proportion of cells that lack H2AK119u1 deposition on the inactive chromosome fits with the result obtained in Chapter 3 that shows a complete dependency of PRC2 recruitment/ H3K27me3 deposition on PRC1/H2AK119u1. Immunofluorescence of H2AK119u1 and H3K27me3 in the same cells would be required to confirm that H3K27me3 loss occurs in cells that do not have detectable H2AK119u1 on the inactive chromosome.

In some PCGF5 knockout cells, a visible, albeit small, decrease in H3K27me3 accumulation is detected in the absence of H2AK119u1 deposition changes. That depletion of PCGF5 leads to changes in H3K27me3 without visible changes in the accumulation of H2AK119u1 could potentially be explained by the amount of H2AK119u1 that is deposited on the chromosome. Results in Chapter 3 point towards a total dependency for PRC2 recruitment on H2AK119u1. Herein, I speculate that the observed cells lacking H3K27me3 might have lower levels of H2AK119u1 on the inactive domain due to lack of PCGF5-PRC1, then impacting on PRC2 recognition of this histone mark. Further quantitative experiments will be necessary to confirm that this is the case as differences in the amount of H2AK119u1 are challenging to quantify in an immunofluorescence assay. Importantly, PCGF5 was the only PCGF protein identified in recent screenings for Xist interactors. Due to the type of assay used by Chu *et al.* (2015) it is not conclusive if this interaction is direct or mediated by another protein, however, my results show that depletion of PCGF5 on its own is not sufficient to abolish Polycomb recruitment.

FRAP data described in Chapter 4 supports the idea that PCGF3 and PCGF5 have a similar dynamic behavior on the inactive chromosome, that is moreover distinct from other PCGF proteins. The similarity between these two proteins is further supported by alignments of their aminoacid sequences, potential interaction with the same partners (Gao et al. 2012; Gao et al. 2014; Hauri et al. 2016) and the fact that they both contribute to similar and fairly small proportions of PRC1 complexes in ES cells when compared to other PCGF proteins (Kloet et al. 2016). These observations together with results

from single knockout cells led me to consider an overlapping function of PCGF3 and PCGF5 in PRC1 recruitment and Polycomb domain formation on the inactive chromosome territory.

Analysis of double PCGF knockout cell lines showed that PCGF3 and PCGF5 together are necessary to initiate Polycomb domain formation over the Xist territory. Complete lack of H2AK119u1/H3K27me3 deposition on the inactive chromosome was observed in *Pcgf3*$^{-/-}$ *Pcgf5*$^{-/-}$ mESCs, even though deletion of these two genes does not result in major changes in the expression of other Polycomb complex subunits or in the global levels of H2AK119u1 and H3K27me3. Considering that PCGF3 and PCGF5 have previously been shown to interact with some common partners, for example AUTS2 (Gao et al. 2012; Gao et al. 2014; Hauri et al. 2016), it would be interesting to explore if these interactors have any role in the context of Xist-dependent domain formation. Nevertheless, it is important to stress that none of these PCGF3/5 partners was identified in several screenings to identify the Xist RNA interactome (McHugh et al. 2015; Chu et al. 2015).

The use of an inducible Xist system in these experiments was crucial to understand the impact of depleting one or several PCGF proteins in the *de novo* Polycomb recruitment. The advantages of this system have been previously discussed, but there are also some caveats to consider: the use of a randomly integrated inducible Xist transgene might result in an upregulation of Xist to different levels compared to endogenous Xist, both due to the use of an inducible promoter but also because of the number of Xist copies that are

integrated. Furthermore, expression of a random Xist integrant in an autosome might result in a different ability to silence, depending on the features of the chromosome such as gene density, LINE enrichment or chromatin architecture, which might have consequences in protein recruitment (such as Polycomb proteins). To overcome these issues, several clones with different random Xist integrants were tested, and the same result was observed in all of them. Moreover, in an experiment performed by Dr. Osamu Masui, a postdoc in Dr. H. Koseki's lab, in collaboration with Dr. Tatyana Nesterova, a senior research associate in the Brockdorff lab, the absence of Polycomb domain formation in *Pcgf3*$^{-/-}$ *Pcgf5*$^{-/-}$ mESCs was verified on the X chromosome in female *Pcgf3*$^{-/-}$ *Pcgf5*$^{-/-}$ mouse embryos (see Figure 48).

Even though PCGF1 and PCGF6 together account for 80% of all non-canonical PRC1 complexes in mESCs (Kloet et al. 2016), a double knockout for these two proteins had no impact on Polycomb deposition on the inactive chromosome domain. In fact, a more thorough investigation of different PCGF double knockouts led me to conclude that the lack of Polycomb domain formation is a specific consequence of depletion of PCGF3 and PCGF5. Depletion of PCGF3 or PCGF5 in combination with other proteins (PCGF6 and PCGF1, respectively) did not have a significant effect in the initiation of Polycomb domain formation. However, it was surprising to observe that different double PCGF knockout combinations resulted in a reduced accumulation of H3K27me3 regardless of which two PCGF proteins were missing. As speculated for the PCGF5 single knockout, it is reasonable to suggest that this is linked to reduced levels of H2AK119u1 deposited on the

inactive chromosome. However, PRC1 functions independent of H2AK119u1 might also contribute to the stability and activity of PRC2 on the chromosome. PRC1 complexes were proposed to contribute to chromatin compaction (Francis et al. 2004; Eskeland et al. 2010), a feature that has been related to increased PRC2 activity (Yuan et al. 2012). Hence, depletion of some PRC1 complexes in PCGF double knockout cells might affect the ability of PRC2 to place H3K27me3.

CHAPTER 7: General Discussion

Mechanisms of Polycomb recruitment have been widely studied, not only in *Drosophila*, where the first Polycomb gene functions were unveiled, but also in mammalian systems, where the diversity of Polycomb complex variants suggest a diversification and increased specialization in Polycomb function and recruitment mechanisms. Due to the complexity of the Polycomb system, many of these studies were conducted using tethering systems or investigating specific loci important in particular cell types. In this thesis, I emphasized the value of a model for XCI, triggered by induced Xist RNA expression in mESCs, to investigate *de novo* Polycomb recruitment. The ability to trigger *de novo* Polycomb recruitment using a system that mimics a physiological situation, Xist-mediated silencing, allowed me to dissect the role of different non-canonical PRC1 subunits not only in *de novo* recruitment but also in positive feedback pathways involved in Polycomb domain formation on the Xi and potentially in other regions of the genome.

Based on the findings described in this thesis, I proposed a revised model for a stepwise pathway of Xist-mediated Polycomb recruitment as depicted in Figure 46. In summary, Xist RNA expression and coating of the chromosome leads to Polycomb recruitment, initiated by PCGF3-PRC1 and PCGF5-PRC1 variants that deposit H2AK119u1 on the chromosome. Recognition of H2AK119u1 by the RYBP Ubiquitin Binding Domain and its weak and transient binding to this histone modification leads to enrichment of all non-canonical PRC1 variants containing PCGF1 to PCGF6 over the inactive

chromosome. This results in a reinforcing mechanism for the accumulation of PRC1 and H2AK119u1. H2AK119u1 is then recognized by the JARID2 Ubiquitin Interacting Motif, bringing PRC2 complexes to the inactive chromosome to deposit H3K27me3 as proposed by Cooper et al. (2016). In agreement with the initially proposed hierarchical model of Polycomb recruitment, H3K27me3 can be recognized by the chromodomain in CBX proteins promoting accumulation of canonical PRC1 variants on the inactive chromosome domain.

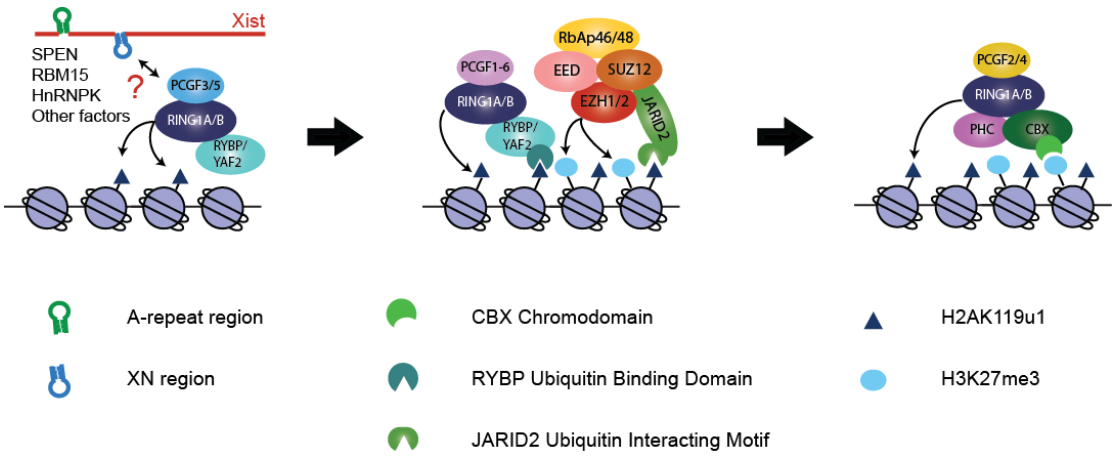


Figure 46: Model summarising key steps in Polycomb recruitment by Xist RNA.

7.1. PCGF3/PCGF5-PRC1 variants and Xist-mediated Polycomb recruitment

The molecular basis for PCGF3/PCGF5-PRC1 interaction with Xist RNA is still unresolved. Deletion of Xist XN region (encompassing the F- and B-repeats) was previously shown to impair PRC2 binding to the inactive

chromosome and JARID2 was proposed to directly interact with this region of the RNA to mediate PRC2 recruitment (da Rocha et al. 2014). However, in this thesis, I was able to show that PRC2 recruitment upon Xist expression is fully dependent on PRC1/H2AK119u1. Furthermore, the XN region of Xist was also shown by others in the Brockdorff lab to be required for PRC1 recruitment (G. Pintacuda, personal communication), which fits well with the aforementioned observation. Further studies will be necessary to pinpoint the exact Xist region necessary for Polycomb recruitment as well as to determine if Polycomb recruitment is mediated by a direct interaction of one of its subunits to the RNA or if an adaptor protein might be involved.

As previously discussed, PCGF3 and PCGF5 are close homologs, showing high similarity in their amino acid sequence. This observation together with the fact that these proteins seem to have overlapping function in Xist-mediated *de novo* Polycomb recruitment leads me to speculate that they might directly interact with Xist to accomplish this function. In fact, PCGF5 was the only PCGF protein identified in several MS studies that aim to characterize the Xist interactome (Chu et al. 2015; McHugh et al. 2015; Minajigi et al. 2015). Importantly, a study that identified Polycomb subunits as direct interactors of Xist relied on crosslinking using formaldehyde (Chu et al. 2015). Formaldehyde is able to cross link proteins in the proximity of the RNA even if not direct interacting with it, leaving open for discussion the possibility of PRC1 being a direct interactor or not. As previously mentioned, a super-resolution microscopy study supports the idea of an adaptor protein mediating Polycomb-Xist interaction. In line with this, data from a UV cross linking based

purification of Xist interactors followed by MS (McHugh et al. 2015) identified putative Xist RNA direct binders and did not retrieve any PRC1 components. The idea of an adaptor protein also fits well with the fact that none of the known PRC1 components, namely PCGF3 and PCGF5, have known RNA-binding motifs. Potential mediators of PRC1 interaction with Xist RNA are the RNA-binding proteins SPEN, HnRNPK and RBM15, as knockdowns of these proteins have been shown to affect Polycomb deposition on the inactive chromosome domain (Moindrot et al. 2015; McHugh et al. 2015; Monfort et al. 2015; Chu et al. 2015). Other PCGF3 and PCGF5 interactors identified in prior studies (Gao et al. 2014; Hauri et al. 2016) may also have a role, although these proteins were not identified as Xist interactors and their role in XCI has not been addressed.

