

Disregulation of PRC2 function caused by the JAZF1-SUZ12 fusion protein

Doctoral thesis

Manuel de Jesus Tavares Cornejo, M.Sc., B.Sc.

PhD Cancer Institute

Supervisor:

Prof. Richard G. Jenner

UCL Cancer Institute

Declaration

I, Manuel de Jesus Tavares Cornejo, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis

Abstract

Polycomb repressive complex 2 (PRC2) maintains developmental genes specific for other cell types in a repressed state through methylation of histone H3 lysine 27 and formation of a repressive chromatin state. PRC2 is frequently mutated or dysregulated in cancer. In addition to binding to chromatin, PRC2 also binds RNA and this prevents its interaction with chromatin. In endometrial stromal sarcoma (ESS), chromosomal translocation fuses the first 128 amino acids of JAZF1 in place of the N-terminal 93 amino acids of the PRC2 subunit SUZ12. However, the effects this has on PRC2 function are unknown. Here, I show that the JAZF1-SUZ12 fusion protein prevents association of PRC2 with the accessory subunits JARID2, EPOP and PALI1. Moreover, I demonstrate that JAZF1-SUZ12 prevents the transfer of PRC2 from RNA to chromatin. Consistent with this defect being due to loss of EPOP and JARID2 interaction, PRC2 is also unable to transfer from RNA to chromatin in JARID2, EPOP or PALI1 deficient embryonic stem cells. Finally, I show that JAZF1-SUZ12 reduces the binding of PRC2 to its target genes and disrupts regulation of key developmental genes during differentiation of mouse embryonic stem cells. Thus, this work reveals molecular defects in PRC2 function caused by the JAZF1-SUZ12 fusion protein and suggests that changes in the balance between PRC2 RNA and chromatin binding may play a role in oncogenesis.

Impact statement

The mechanisms that confer carcinogenicity to cells are complex. In this work I study a mutation, a product of a chromosomal translocation, that is present in around 60% of low grade endometrial stromal sarcomas (LG-ESS), a rare type of uterine cancer. This translocation generates a fusion of two proteins: SUZ12, a subunit of the regulator of embryonic development polycomb repressive complex 2 (PRC2), and Juxtaposed with another zinc finger protein 1 (JAZF1), a transcriptional factor.

By using mouse embryonic stem cells (mESC) as a model of study, I describe potential mechanisms that might explain how this mutation transforms cells into malignant cells. I show that the fusion of JAZF1 to SUZ12 inhibits the formation of holo-PRC2, and that this might be responsible for inhibiting the recruitment of PRC2 to its target genes. Thus, this work reveals the molecular defects in PRC2 function caused by the fusion of JAZF1 to SUZ12, thereby opening the door to potential therapies that could correct these defects in the future.

Additionally, I identified the PRC2 accessory factors that are necessary for the recruitment of PRC2 from RNA to chromatin, which could be targeted to block PRC2 recruitment to genes to control cell state and to inhibit oncogenic or other detrimental gene silencing events.

Also, I studied the physiological effect of three compounds pyridostatin (PDS), 5,10,15,20-tetra(N-methyl-4-pyridyl)porphyrin (TMPyP4) and (4,9-bis((3-(4-methylpiperazin-1-yl)propyl)amino)-2,7-bis (3-morpholinopropyl) benzo[Imn]

[3,8] phenanthroline-1,3,6,8(2H,7H)-tetrone (MM41) on the interaction of PRC2 with RNA in cells. Although each compound binds G-quadruplex RNA, the RNA structure bound most strongly by PRC2, the compounds have different effects on PRC2 in cells. These results could provide insight into the toxic effects of these compounds in cancer models.

Publication

Beltran, M.; Tavares, M. *et al.* (2019) 'G-tract RNA removes Polycomb repressive complex 2 from genes', *Nature Structural & Molecular Biology*. doi: 10.1038/s41594-019-0293-z.

This publication contains Figures 3 G and Supplementary Figures 4E and 4F in Chapter 4.

Acknowledgements

First and foremost, I would like to thank my primary supervisor, Richard, without whom this work could not have been possible. His continuous support and keen observations were essential for the work presented here.

Also, thanks to Prof. Shaun Cowley and Dr. David Michod, for making time in their schedule to act as the examiners of my viva.

I would also like to thank the funding bodies that supported me throughout my PhD, namely the *Consejo Nacional de Ciencia y Tecnologia* (CONACyT) in Mexico, UCL Overseas Research Scholarship and the UCL Cancer Institute.

Special thanks go to Manolo, whose work was seminal for my project, and whose experimental and personal support were fundamental for my PhD.

Thanks also to Maria, my desk buddy, for the technical support and for making my PhD experience merrier.

Also, thanks to all the members of the 507, for sharing their expertise and for their helpful comments in the lab meetings: Arnulf, Garima, Vicky, Lenka and Chris.

Additionally, I would like to thank the UCL Cancer Institute staff and students but specially the fifth-floor crew for all the help and laughs.

Also, I would like to thank my London family, particularly the UCL Mexican Society and members of the Goodenough College, which made me feel like home from day one.

Finally, I would like to thank Ana and my parents. You were truly the foundation of my PhD.

To Ana. Thanks for being my inspiration and my strength.

A mis papas. Por su apoyo y por ser siempre mi modelo a seguir.

Table of contents

1. Chapter 1: Introduction.	22
1.1. Chromatin structure and its modifications.....	22
1.1.1 Canonical and variant nucleosomes	23
1.1.2. DNA methylation as a repressive mark.....	25
1.1.3. Histone post-translational modifications.....	28
1.1.4. Chromatin remodelling.....	31
1.2. Polycomb proteins	33
1.2.1. PRC1 composition	34
1.2.2. PRC2 catalytic core	35
1.2.2.1 EZH1/2	36
1.2.2.2 SUZ12	37
1.2.2.3 EED	38
1.2.3 PRC2 accessory factors	38
1.2.3.1 JARID2	40
1.2.3.2 AEBP2	41
1.2.3.3. EPOP.....	42
1.2.3.4. PCLs.....	42
1.2.3.5. PALI1/2.....	43
1.2.3.6. EZHIP	43
1.2.4. Three-dimensional structure of PRC2.....	45
1.2.5. Recruitment of PRC2.....	45
1.3. RNA and PRC2	48
1.3.1. PRC2 RNA binding specificity.....	48

1.3.1. PRC2 binds RNA through a number of different subunits.....	50
1.3.3. RNA regulates the recruitment of PRC2 to chromatin.....	51
1.4. PRC2 and cancer	54
1.4.1. Mutations in EZH2 are often found in cancers	54
1.4.2 Inhibitors of PRC2 activity have anti-tumorigenic properties.....	55
1.4.3. Endometrial stromal sarcoma	56
1.4.3.1 JAZF1-SUZ12.....	57
1.4.3.1. Endometrial stromal cells and decidualisation	59
1.5. Aims of this thesis	61
2. Chapter 2: Materials and Methods.....	62
2.1. DNA cloning	62
2.1.1. DNA constructs and sequencing primers used in this work	62
2.1.1. PCR amplification of SUZ12 sequences.....	63
2.1.2. Generation of N-terminal FS2-tagged constructs.....	63
2.1.3. Generation of pMY constructs	64
2.1.4. DNA transformation	65
2.1.5 Colony screening by PCR and plasmid purification	65
2.2. Cell culture	66
2.2.1. mESCs.....	66
2.2.2. Other cell lines	67
2.2.3. Primary human endometrial stromal cells (hEnSCs)	67
2.2.4. Generation of stable mESC lines.....	70
2.2.5. Transduction of immortalized fibroblasts.....	72
2.2.6. Immunofluorescence assays	72

2.2.7. Treatment with G4 binding compounds	73
2.2.8. Embryoid Body (EB) Formation	73
2.2.9. Transient transfection of hEnSCs.	74
2.2.10. Transfection, selection and decidualisation of hEnSCs	75
2.3. Protein work.	76
2.3.1. SDS PAGE and Immunoblot.....	76
2.3.2. Protein immunoprecipitation (IP) from cytoplasmic and nucleoplasmic fractions	78
2.3.3. Protein IP from whole cell extracts.....	79
2.3.4. Streptactin pull-down of FS2-tagged proteins.	80
2.3.5. Chromatin Immunoprecipitation (ChIP).....	81
2.3.6. Nuclear fractionation and RNaseA treatment.....	84
2.3.7. Nucleosome pulldowns.	84
2.4.5. CLIP	87
2.4. Measuring gene expression	88
2.4.1. RNA purification	88
2.4.2. cDNA synthesis	89
3. Chapter 3: Interaction of JAZF1-SUZ12 with PRC2 accessory factors. .91	
3.1. Introduction.....	91
3.2. JAZF1-SUZ12 and SUZ12 Δ 93 recover H3K27me3 in Suz12 ^{GT/GT} mESC	93
3.3. JAZF1-SUZ12 and SUZ12 Δ 93 partly localize to the cytoplasm.....	97

3.4. The SUZ12 N-terminus is necessary for interaction with a specific set of accessory factors	99
3.5. JAZF1-SUZ12 does not interact with the NuA4 complex in 3T3 cells ..	104
3.6. hEnSCs express accessory subunits EPOP and JARID2 Δ N	105
3.7. Summary and discussion	107
4. Chapter 4: Regulation of PRC2 by RNA in cells is modulated by accessory subunits.	110
4.1. Introduction.....	110
4.2. JARID2 but not AEBP2 are necessary for recruitment of PRC2 to chromatin upon RNA degradation	112
4.3. EPOP and PALI1 but not PCLs are necessary for recruitment of PRC2.1 to chromatin upon RNA degradation	114
4.4. JAZF1-SUZ12 is not recruited to chromatin upon RNA degradation	116
4.5. Accessory subunits do not affect recruitment of PRC2 to nucleosomes upon RNA degradation	117
4.6. Effect of PDS in PRC2 and RNA binding in mESCs.....	124
4.7. TMPyP4 and MM41 show pleiotropic effects and increased PRC2 RNA binding in mESCs, respectively	129
4.8. Summary and discussion	132
5. Chapter 5: Effects of JAZF1-SUZ12 on cell differentiation.....	135
5.1. Introduction.....	135

5.2. Overexpression of SUZ12, but not JAZF1-SUZ12, generates more colonies in immortalized fibroblasts.....	137
5.3. JAZF1-SUZ12 promotes differentiation towards endodermal lineages during differentiation of ES cells into EBs.....	140
5.4. In vitro decidualization of transfected hEnSCs.	145
5.5. Transfected of hEnSCs with SUZ12 constructs.....	146
5.8. Summary and discussion	149
6. Chapter 6. Conclusions and future work	151
6.1. H3K27me3 is recovered by JAZF1-SUZ12 in Suz12 ^{GT/GT} cells.	152
6.2. Interactions of PRC2 with a subset of accessory factors is dependent on the SUZ12 N-terminus.....	152
6.3. hEnSCs contain EPOP and JARID2 Δ N.....	154
6.4. EPOP, JARID2 and PALI1 are necessary for the recruitment of PRC2 to chromatin when RNA is depleted.	154
6.5. Nucleosomes compete with RNA for binding to core PRC2.	155
6.6. G4 RNA binding compounds modulate PRC2 RNA binding in cells and reduce SUZ12 stability.	156
6.7. JAZF1-SUZ12 alters gene expression during EB formation.	157
6.8. hEnSCs are a promising model for the study of LG-ESS.	158
References.....	160

List of figures

Figure 1.1. Known PRC2 variants and interactions made by accessory subunits.	40
Figure 1.2. Summary of interplay between PRC2.1 and PRC2.2 with PRC1	47
Figure 1.3. Current model of antagonism between RNA and chromatin for PRC2 binding,	53
Figure 1.4. Structure of SUZ12 and JAZF1 and their fusion in LG-ESS.	58
Figure 2.1. EB formation schedule.....	74
Figure 2.2. Schematic of the generation of decidualised transfected hEnSCs..	75
Figure 3.1. JAZF1-SUZ12 rescues H3K27me3 in SUZ12-null cells.....	95
Figure 3.2. JAZF1-SUZ12 and SUZ12 Δ 93 do not bind to PRC2 target genes in <i>Suz12</i> ^{GT/GT}	96
Figure 3.3. JAZF1-SUZ12 and SUZ12 Δ 93 localize to the nucleus and cytoplasm.	98
Figure 3.4. EPOP and JARID2 do not interact with SUZ12 Δ 93 and JAZF1-SUZ12 in mESCs.....	100
Figure 3.5. The first 93 amino acids of SUZ12 are needed for interaction with EPOP, but not with AEBP2 or PCL3.	102
Figure 3.6. PALI1 does not interact with SUZ12 Δ 93 and JAZF1-SUZ12.	103
Figure 3.7. JAZF1-SUZ12 does not bind to the NuA4 complex in 3T3 cells. ...	105
Figure 3.8 Accessory subunits EPOP and JARID2 are expressed in hEnSCs.	106
Figure 3.9 Model describing the main results of Chapter 3.....	109
Figure 4.1. JARID2, but not AEBP2, is necessary for PRC2 recruitment to chromatin upon RNA degradation.....	113

Figure 4.2. EPOP is necessary for recruitment of PRC2.1 to chromatin upon release from RNA.	114
Figure 4.3. Lack of PALI1, but not PCLs, reduces SUZ12 recruitment to chromatin under RnaseA treatment.	115
Figure 4.4. JAZF1-SUZ12 is not recruited to chromatin upon RNA depletion.	116
Figure 4.5 Release of PRC2 from RNA triggers its recruitment to nucleosomes.	118
Figure 4.6 AEBP2 is not necessary for PRC2 to bind nucleosomes upon its release from RNA.	119
Figure 4.7 Accessory subunits JARID2 and AEBP2 are not necessary for PRC2 binding to nucleosomes.	121
Figure 4.8 PCLs are not necessary for PRC2 binding to nucleosomes.	123
Figure 4.9 PDS treatment reduces binding of PRC2 to RNA in mESC.	125
Figure 4.10. PDS affects PRC2 localization in mESCs, but not its binding to target genes.	127
Figure 4.11 CLIP for SUZ12, FUS and HNRNPC in mESCs treated with a different batch of PDS does not recapitulate inhibition of PRC2 RNA binding.	128
Figure 4.12 TMPyP4 induces degradation of SUZ12 in mESCs.	130
Figure 4.13 MM41 increases binding of PRC2 to RNA in mESCs.	131
Figure 4.14 Model of the main findings of Chapter 4.	134
Figure 5.1. Generation of immortalized fibroblast cell lines expressing FLAG-tagged SUZ12 constructs.	138
Figure 5.2. SUZ12 but not JAZF1-SUZ12 promotes colony formation in hTERT and SV40 fibroblasts.	139

Figure 5.3. Defective formation of EBs in ESC expressing SUZ12 Δ 93 or JAZF1-SUZ12.....141

Figure 5.4. JAZF1-SUZ12 and SUZ12 Δ 93 have limited effect on the repression of pluripotency markers during differentiation of *Suz12*^{GT/GT} ESC into EBs. ...142

Figure 5.5. Mesoderm genes need the N-terminus of SUZ12 during EB formation to be expressed.143

Figure 5.6. JAZF1-SUZ12 promotes differentiation of mESC towards an endodermal lineage.144

Figure 5.7. hEnSCs are transfectable and undergo decidualization after selection with puromycin.145

Figure 5.8. SUZ12 transgenes do not affect decidualisation in hEnSCs.147

Figure 5.9. FS2-tagged proteins are only expressed at low levels following transient transfection of hEnSCs.....148

List of tables

Table 1.1 Canonical and variant histones.	25
Table 1.2. Selected histone modifications in vertebrates relevant for this work and their function.	31
Table 1.3. Summary of Polycomb complex subunits and their functions.	44
Table 1.4. Gene fusion events in LG-ESS.	59
Table 2.1. Summary of the constructs used in this work.	62
Table 2.2. Primers used in this work.	62
Table 2.3. Culture media composition.	69
Table 2.4. Antibodies used for immunoblotting.	77
Table 2.5. Primers used for CHIP-qPCR.	83
Table 2.6. Primers used to quantify RNA.	90

Abbreviations

5-mC	5-methylcytosine
AML	Acute myeloid lymphoblastic leukemia
bp	Base pair
BCOR	BCL6 co-repressor
BPTF	Bromodomain and PHD finger-containing transcription factor
BRD	Bromodomain-containing protein
CENPA	Centromere associated protein A
CGI	CpG island
ChIP	Chromatin Immunoprecipitation
CBP	CREB binding protein
DLBCL	Diffuse large cell B-cell lymphoma
DNMT	DNA methyltransferase
EB	Embryoid body
EED	Embryonic ectoderm development
EPOP	Elongin BC and PRC2 associated protein
EST	Endometrial stromal tumour
ESN	Endometrial stromal nodule
ESS	Endometrial stromal sarcoma
EZH	Enhancer of zeste

EZH1P	Enhancer of zeste homologs inhibitory protein
FUS	Fused under sarcoma
HA-tag	Human influenza hemagglutinin tag
HAT	Histone acetyl transferase
hEnSCs	Human endometrial stem cells
HP1	Heterochromatin protein 1
hr	Hour
JARID2	Jumonji/ARID domain-containing protein 2
lncRNA	Long non-coding RNA
LSD1	Lysine specific demethylase 1
mESCs	Mouse embryonic stem cells
MM41	(4,9-bis((3-(4-methylpiperazin-1-yl)propyl)amino)-2,7-bis (3-morpholinopropyl) benzo[lmn][3,8] phenanthroline-1,3,6,8(2H,7H)-tetraone
MLL	Mixed-lineage leukaemia
NuA4	Nucleosome acetyltransferase of histone H4
NURF	Nucleosomes remodelling factor
PAR-CLIP	photoactivatable-ribonucleoside-enhanced crosslinking and immunoprecipitation
PALI	PRC2 associated LCOR isoform

PcG	Polycomb group
PCL	Polycomb like protein
PDS	Pyridostatin
PHD	Plant homeodomain
PRC	Polycomb repressive complex
PSC	Posterior sex combs
RBBP4	retinoblastoma-binding protein
RNAPII	RNA polymerase II
SUZ12	Suppressor of zeste 12
SAHF	Senescent associated heterochromatic foci
TET	Ten-eleven translocation
TMPyP4	5,10,15,20-tetra(N-methyl-4-pyridyl)porphyrin
TrxG	Trithorax group
TSS	Transcription start site

1. Chapter 1: Introduction.

1.1. Chromatin structure and its modifications

Chromatin was first described in 1879 as a threadlike material that was stained by aniline based dyes within the nucleus of salamander embryos (Fleming 1879). Decades later, differentially stained regions of chromatin were reported. These were termed as heterochromatin and euchromatin that described condensed and decondensed states of chromatin, respectively (Heitz 1928).

Afterwards, it was suggested that heterochromatin could be divided into two categories: constitutive and facultative. It was proposed that the latter was developmentally regulated, and that only one allele of an homologous chromosome pair was compacted during X inactivation (Brown, 1966).

The term chromatin was employed later to describe the complex of histones and DNA (Kornberg, 1974). After the discovery that the interaction of proteins with DNA provided a means of compaction, it became evident that DNA-binding proteins acted as obstacles to the transcription machinery to access DNA, and thus chromatin became recognised to regulate gene expression (Li, Carey and Workman, 2007) .

1.1.1 Canonical and variant nucleosomes

The nucleosome is the primary unit of chromatin and the repeated array of nucleosomes and linker DNA in between form the least compacted arrangement of chromatin, the euchromatin fibre (Campbell, Cotter and Pardon, 1978). Each nucleosome is formed from two molecules of histone proteins H2A, H2B, H3 and H4. H3 and H4 form a (H3/H4)₂ tetramer and interacting with this are two heterodimers of H2A/H2B, which dock at the DNA entry and exit sites through the H2A C-terminus docking domain. Additionally, the two H2A histones interact through their L1 loop, and H2B interacts with H4 through a weak four-helix bundle (Weber and Henikoff, 2014).

The histone octamer is then surrounded by 146 bp (1.7 turns) of DNA (Luger *et al.*, 1997), and nucleosomes are spaced along the genome with an average of 20-90 bp of non-nucleosomal linker DNA in between, depending on the species and cell type (van Holde and Zlatanova, 1995). An additional “linker” histone H1 is sometimes present and binds to the “entry” and “exit” point of DNA to the nucleosome and guides the orientation of linker DNA within the nucleosome. When present *in vitro*, H1 also promotes a higher level of condensation by compacting the chromatin fibre even further, giving its characteristic solenoid structure of 30 nm in width (Bednar *et al.*, 1998).

Apart from the canonical histone proteins, variants of these histones can be found exerting specific functions. H2A.Z enrichment correlates with increased expression levels. In human cells, before RNAPII is loaded, H2A is exchanged for H2A.Z as a consequence of promoter remodeling, suggesting a role in RNAPII

recruitment (Hardy *et al.*, 2009). Also, in ESCs H2A.Z is preferentially localized at the TSS of silent developmental regulator genes.

H2A.B lowers nucleosomes stability, and it was shown to inhibit the formation of compact chromatin *in vitro*. Additionally, it was observed that DNA assembled with nucleosomes containing H2A.B showed reduced *in vitro* transcription in HeLa nuclear extracts when compared to H2A nucleosomes (Zhou *et al.*, 2007). When expressed ectopically, this histone was found localized primarily over the bodies of active genes (Tolstorukov *et al.*, 2012), thus demonstrating that localization of this histone variant correlates with an increased expression.

Variant H2A.X is associated with DNA damage repair, as its phosphorylation on Ser139 localizes to sites in which double strand breaks (DSBs) occur. When phosphorylated, this histone variant it is referred to as γ H2AX (Rogakou *et al.*, 1998), and is recognized by MDC1, a scaffold protein that recruits DNA damage response factors (Stucki *et al.*, 2005). Additionally, it has been shown that its incorporation to nucleosomes relaxes the chromatin conformation as it impairs histone H1 binding, thus making easier the access to DNA (Li *et al.*, 2010).

In vertebrates, a specific version of H2A is present, named macroH2A. This H2A variant is enriched on the transcriptionally inactivated female X chromosome (Costanzi and Pehrson, 1998), senescence-associated heterochromatic foci (SAHF) (Zhang *et al.*, 2005), and large transcriptionally silent domains (Gamble *et al.*, 2010), suggesting a role in the repression of transcription.

The H3.3 variant is incorporated at dynamic regions such as gene promoters, the body of active genes, and cis-regulatory elements (Mito, Henikoff and Henikoff, 2007; Ray-Gallet *et al.*, 2011; Schneiderman *et al.*, 2012).

Finally, a variant of H3 assemble centromere protein A (CENPA), that is present in all eukaryotes, was found as a constitutive chromatin component associated with the kinetochore throughout the cell cycle (Palmer *et al.*, 1987).

Nucleosome type	Histone Protein	Function
Canonical	H1	Constriction of chromatin
	H2A	Nucleosome core
	H2B	Nucleosome core
	H3	Nucleosome core
	H4	Nucleosome core
Variants	H2A.Z	Increased gene expression
	H2A.B	Reduced gene expression
	H2A.X	DNA repair
	macroH2A	Gene silencing
	H3.3	Chromatin integrity
	CENPA	Kinetochore formation

Table 1.1 Canonical and variant histones.

1.1.2. DNA methylation as a repressive mark

The first indication that chromatin could be modified came early in the study of the chemical composition of DNA. Additional nucleotides other than A, T, C and G were described, with 5-methylcytosine (5-mC) being the most abundant. 5-mC is the product of the transfer of a methyl group from S-adenosylmethionine (SAM)

to position 5 of cytosine. It was reasoned that in eukaryotes DNA methylation was somehow related with gene repression, through the alteration of the interaction of DNA with proteins (A Razin and AD Riggs, 1980).

Cytosine analogues, such as 5-aza-2'-deoxycytidine (5-azaC), have been shown to inhibit the methylation of newly synthesised DNA (Jones and Taylor, 1980). The drug reactivates the expression of a variety of genes, including tumour suppressors, indirectly showing that 5-mC is a repressive modification (Groudine, Eisenman and Weintraub, 1981).

DNA methylation in mammalian cells is found mainly at CpG dinucleotides and it is established by DNA methyltransferase enzymes (DNMTs). Strikingly, less than 10% of CpGs occur in CG-dense regions that are termed CpG islands, which are mainly localized at transcription start sites of housekeeping and developmental regulator genes and are largely resistant to DNA methylation (Deaton and Bird, 2011).

The first cloned DNMT showed striking similarities in its N-terminus with bacterial methyltransferases (Bestor *et al.*, 1988). To understand the mechanism of mammalian DNA methylation, *Dnmt1* was knocked out in mouse embryonic stem cells (mESCs). This demonstrated that this protein was responsible for maintaining methylation *in vivo*, thus explaining mechanistically how methylation patterns were transferred from mother to daughter cell. However, it was noted that even when this only known DNA methyltransferase was knocked out, cells were still able to methylate DNA of "viral origin" when ectopically introduced (Lei *et al.*, 1996). This was later explained when more members of the family, *Dnmt2* (Yoder, 1998), *Dnmt3a* and *Dnmt3b* (Okano, 1998), were discovered. Strikingly,

DNMT3A and DNMT3B were shown to be able to methylate non-methylated sequences, without the need of hemimethylated DNA, the preferred substrate of DNMT1. These results divided DNMTs into two groups: de novo and maintenance DNMTs. Interestingly, it was later found that even though the level of homology of DNMT2 with DNMT1/3 families was high, this protein was not necessary for DNA methylation and rather was found to methylate tRNA^{ASP} (Goll *et al.*, 2006).

Of note, genome-wide DNA methylation patterns are not static throughout all stages of development. 5-mC was reported to be predominantly localized at CGIs in differentiated cells. However, it was found widespread throughout the entire body of genes in pluripotent cells (Lister *et al.*, 2009), suggesting that changes in 5-mC distribution were caused by developmental cues.

Although it was widely accepted that DNA methylation somehow correlated with repression, the mechanism by which 5-mC affected gene expression regulation was described initially in yeast, when MeCP1 was shown to bind methylated DNA and to repress transcription from a methylated promoter *in vitro* (Cross *et al.*, 1997). This report was followed by the discovery in mammalian cells that MeCP2, which also binds to 5-meC, recruited histone deacetylases through the mSinA3 complex and thus repressed gene expression by constraining the chromatin conformation (Nan *et al.*, 1998).

5-mC was initially described as an irreversible modification, given that loss of methylation had only been described as a passive phenomenon that occurred only when *Dnmt1* was absent during DNA replication (Bestor, 2000). However, it was difficult to explain how changes in the methylome of differentiated tissues

arose, as there was no prior evidence of enzymes that specifically removed 5-mC. This was explained by the discovery of the Ten-eleven translocation (TET) family of DNA hydroxylases. TET1 was initially identified in acute myeloid leukemia (AML) as a fusion partner of the histone H3 Lys 4 (H3K4) methyltransferase mixed-lineage leukemia (MLL) (Ono *et al.*, 2002; Lorsback *et al.*, 2003). TET proteins convert 5-mC to 5-hydroxymethylcytosine (5-hmC), 5-formylcytosine (5-fC), and 5-carboxylcytosine (5-caC) through three consecutive oxidation reactions (Wu and Zhang, 2011).

1.1.3. Histone post-translational modifications.

It is commonly noted that there are two main groups of protein complexes that change the chromatin landscape: those that covalently modify histones and those that reposition nucleosomes (Wilson and Roberts, 2011).

Histone proteins are predominantly globular in structure, except for an unstructured N-terminal domain called the tail, which protrude outside from the octamer. These histone tails have an unusually high proportion of lysines, serines and arginines, which can be post-translationally modified and function as both recognition and docking sites for chromatin binding proteins (Zenter 2013). To date, several different histone modifications, or marks, are known: acetylation, methylation, phosphorylation, ubiquitination, sumoylation, ADP-ribosylation, deamination (citrullination), and proline isomerization, among others (Kouzarides, 2007). The most common and therefore most studied modifications are acetylation and methylation, which are highly dynamic.

Acetylation of lysine residues in histones destabilizes the interactions between histone and DNA, and thus acetylation of histones has been invariably linked to an open conformation of chromatin and thus with an increase in transcription (Kouzarides, 2007; Steunou, Rossetto and Côté, 2014). Histone acetyltransferase 1 (HAT1) was first isolated in yeast in 1995 and found to be responsible for acetylating lysine 12 of histone H4 (Kleff *et al.*, 1995). Acetylation of lysine reduces the basic charge of the histone tail, neutralizing the electrostatic interaction with DNA and thus has an overall effect of chromatin decondensation. Consistent with this, localization of histone acetylation is correlated with transcriptionally active regions of chromatin. In mammals, histone acetyltransferases include the p300 and CREB binding protein (CBP) family, which catalyse H3K27ac (Ogryzko *et al.*, 1996). Also, acetylation of H4K5/8/12/16 is established by Tip60/KAT5, the catalytic subunit of nucleosome acetyltransferase of histone H4 (NuA4) complex, and this correlates with increased gene expression (Frank *et al.*, 2003; Taubert *et al.*, 2004). Acetyl groups are recognized by proteins containing bromodomains which are found in several chromatin associated factors, such the bromodomain-containing proteins (BRD) (Umehara *et al.*, 2010). BRD4 interacts with P-TEFb, which in turns phosphorylates and activates transcriptional elongation by RNA polymerase II (RNAPII) (Moon *et al.*, 2005), thus mechanistically explaining how H3K27ac is linked to transcriptional activation.

Methylation of histone residues that correlate with gene transcription include those deposited by the yeast proteins Set1 and Set2, two members of the SET-domain methyltransferase superfamily. Set1 was shown to establish both H3K4 dimethylation (H3K4me₂) and trimethylation (H3K4me₃). The latter is

exclusively found in active genes, whereas H3K4me2 is found in both active and inactive genes (Santos-Rosa et al 2002). Afterwards, it was described that H3K4me3 was predominantly found at transcription start sites (TSS) (Liang *et al.*, 2004), and that this modification was reversed by Lysine Specific Demethylase 1 (LSD1), making this a transcriptional corepressor (Shi *et al.*, 2004). In mammals, H3K4 is methylated by MLL1, MLL2, MLL3, MLL4, SET1A and SET1B (Milne *et al.*, 2002). H3K4me3 is then recognized by bromodomain and PHD finger-containing transcription factor (BPTF), a subunit of the NURF chromatin-remodeling complex, through its plant homeodomain (PHD) domain (Wysocka *et al.*, 2006).

Set2 (SETD2 in mammals) establishes H3K36me3 and triggers transcription elongation by binding to RNAPII. Interestingly, Set2 can also recognize and bind to its own established mark, thus generating a positive feedback loop (Krogan *et al.*, 2003). H3K36me3 can recruit HDACs through Eaf3 in yeast and MRG15, a subunit of NuA4, in mammals to block ectopic transcription initiation within gene bodies (Doyon *et al.*, 2004; Joshi and Struhl, 2005; Zhang *et al.*, 2006)

H3K9me3 is deposited in mammals by the histone methyltransferases SUV39H1 and SUV39H2 (Jenuwein *et al.*, 2000), EHMT2 (G9A) (Tachibana *et al.*, 2002), EHMT1 (GLP) (Zhang *et al.*, 2016), TRIM28 (KAP1) and SETDB1. H3K9me3 is recognized by heterochromatin protein 1 (HP1) through its chromodomain (Bannister *et al.*, 2001; Jenuwein *et al.*, 2002), which oligomerizes and promotes gene repression through heterochromatin formation.

Histone Protein	Mark	Writer	Reader	Function
H2A	K119ub	RING1A/B	JARID2	Chromatin compaction
H3	K4me1	MLL1-5, SET1A/B, ASH1	WDR5 (MLL1)	Transcription increase, in promoters
	K4me3		BPTF	
	K9me3	Suv39H1/2, GLP, G9A, TRIM28 and SETDB1	HP1	Chromatin compaction
	K27me3	EZH1/2	EED, CBX2/4/6/7/8	Facultative heterochromatin
	K27ac	P300-CBP family	BRD family	Transcription increase
	K36me3	SETD2	MRG15, DNMT3A	In gene bodies, inhibits ectopic transcription.
H4	K5/8/12 /16ac	TIP60	BRD family	Transcription increase

Table 1.2. Selected histone modifications in vertebrates relevant for this work and their function.

1.1.4. Chromatin remodelling

Nucleosome positioning plays a pivotal role in regulating gene expression. Low nucleosome density promotes gene expression by allowing transcription factor access to regulatory sequences (Owen-Hughes and Workman 1994) and passage of RNAPII through gene bodies (Côté *et al.*, 1994).

Nucleosomes are repositioned by families of protein complexes that are highly conserved in eukaryotes. These complexes are DNA-dependent ATPases with a helicase-like domain, which allows them to use ATP hydrolysis as energy source to move nucleosomes. Depending on the subunits of which they are composed, all chromatin remodelling complexes fall within one of four sub-families: SWI/SNF, ISWI, CHD and INO80 (Mirabella, Foster and Bartke, 2016).

The key characteristic that distinguishes the above families is the type of histone modifications that they recognize and the function they exert. For

example, SWI/SNF (mammalian BAF complex) binds to acetylated histones through the bromodomains of its subunits BRM and BRG1 (Hassan *et al.*, 2002).

Imitation SWI (ISWI) family members are composed of 2 to 4 subunits and all complexes in the ISWI family include a subunit that possesses a characteristic SANT domain adjacent to a SLIDE domain. These two domains form a nucleosome recognition module that binds both to unmodified histone tails and to DNA (Grüne *et al.*, 2003). Two different functions were described for ISWI complex in yeast depending on whether the complex contains ISWI1P or ISWI1B. ISWI1P-containing complexes act to repress RNAPII transcribed genes (Moreau *et al.*, 2003) and ISWI1B-containing complexes modulate the passage of RNAPII through the gene (Morillon *et al.*, 2003)

Nucleosome Remodeling and histone Deacetylation (NuRD) complex, as its name indicates, can both displace nucleosomes and deacetylate histones (Torchy, Hamiche and Klaholz, 2015). This complex contains six core subunits, which include either chromodomain and helicase-like domain 3 or 4 (CHD3 or CDH4), HDAC1 or 2, retinoblastoma-binding protein 4 or 7 (RBBP4 or RBBP7), MTA1, MTA2 or MTA3 and p66 α (GATAD2A) or p66 β (GATAD2B). The CHD subunits recognize H3K9me3 through their tandem plant homeodomain fingers (PHDs) (Musselman *et al.*, 2009). Finally, MBD2 and MBD3 are interchangeable proteins in the complex and both bind 5'-mC (Le Guezennec *et al.*, 2006).

1.2. Polycomb proteins

Polycomb group (PcG) proteins are an evolutionary conserved group of proteins responsible for silencing homeotic and other developmental regulator genes by constraining the chromatin through histones modification. Polycomb group proteins form two main complexes: polycomb repressive complex 1 (PRC1) and polycomb repressive complex 2 (PRC2).

PcG genes were originally identified in *Drosophila melanogaster* as regulators of body segmentation by repressing Hox genes (Lewis 1978, Struhl 1981). The Trithorax group (TrxG) of proteins were described as antagonists of PcG proteins, and thus promoters of Hox gene expression during early embryonic development. TrxG proteins include *Drosophila* *trx*, that catalyses H3K4me3. Many of the PcG genes were named after the mutant phenotypes in which extra sex combs appeared on the second and third pair of legs of male flies instead of only on the first pair of legs. This distinctive phenotype is caused by the derepression of the sex combs reduced (*scr*) Hox gene and became a marker for subsequent genetic screens to identify additional PcG and TrxG genes (Kassis et al. 2017).

In mouse embryonic stem cells (mESC), both PRC1 and PRC2 were found localized to the promoters of genes encoding developmental regulators including Hox genes, suggesting that PcGs function during mammalian development in a similar manner to flies (Azuara et al., 2006; Boyer et al., 2006; Lee et al., 2006).

1.2.1. PRC1 composition

In flies, the PRC1 core complex (Shao *et al.*, 1999; Saurin *et al.*, 2001) comprises four PcG proteins: polycomb (PC), a chromodomain-containing protein that binds to H3K27me3 (Fischle *et al.*, 2003; Min, Zhang and Xu, 2003); dRing, the ubiquitin ligase that mono-ubiquitinates H2AK118 (Wang *et al.*, 2004); Polyhomeotic (PH); Posterior Sex Combs (PSC), responsible for chromatin compaction *in vitro* (Francis, Kingston and Woodcock, 2004); and Sex comb on Midleg (SCM).

PRC1 is more complex in mammals (Table 1.3). When purified from HeLa cells, the PRC1 core pulled-down other substoichiometric proteins, chiefly chromodomain proteins (CBX2, 4, 6, 7 and 8), homologous to PC; ubiquitin ligases RING1A and RING1B, homologues of dRing that establish H2AK119ub; Ph homologs PHC1-3; BMI1/PCGF4 and an homolog of Scm, SCM1 (Levine *et al.*, 2002),

Later it was established that there are at least six distinct PRC1 subcomplexes (PRC1.1-PRC1.6), which are defined by the members of the PCGF protein family (PCGF1-6) that they contain. These proteins dictate the recruitment of specific accessory subunits and thus confer to each subcomplex diverse functional properties (Di Croce and Helin, 2013). Reports of redundant activity between complex containing PCGF2/4 and PCGF3/5 suggests that these are biochemically identical (Gao *et al.*, 2012).

Usually, PRC1.2 and PRC1.4 are termed canonical PRC1 complexes, and these are recruited to H3K27me3 (Scelfo, Piunti and Pasini, 2015), deposited by

PRC2, via the chromodomain of CBX2, 4, 6, 7 or 8 (Fischle *et al.*, 2003; Min, Zhang and Xu, 2003). Canonical PRC1 complexes have been proposed to compact chromatin and mediate higher-order chromatin structures through their PHC subunits, and it has been widely postulated that these activities are a central determinant of PRC1-mediated gene repression

The other PRC1 subcomplexes, classified as non-canonical or variant complexes, are defined by the presence of RYBP or YAF1 and are recruited by means other than H3K27me3 binding, reflecting their lack of CBX subunit (Gao *et al.*, 2012). PRC1.1 is recruited to its targets thanks to its KDM2B subunit, which recognizes unmethylated CpG islands through its CXXC domain (Blackledge *et al.*, 2014) or by BCL6 co repressor (BCOR), a newly described BCOR-PRC1.1 complex that binds to BCL6 and is recruited to BCL6 target genes, which was demonstrated to be necessary for the regulation of primed pluripotent state (Wang *et al.*, 2018). Recently it was shown that PRC1.3 interacts with USF1/2 DNA binding transcription factors and that PRC1.6 is recruited by DNA-binding E2F6-DP1 and MGA-MAX heterodimers, and that MGA was necessary for the structural stability of the complex (Scelfo *et al.*, 2019)

1.2.2. PRC2 catalytic core

The core of PRC2 in *D melanogaster* is formed by E(z), Su(z)12, CAF1-p55, and Esc (Pengelly *et al.*, 2013). There are two main forms of PRC2 in flies and are defined by whether they contain the alternative subunits Pcl (Polycomb- like) or

Jumonji/ARID domain-containing protein 2 (Jarid2) (Nekrasov *et al.*, 2007; Herz *et al.*, 2012).

In mammals, the PRC2 core is composed of four subunits: the E(z) ortholog enhancer of zeste 2 (EZH2) or its close homologue EZH1, the Esc ortholog embryonic ectoderm development (EED), the Su(z)12 orthologue suppressor of zeste 12 (SUZ12) and the CAF1 orthologues RBBP4 and RBBP7 (Margueron and Reinberg, 2011) (Table 1.3).

PRC2 is responsible for the mono-, di- and tri-methylation of lysine 27 of histone H3 (H3K27me_{1/2/3}) through its enzymatic subunits EZH2 or EZH1 (Simon and Kingston, 2009). Although established by the same enzyme, all three H3K27 methylation marks are found in different places across the genome. H3K27me₃ accumulates at the TSS of genes, whereas H3K27me₁ is located within the body of highly expressed genes and H3K27me₂ is broadly distributed throughout the genome and deposited on approximately 70% of all H3 proteins in mESCs (Ferrari *et al.*, 2014). Recent studies in which *Suz12* and *Ezh2* were knocked out and then the genes reconstituted, showed that H3K27me₃ can be established *de novo* without previous H3K27me₃, showing that H3K27me₃ patterns in mESCs are not dependent on inheritance to be properly established (Højfeldt *et al.*, 2018)

1.2.2.1 EZH1/2

EZH2 and SUZ12 expression are activated by E2F and thus are more highly expressed in proliferating cells (Bracken 2003). EZH1, EZH2 close orthologue,

is more highly expressed in differentiated cells and although it is less active than EZH2 (Margueron *et al.*, 2008), EZH1 is necessary for protecting slow-cycling, undifferentiated hematopoietic stem cells from senescence (Hidalgo *et al.*, 2012).

EZH2 and EZH1 share similar structures. Both contain an N-terminal catalytic SET domain (Müller *et al.*, 2002) and a minimal EED binding domain is located between residues 39-68 (Han *et al.*, 2007). The catalytic activity of EZH2 is regulated by binding to EED, as this interaction stabilizes the SET domain (Justin *et al.*, 2016).

1.2.2.2 SUZ12

SUZ12 has no catalytic domains but is necessary for EZH2 histone methyltransferase activity *in vitro* and in cells (Cao and Zhang, 2004; Montgomery *et al.*, 2005; Pasini *et al.*, 2007). *Suz12^{GT/GT}* mESC show no detectable H3K27me3 and strongly reduced levels of H2K27me2. Also, EZH2 is not observable by immunoblotting in the absence of SUZ12, suggesting that lack of SUZ12 causes EZH2 proteolytic degradation (Pasini *et al.*, 2007). SUZ12 contains two discrete domains that are conserved throughout evolution. One of them is a zinc finger at its N-terminus, which when removed does not affect the catalytic activity of PRC2, and a C-terminal domain, called the VRN2, EMF2, FIS2, and Su(z)12 (VEFS), which is necessary for interaction with EZH2 (Rai *et al.*, 2013). More recently, SUZ12 was reported to contain two separate WD40-binding (WDB) domains that interact with RBBP4, a N-terminal helix that,

together with the zinc finger (NtH + ZnF), creates a binding site for JARID2 and EPOP, and a C2 domain that AEBP2 binds and stabilizes (Chen *et al.*, 2018a).

1.2.2.3 EED

EED is also necessary for the formation of the PRC2 catalytic core. Although *Eed*^{-/-} mESC can be expanded in tissue culture, these cells show reduced H3K27 methylation levels and also an increased expression of differentiation-specific genes, associated with loss of pluripotency (Boyer *et al.*, 2006). Also, interaction of EZH2 with EED increases the methylation activity of PRC2 through conformational changes in the complex (Shu *et al.* 2010). Like, *Suz12*^{-/-} cells, *Eed*^{-/-} lines show no detectable levels of EZH2 protein due to proteolytic degradation (Montgomery *et al.*, 2005).

1.2.3 PRC2 accessory factors

Evidence gained over the last few years has revealed additional proteins forming part of an extended PRC2 holo-complex (Table 1.3). These are all considered as accessory factors that define two variants of PRC2, usually called PRC2.1 and PRC2.2 (Grijzenhout *et al.*, 2016). PRC2.1 is composed of the core plus one of three polycomb like proteins (PCL), PCL1, 2 or and either Elongin BC and PRC2 associated protein (EPOP) (Beringer *et al.*, 2016) or PRC2 associated LCOR isoform 1 (PALI1) or isoform 2 (PALI2) (Conway *et al.*, 2018). PRC2.2 is less complex and is only formed by the PRC2 core plus adipocyte enhancer-

binding protein 2 (AEBP2) and JARID2 (Hauri *et al.*, 2016; Conway *et al.*, 2018) (Fig. 1.1).

