The function of PRC2 in MLL-AF4 Acute Lymphoid Leukaemia

Marta Cecylia Tapia

St Peter's College

Thesis submitted in partial fulfilment of the requirements

for the degree of Doctor of Philosophy

Radcliffe Department of Medicine

University of Oxford

MT 2015-10-09

Declaration of Authorship

I hereby acknowledge that this thesis is my own work, except where clearly indicated otherwise.

Abstract

The MLL-AF4 fusion protein is the most prevalent MLL rearrangement in acute lymphoid leukaemia. It is associated with a very poor prognosis with median overall survival of 10 months in children. Epigenetic therapy is a promising new approach for the treatment haematological malignancies. Predicting response to epigenetic therapy has been difficult, in part due to the lack of sufficient knowledge regarding the molecular mechanism by which these inhibitors exert their anti-leukaemic effects. In this study, inhibition of the PRC2 complex by the EZH2/EZH1 specific small molecule inhibitor UNC1999 in an MLL-AF4 cell line, resulted in a potent anti-leukaemic effect *in vitro*. Quantitative ChIP-seq (ChIP-Rx) revealed a dose-dependent decrease in H3K27me3 levels following UNC1999 treatment of an MLL-AF4 leukaemic cell line, which was accompanied by a dose-dependent increase in gene expression. A combination of ATACseq, ChIPseq, and Nascent-RNAseq identified a signature of PRC2 target genes that are susceptible to upregulation following inhibition of PRC2; these were genes that were marked by H3K4me3, bound by both RING1B and PRC2 and were expressed at low levels. Knowledge of a signature of genes responsive to PRC2 inhibition may help in predicting responses to epigenetic therapies, which target this complex.

Acknowledgments

I would like to thank my supervisor Dr. Thomas Milne for all his help and support during my PhD.

I would also like to thank all the members of the Milne Lab, especially Phill North for his guidance, humour and good banter. I would also like to thank Jon Kerry for writing many of the scripts used in the bioinformatics analysis.

Thank you to my collaborators, Prof. Jian Jin and members of his lab at Mount Sinai for kindly providing the UNC1999 and UNC2400 inhibitors. Thanks to Prof. Olaf |Heidenreich and Dr. Frida Ponthan at Newcastle University for the *in vivo* experiments.

A special thank you to Ghadeer Almuhaini for being a great friend. Thank you for all the fun times we shared in Oxford and for feeding me and keeping me sane especially towards the end of my Phd. A special thank you to my family for their unconditional support.

Table of Contents

Table of Figures

List of abbreviations

ALL	Acute lymphoid leukaemia
AML	Acute myeloid leukaemia
B-ALL	B-cell acute lymphoid leukaemia
bp	Base pairs
CBX	Chromobox homolog
CDB	ChIP dilution buffer
cDNA	Complementary DNA
ChIP	Chromatin immunoprecipitation
COMPASS	Complex proteins associated with SET1
CPM	Count per million
CTD	C-terminal repeat domain
DM	Double Marked (H3K4me3 + H3K27me3)
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
DSG	Disuccinimidyl glutarate
DTT	Dithiothreitol
EAP	ENL associated protein complex
EC50	Half maximal effective concentration
EDTA	Ethylenediaminetetraacetic acid
EED	Emrbyonic ectoderm development

ES	Embryonic stem cell
EZH1	Enhancer of zeste 1
EZH2	Enhancer of zeste 2
FA	Formaldehyde
FDR	False discovery rate
Fli1	Friend leukemia integration 1 transcription factor
FPKM	Fragments per kilobase millions
FYRC	FY rich domain c-terminal
FYRN	FyYrich domain n-terminal
GC	Glucocorticoids
GFP	Green fluorescent protein
GR	Glucocorticoid receptor
HAT	Histone acetyl transferase
HDAC	Histone deacetylase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HSC	Haematopoietic stem cell
HSPC	Haematopoietic stem and progenitor cell
JMJD	Jumonji domain
KDM	Lysine (k)-specific demethylase
KMT	Lysine (k) - specific methyltransferase
Kn	Knock-down
KO	Knock-out
MDS	Myelodysplastic syndrome

mES	Mouse embryonic stem cell
MLL	Mixed lineage leukaemia (KMT2A)
MLL-FP	Mixed lineage leukaemia fusion protein
MLL2	Mixed lineage leukaemia 2 (KMT2B)
MLLr	Mixed lineage rearranged leukaemia
MOPS	3-(n-morpholino)propanesulfonic acid
mRNA	Messenger RNA
NGS	Next generation sequencing
NuRD	Nucleosome remodelling and deacetylation complex
P-TEFb	Positive transcription elongation factor b
PAF1C	RNA Polymerase associated factor 1 complex
PBS	Phosphate-buffered saline
PcG	Polycomb group
PCGF	Polycomb group ring finger protein
Pcl	Polycomb like (PHF)
PH	Polyhomeotic
PHD	Plant homeodomain
PHF	Phd finger protein (pcl)
PIC	Pre-inititaion complex
PRC	Polycomb repressive complex
PTM	Post translational modification
RING1	Ring finger protein 1
RIPA	Radioimmunoprecipitation assay buffer

RNA	Ribonucleic acid
RNAP2	CTD serine 2 phosphorylated RNA polymerase II
RNAP5	CTD serine 5 phosphorylated RNA polymerase II
RNAPII	RNA polymerase II
RRPM	Reference normalised reads per ten million
rt-qPCR	Real-time quantitative PCR
SAM	Sterile alpha motif
SD	Standard deviation
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Super elongation complex
seq	Sequencing
SET	Su(var), enhnacer of zeste, trithorax
shRNA	Short hairpin RNA
siRNA	Small interfering RNA
SUZ12	Suppressor of zeste 12
TE	Tris-EDTA buffer
TF	Transcription factor
TrxG	Trithorax group
TSS	Transcription start site
UN	Unchanged
UP	Upregulated
UTX	Ubiquitously transcribed tetratricopeptide repeat, x

	chromosome
WB	Western blotting
XCI	X chromosome inactivation
XIST	X inactive specific transcript

1 Chapter 1 Introduction

1.1 MLL is an epigenetic protein deregulated in cancer

Acute leukaemia with 11q23 abnormalities involving the *mixed-lineage leukaemia* (*MLL*) gene, comprises one category of recurring genetic abnormalities as defined by the WHO classification (Vardiman et al., 2009). Chromosomal translocations involving *MLL* lead to the production of chimeric proteins where the amino-terminal portion of the *MLL* gene is fused to the carboxy-terminal portion of a fusion partner gene (Meyer et al., 2009). The novel MLL-fusion proteins (MLL-FPs) created are potent oncogenes that function as positive regulators of gene expression, leading to the development of both acute myeloid leukaemia (AML) and acute lymphoid leukaemia (ALL).

MLL is an epigenetic protein belonging to the trithorax group (TrxG) of epigenetic regulators. TrxG proteins function antagonistically to Polycomb group (PcG) proteins, as positive regulators of gene expression in early embryonic development and haematopoiesis (Jude et al., 2007; McMahon et al., 2007). Epigenetic regulation is crucial for the correct temporal and spatial gene expression patterns necessary for proper development and cell differentiation, and while the term can be used in multiple ways, for the purpose of this thesis it will be defined as "the structural adaptation of chromosomal regions so as to register, signal or perpetuate altered activity states" (Bird, 2007). Epigenetic events that modulate gene expression include: DNA methylation, ATP-dependent

chromatin remodeling, RNA-interference and important for this thesis, post-translational modifications of histone proteins (PTMs).

The enzymatic binding pockets of chromatin modifying proteins can be targeted with small molecule inhibitors, which has important therapeutic implications. Driven by the therapeutic success of DNA methyltransferase (DNMT) inhibitors and histone deacetylase (HDAC) inhibitors for the treatment of myelodysplastic syndrome (MDS) and T-cell lymphoma patients respectively (Fenaux et al., 2009; Issa and Kantarjian, 2009; Khan and La Thangue, 2012; Silverman and Mufti, 2005; Wagner et al., 2010), there has recently been an explosion in the development of highly specific small molecule inhibitors that target chromatin proteins. Some of these are in clinical trials for the treatment of both haematological malignancies and solid tumours. Despite the early successes of epigenetic therapy the exact molecular mechanism by which they function are not fully understood. Unravelling these will be beneficial for predicting patient responses as well as guiding rational therapeutic approaches for combination therapy with traditional cancer therapeutic agents.

1.2 Nucleosomes and higher order chromatin structure

In humans, single contiguous DNA molecules consisting of up to 250 million base pairs needs to be compacted by four orders of magnitude in order to fit into a micrometer-size nucleus. Compaction of DNA is achieved by wrapping ~145-147 bp of DNA around a histone octamer core to form the basic building block of

chromatin called the nucleosome (Luger et al., 1997). The octamer core consists of two copies of histones H2A, H2B, H3 and H4 all of which share a structural histone fold domain composed of 3 α-helices separated by two loops which allow for heterodimerization of H2A with H2B and H3 with H4 (Kornberg, 1974; Luger et al., 1997). The histone pairs are then brought together by the interactions of H3 α-helices with each other and those of H2B and H4 (Luger et al., 1997; Smith and Stillman, 1991). Arrays of nucleosomes spaced at ~10-60bp intervals form 10ηm chromatin fibres termed "beads-on-a string" which further compact into 30ηm chromatin fibres (Robinson and Rhodes, 2006). Early studies showed that the 30ηm structure was dependent on the linker histone H1 which organises a further 20bp of DNA completing the nucleosome (Marsden and Laemmli, 1979; Zhou et al., 1998). Two main models for the structure of the 30ηm fibres prevail, a two-start helix and a one-start solenoid although the universal existence of these in all cell types *in vivo* is debated and a new model proposes that chromosome condensation occurs through packaging of the 10ηm chromatin fibres in a fractal manner (Hansen, 2012; Nishino et al., 2012).

The function of chromatin extends beyond compaction of DNA and is tightly linked to the control of gene expression. Early *in vitro* studies showed that nucleosomes impede transcription initiation (Knezetic and Luse, 1986) but once initiated (from a nucleosome free region) elongation could proceed across nucleosomes which became displaced (Lorch et al., 1987). Nucleosomes have now been established as regulatory units and core histones can be replaced by histone variants affecting their stability and thus transcription. Importantly,

histones can be post-translationally modified either at their structured globular domains or at their N-terminal flexible unstructured tails, which protrude from the surface of the nucleosome.

1.3 Post-translational histone modifications

Methylation and acetylation of histone tails were the first reported post-translational modifications (PTMs) of histones (Allfrey et al., 1964). Since then histones have been shown to be ubiquitinated, biotinylated, sumoylated ADP-ribosylated, phosphorylated among others.

Genome wide analysis of PTMs has revealed that these marks are not uniformly distributed and are instead associated with specific cis-regulatory elements. H3K4me3 is associated with gene promoters (Guenther et al., 2007; Mikkelsen et al., 2007), H3K36me3 is associated with actively elongating genes (Mikkelsen et al., 2007), H3K27me3 is associated with repressed genes (Peters et al. 2003), while high levels of H3K4me1 with low levels of H3K4me3 mark enhancers which can be further separated into active or poised state by the presence of H3K27Ac (Creyghton et al., 2010).

1.3.1 Histone lysine acetylation

Histone lysine acetylation is associated with active gene transcription (Allfrey et al., 1964; Hebbes et al., 1988; Wang et al., 2008) as well as with regions of "open" chromatin. Histone acetylation is mediated by three main families of histone acetyl transferases (HATs): GNAT, MYST and CBP/p300. Although these all function by transferring the acetyl group from the acetyl-CoA cofactor to the ε-amino group of the lysine side chain, they do so by distinct mechanisms.

As mentioned previously, histone acetylation neutralises the positive charge of the lysine residues, thus reducing the affinity of DNA for histones (Hong et al., 1993). This results in a more accessible chromatin architecture which is thought to aid transcription by allowing for binding of transcription factors and the transcriptional apparatus to DNA (Dion et al., 2005; Megee et al., 1995). Acetylation marks can also act as binding sites for bromodomain containing proteins. Human bromodomain containing proteins can be clustered into 8 families (based on sequence length and similarity of the bromodomains) (Marmorstein and Zhou, 2014) . Bromodomain containing proteins include the HATs, nucleosome remodelling complexes as well as the BET proteins (BRD2, BRD3, BRD4). The latter play important roles in the promotion of transcriptional elongation through recruitment of P-TEFb, while nucleosome remodelling complexes alter the structure of chromatin and facilitate binding of important transcription activators.

Acetylation can also contribute to the active state of a gene by antagonising repressive complexes. This is the case for H3K27Ac deposited by CBP/300 (Pasini et al., 2010b; Tie et al., 2009) which is inversely correlated with Polycomb Repressive Complex 2 (PRC2) occupancy on chromatin (Pasini et al., 2010b), a complex associated with repressed genes (discussed further in section 1.6). Binding and repression of a subset of PRC2 gene targets was shown to require the activity of nucleosome remodelling and deacetylation complex (NuRD), particularly its deacetylase activity. Thus the removal of histone acetylation is related to gene repression. Removal of the acetyl mark is mediated by histone deacetylases (HDACs). This family of proteins is of particular interest, as they represent the first targets of epigenetic therapy to have reached clinical use (Richon et al., 1998).

1.3.2 Histone lysine methylation

1.3.2.1 H3K4me3

H3K4me3 has been shown to be present at virtually all CpG promoters irrespective of their transcriptional status (Guenther et al., 2007). Nevertheless, overall levels of H3K4me3 are positively correlated with levels of transcription and the mark is enriched at actively transcribed genes. In yeast the sole H3K4 methyltransferase, SET1 was purified as part of a complex termed COMPASS (Complex Proteins Associated with SET1) (Miller et al., 2001; Roguev et al., 2001) and catalyses the addition of the mono-, di- and tri-methyl marks (Briggs

et al., 2001; Nagy et al., 2002; Roguev et al., 2001). In mammals COMPASS-like complexes containing SET1A/SET1B are responsible for bulk H3K4 methylation while COMPASS-like complexes of MLL1-4 are important for methylation at specific target loci.

H3K4me3 has been shown to act as a binding site for the PHD domain containing TAF3 subunit of the basal transcription factor TFIID (Vermeulen et al., 2007). This interaction was shown to be important for assembly of the Pre-Initiation Complex (PIC) and expression of a subset of genes (Lauberth et al., 2013). Additionally, H3K4me3 has been shown to act as a binding site for ATP-dependent chromatin remodellers. The BPTF subunit of the NURF chromatin-remodelling complexes contains a PHD finger as well as a Bromodomain at its C-terminal. The PHD finger was shown to bind H3K4me3 and this interaction was necessary for expression of the *HOXC8* promoter as well as for the rescue of phenotypes associated with knockdowns of BPTF in *Xenopus laevis* embryos (Wysocka et al., 2006). Interestingly mutations affecting both the PHD finger and bromodomain of BPTF were less efficient at rescuing the BPTF knockdown (Wysocka et al., 2006) and it was subsequently shown that BPTF binding to chromatin is higher in the presence of both H3K4me3 and H3K16ac (Ruthenburg et al., 2011). The mark was also found to stabilise the interaction of CHD1 (previously linked to transcriptional elongation and deposition of the H3.3 variant) with chromatin *in vitro* with recruitment attributed to the Mediator complex. Thus H3K4me3 serves as one of the contact points of various effector

proteins as well as the transcriptional machinery to chromatin, promoting stable binding and transcription.

H3K4me3 has also been found to promote association of HATs such as NuA3, SAGA as well as the H3K9me3 demethylase JMJD2A further contributing to the active state of genes. It also serves to antagonise binding of repressive complexes like the nucleosome remodelling and deacetylase complex (NuRD) (Musselman et al., 2009) and the DNA methyltransferases DNMT3A and DNMT3B (Ooi et al., 2007). *In vitro* experiments have also shown it to antagonise binding of PRC2 and deposition of H3K27me3 in *cis* (Schmitges et al., 2011).

Of note, MLL can bind to H3K4me3 through its PHD finger and SET1A/B can bind this mark through association with CFP1. Loss of this binding leads to reduced H3K4me3 suggesting a feed forward mechanism.

1.3.2.2 H3K36me3

Enrichment of H3K36me3 along the gene body has been found to increase toward the 3' ends of genes and to correlate with active transcription (Miao and Natarajan, 2005; Mikkelsen et al., 2007; Pokholok et al., 2005; Rao et al., 2005). In yeast, SET2 methylates all forms of H3K36 (Strahl et al., 2002) while in humans, 8 proteins with either *in vitro* or *in vivo* H3K36 methyltransferase activity have been identified (Wagner and Carpenter, 2012). Yeast SET2 and Human SETD2 have been found to interact with the elongating form of RNAPII

and deposit H3K36me3 during transcription, however the exact mechanism by which H3K36me3 contributes to transcriptional elongation has not been defined. While some studies have found that abrogation of this interaction in yeast abolishes H3K36me3 and leads to decreased PolII occupancy another study did not see any changes in transcription and PolII occupancy upon deletion of SETD2.

Interestingly while H3K36 methylation has been shown to inhibit H3K27 methylation by PRC2 *in vitro* (Yuan et al., 2011), recent studies have found that PRC2 associated proteins PHF1 and PHF19 can bind H36me2/3 through their Tudor domain (Ballare et al., 2012; Cai et al., 2013). This binding was shown to be important for targeting of PRC2 to chromatin as well as spreading of the H3K27me3 into gene bodies (Cai et al., 2013). The deposition of H3K27me3 upon gene repression coincided with loss of H3K36me3 and the authors stated that this may be a result of PHF19 dependent recruitment of the H3K36me3 demethylase KDM2B, although the data for this was not shown. This is in contrast to studies on the TrxG protein Absent, small or homeotic discs 1-like (ASH1L) which was shown to methylate H3K36me2/me3 in a transcriptionally independent manner and antagonise binding of PRC1 complex components (Miyazaki et al., 2013).

1.3.2.3 H3K27 methylation

The PRC2 complex mediates the deposition of a mono- di- or tri-methyl group to lysine 27 of histone 3 (H3K27me1/2/3). *In vitro* studies have shown that un-

methylated and mono-methylated H3K27 are better substrates for PRC2 (McCabe et al., 2012a). In ES cells 15 % of H3 is trimethylated at lysine 27, 50% is dimethylated and 15% is monomethylated (Peters et al., 2003). This makes H3K27me2 the main catalytic output of the PRC2 complex. The three marks form spatially defined genomic domains. H3K27me1 has been found enriched at sites of constitutive heterochromatin as well as in the body of actively transcribed genes while H3K27me2 is more broadly distributed and is found in intra-genic and inter-genic regions (Ferrari et al., 2014). The function of H3K27me2 is not fully elucidated although it may serve to antagonise H3K27Ac at enhancers (Ferrari et al., 2014). Because H3K27me3 is found to be enriched at TSS of PRC2 repressed genes, and unlike H3K27me1 and H3K27me2 co-localises with members of the PRC2 complex, most research has focus on the study of this mark, although recent findings point to important roles for H3K27me1/2 as well.

H3K27me3 enrichment has been found to correlate with gene silencing (Peters et al., 2003). In mammals H3K27me3 can form large domains of more than 100kb, or smaller domains of a few kilo-base pairs (Boyer et al., 2006; Hawkins et al., 2010; Mikkelsen et al., 2007; Zhao et al., 2007). ES cells have smaller H3K27me3 domains than more differentiated cell types perhaps suggestive of an expansion of the silencing mark during differentiation (Hawkins et al., 2010). Enrichment levels of H3K27me3 peak just downstream of TSS (Pan et al., 2007; Zhao et al., 2007), and are associated with developmental genes (Bracken et al., 2006; Lee et al., 2006). Recently, analysis carried out by the Roadmap Epigenomics Consortium comparing 127 epigenomes revealed that repressed

states of chromatin marked by H3K27me3 were highly lineage specific in contrast to actively transcribed regions marked by H3K4me3 for example. Furthermore, five distinct cancer cell lines (distinct lineages) were shown to cluster together based on their H3K27me3 profiles (representing all cancer cell lines analysed). This provides further evidence of the importance of H3K27me3 in development and suggest that despite the heterogeneity of cancers a common mechanism mediated by PRC2 repression may exist (Kundaje et al., 2015).

Two H3K27me2/3 demethylases have been identified to date, the Ubiquitosuly Transcribed Tetratricopeptide Repeat on chromosome X (UTX/KDM6A) and Jumonji D3 (JMJD3/KDM6A) (Agger et al., 2007; De Santa et al., 2007; Lan et al., 2007; Lee et al., 2007). UTX interacts with the MLL2 H3K4me3 methyltransferase complex (Cho et al., 2007; Issaeva et al., 2007), and has been shown to be important for embryonic development as well as haematopoiesis (reviewed in (Van der Meulen et al., 2014) . Its interaction with MLL2 points to an important role in removing H3K27me3 for proper expression of genes. While UTX is expressed ubiquitously, JMJD3 has been shown to be induced in macrophages following lipopolysaccharide stimulation (De Santa et al., 2007; 2009), and this results in the demethylation and activation of inflammation responsive genes (De Santa et al., 2009)

1.3.2.4 Bivalent domains

Bivalent domains are chromatin regions that are marked by H3K27me3 and H3K4me3 and have been found in stem cells and differentiated cells (Azuara et al., 2006; Bernstein et al., 2006; Mikkelsen et al., 2007; Mohn et al., 2008; Pan et al., 2007; Zhao et al., 2007). Genome-wide analysis of H3K4me3 and H3K27me3 in mouse embryonic stem cells (mES) showed that 22% of H3K4me3 marked promoters were bivalent and were expressed at low levels, while 4% of CpG promoters were bivalent in mouse embryonic fibroblasts (MEFs) (Mikkelsen et al., 2007). Upon differentiation of mES, bivalent domains can resolve to monovalent states or maintain their bivalency while new bivalent domains are also established (Mohn et al., 2008). Sequential ChIP on mononucleosomes as well as mononuclesome ChIP combined with mass spectrometry has demonstrated the existence of these modifications on the same nucleosome (Azuara et al., 2006; Seenundun et al., 2010; Voigt et al., 2012). Bivalent domains seem to represent asymmetrical methylation of the two copies of histone 3 that form part of the nucleosome (Voigt et al., 2012). While previous research had shown inhibition of the methyltransferase activity of PRC2 by H3K4me3 *in vitro*, Voigt et al. (2012) extended this research to show that this inhibition is dependent on the modifications being present on the same histone tail and that PRC2 is able to methylate H3K27 assymetrically to H3K4me3. Despite evidence for the existence of true bivalency, most published work relies on ChIP approaches to identify these regions and it is important to bear in mind that

these regions may result from the heterogeneity of cell populations or allele specific modifications.

In ES cells bivalent domain are enriched for TFs and developmental genes. They have been proposed to prime genes important for differentiation by allowing their rapid induction (Voigt et al., 2013). However, recent findings have identified MLL2 as the methyltransferase responsible for H3K4me3 at bivalent genes (Denissov et al., 2014; Hu et al., 2013). By inducing differentiation in ES through retinoic acid treatment of MLL2 null cells, two groups showed that despite the loss of bivalency, rapid induction of transcription was not affected at all bivalent genes (Denissov et al., 2014; Hu et al., 2013). While half of all the genes that lost bivalency in MLL2 null cells showed impaired induction upon all trans- retinoic acid (ATRA) mediated differentiation, the other half were not significantly different to their wild-type counterparts. This suggested that other methyltransferases are responsible for correct expression of bivalent genes upon differentiation. MLL was found to be responsible for proper induction of 39% of genes that showed a normal response (Denissov et al., 2014). These findings argue against the idea that bivalent genes are poised for rapid activation, as does the observation that some bivalent domains resolve to a repressed H3K27me3 monovalent state. Recently a responsive model for PcG recruitment has been proposed (Klose et al., 2013). According to this model, bivalency is the result of both TrxG and PcG sampling the same loci, unable to establish a monovalent state as a result of the presence of activating signals that on the one hand inhibit PcG domain formation, but are not strong enough to establish a TrxG domain.

1.3.2.5 CpG islands

In mammals, approximately 70-80% of CpG dinucleotides are methylated at the 5 position of the cytosine pyrimidine ring (Ehrlich et al., 1982). Methylation of CpGs at promoters is associated with gene repression and is catalysed by DNA methyltransferases (DNMTs). DNMT3a and DNMT3b are *de novo* methyltransferases and are the main DNMTs at previously unmethylated CpGs while DNMT1 catalyses DNA-methylation during DNA replication of methylated strands. DNA methylation may contribute to repression of genes by inhibiting binding of transcription factors to their cognate sites as well as through the recruitment of repressive complexes that possess methyl-binding domains (MBD).

Roughly 70% of mammalian gene promoters are associated with so called CpG islands (Saxonov et al., 2006). CpG islands are on average 1kb long, have an above average CpG content, are associated with regions of reduced nucleosome occupancy, and importantly lack DNA methylation (Bird et al., 1985; Fenouil et al., 2012). Non-methylated DNA islands (CpG island defined through an experimental approach rather than through a computational one), have recently been found to be a conserved feature of vertebrate promoters supporting a functional role of these elements in gene regulation (Long et al., 2013).

CpG islands have been proposed to act as a platform for the binding of proteins that modulate the chromatin state surrounding the island (Blackledge and Klose,

2011). Specifically, proteins containing a ZF-CXXC domain can bind to non-methylated CpG dinucleotides. Two such proteins, CFP1 (associated with the SETD1 H3K4 methyltransferase complex) and KDM2A (a H3K36me2 demethylase) have been found to be associated with most unmethylated CpG dinucleotides *in vivo.* Consequently, most CpG islands are found in a "permissive" state for transcription which is thought to allow both activating and repressive complexes to bind and modulate transcriptional output (Blackledge and Klose, 2011). Other ZF-CXXC containing proteins like MLL for example, are found at a subset of CpG island promoters, and are associated with active gene transcription, while repressed states are associated with binding of PcG proteins in part through their association with the ZF-CXXC domain containing KDM2B.

1.4 PcG and TrxG proteins

The expression patterns of homeotic genes established during early embryogenesis, control segment identity along the antero-posterior axis. PcG and TrxG proteins were discovered in *D.melanogaster* as trans-acting factors that either maintain the repressed (PcG) or active (TrxG) state of homeotic genes. The first PcG protein to be discovered was *Polycomb (Pc),* where a dominant mutation in the gene resulted in ectopic sex combs on the second and third legs of adult male flies as a result of the overexpression of the homeotic gene *Sex combs reduces (Scr)* (Lewis, 1978). TrxG proteins were subsequently defined by their ability to maintain homeotic gene expression and counteract the role of PcG proteins. The first TrxG protein identified was *trithorax* (*trx*). Mutations in *trx* led

to the partial transformation of halters to wings as well as transformation of posterior abdominal segments to anterior ones (Ingham, 1983; 1998). These transformations were linked to decreased expression of *Ubx* as well as *abdA* and *AbdB* respectively (Ingham, 1983; 1998) .

Since their discovery, TrxG and PcG proteins have been implicated in the control of genes belonging to multiple developmental pathways in addition to the regulation of *Hox* genes (Hanson et al., 1999; Jude et al., 2007; Lim et al., 2009; McMahon et al., 2007; Yu et al., 1998; 1995). Due to their important developmental roles it is perhaps not surprising that deregulation of these factors are associated with multiple forms of cancer.

1.5 MLL and leukaemia

The cells that make up the haematopoietic system are derived from a single multipotent progenitor cell type, the haematopoietic stem cell (HSC). Differentiation of the HSC to mature blood cells is a step-wise process of lineage commitment which is mediated by the coordinate expression of lineage specific transcription factors and the epigenetic landscape of the cell (Rice et al., 2007; Zhu and Emerson, 2002). The hallmark of leukaemic transformation is a block in terminal differentiation of haematopoietic cells. TFs and epigenetic proteins both function to regulate transcriptional programs and mutations in these proteins, have been found to either drive or contribute to the development of multiple haematological malignancies.

In the early 1990s, cloning of the leukaemia associated translocation breakpoint at 11q23 led to the discovery of the *MLL* gene (Djabali et al., 1992; Gu et al., 1992; Tkachuk et al., 1992; Zieminvanderpoel et al., 1991). Sequence comparison revealed the gene to be a homolog of the *Drosophila melanogaster trithorax* gene (*Trx*) (Gu et al., 1992; McCabe et al., 1992; Tkachuk et al., 1992) already known to be a regulator of homeotic gene expression (Ingham, 1985) and antagonist of the function of Polycomb (Kennison and Tamkun, 1988).

MLL rearrangments are associated with over 70% of childhood acute lymphopblasitc (ALL) and 5-10% of adult acute myeloid leukaemia (AML) and ALL cases (Marschalek, 2010). A small portion of MLL leukaemias is accounted for by small internal deletions, partial tandem duplications (MLL-PTD) and whole gene duplications of MLL. 90% of all MLL leukaemia cases involve in frame translocation of the *MLL* gene fusing it to either *AF9, ENL, ELL, AF10, AF6* or *AF4* (Meyer et al., 2013). These are associated with acute myeloid leukaemia (AML), acute lymphoid leukaemia (ALL) and biphenotypic leukaemias and a generally poor prognosis (Meyer et al., 2013).

1.5.1 Role of MLL in haematopoiesis

Homozygous deletions of *MLL* result in embryonic lethality of mice between embryonic day 11.5 and 14.5 (Yagi et al., 1998; Yu et al., 1995). Haematopoietic cells were found to be reduced in the fetal livers of *MLL* null embryos and showed retarded growth in colony forming assays. This suggested a role for MLL

in proliferation of HSC precursors (Yagi et al., 1998). Heterozygous mutations of *MLL* led to retarded growth of mice, bidirectional homeotic transformations of the axial skeleton as well as sternal malformations. These mice were anaemic with decreased platelet numbers and consistently lower numbers of B-cells compared to wild-type mice (Yu et al., 1995). Conditional deletion of *MLL* in the haematopoietic compartment showed that MLL is not necessary for normal steady-state haematopoiesis or for the differentiation of haematopoietic cells in adult mice. Competitive transplantation of $Mll^{-/-}$ with $MLL^{+/+}$ HSCs into lethally irradiated recipient mice revealed a severe defect in repopulating potential of $Mll^{-/-}$ HSC which was not a result of defective homing (McMahon et al., 2007), indicating a role for MLL in self-renewal of HSC. The function of MLL in HSC maintenance was recently shown to be independent of its methyltransferase activity (Mishra et al., 2014) and relies instead on its association with the acetyltransferase MOF (Mishra et al., 2014).

1.5.2 The wild-type MLL protein complex

MLL contains three AT-hooks that can bind DNA (Zeleznik-Le et al., 1994), a CXXC domain (Allen et al., 2006; Birke et al., 2002; Cierpicki et al., 2010; Xu et al., 2011), four PHD fingers (Fair et al., 2001; Hom et al., 2010; Park et al., 2010; Wang et al., 2010), an atypical bromodomain (Wang et al., 2010), a FYRN and a FYNC domain as well as a SET domain at its C-terminus which mediates the methylation of H3K4 (Milne et al., 2002; Nakamura et al., 2002). MLL is cleaved

by Taspase 1 into a 320kDa N-terminal (MLL-N) and an 182kDa C-terminal (MLL-C) fragment, which are able to dimerise through the FYRN and FYRC domains (Dou et al., 2005; Hsieh et al., 2003; Nakamura et al., 2002; Yokoyama et al., 2004). MLL interacts with three additional components which are important for its methyltransferase activity: WD repeat-containing protein 5 (WDR5) (Wysocka et al., 2005), ASHL2 and retinoblastoma binding protein 5 (RBBP5) (Yokoyama et al., 2004). WDR5 can bind H3K4me2 while RBBP5 stabilises the interaction between MLL and WDR5 and ASHL1 Wysocka:2005iu}(Yokoyama et al., 2004). Biochemical assays have also identified Menin (MEN1) as a component of this complex (Hughes et al., 2004; Yokoyama et al., 2005). In addition to these components the MLL complex interacts with the acetyltransferase MOF through its C-terminal region (Dou et al., 2005). Acetylation of H4K16 by MOF can stimulate the methyltransferase ability of MLL. The acetyltransferase CBP is also associated with the complex (Ernst et al., 2001) and as mentioned previously, may serve to antagonise PRC2 mediated H3K27 methylation. Because WDR5 as well as MLL can bind to H3K4me2 it has been proposed that once methylation is initiated, these interactions result in further methylation of H3K4 in a positive-feedback loop (Dou et al., 2006; Wysocka et al., 2005).

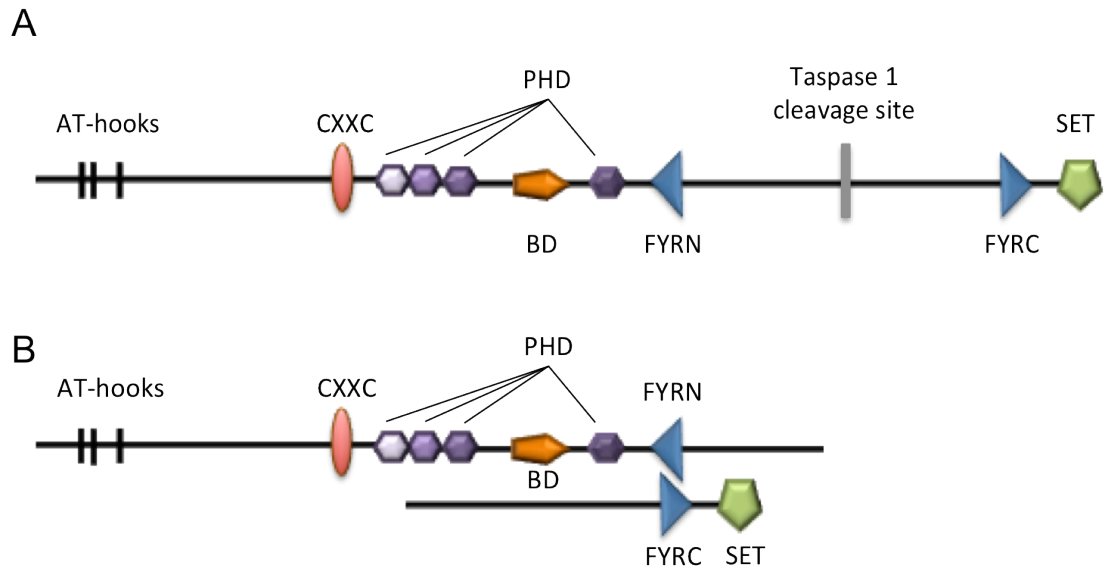


Figure 1 Wild-type MLL.

Wild-type MLL has three AT-hooks (black bars), a CXXC domain (red oval), four PHD fingers (purple hexagons), an atypical bromodomain (orange pentagon), a FYRN and FYRC domain (blue tirangles) and a SET domain (green pentagon). Taspase 1 cleaves MLL into two fragments which can dimerise through their FYRN and FYRC domains.

While the methyltransferase activity of MLL was shown to be important for the expression of *HOX* genes in mouse embryo fibroblasts (MEFs) and during embryogenesis (Terranova et al., 2006) recently it was shown to be dispensable for the expression of *HOXA9*, *Evi1* and *PRMD16* in mouse HSCs (Mishra et al., 2014). The expression of these targets was instead found to rely on the acetylation of H4K16 by MOF (Mishra et al., 2014). While deletions of *MLL* in HSPCs affected gene expression dramatically, these changes were not accompanied by reduction of H3K4me3, and therefore the exact mechanism by which MLL maintains expression of its target genes remains to be determined (Mishra et al., 2014).

1.5.3 Recruitment of MLL

MLL interacts with MEN1 and lens epithelium-derived growth factor (LEDGF) through its N-terminus (Hughes et al., 2004). The function of MEN1 is to bridge the interaction with LEDGF, which contains a PWWP domain. The PWWP domain has been proposed to mediate binding to H3K36me2/3 and thus recruit MLL to its targets, but the specificity of this interaction is controversial. Importantly a minimal domain required for the recruitment of MLL to *HOXA9* does not include the interaction with MEN1 and is instead mediated by the CXXC domain and the third PHD finger (Milne et al., 2010). The CXXC domain can bind to un-methylated CpG dinucleotides as well as mediate interaction with the polymerase associated factor 1 elongation complex (PAF1C). Point mutations that specifically disrupt interactions with either DNA or PAF1C reduced binding of MLL substantially but complete loss of binding was only observed when point mutations that abrogate binding of MLL to DNA, PAF1C and H3K4me2/3 mediated through the third PHD finger were introduced simultaneously (Milne et al., 2010). Immunofluorescence experiments showed that PAF1C and MLL only co-localise at a few loci, indicating that recruitment of MLL may vary at different gene targets. Importantly these experiments demonstrate that multiple interactions are necessary for recruitment/stable binding of MLL (Milne et al., 2010).

1.5.4 MLL-FPs

Translocation events leading to the formation of MLL-FPs predominantly occur over an 8.3kb region of the gene upstream of the first PHD finger (Stanulla et al., 2001; Thirman et al., 1993), and thus MLL-FPs only retain the AT-hooks and CXXC domains of the wild-type MLL protein, and do not possess any H3K4 methyltransferase activity. The retention of certain domains and loss of others has proven to be crucial for MLL mediated leukaemogenesis. MLL-FPs maintain the CXXC domain, which mediate interaction with PAF1C, but unlike wild-type MLL, disruption of the DNA binding ability of the CXXC domain abrogates recruitment of MLL-FPs completely. Interestingly, introduction of the PHD motifs blocks leukemic transformation by MLL-AF9 and MLL-ENL (Chen et al., 2008; Muntean et al., 2008).

Retention of the N-terminal Menin interacting domain is also important as mutations within this domain prevent development of leukaemia in mice (Caslini et al., 2007; Yokoyama et al., 2005). The interaction with menin is important for upregulation of key targets such as *HOXA9* and *MEIS1* and for proliferation of MLL-FP transformed cells (Chen et al., 2006b; Yokoyama et al., 2005). As a result novel inhibitors targeting the interaction between Menin, LEDGF and MLL-FPs have been developed and these have shown to inhibit the transforming ability of MLL-AF9 as wells as the proliferative ability of MLL-FP leukaemic cell lines (Cermakova et al., 2014; Grembecka et al., 2012).

While studies have indicated a need for wild-type MLL in the development of MLL-FP leukaemias (Thiel et al., 2010), and a role for it in recruitment of MLL-FPs to specific gene targets (Milne et al., 2002), it was possible that loss of the SET domain in MLL-FPs was compensated for by wild-type MLL. Consistent with this, pharmacological inhibition of the catalytic activity of MLL was detrimental to MLL-FP leukaemogenesis through induction of cell cycle arrest and apoptosis (Cao et al., 2014a). However, in another study, deletion of the SET domain of MLL was shown to be dispensable for MLL-mediated leukaemia (Mishra et al., 2014). It is possible that pharmacological inhibition of the catalytic activity of MLL may have unspecific functions when used *in vivo,* providing a possible explanation for the discrepancy of these results.

1.5.4.1 Gene targets

Multiple groups have shown that MLL-leukaemias define a distinct group of leukaemias based on their gene expression profiles. Expression of *HOXA* cluster genes is often elevated in MLL-FP leukaemias, particularly *HOXA9* as wells as its co-factor *MEIS1*. However, distinct gene expression signatures exist for different MLL-FP proteins (Stam et al., 2010) in line with the minimal gene target overlap of MLL-AF4 and MLL-AF9 as determined through ChIPseq experiments (Bernt et al., 2011; Guenther et al., 2008). Additionally, despite the strong association of MLL-leukaemias with *HOXA9* expression, roughly 50% of infant ALL patients expressing MLL-AF4 do not express *HOXA9* and define a sub-group of patients with an extreme high-risk of relapse (Stam et al., 2010). These findings are in

line with experiments showing that in an MLL-AF9 knock–in mouse model leukaemia development proceeded in a *HOXA9* null background.

Many other direct targets of MLL-FPs have been shown to be important for MLL leukaemic growth. C-myc for instance was shown to be necessary for the transformation of haematopoietic progenitors by MLL-AF9 (Schreiner et al., 2001) and the down regulation of *c-myc* was proposed as the mechanism by which inhibition of BRD4 leads to inhibition of MLL-leukaemias (Dawson et al., 2011; Zuber et al., 2011b). Other direct MLL-FPs targets shown to be important for MLL-mediated leukaemogenesis include *RUNX1* (Wilkinson et al., 2013)*, MYB* (Hess et al., 2006) *, MEF2C* (Krivtsov et al., 2006) *, Eya1, Six1* (Wang et al., 2011).

1.5.4.2 Molecular Mechanism of Transformation

Transcription in eukaryotes can be broadly divided into three stages: inititation, elongation and termination. The main regulatory step was traditionally thought to occur during the initiation phase of transcription with the assembly of the pre-initiation complex (PIC). Since than it has been shown that regulation of the elongation step is also important and many developmental genes are controlled by release of paused Pol II. It is currently estimated that at least half of all developmental genes are controlled at the level of transcriptional elongation (Levine, 2011) including the *Hox* genes (Chopra et al., 2009; Muse et al., 2007; Zeitlinger et al., 2007).

Following PIC assembly and promoter clearance, Pol II pauses ~30-50 nucleotides downstream of the transcription start site (TSS). Release to productive elongation involves the phosphorylation of the C-terminal domain (CTD) of RNAPII on serine 2 as well as the phosphorylation of Spt5, a subunit of DSIF by P-TEFB leading to the dissociation of the negative elongation factor (NELF). P-TEFb has been found to be in a complex with AF4, AF9, ENL and AF10 (Bitoun et al., 2007; Mueller et al., 2007), which as mention before are some of the most common fusion partners of MLL in leukaemia. The association of AF9 with P-TEFb is mediated by AF4 and AF5 (MCEF, AFF4) while the association of P-TEFb with ELL (identified in 1996 as a Pol II elongation factor (Shilatifard et al., 1996)) mediates the interaction with AF4 (Bitoun et al., 2007; Mueller et al., 2007). Members of the AFF family (AF4 and AF5) are considered to be the scaffold for the assembly of the complex, as they can interact with P-TEFb, AF9, ENL and ELL. Multiple other proteins have been found to be part of the EAP/SEC complexes (different combinations of components can be found in various sub-complexes) and besides interacting with P-TEFb they have been shown to stimulate its kinase activity *in vitro* (Luo et al., 2012). The stimulation of P-TEFb increases gene expression and consequently alterations in members of the SEC influence transcription. For instance, overexpression of AF4, AF9 and ENL stimulates the expression of a luciferase reporter gene (Bitoun et al., 2007) and AF4 overexpression in fibroblast cells leads to a three-fold increase in total RNA (Marschalek, 2015). Artificial recruitment of AF5 and AF5 to a luciferase reporter gene also lead to increased expression in 293T cells (Mueller et al., 2007). Evidence for the importance of these interaction in MLL-FPs is provided by the

observations that disruption of the AF4-AF9 interaction by a synthetic peptide results in the death of MLL-AF4 leukaemic cells (Palermo et al., 2008; Srinivasan et al., 2004).

MLL-FPs also interact with the chromatin proteins Disruptor of telomeric silencing 1-like (DOT1L) (Bitoun et al., 2007; Mueller et al., 2007; Okada et al., 2005; Zeisig et al., 2005) and BRD4 (Dawson et al., 2011). DOT1L is the only known H3K79 methyltranferase (Lacoste et al., 2002; van Leeuwen et al., 2002) and plays important roles in embryonic development (Jones et al., 2008) and adult haematopoiesis (Jo et al., 2011; Nguyen et al., 2011). H3K79 can be mono-, di-, or trimethylated, and the conversion of H3K79me1 to the di and tri methylated forms has been correlated with the transition of gene expression from low to high (Steger et al., 2008). Identifying the mediators of the interaction between DOT1L and MLL-FPs is an active area of investigation, but for now it is known that DOT1L can interact directly with AF9 in a mutually exclusive manner to AF4 and P-TEFb (Biswas et al., 2011). It can also interact with AF10 through the octapeptide motif-luecine zipper region of AF10, a region shown to be required for MLL-AF10 leukaemogenesis (Okada et al., 2005). Genome wide studies revealed that MLL-FPs targets are associated with H3K79me2/me3 and that MLL-FP leukaemias can be distinguished from other acute leukaemias based on their H3K79me2/me3 profiles (Bernt et al., 2011; Guenther et al., 2007; 2008; Krivtsov et al., 2008). Because of these findings aberrant H3K79 methyaltion patterns are considered a hallmark of MLL-leukaemias. Genetic inactivation of DOT1L as well as its inhibition through small molecule inhibitors results in

impaired leukaemic growth, both in *vitro* and *in vivo* (Bernt et al., 2011; Daigle et al., 2013; 2011). This has led to the initiation of clinical trials to determine the safety and efficacy of the EPZ-5676 DOT1L inhibitor for treatment of MLLr leukaemias (Daigle et al., 2013; Helin and Dhanak, 2013).

A rigorous global proteomic survey recently identified BRD4 as an interaction partner of PAFc and members of the EAP/SEC (Dawson et al., 2011). As mentioned previously, BRD4 is a bromodomain containing protein able to bind acetylated histone tails. Importantly, BRD4 was shown to sequester P-TEFb from the snRNP complex, where its kinase activity is inhibited by HEXIM1. Association with BRD4 results in increased binding of P-TEFb to gene targets as well as stimulation of its kinase activity (Jang et al., 2005; Schroeder et al., 2012; Yang et al., 2005) . Because if its association with PAFc and EAP/SEC a small molecule inhibitor targeting the bromodomain pocket of BRD4 was tested for its ability to inhibit MLL-FP mediated leukaemogenesis. Inhibition of BRD4 led to decreased expression of MLL-FPs gene targets with a concurrent loss of PAFc and SEC components from chromatin (Dawson et al., 2011). BRD4 inhibition was shown to effectively inhibit proliferation of MLL-FPs leukaemic cells lines and patient samples *in vitro* as well as the progression of leukaemia *in vivo* (Dawson et al., 2011; Herrmann et al., 2012; Zuber et al., 2011a).

A unifying model for the molecular mechanism by which MLL-FPs transform cells has been proposed in which BRD4 and PAF1 are thought to recruit MLL-FPs and the SEC to target genes leading to increased transcriptional elongation of

these and subsequent transformation of cells. However, work using mouse models has shown that the fusion partners of MLL have a role in determining the type of leukaemia that arises arguing against the existence of a common molecular mechanism. Rather they support an instructive model of tumour tropism for these fusion proteins. For example, MLL-AF4 could only produce B-cell lymphomas when expressed in T-cells, while MLL-ENL could cause both AML and T-ALL but not B-ALL when expressed in the same cell type (Collins et al., 2000; Drynan et al., 2005; Forster et al., 2003; Metzler et al., 2006). Whether the fusion partner is influencing recruitment of the fusion protein or the mechanism by which it regulates its targets genes is still unknown. Furthermore, the cytoplasmic fusion partner AF6 does not interact directly with members of the EAP complex yet ChIPseq experiments have revealed co-localisations of these with MLL-AF6 genome wide, suggesting that these may be recruited to MLL-FP targets regardless of the identity of the fusion partner.

1.5.4.3 MLL-ENL and MLL-AF9 interact with CBX

Another component found to interact with the EAP/SEC complex is the PRC1 subunit CBX8 (Garcia-Cuellar et al., 2000; Malik and Hemenway, 2013; Monroe et al., 2011; Mueller et al., 2007). Two studies have addressed the function of CBX8 in MLL-FPs leukaemias and found opposing roles for it. In the study by Tan et al. (2011) CBX8 was found to function as a co-factor of MLL-AF9. AF9 was shown to interact with CBX8 via its C-terminal and this interaction was important for the

leukaemic transformation of lineage negative mouse bone marrow cells. By employing a conditional Cbx8 knockout mouse model the authors further demonstrated the requirement for Cbx8 in both initiation and maintenance of MLL-AF9 *in vitro* and *in vivo* and for MLL-ENL mediated leukaemia *in vitro*. This was attributed to CBX8s interaction with the acetyltransferase TIP60 shown to be important for the expression of MLL-AF9 targets like *HOXA9*, and independent of CBX8s function as part of the PRC1 complex (Tan et al., 2011). Knockdown of CBX8 in this model did not lead to the upregulation of *CDKN2A*.

In the study by Maethner et al. on the other hand, CBX8s was shown to repress MLL-ENL gene targets through a PRC1 independent manner. Overexpression of CBX8 reduced the transformation ability of haematopoietic progenitor cells by MLL-ENL and resulted in decreased expression of MLL-ENL targets including *HOXA9.* ENL was shown to interact directly with CBX8 and this interaction was important for the inhibition of the repressive function of CBX8, as shown through the use of a point mutation in MLL-ENL abrogating this interaction coupled to a luciferase reporter assay. Mutant MLL-ENL was further shown to be unable to transform HSPCs. Therefore, in this system ENL seems to play a crucial role in sequestering CBX8 and thus inhibiting its repressive function allowing for the expression of targets important for leukaemic transformation (Maethner et al., 2013). In this model, knockdown of CBX8 led to the upregulation of *CDKN2A*.

While it is difficult to reconcile these findings, as noted by Maethner et al., AF9 and ENL have previously been shown to interact with different compositions of

P-TEFb, either with CDK9/CYCT2 for ENL and CDK9/CYCT1 for AF9 establishing a precedent for unique interactions of the proteins.

1.5.5 MLL-AF4

The MLL-AF4 fusion protein is the most prevalent MLL rearrangement in acute lymphoid leukaemia (ALL) and accounts for over 50% of ALL cases in infants less than 6 months of age, 10-20% in older infants, 2% in children and 7% in adults (Tamai and Inokuchi, 2010). Studies in monozygotic twins have found an almost 100% concordance for development of leukaemia in (t4;11) infants which coupled with retrospective analysis of neonatal blood spots of patients have indicate the *in utero* development of this malignancy. Due to the rapid and aggressive development of these leukaemias it is perhaps not surprising that they are associated with a very poor prognosis with median overall survival (OS) of 10 months in children and 7 months in adults. While treatment strategies for ALL as a group have improved substantially in recent times, the refractory nature of (t4;11) leukaemias to these highlights the importance of understanding the molecular mechanism of MLL-AF4 leukaemogenesis in order to identify novel therapeutic strategies.

1.5.5.1 In vivo models of MLL-AF4 leukaemia

Despite multiple attempts, modelling MLL-AF4 leukaemias *in vivo* has proven difficult and no model has been able to recapitulate the human disease phenotype. In 2006, a knock-in and an invertor mouse model for the expression of the MLL-AF4 protein were established (Chen et al., 2006a; Metzler et al., 2006). These showed, for the first time that the MLL-AF4 protein was able to induce leukaemia *in vivo.* However, the resulting phenotype was predominantly that of a B-cell lymphoma, rather than the more immature pro B-cell phenotype seen in human ALL patients. Subsequently, a similar approach using a conditional mouse knock-in model of MLL-AF4 led to the development of pre-B cell leukaemias in mice and these had similar gene expression profiles and chromatin landscapes to human (t4;11) patient samples. This model provided valuable insights into molecular mechanism by which leukaemogenesis is achieved (Krivtsov et al., 2008) although it is important to note that, the majority of mice developed AML rather than ALL, and the pro-B cell phenotype was not seen.

Leukaemic development using these models proceeded with a very long latency, in contrast to human (t4;11) leukaemias, which are though to arise *in utero* and develop rapidly often detected at or shortly after birth (Greaves, 2005). The long latency for leukaemic development led to the proposition that a " second hit" was necessary for leukaemic transformation. However, with the exception of *RAS* and *FLT3* mutations reported in ~25-40% and 3-21% of (t4;11) patients

respectively (Armstrong et al., 2004; Driessen et al., 2013; Liang et al., 2006; Stam et al., 2007; Taketani et al., 2004), SNP array analyses and whole genome sequencing approaches have identified few cooperating mutations in these leukaemias (Andersson et al., 2011; Bardini et al., 2011; 2010).

While (t4;11) leukaemias lack additional mutations, 80% of patients express the reciprocal translocation protein AF4-MLL. In 2010, Bursen et al. demonstrated that the reciprocal AF4-MLL protein was able to induce pro-B cell like leukaemia when overexpressed in mouse haematopoietic stem/progenitor cells, either alone or in conjunction with MLL-AF4 (Bursen et al., 2010). Surprisingly none of these leukaemias expressed *HoxA9*, a well established target of MLL-FPs. Nonetheless, this model suggested that perhaps the "second hit" might be provided by the AF4-MLL protein. Surprisingly, MLL-AF4 alone was not able to induce leukaemia in this model which contrasts with research conducted by Tamai et al. (2011) where overexpression of human MLL-AF4 in mouse HSC led to leukaemic development, albeit with a very long latency and a B-cell lymphoma phenotype (Tamai et al., 2011). Despite the potential need of AF4-MLL in initiating leukaemia, knockdown of MLL-AF4 has proven sufficient to disrupt leukaemic growth of an MLL-AF4 patient cell line both *in vitro* and in an *in vivo* xenograft model (Thomas et al., 2005), suggesting that targeting key aspects of MLL-AF4 activity alone is sufficient to disrupt leukaemogenesis.

Since murine models were not successful at recapitulating the human leukaemic phenotype, attention switched to transformation of human cells. A long standing

idea in the field had been that perhaps the right cell targets for transformation were not being used. Since (t4;11) leukaemias arise in utero, two groups attempted modelling it by transforming either human ESCs (Bueno et al., 2012) or CD34+ cord blood cells (Montes et al., 2011) with MLL-AF4. Unfortunately, neither of these models led to leukaemic development, and perhaps the right cell target still remains to be identified. Given the early onset of (t4;11) leukaemias it has been proposed that rather than targeting haematopoietic cells derived from the bone marrow, attention should be switched to haematopoietic precursors of the fetal liver (Stam, 2012).

1.5.5.2 In vitro models of MLL-AF4

The use of established cell lines in the study of MLL-AF4 have provided useful insights into the molecular mechanism of these leukaemias. For instance, more than 2/3 of direct MLL-AF4 targets identified through ChIPseq approaches in patient cell lines, were found to be overexpressed in (t4;11) patient samples compared to non-MLLr leukaemias (Guenther et al., 2008; Wilkinson et al., 2013). ChIPseq analysis carried out by Guenther et al., in patient cell lines also revealed the association of MLL-AF4 with H3K279me2 which as discussed previously is a hallmark of MLL-leukaemias and its inhibition is a potential treatment strategy for patients with t(4;11) translocations. Similarly, ChIPseq performed by Wilkinson et al. identified the *RUNX1* transcription factor as a direct target of MLL-AF4. The authors demonstrated the importance of RUNX1 for leukaemic growth *in vitro* and analysis of patient data revealed *RUNX1*

expression to be a negative prognostic factor in t(4;11) leukaemias (Wilkinson et al., 2013).

Additionally, patient cell lines have the ability to engraft and cause leukaemia in mice (Thomas et al., 2005). Importantly these maintain addiction to the fusion protein (Thomas et al., 2005) as well as its downstream targets (Faber et al., 2009; Orlovsky et al., 2011).

1.6 PcG proteins

Polycomb group proteins, as mentioned previously, were discovered in Drosophila as negative regulators of homeotic genes (Lewis, 1978). Biochemical characterisation of PcG proteins has revealed that they exist as two distinct complexes, PRC1 and PRC2, and multiple subcomplex versions. Extensive overlap of these complexes on chromatin suggested that gene repression by PcG proteins is achieved by their concerted action (Boyer et al., 2006; Bracken et al., 2006; Ku et al., 2008; Lee et al., 2006; Schwartz et al., 2006). However, PRC1 and PRC2 have also been shown to bind to distinct targets and KO studies of different complex components result in different phenotypes of ES cells (Chamberlain et al., 2008; Endoh et al., 2008; Shen et al., 2008).

EZH2, EED, SUZ12 are essential for development since KO mice lacking any of these proteins fail to develop and die around embryonic day 7.5 to 8.5 (Faust et al., 1998; O'Carroll et al., 2001; Pasini et al., 2004)

PRC2 required for repression of target genes in mESCs (Boyer et al., 2006; Chamberlain et al., 2008; Leeb et al., 2010; Montgomery et al., 2005; Pasini et al., 2007; Shen et al., 2008)

EZH2 contributes to cell fate decisions, orchestrating gene expression to control the balance between self renewal and differentiation (Bracken et al., 2006)

1.6.1 Regulation of PRC2 function by Core Components

The functional human PRC2 complex was purified from HeLa cells and consists 3 core components, EZH2, EED and SUZ12 as well as the nucleosome remodelling factor RbAp46/RbAp48. (Cao et al., 2002; Kuzmichev et al., 2002; Margueron et al., 2008; Shen et al., 2008). These proteins modulate the function of PRC2 by affecting both its assembly/stability and its enzymatic properties, which have important consequences for the regulation of gene expression by the complex.

The catalytic subunits of PRC2 are EZH2 or its homolog EZH1 (Cao et al., 2002; Joshi et al., 2008; Muller et al., 2002) which require EED and SUZ12 for their enzymatic activity (Cao and Zhang, 2004; Ketel et al., 2005; Nekrasov et al., 2005; Pasini et al., 2004). EZH2/EZH1 contain the highly conserved SET domain and utilise S-adenosylmethionine (SAM) as a cofactor for the addition of a mono- di- or tri methyl group on lysine 27 of histone 3 (Cao et al., 2002; Czermin et al., 2002; Kuzmichev et al., 2002; Muller et al., 2002). EZH1/EZH2 also contain two SANT domains which may mediate binding to SUZ12 (Ciferri et al., 2012) as well as histone tails (Boyer et al., 2004) and a cysteine rich CXC domain proposed to

stabilise the SET domain (Ciferri et al., 2012). EZH1 and EZH2 compete for binding to EED and SUZ12 and PRC2 complexes containing EZH2 have higher methyltransferase activity than those containing EZH1 (Margueron et al., 2008; Shen et al., 2008). Even though functional complementation of EZH1 and EZH2 has been observed in different cell types (Ezhkova et al., 2011; Mochizuki-Kashio, 2012; Shen et al., 2008) non-redundant functions have also been reported (Stojic et al., 2011). Additionally, PRC2 containing EZH1 has been shown to compact chromatin *in vitro* consistent with its repressive function, but has surprisingly also been associated with active gene transcription and suggested to stimulate transcriptional elongation possibly through the deposition of H3K27me1 (Margueron et al., 2008; Mousavi et al., 2012; Stojic et al., 2011). Further work is required to fully understand the different properties of the two PRC2 complexes, as despite their close homology, it is becoming increasingly clear that the catalytic subunit of PRC2 has an important role in determining the function and gene regulatory properties of the complex.