Further studies will be crucial to define the molecular mechanism involved in PCGF3/PCGF5-PRC1 complex recruitment to the inactive chromosome. Understanding this mechanism will be crucial not only to completely define Polycomb domain formation in the process of XCI, but also to understand if PCGF3/5-PRC1 targeting might be important for *de novo* Polycomb recruitment to targets genome-wide, for example through binding to other lncRNAs. Given the recent studies in *Drosophila*, which emphasize the contribution of L3(7)3AH, a close homolog of PCGF3, to H2AK118u1 deposition genome-wide, it will also be interesting to understand if there is a link between the function played by these proteins in *Drosophila* and in mammals, respectively.

Importantly, a collaborative study between the Koseki and Brockdorff laboratories allowed us to confirm that depletion of PCGF3 and PCGF5 impairs Polycomb domain formation over the endogenous inactive X chromosome in female mouse embryos. Future studies using *Pcgf3*$^{-/-}$ *Pcgf5*$^{-/-}$ embryos or cell lines with hybrid genetic backgrounds that allow allele discrimination due to the presence of strain specific SNPs (as used before in Pinter et al. 2012; da Rocha et al. 2014; Marks et al. 2015) will be pivotal in determining the contribution of Polycomb for Xist-mediated silencing, without altering Xist RNA sequence. Analysis of embryos resulting from crosses of heterozygous mice for *Pcgf3* and *Pcgf5* (*Pcgf3*$^{+/-}$ *Pcgf5*$^{+/-}$) suggested that these genes are required for development since no *Pcgf3*$^{-/-}$ *Pcgf5*$^{-/-}$ mice were born (data not shown). Analysis of embryos at days 9.5 and 10.5 of embryonic development suggest that the depletion might have a stronger impact in females, as they seem to be underrepresented at this stage when compared to males (data not shown), potentially due to defective X chromosome inactivation. Although recovery of *Pcgf3*$^{-/-}$ *Pcgf5*$^{-/-}$ embryos has proven challenging, analysis of these embryos could also provide some clues on the role of these complexes genome-wide.

Another open question is how important PCGF3/PCGF5-PRC1 variants are for the maintenance of Polycomb on the Xi, after establishment of XCI. Ongoing studies using trophoblast stem (TS) cells that possess an imprinted paternal inactive X chromosome have shown that depletion of PCGF3 and PCGF5 does not produce such a drastic effect as observed in mESCs. Most TS cells maintain H2AK119u1 enrichment on the imprinted Xi concomitantly

to H3K27me3, 48 hours after depletion of PCGF3 and PCGF5. Longer term deletion analysis of this mutant TS cells show that H2AK119u1 is lost in most of the cells, although H3K27me3 still accumulates on the Xi in a large proportion of cells after 120h after deletion (data not shown).

Loss of H2AK119u1 is surprising in the light of the feedback mechanism in which RYBP recognizes this histone modification. However, RYBP binding to H2AK119u1 seems to be highly dynamic as analyzed by FRAP, perhaps rendering the complexes recruited in a RYBP-dependent manner less efficient in catalyzing H2AK119u1 on the Xi due to their fast turnover on chromatin. A second aspect to consider is the rate of H2AK119u1 turnover. When the rate of H2AK119u1 removal by deubiquitinases exceeds the rate of deposition due to loss of *de novo* PRC1 recruitment, the levels of H2AK119u1 fall in the absence of the signal that initiates recruitment, in this case PCGF3 and PCGF5. Moreover, the result obtained in TS cells supports the idea that PRC1 self-reinforcing mechanism is less efficient than EED recognition of PRC2 and perpetuation of H3K27me3. Further investigation is needed to understand how long the latter mechanism is capable of maintaining PRC2 enrichment on the Xi in the absence of H2AK119u1, determining the relevance of the initiating stimulus (H2AK119u1) for PRC2 domain maintenance. These conditional knockout TS cells might also be useful to address if canonical PRC1 complexes are still present on the Xi in the absence of H2AK119u1 and what their contribution towards silencing is. Even though this system might be useful to further explore mechanisms of Polycomb recruitment, it is important to keep in mind that, differences in cell

type, duration of cell cycle and X inactivation type might account for some discrepancies with the data obtained in mESCs using an inducible Xist system.

7.2. Relevance of H2AK119u1 in Polycomb domain formation

Since the discovery of the interaction of PRC2 with H2AK119u1 (Kalb et al. 2014; Cooper et al. 2016), a key question has been whether this serves a primary recruitment mechanism or a mechanism to reinforce Polycomb silencing after initiation. Studies on Polycomb occupancy in mESCs have given ambiguous results (Blackledge et al. 2014), possibly reflecting feedback mechanisms at previously established Polycomb domains. In this work I demonstrate unequivocally that in the XCI model, PRC1 mediated H2AK119u1 initiates Polycomb domain formation. Whether or not this is the case at other Polycomb targets remains an open question.

As previously mentioned, deposition of H2AK119u1 has been associated with Polycomb-mediated silencing, even though its role has recently been challenged (Illingworth et al. 2015; Pengelly et al. 2015). Regardless of the direct role that H2AK119u1 might play in Xist-mediated silencing, the fact that recognition of this histone modification is necessary for enrichment of all non-canonical PRC1, PRC2 and consequently canonical PRC1 complexes on the Xi, emphasizes its relevance in Polycomb domain formation.

RYBP/YAF2 recognition of H2AK119u1 provides a mechanism for non-canonical PRC1 reinforcement on the Xi and potentially genome-wide (as discussed in chapter 5). Although it is not known how much RYBP-dependent recruited PRC1 complexes contribute to H2AK119u1 (as discussed in section 7.1), it is important to consider that these complexes might exert other functions, contributing to silencing independently of H2AK119u1 (by mechanisms discussed in section 1.2.3.2. or others yet unknown). Further studies will be necessary to disentangle potential specific roles for each PRC1 variant as well as for subcomplexes including substoichiometric components.

JARID2, the subtoichiometric component of PRC2 that has been proposed to mediate PRC2 recruitment to the Xi through direct interaction with Xist RNA, has been recently shown to recognize H2AK119u1. In fact, recognition of H2AK119u1 by JARID2 was proposed to be the mechanism by which this protein contributes to PRC2 recruitment to the Xi (Cooper et al. 2016). The current model depicted in Figure 46 considers this mechanism for PRC2 recruitment. However, it is clear that other mechanisms are in place, at least on the Xi, since JARID2 knockout does not lead to complete depletion of H3K27me3 on the X chromosome. Further research will be essential to establish how other subunits of PRC2 recognize H2AK119u1 or conformational changes resulting from the deposition of this mark, to recruit the complex and allow H3K27me3 accumulation.

Studies addressing the stability of these positive feedback mechanisms (mediated by RYBP and JARID2 recognition of H2AK119u1) for Polycomb

domain formation genome-wide might shed light on how they contribute to their maintenance and ensures a specific transcriptional program.

CHAPTER 8: Bibliography

Akasaka T, Takahashi N, Suzuki M, Koseki H, Bodmer R, Koga H. 2002. MBLR, a new RING finger protein resembling mammalian Polycomb gene products, is regulated by cell cycle-dependent phosphorylation. *Genes Cells* **7**: 835–850.

Alekseyenko AA, Gorchakov AA, Kharchenko PV, Kuroda MI. 2014. Reciprocal interactions of human C10orf12 and C17orf96 with PRC2 revealed by BioTAP-XL cross-linking and affinity purification. *Proc Natl Acad Sci USA* **111**: 2488–2493.

Alkema MJ, Bronk M, Verhoeven E, Otte A, van 't Veer LJ, Berns A, van Lohuizen M. 1997. Identification of Bmi1-interacting proteins as constituents of a multimeric mammalian polycomb complex. *Genes Dev* **11**: 226–240.

Alqarni SSM, Murthy A, Zhang W, Przewloka MR, Silva APG, Watson AA, Lejon S, Pei XY, Smits AH, Kloet SL, et al. 2014. Insight into the architecture of the NuRD complex: structure of the RbAp48-MTA1 subcomplex. *J Biol Chem* **289**: 21844–21855.

Anderson AE, Karandikar UC, Pepple KL, Chen Z, Bergmann A, Mardon G. 2011. The enhancer of trithorax and polycomb gene Caf1/p55 is essential for cell survival and patterning in Drosophila development. *Development* **138**: 1957–1966.

Aranda S, Mas G, Di Croce L. 2015. Regulation of gene transcription by Polycomb proteins. *Sci Adv* **1**: e1500737–e1500737.

Arrigoni R, Alam SL, Wamstad JA, Bardwell VJ, Sundquist WI, Schreiber-Agus N. 2006. The Polycomb-associated protein Rybp is a ubiquitin binding protein. *FEBS Lett* **580**: 6233–6241.

Aslanidis C, de Jong PJ. 1990. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res* **18**: 6069–6074.

Atchison L, Ghias A, Wilkinson F, Bonini N, Atchison ML. 2003. Transcription factor YY1 functions as a PcG protein in vivo. *EMBO J* **22**: 1347–1358.

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

Banaszynski LA, Wen D, Dewell S, Whitcomb SJ, Lin M, Diaz N, Elsässer SJ, Chapgier A, Goldberg AD, Canaani E, et al. 2013. Hira-dependent histone H3.3 deposition facilitates PRC2 recruitment at developmental loci in ES cells. *Cell* **155**: 107–120.

Bannister AJ, Zegerman P, Partridge JF, Miska EA, Thomas JO, Allshire RC, Kouzarides T. 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**: 120–124.

Bantignies F, Roure V, Comet I, Leblanc B, Schuettengruber B, Bonnet J, Tixier V, Mas A, Cavalli G. 2011. Polycomb-dependent regulatory contacts between distant Hox loci in Drosophila. *Cell* **144**: 214–226.

Barr ML, Bertram EG. 1949. A Morphological Distinction between Neurones of the Male and Female, and the Behaviour of the Nucleolar Satellite during

Accelerated Nucleoprotein Synthesis. *Nature* **163**: 676–677.

Batta K, Zhang Z, Yen K, Goffman DB, Pugh BF. 2011. Genome-wide function of H2B ubiquitylation in promoter and genic regions. *Genes Dev* **25**: 2254–2265.

Baubec T, Colombo DF, Wirbelauer C, Schmidt J, Burger L, Krebs AR, Akalin A, Schübeler D. 2015. Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature* **520**: 243–247.

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

Beck DB, Bonasio R, Kaneko S, Li G, Margueron R, Oda H, Sarma K, Sims RJ, Son J, Trojer P, et al. 2010. Chromatin in the nuclear landscape. *Cold Spring Harbor Symposia on Quantitative Biology* **75**: 11–22.

Belote JM, Hoffmann FM, McKeown M, Chorsky RL, Baker BS. 1990. Cytogenetic analysis of chromosome region 73AD of Drosophila melanogaster. *Genetics* **125**: 783–793.

Bender LB, Cao R, Zhang Y, Strome S. 2004. The MES-2/MES-3/MES-6 Complex and Regulation of Histone H3 Methylation in C. elegans. *Current Biology* **14**: 1639–1643.

Beringer M, Pisano P, Di Carlo V, Blanco E, Chammas P, Vizán P, Gutiérrez A, Aranda S, Payer B, Wierer M, et al. 2016. EPOP Functionally Links Elongin and Polycomb in Pluripotent Stem Cells. *Mol Cell* **64**: 645–658.

Bernstein BE, Meissner A, Lander ES. 2007. The Mammalian Epigenome.

Cell **128**: 669–681.

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, et al. 2006a. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**: 315–326.

Bernstein E, Duncan EM, Masui O, Gil J, Heard E, Allis CD. 2006b. Mouse Polycomb Proteins Bind Differentially to Methylated Histone H3 and RNA and Are Enriched in Facultative Heterochromatin. *Mol Cell Biol* **26**: 2560–2569.

Blackledge NP, Farcas AM, Kondo T, King HW, McGouran JF, Hanssen LLP, Ito S, Cooper S, Kondo K, Koseki Y, et al. 2014. Variant PRC1 Complex-Dependent H2A Ubiquitylation Drives PRC2 Recruitment and Polycomb Domain Formation. *Cell* **157**: 1445–1459.

Blackledge NP, Rose NR, Klose RJ. 2015. Targeting Polycomb systems to regulate gene expression: modifications to a complex story. *Nat Rev Mol Cell Biol* **16**: 643–649.

Blewitt ME, Gendrel A-V, Pang Z, Sparrow DB, Whitelaw N, Craig JM, Apedaile A, Hilton DJ, Dunwoodie SL, Brockdorff N, et al. 2008. SmcHD1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation. *Nat Genet* **40**: 663–669.