The first compelling evidence that there are two separate PRC2 subcomplexes was provided by labelling both EPOP and PALI1 by BIO-PAX and performing mass spectrometry which showed that both proteins interacted with the core PRC2 components bound to both proteins but not with the AEBP2 and JARID2 (Alekseyenko *et al.*, 2014). Afterwards, mass spectrometry of immunoprecipitated AEBP2 showed exclusively interaction with the core subunits of PRC2 and JARID2, but not EPOP, PALI1/2 nor any of the PCLs (Grijzenhout *et al.*, 2016). Besides having different subunits, PRC2.1 and PRC2.2 are catalytically discernible complexes, the latter being the most highly active (Beringer *et al.*, 2016; Conway *et al.*, 2018), suggesting a mechanism in which deposition of H3K27me3 is finely tuned by presence and absence of PRC2 accessory factors.

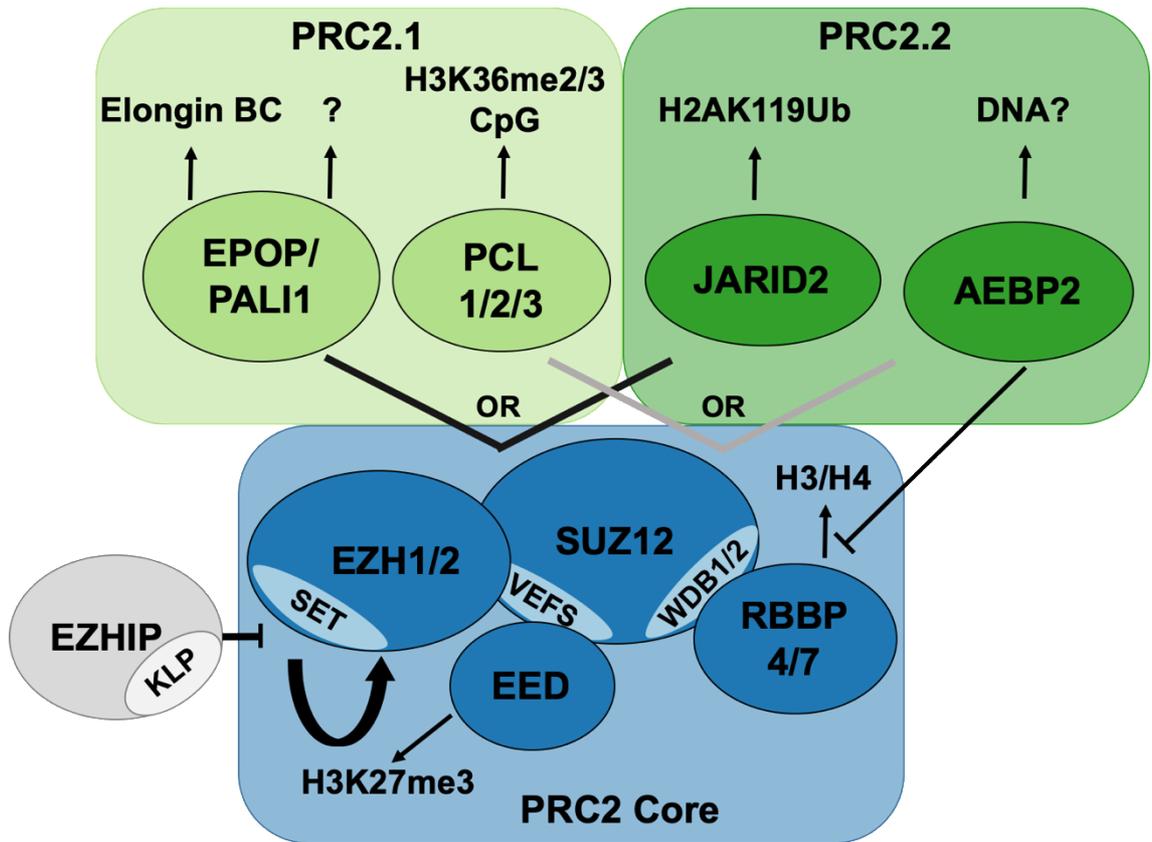


Figure 1.1. Known PRC2 variants and interactions made by accessory subunits.

Mutually exclusive components are noted with a dash. The core components of PRC2 are in blue. PRC2.1 subunits are in light green and PRC2.2 subunits in dark green. EZHIP is depicted in grey as it binds both PRC2.1 and PRC2.2 but is not a core subunit. Updated from Vann & Kutateladze, 2018.

1.2.3.1 JARID2

JARID2 belongs to the JmjC domain containing protein family, which removes histone methylation by iron- and α -ketoglutarate (α -KG) dependent oxidation. However, no demethylation activity has been reported for JARID2, due to substitution of the residues that bind iron and α -KG in other members of the protein family (Zhang *et al.*, 2011).

JARID2 has been described as necessary for the proper differentiation of mESCs and establishment of mesodermal and neural lineages, in embryoid body formation by leukaemia inhibitor factor (LIF) removal (Landeira *et al.*, 2010).

Recently, an additional JARID2 isoform has been described that is lacking the N-terminus. This form is not able to bind to PRC2, given that the domain responsible for this is localized to the N-terminus. Moreover, it was shown that this cleaved version of JARID2 was necessary for cell differentiation (Al-Raawi *et al.*, 2019).

1.2.3.2 AEBP2

AEBP2 is a Gli-type zinc finger protein, which was originally identified due to its capability for binding to the promoter region of adipose P2 (aP2) gene in vitro (Mitsui *et al.*, 1999). Afterwards, independent studies suggested that AEBP2 was binding to a specific DNA sequence similar to the GAGA motif and that this protein co-occupied loci with SUZ12, potentially showing an interaction with PRC2 (Kim *et al.*, 2009).

Finally, by co-immunoprecipitation and mass-spectrometry of FS2-tagged AEBP2 in mESCs, it was revealed that AEBP2 interacted with the PRC2 core subunits and with JARID2. In the same study it was described that mice with ablated *Aebp2* exhibited a posterior transformation of the skeleton, observed previously by misexpression of *Hox* genes (Grijzenhout *et al.*, 2016).

1.2.3.3. EPOP

Initially known as esPRC2p48, EPOP was first described in mESCs as a protein associated with EZH2 and SUZ12. This interaction was demonstrated by glycerol gradient, mass spectrometry assays and reciprocal co-immunoprecipitations. In this same study, it was reported that EPOP increased PRC2 activity *in vitro* (Zhang *et al.*, 2011). Later, it was shown that EPOP was highly expressed in mESCs and that its presence was needed for neuronal and glial differentiation (De Cegli *et al.*, 2013). Ablation of *Epop* in mESC prompted an increase of H3K27me3 and in PRC2 occupancy in the TSS of PRC2 targets. Also, it was shown through mass spectrometry of EPOP interactors that Elongin BC was associated with EPOP, suggesting that this protein generates a bridge between PRC2 and Elongin BC complexes (Beringer *et al.*, 2016; Liefke, Karwacki-Neisius and Shi, 2016).

1.2.3.4. PCLs

PCL1 was the first Pcl orthologue to be described in vertebrates. PCL1 was described as stimulating PRC2 catalysis *in vitro* and was shown to bind to the *Hoxa* locus (Cao *et al.*, 2008). Furthermore, knockdown of *Pcl1* causes a decrease in levels of H3K27me3, and initially it suggested that PCL1 was necessary for EZH2 to establish H3K27me3 at the *Hoxa* locus (Sarma *et al.*, 2008). PCL proteins are not ubiquitously expressed; PCL1 is mostly abundant in quiescent cells when compared with PCL2 and PCL3 (Brien *et al.*, 2012),

suggesting that although highly homologous these proteins potentially have different functions. Consistent with this, PCL1 interacts with p53 through a region unique to this protein (Brien *et al.*, 2015).

1.2.3.5. PALI1/2

Two proteins of a vertebrate specific family named PALI1 and PALI2 were recently revealed as additional subunits of PRC2.1. These are encoded by splice variants of the genes *Lcor* and *Lcorl*, respectively, and were found to increase PRC2 catalytic activity *in vitro* (Conway *et al.*, 2018). Additionally, mass spectrometry analysis revealed that PALI1 was associated with chromatin regulators like SET protein, deubiquitinases and the G9A/GLP complex, suggesting a role in coupling PRC2 to other chromatin modifiers.

1.2.3.6. EZHIP

The most recently discovered member of PRC2 is enhancer of zeste homologs inhibitory protein (EZHIP), also called CATACOMB or CXorf67 (Pajtler *et al.*, 2018; Jain *et al.*, 2019; Piunti *et al.*, 2019; Ragazzini *et al.*, 2019). EZHIP was found to be an inhibitor of PRC2 activity, as this protein contains a K27M-like peptide (KLP) in its C-terminus that acts as an inhibitor of EZH2, mimicking the histone H3K27M mutation, which engages the EZH2 active site but cannot be methylated (Jain *et al.*, 2019).

The *EZH1P* gene is kept silenced by DNA methylation in its promoter and treatment with 5-AzaC activates *EZH1P*, suggesting a link between DNA-methylation status and H3K27me3 regulation (Piunti *et al.*, 2019). Additionally, it has been described that EZHIP is necessary for female fertility, as *Ezh1p*^{-/-} mice showed impaired oocyte maturation (Ragazzini *et al.*, 2019).

Complex	Subcomplex	Subunit	Function
PRC2	Core	EZH1/2	Through its SET domain catalyses H3K27me3
		EED	Necessary for catalytic activity, binding to H3K27me3
		SUZ12	Necessary for catalytic activity, adaptor protein
		RBBP4/7	Core histone binding, H3 and H4
	PRC2.1	EPOP	Stimulation of <i>in vitro</i> activity, binding to Elongin BC
		PAL1	Stimulation of <i>in vitro</i> activity
		PCL1-3	Stimulation of <i>in vitro</i> activity, binding to unmethylated CGIs and to H3K36me3
	PRC2.2	JARID2	Stimulation of <i>in vitro</i> activity and binding to H2AK119ub
		AEBP2	Stimulation of <i>in vitro</i> activity
?	EZH1P	Inhibits catalytic activity of PRC2 by binding of KLP	
PRC1	Core	RING1A/B	Ubiquitin ligase, establishes H2AK119ub
	Non canonical core	RYBP	Stimulation of RING1 activity
	cPRC1.2/ cPRC1.4	CBX2/4/6/8	Binding to H3K27me3
		HPH1-3	Chromatin compaction
		PCGF2/4	Stimulation of RING1 activity
	ncPRC1.3/ ncPRC1.5	AUTS	p300 recruitment
		FRS2	-
		CK2	Inhibits RING1B activity
		PCGF3/5	Stimulation of RING1 activity
	ncPRC1.1	USP7	Regulates PRC1 ubiquitination
		SKP1	Part of ubiquitin ligase complex
		BCOR/BCORL	Binding to BCL6
		KDM2B	Binding to unmethylated CGIs
		PCGF1	Stimulation of RING1 activity
	ncPRC1.6	HDAC1/2	Histone deacetylation
		E2F-DP1	DNA binding to E-box elements
		MGA-MAX	DNA binding to E-box elements
		USF1/2	DNA binding to E-box elements
		L3MBTL2	Histone binding
		PCGF6	Stimulation of RING1 activity

Table 1.3. Summary of Polycomb complex subunits and their functions.

1.2.4. Three-dimensional structure of PRC2

Obtaining the three-dimensional structure of the PRC2 holo-complex had proven hard to obtain until recently. The first crystal structure of PRC2 with 2.3 angstroms resolution was obtained by expressing EZH2 fused with the SUZ12 VEFS domain and EED from a thermophilic yeast, *Chaetomium thermophilum*. This generated a ternary “minimal” complex capable of methylating H3K27 peptides. Moreover, the recent crystal structures of an isolated inactive catalytic domain of EZH2 revealed an autoinhibited conformation, implying that structural rearrangement of this domain is likely required for an active PRC2, and mechanistically explaining why SUZ12 is needed for EZH2 catalytic activity (Jiao and Liu, 2015).

Afterwards, Kasinath and colleagues succeeded in generating structures of the PRC2.2 holo-complex by cryogenic electron microscopy (Cryo-EM). They described both inactive and active states of the complex. In this work it was shown that both JARID2 and AEBP2 were needed to generate a structurally stable PRC2, in which AEBP2 N-terminus mimics the H3 protein tail, promoting a conformational change that enhances methylation (Kasinath *et al.*, 2018).

1.2.5. Recruitment of PRC2

The first mechanism of PRC2 recruitment was described in *D. melanogaster*. In flies, PRC2 is recruited to specific DNA sequences called Polycomb responsive elements (PRE) (Strutt and Paro, 1997; Strutt, Cavalli and Paro, 1997; Orlando

et al., 1998). These sequences are bound by specific transcription factors such as PHO (Brown *et al.*, 1998), GAGA factor (GAF), Pipsqueak (Psq), Zeste, Grainyhead/NTF-1, Dsp1 and Sp1/KLF family members, which act like anchors for PcG proteins (Müller and Kassis, 2006).

However, it is still controversial whether PREs, or similar sequences, exist in vertebrates. In mESC, both PRC1 and PRC2 are associated mainly with large CGIs depleted of activating transcription factor motifs (Tanay *et al.*, 2007; Ku *et al.*, 2008). It was described that insertion of CGIs lacking promoter activity or binding sites for transcription factors can induce specific binding of PRC2 (Mendenhall *et al.*, 2010; Jermann *et al.*, 2014). Later it was suggested that PCLs were responsible for this, given that these bind to unmethylated CGIs thanks to a cassette that recognizes a 12-base pair palindromic sequence that contains 2 CpG motifs (Li *et al.*, 2017; Perino *et al.*, 2018).

PRC2 subunits recognize both modified and unmodified histones, generating crosstalk with other chromatin modifiers. EED binds to H3K27me₃, thus generating a positive feedback loop that was thought to propagate the H3K27me₃ mark to daughter cells (Hansen *et al.* 2008, Margueron *et al.* 2009). The core subunits RBBP4/7 bind to the unmodified histones H3 and H4, and also increases the catalytic activity of the complex (Schmitges *et al.*, 2011). Additionally, JARID2 binds to H2AK119ub through its ubiquitin interaction motif (UIM) localized between amino acids 24-43, making this subunit the link between PRC1 and PRC2 (Kalb *et al.*, 2014b; Cooper *et al.*, 2016) (Fig. 1.2). AEBP2 can also bind DNA, but no sequence specificity has been reported yet (Grijzenhout *et al.*, 2016).

Another factor that affects PRC2 recruitment is the conformation of the chromatin. By using different types of nucleosomal arrays, in which the nucleosome density was varied but not the total number of nucleosomes, it was shown that the preferred substrate of PRC2 were densely packed nucleosomes (Yuan *et al.*, 2012). Ablation of *Dnmt3a* in mESCs causes an increase of SUZ12 and EZH2 binding to DNMT3A target genes, suggesting that methylation of CpGs inhibits PRC2 recruitment, which was corroborated *in vitro* (Bartke *et al.*, 2010; Wu *et al.*, 2010).

Histone modifications also alter PRC2 recruitment. It has been demonstrated that H3K4me3 and H3K36me2/3 inhibit PRC2 activity *in vitro*, potentially showing that these marks exclude PRC2 from transcribed regions (Schmitges *et al.*, 2011). However, it has been shown *in vitro* and *in vivo* that PCL proteins bind to H3K36me3 (Ballaré *et al.*, 2012; Brien *et al.*, 2012), suggesting that this mark tunes PRC2 activity.

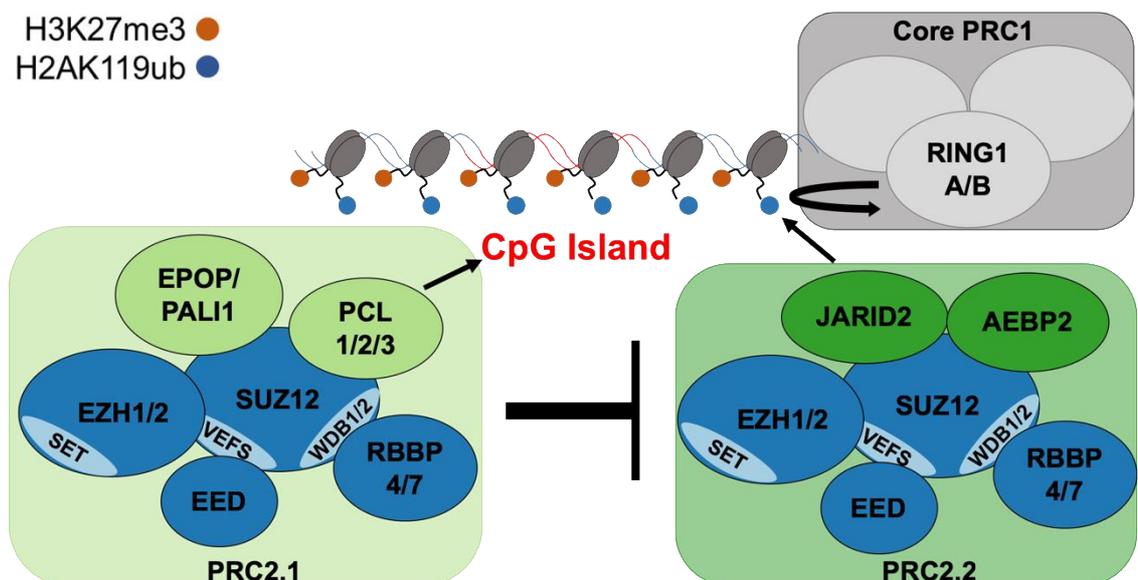


Figure 1.2. Summary of interplay between PRC2.1 and PRC2.2 with PRC1

PRC1 catalyses H2AK119ub by its RING1A/B subunit, which is recognized by JARID2, and thus PRC2.2. Modified from Brien *et al* 2016 and Healy *et al* 2019.

1.3. RNA and PRC2

1.3.1. PRC2 RNA binding specificity

PRC2 interacts with RNA in addition to interacting with chromatin. Identification of HOTAIR, a long non-coding RNA (lncRNA) product of the HOXC locus, was followed by the discovery that this RNA interacted with PRC2 in co-precipitation experiments (Rinn *et al.*, 2007). A model was proposed in which HOTAIR recruited PRC2 to the HOXD locus *in trans* (Rinn *et al.*, 2007). The lncRNA Xist is necessary for the process of X-inactivation, as it spreads in cis across the future inactive X chromosome which forms a transcriptionally silent nuclear compartment, enriched with repressive chromatin marks, such as H3K27me3 (Plath *et al.*, 2003).

Similar to what was found for HOTAIR, an Xist isoform, RepA, was found to bind to PRC2 by electrophoretic shift assay (EMSA) and it was proposed that Xist RepA functioned to recruit PRC2 to the inactive X chromosome *in cis* (Zhao *et al.*, 2008). These studies were followed by others identifying further lncRNAs that interacted with PRC2. The specific interaction of PRC2 with lncRNAs supported by native RNA immunoprecipitation (RIP) followed by microarray analysis (RIP-Chip), which showed that PRC2 was binding to 20% of all lncRNAs produced in different cell types (Khalil *et al.*, 2009). Afterwards, a study using native RIP using EZH2 as bait followed by sequencing, also reported a set of lncRNAs that were specifically bound by PRC2 (Zhao *et al.*, 2010). However, PRC2 was also found to interact with RNAs in addition to lncRNAs. It was found

that PRC2 bound short RNAs, between 50-200 nucleotides in length, that were transcribed from 5' end of PcG target genes in T cells and mESCs (Kanhere *et al.*, 2010). This report, plus reanalysis of previous RIP-seq analysis and *in vitro* binding assays using different RNAs, led to the model that PRC2-RNA binding was promiscuous and that affinity towards RNA was dependent on length (Davidovich *et al.*, 2013, 2015). However, later reports using individual-nucleotide resolution UV crosslinking and immunoprecipitation sequencing (iCLIP-seq), with SUZ12 as bait in mESCs, identified that PRC2 bound to essentially all nascent pre-mRNA and lncRNAs in mESC (Beltran *et al.*, 2016).

Earlier work suggested that the three-dimensional conformation of Xist RepA RNA was important for PRC2 RNA binding specificity (Zhao *et al.*, 2008). That PRC2 exhibits preferential binding to specific RNA sequences was shown using recombinant JARID2-containing PRC2 *in vitro*. It was found that PRC2 exhibited preferential binding to poly(G) DNA, which can fold into a G-quadruplex (G4) structure, versus poly(A) DNA of the same length, that forms an extended right-hand helix (Kaneko *et al.*, 2014). The high affinity of PRC2 for G4 RNA was further confirmed with EMSA using a synthetic RNA sequence of 40mers composed of repeated G tracts and a physiological 40mer sequence contained in the ncRNA TERRA, both of which reduced the mobility of recombinant PRC2. This was only achieved when RNAs were folded in the presence in K⁺-containing buffer, which supports G4 formation, and not in Li⁺-containing buffer, which does not (Wang *et al.*, 2017).

1.3.1. PRC2 binds RNA through a number of different subunits

None of the PRC2 subunits contain a known RNA binding domain so the search has been on to identify how PRC2 interacts with RNA. Initial studies using (EMSA) showed that an EZH2-EED complex and EZH2 alone could interact with sequences from Xist RepA (Zhao *et al.*, 2008). Subsequently, SUZ12 was also found to bind to this Xist Rep sequence and to other RNAs (Kanhere *et al.*, 2010). This was later confirmed by EMSAs which demonstrated that the affinity of SUZ12 and EZH2 for Xist RepA and HOTAIR was comparable, while EED bound poorly to both RNAs (Cifuentes-Rojas *et al.*, 2014), thus showing that at least two core subunits were able to bind RNA.

Application of UV cross-linking–based methods later revealed that several PRC2 components directly interact with RNA in cells. Using photoactivatable-ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP), with HA-tagged version of EZH2 as bait, indicated that this protein interacted with RNA in cells (Kaneko *et al.*, 2013). This same technique was used to discover that JARID2 was also crosslinking with RNA in mESC (Kaneko, Bonasio, *et al.*, 2014), showing that RNA binding activity is not limited to the core PRC2 subunits. Later, it was shown by individual-nucleotide resolution UV crosslinking and immunoprecipitation (iCLIP) that SUZ12 could recapitulate PRC2 binding to RNA in cells, even in the absence of EZH2, EED and JARID2 (Beltran *et al.*, 2016). This suggested that either SUZ12 defined binding of RNA or that there was redundancy between EZH2 and SUZ12 in RNA binding.

To try to assess the relative importance of each subunit for RNA binding, recombinant PRC2 comprised of SUZ12, EZH2, EED, RBBP4 and AEBP2 was crosslinked with radioactively labelled RNA *in vitro* and the crosslinked proteins resolved by SDS-PAGE. It was found that EZH2 and SUZ12 bound RNA to similar extents, followed by AEBP2 and EED, while RBBP4 showed no binding at all (Wang, Goodrich, *et al.*, 2017).

Using recombinant minimal PRC2 complex, comprised of EZH2, EED from 81-441 residues and the VEFS domain of SUZ12, it was established that residues dispersed through EZH2, including R34, K39 and R494, were important for RNA binding. Interestingly, the addition of JARID2 and AEBP2 to this recombinant complex increased the affinity of the complex towards RNA, but did not change the specificity (Long *et al.*, 2017), suggesting that the catalytic core of PRC2 is the source of its specificity for G4 RNA.

1.3.3. RNA regulates the recruitment of PRC2 to chromatin.

When PRC2 was discovered to bind lncRNAs, it was proposed that PRC2 was recruited to specific loci by these RNA species. Particularly, PRC2 was reported to be recruited to the future inactive X chromosome *in cis* by Xist RepA RNA (Jing Zhao *et al.*, 2008) and to the *HOXD* locus *in trans* by HOTAIR RNA (Rinn *et al.*, 2007). Recent studies have led to the "junk mail" model for PRC2-RNA binding in which promiscuous binding to nascent RNA allows PRC2 recruitment to the subset of Polycomb target genes that have escaped repression (Davidovich *et al.*, 2013, 2015).

However, subsequent reports showed that inhibition of RNAPII transcription using 5,6-dichoro-1-D-ribofurano-sylbenzimidazole riboside (DRB) or triptolide, resulted in an increase of PRC2 binding to CGIs at active genes (Riising *et al.*, 2014). Similarly, depletion of RNA in permeabilised cells by RNaseA treatment was sufficient to increase the association of SUZ12 and EZH2 with the chromatin fraction and induce PRC2 binding to CGI at active genes (Beltran *et al.*, 2016). Also, insertion of a CGI between the enhancer and promoter of *Utf1* gene in mESCs did not show H3K27me3 deposition until differentiation into neuronal precursors, consistent with loss of transcriptional activity (Jermann *et al.*, 2014). Similarly, deletion of the *Ephx1* TSS in 3T3 cells, and thus ablating its transcription, induced H3K27me3 deposition through the gene body (Hosogane *et al.*, 2016).

That RNA might not recruit PRC2 to chromatin was suggested by experiments showing that RNA inhibited PRC2 histone methyltransferase activity *in vitro*. It was also found that poly(G) DNA inhibited PRC2 activity while poly(A) DNA did not, consistent with findings that PRC2 bound more efficiently to poly(G) DNA (Kaneko, Son, *et al.*, 2014).

Also, increasing concentration of nuclear RNA and yeast tRNA *in vitro* reduced binding of recombinant PRC2 to nucleosomes, suggesting that RNA and nucleosomes were competing with each other for PRC2 binding (Beltran *et al.*, 2016). Moreover, competition assays with protein-free linker DNA and G4 RNA suggested that the competition was between RNA and linker DNA, rather than the core nucleosome particle (Wang, Paucek, *et al.*, 2017a). RNA was shown to allosterically inhibit PRC2 activity by binding to a region that is exposed in both PRC2.1 and PRC2.2, thus revealing an additional aspect of the competition

between RNA and chromatin for holo-PRC2 binding (Zhang *et al.*, 2019). Thus, taken together with iCLIP data showing that PRC2 interacts with nascent pre-mRNA in cells, these results suggest that nascent pre-mRNA inhibits the recruitment of PRC2 to chromatin at genes (Fig. 1.3).

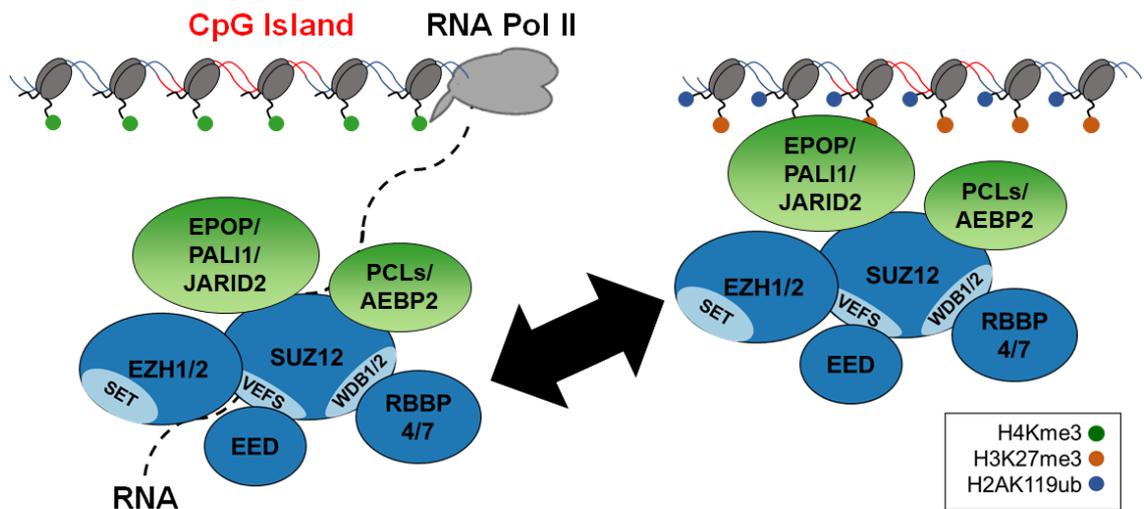


Figure 1.3. Current model of antagonism between RNA and chromatin for PRC2 binding,

At active genes, nascent RNA competes with chromatin for PRC2 binding, evicting existing PRC2 from chromatin and preventing other PRC2 molecules from being stably recruited. On the contrary, when transcription is reduced, PRC2 is released from RNA, and can then bind to chromatin and maintain its target genes in a repressed state. Figure modified from Beltran *et al.*, 2016.

1.4. PRC2 and cancer

1.4.1. Mutations in EZH2 are often found in cancers

Consistent with the fact that PRC2 regulates developmental genes, mutation and consequent deregulation of their components can lead to a wide array of pathological states, ranging from developmental disorders to several types of cancer.

The first direct evidence that dysregulation of PcG proteins could play a role on oncogenesis was the discovery that repression of the INK4A-ARF tumour suppressor locus by BMI1 (PCGF4) (Jacobs *et al.*, 1999) was dependent on EZH2 (Bracken *et al.*, 2007). Afterwards, through gene expression profiling it was described that in individuals with increased EZH2 expression showed a direct correlation with progression and poor prognosis in prostate cancer. Additionally, ectopic expression of EZH2 caused increased proliferation of two different prostate cancer cell lines (Varambally *et al.*, 2002), thus suggesting that EZH2 can act as an oncoprotein when overexpressed.

Additionally, several cancer types exhibit common mutations in EZH2. For example, a recurrent heterozygous point mutation in EZH2 that results in the substitution of residues within the SET domain of EZH2, primarily Y641 and A677, is present in over 20% of cases of diffuse large B-cell lymphoma (DLBCL) and 7% of case of follicular lymphoma (FL) (Sneeringer *et al.*, 2010). It was later revealed that these mutations conferred a gain-of-function phenotype on EZH2, as it was shown that the mutant enzyme is defective in establishing mono-

methylation of H3K27, but more efficiently catalysed H3K27me_{2/3} (McCabe *et al.*, 2012). Moreover, ectopic expression of EZH2^{Y641F} caused a global increase of H3K27me₃, generating lymphoma in mice and cooperating with Braf in melanoma progression (Souroullas *et al.*, 2016).

Mutations in histone H3, the substrate of EZH2, are also found in cancer. In a loss of function mechanism, it was found that a high proportion of cases of diffuse intrinsic pontine glioma (DIPG) contained mutation of histone H3 Lys27 to methionine, and that this was associated with a drastic decrease in H3K27me₃ in paediatric glioblastoma (Lewis *et al.*, 2013). This mutation causes that the methionine, product of the mutation, is placed into the 'lysine' access channel in the active site of EZH2 SET domain. This blocks the binding of other lysines, as this channel has a higher affinity towards methionine (Justin *et al.*, 2016)

1.4.2 Inhibitors of PRC2 activity have anti-tumorigenic properties

A number of academic groups and companies have sought to develop compounds that block PRC2 activity as potential anti-cancer therapies. A high-throughput screening using recombinant PRC2 identified EI1, which inhibited methylation by both WT and Y641F EZH2 by competitive inhibition with SAM. In cell culture, EI1 reduced colony formation and proliferation by the DLBCL cell line WSU-DLCL2 harbouring *Ezh2*^{Y641F} (Qi *et al.*, 2012) and it was later shown that it reduced also proliferation and promoted senescence in malignant rhabdoid tumours (Knutson *et al.*, 2014).

EED inhibitors have also been developed to hinder PRC2 activity. EED226 was identified as an allosteric inhibitor of PRC2, competing with the EED H3K27me3 binding pocket. This has a similar effect on gene expression as EI1 in a human lymphoma carrying a gain of function mutation of EZH2, and also reduced cell proliferation tumour size in xenograft assays (Qi *et al.*, 2017). Another EED binding compound that inhibits PRC2 activity is A-395, which also competes for EED binding to H3K27me3. Like EED225, A-395 inhibits the growth of DLBCL cells *in vitro* and in mouse xenografts (He *et al.*, 2017).

1.4.3. Endometrial stromal sarcoma

Endometrial stromal tumours (EST) are among the rarest forms of uterine malignancies, as they comprise less than 10% of all uterine sarcomas and less than 1% of all primary malignant tumours of the uterus. EST are often associated with endometriosis, occur mostly in the uterus and occasionally in the ovary and peritoneum (Conklin and Longacre, 2014). EST were recently reclassified by the WHO into four subtypes: endometrial stromal nodule (ESN), low grade endometrial stromal sarcoma (LG-ESS), high-grade endometrial stromal sarcoma (HG-ESS) and undifferentiated uterine sarcoma (UUS). These are ordered by increasing malignancy, with UUS having the poorest prognosis and ESN being benign tumours.

LG-ESS is a genetically heterogenous group of sarcomas, most of which contain genomic aberrations resulting in the fusion of several genes. To date, eight translocations that form chimeric genes have been described in LG-ESS:

JAZF1-SUZ12 (Koontz *et al.*, 2001), JAZF1-PHF1, EPC1-PHF1 (Micci *et al.*, 2006), MEAF6-PHF1 (Panagopoulos *et al.*, 2012), MBTD1-CXorf67/EZHIP (Dewaele *et al.*, 2014), ZC3H7-BCOR (Panagopoulos *et al.*, 2013), BRD8-PHF1 (Davidson and Micci, 2017) and MEAF6-SUZ12 (Makise *et al.*, 2019). Notably, a number of these fusion events pair a NuA4 complex subunit with a PRC2 subunit.

1.4.3.1 JAZF1-SUZ12

JAZF1-SUZ12 fusion is generated by the translocation t(7;17)(p15;q21), which fuses the 128 N-terminal amino acids from JAZF1 in place of the first 93 amino acids of SUZ12 (Koontz *et al.*, 2001) (Fig. 1.4). Afterwards, it was reported that ectopic expression of JAZF1-SUZ12 and knockdown of endogenous SUZ12 using small interfering RNAs (siRNAs) in 293 cells increased cell proliferation rate and increased resistance to hypoxic conditions (Li *et al.*, 2007). In this study it was also reported that wild type SUZ12 mRNA was only observed in ESN but not in LG-ESS, suggesting that only one allele is mutated and that silencing of the WT allele (allelic exclusion) generates a more malignant phenotype. It was later described that an mRNA that coded for JAZF1-SUZ12 could be found in normal human endometrial stromal cells (hEnSCs) due to transplicing without the need for chromosomal translocation (Li *et al.*, 2008). However, independent studies have not been able to confirm the existence of this chimeric mRNA (Panagopoulos, 2010).

Information about the biological functions of JAZF1 (also known as TIP27) is limited. It has been described that JAZF1 is a TAK1 orphan receptor, a NR2C2-

selective transcriptional co-factor, and that it may play an important role in mediating transcriptional repression by NR2C2, by interacting specifically through the TAK1-interaction domain (TID) which spans from Asp29 to Lys79 (Nakajima *et al.*, 2004).

Initially found in endometrial malignancies, JAZF1-SUZ12 has also been identified in extrauterine tumours (Sato *et al.*, 2007; Amador-ortiz *et al.*, 2011). Moreover, other non-uterine cancer cases containing JAZF1-SUZ12 have recently been reported, one being a low-grade endometrioid stromal sarcoma of the paratestis (Agaimy *et al.*, 2018). This all suggests that fusion of JAZF1 to SUZ12 might not be specific of uterine malignancies, and could also confer cellular transformation to other tissues.

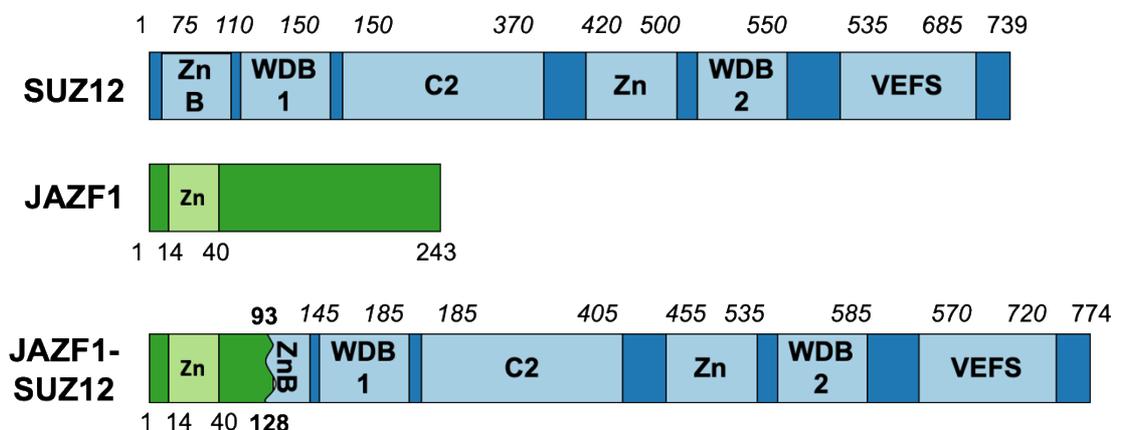


Figure 1.4. Structure of SUZ12 and JAZF1 and their fusion in LG-ESS.

In LG-ESS, the first 93 amino acids of SUZ12 are lost and replaced by the N-terminal 128 amino acids of JAZF1. Domains for SUZ12 are: Zinc finger B (ZnB), WD40-binding (WDB) 1 and 2 and VEFS. JAZF1 only known domain is a Zinc finger (Zn)

That JAZF1-SUZ12 and the other fusion events observed in LG-ESS might have similar effects on endometrial cells was suggested by comparison of the gene expression signatures of LG-ESS that contain JAZF1-SUZ12 versus

those that contain rearrangements of the PHF1 gene, as no significant differences were found (Micci *et al.*, 2016). This suggested that, although different genes are involved in the various translocations in LG-ESS, at least these two translocations must generate biologically and clinically equivalent oncogenic events. Recently, ectopic expression of tagged JAZF1-SUZ12 in human endometrial stromal cells (hEnSCs) coupled with mass spectrometry revealed that JAZF1-SUZ12 interacted with members of the NuA4 HAT complex (Piunti *et al.*, 2019). Therefore, as for the other fusion events identified in ESN and LG-ESS, JAZF1-SUZ12 represents fusion between NuA4 and PRC2 subunits (Table 1.4).

Fusion genes	ESN	LG-ESS	PcG protein	NuA4
JAZF1-SUZ12	65%	48%	SUZ12	JAZF1
PCL1-JAZF1	11%	6%	PCL1	JAZF1
MEAF6-PCL1	-	1	PCL1	MEAF6
EPC1-PCL1	-	1	PCL1	EPC1
MBTD1-EZH1P	-	2	EZH1P	MBTD1
BRD8-PCL1	-	1	PCL1	BRD8
ZC3H7-BCOR	-	1	BCOR	?
MEAF6-SUZ12	-	1	SUZ12	MEAF6

Table 1.4. Gene fusion events in LG-ESS.

Gene fusion events caused by translocations observed in ESN and LG-ESS and their percentage or number of cases. Updated from (Conklin and Longacre, 2014).

1.4.3.1. Endometrial stromal cells and decidualisation

Decidualisation is the process in which the stromal compartment of the endometrium undergoes differentiation to accommodate pregnancy (Gellersen

and Brosens, 2014). Although the human endometrium contains numerous cell types, decidualisation is carried out by a subset population of cells that comprise endometrial stromal fibroblasts (ESF) and decidual stromal cells (DSB), which collectively are known as human endometrial stromal cells (hEnSCs) (Kin *et al.*, 2015). ESF differentiate (decidualise) to become decidual stromal cells (DSCs) during the menstrual cycle (Gellersen *et al.*, 2007; Ramathal *et al.*, 2010). Decidualisation can also be induced *in vitro* using a mixture of 8-Br-cAMP and medroxy-progesterone-acetate (MPA) (Gellersen and Brosens, 2014), which induce the transcription of the decidualisation markers *IGFBP1* (Brar *et al.*, 1997) and *PRL* (Brosens, 1999). Upon decidualisation, EZH2 is downregulated and H3K27me3 reduced at the TSS of *PRL* and *IGFBP1*. Conversely, these genes exhibit a gain of H3K27ac, thus indicating that expression of these genes is regulated by PRC2 (Grimaldi *et al.*, 2011). The common mutation of PRC2 subunits in EST thus suggests that oncogenesis may be related to defects in ESF differentiation due to PRC2 dysregulation.

1.5. Aims of this thesis

There is currently little understanding of how JAZF1-SUZ12 affects PRC2 function. Thus, the aim of this thesis is to establish how the functions of PRC2 are disrupted by the fusion of JAZF1 to SUZ12, with specific focus on the interaction of PRC2 with its accessory subunits and with RNA.

To address this, in Chapter 3 my aim is to establish the composition of PRC2 containing JAZF1-SUZ12, and how this affects PRC2 catalytic activity, cell localization and binding of accessory factors, primarily using *Suz12^{GT/GT}* mouse ESCs as model of study.

Chapter 4 focuses on studying the impact of JAZF1-SUZ12 on the competition between RNA and chromatin for PRC2 binding. To address this, I use mESCs lacking specific PRC2 accessory subunits and use RNaseA treatment to degrade RNA and cell fractionation and nucleosome pull-downs to measure PRC2 interaction with chromatin. Additionally, I test the importance of G4 RNA on PRC2 binding to RNA in mESCs.

Finally, in Chapter 5, I address the consequences of JAZF1-SUZ12 on PRC2-dependent gene regulation and cell state. First, I use primary immortalized fibroblasts to establish whether JAZF1-SUZ12 is oncogenic. Also, I use embryoid body (EB) formation to model dysregulation of gene expression and cell differentiation caused by JAZF1-SUZ12. Finally, by transfection and selection of hEnSCs, I address whether ectopic expression of JAZF1-SUZ12 affects decidualisation in these cells.

2. Chapter 2: Materials and Methods

2.1. DNA cloning

2.1.1. DNA constructs and sequencing primers used in this work

Construct	Promoter	Selection	Source
pCAG-SUZ12-FLAG	Chicken b-actin promoter with CMV enhancer	Puromycin	This work
pCAG-SUZ12D93-FLAG	Chicken b-actin promoter with CMV enhancer	Puromycin	This work
pCAG-LIC-FS2-SUZ12D93	Chicken b-actin promoter with CMV enhancer	Puromycin	This work
pCAG-LIC-FS2-JAZF1-SUZ12	Chicken b-actin promoter with CMV enhancer	Puromycin	This work
pCAG-LIC-FS2-GFP	Chicken b-actin promoter with CMV enhancer	Puromycin	F. Reis
pCAG-LIC-FS2-SUZ12	Chicken b-actin promoter with CMV enhancer	Puromycin	F. Reis
pCAG-LIC-FS2-JAZF1	Chicken b-actin promoter with CMV enhancer	Puromycin	F. Reis
pCAG-GFP-FLAG	Chicken b-actin promoter with CMV enhancer	Puromycin	F. Reis
pCAG-JAZF1-SUZ12-FLAG	Chicken b-actin promoter with CMV enhancer	Puromycin	F. Reis
pMY-SUZ12-FLAG	Hybrid LTRs, with elements of MMLV and MPSV	Blasticidin	This work
pMY-JAZF1-SUZ12-FLAG	Hybrid LTRs, with elements of MMLV and MPSV	Blasticidin	This work
pMY-JAZF1-FLAG	Hybrid LTRs, with elements of MMLV and MPSV	Blasticidin	This work
pCBA-SUZ12-HA	Chicken b-actin promoter with CMV enhancer	Blasticidin	K. Virii
pCBA-SUZ12D93-HA	Chicken b-actin promoter with CMV enhancer	Blasticidin	K. Virii
pCBA-JAZF1-SUZ12-HA	Chicken b-actin promoter with CMV enhancer	Blasticidin	K. Virii
pCBA-JAZF1-HA	Chicken b-actin promoter with CMV enhancer	Blasticidin	K. Virii
pCAG-FS2-AEBP2	Chicken b-actin promoter with CMV enhancer	Puromycin	N. Brockdorff
pCAG-FS2-EPOP	Chicken b-actin promoter with CMV enhancer	Puromycin	N. Brockdorff
AAVS- EPC1 F3S2	hPGK1	Puromycin	J. Cote
PCL3-V5	-	Puromycin	J. Reiter

Table 2.1. Summary of the constructs used in this work.