SUZ12 is important for both the assembly of the PRC2 complex as well as for its methyltransferase activity. Deletions of the C-terminal VRN2-EMF2-FIS2-SUZ12 (VEFS) domain has been shown to affect PRC2 complex assembly in both Drosophila and mammals. Mutations of a conserved and highly charged sub-element of the VEFS domain has been shown to affect the methyltransferase activity of PRC2 in Drosophila (Ketel et al., 2005; Yamamoto et al., 2004). The C-terminal domain of SUZ12 has also been shown to mediate the inhibition of PRC2 activity by H3K4me3 (Schmitges et al., 2011) although the existence of bivalent

domains suggests that *in vivo* this inhibition might not be sufficient to inhibit the function of PRC2.

In mammals the drosophila homolog of ESC is EED and exists as four translational start site isoforms (Kuzmichev et al., 2004). All isoforms contain WD40 repeats that fold into a seven bladed β-propeller donut-like structure (Han et al., 2007; Margueron et al., 2009; Xu et al., 2010). This structure is involved in providing the scaffold for the interaction with EZH2 (Han et al., 2007; Margueron et al., 2009) and mutations of it reduce the levels of H3K27me3 in Drosophila (Margueron et al., 2009). The unstructured N-terminal tail of EED was found to mediate the interaction with the histone-fold domain of H3 (Tie et al., 2007). This interaction was shown to be dispensable for binding to the *Ubx* response element in Drosophila but crucial for the H3K27me3 activity of PRC2 (Tie et al., 2007). EED also mediates the association of PRC2 with chromatin through a central pocket on the top of the β-propeller that can bind both H3K27me3 as well as H3K9me3 (Margueron et al., 2009; Xu et al., 2010). Even though EED was able to bind to both H2K27me3 and H3K9me3, binding to H3K27me3 stimulated the methyltransferase ability of PRC2 substantially more than H3K9me3. Furthermore, in an elegant system for the transient recruitment of PRC2 to a luciferase reporter gene, it was shown that deposition of H3K27me3 led to stable silencing and maintenance of the repressed state over multiple cell divisions after withdrawal of the recruitment signal (Hansen et al., 2008). Additionally, spreading of the H3K27me3 was observed up to 1600 nucleotides downstream of the recruitment sites. Interestingly, recruitment of a catalytically

inactive EZH2 to the luciferase reporter gene also resulted in transcriptional repression although the silenced state was not maintained across cell divisions, thus attributing a crucial role for H3K27me3 in maintaining stable repression states. Taken together these findings led to the proposition of a model where PRC2 is recruited to its own site of methylation which in turn stimulate its methyltransferase ability resulting in spreading of the H3K27me3 mark and maintenance of gene expression profiles across cell divisions (Hansen et al., 2008; Margueron et al., 2009).

RBBP4/7 are orthologues of *Drosophila* NURF55 and like EED they contain WD40 repeats and fold into a B-propeller (Murzina et al., 2008; Song et al., 2008). NURF55 was found to enhance PRC2 enzymatic activity by roughly 3-fold but since PRC2 complex lacking NURF55 maintains enzymatic activity they are not strictly considered core members of PRC2 (Ketel et al., 2005). Additionally, RBBP4/7 is not unique to PRC2 as it associates with additional complexes including histone acetyltransferases and histone remodelling complexes (Barak et al., 2003; Smith and Stillman, 1989; Verreault et al., 1998) . RBBP4/7 are able to bind the histone H3-H4 complex (Murzina et al., 2008; Song et al., 2008; Zhang et al., 2013). ln the context of PRC2 the Drosophila NURF55 uses its H4 binding domain to bind SUZ12 and contributes to the binding of PRC2 to chromatin through its interaction with H3 (Nekrasov et al., 2005; Nowak et al., 2011).

1.6.2 PRC2 associated factors

In addition to the core components, PRC2 biochemical analysis has shown that PRC2 interacts with various other proteins including adipocyte enhancer-binding protein 2 (AEBP2) (Cao et al., 2002), Jumonji-ARID-domain-containing protein 2 (JARID2) and PCL1, 2 and 3.

PHF19/PCL3, PHF1/PCL1 and MTF2/PCL2 are the mammalian orthologues of Drosophila PCL. They contain two PHD domains that mediate the interaction with the PRC2 complex (Casanova et al., 2011) and a Tudor domain that can bind to H3K36me3 (Ballare et al., 2012; Brien et al., 2012; Cai et al., 2013; Musselman et al., 2012). In mouse ES cells PHF19 has been found to co-localise with PRC2 and H3K27me3 (Ballare et al., 2012; Brien et al., 2012; Hunkapiller et al., 2012), and knockdowns of PHF19 lead to loss of PRC2 binding and decreased H3K27me3 levels at ~60-70% of PHF19-PRC2 bound genes (Ballare et al., 2012; Hunkapiller et al., 2012). PRC2 depleted cells also show decreased binding of PHF19 suggesting a role for PHF19 in stabilising binding of PRC2 to chromatin, rather than in recruitment (Ballare et al., 2012).

JARID2 is the founding member of the Jumonji C (JmjC) domain protein family of histone demethylases, but is assumed to be catalytically inactive as it lacks key amino acids in the enzymatic pocket. In ES cells JARID2 has been found to co-localise with PRC2 and their recruitment is interdependent (Landeira et al., 2010; Li et al., 2010; Pasini et al., 2010a; Shen et al., 2009). The effect on

H3K27me3 in JARID2 knock-down or knock-out cells is however not clear as H3K27me3 levels were reported to decrease (Landeira et al., 2010; Li et al., 2010; Pasini et al., 2010a), increase (Shen et al., 2009) or remain unchanged (Peng et al., 2009).

AEBP2 has been found to physically interact with EZH2, SUZ12 EED and RBBP4/7 (Cao et al., 2002; Ciferri et al., 2012). Cryo-electron microscopy studies revealed that AEBP2 interacts with EZH2, SUZ12 and EED while being more distant from RBBP4 (Ciferri et al., 2012) . These interactions were proposed to stabilise the complex as well as promote the catalytic activity of PRC2 by bridging the chromatin-binding and enzymatic domains of PRC2 (Ciferri et al., 2012). While AEBP2 was shown to stimulate PRC2 tri-methylation of H3K27 (Cao and Zhang, 2004) recently AEBP2 was shown to stimulate PRC2 activity even further in the presence of H2AK119ub, thus providing a link between H2AK119ub and H3K27me3 deposition (Kalb et al., 2014). This activity was further stimulated in the presence of AEBP2-JARID2.

1.6.3 PRC1

At the core of all PRC1 complexes is an E3-ubiquitin ligase, either RING1B or RING1A (Cao et al., 2005; Wang et al., 2004a) responsible for the H2AK119 ubiquitination by PRC1. Association of these factors with one of six PCGF (PCGF1-6) proteins define the 6 different sub-complexes of PRC1 identified to date (PRC1 sub-complexes can be classified in different ways, here the

classification by Gao et al. is used) (Gao et al., 2012). RING1A/RING1B catalytic activity is stimulated by interaction with PCGF proteins (Elderkin et al., 2007; Li et al., 2006). Interestingly ChIP-seq experiments in human 293T-Rex cells, revealed that while the majority of RING1B (and RYBP) binding sites overlapped that of either one of the five PCGF proteins analysed, the different PCGF proteins only showed a very small overlap with each other (Gao et al., 2012). In line with the role of distinct PRC1 complexes regulating specific genes PCGF4 has an established role in HSC self-renewal, most likely mediated through repression of the INK4a-ARF locus (Park et al., 2003), while PCGF2 is important for B-cell development (Akasaka et al., 1996) .

PRC1-PCGF2/4 can associate with one of five CBX proteins (CBX,2,4,6,7,8) adding further diversity to the PRC1 complexes. Similar to PCGF proteins, CBX proteins exhibit different expression patterns and can regulate distinct gene targets. For example, while overexpression of CBX7 in HSC enhances self-renewal, overexpression of CBX8 results in differentiation and exhaustion of HSCs (Klauke et al., 2013). CBX proteins are able to bind H3K27me3 through their chromo-domain (Cao et al., 2002; Fischle et al., 2003; Min et al., 2003) providing a nice model for PRC2 mediated recruitment of PRC1 (Boyer et al., 2006; Cao et al., 2002; Wang et al., 2004b). PRC1 complexes containing CBX proteins also contain polyhomeotic proteins (PHC1/2/3) (Gao et al., 2012; Levine et al., 2002) and are referred to as canonical PRC1. PH proteins have been shown to polymerise through their Sterile Alpha Motif (SAM) domain and mutations of this domain affect binding of both PRC1 and PRC2 to chromatin and repression of targets

(Isono et al., 2013). Conflicting data exists regarding the ubiquitination ability of canonical PRC1. Knock-down of CBX7 in mESCs for instance has been shown to reduce H2AK119ub levels as well as RING1B binding to PcG targets not bound by RYBP (Morey et al., 2013) while artificial tethering of PCGF2/4 to chromatin resulted in recruitment of RING1B without ubiquitination of H2AK119. This is in contrast to artificial tethering of PCGF1/3/5 which all resulted in robust ubiquitination (Blackledge et al., 2014).

All PCGF proteins can associate with RYBP or its paralog YAF2 to form variant PRC1 complexes. In the case of PCGF2/4 these interactions are mutually exclusive with CBX proteins. PCGF1 is found in a complex with KDM2B, a ZF-CXXC containing H3K26me2/me3 demethylase. While KDM2B binds independently of PRC1 at thousands of sites it has recently been shown to mediate recruitment of PRC1 to un-methylated CpG islands through its ZF-CXXC domain.

Recently EED was shown to interact with the PRC1 component RING1B both *in vitro* and *in vivo*. The *in vitro* affinity of EED and RING1B was similar to that of EZH2 and EED. RING1B and PCGF4 were shown to compete with EZH2 for binding of EED and EED stimulated the ubiquitination activity of RING1B-PCGF4 (Cao et al., 2014b).

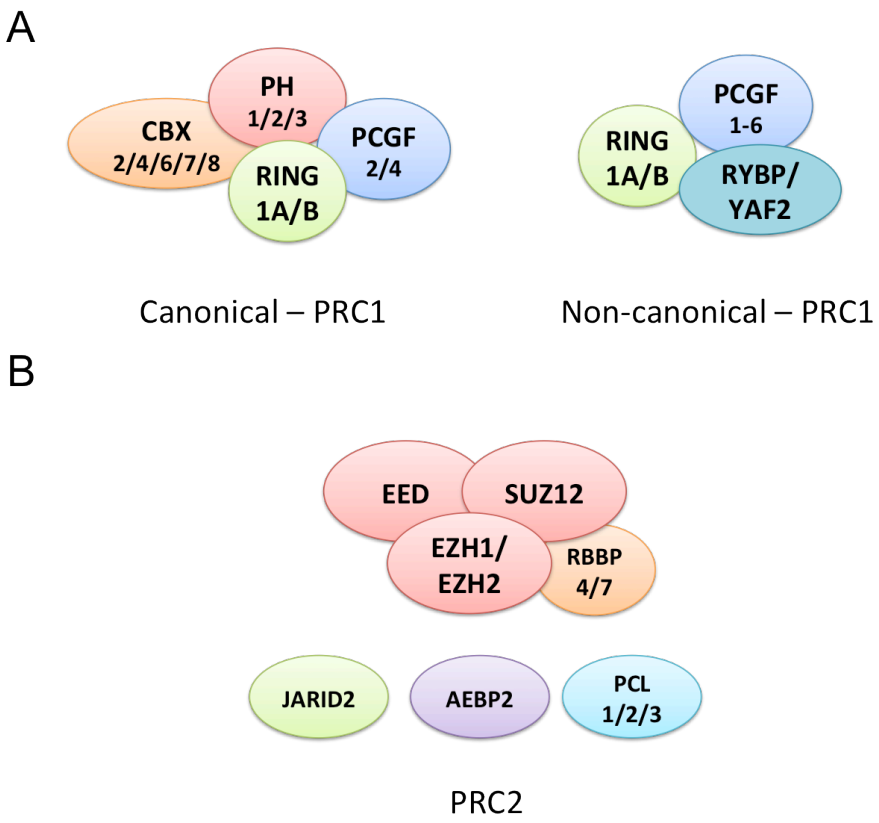


Figure 2 The PRC1 and PRC2 complexes

A) The Canonical-PRC1 complex is composed of RING1A or B, PH1, 2 or 3, CBX2, 4, 6, 7 or 8 and either PCGF2 or 4. The non-canonical PRC1 complex consists of RING1A or B, PCGF1, 2, 3, 4, 5 or 6, RYBP or YAF2 as well as additional proteins. B) The core subunits of PRC2 are EED, SUZ12 and either EZH1 or 2. PRC2 can also associate with RBBP4 or 7 as well as JARID2, AEBP2 and PCL1, 2 or 3.

1.6.4 Molecular mechanism of PcG repression

The exact molecular mechanism by which PcG proteins mediate gene silencing is still unknown. Two broad mechanisms have been proposed: chromatin compaction and interference with the transcriptional machinery. The first evidence of PRC1 mediated nucleosome compaction came from electron microscopy studies using nucleosome arrays *in vitro* (Francis et al., 2004). This

function was later shown to be independent of the catalytic activity of RING1A/RING1B as catalytic mutants were able to compact the *HoxB* locus *in vivo,* although it was shown to be necessary for the proper repression of the locus (Endoh et al., 2012). PRC2 containing EZH2 has little chromatin compaction properties but as mentioned previously, PRC2-EZH1 has been shown to compact nucleosomes independently of its catalytic activity *in vitro* (Mousavi et al., 2012). However, although the catalytic activity of PRC2 has recently been shown to be stimulated by dense chromatin *in vitro,* compaction at the *CYP26a1* gene in mESCs preceded the deposition of H3K27me3 (Yuan et al., 2012). Compaction of chromatin is thought to impede binding of transcription factors or the transcriptional machinery to DNA. Consistent with this, Chopra et al. showed through a ChIP-seq approach in *D. melanogaster esc* mutant embryos that, concomitant to a decrease in H3K27me3, 2100 fly genes acquired paused RNAP compared to wild-type embryos. However, recently PcG target genes have been shown to be associated with regions of DNAse hypersensitivity and therefore chromatin compaction is unlikely to be the sole mechanism by which PcG proteins mediated gene repression (Beck et al., 2014).

Despite evidence for compaction of nucleosomes by PcG proteins, early studies in *D. melanogaster* demonstrated that binding and repression by PcG proteins did not exclude TATA-binding protein (TBP) and TBP-associated factors (TAFs) from binding at the promoter (Breiling et al., 2001; Saurin et al., 2001). This led to the proposition that PcG proteins may interfere with transcriptional elongation. Subsequently, paused polymerase was observed at "bivalent"

domains, marked by both H3K4me3 and H3K27me3 (Brookes et al., 2012; Guenther et al., 2007; Stock et al., 2007) supporting this notion. Conditional deletion of RING1B in RING1A null ESC showed that even though levels of RNAP5 and RNAP2 did not change at these genes the levels of RNAP detected with 8WG16 (thought to recognise the initiating form of polymerase) increased and this correlated well with changes in gene expression (Stock et al., 2007). This led the authors to conclude that RNAP may have an unusual structural conformation at bivalent genes that prevented the detection with 8WG16. Loss of RING1B was proposed to lead to structural changes in RNAP and gene transcription. However, if RNAP is truly paused at bivalent genes as a result of PcG the lack of RNAPS2 increases following deletion of *RING1B* and increases in gene transcription is difficult to reconcile. Since deletion of RING1B led to complete loss of H2ABK119 but did not change the levels of H3K27me3 or the binding of EZH2 substantially the transcriptional changes observed were therefore attributed to H2ABK119ub. This is consistent with findings that non –PRC1 mediated H2ABK119ub impedes the recruitment of FACT and thus transcriptional elongation at the RANTES gene in macrophages (Zhou et al., 2011).

1.6.5 Function of H3K27me3

Given that CBX proteins are able to bind to H3K27me3 a role for this chromatin modification in the recruitment of PRC1 has been proposed (Fischle et al., 2003; Min et al., 2003). Knockout and knockdown studies of multiple components of PRC2 have been shown to reduce binding of PRC1 at specific targets and result in

de-repression of these (Boyer et al., 2006; Cao et al., 2002; 2005; Wang et al., 2004b). However, while the levels of H3K27me3 are reduced in these experiments so are the levels of the PRC2 component being targeted and therefore the contribution of H3K27me3 to PRC1 recruitment cannot be teased apart in experiments of this type. However, alternative approaches affecting levels of H3K27me3 have contributed to this model. The *Pramecium bursaria* chlorella virus 1 (PBCV-1) encodes a functional H3K27me2/3 SET domain histone methyltransferase (vSET). Expression of vSET in HeLa cells treated with an siRNA targeting *EZH2* was able to restore H3K27me3 levels at multiple *Hox* loci and their repression (Mujtaba et al., 2008). Methylation by vSET led to increased levels of CBX8 at *Hox* loci further supporting the idea that H3K27me3 serves as a docking site for PRC1 (Mujtaba et al., 2008).

Further support for the importance of the H3K27me3 modification in gene regulation has come from the observation that mutations in the SET domain of EZH2 are associated with multiple forms of cancer (reviewed in Koppens et al. 2015). In a subset of follicular lymphoma (FL) and diffuse large B-cell lymphoma (DLBCL) patients mutations within the SET domain Tyrosine 641 or Alanine 677 result in increased tri-methylation activity of *EZH2* (Morin et al., 2010; Sneeringer et al., 2010; Yap et al., 2011) . The gain of function and increased levels of H3K27me3, although not the drivers of these lymphomas, contribute to the transformation of cells (Beguelin et al., 2013; Berg et al., 2014). Inhibition of the catalytic activity of EZH2 through small molecule inhibitors reduces the levels of H3K27me3 and inhibits the proliferation of EZH2 mutant DLBCL cell

lines both *in vitro* and *in vivo* xenograft models (McCabe et al., 2012b; Qi et al., 2012). On the other hand inactivating mutations of EZH2 are associated with 25% of T-cell acute lymphoblastic leukemia (T-ALL) cases (Ntziachristos et al., 2012). The main drivers of these leukaemias is NOTCH signalling, which was shown to lead to reduced levels of H3K27me3. In this scenario PRC2 serves as a tumour suppressor and inactivating mutations in EZH2 contributes to the progression of the disease (Ntziachristos et al., 2012). Therefore, it seems that the H3K27me3 has important biological functions.

1.6.6 Recruitment of PcG protein

In Drosophila, recruitment of PcG proteins is mediated by transcription factors recognising Polycomb response elements. In mammals, PREs have proven elusive to find with only a few PRE elements identified and evidence exist for recruitment of PcG proteins through sequence specific transcription factors, noncoding RNAs as well as un-methylated CpG islands.

1.6.6.1 Transcription factor mediated recruitment

The transcription factor REST has been biochemically shown to interact with members of both canonical PRC1 and PRC2 and proposed to recruit these to a subset of targets (Dietrich et al., 2012; Ren and Kerppola, 2011). However, binding of RING1B and SUZ12 in REST null mES both increased and decreased

depending on the target (Dietrich et al., 2012) Similarly RUNX1 was shown to interact with PCGF4 in megakaryocytic and T lymphocytic cells and co-localise with 50% of RING1B sites (Yu et al., 2012). ChIP-seq analysis of RING1B binding in RUNX1 null thymocytes showed decreased binding of RING1B at a subset of RUNX1 and RING1B targets (Yu et al., 2012). The contribution of these TF to PcG recruitment seems to be gene target specific and cannot explain recruitment fully. Additionally, recent work by Riising et al. has shown that transcription inhibition in mES is sufficient to induce genome wide ectopic recruitment of PRC2 to nucleosome free CpG islands (Riising et al., 2014). Therefore it is possible that the change in transcriptional status of the genes is responsible for the loss of PcG binding rather than the interaction with the TFs.

1.6.6.2 Long non-coding RNA mediated recruitment

Xist RNA is the first lncRNA to be implicated in PRC2 recruitment (Mak et al., 2002; Plath et al., 2003; Silva et al., 2003; Zhao et al., 2008). In mammals, X inactivation in females is important to ensure correct dosage of X-linked genes between the two sexes (Lyon, 1961). Xist RNA is transcribed from the inactive X chromosome (Xi) and spreads in *cis* to coat the Xi in its entirety (Brockdorff et al., 1992; Brown et al., 1992) and initiate silencing (Herzing et al., 1997; Lee et al., 1996; Penny et al., 1996; Wutz and Jaenisch, 2000). The inactive X-chromosome is also bound by PRC2 and PRC1 components as well as marked by H3K27me3 and H2AK119ub (de Napoles et al., 2004; Kohlmaier et al., 2004; Plath et al., 2003; Silva et al., 2003) . Since recruitment of PRC2 to the Xi occurs at the onset

of Xist expression (Plath et al., 2003; Silva et al., 2003), and is dependent on its continued expression for binding (Kohlmaier et al., 2004; Mak et al., 2004; Plath et al., 2003), it was suggested to recruit PRC2 directly. This recruitment could potentially be mediated by a stem-loop structure formed by the A-repeat domain Xist, found to interact with EZH2 and SUZ12 (Kohlmaier et al., 2004; Zhao et al., 2008). Recently however, superresolution microscopy revealed that Xist and PRC2 do not colocalise on Xist coated chromosomes arguing against direct recruitment of PRC2 by Xist (Cerase et al., 2014).

Another example of PRC2 interaction with lncRNAs is given by the lncRNA HOTAIR which is transcribed from the *HOXC* locus (Rinn et al., 2007). Rinn et al., showed that knockdown of HOTAIR results in loss of binding of PRC2 and H3K27me3 as well as upregulation of the HOXD genes. Genome-wide mapping of HOTAIR in breast cancer cells revealed co-occupancy of PRC2 with HOTAIR. Importantly HOTAIR occupancy was not altered in EZH2 deficient cells suggesting that HOTAIR may bind to targets first and then recruit PRC2 (Rinn et al., 2007).

1.6.6.3 CpG island mediated recruitment

Multiple studies have reported a correlation between unmethylated CpG islands and PcG occupancy (Bernstein et al., 2006; Boyer et al., 2006; Ku et al., 2008; Mikkelsen et al., 2007; Tanay et al., 2007). Integration of *E.coli* GC rich sequences into mES cells results in recruitment of PRC2 and H3K27me3 (Mendenhall et al.,

2010). Similarly, exchange of short DNA GC-rich fragments of the Polycomb bound human alpha globin locus into the orthologues non-Polycomb bound site in mouse ES cells, results in recruitment of PRC2 and H3K27me3 (Lynch et al., 2012). Of note, is that in both of these experiments H3K4me3 deposition was also observed. Consistent with the recruitment of PcG proteins to CGIs, multiple studies have reported re-distribution of PcG proteins to CGIs in DNA methylation-deficient ESC (Brinkman et al., 2012; Cooper et al., 2014; Hagarman et al., 2013; Lynch et al., 2012; Reddington et al., 2013). Furthermore, *in vitro* experiments demonstrated that the catalytic activities of PRC1 and PRC2 are not affected by DNA methylation, and artificial recruitment of PcG proteins to sites of DNA methylation *in vivo* resulted in both H3K27me3 and H2AbK119ub (Cooper et al., 2014). Therefore it seems that the methylation status of CpG islands affects recruitment of PcG proteins rather than modulating its catalytic activity.

A link between PRC1 recruitment and CGIs is provided by the variant PRC1-PCGF1 complex, which is associated with the ZF-CXXC domain containing KDM2B. While KDM2B has been found to associate with thousands of sites genome wide (irrespective of their expression status) a subset of its targets are also bound by PRC1. Knockdown of KDM2B results in loss of RING1B occupancy genome-wide as well as in de-represssion of some PcG bound genes (Farcas et al., 2012; He et al., 2013; Wu et al., 2013), thus establishing a role for KDM2B in the recruitment of PRC1 and PcG repression.

Additionally, artificial tethering of KDM2B to chromatin was shown by two groups to not only result in H2AbK119ub but importantly also recruitment of PRC2 and H3K27me3 (Blackledge et al., 2014; Cooper et al., 2014). While PRC2 has been found to recruit PRC1 to many targets, independent recruitment of PRC1 has also been observed. Genome wide analysis of RING1B and H2AK119ub in EED null ES cells for instance, revealed no significant alterations of PRC1 binding or function even though CBX7 occupancy was affected (Tavares et al., 2012). Similarly, in PRC2 depleted mESCs and differentiated cells PRC1 binding was retained at specific sites (Leeb et al., 2010; Pasini et al., 2004) while global levels of H2AK119ub were not altered in EED null ES cells, nor recruitment of RING1B to the Xi (Schoeftner et al., 2006). Furthermore, genome-wide analysis in *Drosophila* revealed that although PRC2 and PRC1 were bound to the same targets their binding profiles were different with PRC1 enriched at PRE regions and PRC2 spread across the genes suggestive of a PRC2 independent recruitment mechanism of PRC1 (Schwartz et al., 2006).

Recruitment of PRC2 by artificial tethering of KDM2B was dependent on an intact ZF-CXXC domain while the demethylase activity of KDM2B was dispensable (Blackledge et al., 2014; Cooper et al., 2014). H2AK119ub was shown to be sufficient for recruitment of PRC2 and H3K27me3, as demonstrated by the tethering of a minimal RING1B/PCGF4 catalytic domain (RPCD) that is unable to interact with other PRC1 components but retains E3 ligase activity. Mutations abrogating the catalytic activity of RPCD failed to recruit PRC2 to chromatin. As mentioned previously, AEBP2-JARID2 was found to stimulate the

catalytic activity of PRC2 in the presence of H2ABK119ub nucleosomes *in vitro* providing a link for the recognition of H2ABK119ub by PRC2 (Kalb et al., 2014).

The role of PRC1 in recruiting PRC2 was further confirmed through tamoxifen-induced deletion of RING1B in RING1A null mES cells. This resulted in loss of SUZ12 binding at almost all its target sites, suggesting that PRC1 is the main recruiter of PRC2. However, genome-wide data from previous studies in multiple cell lines including ES cells have shown that PRC1 and PRC2 do not always overlap and whether PRC1 can recruit PRC2 in different systems and all targets sites has not been demonstrated.

1.6.6.4 Modulation of PcG recruitment by active transcription

While PcG proteins are associated with CpG islands not all CpG islands are bound by PcG proteins, despite the presence of KDM2B (Farcas et al., 2012) . Bioinformatics analysis of PRC1 and PRC2 binding sites in ES cells revealed that PcG proteins bind to CpG island promoters lacking in activating TF motifs (AMs) suggesting that PcG proteins are recruited to transcriptionally silent genes (Ku et al., 2008). This was evaluated experimentally by insertion of two BAC constructs into a gene desert region in ES cells. Each construct contained a portion of a CpG island associated with two ubiquitously expressed genes. One construct contained multiple AMs and resulted in H3K4me3 while the other construct lacked AMs and resulted in H3K27me3 deposition (Mendenhall et al., 2010). Similar results were obtained by Lynch et al. where the chromatin landscape of a

construct containing an active gene and marked by both H3K4me3 and H3K27me3 could be altered through either deletion or addition of activating motifs and thus changes in transcriptional output. Deletion of the promoter region resulted in PRC2 recruitment and H3K27me3 with loss of H3K4me3 while addition of an additional promoter resulted in increased transcription and levels of H3K4me3 with loss of H3K27me3 (Lynch et al., 2012). The effects of transcription on recruitment of PcG proteins was evaluated genome-wide and consistent with the findings of Lynch et al. and Mendenhall et al. PcG proteins were recruited to unmethylated CpG islands genome-wide following inhibition of transcriptional elongation (Riising et al., 2014). These findings are consistent with the observed antagonism between H3K27me3 and the H3K36me3 mark discussed previously.

1.6.6.5 A responsive model for Recruitment of PcG proteins and the interplay with TrxG proteins

A model for the recruitment of PcG proteins has recently been proposed by Klose et al. (2013). In this model PcG proteins are envisioned as responsive to the transcriptional status and chromatin landscape of target genes rather than instructing the expression status through their recruitment by either lncRNAs or TFs. PcG proteins were proposed to sample chromatin through their association with KDM2B and establish PcG domains in the absence of transcription. Chromatin states would be stabilised by positive feedback mechanism such that methylation of H3K27me3 would result in binding of PRC1 as well as

propagation of the H3K27me3 mark through EED binding to H3K27me3 or stimulation of PRC2 catalytic activity by chromatin compaction through PRC1. Conversely, at actively transcribed genes the establishment of PcG domain would be antagonised by the action of TrxG proteins establishing H3K4me3 and H3K36me3 marks know to antagonise Polycomb activity. In addition, as actively transcribed genes are enriched for H3K27Ac this would further antagonise PcG binding TrxG proteins would also sample gene targets through ZF-CXXC containing components and similar to PcG proteins would stabilise active chromatin states through positive feedback loops.

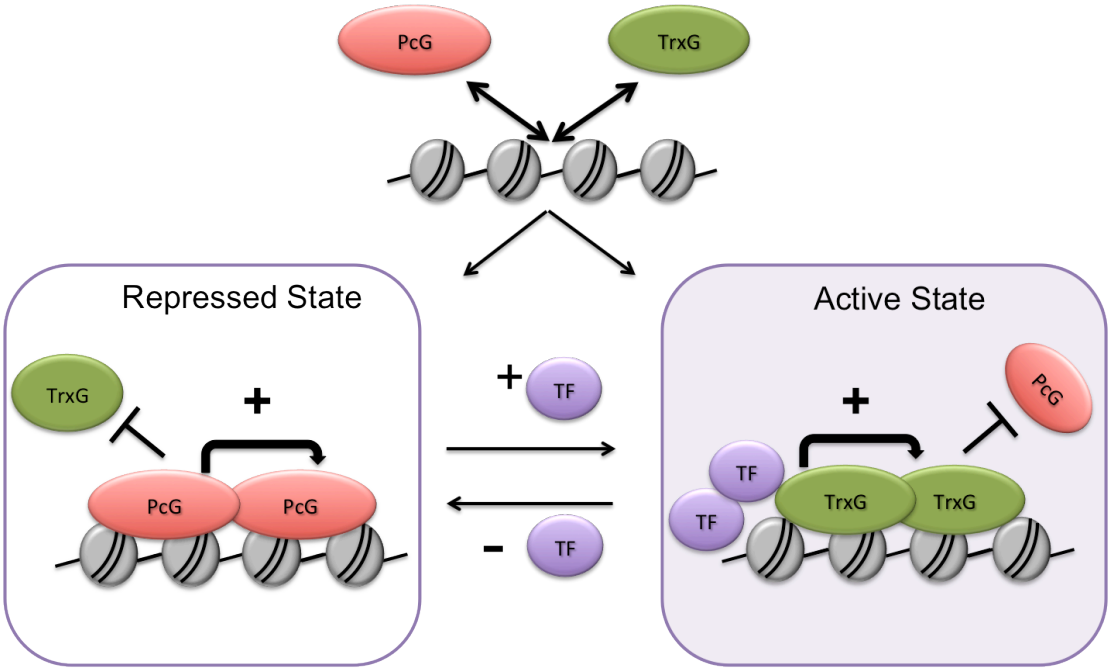


Figure 3 Responsive model of PcG recruitment.

Pcg and TrxG protiens are proposed to sample the chromatin environment of a gene. In the absence of transcription and activating signals, such as TFs, PcG proteins are able to bind chromatin and form a Polycomb domain. These domains are thought to be stabilised through a positive feedback mechanism and serves to antagonise binding of TrxG proteins, maintaining the repressed state of a gene. Conversely, in the presence of transcription and activating signals (TFs), TrxG proteins bind chromatin. TrxG binding is thought to be stabilised through positive feedback mechanisms and serves to antagonise PcG binding, maintaining the gene in an active transcriptional state. Switches between the two states are thought to occur upon increases or decreases of activating signals, for examples TFs.

Through sampling, PcG proteins are suggested to silence and establish stable PcG domains in the absence of transcription and activating signals. This is suggested to be important to prevent stochastic expression of genes through development, as a threshold level of activating signals would be necessary to overcome repression. The observation that relatively few genes are affected by loss of PcG

proteins in the absence of differentiation, suggests that this might in fact be the case.

Bivalent domains in this scenario, as mentioned previously, are the result of sampling by both TrxG and PcG proteins. This is thought to be the result of low level activating signals (possibly TFs) which on the one hand prevent the establishment of PcG domains and on the other hand are not sufficient for the establishment of TrxG domains. Binding of PcG proteins at these domains serves to antagonise the function of TrxG domains and vice versa, resulting in a "balanced" state. This interplay between TrxG and PcG proteins may be important in preventing expression of genes as a result of the stochasticity of expression of TFs.

1.6.7 Functions of PcG proteins in development and differentiation

PcG proteins have important roles in development and differentiation. Knockout *Ring1b, Rybp*, *EZH2, Eed Suz12* and *Jarid2* in mice, results in early embryonic lethality while knockouts of other components of PRC1 lead to homeotic transformations, perinatal lethality and immune defects (reviewed in (Laugesen and Helin, 2014)).

The function of PcG proteins has been studied in great detail in ESCs where they have been shown to be important for repression of genes involved in differentiation. While deletion of the PRC2 component *EED* in ESCs does not

affect the self-renewal ability of ESCs (Chamberlain et al., 2008), double knockouts of RING1A/RING1B do (Endoh et al., 2008). Knockouts of *EZH2* (Shen et al., 2008) and *SUZ12* (Pasini et al., 2004) result in impaired mesodermal and endodermal lineage specification of ESCs respectively. Conversely, in other cell types, inactivation of PcG proteins have been found to promote differentiation for example in myogenesis (Caretti et al., 2005) and epidermis formation (Ezhkova et al., 2009).

In haematopoiesis *PCGF4* and *Cbx 7* have been shown to be important for HSC self-renewal (Klauke et al., 2013; Park et al., 2003). EZH1 is important for the maintenance of adult HSCs (Hidalgo et al., 2012) while EZH2 is dispensable (Xie et al., 2014). While reports have indicated the requirement of EZH2 in fetal liver HSCs using conditional deletions driven by $Tie2^{Cre}$ (Mochizuki-Kashio, 2012), subsequent research demonstrated that conditional knockouts of *EZH2* did not alter fetal liver HSCs through use of Vav^{Cre} mediated excision (Xie et al., 2014). Since Tie2 is also expressed in endothelial cells whereas Vav expression is restricted to haematopoietic cells, it is likely that the dependency of fetal liver HSCs is not cell-autonomos (Xie et al., 2014). Overexpression of *EZH2* on the other hand, has been found to prevent HSCs exhaustion upon serial replating (Kamminga et al., 2006) and to enhance HSCs self-renewal capacity in a knock-in mouse model of *EZH2* (Herrera-Merchan et al., 2012). Inactivation of SUZ12 on the other hand has been found to enhance HSCs activity *in vivo* (Majewski et al., 2008). Taken together, these results indicate that the function of PRC2 may be dosage dependent.

While EZH2 deficient mice do not show alterations in HSCs, the lymphoid compartment is affected (O'Carroll et al., 2001). EZH2 has been found to be required for the proper differentiation of Pro-B to Pre-B cells through regulation of V_HJ558 rearrangments (Su et al., 2003) and in germinal centre formation (Beguelin et al., 2013).

1.6.8 PRC2 in MLLr leukaemia

Contrasting the function of PcG proteins as antagonists of TrxG proteins in development, PcG proteins have been found to cooperate with MLL-FPs in establishing expression programmes that support leukaemic development (Neff et al., 2012; Shi et al., 2013; Smith et al., 2011; Tan et al., 2011; Tanaka et al., 2012; Thiel et al., 2013; Xu et al., 2014) . A common finding among almost all the studies addressing the function of PcG proteins in MLLr leaukaemias, is their function as repressors of the *CDKN2A* locus which encodes the tumour supressors $p16^{INK4a}$ and $p19^{ARF}$. These proteins are important for proper cell-cycle progression. In the absence of $p16^{INK4a}$, CDK4 hyperphosphorylates pRb, which prevents its interaction with E2F proteins and thus prevents appropriate G1/S transition. $p19^{ARF}$ binds the MDM2 protein and thus prevents the degradation of the p53 tumour suppressor.

1.6.8.1 Function of PRC2 in MLLr leukaemias

In 2012, two studies addressed the role of EZH2 in mouse models of MLL-AF9 leukaemia. While Tanaka et al. (2012) were only able to show modest effects on the development of primary leukaemias following deletion of *EZH2 in vivo*, this effect was more pronounced in secondary leukaemias due to decreased numbers of leukaemia initiating cells (LICs). In addition to upregulation of $p16^{INK4a}$ and reduced cycling of *EZH2* deleted cells, they demonstrated upregulation of the TF *EGR1,* which mediated differentiation of MLL-AF9 leukaemic cells. However, overall the deletion of *EZH2* in this model was compatible with leukaemic development (Tanaka et al., 2012). This compatibility was confirmed by Neff et al. (2012) in which primary leukaemias from MLL-AF9 transformed mouse progenitor cells were not affected by deletion of *EZH2 in vivo*. Deletion of *EZH2* led to decreased but detectable levels of H3K27me3. To inactive the function of PRC2 completely, Neff et al. (2012) induced deletion of the core PRC2 component *EED in vivo.* This led to a complete loss of H3K27me3 as detected by immunoblots and was incompatible with leukaemic progression. While both *EZH2* null cells and *EED* null cells upregulated the CDKN2A locus, the upregulation was higher in *EED* null cells, perhaps as a result of the complete lack of H3K27me3. Subsequently, an shRNA screen using an MLL-AF9;$NRas^{G12D}$ mouse model identified *EED* and *SUZ12* as important targets in AML leukaemias. shRNA mediated knockdowns of *EED* and *SUZ12* led to the inhibition of leukaemic growth *in vitro* and *in vivo.* Consistent with the findings of Neff et al. (2012) inhibition of PRC2 led to the upregulation of the *CDKN2A* locus, and

induced differentiation of leukaemic cells. Importantly, shRNA knockdown of *SUZ12* and *EED* did not affect the *in vitro* growth of an immortalised non-leukaemic cell line or the growth of the non-transformed G1E and EML haematopoietic cell lines. This indicates that inhibition of PRC2 may be a feasible therapeutic strategy.

Pharmacological inhibition of PRC2 through the small molecule inhibitor UNC1999 has also been shown to be effective in inhibiting leukaemic growth *in vitro* and *in vivo* (Xu et al., 2014). The UNC1999 inhibitor specifically inhibits the methyltransferase activity of EZH2 and EZH1 leading to decreased levels of H3K27me3/2 (Konze et al., 2013) . Inhibition of PRC2 by this inhibitor affected the proliferation of AML patient cell lines expressing MLL-AF9, MLL-ENL, MLL-PTD and MLL-AF4 while it had no effect on the erythroleukaemic cell line K562 or the T-ALL cell line LOUCY which does not express EZH2 (Xu et al., 2014). Consistent with the studies described above, treatment of MLL-AF9 transformed mouse progenitor cells resulted in differentiation as well as a *CDKN2A* mediated cell cycle arrest. UNC1999 was also shown to be effective at inhibiting leukaemic progression in an *in vivo* MLL-AF9 model, albeit the survival advantage of mice treated with UNC1999 was modest.

1.6.9 Epigenetic Inhibitors of PRC2

Small molecule inhibitors of chromatin proteins have emerged as novel therapeutic agents for the treatment of cancer. Among inhibitors of epigenetic

proteins that have been approved for clinical use are the two DNA-methyltransferase inhibitors azacitidine and decitabine. These are used as first line treatment strategies in myelodysplastic syndrome patients (Fenaux et al., 2009; Issa and Kantarjian, 2009; Silverman and Mufti, 2005). HDAC inhibitors have also been approved, including vorinostat and romidepsin for the treatment of T-cell lymphoma patients (Khan and La Thangue, 2012; Wagner et al., 2010).

The first inhibitor of PRC2 3-decaneplanocin-A(DZNep) was identified by Tan et al., and has shown activity against various cancer types (Tan et al., 2007). It reduces the proteins levels of PRC2 complex components and leads to the upregulation of PRC2 targets (Tan et al., 2007). However, as it reduces levels of PRC2 proteins it does now allow for the study of the catalytic function of PRC2. Additionally, it is not specific for EZH2/EZH1 and inhibits methylation broadly (Miranda et al., 2009)

Recently small molecule inhibitors specific for EZH2/EZH1 have been developed which are competitive with SAM and non-competittive with the peptide substrate (Campbell et al., 2015; Knutson et al., 2012; Konze et al., 2013; McCabe et al., 2012b; Qi et al., 2012) . EPZ005678 has shown efficacy against B-cell lymphomas with tyrosine 641 or alanine 677 mutations in the SET domain of EZH2 (Knutson et al., 2012). Further development of the compound yielded EPZ-6438, which is currently in phase 1 and 2 clinical trials for patients with relapsed or refractory B-cell lymphoma (Campbell et al., 2015). Of particular interest for

this thesis is the UNC1999 inhibitor developed by Konze et al. which is discussed in further detail below (Konze et al., 2013)

1.6.10 The UNC1999 inhibitor

UNC1999 is a specific small molecule inhibitor of EZH2/EZH1 which is competititve with SAM and not competitive with the peptide substrate (Konze et al., 2013). It was developed by Konze et al. and displays a 10-15 fold selectivity for EZH2 over EZH1 and a > 10 000 fold selectivity for EZH2 over 15 other methyltransferases. It demonstrated high selectivity over more than 90 kinases, GPCRs, transporters and ion channels, with the exception of sigma2. The half maximal inhibitory concentration for EZH2 was measured at < 10ηM and 45ηM for EZH1. UNC1999 is orally bioavailable and a single 50mg/kg oral dose has a Cmax of 4.7μM in Swiss albino mice. A di-methylated analogue of UNC1999, UNC2400 is 1 000-fold less potent for EZH2 than UNC1999 in biochemical assays (Konze et al., 2013) and is therefore a useful negative control for *in vitro* experimentation.

Mass spectrometry analysis of histone modification levels in an MLL-ENL murine leukaemic line following treatment with UNC1999 or DMSO revealed that only peptides covering the H3 residues 27-40 were altered significantly. The greatest decreases were seen in peptides with single H3K27me2/3 modifications (Konze et al., 2013).

72 hour 5μM UNC1999 treatment of MCF7 cells results in complete loss of H3K27me3 levels as detected by immunoblotting without affecting the levels of EZH2 (Konze et al., 2013). A four-day treatment of MLL-AF9 transformed leukaemic cells with 3μM UNC1999, similarly results in almost complete loss of H3K27me2/3 while levels of H3K27me1, H3K36me1/2 and H3K4me3 are unaffected. Conversely, treatment with 3μM UNC2400 has no effect on any of the epigenetic modifications listed above (Xu et al., 2014).

1.7 Aims of the Thesis

1. Identify/validate targets important for MLL-AF4 leukaemia
2. Use the small molecule inhibitor UNC1999
    a. To determine the molecular effects of PRC2 inhibition in MLL-AF4 leukaemia
    b. Investigate the efficacy of the PRC2 inhibitor in a xenograft model of MLL-AF4 leukaemia

2 Chapter 2 Methods

2.1 Cell Culture

Cells used in this study and the type of media they were grown in are listed in Table 1. All cell lines were maintained at 37˚C with 5% CO_2. Cells were grown between $2x10^5$-$2x10^6$ cells/ml. Live cell counts were determined using 0.4% Trypan blue (GIBCO, Life technologies) and a Neubauer haemocytometer.

All cell lines grew in suspension except HEK293T cells which are adherent and were grown in flat-bottomed plates. HEK293T cells were passaged at 80% confluency. Cells were detached from the plates through addition of 1X Trypsin (0.5% Trypsin-EDTA (10X), #15400-054, Gibco, Life Technologies) in Phosphate Buffered Saline (PBS pH7.4, #10010-015, Gibco, Life Technologies) followed by 5 min incubation at 37˚C. Trypsin was inactivated with FBS-supplemented DMEM.

Iscove's Modified Dulbecco's Medium (IMDM) (500ml, [+]L-Glutamine, [+]25mM HEPES (#21980-032, Gibco, Life Technologies) was supplemented with 20% Heat-Inactivated Fetal Bovine Serum (Sigma-Aldrich) and 1% GlutaMAX™-I CTS™ (100X) (#A12860-01, Gibco, Life Technologies). RPMI Medium 1640 (RPMI) (500ml, [+]L-Glutamine (#21875-034, Gibco, Life Technologies) was supplemented with 10% Heat-Inactivated Fetal Bovine Serum (FBS, #10500-064, Gibco, Life Technologies) and 1% GlutaMAX™-I CTS™ (100X) (#A12860-

01, Gibco, Life Technologies). Dulbecco's Modified Eagle's Medium (DMEM) (500ml, [+]4.5g/L D-Glucose, [+]L-Glutamine, [-]Pyruvate DMEM (#41965-039, Gibco, Life Technologies) supplemented with 10% Heat-Inactivated Fetal Bovine Serum (FBS, #10500-064, Gibco, Life Technologies) and 1% GlutaMAX™-I CTS™ (100X) (#A12860-01, Gibco, Life Technologies).

Cell line	Media	Translocation	Leukeamic phenotype
SEM	IMDM	MLL-AF4	Pro-B cell
RCH-ACV	RPMI	E2A-PBX	Pre – B cell
MV4;11	RPMI	MLL-AF4	AML
THP1	RPMI	MLL-AF9	AML
K562	RPMI	BCR-ABL	Erythroleukaemias
HEK293T	DMEM		

Table 1 Cell lines used in this study. Growth media, translocation and leukaemic phenotype are indicated for each cell line.

2.1.1 Mycoplasma Testing

Supernatants of cell cultures were routinely tested for *Mycoplasma* contamination using MycoAlert Mycoplasma Detection Assay (Lonza) as per the manufacturer's instructions.

2.1.2 Storage and Recovery of Cell Lines

Cells were frozen at 80°C over night in 1ml 90% FBS, 10% Dimethyl-Sulphoxide (DMSO Hybri-MAX®, #D2650, Sigma, Life Science) using a Cryo 1˚C Freezing Container (#5100-0001, Nalgene). For long term storage cells were kept in liquid nitrogen.

Cells were thawed by brief incubation at 37˚C followed by addition of 5ml suitable media. Cells were spun for 5min at 1,200 rpm and cell pellets were resuspended in fresh media and maintained as described previously.

2.1.3 siRNA transfection of cells

2×10^6 SEM or MV4;11 cells or 5×10^5 THP-1 cells in log phase growth were transfected with 1-2 µg siRNA (optimised according to cell type and target) using an Amaxa Nucleofector (Lonza AG) with Nucleofector Kit V, program T-020 (for SEM cells), V-001 (for THP-1 cells) or Nucleofector Kit L program Q-001 (for MV4;11 cells). Cells were cultured for 24hr (for *Fli-1* knockdownds) or 48 hr (for *EZH2*, *UTX*, *EED* and *MLL-AF4* knockdowns) and re-transfected as described above. Cells were collected 48h post the second transfection for downstream applications.

The following specific siRNAs were used: MLL-AF4 siRNA and control sequences, were siMM (scrambled control, siMA6 (MLL-AF4 siRNA in SEM cells) are from (Thomas et al., 2005). EZH2 siRNA (Dharmacon L-004218-00-0005), UTX siRNA (Dharmacon L-014140-01-0005), EED siRNA (Dharmacon L-014140-01-0005) are all SMARTpool: ON-target plus (Dharmacon) siRNAs and knockdown with these were compared to control knockdowns of a non-targeting smartpool control (Dharmacon On Target plus non targeting pool D001801020); *Fli-1* siRNA: (Life Technologies stealth siRNA, sequence from (Suzuki et al., 2009))

versus a non targeting control (Life Technologies, Stealth negative Control Medium GC, 12935).

2.1.4 Leukaemic growth assays

2.1.4.1 Colony assays coupled to siRNA knockdowns

48 hours after the second transfection, THP1, SEM cells were plated at a density of 1×10^5 cells/ml, in triplicate, in IMDM MethoCult media (H4100; STEMCELL Technologies) supplemented with 20% fetal calf serum and cultured for 14 days (37°C, 5% CO_2) before counting.

48 hr post second transfection MV4;11 cells were plated at a density of 2.5×10^5 cells per ml, in triplicate in 30% FBS with 10^{-4} M (0.1mM) 2-mercaptoetanol supplemented IMDM Methocult media (H4100; STEMCELL Technologies) and cultured for 14 days (37°C, 5% CO_2) before counting.

2.1.5 Colony assay coupled to small molecule inhibitors

SEM and K562 cells were plated at a density of $0.75x\ 10^5$ cells/ml in triplicate in IMDM MethoCult media (H4100; STEMCELL Technologies) supplemented with 20% fetal calf serum. RCH-ACV cells were plated at a density of $0.75x\ 10^5$ cells/ml in triplicate in IMDM MethoCult media (H4100; STEMCELL Technologies) supplemented with 30% fetal calf serum and 10^{-4} M (0.1mM) 2-mercaptoetanol.

UNC1999, UCN2400 and GSK-J4 were added on the day of plating directly to the methycult media mixed with cells. Final concentration of DMSO was 0.01% for all inhibitor concentrations and the control cells.

All colonies were cultured for 14 days (37°C, 5% CO_2) before counting.

2.1.6 Colony assays on CD34+ cord and adult blood cells

Colony assays on CD34+ and adult blood cells were performed and analysed by Dr. Phillip North. Briefly, blood buffy coats containing CD34+ cells were purchased from NHS-BT. CD34+ cells were separated using MaxMacs separation kit (Mylteny Biotech). Cells were plated in triplicate at a density of 250 cells/ml in MethoCult™ H4435 Enriched semi-solid media and cultured for 14 days (37°C, 5% CO_2) before counting.

2.1.7 Liquid proliferation assays

The CyQUANT® Direct Cell Proliferation Assay (Thermo Fishcer Scientific) was used to asses proliferation. Briefly, addition of a fluorescent nucleic acid stain (ex/em 507/527) and a cell impermeable fluorescence suppressor dye allows for determination of cell viability. Standard curves were generated by serial dilution of cells in 96 well plates to determine the linear range of the signal, and all experiments were carried out within these limits. Briefly, cells were mixed with the desired dose of inhibitor and plated at a density of 2 x 10^5 cells/ml in

quadruplicate on 96 well plates. After 5 days, CyQUANT dye was added in a 1:1 ratio as per manufacturers instructions and incubated for 1h at 37C.

2.1.8 Liquid proliferation assay coupled to dexamethasone treatment

10 point 3-fold serial dilutions starting at 1000μM of dexamethasone (Sigma-Aldrich D4902-100MG) were performed in quadruplicate on 96 well plates. SEM cells, pre-mixed with varying doses of the UNC1999 inhibitor or DMSO, were added to the plates to a final cell density of $2x10^5$cells/ml. Cell viability was measure using the CyQUANT® Direct Cell Proliferation Assay (Thermo Fishcer Scientific). EC50 were calculated using a four-parameter-dose-response-curve model.

2.2 Gene expression methods

2.2.1 RNA Extraction

RNA was extracted using the Qiagen RNeasy Mini Kit (#74106) as per manufacturer's instructions. RNA was stored at -80˚C.

2.2.2 Generation of cDNA

100-500ηg of RNA was mixed with 1μL dNTP Mix (10mM, #18427-013, Invitrogen) and 1μL random hexamers (50μM, #S06405, Roche), incubated at 65˚C for 1min and cooled on ice for 1min. 4μL First-Strand Buffer (5X, #y02321, Invitrogen), 2μL DTT (0.1M #y00147, Invitrogen), 1μL RNaseOUT (40units/μL, #51535, Invitrogen) and 1μL SuperScript III Reverse Transcriptase

(200units/µL, #56575, Invitrogen) was added to the samples. Samples were made up to 20µL with water. The reaction was incubated on a PCR machine with the following conditions: 25˚C for 10min (primer extension), 50˚C for 50min (cDNA synthesis), 85˚C for 5min (Termination). cDNA was diluted to 200µL with water and stored at -20˚C for downstream analysis.

2.2.3 RT-qPCR

ChIP DNA or cDNA was used in Real Time Quantitative PCR (RT-qPCR) for the detection of gene targets or gene expression, respectively. For SYBR green primers, 5µL sample was added to 20µL reaction mix (12.5µL Fast SYBR® Green Master Mix (#4385612, Applied Biosystems); 0.25µL forward primer (10µM); 0.25µL reverse primer (10µM); 7µL water) in a single well of a 96-well PCR plate (MicroAmp Fast Optical 96-well Reaction Plate, #4346906, Applied Biosystems). For TaqMan primers, 5µL sample was added to 15µL reaction mix (10µL TaqMan® Fast Advanced Master Mix (#4444557, Applied Biosystems); 1µl primer/probe mix (Invitrogen) 4µL water) in a single well. Plates were sealed with film (MicroAmp Adhesive Film, #4311971, Applied Biosystems) and spun at 1000rpm for 1 minute. Reactions were performed on the 7500 Fast Real Time PCR System (Applied Biosystems). Table 2 list cycling conditions used.

Table 2.3 RT-qPCR Run Conditions

Temperature (˚C)	Time (s)	Cycles
95	20	1
95	3	40
60	30	
95	15	Continuous
60	60	
95	15	
60	15	

2.2.4 Primer Design for SYBR Green qPCR

Primers were designed using Primer-Blast based on sequences obtained from the UCSC genome browser. The product size range was limited to 60-150bp, with primer size between 20-25bp, melting temperatures between 58-61C and a maximum Tm difference of 1C. Primers for detection of cDNA through rt-qPCR were designed to span an intron. Primers were order from IDT or Sigma-Aldrich. Primers for amplification of ChIP-qPCR material were tested for efficiency by performing a serial 10-fold dilution curve on input material. Primers for cDNA detection were tested for efficiency by performing a serial 5-fold dilution curve. Only primers with good PCR efficiency yielding a single amplification product (as assessed through melt curve analysis) were used. Due to the large number of primer pairs used in this study their sequences are not included but are available upon request.

2.2.5 Nascent-RNAseq extraction

The protocol for 4sU-tagging, thiol-specific biotinylation and separation of nascent-RNA was adapted from Raedle et al. (2013).

RNA labelling:

1x10^7 cells treated with either DMOS, 5μM UNC2400 or 0.25μM, 1μM or 5μM UNC1999 for 5 days were used for nascent-RNA extractions.

4-thiouridine (Sigma, T4509 4-thiouridine, 100mg) was added to cell culture medium to a final concentration of 50μM. Cells were incubated for 1 hour at 37°C in 5% $C0_2$. Media was removed and 1ml Trizol (Invitrogen) was added, shaken vigourously and incubated at rt for 5min. 200μl Chloroform was added and shaken vigorously for 15 sec, followed by a 3min incubation at rt. Samples were centrifuged at 13,000 g for 15 min at 4°C. The aqueous upper phase was transfer to a new 1.5ml eppendorf tube and 1x volume of isopropanol was added. Samples were mixed and incubated at rt for 10 minutes, followed by centrifugation for 10 min at 4°C. The supernatant was removed, and the pellet was washed with 1ml ethanol until it dissolved. Samples were centrifuged 13,000 g for 10 min at 4°C and the ethanol was removed. Samples were resuspended in 1xTE and incubated at 65°C for 10 min. Samples were treated with the TURBO DNA-*free*™ Kit (Thermo-fisher) as per manufactures instructions .

20μg of RNA was used for labelling reactions, incubated for 1.5 hours with rotation:

- 2 µl Biotin-HPDP (Pierce, 50mg EZ-Link Biotin-HPDP, Cat. Nr. 21341) (1mg / ml DMF) per 1 µg RNA
- 1 µl 10x Biotinylation Buffer (100 mM Tris pH 7.4, 10 mM EDTA) per 1 µg RNA
- Water up to 500µL

An equal volume of Chloroform/Isoamylacohol (24:1) was added, mixed vigorously and incubate for 2 – 3 minutes at rt in phase lock gel heavy tubes Eppendorf (Order No. 0032 005.152). Samples were centrifuged at 13 0000g for 5min. Superntatnat was transferred into a new 1.5 ml tube. RNA was precipitated by addition of 1/10 reaction volume 5M NaCL and 1x volume isopropanol and centrifuged at 13 0000g for 20 min at 4C. Pellets were washed with 75% ethanol and resuspended in 100 µl 1x TE. Samples were mixed with 100µL streptavidin beads (µMacs Streptavidin Kit (Miltenyi, Order No. 130-074-101) and incubated with rotation for 15min. Samples were added to pre-equilibrated columns µMacs columns, and the columns were washed 3x with 65C wash buffer (100 mM Tris pH 7.5, 10 mM EDTA1 M NaCl, 0.1% Tween20) and 3x with room temperature buffer. Samples were eluted directly into 700µl RLT buffer by addition of 100µl 100mM DTT to the columns, twice.

The Rnease MiniElute Cleanup kit was used as indicated in (Raedle et al., 2013).

2.3 Protein techniques

2.3.1 Nuclear extracts

Nuclear Extracts were prepared from NOMO-1 and SHI-1 cells using a modified Dignam protocol (Dignam et al., 1983). Nuclear extracts for all other cell lines were previously available in the lab. Briefly, 5 l of cells were grown up to a density of 1-2 X 10^6/mL. Cell pellets were rinsed in PBS and dounce homogenized in a hypotonic buffer containing 10mM Tris-HCl pH 7.5 (pH 7.3 at 4°), 10mM KCl, 1.5mM $MgCl_2$ in the presence of Roche protease inhibitors (catalog# 05056489001). Once pelleted the nuclei were resuspended in a low salt buffer (20 mM HEPES, pH 7.9 at 4°C, 25% glycerol, 1.5 mM $MgCl_2$, 0.02 M KCl, 0.2 mM EDTA, 0.5 mM DTT and Roche protease inhibitors) then a high salt buffer (20 mM HEPES, pH 7.9 @ 4°C, 25% glycerol, 1.5 mM $MgCl_2$, 1.2 M KCl, 0.2 mM EDTA, 0.5 mM DTT) was added, dropwise, with stirring. Nuclear extract supernatant was collected after pelleting the chromatin and quantified using a BradfordUltra (expedeon) solution and compared to a standard curve generated using known concentrations of BSA (New England Bioloabs).

2.3.2 Whole cell extracts for

Whole cell extracts were prepared by collecting 1x10^6cells, washing them with PBS and resupending them in 1x NuPAGE® LDS Sample Buffer diluted in PBS. Samples were bolied at 95°C for 5 min, cooled and sonicated on a Bioruptor

(Diagenoned) at a high frequency, with 30s on 30s off for 8 min. Samples were used for downstream applications or stored a -20°C.

2.3.3 Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

15µl whole cell extracts were loaded onto pre-cast gels. For detection of all proteins except MLL-AF4 and MLL, pre-cast 4-12% Bis-Tris gels (Novex, Life Technologies) were used. For detection of MLL-AF4 pre-cast 3-8% Tris-Acetate gels (Novex, Life Technologies) were used. 180V were supplied to the gels assembled in a tank filled with either 1X MOPS SDS Running Buffer (NuPAGE® MOPS SDS Running Buffer (20X), #NP0001, Novex, Life Technologies) (Bis-Tris gels) or 1X Tris-Acetate SDS Running Buffer (#LA0041, Novex, Life Technologies) (Tris-Acetate gels), diluted in water. Gels were run until the bromophenol blue had migrated off the gel.