Boettiger AN, Bintu B, Moffitt JR, Wang S, Beliveau BJ, Fudenberg G, Imakaev M, Mirny LA, Wu C-T, Zhuang X. 2016. Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* **529**: 418–422.

Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, Levine SS, Wernig M, Tajonar A, Ray MK, et al. 2006. Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441**: 349–353.

Brideau NJ, Coker H, Gendrel A-V, Siebert CA, Bezstarosti K, Demmers J, Poot RA, Nesterova TB, Brockdorff N. 2015. Independent Mechanisms Target SMCHD1 to Trimethylated Histone H3 Lysine 9-Modified Chromatin and the Inactive X Chromosome. *Mol Cell Biol* **35**: 4053–4068.

Brockdorff N. 2013. Noncoding RNA and Polycomb recruitment. *RNA* **19**: 429–442.

Brockdorff N, Ashworth A, Kay GF, Cooper P, Smith S, McCabe VM, Norris DP, Penny GD, Patel D, Rastan S. 1991. Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. *Nature* **351**: 329–331.

Brockdorff N, Ashworth A, Kay GF, McCabe VM, Norris DP, Cooper PJ, Swift S, Rastan S. 1992. The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell* **71**: 515–526.

Brown CJ, Ballabio A, Rupert JL, Lafrenière RG. 1991. A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* **349**: 38–44.

Brown CJ, Hendrich BD, Rupert JL, Lafrenière RG, Xing Y, Lawrence J, Willard HF. 1992. The human XIST gene: Analysis of a 17 kb inactive X-

specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell* **71**: 527–542.

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

Busturia A, Wightman CD, Sakonju S. 1997. A silencer is required for maintenance of transcriptional repression throughout Drosophila development. *Development* **124**: 4343–4350.

Cao R, Tsukada Y-I, Zhang Y. 2005. Role of Bmi-1 and Ring1A in H2A Ubiquitylation and Hox Gene Silencing. *Mol Cell* **20**: 845–854.

Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS, Zhang Y. 2002. Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing. *Science* **298**: 1039–1043.

Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK, Shia W-J, Anderson S, Yates J, Washburn MP, et al. 2005. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123**: 581–592.

Casanova M, Preissner T, Cerase A, Poot R, Yamada D, Li X, Appanah R, Bezstarosti K, Demmers J, Koseki H, et al. 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**: 1471–1482.

Chaligné R, Heard E. 2014. X-chromosome inactivation in development and cancer. *FEBS Lett* **588**: 2514–2522.

Chaumeil J, Okamoto I, Guggiari M, Heard E. 2002. Integrated kinetics of X chromosome inactivation in differentiating embryonic stem cells. *Cytogenet Genome Res* **99**: 75–84.

Chen CK, Blanco M, Jackson C, Aznauryan E, Ollikainen N, Surka C, Chow A, Cerase A, McDonel P, Guttman M. 2016. Xist recruits the X chromosome to the nuclear lamina to enable chromosome-wide silencing. *Science* **354**: 468–472.

Chen J, Nikolaitchik O, Singh J, Wright A, Bencsics CE, Coffin JM, Ni N, Lockett S, Pathak VK, Hu W-S. 2009. High efficiency of HIV-1 genomic RNA packaging and heterozygote formation revealed by single virion analysis. *Proc Natl Acad Sci USA* **106**: 13535–13540.

Chen S, Ma J, Wu F, Xiong LJ, Ma H, Xu W, Lv R, Li X, Villen J, Gygi SP, et al. 2012. The histone H3 Lys 27 demethylase JMJD3 regulates gene expression by impacting transcriptional elongation. *Genes Dev* **26**: 1364–1375.

Cheutin T, McNairn AJ, Jenuwein T, Gilbert DM, Singh PB, Misteli T. 2003. Maintenance of stable heterochromatin domains by dynamic HP1 binding. *Science* **299**: 721–725.

Choo JH, Kim JD, Chung JH, Stubbs L, Kim J. 2006. Allele-specific deposition of macroH2A1 in imprinting control regions. *Hum Mol Genet* **15**: 717–724.

Chu C, Zhang QC, da Rocha ST, Flynn RA, Bharadwaj M, Calabrese JM, Magnuson T, Heard E, Chang HY. 2015. Systematic Discovery of Xist RNA Binding Proteins. *Cell* **161**: 404–416.

Clemson CM, McNeil JA, Willard HF, Lawrence JB. 1996. XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. *J Cell Biol* **132**: 259–275.

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

Comet I, Schuettengruber B, Sexton T, Cavalli G. 2011. A chromatin insulator driving three-dimensional Polycomb response element (PRE) contacts and Polycomb association with the chromatin fiber. *Proc Natl Acad Sci USA* **108**: 2294–2299.

Cooper S, Dienstbier M, Hassan R, Schermelleh L, Sharif J, Blackledge NP, De Marco V, Elderkin S, Koseki H, Klose R, et al. 2014. Targeting Polycomb to Pericentric Heterochromatin in Embryonic Stem Cells Reveals a Role for H2AK119u1 in PRC2 Recruitment. *Cell Rep* **7**: 1456–1470.

Cooper S, Grijzenhout A, Underwood E, Ancelin K, Zhang T, Nesterova TB, Anil-Kirmizitas B, Bassett A, Kooistra SM, Agger K, et al. 2016. Jarid2 binds mono-ubiquitylated H2A lysine 119 to mediate crosstalk between Polycomb complexes PRC1 and PRC2. *Nat Commun* **7**: 13661.

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

Costanzi C, Pehrson JR. 1998. Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature* **393**: 599–601.

Creppe C, Palau A, Malinverni R, Valero V, Buschbeck M. 2014. A Cbx8-Containing Polycomb Complex Facilitates the Transition to Gene Activation during ES Cell Differentiation ed. R. Eskeland. *PLoS Genet* **10**: e1004851.

Csankovszki G, Panning B, Bates B, Pehrson JR, Jaenisch R. 1999. Conditional deletion of Xist disrupts histone macroH2A localization but not maintenance of X inactivation. *Nat Genet* **22**: 323–324.

Czypionka A, de los Paños OR, Mateu MG, Barrera FN, Hurtado-Gómez E, Gómez J, Vidal M, Neira JL. 2007. The isolated C-terminal domain of Ring1B is a dimer made of stable, well-structured monomers. *Biochemistry* **46**: 12764–12776.

da Rocha ST, Boeva V, Escamilla-Del-Arenal M, Ancelin K, Granier C, Matias NR, Sanulli S, Chow J, Schulz E, Picard C, et al. 2014. Jarid2 Is Implicated in the Initial Xist-Induced Targeting of PRC2 to the Inactive X Chromosome. *Mol Cell* **53**: 301–316.

Daigle N, Ellenberg J. 2007. LambdaN-GFP: an RNA reporter system for live-cell imaging. *Nat Methods* **4**: 633–636.

Davidovich C, Zheng L, Goodrich KJ, Cech TR. 2013. Promiscuous RNA binding by Polycomb repressive complex 2. *Nat Struct Mol Biol* **20**: 1250–1257.

de Napoles M, Mermoud JE, Wakao R, Tang YA, Endoh M, Appanah R,

Nesterova TB, Silva J, Otte AP, Vidal M, et al. 2004. Polycomb Group Proteins Ring1A/B Link Ubiquitylation of Histone H2A to Heritable Gene Silencing and X Inactivation. *Dev Cell* **7**: 663–676.

Deaton AM, Bird A. 2011. CpG islands and the regulation of transcription. *Genes Dev* **25**: 1010–1022.

Declerck N, Minh NL, Yang Y, Bloch V, Kochoyan M, Aymerich S. 2002. RNA recognition by transcriptional antiterminators of the BglG/SacY family: mapping of SacY RNA binding site. *J Mol Biol* **319**: 1035–1048.

Denholtz M, Bonora G, Chronis C, Splinter E, de Laat W, Ernst J, Pellegrini M, Plath K. 2013. Long-range chromatin contacts in embryonic stem cells reveal a role for pluripotency factors and polycomb proteins in genome organization. *Cell Stem Cell* **13**: 602–616.

Diaz-Perez SV, Ferguson DO, Wang C, Csankovszki G, Wang C, Tsai S-C, Dutta D, Perez V, Kim S, Eller CD, et al. 2006. A deletion at the mouse Xist gene exposes trans-effects that alter the heterochromatin of the inactive X chromosome and the replication time and DNA stability of both X chromosomes. *Genetics* **174**: 1115–1133.

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

Dignam JD, Lebovitz RM, Roeder RG. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei.

Nucleic Acids Res **11**: 1475–1489.

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

Dowen JM, Fan ZP, Hnisz D, Ren G, Abraham BJ, Zhang LN, Weintraub AS, Schuijers J, Lee TI, Zhao K, et al. 2014. Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* **159**: 374–387.

Doyen C-M, An W, Angelov D, Bondarenko V, Mietton F, Studitsky VM, Hamiche A, Roeder RG, Bouvet P, Dimitrov S. 2006. Mechanism of polymerase II transcription repression by the histone variant macroH2A. *Mol Cell Biol* **26**: 1156–1164.

Duncan IM. 1982. Polycomblike: a gene that appears to be required for the normal expression of the bithorax and antennapedia gene complexes of Drosophila melanogaster. *Genetics* **102**: 49–70.

Dupont C, Gribnau J. 2013. Different flavors of X-chromosome inactivation in mammals. *Curr Opin Cell Biol* **25**: 314–321.

Duthie S. 1999. Xist RNA exhibits a banded localization on the inactive X chromosome and is excluded from autosomal material in cis. *Hum Mol Genet* **8**: 195–204.

Eils R, Dietzel S, Bertin E, Schröck E, Speicher MR, Ried T, Robert-Nicoud M, Cremer C, Cremer T. 1996. Three-dimensional reconstruction of painted human interphase chromosomes: active and inactive X chromosome

territories have similar volumes but differ in shape and surface structure. *J Cell Biol* **135**: 1427–1440.

Elderkin S, Maertens GN, Endoh M, Mallery DL, Morrice N, Koseki H, Peters G, Brockdorff N, Hiom K. 2007. A Phosphorylated Form of Mel-18 Targets the Ring1B Histone H2A Ubiquitin Ligase to Chromatin. *Mol Cell* **28**: 107–120.

Endoh M, Endo TA, Endoh T, Fujimura YI, Ohara O, Toyoda T, Otte AP, Okano M, Brockdorff N, Vidal M, et al. 2008. Polycomb group proteins Ring1A/B are functionally linked to the core transcriptional regulatory circuitry to maintain ES cell identity. *Development* **135**: 1513–1524.

Endoh M, Endo TA, Endoh T, Isono K-I, Sharif J, Ohara O, Toyoda T, Ito T, Eskeland R, Bickmore WA, et al. 2012. Histone H2A mono-ubiquitination is a crucial step to mediate PRC1-dependent repression of developmental genes to maintain ES cell identity. ed. D. Schübeler. *PLoS Genet* **8**: e1002774.

Engreitz JM, Ollikainen N, Guttman M. 2016. Long non-coding RNAs: spatial amplifiers that control nuclear structure and gene expression. *Nat Rev Mol Cell Biol* **17**: 756–770.

Engreitz JM, Pandya-Jones A, McDonel P, Shishkin A, Sirokman K, Surka C, Kadri S, Xing J, Goren A, Lander ES, et al. 2013. The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* **341**: 1237973–1237973.

Escamilla-Del-Arenal M, da Rocha ST, Spruijt CG, Masui O, Renaud O, Smits AH, Margueron R, Vermeulen M, Heard E. 2013. Cdyl, a New Partner of the Inactive X Chromosome and Potential Reader of H3K27me3 and H3K9me2.

Mol Cell Biol **33**: 5005–5020.

Eskeland R, Leeb M, Grimes GR, Kress C, Boyle S, Sproul D, Gilbert N, Fan Y, Skoultchi AI, Wutz A, et al. 2010. Ring1B Compacts Chromatin Structure and Represses Gene Expression Independent of Histone Ubiquitination. *Mol Cell* **38**: 452–464.