Cloning primers		
Primer	Sequence	End
pCAGFw-SUZ12-FLAG	GCGCTTCCTCGAGGCCACCATG	N-terminus of SUZ12
pCAGFw-D93-FLAG	GCGCGAATT <u>CGCCACC</u> ATGCCAACACAGATCTATAGA	N-terminus to generate SUZ12D93
pCAGRv-FLAG tag	GCGCGAATTCTCACTTATCGTCGTCATC	C-terminus to generate SUZ12 and SUZ12D93 in pCAG
pCAGLICFW_JAZF1/SUZ12	TACTTCCAATCCATGACAGGCATCGCCGCC	N-terminus to generate FS2-JAZF1-SUZ12
pCAGLICFW_SUZ12D93	TACTTCCAATCCATGCCAACACAGATCTATAG	5' to generate FS2-SUZ12D93
pCAGLICRV_JAZF1/SUZ12	TATCCACCTTACTGTTCATGATGAGCTCGAAGC	C-terminus to generate FS2-JAZF1-SUZ12 and SUZ12D93
Sequencing primers		
pCAG Fw	AGCCTCTGCTAACCATGTTC	Seeded in pCAG
pCAGFS2LIC Fw	CGTGCTGGTTATTGTGCTGTC	Seeded in pCAGLIC
pCAGFS2LIC Rv	GCCTTATTCCAAGCGGCTTC	Seeded in pCAGLIC
SUZ12 PFWD93	CGGGGAUACACAGATCTATAGATTCTTCTCGAACTCGG	Seeded after 93 residue

Table 2.2. Primers used in this work.

2.1.1. PCR amplification of SUZ12 sequences.

For generation of SUZ12 expression constructs, all inserts were PCR-amplified from what previous constructs available in the lab using high fidelity Accuprime *Pfx* DNA polymerase (Invitrogen). All reactions were carried out using 5 μ l of Accuprime *Pfx* Reaction Mix, 10 μ M of primer mix, 10 ng of plasmid DNA, 1 μ l of Accuprime *Pfx* DNA polymerase and water to a final volume of 50 μ l. Thermal cycling was then performed incubating reactions at 95°C for 2 min, then 28 cycles of 2 min at 95 °C, 30 s at 55-64°C and 1 min/kb at 68°C, followed by final extension step of 68 °C of 3 min. The PCR products were resolved by agarose gel electrophoresis, purified with Genron's PCR purification kit (BS664) and eluted in 30 μ l of EB buffer.

2.1.2. Generation of N-terminal FS2-tagged constructs.

5 μ g of pCAG-LIC-FS2 (gift from Neil Brockdorff, Table 2.1) plasmid was linearised with the restriction enzyme *Bae*I (Promega) by incubation overnight at 25°C. The entire digested sample was resolved on a 1% agarose gel, to separate the digested and undigested DNA. The band corresponding to the size of the empty vector was excised and purified by gel extraction (Genron, BS654). The vector was eluted in 30 μ l of EB buffer. In parallel, the insert genes were amplified by PCR with *Pfx* and primers carrying the LIC cloning hybridisation sequence. Single-stranded G overhangs were added to the vector and C overhangs to the amplified inserts by incubation with T4 DNA polymerase (Promega, cat Nno.

M4211) and either dGTP or dCTP (4 nmol, 2 µl of 2 mM stock) for 30 min at 22°C followed by 20 min at 75°C in a 50 µl final volume. Products were then gel-purified (Genron, BS654) in 30 µl of EB buffer. The vector and inserts were hybridised together using different ratios (0.2, 1 and 2µl of insert with 1 µl of vector) for 30 mins at RT. 1 µl of hybridised plasmid was used to transform *E. coli* competent cells as explained below.

2.1.3. Generation of pMY constructs

To subclone into pMY-IRES plasmids, inserts from pCAG constructs were obtained by double digestion using *EcoRI* and *NotI* (NEB), which generated compatible ends in the receiving vector. To increase cloning ratio, the digested vector was treated with calf intestinal alkaline phosphatase (NEB) for 30 min. Afterwards, the products of the digestions were resolved in 1% Agarose gel in tris acetate buffer, and recovered using a gel extraction kit (Genron, BS654). The recovered DNA products of inserts and vectors were ligated using at least 1 µg of vector and 1:3 ratio of insert, and then incubated with T4 ligase (Promega, M1801) overnight at 4°C. The resulting ligation was transformed as explained below.

2.1.4. DNA transformation

Transformation of chemically competent XL-1 Blue *Escherichia coli* strain (prepared in house) was performed as described (Sambrook, 2001), with minor modification. An aliquot of competent cells were thawed on ice and 10 μ l of KCM, 37.5 μ l of water and 5 μ l of ligation reaction were added, and the cells gently mixed. Bacteria were then incubated for 20 mins on ice, and 10 min at RT. Afterwards, 600 μ l of LB was added and incubated for 45 mins at 37°C, with rocking. Finally, the liquid was dispersed throughout the surface of a LB-agar plate (50 μ g/ml ampicilin), and incubated overnight at 37°C.

2.1.5 Colony screening by PCR and plasmid purification

Colonies were screened for inserts by PCR using GoTaq green Master Mix (Promega). Single colonies were picked with a pipette tip, then inoculated in sterile LB by placing the pipette tip briefly, and then pipetted into the PCR mix, which consisted of 12.5 μ l of GoTaq Green Master Mix (2X), 5 μ l of primer mix (10 μ M each) and 2.5 μ l of water. PCR amplification was performed in a thermocycler using the following program: 2 min at 95 °C, then 28 cycles of 95 °C for 30s, followed by 55 °C for 30s and then 72 °C 1 min, followed by 72°C 5 min of final extension. At least 3 positive clones per construct were used to inoculate 3 ml of LB and cultured overnight. Plasmids were purified using the GeneJET mini-prep kit K0503. DNA was quantified by absorbance at 260 nm using a Nanodrop, and 1 μ l of each mini-prepped clone were digested with an

appropriate restriction enzyme (e.g. at least one cut within the insert and one outside) to test for the presence of the gene of interest in the plasmid. Afterwards, inserts were sequenced (by Genewiz) through the entire gene to confirm the correct sequence.

2.2. Cell culture

2.2.1. mESCs

All mESC lines used in this work were grown in feeder-free conditions and seeded in gelatin coated plates (0.1% gelatin solution in PBS). mESC were reseeded every 2nd day to avoid confluence, and their media changed in between.

Suz12^{GT/GT} (gift from D. Pasini) (Pasini *et al.*, 2007), E14, *Aebp2*^{WT/WT} and *Aebp2*^{GT/GT} (kind gifts from N. Brockdorff) (Grijzenhout *et al.*, 2016) and *Jarid2*^{GT/GT} (gift from A. Fisher) (Landeira *et al.*, 2010) mouse ESCs were maintained on 0.1% gelatin in KO-DMEM, 10% FCS, 5% knockout serum replacement, non-essential amino acids, L-glutamine, 2-mercaptoethanol, penicillin-streptomycin and 1000 U/ml leukemia inhibitory factor (03-0011-100, Stemgent). *Pcl1-3* tko, *Pcl2*^{WT/WT}, *Pcl2*^{GT/GT} (Healy *et al.*, 2019) and *Pali1*^{-/-} (Conway *et al.*, 2018) (gifts from A. Bracken) were maintained in GMEM with the same supplements, except with no serum replacement and replacing L-glutamine with GlutaMAX. *Epop*^{GT/GT} and *Epop*^{WT/WT} cell lines (gift from L. Di Croce) (Beringer *et al.*, 2016) were maintained in the same media and supplements as *Pcl*^{WT/WT}, except with 20% FBS.

2.2.2. Other cell lines

Lenti-X 293T (Takara Bio Europe) and NIH-3T3 cells (gift from Barts Vanesbroeck) were grown in DMEM GlutaMAX (Gibco) supplemented with 10% FBS and penicillin-streptomycin. Cells were reseeded when reaching 80% confluency. Immortalized primary human fibroblasts lines expressing human telomere reverse transcriptase (hTERT), hTERT and the SV40 large and small T antigens (SV40) or hTERT, SV40 large and small T antigens and oncogenic HRas (66+++)⁶⁶ (gifts of Paola Scaffidi) (Scaffidi and Misteli, 2011).were grown in MEM (Gibco) supplemented with 15% FBS, penicillin streptomycin and 2mM L-glutamine, and were reseeded once 80% confluency was reached.

2.2.3. Primary human endometrial stromal cells (hEnSCs)

hEnSCs were harvested as previously described (Barros, Brosens and Brighton, 2016) from endometrial biopsies obtained from women attending the Implantation Clinic at University Hospitals Coventry and Warwickshire National Health Service Trust and provided to us by Jan Brosens at Warwick Medical School and University Hospitals Coventry and Warwickshire NHS Trust. All research was undertaken with full ethical approval and with written informed consent obtained from all participants in accordance with the guidelines in The Declaration of Helsinki 2000. In brief, cells were thawed and seeded in either T25 or T75 flasks, depending of the number of cells obtained in the biopsy, in DCC10% media

(DMEM/F12, 10% dextran coated charcoal striped FBS, insulin 2 $\mu\text{g/ml}$, L-glutamine, antimycotic-antibiotic solution and β -estradiol 1 nM). Cells were reseeded when confluence was reached, always using a 1:2 or 1:3 ratio. All experiments were carried out before reaching the 5th passage.

Media	Component	Cat. No.	Concentration	Cells
Immortalized human fibroblasts, Scaffidi 2011	MEM	11095-080	-	hTERT, SV40's and 66+++
	FBS	10270-106	15% (75 ml)	
	P/S	15140-122	1X (5 ml)	
	L-glutamine	25030-024	1X (5 ml)	
mESC, Morey 2012	GMEM	G5154	-	WT (CRSPR), 1A (EPOP -/-), 3A (EPOP -/-), WT (0) and WT + G418 (FLAG-EPOP)
	ES FBS	16141079	20% (100 ml)	
	GlutaMAX	35050038	1X (6 ml)	
	P/S	15140-122	1X (6 ml)	
	NEAA	11140035	1X (6 ml)	
	Sodium Piruvate	11360039	1X (6 ml)	
	b-mercaptoethanol	31350010	0.05 mM	
	LIF	-	-	
mESC, Kanhere 2010	KnockOUT DMEM	10829018	-	SUZ12-/- (Pasini), EZH2 fl/fl, E14, 366 (WT), 368 (AEBP2 -/-), JM8 JARID2 -/-
	ES FBS	16141079	10% (50 ml)	
	KSR	10828028	5% (25ml)	
	L-glutamine	25030-024	1X (6 ml)	
	P/S	15140-122	1X (6 ml)	
	NEAA	11140035	1X (6 ml)	
	Sodium Piruvate	11360039	1X (6 ml)	
	b-mercaptoethanol	31350010	0.05 mM	
mESC, Bracken	GMEM	G5154	-	Pc12 WT/WT, Pc12 GT/GT, PCL1-3 cKO, Gm 340fl/fl -TAT Cre (WT) and Gm 340fl/fl +TAT Cre (Pali1-/-)
	ES FBS	16141079	10%	
	GlutaMAX	35050038	1X (6 ml)	
	P/S	15140-122	1X (6 ml)	
	NEAA	11140035	1X (6 ml)	
	Sodium Piruvate	11360039	1X (6 ml)	
	b-mercaptoethanol	31350010	0.05 mM	
	LIF	-	-	
10% DCC (Jan Brosens Lab SOP WMS JB002)	DMEM/F12	11330	-	Primary Human Endometrial Stromal Cells (HESCs)
	FBS DC stripped	House made	10%	
	Insulin solution	sc-360248	0.002 mg/ml	
	L-glutamine	25030	1X (5 ml)	
	A-A solution	15240062	1X (5 ml)	
	Estradiol	E2758-250MG	1 nM	
2% DCC (Jan Brosens Lab SOP WMS JB002)	DMEM/F12	11330	-	Decidualization Primary Human Endometrial Stromal Cells (HESCs)
	FBS DC stripped	House made	2% (10ml)	
	L-glutamine	25030	1X (5 ml)	
	A-A solution	15240062	1X (5 ml)	

Table 2.3. Culture media composition.

2.5.5. Cell freezing and thawing.

Once cells were ready to freeze (usually at 80% of confluence), cells were trypsinised and counted. Between 5-10 million cells were stored per cryo-vial. Cells were centrifuged at 290 G for 5 min, media was removed and then cells were resuspended in chilled media plus an additional 10% of FBS and 10% of DMSO (Sigma, D2650). Immediately after, cells were stored in isopropanol chamber, pre incubated at RT for 1 hr, and transferred to -80°C.

2.2.4. Generation of stable mESC lines.

mESC were transfected with Effectene (Qiagene, 301425), using a modified version of the manufacturer's protocol. 5×10^6 mESC were seeded onto 6 cm dishes coated with 0.1% gelatin, and left in the incubator for at least 8 hours to allow attachment. 1 µg of plasmid DNA was then topped up to 150 µl with Buffer EC. Afterwards, 16 µl of Enhancer was added, and the mix was vortexed for 1 second. The mix was incubated at room temperature for 2-5 mins and afterward, centrifuged at maximum speed to bring down any liquid. Finally, Effectene reagent was added in a 25:1 ratio (25 µl), and mixed by pipetting up and down for 5 times. The mix was incubated at RT for 5-10 mins to allow transfection-complex formation. During this time, media was removed from cells and changed for 4 ml of fresh mESC media. Afterwards, the Effectene-DNA mix was given a quick centrifuge pulse to bring down any liquid and 1 ml of mESC media was added. The mix was pipetted up & down twice to mix. Immediately afterwards,

the mix was added drop-wise to cells, while swirling the plate. Cells were returned to incubator and left overnight.

Next day (16 hr later), cells were trypsinised and re-plated in a 1:3 ratio on two 10 cm plates, pre-coated with gelatin 0.1%. The remaining volume was resuspended in 100 μ l of Laemmli Buffer 1X to check transfection by immunoblotting. The day after the second seeding, cells were selected with 2 μ g/ml puromycin.

8 to 12 days after starting selection with puromycin, colonies were reseeded on 24-well plates and cultured until reach confluency. Cells were then reseeded on 6-well plates and, once confluency was reached again, cells were trypsinised and cells harvested for cryostorage and verification of transgene expression by immunoblotting.

2.2.7 Transient transfection of NIH-3T3 cells.

1 million cells of NIH-3T3 were plated 10cm plates the day before transient transfection and day afterwards media was changed for 7 ml of fresh media. 7.8 μ g of plasmid was added to a final volume of 362 μ l of OptiMEM and mixed thoroughly. Then 23 μ l of polyethyleneimine (PEI) solution (1 mg/ml) was added, mixed and incubated for 15 min at RT. 350 μ l of the solution was added dropwise to the plates, and then mixed thoroughly by swirling the plate. Media was changed the next day to remove PEI. Cells were harvested 48 hr later.

2.2.5. Transduction of immortalized fibroblasts

A confluent plate of Lenti-293T was split into 4 10cm plates the day prior to transfection. The day after, cells were transfected with a DNA mix comprising 1 µg of pCMVi (gag-pol), 1 µg of pMDG (VSV-G) and 1.5 µg of vector (pMY-IRES) in 200 µl of OptiMEM. To this mix, 10 µl of Fugene HD was added, and then mixed. The solution was then spun down and incubated for 15 min at RT. Media was changed for 8 ml of fresh media, and then the DNA mix added dropwise. Plates were returned to the incubator and the day after new media was added. 48 hrs after transfection, viral supernatant was collected and filtered through a 0.45 µm polyethersulphone (PES) filter (Milipore).

5x10⁵ hTERT, SV40 and 66+++ cells were seeded in 10 cm plates on the day prior to infection. The day after, 3 ml of virus was mixed with 3 ml of MEM media, supplemented with 8 µg/ml of polybrene (Merc, TR-1003) and added to each plate. The media was changed 24 hrs later and selection with 2 µg/ml of blasticidin begun 24 hrs after that. Selection continued until non-transduced control cells were all dead. Expression of the transduced genes was tested by SDS-PAGE and immunoblotting.

2.2.6. Immunofluorescence assays

Suz12^{GT/GT} were plated on 0.1% gelatin-covered coverslips (Sarsted) and then fixed in 4% formaldehyde diluted in warm PBS. Samples were blocked in blocking buffer (1X PBS, BSA 0.3%, Triton X-100) for 1 hr at RT. Coverslips were

incubated with anti-FLAG (M2, F3165) at 1:1500 in blocking buffer for 1 hr. Afterwards, samples were washed twice for 10 minutes with PBS and incubated with a mix of a 1:2000 dilution anti-mouse Alexa 488 (ThermoFisher), Hoesch 1:2500 and a 1:2500 dilution of Phalloidin Rhodamine (R415, ThermoFisher Scientific). Slides were then visualised with a Zeiss LSM 880 and Airyscan confocal microscope and images were processed with ZEN software.

2.2.7. Treatment with G4 binding compounds

E14 cells were reseeded at least once after thawing, prior to be treated with compounds. Once E14 cells were at 80% of confluency, these were trypsinised and reseeded in a 1:10 ratio, and media was changed the day after. The day on the treatment, addition of the compound was added along fresh new media (15 ml per 15 cm plates), and at the appropriate dilution. Once the compound was added, plates were swirled 8 times and returned to the incubator for 4 h.

2.2.8. Embryoid Body (EB) Formation

mESCs were differentiated to form embryoid bodies as described previously (Brien *et al.*, 2012). Briefly, mESCs were cultured and reseeded twice prior to forming EBs. Before seeding, cells were filtered using a 70 mm cell strainer, and washed twice with LIF-free mESCs media. Afterwards, cells were seeded in non-adherent bacterial culture plates in a single cell suspension of 10^6 cells per ml of

media. Media was changed every other day and EBs were harvested at days 4 and 8.

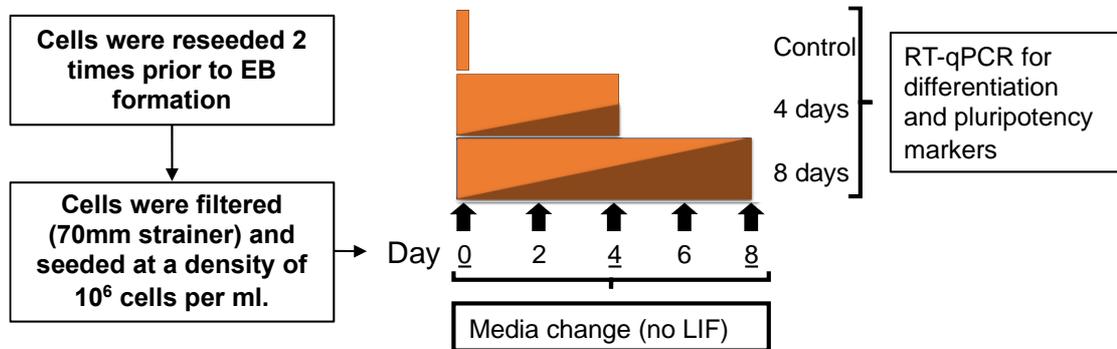


Figure 2.1. EB formation schedule.

2.2.9. Transient transfection of hEnSCs.

Two days prior to transfection, cells were seeded at 2.4 million per 10 cm plate. On the day of transfection, 6 ml of fresh media was added to cells. Cells were then transfected with Effectene using 1.2 μg of DNA in 180 μl of buffer EC, 19.2 μl of Enhancer and 30 μl of Effectene. 16 hours later, cells were washed twice with fresh media and left with 10 ml of fresh media. Two days after that, cells were assayed by immunoblotting.

2.2.10. Transfection, selection and decidualisation of hEnSCs

To generate an enriched population of cells expressing FS2 constructs, hEnSCs were transfected as in protocol 2.2.9. In brief, cells in near confluency (>80%) were transfected using Effectene as above. 16 hrs later, cells were washed twice with fresh media and incubated for 24 hrs with 10% DCC media. Transfected cells were then selected for 3 days with puromycin at 800 ng/ml. After the 3rd day, cells were recovered for two weeks by culture in DCC media with 10% FBS without puromycin, until they reached near confluency. Afterwards, cells were reseeded and two days later the confluent cells were decidualised for up to 8 days by adding DCC media with 2% FBS supplemented with antibiotic antimycotic solution, l-glutamine (1X, Thermofisher), 50 μ M 8-Bromoadenosine 3', 5' cyclic mono-phosphate (8-Br-cAMP, Merck B5386) and 1 μ M methylprogesterone (MPA, M1629). The media was changed every second day.

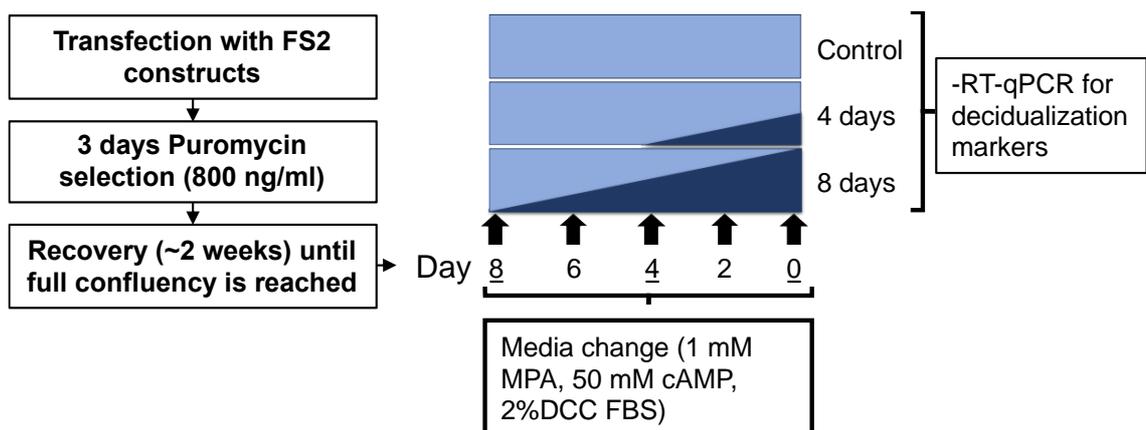


Figure 2.2. Schematic of the generation of decidualised transfected hEnSCs.

2.3. Protein work.

2.3.1. SDS PAGE and Immunoblot

Proteins were separated according to their molecular weight by gel electrophoresis (Lemmli, 1970). The stacking gel was composed of 5% Acrylamide/Bis solution (Biorad), 0.1% SDS, 0.1 % ammonium persulphate (APS), 0.1 N,N,N',N'-Tetramethylethylenediamine (TEMED, Sigma) and 125 mM Tris-HCl pH 6.8. The separating gel had an acrylamide concentration corresponding to the size of the protein of interest (for proteins smaller than 25 kDa, 12%; for a range of 25-130 kDa, 8%; for proteins above 130 kDa, 6%), 0.1% SDS, 0.1% APS, 0.1% TEMED and 400 mM TrisHCl. Protein samples were boiled in a final solution of Laemmli Buffer 1X (2%SDS, 120 mM Tris pH 6.7, 1% beta-mercaptoethanol, 10% glycerol and 0.1% bromophenol blue) for 5 min at 95C. Afterwards, samples were resolved with the Mini-PROTEAN Tetra Cell system, using SDS-PAGE running buffer (200mM Glycine, 24 mM Tris Base and 0.1% SDS) and running at 150 V for at least 1 hr (depending on the concentration of the separating gel).

Gels were then transferred to 8x6 cm of 0.45 μ M nitrocellulose membrane (GE Healthcare Lifescience, 15269794). Nitrocellulose membrane was first soaked in transfer buffer (25mM Tris Base, 192 mM glycine and 20% methanol) and placed on top of two layers of 0.3 mm Whatman filter paper over sponge pads, all pre-soaked in the same buffer. After placing the gel on top of the membrane, two additional layers of pre-soaked filter paper were added on

top. The mini trans-blot transfer cell (Biorad) was then mounted and gels were transferred at 350 mA for 2 hrs.

After the transfer was finished, membranes were blocked with 5% non-fat dried milk plus 0.1% Tween (Sigma) in PBS (PBST) for 1 hr at RT. Then, membranes were incubated with primary antibody (Table 2.3) for how long at what temperature. Membranes were washed three times with PBST for 5 min with rocking, and then incubated with HRP-conjugated secondary antibody (goat anti-mouse (Dako, P0447) or goat anti-rabbit, (Dako, P0448)) at a 1:10,000 concentration in blocking solution for 1 hr at RT. Membranes were washed 3 times for 5 mins each in PBST and then laid out on Saran wrap. Equal amounts of ECL solution peroxide reagent and luminol/enhancer reagent (Clarity, Biorad) were added to the membrane, and allowed to react for 5 min. Images were analysed and quantified using an ImageQuantLAS 4000 imager and ImageQuantTL (GE).

Protein	Company	Catalogue number	Specie	Concentration
FLAG (M2)	Sigma	A8592	Mouse	1:2,000
HA 3F10	Roche (sigma)	12013819001	Rat	1:2,000
V5	Abcam	ab15828	Rabbit	1:2,000
Suz12 (P15)	Sta. Cruz	sc46264	Goat	1:1,000
Ezh2 AC22	Cell Signaling	3147	Mouse	1:1,000
EED (AA19)	Bracken's lab	-	Rabbit	1:10,
EPOP	di Croce lab	-	Rabbit	1:1,000
JARID2	Cell signalling	D6M9X	Rabbit	1:1,000
H3K27me3	Abcam	ab192985	Rabbit	1:5,000
AEBP2	Cell Signalling	D7C6X	Rabbit	1:1,000
MTF2(PCL2)	Proteintech	16208-1-AP	Rabbit	1:300,
FUS	Novus	100-565	Rabbit	1:2,000
HNRNPC	Abcam	ab75822	Rabbit	1:1,000
B-Actin	Cell Signaling	4967	Rabbit	1:1,000
HMGN1	Bethyl	A302-363	Rabbit	1:1,000
H3	Abcam	ab1791	Rabbit	1:5,000
alpha-tubulin	Cell Signaling	2144	Rabbit	1:2,000

Table 2.4. Antibodies used for immunoblotting.

2.3.2. Protein immunoprecipitation (IP) from cytoplasmic and nucleoplasmic fractions

Cells were trypsinised and 40 million cells pelleted at 290xg for 5 min. Cells were washed with PBS and pelleted again. Supernatant was removed and resuspended 500 µl of buffer A (10mM Hepes pH7.9, 1.5mM MgCl₂, 10mM KCl, 0.5 mM DTT and complete protease inhibitor). Cells were incubated on ice for 10 minutes and cells pelleted at 1500xg for 5 mins. The supernatant was removed and 500 µl volumes of buffer A supplemented with 0.1% NP-40 added. Cells were then incubated again on ice for 10 mins with gentle agitation. Nuclei were recovered by centrifugation at 3000 g for 5 mins at 4 °C. The cytoplasmic fraction and is stored for further analysis. Nuclei were resuspended in 150 µl of buffer B (5mM Hepes pH7.9, 26% glycerol, 1.5 mM MgCl₂, 0.2 mM EDTA, 250 mM NaCl, 0.5 mM DTT and complete protease inhibitor) and the concentration of NaCl gradually increased to 400 mM by the dropwise addition of 4.5 µl of a 5M solution of NaCl with continuous mixing. Nuclear extraction was carried out by incubating for 1 hr on ice with occasional agitation. Afterwards, samples were pelleted at 17,000xG for 20 mins at 4°C and the nuclear extract supernatant harvested.

For each IP, the cytoplasmic and nuclear fractions were topped up to 1 ml using buffer BC150 (150mM KCl, 10% glycerol, 50mM Hepes, 0.5mM EDTA, 0.5 mM DTT and complete protease inhibitor). 50 µl of each fraction was saved as input at -20°C..

2 µg of FLAG antibody was added to the remaining 950 µl of each fraction and incubated overnight at 4°C with rotation. The next morning 50 µl of Protein

G Dynabeads were incubated for 2 hrs at 4 °C with rotation. Beads were washed 3 times with buffer BC300 (300mM KCl, 10% glycerol, 50mM Hepes, 0.5 mM DTT complete protease inhibitor). After the last wash, beads were resuspended in 100 ul of 1X Laemli buffer. Input samples were thawed, 50 ul 2x Laemli buffer added and samples boiled for 5 min at 95°C. Samples were briefly spinned down and analysed by SDS-PAGE and immunoblotting.

2.3.3. Protein IP from whole cell extracts.

IPs from whole cell extracts were performed as described (Conway *et al.*, 2018), with small modifications. Briefly, 50 million cells per IP were washed three times with cold PBS, recovered by centrifugation at 290xg, transferred to a 1.5 ml centrifuge tube and then resuspended in 500 µl of high salt buffer (50mM Tris-HCl, pH 7.2, 300mM NaCl, 0.5% (v/v) NP-40, 1mM EDTA pH7.4, complete and DTT 1mM). Cells were then sonicated for 10 seconds three times in a Bioruptor Pico and then rotated at 4°C for 20 minutes before the lysates were diluted with 500 µl of no salt buffer (50mM Tris-HCl, pH 7.2, 0.5% (v/v) NP-40, 1mM EDTA pH7.4, cOmplete inhibitor and freshly added 1µM DTT). Afterwards, lysates were clarified by centrifugation for 10 min at 17,000xg at 4°C and 50 µl were taken from the sample for use as input. The remainder of the lysate was incubated with 2 µg of anti-FLAG M2 antibody (Sigma), or SUZ12 (Cell signalling technologies, 3737), overnight in the presence of 250U/mL Benzonase (with 2mM MgCl₂). The morning after, 50 µl Protein G Dynabeads (Invitrogen, 10003D) per sample were washed 3 times with wash buffer (1:1 dilution of high salt: no salt buffer), and

resuspended in their initial volume. Afterwards, beads were added to protein lysates and incubated for 2 hrs with rotation at 4°C. The flow-through was removed and immunocomplexed beads washed 5 times with 1 ml of wash buffer. After the last wash, the beads were pelleted at 1000xg for 5 min at 4°C and remaining liquid removed. Beads were resuspended in 100 µl of a 1:1 mix of 2X Laemmli buffer and 10 mM Tris HCl for 5 min and 50 µl of 2X Laemmli Buffer added to input samples and both boiled for 5 min at 95°C. Finally, 10 µl of each IP and input sample were resolved by SDS-PAGE.

2.3.4. Streptactin pull-down of FS2-tagged proteins.

Cells were harvested, recovered in a 1.5 ml tubes and washed with PBS. Cells were then resuspended in 500 µl high salt buffer (50mM Tris-HCl, pH 7.2, 300mM NaCl, 0.5% (v/v) NP-40, 1mM EDTA pH7.4, complete and DTT). Cells were then sonicated once for 10 seconds in a Bioruptor Pico (Diagenode). Afterwards, samples were left rotating at 4°C for 20 minutes before diluting with 500 µl of No Salt buffer (50mM Tris-HCl, pH 7.2, 0.5% (v/v) NP-40, 1mM EDTA pH7.4, cOmplete inhibitor and freshly added DTT). Lysates were then clarified by centrifugation for 10 min at 17,000 G at 4°C.

To block any biotinylated proteins binding to streptactin, 10 µg of avidin (Cat No. 2-0204-015) and Benzonase (125 U/ml) were added to the extracts. Cell extracts were then incubated for 30 min at 4°C with rotation, and centrifuged for 5 min at 20,817xg at 4°C. The supernatant was recovered, and 50 µl of the supernatant processed as input controls by adding 50 µl of SMASH buffer (50mM

Tris-HCl pH 6.8, 10% Glycerol, 2% SDS, 0.02% bromophenol blue, 1% beta-mercaptoethanol) and boiling for 5 min at 95°C. Supernatant was added to 10 µl of StrepTactin superflow high-capacity resin (Cat no. , 2-1208-002) previously washed three times with 1 ml of wash buffer (1:1 dilution of high salt: no salt buffer) and the resin recovered by centrifugation at 1000xg for 5 min at 4°C. Resin and cell extract were incubated with rotation for 4 h at 4°C. Then, resin was washed 5 times with wash buffer, each time pelleting the resin at 1000xg for 5 min at 4°C and removing the supernatant. After the last wash, the resin was pelleted again and the remaining liquid removed. Samples were directly processed by adding 100 µl of a 1:1 mix of SMASH buffer and 10 mM Tris-HCl pH 7.5 and boiled for 5 min at 98°C. 10 µl of IP and input samples were then loaded onto a polyacrylamide gel.

2.3.5. Chromatin Immunoprecipitation (ChIP)

ChIP was performed as described (Kanhare *et al.*, 2012). Briefly, cells were trypsinised, resuspended in PBS at 1 million cells per ml and then crosslinked by adding 1/10 of cross-linking solution (11% formaldehyde, 0.1 M NaCl, 1mM EDTA pH 8, 0.5 mM EGTA pH 8, 50 mM HEPES pH 8) for 15 min in suspension. Formaldehyde was quenched by adding glycine to a final concentration of 1.25 mM. Cells were then washed twice with ice-cold PBS, centrifuging at 290G at 4 C for 10 min. After last wash, cells were flash-frozen and stored at -80 C

To carry out ChIP, cells were thawed in 5 ml per 3x10⁷ lysis buffer 1 (50 mM HEPES pH 7.5, 140 mM NaCl, 1 mM EDTA, 10 % glycerol, 0.5 % IGEPAL

CA-630, 0.25 % Triton X-100). Then cells were incubated at 4°C for 10 min with rocking. Cells were recovered by centrifuging at 290 C for 10 min at 4 C. Supernatant was discarded, and nuclei were resuspended in the same volume of lysis buffer 2 (10 mM Tris pH 8, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, supplemented with 0.1 DTT and complete protease inhibitor) and incubated at 4°C for 10 min with rocking. Nuclei were pelleted again at 290 G for 10 min at 4°C, and resuspended in 100 µl of lysis buffer 3 (10mM Tris pH 8, 100mM NaCl, 1mM EDTA, 0.5mM EGTA, 0.1% sodium deoxycholate, 0.5% N-lauryl sarcosine, 0.2 % SDS, supplemented with 0.1 DTT and complete protease inhibitor) per 5×10^6 cells. Cells were incubated for 15 min on ice, to allow lysis.

Afterwards, lysates were sonicated for 5 cycles of 30s on and 30s off, using a Diagenode Bioruptor Pico. The resulting whole-cell extract, equivalent for 4×10^7 cells, was topped up to 900ul per IP. Then, 100 µl of Triton X-100 10% solution were added, and insoluble material was span down at 17,000 G for 20 mins at 4°C. 2% of the lysate was stored at -20 °C to be used as input and whole cell extract and the remaining lysate was incubated overnight at 4 °C with Protein G Dynabeads preincubated for at least 4 hr with FLAG (Sigma, A8592), SUZ12 (Cell Signalling Technologies, 3737), H3K27me3 (Abcam, ab6002), or H3 (Abcam, ab1791) antibodies. Beads were washed and bound complexes eluted and crosslinks reversed by heating at 65 °C. IP and input DNA were then purified by treatment with RNAseA, proteinase K and phenol:chloroform extraction. DNA was resuspended in 60 µl of 10 mM Tris pH 8 and quantified by Nanodrop, with yields ranging from 5 ng/µl for FLAG to 20 ng/µl for H3 and H3K27me3.

Enrichment of specific sequences in ChIP versus input DNA was measured by quantitative PCR (qPCR). Three replicate reactions were performed for each sample. For each reaction, 0.2 μ l of cDNA, 6 μ l of QuantiTECT SYBR Green master mix (Qiagen), 1.44 μ l of 5 μ M of primer mix (Table 2.4.) and 4.36 μ l of water were mixed. Samples were analyzed with an Applied Biosystems 7500 Fast, with the following PCR conditions: 10 min at 95 C followed by 40 cycles of 95°C for 10s and 1 min at 60°C. Production of the PCR product was monitored as the increase in SYBR fluorescence, using as ROX as reference. The cycle threshold (Ct) of each sample was defined automatically by the software. The difference in Cts (Δ Ct) was calculated for each ChIP sample relative to its input sample, by adjusting Ct Input - 6.644 (100 %).

Gene	5'	3'
<i>Hoxd11</i>	GGCACAGCGCCTGTCCAACA	TCTTCCCTGCAGAGCCTACCG
<i>Bmp6</i>	AGCCGCCTCTGAGGGTTC	GCCAGGTGTGTCTCCTAGGCAG
<i>Sox7</i>	CAAGATGCACAACCTCGGAGATC	CTCGGACATGACCTTCCACTC
<i>Fgf4</i>	TCTACTGCAACGTGGGCATC	AGCCCCCGAGACTACTACTG
Intergenic	CCGTGCCCCAGAATTATCAG	GCCGTCCATATCCACCTAAGAA

Table 2.5. Primers used for ChIP-qPCR.

2.3.6. Nuclear fractionation and RNaseA treatment.

RNase A treatment and cell fractionation were performed as described (Zoabi *et al.*, 2014). mESCs were trypsinized, washed twice with PBS, permeabilized with 0.05% Tween-20 in PBS for 10 min on ice, washed once, resuspended with PBS and mock-treated or treated with 1 mg/ml RNase A (Sigma) for 30 min at RT. Cells were centrifuged at 1200 rpm, washed twice with PBS and crosslinked for ChIP or resuspended in 1 ml of buffer A (10 mM HEPES (pH 7.9), 10 mM KCl, 1.5 mM MgCl₂, 0.34 M sucrose, 10% glycerol, 1 mM DTT with Complete protease inhibitor). Triton X-100 (0.1%) was added, and the cells were incubated for 5 min on ice. Nuclei were collected by low-speed centrifugation (4 min, 1,300xg, 4°C). The supernatant (cytoplasmic fraction) was further clarified by high-speed centrifugation (15 min, 20,000g, 4°C). Nuclei were washed twice in buffer A, and then lysed in buffer B (3 mM EDTA, 0.2 mM EGTA, 1 mM DTT, Complete protease inhibitor). Insoluble chromatin was collected by centrifugation (4 min, 1,700xg, 4°C), and the supernatant (nucleoplasmic fraction) was recovered. The final chromatin pellet (chromatin fraction) was washed twice with buffer B and resuspended in 1X Laemmli buffer and sonicated (Diagenode Bioruptor Pico). Samples were analyzed by SDS-PAGE and immunoblotting.

2.3.7. Nucleosome pulldowns.

Everything was carried out RNase-free conditions. 500 μ l of buffer A (10 mM Hepes pH 7.9, 1.5 mM MgCl₂, 10 mM KCl, supplemented with 0.5 mM DTT and cOmplete protease inhibitor) were added to 20 million pre-washed cells. Cells were then incubated on ice for 10 minutes and pelleted at 1500 rpm for 5 min at 4°C. Afterwards, supernatant was removed and 150 ml of buffer A supplemented with 0.1% NP-40 (1.5 μ l of 10% stock solution) added. Cells were incubated on ice for 10 minutes with gentle agitation and recovered by centrifugation at 3000 g for 5 mins at 4°C. Cells were then resuspended in 150 μ l of buffer A supplemented with 150 mM NaCl (4.5 μ l of a 5M stock solution) and incubated on ice for 10 minutes with gentle agitation. Nuclei were recovered by centrifugation 3000 G for 5 mins at 4°C and the cytoplasmic fraction supernatant stored.

Finally, nuclei were resuspended in 150 μ l of buffer B (5mM Hepes pH7.9, 26% glycerol, 1.5 mM MgCl₂, 0.2 mM EDTA, 250 mM NaCl, supplemented with 0.5 mM DTT and cOmplete protease inhibitor) and 4.5 μ l of a 5M NaCl stock solution added gradually to increase the salt concentration to 400 mM) while mixing. Nuclear extraction was carried out by incubating for 1 hr on ice with occasional agitation. Samples were centrifuged at 13,000rpm for 20 mins at 4°C and the supernatant (nuclear extract) harvested. Extracts were quantified by Pierce BCA quantification assays and stored at -80°C. Aliquots were prepared to avoid thaw-freeze cycles.

Nuclear extract was treated with RNase A at a final concentration of 1 μ g/ μ l or mock-treated with the same amount of RNase-free water and RNaseOut (1ul). Both samples were incubated in thermomixer at 37 °C degrees for 30' at

1000 rpm. 5% of the sample was recovered for WCE and the rest was used for pull-down.

Recombinant human histones were expressed in *E. coli* and nucleosomes assembled by the Bartke lab. The nucleosomes were assembled using 147 bp or 185 bp of 601 sequence-containing DNA (Bartke *et al.*, 2010) and containing either HA-tagged H3 or biotinylated H2A (Abcam ab200286). To bind nucleosomes to beads, 50 nM nucleosomes containing biotinylated H2A were incubated with Streptavidin T1 Dynabeads (ThermoFisher) or, for nucleosomes with HA-tagged H3, with pre-coupled anti-HA Dynabeads (ThermoFisher) for 1 hr in nucleosome pull-down buffer (10 mM HEPES pH 7.9, 150 mM NaCl, 0.25 mM EDTA (pH 8.0), 1 mM DTT, 5% glycerol, 0.05% IGEPAL CA-630 and Complete protease inhibitor). Bound nucleosomes were then incubated with 0.2 $\mu\text{g}/\mu\text{l}$ (50 μg in total) or at 0.8 $\mu\text{g}/\mu\text{l}$ (200 μg in total) mESC nuclear extract in nucleosome pull-down buffer for 3 hrs at 4°C with rocking. Afterwards, beads were washed twice in pull-down buffer with 1 M NaCl and then twice in pull-down buffer with 150 mM NaCl. Samples were then resuspended in 1x LDS (ThermoFisher) buffer, heated at 80°C for 10 min and resolved in pre-cast Bis-Tris 12%-6% gradient NuPAGE gels. Immunoblotting was then performed and the difference in SUZ12 levels in the mock and RNaseA treated pulldown samples quantified using ImageQuantTL (GE).

2.4.5. Individual nucleotide-resolution cross-linking and immunoprecipitation (iCLIP)

iCLIP was performed as described (Huppertz et al., 2014). Briefly, cells were irradiated with 0.15 J/cm² of 254 nm UV light in a Stratalinker 2400 (Stratagene). 10⁷ cells were used per IP and were lysed in 1 ml of lysis buffer (50 mM Tris-HCl, pH 7.4; 100 mM NaCl; 1% NP-40; 0.1% SDS; 0.5% sodium deoxycholate and Complete protease inhibitor). Lysates were passed 10 times through a 27 G needle, and then 4 U/ml of DNase Turbo (Ambion) and Rnase I (Ambion, 20 U/ml) were added, and incubated in a thermomixer at 37 °C and 1100 rpm for 3 minutes. Lysates were then clarified by centrifugation for 10 min at 17,000 G at 4°C. 5 % of the lysate was used as input, and the remaining was used for CLIP.

50 µl protein G Dynabeads (ThermoFisher) per IP were washed twice with lysis buffer and then incubated with 5 µg of antibody (SUZ12, 3737 ;IgG, ab171870; FUS, NB100-565; HNRNPC, sc-32308) per IP for 1 hour at RT. Antibody-bound beads were then incubated with lysates on a rotator for 4 hours at 4 °C. Beads were washed 3 times with 900 µl of high-salt buffer (50 mM Tris-HCl, pH 7.4; 1 M NaCl; 1 mM EDTA; 1% NP-40; 0.1% SDS; 0.5% sodium deoxycholate) supplemented with 1 M Urea, and twice with low salt buffer (20 mM Tris-HCl, pH 7.4; 10 mM MgCl₂; 0.2% Tween-20). Beads containing RNA-protein complexes were resuspended in 8 µl of hot PNK mix (0.4 µl PNK (NEB); 0.8 µl 32P-γ-ATP; 0.8 µl 10x PNK buffer (NEB); 6 µl water) and incubated for 5 min at 37°C. Radioactive mix was removed and beads were resuspended in 1x LDS loading buffer (ThermoFisher) and heated at 70°C for 10 min. Samples were loaded on a 4-12% NuPAGE Bis-Tris gel (Invitrogen) according to the

manufacturer's instructions. Protein-RNA complexes were transferred from the gel to a nitrocellulose membrane (0.25 μm , Amersham). After transfer, the membrane was washed twice with 1x PBS and exposed overnight to film (Fuji) at -80°C . Crosslinked RNA was quantified by autoradiography using a Typhoon phosphorimager (GE) and ImageQuantTL (GE). Immunoblotting was then performed as described above and proteins quantified using an ImageQuantLAS 4000 imager and ImageQuantTL (GE).

2.4. Measuring gene expression

2.4.1. RNA purification

RNA was purified using RNase-free tubes and filter tips in an area specifically reserved for RNA work, which was cleaned beforehand with RNase-Away.