Proteins were transferred to a PVDF membrane (Immobilon®-P Transfer Membrane, #IPVH00010, EMD Millipore) using a wet transfer setup. Transfer buffer: 192mM glycine; 25mM Trizma® base; 10% methanol. 100V was supplied for 1 hour.

2.3.4 Western Blotting (WB)

All antibody dilution were prepared in 5% milk-TBS except EZH2 and SUZ12 which were prepared in 5% BSA -TBS with 0.01% Tween-20. Membranes were

incubated with antibody overnight at 4˚C with gentle rocking, followed by three 30min washes with 0.05% TBS-Tween. Membranes were incubated with 1:10,000 appropriate HRP-conjugated secondary antibody diluted in 5% milk at room temperature for 1h with gentle rocking, followed by three 30min washes with 0.05% TBS-Tween. Proteins were detected using either the ECL Prime Western Blotting Detection Reagent, #RPN2232 (Amersham, GE Healthcare) or the Pierce ECL Western Blotting Substrate, #32106 (Thermo Scientific) as per manufacturers instructions. Membranes were developed using X-ray film (Amersham Hyperfilm ECL, #28906837, GE Healthcare) and an X-ray developing machine (Xograph Imaging).

2.3.5 Chromatin Immunoprecipitation (ChIP)

For ChIPqPCR applications in SEM cells $1x10^7$cells were used. For ChIPseq applications $4x10^7$ SEM cells were used .For ChIPqPCR of mouse spleen samples, $5x10^5$ cells were used.

For detection of proteins, cells were fixed with 2mM disuccinimidyl glutarate (DSG, #80424-50MG-F, Sigma) for 30min at room temperature (rt) with rotation followed by fixation with 1% formaldehyde (FA) for 30 min at rt with rotation. For detection of histone modifications cells were fixed with 1% FA for 10 min at rt with rotation. Fixed material was spun at 13,000rpm for 1 minute, washed with PBS and stored at -80°C.

Fixed cell pellets were lysed with SDS lysis buffer (50mM Tris-HCl pH8.1; 10mM EDTA; 1% SDS) (100µl lysis buffer/1x10^7 cells) supplemented with 1:200 protease inhibitor cocktail (ROCHE) and syringe-passaged through a 27-gage (27G). Samples were sonicated in AFA Fiber Pre-Slit Snap-CAP 6x16mm microTUBEs (#520045, Covaris Inc.) using a S220 Covaris Inc. Sonicator. Sonication conditions were as follows: temperature = 4°C, duty cycle = 10%, intensity = 5, cycles/burst = 200, duration = 300s. Single fixed material was sonicated for 5min while double fixed material was sonicated for 10min.

Samples were spun at 13,000rpm for 10min at 4°C and the supernatant was diluted 12x with ChIP dilution buffer (CDB, 16.7mM Tris-HCl pH8.1; 167mM NaCl; 1.2mM EDTA; 0.01% SDS; 1.1% Triton X-100 (#X100-100ML, Sigma). To reduce background signal samples were incubated with 1:600 Protein A/G Dynabeads resuspended in CDB buffer for 30min at 4°C with rotation. The dynabeads were pelleted and the supernatant was transferred to a fresh tube. This was repeated twice. 1ml of samples were incubated with rotation over night at 4°C with desired antibody. 3x 50µl aliquots of samples were taken for inputs and for detection of sonication efficiency.

ChIP samples were incubated with 10uL of Protein A/G Dynabeads resuspended in CDB for 5h at 4˚C with rotation. The dynabeads were washed 2x with cold RIPA buffer (50mM HEPES-KOH pH7.6; 500mM LiCl; 1mM EDTA; 1% NP-40; 0.7% Na-Deoxycholate) using a magnet, followed by a wash with TE + 50mMNaCl. Samples were spun at 3,200rpm for 3 minutes and beads were re-

suspended in 110µL in the SDS lysis buffer (50mM Tris-HCl pH8.1; 10mM EDTA; 1% SDS). Chromatin was eluted by incubation at 65°C for 30min. Samples were spun at 13,000 rpm for 1 min and 100µL of the supernatant was collected. Cross-linking was reversed through incubation at 65°C over night.

RNA was digested through addition of 1µL RNAse A (10mg/mL, #EN0531, Thermo Scientific) to samples followed by incubation at 37˚C for 1 hour. Samples were then treated with 1µL Proteinase K (20mg/mL, #EO0491, Thermo Scientific) at 45˚C for 1 hour. DNA was purified using a PCR Purification kit (#28106, Qiagen), and eluted in 60-140µL water. Samples were stored at -20˚C for further use.

2.3.6 ChIP-Rx

ChIP-Rx was performed exactly as described for ChIPseq experiments (2.3.5) with the following amendment: FA Fixed *Drosophila melanogaster* S2 cells were lysed (fixation and lysing conditions as described in 2.3.5) and added to human samples treated the same way, in a 1:4 ratio. The normal ChIPseq protocol was followed from that point onwards. *Drosophila melanogaster* S2 cells, were kindly provided by Dr. Qianxin Wu.

2.3.7 Preparation of samples for ChIP-RX

SEM cells were seeded at $2x10^5$ cells/ml with 0.25μM, 1μM 5μM UNC1999, 5μM UNC2400 or DMSO as a control, and incubated for 72h. Cells were counted and seeded at $5x10^5$ cells/ml with media containing the appropriate concentrations of UNC1999, UNC2400 or DMSO. 48 hours later cells were harvested and fixed for ChIP-Rx.

2.4 Next-Generation Sequencing (NGS) Techniques

2.4.1 ATACseq

The ATACseq protocol was adapted from (Buenrostro et al., 2013). Briefly, 50 000 SEM cells were rinsed in cold PBS and resuspended in cold lysis buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl2, 0.1% IGEPAL CA-630). Cell were spun down immediately at 500xg for 10min at 4C. The supernatant was discarded and the pellet was resuspended in: 25 μL 2x TD Buffer (Illumina Cat #FC-121-1030), 2.5 μL Tn5 Transposes (Illumina Cat #FC-121-1030), 22.5 μL Nuclease Free H2O. The reaction was incubated for 30min at 37°C. DNA was purified using a Qiagen MinEluteKit as per manufactures instruction. DNA was eluted in 10μL 10mM Tris buffer, pH8. Samples were stored at -20C for downstream applications (ATACseq).

2.4.2 Preparation of libraries for ChIP-Rx and ATACseq

The NEBNext® Ultra™ DNA Library Prep Kit for Illumina® was used for preparing all ChIPseq and ChIP-Rx libraries, as per manufactures instruction. Libraries prepared in this study include: ChIP-RX for H3K27me3, H3K4me3, and H3K27Ac in SEM cells. ATACseq in SEM cells.

2.4.3 Nascent-RNA seq library preparation

The NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina® (NEB) was used for preparation of nascent-RNAseq libraries, per manufacturers instruction.

2.4.4 Sequencing

Prior to sequencing all libraries were quality controlled by analysis on a 2200 TapeStation instrument (Agilent technologies). Libraries were quantified using KAPPA Universal Complete Kit (KK4824, Kappa Biosystems) as per manufactures instruction. All libraries were sequenced on a NextSEQ500 (Illumina).

2.5 Bioinformatics Analysis of NGS data

All in house scripts were written by Jonatahn Kerry. Table 2 lists ChIPseq data sets that were mapped and analysed for the purpose of this thesis but were generated by someone else.

Protein used for ChIPseq	Person who performed the ChIPseq experiment
EZH2	Dr. Thomas Milne
SUZ12	Dr. Thomas Milne
RING1B	Dr. Thomas Milne
MLL-N	Dr Erica Ballabio
AF4-C	Dr Erica Ballabio

Table 2 ChIPseq data generated by someone else

Pair-ended ChIPseq/ATACseq data was mapped with a pipeline developed by Jelena Telenius and hosted by the Genome Biology Research Group. Version 10.1 of the "DnaseANDChIP" pipeline was used. Briefly, the pipeline performs quality control checks of ChIPseq fastq files, trims adaptors and merges very short reads if needed. Reads are then mapped to a selected reference genome and duplicate reads are removed. It employs the following modules: samtools/0.1.19, bedtools/2.17.0, fastqc/0.10.1 bowtie/1.0.0, cutadapt/1.2.1 and trim_galore/0.3.1. ChIPseq data was mapped against the human genome (hg19) and ChIP-Rx data was mapped against both the human (hg19) and the Droshophila genomes (dm3) allowing for two mapping events per fragment,

with distance between paired end reads set to 350bp. ATACseq data was mapped against the human (hg19) genome with the same parameters except the distance between paired end reads was set to 700bp.

2.5.1 Visualisation of ChIPseq/ATACseq tracks

The HOMER tool (vs 4.7) (Heinz et al., 2010) was used for generation of tracks for visualisation in the UCSC genome browser. To allow for processing of files by HOMER, Tag directories had to be created from mapped ChIPseq bam files. This was performed with default parameters of the makeTagDirectory tool except for the flag –keepAll (in order to retain all the reads in the bam file). To visualise tracks, bed graphs were generated from the Tag directories and normalised to 10^7 million reads, using the makeUCSCfile tool.

2.5.2 Identification of enriched regions (Peaks) and associated genes

To identify regions of enrichment in the ChIPseq/ATACseq data the HOMER tool findPeaks was used. For ChIPseq the –style region function was employed to allow for distance and size of peaks to be defined. All peaks were defined taking into account the corresponding input tracks of each ChIPseq experiment. The peak finding parameters (with respect to size of peaks, merging distance between two peaks and enrichment over input) varied for each data set analysed. Fold enrichment over input was always set to at least 4, the poisson p-value threshold relative to input tag count was 0.0001 (default) and the false

discovery rate threshold was set to 0.0001 (default). To remove false peaks, peak files were uploaded to the Multi-Image Genome Viewer (McGowan et al., 2013) were they were ranked by tag count (of each peak) and visually inspected. For ATACseq the –factor function of *findPeaks* was employed, which uses a fixed peak size. All other parameters were kept the same.

Gene targets of proteins/histone modifications were determined by matching peak sets to the hg19 *RefSeq* gene list using the in house *Peak_PromoterGeneTargets_hg19.pl Perl*. This determined gene targets if the peak co-ordinates overlapped the gene at +/- 2kb of the TSS (from now on referred to as promoters). For overlap analysis of gene targets identified in separate experiments, the in house *OverlapPeakGeneList.pl Perl* script was used, which overlaps targets based on the gene name.

2.5.3 Enrichment analysis

Histograms of ChIPseq data were generated using HOMERs *annotatePeaks.pl* tool with the –hist flag. Interval/region size was set to 25bp for every analysis. Resulting units represent reads per base-pair per region (25bp). When traditional normalisation was employed these the units were normalised to reads per 10^7 reads. For ChIP-Rx normalisation the –noadj flag was added to prevent normalisation of reads. These resulting enrichments were then normalised with the reference derived normalisation factor. The same parameters were used for generating heatmaps using the –ghist flag instead. Heatmaps were visualised using treeview

To calculate fold loss/gain of H3K27me3 and H3K27Ac the number of reads at +/- 2kb of TSS of selected genes was used. These were extracted from bam files using the in house *ReadsUnderHG19genes.pl Perl* script. Reads were normalised to 10^7 reads (traditional normalisation ChIPseq) or the raw read numbers were normalised using the reference derived normalisation factor (ChIP-Rx). For calculation of H3K4me3 enrichment the same script was used but the enrichment under H3K4me3 peaks at specific gene targets was analysed rather than the enrichment under the promoter.

2.5.3.1 Determination of the Normalization Factor in ChIP-Rx.

For a detailed description of the basis and derivation of the ChIP-Rx normalisation factors, please refer to (Orlando et al., 2014). In brief, Orlando et al. derived a normalization constant, α, such that after normalization the signal per-reference cell (β) is the same across all samples. The total ChIP-seq signal derived from reference cells is simply the count of reads (in millions) aligning to the *Drosophila* genome, which is represent as N_d. Because the percentage of reference cells as a fraction of the total number of cells is constant and the assumption is made that the epigenome of the reference cells does not vary appreciably, α can be derived as

$$\alpha * N_d = \beta$$

Because β is a constant, we can simply rewrite this as

$$\alpha * N_d = 1$$

or

$$\alpha = \frac{1}{N_d},$$

multiplying the read counts by α produces a normalized read count in normalized RRPM.

2.6 Nascent-RNAseq Analysis

2.6.1 Mapping and identification of differentially expressed genes

Following QC analysis with the fastQC package (http://www.bioinformatics.babraham.ac.uk/projects/fastqc), reads were aligned using STAR (Dobin et al., 2013) against the human genome assembly (NCBI build37 (hg19) UCSC transcripts). Non-uniquely mapped reads and reads that were identified as PCR duplicates using Samtools (Li H et al., 2009) were discarded. Gene expression levels were quantified as read counts using the featureCounts function (Liao et al., 2014) from the Subread package (Liao et al., 2013) with default parameters. The read counts were used for the identification of global differential gene expression between specified populations using the edgeR package (Robinson et al., 2010). RPKM values were also generated using the edgeR package. Genes were considered differentially expressed between populations if they had an adjusted *p*-value (FDR) of less than 0.05.

2.7 *In vivo* mouse work

All *in vivo* work was performed by Dr Frida Ponthan as collaboration with the group of Prof. Heidenreich (Newcatlse University). Xenografts and bioluminescence imaging were performed as described in (Bomken et al., 2013). Briefly, 10 000 SEM (pSLIEW clone 13 expressing GFP and luciferase), cells were injected intra-femurally into NSG mice. Seven days later treatment of UNC1999 was started. Mice were given 50mg/kg UNC1999 once per day for 5 days followed by two days rest, until mice had to be culled at day 35. Leukaemic burdn was assessed by bioimagingin every 7 days, as described previously (Bomken et al., 2013).

2.7.1 Toxicology studies

UNC1999 toxicology studies in NSG mice and RAG2 KO mice used 10 mice each. 5 were given vehicle and 5 were given UNC1999. Mice were given increasing doses of UNC1999 for 3 cycles. Each cycle consisted 5 days of dosing and 2 days rest.
The amount of inhibitor used in NSG toxicologyy studies was : 50mg/kg, 100mg/kg and 150mg/kg. Vehicle was 0.5% NaCMC, 0.1% TWEEN-80 in water.

Amount of inhibitor used in RAG2 KO toxicology studies: 150mg/kg, 200mg/kg. Vehicle was: 10% NMP (N-Methyl-2-pyrrolidone) in 75mM sodium citrate.

3 Chapter 3

The poor prognosis of patients with t(4;11) leukaemias (with a median overall survival of 10 months in children and 7 months in adults) means that new treatment strategies are needed. The rapid onset of t(4;11) leukaemias in infants, the high concordance rate of leukaemia in monozygotic twins and the lack of additional mutations in t(4;11) patients (Dobbins et al., 2013; Greaves, 2005; 1996), suggests that the t(4;11) translocation is sufficient for the leukaemic transformation of cells. Since knocking down MLL-AF4 is sufficient to stop leukaemic growth (Thomas et al., 2005), it is reasonable to suggest that everything we need to know about stopping the leukaemia from growing is contained within the activity of the MLL-AF4 fusion protein; specifically, how MLL-AF4 activates gene targets and what key gene targets it activates. The study of the direct targets of MLL-AF4 is therefore an important area of research as it is these that mediate leukaemic development. Previously, direct targets of MLL-AF4 have been identified in patient cell lines and mouse models (Guenther et al., 2008; Krivtsov et al., 2008; Wilkinson et al., 2013) including many haematopoietic transcription factors as well as chromatin proteins, although the function of only a few of these have been studied to date.

During haematopoietic development TFs serve to establish transcriptional networks that drive terminal differentiation by activating genes important for a specific lineage and restricting alternative cell fates (Zhu and Emerson, 2002). Since leukaemia is characterised by a block in differentiation, TFs regulated by

MLL-AF4 may be aberrantly expressed and thus contribute to leukaemogenesis by re-enforcing a differentiation block. For example, the TF *RUNX1* is a direct target of MLL-AF4 and has been shown to be important for leukaemic growth of t(4;11) patient cell lines (Wilkinson et al., 2013).

Chromatin proteins are interesting targets as not only are they involved in the regulation of gene expression, but they are also excellent targets for small molecule inhibitors. Identification of chromatin proteins with important functions in t(4;11) could lead to the development of novel therapeutic strategies. For instance, inhibitors of DOT1L have been shown to block the expression of MLL-FP target genes and consequently impair leukaemic progression (Daigle et al., 2011; 2013). This has led to the initiation of clinical trials to determine the safety and efficacy of the EPZ-5676 inhibitor for treatment of MLLr leukaemias (Daigle et al., 2013; Helin and Dhanak, 2013).

A candidate approach was chosen to identify MLL-AF4 targets that contribute to leukaemogenesis. Identification of direct MLL-AF4 targets was achieved by analysis of a previously published ChIPseq data set of MLL-AF4 targets (Guenther et al., 2008) with one available in the Milne Lab. Both data sets identified targets in the t(4;11) patient derived ALL cell line, SEM. Targets were validated as important for leukaemogenesis by combining proliferation assays with siRNA mediated knockdowns and small molecule mediated inhibition of selected targets.

3.1 Identification of MLL-AF4 targets

Previously, Guenther et al. identified 2,491 MLL-AF4 targets in the SEM patient cell line. Overlap analysis of these, with the 3,437 MLL-AF4 targets identified in the same cell line by the Milne Lab (unpublished data) was performed which yielded 1,738 high confidence gene targets bound by MLL-AF4. These included at least 27 chromatin proteins as well as 149 TFs (Figure 4). Among these was the Polycomb group H3K27me3 methyltransferase *EZH2*, the H3K27me3 demethylase *UTX* and the transcription factor *Fli-1*. It is worth mentioning that these targets were also bound by MLL-AF4 in another t(4;11) leukaemic patient derived cell line, RS4;11.

Fli-1 was chosen for further analysis as it has been found to act at the top of regulatory networks driving blood and endothelial development (Liu et al., 2008). It has been shown to play an important role in early haematopoiesis during hemangioblast formation as well as in haematopoietic stem cell maintenance (Kruse et al., 2009). In mice, overexpression of *Fli-1* leads to the development of erythroleukaemia (Truong and Ben-David, 2000) while in humans a translocation event fusing *Fli-1* to the *EWS* gene leads to the development of Ewing's Sarcoma (Delattre et al., 1992) . Additionally in AML the aberrant expression of *Fli-1* is associated with an adverse prognosis (Kornblau et al., 2011). The established role of Fli-1 as a factor involved in the malignant transformation of cells coupled with the lack of research of this gene in the

development of lymphoid leukaemia made it a good target for further investigation.

EZH2 is the catalytic component of the PRC2 complex, which has previously been shown to be important for MLLr AMLs (Smith et al., 2011; Tan et al., 2011; Neff et al., 2012; Tanaka et al. 2012; Shi et al., 2013; Thiel et al., 2013; Xu et al., 2014). However, the role of PRC2 in t(4;11) leukaemias had not been investigated and as such it was chosen for further study. Similarly, UTX has been shown to be important for the leukaemic growth of MLL-AF9 patient cell lines, but its role in t(4;11) leukaemias had not been investigated.

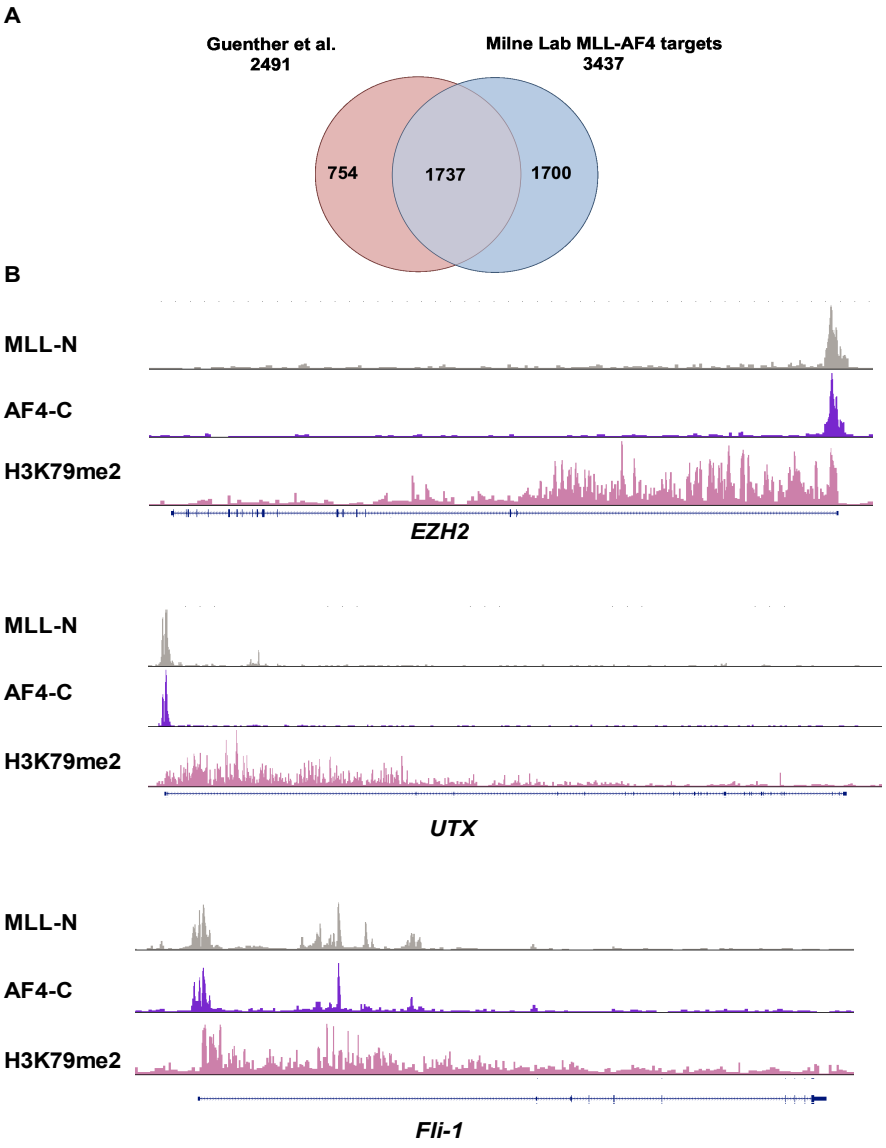


Figure 4 MLL-AF4 binds to *EZH2, UTX* and *Fli-1.*

A. Overlap analysis of MLL-AF4 targets identified by Guenther et al. and ChIPseq experiments performed in the Milne Lab. B. UCSC tracks showing binding of MLL-N, AF4-C to the promoters of *EZH2* (top panel), *UTX* (middle panel) and *Fli-1* (bottom panel), as well as H3K79me2, generated in the Milne Lab.

3.2 MLL-AF4 regulates EZH2, UTX and Fli-1

To determine whether MLL-AF4 was regulating the expression of *EZH2, UTX* and *Fli-1,* SEM cells were treated with an siRNA targeting *MLL-AF4* or a scrambled control siRNA. Knockdown of MLL-AF4 was confirmed through western blotting with a concomitant decrease in the levels of EZH2 and Fli-1. mRNA levels of *EZH2* and *UTX* were reduced by ~35% while those of *Fli-1* and *HoxA9* (included as a positive control) were reduced by ~60% (Figure 5 A and B). This suggests that MLL-AF4 directly regulates these targets.

To determine whether the expression of these proteins was unique or elevated in MLL-AF4 leukaemic cell lines, immunoblot analysis in nuclear extracts of multiple MLLr patient cell lines as well as in two non MLL-r lines was performed. Western blots were probed with antibodies against Fli-1, EZH2 and UTX (Figure 5C). Levels of Fli-1 and UTX were higher in (t4;11) cell lines compared to most other MLLr cells lines while the levels of EZH2 were high across most cell lines tested. Two non-MLL-FP leukaemic ALL cell lines, RCH-ACV and CCRF-CEM showed robust expression of all three proteins, suggesting that while MLL-FP may function to regulate the expression of these proteins other mechanism may be in place to maintain the expression of these gene targets in other leukaemias. Of note is that two MLLr cell lines only expressed these three proteins at low levels compared to the other cell lines examined. Additionally the ML2 cell line which express the MLL-AF6 fusion protein, had no detectable levels of Fli-1 and UTX while it expressed EZH2 at low levels. Altogether, these results suggest that

Fli-1,UTX and EZH2 may not be general MLL-FP targets, but all three are expressed in the three MLL-AF4 lines tested. Since *Fli-1, UTX* and *EZH2* all appear to be targets of MLL-AF4, their contribution to MLL-AF4 leukemic growth was assessed.

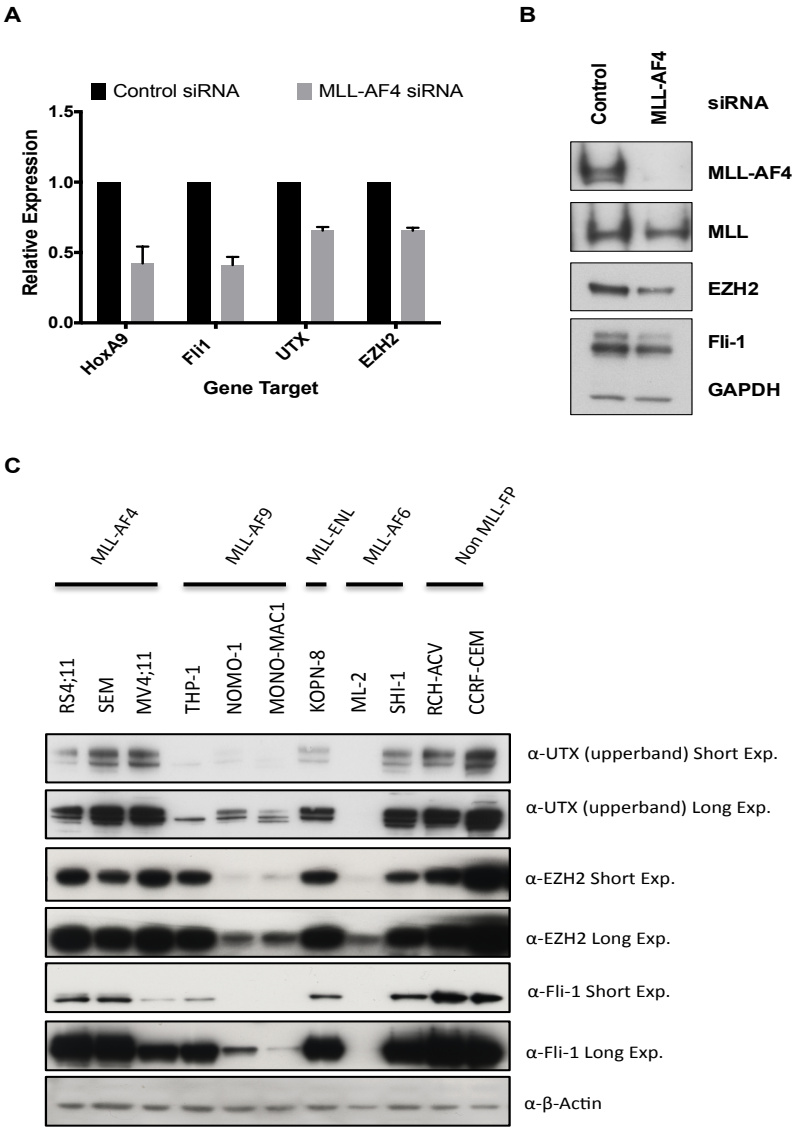


Figure 5 Knockdown of MLL-AF4 decreases expression of *EZH2, UTX* and *Fli-1.*

A. rt-qPCR expression of MLL-AF4 targets 96h after knockdown with either a scrambled control siRNA (black bars) or an MLL-AF4 siRNA (grey bars). Data are the mean +/- SD of two independent experiments. In each experiment control values were set to 1. Expression was normalised to *GAPDH.* B. Example western

blot of MLL-AF4 and its targets 96h after knockdown with either control or MLL-AF4 siRNA. Membranes were probed with the indicated antibodies except in the case of MLL-AF4, which was probed, with antibodies against MLL-N and AF-C (n=2). C. Western blots of EZH2, UTX and Fli-1 in nuclear extracts of indicated cell lines. Short exposures (Short. Exp) and Long exposures (Long. Exp) of membranes are shown. B-actin served as a loading control (n=1).

3.2.1 Role of Fli-1 in leukaemic growth

To determine whether or not Fli-1 was important for the leukaemic growth of t(4;11) cells, a methylcellulose assay was performed using the SEM cell line comparing the colony forming capacity between cells treated with either a non- target control siRNA or an siRNA targeting *Fli-1*. Treatment of SEM cells with the Fli-1 siRNA resulted in a ~90% knockdown of *Fli-1* mRNA levels as well as a reduction of Fli-1 at the protein level compared to cells treated with a non-targeting control siRNA (Figure 6A and B). This knockdown affected the colony forming capacity of SEM cells dramatically with Fli1-1 siRNA treated cells exhibiting a 90% decrease in colony forming capacity (Figure 6C). Similarly, knockdown of Fli-1 inhibited the proliferation of SEM cells in liquid culture with a 70% reduction in cell numbers after 5 days of growth (Figure 6E).

siRNA mediated knockdown of *Fli1* in the myeloid cell line THP1, which harbors the MLL-AF9 fusion protein also led to reduced colony formation suggesting that Fli-1 may play an important role in other MLL-FPs as well (Figure 6).

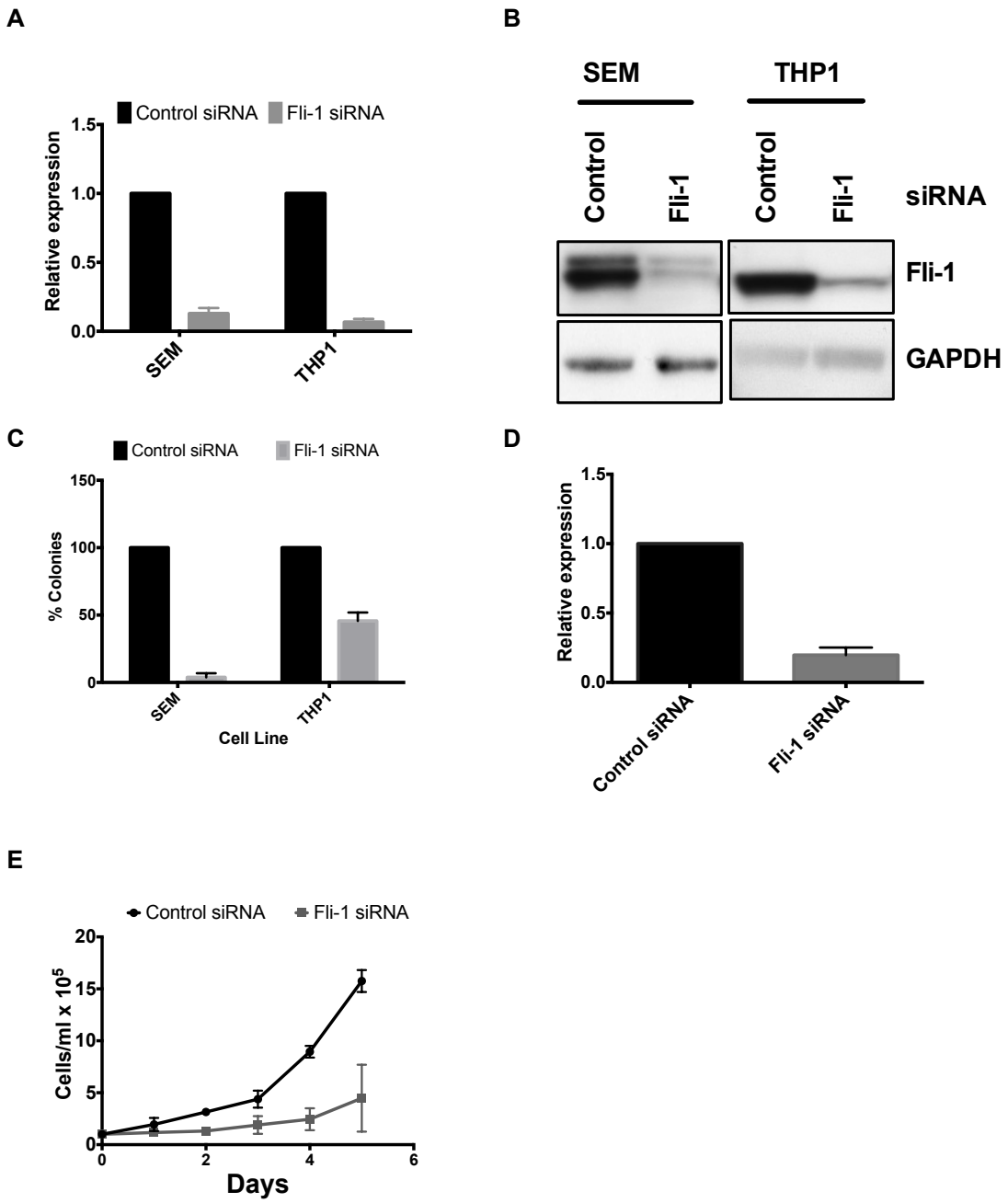


Figure 6 Fli-1 kockdowns affect the leukaemic capacity of MLL-AF4 and MLL-AF9 leukaemic cell lines.

SEM (MLL-AF4) and THP1 (MLL-AF9) cells were treated with either a Fli1 siRNA or a non-targeting control siRNA and 72h post transfection cells were used in colony assays or proliferation assays. A) rt-qPCR of *Fli-1* expression 72h post siRNA treatment, relative to control siRNA treated cells, normalised to GAPDH, related to colony assays shown in (C) B) Representative western blot of Fli-1 knockdowns relating to (A) and (D). C) Percent colonies formed following Fli1 siRNA knockdown relative to DMSO. E) rt-qPCR of *Fli-1* expression 72h post siRNA treatment, relative to control siRNA treated cells, normalised to GAPDH, related

to the proliferation assay shown in (E) Cumulative frequency curve of SEM cell numbers following knockdown of Fli-1. Error bars represent SD of 2-3 biological replicates. (n=3 SEM cells, n=2 THP1 cells).

To confirm the importance of Fli-1 in the SEM cell line, three stable cell lines expressing distinct IPTG-inducible *Fli-1* shRNA constructs were generated through lentiviral transduction of SEM cells. Proliferation of these was assessed following IPTG treatment and compared to untreated controls. Unlike treatment with siRNA, knockdown of *Fli-1* did not affect proliferation of SEM cells in colony forming assays (Figure 7C). Surprisingly, the 5325 stable cell line showed an increase in colony numbers upon knockdown of *Fli-1*. The lack of concordance between siRNA and shRNA mediated knockdowns of *Fli-1* suggest that additional events are perhaps responsible for the observed proliferation changes.

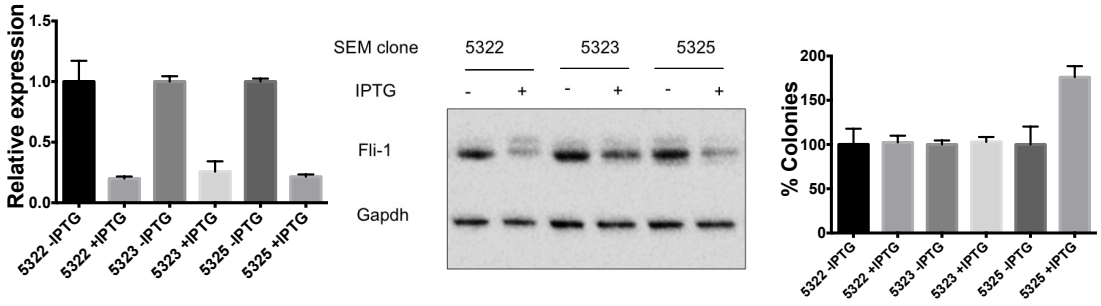


Figure 7 shRNA mediated knockdown of *Fli-1* does not affect leukaemic growth of SEM cells

Expression of Fli1 shRNA in stable cell lines (5322, 5323, 5325) was induced by addition of IPTG. 120h later cells were harvested and used in colony assays. A). rt-qPCR of *Fli-1* expression following induction of Fli1 shRNA with IPTG (+IPTG) compared to non-induced controls (-IPTG), expression levels were normalised to GAPDH, error bars represent SD of three rt-qPCR technical replicates. B) Western blot probed with Fli-1 and GAPDH antibodies C) Percent colonies formed in Fli1 shRNA induced (+IPTG) compared to non-induced controls (-IPTG). Error bars represent SD of three technical replicates.

3.3 Role of UTX in leukaemic growth

In order to assess whether UTX was contributing to the leukaemic growth of t(4;11) leukaemias, siRNA mediated knockdowns of *UTX* coupled to colony forming assays and liquid culture proliferation assays were performed in the SEM and MV4;11 (MLL-AF4, AML) cell lines. Proliferation in liquid culture was assessed by measuring fluorescence after addition of a fluorescent nucleic acid stain. Knockdowns of UTX resulted in a ~75% reduction in mRNA levels in both SEM and MV4;11 cells compared to controls, however no substantial changes in protein levels were observed. Consequently the ability of cells to form colonies in methylcellulose or to proliferate in liquid culture was not affected (Figure 8).

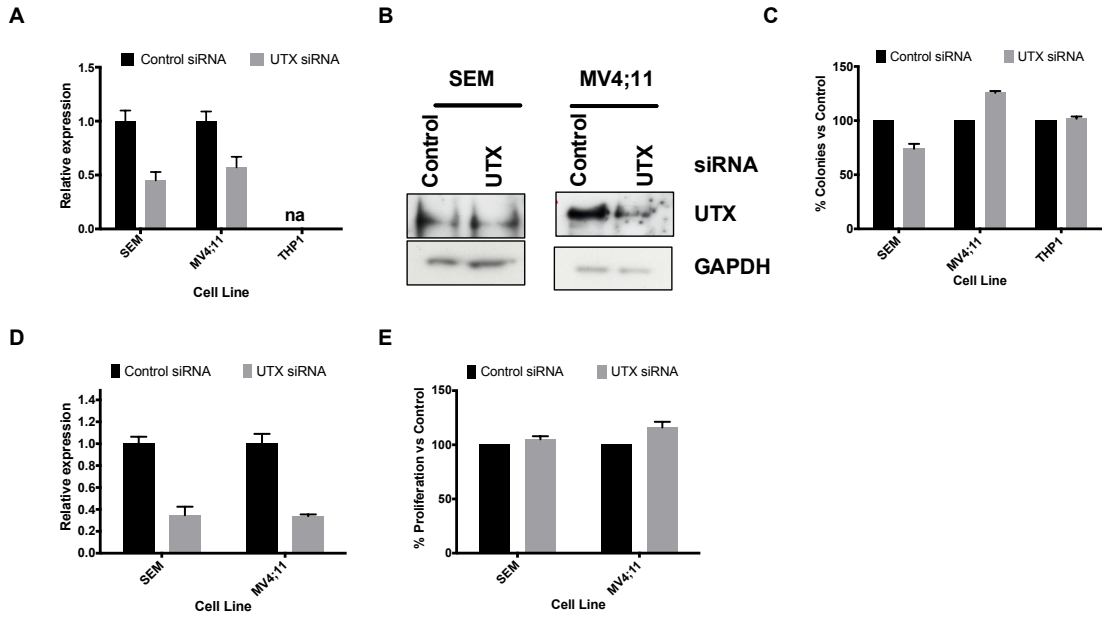


Figure 8 Knockdown of UTX in t(4;11) ALL and AML cell lines does not affect UTX protein levels or proliferative capacity of the leukaemias

SEM (MLL-AF4) and MV4;11 (MLL-AF4) cells were treated with either a UTX siRNA or a non-targeting control siRNA and 96h post transfection cells were used in colony assays or proliferation assays. A) rt-qPCR of *UTX* expression 120h post knockdown, relative to control siRNA treated cells, normalised to GAPDH, related to colony assays shown in (C) B) Representative western blot of UTX knockdowns relating to (A) and (E). C) Percent colonies formed following UTX siRNA knockdown relative to non-targeting control siRNA treated cells. E) rt-qPCR of *UTX* expression 120h post siRNA treatment, relative to control siRNA treated cells, normalised to GAPDH, related to the proliferation assay shown in (E) Percent proliferation cells following treatment with UTX siRNA compared to non-targeting control siRNA treated cells. Error bars represent SD of three technical repeats (n=1).

To assess whether pharmacological inhibition of UTX would be detrimental to leukaemic growth in liquid culture, SEM and MV4;11 cells were treated with two broad acting demethylase inhibitors, dMOG and dMPDCA. Treatment with dMPDCA was not effective at blocking proliferation even at concentrations of 100μM while SEM cells were sensitive to treatment of dMPDCA only at the

highest dose of 100μM (Figure 9A and B). Since theses inhibitors are not specific for UTX a third inhibitor, GSK-J4, reported to be specific for the H3K27me3 demethylase was assessed for its ability to affect colony formation of SEM cells. Colony formation was impaired in a dose-dependent manner with concentrations of GSK-J4 >1μM, with no colonies detected at 2μM treatments (Figure 9C). Further characterization of the GSK-J4 inhibitor by the SGC consortium showed it to be effective at inhibiting the H3K4me3/2/1 demethylases JARIDB/C as well. This complicates interpretation of any results with this inhibitor as the observed effect could be due to a combined effect of UTX/JMJD3 and/or JARIDB/C inhibition.

Since the siRNA mediated knockdowns of *UTX* were not able to reduce protein levels and did not inhibit leukaemic growth, combined with the lack of a specific UTX/JMJD3 inhibitor, research on the role of UTX in MLL-AF4 leukaemias was not pursued.

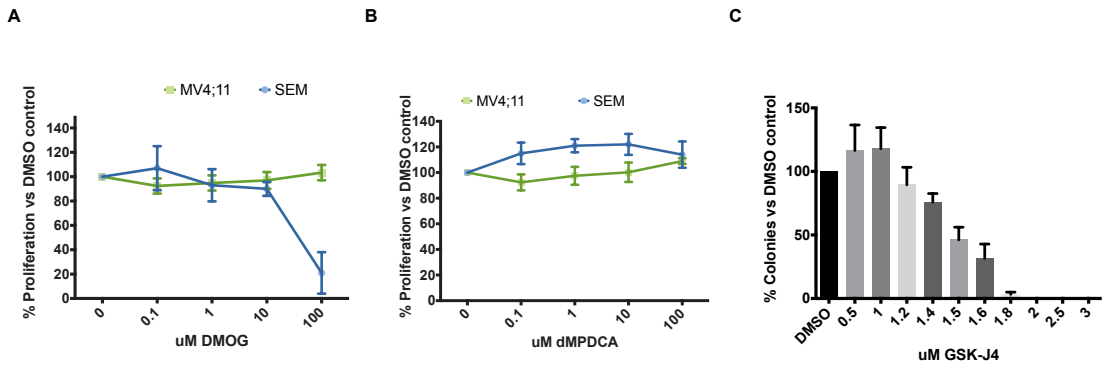


Figure 9 GSK-J4 affects the clonoegenic capacity of SEM cells

Percent proliferating cells following 5 day treatment with A) DMOG and B)dMPDCA relative to DMSO treated cells in the SEM (blue) and MV4;11 (green) cell lines. Error bars represent SD of quadruplicate

technical repeats (n=1) C) Percent colonies formed following GSK-J4 treatment of SEM cells relative to DMSO treated cells. Error bars represent SD of three technical replicates, (n=1).

3.4 Role of PRC2 in leukaemic growth

3.4.1 Genetic inactivation of PRC2

To determine whether EZH2 contributes to colony forming capacity of MLLr leukaemic cells lines, SEM, MV4;11 and THP1 cells were transfected with either a non-targeting control siRNA or an *EZH2* siRNA. This resulted in reduction of *EZH2* expression by at least 65% in SEM, MV4;11 and THP1 cells (Figure10A and B). Protein levels of EZH2 were also reduced in all three-cell lines. The colony forming potential of SEM was not affected significantly, while MV4;11 and THP1 cells showed a 44% and 31% reduction in colony numbers respectively, compared to cells treated with a non-targeting siRNA (Figure 10C).

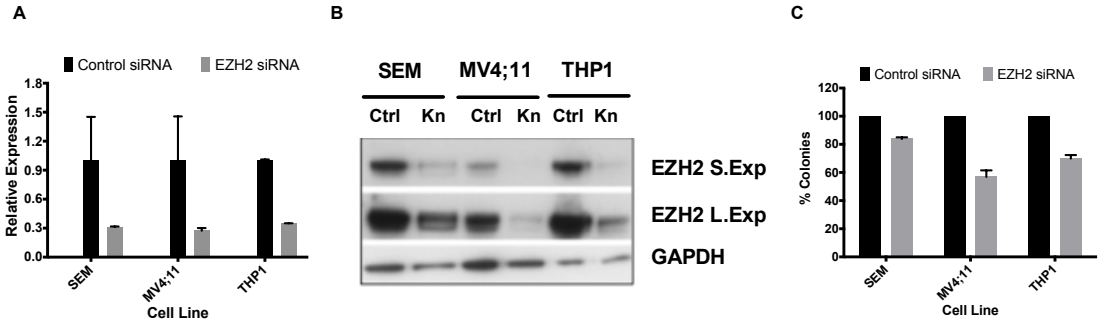


Figure 10 Knockdown of EZH2 does not affect the clonogenic capacity of SEM cells significantly

SEM (MLL-AF4), MV4;11 (MLL-AF4) and THP1 (MLL-AF4) cells were treated with either an EZH2 siRNA or a non-targeting control siRNA and 96h post transfection cells were used in colony assays A) rt-qPCR of *EZH2* expression 120h post knockdown, relative to control siRNA treated cells, normalised to GAPDH, related to colony assays shown in (C) B) Representative western blot of UTX knockdowns relating to (A). C) Percent

colonies formed following UTX siRNA knockdown relative to non-targeting control siRNA treated cells. Error bars represent SD of three technical repeats for MV4;11 and THP1 cells (n=1). Error bars represent SD of 2 biological repeats (n=2).

Previous research has shown that knockdown of either *EZH2* or *EZH1* individually is compatible with leukaemic growth of MLL-AF9 expressing cells, while knockdown of either *SUZ12, EED* or double knockdowns of EZH2 and EZH1 were not. Therefore, to inactivate both PRC2-EZH1 and PRC2-EZH2 siRNA knockdowns of *EED* were performed and these were coupled to colony forming assays. Knockdown of *EED* resulted in a ~60% reduction in *EED* mRNA levels compared to control siRNA treated cells (Figure 11A). EED protein levels were also reduced while H3K27me3 levels were undetectable by western blotting (Figure 11B). The successful inactivation of the methyltransferase activity of PRC2 by this knockdown resulted in a ~35% reduction in the colony forming capacity of SEM cells (Figure 11C).

To validate the requirement of PRC2 in the leukaemic growth of SEM cells a stable cell line expressing an shRNA targeting *EED* under the control of the tetracycline promoter was generated. The colony forming capacity as well as the proliferative capacity of this cell line in liquid culture was assessed following induction of the *EED* shRNA through doxycycline treatment and compared to an un-induced control. Treatment with doxycycline reduced *EED* mRNA levels by ~40% compared to untreated cells (Figure 11D). While no changes in the protein levels of EED were detectable by western blotting, H3K27me3 levels were reduced indicating that the function of PRC2 was inhibited by this knockdown

(Figure 11E). Surprisingly, the colony forming capacity of SEM cells was not affected by this knockdown and neither was the proliferation in liquid culture after 14 days (Figure 11F and G). Since the siRNA mediated knockdown of *EED* resulted in a more dramatic decrease in H3K27me3 levels, it was possible that the shRNA mediated knockdown of *EED* was not robust enough to impair the proliferative capacity of SEM cells (Figure 11B). Therefore two additional stable cell lines containing distinct *EED* shRNAs under the control of the tetracycline promoter were generated. Unfortunately, induction of the shRNA in these cell lines did not result in any changes in *EED* mRNA levels (data not shown). Together, these results suggest that t(4;11) cells are sensitive to the inhibition of PRC2 activity, but consistent with previous published results in MLL-AF9 leukaemias, this inhibitory effect is highly dose sensitive and thus is highly dependent on how efficient the knock downs are (Shi et al., 2013). To get around some of the defects inherent in siRNA mediated approaches the effects of pharmacological inhibition of PRC2 activity were assessed.

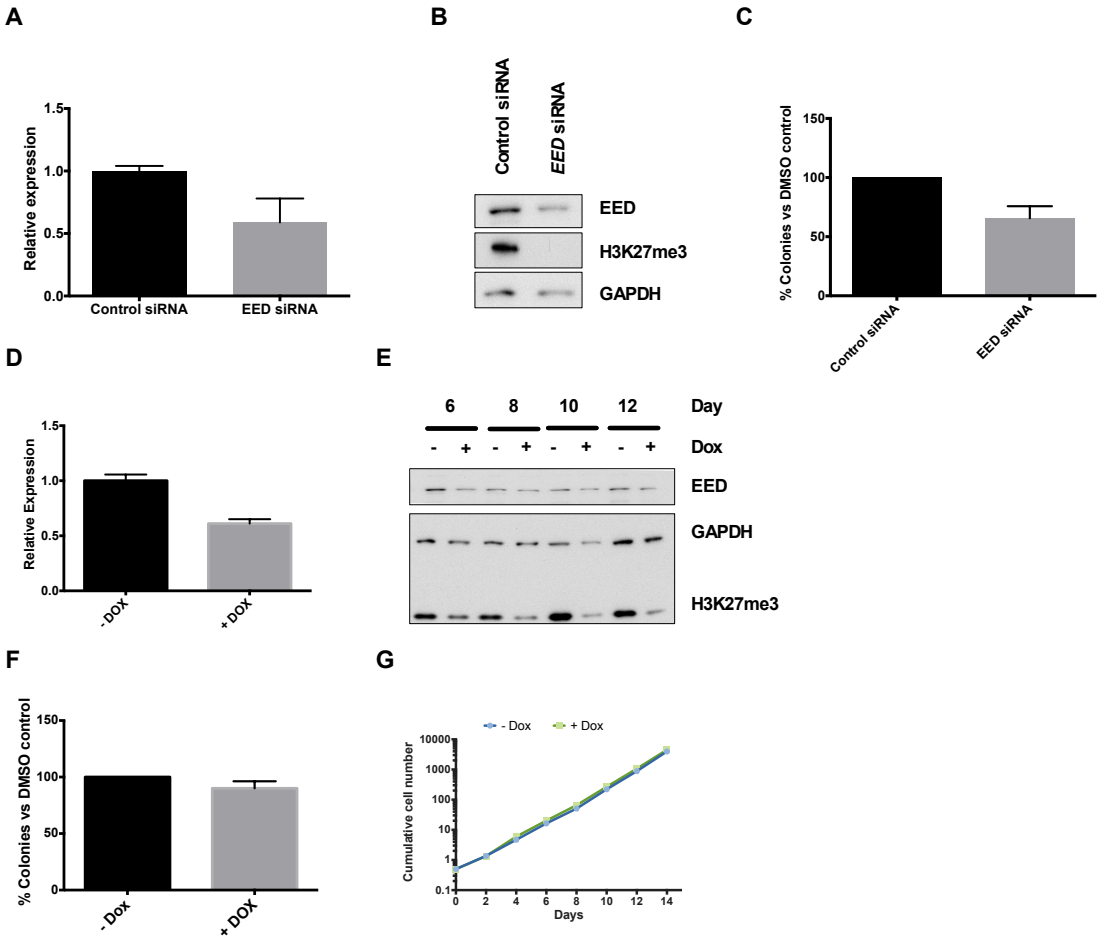


Figure 11 shRNA mediated knockdown of *EED* does not affect leukaemic growth of SEM cells

A,B and C) Cells were treated with either an EED siRNA or a non-targeting control siRNA and 96h post transfection cells were used in colony assays A) rt-qPCR of *EED* expression 120h post knockdown, relative to control siRNA treated cells, normalised to GAPDH, related to colony assays shown in (C) B) Western blot of *EED* knockdowns relating to (A). C) Percent colonies formed following *EED* siRNA knockdown relative to non-targeting control siRNA treated cells. D,E,F and G) Expression of EED shRNA in a stable cell line was induced by addition of Doxycycline. 120h later cells were harvested and used in colony assays and proliferation assays. A). rt-qPCR of *EED* expression following induction of EED shRNA with doxycycline (+Dox) compared to non-induced controls (-Dox), expression levels were normalised to GAPDH, error bars represent SD of three rt-qPCR technical replicates. B) Western blot probed with EED , GAPDH and H3K27me3 antibodies C) Percent colonies formed in EED shRNA induced (+IPTG) compared to non-induced control (-IPTG) cells. Error bars represent SD of three technical replicates.

3.4.2 Pharmacological inhibition of PRC2

Recently, the UNC1999 small molecule inhibitor which blocks the catalytic activity of EZH2/EZH1 in a highly specific manner, has been used to study the function of PRC2 in MLLr AML (Xu et al., 2014). Therefore, to assess the contribution of the PRC2 complex to MLL-AF4 leukaemias, SEM cells were treated with the UNC1999 inhibitor and their colony forming potential was assessed. The clonogenic capacity of SEM cells was inhibited in a dose-dependent manner, with a 42% reduction at the 100ηM treatment dose and an almost complete loss of colony forming ability at doses > 3μM (Figure 12A) Importantly treatment of SEM cells with the inactive UNC2400 analog did not affect the clonogenic capacity of SEM cells even at doses as high as 5μM (Figure 12B) The colony forming potential of the non-MLLr pre-B leukaemic RCH-ACV cell line was also impaired in a dose-dependent manner following treatment with UNC1999 (Figure 12C). Conversely, the erythroleukaemic cell line K562 showed a 40% impairment in colony forming ability at the highest dose of 5μM, a dose at which both SEM and RCH-ACV lines are unable to form colonies (Figure 12D). Therefore the effects of the UNC1999 inhibitor are not due to general toxicity of the compound.

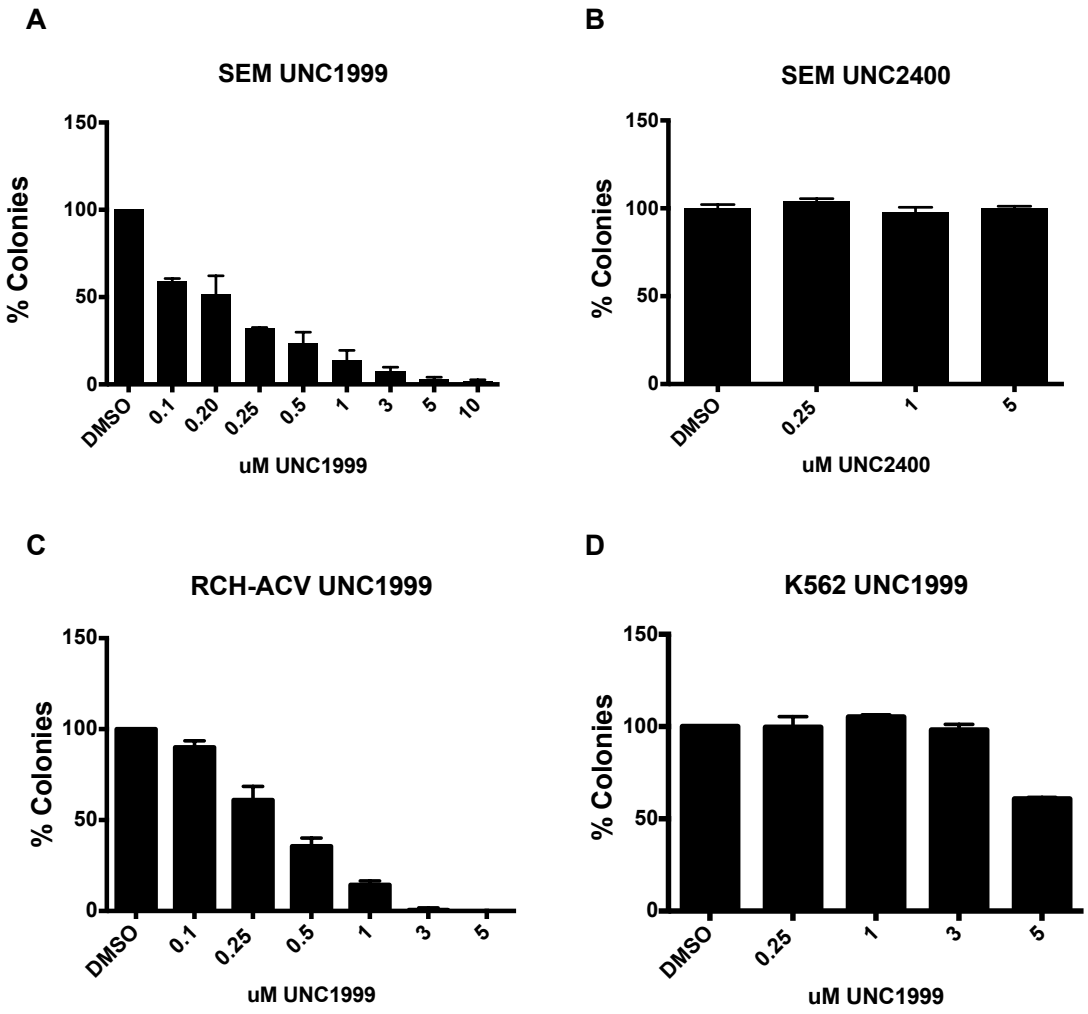


Figure 12 UNC1999 affects the clonogenic capacity of the SEM and RCH-ACV leukaemic cell lines

A and B) Percent colonies compared to DMSO treated controls following treatment SEM cells with A)UNC1999, B)UNC2400 (inactive analog). C and D) Percent colonies compared to DMSO treated controls following treatment with the indicated doses of UNC1999 for C)RCH-ACV (E2A-PBX, pre-B cell line) D) K562 (BRC-ABL, erythroleukaemic line). Error bars represent SD of three biological replicates for SEM cells and two biological replicates for RCH-ACV and K562 cells.

The leukaemic growth of SEM cells was also impaired in liquid culture. While the proliferation of SEM cells treated with 5μM UNC1999 was reduced by ~60% compared to DMSO treated cells after 5 days, the proliferation of SEM cells treated with 1μM UNC1999 were not affected until 10 days post treatment while

those treated with 0.25uM UNC1999 were not affected until 15 days after initial treatment (Figure 13A) . Importantly, treatment with 5uM of UNC2400 did not affect the proliferative capacity of SEM cells even after 15 days of treatment.