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

Fackelmayer FO. 2004. A Stable Proteinaceous Structure in the Territory of Inactive X Chromosomes. *J Biol Chem* **280**: 1720–1723.

Fang J, Chen T, Chadwick B, Li E, Zhang Y. 2004. Ring1b-mediated H2A ubiquitination associates with inactive X chromosomes and is involved in initiation of X inactivation. *J Biol Chem* **279**: 52812–52815.

Farcas AM, Blackledge NP, Sudbery I, Long HK, McGouran JF, Rose NR, Lee S, Sims D, Cerase A, Sheahan TW, et al. 2012. KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands. *Elife* **1**: e00205.

Ferrari KJ, Scelfo A, Jammula S, Cuomo A, Barozzi I, Stützer A, Fischle W, Bonaldi T, Pasini D. 2014. Polycomb-dependent H3K27me1 and H3K27me2 regulate active transcription and enhancer fidelity. *Mol Cell* **53**: 49–62.

Fischle W, Wang Y, Jacobs SA, Kim Y, Allis CD, Khorasanizadeh S. 2003. Molecular basis for the discrimination of repressive methyl-lysine marks in

histone H3 by Polycomb and HP1 chromodomains. *Genes Dev* **17**: 1870–1881.

Fleming AB, Kao C-F, Hillyer C, Pikaart M, Osley MA. 2008. H2B Ubiquitylation Plays a Role in Nucleosome Dynamics during Transcription Elongation. *Mol Cell* **31**: 57–66.

Francis NJ, Kingston RE, Woodcock CL. 2004. Chromatin compaction by a polycomb group protein complex. *Science* **306**: 1574–1577.

Frenster JH, Allfrey VG, Mirsky AE. 1963. Repressed and Active Chromatin Isolated From Interphase Lymphocytes. *Proc Natl Acad Sci USA* **50**: 1026–1032.

Frey F, Sheahan T, Finkl K, Stoehr G, Mann M, Benda C, Müller J. 2016. Molecular basis of PRC1 targeting to Polycomb response elements by PhoRC. *Genes Dev* **30**: 1116–1127.

Fuks F, Hurd PJ, Deplus R, Kouzarides T. 2003. The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Res* **31**: 2305–2312.

Galupa R, Heard E. 2015. X-chromosome inactivation: new insights into cis and trans regulation. *Current Opinion in Genetics & Development* **31**: 57–66.

Gao Z, Lee P, Stafford JM, Schimmelmann von M, Schaefer A, Reinberg D. 2014. An AUTS2–Polycomb complex activates gene expression in the CNS. *Nature* **516**: 349–354.

Gao Z, Zhang J, Bonasio R, Strino F, Sawai A, Parisi F, Kluger Y, Reinberg

D. 2012. PCGF Homologs, CBX Proteins, and RYBP Define Functionally Distinct PRC1 Family Complexes. *Mol Cell* **45**: 344–356.

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

Geisler SJ, Paro R. 2015. Trithorax and Polycomb group-dependent regulation: a tale of opposing activities. *Development* **142**: 2876–2887.

Gendrel A-V, Apedaile A, Coker H, Termanis A, Zvetkova I, Godwin J, Tang YA, Huntley D, Montana G, Taylor S, et al. 2012. Smchd1-dependent and -independent pathways determine developmental dynamics of CpG island methylation on the inactive X chromosome. *Dev Cell* **23**: 265–279.

Gonzalez I, Aparicio R, Busturia A. 2008. Functional Characterization of the dRYBP Gene in Drosophila. *Genetics* **179**: 1373–1388.

Gonzalez I, Mateos-Langerak J, Thomas A, Cheutin T, Cavalli G. 2014. Identification of regulators of the three-dimensional polycomb organization by a microscopy-based genome-wide RNAi screen. *Mol Cell* **54**: 485–499.

Grau DJ, Chapman BA, Garlick JD, Borowsky M, Francis NJ, Kingston RE. 2011. Compaction of chromatin by diverse Polycomb group proteins requires localized regions of high charge. *Genes Dev* **25**: 2210–2221.

Grijzenhout A, Godwin J, Koseki H, Gdula MR, Szumska D, McGouran JF, Bhattacharya S, Kessler BM, Brockdorff N, Cooper S. 2016. Functional

analysis of AEBP2, a PRC2 Polycomb protein, reveals a Trithorax phenotype in embryonic development and in ESCs. *Development* **143**: 2716–2723.

Grossniklaus U, Paro R. 2014. Transcriptional silencing by polycomb-group proteins. *Cold Spring Harb Perspect Biol* **6**: a019331–a019331.

Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, Eussen BH, de Klein A, Wessels L, de Laat W, et al. 2008. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* **453**: 948–951.

Hansen KH, Bracken AP, Pasini D, Dietrich N, Gehani SS, Monrad A, Rappsilber J, Lerdrup M, Helin K. 2008. A model for transmission of the H3K27me3 epigenetic mark. *Nat Cell Biol* **10**: 1291–1300.

Hauri S, Comoglio F, Seimiya M, Gerstung M, Glatter T, Hansen K, Aebersold R, Paro R, Gstaiger M, Beisel C. 2016. A High-Density Map for Navigating the Human Polycomb Complexome. *Cell Rep* **17**: 583–595.

He J, Shen L, Wan M, Taranova O, Wu H, Zhang Y. 2013. Kdm2b maintains murine embryonic stem cell status by recruiting PRC1 complex to CpG islands of developmental genes. *Nat Cell Biol* **15**: 373–384.

Heard E, Rougeulle C, Arnaud D, Avner P, Allis CD, Spector DL. 2001. Methylation of histone H3 at Lys-9 is an early mark on the X chromosome during X inactivation. *Cell* **107**: 727–738.

Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA, et al. 2007. Distinct and predictive chromatin

signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet* **39**: 311–318.

Hennig L, Bouveret R, Gruissem W. 2005. MSI1-like proteins: an escort service for chromatin assembly and remodeling complexes. *Trends in Cell Biology* **15**: 295–302.

Hernandez-Munoz I, Taghavi P, Kuijl C, Neefjes J, van Lohuizen M. 2005. Association of BMI1 with Polycomb Bodies Is Dynamic and Requires PRC2/EZH2 and the Maintenance DNA Methyltransferase DNMT1. *Mol Cell Biol* **25**: 11047–11058.

Herranz N, Pasini D, Díaz VM, Francí C, Gutiérrez A, Dave N, Escrivà M, Hernandez-Muñoz I, Di Croce L, Helin K, et al. 2008. Polycomb complex 2 is required for E-cadherin repression by the Snail1 transcription factor. *Mol Cell Biol* **28**: 4772–4781.

Hisada K, Sanchez C, Endo TA, Endoh M, Roman-Trufero M, Sharif J, Koseki H, Vidal M. 2012. RYBP Represses Endogenous Retroviruses and Preimplantation- and Germ Line-Specific Genes in Mouse Embryonic Stem Cells. *Mol Cell Biol* **32**: 1139–1149.

Holdeman R, Nehrt S, Strome S. 1998. MES-2, a maternal protein essential for viability of the germline in Caenorhabditis elegans, is homologous to a Drosophila Polycomb group protein. *Development* **125**: 2457–2467.

Hon GC, Hawkins RD, Ren B. 2009. Predictive chromatin signatures in the mammalian genome. *Hum Mol Genet* **18**: R195–R201.

Hunkapiller J, Shen Y, Diaz A, Cagney G, McCleary D, Ramalho-Santos M, Krogan N, Ren B, Song JS, Reiter JF. 2012. Polycomb-like 3 promotes polycomb repressive complex 2 binding to CpG islands and embryonic stem cell self-renewal. ed. G. Cavalli. *PLoS Genet* **8**: e1002576.

Illingworth RS, Moffat M, Mann AR, Read D, Hunter CJ, Pradeepa MM, Adams IR, Bickmore WA. 2015. The E3 ubiquitin ligase activity of RING1B is not essential for early mouse development. *Genes Dev* **29**: 1897–1902.

Irminger-Finger I, Nöthiger R. 1995. The Drosophila melanogaster gene lethal(3)73Ah encodes a ring finger protein homologous to the oncoproteins MEL-18 and BMI-1. *Gene* **163**: 203–208.

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

Jackle H, Hoch M, Pankratz MJ, Gerwin N, Sauer F, Bronner G. 1992. Transcriptional control by Drosophila gap genes. *Journal of Cell Science* **1992**: 39–51.

Jaenisch R, Bird A. 2003. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* **33**: 245–254.

Jiao L, Liu X. 2015. Structural basis of histone H3K27 trimethylation by an active polycomb repressive complex 2. *Science* **350**: aac4383–aac4383.

Junco SE, Wang R, Gaipa JC, Taylor AB, Schirf V, Gearhart MD, Bardwell

VJ, Demeler B, Hart PJ, Kim CA. 2013. Structure of the polycomb group protein PCGF1 in complex with BCOR reveals basis for binding selectivity of PCGF homologs. *Structure* **21**: 665–671.

Justin N, Zhang Y, Tarricone C, Martin SR, Chen S, Underwood E, De Marco V, Haire LF, Walker PA, Reinberg D, et al. 2016. Structural basis of oncogenic histone H3K27M inhibition of human polycomb repressive complex 2. *Nat Commun* **7**: 11316.

Jürgens G. 1985. A group of genes controlling the spatial expression of the bithorax complex in Drosophila. *Nature* **316**: 153–155.

Kahn TG, Dorafshan E, Schultheis D, Zare A, Stenberg P, Reim I, Pirrotta V, Schwartz YB. 2016. Interdependence of PRC1 and PRC2 for recruitment to Polycomb Response Elements. *Nucleic Acids Res* gkw701.

Kalantry S, Mills KC, Yee D, Otte AP, Panning B, Magnuson T. 2006. The Polycomb group protein Eed protects the inactive X-chromosome from differentiation-induced reactivation. *Nat Cell Biol* **8**: 195–202.

Kalb R, Latwiel S, Baymaz HI, Jansen PWTC, Müller CW, Vermeulen M, Müller J. 2014. Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression. *Nat Struct Mol Biol* **21**: 569–571.

Kaneko S, Son J, Bonasio R, Shen SS, Reinberg D. 2014. Nascent RNA interaction keeps PRC2 activity poised and in check. *Genes Dev* **28**: 1983–1988.

Kaneko S, Son J, Shen SS, Reinberg D, Bonasio R. 2013. PRC2 binds active

promoters and contacts nascent RNAs in embryonic stem cells. *Nat Struct Mol Biol* **20**: 1258–1264.

Kanhere A, Viiri K, Araújo CC, Rasaiyaah J, Bouwman RD, Whyte WA, Pereira CF, Brookes E, Walker K, Bell GW, et al. 2010. Short RNAs Are Transcribed from Repressed Polycomb Target Genes and Interact with Polycomb Repressive Complex-2. *Mol Cell* **38**: 675–688.

Kassis JA, Brown JL. 2013. Polycomb group response elements in Drosophila and vertebrates. *Adv Genet* **81**: 83–118.

Kennison JA. 2004. Introduction to Trx-G and Pc-G genes. *Meth Enzymol* **377**: 61–70.

Kennison JA. 1995. The Polycomb and trithorax group proteins of Drosophila: trans-regulators of homeotic gene function. *Annu Rev Genet* **29**: 289–303.

Kim D-H, Sung S. 2014. Polycomb-Mediated Gene Silencing in Arabidopsis thaliana. *Molecules and Cells* **37**: 841–850.

Kim H, Kang K, Kim J. 2009. AEBP2 as a potential targeting protein for Polycomb Repression Complex PRC2. *Nucleic Acids Res* **37**: 2940–2950.

Kim T-G, Kraus JC, Chen J, Lee Y. 2003. JUMONJI, a critical factor for cardiac development, functions as a transcriptional repressor. *J Biol Chem* **278**: 42247–42255.

Kloet SL, Makowski MM, Baymaz HI, van Voorthuijsen L, Karemaker ID, Santanach A, Jansen PWTC, Di Croce L, Vermeulen M. 2016. The dynamic interactome and genomic targets of Polycomb complexes during stem-cell

differentiation. *Nat Struct Mol Biol* **23**: 682–690.