Cells were pelleted, washed in 1x PBS and resuspended in 1 ml TRIsure (Bioline, cat. no. BIO-38033). The pellet was broken up both by continuous pipetting and vortexing and the suspension either frozen at -80°C or used immediately for RNA purification. 200 μl of chloroform was added and the tube inverted vigorously by hand for 15s. To separate the phases, the samples were centrifuged at 12,000xg for 15 min at 4°C . The aqueous phase (containing RNA) was transferred to a new RNase-free 1.5 ml tube, 500 μl isopropanol and 1 μl (15 μg) of Glycoblue (ThermoFisher) added, mixed by vortexing and then incubated at RT for 10 min to precipitate RNA. RNA was pelleted at 10,000xg for 10min at 4°C and the supernatant was removed, without disturbing the pellet. The pellet

was then washed with 1 ml of 75% ethanol (room temp, made with RNase-free ddH₂O) and vortexed (making sure that the pellet dislodged from the tube). RNA was then re-pelleted at 6,000xg for 5 min at 4°C, and the supernatant carefully removed. RNA pellets were dried by incubating the tubes on the bench for 5 min with their caps open. RNA was resuspended in 25 µl RNase-free dH₂O by vortexing briefly and incubating for 5 min at RT. RNA was quantified using a Nanodrop spectrophotometer.

2.4.2. cDNA synthesis

Reverse transcriptase-PCR (RT-PCR) was performed using the Im-Prom-II Reverse Transcription System (Promega). Two reactions were performed for each sample: RT+ (200 ng final) and RT- (10 ng final). The purified RNA was topped up to 4 µl and 1 µl of random primers added. Samples were then heated to 70°C for 5 mins and then immediately chilled on ice. The reverse transcription reaction mix was prepared consisting of 6.1 µl Improm II reaction buffer (5X), 2.4 µl of MgCl₂ (25 mM), 1 µl of 10 mM dNTPs, 0.5 ml of RNaseIn (ThermoFisher) and 1 µl of Improm II RT enzyme for RT+ samples or 1 µl of H₂O for RT- samples, giving a total volume of 15 µl. The mix was added to the 5 µl of RNA, and samples were incubated for 10 min at 25°C, then at 50°C for 50 min and then at 70°C for 15 min, to inactivate the reaction. Samples were then stored at -20°C.

qPCR was performed as described in section 2.5.6 using primers shown in Table 2.5 and Δ Ct calculated relative to *Gapdh* or spliced *Actb*.

Species	gene	5'	3'
Human	<i>IGFBP1</i>	CGAAGGCTCTCCATGTCACCA	TGTCTCCTGTGCCTTGGCTAAAC
	<i>PRL</i>	AAGCTGTAGAGATTGAGGAGCAAAC	TCAGGATGAACCTGGCTGACTA
	5S RNA	TTCATTCCGGAGCAGCACTC	TGGTGCTATGAGATTCCGAGTT
Mouse	<i>Oct4</i>	GCTCACCTGGGCGTTCTC	GGCCGCAGCTTACACATGTTT
	<i>Nanog</i>	CCTCCAGCAGATGCAAGAACTC	CTTCAACCACTGGTTTTTCTGCC
	<i>Utf1</i>	ACCAGATCCGCCAACTCATGGG	TCGTCGTGGAAGAAGTAA
	<i>Fgf4</i>	CTACTGCAACGTGGGCATCG	CGCTGCACCGGAGAGAGC
	<i>T</i>	TTTCTTGCTGGACTTCGTGACG	CCACTCCCCGTTTACATATTTT
	<i>Pax3</i>	TCCCATGGTTGCGTCTCTAAG	CTCCACGTGAGGCGTTGTC
	<i>Gata4</i>	CACAAGATGAACGGCATCAACC	CAGCGTGGTGGTGGTAGTCTG
	<i>Fgf5</i>	CTGCAAGATGCACTTAGGACCC	TGAGGAAGAGCAAGGACAGGC
	<i>Actb</i> spliced	AGCACAGAGCCTCGCCTTT	TCATCCATGGTGGTGGTGGC
	<i>Actb</i> unspliced	CCACCCGCGAGCACA	CCGGCGTCCCTGCTTAC
	<i>Gapdh</i>	ATGATGCGCAAAGGTATGCA	CCCCATCTCCCCCTTCT
Transgenes	FS2-JAZF1	CCACCCGAGTTCGAGAAA	TTGTCTCGATGTGCTCGAT
	SUZ12 (c)	AAGGCAGTAACTAAGCTCCGTG	CGTTTGCAGGGGAAGCAGAT

Table 2.6. Primers used to quantify RNA.

3. Chapter 3: Interaction of JAZF1-SUZ12 with PRC2 accessory factors.

3.1. Introduction

One of the major questions in the field is how exactly is PRC2 recruited to the chromatin, and growing evidence points towards a pivotal role from accessory subunits in the distribution of PRC2 (Beringer *et al.*, 2016; Conway *et al.*, 2018).

Recently, it was shown that PRC2 distribution throughout the chromatin is dependent on the core subunit SUZ12 in mESCs. Through knock out of endogenous genes and reconstitution with tagged versions of SUZ12 and EZH2, it was shown that PRC2 was sufficient to re-establish H3K27 methylation patterns even in the absence of prior H3K27me3 (Højfeldt *et al.*, 2018). More importantly, it was also noted that the N-terminus of SUZ12 (SUZ12 Δ VEFS) bound to its target CGIs, but was not able to recover global H3K27me3 levels, whereas SUZ12 VEFS domain recovered H3K27me3 mark in aberrant places as it was unable to interact with CGIs (Højfeldt *et al.*, 2018). Therefore, the N-terminus of SUZ12 is necessary for the recruitment of PRC2.

PRC2 catalytic activity is modulated by its composition. Several mass spectrometry assays have revealed two distinct PRC2 holo-complexes, termed PRC2.1 and PRC2.2 (Alekseyenko *et al.*, 2014; Grijzenhout *et al.*, 2016; Hauri *et al.*, 2016). PRC2.2 is defined by the core PRC2 plus accessory subunits JARID2 and AEBP2 (Grijzenhout *et al.*, 2016), while PRC2.1 is more complex and

contains mutually exclusive subunits EPOP or PALI1/2 (Alekseyenko *et al.*, 2014), plus any of the PCL proteins (Conway *et al.*, 2018). Of note, each PRC2 subcomplex has different catalytic activity, with PRC2.2 being the most active (Beringer *et al.*, 2016; Conway *et al.*, 2018), as PRC2 binding to AEBP2 and a methylated JARID2 generates a conformational change that allosterically enhances EZH2 activity (Kasinath *et al.*, 2018). Interestingly, a new Δ N-JARID2 has been described predominantly in differentiated cells and is unable to bind to PRC2, due to a truncation in its N-terminus. However, this alternate version of JARID2 is still able to bind to its target genes, suggesting that this form can sterically block PRC2 binding to its target genes (Al-Raawi *et al.*, 2019).

It has been proposed that fusion of JAZF1 to SUZ12 in LG-ESS generates a bridge between PRC2 and NuA4, potentially explaining the mechanism behind the pathological phenotype of the translocation (Piunti *et al.*, 2019). Also, it has been suggested that JAZF1-SUZ12 disassembles the core of PRC2 and reduces its catalytic activity (Ma *et al.*, 2017).

In this chapter I sought to establish how the loss of the first 93 residues of SUZ12 in LG-ESS due to fusion with JAZF1 affects PRC2 function, focusing in its catalytic activity, its interaction with target genes and interaction with its accessory factors. Also, I suggest a potential mechanism behind the carcinogenicity of JAZF1-SUZ12 in LG-ESS.

3.2. JAZF1-SUZ12 and SUZ12 Δ 93 recover H3K27me3 in *Suz12*^{GT/GT} mESC

PRC2 catalyses H3K27me3 in mESCs to control transcription of developmental genes (Margueron and Reinberg, 2011). However, it has not been established if JAZF1-SUZ12 affects PRC2 catalytic activity and recruitment, and thus the distribution of H3K27me3 in the genome.

To address this, SUZ12, JAZF1-SUZ12, SUZ12 Δ 93, JAZF1 and GFP expression constructs were generated, each containing a C-terminal FLAG tag (Fig. 3.2 A). The aim of generating this panel of constructs was to establish with SUZ12 Δ 93 if any changes in the cell phenotype was due to the fusion of JAZF1 to SUZ12, and not due the loss of the first 93 aminoacids of SUZ12. Additionally, JAZF1 would show any ectopic binding partners, while GFP was used as a negative control. Using these, I generated single cell clones of *Suz12*^{GT/GT} cells and by immunoblotting I confirmed that the level of expression of all tagged proteins was similar (Fig. 3.1 B). This effectively eliminated the possibility that any phenotype observed is due to the differences in expression of the transgenes. Furthermore, the expression of SUZ12-FLAG was comparable to the endogenous protein, revealing the proteins to be expressed at endogenous levels. However, faint SUZ12 bands were observed in all *Suz12*^{GT/GT} cell lines showing that endogenous SUZ12 expression is not completely ablated in this cell line (Thornton *et al.*, 2014).

Afterwards, I analysed whether the expression of core PRC2 subunits were affected by expression of the SUZ12 constructs, focusing mainly on EZH2 stability which is reduced in the absence of interaction with SUZ12(Pasini *et al.*,

2007). I found that SUZ12 Δ 93 and JAZF1-SUZ12 rescued expression of EZH2, consistent with these proteins being incorporated to PRC2. I also checked whether there was a change in the level of EED, the other PRC2 core component that is also necessary for the establishment of the H3K27me3. However, no consistent changes in EED expression was found with any of the SUZ12 constructs. I next sought to determine the effect of SUZ12 Δ 93 and JAZF1-SUZ12 on PRC2 catalytic activity by measuring the level of H3K27me3. I observed that H3K27me3 was rescued by SUZ12, SUZ12 Δ 93 and JAZF1-SUZ12, but as expected, not by GFP nor JAZF1. However, I noticed that SUZ12 Δ 93, but not JAZF1-SUZ12, promoted a higher level of H3K27me3. This has been previously reported in mESCs when AEBP2 is knocked out, which generates a hybrid JARID2-PCL2 complex that promotes higher levels of H3K27me3 (Grijzenhout *et al.*, 2016).

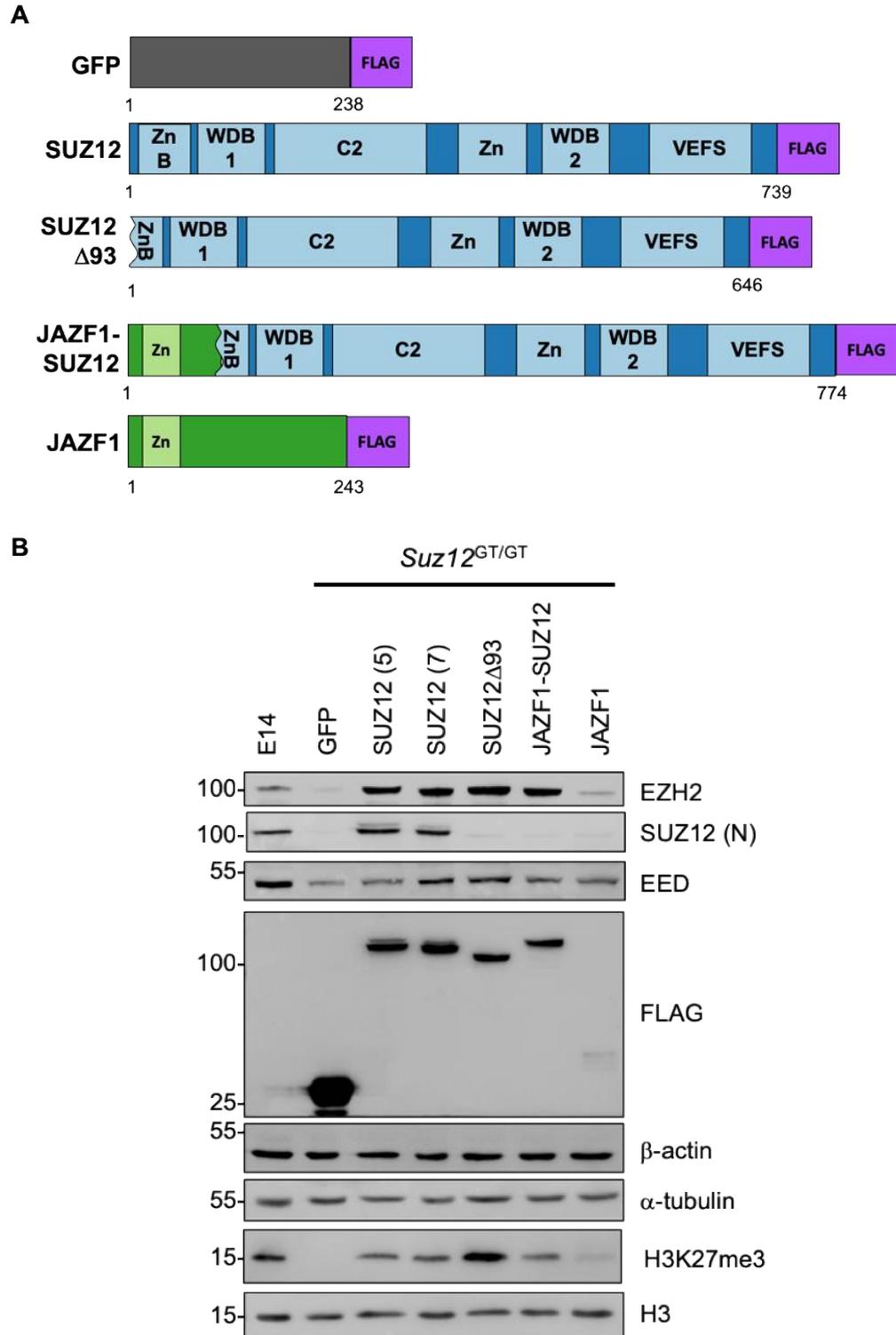


Figure 3.1. JAZF1-SUZ12 rescues H3K27me3 in SUZ12-null cells.

A Diagram of FLAG tagged constructs used to generate stably expressing single cell clones of *Suz12^{GT/GT}* cells.

B Immunoblots for FLAG, EZH2, SUZ12 (antibody to the N-terminus), EED, H3K27me3, β -actin, α -tubulin and H3 in WT E14 cells and in *Suz12^{GT/GT}* cell lines stably expressing different constructs. Two different single cell clones containing full-length SUZ12 were used (5 and 7).

Given the presence of H3K27me3 in JAZF1-SUZ12 cells, I sought to establish whether the fusion protein was recruited to PRC2 target genes. I performed ChIP-qPCR at the well-established PRC2 target genes *Hoxd11* and *Sox7*. I first tested that the immunoprecipitations had successfully enriched for the FLAG tagged proteins by performing immunoblotting (Fig. 3.2 A). Then, ChIP-qPCR assays showed that both the fusion protein and SUZ12 Δ 93 exhibited reduced binding compared to full-length SUZ12 in *Suz12*^{GT/GT} and endogenous SUZ12 in E14 cells (Fig. 3.2 B). This suggests that although JAZF1-SUZ12 and SUZ12 Δ 93 reconstitute catalytically active complexes, these are recruited less efficiently to PRC2 target genes in mESCs.

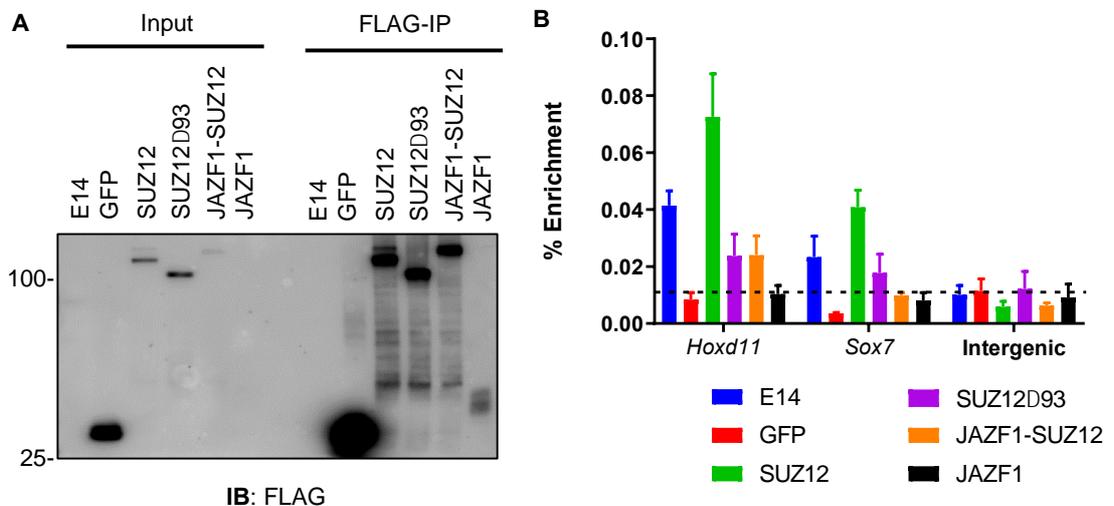


Figure 3.2. JAZF1-SUZ12 and SUZ12 Δ 93 do not bind to PRC2 target genes in *Suz12*^{GT/GT}.

A Immunoblot for FLAG in ChIP input and FLAG-IP of *Suz12*^{GT/GT} cells expressing GFP, SUZ12, SUZ12 Δ 93 and JAZF1, and SUZ12 IP in mESCs E14 cells

B ChIP-qPCR for *Hoxd11* and *Sox7*. A region upstream of *β -actin* was used as ChIP control. (media and s.d of triplicate wells, representative n=2)

3.3. JAZF1-SUZ12 and SUZ12 Δ 93 partly localize to the cytoplasm

PRC2 components are localized to the nucleus. However, the reduction in binding to chromatin observed by ChIP-qPCR, while showing comparable protein levels by immunoblotting, suggested that JAZF1-SUZ12 and SUZ12 Δ 93 might be localized to other places than the nucleus.

To test this hypothesis, I visualized cells by confocal microscopy, using immunofluorescence staining for FLAG, DAPI and phalloidin (Fig. 3A). This experiment showed that in both SUZ12 and JAZF1-SUZ12 proteins were found primarily in the nucleus. Interestingly, SUZ12 Δ 93 protein seemed to have a different distribution as some of the protein was found in the cytoplasm. To confirm these localization patterns, I fractionated cells and performed immunoprecipitation of FLAG-tagged proteins. As expected, all three SUZ12 proteins bound EZH2 in the nuclear fraction. Interestingly, both SUZ12 Δ 93 and JAZF1-SUZ12 proteins were found more abundantly in the cytoplasm than the full-length protein. However, EZH2 was only present in the nuclear fraction, showing that other components of PRC2 were not being exported with the SUZ12 mutants. These results suggested a crucial role of the N-terminus of SUZ12 in the localization to the nucleus of SUZ12, but not for the other components of PRC2.

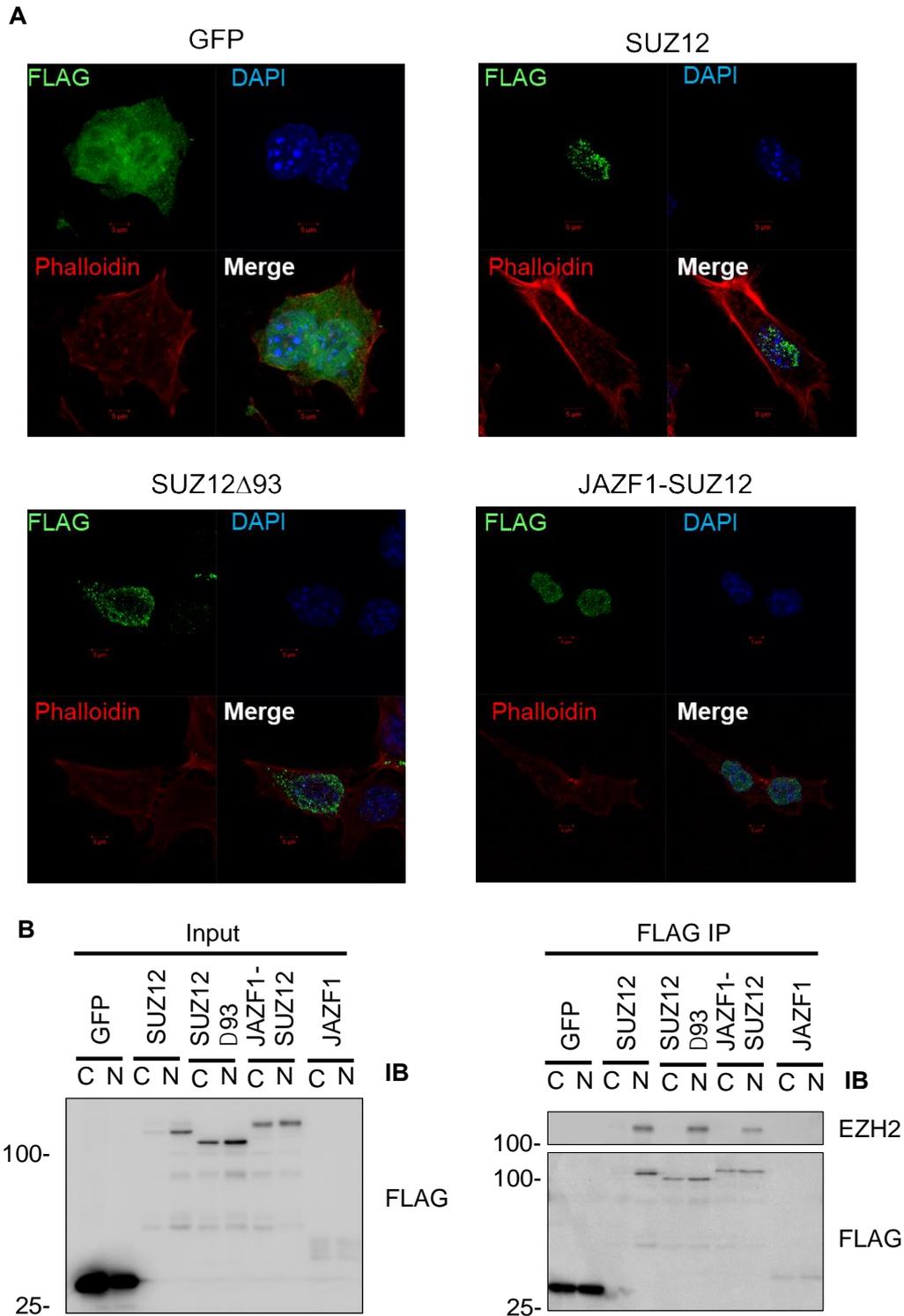


Figure 3.3. JAZF1-SUZ12 and SUZ12 Δ 93 localize to the nucleus and cytoplasm.

A. Representative image for immunofluorescence staining for FLAG, DAPI and Phalloidin in *Suz12*^{GT/GT} cells expressing GFP, SUZ12, SUZ12 Δ 93 and JAZF1-SUZ12. Red bar represents 5 μ m.

B. Left: Immunoblots for FLAG in inputs from cytoplasmic (C) and nuclear (N) extracts from *Suz12*^{GT/GT} cell lines stably expressing GFP, SUZ12, SUZ12 Δ 93, JAZF1-SUZ12 or JAZF1. Right: As left, except immunoblots for FLAG and EZH2 in FLAG-IPs.

3.4. The SUZ12 N-terminus is necessary for interaction with a specific set of accessory factors

PRC2 accessory factors play a pivotal role in the regulation and recruitment of PRC2, as well as in determining its catalytic activity (Beringer *et al.*, 2016; Conway *et al.*, 2018). Therefore, it was important to establish the effect of JAZF1 fusion with SUZ12 on the interaction with these factors. I carried out FLAG immunoprecipitations of benzonase-treated whole cell lysates from each of the *Suz12^{GT/GT}* cell lines. Additionally, I carried out immunoprecipitation of endogenous SUZ12 from E14 mESCs as a positive control (Fig. 3.4). This showed that SUZ12, SUZ12 Δ 93 and JAZF1-SUZ12 proteins interacted with EZH2, corroborating my previous findings that the PRC2 core was being formed. However, the interaction of JARID2 and EPOP with both SUZ12 Δ 93 and JAZF1-SUZ12 was lost, when compared with the full-length SUZ12. Also, slightly reduced interaction with AEBP2 was observed for both SUZ12 Δ 93 and JAZF1-SUZ12 when compared with full-length SUZ12. These results showed that even though the core of PRC2 is still formed with SUZ12 Δ 93 and JAZF1-SUZ12, interaction of PRC2 with JARID2 and EPOP requires the first 93 residues of SUZ12.

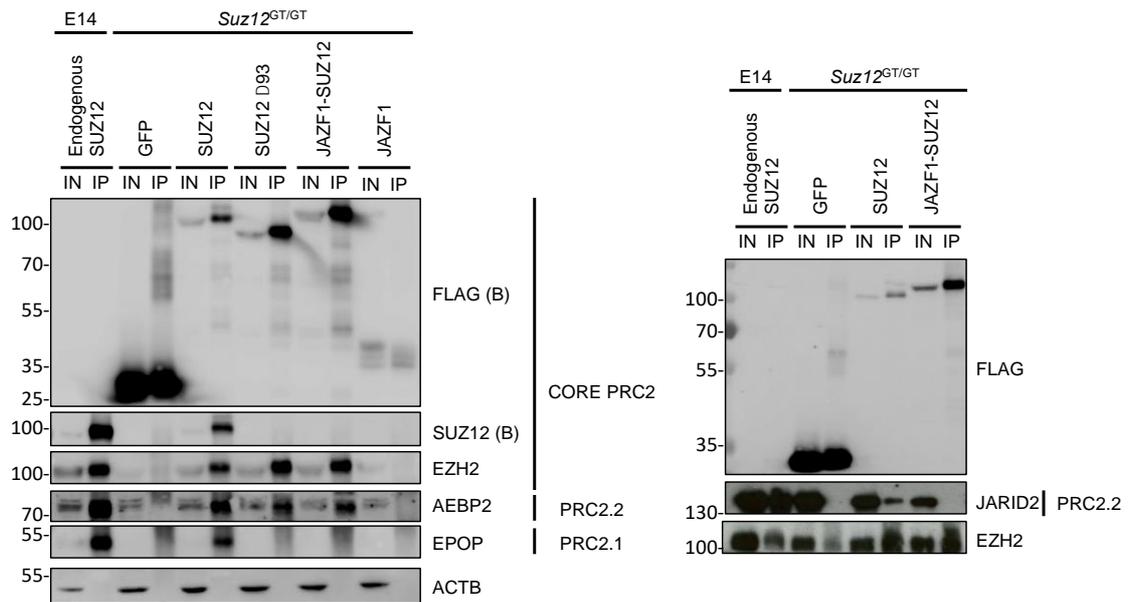


Figure 3.4. EPOP and JARID2 do not interact with SUZ12 Δ 93 and JAZF1-SUZ12 in mESCs.

Left: Immunoblots for FLAG, SUZ12 (N-terminal), EZH2, AEBP2, EPOP and ACTB in input (IN) and FLAG or SUZ12 IPs (IP) from E14 cells or *Suz12*^{GT/GT} cell clones stably expressing GFP, SUZ12, SUZ12 Δ 93, JAZF1-SUZ12 or JAZF1.

Right: As left, except blotting for FLAG, JARID2 and EZH2 in E14 and GFP, SUZ12 and JAZF1-SUZ12 *Suz12*^{GT/GT} cell lines.

After establishing that the endogenously expressed accessory factors EPOP and JARID2 did not interact with JAZF1-SUZ12, I further tested whether lack of the first 93 residues from SUZ12 disrupted interaction with other PRC2 accessory factors. However, given that there were no available antibodies which recognized the remaining PRC2 accessory subunits, I instead transiently co-transfected 3T3 cell lines with HA tagged constructs of SUZ12, SUZ12 Δ 93, JAZF1-SUZ12 and JAZF1 along with tagged PRC2 accessory subunits PCL3 or PALI1. Additionally, I carried out co-transfections of the SUZ12 constructs with FS2-EPOP and FS2-AEBP2 to verify by reciprocal immunoprecipitation the

previously observed loss of JAZF1-SUZ12 interaction with EPOP and reduced interaction with AEBP2.

Immunoprecipitation of V5-PCL3 showed that both JAZF1-SUZ12 and SUZ12 Δ 93 co-precipitated with PCL3, showing that the first 93 residues of SUZ12 are not necessary for PCL3 to bind to PRC2, and possibly for all PCL proteins due to their level of homology (Fig 3.5A). Pull-down of FLAG-AEBP2 and blotting for HA revealed that AEBP2 interacted with all SUZ12 constructs (Fig 3.5B), which showed that if there was a reduction of interaction between JAZF1-SUZ12 and AEBP2, this was not observable using ectopic expression. Finally, to verify the loss of interaction between EPOP and PRC2 containing JAZF1-SUZ12, I immunoprecipitated FS2-tagged EPOP (Fig. 3.4C). This revealed a significant reduction in EPOP binding to JAZF1-SUZ12 and SUZ12 Δ 93 but equal levels of EZH2 binding, confirming what I had observed previously by pull down of FLAG-tagged proteins in *Suz12*^{GT/GT} cells.

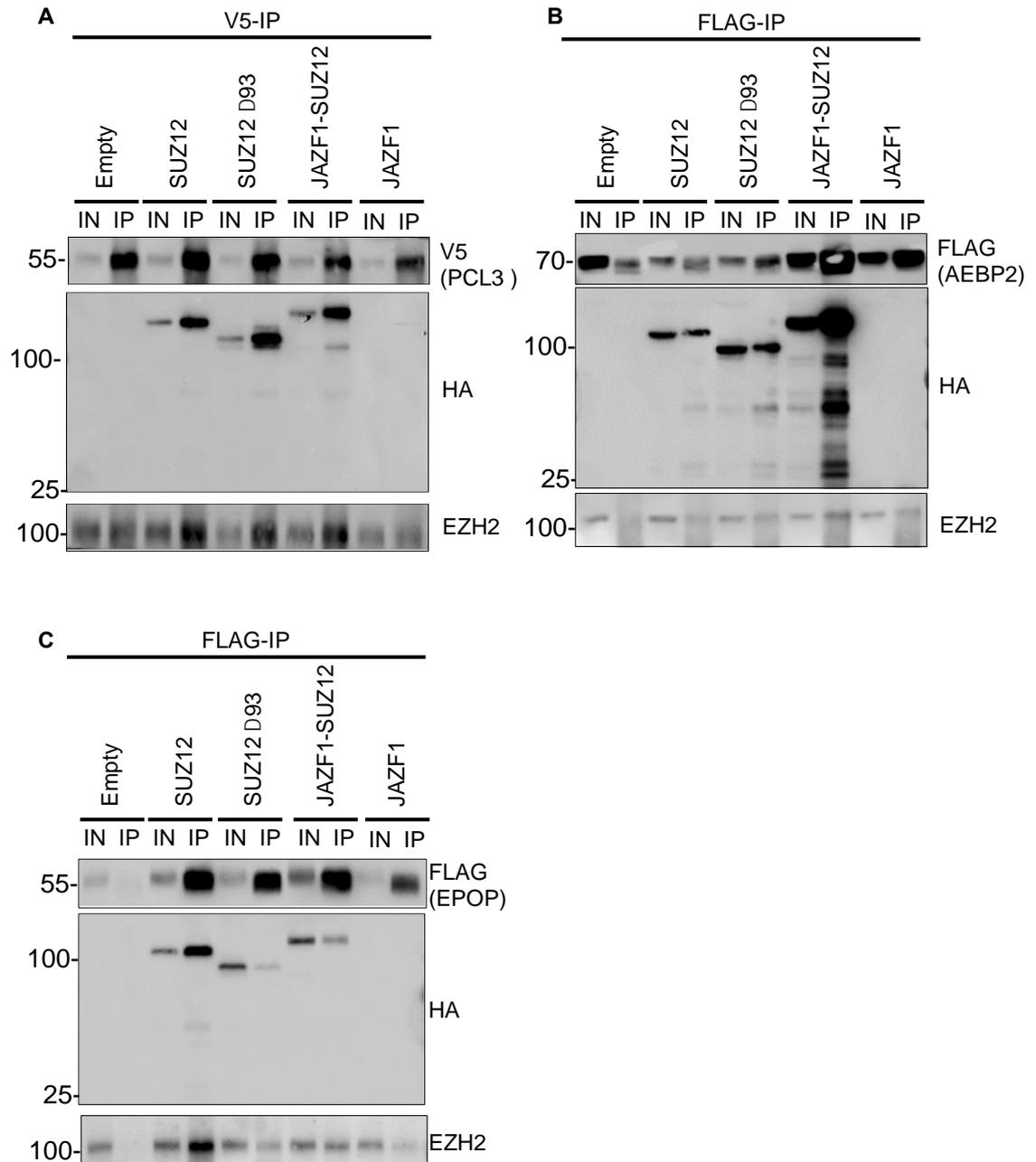


Figure 3.5. The first 93 amino acids of SUZ12 are needed for interaction with EPOP, but not with AEBP2 or PCL3.

A. Immunoblots for FLAG, V5, HA and EZH2 in 3T3 cells co-transfected with V5-tagged PCL3 along with HA-tagged constructs.

B. As A, except immunoblots for FLAG, HA and EZH2 in 3T3 cells co-transfected with FS2-AEBP2 and HA-tagged constructs

C. As B, except in 3T3 cells co-transfected with FS2-EPOP and HA constructs

Considering that the PRC2 accessory subunit PALI1 was tagged with both FLAG and HA, I generated the same panel of SUZ12 constructs using a N-terminal FS2 tag instead, which can be pull-down with streptactin resin.

Once I generated the FS2 SUZ12 constructs, I co-transfected FS2-tagged SUZ12 constructs with HA/FLAG-tagged PALI1 subunit in 3T3 cells and pulled-down with streptactin resin to test whether PALI1 bound to SUZ12 N-terminus (Fig. 3.6). This showed that interaction of PALI1 with SUZ12 Δ 93 and JAZF1-SUZ12 was reduced compared to full-length SUZ12, thus following the same pattern as EPOP and JARID2. Therefore, the N-terminus of SUZ12 is necessary for interaction with EPOP, JARID2 and PALI1.

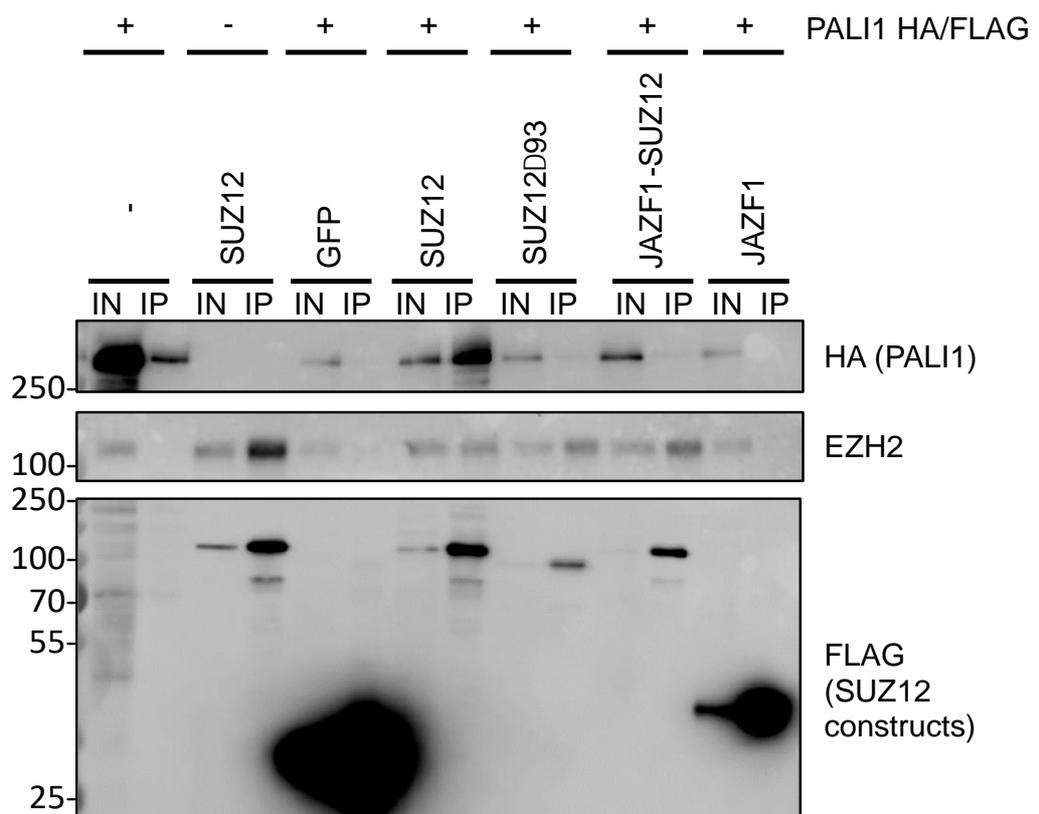


Figure 3.6. PALI1 does not interact with SUZ12 Δ 93 and JAZF1-SUZ12.

Streptactin pull down of FS2 tagged proteins in 3T3 cells co-transfected with HA/FLAG-PALI1 and FS2 SUZ12 constructs. Immunoblots for HA, EZH2 and FLAG of inputs (IN) and pulldown (IP) of each transient co-transfection.

3.5. JAZF1-SUZ12 does not interact with the NuA4 complex in 3T3 cells

It has been proposed that fusion of JAZF1 to SUZ12 in LG-ESS generates a bridge between PRC2 and NuA4, potentially explaining the mechanism behind the pathological phenotype of the translocation (Piunti *et al.*, 2019). To validate these findings and gain further understanding of the consequences of the fusion event, HA constructs were co-transfected along with FS2-tagged EPC1 and pulled-down using streptactin resin. Testing ectopic binding between JAZF1-SUZ12 and EPC1 was of particular interest, given that EPC1 forms part of the minimal catalytic subcomplex of NuA4 that shows robust HAT activity *in vitro* (Doyon *et al.*, 2004). Immunoblotting for HA and FLAG showed that both proteins were properly expressed in the cells. However, streptactin pull-down of EPC1 did not identify an interaction between JAZF1-SUZ12 and EPC1, raising the possibility that JAZF1 does not interact directly with EPC1 in 3T3 cells. This suggests that JAZF1 requires a protein absent in 3T3 cells to interact with the NuA4 complex and that loss of interaction of JAZF1-SUZ12 with PRC2 accessory factors is not due to interaction with NuA4.

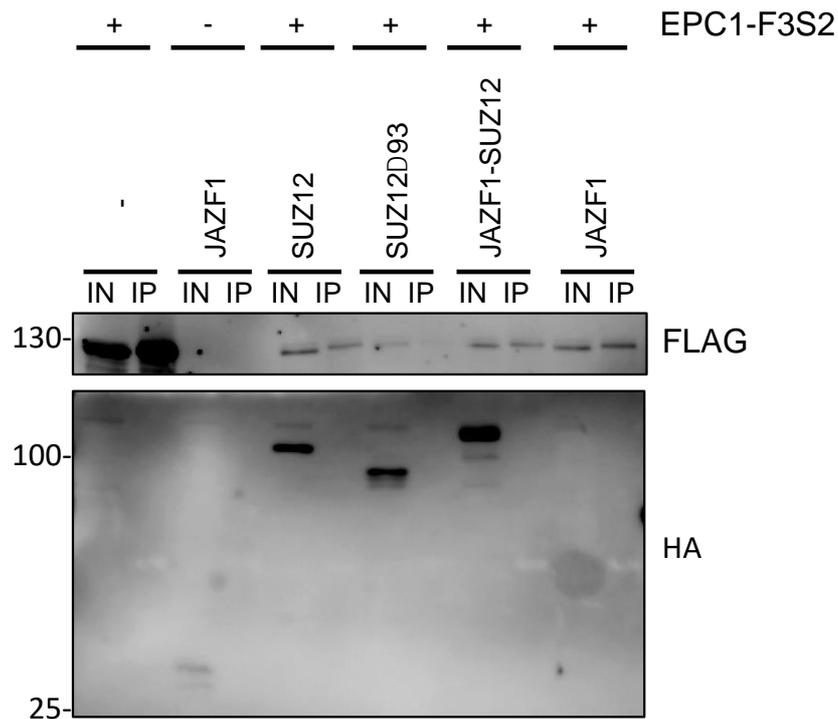


Figure 3.7. JAZF1-SUZ12 does not bind to the NuA4 complex in 3T3 cells.

Immunoblots for FLAG and HA in streptactin pull down of 3T3 cells co-transfected with FLAG-tagged EPC1 and HA-tagged SUZ12 constructs. Input (IN) and streptactin pull down (IP).

3.6. hEnSCs express accessory subunits EPOP and JARID2ΔN

JAZF1-SUZ12 is found in benign ESN and malignant ESS. Phenotypically, ESN and ESS resemble normal proliferative-phase endometrial stroma, with the key difference being that the latter presents myometrial invasion (Conklin and Longacre, 2014). Thus, primary hEnSCs cultured *in vitro* could potentially be used to model the onset of both tumours harbouring JAZF1-SUZ12 (Barros et al. 2016). To establish the expression of PRC2 components in primary hEnSCs, and thus verify the validity of the model, I carried out comparative immunoblots of mouse E14 and human 293T cell lines along with primary hEnSCs cells.

Immunoblots showed expression of SUZ12 in steady-state hEnSCs. Interestingly, both JARID2 and EPOP were expressed in hEnSCs (Fig. 3.8). Human EPOP exhibited a lower molecular weight than in mouse, which has previously been reported (Beringer *et al.*, 2016). However, a shorter version of JARID2 of approximately 80 kDa was found in hEnSCs. This shorter form of JARID2 has previously been reported to be expressed in predominantly in differentiated cells, such as HUVECs and lymphocytes, and is unable to bind to PRC2, due to the cleavage of its N-terminal SUZ12 interacting domain (Al-Raawi *et al.*, 2019). Thus, lack of a full length JARID2 with the potential to bind to PRC2, and presence of EPOP in hEnSCs, suggests that JAZF1 fusion disrupts only the interaction of SUZ12 with EPOP and perhaps PALI1, but not with JARID2, as this last interaction is not present in hEnSCs.

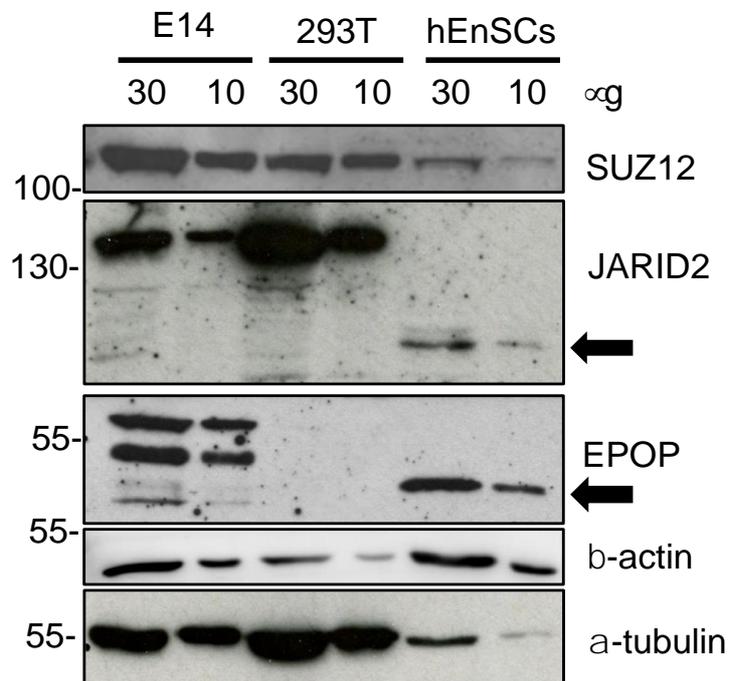


Figure 3.8 Accessory subunits EPOP and JARID2 are expressed in hEnSCs.

Immunoblots for SUZ12, JARID2, EPOP, β -actin and α -tubulin in E14 mESCs, 293T and hEnSCs. Arrows mark JARID2 Δ N and human EPOP.