To assess the effect of the UNC1999 inhibitor on the levels of H3K27me3, SEM cells were treated with 0.25μM and 5μM UNC1999 inhibitor and the levels of H3K27me3 were assessed after 24 and 120 hours. The levels of H3K27me3 decreased in a time dependent manner and the lowest dose of 0.25μM was as effective at reducing global levels of H3K27me3 as was the highest dose of 5μM compared to DMSO treated cells (Figure 13B). To assess the effects of the inhibitor on the levels EZH2 and SUZ12 as well as H3K27Ac, SEM cells were treated with 0.25μM, 1μM or 5μM UNC1999 and western blots were performed. Concomitant with a decrease in H3K27me3 there was an increase in the levels of H3K27Ac while the levels of EZH2 and SUZ12 were not affected (Figure 13C). The UNC2400 inhibitor did not affect the levels of H3K27me3 or EZH2 and SUZ12 at a dose of 5μM. This data confirms that the UNC1999 inhibitor affects the catalytic activity of the PRC2 complex without affecting the levels of PRC2 complex components.

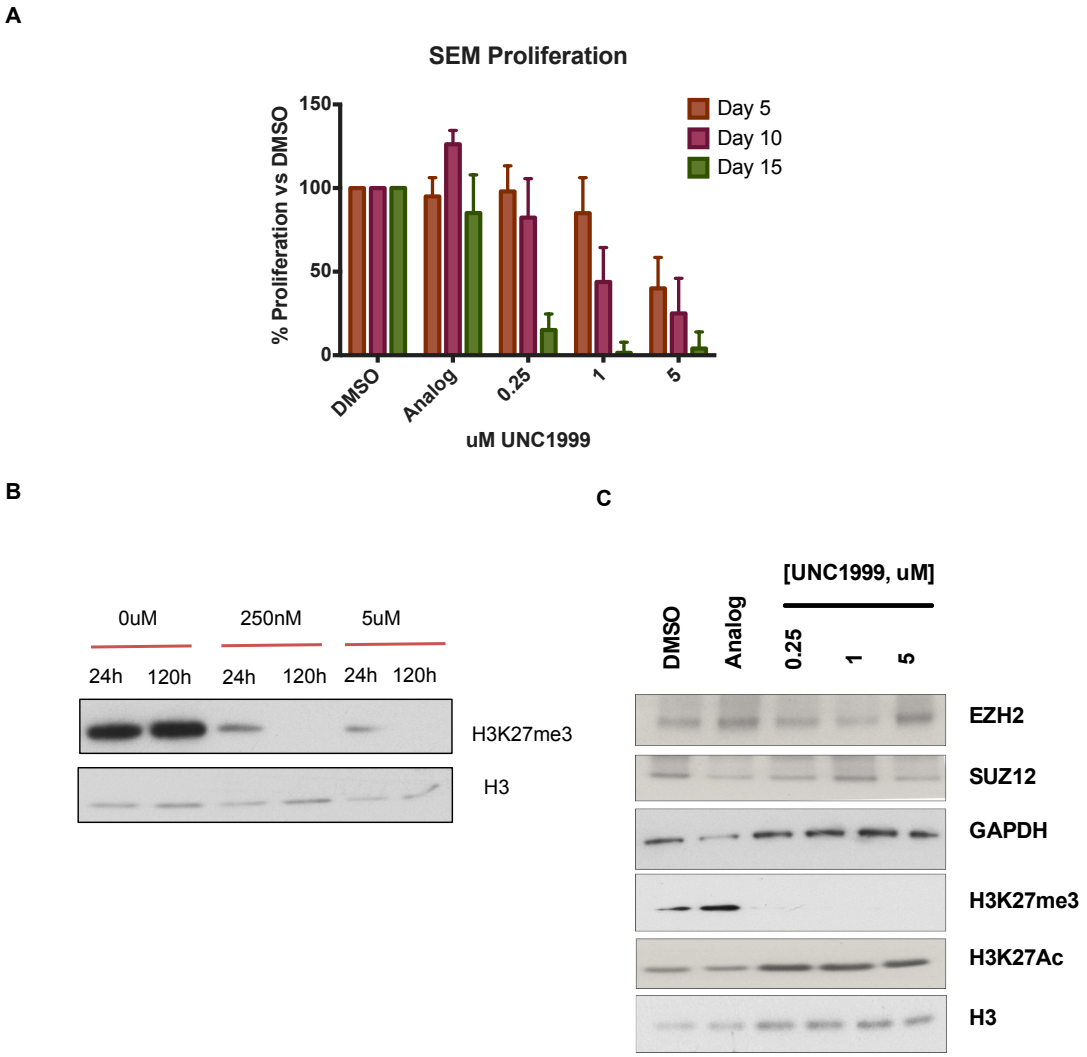


Figure 13 Proliferation of SEM cells is inhibited by UNC1999

A)% Proliferation of SEM cells treated with the indicated doses of UNC1999 or 5μM of the inactive analog UNC2400 (Analog) relative to DMSO treated cells. Error bars represent SD of three biological replicates. B) Representative western blots of whole cell extracts following treatment with DMSO, and the indicated doses of UNC1999 for 24 or 120 hours, probed with the indicated antibodies (h) (n=3) C) Representative western blots of whole cell extracts of cells following treatment with DMSO, 5uM UNC2400 (Analgo) and the indicated doses of UNC1999 for 120 hours, probed with the indicated antibodies. (n=3).

3.4.3 UNC1999 treatment induces a G1 arrest

Previous research has shown that inhibition of PRC2 results in cell cycle arrest at the G1-S transition in MLL-AF9 and MLL-ENL leukaemic models. This was shown to be mediated through de-repression of the CDKN2A locus. To investigate whether the observed anti-proliferative effect of t(4;11) cells following inhibition of PRC2 was a result of induced senescence, a cell cycle analysis using Propidium Iodide (PI) was performed. SEM cells were treated with a low dose of the inhibitor (Low, 0.25μM), a medium dose (Med, 1μM) and a high dose (High, 5μM). Cells treated with DMSO or 5μM of the UNC2400 inhibitor served as controls and the percentage of cells in each phase of the cell cycle was assessed after 5, 10 and 15 days by FACS. Treatment of SEM cells with the high dose of UNC1999 resulted in a time dependent arrest at the G1-S transition. 71%, 78%, 87.5% of cells were found in the G1 phase of the cell cycle in cells treated with the high dose of UNC1999 after 5, 10 and 15 days of culture respectively (
Figure 14). In contrast between 44%-54% of cells treated with either DMSO or UNC2400 were found to be at G1 at either day 5, 10 or 15. A small but consistent increase in cell number at the G1 stage was detected after treatment with the Med dose of UNC1999 at day 5 and 10 with the most pronounced effect seen at day 15 when ~67% of cells were found in G1. No obvious changes in cell cycle progression were observed for cells treated with the low dose of UNC1999 (
Figure 14).

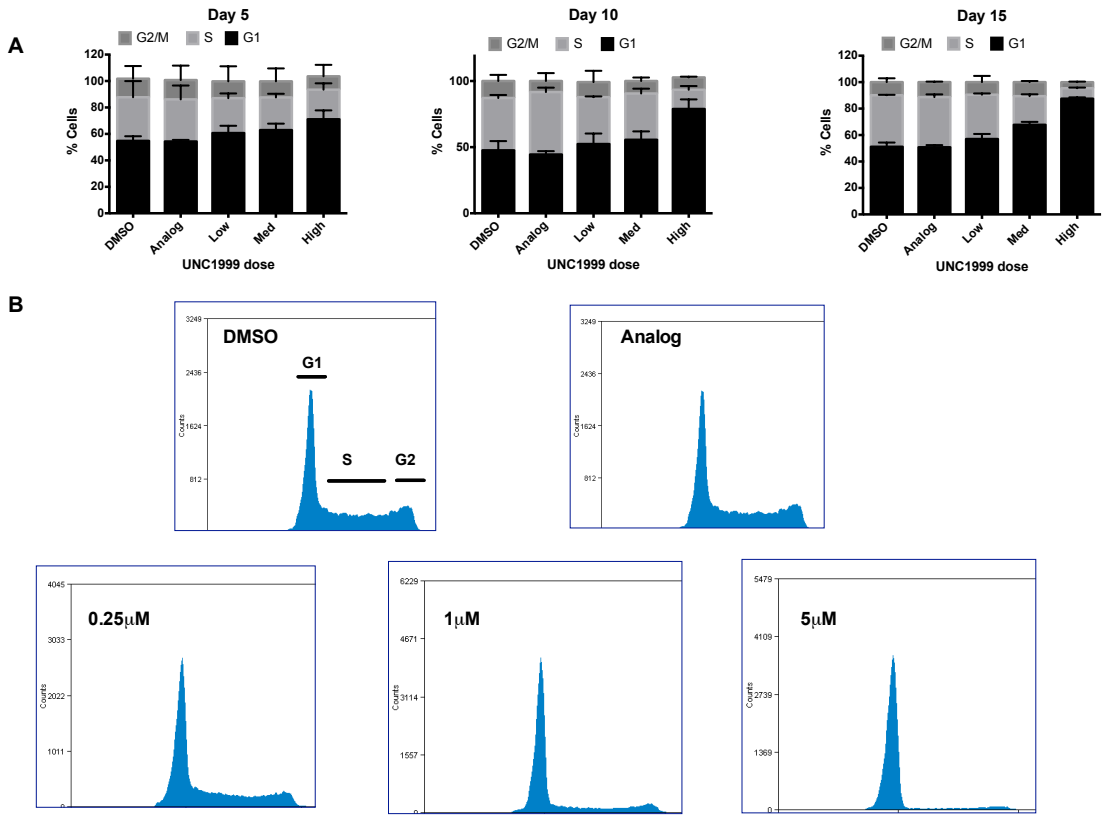


Figure 14 UNC1999 induces cell cycle arrest in SEM cells

SEM cells were treated for 15 days with 0.25μM (Low), 1μM (Med) or 5μM (High) UNC1999 as well as 5μM Analog. Control cells were treated with DMSO. A) Percentage of cells compared in G1, S or G2/M phases of the cell cycle. B) Representative histograms of DNA contents measured by PI staining following 10 days of treatment with the UNC1999 inhibtor at the indicated doses, or treated with DMSO or 5μM UNC2400. Error bars represent SD of three biological experiments.

To determine whether the observed G1-S transition was mediated through de-repression of the CDKN2A locus, the mRNA levels of *p16Ink4a* were measured after treatment of cells with UNC1999 for 5 days. No changes in gene expression of p16Ink4A were detected in contrast to upregulation of GATA6, a known target of PRC2 (Figure 15).

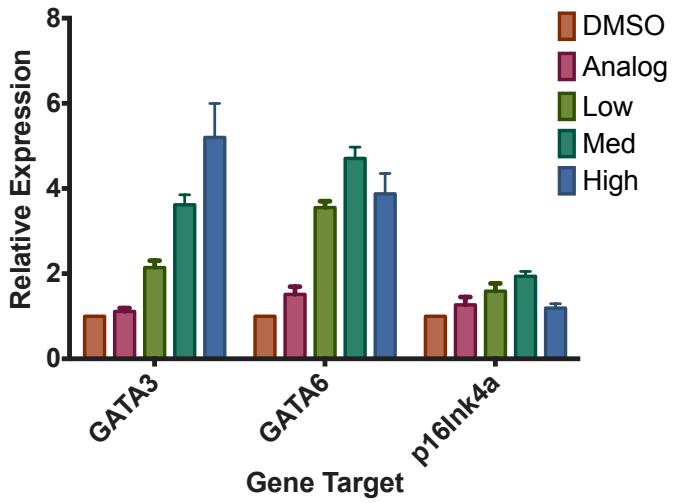


Figure 15 PRC2 inhibition does not result in the upregulation of p16^{INK4A}

SEM cells were treated for five days with 0.25μM (Low), 1μM (Med) or 5μM (High) UNC1999, 5μM UNC2400 (Analog) and DMSO. Gene expression was assessed by rt-qPCR, and expressed relative to DMSO. Expression was normalised to GAPDH. Error bars represent SD of two biological replicates.

Previously it has been demonstrated that following cell cycle arrest, treatment of cells with UNC1999 resulted in increased levels of apoptosis. To investigate whether the anti-proliferative effects of UNC1999 were a result of increased cell death, SEM cells were treated with a low, med and high dose of UNC1999 for 5 days and stained for the expression of Annexin V. Cells treated with DMSO and 5μM of UNC24000 served as control. No increase in apoptotic cells was detected, although a slight increase in dead cells was seen in SEM cells treated with the high dose of UNC1999 (9% dead cells compared to 4% in DMSO treated cells) (Figure 16).

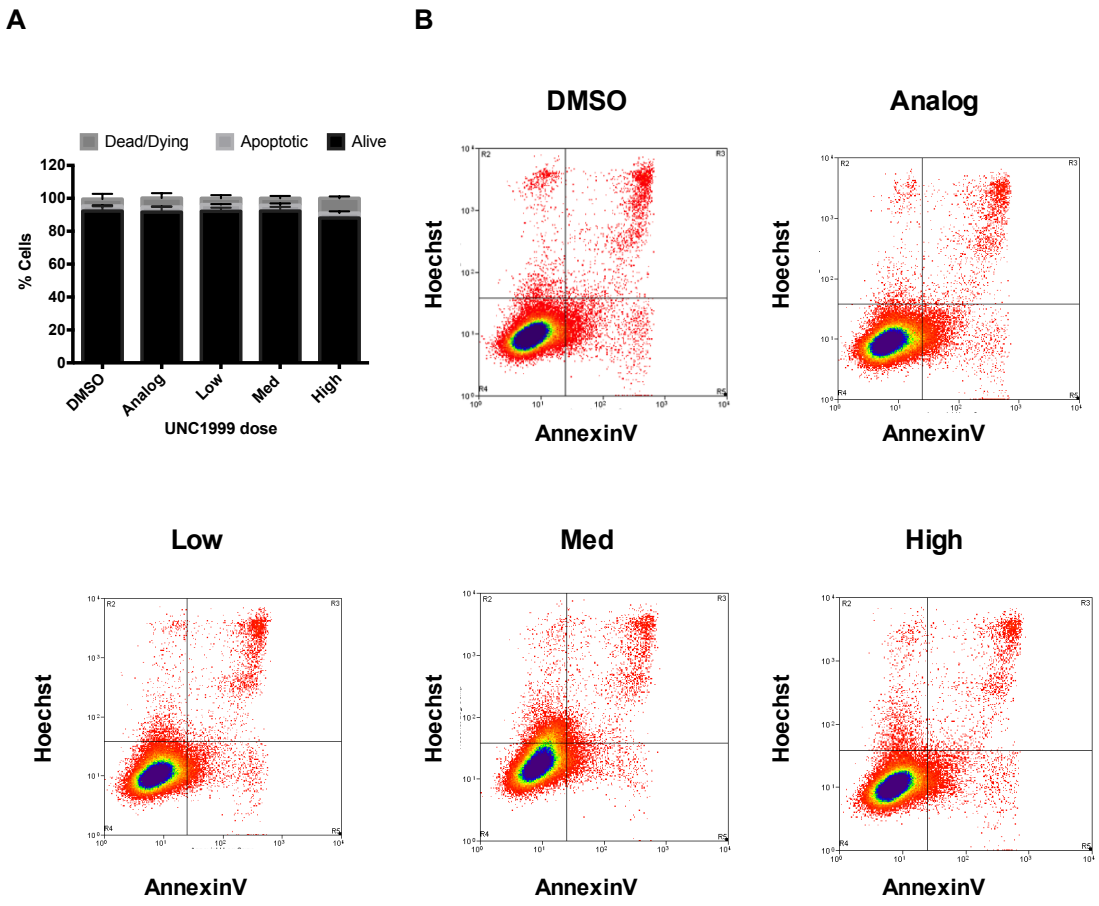


Figure 16 UNC1999 does not cause apoptosis in SEM cells

SEM cells were treated for five days with 0.25μM (Low), 1μM (Med) or 5μM (High) UNC1999, 5μM UNC2400 (Analog) and DMSO and stained with Annexin V and Hoechst (dead/live stain). Bar graph represents percentage of cells that are alive, apoptotic and dead/dying. Error bars represent SD of three biological repeats. B) Representative staining profiles of Hoechst and AnnexinV of SEM cells treated with the indicated doses for five days.

3.4.4 Discussion

Analysis of MLL-AF4 targets from two independent experiments identified the TF *Fli-1* and chromatin proteins *EZH2* and *UTX* as direct targets of MLL-AF4. Their deregulation has been associated with multiple forms of cancer and

therefore they represented interesting downstream targets of MLL-AF4 with possible functions in leukaemic maintenance.

3.4.5 Fli-1

Computational analysis of complex binding patterns identified *Fli-1* as one of ten TFs important for the maintenance and differentiation of HSCs (Wilson et al., 2010). In the paper published by Wilson et al., protein-protein interactions were identified among 7 TFs which included RUNX1 and Fli-1. Importantly TFs were found to bind and regulate genes in a combinatorial fashion. Given the established role of RUNX1 in MLL-AF4 leukaemia it was therefore possible that Fli-1 was also important. In addition to its role in HSC self-renewal and differentiation, Fli-1 is important for erythroid development, T and B-cell development and megakaryocyte development (reviewed in (Li et al., 2014)) Fli-1 over-expression has been associated with eryhtroleukaemia, pre-T-cell lymphoblastic lymphoma, a lupus like auto-immune disease as well as in the progression of AML.

Despite the established role of Fli-1 in multiple forms of cancer its role in ALL and particularly in t(4;11) ALL had not been investigated. Immunoblots of nuclear extracts prepared from multiple MLLr leukaemic lines as well as two non-MLLr ALL cell lines showed that Fli-1 is expressed at highest levels in ALL cell lines independently of MLLr status (SEM, RS4;11, RCH-ACV, CCRF-CEM). The expression pattern of Fli-1 was almost identical to that of

RUNX1 investigated in the same cell lines (Wilkinson et al., 2013), suggesting that co-operation of these TFs may extend from HSC maintenance to leukaemic maintenance (Wilson et al., 2010).

Fli-1 is known to regulate its own expression and it was therefore not surprising that siRNA mediated knockdowns resulted in ~90% reduction in *Fli-1* mRNA levels in SEM cells. This knockdown resulted in robust decreases at the protein level, probably aided by the short half-life of Fli-1 (~50h). Initial results using siRNA mediated knockdowns of *Fli-1* demonstrated a pivotal role for this factor for the clonogenicity of SEM and THP1 cells. Additionally, the proliferation of SEM cells was dramatically inhibited in liquid culture. However, inducible shRNA mediated knockdown of *Fli-1* by three independent shRNAs did not decrease the clonogenic capacity of SEM cells. While knockdown of Fli-1 in the stable cell line 5323 did not result in substantial decreases at the protein level of Fli-1, knockdowns in the 5322 and 5325 cell lines did. However, the clonogenic capacity of the 5322 cell line was not altered following induction of Fli-1 shRNA compared to a non-induced control. On the other hand, knockdown of Fli-1 in the 5325 cell line led to an increase in clonogenic capacity of the cells (~70%).

Taken together, a clear role for the function of Fli-1 in t(4;11) leukaemia cannot be established from these results. The failure of shRNA mediated knockdowns to mimic the results seen with siRNA-mediated knockdowns of *Fli-1* suggest that the results obtained may be a due of off-target effects of the siRNA. Alternatively, since stable cell lines were generated through lenti-viral transduction of SEM cells, it is possible that lenti-viral integration might have

occurred in genomic regions that alter the requirement of Fli-1 in leukaemic growth. It is also important to note that it can be difficult to accurately quantify the extent of a knockdown on the protein level. Although Fli-1 knockdowns in lines 5322 and 5325 appear to be substantial, one other possible explanation is that the shRNA knockdowns do not reduce Fli-1 protein levels as much as the siRNA knockdowns do, thus possible explaining the difference observed in colony forming potential between the two experiments. Importantly, the shRNA experiments were only performed once and it is therefore difficult interpret them conclusively. Recently, a role for Fli-1 in the recruitment of BRD4 to targets important for the leukaemic growth of a murine AML MLL-AF9^{NRAS} cell line has been reported (Roe et al., 2015). shRNA mediate knockdowns in this cell line reduced proliferation, consistent with results generated here inhibiting *Fli-1* in the AML line THP1. It would therefore be useful to try and validate the function of Fli-1 through an alternative approach, perhaps through inducible CRISPR technology.

3.4.6 UTX

UTX is a H3K27me2/3 histone demethylase (Agger et al., 2007; De Santa et al., 2007; Lan et al., 2007; Lee et al., 2007) previously shown to be important for the expression of homeobox genes (Agger et al. 2007). In haematopoiesis the expression of UTX is highest in stem and progenitor cells and decreases during differentiation (Liu et al., 2012). *In vitro* experiments have demonstrated an important role for UTX in cell proliferation and homeostasis of haematopoietic cells (Liu et al., 2012). The regulation of UTX by MLL-AF4 may therefore

contribute to leukaemogenesis by preventing the repression of genes important for differentiation as well as self-renewal. Inactivating mutations of UTX are the first mutations of a histone demethylase associated with cancer, and since then somatic mutations of UTX have been described in many haematological malignancies. Interestingly, UTX has been found to be directly associated with the promoters of *Mll* and *Runx1* and positively regulate their expression in primary murine bone marrow cells through its demethylating activity (Liu et al., 2012). shRNA mediated knockdown of UTX has been shown to impair the proliferation of MLL-AF9 leukaemic cell lines and it was therefore likely that UTX was an important downstream target of MLL-AF4 (Liu et al., 2012).

Unfortunately, siRNA mediated knockdown of UTX did not result in substantial changes at the protein level of UTX preventing assessment of its function. Additionally, treatment of SEM cells and the AML t(4;11) MV4;11 cell line with two broad acting demethylase inhibitors did not affect their proliferation in liquid culture, suggesting that the inhibition of de-methylation in these lines is not important for leukaemogenesis. These experiments were only performed once and would need to be repeated for any strong conclusions to be drawn. Additionally, since they were intended as a quick screens for the efficacy of these compounds in inhibiting leukaemic growth, validation of their function through immunoblots of histone methylation levels were not performed. Therefore no conclusions as to the function of these compounds on a molecular level in the cell lines tested can be derived. The GSK-J4 inhibitor, previously reported to be specific for UTX and its homolog JMJD3, did show efficacy at inhibiting

clonogenic capacity of SEM cells, although this experiment was only carried out once. Since then further work by the SGC consortium has shown that it also inhibits the function of the H3K4me1/2/3 demethylases JARID1B/C complicating any attempts to define specific effects of inhibiting H3K27me3 de-methylation. However, due to the preliminary findings that this compound affects the clonogenic capacity of a t(4;11) ALL cell line the function of this inhibitor warrants further characterization.

3.4.7 PRC2

Since both MLL-AF4 and H3K79me2 are detected at the EZH2 locus it is likely that MLL-AF4 regulates the expression of EZH2 directly. Knockdown of MLL-AF4 was shown to result in decreased expression of EZH2 at the mRNA and protein levels suggesting this to be the case. However, further experiments in MLL-AF4 knockdown cells linking the loss of binding at the EZH2 locus of MLL-AF4 to gene expression changes are needed to establish this with more certainty. These experiments have been carried out in MLL-AF9 transformed bone marrow cells where the fusion protein, in conjunction with menin, were shown to directly regulated the expression of *EZH2* (Thiel et al., 2010) and it is therefore possible that a similar mechanism of EZH2 regulation is employed by the MLL-AF4 protein.

An initial screen using siRNA mediated knockdown of *EZH2* only affected the clonogenic capacity of SEM cells modestly (by 20%), with more pronounced

effects observed in the AML MLLr cell lines Mv4;11 and THP1. Previous work has demonstrated a need for the complete inactivation of the PRC2 complex for the leukaemic growth of MLL-AF9 expressing cells (Neff et al., 2012; Shi et al., 2013) and these results were therefore not surprising. siRNA mediated knockdown of *EED* on the other hand resulted in a complete loss of H3K27me3 as assessed by western blotting and showed a 35% reduction in the colony forming ability of SEM cells. siRNA knockdowns of EED were only performed once and therefore neeed to be repeated. Surprisingly, shRNA mediated knockdown of *EED* did not affect the colony forming ability of SEM cells or their proliferation in liquid culture. H3K27me3 levels were reduced upon knockdown of *EED* although these were not as pronounced as those seen with the siRNA mediated knockdown perhaps explaining the different effects observed. Previously Shi et al. demonstrated that a knockdown efficiency of at least 80% of SUZ12 or EED was necessary for complete loss of H3K27me3 and inhibition of leukaemic growth of MLL-AF9^{NRAS} transformed murine cells, and it is therefore possible that the mild effects observed following inhibition of both *EZH2* and *EED* are due to insufficient inactivation of PRC2.

To inhibit the activity of PRC2 more robustly, SEM cells were treated with the EZH2/EZH1 specific inhibitor UNC1999 or its inactive analog UNC2400 and DMSO as controls. Treatment of SEM cells with the inhibitor resulted in a complete loss of H3K27me3 levels as detected by western blotting. Inhibition of PRC2 by the UNC1999 inhibitor resulted in a dose-dependent decrease in colony forming potential and proliferative capacity in liquid culture of SEM cells.

Importantly, the inactive analog UNC2400 did not affect leukaemic growth or levels of H3K27me3 indicating that the general toxicity of UNC1999 is negligible and observed effects are a result of inhibition of the catalytic activity of PRC2. Therefore, it can be concluded that the catalytic output of the PRC2 complex is necessary for leukaemic growth of SEM cells. Similarly, inhibition of the proliferative capacity of another t(4;11) cell line, RS4;11, has previously been shown to be mediated by the UNC1999 inhibitor (Xu et al., 2014). Therefore it seems that inhibition of PRC2 may be detrimental to t(4;11) ALL leukaemias in general.

The UNC1999 inhibitor also affected the clonogenic capacity of the pre-B cell leukaemic cell line RCH-ACV, although to a lesser degree while the K562 cell line only responded to inhibitor treatment at the highest dose of 5μM consistent with previous reports (Xu et al., 2014). While this data shows that the requirement of PRC2 in ALL is not limited to t(4;11) leukaemias, its higher sensitivity may be important in clinical settings. It is also important to note here that although our data suggests that expression of the EZH2 gene is maintained in part through the activity of the MLL-AF4 fusion protein in t(4;11) leukemias, increased expression of EZH2 is not something that is unique to just MLL-AF4 expressing cells. Instead, high levels of expression of EZH2 is perhaps a more general phenomenon in acute leukemias that can be maintained through a range of different mechanisms.

In accordance with published data, inhibition of PRC2 resulted in inhibition of G1-S transition. However, this effect was not mediated by upregulation of the CDKN2A locus product $p16^{INK4a}$, shown to mediate the anti-leuakemic effects of PRC2 inactivation in multiple studies. Therefore the molecular mechanism by which PRC2 inhibition functions in t(4;11) leukaemia appears to be distinct from other leukaemias. This may be mediated by de-repression of other targets that mediate cell cycle arrest or functions unrelated to transcriptional regulation. Recently it has been shown that EZH2 is important for proper replication of DNA, and deletion of *EZH2* in *CDKN2A* null MEF cells, induces cell cycle arrest through increased DNA replication fork stalling (Piunti et al., 2014). Since the levels of $p19^{ARF}$ were not measured it is also possible that cell cycle arrest is mediated through $p19^{ARF}$.

In contrast to published data showing increased levels of apoptosis following inhibition of PRC2 in AML, no significant increase in apoptotic cell death was detected. It will be interesting to determine whether cell death increases upon longer treatment with the inhibitor.

Surprisingly, treatment of SEM cells with 0.25μM UNC1999 inhibitor did not result in pronounced changes in cell cycle compared to DMSO treated controls, although small changes were observed, even though proliferation in liquid culture was severely affected after 15 days. These results are difficult to reconcile, and repeating these experiments (data presented are mean of 2 independent experiments) may resolve this issue. That is, it is possible that the

very small changes in cell cycle observed with 0.25uM are statistically significant, something that might only become apparent with multiple biological replicates and increased statistical power.

An important observation from the data generated is that inhibition of the clonogenic capacity of SEM cells is dose dependent. Additionally, the UNC1999 inhibitor targets the catalytic activity of PRC2 without affecting the protein levels of its core component, indicating that the observed inhibition of clonogenic capacity is specifically a result of reduced H3K27me3 levels. Therefore, by treating SEM cells with different doses of inhibitor the contribution of H3K27me3 to gene silencing can be investigated.

3.4.8 Conclusion

1 738 targets were found to be bound by MLL-AF4 in the t(4:11) patient cell line SEM. The expression of three of these targets, *Fli-1*, *UTX* and *EZH2* was shown to be dependent on continued expression of MLL-AF4. Preliminary data suggest that *Fli-1* contributes to the leukaemic growth of SEM cells but further validation of this is necessary. The function of *UTX* in this cell line is still unknown as siRNA-mediated knockdowns were inefficient at reducing the protein levels of UTX and no specific UTX inhibitors exists.

Inhibition of the PRC2 complex was found to be important for SEM leukaemogenesis and resulted in an arrest at the G1-S transition. Surprisingly,

$p16^{INK4a}$ was not upregulated in SEM cells and therefore the molecular mechanism by which PRC2 contributes to leukaemogenesis in t(4;11) ALL is distinct from that of MLLr AML. Using the SEM cells as a model, the contribution of the catalytic output of the PRC2 complex, namely H3K27me3, to gene silencing can be investigated. Deducing the molecular mechanism by which PRC2 contributes to t(4;11) leukaemogenesis can also aid in the rational design of therapies for this aggressive leukaemia.

4 Chapter 4 Molecular Mechanism of PRC2 function in t(4;11) leukaemias

In contrast to the function of PcG as antagonist of TrxG proteins (Geisler and Paro, 2015; Ingham, 1983; 1998), PRC2 has been found to co-operate with MLL-AF9 fusion proteins in establishing transcriptional expression programmes important for leukaemogenesis (Neff et al., 2012; Shi et al., 2013; Smith et al., 2011; Tan et al., 2011; Tanaka et al., 2012; Thiel et al., 2013; Xu et al., 2014). Repression of the *CDKN2A* locus by PRC2 as well as pro-differentiation TFs such as *ERG1* have been found in multiple cases to mediate the pro-leukaemic effects of PRC2 in these leukaemias (Neff et al., 2012; Tanaka et al., 2012; Xu et al., 2010). In Chapter 3, it was shown that PRC2 inhibition is not compatible with leukaemic growth in a patient cell line driven by the MLL-AF4 fusion protein. Surprisingly, inhibition of PRC2 did not lead to the upregulation of the *CDKN2A* locus and therefore the molecular mechanism by which PRC2 contributes to t(4;11) leukaemogenesis is distinct from its role in AML.

The exact molecular mechanism by which PRC2 contributes to gene silencing is not fully understood. An important function of PRC2 in mediating gene silencing is the methylation of H3K27me3. This mark has been found to serve as a docking site for CBX proteins and consequently PRC2 is able to recruit PRC1 to its targets (Boyer et al., 2006; Cao et al., 2002; 2005; Wang et al., 2004b). As discussed previously in the general introduction, PRC1 may mediate chromatin compaction

thus preventing the association of the transcriptional machinery with DNA or it may serve to prevent the progression from transcription initiation to productive elongation. However, not all PRC2 targets are co-bound by PRC1 (Ku et al., 2008) and the mechanism by which these targets are maintained in the repressed state is not fully understood.

Recently, small molecule inhibitors specific for EZH2/EZH1 have been developed (Campbell et al., 2015; Knutson et al., 2012; Konze et al., 2013; McCabe et al., 2012b; Qi et al., 2012). These allow for the inhibition of the catalytic function of PRC2 without affecting the levels of the core protein components that make up the complex. Consequently, they allow for the investigation of the function of H3K27me3 in gene repression.

To investigate the molecular mechanism by which PRC2 contributes to t(4;11) leukaemogenesis, the SEM cell line was used as a model. Through use of the small molecule inhibitor UNC1999, the specific contribution of the H3K27me3 PTM to gene silencing was assessed. Previous research has shown that only a subset of PRC2 genes become upregulated following its inactivation and therefore small changes in gene expression were expected to occur. To define why only small subsets of PRC2 targets become upregulated analysis of chromatin states of PRC2 targets was undertaken, specifically relating to the proposed molecular functions by which PcG proteins silence genes. To this end the association of PRC2 with components of the transcriptional machinery, PRC1 as well as well as with regions of hypersensitivity was analysed. Finally, by

specifically inhibiting the methyltransferase activity of PRC2 the effect of H3K27me3 in antagonising the function of H3K4me3 methyltransferases (TrxG proteins) was assessed.

4.1 Identification of PRC2 targets in a t(4;11) leukaemic cell line

To identify PRC2 targets in the SEM cell line a ChIPseq experiment using an antibody against H3K27me3 was performed. 11,159 peaks were identified through Homer's "findPeaks" tool and CBRGs MIG tool (used to filter out peaks with low read counts). Genes marked by H3K27me3 near promoters were identified by overlapping the H3K27me3 peak list to promoters of Refseq annotated genes (defined as +/- 2kb of the TSS) using an in house script. This analysis yielded a set of 4,001 gene targets enriched for H3K27me3 at their promoters (Figure 17). The mapping and analysis procedure was repeated with an already available ChIPseq track for EZH2 which identified 5,631 peaks associated with the promoters of 3,837 genes (Figure 17). 3,231 genes were found to be enriched for both H3K27me3 and EZH2, from now on referred to as PRC2 target genes (Figure 17). As expected PRC2 targets were found to be enriched at regions of unmethylated CpGs (as determined by biotinylated CxxC affinity purification (Bio-CAP) (Blackledge et al., 2012)), SUZ12 as well as depletion of H3K27Ac (Figure 17).

Pathway analysis of PRC2 targets showed that they were involved in developmental and differentiation processes and that targets were enriched for

those that show sequence specific DNA binding activity (Figure 17). This is consistent with the established role of PRC2 in repressing genes important for development and lineage specification (Bracken et al., 2006; Obier et al., 2014; Pasini et al., 2007).

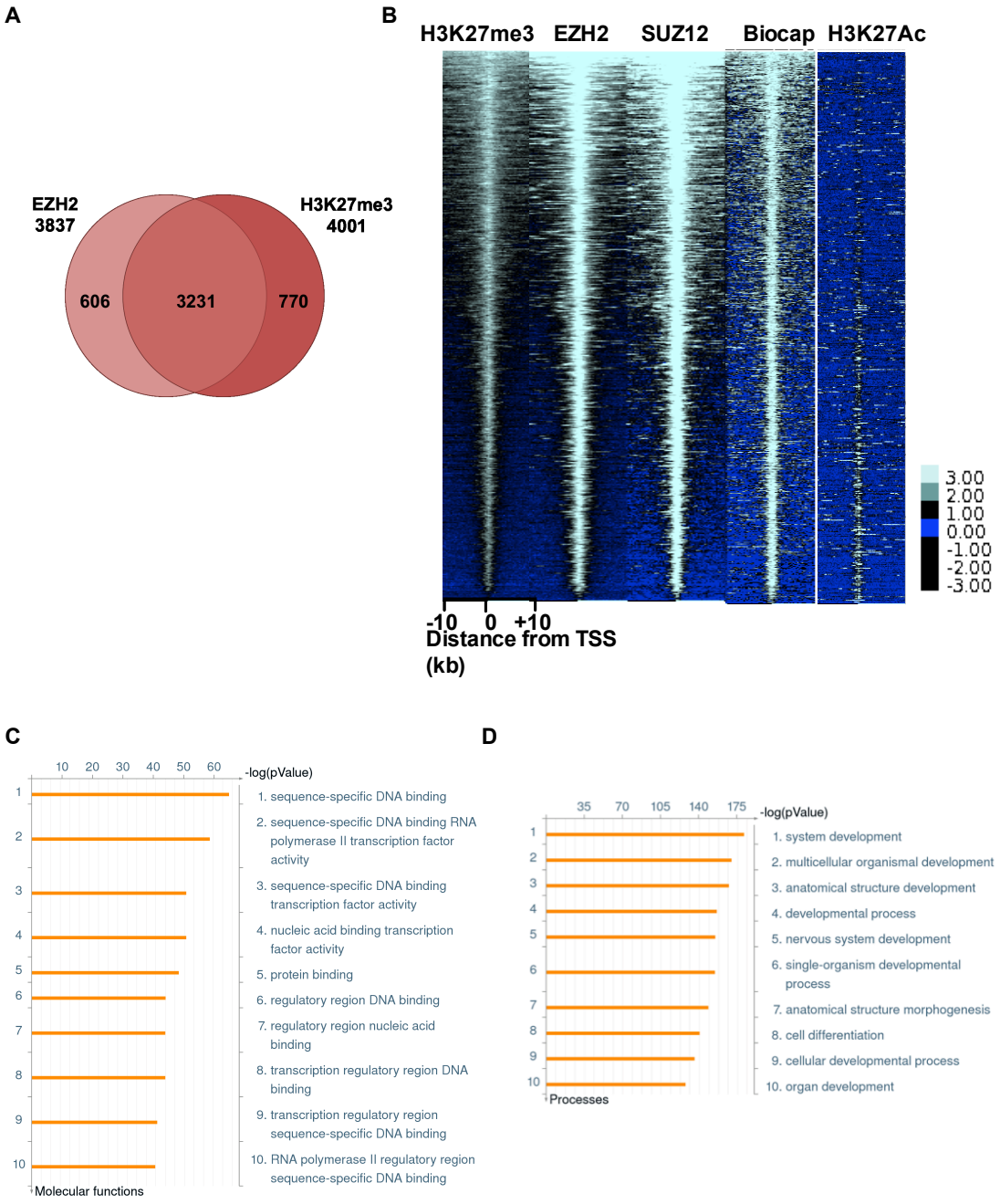


Figure 17 PRC2 targets are enriched for developmental regulators.

A) Overlap of gene targets enriched for EZH2 and H3K27me3 at promoters. B) Heatmap of ChIPseq enrichment at PRC2 targets with indicated antibodies as well as Bio-CAP reads (unmethyladed CpGs). PRC2 targets were ordered by decreasing H3K27me3 enrichment and centred on the TSS B) Metacore molecular function analysis of PRC2 targets. C) Metacore molecular function analysis of PRC2 targets. D) Metacore process analysis of PRC2 targets.

4.2 A subset of PRC2 targets become upregulated following inhibition of PRC2

In order to identify genes that become upregulated following inhibition of PRC2, nascent-RNA was extracted from cells treated with 5μM of the inactive UNC1999 analog UNC2400 (Analog), 0.25μM (Low), 1μM (Med) or 5μM (High) UNC1999 and their expression profiles were compared to cells treated with DMSO. Since nascent RNA-seq provides a meassure of RNA transcripts produced over a defined amount of labelling time it comes closer to measuring actual transcriptional changes and minimises changes masked by the stability/turnover of mRNA. Three biological repeats were performed to allow for differential expression analysis of the data using the Feature counts tool of the R package.

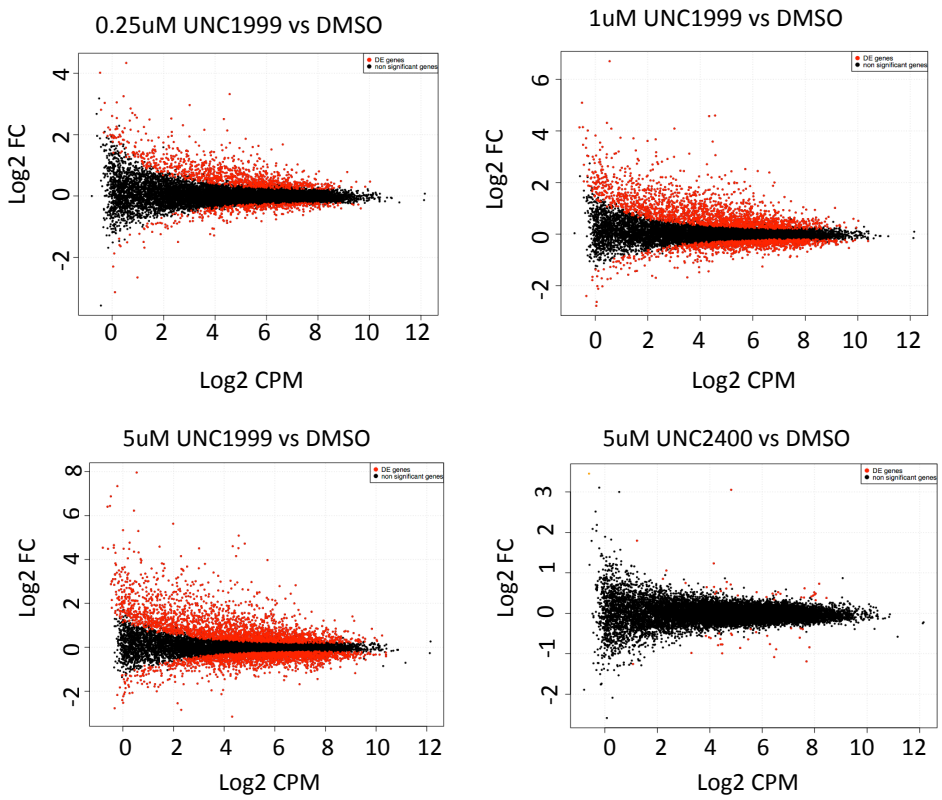


Figure 18 UNC1999 treatments results in upregulation and downregulation of genes

SEM cells were treated with 0.25μM, 1μM, 5μM UNC1999, 5μM UNC2400 and DMSO for 5 days. Nascent-RNA was extracted sequenced and analysed to identify differentially expressed genes. Correlation plots of the expression Log2 Fold change (Fc) vs Log2 counts per million (CPM) for the indicated treatment conditions compared to DMSO . Red: differentially expressed genes (FDR <0.05), black:

4,815 genes were found to be differentially expressed (DE, FDR: <0.05) between the DMSO control and the Low, Med and High conditions. Only 60 genes were found to be statistically different between the DMSO control and samples treated with the inactive Analog, demonstrating that the observed effects are a result of the inhibition of the catalytic activity of the PRC2 complex. Figure 18 shows the

log2 fold change of genes vs log2 count per millions (CPM) for all treatment conditions.

Of the 4,815 DE genes 1,904 genes were found to be upregulated and 816 genes were found to be downregulated by more than 0.5Log2 fold in all three treatment conditions (from now on referred to as either "DEup" or "DEdown" genes). Of these 665, 1,321 and 1,647 were found to be upregulated and 114, 288, 740 downregulated in the Low, Med and High conditions respectively (Figure 19). Of the genes upregulated in the Low condition, 97% were found to be upregulated in the Med and High conditions, and 92% of the genes upregulated in the Med condition were found to be upregulated in the Low or High conditions (Figure 19). The High condition showed the highest number of "unique" upregulated genes at 28%, while the Med and Low conditions only had 8% and 3% unique upregulated genes. The substantial overlap of upregulated genes across the three treatment conditions further serves to validate them as true up-regulated targets.

By overlapping the DEup and DEdown genes with the PRC2 targets identified through ChIPseq experiments, it was found that 744 DEup and 26 DEdown genes were PRC2 targets (from now on referred to as either upregulated and downregulated genes), representing 39% and 3 % of all up- or downregulated genes respectively (Figure 19). Since cells were treated for 5 days with UNC1999 it is expected that not all DEup/down genes will be PRC2 targets. Downstream effects of direct PRC2 targets may be exerting an effect on gene expression of non

PRC2 targets. Additionally, of the 60 DE genes in the 5uM UNC2400 condition only 18 and 22 genes were upregulated and downregulated respectively, by more than 0.5log2 fold and of these only 4 and 3 were PRC2 targets (Figure 19). Therefore, the observed gene expression changes are most likely a consequence of the inhibition of the catalytic activity of the PRC2 complex by the UNC1999 inhibitor.

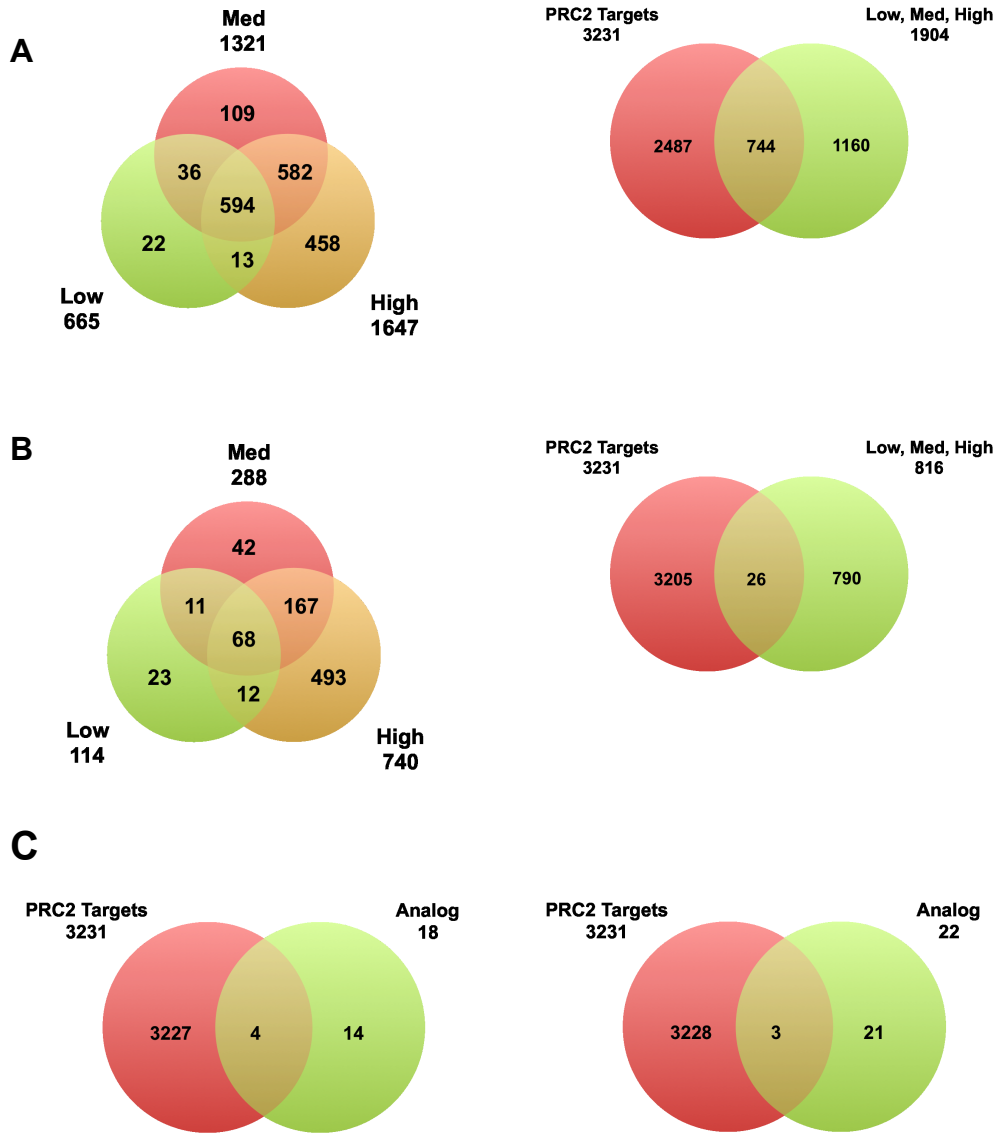


Figure 19 UNC1999 treatment of SEM cells leads to upregulation of PRC2 targets.

A) Left diagram: overlap of genes upregulated in the Low, Med and High conditions. Right diagram: Overlap between all genes upregulated in the Low, Med and High conditions with PRC2 targets. B) Left diagram: overlap of genes downregulated in the Low, Med and High conditions. Right diagram: Overlap between all genes downregulated in the Low, Med and High conditions with PRC2 targets. C) Overlap between genes differentially expressed in the Analog condition and PRC2 targets, left diagram: upregulated genes, right diagram: downregulated genes. Analog: 5uM UNC2400, Low: 0.25uM UNC1999, Med: 1uM UNC1999, High: 5uM UNC1999.

Of the 744 upregulated PRC2 targets, 304, 596 and 718 were found in the Low, Med and High conditions respectively . 293 genes were found to be upregulated in all three conditions (LMH common) (representing 96% of the upregulated genes in the Low condition) and 574 genes were shared between the Med and High conditions (MH common) (representing 96% of the genes upregulated in the Med condition). However, only 9% (Low), 18% (Med) and 22% (High) of all PRC2 targets became upregulated. This is consistent with previously published data where genetic inactivation of PRC2 or its inhibition has resulted in few expression changes (Bracken et al., 2006; Endoh et al., 2008; Ku et al., 2008; Leeb et al., 2010; Riising et al., 2014; Xie et al., 2014; Xu et al., 2014).

The level of expression of upregulated genes increased with increasing concentration with the mean log2 fold change being 1.7, 2.0 and 2.2 for the Low, Med and High conditions respectively (Figure 20). This indicated that the levels of H3K27me3 were determining both which targets became upregulated and the level of upregulation. The mean level of expression of upregulated targets was 2.732, 2.846 and 2.973 (Log2 (FPKM+1)) in the Low, Med and High treatment

conditions respectively (Figure 20). This is substantially lower than the average mean expression of all non-PRC2 targets, which was 4.728 (Log2 (FPKM+1)) and corresponds to the expression level of the lowest 20% of non-PRC2 targets (Figure 20). While a few targets were expressed at high levels, this suggests that loss of H3K27me3 induces only small transcriptional changes.

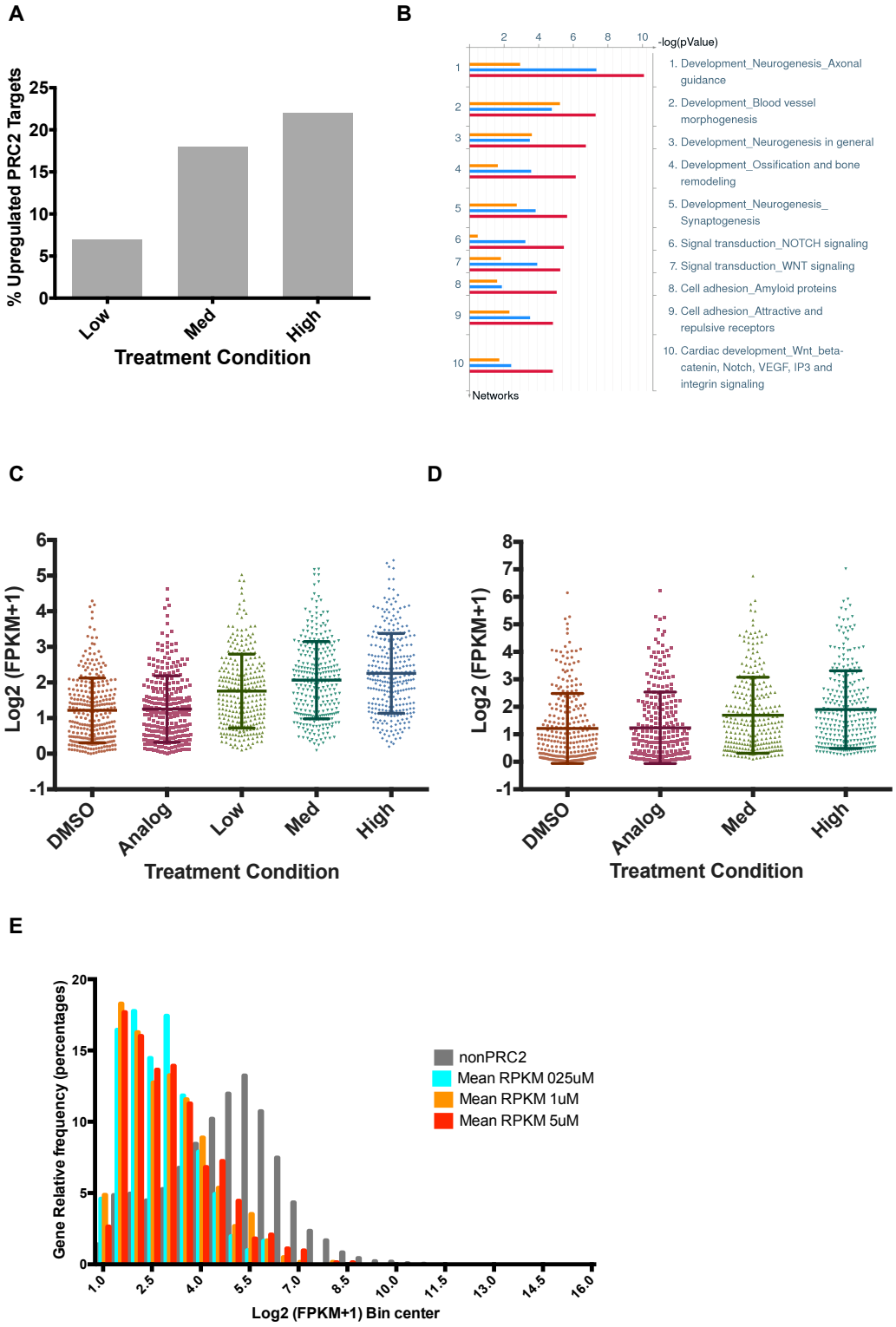


Figure 20 Upregulation of PRC2 targets is dose dependent.

A). Percent upregulated PRC2 targets after 5 day treatment with UNC1999 with the indicated doses. B). Metacore Network analysis of genes upregulated after treatment with a Low (yellow), Medium (blue) or High (red) dose of UNC1999 for 5 days. C). Level of expression of the 293 genes upregulated in all UNC1999 treatment conditions, bars represent median with interquartile ranges. D) Level of expression of the 574 genes upregulated in samples treated with Med and High doses of UNC1999, bars represent median with interquartile ranges. E) Frequency distribution of expression levels of upregulated genes following UNC1999 treatment for 5 days compared to expression levels of non-PRC2 targets.

To determine whether changes in the levels of nascent-RNA resulted in changes at the mRNA level, and to validate the nascent RNA-seq data, ChIP-qPCR was performed for five genes (in addition to GATA3 and GATA6 which were previously shown to be upregulated in chapter 3). All of the genes were upregulated in a dose dependent manner (Figure 21). *HOXA9* which is not a target of PRC2 in the SEM cell line did not show any changes in gene expression (neither did *CDKN2A* as shown in chapter 3) (Figure 21). Additionally, the expression of unchanged genes was assessed, and while three gene targets had no detectable transcript, two genes (*SOX13* and *CCNA1*) showed upregulation by qPCR (Figure 21). Overall this serves to validate the nacent-RNAseq data, although the number of upregulated gene targets may be underestimated due to sensitivity of the method.

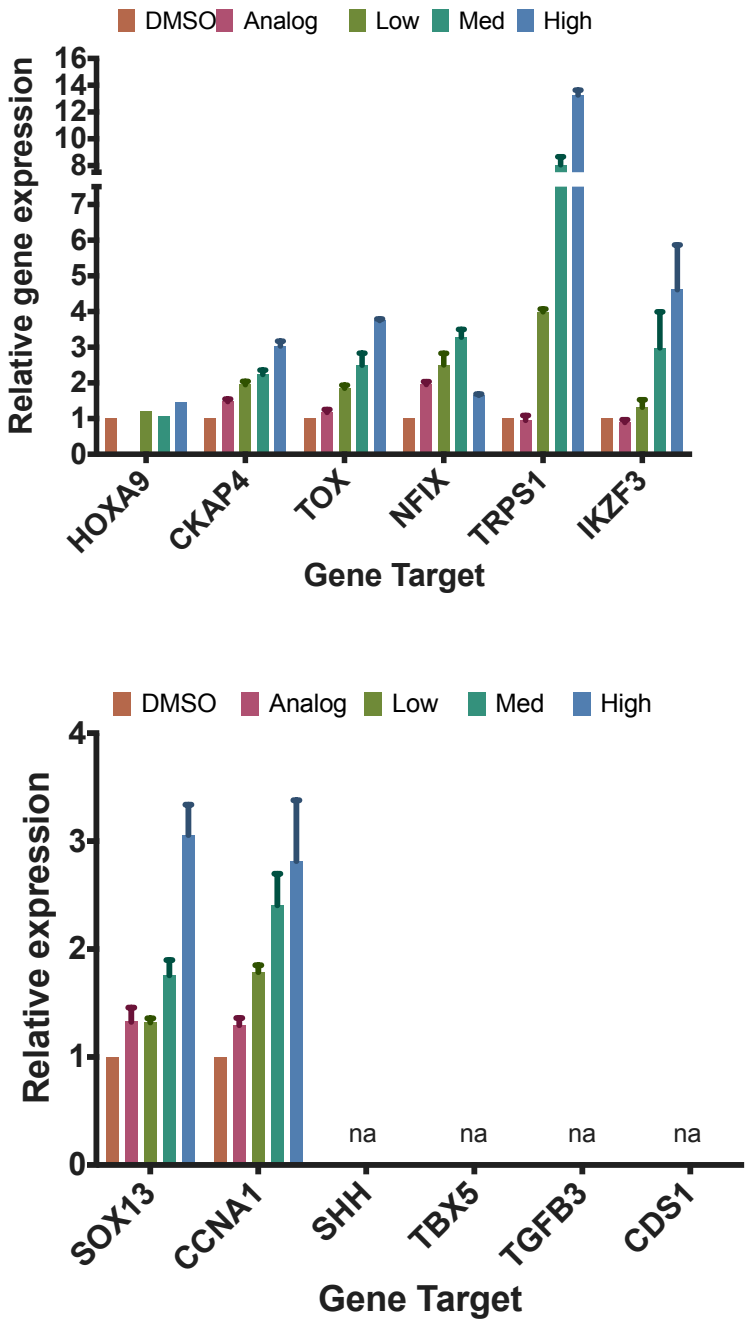


Figure 21 mRNA levels are upregulated in SEM cells following UNC1999 treatment

To validate the expression changes derived from Nascent-RNAseq experiments, cells were treated with DMOS, Analog or Low, Med and High doses of UNC1999 for five days and gene target expression was assessed by rt-qPCR. Top graph: genes shown to be upregulated through nascent-RNAseq analysis, Bottom graph: genes shown to be unchanged trough nascent-RNAseq analysis. Expression levels were normalised to GAPDH and compared to DMSO. Error bars represent the SD of 2-3 biological replicates. Analog: 5μM UNC2400, Low: 0.25μM UNC1999, Med: 1μM UNC1999, High: 5μM UNC1999.

4.2.1 Upregulated genes are enriched for developmental processess

Pathway analysis revealed that upregulated genes were enriched for targets involved in developmental processes including neurogenesis, cardiac development and blood vessel morphogenesis as well as targets involved in cell adhesion and NOTCH signalling (Figure 20) consistent with PRC2s role in these processes.

In accordance with qPCR results, no changes in the levels of expression from the *CDKN2A* locus were detected. Unexpectedly, visual inspection of the nascent-RNAseq revealed that this locus is already expressed in the SEM cell line (FPKM+1 value: 5.6) at similar levels to MLL-AF4 targets such as *RUNX2*, *HOXA9* and *Fli-1* (FPKM+1 values of 5.8, 5.09 and 5.11 respectively). Rather than being repressed by PRC2 this locus may be regulated by MLL-AF4, as it is one of the targets of the fusion protein identified in Chapter 3.