Kohlmaier A, Savarese F, Lachner M, Martens J, Jenuwein T, Wutz A. 2004. A Chromosomal Memory Triggered by Xist Regulates Histone Methylation in X Inactivation ed. Peter Becker. *PLoS Biol* **2**: e171.

Kondo T, Isono K, Kondo K, Endo TA, Itohara S, Vidal M, Koseki H. 2014. Polycomb potentiates meis2 activation in midbrain by mediating interaction of the promoter with a tissue-specific enhancer. *Dev Cell* **28**: 94–101.

Kouzarides T. 2007. Chromatin Modifications and Their Function. *Cell* **128**: 693–705.

Kuzmichev A, Jenuwein T, Tempst P, Reinberg D. 2004. Different EZH2-containing complexes target methylation of histone H1 or nucleosomal histone H3. *Mol Cell* **14**: 183–193.

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

Lagarou A, Mohd-Sarip A, Moshkin YM, Chalkley GE, Bezstarosti K, Demmers JAA, Verrijzer CP. 2008. dKDM2 couples histone H2A ubiquitylation to histone H3 demethylation during Polycomb group silencing. *Genes Dev* **22**: 2799–2810.

Landeira D, Sauer S, Poot R, Dvorkina M, Mazzarella L, Jørgensen HF, Pereira CF, Leleu M, Piccolo FM, Spivakov M, 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. *Nat Cell Biol* **12**: 618–624.

Lanzuolo C, Roure V, Dekker J, Bantignies F, Orlando V. 2007. Polycomb response elements mediate the formation of chromosome higher-order structures in the bithorax complex. *Nat Cell Biol* **9**: 1167–1174.

Lee C-K, Shibata Y, Rao B, Strahl BD, Lieb JD. 2004. Evidence for nucleosome depletion at active regulatory regions genome-wide. *Nat Genet* **36**: 900–905.

Lee H-G, Kahn TG, Simcox A, Schwartz YB, Pirrotta V. 2015. Genome-wide activities of Polycomb complexes control pervasive transcription. *Genome Research* **25**: 1170–1181.

Lee JT, Jaenisch R. 1997. Long-range cis effects of ectopic X-inactivation centres on a mouse autosome. *Nature* **386**: 275–279.

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

Leeb M, Wutz A. 2007. Ring1B is crucial for the regulation of developmental control genes and PRC1 proteins but not X inactivation in embryonic cells. *J Cell Biol* **178**: 219–229.

Lehnertz B, Ueda Y, Derijck AAHA, Braunschweig U, Perez-Burgos L, Kubicek S, Chen T, Li E, Jenuwein T, Peters AHFM. 2003. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite

repeats at pericentric heterochromatin. *Curr Biol* **13**: 1192–1200.

Lewis EB. 1978. A gene complex controlling segmentation in Drosophila. *Nature* **276**: 565–570.

Lewis PW, Müller MM, Koletsky MS, Cordero F, Lin S, Banaszynski LA, Garcia BA, Muir TW, Becher OJ, Allis CD. 2013. Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma. *Science* **340**: 857–861.

Li G, Margueron R, Ku M, Chambon P, Bernstein BE, Reinberg D. 2010. Jarid2 and PRC2, partners in regulating gene expression. *Genes Dev* **24**: 368–380.

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

Li Z, Cao R, Wang M, Myers MP, Zhang Y, Xu R-M. 2006. Structure of a Bmi-1-Ring1B polycomb group ubiquitin ligase complex. *J Biol Chem* **281**: 20643–20649.

Liefke R, Karwacki-Neisius V, Shi Y. 2016. EPOP Interacts with Elongin BC and USP7 to Modulate the Chromatin Landscape. *Mol Cell* **64**: 659–672.

Lo SM, Francis NJ. 2010. Inhibition of chromatin remodeling by polycomb group protein posterior sex combs is mechanistically distinct from nucleosome

binding. *Biochemistry* **49**: 9438–9448.

Lu X, Simon MD, Chodaparambil JV, Hansen JC, Shokat KM, Luger K. 2008. The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure. *Nat Struct Mol Biol* **15**: 1122–1124.

Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**: 251–260.

Lyon MF. 1961. Gene action in the X-chromosome of the mouse (Mus musculus L.). *Nature* **190**: 372–373.

Maenner S, Blaud M, Fouillen L, Savoye A, Marchand V, Dubois A, Sanglier-Cianférani S, Van Dorsselaer A, Clerc P, Avner P, et al. 2010. 2-D Structure of the A Region of Xist RNA and Its Implication for PRC2 Association ed. K. Hall. *PLoS Biol* **8**: e1000276.

Mak W, Baxter J, Silva J, Newall AE, Otte AP, Brockdorff N. 2002. Mitotically Stable Association of Polycomb Group Proteins Eed and Enx1 with the Inactive X Chromosome in Trophoblast Stem Cells. *Current Biology* **12**: 1016–1020.

Mak W, Nesterova TB, de Napoles M, Appanah R, Yamanaka S, Otte AP, Brockdorff N. 2004. Reactivation of the paternal X chromosome in early mouse embryos. *Science* **303**: 666–669.

Marahrens Y, Panning B, Dausman J, Strauss W, Jaenisch R. 1997. Xist-deficient mice are defective in dosage compensation but not

spermatogenesis. *Genes Dev* **11**: 156–166.

Margueron R, Justin N, Ohno K, Sharpe ML, Son J, Drury WJ, Voigt P, Martin SR, Taylor WR, De Marco V, et al. 2009. Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature* **461**: 762–767.

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

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

Marks H, Kerstens HHD, Barakat TS, Splinter E, Dirks RAM, van Mierlo G, Joshi O, Wang S-Y, Babak T, Albers CA, et al. 2015. Dynamics of gene silencing during X inactivation using allele-specific RNA-seq. *Genome Biol* 1–20.

Martin C, Cao R, Zhang Y. 2006. Substrate preferences of the EZH2 histone methyltransferase complex. *J Biol Chem* **281**: 8365–8370.

Maunakea AK, Nagarajan RP, Bilenky M, Ballinger TJ, D'Souza C, Fouse SD, Johnson BE, Hong C, Nielsen C, Zhao Y, et al. 2010. Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* **466**: 253–257.

McGinty RK, Henrici RC, Tan S. 2014. Crystal structure of the PRC1 ubiquitylation module bound to the nucleosome. *Nature* **514**: 591–596.

McHugh CA, Chen C-K, Chow A, Surka CF, Tran C, McDonel P, Pandya-

Jones A, Blanco M, Burghard C, Moradian A, et al. 2015. The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* **521**: 232–236.

Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB, et al. 2008. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature*.

Mendenhall EM, Koche RP, Truong T, Zhou VW, Issac B, Chi AS, Ku M, Bernstein BE. 2010. GC-rich sequence elements recruit PRC2 in mammalian ES cells. ed. H.D. Madhani. *PLoS Genet* **6**: e1001244.

Meshorer E, Misteli T. 2006. Chromatin in pluripotent embryonic stem cells and differentiation. *Nat Rev Mol Cell Biol* **7**: 540–546.

Meshorer E, Yellajoshula D, George E, Scambler PJ, Brown DT, Misteli T. 2006. Hyperdynamic Plasticity of Chromatin Proteins in Pluripotent Embryonic Stem Cells. *Dev Cell* **10**: 105–116.

Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim T-K, Koche RP, et al. 2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**: 553–560.

Minajigi A, Froberg JE, Wei C, Sunwoo H, Kesner B, Colognori D, Lessing D, Payer B, Boukhali M, Haas W, et al. 2015. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science* **349**.

Minks J, Baldry SE, Yang C, Cotton AM, Brown CJ. 2013. XIST-induced silencing of flanking genes is achieved by additive action of repeat a monomers in human somatic cells. *Epigenetics Chromatin* **6**: 23.

Moazed D. 2013. Chromatin: a tail of repression. *Curr Biol* **23**: R456–9.

Mohd-Sarip A, Venturini F, Chalkley GE, Verrijzer CP. 2002. Pleiohomeotic can link polycomb to DNA and mediate transcriptional repression. *Mol Cell Biol* **22**: 7473–7483.

Mohn F, Weber M, Rebhan M, Roloff TC, Richter J, Stadler MB, Bibel M, Schübeler D. 2008. Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol Cell* **30**: 755–766.

Moindrot B, Cerase A, Coker H, Masui O, Grijzenhout A, Pintacuda G, Schermelleh L, Nesterova TB, Brockdorff N. 2015. A Pooled shRNA Screen Identifies Rbm15, Spen, and Wtap as Factors Required for Xist RNA-Mediated Silencing. *Cell Rep* **12**: 562–572.

Monfort A, Di Minin G, Postlmayr A, Freimann R, Arieti F, Thore S, Wutz A. 2015. Identification of Spen as a Crucial Factor for Xist Function through Forward Genetic Screening in Haploid Embryonic Stem Cells. *Cell Rep* **12**: 554–561.

Morey L, Aloia L, Cozzuto L, Benitah SA, Di Croce L. 2013. RYBP and Cbx7 Define Specific Biological Functions of Polycomb Complexes in Mouse Embryonic Stem Cells. *Cell Rep* **3**: 60–69.

Morey L, Brenner C, Fazi F, Villa R, Gutiérrez A, Buschbeck M, Nervi C, Minucci S, Fuks F, Di Croce L. 2008. MBD3, a component of the NuRD complex, facilitates chromatin alteration and deposition of epigenetic marks. *Mol Cell Biol* **28**: 5912–5923.

Morey L, Pascual G, Cozzuto L, Roma G, Wutz A, Benitah SA, Di Croce L. 2012. Nonoverlapping functions of the Polycomb group Cbx family of proteins in embryonic stem cells. *Cell Stem Cell* **10**: 47–62.

Morey L, Santanach A, Blanco E, Aloia L, Nora EP, Bruneau BG, Di Croce L. 2015. Polycomb Regulates Mesoderm Cell Fate-Specification in Embryonic Stem Cells through Activation and Repression Mechanisms. *Cell Stem Cell* **17**: 300–315.

Mousavi K, Zare H, Wang AH, Sartorelli V. 2012. Polycomb Protein Ezh1 Promotes RNA Polymerase II Elongation. *Mol Cell* **45**: 255–262.

Mozzetta C, Pontis J, Fritsch L, Robin P, Portoso M, Proux C, Margueron R, Ait-Si-Ali S. 2014. The histone H3 lysine 9 methyltransferases G9a and GLP regulate polycomb repressive complex 2-mediated gene silencing. *Mol Cell* **53**: 277–289.

Mueller F, Karpova TS, Mazza D, McNally JG. 2012. Monitoring dynamic binding of chromatin proteins in vivo by fluorescence recovery after photobleaching. *Methods Mol Biol* **833**: 153–176.

Murzina NV, Pei X-Y, Zhang W, Sparkes M, Vicente-Garcia J, Pratap JV, McLaughlin SH, Ben-Shahar TR, Verreault A, Luisi BF, et al. 2008. Structural basis for the recognition of histone H4 by the histone-chaperone RbAp46.

Structure **16**: 1077–1085.

Müller J, Kassis JA. 2006. Polycomb response elements and targeting of Polycomb group proteins in Drosophila. *Current Opinion in Genetics & Development* **16**: 476–484.

Nakagawa T, Kajitani T, Togo S, Masuko N, Ohdan H, Hishikawa Y, Koji T, Matsuyama T, Ikura T, Muramatsu M, et al. 2008. Deubiquitylation of histone H2A activates transcriptional initiation via trans-histone cross-talk with H3K4 di- and trimethylation. *Genes Dev* **22**: 37–49.

Nakanishi S, Lee JS, Gardner KE, Gardner JM, Takahashi Y-H, Chandrasekharan MB, Sun Z-W, Osley MA, Strahl BD, Jaspersen SL, et al. 2009. Histone H2BK123 monoubiquitination is the critical determinant for H3K4 and H3K79 trimethylation by COMPASS and Dot1. *J Cell Biol* **186**: 371–377.

Nekrasov M, Wild B, Müller J. 2005. Nucleosome binding and histone methyltransferase activity of Drosophila PRC2. *EMBO Rep* **6**: 348–353.

Ng HH, Ciccone DN, Morshead KB, Oettinger MA, Struhl K. 2003. Lysine-79 of histone H3 is hypomethylated at silenced loci in yeast and mammalian cells: A potential mechanism for position-effect variegation. *Proc Natl Acad Sci USA* **100**: 1820–1825.