3.7. Summary and discussion

In this chapter I show the importance of the N-terminus of SUZ12 in its nuclear localization, the recruitment of PRC2 to its target genes and in establishing the composition of a holo-PRC2.

Regarding the catalytic activity of PRC2, I found that JAZF1-SUZ12 and SUZ12 Δ 93 can recover H3K27me3 methylation, showing that there is no need for the SUZ12 N-terminus to recover H3K27me3, as has been shown before (Højfeldt *et al.*, 2018). This correlates with the fact that the VEFS domain remains unaffected by the translocation observed in LG-ESS, and therefore does not hinder the stability of the catalytic subunit EZH2. However, I observed that JAZF1-SUZ12 reduced PRC2 binding to its target genes *Hoxd11* and *Sox7* in mESCs. This suggests that in LG-ESS that harbour JAZF1-SUZ12 there might be dysregulation of developmental genes usually repressed by PRC2. Further experiments will be necessary to establish whether JAZF1-SUZ12 also exhibits a recruitment defect to its target genes in hEnSCs and the relevance of this for ESS tumourigenesis.

I also found that the N-terminus of SUZ12 is important for nuclear localization, given that both SUZ12 Δ 93 and JAZF1-SUZ12 proteins were found in both the nucleus and cytoplasm. However, neither mutant protein caused translocation of the PRC2 to the cytoplasm. Further experiments are required to determine whether this phenotype could have relevance for JAZF1-SUZ12-associated pathology.

I showed by co-IP and streptactin pull-down that the first 93 residues of SUZ12 are necessary for PRC2 interaction with the accessory subunits EPOP, PALI1 and JARID2, but not for the interaction with PCL3 nor AEBP2. This suggests that PCL proteins and AEBP2 interact with another part of SUZ12 or with another PRC2 subunit. Co-temporal reports to this work describe that although mutually exclusive, the AEBP2 binding domain in SUZ12 is localized between residues 95-106, while PCL2 binding domain spans SUZ12 residues 338-353 (Youmans, Schmidt and Cech, 2018), thus a lack of the first 93 residues of SUZ12 should not affect its binding to PRC2, which I verified in this work. EPOP and PALI1 are reported to be mutually exclusive within PRC2.1 (Hauri *et al.*, 2016), and these two accessory factors are mutually exclusive with JARID2 (Grijzenhout *et al.*, 2016; Conway *et al.*, 2018). Therefore my findings provide an explanation for this, revealing that the first 93 residues of SUZ12 are necessary for these factors to be assembled with PRC2 (Fig 3.9). In an effort to understand the effect of this translocation, ectopic transient co-expression of tagged versions of SUZ12 and JAZF1-SUZ12 in 293T cells, followed by immunoprecipitation of endogenous proteins EPOP and JARID2, showed that both subunits were bound preferentially to the WT protein (Chen *et al.*, 2018b).

It has been reported that the fusion of JAZF1 to SUZ12 generates ectopic PRC2 binding partners belonging to the NuA4 complex (Piunti *et al.*, 2019). However, expression of JAZF1-SUZ12 and EPC1 in 3T3 suggests that this interaction is not direct and another protein missing in 3T3 cells might be required. Further studies incorporating a control sample in which SUZ12 and EPC1 are known to interact are required to identify the NuA4 subunit(s) that interact with JAZF1-SUZ12 and the role of the interaction between PRC2 and NuA4 in LG-

ESS carcinogenesis. Also, mass spectrometry assays using nuclear extract of transduced endometrial stromal cells with JAZF1-SUZ12 HA/FLAG tagged exhibited that accessory subunit AEBP2 interaction from PRC2 was lost (Piunti *et al.*, 2019).

Finally, I showed that hEnSC express EPOP and thus PRC2.1 can be formed, but express a truncated version of JARID2 that is known to lack the PRC2 interaction domain (Al-Raawi *et al.*, 2019). Thus hEnSC are reliant on PRC2.1 for PRC2 function. The lack of interaction of JAZF1-SUZ12 with EPOP and PALI1 thus suggests that specific dysregulation of PRC2.1 function could be the driver of ESN and LG-ESS tumourigenesis. This is further supported by clinical reports of other translocations in LG-ESS that generate fusions between genes encoding NuA4 subunits with genes encoding accessory factors of PRC2.1, but not PRC2.2 (Micci *et al.*, 2006, 2014; Panagopoulos, Mertens and Griffin, 2008; Panagopoulos *et al.*, 2012; Dickson *et al.*, 2018). The expression of PALI1 and PALI2 in hEnSC would need to be measured to assess whether loss of JAZF1-SUZ12 interaction with these proteins could also potentially contribute to the carcinogenicity in LG-ESS.

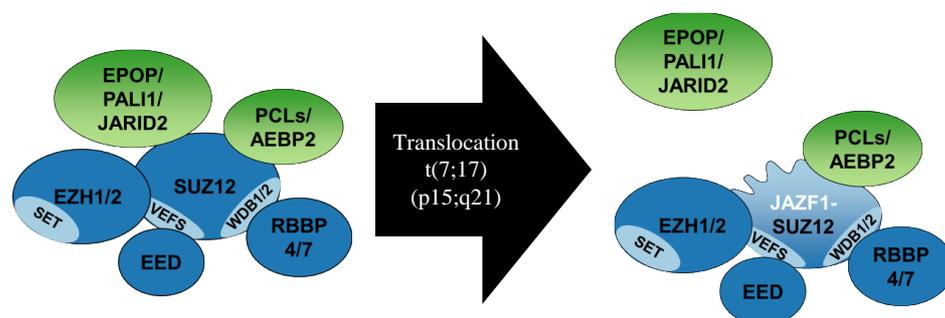


Figure 3.9 Model describing the main results of Chapter 3.

PRC2 interaction with JARID2, EPOP and PALI1 subunits is lost due to lack of the first 93 aas of SUZ12 in the JAZF1-SUZ12 fusion protein generated by the t(7;17) translocation in LG-ESS.

4. Chapter 4: Regulation of PRC2 by RNA in cells is modulated by accessory subunits.

4.1. Introduction

Binding of PRC2 to chromatin is mediated by several of its components. The core of PRC2 interacts with nucleosomes through EZH1/2 binding to H3K27, EED interactions with H3K27me₃, and RBBP4/7 binding to unmodified H3 (Laugesen, Højfeldt and Helin, 2019). Moreover, PRC2 is most likely recruited to CGIs by PCL proteins (Li *et al.*, 2017; Perino *et al.*, 2018), and by JARID2 to H2AK119ub, deposited by PRC1 (Blackledge *et al.*, 2014; Kalb *et al.*, 2014a; Cooper *et al.*, 2016).

Chromatin structure also determines PRC2 recruitment and activity. *In vitro* methylation assays showed that dense nucleosomal arrays were modified more than dispersed nucleosomes, suggesting that the preferred substrate of PRC2 is compacted chromatin (Yuan *et al.*, 2012). Also, ablation of *Dnmt3a* resulted in an increase in H3K27me₃ levels across large genomic regions associated with repression of genes, showing that methylated CpG antagonizes PRC2 binding to chromatin (Wu *et al.*, 2010).

Recent studies identified a major role for RNA in recruitment of PRC2 to chromatin. Global removal of RNA by RNaseA (Beltran *et al.*, 2016) or inhibition of RNA Polymerase II (Riising *et al.*, 2014) caused recruitment of PRC2 to chromatin at active genes. This revealed an antagonism between PRC2 binding to RNA and chromatin. Evidence from multiple groups shows that PRC2 interacts

with RNA through the subunits SUZ12 (Kanhere *et al.*, 2010; Beltran *et al.*, 2016) and EZH2 (Kaneko *et al.*, 2013; Kaneko, Son, *et al.*, 2014). More recent data has also identified the PRC2 catalytic core as the source of the specificity of PRC2 for G4 RNA (Long *et al.*, 2017). However, accessory subunits have also been found to directly interact with RNA, chiefly JARID2 (Cifuentes-Rojas *et al.*, 2014; Kaneko, Bonasio, *et al.*, 2014) and AEBP2 (Wang, Goodrich, *et al.*, 2017), suggesting multiple points of RNA interaction within holo-PRC2.2.

UV-crosslinking methods have shown that nascent pre-mRNAs and nascent non-coding RNAs are bound by PRC2 in mESCs (Kaneko *et al.*, 2013; Beltran *et al.*, 2016). This suggested that PRC2 binds promiscuously to all nascent RNA. However, recent studies using recombinant PRC2 consisting of SUZ12, EZH2, EED, RBBP4 and AEBP2, revealed a preference for PRC2 RNA sequences comprised of G nucleotides. These RNA containing repeats of G-tracts form *in vitro* three-dimensional structures known as G-quadruplexes (G4), coordinated by K⁺ ions. The interaction between G4 RNA and PRC2 *in vitro* is reduced by steric hindrance from pyridostatin (PDS), a stabilizer of G4 structures, and by destabilisation of G4 structures generated by the cationic porphyrin, 5,10,15,20-tetra(N-methyl-4-pyridyl)porphyrin TMPyP4 (Wang, Goodrich, *et al.*, 2017).

In this chapter, I address the role of PRC2 accessory subunits in the antagonism between RNA and chromatin for PRC2 binding. I also sought to determine the role of G-quadruplexes in PRC2 binding to RNA in cells.

4.2. JARID2 but not AEBP2 are necessary for increased recruitment of PRC2 to chromatin upon RNA degradation

Recent reports have shown that PRC2.2 subunits JARID2 and AEBP2 also bind RNA (Kaneko, Bonasio, *et al.*, 2014; Wang, Goodrich, *et al.*, 2017). However, the role that these and other PRC2 accessory subunits play in the antagonism between RNA and chromatin for PRC2 binding has not been explored.

To establish whether AEBP2 or JARID2 modulated the antagonism between RNA and chromatin for PRC2, *Aebp2*^{WT/WT}, *Aebp2*^{GT/GT} and *Jarid2*^{GT/GT} cells were permeabilized, RNaseA treated and fractionated (Fig 4.1). This showed that the levels of both SUZ12 and EZH2 in the chromatin fraction of *Aebp2*^{WT/WT} cells increased when compared to mock treated cells, confirming previous findings (Beltran *et al.*, 2016). This increase in PRC2 binding to chromatin upon RNase treatment was also observed in *Aebp2*^{GT/GT} cells, showing that lack of AEBP2 does not affect the recruitment of PRC2 to chromatin when RNA is depleted. However, when the JARID2 defective cell line was treated under the same conditions, no increase of PRC2 subunits was observed on chromatin (Fig. 4.1). These results suggest that although belonging to the same complex, JARID2 is necessary for chromatin enrichment when RNA is depleted, while AEBP2 is not.

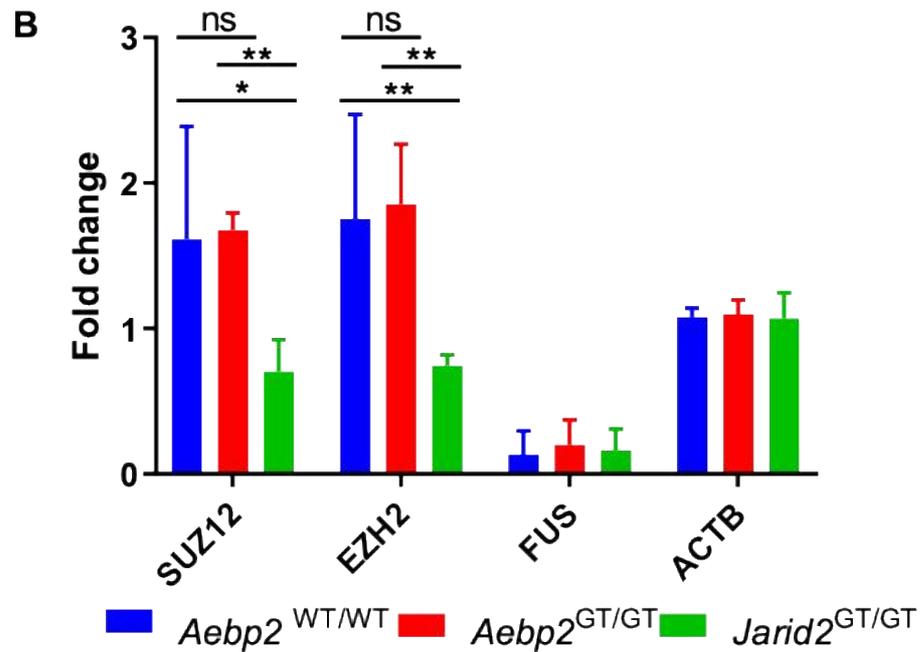
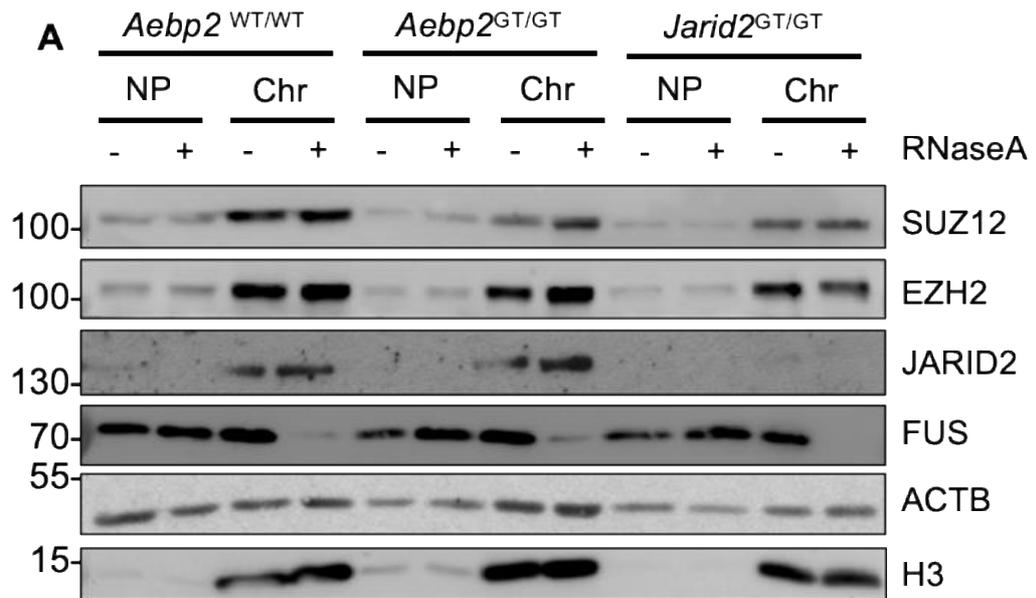


Figure 4.1. JARID2, but not AEBP2, is necessary for PRC2 recruitment to chromatin upon RNA degradation.

A. Immunoblots for SUZ12, EZH2, JARID2, FUS, ACTB and H3 in nucleoplasm (NP) and chromatin (Chr) fractions derived from mock or RNaseA-treated permeabilised WT, AEBP2-null and JARID2-null cells. The loss of FUS from the chromatin fraction controls for RNA degradation; H3 controls for fractionation and ACTB controls for gel loading. Representative of three independent experiments.

B. Fold change in SUZ12, EZH2, FUS and ACTB in the chromatin fraction upon RNaseA treatment (mean and s.d. of 3 independent experiments, * $p < 0.05$, ** $p < 0.005$).

4.3. EPOP and PALI1 but not PCLs are necessary for increased recruitment of PRC2.1 to chromatin upon RNA degradation

To establish whether this phenotype was exclusively observed in mESCs lacking PRC2.2 subunits, I tested the effect of RNaseA treatment in mESCs where *Epop*, *Pali1* and *Pcl1/2/3* genes were ablated. Strikingly, in cells lacking EPOP there was no enrichment of SUZ12 and EZH2 in the chromatin fraction with RNaseA treatment (Fig 4.2). In fact, a decrease of SUZ12 in chromatin was observed when compared to the mock treated cells. Furthermore, mock treated *Epop*^{GT/GT} cells showed increased SUZ12 and EZH2 in the chromatin fraction compared to mock treated *Epop*^{WT/WT} cells. These results suggested that EPOP acts within PRC2.1 in an analogous manner to JARDI2 in PRC2.2 to increase the recruitment of PRC2 to chromatin upon release of PRC2 from RNA.

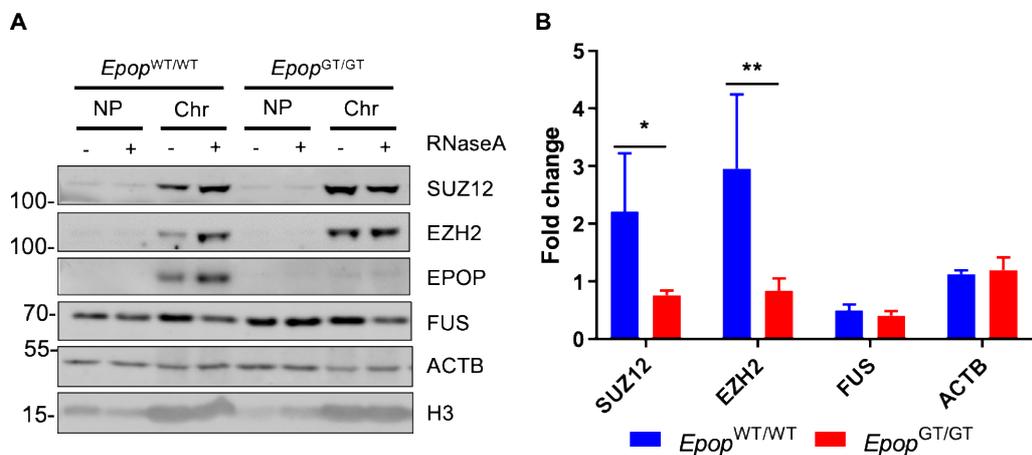


Figure 4.2. EPOP is necessary for recruitment of PRC2.1 to chromatin upon release from RNA.

A. Immunoblots for SUZ12, EZH2, EPOP, FUS, ACTB and H3 in nucleoplasm (NP) and chromatin (Chr) fractions derived from mock or RNaseA-treated permeabilised WT and EPOP-null cells. The loss of FUS from the chromatin fraction controls for RNA degradation; H3 controls for fractionation and ACTB controls for gel loading. Representative of three independent experiments.

B. Fold change in SUZ12, EZH2, FUS and ACTB in the chromatin fraction upon RNaseA treatment (mean and s.d. of 3 independent experiments, * $p < 0.05$, ** $p < 0.005$).

I then sought to determine whether the other PRC2.1 subunit PALI1, which can bind to PRC2 in place of EPOP, and PCL1/2/3 are also necessary for the increase in PRC2 chromatin binding upon RNA degradation. I found that loss of PALI1 mimicked the phenotype in *Jarid2*^{GT/GT} and *Epop*^{GT/GT} cells, with SUZ12 showing a reduction in increased chromatin binding when cells were treated with RnaseA. Surprisingly, loss of PCL1,2 and 3 did not cause any significant difference in the increase in PRC2 chromatin association under the same treatment (Figure 4.3). This suggests that PALI1, EPOP and JARID2 may play a similar role in the recruitment of PRC2 to the chromatin when RNA is absent, while AEBP2 and PCLs do not.

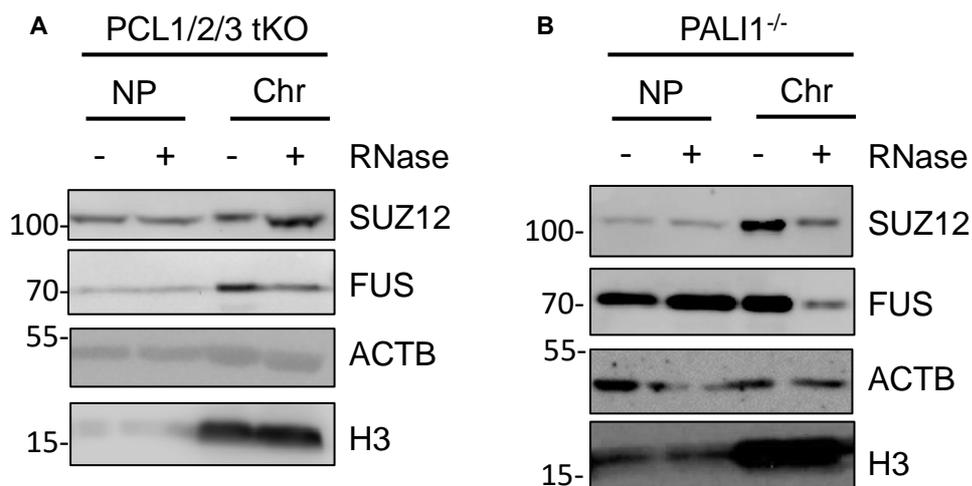


Figure 4.3. Lack of PALI1, but not PCLs, reduces SUZ12 recruitment to chromatin under RnaseA treatment.

A. Immunoblots for SUZ12, FUS, ACTB and H3 in nucleoplasm (NP) and chromatin (Chr) fractions derived from mock or RNaseA-treated permeabilised PCL1/2/3 tKO cells. The loss of FUS from the chromatin fraction controls for RNA degradation; H3 controls for fractionation and ACTB controls for gel loading (n=1).

B. As A, except in fractions derived from mock or RNaseA-treated permeabilised PALI1^{-/-} cells (n=1).

4.4. JAZF1-SUZ12 is not recruited to chromatin upon RNA degradation

I had previously established that JAZF1-SUZ12 does not interact with EPOP, JARID2 and PALI1 but still interacts with AEBP2 and PCLs. Consequently, I sought to establish whether lack of recruitment of PRC2 to chromatin upon RNA degradation was also observed with JAZF1-SUZ12. To this end, I treated *Suz12*^{GT/GT} cells with RNaseA and purified the chromatin fraction (Fig 4.4). Interestingly, PRC2 in cells expressing JAZF1-SUZ12 phenocopied PRC2 in *Epop*^{GT/GT}, *Pali1*^{GT/GT} and *Jarid2*^{GT/GT} cells, as PRC2 increase in chromatin was no longer observed when cells were treated with RNaseA. These results suggest that JAZF1-SUZ12-containing PRC2 is unable to increase its recruitment to chromatin when RNA is degraded, due to the lack of interaction with PALI1, EPOP and JARID2.

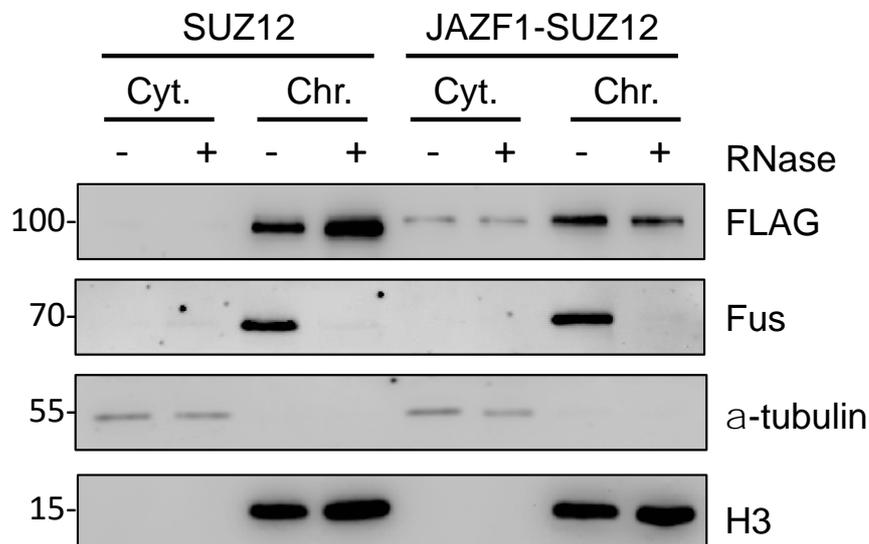


Figure 4.4. JAZF1-SUZ12 is not recruited to chromatin upon RNA depletion.

Immunoblots for FLAG, FUS, α -tubulin and H3 in cytoplasm (Cyt) and chromatin (Chr) fractions derived from mock or RNaseA-treated permeabilised *Suz12*^{GT/GT} cells expressing SUZ12 or JAZF1-SUZ12. The loss of FUS from the chromatin fraction controls for RNA degradation; α -tubulin and H3 controls for fractionation and loading. Representative of 2 experiments.

4.5. Accessory subunits do not affect recruitment of PRC2 to nucleosomes upon RNA degradation

Assemblies of recombinant nucleosomes are a useful tool for modelling interactions between chromatin and its associated proteins (Bartke *et al.*, 2010). For example, it has been previously reported that RNA can evict the PRC2 catalytic core from nucleosomes *in vitro* (Kaneko *et al.*, 2013; Wang, Paucek, *et al.*, 2017a). However, the role of PRC2 accessory subunits in the competition between nucleosomes and RNA for PRC2 binding has not been addressed. To this end, I performed nucleosome pull-downs using nuclear extract from mESC lacking specific accessory factors.

Firstly, I sought to identify the optimum nuclear extract concentration in order to observe the competitive effect of RNA on PRC2 nucleosome binding. Consequently, I tested two concentrations (200 and 50 μg) of nuclear extract protein from E14 mESCs followed by pull-down with nucleosomes with HA-tagged H3. Additionally, to test if PRC2 was recruited to the core nucleosome particle or to linker DNA, these nucleosomes were formulated either as di-nucleosomes or as mono-nucleosomes, which were generated with (reconstituted with 183 bp DNA) or without (reconstituted with 147 bp DNA) linker DNA. To validate the specificity of the assay, HMGN1 was used as a negative control (Fig 4.5A). I found that RNA depletion from nuclear extracts increased PRC2 binding to nucleosomes, regardless of the nucleosome conformation. Increases in PRC2 binding were more apparent in pull-downs that were performed with 50 μg of nuclear extract protein, likely because 200 μg of protein saturated the nucleosomes. Additionally, I found that HMGN1 was not binding to

nucleosomes under our experimental conditions, demonstrating the stringency of our assay. PRC2 exhibited a stronger increase in binding to nucleosomes lacking linker DNA (Fig 4.5B). This suggest that linker DNA plays an important role in the binding of PRC2 to nucleosomes in steady-state, corroborating previous results *in vitro*. However, PRC2 binds to nucleosomes upon release from RNA regardless of linker DNA, which means that RNA competes with binding to the core nucleosome particle, and not with linker DNA, contradicting previous findings(Wang, *et al.*, 2017b).

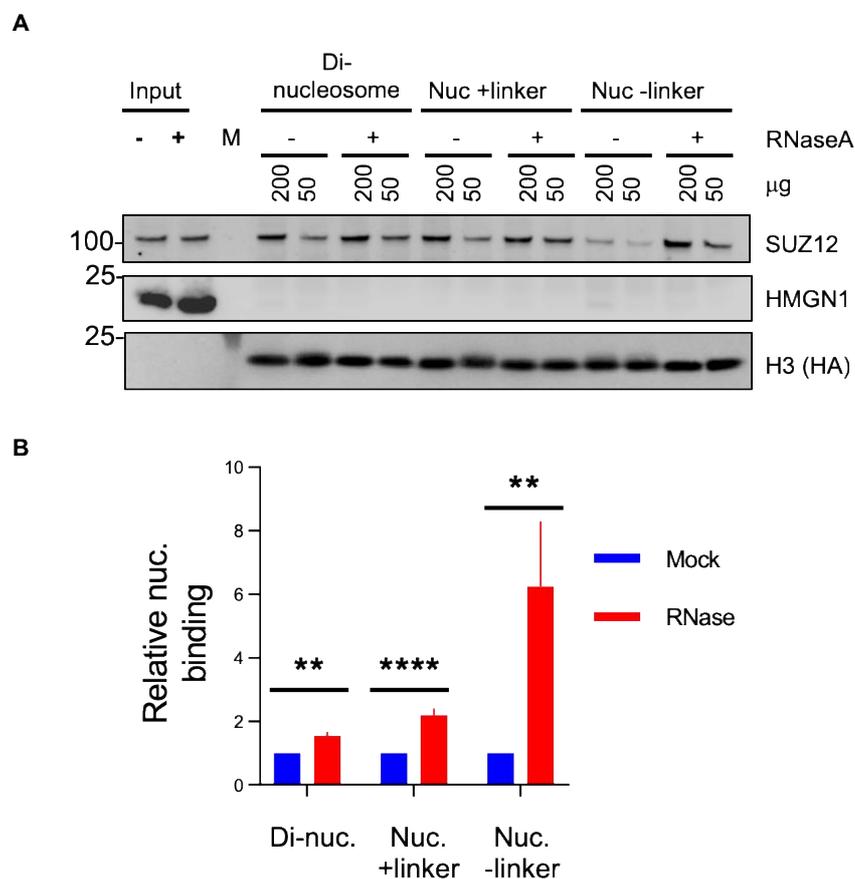


Figure 4.5 Release of PRC2 from RNA triggers its recruitment to nucleosomes.

A Immunoblots for SUZ12, HMGN1 and HA in inputs and pull-down with dinucleosomes, mono nucleosomes with and without linker DNA, using 200 or 50 µg of protein from nuclear extract of E14 cells, mock or RNaseA treated. Representative of 3 independent experiments

B Relative binding of PRC2 to nucleosomes with 50 µg of mock and RNaseA treated nuclear extracts protein (mean and s.d. of 3 independent experiments, **p<0.005, ****p<0.00005).

Considering that AEBP2 binds to DNA (Kim, Kang and Kim, 2009), I hypothesized that this might be necessary for PRC2 to bind nucleosomes with linker DNA. To test this, I carried out the same RNaseA treatment of nuclear extracts from *Aebp2*^{GT/GT} cells along with its matched WT control cell line, followed by nucleosome pull-down (Fig. 4.6). Interestingly, lack of AEBP2 did not cause any noticeable differences in the enrichment of PRC2 on chromatin induced by treatment, nor in the binding to nucleosomes with linker DNA, showing that AEBP2 is not necessary for recruitment of PRC2 to nucleosomes upon its release from RNA. This is consistent with my findings using RNaseA treatment of permeabilised cells (Fig 4.1).

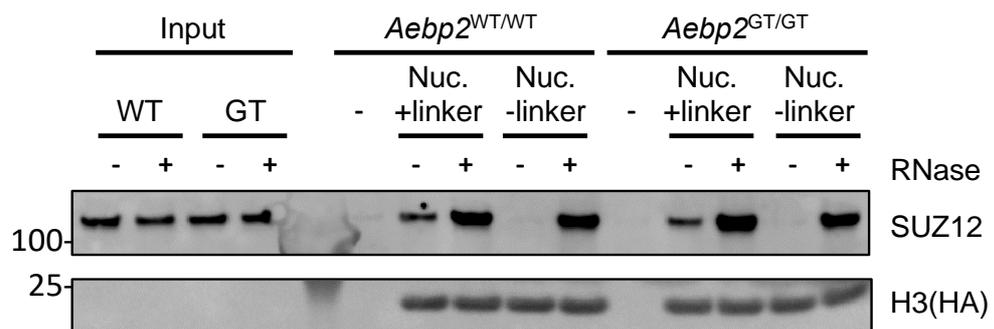


Figure 4.6 AEBP2 is not necessary for PRC2 to bind nucleosomes upon its release from RNA.

Immunoblots for SUZ12 and HA-H3 in inputs and nucleosome pull-downs from mock or RNaseA treated nuclear extract from WT and AEBP2-null cells. Representative of 2 independent experiments.

To validate these findings with a more stringent pull-down, I repeated the experiments using nucleosomes in which H2A was biotinylated. In this assay, immunoblotting for H3 was used to verify equal loading of nucleosomes, and to confirm that the purified nucleosomes were still assembled (Fig. 4.7A). This assay confirmed that there was no significant difference in the recruitment of

PRC2 from RNA to nucleosomes in cells lacking AEBP2. Given the requirement for JARID2 in the recruitment of PRC2 binding to chromatin, I next tested whether JARID2 was necessary for recruitment to nucleosomes when RNA is degraded in nuclear extracts (Fig 4.7 B). I found that the loss of JARID2 did not cause any significant differences in nucleosome binding. This shows that JARID2 is not necessary for recruitment from RNA to unmodified nucleosomes. The difference in these results with my previous findings measuring recruitment to chromatin upon RNase treatment (Fig 4.1) suggests that JARID2 may only be necessary for recruitment to nucleosomes containing modified histones.

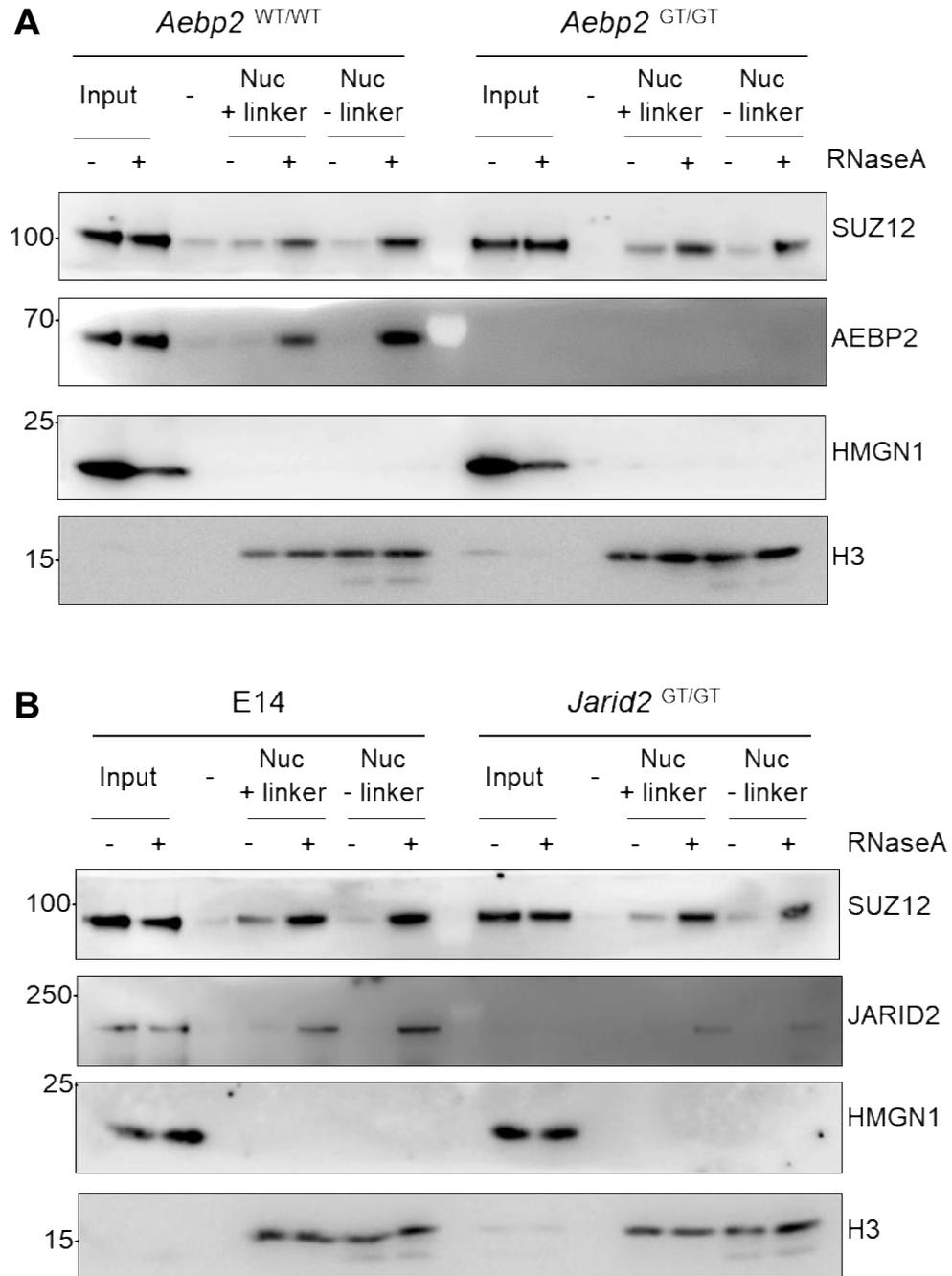


Figure 4.7 Accessory subunits JARID2 and AEBP2 are not necessary for PRC2 binding to nucleosomes.

A Immunoblots of SUZ12, AEBP2, HMGN1 and H3 in input and nucleosome pull-downs from mock and RNaseA treated nuclear extracts of WT and AEBP2-null cells. Representative of 2 independent experiments.

B As A, except for WT E14 and JARID2-null cells. Immunoblots for JARID2 confirms loss of this protein. Representative of 2 independent experiments.

Finally, I asked whether PCLs were necessary for the recruitment of PRC2 from RNA to nucleosomes upon RNA degradation. Nucleosome pull-down from *Pcl2*^{GT/GT} and *Pcl2*^{WT/WT} cell nuclear extracts did not reveal any difference in the enrichment of PRC2 on nucleosomes upon RNaseA treatment (Fig 4.8 A), demonstrating that PCL2 is not necessary for PRC2 to bind to nucleosomes when released from RNA. However, in mock treated cells, there was a slight decrease in the binding of PRC2 to nucleosomes in *Pcl2*^{GT/GT} nuclear extracts when compared to extracts from WT cells. This difference might suggest that PCL2 enhances affinity towards nucleosomes in steady-state but does not affect the competition between nucleosomes and RNA.

To determine whether the binding in steady state of PRC2 that remained to nucleosomes was due to any of the remaining PCL proteins, I carried out nucleosome pull-downs from nuclear extracts of PCL2-null and PCL1/2/3-null cells (Fig 4.8B). This assay showed no observable difference between PRC2 nucleosome binding in the presence or absence of PCLs 1 and 3. I conclude from this that PCLs are not necessary for PRC2 to bind to nucleosomes in steady state.

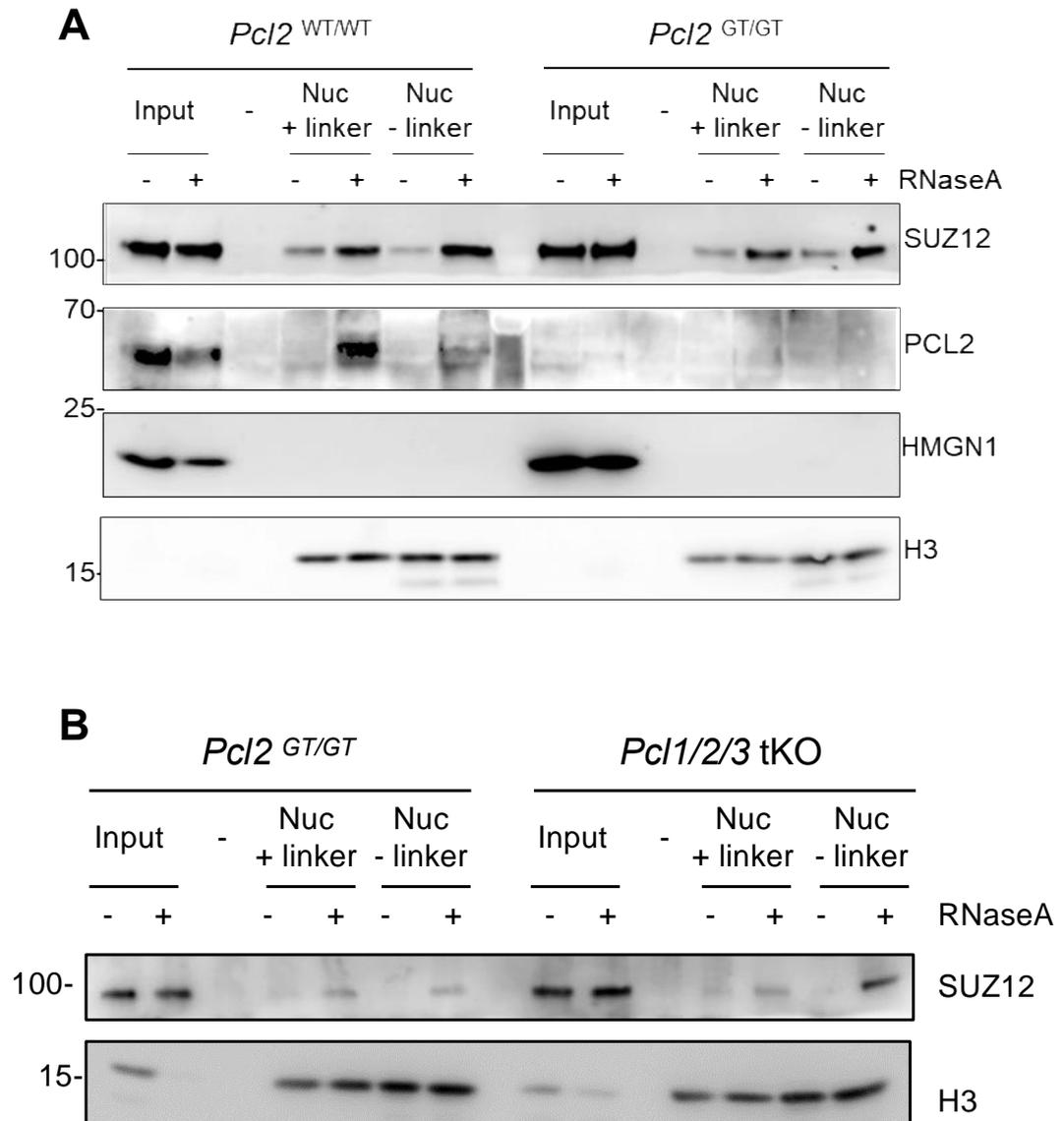


Figure 4.8 PCLs are not necessary for PRC2 binding to nucleosomes.

A Immunoblots of SUZ12, AEBP2 and HMGN1 and H3 as specificity and loading controls, respectively, in nucleosome pull down assays of nuclear extracts using AEBP2-null and its WT match cell lines, mock and RNaseA treated. Representative of 2 independent experiments

B As A, except using nuclear extracts from PCL2-null compared to PCL1/2/3-null cells.

4.6. Effect of PDS in PRC2 and RNA binding in mESCs

PRC2 preferentially binds RNA that contains G-tracts, which form G4 structures. This interaction is hindered by PDS *in vitro*, a compound that competes with PRC2 for binding to G4 RNA (Wang, Goodrich, *et al.*, 2017). Additionally, data from our lab revealed repeated G-tract sequences predicted to form G4 structures were the most enriched at PRC2-RNA crosslink sites in mESCs. Thus, to determine whether G4 structures promoted PRC2 RNA binding in cells, I performed CLIP for the core PRC2 subunit SUZ12, FUS and HNRNPC in cells treated with or without 10 μ M of PDS for 4 hrs. Recapitulating its effect *in vitro*, I found that PDS significantly reduced the binding of PRC2 to RNA in cells (Fig. 4.9A). PDS also reduced RNA binding by FUS, another protein that is known to bind preferentially to G4 RNA, but had no effect on HNRNPC, a poly-U binding protein (Fig. 4.9 B). To verify that this effect did not reflect inhibition of transcription, I used qPCR to quantify both spliced and un-spliced β -actin mRNA (Fig. 4.9 C). There was no change in the ratio between spliced and unspliced β -actin RNA after PDS treatment, thus confirming that the observed reduction of RNA binding was due to competition of PDS with PRC2 for RNA binding and not due to reduction in transcription.