Pathway analysis of upregulated genes coupled with a literature search, led to the identification of a few targets with known tumour suppressor roles and cell-cycle functions. *Sestrin 3 (SESN3)* for instance, has been shown to inhibit cell proliferation through its inhibition of the cell cycle activator mTORC1 (Chen et al., 2010). Interestingly, *SESN3* is regulated by *FOXO1*, another upregulated PRC2 target which acts as a tumour suppressor in B-cell lymphomas through induction of growth arrest and apoptosis (Xie et al., 2012). Additionally, B-cell development in FOXO1 null mice is blocked at the pro-B cell stage (Amin and

Schlissel, 2008; Dengler et al., 2008) suggesting that repression of this transcription factor by PRC2 in t(4;11) leukaemias may be contributing to the differentiation block of this leukaemia. Two other targets with documented tumour suppressive functions identified were *early B-cell factor 3* (*EBF3) and RUNX3*. *EBF3* has been found to act as a tumour suppressor in AML as well as in brain, breast, colorectal, liver and bone tumour cell lines, through activation of genes that induce cell cycle arrest and apoptosis (Tao et al., 2015; Zhao et al., 2006). RUNX3 contributes to breast cancer pathogenesis, and over-expression of RUNX3 results in inhibition of proliferation of breast cancer cells as well as their invasiveness (Lau et al., 2006). t(4;11)

The dose-dependent increase in gene expression observed with the Nascent-RNAseq data suggests that the levels of H3K27me3 are influencing expression levels of PRC2 targets. To investigate how H3K27me3 levels were related to PRC2 target de-repression ChIP-Rx was performed.

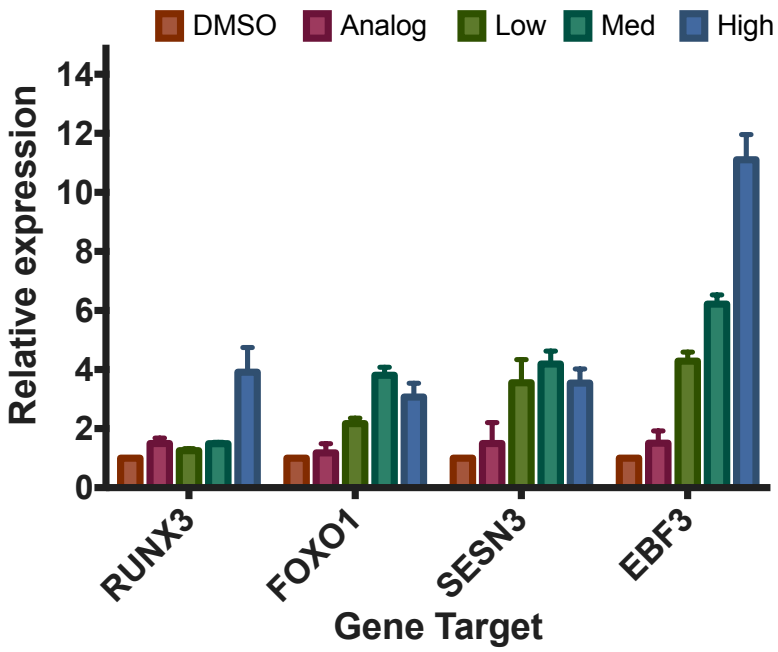


Figure 22 Genes with tumour suppressive functions are upregulated following UNC1999 treatment.

SEM cells were treated with either DMSO, Analog and Low, Med and High doses of UNC1999 for five days and mRNA levels were assessed by rt-qPCR. Expression levels were normalised to GAPDH and compared to DMSO. (n=2).

4.3 ChIP-Rx for H3K27me3 in UNC1999 treated cells

Quantitation of ChIPseq signals using the traditional "reads per million" (RPM) does not allow for accurate comparison across different experiments even when performed with the same antibody. For example, if experiment A containing 100% of a particular PTM is compared to experiment B where 50% of the mark has been lost in a uniform manner, RPM normalisation will result in both samples looking the same because the signal is quantified as a percentage of all mapped reads (Orlando et al., 2014). Furthermore, if for example a specific modification is not lost in a uniform manner, such that it if it is lost completely

from 50% of sites and only reduced to varying degrees at the remaining 50% of sites, traditional normalisation may result in increased signals at sites retaining the mark as a result of the depth of sequencing when compared to a sample that has 100% of the mark. The signal obtained in ChIPseq experiments is further influenced by technical variations such as the efficiency of the chromatin pull down and the fragment size of input chromatin. Therefore, a method that allows for normalisation with regards to the depth of sequencing of two samples while controlling for the efficiency of the immune-precipitation step allows for more accurate quantitation of ChIP-seq experiments across samples.

Therefore, in order to determine the global changes in H3K27me3 following treatment of cells with the UNC1999 inhibitor a novel ChIPseq method, termed ChIP-Rx was employed (Orlando et al., 2014). In this method, developed by Orlando et al. (2014), a spike in of *D. melanogaster* S2 cells at the lysis step of the ChIP procedure allows for subsequent normalisation of sequencing depth. This is achieved by deriving a normalisation factor, which is based on the numbers of reads that map to the exogenous Drosophila DNA. Since the same Drosophila cells are added to treated and control samples, the number of reads mapping to Drosophila should be the same across experiments. Therefore, a factor derived from the Drosophila mapped reads allows for control of depth of sequencing. Since the spike-in takes place at the lysis step of the ChIPseq procedure it also controls for any technical variation across the samples.

Before the method was employed, it was important to determine that the Spike-in of S2 cells did not affect the H3K27me3 sequencing signal and to determine the technical variation of the method. To this end optimisation experiments were performed for the ChIP-Rx method.

4.3.1 Optimisation of H3K27me3 ChIP-RX

SEM cells were treated for five days with either DMSO or 0.25μM UNC1999. ChIP-Rx experiments were carried out with a 4:1 ratio of human to drosophila cells. Two technical replicates for each condition were performed to assess the technical variation of the method. In order to determine whether addition of the *D. melanogaster* genome affected the signal of the ChIPseq data, traditional ChIPseq without addition of S2 cells was performed on the DMSO treated sample in accordance with the validation performed by Orlando et al. (2014).

The data was mapped to both the human (hg19) and the drosophila (dm3) genomes. To determine whether the Spike-in procedure affected the efficiency of the ChIP protocol DMSO treated samples with and without spiked in Drosophila cells were analysed with the traditional normalisation method. The average enrichment levels of H3K27me3 at all pre-defined H3K27me3 peaks and at promoters of PRC2 targets did not show any large-scale differences (spearman's correlation coefficient = 0.9751) (Figure 23). It was therefore concluded that addition of the Drosophila genome did not affect the detection of H3K27me3.

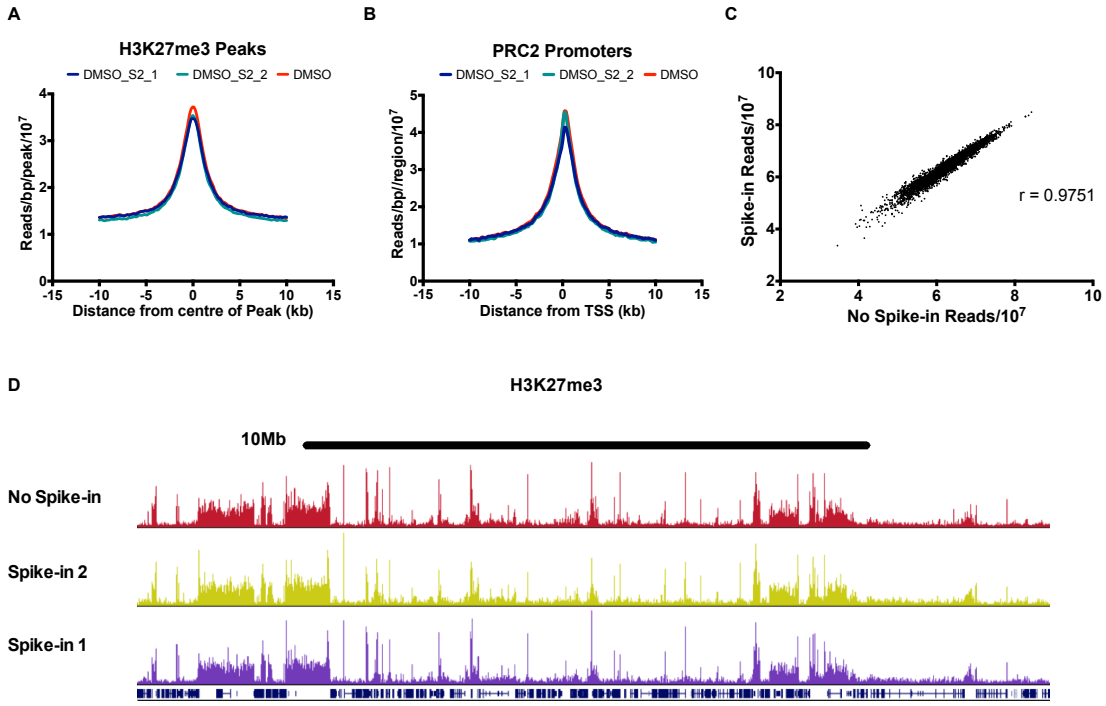


Figure 23 Spike-in of an exogenous genome does not affect the ChIP-seq signal of H3K27me3.

A) Histogram of H3K27me3 enrichment centered on H3K27me3 peaks of ChIPseq samples with S2 cell spike-in (DMSO_S2_1 and DMSO_S2) or without (DMSO). B) Histogram of H3K27me3 enrichment centred on PRC2 targets of ChIPseq samples with S2 cell spike-in (DMSO_S2_1 and DMSO_S2) or without (DMSO). C) Correlation of ChIPseq reads of the average signal of the two technical DMSO spike-in repeats against the DMSO sample without spike-in. D). UCSC track showing H3K27me3 profiles of DMSO sample without spike in (No Spike-in, red) and two technical replicates with spike-in (Spike-in 1 yellow, Spike-in 2 purple).

Normalisation with the reference derived normalisation factor (Ref), in contrast to traditional normalisation (Trad), revealed a substantial loss of H3K27me3 at PRC2 targets (Figure 24). The mean log2 fold change of H3K27me3 signal in the Low treated samples compared to average of the two DMSO treated samples was -2.116 and -2.230 for replicates number 1 and 2 respectively (Figure 24). This is congruent with the observed loss of global H3K27me3 by western blotting.

Spearman's correlation coefficient between technical replicates was 0.9564 and 0.9703 for the two DMSO and UNC1999 treated samples respectively, indicating very the low technical variability of the method (Figure 24).

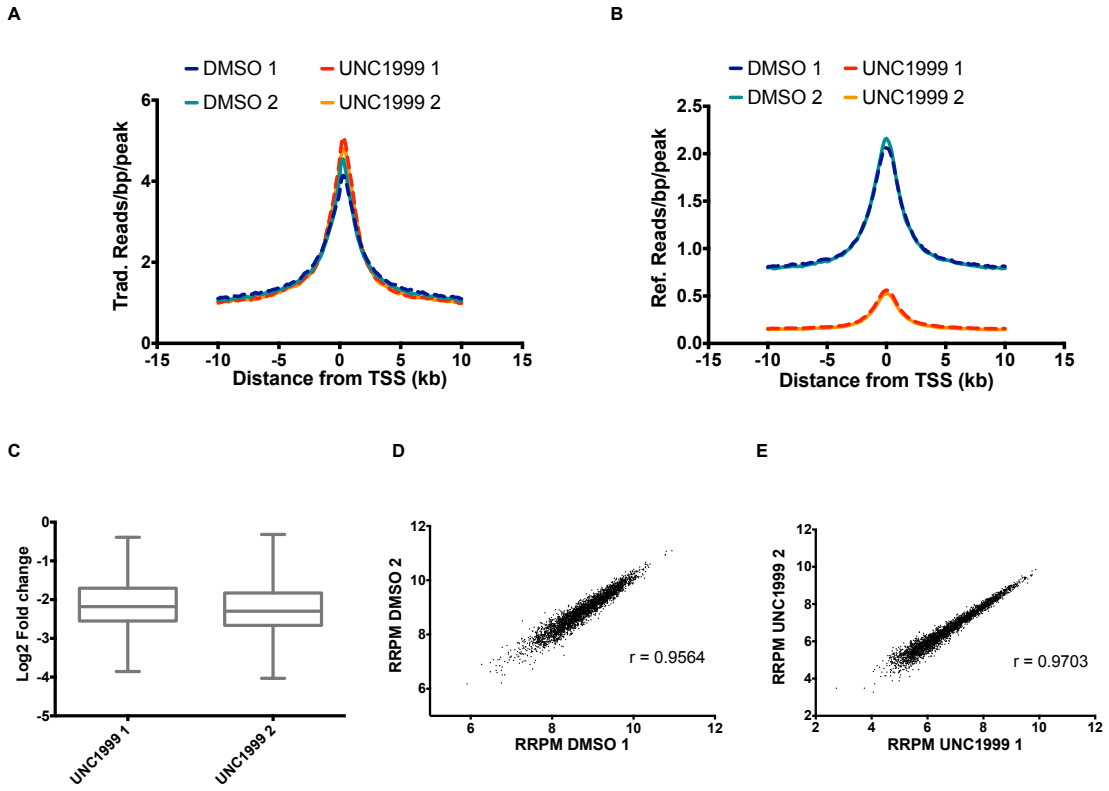


Figure 24 ChIP-rx reveals changes in H3K27me3 obscured by the traditional ChIPseq approach.

A and B) Enrichment levels of H3K27me3 centred on PRC2 targets for samples treated with either DMSO or 0.25μM UNC1999 for 5 days normalised with: traditional normalisation (A) or reference normalisation (B). C) Log2 fold change of reference normalised H3K27me3 enrichment levels at all PRC2 targets of two technical replicates treated with 0.25μM UNC1999 for days relative to the average of the two DMSO treated samples. Median and interquartile ranges are shown as boxes, error bars represent minimum and maximum values. D and E) Correlation plots of reference normalised H3K27me3 reads at PRC2 targets between technical replicates treated for 5 days with DMSO (D) or 0.25μM UNC1999 (E). r = spearman's coefficient, RRPM: reference normalised reads per ten million. DMSO1 and 2 and UNC1999 1 and 2 are technical replicates.

A subsequent H3K27me3 ChIPseq experiment (4.1) in SEM cells treated with either DMSO or 0.25μM (Low) UNC1999 allowed for the biological variability of the technique to be assessed. The average levels of H3K27me3 detected in the two DMSO and UNC1999 treated samples described above were compared to the new data set. The log2 mean fold loss of H3K27me3 was -2.167 and -2.543 for the two biological replicates respectively, compared to DMSO treated cells (Figure 25). This is slightly higher than the observed technical variability. Overall the variability across independent experiments was very low with spearman's correlation coefficients of 0.9598 and 0.971 for DMSO and UNC1999 treated samples respectively (Figure 25).

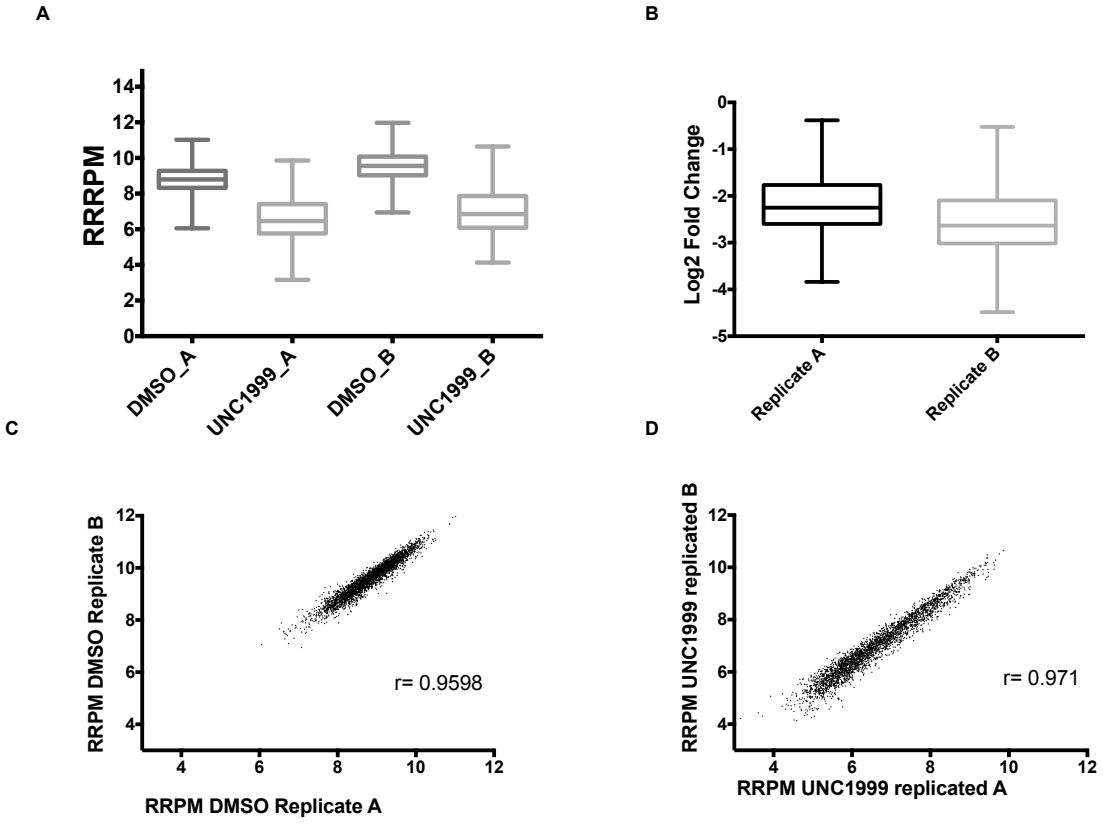
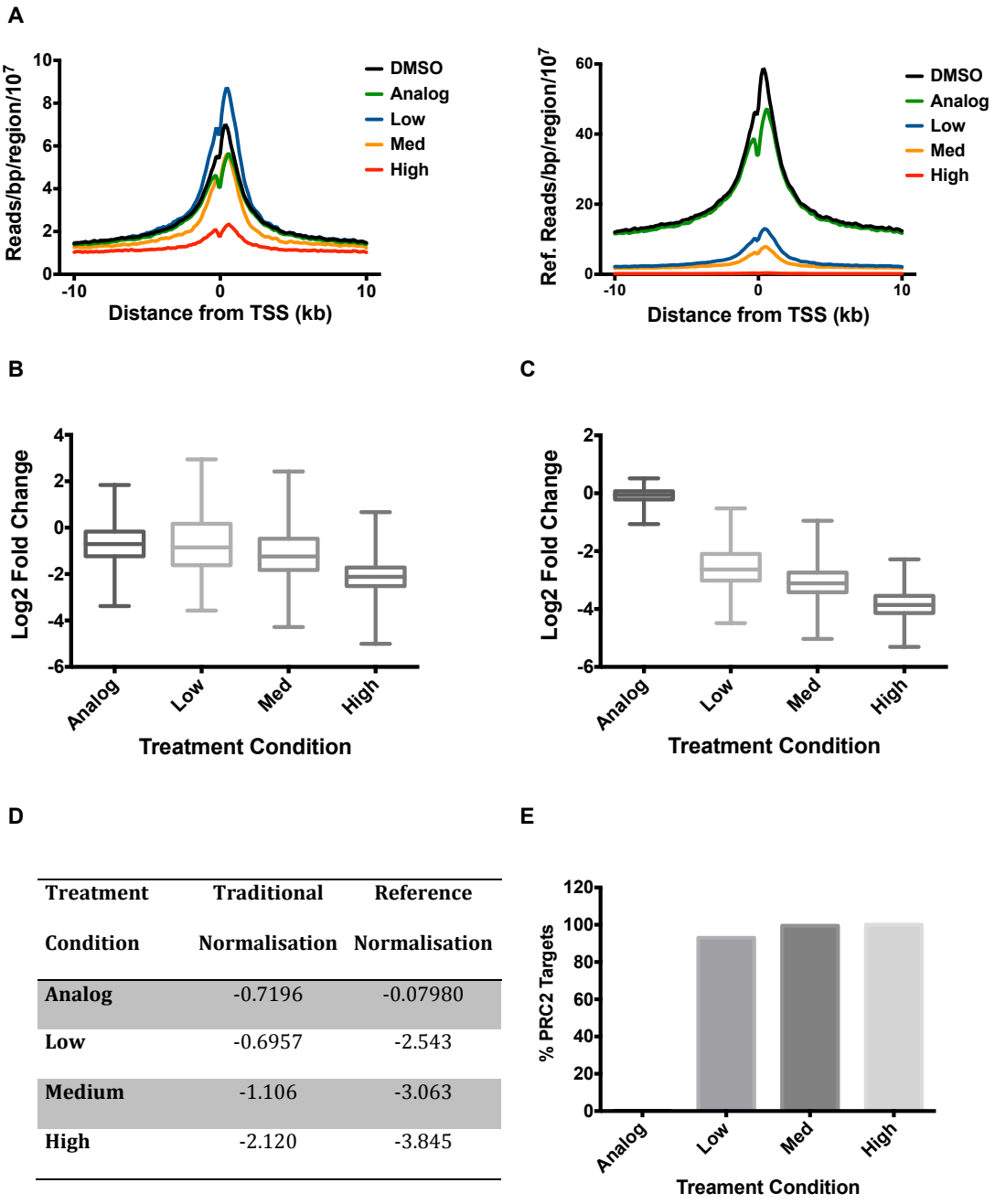


Figure 25 Minimal biological variability of the ChIP-Rx method

A) H3K27me3 reads of two biological replicate samples treated with either DMSO (DMSO_A, DMSO_B) or 0.25μM UNC1999 (UNC1999_A, UNC1999_B) for 5 days. B) Fold loss of H3K27me3 in 0.25μM UNC1999 treated samples compared to DMSO controls. Median and interquartile ranges are shown as boxes, error bars represent minimum and maximum values. C and D) Correlation plots of reference normalised H3K27me3 reads at PRC2 targets between biological replicates treated for 5 days with DMSO (C) or 0.25μM UNC1999.(D) r = spearman's coefficient, RRPM: reference normalised reads per ten million.

4.4 H3K27me3 levels correlate with transcriptional changes at a subset of loci

To determine how the levels of H3K27me3 were affected by treatment with UNC1999 on a global scale, ChIP-Rx using an antibody against H3K27me3 was performed on fixed cells treated with either DMSO, Analog, Low Medium or a High concentration of UNC1999. A dose dependent decrease of H3K27me3 signal at PRC2 targets was only apparent when the data was normalised with REF (described in section 4.3) (Figure 26). 0.05%, 93%, 99.5% and 100% of PRC2 targets in the Analog, Low, Med and High treated samples respectively showed at least a 1.5 log2 fold change in H3K27me3 levels (Figure 26). The mean fold loss of H3K27me3 at promoters of PRC2 targets compared to DMSO increased with increasing inhibitor concentration from -2.53 in the Low treated samples to -3.063 and -3.845 in the Med and High treated samples respectively (Figure 26). From now on all references to H3K27me3 levels are based on analysis using the reference normalisation.



Treatment Condition	Traditional Normalisation	Reference Normalisation
Analog	-0.7196	-0.07980
Low	-0.6957	-2.543
Medium	-1.106	-3.063
High	-2.120	-3.845

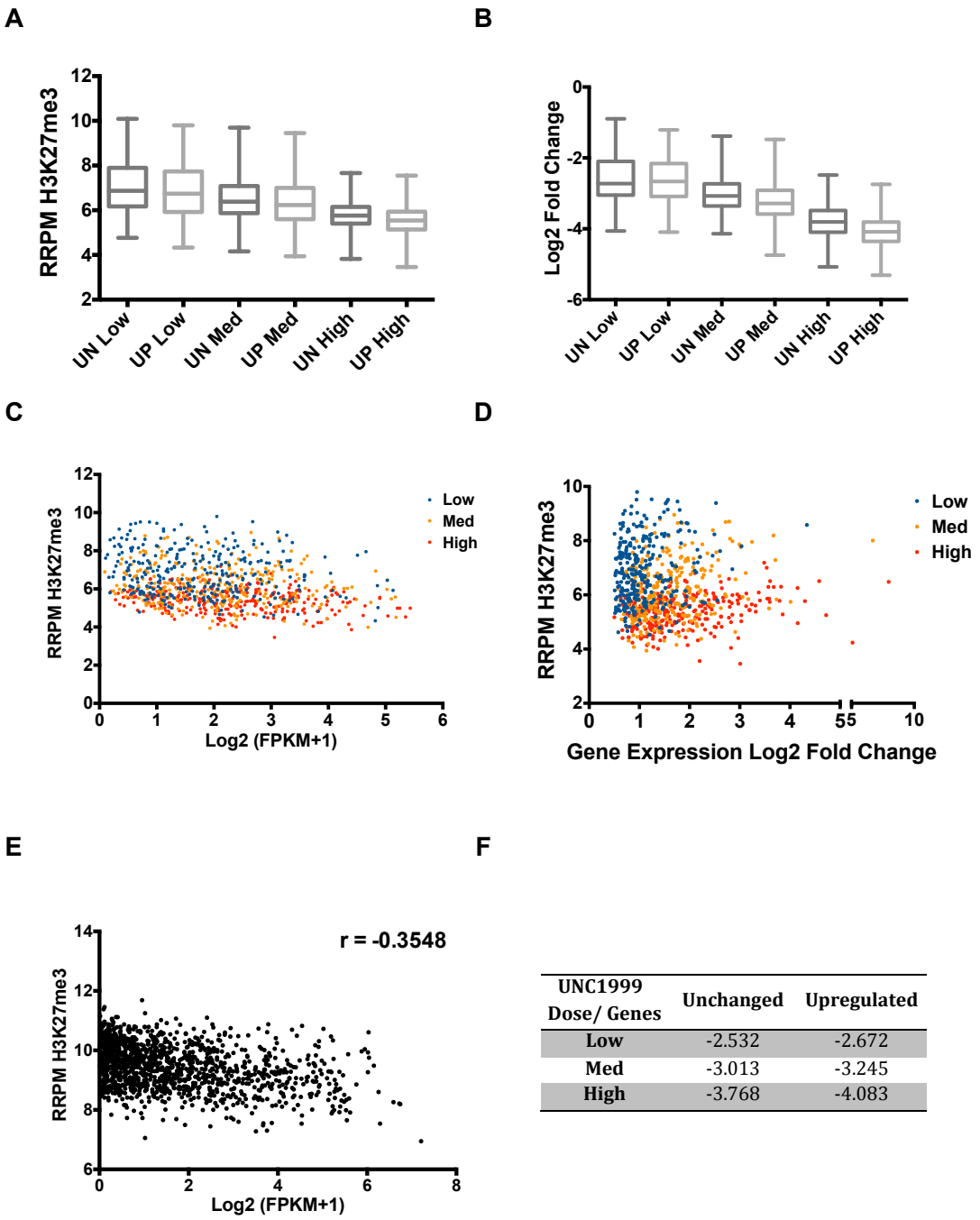
Figure 26 UNC1999 treatments leads to a dose-dependent decrease in H3K27me3.

SEM cells were treated with DMSO, Analog and Low, Medium and High doses of UNC1999 for five days followed by ChIP-Rx using an antibody against H3K27me3. A) Histograms showing enrichment of H3K27me3 at PRC2 target genes with the indicated doses. Left histogram: traditional normalisation, Right Histogram: Reference normalisation. B) Log2 fold change of H3K27me3 levels compared to DMSO around promoters of PRC2 targets normalised with the traditional normalisation method. C) Reference normalised

Log2 fold change of H3K27me3 levels at promoters of PRC2 targets. D) Table showing mean Log2 fold changes of H3K27me3 in the indicated samples normalised with either the traditional or reference derived methods. E) Percentage of PRC2 targets with at least 1.5 Log2 fold change in H3K27me3 levels. Analog: 5uM UNC2400, Low: 0.25uM UNC1999, Med: 1uM UNC1999, High: 5uM UNC1999.

Both unchanged and upregulated gene targets showed loss of H3K27me3 with upregulated genes showing a slightly higher fold loss (Figure 27), possibly as a result of further inhibition of H3K27me3 by activating histone modifications, such as H3K4me3 and H3K36me3.

Gene expression levels increased with decreasing H3K27me3 on a gene-by-gene basis. Of the genes that showed a dose dependent gene expression response across UNC1999 treatment conditions, 78% also showed a dose-dependent decrease of H3K27me3 levels (Figure 27). However, unchanged genes and upregulated genes showed similar loss/levels of H3K27me3 and therefore the levels of H3K27me3 are not predictive of gene expression across distinct genes. 60% of PRC2 targets had no detectable transcripts in the DMSO treated sample, so the correlation between levels of H3K27me3 and gene repression could not be determined at these targets. The H3K27me3 levels of the remaining 40% (1284 genes) showed a moderate negative correlation with gene expression (spearmans coefficient = -0.3548) (Figure 27).



UNC1999 Dose/ Genes	Unchanged	Upregulated
Low	-2.532	-2.672
Med	-3.013	-3.245
High	-3.768	-4.083

Figure 27 Reduction of H3K27me3 is similar at unchanged and upregualated genes, following UNC1999 treatment

SEM cells were treated with DMSO, Analog and Low, Medium and High doses of UNC1999 for five days followed by ChIP-Rx using an antibody against H3K27me3. The levels of H3K27me3 were correlated with expression data derived from the nascent-RNAseq experiment of cells treated the same way. A) Promoter reads at Unchanged (UN) and Upregulated (UP) PRC2 targets. B) Log2 fold change of H3K27me3 reads of

Unchanged (UN) and Upregulated (UP) PRC2 targets. C) Correlation of expression levels with H3K27me3 levels at promoters of the 293 LMH common upregulated targets. D) Correlation of expression log2 fold change with H3K27me3 levels at promoters of the 293 LMH common upregulated targets. E) Correlation between H3K27me3 levels in DMSO treated samples with gene expression levels. F) Table showing the log2 fold change in H3K27me3 at promoters of genes that either become upregulated or remain unchanged in the indicated condition. r= Spearmans coefficient. Low: 0.25uM UNC1999, Med: 1uM UNC1999, High: 5uM UNC1999.

To validate the observed results, ChIPq-PCR was perfomed on samples treated with DMSO and Analog and Low, Med and High doses of UNC1999 for 5 days. H3K27me3 levels were assessed at two loci that showed a dose-dependent decrease of H3K27me3 and were upregulated (*CKAP4* and *TOX*) as well two loci that lost H3K27me3 but were not upregulated (*CDS1, HOXC8*). The levels of H3K27me3 at *HOXA9* were also assessed and as expected, enrichment of H3K27me3 was similar to that of the background control region (negative control).

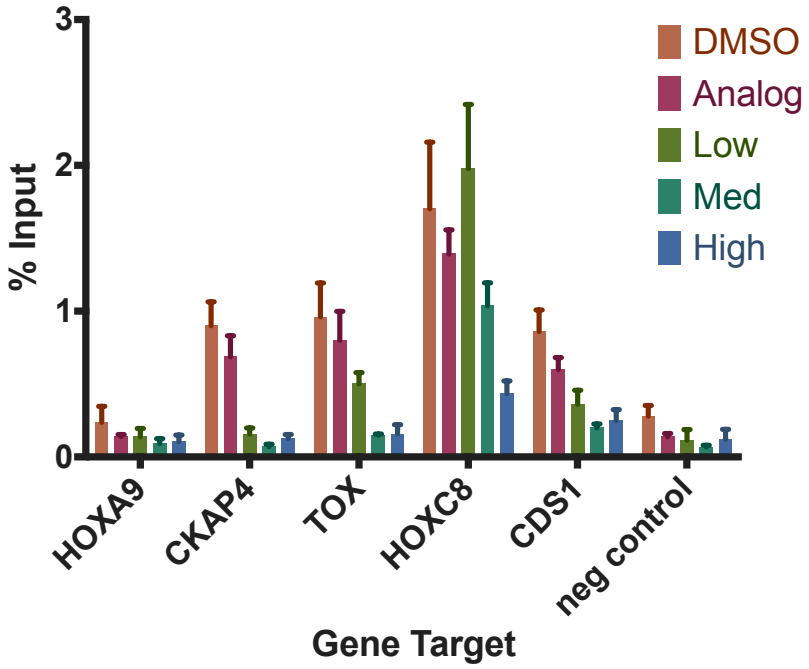


Figure 28 Dose-depenent decrease of H3K27me3 at unchanged and upregulated targets.

ChIP-qPCR using an antibody against H3K27me3 in cells treated with DMSO, Analog or Low, Med and High doses of UNC1999 for 5 days. Enrichment is represented as percent input. Neg control: a negative control region amplifying a sequence 18205kb upstream of the *c-myc* TSS. Error bars represent SD of three biological replicate experiments

4.4.1 Acetylation of H3K27 increases with decreasing levels of H3K27me3

Western blot analysis of samples treated with the UNC1999 inhibitor showed that levels of H3K27Ac increased following UNC1999 treatment. H3K27me3 and H3K27Ac are inversely correlated with each other and H3K27Ac levels are associated with promoters of actively transcribed genes (Pasini, Malatesta, et al. 2010). Previously it has been shown that loss of H3K27me3 results in robust increases in H3K27Ac and consequently, one of the mechanisms by which

H3K27me3 contributes to gene silencing is thought to be its antagonism of H3K27Ac (Pasini et al., 2010b).

To assess the changes in H3K27Ac genome wide and correlate them with the observed gene expression changes ChIP-seq for H3K27Ac in samples treated with DMSO, Analog, Low, Med and High doses of UNC1999 for five days were performed.

H3K27Ac changes were not as pronounced as H3K27me3. Levels of H3K27ac were increased by at least 1.5 log2 fold at 161, 392 and 1,168 PRC2 gene targets in the Low, Med and High treatment conditions respectively (

Figure 29). Only 19 genes were upregulated by 1.5 log2 fold in the sample treated with the inactive Analog (

Figure 29). While ChIP-Rx was attempted to accurately quantify the changes of H3K27Ac, the H3K27Ac antibody used failed to immune-precipitate Drosophila chromatin (data not shown) and therefore a traditional approach of normalisation had to be employed. The relatively few changes observed in H3K27ac levels in the Low and Med conditions may therefore be a result of the inherent limitations of this approach discussed previously (ChIP-Rx for H3K27me3 in UNC1999 treated cells).

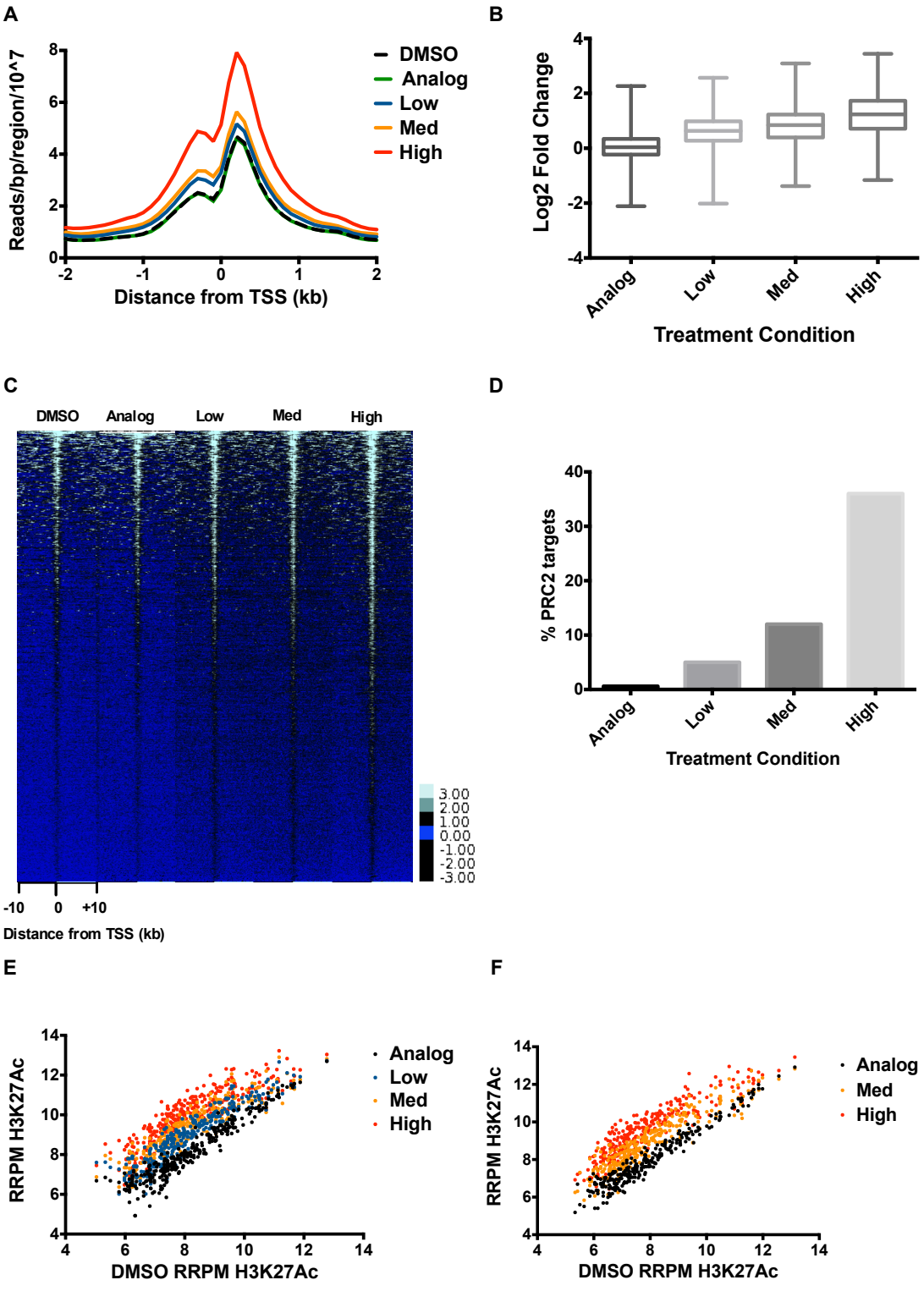


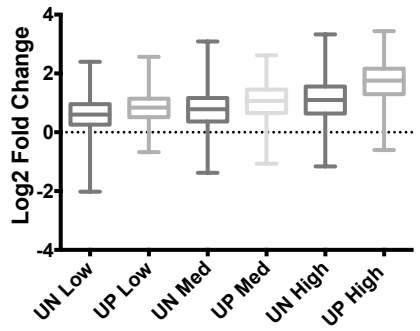
Figure 29 Levels of H3K27Ac increase following PRC2 inhibition.

SEM cells were treated with DMSO, Analog and Low, Medium and High doses of UNC1999 for five days followed by ChIP-seq using an antibody against H3K27Ac A) Histogram of H3K27Ac levels at PRC2 target genes centred on the TSS in the indicted samples. B). C) Log 2 fold change of H327Ac levels compared to DMSO at promoters of PRC2 targets C) Heatmap of H3K27Ac signal at PRC2 target genes ordered by decreasing levels of the High sample. E and F) Correlation of H3K27Ac levels at promoters of the: 293 LMH common upregulated targets (E) and the 574 MH common upregulated targets (F) Analog: 5uM UNC2400, Low: 0.25uM UNC1999, Med: 1uM UNC1999, High: 5uM UNC1999.

Despite this limitation, a dose dependent enrichment of H3K27Ac at PRC2 targets was observed (Figure 30 H3K27Ac are higher at upregulated compared to unchanged targets. The 304 genes upregulated in all three conditions, showed higher levels of H3K27Ac, with increasing concentration of UNC1999 (and therefore decreasing levels of H3K27me3). The same occurred at gene targets upregulated in both the Med and High conditions (Figure 30). The same genes showed no changes in H3K27Ac in the sample treated with the inactive Analog (Figure 30).

While accurate quantitation of H3K27Ac changes across samples was not feasible, comparison of H3K27Ac levels within samples was. This revealed that upregulated genes acquired higher levels of H3K27Ac than unchanged genes. This change was small in samples treated with the Low and Med doses and higher in the sample treated with the high dose of UNC1999 (Figure 30).

A



B

UNC1999 dose/Genes	Unchanged	Upregulated
Low	0.597	0.8379
Med	0.7465	1.032
High	1.083	1.701

Figure 30 H3K27Ac are higher at upregulated compared to unchanged targets

A) Log2 fold change of H3K27Ac reads of Unchanged (UN) and Upregulated (UP) PRC2 targets compared to DMSO. B) Table showing the log2 fold change in H3K27Ac levels at promoters of genes that either become upregulated or remain unchanged in the indicated condition.

To validate the observed results, ChIP-qPCR on samples treated with DMSO, Analog and Low, Med and High doses of UNC1999 for five days was performed. Of two upregulated targets (that showed loss of H3K27me3), *CKAP4* and *TOX* only *CKAP4* showed some increase in H3K27Ac levels (Figure 31). Genes that lost H3K27me3 but did not become upregulated did not acquire H3K27Ac (*HOXC8*, *CDS1*). Therefore it seems that loss of H3K27me3 is not sufficient for robust increases in H3K27Ac levels (Figure 31), although this should be verified at more target genes.

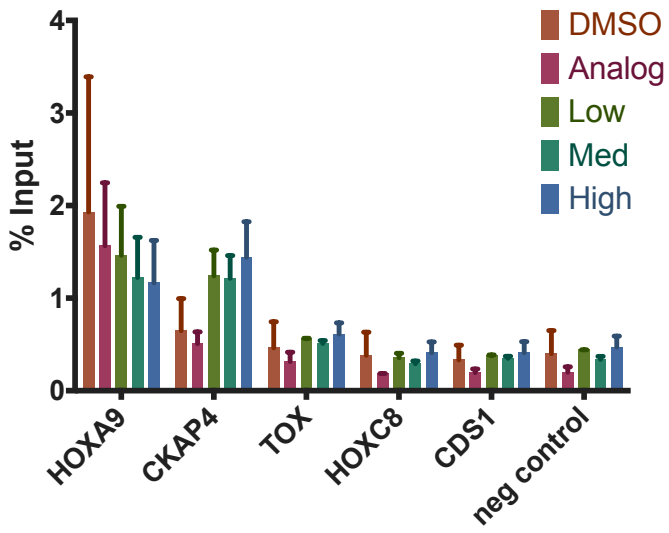


Figure 31 H3K27Ac are not altered robustly following UNC1999 treatment of SEM cells.

ChIP-qPCR using an antibody against H3K27Ac in cells treated with DMSO, Analog or Low, Med and High doses of UNC1999 for 5 days. Enrichment is represented as percent input. Neg control: a negative control region amplifying a sequence 18205kb upstream of the *c-myc* TSS. Error bars represent SD of two biological replicate expereiments

4.5 Defining a signature of upregulated genes

Levels/loss of H3K27me3 did not differ substantially across unchanged and upregulated genes indicating that other factors are involved in determining which PRC2 targets are susceptible to de-repression. To define a signature that would predict/explain why a subset of PRC2 become upregulated following inhibition of PRC2, the levels of H3K4me3, chromatin accessibility, transcription factor occupancy as well as factors associated with distinct stages of transcription were assessed.

4.5.1 H3K4me3 marks genes susceptible to upregulation following PRC2 inhibition

As discussed in the general introduction, a subset of PRC2 targets are also enriched for H3K4me3 and are so called "bivalent domains". The recently proposed responsive model poses that bivalent domains are a read out of the antagonistic action of TrxG and PcG proteins (Klose et al., 2013). In this model, it is envisioned that low levels of activating signal (most likely due to the activity of transcription factors) inhibit PcG domain formation but are insufficient to establish a TrxG domain (Klose et al., 2013). Removal of the inhibiting PcG signal could therefore preferentially allow for the upregulation of this subset of genes, as they already possess some activating signals.

To determine if genes marked by H3K4me3 defined PRC2 targets that are susceptible to de-repression, ChIPseq using an antibody against H3K4me3 in SEM cells was performed. 13 595 gene targets were found to have H3K4me3 at their promoters (Figure 32). Overlap of these with the PRC2 targets revealed that 1 664 genes were enriched for both marks (representing 52% of all PRC2 targets) (Figure 32). The levels of H3K4me3 at PRC2 targets were however substantially lower compared to non-PRC2 H3K4me3 marked genes (Figure 32). (Genes enriched for both H3K27me3 and H3K4me3 will be referred to as "Double Marked (DM) genes, as ChIPseq does not allow the determination of true bivalency).

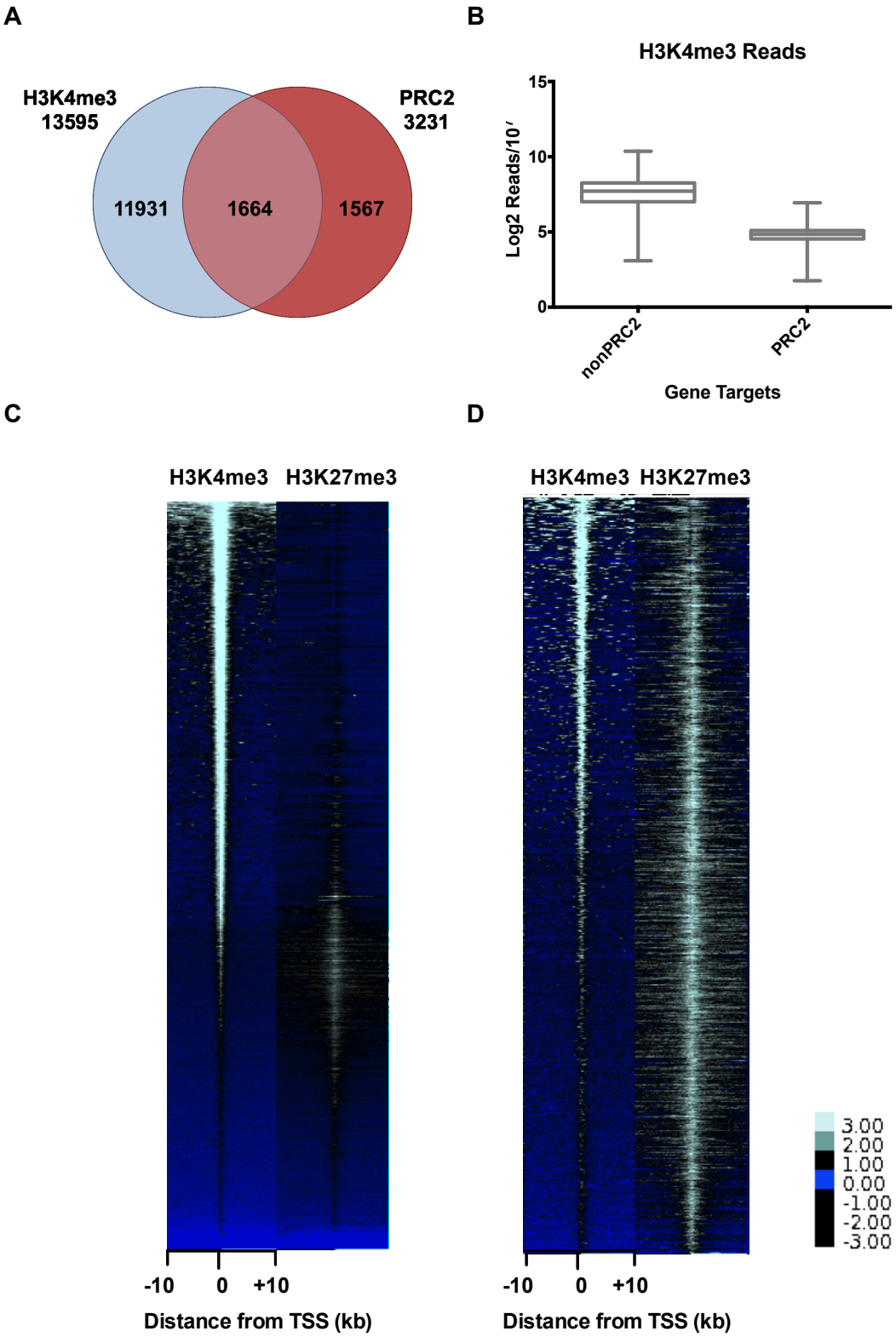


Figure 32 PRC2 targets are marked by H3K4me3.

A) Overlap between PRC2 targets and genes enriched for H3K4me3 at promoters. B) H3K4me3 reads at promoters of H3K4me3 marked genes either bound by PRC2 or not (non-PRC2). Boxes represent median with interquartiale ranges, error bars represent minimum and maximum values. C) Heatmap of H3K4me3 and H3K27me3 enrichment at all genes ordered by decreasing H3K4me3 levels. D) Heatmap of H3K27me3 and H3K4me3 enrichment at PRC2 targets ordered by decreasing H3K4me3 levels.

As expected the majority of upregulated genes were marked by H3K4me3. 92%, 90% and 89% of genes upregulated in the Low, Medium and High conditions respectively, were marked by H3K4me3 (Figure 33). However, not all H3K4me3 marked PRC2 genes became de-repressed, with 83%, 73% and 62% of these genes remaining repressed in the Low, Medium and High treatment conditions respectively. The levels of H3K4me3 at unchanged and upregulated PRC2 targets did not differ and therefore the levels of H3K4me3 are not predictive of susceptibility to de-repression (Figure 33).

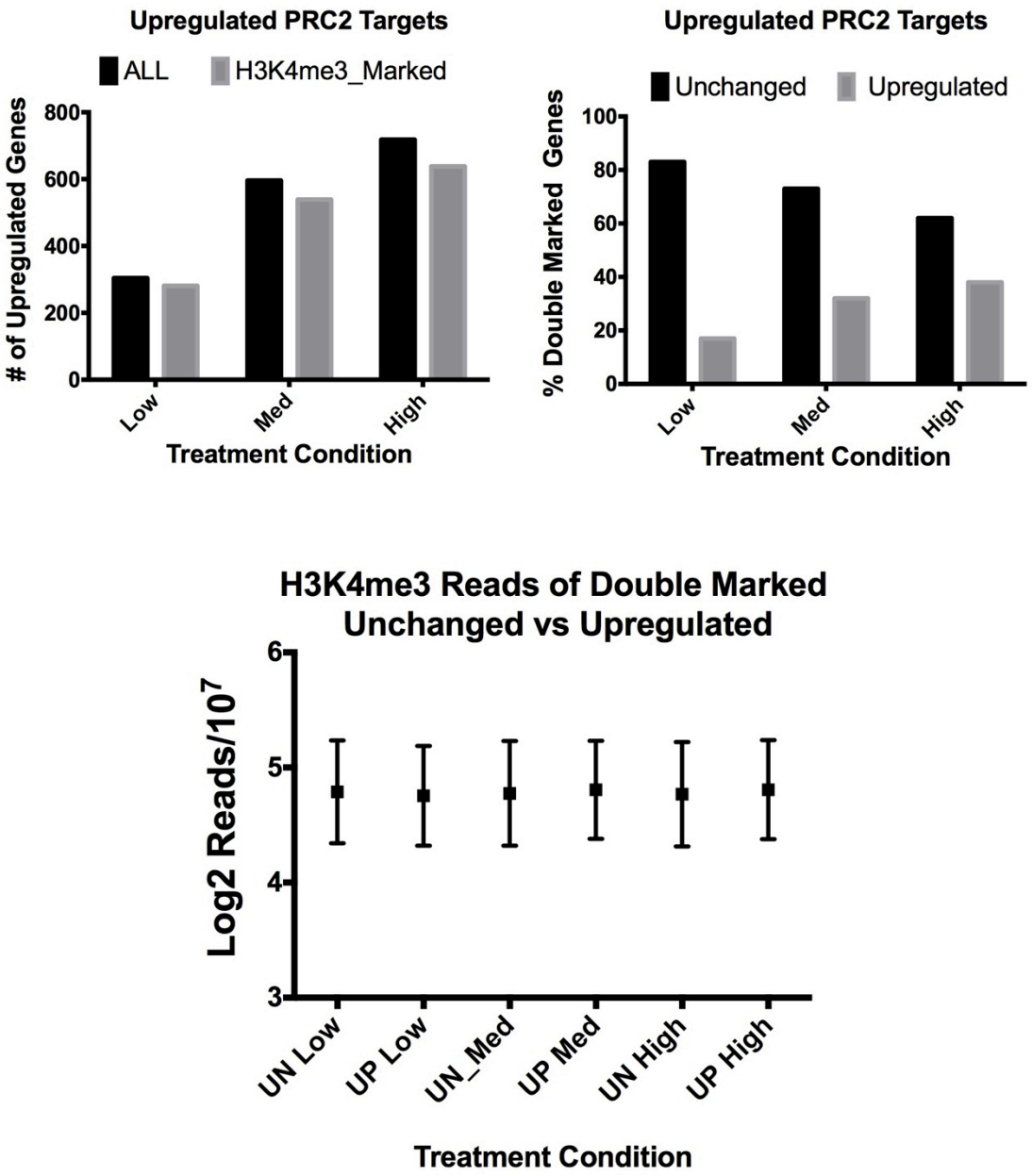


Figure 33 PRC2 targets that are susceptible to derepression are marked by H3K4me3.

A) Number of genes upregulated following five day treatment with the indicated concentrations (ALL, black bars), number of upregulated genes marked by H3K4me3 (H3K4me3_Marked, grey bars). B) Percent of double marked that remain unchanged (black bars) or become upregulated (grey bars) following a five day treatement with the indicated doses of UNC1999. C) Number of reads at promoters of DM genes that either remain unchanged (UN) or become upregulated (UP) following a five day treatment with the indicated doses of UNC1999. Squares represent mean number of reads, error bars represent interquartile ranges.

While H3K4me3 correlates with active transcription it is also associated with promoters that are not expressed. Consistent with this, previous research has shown that while some bivalent genes are expressed at low levels others are not expressed. To determine the expression status of the identified double marked genes and perhaps resolve why only a subset of these became upregulated, the expression status of these was assessed.

4.5.2 PRC2 targets susceptible to upregulation following loss of H3K27me3 are transcribed at low levels.

Transcripts for 13 815 genes were detected in the DMSO treated SEM cells from the previously described Nascent RNAseq experiment. Overlap of these with PRC2 targets showed that 1 286 PRC2 targets were expressed (1 114 of these were double marked) (Figure 34). However, the level of expression of these was substantially lower than the average of all expressed genes (Figure 34). 97% of genes that became upregulated following UNC1999 treatment were found to be expressed in the DMSO treated cells. Therefore it seems that genes that are susceptible to upregulation are those that are expressed (albeit at low levels) and marked by H3K4me3.

However, only 58% of all expressed PRC2 genes became upregulated in either the Low, Medium and High treatment conditions. To determine whether the expression level of genes was related to its susceptibility to upregulation

following inhibition of PRC2, the two groups were compared. Frequency distribution of the two groups showed that, although the unchanged group had a lower percentage of the most lowly expressed genes and a slightly higher proportion of "average" expressed genes (Figure 34), there were no gross difference that could explain the differential response of these genes. It can therefore be concluded that PRC2 targets susceptible to upregulation are those that are marked by H3K4me3 and transcribed, although enrichment of H3K4me3 and expression status do not guarantee de-repression.

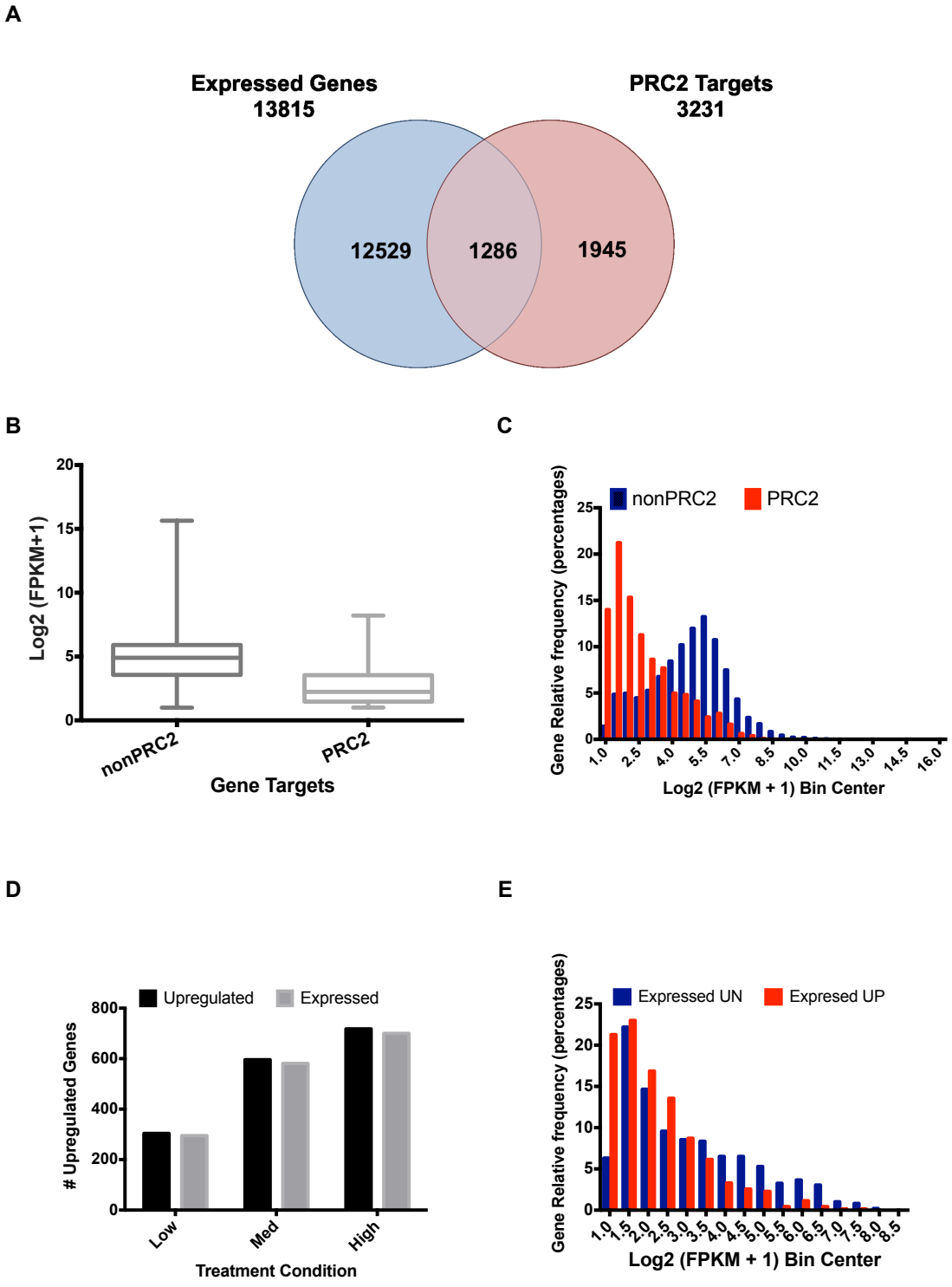


Figure 34 Genes susceptible to upregulation are expressed at low levels

A) Overlap between all expressed genes and PRC2 targets in DMSO treated cells. B) Level of expression of non-PRC2 bound genes and PRC2 target genes. Boxes represent median with interquartile ranges, error bars represent minimum and maximum values. C) Frequency distribution of expression values of nonPRC2 and PRC2 targets. Bin size: 0.5. D) Total number of upregulated genes in the indicated conditions (black bars) and number of upregulated genes that have detectable transcripts in DMSO treated cells (grey bars). E) Frequency distribution of expression levels of genes with detectable transcripts in the DMSO treated cells that remain unchanged (blue bars) or become upregulated (red bars) in samples treated with High doses of UNC1999. Bin size: 0.5.