Ng K, Daigle N, Bancaud A, Ohhata T, Humphreys P, Walker R, Ellenberg J, Wutz A. 2011. A system for imaging the regulatory noncoding Xist RNA in living mouse embryonic stem cells. *Molecular Biology of the Cell* **22**: 2634–2645.

Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piolot T, van Berkum NL, Meisig J, Sedat J, et al. 2012. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**: 381–385.

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

Ogawa H, Ishiguro K-I, Gaubatz S, Livingston DM, Nakatani Y. 2002. A complex with chromatin modifiers that occupies E2F- and Myc-responsive genes in G0 cells. *Science* **296**: 1132–1136.

Okamoto I, Otte AP, Allis CD, Reinberg D, Heard E. 2004. Epigenetic dynamics of imprinted X inactivation during early mouse development. *Science* **303**: 644–649.

Oktaba K, Gutiérrez L, Gagneur J, Girardot C, Sengupta AK, Furlong EEM, Müller J. 2008. Dynamic regulation by polycomb group protein complexes controls pattern formation and the cell cycle in Drosophila. *Dev Cell* **15**: 877–889.

Papp B, Müller J. 2006. Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins. *Genes Dev* **20**: 2041–2054.

Pasini D, Bracken AP, Jensen MR, Lazzerini Denchi E, Helin K. 2004. Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *EMBO J* **23**: 4061–4071.

Pasini D, Cloos PAC, Walfridsson J, Olsson L, Bukowski J-P, Johansen JV, Bak M, Tommerup N, Rappsilber J, Helin K. 2010a. JARID2 regulates binding of the Polycomb repressive complex 2 to target genes in ES cells. *Nature* **464**: 306–310.

Pasini D, Hansen KH, Christensen J, Agger K, Cloos PAC, Helin K. 2008. Coordinated regulation of transcriptional repression by the RBP2 H3K4 demethylase and Polycomb-Repressive Complex 2. *Genes Dev* **22**: 1345–1355.

Pasini D, Malatesta M, Jung HR, Walfridsson J, Willer A, Olsson L, Skotte J, Wutz A, Porse B, Jensen ON, et al. 2010b. Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes. *Nucleic Acids Res* **38**: 4958–4969.

Pasque V, Gillich A, Garrett N, Gurdon JB. 2011. Histone variant macroH2A confers resistance to nuclear reprogramming. *EMBO J* **30**: 2373–2387.

Passarge E. 1979. *Emil Heitz and the concept of heterochromatin: longitudinal chromosome differentiation was recognized fifty years ago.*

Patil DP, Chen C-K, Pickering BF, Chow A, Jackson C, Guttman M, Jaffrey SR. 2016. m(6)A RNA methylation promotes XIST-mediated transcriptional repression. *Nature* **537**: 369–373.

Patrat C, Okamoto I, Diabangouaya P, Vialon V, Le Baccon P, Chow J, Heard E. 2009. Dynamic changes in paternal X-chromosome activity during imprinted X-chromosome inactivation in mice. *Proc Natl Acad Sci USA* **106**:

5198–5203.

Pemberton H, Anderton E, Patel H, Brookes S, Chandler H, Palermo R, Stock J, Rodriguez-Niedenführ M, Racek T, de Breed L, et al. 2014. Genome-wide co-localization of Polycomb orthologs and their effects on gene expression in human fibroblasts. *Genome Biol* **15**: R23.

Peng JC, Valouev A, Swigut T, Zhang J, Zhao Y, Sidow A, Wysocka J. 2009. Jarid2/Jumonji coordinates control of PRC2 enzymatic activity and target gene occupancy in pluripotent cells. *Cell* **139**: 1290–1302.

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

Pengelly AR, Kalb R, Finkl K, Müller J. 2015. Transcriptional repression by PRC1 in the absence of H2A monoubiquitylation. *Genes Dev* **29**: 1487–1492.

Penny GD, Kay GF, Sheardown SA, Rastan S, Brockdorff N. 1996. Requirement for Xist in X chromosome inactivation. *Nature* **379**: 131–137.

Peric-Hupkes D, Meuleman W, Pagie L, Bruggeman SWM, Solovei I, Brugman W, Gräf S, Flicek P, Kerkhoven RM, van Lohuizen M, et al. 2010. Molecular Maps of the Reorganization of Genome-Nuclear Lamina Interactions during Differentiation. *Mol Cell* **38**: 603–613.

Peters AH, O'Carroll D, Scherthan H, Mechtler K, Sauer S, Schöfer C, Weipoltshammer K, Pagani M, Lachner M, Kohlmaier A, et al. 2001. Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin

and genome stability. *Cell* **107**: 323–337.

Peters AHFM, Kubicek S, Mechtler K, O'Sullivan RJ, Derijck AAHA, Perez-Burgos L, Kohlmaier A, Opravil S, Tachibana M, Shinkai Y, et al. 2003. Partitioning and Plasticity of Repressive Histone Methylation States in Mammalian Chromatin. *Mol Cell* **12**: 1577–1589.

Phair RD, Misteli T. 2000. High mobility of proteins in the mammalian cell nucleus. *Nature* **404**: 604–609.

Phair RD, Scaffidi P, Elbi C, Vecerova J, Dey A, Ozato K, Brown DT, Hager G, Bustin M, Misteli T. 2004. Global Nature of Dynamic Protein-Chromatin Interactions In Vivo: Three-Dimensional Genome Scanning and Dynamic Interaction Networks of Chromatin Proteins. *Mol Cell Biol* **24**: 6393–6402.

Pinter SF. 2016. A Tale of Two Cities: How Xist and its partners localize to and silence the bicompartmental X. *Seminars in Cell and Developmental Biology* **56**: 19–34.

Pinter SF, Sadreyev RI, Yildirim E, Jeon Y, Ohsumi TK, Borowsky M, Lee JT. 2012. Spreading of X chromosome inactivation via a hierarchy of defined Polycomb stations. *Genome Research* **22**: 1864–1876.

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

Plath K, Fang J, Mlynarczyk-Evans SK, Cao R, Worringer KA, Wang H, la Cruz de CC, Otte AP, Panning B, Zhang Y. 2003. Role of histone H3 lysine 27

methylation in X inactivation. *Science* **300**: 131–135.

Plath K, Talbot D, Hamer KM, Otte AP, Yang TP, Jaenisch R, Panning B. 2004. Developmentally regulated alterations in Polycomb repressive complex 1 proteins on the inactive X chromosome. *J Cell Biol* **167**: 1025–1035.

Pullirsch D, Hartel R, Kishimoto H, Leeb M, Steiner G, Wutz A. 2010. The Trithorax group protein Ash2l and Saf-A are recruited to the inactive X chromosome at the onset of stable X inactivation. *Development* **137**: 935–943.

Ran FA, Hsu PD, Lin C-Y, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y, et al. 2013a. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**: 1380–1389.

Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. 2013b. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8**: 2281–2308.

Rastan S, Brown SDM. 2009. The search for the mouse X-chromosome inactivation centre. *Genet Res* **56**: 99.

Reddington JP, Perricone SM, Nestor CE, Reichmann J, Youngson NA, Suzuki M, Reinhardt D, Dunican DS, Prendergast JG, Mjoseng H, et al. 2013. Redistribution of H3K27me3 upon DNA hypomethylation results in de-repression of Polycomb target genes. *Genome Biol* **14**: R25.

Reddy KL, Zullo JM, Bertolino E, Singh H. 2008. Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature* **452**: 243–247.

Rego A, Sinclair PB, Tao W, Kireev I, Belmont AS. 2008. The facultative heterochromatin of the inactive X chromosome has a distinctive condensed ultrastructure. *Journal of Cell Science* **121**: 1119–1127.

Ren X, Vincenz C, Kerppola TK. 2008. Changes in the Distributions and Dynamics of Polycomb Repressive Complexes during Embryonic Stem Cell Differentiation. *Mol Cell Biol* **28**: 2884–2895.

Reynolds N, Salmon-Divon M, Dvinge H, Hynes-Allen A, Balasooriya G, Leaford D, Behrens A, Bertone P, Hendrich B. 2012. NuRD-mediated deacetylation of H3K27 facilitates recruitment of Polycomb Repressive Complex 2 to direct gene repression. *EMBO J* **31**: 593–605.

Riising EM, Comet I, Leblanc B, Wu X, Johansen JV, Helin K. 2014. Gene silencing triggers polycomb repressive complex 2 recruitment to CpG islands genome wide. *Mol Cell* **55**: 347–360.

Riising EM, Helin K. 2012. A new role for the polycomb group protein Ezh1 in promoting transcription. *Mol Cell* **45**: 145–146.

Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E, et al. 2007. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129**: 1311–1323.

Rose NR, King HW, Blackledge NP, Fursova NA, Ember KJ, Fischer R, Kessler BM, Klose RJ. 2016. RYBP stimulates PRC1 to shape chromatin-based communication between Polycomb repressive complexes. *Elife* **5**: 524.

Rougeulle C, Chaumeil J, Sarma K, Allis CD, Reinberg D, Avner P, Heard E. 2004. Differential histone H3 Lys-9 and Lys-27 methylation profiles on the X chromosome. *Mol Cell Biol* **24**: 5475–5484.

Ryba T, Hiratani I, Lu J, Itoh M, Kulik M, Zhang J, Schulz TC, Robins AJ, Dalton S, Gilbert DM. 2010. Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome Research* **20**: 761–770.

Sanchez-Pulido L, Devos D, Sung ZR, Calonje M. 2008. RAWUL: a new ubiquitin-like domain in PRC1 ring finger proteins that unveils putative plant and worm PRC1 orthologs. *BMC Genomics* **9**: 308.

Sarma K, Cifuentes-Rojas C, Ergun A, del Rosario A, Jeon Y, White F, Sadreyev R, Lee JT. 2014. ATRX Directs Binding of PRC2 to Xist RNA and Polycomb Targets. *Cell* **159**: 869–883.

Sarma K, Levasseur P, Aristarkhov A, Lee JT. 2010. Locked nucleic acids (LNAs) reveal sequence requirements and kinetics of Xist RNA localization to the X chromosome. *Proc Natl Acad Sci USA* **107**: 22196–22201.

Sarma K, Margueron R, Ivanov A, Pirrotta V, Reinberg D. 2008. Ezh2 requires PHF1 to efficiently catalyze H3 lysine 27 trimethylation in vivo. *Mol Cell Biol* **28**: 2718–2731.

Schmiedeberg L, Skene P, Deaton A, Bird A. 2009. A Temporal Threshold for Formaldehyde Crosslinking and Fixation ed. J.Z. Rappoport. *PLoS ONE* **4**: e4636.

Schmitges FW, Prusty AB, Faty M, Stützer A, Lingaraju GM, Aiwazian J, Sack R, Hess D, Li L, Zhou S, et al. 2011. Histone methylation by PRC2 is inhibited by active chromatin marks. *Mol Cell* **42**: 330–341.

Schoeftner S, Sengupta AK, Kubicek S, Mechtler K, Spahn L, Koseki H, Jenuwein T, Wutz A. 2006. Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. *EMBO J* **25**: 3110–3122.

Schoenfelder S, Sugar R, Dimond A, Javierre B-M, Armstrong H, Mifsud B, Dimitrova E, Matheson L, Tavares-Cadete F, Furlan-Magaril M, et al. 2015. Polycomb repressive complex PRC1 spatially constrains the mouse embryonic stem cell genome. *Nat Genet* **47**: 1179–1186.

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

Schuettengruber B, Ganapathi M, Leblanc B, Portoso M, Jaschek R, Tolhuis B, van Lohuizen M, Tanay A, Cavalli G. 2009. Functional anatomy of polycomb and trithorax chromatin landscapes in Drosophila embryos. ed. R. Kingston. *PLoS Biol* **7**: e13.

Schumacher A, Faust C, Magnuson T. 1996. Positional cloning of a global regulator of anterior-posterior patterning in mice. *Nature* **384**: 648–648.

Seenundun S, Rampalli S, Liu Q-C, Aziz A, Palii C, Hong S, Blais A, Brand M, Ge K, Dilworth FJ. 2010. UTX mediates demethylation of H3K27me3 at muscle-specific genes during myogenesis. *EMBO J* **29**: 1401–1411.