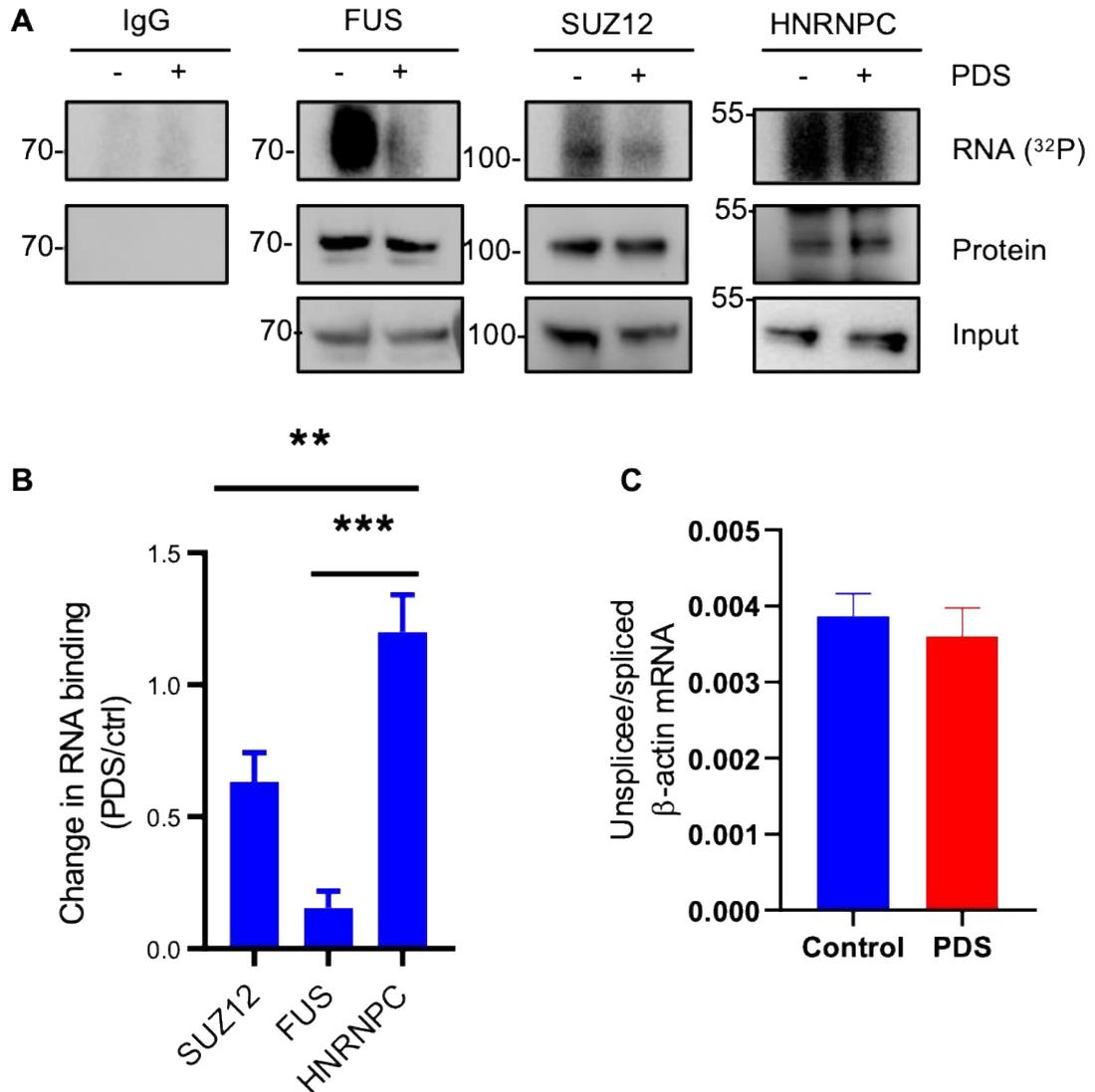


Figure 4.9 PDS treatment reduces binding of PRC2 to RNA in mESC.

A Top: ³²P-labeled RNA crosslinked to SUZ12, FUS and HNRNPC in E14 cells with or without treatment with 10 μ M PDS for 4 hrs. Immunoblots of SUZ12, FUS and HNRNPC in immunoprecipitated protein (centre) and whole cell extract (bottom). Representative of 3 independent experiments.

B Change of RNA binding by PDS treatment of SUZ12, FUS and HNRNPC. Mean and s.d. of 3 independent experiments, ** $p < 0.005$, *** $p < 0.0005$.

C Ratio of unspliced versus spliced β -actin RNAs measured by qRT-PCR under PDS treatment (mean and s.d. of triplicate wells).

Given that a reduction in RNA binding to PRC2 was observed in cells treated with PDS, I then tested the effect of PDS on PRC2 association with chromatin, to determine if the reduction of PRC2 binding to RNA caused

enrichment of PRC2 in the chromatin. Immunoblotting for SUZ12 and EZH2 in cytoplasm, nucleoplasm and chromatin fractions of E14 cells after 4 hours of PDS treatment revealed a decrease in PRC2 levels in the nucleoplasm. Interestingly, enrichment of neither SUZ12 nor EZH2 was noticeable in the chromatin fraction under PDS treatment (Fig 4.10 A), potentially indicating loss of RNA binding but no corresponding increase in the chromatin fraction. Nonetheless, to confirm that there was no enrichment of PRC2 in chromatin with PDS treatment, I performed ChIP-qPCR for SUZ12 and H3K27me3. No significant changes in PRC2 occupancy was observed at any of the genes tested in cells treated with PDS (Fig 4.10 B). This suggested that although PRC2 was removed from the nucleoplasm by PDS due to loss of interaction with RNA, PRC2 was not recruited to chromatin.

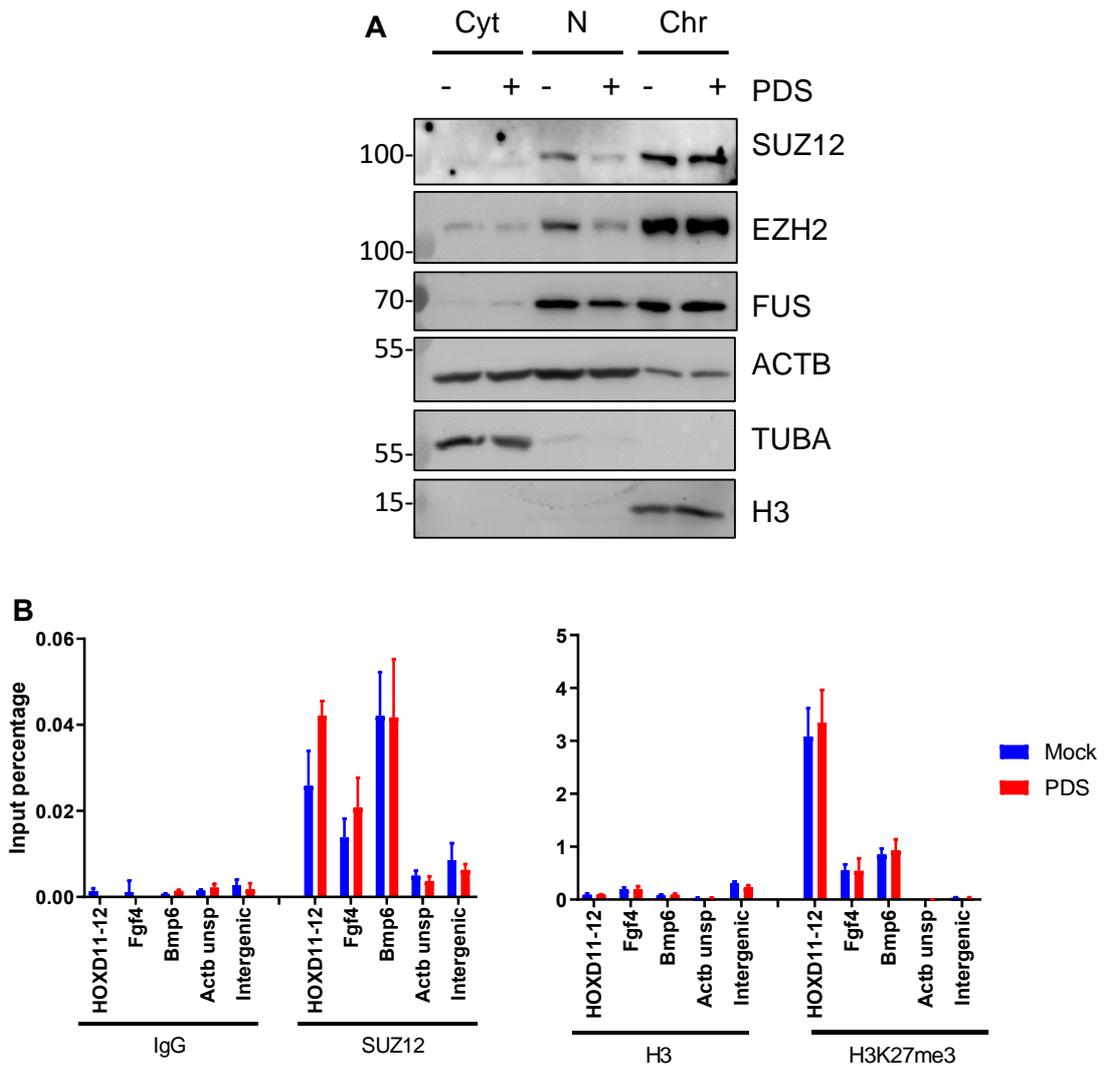


Figure 4.10. PDS affects PRC2 localization in mESCs, but not its binding to target genes.

A Immunoblots for SUZ12, EZH2 and FUS, with ACTB, TUBA and H3 acting as fractionation and loading controls, in cytoplasm (Cyt), nucleoplasm (N) and chromatin (Chr) fractions of E14 cells with and without treatment with PDS.

B ChIP-qPCR measuring SUZ12, H3K27me3 and H3 occupancy at the PRC2 target genes *Hoxd11*, *Fgf4* and *Bmp6* in E14 cells with and without treatment with PDS. *Actb* TSS and an intergenic locus 2 kb upstream of β -actin gene was used as non-PRC2 target negative controls (mean and s.d. of triplicate wells, representative of 2 independent experiments).

I next sought to establish the range of concentrations with which PDS exerted competition with PRC2 for RNA G-quadruplexes. To address this, mESCs were treated with a five-fold increasing concentration of PDS from 2 μ M to 50 μ M (Fig. 4.11). To my surprise, this assay did not recapitulate the previous

results, even at high concentrations. This is potentially due to differences in the batches of PDS between the two experiments.

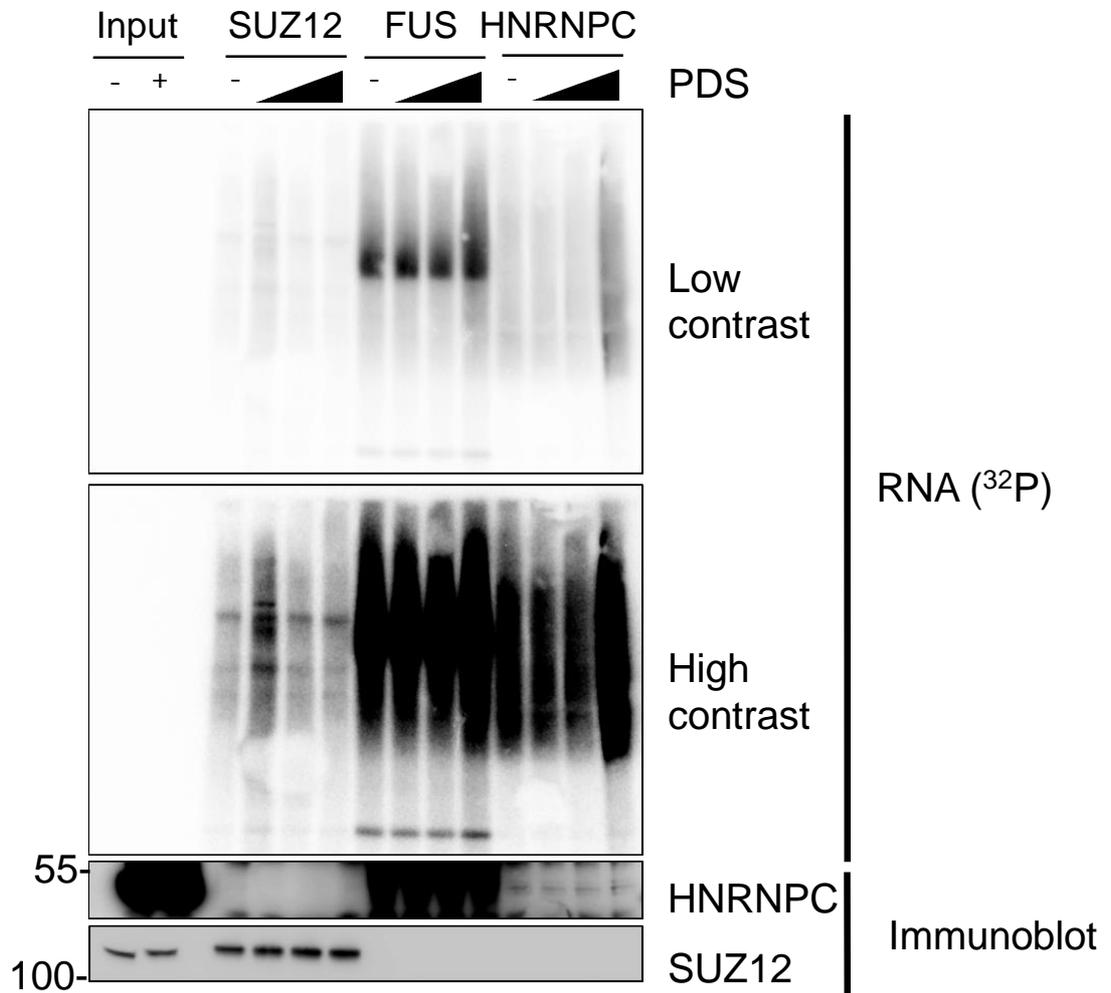


Figure 4.11 CLIP for SUZ12, FUS and HNRNPC in mESCs treated with a different batch of PDS does not recapitulate inhibition of PRC2 RNA binding.

Top: ³²P-labeled RNA crosslinked to input protein and immunoprecipitated SUZ12, FUS and HNRNPC in mock-treated E14 mESCs or cells treated with 2-50 mM PDS for 4 hrs (low and high contrast autoradiograms). Below: Corresponding immunoblots for SUZ12 and HNRNPC in immunoprecipitates and input samples.

4.7. TMPyP4 and MM41 show pleiotropic effects and increased PRC2 RNA binding in mESCs, respectively

A number of other compounds bind with high affinity to G4 structures and alter their stability. One such compound is the cationic porphyrin TMPyP4, which inhibits the formation of the highly stable G4 structures found in the 5'UTR region of matrix metallo-proteinase MT3 mRNA (Morris *et al.*, 2012). Another example is MM41, a tetra-substituted naphthalene-di-imide derivative that stabilizes G4 DNA found at several gene promoters (Ohnmacht *et al.*, 2015). Considering that I had conflicting results with E14 cells treated with PDS, I sought to establish whether any of these compounds affected PRC2 binding to RNA in cells, which would further corroborate the physiological relevance of G4 RNA structures in antagonism between chromatin and RNA for PRC2 binding.

To address whether TMPyP4 affected PRC2 RNA binding, I treated E14 cells with TMPyP4 for 4 h. However, I noticed that both SUZ12 and FUS proteins were degraded as TMPyP4 concentration increased, whereas HNRNPC expression increased (Fig 4.12). This suggested that TMPyP4 has pleiotropic effects in mESCs, negating its use to test PRC2 G4 RNA binding in cells.

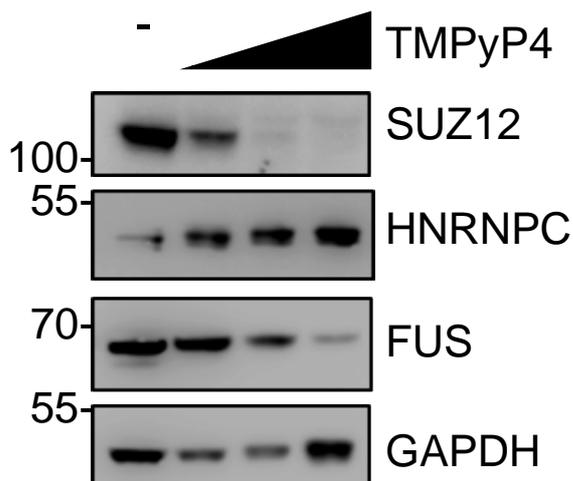


Figure 4.12 TMPyP4 induces degradation of SUZ12 in mESCs.

Immunoblot of SUZ12, HNRNPC, FUS and GAPDH in whole cell lysates of E14 cells treated with 0.8 μ M to 20 μ M of TMPyP4.

I next sought to establish whether MM41 had the same effect as either PDS or TMPyP4. To address this, E14 cells were treated with MM41 for 4 hrs with concentrations increasing 10-fold from 1 nM to 1 μ M (Fig. 4.13). CLIP for SUZ12 demonstrated a proportional increase in the binding to RNA as the concentration of MM41 increased. Of note, CLIP for HNRNPC exhibited the same degree of RNA binding regardless of the concentration of MM41. These data indicated that MM41 may be stabilizing G4 RNA in cells without competing with PRC2, leading to increased PRC2 RNA binding.

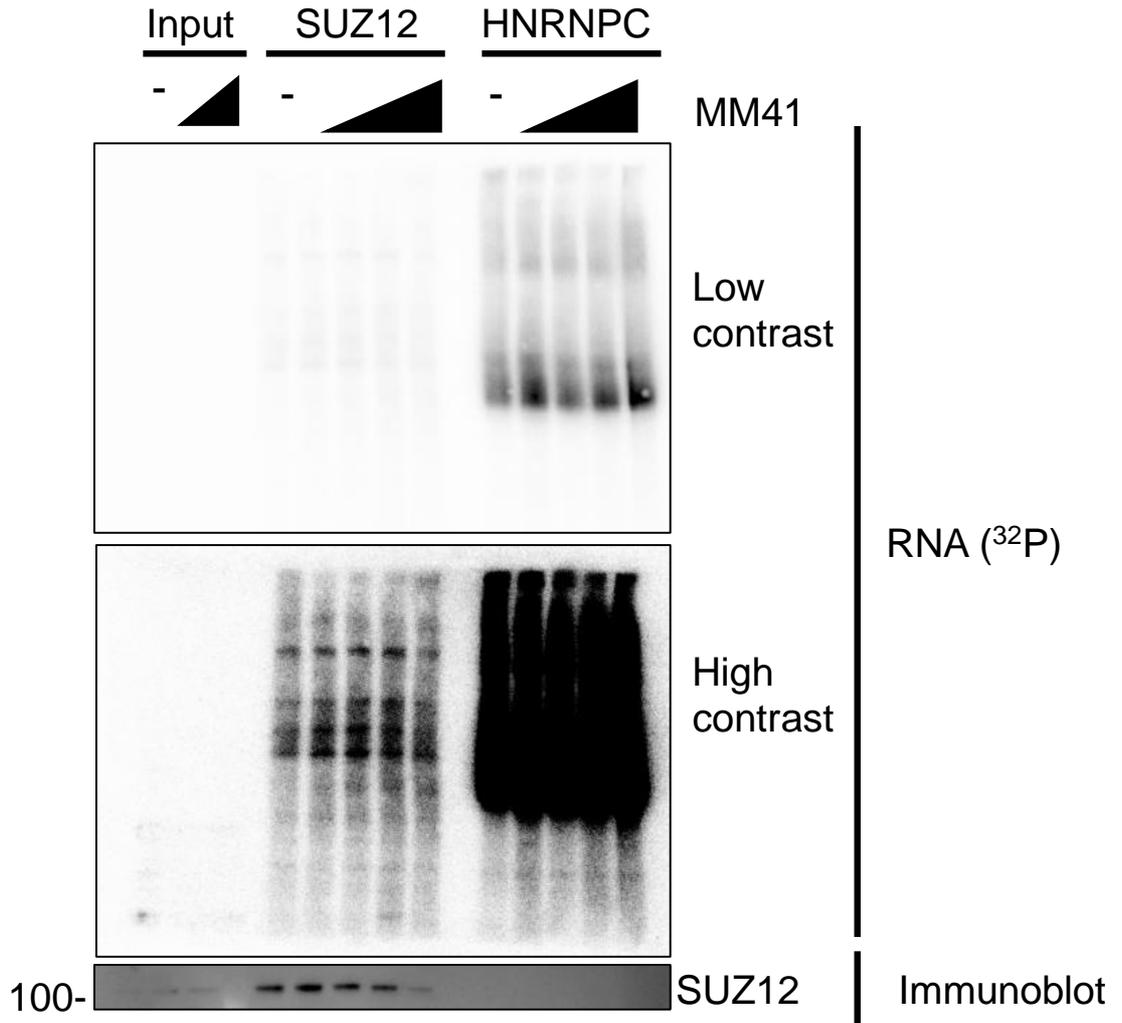


Figure 4.13 MM41 increases binding of PRC2 to RNA in mESCs.

Top: ³²P-labeled RNA crosslinked to input protein and immunoprecipitated SUZ12 and HNRNPC in mock-treated E14 mESCs or cells treated with 1 nM-1 μ M MM41 for 4 hrs (low and high contrast autoradiograms). Below: Corresponding immunoblot for SUZ12 in immunoprecipitates and input samples.

4.8. Summary and discussion

In this chapter I showed the role of PRC2 accessory subunits for the competition between RNA and chromatin for PRC2 binding in cells. I show that the PRC2 subunits EPOP, JARID2 and PALI1 are necessary for the increase in recruitment of PRC2 to chromatin when RNA is absent in mESCs. This was of particular interest given that even though both PRC2.2 subunits bind to chromatin, AEBP2 through its Zn finger to DNA (Kim, Kang and Kim, 2009) and JARID2 via H3K119ub (Landeira *et al.*, 2010; Cooper *et al.*, 2016), my findings show that only the latter is relevant in the context of recruitment to chromatin upon release from RNA. Furthermore, I show that this effect is not exclusive to cells lacking PRC2.2 accessory factors, as EPOP and PALI1-null exhibit the same phenotype. This suggests a common mechanism of RNA and chromatin antagonism that affects both PRC2.1/2 holo-complexes. Of note, the observed increase in PRC2 binding to chromatin in EPOP-null cells is consistent with previous results in which increased PRC2 chromatin occupancy was observed by ChIP when *Epop* was ablated (Beringer *et al.*, 2016; Liefke, Karwacki-Neisius and Shi, 2016). The loss of PRC2 binding upon RNA degradation is unclear, but could further reflect a role for EPOP in RNA Pol II regulation and transcription elongation (Liefke, Karwacki-Neisius and Shi, 2016). Further experiments are needed to determine why EPOP is required for the recruitment of PRC2 to chromatin when RNA is depleted.

More importantly, I showed that JAZF1-SUZ12 exhibits a similar phenotype to that observed in mESC lacking JARID2, EPOP and PALI1, probably due to the lack of interaction of JAZF1-SUZ12 with these accessory subunits due to loss of the N-terminus of SUZ12. This phenotype reveals a deficiency in the

function of JAZF1-SUZ12 that could contribute to the oncogenic effects of the fusion protein in ESS. (Fig. 4.14).

Additionally, I tested whether recruitment of PRC2 to unmodified nucleosomes upon RNA degradation was affected by the lack of accessory subunits. Using nucleosome pull-downs, I discovered that regardless of its composition, binding of PRC2 to nucleosomes increases when RNA was depleted. However, this assay has several limiting factors that could be addressed in the future. Chiefly, the recombinant nucleosomes used do not include modified histones, which might explain the lack of differences of nucleosome binding when *Jarid2* is ablated (Landeira *et al.*, 2010; Cooper *et al.*, 2016). Even with limitations, this assay showed that PCL proteins are not necessary for PRC2 to bind to nucleosomes in mESCs extract. Although only one PCL was deleted, PCL2 is the most abundant PCL protein in mESCs, comprising 80% of all PCLs (Adrian Bracken, personal communication) and thus explains why there was also no observable difference when comparing nucleosome pull-down of nuclear extracts from PCL2-null and PCL1/2/3-null cells. These results reveal that PCL proteins are not necessary for PRC2 binding to nucleosomes upon release from RNA, but might enhance steady-state binding through interaction with linker DNA.

To understand the interplay between PRC2 and G4 RNA structures in cells, I tested the effect of G4 RNA-binding small molecules on PRC2 RNA binding in cells. I showed that PDS reduces PRC2 binding to G4 RNA in cells, confirming previous observations carried out *in vitro* (Wang, Goodrich, *et al.*, 2017). However, the differential effect of one batch compared to another needs to be addressed.

Finally, I show that other compounds that affect G4 also have effects on PRC2. TMPyP4 treatment caused loss of SUZ12 from cells, which suggests a possible role for G4 RNA structural integrity in the stability of SUZ12 protein in mESCs or a more general effect on cellular viability. Also, I showed that MM41 increases RNA binding to PRC2 in mESCs. Contrary to what we were expecting, these findings suggest that MM41 stabilizes G4 structures without competing with PRC2 for RNA binding, potentially implying binding to different sites on the RNA.

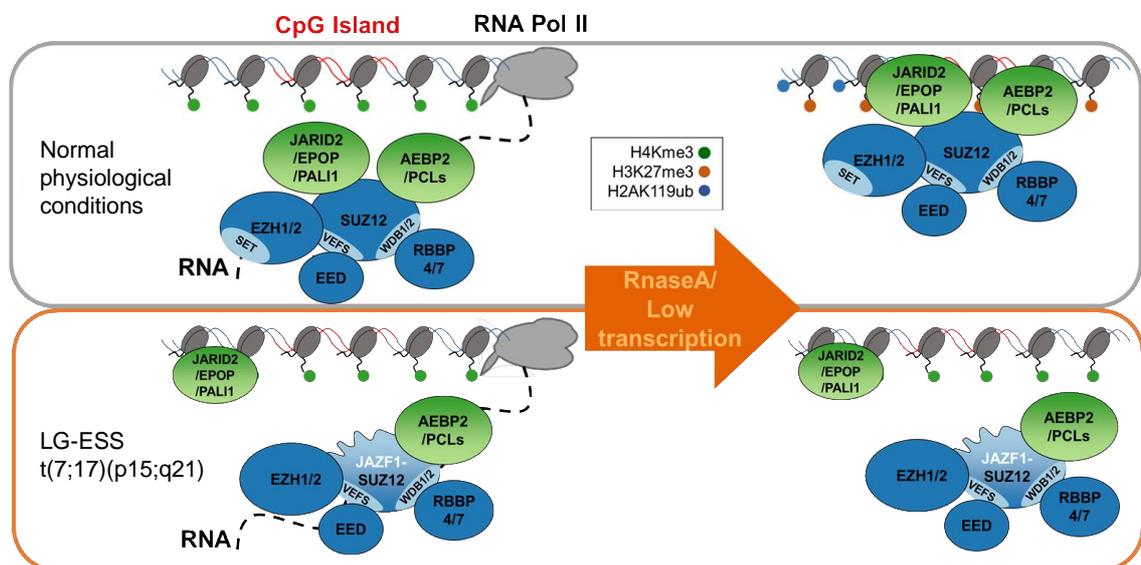


Figure 4.14 Model of the main findings of Chapter 4.

In normal physiological conditions, PRC2 can transition from binding RNA to binding chromatin when there is loss of transcription, which is mimicked by RNaseA treatment in permeabilized cells. With the t(7;17)(p15;q21) translocation, JAZF1-SUZ12 loses its interaction with JARID2, EPOP and PAL1, and this prevents recruitment of PRC2 to chromatin.

5. Chapter 5: Effects of JAZF1-SUZ12 on cell differentiation.

5.1. Introduction.

SUZ12 is necessary for the assembly of core PRC2 and for proper embryonic development. This was revealed by studies which reported that mice lacking SUZ12 were not viable and died during embryogenesis. Moreover, these embryos showed developmental defects as early as 7.5 days post coitus, significantly smaller size, and lack of H3K27me3 (Pasini *et al.*, 2004), which recapitulated the phenotype of mice deficient for *Ezh2* or *Eed* (Faust *et al.*, 1995; O'Carroll *et al.*, 2001). Later, it was revealed that loss of *Suz12* affected differentiation into embryoid bodies (EBs), which showed lack of repression of gastrulation markers, such as *Pax3* and *T (Brachyury)*. In the same study, CHIP-qPCR revealed that PRC2 was recruited to TSS of pluripotency markers during mESCs differentiation, and that ablation of *Suz12* reduced the efficiency of the resultant gene repression. This revealed a mechanism by which PRC2 allows differentiation in mESCs through its recruitment and repression of pluripotency genes. Reciprocally, ES cell differentiation is accompanied by the loss of PRC2 from differentiation-specific genes, thus allowing their activation (Pasini *et al.*, 2007). Afterwards, it was confirmed that knockout of PRC2 subunits SUZ12 and EZH2 in mESCs did not affect gene repression in steady state, but when these cells were differentiated into EB, these showed lack expression of differentiation genes at appropriate times (Rising *et al.*, 2014).

Human endometrium is a complex organ that requires cellular plasticity to be able to overcome menstrual cycle, embryo implantation and other potential environmental insults (Gellersen and Brosens, 2014). It has been previously reported that PRC2 regulates the expression of genes responsible of luteum body development, such as Insulin-like growth factor-binding protein 1 (*IGFBP1*) and anterior pituitary hormone gene prolactin (*PRL*), as reduction of H3K27me3 and increase of H3K4ac is observed at these genes when hEnSCs are decidualised *in vitro* (Grimaldi *et al.*, 2011). Whether this dynamic regulation of PRC2 recruitment during cell differentiation is affected by fusion of JAZF1 to SUZ12 has not been assessed but could underlie the oncogenic nature of the fusion protein in ESS.

It has been reported that ectopic expression of JAZF1-SUZ12 in HEK293 cells, coupled with knockdown of endogenous SUZ12, increases cellular proliferation and resistance to hypoxic conditions, possibly explaining the functions of the JAZF1-SUZ12 functions in the establishment of carcinogenic functions (Li *et al.*, 2007).

In this chapter I assess the consequences of JAZF1-SUZ12 expression in a tumorigenesis model. Also, I sought to determine whether JAZF1-SUZ12 affects differentiation in an EB formation model. Finally, I sought to establish the effects of JAZF1-SUZ12 on decidualisation of hEnSCs.

5.2. Overexpression of SUZ12, but not JAZF1-SUZ12, generates more colonies in immortalized fibroblasts.

JAZF1-SUZ12 ectopic expression in HEK293 cells, coupled with knockdown of endogenous SUZ12, increases cellular proliferation and resistance to hypoxic conditions, possibly explaining the carcinogenic effect of JAZF1-SUZ12 (Li *et al.*, 2007). However, tumorigenic potential was not assessed in this study, thus limiting its scope and significance. To address whether JAZF1-SUZ12 is carcinogenic, I sought to model the onset of ESN in which WT SUZ12 is still expressed from the other, non-mutated, allele (Li *et al.*, 2007). To do this, I transduced three immortalized fibroblast cell lines with C-terminally FLAG-tagged SUZ12, JAZF1-SUZ12 or JAZF1 constructs. The immortalised cell lines contain either hTERT alone, hTERT and SV40 large T antigen, or hTERT, SV40 large T and oncogenic Ras (66+++), and thus have increasing oncogenic potential (Scaffidi and Misteli, 2011).

To test whether the stable lines expressed the transgenes at equivalent levels, I performed immunoblotting for FLAG-tagged SUZ12, JAZF1-SUZ12 and JAZF1 (Fig 5.1). Interestingly, although FLAG-SUZ12 was expressed at similar levels in all three lines, FLAG-JAZF1-SUZ12 showed progressively lower expression in the more transformed cell lines, potentially showing a deleterious effect of this protein in the context of cell transformation.

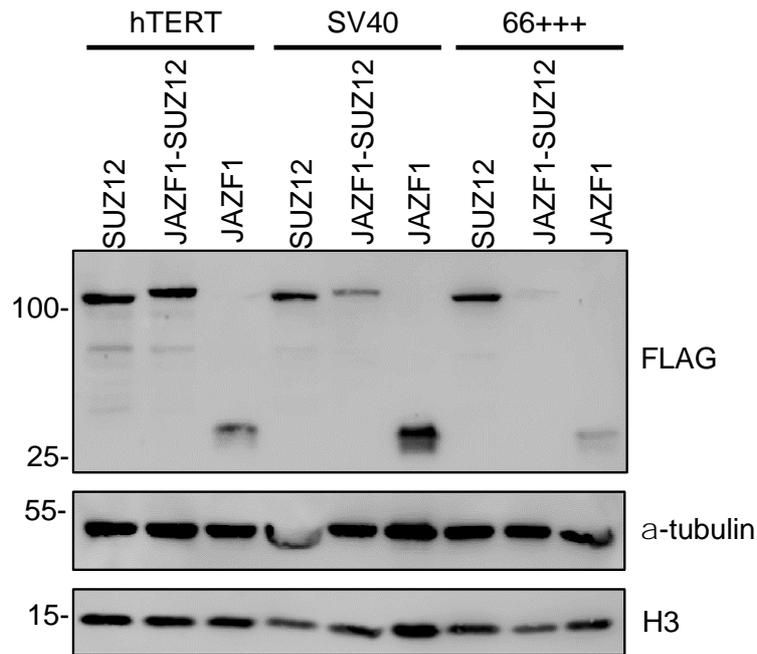


Figure 5.1. Generation of immortalized fibroblast cell lines expressing FLAG-tagged SUZ12 constructs.

Immunoblots for FLAG, α -tubulin and H3 in whole cell lysates of hTERT, SV40 and 66+++ cell lines, stably expressing SUZ12, JAZF1-SUZ12 or JAZF1.

I next sought to establish whether JAZF1-SUZ12 promoted cell transformation, measured by the number of colonies formed in soft agar assays (Figure 5.2). Strikingly, this assay showed that SUZ12-expressing hTERT cells formed colonies, while cells expressing JAZF1-SUZ12 or JAZF1 did not. Additionally, SUZ12-expressing SV40 cells exhibited increased colony formation compared to cells expressing JAZF1-SUZ12 or JAZF1. Finally, in 66+++ cells, all of the proteins tested showed no differences in colony formation when compared with the control, showing that none of the transgenes were tumorigenic in the 66+++ cells. These results revealed that this model of immortalized fibroblasts does not recapitulate the presumed oncogenic effects of JAZF1-SUZ12.

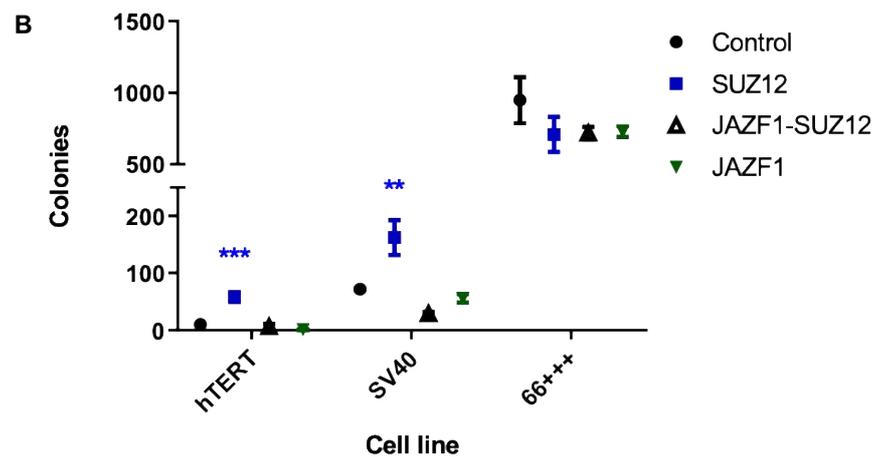
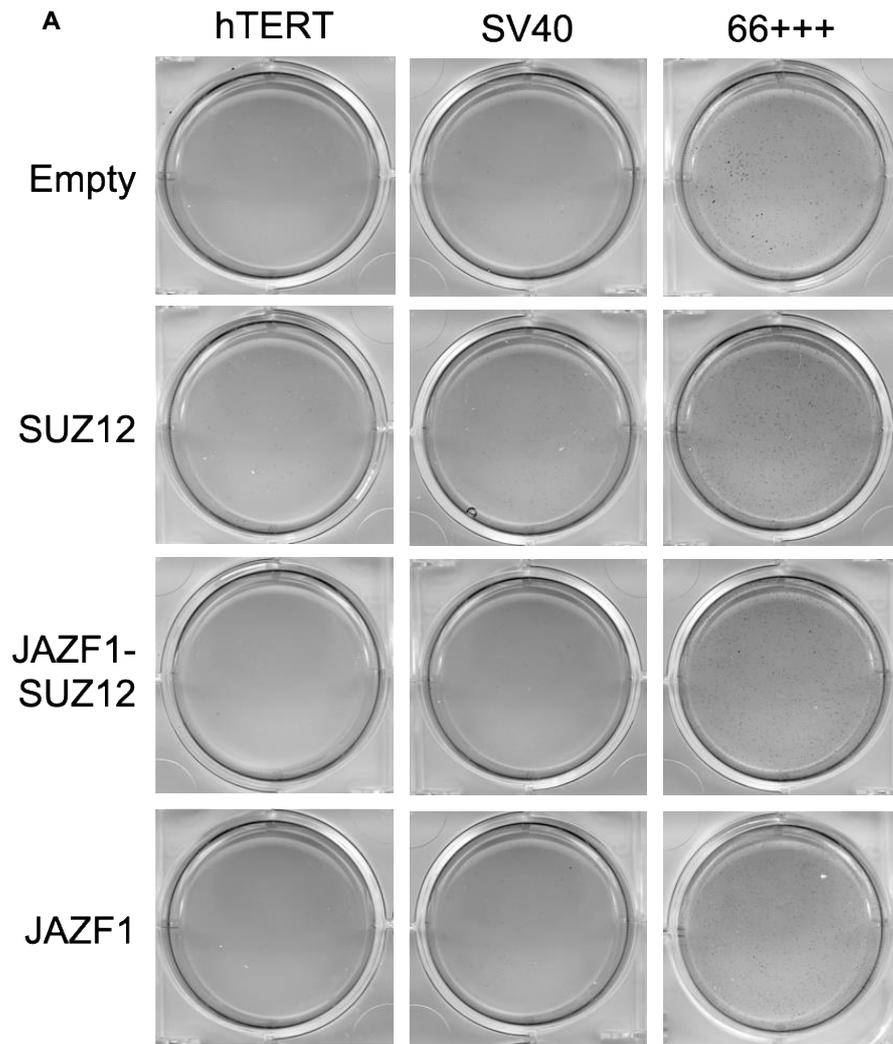


Figure 5.2. SUZ12 but not JAZF1-SUZ12 promotes colony formation in hTERT and SV40 fibroblasts.

A Representative wells of colonies formed by hTERT, SV40 and 66+++ cells either transduced with control vector, SUZ12, JAZF1-SUZ12 or JAZF1.

B Next page: number of colonies formed by hTERT, SV40 and 66+++ cells transduced with control vector, SUZ12, JAZF1-SUZ12 or JAZF1 (mean and s.d. of 3 wells, ** $p < 0.005$, *** $p < 0.0005$).

5.3. JAZF1-SUZ12 promotes differentiation towards endodermal lineages during differentiation of ES cells into EBs.

SUZ12 is necessary to ensure developmental genes maintain repressed until the appropriate point during cell differentiation (Pasini *et al.*, 2007). Considering that JAZF1-SUZ12 is not recruited to PRC2 target genes upon release from RNA due to the loss of interaction with EPOP, PALI1/2 and JARID2, I tested whether this disrupted differentiation of ESC into EBs.

To address this, I differentiated *Suz12^{GT/GT}* cells stably expressing GFP, SUZ12, SUZ12 Δ 93 or JAZF1-SUZ12, and observed the resulting EBs after 2, 4 and 8 days (Fig. 5.3). EBs formed by cells expressing SUZ12 expanded in size and exhibited differentiation features (dense cellular mass) from day 2, as expected, whereas EBs formed by GFP, SUZ12 Δ 93 and JAZF1-SUZ12-expressing cells did not. This effect was more evident at days 4 and 8. These results reveal that SUZ12 is necessary for EB formation and that JAZF1-SUZ12 and SUZ12 Δ 93 are not able to recover this function, suggesting that the N-terminus of SUZ12 plays an important role in early embryonic development.

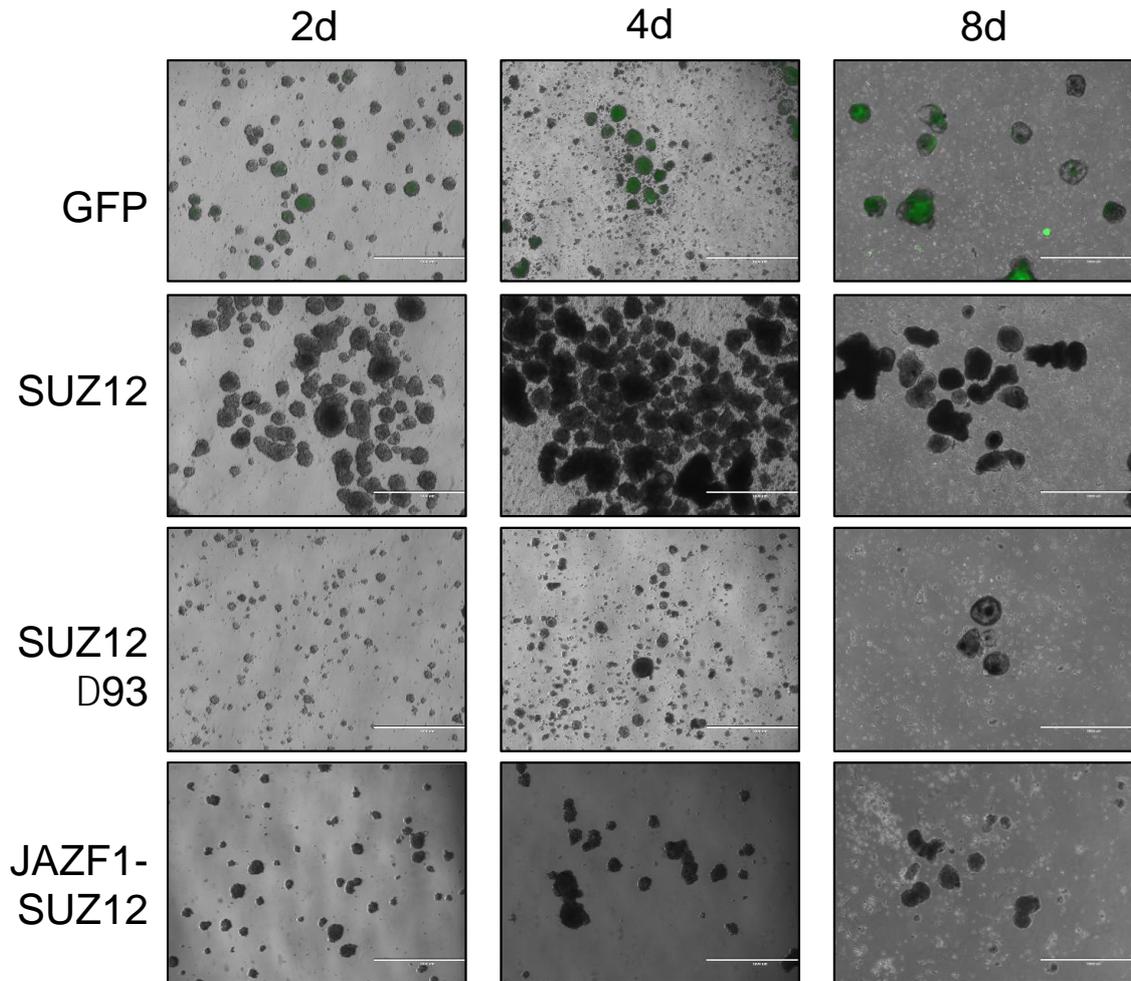


Figure 5.3. Defective formation of EBs in ESC expressing SUZ12 Δ 93 or JAZF1-SUZ12.

Representative images of *Suz12*^{GT/GT} cells expressing GFP, SUZ12, SUZ12 Δ 93 and JAZF1-SUZ12 taken at 2, 4 and 8 days after initiation of EB formation. Images are overlaid with an 488nm channel image detecting GFP. The white reference bar marks 1 mm.

Considering the differences in EB formation, I asked whether this phenotype was associated with failure of SUZ12 Δ 93 and JAZF1-SUZ12 to repress pluripotency genes. Thus, I measured expression of the pluripotency genes *Oct4*, *Fgf4*, *Nanog* and *Utf1* in *Suz12*^{GT/GT} cells expressing GFP, SUZ12, SUZ12 Δ 93 or JAZF1 at 0, 4 and 8 days of EB formation (Figure 5.4). This showed that pluripotency genes were repressed in all cell lines during EB formation

regardless of whether or not the cells expressed WT SUZ12. Although a trend in which cells expressing WT SUZ12 show more efficient repression of pluripotency genes than cells expressing mutant forms of the protein, this was not statistically significant. This lack of difference between different cell lines might reflect the presence of low levels of endogenous SUZ12 in *Suz12*^{GT/GT} (Fig 3.1; (Thornton *et al.*, 2014), which might be sufficient to repress pluripotency genes in this assay.