4.5.3 Association of PRC2 genes with compaction and transcription stage

Chromatin compaction is one of the proposed mechanisms by which PcG proteins are thought exert their repressive activity (Endoh et al., 2012; Francis et al., 2004; Mousavi et al., 2012). Compaction of chromatin is thought to impede binding of transcription factors or the transcriptional machinery to DNA (Chopra et al., 2011). The exact molecular mechanism by which compaction is mediated is not fully understood. It is predominantly attributed to the function of PRC1 complexes, although reports of PRC2-EZH1 mediated chromatin compaction also exist (Endoh et al., 2012; Margueron et al., 2008). On the other hand, reports have found an association between bivalent domains with the paused form of polymerase II (Brookes et al., 2012; Guenther et al., 2007; Stock et al., 2007). Consequently PRC2 has been proposed to exert its repressive function by inhibiting transcriptional elongation.

In view of the findings mentioned above, it is possible that PcG mediated repression can be achieved by either compaction or inhibition of transcriptional elongation in a locus specific manner. To investigate whether the susceptibility of PRC2 targets to derepression following inhibition of its methyltransferase activity was determined by the "mode" of repression, the chromatin accessibility and polymerase occupancy of PRC2 targets was analysed. Since upregulated genes were marked by H3K4me3 and showed low level expression, it was likely that that these would be associated with paused polymerase II.

4.5.3.1 PRC2 targets susceptible to upregulation are associated with regions of accessible chromatin

To assess the chromatin accessibility of PRC2 targets, ATACseq was carried out in the SEM cell line. 13 746 genes were found to be in regions of accessible chromatin (+/-2kb of TSS) and overlap of these with PRC2 targets identified 1,344 PRC2 targets associated with accessible chromatin (Figure 35). Comparison of upregulated and unchanged PRC2 targets revealed higher ATACseq peaks at genes that become upregulated although the enrichment was substantially lower compared to MLL targets (known to be expressed at high levels) (Figure 35). DM unchanged genes showed similar enrichment of ATACseq reads to DM upregulated genes. 76%, 73% and 72% of upregulated genes in the Low, Medium and High conditions were associated with accessible chromatin. While all upregulated genes are presumably associated with accessible chromatin, a possible explanation for the low level enrichment of ATACseq peaks

is cell population heterogeneity. It may be that relatively few cells are actively transcribing a given locus which would be associated with accessible chromatin and that the due to the sensitivity of the technique these are not detected. As only 7 million unique reads were generated in the ATACseq experiments it may be that the depth of sequencing is insufficient to identify these regions.

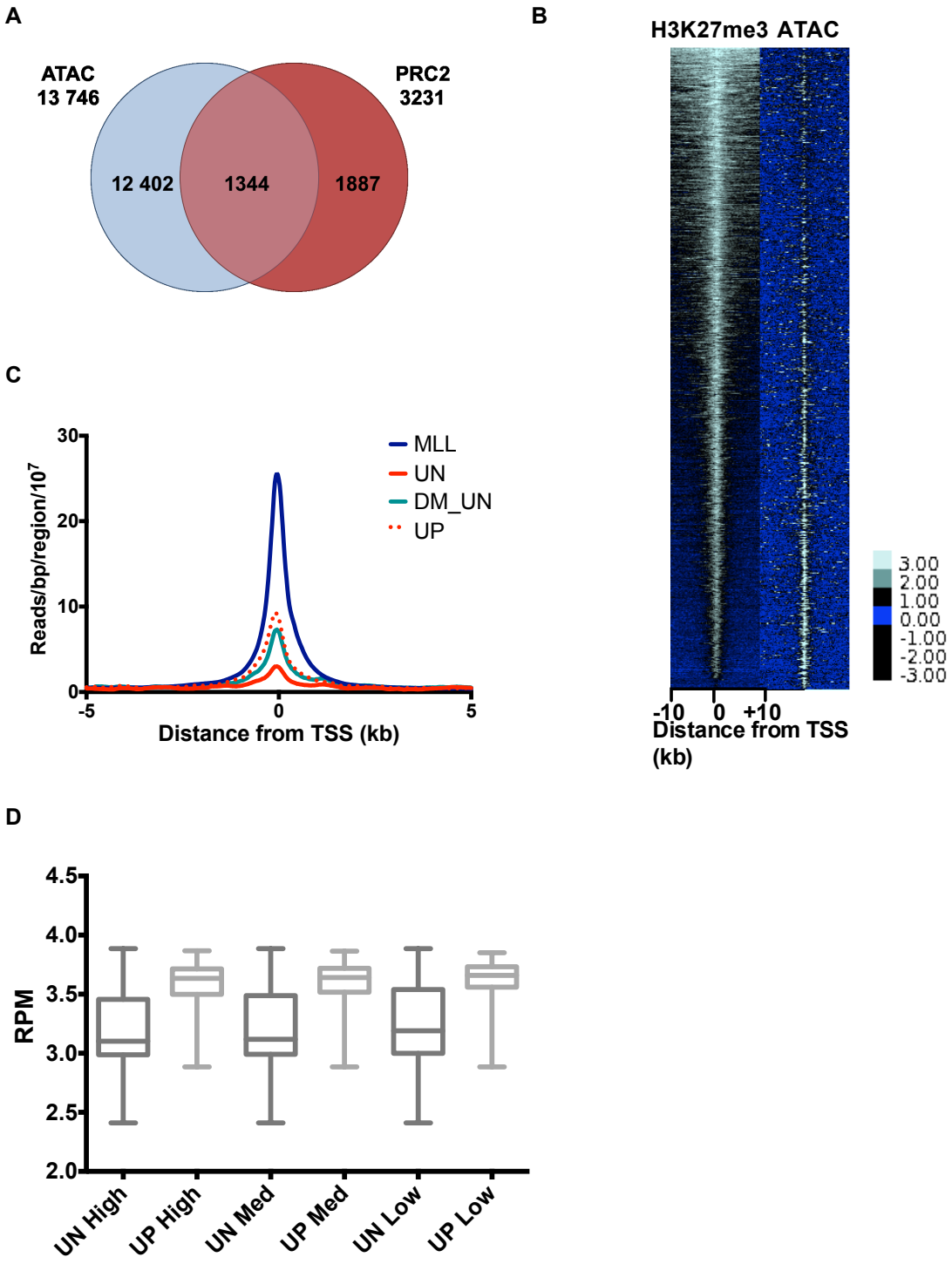


Figure 35 Upregulated PRC2 targets are in regions of accessible chromatin

A) Overlap analysis of regions of accessible chromatin defined by ATACseq and PRC2 targets. B) Heatmap of PRC2 targets ordered by decreasing H3K27me3 enrichment showing ATAC

enrichment. C) Histogram of the mean ATAC peak enrichment at upregulated, unchanged and double marked unchanged PRC2 targets. Erichment of ATAC peaks at MLL targets are shown for comparison purposes. D) Enrichment of ATAC peaks in DMSO treated cells at genes that become upregulated (UP) or remain unchanged (UN) following UNC1999 treatment with 0.25μM (Low), 1μM (Med) and 5μM (High) UNC1999.

4.5.3.2 PRC2 targets susceptible to upregulation are not associated with paused polymerase

To determine whether PRC2 genes susceptible to upregulation were associated with a particular stage of transcription, ChIPseq profiles of components involved in transcription initiation and elongation were compared to those of PRC2.

To determine whether PRC2 targets were associated with the pre-initition complex (PIC), MED1 ChIPseq was analysed. MED1 is a component of the Mediator complex, which serves as a scaffold for the assembly of the PIC, and remains bound during transcription elongation (reviewed in (Allen and Taatjes, 2015) . To determine the association of PRC2 targets with factors involved in transcriptional elongation, ChIPseq binding patterns of two members of the elongation complex PAF1C were analysed, namely LEO1 and PAF1. Finally, to determine whether PRC2 targets were associated with paused polymerase ChIPseq for polymerase II phosphorylated on serine 5 of the CTD heptapeptide repeats (PolS5) was analysed.

Direct targets of the above mentioned proteins were defined by enrichment of the proteins at gene promoters. Overall ~6 500 gene targets were identified for each component and the overlap with PRC2 was minimal ranging from ~5.8%-13%. While this overlap was small, given the relatively few genes that become upregulated following inhibition of PRC2 with the UNC1999 inhibitor, it was possible that they were marked by these factors. However, this this was not the case. MED1 was bound at ~25% of upregulated gene targets in each of the three treatment conditions of UNC199, while PolS5 was bound at ~11% and LEO1 and PAF1 were found at 3-9% of gene targets. Since identification of low levels of enrichment through peak calling of ChIPseq tracks is not very sensitive, a second approach was chosen to investigate the relationship of these factors with PRC2 targets, that does not rely on peak-calling.

A

Protein	Gene Targets	PRC2 Targets
MED1	6 686	423
PolS5	6 567	189
PAF1	6 341	218
LEO1	6 178	276

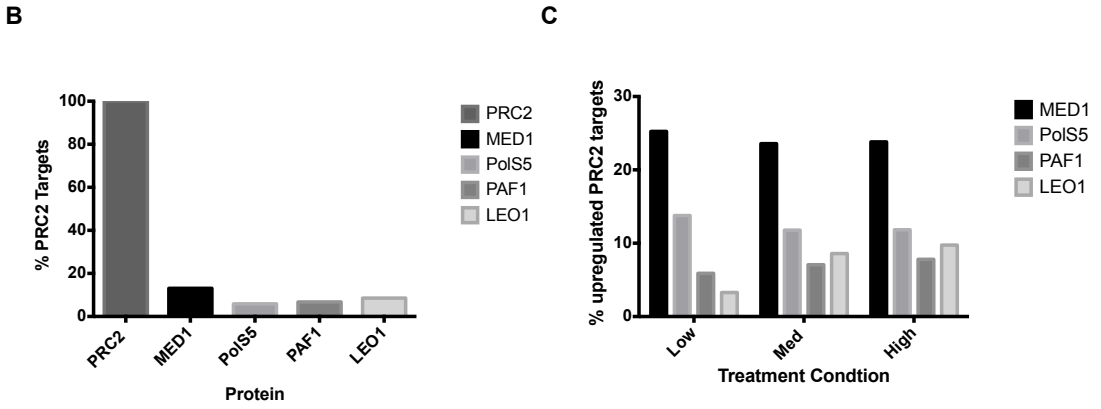
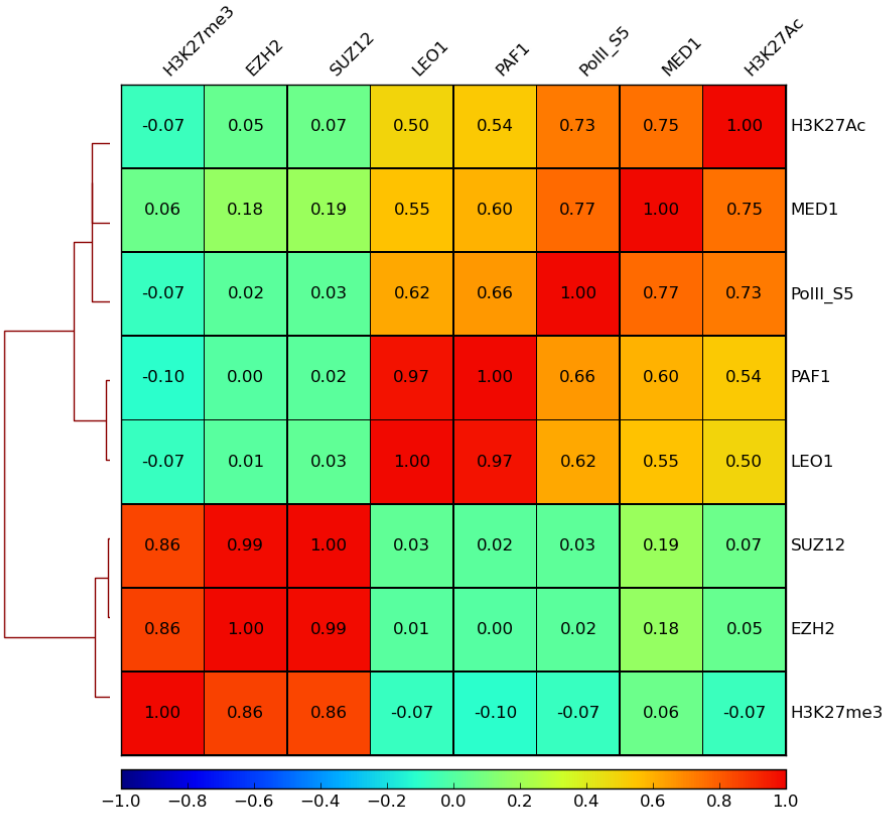
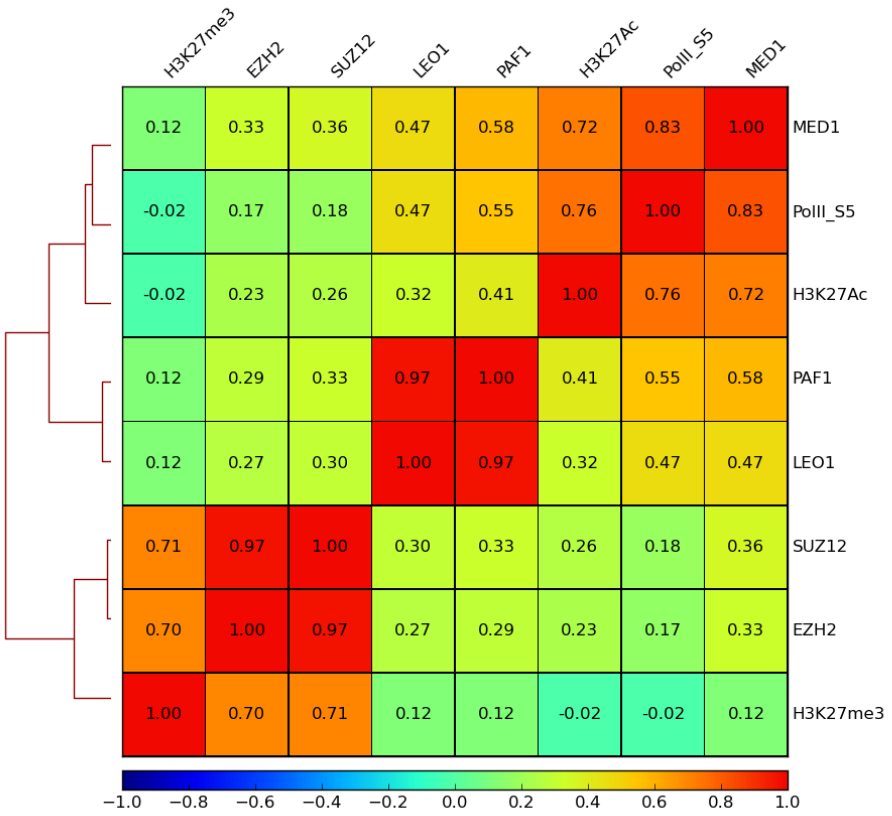


Figure 36 Mnimal overlap between PRC2 targets and components involved in transcription

A) Table summarising the number of gene targets identified for each protein in the SEM cell line, and the overlap with PRC2 targets. B) Percent of all PRC2 targets bound by indicated proteins C) Percent of upregulated genes in the indicated condition bound by the indicated proteins.

To assess the relationship between these components the bam correlate tool was used. Briefly, this tool calculates the similarity of multiple ChIPseq samples, i.e it determines whether regions enriched in one sample are also enriched in another one. The similarity of enrichment is outputted as a correlation coefficient. To determine whether PRC2 targets were enriched for LEO1, PAF1, PolIIS5 and MED1 a correlation analysis was performed which included ChIPseq data for EZH2, SUZ12, H3K27me3 as well as H3K27Ac. This revealed that LEO1, PAF1, PolS5 and MED1 and H3K27Ac correlated with each other but not with EZH2, SUZ12 or H3K27me3 (Figure 37). The same type of analyses was repeated but limited to the 718 gene targets that are susceptible to upregulation following treatment with a high dose of UNC1999 (Figure 37). This was compared to analysis of 718 randomly selected unchanged genes (i.e. genes that are not susceptible to upregulation following UNC1999 treatment) (Figure 37). The most obvious difference between the groups was the closer grouping of the elongation factors LEO1 and PAF1 to MED1, PolS5 and H3K27Ac at the upregulated genes, suggesting that genes with associated elongating factors are the ones susceptible to upregulation, consistent with the previous results showing low level expression of these genes.


H3K27me3
EZH2
SUZ12
LEO1
PAF1
PolII_S5
MED1
H3K27Ac
-0.07 0.05 0.07 0.50 0.54 0.73 0.75 1.00 H3K27Ac
0.06 0.18 0.19 0.55 0.60 0.77 1.00 0.75 MED1
-0.07 0.02 0.03 0.62 0.66 1.00 0.77 0.73 PolII_S5
-0.10 0.00 0.02 0.97 1.00 0.66 0.60 0.54 PAF1
-0.07 0.01 0.03 1.00 0.97 0.62 0.55 0.50 LEO1
0.86 0.99 1.00 0.03 0.02 0.03 0.19 0.07 SUZ12
0.86 1.00 0.99 0.01 0.00 0.02 0.18 0.05 EZH2
1.00 0.86 0.86 -0.07 -0.10 -0.07 0.06 -0.07 H3K27me3
−1.0 −0.8 −0.6 −0.4 −0.2 0.0 0.2 0.4 0.6 0.8 1.0


H3K27me3
EZH2
SUZ12
LEO1
PAF1
H3K27Ac
PolII_S5
MED1
0.12 0.33 0.36 0.47 0.58 0.72 0.83 1.00 MED1
-0.02 0.17 0.18 0.47 0.55 0.76 1.00 0.83 PolII_S5
-0.02 0.23 0.26 0.32 0.41 1.00 0.76 0.72 H3K27Ac
0.12 0.29 0.33 0.97 1.00 0.41 0.55 0.58 PAF1
0.12 0.27 0.30 1.00 0.97 0.32 0.47 0.47 LEO1
0.71 0.97 1.00 0.30 0.33 0.26 0.18 0.36 SUZ12
0.70 1.00 0.97 0.27 0.29 0.23 0.17 0.33 EZH2
1.00 0.70 0.71 0.12 0.12 -0.02 -0.02 0.12 H3K27me3
−1.0 −0.8 −0.6 −0.4 −0.2 0.0 0.2 0.4 0.6 0.8 1.0

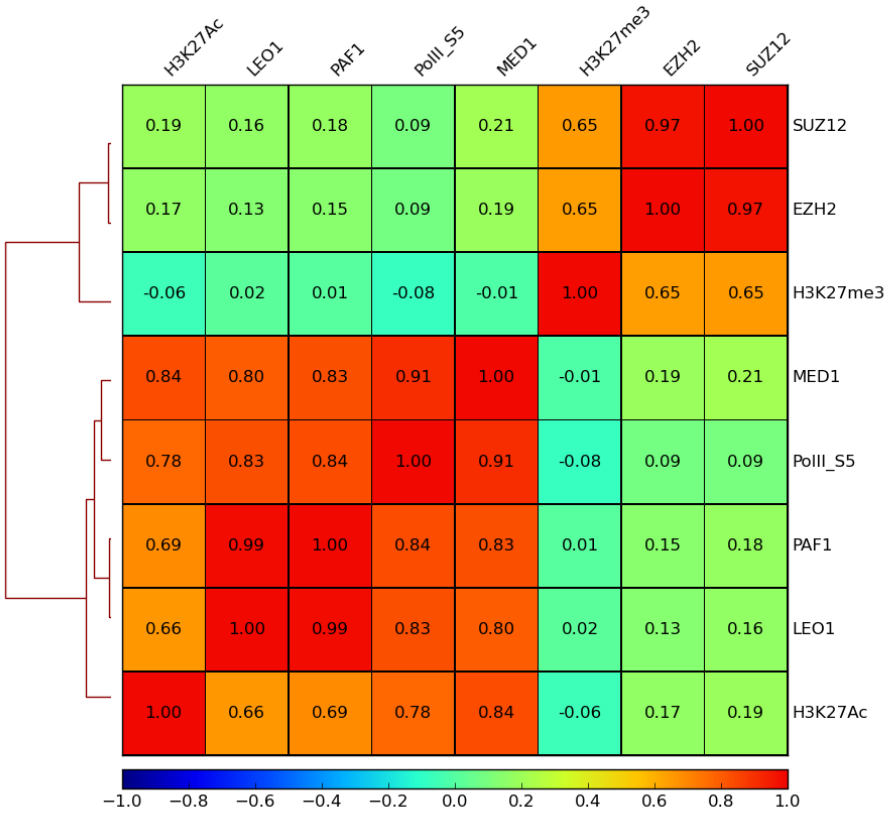


Figure 37 Upregulated genes are not associated with paused polymerase

Correlation plots showing similarity of indicated ChIPseq experiments at PRC2 targets in untreated (MED1, PolIIS5, PAF1, LEO1) or DMSO treated cells (EZH2, SUZ12, H3K27me3, H3K27Ac). Analysis was limited to either, Top: All PRC2 targets, Middle: unchanged PRC2 targets following UNC1999 treatment with the High dose, Bottom: PRC2 targets upregulated in the High treatment conditions. Numbers indicate Pearson correlation coefficient.

4.5.4 Genes susceptible to upregulation are bound by RING1B

RING1B was bound at promoters of 10,011 gene targets. 2 190 of these overlapped with PRC2 targets (representing 68% of PRC2 targets) (Figure 38). 72% of these were marked by H3K4me3, representing 95% of all DM genes. 97%, 95% and 96% of genes upregulated in the Low, Med and High conditions

were also targets of RING1B. Taken together the data indicate that the majority of genes bound by both PRC2 and PRC1 are DM and consequently genes that are susceptible to upregulation are co-bound by PRC1 and PRC2 while genes that do not become upregulated are more likely to be bound solely by PRC2.

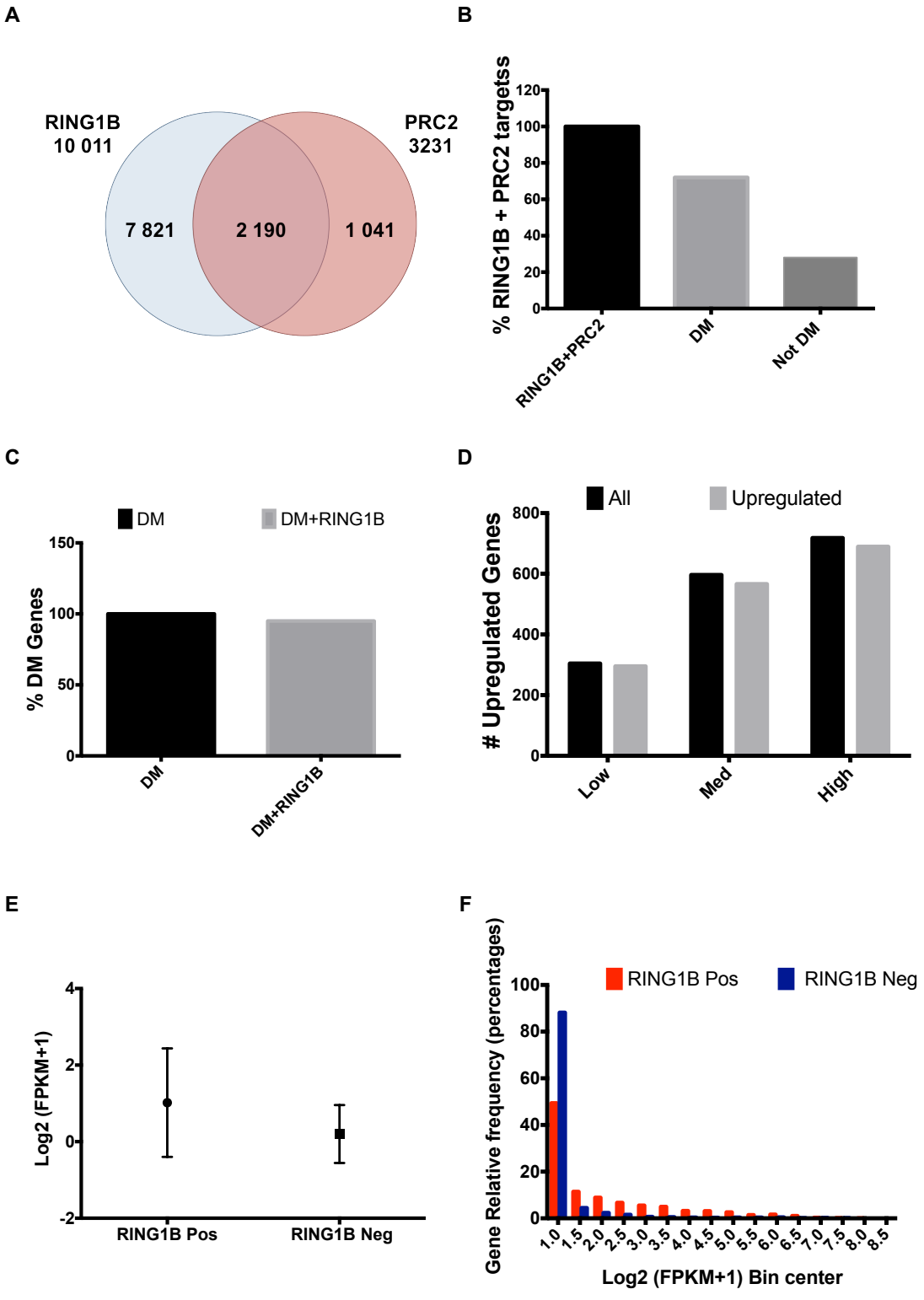


Figure 38 RING1B is bound at upregulated genes

RING1B promoter targets were identified through ChIPseq (data was available in the lab) and were overlapped with PRC2 targets. A) Overlap between RING1B and PRC2 promoter targets B) Percentage of RING1B+PRC1 sites that are double marked (DM) or not double marked (Not dm). C) Percentage of DM marks that are co-bound by RING1B (DM+RING1B), DM: all DM targets D). Number of genes that are upregulated in the indicated treatment condition and are co-bound by RING1B (grey bars), all upregulated genes (black bars). D) Mean gene expression value of PRC2 targets cobound by RING1B (RING1B Pos), and PRC2 only targets (RING1B Neg), lines represent interquartile ranges. E) Frequency distribution of nascent-RNAseq expression levels of PRC2 site not bound by RING1B (RING1B Neg, blue bars) and those bound by RING1B (RING1B pos). bin size= 0.5.

4.6 Effect of inhibition of PRC2 on TrxG function

As discussed previously in the general introduction, the responsive model of PcG recruitment envisions that PcG and TrxG proteins sample chromatin possibly through proteins containing CXXC-domains (Klose et al., 2013). In the absence of active transcription formation of PcG domains occurs while active transcription favours binding of TrxG proteins. Importantly, it is proposed that binding of PcG proteins antagonises TrxG binding and vice versa. It therefore follows, that loss of PcG binding should result in increased binding of TrxG proteins. However, this model also acknowledges the importance of activating signals such as TF in permitting the establishment of TrxG domains and consequently not all sites that lose PcG binding are expected to gain TrxG. Since bivalent domains are bound by both PcG and TrxG proteins they are thought to be the result of low level activating signals that on the one hand prevent PcG domain formation but are

insufficient to permit the establishment of a proper TrxG domain. Therefore bivalent domains provide excellent target sites to test the antagonism between PcG and TrxG.

Recently the MLL2 protein has been identified as important for the generation of bivalent domains in ES cells (Denissov et al., 2014). Additionally, a subset of bivalent targets was found to be dependent on MLL2 for proper expression following ATRA induced differentiation. Similarly, MLL was also found to be important for correct expression of a subset of genes, as induction of differentiation in double MLL2 MLL knockout cells additionally perturbed the upregulation of 39% of genes not altered in MLL2 single knockout cells (Denissov et al., 2014) .

Since MLL ChIPseq data was available for the SEM cell line it was used to determine whether any double marked genes were bound by this protein and investigate whether any changes in enrichment occurred following loss of H3K27me3, as an example TrxG protein. Similarly, levels of H3K4me3 following inhibition of PRC2 were also analysed.

4.6.1 MLL binds to a subset of PRC2 targets

Overlap analysis of MLL and PRC2 targets identified 828 targets co-bound by both proteins (+/- 2kb of TSS), representing 25% of all PRC2 targets and 7.9% of

all MLL targets. As expected, 99% of co-bound genes were marked by both H3K4me3 and H3K27me3 (Figure 39). Of the genes susceptible to upregulation as a result of PRC2 inhibition ~50% were bound by MLL. To test whether MLL binding increased following inhibition of PRC2 at two of these targets, *TOX* and *CKAP4,* ChIPqpcr was performed on samples treated with DMSO, Analog or Low, Med and High doses of UNC1999. These samples were collected from the same experiment where a dose-dependent decrease of H3K27me3 was seen. This revealed an increase in MLL binding at the *TOX* locus but no substantial change at the *CKAP4* locus

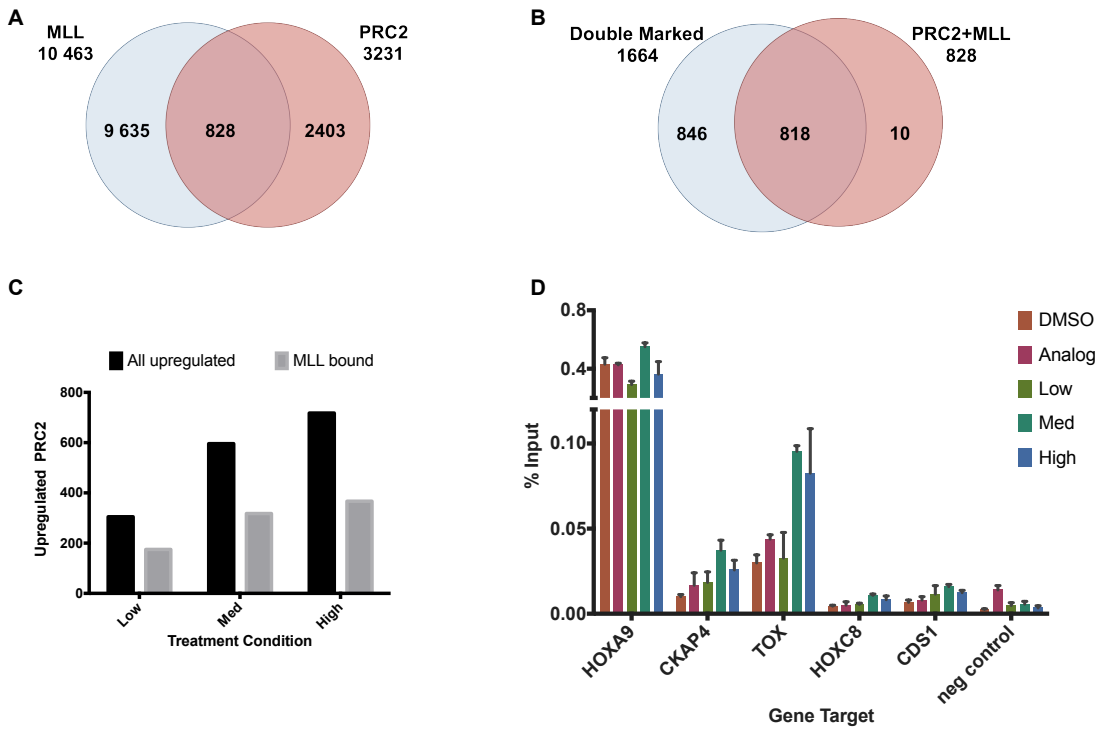


Figure 39 MLL and PRC2 co-bind a subset of DM sites

A) Overlap analysis of MLL and PRC2 promoter targets. B) Overlap of double marked genes with PRC2 and MLL cobound genes (PRC+MLL) C) Number of upregulated PRC2 genes following UNC1999 in the indicated conditions, all upregulated genes (black bars), MLL bound genes (grey

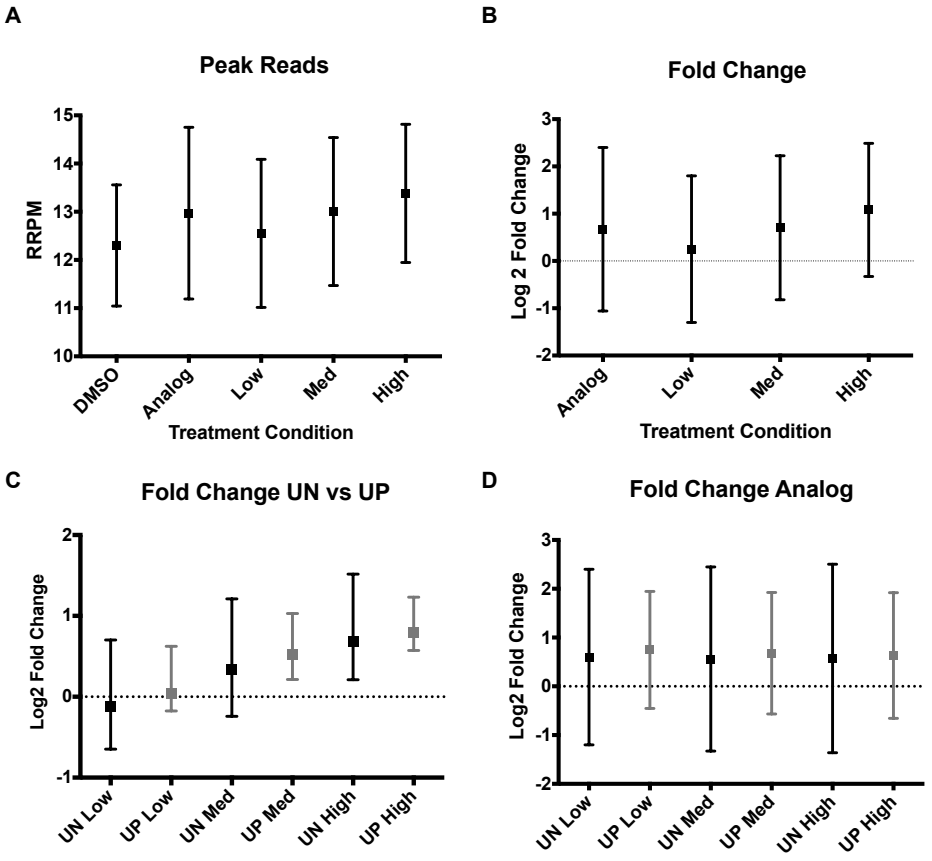
bounds). D) ChIP-qPCR using an antibody against MLL, following treatment of cells with the indicated doses of UNC1999 , Analog and DMSO. Error bars represent SD of two biological replicates.

4.6.2 H3K4me3 levels increase at PRC2 targets following UNC1999

To further test the antagonism of PcG and TrxG proteins, ChIP-Rx was performed for H3K4me3 in samples treated with Low, Med and High doses of UNC1999 and compared to samples treated with DMSO or Analog (Samples where derived from the same experiment used for H3K27me3 ChIP-Rx and H3K27Ac ChIP-seq). While TrxG proteins have functions beyond methylation of H3K4, since this is one of the marks that defines bivalent domains it can be used as a proxy for determining whether decreased PcG activity results in increased TrxG activity at these sites.

To assess the levels of H3K4me3 at double marked genes, the number of reads at H3K4me3 peaks of these was normalised with the reference derived normalisation factor. Surprisingly, only small changes in UNC1999 treated samples, compared to the DMSO control were observed when this analysis was performed on all double marked targets. 389, 292, 362 and 408 genes showed at least a 2 fold log change in H3K4me3 (Figure 40). While a small dose dependent increase in H3K4me3 levels was observed between the Med and High treated samples, the levels of H3K4me3 increased to the same degree in the Analog treated samples as in the Med treated sample, suggesting that this may be noise

or some low level activity of the Analog (Figure 40). The levels of H3K4me3 did not change in the Low treated samples, compared to the DMSO control. Genes that became upregulated following inhibition of UNC1999 showed slightly higher levels of H3K4me3 than did unchanged genes, although a similar effect was observed in samples treated with the inactive Analog.



E

Peak Fold Change

UNC1999 dose	Log2 Fold Change
Analog	0.67
Low	0.25
Med	0.7
High	1.1

F

UNC1999 Fold Change

UNC1999 dose	Unchanged	Upregulated
Low	0.6	0.75
Med	0.56	0.68
High	0.57	0.63

G

Analog Fold Change

UNC1999 dose/Genes	Unchanged	Upregulated
Low	0.18	0.34
Med	0.52	0.76
High	1	1.1

Figure 40 Small increases in H3K4me3 levels following inhibition of UNC1999 treatment

ChIP-Rx was performed on cell treated with Low, Med, High doses UNC1999 , Analog and DMSO for five days. Levels of enrichment were measured at H3K4me3 peaks of double marked genes. All enrichment levels are normalised with the reference derived normalisation factor A) Reads

under H3K4me3 peaks at DM gene B) H3K4me3 enrichment level fold change following treatement with the indicted doses compared to DMSO at all DM targets. C) H3K4me3 enrichment level fold change following treatement with the indicted doses compared to DMSO for unchanged (UN) and (UP) genes. D) H3K4me3 enrichment level fold change in the Analog treated sample, at genes upregulated (UP) or unchanged (UN) in the indicated conditions. E) Table summarising peak fold change data, related to B. F) Table summarising fold change of H3K4me3 following treatment with UNC1999, related to C. G) Table summarising fold change of H3K4me3 following treatment with UNC2400, related to D.

To confirm these results ChIP-qPCR on samples treated with either DMSO and Analog and Low, Med and High doses of the UNC1999 for 5 days was performed. Consistent with the ChIPseq results, no changes in H3K4me3 levels were observed at two double marked loci, previously shown to lose H3K27me3 and increase their expression in a dose dependent manner, namely CKAP4 and TOX (ChIP-Rx log2 fold change of 1 and 2 respectively) (Figure 41). Levels of H3K4me3 were higher at these loci than at two targets marked only by H3K27me3, HOXC8 and CDS1 (the expression of these genes is unaltered following inhibition of UNC1999) confirming them as double marked genes. Levels of H3K4me3 were unaltered at the HOXA9 locus used as a positive control (Figure 41). Importantly, H3K4me3 levels were not altered at TOX, CKAP4 or HOXA9 in samples treated with the Analog (Figure 41). Therefore the data suggests that inhibition of the catalytic activity of PRC2 does not result in large increases of H3K4me3. However, more targets need to be analysed by ChIP-qPCR

to verify this. While this data is surprising, it is consistent with the low level expression of genes that become upregulated following inhibition of PRC2.

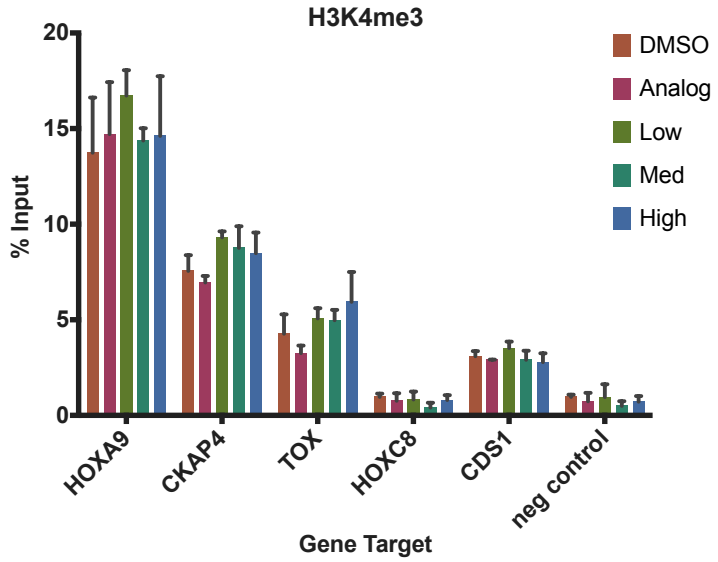


Figure 41 H3K4me3 levels do not change susbstantially following treatment with UNC1999 (ChIPqPCR)

ChIP-qPCR using an antibody against MLL, following treatment of cells with the indicated doses of UNC1999 , Analog and DMSO. Error bars represent SD of two biological replicates.

4.6.3 Discussion

Epigenetic inhibitors are promising agents for the treatment of haematological malignancies, as well as solid tumours. However, while two DNMT inhibitors have been approved by the FDA for the treatment of MDS and AML, only 50% of patients show response to treatment. A similar scenario is seen for HDAC inhibitors. Predicting response to epigenetic therapy has been difficult, in part through the lack of sufficient knowledge regarding the molecular mechanism by which these inhibitors exert their anti-leukaemic effects. Therefore, detailed knowledge of the molecular mechanism by which chromatin proteins, and consequently how inhibition of these, affect gene regulation can aid in the design of rational therapeutic approaches.

In this study, inhibition of PRC2 was found to exert a potent anti-leukaemic effect on a t(4;11) patient derived cell line, *in vitro.* Use of the small molecule inhibitor UNC1999, has allowed for the study of the effects of one aspect of PRC2 function, namely its catalytic activity. This approach allowed for the dissection of how the H3K27me3 PTM contributes to gene repression and the leukaemic properties of the SEM t(4;11) cell line.

4.6.4 PRC2 inhibition leads to low level upregulation of genes

Studies using either shRNA-mediated knockdowns, pharmacological inhibition of PRC2 or knockout cells of various PRC2 components, have reported only a few

gene expression changes. These commonly range between ~173-500 genes based on mRNA levels detected either by microarray analysis or through RNAseq. In contrast in this study, by analysing nascent-RNAseq, up to 1 647 genes were identified as upregulated following inhibition of PRC2. While it is possible that SEM cells are particularly sensitive to PRC2 inhibition, this is more likely a direct result of the sensitivity of the approaches used in detecting transcriptional changes. Importantly, 12 gene targets identified as upregulated by nascent-RNAseq showed at least a 2-fold increase in mRNA levels as detected by rt-qPCR. This indicates that the observed changes in nascent transcripts result in detectable increases in mRNA and potentially at the protein level, although this was not assessed.

Similar to previously published data, only a subset of PRC2 targets became upregulated following loss of H3K27me3. The mean level of expression of upregulated genes corresponded to the expression levels of the lowest 20% of non-PRC2 targets. Although statistical analysis to determine the likelihood of this occurring by chance was not performed, this suggest that loss of H3K27me3 is not sufficient to induce robust transcriptional changes (possible reasons for this are discussed below). Nevertheless, given the profound defects in proliferative ability of SEM cells following PRC2 inhibition, this low level upregulation of genes appears sufficient to induce a significant biological changes.

Pathway analysis of upregulated PRC2 targets identified a few candidates that may be mediating the cell-cycle arrest seen in SEM cells, following inhibition of

PRC2. These included *EBF3*, *FOXO1*, *SESN3* and *RUNX3*. Their upregulation was confirmed by rt-qPCR but not at the protein level. Validation of these targets as tumour suppressor in the context of t(4;11) leukaemias could be achieved by their overexpression, either through use of plasmids or CRISPR- technology (CRISPR-VP64). Conversely, knockdown or knockouts of these could be coupled to treatment with UNC1999 followed by assessment of their proliferative capacity and cell cycle profiles. Additionally, to identify other gene targets that may be mediating the anti-leukaemic effects observed following inhibition of PRC2, a positive shRNA screen could be performed.

4.6.5 ChIP-Rx reveals global decrease in H3K27me3

Combination of different doses of the UNC1999 inhibitor coupled to ChIP-Rx has allowed for the first time, the genome-wide assessment of how different levels of H3K27me3 contribute to gene silencing. The UNC1999 inhibitor resulted in a dose dependent decrease of H3K27me3 globally, an observation that was only detected when ChIP-Rx was employed. Analysis by the traditional ChIPseq method attenuated the observed decreases of H3K27me3 substantially. Samples treated with the Low dose of UNC1999 even seemed to have acquired H3K27me3, a result that is not consistent with either immunoblot analysis or ChIP-qPCR. Therefore, caution should be taken when interpreting ChIPseq data from experiments where the function/levels of protein complexes are attenuated (reasons for this were discussed in section 4.3)

For instance, previously, the effects of the UNC1999 inhibitor on H3K27me3 levels has been assessed in MLL-AF9 transformed mouse progenitor cells (leading to AML) (Xu et al., 2014). Traditional ChIPseq of H3K27me3 resulted in the loss of H3K27me3 at 66% of all its sites and increases in 8% of sites. Furthermore, in the study by Xu et al. (2014) the authors observed a preferential loss of H3K27me3 at sites of little enrichment, with few losses observed at proximal promoter regions. When loss of H3K27me3 was detected at promoters, it seemed to occur at non- CpG promoters. From the data the authors concluded treatment of UNC1999 preferentially affects H3K27me3 erasure at distal regulatory regions, such as enhancers. In contrast, in this study, by employing ChIP-Rx H3K27me3 levels were lost at 93%, 99.5% and 100% of PRC2 promoter targets in samples treated with Low, Med and High doses of UNC1999 respectively, and no increases in H3K27me3 were detected. The preferential loss at sites of little enrichment observed by Xu et al through the traditional ChIPseq approach, fits well with the idea that ChIPseq signal is amplified at sites that retain low levels of H3K27me3 when library complexity is affected, as the levels are expressed as a simple percentage of all mapped reads (Orlando et al., 2014). The preferential erasure of H3K27me3 at non-CpG promoters is probably explained by lower levels of H3K27me3 at these sites, since PcG proteins have been shown to bind and nucleate at CpG islands, the density of which has been correlated with PcG binding.

4.6.6 H3K27me3 loss result in small dose-dependent increases in gene expression

UNC1999 treatment with three different concentrations of UNC1999 led to dose-dependent decreases of H3K27me3 levels. On a gene-by-gene basis these levels were correlated with gene expression changes, indicating that the levels of H3K27me3 are directly modulating expression of genes. However, similar levels of H3K27me3 had different effects on different genes reflecting the complexity of transcriptional regulation. When gene expression levels were detected, these were small, consistent with the low level acquisition of H3K4me3 at these targets (H3K4me3 levels correlate positively with the level of transcription). This may be a result of the presence of low levels of activating signals or perhaps the presence of PRC1 components despite the loss of H3K27me3. While canonical PRC1 complexes can bind H3K27me3 through the chromo-domain containing CBX proteins and their binding is affected by the loss of PRC2 catalysed H3K27me3, recently it has been shown that variant PRC1 can bind independently of PRC2 and even serve to recruit it. Since almost all PRC2 upregulated genes were also bound by RING1B, it is possible that despite loss of H3K27me3 variant-PRC1 complexes may serve to prevent “robust” expression of these genes. ChIP-qPCR experiments following treatment with UNC1999 for PRC1 components would be useful in determining the contribution of PRC1 to repression of these genes in the absence of H3K27me3.

4.6.7 A signature of genes susceptible to upregulation

Recently, PcG domains have been shown to be recruited to thousands of sites in response to transcriptional inhibition. This fits well with the proposed responsive model of recruitment, whereby PcG and TrxG proteins are proposed to sample and respond to the chromatin environment of genes. Therefore, establishment of PcG domains seems to occur when there is a lack of activating signals, while bivalent domains are a result of low level activating signals that allow for both PcG and TrxG binding. This may serve to explain why only double marked or so called bivalent genes became upregulated following inhibition of PRC2, as these would be the only ones where activating signals are present. Supporting this notion further is the observation that all upregulated genes showed low-level transcription in the samples treated with DMSO, and previous reports have reported the regulation of transcribed genes by the PRC1 member RING1B (Brookes et al., 2012).

However a proportion of genes that showed low level transcription and a large fraction of double marked genes were not upregulated following inhibition of PRC2. This may be a consequence of the detection limits of the nascent-RNAseq as two genes not detected as differentially expressed were found to be upregulated by rt-qPCR (CCNA1, SOX13). However, two other double marked genes (TGFB3, CDS1) were not found to be expressed even by rt-qPCR, although it is still possible that some low level upregulation is taking place. Alternatively,

as discussed previously, this may represent genes that rely on the activity of other repressive complexes, such as PRC1.

Contradicting reports of bivalent domain association with paused polymerase, in this study double marked genes were not found to be associated with transcriptional pausing, and neither were genes susceptible to upregulation. Instead comparison of unchanged and upregulated genes revealed the closer association of factors involved in transcription initiation (MED1) with elongation factors (LEO1, PAF1) at these genes, consistent with the production of low level of transcripts.

Interestingly, while not all RING1B PRC2 co-bound genes became upregulated, PRC2 only genes were not susceptible to upregulation at all. This contrast suggests that bivalent domains bound by both PRC2 and PRC1 are more efficient at maintaining the repressed state, based on the observation that these sites preferentially retain PcG domains following differentiation of ES cells.

4.6.8 Antagonism of PcG and TrxG proteins

Given that the UNC1999 inhibitor specifically targets the catalytic function of the PRC2 complex and leads to the upregulation of double marked genes, it provided a useful tool to assess the contribution of H3K27me3 in antagonising TrxG proteins. If the chromatin-sampling model is correct, one could expect an increase in the binding of TrxG proteins following loss of PcG binding.

Additionally, by using different doses of the UNC1999 inhibitor and consequently reducing the levels of H3K27me3 to varying degrees one would expect a similar dose-dependent increasing in binding of TrxG proteins.

The methyltransferases MLL2 and MLL1 have recently been shown to regulate bivalent genes following induction of differentiation in ES cells (Denissov et al., 2014) Therefore ChIP-Rx for the H3K4me3 mark was performed on samples treated with varying doses of the UNC1999 inhibitor, as an example of a TrxG modification (mediated by COMPASS complexes).

Surprisingly, only modest changes were observed in the levels of the mark, and these were more pronounced at upregulated genes compared to double marked unchanged genes. This suggests that similar to PcG proteins TrxG proteins respond to the transcriptional status of the gene, and that proposed positive feedback mechanism to maintain H3K4me3 levels are insufficient to establish "robust" H3K4me3. Alternatively, this may simply indicate that decreasing the levels of H3K27me3 is not sufficient to alleviate the inhibition on COMPASS TrxG complexes. The levels of PRC2 complex components at targets sites was not analysed in this study, although preliminary data (not shown) indicates that there is a dose-dependent decrease in the levels of EZH2 and SUZ12 at these sites. However, Xu et al. (2014) reported that levels of SUZ12 did not change at the *CDKN2A* locus following UNC1999 inhibition of PRC2, despite its upregulation. Therefore, the presence of both PRC2 or PRC1 proteins at these

sites may be sufficient to block TrxG recruitment, consistent with both complexes utilising CXXC-domain containing proteins to bind to target sites.

MLL co-bound ~800 PRC2 double marked genes and therefore its binding following UNC1999 treatment was assessed as an example TrxG protein. ChIP-qPCR showed a modest increase at the upregulated double marked *TOX* locus while no increase was detected at the upregulated *CKAP4* locus or the unchanged *CDS1* locus. While this serves to indicate that TrxG binding may respond to loss of H3K27me3, additional targets need to be assessed and optimisiation of the ChIP-qPCR signal would be beneficial.

Taken together, the data indicate that loss of H3K27me3 results in small changes in H3K4me3 levels, indicating increased TrxG activity at these sites. However, whether additional factors remain that inhibit TrxG function or whether TrxG proteins are responding to activating signals in the same way PcG proteins are, remains an open question.

4.7 Conclusion

Results in this chapter indicate that inhibition of PRC2 by the UNC1999 inhibitor leads to a global dose-dependent decrease in H3K27me3 levels at promoters of PRC2 targets. This is accompanied by a dose-dependent upregulation of a subset of PRC2 targets, some of which have known functions as tumour suppressors. Genes susceptible to upregulation following inhibition of PRC2 are co-bound by

RING1B, marked by H3K4me3, found in regions of accessible chromatin and are transcribed at low levels. While not all genes exhibiting these features become upregulated as a result of decreased H3K27me3, genes that do not possess these features always remain unchanged. Further work in determining how the H3K27me3 mark is affecting complex stability of both PRC1 and PRC2 is an interesting area to pursue. Taken together these findings may serve to predict response to epigenetic inhibition of PRC2 in clinical settings.

5 Chapter 5 Inhibition of PRC2 as a potential therapeutic approach

Combination therapy for childhood ALL has improved over recent years with an approximate 80% event free survival rate (Pui et al., 2004). Despite this MLLr related infant ALL remains difficult to treat with 50% survival rates (Greaves, 1996; Pieters et al., 2007). While combination therapy results in complete remission in ~95% of MLLr ALL infant cases (Frankel et al., 1997; Reaman et al., 1999), most of these patients relapse on treatment within two years of the initial diagnosis. Relapse MLLr patients exhibit resistance to standard chemotherapy, in particular to synthetic glucorticoids (GC) (Stam et al., 2010). There is therefore a need for the development of novel therapeutic strategies for the treatment of this disease.

Epigenetic therapy is a promising new approach for the treatment of multiple forms of cancer. Major successes in the field include the approval of DNA methyltransferase inhibitors and HDAC inhibitors for the treatment of MDS and T-cell lymphoma patients respectively (Fenaux et al., 2009; Issa and Kantarjian, 2009; Khan and La Thangue, 2012; Silverman and Mufti, 2005; Wagner et al., 2010). Recently, clinical trials of the DOT1L inhibitor EPZ-5676 (Daigle et al., 2013) have begun to assess the safety and efficacy of this compound for the treatment of MLLr leukaemias, representing the first human study of a "targeted" HMT inhibitor (Helin and Dhanak, 2013).

Small molecule inhibitors that target EZH2/EZH1 represented potential novel treatment strategies for cancer. The UNC1999 inhibitor is orally bio-available and a single oral dose of 50mg/kg results in UNC1999 plasma concentrations >100ηM for 20h in Swiss albino mice (this is above the cellular IC50 for UNC1999 meassured in MCF10A cells). Additionally, UNC1999 has previously been show to prolong the survival of mice in an MLL-AF9 AML xenograft model (Xu et al., 2014).

Since the UNC1999 inhibitor showed a potent anti-leukaemic effect *in vitro* and was developed for *in vivo* applications, its efficacy as an anti-leukaemic agent of t(4;11) leukaemias was assessed *in vivo.* Additionally, to test whether inhibition of PRC2 would affect normal human haematopoiesis its effect on cord and adult CD34+ cells was assessed.

5.1 UNC1999 only shows mild toxicity in human cord and adult CD34+ cells

To assess the effects of UNC1999 on human haematopoietic cells, colony assays with human cord and adult CD34+ cells were performed. Using appropriated media, these cells are capable of forming erythroid, granulocyte/macrophage and multi-lineage progenitor colonies. Colony assays were performed in combination with Low, Med and High doses of UNC1999. Colony counts were compared to cells treated with DMSO. Treatment of CD34+ cord blood cells with the Low and Med doses of UNC1999 did not affect the total number of colonies

formed significantly, while treatment with the high dose resulted in a ~40% decrease in colony forming ability. This decrease was a result of the loss of erythroid potential as no erythroid colonies were detected. The same experiment carried out on adult CD34+ cells resulted in a 70% decrease in colony forming potential in the High treatment condition compared to DMSO, mostly mediated by the complete loss of erythroid cells (erythroid colonies made up 60% of all colonies in the DMSO treated cells). A small reduction in the formation of erythroid colonies (10%) was observed in the Low and Med conditions as well as a small reduction in granulocyte/macrophage and multi-lineage progenitors cells (10%) in the Med and High treatment conditions. Treatment of CD34+ cord and adult cells did not affect their colony forming ability. The data therefore indicate that inhibition of PRC2 through the UNC1999 inhibitor at doses Low and Med doses does not affect their colony forming ability significantly, suggesting that it may be a possible candidate for therapeutic use.

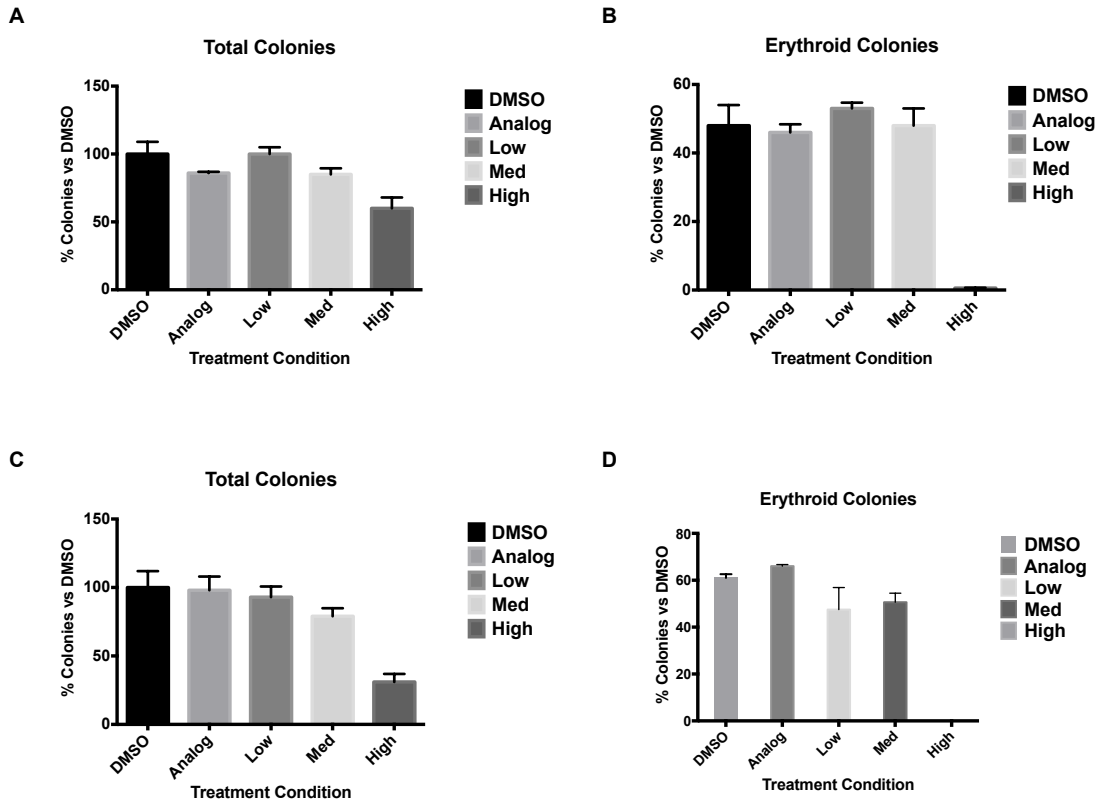


Figure 42 UNC1999 treatment is compatible with colony formation of CD34+ cord and adult cells

CD34+ cord and adult cells were plated on methycellulose media with DMSO, Analog, or low, med and high doses of UNC1999. A and B) Percent colony formation of CD34+ cord bloods compared to DMSO, total colonies (A), eryhtroid colonies (B). Percent colony formation of CD34+ adult cells compared to DMSO, total colonies (C), eryhtroid colonies (D). errors bars represent SD of two biological replicate experiments.