Seibler J, Zevnik B, Küter-Luks B, Andreas S, Kern H, Hennek T, Rode A,

Heimann C, Faust N, Kauselmann G, et al. 2003. Rapid generation of inducible mouse mutants. *Nucleic Acids Res* **31**: e12.

Senner CE, Nesterova TB, Norton S, Dewchand H, Godwin J, Mak W, Brockdorff N. 2011. Disruption of a conserved region of Xist exon 1 impairs Xist RNA localisation and X-linked gene silencing during random and imprinted X chromosome inactivation. *Development* **138**: 1541–1550.

Shahbazian MD, Zhang K, Grunstein M. 2005. Histone H2B ubiquitylation controls processive methylation but not monomethylation by Dot1 and Set1. *Mol Cell* **19**: 271–277.

Shao Z, Raible F, Mollaaghababa R, Guyon JR, Wu CT, Bender W, Kingston RE. 1999. Stabilization of chromatin structure by PRC1, a Polycomb complex. *Cell* **98**: 37–46.

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

Silva J, Mak W, Zvetkova I, Appanah R, Nesterova TB, Webster Z, Peters AHFM, Jenuwein T, Otte AP, Brockdorff N. 2003. Establishment of Histone H3 Methylation on the Inactive X Chromosome Requires Transient Recruitment of Eed-Enx1 Polycomb Group Complexes. *Dev Cell* **4**: 481–495.

Simon J, Chiang A, Bender W, Shimell MJ, O'Connor M. 1993. Elements of the Drosophila bithorax complex that mediate repression by Polycomb group products. *Developmental Biology* **158**: 131–144.

Simon MD, Pinter SF, Fang R, Sarma K, Rutenberg-Schoenberg M, Bowman SK, Kesner BA, Maier VK, Kingston RE, Lee JT. 2013. High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. *Nature* **504**: 465–469.

Sing A, Pannell D, Karaiskakis A, Sturgeon K, Djabali M, Ellis J, Lipshitz HD, Cordes SP. 2009. A vertebrate Polycomb response element governs segmentation of the posterior hindbrain. *Cell* **138**: 885–897.

Smeets D, Markaki Y, Schmid VJ, Kraus F, Tattermusch A, Cerase A, Sterr M, Fiedler S, Demmerle J, Popken J, et al. 2014. Three-dimensional super-resolution microscopy of the inactive X chromosome territory reveals a collapse of its active nuclear compartment harboring distinct Xist RNA foci. *Epigenetics Chromatin* **7**: 8.

Sprague BL, Pego RL, Stavreva DA, McNally JG. 2004. Analysis of binding reactions by fluorescence recovery after photobleaching. *Biophys J* **86**: 3473–3495.

Srinivasan L, Pan X, Atchison ML. 2005. Transient requirements of YY1 expression for PcG transcriptional repression and phenotypic rescue. *J Cell Biochem* **96**: 689–699.

Stock JK, Giadrossi S, Casanova M, Brookes E, Vidal M, Koseki H, Brockdorff N, Fisher AG, Pombo A. 2007. Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. *Nat Cell Biol* **9**: 1428–1435.

Struhl G, Akam M. 1985. Altered distributions of Ultrabithorax transcripts in

extra sex combs mutant embryos of Drosophila. *EMBO J* **4**: 3259–3264.

Sun Z-W, Allis CD. 2002. Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* **418**: 104–108.

Sunwoo H, Wu JY, Lee JT. 2015. The Xist RNA-PRC2 complex at 20-nm resolution reveals a low Xist stoichiometry and suggests a hit-and-run mechanism in mouse cells. *Proc Natl Acad Sci USA* **112**: E4216–E4225.

Taherbhoy AM, Huang OW, Cochran AG. 2015. BMI1–RING1B is an autoinhibited RING E3 ubiquitin ligase. *Nat Commun* **6**: 7621.

Tang YA, Huntley D, Montana G, Cerase A, Nesterova TB, Brockdorff N. 2010. Efficiency of Xist-mediated silencing on autosomes is linked to chromosomal domain organisation. *Epigenetics Chromatin* **3**: 10.

Tavares L, Dimitrova E, Oxley D, Webster J, Poot R, Demmers J, Bezstarosti K, Taylor S, Ura H, Koide H, et al. 2012. RYBP-PRC1 Complexes Mediate H2A Ubiquitylation at Polycomb Target Sites Independently of PRC2 and H3K27me3. *Cell* **148**: 664–678.

Terranova R, Yokobayashi S, Stadler MB, Otte AP, van Lohuizen M, Orkin SH, Peters AHFM. 2008. Polycomb group proteins Ezh2 and Rnf2 direct genomic contraction and imprinted repression in early mouse embryos. *Dev Cell* **15**: 668–679.

Thiru A, Nietlispach D, Mott HR, Okuwaki M, Lyon D, Nielsen PR, Hirshberg M, Verreault A, Murzina NV, Laue ED. 2004. Structural basis of HP1/PXVXL motif peptide interactions and HP1 localisation to heterochromatin. *EMBO J*

23: 489–499.

Tie F, Banerjee R, Fu C, Stratton CA, Fang M, Harte PJ. 2016. Polycomb inhibits histone acetylation by CBP by binding directly to its catalytic domain. *Proc Natl Acad Sci USA* **113**: E744–53.

Tiwari VK, Cope L, McGarvey KM, Ohm JE, Baylin SB. 2008. A novel 6C assay uncovers Polycomb-mediated higher order chromatin conformations. *Genome Research* **18**: 1171–1179.

Trojer P, Cao AR, Gao Z, Li Y, Zhang J, Xu X, Li G, Losson R, Erdjument-Bromage H, Tempst P, et al. 2011. L3MBTL2 Protein Acts in Concert with PcG Protein-Mediated Monoubiquitination of H2A to Establish a Repressive Chromatin Structure. *Mol Cell* **42**: 438–450.

Trojer P, Reinberg D. 2007. Facultative Heterochromatin: Is There a Distinctive Molecular Signature? *Mol Cell* **28**: 1–13.

Tsai M-C, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, Shi Y, Segal E, Chang HY. 2010. Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **329**: 689–693.

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

van Lohuizen M, Frasch M, Wientjens E, Berns A. 1991. Sequence similarity

between the mammalian bmi-1 proto-oncogene and the Drosophila regulatory genes Psc and Su(z)2. *Nature* **353**: 353–355.

Verreault A, Kaufman PD, Kobayashi R, Stillman B. 1996. Nucleosome Assembly by a Complex of CAF-1 and Acetylated Histones H3/H4. *Cell* **87**: 95–104.

Vethantham V, Yang Y, Bowman C, Asp P, Lee J-H, Skalnik DG, Dynlacht BD. 2012. Dynamic loss of H2B ubiquitylation without corresponding changes in H3K4 trimethylation during myogenic differentiation. *Mol Cell Biol* **32**: 1044–1055.

Voigt P, Tee WW, Reinberg D. 2013. A double take on bivalent promoters. *Genes Dev* **27**: 1318–1338.

Voncken JW, Roelen BAJ, Roefs M, de Vries S, Verhoeven E, Marino S, Deschamps J, van Lohuizen M. 2003. Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. *Proc Natl Acad Sci USA* **100**: 2468–2473.

Wang H, Wang L, Erdjument-Bromage H, Vidal M, Tempst P, Jones RS, Zhang Y. 2004a. Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431**: 873–878.

Wang J, Mager J, Chen Y, Schneider E, Cross JC, Nagy A, Magnuson T. 2001. Imprinted X inactivation maintained by a mouse Polycomb group gene. *Nat Genet* **28**: 371–375.

Wang L, Brown JL, Cao R, Zhang Y, Kassis JA, Jones RS. 2004b.

Hierarchical Recruitment of Polycomb Group Silencing Complexes. *Mol Cell* **14**: 637–646.

Wang R, Taylor AB, Leal BZ, Chadwell LV, Ilangovan U, Robinson AK, Schirf V, Hart PJ, Lafer EM, Demeler B, et al. 2010. Polycomb Group Targeting through Different Binding Partners of RING1B C-Terminal Domain. *Structure* **18**: 966–975.

Wani AH, Boettiger AN, Schorderet P, Ergun A, Münger C, Sadreyev RI, Zhuang X, Kingston RE, Francis NJ. 2016. Chromatin topology is coupled to Polycomb group protein subnuclear organization. *Nat Commun* **7**: 10291.

Wong SJ, Gearhart MD, Taylor AB, Nanyes DR, Ha DJ, Robinson AK, Artigas JA, Lee OJ, Demeler B, Hart PJ, et al. 2016. KDM2B Recruitment of the Polycomb Group Complex, PRC1.1, Requires Cooperation between PCGF1 and BCORL1. *Structure* **24**: 1795–1801.

Woo CJ, Kharchenko PV, Daheron L, Park PJ, Kingston RE. 2010. A region of the human HOXD cluster that confers polycomb-group responsiveness. *Cell* **140**: 99–110.

Wu X, Johansen JV, Helin K. 2013. Fbxl10/Kdm2b recruits polycomb repressive complex 1 to CpG islands and regulates H2A ubiquitylation. *Mol Cell* **49**: 1134–1146.

Wutz A, Jaenisch R. 2000. A Shift from Reversible to Irreversible X Inactivation Is Triggered during ES Cell Differentiation. *Mol Cell* **5**: 695–705.

Wutz A, Rasmussen TP, Jaenisch R. 2002. Chromosomal silencing and

localization are mediated by different domains of Xist RNA. *Nat Genet* **30**: 167–174.

Yap KL, Li S, Muñoz-Cabello AM, Raguz S, Zeng L, Mujtaba S, Gil J, Walsh MJ, Zhou M-M. 2010. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol Cell* **38**: 662–674.

Yu M, Mazor T, Huang H, Huang H-T, Kathrein KL, Woo AJ, Chouinard CR, Labadorf A, Akie TE, Moran TB, et al. 2012. Direct recruitment of polycomb repressive complex 1 to chromatin by core binding transcription factors. *Mol Cell* **45**: 330–343.

Yuan W, Wu T, Fu H, Dai C, Wu H, Liu N, Li X, Xu M, Zhang Z, Niu T, et al. 2012. Dense chromatin activates Polycomb repressive complex 2 to regulate H3 lysine 27 methylation. *Science* **337**: 971–975.

Zhang L-F, Huynh KD, Lee JT. 2007. Perinucleolar targeting of the inactive X during S phase: evidence for a role in the maintenance of silencing. *Cell* **129**: 693–706.

Zhang T, Cooper S, Brockdorff N. 2015. The interplay of histone modifications - writers that read. *EMBO Rep* **16**: 1467–1481.

Zhang W, Tyl M, Ward R, Sobott F, Maman J, Murthy AS, Watson AA, Fedorov O, Bowman A, Owen-Hughes T, et al. 2013. Structural plasticity of histones H3-H4 facilitates their allosteric exchange between RbAp48 and ASF1. *Nat Struct Mol Biol* **20**: 29–35.

Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT. 2008. Polycomb Proteins Targeted by a Short Repeat RNA to the Mouse X Chromosome. *Science* **322**: 750–756.

Zhou W, Zhu P, Wang J, Pascual G, Ohgi KA, Lozach J, Glass CK, Rosenfeld MG. 2008. Histone H2A Monoubiquitination Represses Transcription by Inhibiting RNA Polymerase II Transcriptional Elongation. *Mol Cell* **29**: 69–80.