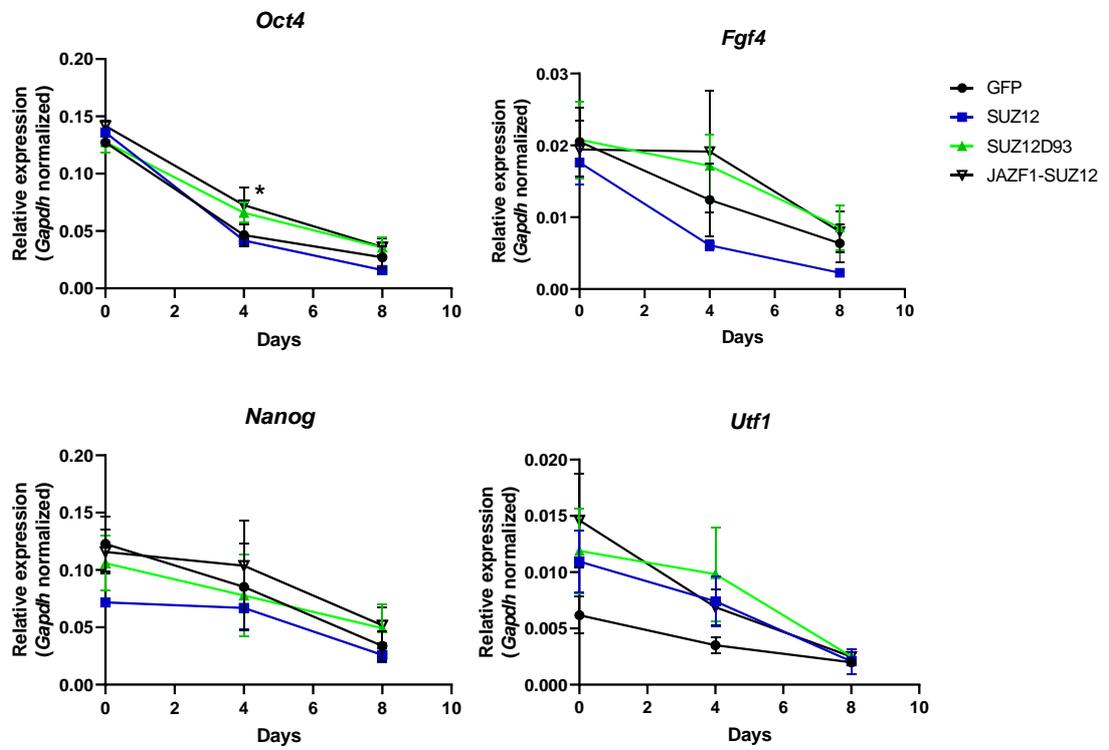


Figure 5.4. JAZF1-SUZ12 and SUZ12 Δ 93 have limited effect on the repression of pluripotency markers during differentiation of *Suz12*^{GT/GT} ESC into EBs.

Expression of *Oct4*, *Fgf4*, *Nanog* and *Utf1* measured by RT-qPCR, normalized to *Gapdh*, at 0, 4 and 8 days of EB formation in *Suz12*^{GT/GT} cells expressing GFP, SUZ12, SUZ12 Δ 93 or JAZF1-SUZ12 (mean and s.d. of 4 independent experiments; *p<0.05, determined by 2-way ANOVA).

Given that no significant difference was observed in the repression of pluripotency genes during differentiation of cells expressing JAZF1-SUZ12 or

SUZ12 Δ 93 compared with SUZ12, I sought to establish whether the defects in EB morphology (Fig 5.3) instead reflected dysregulation of differentiation marker expression. To test this, I carried out RT-qPCR to measure the expression of the mesodermal and early gastrulation markers *Pax3* and *T* (Fig. 5.5). Interestingly, I discovered that although these genes were upregulated cells expressing WT SUZ12 by the eighth day of EB formation, as reported before (Pasini *et al.*, 2007), this did not occur in cells expressing JAZF1-SUZ12 or SUZ12 Δ 93. This suggested that expression of genes that promote lineage commitment towards mesoderm and early gastrulation requires the N-terminus of SUZ12. This is likely due to the requirement for this region for interaction with EPOP, JARID2 or PALI1/2 and recruitment from RNA to chromatin.

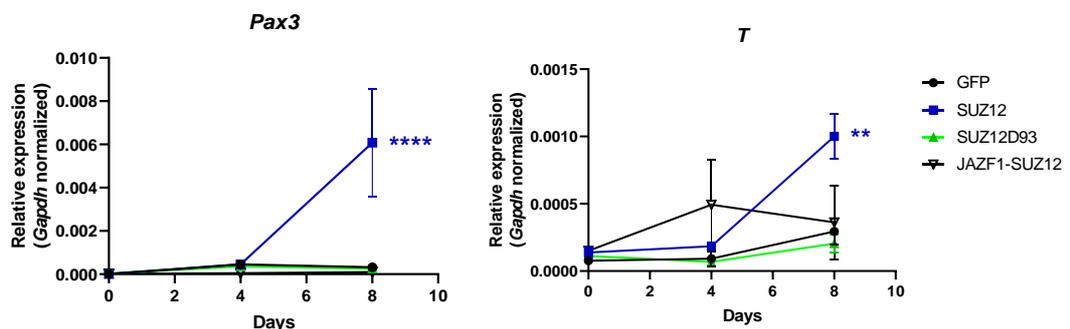


Figure 5.5. Mesoderm genes need the N-terminus of SUZ12 during EB formation to be expressed.

Expression of *Pax3* and *T* measured by RT-qPCR, normalized to *Gapdh*, at 0, 4 and 8 days of EB formation in *Suz12*^{GT/GT} cells expressing GFP, SUZ12, SUZ12 Δ 93 or JAZF1-SUZ12 (mean and s.d. of 4 independent experiments; **p<0.005 and ****p<0.00005, determined by 2-way ANOVA).

Having showed that the N-terminus of SUZ12 is necessary for expression of mesoderm markers during EB formation, I next sought to establish the effects of the expression of JAZF1-SUZ12 in the expression of ectodermal marker *Fgf5*

and the endodermal marker *Gata4*, to test if commitment towards other germ layers was affected (Figure 5.6). Interestingly, *Fgf5* mRNA was upregulated at day 4 of EB formation in cells expressing SUZ12 or SUZ12 Δ 93 but this increase was abrogated in cells expressing JAZF1-SUZ1. In contrast, upregulated of *Gata4* was greater in cells expressing JAZF1-SUZ12 than the other cell lines. These results suggest that JAZF1-SUZ12 promotes differentiation towards the endodermal lineage. That this effect is not observed for SUZ12 Δ 93 suggests that this is not due to the loss of interaction with EPOP, PALI1 and JARID2 but could instead be due to a gain of function caused by fusion of JAZF1 to SUZ12.

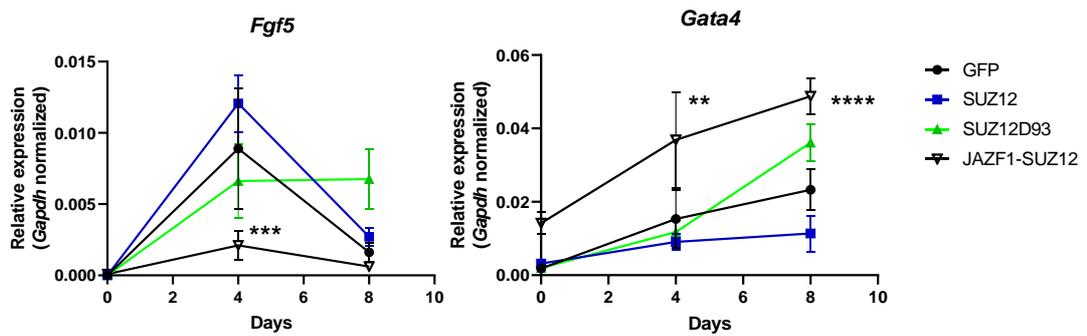


Figure 5.6. JAZF1-SUZ12 promotes differentiation of mESC towards an endodermal lineage.

Expression of *Fgf5* and *Gata4* measured by RT-qPCR, normalized to *Gapdh*, at 0, 4 and 8 days of EB formation in *Suz12*^{GT/GT} cells expressing GFP, SUZ12, SUZ12 Δ 93 or JAZF1-SUZ12 (mean and s.d. of 4 independent experiments; **p<0.005, ***p<0.0005 and ****p<0.00005, determined by 2-way ANOVA).

5.4. *In vitro* decidualization of transfected hEnSCs.

Primary human endometrial cells (hEnSCs) can be isolated by treating biopsies with DNase I to remove the extracellular matrix (Barros et al. 2016). Once hEnSCs are recovered, the cells can be decidualised *in vitro* by cultivation in media containing MPA and 8-Br-cAMP for up to 10 days (Gellersen and Brosens, 2014). Considering that my results in ESC suggested that JAZF1-SUZ12 might only impact gene expression during cell differentiation, and that hEnSCs could model more closely what could be happening in LG-ESS, I sought to establish whether I was able to transfect, select and decidualise hEnSCs. To test this, I transfected hEnSCs with FS2-GFP, selected with puromycin and observed whether GFP was still expressed after cells were decidualized for 8 days (Figure 5.7). This showed that the FS2-GFP was present at all timepoints of decidualisation, and more importantly, that the hEnSCs transitioned from a fibroblastoid to an epithelioid phenotype, characteristic of decidualisation.

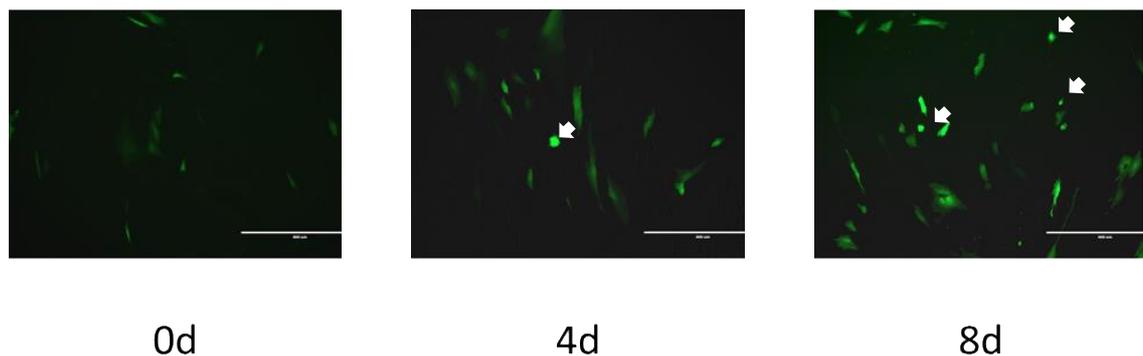


Figure 5.7. hEnSCs are transfectable and undergo decidualization after selection with puromycin.

Representative image of FS2-GFP-transfected and selected hEnSCs after decidualization in cAMP and MPA for 0, 4 and 8 days. White arrows indicate cells with an epithelioid phenotype. The white bar marks 1 μm.

5.5. Transfected of hEnSCs with SUZ12 constructs

Once I had established that hEnSCs could be transfected with FS2-GFP, selected and subsequently decidualised, I sought to test whether JAZF1-SUZ12 impacted differentiation in this model by transfecting hEnSCs with FS2-tagged GFP, SUZ12, SUZ12 Δ 93, JAZF1-SUZ12 and JAZF1 and selecting stable lines.

Interestingly, after transfection of SUZ12 Δ 93 all cells died with selection media. Regardless of this, I first confirmed that the transgenes were expressed by RT-qPCR, using primers specific for the different constructs (Fig. 5.8A). This showed that SUZ12, JAZF1-SUZ12 and JAZF1 were expressed at all timepoints of decidualisation. Afterwards, I carried out RT-qPCR to measure expression of the decidualisation markers *IGFBP1* and *PRL*, to establish if JAZF1-SUZ12 impaired hEnSCs decidualisation (Fig. 5.8B). This showed no significant difference between cells expressing SUZ12 and JAZF1-SUZ12, suggesting that expression of these genes is not affected by the presence of JAZF1-SUZ12.

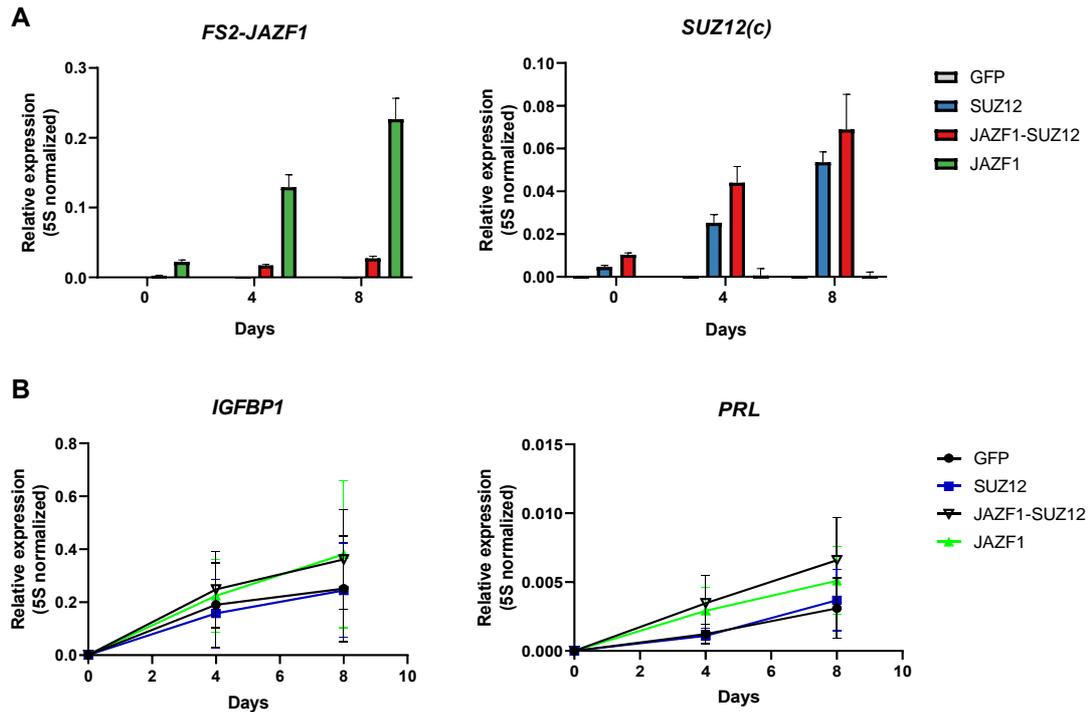


Figure 5.8. SUZ12 transgenes do not affect decidualisation in hEnSCs.

A. Left: expression of transcripts of the fusion of the FS2 tag with JAZF1 coding sequence in transfected and selected hEnSCs expressing GFP, SUZ12, JAZF1-SUZ12 or JAZF1, normalized to 5S rRNA. Right: as left, but quantifying a SUZ12 transgene-specific sequence. Mean and s.d. of triplicate wells, representative of 3 independent experiments.

B. As A, but expression of *IGFBP1* (left) and *PRL* (right) mRNAs, normalized with 5S rRNA. Mean and s.d. of 3 independent replicates.

Considering that no significant difference was observed between cells expressing the transgenes, I verified by immunoblot whether the protein products of the transgenes could be detected in hEnSCs after transient transfection (Fig. 5.9A). This showed that only low levels of SUZ12 and JAZF1-SUZ12 could be detected, when compared to GFP and JAZF1. Moreover, the production of FS2-GFP from the pCAG vector was low compared to GFP produced from the control vector pMAX (Fig. 5.9.B). These results suggest that FS2-tagged transgenes are expressed, but the proteins are only synthesised at low levels in hEnSCs.

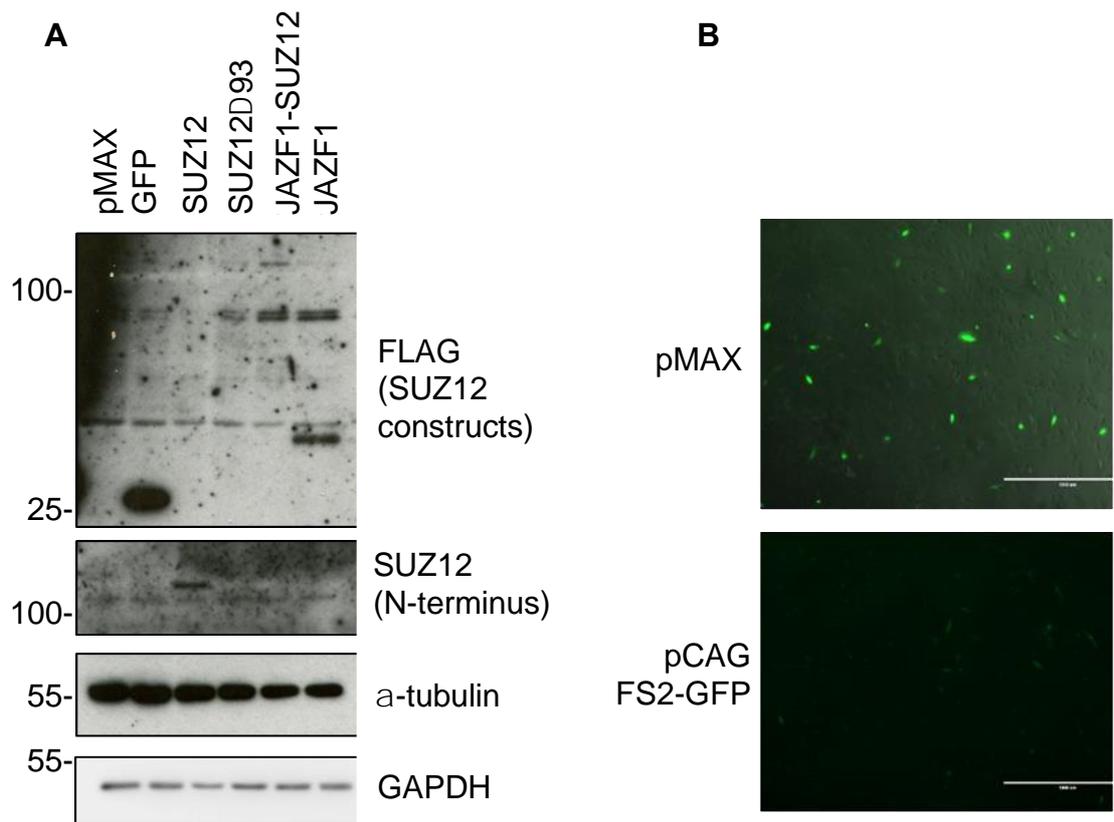


Figure 5.9. FS2-tagged proteins are only expressed at low levels following transient transfection of hEnSCs.

A. Immunoblots for FLAG, SUZ12 (N-terminus), α -tubulin and GAPDH in whole cell lysates of hEnSCs transiently transfected with pMAX-GFP and FS2-tagged GFP, SUZ12, SUZ12 Δ 93, JAZF1-SUZ12 and JAZF1.

B. Images of hEnSCs transiently transfected with pMAX-GFP (top) and pCAG FS2-GFP (bottom), two days after transfection.

5.8. Summary and discussion

In this chapter I tested the effect of JAZF1-SUZ12 on cell transformation and differentiation state. I found that SUZ12 but not JAZF1-SUZ12 can induce colony formation in immortalized fibroblasts. Additionally, I demonstrated that JAZF1-SUZ12 promotes an ectodermal phenotype by altering the expression of differentiation markers during differentiation of mESC to EBs, and that expression of SUZ12 Δ 93 is not able to rescue the differentiation defects of *Suz12*^{GT/GT} cells. Finally, I show that hEnSCs are transfectable and can be selected and decidualised.

I used primary transformed human fibroblasts to model ESN tumorigenesis but found that JAZF1-SUZ12 had no effect on the cells' tumorigenic potential. However, SUZ12 stimulated colony formation by hTERT cells and increased colony number of SV40 cells. This is surprising, but might be explained by the fact that it has been previously shown that SUZ12 is necessary for proliferation in certain cell types, such as U2OS cells, in which knock down of SUZ12 reduced colony formation in agar assays (Bracken *et al.*, 2007). Lack of a tumorigenic effect of JAZF1-SUZ12 in all three cell lines might reflect the presence of endogenous SUZ12, loss of which might be needed to observe a malignant phenotype, as previously reported (Nucci *et al.*, 2007).

Also, using EB formation to model the role of PRC2 in regulating gene expression during cell differentiation, I showed that the N-terminus of SUZ12 is necessary for EB formation and for the expression of *Pax3* and *T*, both genes markers of early gastrulation. Moreover, in this same model, I showed that

JAZF1-SUZ12 presence has pleiotropic effects on the expression of differentiation genes, as it inhibits the proper expression of *Fgf5*, and promotes over expression of *Gata4*. My findings suggests that, as important as the N-terminus of SUZ12 is for early development, the fusion of JAZF1 to SUZ12 also exerts its own effects separate from SUZ12 Δ 93, potentially exhibiting an additional mechanism of dysregulation of gene expression to that mentioned in Chapters 3 (Fig. 3.9) and 4 (Fig. 4.14) of this work. This may reflect interaction of JAZF1-SUZ12 with the NuA4 complex (Piunti *et al.*, 2019), which would cause aberrant gene expression, given that both chromatin-modifying complexes have opposing functions, as NuA4 drives openness of chromatin by histone acetylation, while PRC2 repress gene expression through H3K27me3. How these findings relate to the oncogenic effects of JAZF1-SUZ12 in LG-ESS still needs to be addressed.

Finally, I attempted to use decidualisation of hEnSCs as a model to study the effect of JAZF1-SUZ12 on PRC2-mediated gene repression during cell differentiation. I succeeded in transfecting, selecting and decidualizing hEnSCs, testing this with FS2-GFP constructs. However, only very low expression of SUZ12 and JAZF1-SUZ12 was achieved, thus suggesting that the lack of an effect of JAZF1-SUZ12 on *IGFBP1* and *PRL* expression is due to insufficient protein. Others have used lentiviral transduction of hEnSCs to express JAZF1-SUZ12 with very limited success, as not much product of the transgenes were observed (Piunti *et al.*, 2019). This is something that could be optimised for further study of the effect of the fusion protein on hEnSCs differentiation.

6. Chapter 6. Conclusions and future work

In this work I have discovered specific PRC2 functions that are affected by the fusion of JAZF1 to SUZ12, a mutation that is present in LG-ESS, and I showed potential mechanisms that might explain how this fusion of genes generates cell transformation. I showed that JAZF1-SUZ12 reduced PRC2 binding to its target genes in mESC. Also, I discovered that interaction of SUZ12 with EPOP, PALI1 and JARID2 is lost due to the deletion of its N-terminus. Also, I showed that EPOP, but not full length JARID2, is expressed in hEnSCs.

Additionally, I showed that EPOP, JARID2 and PALI1, are necessary for recruitment of PRC2 to chromatin when RNA is depleted, and that this phenotype is recapitulated by JAZF1-SUZ12. Also, I discovered that binding of PRC2 to nucleosomes increased after RNA depletion, that this occurs even in the absence of AEBP2, JARID2 or PCL proteins and occurs regardless of the length of linker DNA. Also, I showed that G4 RNA-binding compounds can either reduce or increase the binding of PRC2 to RNA in cells, and also affect the SUZ12 protein stability.

Finally, my findings demonstrate that JAZF1-SUZ12 is not a driver mutation in the context of immortalized fibroblasts. Also, using differentiation of mESC to EB as a model, I showed that the N-terminus of SUZ12 is necessary for changes in gene expression upon cell differentiation and that EB formation is further dysregulated by JAZF1-SUZ12. These data are consistent with a model in which JAZF1-SUZ12 disrupts endometrial stromal cell differentiation and that this can promote cell transformation.

6.1. H3K27me3 is recovered by JAZF1-SUZ12 in *Suz12*^{GT/GT} cells.

Previous reports suggested that JAZF1-SUZ12 inhibits the assembly of core PRC2, and reduces HMTase activity of PRC2 *in vitro* (Ma *et al.*, 2017). Thus, one of the major questions to address in this work was whether PRC2 catalytic activity was impacted by JAZF1-SUZ12 in cells. I have demonstrated that global levels of H3K27me3 can be recovered by SUZ12 Δ 93 and JAZF1-SUZ12, which is consistent with the observed restoration in H3K27me3 levels upon expression of the SUZ12 VEFS domain in SUZ12 knockout cells (Højfeldt *et al.*, 2018). However, I also found that JAZF1-SUZ12 and SUZ12 Δ 93 showed reduced recruitment to their target genes. Thus, it will be informative to establish the location of H3K27me3 deposited by JAZF1-SUZ12-containing PRC2. To address this, CHIP-seq for H3K27me3 and the FLAG-tagged SUZ12 constructs could be performed in *Suz12*^{GT/GT} mESC to establish to where JAZF1-SUZ12 and SUZ12 Δ 93 containing PRC2 is recruited.

6.2. Interactions of PRC2 with a subset of accessory factors is dependent on the SUZ12 N-terminus.

The N-terminus of SUZ12 is important for binding to accessory subunits (Chen *et al.*, 2018b; Youmans *et al.*, 2018). Therefore, loss of the first 93 residues of SUZ12 suggested that fusion with JAZF1 may disrupt some of the interactions with PRC2 accessory factors. I discovered that essential subunits were lost from both PRC2 holocomplexes, PRC2.1 and PRC2.2, in the presence of JAZF1-

SUZ12. These factors included JARID2 and EPOP, confirming results published while the work reported here was ongoing (Chen *et al.*, 2018b). Additionally, I discovered that a recently discovered PRC2 subunit, PALI1 (Conway *et al.*, 2018), interacts with the N-terminus of SUZ12 and that this interaction is lost upon fusion of JAZF1 to SUZ12.

Considering that it has been reported that EPOP, JARID2 and PALI1 all confer increased PRC2 catalytic activity (Zhang *et al.*, 2011; Conway *et al.*, 2018; Kasinath *et al.*, 2018), it would be necessary to establish in more depth whether the catalytic activity of JAZF1-SUZ12-containing PRC2 is affected by the loss of PALI1, JARID2 and EPOP. The effect of JAZF1-SUZ12 on the HMTase activity of PRC2 has been addressed by other labs (Ma *et al.*, 2017), however this study only tested the effect of JAZF1-SUZ12 on minimal complex comprised of SUZ12, EZH2 and EED.

Finally, it will be interesting to test whether JAZF1-SUZ12 affects the interaction of PRC2 with the newly discovered PRC2 subunit, EZHIP (Jain *et al.*, 2019; Piunti *et al.*, 2019; Ragazzini *et al.*, 2019). This is of particular interest given that this protein is also the subject of translocation events in ESS (Piunti *et al.*, 2019), is necessary in female ovarian development and is predominantly expressed in the gonads (Ragazzini *et al.*, 2019).

6.3. hEnSCs contain EPOP and JARID2 Δ N.

I showed by immunoblotting that EPOP and JARID2 Δ N, but not full-length JARDI2, are present in hEnSCs. JARID2 Δ N cannot interact with SUZ12 due to the lack of its N-terminus (Al-Raawi *et al.*, 2019). Thus, this indicates that the effects of JAZF1-SUZ12 on PRC2 function are focused on its failure to interact with PRC2.1 subunits. Consistent with this, most translocations in ESN and LG-ESS reported to date involve either SUZ12 or PCL1 and never JARDI2 or AEBP2 (Table 1.3.). To gain a more complete picture, it will be necessary to establish first whether PALI1 or PALI2 are also expressed in hEnSCs, and secondly, to identify the genes regulated by PRC2.1 in hEnSCs and how these are dysregulated by JAZF1-SUZ12 and other PRC2 fusion proteins in LG-ESS.

6.4. EPOP, JARID2 and PALI1 are necessary for the recruitment of PRC2 to chromatin when RNA is depleted.

RNA inhibits the interaction of PRC2 with nucleosomes and PRC2 methyltransferase activity. (Cifuentes-Rojas *et al.*, 2014; Kaneko, Son, *et al.*, 2014; Beltran *et al.*, 2016; Wang *et al.*, 2017a). It has previously been reported that the PRC2 catalytic core is the source of the specificity of PRC2 for G4 RNA (Long *et al.*, 2019) but that JARID2 and AEBP2 also bind RNA. The function of the different PRC2 accessory factors in the recruitment of PRC2 from RNA to chromatin was not known. I discovered that PALI1, EPOP and JARID2 are necessary for the increased recruitment of PRC2 to chromatin when RNA is

depleted. I also found that JAZF1-SUZ12 also fails to increase the recruitment of PRC2 to chromatin upon RNA depletion, recapitulating the phenotype, consistent with the loss of interaction of JAZF1-SUZ12 with these factors. Thus PRC2 accessory factors regulate the transfer of PRC2 from RNA to chromatin upon RNA depletion. To further confirm the role of accessory factors in regulating chromatin recruitment of PRC2, the effect of Pol II inhibition on PRC2 recruitment to chromatin could be tested in JARID2-null, PALI1-null and EPOP-null cells, as well as *Suz12*^{GT/GT} cells expressing JAZF1-SUZ12.

6.5. Nucleosomes compete with RNA for binding to core PRC2.

In this work, I have shown that RNA competes with nucleosomes containing non-modified histones for binding to the PRC2 core. I found RNaseA treatment increased nucleosome binding by PRC2 in JARID2-null, AEBP2-null and PCL1/2/3-null nuclear extracts. However, several questions remain about the regulation of PRC2 binding to nucleosomes by PRC2 subunits. Considering that PCLs bind to CpG-rich DNA (Li *et al.*, 2017), it would be important to establish whether CpG-rich linker DNA has any effect on PRC2 binding to nucleosomes in steady state. Additionally it could be tested whether histone modifications that inhibit PRC2 nucleosome binding, such as H3K36me3 or H3K4me3 (Schmitges *et al.*, 2011), histone modifications that promote PRC2 nucleosome binding, such as H3K27me3 (Cao and Zhang, 2004; Margueron and Reinberg, 2011) and H2AK119ub (Cooper *et al.*, 2016), impact the antagonistic effect of RNA on PRC2 nucleosome binding.

6.6. G4 RNA binding compounds modulate PRC2 RNA binding in cells and reduce SUZ12 stability.

Given that PRC2 preferentially binds G4 RNA structures *in vitro* (Wang, Goodrich, *et al.*, 2017), I sought to confirm whether G4 structures promote PRC2 binding in cells by testing the effect of 3 compounds that bind with high affinity to G4 RNA structures in PRC2 binding to RNA in cells.

PDS has previously been shown to compete with PRC2 for G4 RNA binding *in vitro* (Wang 2017). PDS has been used to induce toxicity in BRCA1 and BRCA2 deficient cells by blocking DNA-replication, due to its affinity for G4 DNA structures (Zimmer *et al.*, 2016), suggesting potential uses for this compound as a cancer therapy. I showed that treating E14 cells with PDS reduced PRC2 and FUS RNA binding, but not HNRNPC RNA binding. However, this was only observed with one batch of PDS, thus it will be important to establish whether the phenotype described in this work can be recapitulated using other sources of PDS.

Interestingly, treatment of E14 cells with TMPyP4 caused loss of SUZ12 and FUS from the cells, proteins that regulate development and transcription (Yang *et al.*, 2014). It is unclear why this might be but may be consistent with the anti-tumour effects of this compound (Grand *et al.*, 2002). Considering that SUZ12 is necessary for EZH2 stability (Bracken *et al.*, 2007), and that EZH2 overexpression is correlated with poor prognosis in some types of cancer (Varambally *et al.*, 2002), it would be important to test whether TMPyP4 treatment of cancer cells depletes SUZ12 and EZH2 proteins.

Finally, I showed that MM41 increases PRC2 binding to RNA. It will be interesting to test whether this increase in binding to RNA is reflected by a corresponding loss of PRC2 binding to chromatin. MM41 has been shown to reduce mortality in a pancreatic cancer xenograft model (Ohnmacht *et al.*, 2015). Whether or not MM41 induces the removal of PRC2 from chromatin at its target genes could be tested in this model.

6.7. JAZF1-SUZ12 alters gene expression during EB formation.

PRC2 is essential for the differentiation of cells into defined lineages and thus JAZF1-SUZ12 may alter the process of decidualization. The lack of recruitment of JAZF1-SUZ12 to chromatin upon release from RNA, mimicking the recruitment of PRC2 to chromatin upon gene silencing, supports this hypothesis. Consistent with this, I discovered that JAZF1-SUZ12 disrupts differentiation of mESC to EBs. Specifically, JAZF1-SUZ12 promoted increased expression of *Gata4* and reduced expression of *Fgf5*, when compared with WT SUZ12. This effect was not shared with SUZ12 Δ 93 suggesting it may reflect a gain of function caused by fusion to JAZF1, possibly generating ectopic binding sites for PRC2 in target genes of NuA4, generating aberrant repression of genes or by recruiting NuA4 to target genes of PRC2, generating overexpression of genes. EB formation by *Suz12*^{GT/GT} mESC expressing JAZF1-SUZ12 may be a useful model for LG-ESS as it mimics the state of allelic exclusion that is prevalent in this disease, which is the most malignant form of EST to express the JAZF1-SUZ12 fusion gene (Li *et al.*, 2007).

To provide more confidence in the effect of JAZF1-SUZ12 on ESC differentiation, it will be necessary to establish whether other genes are dysregulated, for example using RNA-seq. To determine whether dysregulation of these genes is due to effects of JAZF1-SUZ12 on PRC2 recruitment, ChIP-qPCR for FLAG and H3K27me3 could be performed in ESCs and EBs. This would show whether genes over-expression is caused by failure of PRC2-mediated repression and whether failure to upregulate genes is due to ectopic binding by JAZF1-SUZ12 PRC2, potentially due to interaction of JAZF1 with NuA4 (Piunti *et al.*, 2019).

6.8. hEnSCs are a promising model for the study of LG-ESS.

Generating a LG-ESS model is necessary to understand the processes that lead to cell transformation and oncogenesis. Oncogenesis is often linked to inhibition of cell differentiation. Endometrial cells differentiate in the process of decidualisation, which is associated with dynamic changes in H3K27me3 at gene promoters (Grimaldi *et al.*, 2011).

In this work, I showed that hEnSCs can be transfected, selected and decidualized, thus providing a potential model for the study of the effect of JAZF1-SUZ12 and other fusion proteins on endometrial cell state. However, the selected population of hEnSCs did not show significant expression of SUZ12 or JAZF1-SUZ12 and did not exhibit any differences in the expression of the decidualization markers *PRL* and *IGFBP1*. Further work is needed to establish an appropriate protocol that allows sufficiently robust expression of transgenes in hEnSCs. Other

groups have transduced these cells with lentiviruses and selected the cells with puromycin (Piunti *et al.*, 2019). Lentiviral transduction could also be coupled with selection by GFP to minimize the time the cells are in culture. Once such a protocol is in place, the effect of JAZF1-SUZ12 on gene expression and PRC2 occupancy in steady-state and in a differentiation model, such as decidualisation, could be addressed.

References

- A Razin and AD Riggs (1980) 'DNA methylation and gene function', *Science*, 210(November), pp. 604–610. doi: 10.1126/SCIENCE.6254144.
- Agaimy, A. *et al.* (2018) 'Low-grade Endometrioid Stromal Sarcoma of the Paratestis', *American Journal of Surgical Pathology*, 42(5), pp. 695–700. doi: 10.1097/PAS.0000000000001030.
- Al-Raawi, D. *et al.* (2019) 'A novel form of JARID2 is required for differentiation in lineage-committed cells', *The EMBO Journal*, 38(3). doi: 10.15252/embj.201798449.
- Alekseyenko, A. a *et al.* (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(7), pp. 2488–93. doi: 10.1073/pnas.1400648111.
- Amador-ortiz, C. *et al.* (2011) 'JAZF1 and JJAZ1 gene fusion in primary extrauterine endometrial stromal sarcoma', *Human Pathology*. Elsevier Inc., 42(7), pp. 939–946. doi: 10.1016/j.humpath.2010.11.001.
- Azuara, V. *et al.* (2006) 'Chromatin signatures of pluripotent cell lines', *Nature Cell Biology*, 8(5), pp. 532–538. doi: 10.1038/ncb1403.
- Ballaré, C. *et al.* (2012) 'Phf19 links methylated Lys36 of histone H3 to regulation of Polycomb activity', *Nature Structural and Molecular Biology*. Nature Publishing Group, 19(12), pp. 1257–1265. doi: 10.1038/nsmb.2434.
- Bannister, A. J. *et al.* (2001) 'Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain', *Nature*, 410(6824), pp. 120–124. doi: 10.1038/35065138.
- Barros, F. S. V., Brosens, J. J. and Brighton, P. J. (2016) 'Isolation and Primary

Culture of Various Cell Types from Whole Human Endometrial Biopsies', *Bio-protocol*. Bio-protocol LLC., 6(22), p. e2028. doi: 10.21769/BioProtoc.2028.

Bartke, T. *et al.* (2010) 'Nucleosome-interacting proteins regulated by DNA and histone methylation', *Cell*. Elsevier Inc., 143(3), pp. 470–484. doi: 10.1016/j.cell.2010.10.012.

Bednar, J. *et al.* (1998) 'Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin.', *Proceedings of the National Academy of Sciences of the United States of America*, 95(24), pp. 14173–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9826673><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC24346>.

Beltran, M. *et al.* (2016) 'The interaction of PRC2 with RNA or chromatin is mutually antagonistic', *Genome Research*, 26(7), pp. 896–907. doi: 10.1101/gr.197632.115.

Beringer, M. *et al.* (2016) 'EPOP Functionally Links Elongin and Polycomb in Pluripotent Stem Cells', *Molecular Cell*, 64(4), pp. 645–658. doi: 10.1016/j.molcel.2016.10.018.

Bestor, T. *et al.* (1988) 'Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases', *Journal of Molecular Biology*, 203(4), pp. 971–983. doi: 10.1016/0022-2836(88)90122-2.

Bestor, T. H. (2000) 'The DNA methyltransferases of mammals', *Human Molecular Genetics*, 9(16), pp. 2395–2402. doi: 10.1093/hmg/9.16.2395.

Blackledge, N. P. *et al.* (2014) 'Variant PRC1 complex-dependent H2A

ubiquitylation drives PRC2 recruitment and polycomb domain formation', *Cell*. The Authors, 157(6), pp. 1445–1459. doi: 10.1016/j.cell.2014.05.004.

Boyer, Laurie A *et al.* (2006) 'Polycomb complexes repress developmental regulators in murine embryonic stem cells.', *Nature*, 441(7091), pp. 349–53. doi: 10.1038/nature04733.

Boyer, Laurie A. *et al.* (2006) 'Polycomb complexes repress developmental regulators in murine embryonic stem cells', *Nature*, 441(7091), pp. 349–353. doi: 10.1038/nature04733.

Bracken, A. P. *et al.* (2007) 'The Polycomb group proteins bind throughout the INK4A-ARF locus and are disassociated in senescent cells', *Genes & Development*, 21(5), pp. 525–530. doi: 10.1101/gad.415507.

Brar, A. K. *et al.* (1997) 'Progesterone-dependent decidualization of the human endometrium is mediated by cAMP.', *Endocrine*, 6(3), pp. 301–7. doi: 10.1007/BF02820507.

Brien, G. L. *et al.* (2012) 'Polycomb PHF19 binds H3K36me3 and recruits PRC2 and demethylase NO66 to embryonic stem cell genes during differentiation', *Nat Struct Mol Biol*. Nature Publishing Group, 19(12), pp. 1273–1281. doi: 10.1038/nsmb.2449.

Brien, G. L. *et al.* (2015) 'A chromatin-independent role of Polycomb-like 1 to stabilize p53 and promote cellular quiescence', *Genes & Development*, 29(21), pp. 2231–2243. doi: 10.1101/gad.267930.115.

Brosens, J. J. (1999) 'Progesterone Receptor Regulates Decidual Prolactin Expression in Differentiating Human Endometrial Stromal Cells', *Endocrinology*, 140(10), pp. 4809–4820. doi: 10.1210/en.140.10.4809.

Brown, J. L. *et al.* (1998) 'The Drosophila polycomb group gene pleiohomeotic

encodes a DNA binding protein with homology to the transcription factor YY1', *Molecular Cell*, 1(7), pp. 1057–1064. doi: 10.1016/S1097-2765(00)80106-9.

Brown, S. W. (1966) 'Heterochromatin.', *Science (New York, N.Y.)*, 151(3709), pp. 417–25. doi: 10.1126/science.151.3709.417.

Campbell, A., Cotter, R. and Pardon, J. (1978) 'Light scattering measurements supporting helical structures for chromatin in solution', 5(5), pp. 1571–1580.

Cao, R. *et al.* (2008) 'Role of hPHF1 in H3K27 Methylation and Hox Gene Silencing', *Molecular and Cellular Biology*, 28(5), pp. 1862–1872. doi: 10.1128/mcb.01589-07.

Cao, R. and Zhang, Y. (2004) 'SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex', *Molecular Cell*, 15(1), pp. 57–67. doi: 10.1016/j.molcel.2004.06.020.

De Cegli, R. *et al.* (2013) 'Reverse engineering a mouse embryonic stem cell-specific transcriptional network reveals a new modulator of neuronal differentiation', *Nucleic Acids Research*, 41(2), pp. 711–726. doi: 10.1093/nar/gks1136.

Chen, S. *et al.* (2018a) 'Unique Structural Platforms of Suz12 Dictate Distinct Classes of PRC2 for Chromatin Binding', *Molecular Cell*. Elsevier Inc., 69(5), pp. 840-852.e5. doi: 10.1016/j.molcel.2018.01.039.

Chen, S. *et al.* (2018b) 'Unique Structural Platforms of Suz12 Dictate Distinct Classes of PRC2 for Chromatin Binding', *Molecular Cell*, 69(5), pp. 840-852.e5. doi: 10.1016/j.molcel.2018.01.039.

Cifuentes-Rojas, C. *et al.* (2014) 'Regulatory Interactions between RNA and Polycomb Repressive Complex 2', *Molecular Cell*. Elsevier Inc., 55(2), pp. 171–185. doi: 10.1016/j.molcel.2014.05.009.

- Conklin, C. M. J. and Longacre, T. A. (2014) 'Endometrial stromal tumors: the new WHO classification.', *Advances in anatomic pathology*, 21(6), pp. 383–93. doi: 10.1097/PAP.0000000000000046.
- Conway, E. *et al.* (2018) 'A Family of Vertebrate-Specific Polycombs Encoded by the LCOR/LCORL Genes Balance PRC2 Subtype Activities', *Molecular Cell*, 70(3), pp. 408-421.e8. doi: 10.1016/j.molcel.2018.03.005.
- Cooper, S. *et al.* (2016) 'Jarid2 binds mono-ubiquitylated H2A lysine 119 to mediate crosstalk between Polycomb complexes PRC1 and PRC2', *Nature Communications*. Nature Publishing Group, 7, pp. 1–8. doi: 10.1038/ncomms13661.
- Costanzi, C. and Pehrson, J. R. (1998) 'Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals', *Nature*, 393(6685), pp. 599–601. doi: 10.1038/31275.
- Côté, J. *et al.* (1994) 'Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex', *Science*, 265(5168), pp. 53–60. doi: 10.1126/science.8016655.
- Di Croce, L. and Helin, K. (2013) 'Transcriptional regulation by Polycomb group proteins', *Nature Structural and Molecular Biology*. Nature Publishing Group, 20(10), pp. 1147–1155. doi: 10.1038/nsmb.2669.
- Cross, S. H. *et al.* (1997) 'A component of the transcriptional repressor MeCP1 shares a motif with DNA methyltransferase and HRX proteins', *Nature Genetics*, 16(3), pp. 256–259. doi: 10.1038/ng0797-256.
- Davidovich, C. *et al.* (2013) 'Promiscuous RNA binding by Polycomb repressive complex 2.', *Nature Structural & Molecular Biology*, 20(11), pp. 1250–1257. doi: 10.1038/nsmb.2679.