5.2 *In vivo* efficacy of the UNC1999 inhibitor

All *in vivo* experiments were performed by Dr. Frida Ponthan in a collaboration with the group of Prof. Heidenreich, Newcastle University.

SEM cells are capable of engrafting and causing leukaemia in a xenograft model using NSG mice (Thomas et al., 2005). In order to the test the efficacy of UNC1999 in inhibiting leukaemia in this model, the toxicity of the compound had to be assessed in NSG mice, as this had not been done previously. This was performed by administering UNC1999 by oral gavage once per day, with escalating doses of the inhibitor, and monitoring their well-being and weight. Five mice were treated for five days with 50mg/kg UNC1999 followed by two days rest. The dose was then increased to 100mg/kg with the same treatment plan and finally increased to 150mg/kg. The maximum tolerated dose was 100 mg/kg as mice treated with 150mg/kg lost >10% of their weight and consequently treatment had to be stopped (Figure 42).

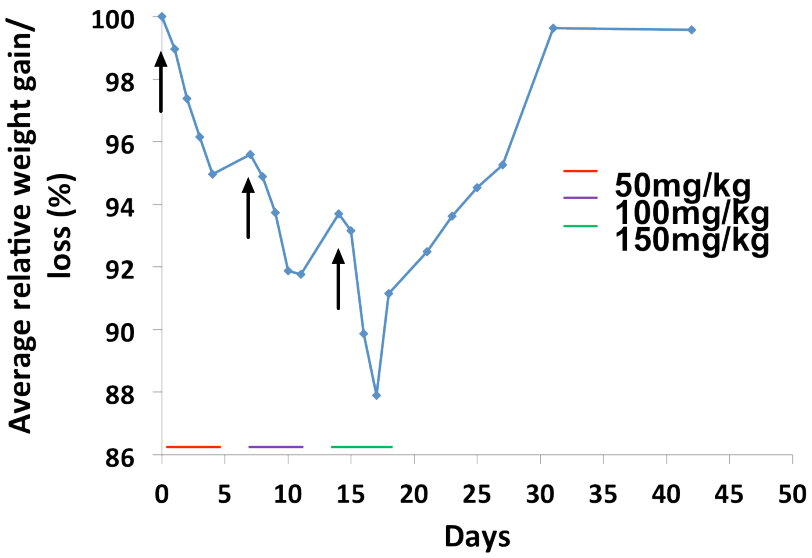


Figure 43 Toxicity assessment of UNC1999 in NSG mice.

UNC1999 was administered by oral gavage in increasing doses. One dose per day was administered for 5 days followed by two days rest. On days 0-4 mice were dosed with 50mg/kg UNC1999, 7-11 with 100mg/kg UNC1999 and 14-18 with 150mg/kg UNC1999. Graph shows the average weight of 5 mice during the treatment period. Arrows indicate dosing at day 0, 7 and 14.

To determine if UNC1999 could inhibit the progression of t(4;11) leukaemia 10,000 SEM cells were injected into the femurs of ten mice. The SEM cell line used, was pSLIEW clone13 which stably expresses firefly luciferase and GFP. Treatment was started 7 days post injection, at which time engraftment was detectable by bioluminescence (data not shown). Five mice were treated with a vehicle control and five mice were treated with 100mg/kg UNC1999 by oral gavage for five days followed by two days rest. Of note is that during the first round of treatment, mice were only treated for three days followed by two days rest. Mice treated with the UNC1999 inhibitor did not show decreased leukaemic burden as measured by bioluminescence and all mice had to be culled 35 days post injection of SEM (pSLIEW clone 13) cells due to signs of leukaemic disease.

To assess the function of the inhibitor on a molecular level, spleen samples collected from vehicle treated mice and UNC1999 were used. The expression of three genes previously shown to be upregulated following inhibition of UNC1999, *TRPS1, CKAP4* and *TOX* was assessed. Mice treated with UNC1999 showed mild upregulation of *TPRS1* while the levels of *CKAP4* and *TOX* were not affected. Consistent with this, examination of H3K27me3 levels by ChIP-qPCR did not showed significant reduction in the levels of this mark compared to vehicle treated mice at two PRC2 target genes (*CKAP4, IKZF3*). It was therefore concluded that the lack of efficacy at reducing leukaemic growth *in vivo* was due to insufficient inhibition of PRC2.

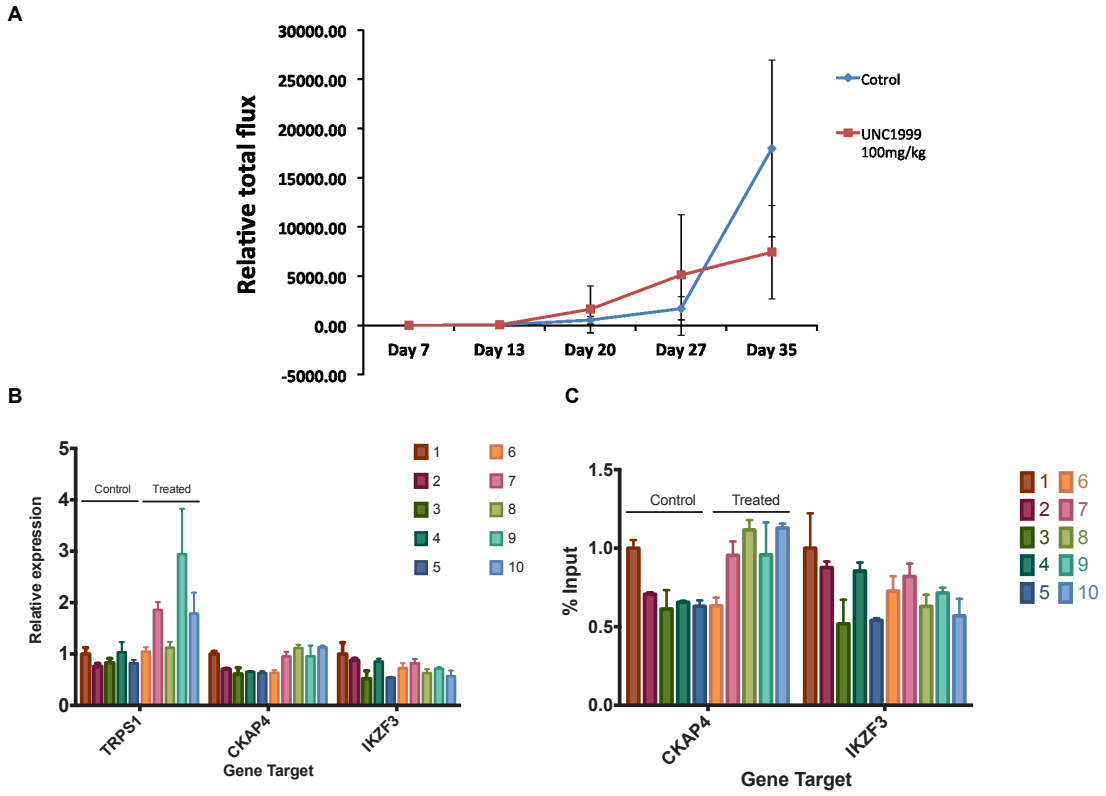


Figure 44 UNC1999 treatment does not confer a survival advantage in a xenograft model of t(4;11) leukaemia.

UNC1999 or vehicle was administered to NSG mice seven days post IF infections of SEM cells. Each group consisted of five mice. A) Graph showing engraftment level of SEM cells measured by bioluminescence. Points indicate the average luminsence of mice in each group. Error bars represent SD of signal between mice. B) rt-qPCR of spleen samples of vehicle treated (1-5) and UNC1999 treated (6-10) mice. C) H3K27me3 ChIP-qPCR of spleen samples of vehicle treated (1-5) and UNC1999 treated (6-10) mice. (n=1)

Since the maximum tolerated dose of UNC1999 in NSG mice was 100mg/kg the toxicity of the inhibitor was assessed in a different mouse strain, RAG KO. The same toxicity treatment plan was implemented as described for the NSG mice, but mice were dosed with 150mg/kg and 200mg/kg instead. While >10% weight loss was observed when mice were treated with 200mg/kg UNC1999 while 150mg/kg doses were well tolerated. Therefore, in future experiments

xenografts using this mouse model, which allows for higher doses of UNC1999 to be administered, may result in inhibition of PRC2 *in vivo.*

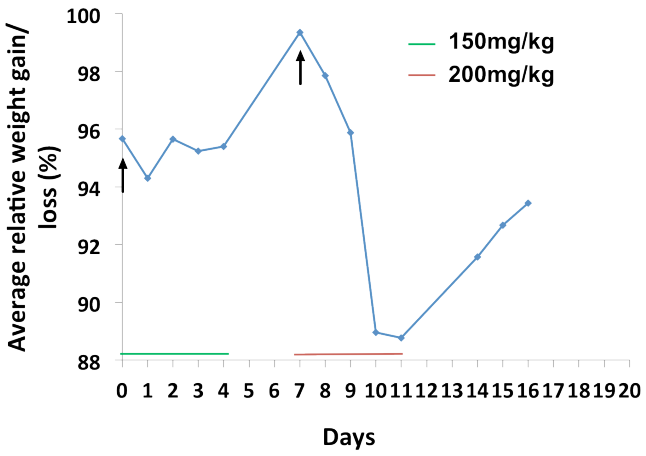


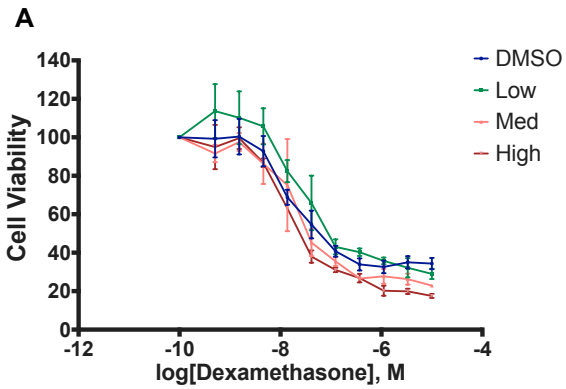
Figure 45 Toxicity of UNC1999 in RAG2 KO mice

UNC1999 was administered by oral gavage in increasing doses. One dose per day was administerd for 5 days followed by two days rest. On days 0-4 mice were dosed with 150mg/kg UNC1999, 7-11 with 200mg/kg UNC1999. Graph shows the average weight of 5 mice during the treatment period. Arrows indicate dosing at day 0 and 7.

5.2.1 UNC1999 sensitises SEM cells to dexamethasone

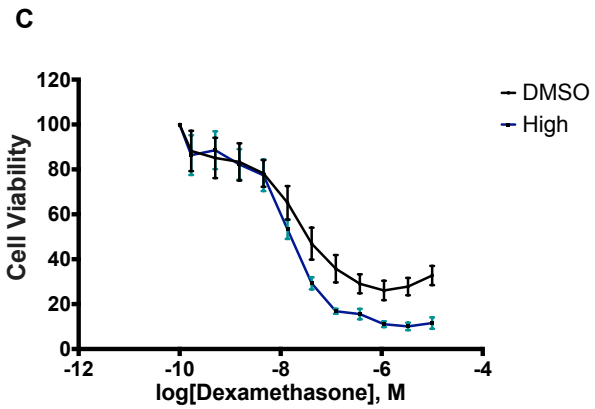
Synthetic glucocorticoids (GC), such as dexamethasone and prednisolone, induce apoptosis and lysis of lymphoblast's and are potent chemotherapeutic agents used standardly in therapy of childhood ALL (Bostrom et al., 2003; Distelhorst, 2002; Galili, 1983; Greenstein et al., 2002). Resistance to GC is a negative prognostic factor in ALL (Boer et al., 2003; Dordelmann et al., 1999; Kaspers et al., 1997) and is often seen at relapse in MLLr patients.

SEM cells have previously been shown to be resistant to dexamethasone (2010). To determine whether inhibition of PRC2 by the UNC1999 inhibitor would sensitise the cells to dexamethasone, the EC50 of dexamethasone of cells treated for two days with either a Low, Med or High dose of UNC1999 was compared to that of control DMSO treated cells. Cells treated with the Low dose showed a higher resistance to dexamethasone than did DMSO cells, but Med and High treated samples were roughly 2-fold and 4-fold more sensitive to dexamethasone respectively. To determine whether pre-treatment of SEM cells would further sensitise the cells to dexamethasone, SEM cells were pre-treated with a high dose of UNC1999. The EC50 of dexamethasone was measured after an additional three days (note that previously the EC50 was measured after 2 days). This resulted in a ~4 fold sensitisation to dexamethasone compared to DMSO treated cells. The data indicated that inhibition of UNC1999 sensitises cells to dexamethasone, which could potentially have significant clinical impacts.



B

Treatment Condition	EC50 ηM
DMSO	1 846
Low	2 926
Med	835
High	478



D

Treatment Condition	EC50 ηM
DMSO	720
High	166

Figure 46 UNC1999 sensitises SEM cells to dexamethasone

A) Cell viability of SEM cells treated with a combination of UNC1999 and dexamethasone for two days. Control cells were treated with DMSO and dexamethasone. B) EC50 values of dexamethasone for cells treated with the indicated doses of UNC1999 or DMSO (related to A). C) Cell viability of SEM cells pre-treated with a high dose of UNC1999 or DMSO for 3 days, and then treated with varying doses of dexamethasone for 3 days. B) EC50 values of dexamethasone for cells pre-treated with DMSO or a high dose of UNC1999. Low: 0.25μM UNC1999, Med: 1μM UNC1999, High: 5μM UNC1999. Error bars represent SD of three biological replicates.

5.3 Discussion

To get an indication of the possible effects of UNC1999 on normal human haematopoiesis, colony assays using cord blood and adult CD34+ cells were

performed. CD34+ cells from cord blood failed to form erythroid colonies when treated with a high dose of UNC1999, while adult CD34+ cells only showed a mild reduction (Figure 42). This is consistent with reports indicating that EZH2 is required for fetal but not adult erythropoiesis (Mochizuki-Kashio, 2012). In primary mouse erythroid cells, PRC2 repression of the *HES1* gene was shown to be important for erythroid differentiation (Ross et al., 2012). The mild defect observed in the formation of erythropoietic colonies from adult CD34+ cells following treatment with Med and High doses of UNC1999 (Figure 42) may in part be mediated by upregulation of this gene.

Overall, the effects of UNC1999 treatment on cord and adult CD34+ cells indicate that Low/Med doses of the inhibitor are compatible with haematopoiesis (Figure 42). However, the proportion of myeloid and multipotent progenitor cells were not analysed and therefore skewing of these lineages may be taking place. Importantly, given the established function of PRC2 in lymphoid development (Beguelin et al., 2013; Su et al., 2003) (O'Carroll et al., 2001), it is possible that the most pronounced effects of UNC1999 toxicity will be observed in this compartment. However, as shown for the DNA methyltransferase inhibitors used for treatment of MDS, it may be that at low doses, epigenetic inhibitors preferentially target transformed cells and affect normal cells to a lesser degree (Fenaux et al., 2009; Issa and Kantarjian, 2009; Silverman and Mufti, 2005). Consistent with this, Xu et al. (2014) did not detect any alterations in red blood cell counts in mice treated with UNC1999 compared to vehicle treated mice. Furthermore, the EPZ-6438 EZH2 inhibitor is in phase 2 clinical trials for the

treatment of B-cell lymphomas, indicating that inhibition of PRC2 may be a viable treatment strategy (Knutson et al., 2013).

Unfortunately, UNC1999 treatment of mice in a xenograft model of t(4;11) leukaemia failed to slow the progression of the disease (Figure 44). This is most likely a result of sub-optimal plasma concentrations of the inhibitor, as ChIP-qPCR failed to detect any decrease in H3K27me3 levels and upregulation of PRC2 targets was not observed. However analysis of more target genes would be beneficial in order to determine this with more certainty as well as pharmacokinetic studies. Previous work by Xu et al. (2014) has shown that the UNC1999 inhibitor decreases H3K27me3 levels in a xenograft model of MLL-AF9 AML. The discrepancy between the experiment carried out by Xu et al. (2014) and those performed here, in terms of the ability of UNC1999 to reduce H3K27me3 levels, can be explained by the treatment plans employed. Xu et al. treated mice twice per day with 50mg/kg (by oral gavage), and as such the plasma concentrations were likely higher than those achieved by a single 150mg/kg dose per day used for these experiments. Additionally, Xu et al. (2014) treated mice every day while in these experiments mice were treated for five days followed by two days rest. It is therefore likely that the lack of efficacy of the UNC1999 inhibitor may have been the result of sub-optimal plasma concentrations.

Furthermore, at the start of treatment mice had already substantial levels of engrafted cells that had disseminated through out the body (data not shown).

During the fist treatment cycle (1 cycle is five days on two days off) mice were only treated for three days with UNC1999 followed by two days rest. Therefore, it is also possible that a combination of "too many" engrafted cells with "too little" inhibitor may have contributed to the lack of efficacy of the compound. Given that the RAG KO mice tolerated higher doses of UNC1999 (Figure 45), in future experiments it would be beneficial to engraft fewer SEM cells, start treatment earlier and increase the dose of UNC1999 it order to reach plasma concentrations that are sufficient to decrease the levels of H3K27me3 *in vivo.* Of note is that it may be possible that t(4;11) patient derived cells become less sensitive to the UNC1999 inhibitor *in vivo*, and knowledge of the pharmacokinetic properties of UNC1999 would be beneficial in this respect.

Interestingly, treatment of SEM cells with Med and High doses of UNC1999 sensitised them to the synthetic GC dexamethasone (Figure 46). GCs are steroid hormones that bind to cytoplasmic glucocorticoid receptor (GR) causing it to relocate to the nucleus (reviewed in (Geley et al., 1996)). Within the nucleus, GR binds to promoters at specific glucocorticoid response elements (GRE) and regulates transcription of its target genes. The exact molecular mechanism by which GR cause apoptosis are not known and may be mediate by regulation of components of both the extrinsic and intrinsic apoptotic pathway (reviewed in (Schmidt et al., 2004)).

Gene expression studies have identified signatures of cells sensitive and resistant to GC treatment composed of a few hundred genes (Holleman et al., 2004;

Ramakers-van Woerden et al., 2004; Spijkers-Hagelstein et al., 2013; Wei et al., 2006). Based on these, bioactive compounds known to induce the "sensitive expression profile" (through use of the ConnectivityMap database; a database of genome wide expression profiles induced by bioactive molecules in multiple cancer cell lines) have been identified including the PI3K inhibitor LY494002 (Spijkers-Hagelstein et al., 2014) and the mTOR inhibitor rapamycin (Wei et al., 2006). The UNC1999 inhibitor results in upregulation (by at least 0.5log2 fold, fdr; 0.05) of 1 321 and 1 647 and downregulation of, 288, 740 genes in SEM cells treated with Med and High doses of UNC1999 respectively. Given that both gene expression and sensitisation to dexamethasone is dose-dependent (Figure 46) the genes mediating this sensitivity are likely all found as differentially expressed in the Med condition. Since "sensitivity profiles" have been identified for prednisolone (GC used in the treatment of infant leukaemia), comparison of the upregulated gene signature following UNC1999 treatment with the published prednisolone "sensitivity gene signature" could predict response to therapy for this GC.

Previously two microRNAs, miR-128b and miR-221, have been found to mediate sensitivity to GC treatment in the SEM cell line (Kotani et al., 2009; 2010). Both micro-RNAs have been found to be downregualed in MLLr ALL compared to other forms of ALL (Lu et al., 2005). It was therefore possible that these had become upregulated following UNC1999 treatment. Examination of the gene expression data generated following UNC1999 in SEM cells (cross-reference)

however did not show any upregulation of either micro-RNA. However, it would be worthwhile to verify this by rt-qPCR.

Combining treatment with a Low dose of UNC1999 with dexamethasone increased resistance to dexamethasone. The low treatment condition resulted in fewer gene expression changes (665 and 114 up and downregulated genes respectively) and may fit better with the resistance signature. However, as the anti-leukaemic effects of UNC1999 mediated inhibition of PRC2 is only apparent after a prolonged treatment with the Low dose, it is likely that pre-treatment with the Low dose for 10-15 days may result in sensitisation to dexamethasone. Verification of this is important, as increasing resistance to GC in clinical settings would have severe consequences in terms of prognosis and overall survival.

5.4 Conclusion

Work in this chapter has demonstrated low toxicity of the UNC1999 inhibitor on normal human haematopoiesis, although the effects of UNC1999 on the lymphoid compartment have not been assessed. Further work is required to determine an appropriate treatment and dosing plan that results in reduction of H3K27me3 *in vivo*. Additionally, UNC1999 resulted in a dose-dependent sensitisation to dexamethasone, which may serve to treat patients at relapse.

6 Concluding Remarks

The MLL-AF4 fusion protein is the most prevalent MLL rearrangement in ALL. Combination therapy for childhood ALL has improved over recent years with an approximate 80% event free survival rate (Pui et al. 2004). Despite this MLL-AF4 related infant ALL remains difficult to treat with 50% survival rates (Greaves 1996; Pieters et al. 2007), and there is therefore a need for novel treatment strategies.

Epigenetic inhibitors are promising agents for the treatment of haematological malignancies with FDA drugs approved for the treatment of MDS and T-cell lymphomas. However, while two DNMT inhibitors have been approved by the FDA for the treatment of MDS and AML, only 50% of patients show response to treatment. Predicting response to epigenetic therapy has been difficult, in part through the lack of sufficient knowledge regarding the molecular mechanism by which these inhibitors exert their anti-leukaemic effects.

In this study, the PRC2 core component EZH2 was identified as a direct target of MLL-AF4. Inhibition of the catalytic activity of PRC2 by the small molecule inhibitor UNC1999 led to a G1 arrest and was incompatible with leukaemic growth. Employment of the novel ChIP-Rx technique allowed for the detection of changes in the levels of H3K27me3 genome-wide, which were occluded in

previous studies employing the traditional ChIPseq approach. By combining changes in nascent-RNA transcripts with ChIPseq for H3K27me3, H3K4me3 and RING1B a signature of genes susceptible to upregulation was identified. Performing similar experiments in multiple cancer cell lines would be useful in determining how generalizable these findings are. Given that EZH2 inhibitors are in clinical trials, predicting genes that will respond to PRC2 inhibition could predict response to treatment. Additionally, since signatures of GR sensitive and resistant leukaemias have been identified, predicting sensitisation to GR by inhibitors of PRC2 could be achieved.

7 References

Agger, K., Cloos, P.A.C., Christensen, J., Pasini, D., Rose, S., Rappsilber, J., Issaeva, I., Canaani, E., Salcini, A.E., and Helin, K. (2007). UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. Nature *449*, 731–U10.

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

Allen, B.L., and Taatjes, D.J. (2015). The Mediator complex: a central integrator of transcription. Nat. Rev. Mol. Cell Biol. *16*, 155–166.

Allen, M.D., Grummitt, C.G., Hilcenko, C., Min, S.Y., Tonkin, L.M., Johnson, C.M., Freund, S.M., Bycroft, M., and Warren, A.J. (2006). Solution structure of the nonmethyl-CpG-binding CXXC domain of the leukaemia-associated MLL histone methyltransferase. Embo J. *25*, 4503–4512.

Allfrey, V.G., Faulkner, R., and Mirsky, A.E. (1964). Acetylation + Methylation of Histones + Their Possible Role in Regulation of Rna Synthesis. Proc. Natl. Acad. Sci. U.S.a. *51*, 786–.

Amin, R.H., and Schlissel, M.S. (2008). Foxo1 directly regulates the transcription of recombination-activating genes during B cell development. Nat. Immunol. *9*, 613–622.

Andersson, A.K., Ma, J., Wang, J., Chen, X., Rusch, M., Wu, G., Easton, J., Parker, M., Raimondi, S.C., Holmfeldt, L., et al. (2011). Whole Genome Sequence Analysis of 22 MLL Rearranged Infant Acute Lymphoblastic Leukemias Reveals Remarkably Few Somatic Mutations: A Report From the St Jude Children's Research Hospital - Washington University Pediatric Cancer Genome Project. Blood *118*, 33–34.

Armstrong, S.A., Mabon, M.E., Silverman, L.B., Li, A.H., Gribben, J.G., Fox, E.A., Sallan, S.E., and Korsmeyer, S.J. (2004). FLT3 mutations in childhood acute lymphoblastic leukemia. Blood *103*, 3544–3546.

Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jorgensen, H.F., John, R.M., Gouti, M., Casanova, M., Warnes, G., Merkenschlager, M., et al. (2006). Chromatin signatures of pluripotent cell lines. Nat. Cell Biol. *8*, 532–U189.

Ballare, C., Lange, M., Lapinaite, A., Martin, G.M., Morey, L., Pascual, G., Liefke, R., Simon, B., Shi, Y., Gozani, O., et al. (2012). Phf19 links methylated Lys36 of histone H3 to regulation of Polycomb activity. Nat. Struct. Mol. Biol. *19*, 1257–.

Barak, O., Lazzaro, M.A., Lane, W.S., Speicher, D.W., Picketts, D.J., and Shiekhattar,

R. (2003). Isolation of human NURF: a regulator of Engrailed gene expression. Embo J. *22*, 6089–6100.

Bardini, M., Galbiati, M., Lettieri, A., Bungaro, S., Gorletta, T.A., Biondi, A., and Cazzaniga, G. (2011). Implementation of array based whole-genome high-resolution technologies confirms the absence of secondary copy-number alterations in MLL-AF4-positive infant ALL patients. Leukemia *25*, 175–178.

Bardini, M., Spinelli, R., Bungaro, S., Mangano, E., Corral, L., Cifola, I., Fazio, G., Giordan, M., Basso, G., De Rossi, G., et al. (2010). DNA copy-number abnormalities do not occur in infant ALL with t(4;11)/MLL-AF4. Leukemia *24*, 169–176.

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

Beguelin, W., Popovic, R., Teater, M., Jiang, Y., Bunting, K.L., Rosen, M., Shen, H., Yang, S.N., Wang, L., Ezponda, T., et al. (2013). EZH2 Is Required for Germinal Center Formation and Somatic EZH2 Mutations Promote Lymphoid Transformation. Cancer Cell *23*, 677–692.

Berg, T., Thoene, S., Yap, D., Wee, T., Schoeler, N., Rosten, P., Lim, E., Bilenky, M., Mungall, A.J., Oellerich, T., et al. (2014). A transgenic mouse model demonstrating the oncogenic role of mutations in the polycomb-group gene EZH2 in lymphomagenesis. Blood *123*, 3914–3924.

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

Bernt, K.M., Zhu, N., Sinha, A.U., Vempati, S., Faber, J., Krivtsov, A.V., Feng, Z., Punt, N., Daigle, A., Bullinger, L., et al. (2011). MLL-Rearranged Leukemia Is Dependent on Aberrant H3K79 Methylation by DOT1L. Cancer Cell *20*, 66–78.

Bird, A., Taggart, M., Frommer, M., Miller, O.J., and Macleod, D. (1985). A Fraction of the Mouse Genome That Is Derived From Islands of Nonmethylated, Cpg-Rich Dna. Cell *40*, 91–99.

Bird, A. (2007). Perceptions of epigenetics. Nature *447*, 396–398.

Birke, M., Schreiner, S., Garcia-Cuellar, M.P., Mahr, K., Titgemeyer, F., and Slany, R.K. (2002). The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation. Nucleic Acids Res. *30*, 958–965.

Biswas, D., Milne, T.A., Basrur, V., Kim, J., Elenitoba-Johnson, K.S.J., Allis, C.D., and Roeder, R.G. (2011). Function of leukemogenic mixed lineage leukemia 1 (MLL) fusion proteins through distinct partner protein complexes. Proc. Natl. Acad. Sci. U.S.a. *108*, 15751–15756.

Bitoun, E., Oliver, P.L., and Davies, K.E. (2007). The mixed-lineage leukemia fusion partner AF4 stimulates RNA polymerase II transcriptional elongation and mediates coordinated chromatin remodeling. Hum. Mol. Genet. *16*, 92–106.

Blackledge, N.P., and Klose, R.J. (2011). CpG island chromatin A platform for gene regulation. Epigenetics *6*, 147–152.

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

Blackledge, N.P., Long, H.K., Zhou, J.C., Kriaucionis, S., Patient, R., and Klose, R.J. (2012). Bio-CAP: a versatile and highly sensitive technique to purify and characterise regions of non-methylated DNA. Nucleic Acids Res. *40*, –e32.

Boer, den, M.L., Harms, D.O., Pieters, R., Kazemier, K.M., Gobel, U., Korholz, D., Graubner, U., Haas, R.J., Jorch, N., Spaar, H.J., et al. (2003). Patient stratification based on prednisolone-vincristine-asparaginase resistance profiles in children with acute lymphoblastic leukemia. J. Clin. Oncol. *21*, 3262–3268.

Bomken, S., Buechler, L., Rehe, K., Ponthan, F., Elder, A., Blair, H., Bacon, C.M., Vormoor, J., and Heidenreich, O. (2013). Lentiviral marking of patient-derived acute lymphoblastic leukaemic cells allows in vivo tracking of disease progression (vol 27, pg 718, 2013). Leukemia *27*, 1792–1792.

Bostrom, B.C., Sensel, M.R., Sather, H.N., Gaynon, P.S., La, M.K., Johnston, K., Erdmann, G.R., Gold, S., Heerema, N.A., Hutchinson, R.J., et al. (2003). Dexamethasone versus prednisone and daily oral versus weekly intravenous mercaptopurine for patients with standard-risk acute lymphoblastic leukemia: a report from the Children's Cancer Group. Blood *101*, 3809–3817.

Boyer, L.A., Latek, R.R., and Peterson, C.L. (2004). The SANT domain: a unique histone-tail-binding module? Nat. Rev. Mol. Cell Biol. *5*, 158–163.

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

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

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

Brien, G.L., Gambero, G., O'Connell, D.J., Jerman, E., Turner, S.A., Egan, C.M., Dunne, E.J., Jurgens, M.C., Wynne, K., Piao, 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. *19*, 1273–.

Briggs, S.D., Bryk, M., Strahl, B.D., Cheung, W.L., Davie, J.K., Dent, S., Winston, F., and Allis, C.D. (2001). Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in Saccharomyces cerevisiae. Genes Dev. *15*, 3286–3295.

Brinkman, A.B., Gu, H., Bartels, S.J.J., Zhang, Y., Matarese, F., Simmer, F., Marks, H., Bock, C., Gnirke, A., Meissner, A., et al. (2012). Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. Genome Res. *22*, 1128–1138.

Brockdorff, N., ASHWORTH, A., Kay, G.F., MCCABE, V.M., NORRIS, D.P., COOPER, P.J., SWIFT, S., and RASTAN, S. (1992). The Product of the Mouse Xist Gene Is a 15 Kb Inactive X-Specific Transcript Containing No Conserved Orf and Located in the Nucleus. Cell *71*, 515–526.

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

Brown, C.J., Hendrich, B.D., Rupert, J.L., Lafreniere, R.G., Xing, Y., Lawrence, J., and Willard, H.F. (1992). The Human Xist Gene - Analysis of a 17 Kb Inactive X-Specific Rna That Contains Conserved Repeats and Is Highly Localized Within the Nucleus. Cell *71*, 527–542.

Bueno, C., Montes, R., Melen, G.J., Ramos-Mejia, V., Real, P.J., Ayllón, V., Sanchez, L., Ligero, G., Gutierrez-Aranda, I., Fernández, A.F., et al. (2012). A human ESC model for MLL-AF4 leukemic fusion gene reveals an impaired early hematopoietic-endothelial specification. Cell Res. *22*, 986–1002.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat. Methods *10*, 1213–.

Bursen, A., Schwabe, K., Ruester, B., Henschler, R., Ruthardt, M., Dingermann, T., and Marschalek, R. (2010). The AF4.MLL fusion protein is capable of inducing ALL in mice without requirement of MLL.AF4. Blood *115*, 3570–3579.

Cai, L., Rothbart, S.B., Lu, R., Xu, B., Chen, W.-Y., Tripathy, A., Rockowitz, S., Zheng, D., Patel, D.J., Allis, C.D., et al. (2013). An H3K36 Methylation-Engaging Tudor Motif of Polycomb-like Proteins Mediates PRC2 Complex Targeting. Mol. Cell *49*, 571–582.

Campbell, J.E., Kuntz, K.W., Knutson, S.K., Warholic, N.M., Keilhack, H., Wigle, T.J., Raimondi, A., Klaus, C.R., Rioux, N., Yokoi, A., et al. (2015). EPZ011989, A Potent,

Orally-Available EZH2 Inhibitor with Robust in Vivo Activity. Acs Medicinal Chemistry Letters *6*, 491–495.

Cao, F., Townsend, E.C., Karatas, H., Xu, J., Li, L., Lee, S., Liu, L., Chen, Y., Ouillette, P., Zhu, J., et al. (2014a). Targeting MLL1 H3K4 methyltransferase activity in mixed-lineage leukemia. Mol. Cell *53*, 247–261.

Cao, Q., Wang, X., Zhao, M., Yang, R., Malik, R., Qiao, Y., Poliakov, A., Yocum, A.K., Li, Y., Chen, W., et al. (2014b). The central role of EED in the orchestration of polycomb group complexes. Nat Commun *5*, 3127.

Cao, R., Tsukada, Y., and Zhang, Y. (2005). Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. Mol. Cell *20*, 845–854.

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

Cao, R., and Zhang, Y. (2004). SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. Mol. Cell *15*, 57–67.

Caretti, G., Di Padova, M., Micales, B., Lyons, G.E., and Sartorelli, V. (2005). The Polycomb Ezh2 methyltransferase regulates muscle gene expression and skeletal muscle differentiation (vol 18, pg 2627, 2004). Genes Dev. *19*, 768–768.

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

Caslini, C., Yang, Z., El-Osta, M., Milne, T.A., Slany, R.K., and Hess, J.L. (2007). Interaction of MLL amino terminal sequences with menin is required for transformation. Cancer Res. *67*, 7275–7283.

Cerase, A., Smeets, D., Tang, Y.A., Gdula, M., Kraus, F., Spivakov, M., Moindrot, B., Leleu, M., Tattermusch, A., Demmerle, J., et al. (2014). Spatial separation of Xist RNA and polycomb proteins revealed by superresolution microscopy. Proc. Natl. Acad. Sci. U.S.a. *111*, 2235–2240.

Cermakova, K., Tesina, P., Demeulemeester, J., Ashkar, El, S., Mereau, H., Schwaller, J., Rezacova, P., Veverka, V., and De Rijck, J. (2014). Validation and Structural Characterization of the LEDGF/p75-MLL Interface as a New Target for the Treatment of MLL-Dependent Leukemia. Cancer Res. *74*, 5139–5151.

Chamberlain, S.J., Yee, D., and Magnuson, T. (2008). Polycomb repressive complex 2 is dispensable for maintenance of embryonic stem cell pluripotency. Stem Cells *26*, 1496–1505.

Chen, C.-C., Jeon, S.-M., Bhaskar, P.T., Nogueira, V., Sundararajan, D., Tonic, I., Park, Y., and Hay, N. (2010). FoxOs inhibit mTORC1 and activate Akt by inducing the expression of Sestrin3 and Rictor. Dev. Cell *18*, 592–604.

Chen, J., Santillan, D.A., Koonce, M., Wei, W., Luo, R., Thirman, M.J., Zeleznik-Le, N.J., and Diaz, M.O. (2008). Loss of MLL PHD finger 3 is necessary for MLL-ENL-induced hematopoietic stem cell immortalization. Cancer Res. *68*, 6199–6207.

Chen, W., Li, Q., Hudson, W.A., Kumar, A., Kirchhof, N., and Kersey, J.H. (2006a). A murine Mll-AF4 knock-in model results in lymphoid and myeloid deregulation and hematologic malignancy. Blood *108*, 669–677.

Chen, Y.X., Yan, J.Z., Keeshan, K., Tubbs, A.T., Wang, H.R., Silva, A., Brown, E.J., Hess, J.L., Pear, W.S., and Hua, X.X. (2006b). The tumor suppressor menin regulates hematopoiesis and myeloid transformation by influencing Hox gene expression. Proc. Natl. Acad. Sci. U.S.a. *103*, 1018–1023.

Cho, Y.-W., Hong, T., Hong, S., Guo, H., Yu, H., Kim, D., Guszczynski, T., Dressler, G.R., Copeland, T.D., Kalkum, M., et al. (2007). PTIP associates with MLL3-and MLL4-containing histone H3 lysine 4 methyltransferase complex. J. Biol. Chem. *282*, 20395–20406.

Chopra, V.S., Hendrix, D.A., Core, L.J., Tsui, C., Lis, J.T., and Levine, M. (2011). The Polycomb Group Mutant esc Leads to Augmented Levels of Paused Pol II in the Drosophila Embryo. Mol. Cell *42*, 837–844.

Chopra, V.S., Hong, J.-W., and Levine, M. (2009). Regulation of Hox Gene Activity by Transcriptional Elongation in Drosophila. Curr. Biol. *19*, 688–693.

Cierpicki, T., Risner, L.E., Grembecka, J., Lukasik, S.M., Popovic, R., Omonkowska, M., Shultis, D.D., Zeleznik-Le, N.J., and Bushweller, J.H. (2010). Structure of the MLL CXXC domain-DNA complex and its functional role in MLL-AF9 leukemia. Nat. Struct. Mol. Biol. *17*, 62–U82.

Ciferri, C., Lander, G.C., Maiolica, A., Herzog, F., Aebersold, R., and Nogales, E. (2012). Molecular architecture of human polycomb repressive complex 2. Elife *1*.

Collins, E.C., Pannell, R., Simpson, E.M., Forster, A., and Rabbitts, T.H. (2000). Inter-chromosomal recombination of Mll and Af9 genes mediated by cre-loxP in mouse development. EMBO Rep. *1*, 127–132.

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

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental

state. Proc. Natl. Acad. Sci. U.S.a. *107*, 21931–21936.

Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A., and Pirrotta, V. (2002). Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal polycomb sites. Cell *111*, 185–196.

Daigle, S.R., Olhava, E.J., Therkelsen, C.A., Basavapathruni, A., Jin, L., Boriack-Sjodin, P.A., Allain, C.J., Klaus, C.R., Raimondi, A., Scott, M.P., et al. (2013). Potent inhibition of DOT1L as treatment of MLL-fusion leukemia. Blood *122*, 1017–1025.

Daigle, S.R., Olhava, E.J., Therkelsen, C.A., Majer, C.R., Sneeringer, C.J., Song, J., Johnston, L.D., Scott, M.P., Smith, J.J., Xiao, Y., et al. (2011). Selective Killing of Mixed Lineage Leukemia Cells by a Potent Small-Molecule DOT1L Inhibitor. Cancer Cell *20*, 53–65.

Dawson, M.A., Prinjha, R.K., Dittmann, A., Giotopoulos, G., Bantscheff, M., Chan, W.-I., Robson, S.C., Chung, C.-W., Hopf, C., Savitski, M.M., et al. (2011). Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. Nature *478*, 529–533.

de Napoles, M., Mermoud, J.E., Wakao, R., Tang, Y.A., Endoh, M., Appanah, R., Nesterova, T.B., Silva, J., Otte, A.P., Vidal, M., et al. (2004). Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. Dev. Cell *7*, 663–676.

De Santa, F., Narang, V., Yap, Z.H., Tusi, B.K., Burgold, T., Austenaa, L., Bucci, G., Caganova, M., Notarbartolo, S., Casola, S., et al. (2009). Jmjd3 contributes to the control of gene expression in LPS-activated macrophages. Embo J. *28*, 3341–3352.

De Santa, F., Totaro, M.G., Prosperini, E., Notarbartolo, S., Testa, G., and Natoli, G. (2007). The histone H3 lysine-27 demethylase Jmjd3 links inflammation to inhibition of polycomb-mediated gene silencing. Cell *130*, 1083–1094.

Delattre, O., Zucman, J., Plougastel, B., Desmaze, C., Melot, T., Peter, M., Kovar, H., Joubert, I., Dejong, P., Rouleau, G., et al. (1992). Gene Fusion with an Ets Dna-Binding Domain Caused by Chromosome-Translocation in Human Tumors. Nature *359*, 162–165.

Dengler, H.S., Baracho, G.V., Omori, S.A., Bruckner, S., Arden, K.C., Castrillon, D.H., DePinho, R.A., and Rickert, R.C. (2008). Distinct functions for the transcription factor Foxo1 at various stages of B cell differentiation. Nat. Immunol. *9*, 1388–1398.

Denissov, S., Hofemeister, H., Marks, H., Kranz, A., Ciotta, G., Singh, S., Anastassiadis, K., Stunnenberg, H.G., and Stewart, A.F. (2014). Mll2 is required for H3K4 trimethylation on bivalent promoters in embryonic stem cells, whereas

Mll1 is redundant. Development *141*, 526–537.

Dietrich, N., Lerdrup, M., Landt, E., and Agrawal-Singh, S. (2012). REST–mediated recruitment of Polycomb repressor complexes in mammalian cells. PLoS Genet.

Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983). Accurate Transcription Initiation by Rna Polymerase-Ii in a Soluble Extract From Isolated Mammalian Nuclei. Nucleic Acids Res. *11*, 1475–1489.

Dion, M.F., Altschuler, S.J., Wu, L.F., and Rando, O.J. (2005). Genomic characterization reveals a simple histone H4 acetylation code. Proc. Natl. Acad. Sci. U.S.a. *102*, 5501–5506.

Distelhorst, C.W. (2002). Recent insights into the mechanism of glucocorticosteroid-induced apoptosis. Cell Death Differ. *9*, 6–19.

Djabali, M., Selleri, L., Parry, P., Bower, M., Young, B.D., and Evans, G.A. (1992). A Trithorax-Like Gene Is Interrupted by Chromosome 11q23 Translocations in Acute Leukemias. Nat. Genet. *2*, 113–118.

Dobbins, S.E., Sherborne, A.L., Ma, Y.P., Bardini, M., Biondi, A., Cazzaniga, G., Lloyd, A., Chubb, D., Greaves, M.F., and Houlston, R.S. (2013). The silent mutational landscape of infant MLL-AF4 pro-B acute lymphoblastic leukemia. Genes Chromosomes Cancer *52*, 954–960.

Dordelmann, M., Reiter, A., Borkhardt, A., Ludwig, W.D., Gotz, N., Viehmann, S., Gadner, H., Riehm, H., Schrappe, M., and Grp, A.-B. (1999). Prednisone response is the strongest predictor of treatment outcome in infant acute lymphoblastic leukemia. Blood *94*, 1209–1217.

Dou, Y., Milne, T.A., Ruthenburg, A.J., Lee, S., Lee, J.W., Verdine, G.L., Allis, C.D., and Roeder, R.G. (2006). Regulation of MLL1 H3K4 methyltransferase activity by its core components. Nat. Struct. Mol. Biol. *13*, 713–719.

Dou, Y., Milne, T.A., Tackett, A.J., Smith, E.R., Fukuda, A., Wysocka, J., Allis, C.D., Chait, B.T., Hess, J.L., and Roeder, R.G. (2005). Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. Cell *121*, 873–885.

Driessen, E.M.C., van Roon, E.H.J., Spijkers-Hagelstein, J.A.P., Schneider, P., De Lorenzo, P., Valsecchi, M.G., Pieters, R., and Stam, R.W. (2013). Frequencies and prognostic impact of RAS mutations in MLL-rearranged acute lymphoblastic leukemia in infants. Haematologica *98*, 937–944.

Drynan, L.F., Pannell, R., Forster, A., Chan, N., Cano, F., Daser, A., and Rabbitts, T.H. (2005). Mll fusions generated by Cre-loxP-mediated de novo translocations can induce lineage reassignment in tumorigenesis. Embo J. *24*, 3136–3146.

Ehrlich, M., Gamasosa, M.A., Huang, L.H., Midgett, R.M., Kuo, K.C., Mccune, R.A.,

and Gehrke, C. (1982). Amount and Distribution of 5-Methylcytosine in Human Dna From Different Types of Tissues or Cells. Nucleic Acids Res. *10*, 2709–2721.

Elderkin, S., Maertens, G.N., Endoh, M., Mallery, D.L., Morrice, N., Koseki, H., Peters, G., Brockdorff, N., and Hiom, K. (2007). A phosphorylated form of mel-18 targets the Ring1B histone H2A ubliquitin ligase to chromatin. Mol. Cell *28*, 107–120.

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

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

Ernst, P., Wang, J., Huang, M., Goodman, R.H., and Korsmeyer, S.J. (2001). MLL and CREB bind cooperatively to the nuclear coactivator CREB-binding protein. Mol. Cell. Biol. *21*, 2249–2258.

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

Ezhkova, E., Pasolli, H.A., Parker, J.S., Stokes, N., Su, I.-H., Hannon, G., Tarakhovsky, A., and Fuchs, E. (2009). Ezh2 Orchestrates Gene Expression for the Stepwise Differentiation of Tissue-Specific Stem Cells. Cell *136*, 1122–1135.

Faber, J., Krivtsov, A.V., Stubbs, M.C., Wright, R., Davis, T.N., van den Heuvel-Eibrink, M., Zwaan, C.M., Kung, A.L., and Armstrong, S.A. (2009). HOXA9 is required for survival in human MLL-rearranged acute leukemias. Blood *113*, 2375–2385.

Fair, K., Anderson, M., Bulanova, E., Mi, H.F., Tropschug, M., and Diaz, M.O. (2001). Protein interactions of the MLL PHD fingers modulate MLL target gene regulation in human cells. Mol. Cell. Biol. *21*, 3589–3597.

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

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

Fenaux, P., Mufti, G.J., Hellstrom-Lindberg, E., Santini, V., Finelli, C., Giagounidis, A., Schoch, R., Gattermann, N., Sanz, G., List, A., et al. (2009). Efficacy of azacitidine

compared with that of conventional care regimens in the treatment of higher-risk myelodysplastic syndromes: a randomised, open-label, phase III study. Lancet Oncol. *10*, 223–232.

Fenouil, R., Cauchy, P., Koch, F., Descostes, N., Cabeza, J.Z., Innocenti, C., Ferrier, P., Spicuglia, S., Gut, M., Gut, I., et al. (2012). CpG islands and GC content dictate nucleosome depletion in a transcription-independent manner at mammalian promoters. Genome Res. *22*, 2399–2408.

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

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

Forster, A., Pannell, R., Drynan, L.F., McCormack, M., Collins, E.C., Daser, A., and Rabbitts, T.H. (2003). Engineering de novo reciprocal chromosomal translocations associated with Mll to replicate primary events of human cancer. Cancer Cell *3*, 449–458.

Francis, N.J., Kingston, R.E., and Woodcock, C.L. (2004). Chromatin compaction by a polycomb group protein complex. Science *306*, 1574–1577.

Frankel, L.S., Ochs, J., Shuster, J.J., Dubowy, R., Bowman, W.P., HockenberryEaton, M., Borowitz, M., Carroll, A.J., Steuber, C.P., and Pullen, D.J. (1997). Therapeutic trial for infant acute lymphoblastic leukemia: The pediatric oncology group experience (POG 8493). J. Pediatr. Hematol. Oncol. *19*, 35–42.

Galili, U. (1983). Glucocorticoid Induced Cytolysis of Human Normal and Malignant Lymphocytes. J. Steroid Biochem. *19*, 483–490.

Gao, Z., Zhang, J., Bonasio, R., Strino, F., Sawai, A., Parisi, F., Kluger, Y., and Reinberg, D. (2012). PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes. Mol. Cell *45*, 344–356.

Garcia-Cuellar, M.P., Schreiner, S.A., Birke, M., Hamacher, M., Fey, G.H., and Slany, R.K. (2000). ENL, the MLL fusion partner in t(11;19), binds to the c-Abl interactor protein 1 (ABI1) that is fused to MLL in t(10;11)(+). Oncogene *19*, 1744–1751.

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

Geley, S., Fiegl, M., Hartmann, B.L., and Kofler, R. (1996). Genes mediating glucocorticoid effects and mechanisms of their regulation. Rev. Physiol. Biochem. Pharmacol. *128*, 1–97.

Greaves, M. (2005). In utero origins of childhood leukaemia. Early Hum. Dev. *81*, 123–129.

Greaves, M.F. (1996). Infant leukaemia biology, aetiology and treatment. Leukemia *10*, 372–377.

Greenstein, S., Ghias, K., Krett, N.L., and Rosen, S.T. (2002). Mechanisms of glucocorticoid-mediated apoptosis in hematological malignancies. Clin. Cancer Res. *8*, 1681–1694.

Grembecka, J., He, S., Shi, A., Purohit, T., Muntean, A.G., Sorenson, R.J., Showalter, H.D., Murai, M.J., Belcher, A.M., Hartley, T., et al. (2012). Menin-MLL inhibitors reverse oncogenic activity of MLL fusion proteins in leukemia. Nat. Chem. Biol. *8*, 277–284.

Gu, Y., Nakamura, T., Alder, H., Prasad, R., Canaani, O., Cimino, G., Croce, C.M., and Canaani, E. (1992). The t(4;11) chromosome translocation of human acute leukemias fuses the ALL-1 gene, related to Drosophila trithorax, to the AF-4 gene. Cell *71*, 701–708.

Guenther, M.G., Lawton, L.N., Rozovskaia, T., Frampton, G.M., Levine, S.S., Volkert, T.L., Croce, C.M., Nakamura, T., Canaani, E., and Young, R.A. (2008). Aberrant chromatin at genes encoding stem cell regulators in human mixed-lineage leukemia. Genes Dev. *22*, 3403–3408.

Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. Cell *130*, 77–88.

Hagarman, J.A., Motley, M.P., Kristjansdottir, K., and Soloway, P.D. (2013). Coordinate Regulation of DNA Methylation and H3K27me3 in Mouse Embryonic Stem Cells. PLoS ONE *8*.

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

Hansen, J.C. (2012). Human mitotic chromosome structure: what happened to the 30-nm fibre? Embo J. *31*, 1621–1623.

Hansen, K.H., Bracken, A.P., Pasini, D., Dietrich, N., Gehani, S.S., Monrad, A., Rappsilber, J., Lerdrup, M., and Helin, K. (2008). A model for transmission of the H3K27me3 epigenetic mark. Nat. Cell Biol. *10*, 1291–U89.

Hanson, R.D., Hess, J.L., Yu, B.D., Ernst, P., van Lohuizen, M., BERNS, A., van der Lugt, N.M., Shashikant, C.S., Ruddle, F.H., Seto, M., et al. (1999). Mammalian Trithorax and polycomb-group homologues are antagonistic regulators of homeotic development. Proc. Natl. Acad. Sci. U.S.a. *96*, 14372–14377.

Hawkins, R.D., Hon, G.C., Lee, L.K., Ngo, Q., Lister, R., Pelizzola, M., Edsall, L.E.,

Kuan, S., Luu, Y., Klugman, S., et al. (2010). Distinct Epigenomic Landscapes of Pluripotent and Lineage-Committed Human Cells. Cell Stem Cell *6*, 479–491.

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

Hebbes, T.R., Thorne, A.W., and Cranerobinson, C. (1988). A Direct Link Between Core Histone Acetylation and Transcriptionally Active Chromatin. Embo J. *7*, 1395–1402.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. Mol. Cell *38*, 576–589.

Helin, K., and Dhanak, D. (2013). Chromatin proteins and modifications as drug targets. Nature *502*, 480–488.

Herrera-Merchan, A., Arranz, L., Ligos, J.M., de Molina, A., Dominguez, O., and Gonzalez, S. (2012). Ectopic expression of the histone methyltransferase Ezh2 in haematopoietic stem cells causes myeloproliferative disease. Nat Commun *3*.

Herrmann, H., Blatt, K., Shi, J., Gleixner, K.V., Cerny-Reiterer, S., Muellauer, L., Vakoc, C.R., Sperr, W.R., Horny, H.-P., Bradner, J.E., et al. (2012). Small-molecule inhibition of BRD4 as a new potent approach to eliminate leukemic stem-and progenitor cells in acute myeloid leukemia (AML). Oncotarget *3*, 1588–1599.

Herzing, L., Romer, J.T., Horn, J.M., and ASHWORTH, A. (1997). Xist has properties of the X-chromosome inactivation centre. Nature *386*, 272–275.

Hess, J.L., Bittner, C.B., Zeisig, D.T., Bach, C., Fuchs, U., Borkhardt, A., Frampton, J., and Slany, R.K. (2006). c-Myb is an essential downstream target for homeobox-mediated transformation of hematopoietic cells. Blood *108*, 297–304.

Hidalgo, I., Herrera-Merchan, A., Manuel Ligos, J., Carramolino, L., Nunez, J., Martinez, F., Dominguez, O., Torres, M., and Gonzalez, S. (2012). Ezh1 Is Required for Hematopoietic Stem Cell Maintenance and Prevents Senescence-like Cell Cycle Arrest. Cell Stem Cell *11*, 649–662.

Holleman, A., Cheok, M.H., Boer, den, M.L., Yang, W.J., Veerman, A., Kazemier, K.M., Pei, D.Q., Cheng, C., Pui, C.H., Relling, M.V., et al. (2004). Gene-expression patterns in drug-resistant acute lymphoblastic leukemia cells and response to treatment. N. Engl. J. Med. *351*, 533–542.

Hom, R.A., Chang, P.-Y., Roy, S., Musselman, C.A., Glass, K.C., Selezneva, A.I., Gozani, O., Ismagilov, R.F., Cleary, M.L., and Kutateladze, T.G. (2010). Molecular Mechanism of MLL PHD3 and RNA Recognition by the Cyp33 RRM Domain. J. Mol. Biol. *400*, 145–154.

Hong, L., Schroth, G.P., Matthews, H.R., Yau, P., and Bradbury, E.M. (1993). Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA. J. Biol. Chem. *268*, 305–314.

Hsieh, J., Ernst, P., Erdjument-Bromage, H., Tempst, P., and Korsmeyer, S.J. (2003). Proteolytic cleavage of MLL generates a complex of N- and C-terminal fragments that confers protein stability and subnuclear localization. Mol. Cell. Biol. *23*, 186–194.

Hu, D., Garruss, A.S., Gao, X., Morgan, M.A., Cook, M., Smith, E.R., and Shilatifard, A. (2013). The Mll2 branch of the COMPASS family regulates bivalent promoters in mouse embryonic stem cells. Nat. Struct. Mol. Biol. *20*, 1093–1097.

Hughes, C.M., Rozenblatt-Rosen, O., Milne, T.A., Copeland, T.D., Levine, S.S., Lee, J.C., Hayes, D.N., Shanmugam, K.S., Bhattacharjee, A., Biondi, C.A., et al. (2004). Menin associates with a trithorax family histone methyltransferase complex and with the Hoxc8 locus. Mol. Cell *13*, 587–597.

Hunkapiller, J., Shen, Y., Diaz, A., Cagney, G., McCleary, D., Ramalho-Santos, M., Krogan, N., Ren, B., Song, J.S., and Reiter, J.F. (2012). Polycomb-Like 3 Promotes Polycomb Repressive Complex 2 Binding to CpG Islands and Embryonic Stem Cell Self-Renewal. PLoS Genet. *8*.

Ingham, P.W. (1983). Differential expression of bithorax complex genes in the absence of the extra sex combs and trithorax genes. Nature *306*, 591–593.

Ingham, P.W. (1985). A Clonal Analysis of the Requirement for the Trithorax Gene in the Diversification of Segments in Drosophila. J Embryol Exp Morphol *89*, 349–365.

Ingham, P.W. (1998). trithorax and the regulation of homeotic gene expression in Drosophila: a historical perspective. Int. J. Dev. Biol. *42*, 423–429.

Isono, K., Endo, T.A., Ku, M., Yamada, D., Suzuki, R., Sharif, J., Ishikura, T., Toyoda, T., Bernstein, B.E., and Koseki, H. (2013). SAM domain polymerization links subnuclear clustering of PRC1 to gene silencing. Dev. Cell *26*, 565–577.

Issa, J.-P.J., and Kantarjian, H.M. (2009). Targeting DNA Methylation. Clin. Cancer Res. *15*, 3938–3946.

Issaeva, I., Zonis, Y., Rozovskaia, T., Orlovsky, K., Croce, C.M., Nakamura, T., Mazo, A., Eisenbach, L., and Canaani, E. (2007). Knockdown of ALR (MLL2) reveals ALR target genes and leads to alterations in cell adhesion and growth. Mol. Cell. Biol. *27*, 1889–1903.

Jang, M.K., Mochizuki, K., Zhou, M.S., Jeong, H.S., Brady, J.N., and Ozato, K. (2005). The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. Mol. Cell *19*, 523–

534.

Jo, S.Y., Granowicz, E.M., Maillard, I., Thomas, D., and Hess, J.L. (2011). Requirement for Dot1l in murine postnatal hematopoiesis and leukemogenesis by MLL translocation. Blood *117*, 4759–4768.

Jones, B., Su, H., Bhat, A., Lei, H., Bajko, J., Hevi, S., Baltus, G.A., Kadam, S., Zhai, H., Valdez, R., et al. (2008). The histone H3K79 methyltransferase Dot1L is essential for mammalian development and heterochromatin structure. PLoS Genet. *4*, e1000190.

Joshi, P., Carrington, E.A., Wang, L., Ketel, C.S., Miller, E.L., Jones, R.S., and Simon, J.A. (2008). Dominant alleles identify SET domain residues required for histone methyltransferase of polycomb repressive complex 2. J. Biol. Chem. *283*, 27757–27766.

Jude, C.D., Climer, L., Xu, D., Artinger, E., Fisher, J.K., and Ernst, P. (2007). Unique and independent roles for MLL in adult hematopoietic stem cells and progenitors. Cell Stem Cell *1*, 324–337.

Kalb, R., Latwiel, S., Baymaz, H.I., Jansen, P.W.T.C., Müller, C.W., Vermeulen, M., and Müller, J. (2014). Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression. Nat. Struct. Mol. Biol. *21*, 569–571.

Kamminga, L.M., Bystrykh, L.V., Boer, A.C., Houwer, S., Douma, J., Weersing, E., Dontje, B., and de Haan, G. (2006). The polycomb group gene Ezh2 prevents hematopoietic stem cell exhaustion. Blood *107*, 2170–2179.