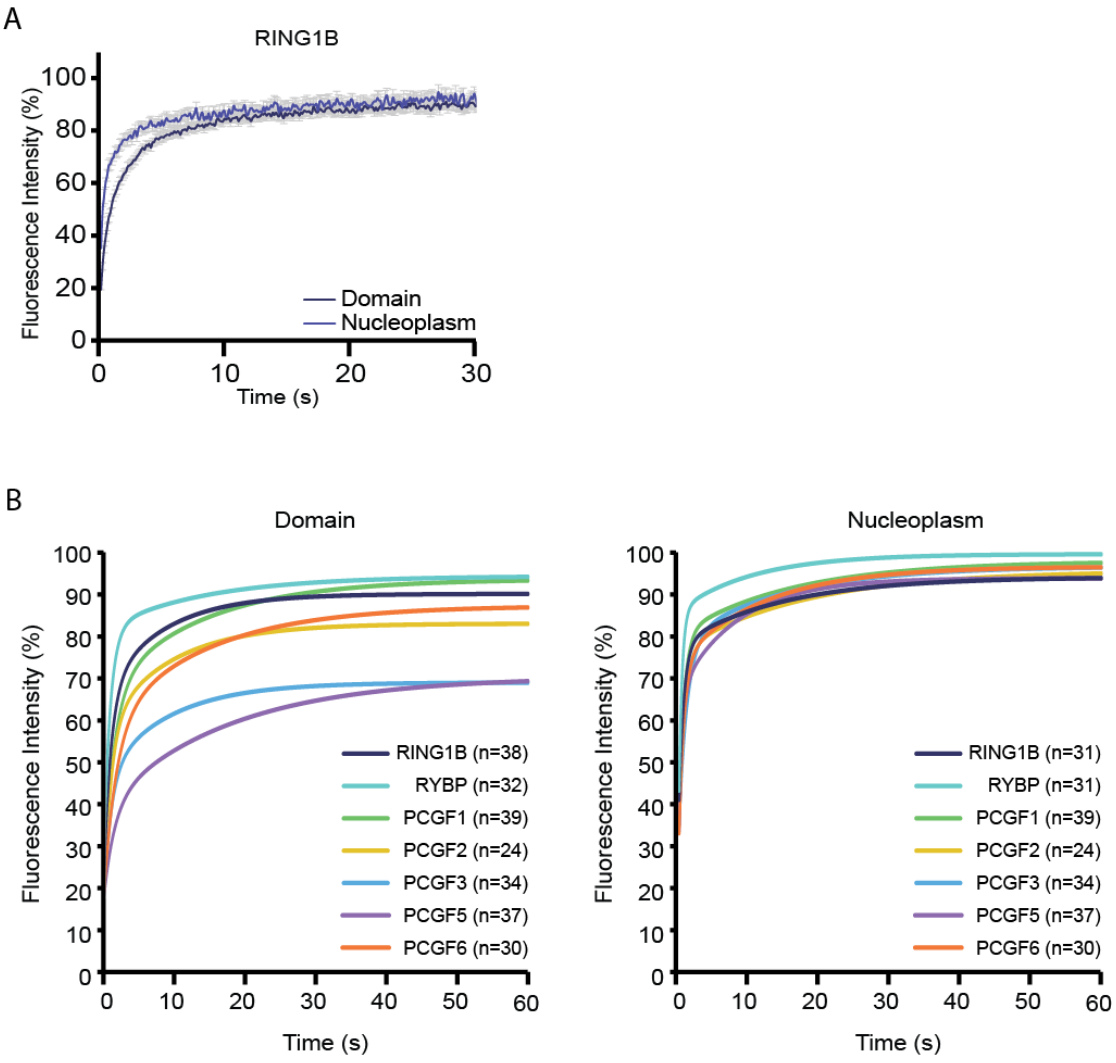


Figure 47: Fluorescence recovery after photobleaching curves for eGFP-RING1B in the nucleoplasm and the inactive chromosome domain. (A) Average fluorescence recovery curves for eGFP-RING1B over the first 30 seconds after photobleaching show an initial slower recovery over the inactive chromosome domain when compared to the nucleoplasm. A minimum of 31 cells were analyzed for each region of interest. Grey bars represent standard error. (B) Fluorescence recovery data for eGFP-RING1B in single cells was fitted to a bi-exponential model for quantitative analysis. Average fitted curves for eGFP-RING1B on the inactive chromosome domain (on the left) and on the nucleoplasm (on the right) are shown in dark blue. eGFP-PCGF and eGFP-RYBP data is shown for comparison. Number of cells analyzed is showed in brackets (n).

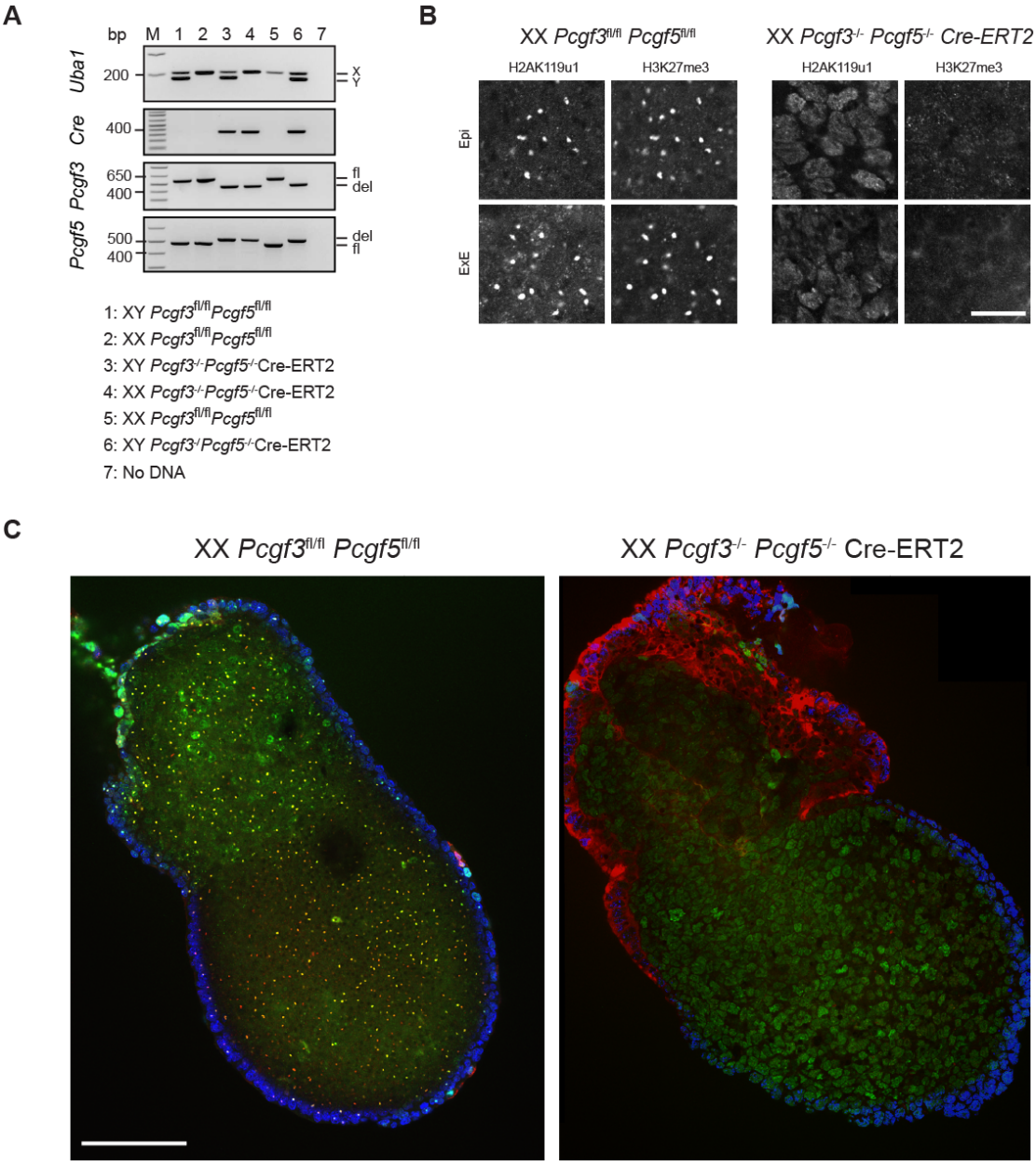


Figure 48: Effect of *Pcgf3/5* knockout in Polycomb recruitment by Xist RNA *in vivo*. (A) Genotyping of E7.5 embryos after immunofluorescence analysis. Representative examples are shown for *Uba1*, *Cre-ERT2* (*Cre*), *Pcgf3* and *Pcgf5*. PCR with *Uba1* primers results in two bands in males (XY) and one band in females (XX). Pooled F1, F2 and R1 primers were used for both *Pcgf3* and *Pcgf5* genotyping (fl – floxed allele; del – deleted allele). (B) Bar chart illustrating representation of different genotypes and the ratio of male and female E7.5 embryos recovered for immunofluorescence analysis after IVF, *in vitro* 4-hydroxytamoxifen treatment and transplantation into foster females. Mixed genotype column demonstrates the number of samples

showing presence of *Pcgf5* deleted allele together with wt allele due to maternal tissue contamination. (C) Immunofluorescence analysis of H3K27me3 and H2AK119u1 domains in *Pcgf3/5* dKO female embryo and its floxed littermate at E7.5. Examples show tiled single focal plane confocal scans through the entire embryo (60x objective). Scale bar is 100 µm. (Figure provided by Dr. Tatyana Nesterova)

Table 6: *p*-values resulting from Student t-test for pairwise comparisons of scoring data for H2AK119u1 and H3K27me3 domains overlapping with Xist in all cell lines generated in *Pcgf1*$^{fl/fl}$ background.

H2AK119u1 / H3K27me3	*Pcgf1*$^{fl/fl}$ +dox	*Pcgf1*$^{fl/fl}$ +OHT +dox	*Pcgf1*$^{fl/fl}$ *Pcgf5*$^{-/-}$ +dox	*Pcgf1*$^{fl/fl}$ *Pcgf5*$^{-/-}$ +OHT +dox	*Pcgf1*$^{fl/fl}$ *Pcgf6*$^{-/-}$ +dox	*Pcgf1*$^{fl/fl}$ *Pcgf6*$^{-/-}$ +OHT +dox
Pcgf1$^{fl/fl}$ +dox		N.S.	N.S.	N.S.	N.S.	N.S.
Pcgf1$^{fl/fl}$ +OHT +dox	N.S.		p<0.05	p<0.05	N.S.	p<0.05
Pcgf1$^{fl/fl}$ *Pcgf5*$^{-/-}$ +dox	p<0.01	p<0.01		N.S.	N.S.	p<0.05
Pcgf1$^{fl/fl}$ *Pcgf5*$^{-/-}$ +OHT +dox	p<0.001	p<0.001	p<0.05		N.S.	N.S.
Pcgf1$^{fl/fl}$ *Pcgf6*$^{-/-}$ +dox	N.S.	N.S.	p<0.05	p<0.001		N.S.
Pcgf1$^{fl/fl}$ *Pcgf6*$^{-/-}$ +OHT +dox	p<0.05	p<0.05	N.S.	p<0.001	N.S.	

Table 7: *p*-values resulting from Student t-test for pairwise comparisons of scoring data for H2AK119u1 and H3K27me3 domains overlapping with Xist in all cell lines generated in *Pcgf3*$^{fl/fl}$ background.

H2AK119u1 / H3K27me3	*Pcgf3*$^{fl/fl}$ +dox	*Pcgf3*$^{fl/fl}$ +OHT +dox	*Pcgf3*$^{fl/fl}$ *Pcgf5*$^{-/-}$ +dox	*Pcgf3*$^{fl/fl}$ *Pcgf5*$^{-/-}$ +OHT +dox	*Pcgf3*$^{fl/fl}$ *Pcgf6*$^{-/-}$ +dox	*Pcgf3*$^{fl/fl}$ *Pcgf6*$^{-/-}$ +OHT +dox
Pcgf3$^{fl/fl}$ +dox		p<0.05	N.S.	p<0.001	N.S.	N.S.
Pcgf3$^{fl/fl}$ +OHT +dox	p<0.001		p<0.01	p<0.001	p<0.01	p<0.05
Pcgf3$^{fl/fl}$ *Pcgf5*$^{-/-}$ +dox	N.S.	p<0.001		p<0.001	N.S.	p<0.05
Pcgf3$^{fl/fl}$ *Pcgf5*$^{-/-}$ +OHT +dox	p<0.001	p<0.001	p<0.001		p<0.001	p<0.001
Pcgf3$^{fl/fl}$ *Pcgf6*$^{-/-}$ +dox	p<0.01	p<0.001	p<0.001	p<0.001		p<0.05
Pcgf3$^{fl/fl}$ *Pcgf6*$^{-/-}$ +OHT +dox	p<0.05	N.S.	p<0.05	p<0.001	p<0.001	