- Davidovich, C. *et al.* (2015) 'Toward a consensus on the binding specificity and promiscuity of PRC2 for RNA', *Molecular Cell*. Elsevier Inc., 57(3), pp. 552–559. doi: 10.1016/j.molcel.2014.12.017.
- Davidson, B. and Micci, F. (2017) 'Molecular characteristics of uterine sarcomas', *Expert Review of Molecular Diagnostics*. Taylor & Francis, 17(5), p. 14737159.2017.1311790. doi: 10.1080/14737159.2017.1311790.
- Deaton, A. M. and Bird, A. (2011) 'CpG islands and the regulation of transcription', *Genes and Development*, 25(10), pp. 1010–1022. doi: 10.1101/gad.2037511.
- Dewaele, B. *et al.* (2014) 'Identification of a novel, recurrent MBTD1-CXorf67 fusion in low-grade endometrial stromal sarcoma', *International Journal of Cancer*, 134(5), pp. 1112–1122. doi: 10.1002/ijc.28440.
- Dickson, B. C. *et al.* (2018) 'Novel *EPC1* gene fusions in endometrial stromal sarcoma', *Genes, Chromosomes and Cancer*, (April), pp. 1–6. doi: 10.1002/gcc.22649.
- Doyon, Y. *et al.* (2004) 'Structural and Functional Conservation of the NuA4 Histone Acetyltransferase Complex from Yeast to Humans', *Molecular and Cellular Biology*, 24(5), pp. 1884–1896. doi: 10.1128/mcb.24.5.1884-1896.2004.
- Faust, C. *et al.* (1995) 'The eed mutation disrupts anterior mesoderm production in mice', *Development*, 121(2), pp. 273–285.
- Ferrari, K. J. *et al.* (2014) 'Polycomb-Dependent H3K27me1 and H3K27me2 Regulate Active Transcription and Enhancer Fidelity', *Molecular Cell*, 53(1), pp. 49–62. doi: 10.1016/j.molcel.2013.10.030.
- Fischle, W. *et al.* (2003) 'Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by polycomb and HP1 chromodomains',

- Genes and Development*, 17(15), pp. 1870–1881. doi: 10.1101/gad.1110503.
- Francis, N. J., Kingston, R. E. and Woodcock, C. L. (2004) 'Chromatin compaction by a polycomb group protein complex', *Science*, 306(5701), pp. 1574–1577. doi: 10.1126/science.1100576.
- Frank, S. R. *et al.* (2003) 'MYC recruits the TIP60 histone acetyltransferase complex to chromatin', *EMBO Reports*, 4(6), pp. 575–580. doi: 10.1038/sj.embor.embor861.
- Gamble, M. J. *et al.* (2010) 'The histone variant macroH2A1 marks repressed autosomal chromatin, but protects a subset of its target genes from silencing', *Genes and Development*, 24(1), pp. 21–32. doi: 10.1101/gad.1876110.
- Gao, Z. *et al.* (2012) 'PCGF Homologs, CBX Proteins, and RYBP Define Functionally Distinct PRC1 Family Complexes', *Molecular Cell*. Elsevier Inc., 45(3), pp. 344–356. doi: 10.1016/j.molcel.2012.01.002.
- Gellersen, B. and Brosens, J. J. (2014) 'Cyclic decidualization of the human endometrium in reproductive health and failure', *Endocrine Reviews*, 35(6), pp. 851–905. doi: 10.1210/er.2014-1045.
- Goll, M. G. *et al.* (2006) 'Methylation of tRNA^{Asp} by the DNA methyltransferase homolog Dnmt2', *Science*, 311(5759), pp. 395–398. doi: 10.1126/science.1120976.
- Grand, C. L. *et al.* (2002) 'The cationic porphyrin TMPyP4 down-regulates c-MYC and human telomerase reverse transcriptase expression and inhibits tumor growth in vivo', *Molecular Cancer Therapeutics*, 1(8), pp. 565–573.
- Grijzenhout, A. *et al.* (2016) 'Functional analysis of AEBP2, a PRC2 Polycomb protein, reveals a Trithorax phenotype in embryonic development and in ES cells.', *Development (Cambridge, England)*, pp. 2716–2723. doi:

10.1242/dev.123935.

Grimaldi, G. *et al.* (2011) 'Down-regulation of the histone methyltransferase EZH2 contributes to the epigenetic programming of decidualizing human endometrial stromal cells', *Molecular Endocrinology*, 25(11), pp. 1892–1903. doi: 10.1210/me.2011-1139.

Groudine, M., Eisenman, R. and Weintraub, H. (1981) 'Chromatin structure of endogenous retroviral genes and activation by an inhibitor of DNA methylation', *Nature*, 292(5821), pp. 311–317. doi: 10.1038/292311a0.

Grüne, T. *et al.* (2003) 'Crystal structure and functional analysis of a nucleosome recognition module of the remodeling factor ISWI', *Molecular Cell*, 12(2), pp. 449–460. doi: 10.1016/S1097-2765(03)00273-9.

Le Guezennec, X. *et al.* (2006) 'MBD2/NuRD and MBD3/NuRD, Two Distinct Complexes with Different', *Molecular and Cellular Biology*, 26(3), pp. 843–851. doi: 10.1128/MCB.26.3.843.

Han, Z. *et al.* (2007) 'Structural Basis of EZH2 Recognition by EED', *Structure*, 15(10), pp. 1306–1315. doi: 10.1016/j.str.2007.08.007.

Hardy, S. *et al.* (2009) 'The euchromatic and heterochromatic landscapes are shaped by antagonizing effects of transcription on H2A.Z deposition', *PLoS Genetics*, 5(10). doi: 10.1371/journal.pgen.1000687.

Hassan, A. H. *et al.* (2002) 'Function and Selectivity of Bromodomains in Anchoring Chromatin-Modifying Complexes to Promoter Nucleosomes Deletion of the Sth1 bromodomain results in a condensation in the other related RSC subunits (Rsc1 and Rsc2', *Cell Biology and Physiology*, 111, p. 63110.

Hauri, S. *et al.* (2016) 'A High-Density Map for Navigating the Human Polycomb Complexome', *Cell Reports*, 17(2), pp. 583–595. doi:

10.1016/j.celrep.2016.08.096.

He, Y. *et al.* (2017) 'The EED protein-protein interaction inhibitor A-395 inactivates the PRC2 complex', *Nature Chemical Biology*, 13(4), pp. 389–395. doi: 10.1038/nchembio.2306.

Healy, E. *et al.* (2019) 'PRC2.1 and PRC2.2 Synergize to Coordinate H3K27 Trimethylation', *Molecular Cell*. Elsevier, 76(3), pp. 437-452.e6. doi: 10.1016/j.molcel.2019.08.012.

Herz, H.-M. *et al.* (2012) 'Polycomb Repressive Complex 2-Dependent and -Independent Functions of Jarid2 in Transcriptional Regulation in *Drosophila*', *Molecular and Cellular Biology*, 32(9), pp. 1683–1693. doi: 10.1128/mcb.06503-11.

Hidalgo, I. *et al.* (2012) 'Ezh1 is required for hematopoietic stem cell maintenance and prevents senescence-like cell cycle arrest', *Cell Stem Cell*, 11(5), pp. 649–662. doi: 10.1016/j.stem.2012.08.001.

Højfeldt, J. W. *et al.* (2018) 'Accurate H3K27 methylation can be established de novo by SUZ12-directed PRC2', *Nature Structural and Molecular Biology*. Springer US, 25(3), pp. 225–232. doi: 10.1038/s41594-018-0036-6.

van Holde, K. and Zlatanova, J. (1995) 'Chromatin Higher Order Structure: Chasing a Mirage?', *Journal of Biological Chemistry*, 270(15), pp. 8373–8376. doi: 10.1074/jbc.270.15.8373.

Hosogane, M. *et al.* (2016) 'Lack of Transcription Triggers H3K27me3 Accumulation in the Gene Body', *Cell Reports*. The Author(s), 16(3), pp. 696–706. doi: 10.1016/j.celrep.2016.06.034.

Huppertz, I. *et al.* (2014) 'iCLIP: Protein-RNA interactions at nucleotide resolution', *Methods*, 65(3), pp. 274–287. doi: 10.1016/j.ymeth.2013.10.011.

Jacobs, J. L. *et al.* (1999) 'The oncogene and Polycombgroup gene bmi-1 regulates cell proliferation and senescence through the ink4a locus', *Nature*, 397(6715), pp. 164–168. doi: 10.1038/16476.

Jain, S. U. *et al.* (2019) 'PFA ependymoma-associated protein EZHIP inhibits PRC2 activity through a H3 K27M-like mechanism', *Nature Communications*. Springer US, 10(1). doi: 10.1038/s41467-019-09981-6.

Jenuwein, T. *et al.* (2000) 'Regulation of chromatin structure by site-specific histone H3 methyltransferases', *Nature*, 406(6796), pp. 593–599. doi: 10.1038/35020506.

Jenuwein, T. *et al.* (2002) 'Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins', *Nature*, 410(6824), pp. 116–120. doi: 10.1038/35065132.

Jermann, P. *et al.* (2014) 'Short sequences can efficiently recruit histone H3 lysine 27 trimethylation in the absence of enhancer activity and DNA methylation.', *Proceedings of the National Academy of Sciences of the United States of America*, 111(33), pp. E3415-21. doi: 10.1073/pnas.1400672111.

Jiao, L. and Liu, X. (2015) 'Structural basis of histone H3K27 trimethylation by an active polycomb repressive complex 2', *Science*, 350(6258), pp. aac4383–aac4383. doi: 10.1126/science.aac4383.

Jones, P. A. and Taylor, S. M. (1980) 'Cellular differentiation, cytidine analogs and DNA methylation.', *Cell*, 20(1), pp. 85–93. doi: 10.1016/0092-8674(80)90237-8.

Joshi, A. A. and Struhl, K. (2005) 'Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to pol II elongation', *Molecular Cell*, 20(6), pp. 971–978. doi: 10.1016/j.molcel.2005.11.021.

Justin, N. *et al.* (2016) 'Structural basis of oncogenic histone H3K27M inhibition

of human polycomb repressive complex 2', *Nature Comm.* Nature Publishing Group, 7, p. 11316. doi: 10.1038/ncomms11316.

Kalb, R. *et al.* (2014a) 'Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression', *Nature Structural & Molecular Biology*, 21(6), pp. 569–571. doi: 10.1038/nsmb.2833.

Kalb, R. *et al.* (2014b) 'SAF-A Regulates Interphase Chromosome Structure through Oligomerization with Chromatin- Associated RNAs', *Nature Structural & Molecular Biology*, 21(6), pp. 569–571. doi: 10.1038/nsmb.2833.

Kaneko, S. *et al.* (2013) 'PRC2 binds active promoters and contacts nascent RNAs in embryonic stem cells', *Nat Struct Mol Biol.* Nature Publishing Group, 20(11), pp. 1258–1264. doi: 10.1038/nsmb.2700.

Kaneko, S., Bonasio, R., *et al.* (2014) 'Interactions between JARID2 and Noncoding RNAs Regulate PRC2 Recruitment to Chromatin', *Molecular Cell*, 53(2), pp. 290–300. doi: 10.1016/j.molcel.2013.11.012.

Kaneko, S., Son, J., *et al.* (2014) 'Nascent RNA interaction keeps PRC2 activity poised and in check', *Genes and Development*, 28(18), pp. 1983–1988. doi: 10.1101/gad.247940.114.

Kanhere, A. *et al.* (2010) 'Short RNAs Are Transcribed from Repressed Polycomb Target Genes and Interact with Polycomb Repressive Complex-2', *Molecular Cell*, 38(5), pp. 675–688. doi: 10.1016/j.molcel.2010.03.019.

Kanhere, A. *et al.* (2012) 'T-bet and GATA3 orchestrate Th1 and Th2 differentiation through lineage-specific targeting of distal regulatory elements.', *Nature communications*, 3, p. 1268. doi: 10.1038/ncomms2260.

Kasinath, V. *et al.* (2018) 'Structures of human PRC2 with its cofactors AEBP2 and JARID2', *Science*, 5700(January), p. eaar5700. doi:

10.1126/science.aar5700.

Khalil, A. M. *et al.* (2009) 'Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression', *Proceedings of the National Academy of Sciences of the United States of America*, 106(28), pp. 11667–11672. doi: 10.1073/pnas.0904715106.

Kim, H., Kang, K. and Kim, J. (2009) 'AEBP2 as a potential targeting protein for Polycomb Repression Complex PRC2', *Nucleic Acids Research*, 37(9), pp. 2940–2950. doi: 10.1093/nar/gkp149.

Kin, K. *et al.* (2015) 'Cell-type phylogenetics and the origin of endometrial stromal cells', *Cell Reports*, 10(8), pp. 1398–1409. doi: 10.1016/j.celrep.2015.01.062.

Kleff, S. *et al.* (1995) 'Identification of a gene encoding a yeast histone H4 acetyltransferase', *Journal of Biological Chemistry*, 270(42), pp. 24674–24677. doi: 10.1074/jbc.270.42.24674.

Knutson, S. K. *et al.* (2014) 'Selective inhibition of EZH2 by EPZ-6438 leads to potent antitumor activity in EZH2-mutant non-Hodgkin lymphoma', *Molecular Cancer Therapeutics*, 13(4), pp. 842–854. doi: 10.1158/1535-7163.MCT-13-0773.

Koontz, J. I. *et al.* (2001) 'Frequent fusion of the JAZF1 and JJAZ1 genes in endometrial stromal tumors.', *Proceedings of the National Academy of Sciences of the United States of America*, 98(11), pp. 6348–53. doi: 10.1073/pnas.101132598.

Kornberg, R. D. (1974) 'Chromatin Structure : A Repeating Unit of Histones and DNA Published by: American Association for the Advancement of Science
Linked references are available on JSTOR for this article : Chromatin Structure : A Repeating Unit of Histones and DNA Chromatin', *Science (New York, N.Y.)*,

184(4139), pp. 868–871.

Kouzarides, T. (2007) 'Chromatin modifications and their function.', *Cell*, 128(4), pp. 693–705. doi: 10.1016/j.cell.2007.02.005.

Krogan, N. J. *et al.* (2003) 'Methylation of Histone H3 by Set2 in *Saccharomyces cerevisiae* Is Linked to Transcriptional Elongation by RNA Polymerase II
Downloaded from <http://mcb.asm.org/> on September 18 , 2012 by IMPERIAL COLLEGE LONDON Methylation of Histone H3 by Set2 in *Saccharom*', 23(12), pp. 4207–4218. doi: 10.1128/MCB.23.12.4207.

Ku, M. *et al.* (2008) 'Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains', *PLoS Genetics*, 4(10). doi: 10.1371/journal.pgen.1000242.

LAEMMLI, U. K. (1970) 'Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4', *Nature*, 227(5259), pp. 680–685. doi: 10.1038/227680a0.

Landeira, D. *et al.* (2010) 'Jarid2 is a PRC2 component in embryonic stem cells required for multi-lineage differentiation and recruitment of PRC1 and RNA Polymerase II to developmental regulators', *Nature Cell Biology*. Nature Publishing Group, 12(6), pp. 618–624. doi: 10.1038/ncb2065.

Laugesen, A., Højfeldt, J. W. and Helin, K. (2019) 'Molecular Mechanisms Directing PRC2 Recruitment and H3K27 Methylation', *Molecular Cell*. Elsevier Inc., 74(1), pp. 8–18. doi: 10.1016/j.molcel.2019.03.011.

Lee, T. I. *et al.* (2006) 'Control of Developmental Regulators by Polycomb in Human Embryonic Stem Cells', *Cell*, 125(2), pp. 301–313. doi: 10.1016/j.cell.2006.02.043.

Lei, H. *et al.* (1996) 'De novo DNA cytosine methyltransferase activities in mouse

embryonic stem cells.', *Development (Cambridge, England)*, 122(10), pp. 3195–205. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8898232>.

Levine, S. S. *et al.* (2002) 'The Core of the Polycomb Repressive Complex Is Compositionally and Functionally Conserved in Flies and Humans', *Molecular and Cellular Biology*, 22(17), pp. 6070–6078. doi: 10.1128/mcb.22.17.6070-6078.2002.

Li, A. *et al.* (2010) 'Phosphorylation of histone H2A.X by DNA-dependent protein kinase is not affected by core histone acetylation, but it alters nucleosome stability and histone H1 binding', *Journal of Biological Chemistry*, 285(23), pp. 17778–17788. doi: 10.1074/jbc.M110.116426.

Li, B., Carey, M. and Workman, J. L. (2007) 'The Role of Chromatin during Transcription', *Cell*, 128(4), pp. 707–719. doi: 10.1016/j.cell.2007.01.015.

Li, H. *et al.* (2007) 'Effects of rearrangement and allelic exclusion of JJAZ1/SUZ12 on cell proliferation and survival.', *Proceedings of the National Academy of Sciences of the United States of America*, 104(50), pp. 20001–6. doi: 10.1073/pnas.0709986104.

Li, H. *et al.* (2008) 'A Neoplastic Gene Fusion Mimics Trans-Splicing of RNAs in Normal Human Cells', *Science*, 321(5894), pp. 1357–1361. doi: 10.1126/science.1156725.

Li, H. *et al.* (2017) 'Polycomb-like proteins link the PRC2 complex to CpG islands', *Nature*. Nature Publishing Group. doi: 10.1038/nature23881.

Liang, G. *et al.* (2004) 'Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome', *Proceedings of the National Academy of Sciences*, 101(19), pp. 7357–7362. doi: 10.1073/pnas.0401866101.

- Liefke, R., Karwacki-Neisius, V. and Shi, Y. (2016) 'EPOP Interacts with Elongin BC and USP7 to Modulate the Chromatin Landscape', *Molecular Cell*. Elsevier Inc., 64(4), pp. 659–672. doi: 10.1016/j.molcel.2016.10.019.
- Lister, R. *et al.* (2009) 'Human DNA methylomes at base resolution show widespread epigenomic differences', *Nature*. Nature Publishing Group, 462(7271), pp. 315–322. doi: 10.1038/nature08514.
- Long, Y. *et al.* (2017) 'Conserved RNA-binding specificity of polycomb repressive complex 2 is achieved by dispersed amino acid patches in EZH2', *eLife*, 6(1), pp. 1–23. doi: 10.7554/eLife.31558.
- Lorsback, R. B. *et al.* (2003) 'TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;23) [3]', *Leukemia*, 17(3), pp. 637–641. doi: 10.1038/sj.leu.2402834.
- Luger, K. *et al.* (1997) 'Crystal structure of the nucleosome core particle at 2.8 Å resolution', *Nature*, 389, pp. 251–260. doi: 10.1038/38444.
- Ma, X. *et al.* (2016) 'The JAZF1/SUZ12 fusion protein disrupts PRC2 complexes and impairs chromatin repression during human endometrial stromal tumorigenesis.', *Oncotarget*. doi: 10.18632/oncotarget.13270.
- Ma, X. *et al.* (2017) 'The JAZF1-SUZ12 fusion protein disrupts PRC2 complexes and impairs chromatin repression during human endometrial stromal tumorigenesis', *Oncotarget*, 8(3), pp. 4062–4078. doi: 10.18632/oncotarget.13270.
- Makise, N. *et al.* (2019) 'Low-grade endometrial stromal sarcoma with a novel MEAF6-SUZ12 fusion', *Virchows Archiv*. Virchows Archiv, pp. 4–8. doi: 10.1007/s00428-019-02588-8.
- Margueron, R. *et al.* (2008) 'Ezh1 and Ezh2 Maintain Repressive Chromatin

through Different Mechanisms', *Molecular Cell*. Elsevier Inc., 32(4), pp. 503–518. doi: 10.1016/j.molcel.2008.11.004.

Margueron, R. and Reinberg, D. (2011) 'The Polycomb complex PRC2 and its mark in life.', *Nature*, 469(7330), pp. 343–9. doi: 10.1038/nature09784.

Mccabe, M. T. *et al.* (2012) 'Mutation of A677 in histone methyltransferase EZH2 in human B-cell lymphoma promotes hypertrimethylation of histone H3 on lysine 27 (H3K27)', 27(18), pp. 1–6. doi: 10.1073/pnas.1116418109.

Mendenhall, E. M. *et al.* (2010) 'GC-rich sequence elements recruit PRC2 in mammalian ES cells', *PLoS Genetics*, 6(12), pp. 1–10. doi: 10.1371/journal.pgen.1001244.

Micci, F. *et al.* (2006) 'Consistent rearrangement of chromosomal band 6p21 with generation of fusion genes JAZF1/PHF1 and EPC1/PHF1 in endometrial stromal sarcoma', *Cancer Research*, 66(1), pp. 107–112. doi: 10.1158/0008-5472.CAN-05-2485.

Micci, F. *et al.* (2014) 'MEAF6/PHF1 is a recurrent gene fusion in endometrial stromal sarcoma', *Cancer Letters*. Elsevier Ireland Ltd, 347(1), pp. 75–78. doi: 10.1016/j.canlet.2014.01.030.

Micci, F. *et al.* (2016) 'Cytogenetic and molecular profile of endometrial stromal sarcoma', *Genes, Chromosomes and Cancer*, 55(11), pp. 834–846. doi: 10.1002/gcc.22380.

Milne, T. A. *et al.* (2002) 'MLL targets SET domain methyltransferase activity to Hox gene promoters', *Molecular Cell*, 10(5), pp. 1107–1117. doi: 10.1016/S1097-2765(02)00741-4.

Min, J., Zhang, Y. and Xu, R. M. (2003) 'Structural basis for specific binding of polycomb chromodomain to histone H3 methylated at Lys 27', *Genes and*

- Development*, 17(15), pp. 1823–1828. doi: 10.1101/gad.269603.
- Mirabella, A. C., Foster, B. M. and Bartke, T. (2016) 'Chromatin deregulation in disease', *Chromosoma*, 125(1), pp. 75–93. doi: 10.1007/s00412-015-0530-0.
- Mito, Y., Henikoff, J. G. and Henikoff, S. (2007) 'Histone Replacement Marks the Boundaries of cis-Regulatory Domains', *Science*, 315(5817), pp. 1408–1411. doi: 10.1126/science.1134004.
- Mitsui, K. *et al.* (1999) 'Cloning and characterization of a novel p21(Cip1/Waf1)-interacting zinc finger protein, Ciz1', *Biochemical and Biophysical Research Communications*, 264(2), pp. 457–464. doi: 10.1006/bbrc.1999.1516.
- Montgomery, N. D. *et al.* (2005) 'The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation', *Current Biology*, 15(10), pp. 942–947. doi: 10.1016/j.cub.2005.04.051.
- Moon, K. J. *et al.* (2005) 'The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription', *Molecular Cell*, 19(4), pp. 523–534. doi: 10.1016/j.molcel.2005.06.027.
- Moreau, J. L. *et al.* (2003) 'Regulated displacement of TBP from the PHO8 promoter in vivo requires Cbf1 and the Isw1 chromatin remodeling complex', *Molecular Cell*, 11(6), pp. 1609–1620. doi: 10.1016/S1097-2765(03)00184-9.
- Morillon, A. *et al.* (2003) 'Isw1 Chromatin Remodeling ATPase Coordinates Transcription Elongation and Termination by RNA Polymerase II', *Cell*, 115(4), pp. 425–435. doi: 10.1016/S0092-8674(03)00880-8.
- Morris, M. J. *et al.* (2012) 'The porphyrin TmPyP4 unfolds the extremely stable G-quadruplex in MT3-MMP mRNA and alleviates its repressive effect to enhance translation in eukaryotic cells', *Nucleic Acids Research*, 40(9), pp. 4137–4145.

doi: 10.1093/nar/gkr1308.

Müller, J. *et al.* (2002) 'Histone methyltransferase activity of a *Drosophila* Polycomb group repressor complex', *Cell*, 111(2), pp. 197–208. doi: 10.1016/S0092-8674(02)00976-5.

Müller, J. and Kassis, J. A. (2006) 'Polycomb response elements and targeting of Polycomb group proteins in *Drosophila*', *Current Opinion in Genetics and Development*, 16(5), pp. 476–484. doi: 10.1016/j.gde.2006.08.005.

Musselman, C. A. *et al.* (2009) 'Binding of the CHD4 PHD2 finger to histone H3 is modulated by covalent modifications', *Biochemical Journal*, 423(2), pp. 179–187. doi: 10.1042/BJ20090870.

Nakajima, T. *et al.* (2004) 'TIP27: A novel repressor of the nuclear orphan receptor TAK1/TR4', *Nucleic Acids Research*, 32(14), pp. 4194–4204. doi: 10.1093/nar/gkh741.

Nan, X. *et al.* (1998) 'Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex', *Nature*, 393(6683), pp. 386–389. doi: 10.1038/30764.

Nekrasov, M. *et al.* (2007) 'Pcl-PRC2 is needed to generate high levels of H3-K27 trimethylation at Polycomb target genes', *EMBO Journal*, 26(18), pp. 4078–4088. doi: 10.1038/sj.emboj.7601837.

Nucci, M. R. *et al.* (2007) 'Molecular analysis of the JAZF1/JJAZ1 gene fusion by RT-PCR and fluorescence in situ hybridisation in endometrial stromal neoplasms', *Am J Surg Pathol*, 31(1), pp. 65–70. doi: 10.1097/01.pas.0000213327.86992.d1.

O'Carroll, D. *et al.* (2001) 'The Polycomb -Group Gene Ezh2 Is Required for Early Mouse Development', *Molecular and cellular biology*, 21(13), pp. 4330–4336. doi: 10.1128/MCB.21.13.4330.

- Ogryzko, V. V. *et al.* (1996) 'The transcriptional coactivators p300 and CBP are histone acetyltransferases', *Cell*, 87(5), pp. 953–959. doi: 10.1016/S0092-8674(00)82001-2.
- Ohnmacht, S. A. *et al.* (2015) 'A G-quadruplex-binding compound showing anti-tumour activity in an in vivo model for pancreatic cancer', *Scientific Reports*. Nature Publishing Group, 5, pp. 1–11. doi: 10.1038/srep11385.
- Okano, M. (1998) 'Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases Non-invasive sexing of preimplantation stage mammalian embryos', *Nature*, 19(july), pp. 219–220. doi: 10.1038/890.
- Ono, R. *et al.* (2002) 'LCX, leukemia-associated protein with a CXXC domain, is fused to MLL in acute myeloid leukemia with trilineage dysplasia having t(10;11)(q22;q23)', *Cancer Research*, 62(14), pp. 4075–4080.
- Orlando, V. *et al.* (1998) 'Binding of Trithorax and Polycomb proteins to the bithorax complex: Dynamic changes during early Drosophila embryogenesis', *EMBO Journal*, 17(17), pp. 5141–5150. doi: 10.1093/emboj/17.17.5141.
- Pajtler, K. W. *et al.* (2018) 'Molecular heterogeneity and CXorf67 alterations in posterior fossa group A (PFA) ependymomas', *Acta Neuropathologica*. Springer Berlin Heidelberg, 136(2), pp. 211–226. doi: 10.1007/s00401-018-1877-0.
- Palmer, D. K. *et al.* (1987) 'A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones', *Journal of Cell Biology*, 104(4), pp. 805–815. doi: 10.1083/jcb.104.4.805.
- Panagopoulos, I. (2010) 'Absence of the JAZF1/SUZ12 chimeric transcript in the immortalized non-neoplastic endometrial stromal cell line T HESCs', *Oncology Letters*, 1(6), pp. 947–950. doi: 10.3892/ol.2010.185.
- Panagopoulos, I. *et al.* (2012) 'Novel fusion of MYST/Esa1-associated factor 6

and PHF1 in endometrial stromal sarcoma', *PLoS ONE*, 7(6), pp. 3–8. doi: 10.1371/journal.pone.0039354.

Panagopoulos, I. *et al.* (2013) 'Fusion of the ZC3H7B and BCOR genes in endometrial stromal sarcomas carrying an X;22-translocation.', *Genes, chromosomes & cancer*, 52(7), pp. 610–8. doi: 10.1002/gcc.22057.

Panagopoulos, I., Mertens, F. and Griffin, C. A. (2008) 'An endometrial stromal sarcoma cell line with the JAZF1/PHF1 chimera', *Cancer Genetics and Cytogenetics*, 185(2), pp. 74–77. doi: 10.1016/j.cancergencyto.2008.04.020.

Pasini, D. *et al.* (2004) 'Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity', *The EMBO Journal*, 23(20), pp. 4061–4071. doi: 10.1038/sj.emboj.7600402.

Pasini, D. *et al.* (2007) 'The Polycomb Group Protein Suz12 Is Required for Embryonic Stem Cell Differentiation □', *Molecular and Cellular Biology*, 27(10), pp. 3769–3779. doi: 10.1128/MCB.01432-06.

Pengelly, A. R. *et al.* (2013) 'A Histone Mutant Reproduces the Phenotype Caused by Loss of Histone-Modifying Factor Polycomb', *Science*, 339(6120), pp. 698–699. doi: 10.1126/science.1231382.

Perino, M. *et al.* (2018) 'MTF2 recruits Polycomb Repressive Complex 2 by helical-shape-selective DNA binding', *Nature Genetics*. Springer US, pp. 1–9. doi: 10.1038/s41588-018-0134-8.

Piunti, A. *et al.* (2019) 'Catacomb: An endogenous inducible gene that antagonizes H3K27 methylation activity of Polycomb repressive complex 2 via an H3K27M-like mechanism', *Science Advances*, 5(7). doi: 10.1126/sciadv.aax2887.

Plath, K. *et al.* (2003) 'Role of histone H3 lysine 27 methylation in X inactivation.',

Science (New York, N.Y.), 300(5616), pp. 131–5. doi: 10.1126/science.1084274.

Qi, W. *et al.* (2012) 'Selective inhibition of Ezh2 by a small molecule inhibitor blocks tumor cells proliferation', *Proceedings of the National Academy of Sciences*, 109(52), pp. 21360–21365. doi: 10.1073/pnas.1210371110.

Qi, W. *et al.* (2017) 'An allosteric PRC2 inhibitor targeting the H3K27me3 binding pocket of EED', *Nature Chemical Biology*, 13(4), pp. 381–388. doi: 10.1038/nchembio.2304.

Ragazzini, R. *et al.* (2019) 'EZHIP constrains Polycomb Repressive Complex 2 activity in germ cells', *bioRxiv*, 33(0), pp. 1–18. doi: 10.1038/s41467-019-11800-x.

Rai, A. N. *et al.* (2013) 'Elements of the Polycomb Repressor SU(Z)12 Needed for Histone H3-K27 Methylation, the Interface with E(Z), and In Vivo Function', *Molecular and Cellular Biology*, 33(24), pp. 4844–4856. doi: 10.1128/mcb.00307-13.

Ray-Gallet, D. *et al.* (2011) 'Dynamics of Histone H3 Deposition In Vivo Reveal a Nucleosome Gap-Filling Mechanism for H3.3 to Maintain Chromatin Integrity', *Molecular Cell*, 44(6), pp. 928–941. doi: 10.1016/j.molcel.2011.12.006.

Riising, E. M. *et al.* (2014) 'Gene silencing triggers polycomb repressive complex 2 recruitment to CpG Islands genome wide', *Molecular Cell*. Elsevier Inc., 55(3), pp. 347–360. doi: 10.1016/j.molcel.2014.06.005.

Rinn, J. L. *et al.* (2007) 'Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs', *Cell*, 129(7), pp. 1311–1323. doi: 10.1016/j.cell.2007.05.022.

Rogakou, E. P. *et al.* (1998) 'DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139', *Journal of Biological Chemistry*, 273(10), pp.

5858–5868. doi: 10.1074/jbc.273.10.5858.

Sarma, K. *et al.* (2008) 'Ezh2 Requires PHF1 To Efficiently Catalyze H3 Lysine 27 Trimethylation In Vivo', *Molecular and Cellular Biology*, 28(8), pp. 2718–2731. doi: 10.1128/mcb.02017-07.

Sato, K. *et al.* (2007) 'Extrauterine endometrial stromal sarcoma with JAZF1/JJAZ1 fusion confirmed by RT-PCR and interphase FISH presenting as an inguinal tumor', *Virchows Archiv*, 450(3), pp. 349–353. doi: 10.1007/s00428-006-0345-8.

Saurin, A. J. *et al.* (2001) 'A drosophila polycomb group complex includes zeste and dTAFII proteins', *Nature*, 412(6847), pp. 655–660. doi: 10.1038/35088096.

Scaffidi, P. and Misteli, T. (2011) 'In vitro generation of human cells with cancer stem cell properties', *Nat Cell Biol.* Nature Publishing Group, 13(9), pp. 1051–61. doi: 10.1038/ncb2308.

Scelfo, A. *et al.* (2019) 'Functional Landscape of PCGF Proteins Reveals Both RING1A/B-Dependent-and RING1A/B-Independent-Specific Activities', *Molecular Cell*, 74(5), pp. 1037-1052.e7. doi: 10.1016/j.molcel.2019.04.002.

Scelfo, A., Piunti, A. and Pasini, D. (2015) 'The controversial role of the Polycomb group proteins in transcription and cancer: How much do we not understand Polycomb proteins?', *FEBS Journal*, 282(9), pp. 1703–1722. doi: 10.1111/febs.13112.

Schmitges, F. W. *et al.* (2011) 'Histone Methylation by PRC2 Is Inhibited by Active Chromatin Marks', *Molecular Cell*, 42(3), pp. 330–341. doi: 10.1016/j.molcel.2011.03.025.

Schneiderman, J. I. *et al.* (2012) 'Nucleosome-depleted chromatin gaps recruit assembly factors for the H3.3 histone variant', *Proceedings of the National*

Academy of Sciences of the United States of America, 109(48), pp. 19721–19726. doi: 10.1073/pnas.1206629109.

Shao, Z. *et al.* (1999) 'Stabilization of chromatin structure by PRC1, a polycomb complex', *Cell*, 98(1), pp. 37–46. doi: 10.1016/S0092-8674(00)80604-2.

Shi, Yujiang *et al.* (2004) 'Histone Demethylation Mediated by the Nuclear Amine Oxidase Homolog LSD1', *Cell*, 119(7), pp. 941–953. doi: 10.1016/j.cell.2004.12.012.

Simon, J. a and Kingston, R. E. (2009) 'Mechanisms of polycomb gene silencing: knowns and unknowns.', *Nature reviews. Molecular cell biology*. Nature Publishing Group, 10(10), pp. 697–708. doi: 10.1038/nrm2763.

Sneeringer, C. J. *et al.* (2010) 'Coordinated activities of wild-type plus mutant EZH2 drive tumor-associated hypertrimethylation of lysine 27 on histone H3 (H3K27) in human B-cell lymphomas', *Proceedings of the National Academy of Sciences of the United States of America*, 107(49), pp. 20980–20985. doi: 10.1073/pnas.1012525107.

Souroullas, G. P. *et al.* (2016) 'An oncogenic Ezh2 mutation induces tumors through global redistribution of histone 3 lysine 27 trimethylation', *Nature Medicine*. Nature Publishing Group, 22(6), pp. 632–640. doi: 10.1038/nm.4092.

Steunou, A.-L., Rossetto, D. and Côté, J. (2014) 'Regulating Chromatin by Histone Acetylation', in *Fundamentals of Chromatin*. New York, NY: Springer New York, pp. 147–212. doi: 10.1007/978-1-4614-8624-4_4.

Strutt, H., Cavalli, G. and Paro, R. (1997) 'Co-localization of Polycomb protein and GAGA factor on regulatory elements responsible for the maintenance of homeotic gene expression', *EMBO Journal*, 16(12), pp. 3621–3632. doi: 10.1093/emboj/16.12.3621.

- Strutt, H. and Paro, R. (1997) 'The polycomb group protein complex of *Drosophila melanogaster* has different compositions at different target genes', *Mol. Cell Biol.*, 17(12), pp. 6773–6783.
- Stucki, M. *et al.* (2005) 'MDC1 directly binds phosphorylated histone H2AX to regulate cellular responses to DNA double-strand breaks', *Cell*, 123(7), pp. 1213–1226. doi: 10.1016/j.cell.2005.09.038.
- Tachibana, M. *et al.* (2002) 'G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis', *Genes and Development*, 16(14), pp. 1779–1791. doi: 10.1101/gad.989402.
- Tanay, A. *et al.* (2007) 'Hyperconserved CpG domains underlie Polycomb-binding sites.', *Proceedings of the National Academy of Sciences of the United States of America*, 104(13), pp. 5521–5526. doi: 10.1073/pnas.0609746104.
- Taubert, S. *et al.* (2004) 'E2F-Dependent Histone Acetylation and Recruitment of the Tip60 Acetyltransferase Complex to Chromatin in Late G1', *Molecular and Cellular Biology*, 24(10), pp. 4546–4556. doi: 10.1128/mcb.24.10.4546-4556.2004.
- Thornton, S. R. *et al.* (2014) 'Polycomb Repressive Complex 2 Regulates Lineage Fidelity during Embryonic Stem Cell Differentiation', *PLoS ONE*. Edited by J. G. Knott, 9(10), p. e110498. doi: 10.1371/journal.pone.0110498.
- Tolstorukov, M. Y. *et al.* (2012) 'Histone Variant H2A.Bbd Is Associated with Active Transcription and mRNA Processing in Human Cells', *Molecular Cell*. Elsevier, 47(4), pp. 596–607. doi: 10.1016/j.molcel.2012.06.011.
- Torchy, M. P., Hamiche, A. and Klaholz, B. P. (2015) 'Structure and function insights into the NuRD chromatin remodeling complex', *Cellular and Molecular*

Life Sciences. Springer Basel, 72(13), pp. 2491–2507. doi: 10.1007/s00018-015-1880-8.

Umehara, T. *et al.* (2010) 'Structural basis for acetylated histone H4 recognition by the human BRD2 bromodomain', *Journal of Biological Chemistry*, 285(10), pp. 7610–7618. doi: 10.1074/jbc.M109.062422.

Varambally, S. *et al.* (2002) 'The polycomb group protein EZH2 is involved in progression of prostate cancer.', *Nature*, pp. 624–9. doi: 10.1038/nature01075.

Wang, H. *et al.* (2004) 'Role of histone H2A ubiquitination in Polycomb silencing', *Letters to Nature*, 431(October), pp. 873–878. doi: 10.1038/nature02966.1.

Wang, X., Paucek, R. D., *et al.* (2017a) 'Molecular analysis of PRC2 recruitment to DNA in chromatin and its inhibition by RNA.', *Nature structural & molecular biology*, 24(12), pp. 1028–1038. doi: 10.1038/nsmb.3487.

Wang, X., Paucek, R. D., *et al.* (2017b) 'Molecular analysis of PRC2 recruitment to DNA in chromatin and its inhibition by RNA', *Nature Structural & Molecular Biology*, 24(12), pp. 1028–1038. doi: 10.1038/nsmb.3487.

Wang, X., Goodrich, K. J., *et al.* (2017) 'Targeting of Polycomb Repressive Complex 2 to RNA by Short Repeats of Consecutive Guanines', *Molecular Cell*. Elsevier Inc., 65(6), pp. 1056-1067.e5. doi: 10.1016/j.molcel.2017.02.003.

Wang, Z. *et al.* (2018) 'A Non-canonical BCOR-PRC1.1 Complex Represses Differentiation Programs in Human ESCs', *Cell Stem Cell*. Elsevier Inc., 22(2), pp. 235-251.e9. doi: 10.1016/j.stem.2017.12.002.

Weber, C. M. and Henikoff, S. (2014) 'Histone variants_dynamic punctuation in transcription_Weber_Henikoff_GenDev_2014_review.p.pdf', pp. 672–682. doi: 10.1101/gad.238873.114.Freely.

Wilson, B. G. and Roberts, C. W. M. (2011) 'SWI/SNF nucleosome remodellers

and cancer', *Nature Reviews Cancer*, 11(7), pp. 481–492. doi: 10.1038/nrc3068.

Wu, H. *et al.* (2010) 'Dnmt3a-Dependent Nonpromoter DNA Methylation Facilitates Transcription of Neurogenic Genes', *Science*, 329(5990), pp. 444–448. doi: 10.1126/science.1190485.

Wu, H. and Zhang, Y. (2011) 'Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation', *Genes and Development*, 25(23), pp. 2436–2452. doi: 10.1101/gad.179184.111.

Wysocka, J. *et al.* (2006) 'A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling', *Nature*, 442(7098), pp. 86–90. doi: 10.1038/nature04815.

Yang, L. *et al.* (2014) 'Self-assembled FUS binds active chromatin and regulates gene transcription.', *Proceedings of the National Academy of Sciences of the United States of America*, 111(50), pp. 17809–14. doi: 10.1073/pnas.1414004111.

Yoder, J. (1998) 'A candidate mammalian DNA methyltransferase related to pmt1p of fission yeast', *Human Molecular Genetics*, 7(2), pp. 279–284. doi: 10.1093/hmg/7.2.279.

Youmans, D. T., Schmidt, J. C. and Cech, T. R. (2018) 'Live-cell imaging reveals the dynamics of PRC2 and recruitment to chromatin by SUZ12-associated subunits.', *Genes & development*, pp. 1–12. doi: 10.1101/gad.311936.118.

Yuan, W. *et al.* (2012) 'Dense chromatin activates polycomb repressive complex 2 to regulate H3 lysine 27 methylation', *Science*, 337(6097), pp. 971–975. doi: 10.1126/science.1225237.

Zhang, P. *et al.* (2006) 'Structure of human MRG15 chromo domain and its binding to Lys36-methylated histone H3', *Nucleic Acids Research*, 34(22), pp.

6621–6628. doi: 10.1093/nar/gkl989.

Zhang, Q. *et al.* (2019) 'RNA exploits an exposed regulatory site to inhibit the enzymatic activity of PRC2', *Nature Structural and Molecular Biology*, 26(3), pp. 237–247. doi: 10.1038/s41594-019-0197-y.

Zhang, R. *et al.* (2005) 'Formation of macroH2A-containing senescence-associated heterochromatin foci and senescence driven by ASF1a and HIRA', *Developmental Cell*, 8(1), pp. 19–30. doi: 10.1016/j.devcel.2004.10.019.

Zhang, T. *et al.* (2016) 'G9a/GLP Complex Maintains Imprinted DNA Methylation in Embryonic Stem Cells', *Cell Reports*, 15(1), pp. 77–85. doi: 10.1016/j.celrep.2016.03.007.

Zhang, Z. *et al.* (2011) 'PRC2 complexes with JARID2, MTF2, and esPRC2p48 in ES cells to modulate ES cell pluripotency and somatic cell reprogramming', *Stem Cells*, 29(2), pp. 229–240. doi: 10.1002/stem.578.

Zhao, Jing *et al.* (2008) 'Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome', *Science*, 322(5902), pp. 750–756. doi: 10.1126/science.1163045.

Zhao, J. *et al.* (2008) 'Polycomb Proteins Targeted by a Short Repeat RNA to the Mouse X Chromosome', *Science*, 322(5902), pp. 750–756. doi: 10.1126/science.1163045.

Zhao, J. *et al.* (2010) 'Genome-wide Identification of Polycomb-Associated RNAs by RIP-seq', *Molecular Cell*. Elsevier Inc., 40(6), pp. 939–953. doi: 10.1016/j.molcel.2010.12.011.

Zhou, J. *et al.* (2007) 'The nucleosome surface regulates chromatin compaction and couples it with transcriptional repression', *Nature Structural and Molecular Biology*, 14(11), pp. 1070–1076. doi: 10.1038/nsmb1323.

Zimmer, J. *et al.* (2016) 'Targeting BRCA1 and BRCA2 Deficiencies with G-Quadruplex-Interacting Compounds', *Molecular Cell*. The Authors, 61(3), pp. 449–460. doi: 10.1016/j.molcel.2015.12.004.

Zoabi, M. *et al.* (2014) 'RNA-dependent chromatin localization of KDM4D lysine demethylase promotes H3K9me3 demethylation', *Nucleic Acids Research*, 42(21), pp. 13026–13038. doi: 10.1093/nar/gku1021.