Kaspers, G., Veerman, A., Pieters, R., VanZantwijk, C.H., Smets, L.A., VanWering, E.R., and DenBerg, A.V. (1997). In vitro cellular drug resistance and prognosis in newly diagnosed childhood acute lymphoblastic leukemia. Blood *90*, 2723–2729.

Kennison, J.A., and Tamkun, J.W. (1988). Dosage-Dependent Modifiers of Polycomb and Antennapedia Mutations in Drosophila. Proc. Natl. Acad. Sci. U.S.a. *85*, 8136–8140.

Ketel, C.S., Andersen, E.F., Vargas, M.L., Su, J., Strome, S., and Simon, J.A. (2005). Subunit contributions to histone methyltransferase, activities of fly and worm Polycomb group complexes. Mol. Cell. Biol. *25*, 6857–6868.

Khan, O., and La Thangue, N.B. (2012). HDAC inhibitors in cancer biology: emerging mechanisms and clinical applications. Immunol. Cell Biol. *90*, 85–94.

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

Klose, R.J., Cooper, S., Farcas, A.M., Blackledge, N.P., and Brockdorff, N. (2013).

Chromatin Sampling-An Emerging Perspective on Targeting Polycomb Repressor Proteins. PLoS Genet. *9*.

Knezetic, J.A., and Luse, D.S. (1986). The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II in vitro. Cell *45*, 95–104.

Knutson, S.K., Warholic, N.M., Wigle, T.J., Klaus, C.R., Allain, C.J., Raimondi, A., Scott, M.P., Chesworth, R., Moyer, M.P., Copeland, R.A., et al. (2013). Durable tumor regression in genetically altered malignant rhabdoid tumors by inhibition of methyltransferase EZH2. Proc. Natl. Acad. Sci. U.S.a. *110*, 7922–7927.

Knutson, S.K., Wigle, T.J., Warholic, N.M., Sneeringer, C.J., Allain, C.J., Klaus, C.R., Sacks, J.D., Raimondi, A., Majer, C.R., Song, J., et al. (2012). A selective inhibitor of EZH2 blocks H3K27 methylation and kills mutant lymphoma cells. Nat. Chem. Biol. *8*, 890–896.

Kohlmaier, A., Savarese, F., Lachner, M., Martens, J., Jenuwein, T., and Wutz, A. (2004). A chromosomal memory triggered by Xist regulates histone methylation in X inactivation. PLoS Biol. *2*, 991–1003.

Konze, K.D., Ma, A., Li, F., Barsyte-Lovejoy, D., Parton, T., MacNevin, C.J., Liu, F., Gao, C., Huang, X.-P., Kuznetsova, E., et al. (2013). An Orally Bioavailable Chemical Probe of the Lysine Methyltransferases EZH2 and EZH1. ACS Chem. Biol. *8*, 1324–1334.

Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. Science *184*, 868–871.

Kornblau, S.M., Qiu, Y.H., Zhang, N., Singh, N., Faderl, S., Ferrajoli, A., York, H., Qutub, A.A., Coombes, K.R., and Watson, D.K. (2011). Abnormal expression of FLI1 protein is an adverse prognostic factor in acute myeloid leukemia. Blood *118*, 5604–5612.

Kotani, A., Ha, D., Hsieh, J., Rao, P.K., Schotte, D., Boer, den, M.L., Armstrong, S.A., and Lodish, H.F. (2009). miR-128b is a potent glucocorticoid sensitizer in MLL-AF4 acute lymphocytic leukemia cells and exerts cooperative effects with miR-221. Blood *114*, 4169–4178.

Krivtsov, A.V., Feng, Z., Lemieux, M.E., Faber, J., Vempati, S., Sinha, A.U., Xia, X., Jesneck, J., Bracken, A.P., Silverman, L.B., et al. (2008). H3K79 Methylation Profiles Define Murine and Human MLL-AF4 Leukemias. Cancer Cell *14*, 355–368.

Krivtsov, A.V., Twomey, D., Feng, Z., Stubbs, M.C., Wang, Y., Faber, J., Levine, J.E., Wang, J., Hahn, W.C., Gilliland, D.G., et al. (2006). Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. Nature *442*, 818–822.

Kruse, E.A., Loughran, S.J., Baldwin, T.M., Josefsson, E.C., Ellis, S., Watson, D.K., Nurden, P., Metcalf, D., Hilton, D.J., Alexander, W.S., et al. (2009). Dual

requirement for the ETS transcription factors Fli-1 and Erg in hematopoietic stem cells and the megakaryocyte lineage. Proc. Natl. Acad. Sci. U.S.a. *106*, 13814–13819.

Ku, M., Koche, R.P., Rheinbay, E., Mendenhall, E.M., Endoh, M., Mikkelsen, T.S., Presser, A., Nusbaum, C., Xie, X., Chi, A.S., et al. (2008). Genomewide Analysis of PRC1 and PRC2 Occupancy Identifies Two Classes of Bivalent Domains. PLoS Genet. *4*.

Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., Ziller, M.J., et al. (2015). Integrative analysis of 111 reference human epigenomes. Nature *518*, –330.

Kuzmichev, A., Jenuwein, T., Tempst, P., and Reinberg, D. (2004). Different Ezh2-containing complexes target methylation of histone H1 or nucleosomal histone H3. Mol. Cell *14*, 183–193.

Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2002). Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. Genes Dev. *16*, 2893–2905.

Lacoste, N., Utley, R.T., Hunter, J.M., Poirier, G.G., and Côte, J. (2002). Disruptor of telomeric silencing-1 is a chromatin-specific histone H3 methyltransferase. J. Biol. Chem. *277*, 30421–30424.

Lan, F., Bayliss, P.E., Rinn, J.L., Whetstine, J.R., Wang, J.K., Chen, S., Iwase, S., Alpatov, R., Issaeva, I., Canaani, E., et al. (2007). A histone H3 lysine 27 demethylase regulates animal posterior development. Nature *449*, 689–U3.

Landeira, D., Sauer, S., Poot, R., Dvorkina, M., Mazzarella, L., Jorgensen, H.F., Pereira, C.F., Leleu, M., Piccolo, F.M., Spivakov, M., et al. (2010). Jarid2 is a PRC2 component in embryonic stem cells required for multi-lineage differentiation and recruitment of PRC1 and RNA Polymerase II to developmental regulators. Nat. Cell Biol. *12*, 618–U214.

Lau, Q.C., Raja, E., Salto-Terez, M., Liu, Q., Ito, K., Inoue, M., Putti, T.C., Loh, M., Ko, T.K., Huang, C., et al. (2006). RUNX3 is frequently inactivated by dual mechanisms of protein mislocalization and promoter hypermethylation in breast cancer. Cancer Res. *66*, 6512–6520.

Lauberth, S.M., Nakayama, T., Wu, X., Ferris, A.L., Tang, Z., Hughes, S.H., and Roeder, R.G. (2013). H3K4me3 Interactions with TAF3 Regulate Preinitiation Complex Assembly and Selective Gene Activation. Cell *152*, 1021–1036.

Laugesen, A., and Helin, K. (2014). Chromatin repressive complexes in stem cells, development, and cancer. Cell Stem Cell *14*, 735–751.

Lee, J.T., Strauss, W.M., Dausman, J.A., and Jaenisch, R. (1996). A 450 kb transgene displays properties of the mammalian X-inactivation center. Cell *86*, 83–94.

Lee, M.G., Villa, R., Trojer, P., Norman, J., Yan, K.-P., Reinberg, D., Di Croce, L., and Shiekhattar, R. (2007). Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. Science *318*, 447–450.

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

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

Levine, M. (2011). Paused RNA Polymerase II as a Developmental Checkpoint. Cell *145*, 502–511.

Levine, S.S., Weiss, A., Erdjument-Bromage, H., Shao, Z., Tempst, P., and Kingston, R.E. (2002). The core of the polycomb repressive complex is compositionally and functionally conserved in flies and humans. Mol. Cell. Biol. *22*, 6070–6078.

Lewis, E.B. (1978). A gene complex controlling segmentation in Drosophila. Nature *276*, 565–570.

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

Li, Y., Luo, H., Liu, T., Zacksenhaus, E., and Ben-David, Y. (2014). The ets transcription factor Fli-1 in development, cancer and disease. Oncogene.

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

Liang, D.C., Shih, L.Y., Fu, J.F., Li, H.Y., Wang, H.I., Hung, I.J., Yang, C.P., Jaing, T.H., Chen, S.H., and Liu, H.C. (2006). K-Ras mutations and N-Ras mutations in childhood acute leukemias with or without mixed-lineage leukemia gene rearrangements. Cancer *106*, 950–956.

Lim, D.A., Huang, Y.-C., Swigut, T., Mirick, A.L., Garcia-Verdugo, J.M., Wysocka, J., Ernst, P., and Alvarez-Buylla, A. (2009). Chromatin remodelling factor Mll1 is essential for neurogenesis from postnatal neural stem cells. Nature *458*, 529–533.

Liu, F., Walmsley, M., Rodaway, A., and Patient, R. (2008). Fli1 acts at the top of the transcriptional network driving blood and endothelial development. Curr. Biol. *18*, 1234–1240.

Liu, J., Mercher, T., Scholl, C., Brumme, K., Gilliland, D.G., and Zhu, N. (2012). A

functional role for the histone demethylase UTX in normal and malignant hematopoietic cells. Exp. Hematol. *40*, 487–498.

Long, H.K., Sims, D., Heger, A., Blackledge, N.P., Kutter, C., Wright, M.L., Grützner, F., Odom, D.T., Patient, R., Ponting, C.P., et al. (2013). Epigenetic conservation at gene regulatory elements revealed by non-methylated DNA profiling in seven vertebrates. Elife *2*, e00348.

Lorch, Y., LaPointe, J.W., and Kornberg, R.D. (1987). Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. Cell *49*, 203–210.

Lu, J., Getz, G., Miska, E.A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B.L., Mak, R.H., Ferrando, A.A., et al. (2005). MicroRNA expression profiles classify human cancers. Nature *435*, 834–838.

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

Luo, Z., Lin, C., Guest, E., Garrett, A.S., Mohaghegh, N., Swanson, S., Marshall, S., Florens, L., Washburn, M.P., and Shilatifard, A. (2012). The Super Elongation Complex Family of RNA Polymerase II Elongation Factors: Gene Target Specificity and Transcriptional Output. Mol. Cell. Biol. *32*, 2608–2617.

Lynch, M.D., Smith, A.J.H., De Gobbi, M., Flenley, M., Hughes, J.R., Vernimmen, D., Ayyub, H., Sharpe, J.A., Sloane-Stanley, J.A., Sutherland, L., et al. (2012). An interspecies analysis reveals a key role for unmethylated CpG dinucleotides in vertebrate Polycomb complex recruitment. Embo J. *31*, 317–329.

Lyon, M.F. (1961). Gene Action in X-Chromosome of Mouse (Mus Musculus L). Nature *190*, 372–&.

Maethner, E., Garcia-Cuellar, M.-P., Breitinger, C., Takacova, S., Divoky, V., Hess, J.L., and Slany, R.K. (2013). MLL-ENL inhibits polycomb repressive complex 1 to achieve efficient transformation of hematopoietic cells. Cell Rep *3*, 1553–1566.

Majewski, I.J., Blewitt, M.E., de Graaf, C.A., McManus, E.J., Bahlo, M., Hilton, A.A., Hyland, C.D., Smyth, G.K., Corbin, J.E., Metcalf, D., et al. (2008). Polycomb repressive complex 2 (PRC2) restricts hematopoietic stem cell activity. PLoS Biol. *6*, 796–809.

Mak, W., Baxter, J., Silva, J., Newall, A.E., Otte, A.P., and Brockdorff, N. (2002). Mitotically stable association of polycomb group proteins Eed and Enx1 with the inactive X chromosome in trophoblast stem cells. Curr. Biol. *12*, 1016–1020.

Mak, W., Nesterova, T.B., de Napoles, M., Appanah, R., Yamanaka, S., Otte, A.P., and Brockdorff, N. (2004). Reactivation of the paternal X chromosome in early mouse embryos. Science *303*, 666–669.

Malik, B., and Hemenway, C.S. (2013). CBX8, a component of the Polycomb PRC1 complex, modulates DOT1L-mediated gene expression through AF9/MLLT3. FEBS Lett. *587*, 3038–3044.

Margueron, R., Justin, N., Ohno, K., Sharpe, M.L., Son, J., Drury, W.J.I., Voigt, P., Martin, S.R., Taylor, W.R., De Marco, V., et al. (2009). Role of the polycomb protein EED in the propagation of repressive histone marks. Nature *461*, 762–U11.

Margueron, R., Li, G., Sarma, K., Blais, A., Zavadil, J., Woodcock, C.L., Dyniacht, B.D., and Reinberg, D. (2008). Ezh1 and Ezh2 Maintain Repressive Chromatin through Different Mechanisms. Mol. Cell *32*, 503–518.

Marmorstein, R., and Zhou, M.-M. (2014). Writers and Readers of Histone Acetylation: Structure, Mechanism, and Inhibition. Cold Spring Harb Perspect Biol *6*, –a018762.

Marschalek, R. (2010). Mixed lineage leukemia: roles in human malignancies and potential therapy. Febs J. *277*, 1822–1831.

Marschalek, R. (2015). MLL leukemia and future treatment strategies. Arch. Pharm. (Weinheim) *348*, 221–228.

Marsden, M.P., and Laemmli, U.K. (1979). Metaphase chromosome structure: evidence for a radial loop model. Cell *17*, 849–858.

McCabe, M.T., Graves, A.P., Ganji, G., Diaz, E., Halsey, W.S., Jiang, Y., Smitheman, K.N., Ott, H.M., Pappalardi, M.B., Allen, K.E., et al. (2012a). Mutation of A677 in histone methyltransferase EZH2 in human B-cell lymphoma promotes hypertrimethylation of histone H3 on lysine 27 (H3K27). Proc. Natl. Acad. Sci. U.S.a. *109*, 2989–2994.

McCabe, M.T., Ott, H.M., Ganji, G., Korenchuk, S., Thompson, C., Van Aller, G.S., Liu, Y., Graves, A.P., Pietra, Della, A.I., Diaz, E., et al. (2012b). EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. Nature *492*, 108–.

McCabe, N.R., Burnett, R.C., Gill, H.J., Thirman, M.J., Mbangkollo, D., Kipiniak, M., van Melle, E., Ziemin-van der Poel, S., Rowley, J.D., and Diaz, M.O. (1992). Cloning of cDNAs of the MLL gene that detect DNA rearrangements and altered RNA transcripts in human leukemic cells with 11q23 translocations. Proc. Natl. Acad. Sci. U.S.a. *89*, 11794–11798.

McGowan, S.J., Hughes, J.R., Han, Z.-P., and Taylor, S. (2013). MIG: Multi-Image Genome viewer. Bioinformatics *29*, 2477–2478.

McMahon, K.A., Hiew, S.Y.-L., Hadjur, S., Veiga-Fernandes, H., Menzel, U., Price, A.J., Kioussis, D., Williams, O., and Brady, H.J.M. (2007). Mll has a critical role in fetal and adult hematopoietic stem cell self-renewal. Cell Stem Cell *1*, 338–345.

Megee, P.C., Morgan, B.A., and Smith, M.M. (1995). Histone H4 and the Maintenance of Genome Integrity. Genes Dev. *9*, 1716–1727.

Mendenhall, E.M., Koche, R.P., Truong, T., Zhou, V.W., Issac, B., Chi, A.S., Ku, M., and Bernstein, B.E. (2010). GC-Rich Sequence Elements Recruit PRC2 in Mammalian ES Cells. PLoS Genet. *6*.

Metzler, M., Forster, A., Pannell, R., Arends, M.J., Daser, A., Lobato, M.N., and Rabbitts, T.H. (2006). A conditional model of MLL-AF4 B-cell tumourigenesis using invertor technology. Oncogene *25*, 3093–3103.

Meyer, C., Hofmann, J., Burmeister, T., Gröger, D., Park, T.S., Emerenciano, M., Pombo de Oliveira, M., Renneville, A., Villarese, P., Macintyre, E., et al. (2013). The MLL recombinome of acute leukemias in 2013. Leukemia *27*, 2165–2176.

Meyer, C., Kowarz, E., Hofmann, J., Renneville, A., Zuna, J., Trka, J., Ben Abdelali, R., Macintyre, E., De Braekeleer, E., De Braekeleer, M., et al. (2009). New insights to the MLL recombinome of acute leukemias. Leukemia *23*, 1490–1499.

Miao, F., and Natarajan, R. (2005). Mapping global histone methylation patterns in the coding regions of human genes. Mol. Cell. Biol. *25*, 4650–4661.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.-K., Koche, R.P., et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553–U2.

Miller, T., Krogan, N.J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J.F., and Shilatifard, A. (2001). COMPASS: A complex of proteins associated with a trithorax-related SET domain protein. Proc. Natl. Acad. Sci. U.S.a. *98*, 12902–12907.

Milne, T.A., Briggs, S.D., Brock, H.W., Martin, M.E., Gibbs, D., Allis, C.D., and Hess, J.L. (2002). MLL targets SET domain methyltransferase activity to Hox gene promoters. Mol. Cell *10*, 1107–1117.

Milne, T.A., Kim, J., Wang, G.G., Stadler, S.C., Basrur, V., Whitcomb, S.J., Wang, Z., Ruthenburg, A.J., Elenitoba-Johnson, K.S.J., Roeder, R.G., et al. (2010). Multiple interactions recruit MLL1 and MLL1 fusion proteins to the HOXA9 locus in leukemogenesis. Mol. Cell *38*, 853–863.

Min, J.R., Zhang, Y., and Xu, R.M. (2003). Structural basis for specific binding of polycomb chromodomain to histone H3 methylated at Lys 27. Genes Dev. *17*, 1823–1828.

Miranda, T.B., Cortez, C.C., Yoo, C.B., Liang, G., Abe, M., Kelly, T.K., Marquez, V.E., and Jones, P.A. (2009). DZNep is a global histone methylation inhibitor that reactivates developmental genes not silenced by DNA methylation. Mol. Cancer Ther. *8*, 1579–1588.

Mishra, B.P., Zaffuto, K.M., Artinger, E.L., Org, T., Mikkola, H.K.A., Cheng, C., Djabali, M., and Ernst, P. (2014). The histone methyltransferase activity of MLL1 is dispensable for hematopoiesis and leukemogenesis. Cell Rep *7*, 1239–1247.

Miyazaki, H., Higashimoto, K., Yada, Y., Endo, T.A., Sharif, J., Komori, T., Matsuda, M., Koseki, Y., Nakayama, M., Soejima, H., et al. (2013). Ash1l Methylates Lys36 of Histone H3 Independently of Transcriptional Elongation to Counteract Polycomb Silencing. PLoS Genet. *9*.

Mochizuki-Kashio, M. (2012). Dependency on the polycomb gene Ezh2 distinguishes fetal from adult hematopoietic stem cells (vol 118, pg 6553, 2011). Blood *120*, 924–924.

Mohn, F., Weber, M., Rebhan, M., Roloff, T.C., Richter, J., Stadler, M.B., Bibel, M., and Schuebeler, D. (2008). Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. Mol. Cell *30*, 755–766.

Monroe, S.C., Jo, S.Y., Sanders, D.S., Basrur, V., Elenitoba-Johnson, K.S., Slany, R.K., and Hess, J.L. (2011). MLL-AF9 and MLL-ENL alter the dynamic association of transcriptional regulators with genes critical for leukemia. Exp. Hematol. *39*, 77–86.e1–5.

Montes, R., Ayllón, V., Gutierrez-Aranda, I., Prat, I., Carmen Hernandez-Lamas, M., Ponce, L., Bresolin, S., Kronnie, te, G., Greaves, M., Bueno, C., et al. (2011). Enforced expression of MLL-AF4 fusion in cord blood CD34(+) cells enhances the hematopoietic repopulating cell function and clonogenic potential but is not sufficient to initiate leukemia. Blood *117*, 4746–4758.

Montgomery, N.D., Yee, D., Chen, A., Kalantry, S., Chamberlain, S.J., Otte, A.P., and Magnuson, T. (2005). The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation. Curr. Biol. *15*, 942–947.

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

Morin, R.D., Johnson, N.A., Severson, T.M., Mungall, A.J., An, J., Goya, R., Paul, J.E., Boyle, M., Woolcock, B.W., Kuchenbauer, F., et al. (2010). Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. Nat. Genet. *42*, 181–U124.

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

Mueller, D., Bach, C., Zeisig, D., Garcia-Cuellar, M.-P., Monroe, S., Sreekumar, A., Zhou, R., Nesvizhskii, A., Chinnaiyan, A., Hess, J.L., et al. (2007). A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification. Blood *110*, 4445–4454.

Mujtaba, S., Manzur, K.L., Gurnon, J.R., Kang, M., Van Etten, J.L., and Zhou, M.-M. (2008). Epigenetic transcriptional repression of cellular genes by a viral SET protein. Nat. Cell Biol. *10*, 1114–1122.

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

Muntean, A.G., Giannola, D., Udager, A.M., and Hess, J.L. (2008). The PHD fingers of MLL block MLL fusion protein-mediated transformation. Blood *112*, 4690–4693.

Murzina, N.V., Pei, X.-Y., Zhang, W., Sparkes, M., Vicente-Garcia, J., Pratap, J.V., McLaughlin, S.H., Ben-Shahar, T.R., Verreault, A., Luisi, B.F., et al. (2008). Structural basis for the recognition of histone H4 by the histone-chaperone RbAp46. Structure *16*, 1077–1085.

Muse, G.W., Gilchrist, D.A., Nechaev, S., Shah, R., Parker, J.S., Grissom, S.F., Zeitlinger, J., and Adelman, K. (2007). RNA polymerase is poised for activation across the genome. Nat. Genet. *39*, 1507–1511.

Musselman, C.A., Avvakumov, N., Watanabe, R., Abraham, C.G., Lalonde, M.-E., Hong, Z., Allen, C., Roy, S., Nunez, J.K., Nickoloff, J., et al. (2012). Molecular basis for H3K36me3 recognition by the Tudor domain of PHF1. Nat. Struct. Mol. Biol. *19*, 1266–.

Musselman, C.A., Mansfield, R.E., Garske, A.L., Davrazou, F., Kwan, A.H., Oliver, S.S., O'Leary, H., Denu, J.M., Mackay, J.P., and Kutateladze, T.G. (2009). Binding of the CHD4 PHD2 finger to histone H3 is modulated by covalent modifications. Biochem. J. *423*, 179–187.

Nagy, P.L., Griesenbeck, J., Kornberg, R.D., and Cleary, M.L. (2002). A trithorax-group complex purified from Saccharomyces cerevisiae is required for methylation of histone H3. Proc. Natl. Acad. Sci. U.S.a. *99*, 90–94.

Nakamura, T., Mori, T., Tada, S., Krajewski, W., Rozovskaia, T., Wassell, R., Dubois, G., Mazo, A., Croce, C.M., and Canaani, E. (2002). ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. Mol. Cell *10*, 1119–1128.

Neff, T., Sinha, A.U., Kluk, M.J., Zhu, N., Khattab, M.H., Stein, L., Xie, H., Orkin, S.H., and Armstrong, S.A. (2012). Polycomb repressive complex 2 is required for MLL-AF9 leukemia. Proc. Natl. Acad. Sci. U.S.a. *109*, 5028–5033.

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

Nguyen, A.T., Taranova, O., He, J., and Zhang, Y. (2011). DOT1L, the H3K79 methyltransferase, is required for MLL-AF9-mediated leukemogenesis. Blood

117, 6912–6922.

Nishino, Y., Eltsov, M., Joti, Y., Ito, K., Takata, H., Takahashi, Y., Hihara, S., Frangakis, A.S., Imamoto, N., Ishikawa, T., et al. (2012). Human mitotic chromosomes consist predominantly of irregularly folded nucleosome fibres without a 30-nm chromatin structure. Embo J. *31*, 1644–1653.

Nowak, A.J., Alfieri, C., Stirnimann, C.U., Rybin, V., Baudin, F., Ly-Hartig, N., Lindner, D., and Mueller, C.W. (2011). Chromatin-modifying Complex Component Nurf55/p55 Associates with Histones H3 and H4 and Polycomb Repressive Complex 2 Subunit Su(z) 12 through Partially Overlapping Binding Sites. J. Biol. Chem. *286*, 23388–23396.

Ntziachristos, P., Tsirigos, A., Van Vlierberghe, P., Nedjic, J., Trimarchi, T., Flaherty, M.S., Ferres-Marco, D., da Ros, V., Tang, Z., Siegle, J., et al. (2012). Genetic inactivation of the polycomb repressive complex 2 in T cell acute lymphoblastic leukemia. Nat. Med. *18*, 298–303.

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

Obier, N., Lin, Q., Cauchy, P., Hornich, V., Zenke, M., Becker, M., and Müller, A.M. (2014). Polycomb Protein EED is Required for Silencing of Pluripotency Genes upon ESC Differentiation. Stem Cell Rev 1–12.

Okada, Y., Feng, Q., Lin, Y.H., Jiang, Q., Li, Y.Q., Coffield, V.M., Su, L.S., Xu, G.L., and Zhang, Y. (2005). hDOT1L links histone methylation to leukemogenesis. Cell *121*, 167–178.

Ooi, S.K.T., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S.-P., Allis, C.D., et al. (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature *448*, 714–717.

Orlando, D.A., Chen, M.W., Brown, V.E., Solanki, S., Choi, Y.J., Olson, E.R., Fritz, C.C., Bradner, J.E., and Guenther, M.G. (2014). Quantitative ChIP-Seq Normalization Reveals Global Modulation of the Epigenome. Cell Rep *9*, 1163–1170.

Orlovsky, K., Kalinkovich, A., Rozovskaia, T., Shezen, E., Itkin, T., Alder, H., Ozer, H.G., Carramusa, L., Avigdor, A., Volinia, S., et al. (2011). Down-regulation of homeobox genes MEIS1 and HOXA in MLL-rearranged acute leukemia impairs engraftment and reduces proliferation. Proc. Natl. Acad. Sci. U.S.a. *108*, 7956–7961.

Palermo, C.M., Bennett, C.A., Winters, A.C., and Hemenway, C.S. (2008). The AF4-mimetic peptide, PFWT, induces necrotic cell death in MV4-11 leukemia cells. Leuk. Res. *32*, 633–642.

Pan, G., Tian, S., Nie, J., Yang, C., Ruotti, V., Wei, H., Jonsdottir, G.A., Stewart, R., and

Thomson, J.A. (2007). Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. Cell Stem Cell *1*, 299–312.

Park, I.K., Qian, D.L., Kiel, M., Becker, M.W., Pihalja, M., Weissman, I.L., Morrison, S.J., and Clarke, M.F. (2003). Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells. Nature *423*, 302–305.

Park, S., Osmers, U., Raman, G., Schwantes, R.H., Diaz, M.O., and Bushweller, J.H. (2010). The PHD3 Domain of MLL Acts as a CYP33-Regulated Switch between MLL-Mediated Activation and Repression. Biochemistry *49*, 6576–6586.

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

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

Pasini, D., Cloos, P.A.C., Walfridsson, J., Olsson, L., Bukowski, J.-P., Johansen, J.V., Bak, M., Tommerup, N., Rappsilber, J., and Helin, K. (2010a). JARID2 regulates binding of the Polycomb repressive complex 2 to target genes in ES cells. Nature *464*, 306–U193.

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

Peng, J.C., Valouev, A., Swigut, T., Zhang, J., Zhao, Y., Sidow, A., and Wysocka, J. (2009). Jarid2/Jumonji Coordinates Control of PRC2 Enzymatic Activity and Target Gene Occupancy in Pluripotent Cells. Cell *139*, 1290–1302.

Penny, G.D., Kay, G.F., Sheardown, S.A., RASTAN, S., and Brockdorff, N. (1996). Requirement for Xist in X chromosome inactivation. Nature *379*, 131–137.

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

Pieters, R., Schrappe, M., De Lorenzo, P., Hann, I., De Rossi, G., Felice, M., Hovi, L., LeBlanc, T., Szczepanski, T., Ferster, A., et al. (2007). A treatment protocol for infants younger than 1 year with acute lymphoblastic leukaemia (Interfant-99): an observational study and a multicentre randomised trial. Lancet *370*, 240–250.

Piunti, A., Rossi, A., Cerutti, A., Albert, M., Jammula, S., Scelfo, A., Cedrone, L., Fragola, G., Olsson, L., Koseki, H., et al. (2014). Polycomb proteins control

proliferation and transformation independently of cell cycle checkpoints by regulating DNA replication. Nat Commun *5*.

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

Pokholok, D.K., Harbison, C.T., Levine, S., Cole, M., Hannett, N.M., Lee, T.I., Bell, G.W., Walker, K., Rolfe, P.A., Herbolsheimer, E., et al. (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. Cell *122*, 517–527.

Pui, C.H., Sandlund, J.T., Pei, D.Q., Campana, D., Rivera, G.K., Ribeiro, R.C., Rubnitz, J.E., Razzouk, B.I., Howard, S.C., Hudson, M.M., et al. (2004). Improved outcome for children with acute lymphoblastic leukemia: results of Total Therapy Study XIIIB at St Jude Children's Research Hospital. Blood *104*, 2690–2696.

Qi, W., Chan, H., Teng, L., Li, L., Chuai, S., Zhang, R., Zeng, J., Li, M., Fan, H., Lin, Y., et al. (2012). Selective inhibition of Ezh2 by a small molecule inhibitor blocks tumor cells proliferation. Proc. Natl. Acad. Sci. U.S.a. *109*, 21360–21365.

Raedle, B., Rutkowski, A.J., Ruzsics, Z., Friedel, C.C., Koszinowski, U.H., and Doelken, L. (2013). Metabolic Labeling of Newly Transcribed RNA for High Resolution Gene Expression Profiling of RNA Synthesis, Processing and Decay in Cell Culture. J Vis Exp –e50195.

Ramakers-van Woerden, N.L., Beverloo, H.B., Veerman, A.J.P., Camitta, B.M., Loonen, A.H., Van Wering, E.R., Slater, R.M., Harbott, J., Boer, den, M.L., Ludwig, W.D., et al. (2004). In vitro drug-resistance profile in infant acute lymphoblastic leukemia in relation to age, MLL rearrangements and immunophenotype. Leukemia *18*, 521–529.

Rao, B., Shibata, Y., Strahl, B.D., and Lieb, J.D. (2005). Dimethylation of histone H3 at lysine 36 demarcates regulatory and nonregulatory chromatin genome-wide. Mol. Cell. Biol. *25*, 9447–9459.

Reaman, G.H., Sposto, R., Sensel, M.G., Lange, B.J., Feusner, J.H., Heerema, N.A., Leonard, M., Holmes, E.J., Sather, H.N., Pendergrass, T.W., et al. (1999). Treatment outcome and prognostic factors for infants with acute lymphoblastic leukemia treated on two consecutive trials of the Children's Cancer Group. J. Clin. Oncol. *17*, 445–455.

Reddington, J.P., Perricone, S.M., Nestor, C.E., Reichmann, J., Youngson, N.A., Suzuki, M., Reinhardt, D., Dunican, D.S., Prendergast, J.G., Mjoseng, H., et al. (2013). Redistribution of H3K27me3 upon DNA hypomethylation results in de-repression of Polycomb target genes. Genome Biol. *14*.

Ren, X., and Kerppola, T.K. (2011). REST Interacts with Cbx Proteins and Regulates Polycomb Repressive Complex 1 Occupancy at RE1 Elements. Mol. Cell. Biol. *31*, 2100–2110.

Rice, K.L., Hormaeche, I., and Licht, J.D. (2007). Epigenetic regulation of normal and malignant hematopoiesis. Oncogene *26*, 6697–6714.

Richon, V.M., Emiliani, S., Verdin, E., Webb, Y., Breslow, R., Rifkind, R.A., and Marks, P.A. (1998). A class of hybrid polar inducers of transformed cell differentiation inhibits histone deacetylases. Proc. Natl. Acad. Sci. U.S.a. *95*, 3003–3007.

Riising, E.M., Comet, I., Leblanc, B., Wu, X., Johansen, J.V., and Helin, K. (2014). Gene Silencing Triggers Polycomb Repressive Complex 2 Recruitment to CpG Islands Genome Wide. Mol. Cell *55*, 347–360.

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

Robinson, P.J.J., and Rhodes, D. (2006). Structure of the “30 nm” chromatin fibre: a key role for the linker histone. Curr. Opin. Struct. Biol. *16*, 336–343.

Roe, J.-S., Mercan, F., Rivera, K., Pappin, D.J., and Vakoc, C.R. (2015). BET Bromodomain Inhibition Suppresses the Function of Hematopoietic Transcription Factors in Acute Myeloid Leukemia. Mol. Cell *58*, 1028–1039.

Roguev, A., Schaft, D., Shevchenko, A., Pijnappel, W., Wilm, M., Aasland, R., and Stewart, A.F. (2001). The Saccharomyces cerevisiae Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. Embo J. *20*, 7137–7148.

Ross, J., Mavoungou, L., Bresnick, E.H., and Milot, E. (2012). GATA-1 Utilizes Ikaros and Polycomb Repressive Complex 2 To Suppress Hes1 and To Promote Erythropoiesis. Mol. Cell. Biol. *32*, 3624–3638.

Ruthenburg, A.J., Li, H., Milne, T.A., Dewell, S., McGinty, R.K., Yuen, M., Ueberheide, B., Dou, Y., Muir, T.W., Patel, D.J., et al. (2011). Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions. Cell *145*, 692–706.

Saurin, A.J., Shao, Z.H., Erdjument-Bromage, H., Tempst, P., and Kingston, R.E. (2001). A Drosophila Polycomb group complex includes Zeste and dTAFII proteins. Nature *412*, 655–660.

Saxonov, S., Berg, P., and Brutlag, D.L. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proc. Natl. Acad. Sci. U.S.a. *103*, 1412–1417.

Schmidt, S., Rainer, J., Ploner, C., Presul, E., Riml, S., and Kofler, R. (2004). Glucocorticoid-induced apoptosis and glucocorticoid resistance: molecular mechanisms and clinical relevance. Cell Death Differ. *11*, S45–S55.

Schmitges, F.W., Prusty, A.B., Faty, M., Stuetzer, A., Lingaraju, G.M., Aiwazian, J., Sack, R., Hess, D., Li, L., Zhou, S., et al. (2011). Histone Methylation by PRC2 Is Inhibited by Active Chromatin Marks. Mol. Cell *42*, 330–341.

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

Schreiner, S., Birke, M., Garcia-Cuellar, M.P., Zilles, O., Greil, J., and Slany, R.K. (2001). MLL-ENL causes a reversible and myc-dependent block of myelomonocytic cell differentiation. Cancer Res. *61*, 6480–6486.

Schroeder, S., Cho, S., Zeng, L., Zhang, Q., Kaehlcke, K., Mak, L., Lau, J., Bisgrove, D., Schnoelzer, M., Verdin, E., et al. (2012). Two-pronged Binding with Bromodomain-containing Protein 4 Liberates Positive Transcription Elongation Factor b from Inactive Ribonucleoprotein Complexes. J. Biol. Chem. *287*, 1090–1099.

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

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

Shen, X., Kim, W., Fujiwara, Y., Simon, M.D., Liu, Y., Mysliwiec, M.R., Yuan, G.-C., Lee, Y., and Orkin, S.H. (2009). Jumonji Modulates Polycomb Activity and Self-Renewal versus Differentiation of Stem Cells. Cell *139*, 1303–1314.

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

Shi, J., Wang, E., Zuber, J., Rappaport, A., Taylor, M., Johns, C., Lowe, S.W., and Vakoc, C.R. (2013). The Polycomb complex PRC2 supports aberrant self-renewal in a mouse model of MLL-AF9;Nras(G12D) acute myeloid leukemia. Oncogene *32*, 930–938.

Shilatifard, A., Lane, W.S., Jackson, K.W., Conaway, R.C., and Conaway, J.W. (1996). An RNA polymerase II elongation factor encoded by the human ELL gene. Science *271*, 1873–1876.

Silva, J., Mak, W., Zvetkova, I., Appanah, R., Nesterova, T.B., Webster, Z., Peters, A., Jenuwein, T., Otte, A.P., and Brockdorff, N. (2003). Establishment of histone H3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 Polycomb group complexes. Dev. Cell *4*, 481–495.

Silverman, L.R., and Mufti, G.J. (2005). Methylation inhibitor therapy in the treatment of myelodysplastic syndrome. Nat Clin Pract Oncol *2 Suppl 1*, S12–S23.

Smith, L.-L., Yeung, J., Zeisig, B.B., Popov, N., Huijbers, I., Barnes, J., Wilson, A.J., Taskesen, E., Delwel, R., Gil, J., et al. (2011). Functional crosstalk between Bmi1 and MLL/Hoxa9 axis in establishment of normal hematopoietic and leukemic stem cells. Cell Stem Cell *8*, 649–662.

Smith, S., and Stillman, B. (1989). Purification and characterization of CAF-I, a human cell factor required for chromatin assembly during DNA replication in vitro. Cell *58*, 15–25.

Smith, S., and Stillman, B. (1991). Stepwise assembly of chromatin during DNA replication in vitro. Embo J. *10*, 971–980.

Sneeringer, C.J., Scott, M.P., Kuntz, K.W., Knutson, S.K., Pollock, R.M., Richon, V.M., and Copeland, R.A. (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. Proc. Natl. Acad. Sci. U.S.a. *107*, 20980–20985.

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

Spijkers-Hagelstein, J.A.P., Pinhancos, S.M., Schneider, P., Pieters, R., and Stam, R.W. (2013). Src kinase-induced phosphorylation of annexin A2 mediates glucocorticoid resistance in MLL-rearranged infant acute lymphoblastic leukemia. Leukemia *27*, 1063–1071.

Spijkers-Hagelstein, J.A.P., Pinhancos, S.S., Schneider, P., Pieters, R., and Stam, R.W. (2014). Chemical genomic screening identifies LY294002 as a modulator of glucocorticoid resistance in MLL-rearranged infant ALL. Leukemia *28*, 761–769.

Srinivasan, R.S., Nesbit, J.B., Marrero, L., Erfurth, F., LaRussa, V.F., and Hemenway, C.S. (2004). The synthetic peptide PFWT disrupts AF4-AF9 protein complexes and induces apoptosis in t(4;11) leukemia cells. Leukemia *18*, 1364–1372.

Stam, R.W. (2012). MLL-AF4 driven leukemogenesis: what are we missing? Cell Res. *22*, 948–949.

Stam, R.W., Boer, den, M.L., Schneider, P., Meier, M., Beverloo, H.B., and Pieters, R. (2007). D-HPLC analysis of the entire FLT3 gene in MLL rearranged and hyperdiploid acute lymphoblastic leukemia. Haematologica *92*, 1565–1568.

Stam, R.W., Schneider, P., Hagelstein, J.A.P., van der Linden, M.H., Stumpel, D.J.P.M., de Menezes, R.X., De Lorenzo, P., Valsecchi, M.G., and Pieters, R. (2010). Gene expression profiling-based dissection of MLL translocated and MLL germline acute lymphoblastic leukemia in infants. Blood *115*, 2835–2844.

Stanulla, M., Chhalliyil, P., Wang, J.J., Jani-Sait, S.N., and Aplan, P.D. (2001).

Mechanisms of MLL gene rearrangement: site-specific DNA cleavage within the breakpoint cluster region is independent of chromosomal context. Hum. Mol. Genet. *10*, 2481–2491.

Steger, D.J., Lefterova, M.I., Ying, L., Stonestrom, A.J., Schupp, M., Zhuo, D., Vakoc, A.L., Kim, J.-E., Chen, J., Lazar, M.A., et al. (2008). DOT1L/KMT4 recruitment and H3K79 methylation are ubiquitously coupled with gene transcription in mammalian cells. Mol. Cell. Biol. *28*, 2825–2839.

Stock, J.K., Giadrossi, S., Casanova, M., Brookes, E., Vidal, M., Koseki, H., Brockdorff, N., Fisher, A.G., and Pombo, A. (2007). Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. Nat. Cell Biol. *9*, 1428–U174.

Stojic, L., Jasencakova, Z., Prezioso, C., Stuetzer, A., Bodega, B., Pasini, D., Klingberg, R., Mozzetta, C., Margueron, R., Puri, P.L., et al. (2011). Chromatin regulated interchange between polycomb repressive complex 2 (PRC2)-Ezh2 and PRC2-Ezh1 complexes controls myogenin activation in skeletal muscle cells. Epigenetics Chromatin *4*.

Strahl, B.D., Grant, P.A., Briggs, S.D., Sun, Z.W., Bone, J.R., Caldwell, J.A., Mollah, S., Cook, R.G., Shabanowitz, J., Hunt, D.F., et al. (2002). Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. Mol. Cell. Biol. *22*, 1298–1306.

Su, I.H., Basavaraj, A., Krutchinsky, A.N., Hobert, O., Ullrich, A., Chait, B.T., and Tarakhovsky, A. (2003). Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement. Nat. Immunol. *4*, 124–131.

Suzuki, H., Forrest, A.R.R., van Nimwegen, E., Daub, C.O., Balwierz, P.J., Irvine, K.M., Lassmann, T., Ravasi, T., Hasegawa, Y., de Hoon, M.J.L., et al. (2009). The transcriptional network that controls growth arrest and differentiation in a human myeloid leukemia cell line. Nat. Genet. *41*, 553–562.

Taketani, T., Taki, T., Sugita, K., Furuichi, Y., Ishii, E., Hanada, R., Tsuchicla, M., Ida, K., and Hayashi, Y. (2004). FLT3 mutations in the activation loop of tyrosine kinase domain are frequently found in infant ALL with MLL rearrangements and pediatric ALL with hyperdiploidy. Blood *103*, 1085–1088.

Tamai, H., Miyake, K., Takatori, M., Miyake, N., Yamaguchi, H., Dan, K., Shimada, T., and Inokuchi, K. (2011). Activated K-Ras protein accelerates human MLL/AF4-induced leukemo-lymphomogenicity in a transgenic mouse model. Leukemia *25*, 888–891.

Tamai, H., and Inokuchi, K. (2010). 11q23/MLL acute leukemia : update of clinical aspects. J Clin Exp Hematop *50*, 91–98.

Tan, J., Jones, M., Koseki, H., Nakayama, M., Muntean, A.G., Maillard, I., and Hess, J.L. (2011). CBX8, a Polycomb Group Protein, Is Essential for MLL-AF9-Induced

Leukemogenesis. Cancer Cell *20*, 563–575.

Tan, J., Yang, X., Zhuang, L., Jiang, X., Chen, W., Lee, P.L., Karuturi, R.K.M., Tan, P.B.O., Liu, E.T., and Yu, Q. (2007). Pharmacologic disruption of polycomb-repressive complex 2-mediated gene repression selectively induces apoptosis in cancer cells. Genes Dev. *21*, 1050–1063.

Tanaka, S., Miyagi, S., Sashida, G., Chiba, T., Yuan, J., Mochizuki-Kashio, M., Suzuki, Y., Sugano, S., Nakaseko, C., Yokote, K., et al. (2012). Ezh2 augments leukemogenicity by reinforcing differentiation blockage in acute myeloid leukemia. Blood *120*, 1107–1117.

Tanay, A., O'Donnell, A.H., Damelin, M., and Bestor, T.H. (2007). Hyperconserved CpG domains underlie Polycomb-binding sites. Proc. Natl. Acad. Sci. U.S.a. *104*, 5521–5526.

Tao, Y.-F., Xu, L.-X., Lu, J., Hu, S.-Y., Fang, F., Cao, L., Xiao, P.-F., Du, X.-J., Sun, L.-C., Li, Z.-H., et al. (2015). Early B-cell factor 3 (EBF3) is a novel tumor suppressor gene with promoter hypermethylation in pediatric acute myeloid leukemia. J. Exp. Clin. Cancer Res. *34*.

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

Terranova, R., Agherbi, H., Boned, A., Meresse, S., and Djabali, M. (2006). Histone and DNA methylation defects at Hox genes in mice expressing a SET domain-truncated form of Mll. Proc. Natl. Acad. Sci. U.S.a. *103*, 6629–6634.

Thiel, A.T., Blessington, P., Zou, T., Feather, D., Wu, X., Yan, J., Zhang, H., Liu, Z., Ernst, P., Koretzky, G.A., et al. (2010). MLL-AF9-Induced Leukemogenesis Requires Coexpression of the Wild-Type Mll Allele. Cancer Cell *17*, 148–159.

Thiel, A.T., Feng, Z., Pant, D.K., Chodosh, L.A., and Hua, X. (2013). The trithorax protein partner menin acts in tandem with EZH2 to suppress C/EBP alpha and differentiation in MLL-AF9 leukemia. Haematologica *98*, 918–927.

Thirman, M.J., Gill, H.J., Burnett, R.C., Mbangkollo, D., McCabe, N.R., KOBAYASHI, H., Zieminvanderpoel, S., KANEKO, Y., MORGAN, R., SANDBERG, A.A., et al. (1993). Rearrangement of the Mll Gene in Acute Lymphoblastic and Acute Myeloid Leukemias with 11q23 Chromosomal Translocations. N. Engl. J. Med. *329*, 909–914.

Thomas, M., Gessner, A., Vornlocher, H.-P., Hadwiger, P., Greil, J., and Heidenreich, O. (2005). Targeting MLL-AF4 with short interfering RNAs inhibits clonogenicity and engraftment of t(4;11)-positive human leukemic cells. Blood *106*, 3559–3566.

Tie, F., Banerjee, R., Stratton, C.A., Prasad-Sinha, J., Stepanik, V., Zlobin, A., Diaz, M.O., Scacheri, P.C., and Harte, P.J. (2009). CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing. Development *136*, 3131–3141.

Tie, F., Stratton, C.A., Kurzhals, R.L., and Harte, P.J. (2007). The N terminus of Drosophila ESC binds directly to histone H3 and is required for E(Z)-dependent trimethylation of H3 lysine 27. Mol. Cell. Biol. *27*, 2014–2026.

Tkachuk, D.C., Kohler, S., and Cleary, M.L. (1992). Involvement of a homolog of Drosophila trithorax by 11q23 chromosomal translocations in acute leukemias. Cell *71*, 691–700.

Truong, A., and Ben-David, Y. (2000). The role of Fli-1 in normal cell function and malignant transformation. Oncogene *19*, 6482–6489.

Van der Meulen, J., Speleman, F., and Van Vlierberghe, P. (2014). The H3K27me3 demethylase UTX in normal development and disease. Epigenetics *9*, 658–668.

van Leeuwen, F., Gafken, P.R., and Gottschling, D.E. (2002). Dot1p modulates silencing in yeast by methylation of the nucleosome core. Cell *109*, 745–756.

Vardiman, J.W., Thiele, J., Arber, D.A., Brunning, R.D., Borowitz, M.J., Porwit, A., Harris, N.L., Le Beau, M.M., Hellstrom-Lindberg, E., Tefferi, A., et al. (2009). The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. Blood *114*, 937–951.

Vermeulen, M., Mulder, K.W., Denissov, S., Pijnappel, W.W.M.P., van Schaik, F.M.A., Varier, R.A., Baltissen, M.P.A., Stunnenberg, H.G., Mann, M., and Timmers, H.T.M. (2007). Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. Cell *131*, 58–69.

Verreault, A., Kaufman, P.D., Kobayashi, R., and Stillman, B. (1998). Nucleosomal DNA regulates the core-histone-binding subunit of the human Hat1 acetyltransferase. Curr. Biol. *8*, 96–108.

Voigt, P., LeRoy, G., Drury, W.J., Zee, B.M., Son, J., Beck, D.B., Young, N.L., Garcia, B.A., and Reinberg, D. (2012). Asymmetrically modified nucleosomes. Cell *151*, 181–193.

Voigt, P., Tee, W.-W., and Reinberg, D. (2013). A double take on bivalent promoters. Genes Dev. *27*, 1318–1338.

Wagner, E.J., and Carpenter, P.B. (2012). Understanding the language of Lys36 methylation at histone H3. Nat. Rev. Mol. Cell Biol. *13*, 115–126.

Wagner, J.M., Hackanson, B., Lübbert, M., and Jung, M. (2010). Histone deacetylase (HDAC) inhibitors in recent clinical trials for cancer therapy. Clin

Epigenetics *1*, 117–136.

Wang, H.B., Wang, L.J., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R.S., and Zhang, Y. (2004a). Role of histone H2A ubiquitination in polycomb silencing. Nature *431*, 873–878.

Wang, L., Brown, J.L., Cao, R., Zhang, Y., Kassis, J.A., and Jones, R.S. (2004b). Hierarchical recruitment of Polycomb group silencing complexes. Mol. Cell *14*, 637–646.

Wang, Q.-F., Wu, G., Mi, S., He, F., Wu, J., Dong, J., Luo, R.T., Mattison, R., Kaberlein, J.J., Prabhakar, S., et al. (2011). MLL fusion proteins preferentially regulate a subset of wild-type MLL target genes in the leukemic genome. Blood *117*, 6895–6905.

Wang, Z., Song, J., Milne, T.A., Wang, G.G., Li, H., Allis, C.D., and Patel, D.J. (2010). Pro isomerization in MLL1 PHD3-bromo cassette connects H3K4me readout to CyP33 and HDAC-mediated repression. Cell *141*, 1183–1194.

Wang, Z., Zang, C., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Peng, W., Zhang, M.Q., et al. (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. Nat. Genet. *40*, 897–903.

Wei, G., Twomey, D., Lamb, J., Schlis, K., Agarwal, J., Stam, R.W., Opferman, J.T., Sallan, S.E., Boer, den, M.L., Pieters, R., et al. (2006). Gene expression-based chemical genomics identifies rapamycin as a modulator of MCL1 and glucocorticoid resistance. Cancer Cell *10*, 331–342.

Wilkinson, A.C., Ballabio, E., Geng, H., North, P., Tapia, M., Kerry, J., Biswas, D., Roeder, R.G., Allis, C.D., Melnick, A., et al. (2013). RUNX1 is a key target in t(4;11) leukemias that contributes to gene activation through an AF4-MLL complex interaction. Cell Rep *3*, 116–127.

Wilson, N.K., Foster, S.D., Wang, X., Knezevic, K., Schütte, J., Kaimakis, P., Chilarska, P.M., Kinston, S., Ouwehand, W.H., Dzierzak, E., et al. (2010). Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. Cell Stem Cell *7*, 532–544.

Wu, X., Johansen, J.V., and Helin, K. (2013). Fbxl10/kdm2b Recruits Polycomb Repressive Complex 1 to CpG Islands and Regulates H2A Ubiquitylation. Mol. Cell *49*, 1134–1146.

Wutz, A., and Jaenisch, R. (2000). A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. Mol. Cell *5*, 695–705.

Wysocka, J., Swigut, T., Milne, T.A., Dou, Y.L., Zhang, X., Burlingame, A.L., Roeder, R.G., Brivanlou, A.H., and Allis, C.D. (2005). WDR5 associates with histone H3 methylated at K4 and is essential for H3K4 methylation and vertebrate development. Cell *121*, 859–872.

Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T., Badenhorst, P., et al. (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature *442*, 86–90.

Xie, H., Xu, J., Hsu, J.H., Nguyen, M., Fujiwara, Y., Peng, C., and Orkin, S.H. (2014). Polycomb Repressive Complex 2 Regulates Normal Hematopoietic Stem Cell Function in a Developmental-Stage-Specific Manner. Cell Stem Cell *14*, 68–80.

Xie, L., Ushmorov, A., Leithaeuser, F., Guan, H., Steidl, C., Faerbinger, J., Pelzer, C., Vogel, M.J., Maier, H.J., Gascoyne, R.D., et al. (2012). FOXO1 is a tumor suppressor in classical Hodgkin lymphoma. Blood *119*, 3503–3511.

Xu, B., On, D.M., Ma, A., Parton, T., Konze, K.D., Pattenden, S.G., Allison, D.F., Cai, L., Rockowitz, S., Liu, S., et al. (2014). Selective inhibition of EZH2 and EZH1 enzymatic activity by a small molecule suppresses MLL-rearranged leukemia. Blood.

Xu, C., Bian, C., Lam, R., Dong, A., and Min, J. (2011). The structural basis for selective binding of non-methylated CpG islands by the CFP1 CXXC domain. Nat Commun *2*.

Xu, C., Bian, C., Yang, W., Galka, M., Ouyang, H., Chen, C., Qiu, W., Liu, H., Jones, A.E., MacKenzie, F., et al. (2010). Binding of different histone marks differentially regulates the activity and specificity of polycomb repressive complex 2 (PRC2). Proc. Natl. Acad. Sci. U.S.a. *107*, 19266–19271.

Yagi, H., Deguchi, K., Aono, A., Tani, Y., Kishimoto, T., and Komori, T. (1998). Growth disturbance in fetal liver hematopoiesis of Mll-mutant mice. Blood *92*, 108–117.

Yamamoto, K., Sonoda, M., Inokuchi, J., Shirasawa, S., and Sasazuki, T. (2004). Polycomb group suppressor of zeste 12 links heterochromatin protein 1 alpha and enhancer of zeste 2. J. Biol. Chem. *279*, 401–406.

Yang, Z.Y., Yik, J., Chen, R.C., He, N.H., Jang, M.K., Ozato, K., and Zhou, Q. (2005). Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein brd4. Mol. Cell *19*, 535–545.

Yap, D.B., Chu, J., Berg, T., Schapira, M., Cheng, S.-W.G., Moradian, A., Morin, R.D., Mungall, A.J., Meissner, B., Boyle, M., et al. (2011). Somatic mutations at EZH2 Y641 act dominantly through a mechanism of selectively altered PRC2 catalytic activity, to increase H3K27 trimethylation. Blood *117*, 2451–2459.

Yokoyama, A., Somervaille, T., Smith, K.S., Rozenblatt-Rosen, O., Meyerson, M., and Cleary, M.L. (2005). The menin tumor suppressor protein is an essential oncogenic cofactor for MLL-associated leukemogenesis. Cell *123*, 207–218.

Yokoyama, A., Wang, Z., Wysocka, J., Sanyal, M., Aufiero, D.J., Kitabayashi, I., Herr,

W., and Cleary, M.L. (2004). Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. Mol. Cell. Biol. *24*, 5639–5649.

Yu, B.D., Hanson, R.D., Hess, J.L., Horning, S.E., and Korsmeyer, S.J. (1998). MLL, a mammalian trithorax-group gene, functions as a transcriptional maintenance factor in morphogenesis. Proc. Natl. Acad. Sci. U.S.a. *95*, 10632–10636.

Yu, B.D., Hess, J.L., Horning, S.E., Brown, G.A., and Korsmeyer, S.J. (1995). Altered Hox expression and segmental identity in Mll-mutant mice. Nature *378*, 505–508.

Yu, M., Mazor, T., Huang, H., Huang, H.-T., Kathrein, K.L., Woo, A.J., Chouinard, C.R., Labadorf, A., Akie, T.E., Moran, T.B., et al. (2012). Direct Recruitment of Polycomb Repressive Complex 1 to Chromatin by Core Binding Transcription Factors. Mol. Cell *45*, 330–343.

Yuan, W., Wu, T., Fu, H., Dai, C., Wu, H., Liu, N., Li, X., Xu, M., Zhang, Z., Niu, T., et al. (2012). Dense Chromatin Activates Polycomb Repressive Complex 2 to Regulate H3 Lysine 27 Methylation. Science *337*, 971–975.

Yuan, W., Xu, M., Huang, C., Liu, N., Chen, S., and Zhu, B. (2011). H3K36 Methylation Antagonizes PRC2-mediated H3K27 Methylation. J. Biol. Chem. *286*, 7983–7989.

Zeisig, D.T., Bittner, C.B., Zeisig, B.B., Garcia-Cuellar, M.P., Hess, J.L., and Slany, R.K. (2005). The eleven-nineteen-leukemia protein ENL connects nuclear MLL fusion partners with chromatin. Oncogene *24*, 5525–5532.

Zeitlinger, J., Stark, A., Kellis, M., Hong, J.-W., Nechaev, S., Adelman, K., Levine, M., and Young, R.A. (2007). RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo. Nat. Genet. *39*, 1512–1516.

Zeleznik-Le, N.J., Harden, A.M., and Rowley, J.D. (1994). 11q23 translocations split the "AT-hook" cruciform DNA-binding region and the transcriptional repression domain from the activation domain of the mixed-lineage leukemia (MLL) gene. Proc. Natl. Acad. Sci. U.S.a. *91*, 10610–10614.

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

Zhao, J., Sun, B.K., Erwin, J.A., Song, J.-J., and Lee, J.T. (2008). Polycomb Proteins Targeted by a Short Repeat RNA to the Mouse X Chromosome. Science *322*, 750–756.

Zhao, L.Y., Niu, Y., Santiago, A., Liu, J., Albert, S.H., Robertson, K.D., and Liao, D. (2006). An EBF3-mediated transcriptional program that induces cell cycle arrest and apoptosis. Cancer Res. *66*, 9445–9452.

Zhao, X.D., Han, X., Chew, J.L., Liu, J., Chiu, K.P., Choo, A., Oriov, Y.L., Sung, W.-K., Shahab, A., Kuznetsov, V.A., et al. (2007). Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. Cell Stem Cell *1*, 286–298.

Zhou, V.W., Goren, A., and Bernstein, B.E. (2011). Charting histone modifications and the functional organization of mammalian genomes. Nat. Rev. Genet. *12*, 7–18.

Zhou, Y.B., Gerchman, S.E., Ramakrishnan, V., Travers, A., and Muyldermans, S. (1998). Position and orientation of the globular domain of linker histone H5 on the nucleosome. Nature *395*, 402–405.

Zhu, J., and Emerson, S.G. (2002). Hematopoietic cytokines, transcription factors and lineage commitment. Oncogene *21*, 3295–3313.

Zieminvanderpoel, S., McCabe, N.R., Gill, H.J., Espinosa, R., Patel, Y., Harden, A., Rubinelli, P., Smith, S.D., Lebeau, M.M., Rowley, J.D., et al. (1991). Identification of a Gene, Mll, That Spans the Breakpoint in 11q23 Translocations Associated with Human Leukemias. Proc. Natl. Acad. Sci. U.S.a. *88*, 10735–10739.

Zuber, J., Rappaport, A.R., Luo, W., Wang, E., Chen, C., Vaseva, A.V., Shi, J., Weissmueller, S., Fellman, C., Taylor, M.J., et al. (2011a). An integrated approach to dissecting oncogene addiction implicates a Myb-coordinated self-renewal program as essential for leukemia maintenance. Genes Dev. *25*, 1628–1640.

Zuber, J., Shi, J., Wang, E., Rappaport, A.R., Herrmann, H., Sison, E.A., Magoon, D., Qi, J., Blatt, K., Wunderlich, M., et al. (2011b). RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. Nature *478*, 524–U124.

(2010). A novel mutation in the miR-128b gene reduces miRNA processing and leads to glucocorticoid resistance of MLL-AF4 acute lymphocytic leukemia cells. Cell Cycle *9*, 1037–1042.